Optimal Control of Stochastic Reaction-Diffusion Equations

vorgelegt von
Lukas Wessels, M.Sc.

an der Fakultät II – Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr. Tobias Breiten
Gutachter:	Prof. Dr. Wilhelm Stannat
Gutachter:	Prof. Dr. Andrzej Święch
Gutachter:	Prof. Dr. Gianmario Tessitore

Tag der wissenschaftlichen Aussprache: 19.07.2022

Berlin 2022

To Melina

Abstract

In this dissertation, we consider finite-horizon optimal control problems governed by stochastic partial differential equations (SPDEs).

In the first part, we derive necessary optimality conditions in terms of adjoint processes as well as the value function. First, we give characterizations of the adjoint processes as solutions of backward stochastic differential equations (BSDEs). In particular, the operator-valued second order adjoint process is represented via its integral kernel, allowing for a characterization as the solution of a function-valued BSDE. Using these BSDEs, we prove Peng's maximum principle for controlled SPDEs. Furthermore, we derive necessary optimality conditions relating the adjoint states to the viscosity differential of the value function evaluated along the optimal trajectory. This extends a well-known relationship between Peng's maximum principle and the dynamic programming approach to the case of controlled SPDEs.

The second part of this dissertation is devoted to sufficient optimality conditions. First, we derive a sufficient optimality condition in terms of the value function. This result exhibits a link between the necessary and the sufficient optimality conditions. Combining this result with a well-known result that identifies the value function as the unique B-continuous viscosity solution of the Hamilton-Jacobi-Bellman equation, we prove a stochastic verification theorem for controlled semilinear SPDEs in the framework of viscosity solutions.

In the last part of this dissertation, we analyze an optimal control problem governed by the stochastic Nagumo model with a view towards efficient numerical approximations. Due to the cubic nonlinearity, our previous results based on global Lipschitz assumptions are not directly applicable in this situation. Therefore, we first investigate the well-posedness of the control problem, and derive a local necessary optimality condition in the spirit of Pontryagin's maximum principle. Next, we show how the restriction to additive noise allows for a simplification of the backward SPDE characterizing the adjoint state to a random backward PDE, which in turn significantly reduces the computational complexity of the approximation of the adjoint state. Finally, we develop a gradient descent method for the approximation of optimal controls and present numerical examples.

Acknowledgements

First and foremost, I am extremely grateful to my advisor Wilhelm Stannat for giving me the opportunity to pursue a PhD, for his guidance throughout my studies, and for creating a productive environment in his work group. In particular, I would like to express my deepest gratitude to him for always taking the time to discuss mathematical problems and advise me on my career.

Also, I wish to thank Andrzej Święch for his invaluable insight into the theory of viscosity solutions and for his constructive feedback on my work. I very much appreciate his agreement to review my thesis, and I am looking forward to working with him during my research stay at the Georgia Institute of Technology this fall.

I would like to extend my sincere thanks to Gianmario Tessitore for helpful discussions on the stochastic maximum principle for controlled SPDEs and for agreeing to be part of the thesis committee.

I am grateful to Barbara Zwicknagl for being my mentor within the Berlin Mathematical School, and for giving guidance throughout the duration of this project.

Many thanks should also go to Tobias Breiten for chairing my PhD defense.

I cannot leave Berlin without mentioning my colleagues at TU Berlin, within the research project CRC 910, and within the Berlin Mathematical School for countless conversations about mathematics and beyond.

Finally, I would like to thank my family and my friends for their support during the time of my studies.

This work was partially funded by Deutsche Forschungsgemeinschaft (DFG) through grant CRC 910 "Control of self-organizing nonlinear systems: Theoretical methods and concepts of application," Project (A10) "Control of stochastic mean-field equations with applications to brain networks."

Contents

List of Figures

List of Acronyms

HJB	Hamilton-Jacobi-Bellman
SDE	Stochastic Differential Equations
BSDE	Backward Stochastic Differential Equation
PDE	Partial Differential Equation
SPDE	Stochastic Partial Differential Equation

List of Symbols

∂_t	partial derivative with respect to the time variable
D^k	k-th order Fréchet derivative with respect to the space variable
$D_t^{1,\pm}$	first order viscosity super- and subdifferential in the time variable, see Definition 3.19
$D_x^{1,\pm}$	first order viscosity super- and subdifferential in the space variable, see Definition 3.15
$D_x^{2,\pm}$	second order viscosity super- and subdifferential in the space variable, see Definition 3.17
$D_{t+,x}^{1,2,\pm}$	parabolic viscosity super- and subdifferential, see Definition 3.3
$\mathcal{H}$	Hamiltonian, see (3.86) and (4.8)
$\mathcal{G}$	generalized Hamiltonian, see (2.104)
$L^p(\Lambda)$	Lebesgue space of p-fold integrable functions on $\Lambda \subset \mathbb{R}^d$
$H^1(\Lambda)$	Sobolev space of order 1 with Neumann boundary conditions on $\Lambda \subset \mathbb{R}^d$
$H_0^\gamma(\Lambda)$	fractional Sobolev space of order $\gamma > 0$ with Dirichlet boundary conditions on $\Lambda \subset \mathbb{R}^d$
$H^{-\gamma}(\Lambda)$	dual space of $H_0^\gamma(\Lambda)$
$L(\mathcal{X}, \mathcal{Y})$	space of linear bounded operators mapping from a Hilbert space $\mathcal{X}$ to a Hilbert space $\mathcal{Y}$
$L(\mathcal{X})$	space of linear bounded operators mapping from $\mathcal{X}$ to itself
$L_1(\mathcal{X}, \mathcal{Y})$	space of trace class operators mapping from $\mathcal{X}$ to $\mathcal{Y}$
$L_1(\mathcal{X})$	space of trace class operators mapping from $\mathcal{X}$ to itself
$L_2(\mathcal{X}, \mathcal{Y})$	space of Hilbert-Schmidt operators mapping from $\mathcal{X}$ to $\mathcal{Y}$
$L_2(\mathcal{X})$	space of Hilbert-Schmidt operators mapping from $\mathcal{X}$ to itself
$\mathcal{S}(\mathcal{X})$	space of bounded, linear, symmetric operators on $\mathcal{X}$
$\mathcal{S}_{\succeq P}(\mathcal{X})$	convex cone of bounded, linear, symmetric, positive operators on $\mathcal{X}$ translated by P, see equation (3.11)

List of Publications

This dissertation is based on the following three articles:

[SW21a] W. Stannat and L. Wessels, *Deterministic control of stochastic reaction-diffusion equations*, Evol. Equ. Control Theory **10** (2021), 701–722.

[SW21b] W. Stannat and L. Wessels, *Peng's maximum principle for stochastic partial differential equations*, SIAM J. Control Optim. **59** (2021), 3552–3573.

[SW22] W. Stannat and L. Wessels, *Necessary and sufficient conditions for optimal control of semilinear stochastic partial differential equations*, preprint, https://arxiv.org/abs/2112.09639, 2022.

1. Introduction

Background

The two great pillars of mathematical control theory are Pontryagin's maximum principle, first developed by Pontryagin et al., see [PBGM62], and the dynamic programming approach developed simultaneously by Richard Bellman, see [Bel57]. Since the development of these theories for controlled ordinary differential equations more than 60 years ago, much research has been geared towards their extension to controlled stochastic ordinary differential equations, controlled partial differential equations (PDEs), and, more recently, controlled stochastic partial differential equations (SPDEs). In this dissertation, we contribute to the generalization of these theories to the class of controlled SPDEs.

A large class of SPDEs widely used in applications is given by stochastic reaction-diffusion equations. Many systems either intrinsically involve randomness, or lower order perturbations are neglected in favor of simplicity. In such situations, noise can be introduced in order to obtain a more realistic mathematical description, formally leading to the equation

$$\partial_t x_t(\lambda) = \Delta x_t(\lambda) + b(x_t(\lambda)) + \xi_t(\lambda), \tag{1.1}$$

where $\lambda \in \Lambda \subset \mathbb{R}^n$ denotes the spatial variable, the Laplace operator Δ models the diffusion, b models a local reaction term, and ξ denotes random fluctuations in space and time. These random fluctuations can be modeled by a Wiener process in the framework of stochastic evolution equations, i.e., for a bounded domain $\Lambda \subset \mathbb{R}^n$, we formulate the equation as a stochastic integral equation in the space $L^2(\Lambda)$ given by

$$\mathrm{d}x_t = [\Delta x_t + B(x_t)]\mathrm{d}t + \Sigma(x_t)\mathrm{d}W_t, \tag{1.2}$$

where $(W_t)_{t\in[s,T]}$ is a cylindrical Wiener process, B corresponds to the reaction term, and Σ models the noise intensity depending on the state of the system.

The objective of control theory is to influence the evolution of the system by an external input in order to achieve a desired outcome. Therefore, we introduce a control function u and a cost functional J into the model. We fix a finite time horizon $T > 0$, an initial time $s \in [0, T)$, and an initial point $x \in L^2(\Lambda)$, and consider the controlled SPDE

$$\begin{cases} \mathrm{d}x_t^u = [\Delta x_t^u + B(x_t^u, u_t)]\mathrm{d}t + \Sigma(x_t^u, u_t)\mathrm{d}W_t, & t \in [s, T] \\ x_s^u = x \in L^2(\Lambda). \end{cases} \tag{1.3}$$

Here, we explicitly denote the dependence of the state x^u on the control u. For precise assumptions on the drift coefficient B and the noise coefficient Σ, see Assumption 2.2 and Assumption 4.1. The usual target for the control problem consists of finding a control u in some set of admissible controls that minimizes a given cost under the expectation, i.e.,

$$J(s,x;u) := \mathbb{E}\left[\int_s^T L(x_t^u, u_t)\mathrm{d}t + H(x_T^u)\right], \tag{1.4}$$

where L and H are real-valued functions called the running costs and the terminal costs, respectively.

In the finite-dimensional deterministic setting, Pontryagin's maximum principle characterizes an optimal control $\bar{u}$ by deriving a necessary condition in terms of an adjoint state p. Moreover, p is characterized as the solution of a backward equation. His approach adapts the direct method from the calculus of variations to optimal control problems. The derivation of the optimality condition is based on a spike variation of an optimal control on a short time interval and a Taylor expansion of the cost functional up to first order.

There are two major obstacles that had to be overcome in order to generalize Pontryagin's result to the stochastic case. The first obstacle is the characterization of the adjoint state by a backward equation with a random terminal condition while still maintaining the adaptedness of the solution process with respect to the underlying filtration. This problem sparked the development of the theory of backward stochastic differential equations (BSDEs) pioneered by Bismut where the solution is a pair of adapted processes (p, q), see [Bis73]. The second obstacle arises when the control enters the noise coefficient. In this case, one has to consider second order Taylor expansions in order to account for the Itô correction term arising due to the spike variation and the unbounded variation of stochastic processes. In order to resolve this issue, Peng introduced a second order adjoint process (P, Q) which he characterized as the solution of a matrix-valued BSDE, see [Pen90].

The problem of necessary conditions for controlled SPDEs was first discussed in [Ben83]. In this work, Bensoussan derives a necessary optimality condition by perturbing an optimal control on the entire time interval, thus avoiding the necessity of the second order adjoint state. However, in contrast to Peng's approach, Bensoussan's approach relies on the convexity of the control domain. For related maximum principles in various situations that only include the first order adjoint state, see [DFT07, FHT18, Gua11, HP90].

The major difficulty in the generalization of Peng's maximum principle (with a control-dependent noise coefficient and general control domain) to the infinite-dimensional case is the characterization of the second order adjoint state. The direct analogue of the matrix-valued BSDE for the second order adjoint state is a BSDE with values in the space of bounded linear operators. However, since this space is merely a Banach space, the classical construction of the stochastic integral, which relies on the Wiener-Itô isometry, fails. Moreover, more general stochastic integration theories for Banach spaces based on the UMD property or the M-type condition do not apply to the space of bounded linear operators either.

In the existing literature, two major approaches have been investigated to circumvent the issue around stochastic integration in Banach spaces. In the first approach, the arising stochastic integrals actually take values in a space of Hilbert-Schmidt operators, which is a Hilbert space and therefore admits a classical stochastic integral. This idea is pursued by Tang and Li in [TL94], however, they need the second Fréchet derivative of the terminal costs to be a Hilbert-Schmidt operator, which in particular excludes the natural choices of quadratic costs or Nemytskii-type costs. In [GT14], Guatteri and Tessitore also construct mild solutions for a class of operator-valued BSDEs. In their mild formulation, the stochastic integral takes values in a space of Hilbert-Schmidt operators between real interpolation spaces associated with the unbounded operator in the equation. Their theory in particular covers the second order adjoint equation with quadratic costs. However, the proof of Peng's maximum principle requires a control of the term (2.87) arising in the duality relation as ε tends to zero, and this analysis has not been executed.

The second approach to avoid the issues around stochastic integration in Banach spaces is to characterize the second order adjoint state based on duality. One of these methods was introduced by Lü and Zhang in [LZ14, LZ15, LZ18] and further studied by Frankowska and Zhang in [FZ20]. In these works, the authors define the notion of transposition solution, which is based on a duality that does not include any stochastic integrals. Nevertheless, they need to impose Lipschitz conditions on the coefficients of the cost functional, which again excludes quadratic costs. Fuhrman, Hu and Tessitore in [FHT12, FHT13], and Du and Meng in [DM13] characterize the second order adjoint state via a so-called stochastic bilinear form. However, while Fuhrman, Hu and Tessitore impose a Lipschitz condition on the coefficients of the cost functional, therefore excluding quadratic costs, Du and Meng assume twice Fréchet differentiability of the coefficients of the state equation excluding Nemytskii-type coefficients. For a more detailed exposition of the existing literature concerning the stochastic maximum principle, see the survey article [Hu19].

In contrast to the direct method, Richard Bellman's idea was to break down the optimization over the time horizon $[s, T]$ into optimization problems over shorter time intervals and exploit the fact that an optimal control also has to be optimal for each of the resulting control problems. To formalize this, he introduced the value function defined as the optimal cost achievable from a starting point (s, x), i.e.,

$$V(s,x) := \inf_u J(s,x;u), \tag{1.5}$$

and derived the dynamic programming principle which in the stochastic case reads

$$V(s,x) = \inf_u \mathbb{E}\left[\int_s^t L(x_r^u, u_r)\mathrm{d}r + V(t, x_t^u)\right], \qquad \forall t \in [s,T]. \tag{1.6}$$

The value function can be used to derive a sufficient optimality condition, the verification theorem, involving the derivatives of V. This in turn enables us to construct optimal controls via a procedure called optimal synthesis. However, obtaining the value function directly from its definition (1.5) is rarely feasible. Therefore, Bellman also derived under smoothness assumptions a differential equation for V, the Hamilton-Jacobi-Bellman (HJB) equation, which in our setting is given as

$$\begin{cases} V_s + \langle \Delta x, DV \rangle_{L^2(\Lambda)} + \inf_u \mathcal{H}(x, u, DV, D^2 V) = 0 \\ V(T, x) = H(x), \end{cases} \tag{1.7}$$

where the Hamiltonian $\mathcal{H} : L^2(\Lambda) \times U \times L^2(\Lambda) \times L(L^2(\Lambda)) \to \mathbb{R}$ is given by

$$\mathcal{H}(x, u, p, P) := L(x, u) + \langle p, B(x, u) \rangle_{L^2(\Lambda)} + \frac{1}{2} \mathrm{tr}\left(\Sigma(x, u)^* P \Sigma(x, u)\right). \tag{1.8}$$

The trace term arises due to the noise in the state equation (1.3) and can be viewed as an infinite-dimensional generalization of the Laplace operator.

In our setting, the HJB equation is a fully nonlinear PDE posed on an infinite dimensional space and involves an unbounded term. Even in the finite-dimensional case, the value function in general does not possess the regularity to satisfy the HJB equation in a classical sense. To overcome this issue, Crandall and Lions introduced a new notion of solution for these equations, called viscosity solutions, see [CL83]. In finite dimensions, there are two equivalent ways to define viscosity solutions: either by introducing a certain class of test functions and replacing the derivatives of the value function by derivatives of the test function, or by replacing the derivatives of the value function by elements in the viscosity differential of the value function. In general, a viscosity solution is merely a continuous function, raising the question how to generalize the verification theorem, which involves derivatives of the value function. Such a generalization was established for finite-dimensional stochastic control problems in [ZYL97, GŚZ05, GŚZ10] by exploiting the equivalence of the two definitions of viscosity solutions.

The extension of the verification theorem and the optimal synthesis to the case of controlled SPDEs is a large and active research field. There are various notions of solutions for second order HJB equations in infinite dimensions. The mild and strong approaches apply to semilinear equations, i.e., when the noise coefficient Σ is independent of the control, and use the linear part of the equation as a smoothing device to obtain solutions that possess higher regularity. This higher regularity can then be exploited to prove the verification theorem, see e.g. [FR17, FG18].

Another way to construct solutions for HJB equations and prove a verification theorem was developed by Fuhrman and Tessitore in [FT02, FT04, FT05]. In this approach, the value function is represented via the solution of a scalar-valued BSDE that is coupled to a forward SPDE. However, currently the available theory is still limited to semilinear HJB equations.

If the noise coefficient Σ depends on the control, the associated HJB equation is fully nonlinear. In this case, the theory of viscosity solutions is well-suited. However, due to the unbounded term in the HJB equation, there is no straightforward generalization of the finite-dimensional theory. In order to handle the unbounded term, Crandall and Lions introduced in [CL90, CL91] the notion of B-continuous viscosity solutions for first order equations, which was generalized by Święch in [Świ94] to second order equations. In this framework, the equivalence between the definition via test functions and the definition via viscosity differentials does not hold anymore. Therefore, the verification theorem in finite dimensions obtained in [ZYL97, GŚZ05, GŚZ10] does not generalize straightforwardly to the infinite-dimensional case involving unbounded terms. In [FGŚ10], Fabbri, Gozzi, and Święch obtained a similar verification theorem for controlled PDEs, imposing the additional assumption that there is an admissible test function at the point in the viscosity differential of the value function. An extension to the infinite-dimensional stochastic setting, still assuming the existence of an admissible test function, can be found in Fabbri's PhD dissertation, see [Fab06]. We would like to emphasize, however, that the proof relies on [YZ99, Chapter 5, Lemma 5.2], which is incorrect as pointed out in [FGG11]. Finally, let us mention that the problem of optimal synthesis in the framework of viscosity solutions is still open. The different approaches to infinite-dimensional HJB equations are presented in the monograph [FGŚ17].

Given these two approaches to optimal control problems, the question of their relationship arises. In the deterministic case, Pontryagin already identified the adjoint state p as the derivative of the value function evaluated along the optimal trajectory. Bismut extended this relationship to the stochastic case by also identifying the second part of the adjoint state, the process q, in terms of the second order derivative of the value function, see [Bis78]. But again, these results rely on the differentiability of the value function. There are various generalizations of the classical relationship between the value function and the adjoint states dispensing with the smoothness assumptions on the value function, both in the deterministic and stochastic setting, as well as in finite and infinite dimensions. In [CF91], Cannarsa and Frankowska derive a connection between the adjoint state and the viscosity differential of the value function in the finite-dimensional deterministic case. This result is generalized to the infinite-dimensional setting in [CF92]. The finite-dimensional stochastic case is treated by Zhou, see [Zho91b]. He also proved a result in the infinite-dimensional stochastic case, but only covered the first order adjoint state, see [Zho91a], which left the full generalization of the relationship between the value function and the adjoint states for the case of controlled SPDEs open.

Many models arising in applications are highly nonlinear and therefore often do not satisfy the usual assumptions imposed on the coefficients of the control problem. As customary in the theory of PDEs, there is no unified mathematical theory that covers control problems associated with all nonlinear state equations. However, in recent years there have been many works devoted to the analysis of control problems for specific highly nonlinear equations. The existence of optimal controls for the Navier-Stokes equation was established by Lisei in [Lis02] and for the FitzHugh-Nagumo system with additive noise by Cordoni and Di Persio in [CDP18]. Necessary first order conditions for optimality are discussed by Fuhrman and Orrieri in [FO16] within the mild approach to dissipative SPDEs with additive noise. In [CDP21] the authors derive first order necessary conditions for the FitzHugh-Nagumo system using a rescaling method which exploits a certain structure of the state equation.

The final step in order to apply the mathematical theory is the development of efficient numerical algorithms for the approximation of optimal controls. There is a rich literature on numerical algorithms for controlled PDEs. Let us mention [BEKT13a, BEKT13b, Ryl17, RLM$^+$16], where the authors analyze deterministic reaction-diffusion models and develop a gradient descent algorithm based on Pontryagin's maximum principle. However, the extension of such deterministic algorithms to the stochastic case poses significant challenges due to the computational complexity of approximating the solution of the backward SPDE. There are few works that actually compute optimal controls for SPDEs via the BSDE approach, see e.g. [DP16, DMPV19]. However, recent years have shown a rising interest in the development of efficient numerical methods for BSDEs, see the survey article [CKSY22], and it is to be expected that these methods will help to develop efficient numerical algorithms to approximate optimal controls in the near future.

Main Results and Outline

Part I: Necessary Optimality Conditions. The first part of this dissertation is devoted to necessary optimality conditions. In Chapter 2 we extend Peng's maximum principle to a certain class of semilinear SPDEs. We study the state equation in the variational setting and impose Nemytskii-type assumptions on the coefficients of the control problem (1.3) and (1.4), i.e., the operator $B : L^2(\Lambda) \times U \to L^2(\Lambda)$ is given by

$$B(x,u)(\lambda) := b(x(\lambda),u), \qquad (x,u) \in L^2(\Lambda) \times U, \tag{1.9}$$

for some function $b : \mathbb{R} \times U \to \mathbb{R}$. Similarly, the operators Σ, L and H are given by Nemytskii operators associated with functions σ, l and h.

In this setting, we are going to prove that an optimal control $\bar{u}$ and the associated trajectory $\bar{x}$ obtained by solving (1.3) corresponding to $u = \bar{u}$, satisfies

$$\mathcal{G}(t,\bar{x}_t,v) \geq \mathcal{G}(t,\bar{x}_t,\bar{u}_t) \tag{1.10}$$

for all $v \in U$, and almost all $(t,\omega) \in [s,T] \times \Omega$, where the generalized Hamiltonian $\mathcal{G} : [s,T] \times L^2(\Lambda) \times U \to \mathbb{R}$ is given by

$$\mathcal{G}(t,x,u) := \mathcal{H}(x,u,p_t,P_t) + \operatorname{tr}\left(\sigma(x,u)^* \left[q_t - P_t\sigma(\bar{x}_t,\bar{u}_t)\right]\right). \tag{1.11}$$

Here, the processes (p,q) and (P,Q) are the first and second order adjoint states, respectively, which we characterize as solutions of certain BSDEs, see (2.54) and (2.70). In particular, the second order adjoint process P, which is an operator-valued process, is in our approach characterized via its integral kernel – again denoted by P – allowing for the derivation of a function-valued BSDE.

In Chapter 3, we use this representation for P to derive additional necessary optimality conditions in terms of the adjoint states and the viscosity differentials of the value function. In particular, we prove that for almost every $t \in [s, T]$,

$$[-\langle \Delta \bar{x}_t, p_t \rangle_{H^{-1}(\Lambda) \times H_0^1(\Lambda)} - \mathcal{G}(t, \bar{x}_t, \bar{u}_t), \infty) \times \{p_t\} \times \mathcal{S}_{\succeq P_t}(L^2(\Lambda)) \subset D^{1,2,+}_{t+,x} V(t, \bar{x}_t), \tag{1.12}$$

$\mathbb{P}$–almost surely, where $\mathcal{S}_{\succeq P_t}(L^2(\Lambda))$ is the convex cone of symmetric, positive operators translated by P_t (see equation (3.11)), and the derivative on the right-hand side is the parabolic viscosity superdifferential of the value function (see Definition 3.3). This means in particular

$$\begin{aligned} \limsup_{\tau \downarrow 0, z \to 0} \frac{1}{\tau + \|z\|^2_{L^2(\Lambda)}} \Big[& V(t+\tau, \bar{x}_t + z) - V(t, \bar{x}_t) \\ & + (\langle \Delta \bar{x}_t, p_t \rangle_{H^{-1}(\Lambda) \times H_0^1(\Lambda)} + \mathcal{G}(t, \bar{x}_t, \bar{u}_t))\tau - \langle p_t, z \rangle_{L^2(\Lambda)} - \frac{1}{2} \langle z, P_t z \rangle_{L^2(\Lambda)} \Big] \leq 0. \end{aligned} \tag{1.13}$$

This relationship between the adjoint states and the value function extends the well-known connection between the dynamic programming approach and Peng's maximum principle to the case of controlled SPDEs. This part is based on the article [SW21b] and the first part of the article [SW22].

Part II: Sufficient Optimality Conditions. The second part of this dissertation is devoted to sufficient optimality conditions. The main result is the verification theorem in the framework of B-continuous viscosity solutions. This result does not rely on the coefficients of the control problem being of Nemytskii-type. Instead we work in the general setting introduced in (1.3) and (1.4). In order to derive the verification theorem, we first prove the following sufficient optimality condition in terms of the value function: Let V be the value function, and let u^* be an admissible control with corresponding state x^*. If there exist processes

$$(G_t, p_t, P_t) \in D^{1,2,+}_{t+,x} V(t, x_t^*) \tag{1.14}$$

such that

$$\mathbb{E}\left[\int_s^T G_t + \langle \Delta x_t^*, p_t \rangle_{H^{-1}(\Lambda) \times H_0^1(\Lambda)} + \mathcal{H}(x_t^*, u_t^*, p_t, P_t) \mathrm{d}t \right] \leq 0, \tag{1.15}$$

then u^* is an optimal control; for the precise statement, see Theorem 4.3. It is well-known that under additional regularity assumptions on the coefficients of the control problem, the value function can be characterized as the unique B-continuous viscosity solution of the HJB equation (1.7). Together with the previous result, this yields a generalization of the classical verification theorem to the case of controlled SPDEs with control-dependent noise. This part is based on the second part of the article [SW22].

Part III: Applications. The last part of this dissertation is devoted to the analysis of the optimal control of the stochastic Nagumo model with a view towards efficient numerical implementations. The Nagumo equation is one of the simplest models exhibiting traveling waves and is therefore studied in various contexts. More specifically, the state is governed by

$$\mathrm{d}x_t = [\Delta x_t + \gamma x_t(x_t - 1)(a - x_t)]\, \mathrm{d}t + \Sigma(t, x_t)\mathrm{d}W_t, \quad t \in [0, T]. \tag{1.16}$$

Due to the cubic nonlinearity, our previous results based on global Lipschitz assumptions are not directly applicable to the associated control problem. Therefore, we first investigate the well-posedness of the control problem, i.e., we prove existence of optimal controls and analyze the regularity of the control-to-state operator. Next, we derive a necessary optimality condition in the spirit of Pontryagin's maximum principle. We also show how the restriction to additive noise allows for a simplification of the backward SPDE characterizing the adjoint state to a random backward PDE. This random backward PDE significantly reduces the computational complexity of the approximation of the adjoint state. Finally, we develop a gradient descent method for the approximation of optimal controls and present numerical examples. This part is based on the article [SW21a].

Part I.

Necessary Optimality Conditions

2. Peng's Maximum Principle

Part I of this dissertation is devoted to necessary optimality conditions for controlled SPDEs. In this chapter we generalize Peng's maximum principle to the case of controlled semilinear SPDEs with Nemytskii-type coefficients. This chapter is based on [SW21b].

2.1. Introduction

In this section, we introduce the precise setting for our control problem.

Let $\Lambda \subset \mathbb{R}$, be a bounded interval. We fix a finite terminal time $T > 0$, an initial time $s \in [0, T)$, and an initial condition $x \in L^2(\Lambda)$, and consider the SPDE

$$\begin{cases} \mathrm{d}x_t^u = [\Delta x_t^u + b(x_t^u, u_t)]\mathrm{d}t + \sigma(x_t^u, u_t)\mathrm{d}W_t, \quad t \in [s, T] \\ x_s^u = x \in L^2(\Lambda), \end{cases} \tag{2.1}$$

where Δ denotes the Laplace operator, b denotes the drift coefficient, σ denotes the noise coefficient, and $(W_t)_{t\in[s,T]}$ is a cylindrical Wiener process. The control problem consists of minimizing the cost functional

$$J(s, x; u) := \mathbb{E}\left[\int_s^T \int_\Lambda l(x_t^u(\lambda), u_t)\mathrm{d}\lambda\mathrm{d}t + \int_\Lambda h(x_T^u(\lambda))\mathrm{d}\lambda\right] \tag{2.2}$$

for running costs l and terminal costs h, subject to the state equation (2.1), over all controls u in some set of admissible controls U_{ad} to be specified below.

For real, separable Hilbert spaces $\mathcal{X}, \mathcal{Y}$, let $L(\mathcal{X}, \mathcal{Y})$ denote the space of bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$, $L_2(\mathcal{X}, \mathcal{Y}) \subset L(\mathcal{X}, \mathcal{Y})$ denote the subspace of Hilbert-Schmidt operators, and $L_1(\mathcal{X}, \mathcal{Y}) \subset L(\mathcal{X}, \mathcal{Y})$ denote the subspace of nuclear operators. Let $L(\mathcal{X}) := L(\mathcal{X}, \mathcal{X})$, $L_2(\mathcal{X}) := L_2(\mathcal{X}, \mathcal{X})$, and $L_1(\mathcal{X}) := L_1(\mathcal{X}, \mathcal{X})$. Furthermore, let $\mathcal{S}(\mathcal{X})$ denote the space of bounded, linear, symmetric operators on $\mathcal{X}$. By an abuse of notation, we are going to use the same notation for a function $P \in L^2(\Lambda^2)$ and the associated operator in $L_2(L^2(\Lambda))$ given by

$$f \mapsto \int_\Lambda P(\cdot, \lambda)f(\lambda)\mathrm{d}\lambda, \tag{2.3}$$

for $f \in L^2(\Lambda)$. Note that $\|P\|_{L^2(\Lambda^2)} = \|P\|_{L_2(L^2(\Lambda))}$.

We impose the following assumptions on the set of admissible controls U_{ad}.

Assumption 2.1. (A1) *Let Ξ be a real, separable Hilbert space and let $(W_t)_{t\in[s,T]}$ be a Ξ-valued cylindrical Wiener process on a filtered probability space*

$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[s,T]}, \mathbb{P})$, where $(\mathcal{F}_t)_{t\in[s,T]}$ is the filtration generated by (W_t) augmented by all $\mathbb{P}$-null sets.

(A2) *Let U be a non-empty subset of a separable Banach space $\mathcal{U}$, and let*

$$U_{ad} := \Big\{ u : [s,T] \times \Omega \to U : \ u \ (\mathcal{F}_t)_{t\in[s,T]} - \text{adapted and} \\ \sup_{t\in[s,T]} \mathbb{E}\left[\|u_t\|_{\mathcal{U}}^k \right] < \infty, \forall k \in \mathbb{N} \Big\} \quad (2.4)$$

be the set of admissible controls. In particular, U_{ad} is not assumed to be convex.

Furthermore, we impose the following assumptions on the coefficients of the control problem.

Assumption 2.2. (B1) *Let $l : \mathbb{R} \times U \to \mathbb{R}$ be twice continuously differentiable with respect to the first variable, and let l, l_x, l_{xx} be continuous in (x,u). Furthermore, assume that there exists a generic constant $C > 0$ such that for all $(x,u) \in \mathbb{R} \times U$ it holds*

$$\begin{cases} |l(x,u)| \le C\left(1 + |x|^2 + \|u\|_{\mathcal{U}}^2\right) \\ |l_x(x,u)| \le C\left(1 + |x| + \|u\|_{\mathcal{U}}\right) \\ |l_{xx}(x,u)| \le C. \end{cases} \quad (2.5)$$

(B2) *Let $h : \mathbb{R} \to \mathbb{R}$ be twice continuously differentiable. Furthermore, assume that there exists a generic constant $C > 0$ such that for all $x \in \mathbb{R}$ it holds*

$$\begin{cases} |h(x)| \le C\left(1 + |x|^2\right) \\ |h_x(x)| \le C\left(1 + |x|\right) \\ |h_{xx}(x)| \le C. \end{cases} \quad (2.6)$$

(B3) *Let $b : \mathbb{R} \times U \to \mathbb{R}$ be twice continuously differentiable with respect to the first variable, and let b, b_x, b_{xx} be continuous in (x,u). Furthermore, assume that there exists a generic constant $C > 0$ such that for all $(x,u) \in \mathbb{R} \times U$ it holds*

$$\begin{cases} |b(x,u)| \le C\left(1 + |x| + \|u\|_{\mathcal{U}}\right) \\ |b_x(x,u)| \le C \\ |b_{xx}(x,u)| \le C. \end{cases} \quad (2.7)$$

(B4) *Let $\sigma : \mathbb{R} \times U \to L_2(\Xi, \mathbb{R})$ be twice continuously differentiable with respect to the first variable, and let $\sigma, \sigma_x, \sigma_{xx}$ be continuous in (x,u). Furthermore, assume that there exists a generic constant $C > 0$ such that for all $(x,u) \in \mathbb{R} \times U$ it holds*

$$\begin{cases} |\sigma(x,u)| \le C\left(1 + |x| + \|u\|_{\mathcal{U}}\right) \\ |\sigma_x(x,u)| \le C \\ |\sigma_{xx}(x,u)| \le C. \end{cases} \quad (2.8)$$

All these coefficients give rise to Nemytskii operators on $L^2(\Lambda)$. For example, we have an operator

$$\begin{aligned} \sigma : & L^2(\Lambda) \times U \to L_2(\Xi, L^2(\Lambda)), \\ & (x, u) \mapsto ((\xi, \lambda) \mapsto \sigma(x(\lambda), u)(\xi)) . \end{aligned} \tag{2.9}$$

Throughout Part I, we are going to use the identification

$$\begin{aligned} L^2(\Lambda; L_2(\Xi, \mathbb{R})) &\cong L_2(\Xi, L^2(\Lambda)) \\ q(\lambda)(\xi) &\leftrightarrow q(\xi)(\lambda). \end{aligned} \tag{2.10}$$

For $\gamma > 0$, let $H_0^\gamma(\Lambda) := W_0^{\gamma,2}(\Lambda)$ be the fractional Sobolev space of order γ with Dirichlet boundary conditions and let $H^{-\gamma}(\Lambda)$ denote its dual space. Under these assumptions, we can solve the state equation in the variational setting, i.e., we work on the Gelfand triple

$$H_0^1(\Lambda) \hookrightarrow L^2(\Lambda) \hookrightarrow H^{-1}(\Lambda), \tag{2.11}$$

and realize $\Delta : H_0^1(\Lambda) \to H^{-1}(\Lambda)$ as a continuous operator, for details see [LR15].

Remark 2.3.
1. The differential operator Δ in the state equation (2.1) can be replaced by the generator $A : \mathcal{D}(A) \subset L^2(\Lambda) \to L^2(\Lambda)$ of the quadratic form

$$\int_\Lambda a(\lambda)(\partial_\lambda x)^2(\lambda) \mathrm{d}\lambda, \quad x \in H_0^1(\Lambda), \tag{2.12}$$

for some $a \in L^\infty(\Lambda)$ with $a(\lambda) \geq a_0 > 0$, that can be formally represented as the second order differential operator in divergence form

$$Ax(\lambda) = \partial_\lambda(a\partial_\lambda x)(\lambda), \tag{2.13}$$

see also Remark 2.13.

2. The restriction to one space-dimension can be relaxed by assuming higher space regularity of the noise coefficient σ, see Remark 2.10.

2.2. Variational Inequality

Following Pontryagin's classical idea, we introduce a so-called spike variation. Throughout this part, let $\bar{u}$ be an optimal control of the control problem (2.2) and (2.1), and let $\bar{x}$ be the associated optimal state. Fix any $v \in U$, $t \in (s, T)$ and $\varepsilon > 0$, and define

$$u_r^\varepsilon := \begin{cases} v, & t \leq r \leq t + \varepsilon \\ \bar{u}_r, & \text{otherwise.} \end{cases} \tag{2.14}$$

Let x^ε denote the state associated with u^ε. Let y^ε denote the solution of the SPDE

$$\begin{cases} \mathrm{d}y_r^\varepsilon = [\Delta y_r^\varepsilon + b_x(\bar{x}_r, \bar{u}_r) y_r^\varepsilon + b(\bar{x}_r, u_r^\varepsilon) - b(\bar{x}_r, \bar{u}_r)]\,\mathrm{d}r \\ \qquad\qquad + [\sigma_x(\bar{x}_r, \bar{u}_r) y_r^\varepsilon + \sigma(\bar{x}_r, u_r^\varepsilon) - \sigma(\bar{x}_r, \bar{u}_r)]\,\mathrm{d}W_r \\ y_s^\varepsilon = 0, \end{cases} \tag{2.15}$$

and let z^ε denote the solution of the SPDE

$$\begin{cases} \mathrm{d}z_r^\varepsilon = \left[\Delta z_r^\varepsilon + b_x(\bar{x}_r, \bar{u}_r) z_r^\varepsilon + \frac{1}{2} b_{xx}(\bar{x}_r, \bar{u}_r) y_r^\varepsilon y_r^\varepsilon + (b_x(\bar{x}_r, u_r^\varepsilon) - b_x(\bar{x}_r, \bar{u}_r)) y_r^\varepsilon\right] \mathrm{d}r \\ \qquad\qquad + \left[\sigma_x(\bar{x}_r, \bar{u}_r) z_r^\varepsilon + \frac{1}{2} \sigma_{xx}(\bar{x}_r, \bar{u}_r) y_r^\varepsilon y_r^\varepsilon + (\sigma_x(\bar{x}_r, u_r^\varepsilon) - \sigma_x(\bar{x}_r, \bar{u}_r)) y_r^\varepsilon\right] \mathrm{d}W_r \\ z_s^\varepsilon = 0. \end{cases} \tag{2.16}$$

These equations are called first and second order variational equations, respectively.

Lemma 2.4. *It holds*

$$\sup_{r \in [s,T]} \mathbb{E}\left[\|x_r^\varepsilon - \bar{x}_r - y_r^\varepsilon - z_r^\varepsilon\|_{L^2(\Lambda)}^2\right] \leq o\left(\varepsilon^2\right) \tag{2.17}$$

as $\varepsilon \downarrow 0$.

Before we prove this lemma, we need the following lemma as a preparation.

Lemma 2.5. *It holds*

$$\sup_{r \in [s,T]} \mathbb{E}\left[\|y_r^\varepsilon\|_{L^2(\Lambda)}^{2k}\right] \leq C\varepsilon^k \tag{2.18}$$

$$\sup_{r \in [s,T]} \mathbb{E}\left[\|z_r^\varepsilon\|_{L^2(\Lambda)}^{k}\right] \leq C\varepsilon^k, \tag{2.19}$$

for $k \in \mathbb{N}$.

Remark 2.6. In Lemma 2.15 below, we prove higher space-regularity for y_T^ε.

Proof. Let us begin with the inequalities for y^ε. By Itô's formula for variational solutions of SPDEs (see [LR15, Theorem 4.2.5]) and elementary estimates, we have

$$\begin{aligned} &\|y_r^\varepsilon\|_{L^2(\Lambda)}^2 \\ &\leq 2\int_s^r (\|b_x\|_\infty + \|\sigma_x\|_\infty^2 + 1)\|y_\theta^\varepsilon\|_{L^2(\Lambda)}^2 \mathrm{d}\theta \\ &\quad + 2\int_s^r \|b(\bar{x}_\theta, u_\theta^\varepsilon) - b(\bar{x}_\theta, \bar{u}_\theta)\|_{L^2(\Lambda)}^2 + \|\sigma(\bar{x}_\theta, u_\theta^\varepsilon) - \sigma(\bar{x}_\theta, \bar{u}_\theta)\|_{L_2(\Xi, L^2(\Lambda))}^2 \mathrm{d}\theta \\ &\quad + 2\int_s^r \langle y_\theta^\varepsilon, \sigma_x(\bar{x}_\theta, \bar{u}_\theta) y_\theta^\varepsilon + \sigma(\bar{x}_\theta, u_\theta^\varepsilon) - \sigma(\bar{x}_\theta, \bar{u}_\theta) \mathrm{d}W_\theta\rangle_{L^2(\Lambda)}. \end{aligned} \tag{2.20}$$

Taking both sides to the power $k \in \mathbb{N}$, and taking the supremum and expectations, we arrive at

$$\begin{aligned}
&\mathbb{E}\left[\sup_{r\in[s,T]}\|y_r^\varepsilon\|_{L^2(\Lambda)}^{2k}\right]\\
&\le C\int_s^T(\|b_x\|_\infty+\|\sigma_x\|_\infty^2+1)^k\mathbb{E}\left[\sup_{\vartheta\in[s,\theta]}\|y_\vartheta^\varepsilon\|_{L^2(\Lambda)}^{2k}\right]\mathrm{d}\theta\\
&\quad+C\mathbb{E}\Bigg[\Bigg(\int_s^T\|b(\bar{x}_\theta,u_\theta^\varepsilon)-b(\bar{x}_\theta,\bar{u}_\theta)\|_{L^2(\Lambda)}^2 \qquad (2.21)\\
&\qquad\qquad+\|\sigma(\bar{x}_\theta,u_\theta^\varepsilon)-\sigma(\bar{x}_\theta,\bar{u}_\theta)\|_{L_2(\Xi,L^2(\Lambda))}^2\mathrm{d}\theta\Bigg)^k\Bigg]\\
&\quad+C\mathbb{E}\left[\sup_{r\in[s,T]}\left|\int_s^r\langle y_\theta^\varepsilon,\sigma_x(\bar{x}_\theta,\bar{u}_\theta)y_\theta^\varepsilon+\sigma(\bar{x}_\theta,u_\theta^\varepsilon)-\sigma(\bar{x}_\theta,\bar{u}_\theta)\mathrm{d}W_\theta\rangle_{L^2(\Lambda)}\right|^k\right].
\end{aligned}$$

Using Burkholder-Davis-Gundy inequality (see e.g. [KS91]), we obtain

$$\begin{aligned}
&\mathbb{E}\left[\sup_{r\in[s,T]}\left|\int_s^r\langle y_\theta^\varepsilon,\sigma_x(\bar{x}_\theta,\bar{u}_\theta)y_\theta^\varepsilon+\sigma(\bar{x}_\theta,u_\theta^\varepsilon)-\sigma(\bar{x}_\theta,\bar{u}_\theta)\mathrm{d}W_\theta\rangle_{L^2(\Lambda)}\right|^k\right]\\
&\le C\mathbb{E}\left[\left\langle\int_s^\cdot\langle y_\theta^\varepsilon,\sigma_x(\bar{x}_\theta,\bar{u}_\theta)y_\theta^\varepsilon+\sigma(\bar{x}_\theta,u_\theta^\varepsilon)-\sigma(\bar{x}_\theta,\bar{u}_\theta)\mathrm{d}W_\theta\rangle_{L^2(\Lambda)}\right\rangle_T^{\frac{k}{2}}\right]\\
&\le C\mathbb{E}\left[\sup_{r\in[s,T]}\|y_r^\varepsilon\|_{L^2(\Lambda)}^{k}\left(\int_s^T\|\sigma_x(\bar{x}_\theta,\bar{u}_\theta)y_\theta^\varepsilon+\sigma(\bar{x}_\theta,u_\theta^\varepsilon)-\sigma(\bar{x}_\theta,\bar{u}_\theta)\|_{L_2(\Xi,L^2(\Lambda))}^2\mathrm{d}\theta\right)^{\frac{k}{2}}\right]\\
&\le C\mathbb{E}\Bigg[\frac{\alpha}{2}\sup_{r\in[s,T]}\|y_r^\varepsilon\|_{L^2(\Lambda)}^{2k}\\
&\qquad+\frac{1}{2\alpha}\left(\int_s^T\|\sigma_x(\bar{x}_\theta,\bar{u}_\theta)y_\theta^\varepsilon+\sigma(\bar{x}_\theta,u_\theta^\varepsilon)-\sigma(\bar{x}_\theta,\bar{u}_\theta)\|_{L_2(\Xi,L^2(\Lambda))}^2\mathrm{d}\theta\right)^k\Bigg],
\end{aligned} \tag{2.22}$$

for every $\alpha>0$. Choosing $\alpha>0$ sufficiently small, we derive from equation (2.21)

$$\begin{aligned}
&\mathbb{E}\left[\sup_{r\in[s,T]}\|y_r^\varepsilon\|_{L^2(\Lambda)}^{2k}\right]\\
&\le C\int_s^T\mathbb{E}\left[\sup_{\vartheta\in[s,\theta]}\|y_\vartheta^\varepsilon\|_{L^2(\Lambda)}^{2k}\right]\mathrm{d}\theta\\
&\quad+C\mathbb{E}\left[\left(\int_s^T\|b(\bar{x}_\theta,u_\theta^\varepsilon)-b(\bar{x}_\theta,\bar{u}_\theta)\|_{L^2(\Lambda)}^2+\|\sigma(\bar{x}_\theta,u_\theta^\varepsilon)-\sigma(\bar{x}_\theta,\bar{u}_\theta)\|_{L_2(\Xi,L^2(\Lambda))}^2\mathrm{d}\theta\right)^k\right].
\end{aligned} \tag{2.23}$$

Using the properties of b, we have

$$\begin{aligned}&\mathbb{E}\left[\left(\int_s^T \|b(\bar{x}_r, u_r^\varepsilon) - b(\bar{x}_r, \bar{u}_r)\|_{L^2(\Lambda)}^2 \mathrm{d}r\right)^k\right] \\ &\leq C\varepsilon^k \mathbb{E}\left[1 + \sup_{r\in[s,T]} \|\bar{x}_r\|_{L^2(\Lambda)}^{2k} + \sup_{r\in[s,T]} \|\bar{u}_r\|_{\mathcal{U}}^{2k} + \|v\|_{\mathcal{U}}^{2k}\right],\end{aligned} \tag{2.24}$$

where the right-hand side is finite using a-priori estimates for variational solutions of SPDEs, see [LR15, Theorem 5.1.3]. Analogously, we obtain the same estimate for the term involving σ. Grönwall's inequality yields the claim for y^ε.

The inequalities for z^ε follow in a similar fashion. The higher order of convergence follows from the fact that the second order expansions in the equation for z^ε satisfy twice the order of the convergence rates of the respective terms in the equation for y^ε. □

Now let us prove Lemma 2.4.

Proof. Applying the first order case of Taylor's theorem for the Gâteaux derivative from [Zei86, Section 4.6] twice, we obtain

$$\begin{aligned}b(\bar{x}_r + y_r^\varepsilon + z_r^\varepsilon, u_r^\varepsilon) =& b(\bar{x}_r, u_r^\varepsilon) + b_x(\bar{x}_r, u_r^\varepsilon)(y_r^\varepsilon + z_r^\varepsilon) \\ &+ \int_0^1\int_0^1 \theta_1 b_{xx}(\bar{x}_r + \theta_1\theta_2(y_r^\varepsilon + z_r^\varepsilon), u_r^\varepsilon)(y_r^\varepsilon + z_r^\varepsilon)(y_r^\varepsilon + z_r^\varepsilon)\mathrm{d}\theta_1\mathrm{d}\theta_2\end{aligned} \tag{2.25}$$

and the same expansion for σ. Using this Taylor expansion and the estimates from Lemma 2.5, the proof is exactly the same as in the finite-dimensional case, see [Pen90]. □

With this result, we can derive the following inequality from the fact that $J(\bar{u}) \leq J(u^\varepsilon)$. This inequality is the basis for deriving the variational inequality.

Lemma 2.7. *It holds*

$$\begin{aligned}&\mathbb{E}\left[\int_s^T\int_\Lambda l_x(\bar{x}_r(\lambda), \bar{u}_r)(y_r^\varepsilon(\lambda) + z_r^\varepsilon(\lambda)) + \frac{1}{2} l_{xx}(\bar{x}_r(\lambda), \bar{u}_r) y_r^\varepsilon(\lambda) y_r^\varepsilon(\lambda)\mathrm{d}\lambda\mathrm{d}r\right] \\ &+\mathbb{E}\left[\int_\Lambda h_x(\bar{x}_T(\lambda))(y_T^\varepsilon(\lambda) + z_T^\varepsilon(\lambda)) + \frac{1}{2} h_{xx}(\bar{x}_T(\lambda)) y_T^\varepsilon(\lambda) y_T^\varepsilon(\lambda)\mathrm{d}\lambda\right] \\ &+\mathbb{E}\left[\int_s^T\int_\Lambda l(\bar{x}_r(\lambda), u_r^\varepsilon) - l(\bar{x}_r(\lambda), \bar{u}_r)\mathrm{d}\lambda\mathrm{d}r\right] \geq o(\varepsilon)\end{aligned} \tag{2.26}$$

as $\varepsilon \downarrow 0$.

Proof. Due to Lemma 2.5, we have

$$\begin{aligned} &l(\bar{x}_r + y_r^\varepsilon + z_r^\varepsilon, \bar{u}_r) - l(\bar{x}_r, \bar{u}_r) \\ &= l_x(\bar{x}_r, \bar{u}_r)(y_r^\varepsilon + z_r^\varepsilon) + \frac{1}{2} l_{xx}(\bar{x}_r, \bar{u}_r)(y_r^\varepsilon + z_r^\varepsilon)(y_r^\varepsilon + z_r^\varepsilon) + o(\varepsilon). \end{aligned} \tag{2.27}$$

Using the same expansion for h, the proof again follows along the same lines as in the finite-dimensional case, see [Pen90]. □

2.3. Adjoint States

In this section, we are going to define the adjoint states using Riesz' representation theorem. We start with the first order adjoint state.

2.3.1. First Order Adjoint State

Consider the SPDE

$$\begin{cases} \mathrm{d}y_r = [\Delta y_r + b_x(\bar{x}_r, \bar{u}_r) y_r + \varphi_r]\,\mathrm{d}r + [\sigma_x(\bar{x}_r, \bar{u}_r) y_r + \psi_r]\,\mathrm{d}W_r \\ y_s = 0, \end{cases} \tag{2.28}$$

where $(\varphi, \psi) \in L^2([s,T] \times \Omega; L^2(\Lambda)) \times L^2([s,T] \times \Omega; L_2(\Xi, L^2(\Lambda)))$. Now, we construct a linear functional on the space $L^2([s,T] \times \Omega; L^2(\Lambda)) \times L^2([s,T] \times \Omega; L_2(\Xi, L^2(\Lambda)))$ as follows

$$\mathcal{T}_1(\varphi, \psi) := \mathbb{E}\left[\int_s^T \int_\Lambda l_x(\bar{x}_r(\lambda), \bar{u}_r) y_r(\lambda) \mathrm{d}\lambda \mathrm{d}r + \int_\Lambda h_x(\bar{x}_T(\lambda)) y_T(\lambda) \mathrm{d}\lambda\right], \tag{2.29}$$

where y denotes the solution of equation (2.28) associated with (φ, ψ). By Riesz' representation theorem, there is a unique pair of adapted processes

$$(p, q) \in L^2([s,T] \times \Omega; L^2(\Lambda)) \times L^2([s,T] \times \Omega; L_2(\Xi, L^2(\Lambda))), \tag{2.30}$$

such that

$$\mathcal{T}_1(\varphi, \psi) = \mathbb{E}\left[\int_s^T \langle \varphi_r, p_r \rangle_{L^2(\Lambda)} + \langle \psi_r, q_r \rangle_{L_2(\Xi, L^2(\Lambda))} \mathrm{d}r\right], \tag{2.31}$$

for all $(\varphi, \psi) \in L^2([s,T] \times \Omega; L^2(\Lambda)) \times L^2([s,T] \times \Omega; L_2(\Xi, L^2(\Lambda)))$. Equation (2.31) is called the first order adjoint state property. We exploit this property once for the process y^ε given by equation (2.15) and once for the process z^ε given by equation (2.16). By choosing (φ, ψ) accordingly, we can rewrite the left-hand side of the inequality in

Lemma 2.7 as

$$\begin{aligned}
&\mathbb{E}\left[\int_s^T \int_\Lambda l_x(\bar{x}_r(\lambda),\bar{u}_r)(y_r^\varepsilon(\lambda)+z_r^\varepsilon(\lambda))+\frac{1}{2}l_{xx}(\bar{x}_r(\lambda),\bar{u}_r)y_r^\varepsilon(\lambda)y_r^\varepsilon(\lambda)\mathrm{d}\lambda\mathrm{d}r\right]\\
&+\mathbb{E}\left[\int_\Lambda h_x(\bar{x}_T(\lambda))(y_T^\varepsilon(\lambda)+z_T^\varepsilon(\lambda))+\frac{1}{2}h_{xx}(\bar{x}_T(\lambda))y_T^\varepsilon(\lambda)y_T^\varepsilon(\lambda)\mathrm{d}\lambda\right]\\
&+\mathbb{E}\left[\int_s^T \int_\Lambda l(\bar{x}_r(\lambda),u_r^\varepsilon)-l(\bar{x}_r(\lambda),\bar{u}_r)\mathrm{d}\lambda\mathrm{d}r\right]\\
&=\mathbb{E}\left[\int_s^T \langle b(\bar{x}_r,u_r^\varepsilon)-b(\bar{x}_r,\bar{u}_r),p_r\rangle_{L^2(\Lambda)}+\langle \sigma(\bar{x}_r,u_r^\varepsilon)-\sigma(\bar{x}_r,\bar{u}_r),q_r\rangle_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}r\right]\\
&\quad+\mathbb{E}\left[\int_s^T \frac{1}{2}\langle b_{xx}(\bar{x}_r,\bar{u}_r)y_r^\varepsilon y_r^\varepsilon,p_r\rangle_{L^2(\Lambda)}+\frac{1}{2}\langle \sigma_{xx}(\bar{x}_r,\bar{u}_r)y_r^\varepsilon y_r^\varepsilon,q_r\rangle_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}r\right]\\
&\quad+\mathbb{E}\left[\int_s^T \int_\Lambda \frac{1}{2}l_{xx}(\bar{x}_r(\lambda),\bar{u}_r)y_r^\varepsilon(\lambda)y_r^\varepsilon(\lambda)\mathrm{d}\lambda+\int_\Lambda l(\bar{x}_r(\lambda),u_r^\varepsilon)-l(\bar{x}_r(\lambda),\bar{u}_r)\mathrm{d}\lambda\mathrm{d}r\right]\\
&\quad+\mathbb{E}\left[\frac{1}{2}\int_\Lambda h_{xx}(\bar{x}_T(\lambda))y_T^\varepsilon(\lambda)y_T^\varepsilon(\lambda)\mathrm{d}\lambda\right]+o(\varepsilon).
\end{aligned}\tag{2.32}$$

Note that the term

$$\begin{aligned}
\mathbb{E}\Bigg[\int_s^T &\langle (b_x(\bar{x}_r,u_r^\varepsilon)-b_x(\bar{x}_r,\bar{u}_r))y_r^\varepsilon,p_r\rangle_{L^2(\Lambda)}\\
&+\langle (\sigma_x(\bar{x}_r,u_r^\varepsilon)-\sigma_x(\bar{x}_r,\bar{u}_r))y_r^\varepsilon,q_r\rangle_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}r\Bigg]
\end{aligned}\tag{2.33}$$

is of order $o(\varepsilon)$ and hence can be omitted.

2.3.2. Mollified Second Order Adjoint State

In order to handle the quadratic terms using the same idea as for the linear terms, we have to turn the bilinear forms into linear forms on the tensor product $L^2(\Lambda)\otimes L^2(\Lambda)\cong L^2(\Lambda^2)$ (see [RS80, Theorem II.10] for the isomorphism).

Proposition 2.8. *The process $Y_r^\varepsilon(\lambda,\mu):=y_r^\varepsilon(\lambda)y_r^\varepsilon(\mu)$, $\lambda,\mu\in\Lambda$, is in the space*

$$L^2([s,T]\times\Omega;H_0^1(\Lambda^2))\cap L^2(\Omega;C([s,T];L^2(\Lambda^2)))\tag{2.34}$$

and satisfies the equation

$$\begin{cases} \mathrm{d}Y_r^\varepsilon(\lambda,\mu) = [\Delta Y_r^\varepsilon(\lambda,\mu) + (b_x(\bar{x}_r(\lambda),\bar{u}_r) + b_x(\bar{x}_r(\mu),\bar{u}_r))Y_r^\varepsilon(\lambda,\mu) \\ \qquad\qquad\qquad + \langle \sigma_x(\bar{x}_r(\lambda),\bar{u}_r), \sigma_x(\bar{x}_r(\mu),\bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})} Y_r^\varepsilon(\lambda,\mu) + \Phi_r^\varepsilon(\lambda,\mu)]\mathrm{d}r \\ \qquad\qquad\qquad + [(\sigma_x(\bar{x}_r(\lambda),\bar{u}_r) + \sigma_x(\bar{x}_r(\mu),\bar{u}_r))Y_r^\varepsilon(\lambda,\mu) + \Psi_r^\varepsilon(\lambda,\mu)]\mathrm{d}W_r \\ Y_s^\varepsilon = 0, \end{cases} \tag{2.35}$$

where

$$(\Phi^\varepsilon,\Psi^\varepsilon) \in L^2([s,T]\times\Omega; L^2(\Lambda^2)) \times L^2([s,T]\times\Omega; L_2(\Xi, L^2(\Lambda^2))) \tag{2.36}$$

are given by

$$\begin{aligned} \Phi_r^\varepsilon(\lambda,\mu) =& y_r^\varepsilon(\lambda)(b(\bar{x}_r(\mu),u_r^\varepsilon) - b(\bar{x}_r(\mu),\bar{u}_r)) + y_r^\varepsilon(\mu)(b(\bar{x}_r(\lambda),u_r^\varepsilon) - b(\bar{x}_r(\lambda),\bar{u}_r)) \\ &+ \langle \sigma_x(\bar{x}_r(\lambda),\bar{u}_r)y_r^\varepsilon(\lambda), \sigma(\bar{x}_r(\mu),u_r^\varepsilon) - \sigma(\bar{x}_r(\mu),\bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})} \\ &+ \langle \sigma_x(\bar{x}_r(\mu),\bar{u}_r)y_r^\varepsilon(\mu), \sigma(\bar{x}_r(\lambda),u_r^\varepsilon) - \sigma(\bar{x}_r(\lambda),\bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})} \\ &+ \langle \sigma(\bar{x}_r(\lambda),u_r^\varepsilon) - \sigma(\bar{x}_r(\lambda),\bar{u}_r), \sigma(\bar{x}_r(\mu),u_r^\varepsilon) - \sigma(\bar{x}_r(\mu),\bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})}, \end{aligned} \tag{2.37}$$

and

$$\Psi_r^\varepsilon(\lambda,\mu) = (\sigma(\bar{x}_r(\lambda),u_r^\varepsilon) - \sigma(\bar{x}_r(\lambda),\bar{u}_r))y_r^\varepsilon(\mu) + (\sigma(\bar{x}_r(\mu),u_r^\varepsilon) - \sigma(\bar{x}_r(\mu),\bar{u}_r))y_r^\varepsilon(\lambda). \tag{2.38}$$

Proof. The regularity of Y^ε follows from the regularity of y^ε. Applying Itô's product rule for real-valued semimartingales to

$$\langle Y_r^\varepsilon, f_1 \otimes f_2\rangle_{L^2(\Lambda^2)} = \langle y_r^\varepsilon, f_1\rangle_{L^2(\Lambda)} \langle y_r^\varepsilon, f_2\rangle_{L^2(\Lambda)}, \tag{2.39}$$

$f_1, f_2 \in H_0^1(\Lambda)$, and using a density argument yields

$$\begin{aligned} \mathrm{d}Y_r^\varepsilon(\lambda,\mu) =& y_r^\varepsilon(\lambda)\mathrm{d}y_r^\varepsilon(\mu) + y_r^\varepsilon(\mu)\mathrm{d}y_r^\varepsilon(\lambda) + \mathrm{d}\langle y^\varepsilon(\lambda), y^\varepsilon(\mu)\rangle_r \\ =& y_r^\varepsilon(\lambda)(\Delta_\mu y_r^\varepsilon(\mu) + b_x(\bar{x}_r(\mu),\bar{u}_r)y_r^\varepsilon(\mu) + b(\bar{x}_r(\mu),u_r^\varepsilon) - b(\bar{x}_r(\mu),\bar{u}_r))\mathrm{d}r \\ &+ y_r^\varepsilon(\mu)(\Delta_\lambda y_r^\varepsilon(\lambda) + b_x(\bar{x}_r(\lambda),\bar{u}_r)y_r^\varepsilon(\lambda) + b(\bar{x}_r(\lambda),u_r^\varepsilon) - b(\bar{x}_r(\lambda),\bar{u}_r))\mathrm{d}r \\ &+ \langle \sigma_x(\bar{x}_r(\lambda),\bar{u}_r)y_r^\varepsilon(\lambda) + \sigma(\bar{x}_r(\lambda),u_r^\varepsilon) - \sigma(\bar{x}_r(\lambda),\bar{u}_r), \\ &\qquad \sigma_x(\bar{x}_r(\mu),\bar{u}_r)y_r^\varepsilon(\mu) + \sigma(\bar{x}_r(\mu),u_r^\varepsilon) - \sigma(\bar{x}_r(\mu),\bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})}\mathrm{d}r \\ &+ y_r^\varepsilon(\lambda)(\sigma_x(\bar{x}_r(\mu),\bar{u}_r)y_r^\varepsilon(\mu) + \sigma(\bar{x}_r(\mu),u_r^\varepsilon) - \sigma(\bar{x}_r(\mu),\bar{u}_r))\mathrm{d}W_r \\ &+ y_r^\varepsilon(\mu)(\sigma_x(\bar{x}_r(\lambda),\bar{u}_r)y_r^\varepsilon(\lambda) + \sigma(\bar{x}_r(\lambda),u_r^\varepsilon) - \sigma(\bar{x}_r(\lambda),\bar{u}_r))\mathrm{d}W_r \end{aligned} \tag{2.40}$$

in $L^2(\Lambda^2)$. Note that

$$y_r^\varepsilon(\lambda)\Delta_\mu y_r^\varepsilon(\mu) + y_r^\varepsilon(\mu)\Delta_\lambda y_r^\varepsilon(\lambda) = \Delta Y_r^\varepsilon(\lambda,\mu), \tag{2.41}$$

where the Laplace operator Δ acting on functions in the two variables (λ, μ) arises via the tensor product

$$\Delta = I \otimes \Delta_\mu + \Delta_\lambda \otimes I, \tag{2.42}$$

see [RS80, Section VIII.10] for more details on tensor products of operators.

Combining the remaining terms in a similar fashion, we end up with

$$\begin{aligned} \mathrm{d}Y_r^\varepsilon(\lambda, \mu) =& [\Delta Y_r^\varepsilon(\lambda, \mu) + (b_x(\bar{x}_r(\lambda), \bar{u}_r) + b_x(\bar{x}_r(\mu), \bar{u}_r))Y_r^\varepsilon(\lambda, \mu) \\ &+ \langle \sigma_x(\bar{x}_r(\lambda), \bar{u}_r), \sigma_x(\bar{x}_r(\mu), \bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})} Y_r^\varepsilon(\lambda, \mu) + \Phi_r^\varepsilon(\lambda, \mu)]\mathrm{d}r \\ &+ [(\sigma_x(\bar{x}_r(\lambda), \bar{u}_r) + \sigma_x(\bar{x}_r(\mu), \bar{u}_r))Y_r^\varepsilon(\lambda, \mu) + \Psi_r^\varepsilon(\lambda, \mu)]\mathrm{d}W_r, \end{aligned} \tag{2.43}$$

for Φ^ε and Ψ^ε as stated in the proposition. This concludes the proof. □

We can now rewrite the quadratic terms in y^ε in the variational inequality into linear terms in Y^ε evaluated on the diagonal in Λ^2.

Proposition 2.9. *It holds*

$$\begin{aligned} &\mathbb{E}\left[\int_s^T \langle b_{xx}(\bar{x}_r, \bar{u}_r)y_r^\varepsilon y_r^\varepsilon, p_r\rangle_{L^2(\Lambda)} + \langle \sigma_{xx}(\bar{x}_r, \bar{u}_r)y_r^\varepsilon y_r^\varepsilon, q_r\rangle_{L_2(\Xi,L^2(\Lambda))}\,\mathrm{d}r\right] \\ &+ \mathbb{E}\left[\int_s^T \int_\Lambda l_{xx}(\bar{x}_r, \bar{u}_r)y_r^\varepsilon y_r^\varepsilon \mathrm{d}\lambda\mathrm{d}r\right] \\ &= \mathbb{E}\left[\int_s^T \int_\Lambda \big(b_{xx}(\bar{x}_r(\lambda), \bar{u}_r)p_r(\lambda) + \langle \sigma_{xx}(\bar{x}_r(\lambda), \bar{u}_r), q_r(\lambda)\rangle_{L_2(\Xi,\mathbb{R})}\big)\, \delta(Y_r^\varepsilon)(\lambda)\mathrm{d}\lambda\mathrm{d}r\right] \\ &\quad + \mathbb{E}\left[\int_s^T \int_\Lambda l_{xx}(\bar{x}_r(\lambda), \bar{u}_r)\delta(Y_r^\varepsilon)(\lambda)\mathrm{d}\lambda\mathrm{d}r\right], \end{aligned} \tag{2.44}$$

where $\delta : H_0^1(\Lambda^2) \to L^2(\Lambda)$ is defined by $\delta(w)(\lambda) := w(\lambda, \lambda)$.

Proof. Let $(\xi_k)_{k\in\mathbb{N}}$ be an orthonormal basis of Ξ. We have

$$\begin{aligned} \langle \sigma_{xx}(\bar{x}_r, \bar{u}_r)y_r^\varepsilon y_r^\varepsilon, q_r\rangle_{L_2(\Xi,L^2(\Lambda))} &= \sum_{k=1}^\infty \langle \sigma_{xx}(\bar{x}_r, \bar{u}_r)(\xi_k)y_r^\varepsilon y_r^\varepsilon, q_r(\xi_k)\rangle_{L^2(\Lambda)} \\ &= \sum_{k=1}^\infty \int_\Lambda \sigma_{xx}(\bar{x}_r(\lambda), \bar{u}_r)(\xi_k)y_r^\varepsilon(\lambda)y_r^\varepsilon(\lambda)q_r(\xi_k)(\lambda)\mathrm{d}\lambda \\ &= \int_\Lambda \delta(Y_r^\varepsilon)(\lambda)\langle \sigma_{xx}(\bar{x}_r(\lambda), \bar{u}_r), q_r(\lambda)\rangle_{L_2(\Xi,\mathbb{R})}\mathrm{d}\lambda. \end{aligned} \tag{2.45}$$

A similar calculation shows the claim for the remaining terms. □

The operator $\delta : H_0^1(\Lambda^2) \to L^2(\Lambda)$, $\Lambda \subset \mathbb{R}$, is continuous due to the Sobolev

Imbedding Theorem, see [AF03, Theorem 4.12]. Since

$$Y^\varepsilon \in L^2([s,T]\times\Omega; H_0^1(\Lambda^2)), \tag{2.46}$$

the right-hand side of equation (2.44) is linear and bounded in Y^ε. However, the spatial regularity of the solution evaluated at the terminal time T is not sufficient for

$$\mathbb{E}\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda))Y_T^\varepsilon(\lambda,\lambda)\mathrm{d}\lambda\right] \tag{2.47}$$

to be continuous in Y_T^ε. In order to obtain a continuous operator in Y_T^ε, we have to mollify the terminal condition. Using the heat kernel, we have

$$\begin{aligned}
&\mathbb{E}\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda))y_T^\varepsilon(\lambda)y_T^\varepsilon(\lambda)\mathrm{d}\lambda\right] \\
&= \lim_{\eta\to 0}\mathbb{E}\left[\int_{\Lambda^2}\frac{1}{2}\left(h_{xx}(\bar{x}_T(\lambda))+h_{xx}(\bar{x}_T(\mu))\right)y_T^\varepsilon(\lambda)y_T^\varepsilon(\mu)\frac{1}{\sqrt{4\pi\eta}}\exp\left(-\frac{|\lambda-\mu|^2}{4\eta}\right)\mathrm{d}\mu\mathrm{d}\lambda\right] \\
&= \lim_{\eta\to 0}\mathbb{E}\left[\int_{\Lambda^2}\frac{1}{2}\left(h_{xx}(\bar{x}_T(\lambda))+h_{xx}(\bar{x}_T(\mu))\right)\frac{1}{\sqrt{4\pi\eta}}\exp\left(-\frac{|\lambda-\mu|^2}{4\eta}\right)Y_T^\varepsilon(\lambda,\mu)\mathrm{d}\lambda\mathrm{d}\mu\right].
\end{aligned} \tag{2.48}$$

We denote

$$h_{xx}^\eta(\lambda,\mu) := \frac{1}{2}\left(h_{xx}(\bar{x}_T(\lambda))+h_{xx}(\bar{x}_T(\mu))\right)\frac{1}{\sqrt{4\pi\eta}}\exp\left(-\frac{|\lambda-\mu|^2}{4\eta}\right)\in L^2(\Lambda^2). \tag{2.49}$$

With this mollification and Proposition 2.9, we construct another bounded, linear functional via

$$\begin{aligned}
&\mathcal{T}_2^\eta(\Phi,\Psi) \\
&:= \mathbb{E}\left[\int_s^T\int_\Lambda \left(b_{xx}(\bar{x}_r(\lambda),\bar{u}_r)p_r(\lambda)+\langle\sigma_{xx}(\bar{x}_r(\lambda),\bar{u}_r),q_r(\lambda)\rangle_{L_2(\Xi,\mathbb{R})}\right)\delta(Y_r)(\lambda)\mathrm{d}\lambda\mathrm{d}r\right] \\
&\quad + \mathbb{E}\left[\int_s^T\int_\Lambda l_{xx}(\bar{x}_r(\lambda),\bar{u}_r)\delta(Y_r)(\lambda)\mathrm{d}\lambda\mathrm{d}r + \int_{\Lambda^2}h_{xx}^\eta(\lambda,\mu)Y_T(\lambda,\mu)\mathrm{d}\lambda\mathrm{d}\mu\right],
\end{aligned} \tag{2.50}$$

where Y denotes the solution of equation (2.35) with $(\Phi^\varepsilon,\Psi^\varepsilon)$ replaced by (Φ,Ψ). By Riesz' representation theorem, there exists a pair

$$(P^\eta,Q^\eta)\in L^2([s,T]\times\Omega;L^2(\Lambda^2))\times L^2([s,T]\times\Omega;L_2(\Xi,L^2(\Lambda^2))) \tag{2.51}$$

such that

$$\mathcal{T}_2^\eta(\Phi,\Psi) = \mathbb{E}\left[\int_s^T\langle P_r^\eta,\Phi_r\rangle_{L^2(\Lambda^2)}+\langle Q_r^\eta,\Psi_r\rangle_{L_2(\Xi,L^2(\Lambda^2))}\mathrm{d}r\right], \tag{2.52}$$

for all $(\Phi, \Psi) \in L^2([s,T] \times \Omega; L^2(\Lambda^2)) \times L^2([s,T] \times \Omega; L_2(\Xi, L^2(\Lambda^2)))$. The pair (P^η, Q^η) is called the mollified second order adjoint state, and (2.52) is called the mollified second order adjoint state property. Choosing $\Phi = \Phi^\varepsilon$ and $\Psi = \Psi^\varepsilon$ as given by equations (2.37) and (2.38), respectively, and using (2.32) we can rewrite the inequality from Lemma 2.7 as

$$\begin{aligned}
&\mathbb{E}\left[\int_s^T \langle b(\bar{x}_r, u_r^\varepsilon) - b(\bar{x}_r, \bar{u}_r), p_r\rangle_{L^2(\Lambda)} + \langle \sigma(\bar{x}_r, u_r^\varepsilon) - \sigma(\bar{x}_r, \bar{u}_r), q_r\rangle_{L_2(\Xi, L^2(\Lambda))}\,\mathrm{d}r\right] \\
&+ \frac{1}{2}\mathbb{E}\left[\int_s^T \langle P_r^\eta, \Phi_r^\varepsilon\rangle_{L^2(\Lambda^2)} + \langle Q_r^\eta, \Psi_r^\varepsilon\rangle_{L_2(\Xi, L^2(\Lambda^2))}\mathrm{d}r\right] \\
&+ \mathbb{E}\left[\int_s^T \int_\Lambda l(\bar{x}_r(\lambda), u_r^\varepsilon) - l(\bar{x}_r(\lambda), \bar{u}_r)\mathrm{d}\lambda\mathrm{d}r\right] \\
&+ \frac{1}{2}\mathbb{E}\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda)) y_T^\varepsilon(\lambda) y_T^\varepsilon(\lambda)\mathrm{d}\lambda - \int_{\Lambda^2} h_{xx}^\eta(\lambda,\mu) Y_T^\varepsilon(\lambda,\mu)\mathrm{d}\lambda\mathrm{d}\mu\right] \geq o(\varepsilon).
\end{aligned} \tag{2.53}$$

Remark 2.10. The restriction to one space-dimension goes back to the required continuity of the operator δ defined in (2.9). For two-dimensional $\Lambda \subset \mathbb{R}^2$, δ maps from $H_0^1(\Lambda^2)$ to $L^2(\Lambda)$, which means that we lose two space-dimensions and therefore lose the continuity of δ. However, continuity can be restored if we have the space regularity $H_0^{1+\epsilon}(\Lambda^2)$, $\epsilon > 0$, see [AF03, Section 7.43]. This can be achieved by assuming higher space-regularity on the noise coefficient σ.

2.4. Adjoint Equations

In this section, we are going to deduce equations for the adjoint states (p, q) and (P^η, Q^η), respectively.

2.4.1. First Order Adjoint Equation

We introduce the following first order adjoint equation

$$\begin{cases} \mathrm{d}p_r = -\left[\Delta p_r + b_x(\bar{x}_r, \bar{u}_r)p_r + \langle \sigma_x(\bar{x}_r, \bar{u}_r), q_r\rangle_{L_2(\Xi,\mathbb{R})} + l_x(\bar{x}_r, \bar{u}_r)\right]\mathrm{d}r + q_r\mathrm{d}W_r \\ p_T = h_x(\bar{x}_T). \end{cases} \tag{2.54}$$

The existence of a unique variational solution (p, q) to this equation, where

$$p \in L^2([s,T] \times \Omega; H_0^1(\Lambda)) \cap L^2(\Omega; C([s,T]; L^2(\Lambda))) \tag{2.55}$$

and

$$q \in L^2([s,T] \times \Omega; L_2(\Xi, L^2(\Lambda))), \tag{2.56}$$

can be found in [Ben83]. In order to verify the adjoint state property (2.31) we need to apply Itô's formula to the process $\langle p_r, y_r\rangle_{L^2(\Lambda)}$, where y denotes the solution of equation (2.28) associated with (φ, ψ). To this end we need Itô's formula for variational solutions of SPDEs (see [Par21, Lemma 2.15] or [Kry13, Section 3]) with $V := H_0^1(\Lambda) \times H_0^1(\Lambda)$, $H := L^2(\Lambda) \times L^2(\Lambda)$, and $F : H \to \mathbb{R}$, $(x, y) \mapsto \langle x, y\rangle_{L^2(\Lambda)}$. This yields

$$\begin{aligned}
\mathrm{d}\langle p_r, y_r\rangle_{L^2(\Lambda)} =& \langle p_r, \mathrm{d}y_r\rangle_{L^2(\Lambda)} + \langle y_r, \mathrm{d}p_r\rangle_{L^2(\Lambda)} + \mathrm{d}\langle p, y\rangle_r \\
=& \langle p_r, b_x(\bar{x}_r, \bar{u}_r)y_r + \varphi_r\rangle_{L^2(\Lambda)}\mathrm{d}r \\
&- \langle y_r, b_x(\bar{x}_r, \bar{u}_r)p_r + \langle \sigma_x(\bar{x}_r, \bar{u}_r), q_r\rangle_{L_2(\Xi,\mathbb{R})} + l_x(\bar{x}_r, \bar{u}_r)\rangle_{L^2(\Lambda)}\mathrm{d}r \\
&+ \langle q_r, \sigma_x(\bar{x}_r, \bar{u}_r)y_r + \psi_r\rangle_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}r \\
&+ \langle (\sigma_x(\bar{x}_r, \bar{u}_r)y_r + \psi_r)^* \, p_r, \mathrm{d}W_r\rangle_{L^2(\Lambda)} + \langle q_r^* y_r, \mathrm{d}W_r\rangle_{L^2(\Lambda)} \\
=& \left[\langle p_r, \varphi_r\rangle_{L^2(\Lambda)} + \langle q_r, \psi_r\rangle_{L_2(\Xi,L^2(\Lambda))} - \langle y_r, l_x(\bar{x}_r, \bar{u}_r)\rangle_{L^2(\Lambda)}\right]\mathrm{d}r \\
&+ \langle (\sigma_x(\bar{x}_r, \bar{u}_r)y_r + \psi_r)^* \, p_r, \mathrm{d}W_r\rangle_{L^2(\Lambda)} + \langle q_r^* y_r, \mathrm{d}W_r\rangle_{L^2(\Lambda)}.
\end{aligned} \tag{2.57}$$

Hence, considering the terminal condition, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\langle h_x(\bar{x}_T), y_T\rangle_{L^2(\Lambda)}\right] \\
&= \mathbb{E}\left[\int_s^T \langle p_r, \varphi_r\rangle_{L^2(\Lambda)} + \langle q_r, \psi_r\rangle_{L_2(\Xi,L^2(\Lambda))} - \langle y_r, l_x(\bar{x}_r, \bar{u}_r)\rangle_{L^2(\Lambda)}\mathrm{d}r\right],
\end{aligned} \tag{2.58}$$

which is the adjoint state property.

The strategy for the mollified second order adjoint state is the same: First, we introduce the mollified second order adjoint equation and show that a solution of that equation exists; then we apply Itô's formula and show that the solution satisfies the mollified adjoint state property (2.52), which characterizes it as the mollified second order adjoint state. Afterwards, in Section 2.5, we pass to the limit to derive an equation for the second order adjoint state.

2.4.2. Mollified Second Order Adjoint Equation

We introduce the mollified second order adjoint equation

$$\begin{cases}
\mathrm{d}P_r^\eta(\lambda,\mu) = -[\Delta P_r^\eta(\lambda,\mu) + (b_x(\bar{x}_r(\lambda), \bar{u}_r) + b_x(\bar{x}_r(\mu), \bar{u}_r))P_r^\eta(\lambda,\mu) \\
\qquad\qquad + \langle \sigma_x(\bar{x}_r(\lambda), \bar{u}_r), \sigma_x(\bar{x}_r(\mu), \bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})} P_r^\eta(\lambda,\mu) \\
\qquad\qquad + \langle \sigma_x(\bar{x}_r(\lambda), \bar{u}_r) + \sigma_x(\bar{x}_r(\mu), \bar{u}_r), Q_r^\eta(\lambda,\mu)\rangle_{L_2(\Xi,\mathbb{R})} \\
\qquad\qquad + \delta^*(l_{xx}(\bar{x}_r(\lambda), \bar{u}_r)) + \delta^*(b_{xx}(\bar{x}_r(\lambda), \bar{u}_r)p_r(\lambda)) \\
\qquad\qquad + \delta^*(\langle \sigma_{xx}(\bar{x}_r(\lambda), \bar{u}_r), q_r\rangle_{L_2(\Xi,\mathbb{R})})]\mathrm{d}r + Q_r^\eta(\lambda,\mu)\mathrm{d}W_r \\
P_T^\eta(\lambda,\mu) = h_{xx}^\eta(\lambda,\mu),
\end{cases} \tag{2.59}$$

where h_{xx}^{η} is given by equation (2.49), and $\delta^* : L^2(\Lambda) \to H^{-1}(\Lambda^2)$ is the adjoint of the operator introduced in (2.9), i.e.,

$$\langle \delta^*(f), w \rangle_{H^{-1}(\Lambda^2)\times H_0^1(\Lambda^2)} := \int_{\Lambda} f(\lambda)\delta(w)(\lambda)\mathrm{d}\lambda = \int_{\Lambda} f(\lambda)w(\lambda,\lambda)\mathrm{d}\lambda, \tag{2.60}$$

for $f \in L^2(\Lambda)$, $w \in H_0^1(\Lambda^2)$.

Proposition 2.11. *The mollified second order adjoint equation* (2.59) *has a unique variational solution* (P^{η}, Q^{η})*, where*

$$P^{\eta} \in L^2([s,T]\times\Omega; H_0^1(\Lambda^2)) \cap L^2(\Omega; C([s,T]; L^2(\Lambda))) \tag{2.61}$$

and

$$Q^{\eta} \in L^2([s,T]\times\Omega; L_2(\Xi, L^2(\Lambda^2))). \tag{2.62}$$

Proof. We apply the result from [Ben83] on the Gelfand triple

$$H_0^1(\Lambda^2) \hookrightarrow L^2(\Lambda^2) \hookrightarrow H^{-1}(\Lambda^2). \tag{2.63}$$

□

2.4.3. Adjoint State Property for the Mollified Second Order Adjoint State

Now, we are going to show that the solution of the mollified second order adjoint equation satisfies the mollified adjoint state property (2.52). To this end let Y denote the solution of the second variational equation (2.35) associated with (Φ, Ψ), and let (P^{η}, Q^{η}) denote the solution of the mollified second order adjoint equation (2.59). We again apply Itô's formula for variational solutions of SPDEs, this time to the expression

$$\langle P_r^{\eta}(\lambda,\mu), Y_r(\lambda,\mu)\rangle_{L^2(\Lambda^2)}. \tag{2.64}$$

Choosing $V := H_0^1(\Lambda^2)\times H_0^1(\Lambda^2)$, $H := L^2(\Lambda^2)\times L^2(\Lambda^2)$, and $F : H \to \mathbb{R}$, $(x,y) \mapsto \langle x, y\rangle_{L^2(\Lambda^2)}$, yields

$$\begin{aligned} &\mathrm{d}\langle P_r^{\eta}(\lambda,\mu), Y_r(\lambda,\mu)\rangle_{L^2(\Lambda^2)} \\ &= \langle P_r^{\eta}(\lambda,\mu), \mathrm{d}Y_r(\lambda,\mu)\rangle_{L^2(\Lambda^2)} + \langle Y_r(\lambda,\mu), \mathrm{d}P_r^{\eta}(\lambda,\mu)\rangle_{L^2(\Lambda^2)} + \mathrm{d}\langle P^{\eta}(\lambda,\mu), Y(\lambda,\mu)\rangle_r. \end{aligned} \tag{2.65}$$

Plugging in the equations for P^η and Y, respectively, we arrive at

$$
\begin{aligned}
&\mathrm{d}\langle P^\eta_r(\lambda,\mu), Y_r(\lambda,\mu)\rangle_{L^2(\Lambda^2)} \\
&= \Big\langle \Delta Y_r(\lambda,\mu) + (b_x(\bar{x}_r(\lambda),\bar{u}_r) + b_x(\bar{x}_r(\mu),\bar{u}_r))Y_r(\lambda,\mu) \\
&\qquad + \langle \sigma_x(\bar{x}_r(\lambda),\bar{u}_r), \sigma_x(\bar{x}_r(\mu),\bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})} Y_r(\lambda,\mu) \\
&\qquad + \Phi_r(\lambda,\mu), P^\eta_r(\lambda,\mu)\Big\rangle_{H^{-1}(\Lambda^2)\times H^1_0(\Lambda^2)} \mathrm{d}r \\
&\quad - \Big\langle \Delta P^\eta_r(\lambda,\mu) + (b_x(\bar{x}_r(\lambda),\bar{u}_r) + b_x(\bar{x}_r(\mu),\bar{u}_r))P^\eta_r(\lambda,\mu) \\
&\qquad + \langle \sigma_x(\bar{x}_r(\lambda),\bar{u}_r), \sigma_x(\bar{x}_r(\mu),\bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})} P^\eta_r(\lambda,\mu) \\
&\qquad + \langle \sigma_x(\bar{x}_r(\lambda),\bar{u}_r) + \sigma_x(\bar{x}_r(\mu),\bar{u}_r), Q^\eta_r(\lambda,\mu)\rangle_{L_2(\Xi,\mathbb{R})} \\
&\qquad + \delta^*(l_{xx}(\bar{x}_r,\bar{u}_r)) + \delta^*(b_{xx}(\bar{x}_r,\bar{u}_r)p_r(\lambda)) \\
&\qquad + \delta^*(\langle \sigma_{xx}(\bar{x}_r,\bar{u}_r), q_r(\lambda)\rangle_{L_2(\Xi,\mathbb{R})}), Y_r(\lambda,\mu)\Big\rangle_{H^{-1}(\Lambda^2)\times H^1_0(\Lambda^2)} \mathrm{d}r \\
&\quad + \langle Q^\eta_r(\lambda,\mu), (\sigma_x(\bar{x}_r(\lambda),\bar{u}_r) + \sigma_x(\bar{x}_r(\mu),\bar{u}_r))Y_r(\lambda,\mu) + \Psi_r(\lambda,\mu)\rangle_{L_2(\Xi,L^2(\Lambda^2))}\mathrm{d}r \\
&\quad + \langle ((\sigma_x(\bar{x}_r(\lambda),\bar{u}_r) + \sigma_x(\bar{x}_r(\mu),\bar{u}_r))Y_r(\lambda,\mu) + \Psi_r(\lambda,\mu))^* Q^\eta_r(\lambda,\mu), \mathrm{d}W_r\rangle_{L^2(\Lambda^2)} \\
&\quad + \langle Q^\eta_r(\lambda,\mu)^*((\sigma_x(\bar{x}_r(\lambda),\bar{u}_r) + \sigma_x(\bar{x}_r(\mu),\bar{u}_r))Y_r(\lambda,\mu) + \Psi_r(\lambda,\mu)), \mathrm{d}W_r\rangle_{L^2(\Lambda^2)}.
\end{aligned}
\tag{2.66}
$$

Integrating the Laplacian by parts yields

$$
\begin{aligned}
&\mathrm{d}\langle P^\eta_r(\lambda,\mu), Y_r(\lambda,\mu)\rangle_{L^2(\Lambda^2)} \\
&= \Big[\langle P^\eta_r(\lambda,\mu), \Phi_r(\lambda,\mu)\rangle_{L^2(\Lambda^2)} + \langle Q^\eta_r(\lambda,\mu), \Psi_r(\lambda,\mu)\rangle_{L_2(\Xi,L^2(\Lambda^2))} \\
&\qquad - \langle \delta^*(l_{xx}(\bar{x}_r,\bar{u}_r)) + \delta^*(b_{xx}(\bar{x}_r,\bar{u}_r)p_r(\lambda)) \\
&\qquad + \delta^*(\langle \sigma_{xx}(\bar{x}_r,\bar{u}_r), q_r(\lambda)\rangle_{L_2(\Xi,\mathbb{R})}), Y_r(\lambda,\mu)\rangle_{H^{-1}(\Lambda^2)\times H^1_0(\Lambda^2)}\Big]\mathrm{d}r \\
&\quad + \langle ((\sigma_x(\bar{x}_r(\lambda),\bar{u}_r) + \sigma_x(\bar{x}_r(\mu),\bar{u}_r))Y_r(\lambda,\mu) + \Psi_r(\lambda,\mu))^* Q^\eta_r(\lambda,\mu), \mathrm{d}W_r\rangle_{L^2(\Lambda^2)} \\
&\quad + \langle Q^\eta_r(\lambda,\mu)^*((\sigma_x(\bar{x}_r(\lambda),\bar{u}_r) + \sigma_x(\bar{x}_r(\mu),\bar{u}_r))Y_r(\lambda,\mu) + \Psi_r(\lambda,\mu)), \mathrm{d}W_r\rangle_{L^2(\Lambda^2)}.
\end{aligned}
\tag{2.67}
$$

Therefore, taking expectations and considering the initial and terminal condition for Y and P^η, respectively, we obtain

$$
\begin{aligned}
&\mathbb{E}\left[\langle h^\eta_{xx}, Y_T\rangle_{L^2(\Lambda^2)}\right] \\
&= \mathbb{E}\Big[\int_s^T \langle P^\eta_r, \Phi_r\rangle_{L^2(\Lambda^2)} + \langle Q^\eta_r, \Psi_r\rangle_{L^2(\Xi,L^2(\Lambda^2))} \\
&\qquad - \langle \delta^*(l_{xx}(\bar{x}_r,\bar{u}_r)) + \delta^*(b_{xx}(\bar{x}_r,\bar{u}_r)p_r(\lambda)) \\
&\qquad + \delta^*(\langle \sigma_{xx}(\bar{x}_r,\bar{u}_r), q_r(\lambda)\rangle_{L_2(\Xi,\mathbb{R})}), Y_r(\lambda,\mu)\rangle_{H^{-1}(\Lambda^2)\times H^1_0(\Lambda^2)}\mathrm{d}r\Big],
\end{aligned}
\tag{2.68}
$$

which is the mollified second order adjoint state property (2.52). Hence, the mollified second order adjoint state is characterized by equation (2.59).

2.5. Passing to the Limit of the Mollified Second Order Adjoint State

In this section, we derive an equation for the second order adjoint state $P = \lim_{\eta\to 0} P^\eta$. Recall that we chose h^η_{xx} in such a way, that

$$\lim_{\eta\to 0} h^\eta_{xx} = \delta^*(h_{xx}(\bar{x}_T(\lambda))) \quad \text{in } H^{-1}(\Lambda^2). \tag{2.69}$$

Theorem 2.12. *The equation*

$$\begin{cases} \mathrm{d}P_r(\lambda,\mu) = -[\Delta P_r(\lambda,\mu) + (b_x(\bar{x}_r(\lambda),\bar{u}_r) + b_x(\bar{x}_r(\mu),\bar{u}_r))P_r(\lambda,\mu) \\ \qquad\qquad\qquad +\langle \sigma_x(\bar{x}_r(\lambda),\bar{u}_r), \sigma_x(\bar{x}_r(\mu),\bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})} P_r(\lambda,\mu) \\ \qquad\qquad\qquad +\langle \sigma_x(\bar{x}_r(\lambda),\bar{u}_r) + \sigma_x(\bar{x}_r(\mu),\bar{u}_r), Q_r(\lambda,\mu)\rangle_{L_2(\Xi,\mathbb{R})} \\ \qquad\qquad\qquad +\delta^*(l_{xx}(\bar{x}_r(\lambda),\bar{u}_r)) + \delta^*(b_{xx}(\bar{x}_r(\lambda),\bar{u}_r)p_r(\lambda)) \\ \qquad\qquad\qquad +\delta^*(\langle \sigma_{xx}(\bar{x}_r(\lambda),\bar{u}_r), q_r\rangle_{L_2(\Xi,\mathbb{R})})]\mathrm{d}r + Q_r(\lambda,\mu)\mathrm{d}W_r \\ P_T(\lambda,\mu) = \delta^*(h_{xx}(\bar{x}_T(\lambda))) \end{cases} \tag{2.70}$$

has a unique adapted solution (P, Q), *where*

$$P \in L^2([s,T]\times\Omega; L^2(\Lambda^2)) \cap L^2(\Omega; C([s,T]; H^{-1}(\Lambda^2))), \tag{2.71}$$

and

$$Q \in L^2([s,T]\times\Omega; L_2(\Xi; H^{-1}(\Lambda^2))). \tag{2.72}$$

Here equation (2.70) *holds in* $L^2([s,T]\times\Omega; H^{-2}(\Lambda^2))$.

Proof. First, we prove existence of a solution. Let (P^η, Q^η) denote the solution of equation (2.59). We define $F : H^{-1}(\Lambda^2) \to \mathbb{R}$, $x \mapsto \|x\|^2_{H^{-1}(\Lambda^2)}$. Since P^η is an $H^{-1}(\Lambda^2)$-valued semimartingale, we can apply the classical version of Itô's formula for

Hilbert space-valued semimartingales (see [DPZ14, Section 4.4]), which yields

$$
\begin{aligned}
&\|P_r^\eta(\lambda,\mu)\|_{H^{-1}(\Lambda^2)}^2 \\
&= \|h_{xx}^\eta\|_{H^{-1}(\Lambda^2)}^2 + 2\int_r^T \langle \Delta P_\theta^\eta(\lambda,\mu), P_\theta^\eta(\lambda,\mu)\rangle_{H^{-1}(\Lambda^2)}\mathrm{d}\theta \\
&\quad + 2\int_r^T \langle (b_x(\bar{x}_\theta(\lambda),\bar{u}_\theta) + b_x(\bar{x}_\theta(\mu),\bar{u}_\theta))P_\theta^\eta(\lambda,\mu), P_\theta^\eta(\lambda,\mu)\rangle_{H^{-1}(\Lambda^2)}\mathrm{d}\theta \\
&\quad + 2\int_r^T \langle \langle \sigma_x(\bar{x}_\theta(\lambda),\bar{u}_\theta), \sigma_x(\bar{x}_\theta(\mu),\bar{u}_\theta)\rangle_{L_2(\Xi,\mathbb{R})} P_\theta^\eta(\lambda,\mu), P_\theta^\eta(\lambda,\mu)\rangle_{H^{-1}(\Lambda^2)}\mathrm{d}\theta \\
&\quad + 2\int_r^T \langle \langle \sigma_x(\bar{x}_\theta(\lambda),\bar{u}_\theta) + \sigma_x(\bar{x}_\theta(\mu),\bar{u}_\theta), Q_\theta^\eta(\lambda,\mu)\rangle_{L_2(\Xi,\mathbb{R})}, P_\theta^\eta(\lambda,\mu)\rangle_{H^{-1}(\Lambda^2)}\mathrm{d}\theta \\
&\quad + 2\int_r^T \langle \delta^*(l_{xx}(\bar{x}_\theta(\lambda),\bar{u}_\theta)) + \delta^*(b_{xx}(\bar{x}_\theta(\lambda),\bar{u}_\theta)p_\theta(\lambda)) \\
&\qquad\qquad + \delta^*(\langle \sigma_{xx}(\bar{x}_\theta(\lambda),\bar{u}_\theta), q_\theta\rangle_{L_2(\Xi,\mathbb{R})}), P_\theta^\eta(\lambda,\mu)\rangle_{H^{-1}(\Lambda^2)}\mathrm{d}\theta \\
&\quad - \int_r^T \|Q_\theta^\eta\|_{L_2(\Xi,H^{-1}(\Lambda^2))}^2\mathrm{d}\theta + 2\int_r^T \langle P_\theta^\eta(\lambda,\mu), Q_\theta^\eta(\lambda,\mu)\mathrm{d}W_\theta\rangle_{H^{-1}(\Lambda^2)}.
\end{aligned}
\tag{2.73}
$$

By [LR15, Lemma 4.1.12], we have

$$
\langle \Delta P_\theta^\eta(\lambda,\mu), P_\theta^\eta(\lambda,\mu)\rangle_{H^{-1}(\Lambda^2)} = -\|P_\theta^\eta(\lambda,\mu)\|_{L^2(\Lambda^2)}^2. \tag{2.74}
$$

Therefore, from equation (2.73) we derive

$$
\begin{aligned}
&\|P_r^\eta(\lambda,\mu)\|_{H^{-1}(\Lambda^2)}^2 + 2\int_r^T \|P_\theta^\eta(\lambda,\mu)\|_{L^2(\Lambda^2)}^2\mathrm{d}\theta + \int_r^T \|Q_\theta^\eta\|_{L_2(\Xi,H^{-1}(\Lambda^2))}^2\mathrm{d}\theta \\
&\le \|h_{xx}^\eta\|_{H^{-1}(\Lambda^2)}^2 + C(b,\sigma,T,l)\left(1 + \int_r^T \|P_\theta^\eta(\lambda,\mu)\|_{H^{-1}(\Lambda^2)}^2\mathrm{d}\theta\right) \\
&\quad + 2\int_r^T \langle P_\theta^\eta(\lambda,\mu), Q_\theta^\eta(\lambda,\mu)\mathrm{d}W_\theta\rangle_{H^{-1}(\Lambda^2)}.
\end{aligned}
\tag{2.75}
$$

Taking the supremum and expectations, using Burkholder-Davis-Gundy inequality for the stochastic integral, and applying Grönwall's inequality, we obtain

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{r\in[s,T]} \|P_r^\eta(\lambda,\mu)\|_{H^{-1}(\Lambda^2)}^2 + 2\int_s^T \|P_r^\eta(\lambda,\mu)\|_{L^2(\Lambda^2)}^2\mathrm{d}r + \int_s^T \|Q_r^\eta\|_{L_2(\Xi,H^{-1}(\Lambda^2))}^2\mathrm{d}r\right] \\
&\le C\left(1 + \mathbb{E}\left[\|h_{xx}^\eta\|_{H^{-1}(\Lambda^2)}^2\right]\right),
\end{aligned}
\tag{2.76}
$$

where the right-hand side is uniformly bounded in η. Therefore, we can extract weakly convergent subsequences

$$P^\eta \rightharpoonup P \quad \text{in } L^2([s,T] \times \Omega; L^2(\Lambda^2)), \tag{2.77}$$

$$Q^\eta \rightharpoonup Q \quad \text{in } L^2([s,T] \times \Omega; L_2(\Xi, H^{-1}(\Lambda^2))), \tag{2.78}$$

which implies

$$\int_\cdot^T Q_r^\eta \mathrm{d}W_r \overset{*}{\rightharpoonup} \int_\cdot^T Q_r \mathrm{d}W_r \quad \text{in } L^\infty([s,T]; L^2(\Omega; H^{-1}(\Lambda^2))). \tag{2.79}$$

Since $\Delta : L^2(\Lambda^2) \to H^{-2}(\Lambda^2)$ is weak-weak continuous, we can test the mollified second order adjoint equation (2.59) with a test function in $H_0^2(\Lambda^2)$ and pass to the limit $\eta \to 0$, which concludes the proof of existence. The continuity of P as a process with values in $H^{-1}(\Lambda^2)$ follows from [LR15, Theorem 4.2.5]. In order to prove uniqueness, we observe that, by the linearity of the equation, the difference of two solutions satisfies the corresponding equation with vanishing inhomogeneity and terminal condition. Hence, by an analogous argument as the one for the a priori bound (2.76), the two solutions must coincide. □

Remark 2.13. In case the state equation (2.1) is governed by the more general uniformly elliptic differential operator A given in equation (2.13), the Laplacian in equation (2.70) is replaced by the operator $\bar{A} : H_0^1(\Lambda^2) \to H^{-1}(\Lambda^2)$,

$$\bar{A}x(\lambda, \mu) := (\partial_\lambda(a(\lambda)\partial_\lambda x) + \partial_\mu(a(\mu)\partial_\mu x))(\lambda, \mu). \tag{2.80}$$

Therefore, we have to consider the functional

$$F : H^{-1}(\Lambda^2) \to \mathbb{R} \tag{2.81}$$

$$x \mapsto \|x\|^2_{\mathcal{D}((-\bar{A})^{-\frac{1}{2}})} = \|(I - \bar{A})^{-\frac{1}{2}} x\|^2_{L^2(\Lambda^2)}. \tag{2.82}$$

In this case, the term

$$\langle \Delta P_\theta^\eta(\lambda, \mu), P_\theta^\eta(\lambda, \mu)\rangle_{H^{-1}(\Lambda^2)} \tag{2.83}$$

is replaced by

$$\langle \bar{A} P_\theta^\eta(\lambda, \mu), P_\theta^\eta(\lambda, \mu)\rangle_{\mathcal{D}((-\bar{A})^{-\frac{1}{2}})} = -\|P_\theta^\eta\|^2_{L^2(\Lambda^2)} + \|P_\theta^\eta\|^2_{\mathcal{D}((-\bar{A})^{-\frac{1}{2}})}. \tag{2.84}$$

Now, using the same arguments as in the preceding proof, we can generalize the previous result mutatis mutandis to the case of uniformly elliptic differential operators.

The following property of the second order adjoint state is not used hereafter, but is of independent interest.

Proposition 2.14. *It holds*

$$\begin{aligned}
&\mathbb{E}\left[\int_s^T \int_\Lambda \left(b_{xx}(\bar{x}_r(\lambda),\bar{u}_r)p_r(\lambda) + \langle \sigma_{xx}(\bar{x}_r(\lambda),\bar{u}_r), q_r(\lambda)\rangle_{L_2(\Xi,\mathbb{R})}\right) y_r^\varepsilon(\lambda) y_r^\varepsilon(\lambda)\mathrm{d}\lambda\mathrm{d}r\right]\\
&+\mathbb{E}\left[\int_s^T \int_\Lambda l_{xx}(\bar{x}_r(\lambda),\bar{u}_r) y_r^\varepsilon(\lambda) y_r^\varepsilon(\lambda)\mathrm{d}\lambda\mathrm{d}r + \int_\Lambda h_{xx}(\bar{x}_T(\lambda)) y_T^\varepsilon(\lambda) y_T^\varepsilon(\lambda)\mathrm{d}\lambda\right]\\
&=\mathbb{E}\left[\int_s^T \langle P_r, \Phi_r^\varepsilon\rangle_{L^2(\Lambda^2)} + \langle Q_r, \Psi_r^\varepsilon\rangle_{L_2(\Xi,H^{-1}(\Lambda^2))\times L_2(\Xi,H_0^1(\Lambda^2))}\mathrm{d}r\right],
\end{aligned} \tag{2.85}$$

where y^ε is the solution of equation (2.15) *and Φ^ε and Ψ^ε are given by equations* (2.37) *and* (2.38)*, respectively.*

Proof. Since for $\eta \to 0$,

$$\mathbb{E}\left[\int_{\Lambda^2} h_{xx}^\eta(\lambda,\mu) Y_T^\varepsilon(\lambda,\mu)\mathrm{d}\lambda\mathrm{d}\mu\right] \to \mathbb{E}\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda)) y_T^\varepsilon(\lambda) y_T^\varepsilon(\lambda)\mathrm{d}\lambda\right], \tag{2.86}$$

taking the limit $\eta \to 0$ in equation (2.52) yields the claim. □

2.6. Peng's Maximum Principle

In order to prove the stochastic maximum principle, we have to take the limits $\eta \to 0$ and $\varepsilon \to 0$ in the variational inequality (2.53). If we take the limit $\eta \to 0$ first, we eliminate the terms involving the terminal condition. However, the remaining term

$$\mathbb{E}\left[\int_s^T \langle P_r, \Phi_r^\varepsilon\rangle_{L^2(\Lambda^2)} + \langle Q_r, \Psi_r^\varepsilon\rangle_{L_2(\Xi,H^{-1}(\Lambda^2))\times L_2(\Xi,H_0^1(\Lambda^2))}\mathrm{d}r\right] \tag{2.87}$$

does not have the asymptotic needed in (2.107). Indeed, the lacking regularity of Q requires us to control Ψ^ε in $L_2(\Xi, H_0^1(\Lambda^2))$. Since we can't expect such a control in general, we have to interchange the limits in η and ε. In order to ensure convergence in the converse order, we need compactness of $y_T^\varepsilon/\sqrt{\varepsilon}$, $\varepsilon > 0$, in $L^2(\Lambda)$.

Lemma 2.15. *For $\gamma \in (0, 1/2)$ and $\varepsilon \in (0, T-t)$, it holds*

$$\mathbb{E}\left[\|y_T^\varepsilon\|^2_{H_0^\gamma(\Lambda)}\right] \le C\varepsilon. \tag{2.88}$$

Proof. Set $\tilde{y}_r^\varepsilon := y_r^\varepsilon/\sqrt{\varepsilon}$, $r \in [s,T]$. Then $\tilde{y}^\varepsilon$ satisfies the equation

$$\begin{cases}
\mathrm{d}\tilde{y}_r^\varepsilon = \left[\Delta\tilde{y}_r^\varepsilon + b_x(\bar{x}_r,\bar{u}_r)\tilde{y}_r^\varepsilon + \frac{1}{\varepsilon}\left(b(\bar{x}_r,u_r^\varepsilon) - b(\bar{x}_r,\bar{u}_r)\right)\right]\mathrm{d}r\\
\qquad\qquad + \left[\sigma_x(\bar{x}_r,\bar{u}_r)\tilde{y}_r^\varepsilon + \frac{1}{\varepsilon}\left(\sigma(\bar{x}_r,u_r^\varepsilon) - \sigma(\bar{x}_r,\bar{u}_r)\right)\right]\mathrm{d}W_r\\
\tilde{y}_s^\varepsilon = 0.
\end{cases} \tag{2.89}$$

Duhamel's formula for mild solutions yields

$$\begin{aligned}\tilde{y}_T^\varepsilon = &\int_s^T S_{T-r}\left(b_x(\bar{x}_r,\bar{u}_r)\tilde{y}_r^\varepsilon\right)\mathrm{d}r + \frac{1}{\varepsilon}\int_t^{t+\varepsilon} S_{T-r}\left(b(\bar{x}_r,v)-b(\bar{x}_r,\bar{u}_r)\right)\mathrm{d}r \\ &+ \int_s^T S_{T-r}\left(\sigma_x(\bar{x}_r,\bar{u}_r)\tilde{y}_r^\varepsilon\right)\mathrm{d}W_r + \frac{1}{\varepsilon}\int_t^{t+\varepsilon} S_{T-r}\left(\sigma(\bar{x}_r,v)-\sigma(\bar{x}_r,\bar{u}_r)\right)\mathrm{d}W_r,\end{aligned} \tag{2.90}$$

where $(S_r)_{r\geq 0}$ denotes the heat semigroup. Notice that the variational solution and the mild solution coincide, see [Hai09, Proposition 5.7]. By analyticity, we have the bound

$$\|S_r f\|^2_{H_0^\gamma(\Lambda)} \leq \frac{C}{r^{2\gamma}}\|f\|^2_{L^2(\Lambda)} \tag{2.91}$$

for any $f \in L^2(\Lambda)$, see [Paz83, Chapter 2, Lemma 6.13]. Using this property and the boundedness of b_x, we can estimate

$$\mathbb{E}\left[\left\|\int_s^T S_{T-r}\left(b_x(\bar{x}_r,\bar{u}_r)\tilde{y}_r^\varepsilon\right)\mathrm{d}r\right\|^2_{H_0^\gamma(\Lambda)}\right] \leq \sup_{r\in[s,T]} \mathbb{E}\left[\|\tilde{y}_r^\varepsilon\|^2_{L^2(\Lambda)}\right]\int_s^T \frac{C}{(T-r)^{2\gamma}}\mathrm{d}r < \infty. \tag{2.92}$$

Furthermore, for the second integral in (2.90), we obtain

$$\begin{aligned}&\mathbb{E}\left[\left\|\frac{1}{\varepsilon}\int_t^{t+\varepsilon} S_{T-r}\left(b(\bar{x}_r,v)-b(\bar{x}_r,\bar{u}_r)\right)\mathrm{d}r\right\|^2_{H_0^\gamma(\Lambda)}\right] \\ &\leq \frac{C}{\varepsilon}\mathbb{E}\left[\int_t^{t+\varepsilon}\frac{1}{(T-r)^{2\gamma}}\|b(\bar{x}_r,v)-b(\bar{x}_r,\bar{u}_r)\|^2_{L^2(\Lambda)}\,\mathrm{d}r\right]\end{aligned} \tag{2.93}$$

Since $t < T$, using the bounds on b we obtain

$$\begin{aligned}&\frac{C}{\varepsilon}\mathbb{E}\left[\int_t^{t+\varepsilon}\frac{1}{(T-r)^{2\gamma}}\|b(\bar{x}_r,v)-b(\bar{x}_r,\bar{u}_r)\|^2_{L^2(\Lambda)}\,\mathrm{d}r\right] \\ &\leq \frac{C}{\varepsilon}\int_t^{t+\varepsilon}\mathbb{E}\left[1+\|\bar{x}_r\|^2_{L^2(\Lambda)}+\|v\|^2_{\mathcal{U}}+\|\bar{u}_r\|^2_{\mathcal{U}}\right]\mathrm{d}r \\ &\leq C\left(1+\sup_{r\in[s,T]}\mathbb{E}\left[\|\bar{x}_r\|^2_{L^2(\Lambda)}\right]+\|v\|^2_{\mathcal{U}}+\sup_{r\in[s,T]}\mathbb{E}\left[\|\bar{u}_r\|^2_{\mathcal{U}}\right]\right) < \infty.\end{aligned} \tag{2.94}$$

Now we consider the first stochastic integral in (2.90). We have

$$\begin{aligned}
&\mathbb{E}\left[\left\|\int_s^T S_{T-r}\left(\sigma_x(\bar{x}_r,\bar{u}_r)\tilde{y}_r^\varepsilon\right)\mathrm{d}W_r\right\|^2_{H_0^\gamma(\Lambda)}\right]\\
&=\mathbb{E}\left[\int_s^T \|S_{T-r}(\sigma_x(\bar{x}_r,\bar{u}_r)\tilde{y}_r^\varepsilon)\|^2_{L_2(\Xi,H_0^\gamma(\Lambda))}\,\mathrm{d}r\right]\\
&\le\mathbb{E}\left[\int_s^T \|S_{T-r}\|^2_{L(L^2(\Lambda),H_0^\gamma(\Lambda))}\|\sigma_x(\bar{x}_r,\bar{u}_r)\|^2_{L(L^2(\Lambda),L_2(\Xi,L^2(\Lambda)))}\|\tilde{y}_r^\varepsilon\|^2_{L^2(\Lambda)}\mathrm{d}r\right],
\end{aligned}\tag{2.95}$$

which can be controlled by the same arguments as for the corresponding term with b, since σ_x is bounded as well. Finally, for the second stochastic integral in (2.90), we have

$$\begin{aligned}
&\mathbb{E}\left[\left\|\frac{1}{\varepsilon}\int_t^{t+\varepsilon} S_{T-r}\left(\sigma(\bar{x}_r,v)-\sigma(\bar{x}_r,\bar{u}_r)\right)\mathrm{d}W_r\right\|^2_{H_0^\gamma(\Lambda)}\right]\\
&=\mathbb{E}\left[\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\|S_{T-r}\left(\sigma(\bar{x}_r,v)-\sigma(\bar{x}_r,\bar{u}_r)\right)\|^2_{L_2(\Xi,H_0^\gamma(\Lambda))}\,\mathrm{d}r\right]
\end{aligned}\tag{2.96}$$

which is again finite by the same arguments as for the corresponding term with b. ☐

Now we are able to prove the main result of this chapter.

Theorem 2.16 (Peng's Maximum Principle)**.** *Let $(\bar{x},\bar{u})$ be an optimal pair of the control problem* (2.2) *and* (2.1)*. Then there exist adapted processes (p,q), where*

$$p\in L^2([s,T]\times\Omega;H_0^1(\Lambda))\cap L^2(\Omega;C([s,T];L^2(\Lambda)))\tag{2.97}$$

and

$$q\in L^2([s,T]\times\Omega;L_2(\Xi,L^2(\Lambda))),\tag{2.98}$$

satisfying the first order adjoint equation

$$\begin{cases}\mathrm{d}p_r=-\left[\Delta p_r+b_x(\bar{x}_r,\bar{u}_r)p_r+\langle\sigma_x(\bar{x}_r,\bar{u}_r),q_r\rangle_{L_2(\Xi,\mathbb{R})}+l_x(\bar{x}_r,\bar{u}_r)\right]\mathrm{d}r+q_r\mathrm{d}W_r\\ p_T=h_x(\bar{x}_T),\end{cases}\tag{2.99}$$

and adapted processes (P,Q), where

$$P\in L^2([s,T]\times\Omega;L^2(\Lambda^2))\cap L^2(\Omega;C([s,T];H^{-1}(\Lambda)))\tag{2.100}$$

and

$$Q\in L^2([s,T]\times\Omega;L_2(\Xi,H^{-1}(\Lambda^2))),\tag{2.101}$$

satisfying the second order adjoint equation

$$\begin{cases} \mathrm{d}P_r(\lambda,\mu) = -[\Delta P_r(\lambda,\mu) + (b_x(\bar{x}_r(\lambda),\bar{u}_r) + b_x(\bar{x}_r(\mu),\bar{u}_r))P_r(\lambda,\mu) \\ \qquad\qquad\qquad + \langle \sigma_x(\bar{x}_r(\lambda),\bar{u}_r), \sigma_x(\bar{x}_r(\mu),\bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})} P_r(\lambda,\mu) \\ \qquad\qquad\qquad + \langle \sigma_x(\bar{x}_r(\lambda),\bar{u}_r) + \sigma_x(\bar{x}_r(\mu),\bar{u}_r), Q_r(\lambda,\mu)\rangle_{L_2(\Xi,\mathbb{R})} \\ \qquad\qquad\qquad + \delta^*(l_{xx}(\bar{x}_r(\lambda),\bar{u}_r)) + \delta^*(b_{xx}(\bar{x}_r(\lambda),\bar{u}_r)p_r(\lambda)) \\ \qquad\qquad\qquad + \delta^*(\langle \sigma_{xx}(\bar{x}_r(\lambda),\bar{u}_r), q_r\rangle_{L_2(\Xi,\mathbb{R})})]\mathrm{d}r + Q_r(\lambda,\mu)\mathrm{d}W_r \\ P_T(\lambda,\mu) = \delta^*(h_{xx}(\bar{x}_T(\lambda))), \end{cases} \tag{2.102}$$

such that

$$\mathcal{G}(t,\bar{x}_t,v) \geq \mathcal{G}(t,\bar{x}_t,\bar{u}_t) \tag{2.103}$$

for all $v \in U$, and almost all $(t,\omega) \in [s,T] \times \Omega$. Here we denote by $\mathcal{G}$ the generalized Hamiltonian, i.e., $\mathcal{G} : [s,T] \times L^2(\Lambda) \times U \to \mathbb{R}$,

$$\mathcal{G}(t,x,u) := \int_\Lambda l(x(\lambda),u)\mathrm{d}\lambda + \langle p_t, b(x,u)\rangle_{L^2(\Lambda)} + \frac{1}{2} tr(\sigma(x,u)^* P_t \sigma(x,u)) \\ + tr(\sigma(x,u)^*[q_t - P_t\sigma(\bar{x}_t,\bar{u}_t)]). \tag{2.104}$$

Remark 2.17. Notice that the generalized Hamiltonian $\mathcal{G}$ consists of the Hamiltonian $\mathcal{H}$ defined in (1.8) and the correction term

$$\mathrm{tr}(\sigma(x,u)^*[q_t - P_t\sigma(\bar{x}_t,\bar{u}_t)]). \tag{2.105}$$

Below, in Example 2.18, we give a simple example in which this correction term does not vanish.

Proof. Inequality (2.53) states

$$\mathbb{E}\left[\int_s^T \langle b(\bar{x}_r,u_r^\varepsilon) - b(\bar{x}_r,\bar{u}_r), p_r\rangle_{L^2(\Lambda)} + \langle \sigma(\bar{x}_r,u_r^\varepsilon) - \sigma(\bar{x}_r,\bar{u}_r), q_r\rangle_{L_2(\Xi,L^2(\Lambda))}\,\mathrm{d}r\right] \\ + \frac{1}{2}\mathbb{E}\left[\int_s^T \langle P_r^\eta, \Phi_r^\varepsilon\rangle_{L^2(\Lambda^2)} + \langle Q_r^\eta, \Psi_r^\varepsilon\rangle_{L_2(\Xi,L^2(\Lambda^2))}\mathrm{d}r\right] \\ + \mathbb{E}\left[\int_s^T \int_\Lambda l(\bar{x}_r(\lambda),u_r^\varepsilon) - l(\bar{x}_r(\lambda),\bar{u}_r)\mathrm{d}\lambda\mathrm{d}r\right] \\ + \frac{1}{2}\mathbb{E}\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda))y_T^\varepsilon(\lambda)y_T^\varepsilon(\lambda)\mathrm{d}\lambda - \int_{\Lambda^2} h_{xx}^\eta(\lambda,\mu)Y_T^\varepsilon(\lambda,\mu)\mathrm{d}\lambda\mathrm{d}\mu\right] \geq o(\varepsilon). \tag{2.106}$$

Localizing by dividing by ε and taking the limit $\varepsilon \to 0$ yields

$$
\begin{aligned}
&\frac{1}{\varepsilon}\mathbb{E}\left[\int_s^T \langle b(\bar{x}_r, u_r^\varepsilon) - b(\bar{x}_r, \bar{u}_r), p_r\rangle_{L^2(\Lambda)} + \langle \sigma(\bar{x}_r, u_r^\varepsilon) - \sigma(\bar{x}_r, \bar{u}_r), q_r\rangle_{L_2(\Xi, L^2(\Lambda))}\,\mathrm{d}r\right] \\
&+\frac{1}{2\varepsilon}\mathbb{E}\left[\int_s^T \langle P_r^\eta, \Phi_r^\varepsilon\rangle_{L^2(\Lambda^2)} + \langle Q_r^\eta, \Psi_r^\varepsilon\rangle_{L_2(\Xi, L^2(\Lambda^2))}\mathrm{d}r\right] \\
&+\frac{1}{\varepsilon}\mathbb{E}\left[\int_s^T\int_\Lambda l(\bar{x}_r(\lambda), u_r^\varepsilon) - l(\bar{x}_r(\lambda), \bar{u}_r)\mathrm{d}\lambda\mathrm{d}r\right] \\
&\to \langle b(\bar{x}_t, v) - b(\bar{x}_t, \bar{u}_t), p_t\rangle_{L^2(\Lambda)} + \langle \sigma(\bar{x}_t, v) - \sigma(\bar{x}_t, \bar{u}_t), q_t\rangle_{L_2(\Xi, L^2(\Lambda))} \\
&\quad + \int_\Lambda l(\bar{x}_t(\lambda), v) - l(\bar{x}_t(\lambda), \bar{u}_t)\mathrm{d}\lambda \\
&\quad + \frac{1}{2}\Big\langle P_t^\eta(\lambda,\mu), \langle \sigma(\bar{x}_t(\lambda), v) - \sigma(\bar{x}_t(\lambda), \bar{u}_t), \sigma(\bar{x}_t(\mu), v) - \sigma(\bar{x}_t(\mu), \bar{u}_t)\rangle_{L_2(\Xi,\mathbb{R})}\Big\rangle_{L^2(\Lambda^2)}.
\end{aligned} \tag{2.107}
$$

Notice that all but the remaining term in

$$
\mathbb{E}\left[\int_s^T \langle P_r^\eta, \Phi_r^\varepsilon\rangle_{L^2(\Lambda^2)} + \langle Q_r^\eta, \Psi_r^\varepsilon\rangle_{L_2(\Xi, L^2(\Lambda^2))}\mathrm{d}r\right] \tag{2.108}
$$

are of order $o(\varepsilon)$. It remains to prove that

$$
\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathbb{E}\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda))y_T^\varepsilon(\lambda)y_T^\varepsilon(\lambda)\mathrm{d}\lambda - \int_{\Lambda^2} h_{xx}^\eta(\lambda,\mu)Y_T^\varepsilon(\lambda,\mu)\mathrm{d}\lambda\mathrm{d}\mu\right] \tag{2.109}
$$

vanishes in the limit $\eta \to 0$. Using Lemma 2.15 and the compact embedding $H_0^\gamma(\Lambda) \subset\subset L^2(\Lambda)$, $\gamma \in (0, 1/2)$ (see e.g. [DD12, Theorem 4.54]), we can extract a subsequence of $y_T^\varepsilon/\sqrt{\varepsilon}$ converging in $L^2(\Lambda)$ to some $\tilde{y}_T \in L^2(\Lambda)$. Therefore

$$
\begin{aligned}
&\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathbb{E}\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda))y_T^\varepsilon(\lambda)y_T^\varepsilon(\lambda)\mathrm{d}\lambda - \int_{\Lambda^2} h_{xx}^\eta(\lambda,\mu)Y_T^\varepsilon(\lambda,\mu)\mathrm{d}\lambda\mathrm{d}\mu\right] \\
&= \mathbb{E}\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda))\tilde{y}_T(\lambda)\tilde{y}_T(\lambda)\mathrm{d}\lambda - \int_{\Lambda^2} h_{xx}^\eta(\lambda,\mu)\tilde{y}_T(\lambda)\tilde{y}_T(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right],
\end{aligned} \tag{2.110}
$$

which vanishes in the limit $\eta \to 0$. This concludes the proof. □

Example 2.18. Let $(W_t)_{t\in[0,T]}$ be a one-dimensional Brownian motion. Consider the scalar-valued controlled state equation

$$
\begin{cases}
\mathrm{d}x_t^u = u_t\mathrm{d}t + u_t\mathrm{d}W_t, \quad t \in [0,T] \\
x_0^u = x \in \mathbb{R}
\end{cases} \tag{2.111}
$$

and the cost functional

$$
J(u) = \mathbb{E}\left[\frac{1}{2}(x_T^u)^2\right]. \tag{2.112}
$$

In this case, the adjoint equations read

$$\begin{cases} \mathrm{d}p_t = q_t \mathrm{d}W_t, & t \in [0,T] \\ p_T = \bar{x}_T \end{cases} \tag{2.113}$$

and

$$\begin{cases} \mathrm{d}P_t = Q_t \mathrm{d}W_t, & t \in [0,T] \\ P_T = 1. \end{cases} \tag{2.114}$$

Therefore, the second order adjoint state is given by $(P_t, Q_t) = (1, 0)$.

Using the theory for linear quadratic control problem (see [YZ99, Chapter 6]), we calculate the optimal feedback control as

$$\bar{u}_t = -\bar{x}_t. \tag{2.115}$$

Making the ansatz $p_t = \theta(t)\bar{x}_t$ for some $\theta \in C^1([0,T])$ and applying Itô's rule yields

$$\mathrm{d}p_t = (\dot{\theta}(t) - \theta(t))\bar{x}_t \mathrm{d}t - \theta(t)\bar{x}_t \mathrm{d}W_t. \tag{2.116}$$

Equation (2.113) yields

$$\begin{cases} \dot{\theta}(t) = \theta(t), & t \in [0,T] \\ \theta(T) = 1, \end{cases} \tag{2.117}$$

thus $\theta(t) = \exp(t - T)$. Therefore, the first order adjoint state is given by $p_t = \exp(t-T)\bar{x}_t$ and $q_t = -\exp(t-T)\bar{x}_t$.

Altogether, we have for $x \neq 0$

$$q_t = -\exp(t-T)\bar{x}_t \neq -\bar{x}_t = \bar{u}_t = P_t \sigma(\bar{x}_t, \bar{u}_t). \tag{2.118}$$

3. Viscosity Differentials of the Value Function

In this chapter, we derive additional necessary optimality conditions for controlled semilinear SPDEs with Nemytskii-type coefficients. In particular, we relate the adjoint states to the viscosity differentials of the value function evaluated along an optimal trajectory. This chapter is based on [SW22].

3.1. Introduction

In this chapter, we work in the same framework as in Chapter 2. However, Assumption 2.1 has to be slightly modified and Assumption 2.2 has to be expanded.

We need the dynamic programming principle, which relies on the weak formulation of the control problem. Therefore, we introduce the following framework. For a more detailed discussion of the weak formulation, see [FGŚ17, Section 2.1.2].

Assumption 3.1. (A1)′ *Let $(W_t^\nu)_{t\in[s,T]}$ be a cylindrical Wiener process on a probability space $(\Omega^\nu, \mathcal{F}^\nu, \mathbb{P}^\nu)$ with values in some real, separable Hilbert space Ξ and $W_s^\nu = 0$ $\mathbb{P}^\nu$-almost surely. Let $(\mathcal{F}_{\nu,t}^s)_{t\in[s,T]}$ be the filtration generated by (W_t^ν) augmented by all $\mathbb{P}^\nu$-null sets. Following [FGŚ17, Definition 2.7], we call $\nu = (\Omega^\nu, \mathcal{F}^\nu, (\mathcal{F}_{\nu,t}^s)_{t\in[s,T]}, \mathbb{P}^\nu, W^\nu)$ a reference probability space. Furthermore, assume that ν is standard in the sense of [FGŚ17, Definition 2.8].*

(A2)′ *Let U be a non-empty subset of a separable Banach space $\mathcal{U}$, and let*

$$\mathcal{U}_s^\nu := \Big\{ u : [s,T]\times\Omega \to U \Big| u \text{ is } (\mathcal{F}_{\nu,t}^s) - \text{progressively measurable and} \\ \sup_{t\in[s,T]} \mathbb{E}\left[\|u_t\|_{\mathcal{U}}^k\right] < \infty, \forall k\in\mathbb{N}\Big\}. \quad (3.1)$$

The set of all admissible controls is given by

$$\mathcal{U}_s := \bigcup_\nu \mathcal{U}_s^\nu, \quad (3.2)$$

where the union is taken over all standard reference probability spaces ν.

We keep the Assumption 2.2, and impose additionally the following assumption on the coefficients of the state equation.

Assumption 3.2. (B3)′ *Assume that there exists a generic constant $C > 0$ such that* *for all* $x, y \in \mathbb{R}$ *and* $u \in U$ *it holds*

$$\begin{cases} |b_x(x,u) - b_x(y,u)| \le C|x-y| \\ |b_{xx}(x,u) - b_{xx}(y,u)| \le C|x-y|. \end{cases} \tag{3.3}$$

(B4)′ *Assume that there exists a generic constant* $C > 0$ *such that for all* $x, y \in \mathbb{R}$ *and* $u \in U$ *it holds*

$$\begin{cases} |\sigma_x(x,u) - \sigma_x(y,u)| \le C|x-y| \\ |\sigma_{xx}(x,u) - \sigma_{xx}(y,u)| \le C|x-y|. \end{cases} \tag{3.4}$$

Now, we define the value function as

$$V(s,x) := \inf_{u \in \mathcal{U}_s} J^\nu(s,x;u), \tag{3.5}$$

where

$$J^\nu(s,x;u) = \mathbb{E}^\nu \left[\int_s^T \int_\Lambda l(x_t^u(\lambda), u_t) \mathrm{d}\lambda \mathrm{d}t + \int_\Lambda h(x_T^u(\lambda)) \mathrm{d}\lambda \right]. \tag{3.6}$$

Note that this value function coincides with the value function obtained by minimizing the cost functional J over all admissible controls defined on any fixed reference probability space (not necessarily standard), see [FGŚ17, Theorem 2.22], and thus we would obtain the same value function by minimizing over all admissible controls and all reference probability spaces. In this setting, the value function satisfies the dynamic programming principle, see [FGŚ17, Theorem 2.24].

Throughout this chapter and the following chapter, we denote by

$$\mathbb{E}_t[\,\cdot\,] := \mathbb{E}[\,\cdot\,|\mathcal{F}_{\nu,t}^s] \tag{3.7}$$

the conditional expectation on a given reference probability space ν.

3.2. Parabolic Derivatives

In this section, we are going to prove a relationship between the parabolic viscosity super- and subdifferentials of the value function on the one hand and the first and second order adjoint state on the other hand.

First, let us recall the definition of the parabolic viscosity super- and subdifferential.

Definition 3.3. For $v \in C([s,T] \times L^2(\Lambda))$ the parabolic viscosity superdifferential of v at $(t,x) \in [s,T) \times L^2(\Lambda)$ is the set

$$\begin{aligned} D_{t+,x}^{1,2,+} v(t,x) := \Bigg\{ (G,p,P) \in \mathbb{R} \times L^2(\Lambda) \times \mathcal{S}(L^2(\Lambda)) \Bigg| \limsup_{\tau \downarrow 0, z \to 0} \frac{1}{\tau + \|z\|_{L^2(\Lambda)}^2} \\ \left[v(t+\tau, x+z) - v(t,x) - G\tau - \langle p, z \rangle_{L^2(\Lambda)} - \frac{1}{2} \langle z, Pz \rangle_{L^2(\Lambda)} \right] \le 0 \Bigg\}. \end{aligned} \tag{3.8}$$

The parabolic viscosity subdifferential $D^{1,2,-}_{t+,x}v$ is defined analogously with the lim sup replaced by lim inf and the $\leq$ replaced by $\geq$.

Now, we are ready to state the main result of this section.

Theorem 3.4 (Parabolic Viscosity Differentials)**.** *Let $(\bar{x}, \bar{u})$ be an optimal pair of the control problem (2.2) and (2.1), (p, q) and (P, Q) be the first and second order adjoint states, respectively, $\mathcal{G}$ be the generalized Hamiltonian defined in (2.104), and V be the value function. Then it holds for almost every $t \in [s, T]$,*

$$[-\langle \Delta\bar{x}_t, p_t\rangle_{H^{-1}(\Lambda)\times H^1_0(\Lambda)} - \mathcal{G}(t, \bar{x}_t, \bar{u}_t), \infty) \times \{p_t\} \times \mathcal{S}_{\succeq P_t}(L^2(\Lambda)) \subset D^{1,2,+}_{t+,x}V(t, \bar{x}_t), \tag{3.9}$$

$\mathbb{P}$–almost surely. Furthermore, for almost every $t \in [s, T]$,

$$D^{1,2,-}_{t+,x}V(t, \bar{x}_t) \subset (-\infty, -\langle \Delta\bar{x}_t, p_t\rangle_{H^{-1}(\Lambda)\times H^1_0(\Lambda)} - \mathcal{G}(t, \bar{x}_t, \bar{u}_t)] \times \{p_t\} \times \mathcal{S}_{\preceq P_t}(L^2(\Lambda)), \tag{3.10}$$

$\mathbb{P}$–almost surely. Here we define

$$\mathcal{S}_{\succeq P_t}(L^2(\Lambda)) := \{S \in \mathcal{S}(L^2(\Lambda)) : S - P_t \text{ is a positive operator}\}, \tag{3.11}$$

and $\mathcal{S}_{\preceq P_t}(L^2(\Lambda))$ mutatis mutandis.

Remark 3.5. Equation (3.9) in particular implies that the parabolic viscosity superdifferential is not empty.

First, we discuss several lemmata that are needed in the proof of Theorem 3.4. We suggest that the reader skip directly to the proof of Theorem 3.4 in Section 3.2.5 and refer to the lemmata as needed.

3.2.1. Variational Equation

In contrast to Chapter 2, where we perturbed the optimal control, in this chapter we perturb the initial condition of the control problem. Nevertheless, the arguments used here are similar to the ones used before. We begin by introducing the appropriate variational equation and by deriving a priori bounds as well as regularity results for the solution.

Lemma 3.6. *Let $\tau \in [0, T - t)$ and $z \in L^2(\Lambda)$, and let*

$$\begin{cases} \mathrm{d}x^{\tau,z}_r = [\Delta x^{\tau,z}_r + b(x^{\tau,z}_r, \bar{u}_r)]\,\mathrm{d}r + \sigma(x^{\tau,z}_r, \bar{u}_r)\mathrm{d}W_r, \quad r \in [t+\tau, T] \\ x^{\tau,z}_{t+\tau} = z + \bar{x}_t \in L^2(\Lambda). \end{cases} \tag{3.12}$$

Define $y^{\tau,z}_r := x^{\tau,z}_r - \bar{x}_r$, i.e.,

$$\begin{cases} \mathrm{d}y^{\tau,z}_r = [\Delta y^{\tau,z}_r + b(x^{\tau,z}_r, \bar{u}_r) - b(\bar{x}_r, \bar{u}_r)]\,\mathrm{d}r \\ \qquad\qquad + [\sigma(x^{\tau,z}_r, \bar{u}_r) - \sigma(\bar{x}_r, \bar{u}_r)]\,\mathrm{d}W_r, \quad r \in [t+\tau, T] \\ y^{\tau,z}_{t+\tau} = z + \bar{x}_t - \bar{x}_{t+\tau} \in L^2(\Lambda). \end{cases} \tag{3.13}$$

Then, for any $k \in \mathbb{N}$ it holds for almost every $t \in [s,T]$,

$$\mathbb{E}_t\left[\left(\int_{t+\tau}^T \|y_r^{\tau,z}\|^2_{H_0^1(\Lambda)}\mathrm{d}r\right)^k + \sup_{r\in[t+\tau,T]}\|y_r^{\tau,z}\|^{2k}_{L^2(\Lambda)}\right] \le C\left(\tau^k + \|z\|^{2k}_{L^2(\Lambda)}\right) \tag{3.14}$$

$\mathbb{P}$–almost surely.

Proof. Using the Lipschitz continuity of b and σ, and standard a priori estimates, we obtain similarly to the proof of Lemma 2.5

$$\begin{aligned}
&\mathbb{E}_t\left[\left(\int_{t+\tau}^T \|y_r^{\tau,z}\|^2_{H_0^1(\Lambda)}\mathrm{d}r\right)^k + \sup_{r\in[t+\tau,T]}\|y_r^{\tau,z}\|^{2k}_{L^2(\Lambda)}\right] \\
&\le C\mathbb{E}_t\left[\|z + \bar{x}_t - \bar{x}_{t+\tau}\|^{2k}_{L^2(\Lambda)}\right] \\
&\le C\left(\|z\|^{2k}_{L^2(\Lambda)} + \mathbb{E}_t\left[\|\bar{x}_{t+\tau} - \bar{x}_t\|^{2k}_{L^2(\Lambda)}\right]\right).
\end{aligned} \tag{3.15}$$

The claim follows from the fact that

$$\mathbb{E}_t\left[\|\bar{x}_{t+\tau} - \bar{x}_t\|^{2k}_{L^2(\Lambda)}\right] \le C\tau^k, \tag{3.16}$$

for every $k \in \mathbb{N}$. □

Next, we derive a Taylor expansion for the variational process $y^{\tau,z}$.

Lemma 3.7. *The variational process $y^{\tau,z}$ given by* (3.13) *satisfies the equations*

$$\begin{cases}
\mathrm{d}y_r^{\tau,z} = \left[\Delta y_r^{\tau,z} + b_x(\bar{x}_r,\bar{u}_r)y_r^{\tau,z} + \varphi_r^1\right]\mathrm{d}r + \left[\sigma_x(\bar{x}_r,\bar{u}_r)y_r^{\tau,z} + \psi_r^1\right]\mathrm{d}W_r \\
y_{t+\tau}^{\tau,z} = z + \bar{x}_t - \bar{x}_{t+\tau} \in L^2(\Lambda),
\end{cases} \tag{3.17}$$

where

$$\begin{aligned}
\varphi_r^1 &:= \int_0^1 \left[b_x(\bar{x}_r + \theta y_r^{\tau,z}, \bar{u}_r) - b_x(\bar{x}_r,\bar{u}_r)\right] y_r^{\tau,z}\mathrm{d}\theta \\
\psi_r^1 &:= \int_0^1 \left[\sigma_x(\bar{x}_r + \theta y_r^{\tau,z}, \bar{u}_r) - \sigma_x(\bar{x}_r,\bar{u}_r)\right] y_r^{\tau,z}\mathrm{d}\theta,
\end{aligned} \tag{3.18}$$

and

$$\begin{cases}
\mathrm{d}y_r^{\tau,z} = \left[\Delta y_r^{\tau,z} + b_x(\bar{x}_r,\bar{u}_r)y_r^{\tau,z} + \frac{1}{2}b_{xx}(\bar{x}_r,\bar{u}_r)y_r^{\tau,z}y_r^{\tau,z} + \varphi_r^2\right]\mathrm{d}r \\
\qquad\qquad + \left[\sigma_x(\bar{x}_r,\bar{u}_r)y_r^{\tau,z} + \frac{1}{2}\sigma_{xx}(\bar{x}_r,\bar{u}_r)y_r^{\tau,z}y_r^{\tau,z} + \psi_r^2\right]\mathrm{d}W_r, \quad r \in [t+\tau,T] \\
y_{t+\tau}^{\tau,z} = z + \bar{x}_t - \bar{x}_{t+\tau} \in L^2(\Lambda),
\end{cases} \tag{3.19}$$

where

$$\begin{aligned}\varphi_r^2 &:= \int_0^1 (1-\theta)\left[b_{xx}(\bar{x}_r + \theta y_r^{\tau,z}, \bar{u}_r) - b_{xx}(\bar{x}_r, \bar{u}_r)\right] y_r^{\tau,z} y_r^{\tau,z} \mathrm{d}\theta \\ \psi_r^2 &:= \int_0^1 (1-\theta)\left[\sigma_{xx}(\bar{x}_r + \theta y_r^{\tau,z}, \bar{u}_r) - \sigma_{xx}(\bar{x}_r, \bar{u}_r)\right] y_r^{\tau,z} y_r^{\tau,z} \mathrm{d}\theta.\end{aligned} \tag{3.20}$$

The remainder terms satisfy for every $k \in \mathbb{N}$, for almost every $t \in [s,T]$,

$$\begin{aligned}\mathbb{E}_t\left[\int_{t+\tau}^T \|\varphi_r^1\|_{L^2(\Lambda)}^{2k} \mathrm{d}r\right] &= o\left(\tau^k + \|z\|_{L^2(\Lambda)}^{2k}\right), \\ \mathbb{E}_t\left[\int_{t+\tau}^T \|\psi_r^1\|_{L_2(\Xi, L^2(\Lambda))}^{2k} \mathrm{d}r\right] &= o\left(\tau^k + \|z\|_{L^2(\Lambda)}^{2k}\right),\end{aligned} \tag{3.21}$$

as $\tau \downarrow 0$, $z \to 0$, $z \in L^2(\Lambda)$, $\mathbb{P}$–almost surely, and for almost every $t \in [s,T]$

$$\begin{aligned}\mathbb{E}_t\left[\int_{t+\tau}^T \|\varphi_r^2\|_{L^2(\Lambda)}^{k} \mathrm{d}r\right] &= o\left(\tau^k + \|z\|_{L^2(\Lambda)}^{2k}\right), \\ \mathbb{E}_t\left[\int_{t+\tau}^T \|\psi_r^2\|_{L_2(\Xi, L^2(\Lambda))}^{k} \mathrm{d}r\right] &= o\left(\tau^k + \|z\|_{L^2(\Lambda)}^{2k}\right),\end{aligned} \tag{3.22}$$

as $\tau \downarrow 0$, $z \to 0$, $z \in L^2(\Lambda)$, $\mathbb{P}$–almost surely.

Proof. The equations follow from the original equation for $y^{\tau,z}$ and Taylor's theorem for the Gâteaux derivative, see [Zei86, Section 4.6].

Now let us prove the first asymptotic in (3.21). By Lipschitz continuity of the derivative of b, we have

$$\begin{aligned}&\mathbb{E}_t\left[\int_{t+\tau}^T \left\|\int_0^1 \left[b_x(\bar{x}_r + \theta y_r^{\tau,z}, \bar{u}_r) - b_x(\bar{x}_r, \bar{u}_r)\right] y_r^{\tau,z} \mathrm{d}\theta\right\|_{L^2(\Lambda)}^{2k} \mathrm{d}r\right] \\ &\le \mathbb{E}_t\left[\int_{t+\tau}^T \int_0^1 \theta^{2k} \|y_r^{\tau,z}\|_{L^2(\Lambda)}^{4k} \mathrm{d}\theta \mathrm{d}r\right] \\ &\le \mathbb{E}_t\left[\int_{t+\tau}^T \|y_r^{\tau,z}\|_{L^2(\Lambda)}^{4k} \mathrm{d}r\right] \\ &\le C\left(\tau^{2k} + \|z\|_{L^2(\Lambda)}^{4k}\right),\end{aligned} \tag{3.23}$$

where we used Lemma 3.6 in the last step. The remaining estimates follow analogously using the Lipschitz continuity of the first derivative of σ and the Lipschitz continuity of the second derivatives of b and σ. □

The following higher regularity of the variational process at the terminal time is needed in order to extract convergent subsequences as $\tau + \|z\|_{L^2(\Lambda)}^2$ tends to zero.

Lemma 3.8. *Let $y^{\tau,z}$ be the variational process given by equation* (3.17). *Then, for any $\gamma \in (0, 1/4)$, we have for almost every $t \in [s, T]$,*

$$\mathbb{E}_t \left[\|y_T^{\tau,z}\|^2_{H_0^\gamma(\Lambda)} \right] \leq C \left(\tau + \|z\|^2_{L^2(\Lambda)} \right) \tag{3.24}$$

$\mathbb{P}$*–almost surely.*

Proof. The proof is similar to the proof of Lemma 2.15. By Lemma 3.7 and Duhamel's formula, we have

$$\begin{aligned} y_T^{\tau,z} =& S_{T-t-\tau}(z + \bar{x}_t - \bar{x}_{t+\tau}) + \int_{t+\tau}^T S_{r-t-\tau} \left(b_x(\bar{x}_r, \bar{u}_r) y_r^{\tau,z} + \varphi_r^1 \right) \mathrm{d}r \\ &+ \int_{t+\tau}^T S_{r-t-\tau} \left(\sigma_x(\bar{x}_r, \bar{u}_r) y_r^{\tau,z} + \psi_r^1 \right) \mathrm{d}W_r, \end{aligned} \tag{3.25}$$

where $(S_r)_{r \geq 0}$ denotes the heat semigroup. Now, let us estimate the $H_0^\gamma(\Lambda)$-norm. For the term involving the initial condition, we have

$$\begin{aligned} \mathbb{E}_t \left[\|S_{T-t-\tau}(z + \bar{x}_t - \bar{x}_{t+\tau})\|^2_{H_0^\gamma(\Lambda)} \right] &\leq C \mathbb{E}_t \left[\|z + \bar{x}_t - \bar{x}_{t+\tau}\|^2_{L^2(\Lambda)} \right] \\ &\leq C \left(\|z\|^2_{L^2(\Lambda)} + \mathbb{E}_t \left[\|\bar{x}_{t+\tau} - \bar{x}_t\|^2_{L^2(\Lambda)} \right] \right). \end{aligned} \tag{3.26}$$

Since

$$\mathbb{E}_t \left[\|\bar{x}_{t+\tau} - \bar{x}_t\|^2_{L^2(\Lambda)} \right] \leq C\tau, \tag{3.27}$$

the term involving the initial condition satisfies the required bound. Now let us consider the first integral in (3.25). We have

$$\begin{aligned} &\mathbb{E}_t \left[\left\| \int_{t+\tau}^T S_{r-t-\tau} \left(b_x(\bar{x}_r, \bar{u}_r) y_r^{\tau,z} \right) \mathrm{d}r \right\|^2_{H_0^\gamma(\Lambda)} \right] \\ &\leq C \mathbb{E}_t \left[\int_{t+\tau}^T \frac{1}{(r-t-\tau)^{2\gamma}} \|b_x(\bar{x}_r, \bar{u}_r) y_r^{\tau,z}\|^2_{L^2(\Lambda)} \mathrm{d}r \right] \\ &\leq C \sup_{t+\tau \leq r \leq T} \mathbb{E}_t \left[\|y_r^{\tau,z}\|^2_{L^2(\Lambda)} \right] \int_{t+\tau}^T \frac{1}{(r-t-\tau)^{2\gamma}} \mathrm{d}r. \end{aligned} \tag{3.28}$$

The required bound now follows from Lemma 3.6. For the second part of the first integral, we have

$$
\begin{aligned}
&\mathbb{E}_t\left[\left\|\int_{t+\tau}^{T} S_{r-t-\tau}\varphi_r^1 \mathrm{d}r\right\|_{H_0^\gamma(\Lambda)}^2\right] \\
&\leq C\mathbb{E}_t\left[\int_{t+\tau}^{T}\frac{\|\varphi_r^1\|_{L^2(\Lambda)}^2}{(r-t-\tau)^{2\gamma}}\mathrm{d}r\right] \\
&\leq C\mathbb{E}_t\left[\int_{t+\tau}^{T}\|\varphi_r^1\|_{L^2(\Lambda)}^4\,\mathrm{d}r\right]^{\frac{1}{2}}\left(\int_{t+\tau}^{T}\frac{1}{(r-t-\tau)^{4\gamma}}\mathrm{d}r\right)^{\frac{1}{2}}.
\end{aligned}
\tag{3.29}
$$

The required bound for this term follows from Lemma 3.7. For the stochastic integral in (3.25), we have

$$
\begin{aligned}
&\mathbb{E}_t\left[\left\|\int_{t+\tau}^{T} S_{r-t-\tau}\left(\sigma_x(\bar{x}_r,\bar{u}_r)y_r^{\tau,z}+\psi_r^1\right)\mathrm{d}W_r\right\|_{H_0^\gamma(\Lambda)}^2\right] \\
&=\mathbb{E}_t\left[\int_{t+\tau}^{T}\left\|S_{r-t-\tau}\left(\sigma_x(\bar{x}_r,\bar{u}_r)y_r^{\tau,z}+\psi_r^1\right)\right\|_{L_2(\Xi,H_0^\gamma(\Lambda))}^2\mathrm{d}r\right] \\
&\leq\mathbb{E}_t\left[\int_{t+\tau}^{T}\frac{C}{(r-t-\tau)^{2\gamma}}\left\|\sigma_x(\bar{x}_r,\bar{u}_r)y_r^{\tau,z}+\psi_r^1\right\|_{L_2(\Xi,L^2(\Lambda))}^2\mathrm{d}r\right].
\end{aligned}
\tag{3.30}
$$

Using the same argument as above yields the claim. □

3.2.2. Duality Relations

Now, we discuss the duality relations for the first and second order adjoint states, respectively.

Lemma 3.9. *Let $y^{\tau,z}$ be the variational process given by* (3.19) *and let p be the first order adjoint state. Then it holds for almost every $t\in[s,T]$,*

$$
\begin{aligned}
&\mathbb{E}_t\left[\int_{t+\tau}^{T}\langle l_x(\bar{x}_r,\bar{u}_r),y_r^{\tau,z}\rangle_{L^2(\Lambda)}\mathrm{d}r+\langle h_x(\bar{x}_T),y_T^{\tau,z}\rangle_{L^2(\Lambda)}\right] \\
&=\mathbb{E}_t\left[\langle p_{t+\tau},y_{t+\tau}^{\tau,z}\rangle_{L^2(\Lambda)}\right] \\
&\quad+\mathbb{E}_t\left[\frac{1}{2}\int_{t+\tau}^{T}\langle p_r,b_{xx}(\bar{x}_r,\bar{u}_r)y_r^{\tau,z}y_r^{\tau,z}\rangle_{L^2(\Lambda)}+\langle q_r,\sigma_{xx}(\bar{x}_r,\bar{u}_r)y_r^{\tau,z}y_r^{\tau,z}\rangle_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}r\right] \\
&\quad+\mathbb{E}_t\left[\int_{t+\tau}^{T}\langle p_r,\varphi_r^2\rangle_{L^2(\Lambda)}+\langle q_r,\psi_r^2\rangle_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}r\right],
\end{aligned}
\tag{3.31}
$$

$\mathbb{P}$*–almost surely. Furthermore, it holds for almost every $t\in[s,T]$,*

$$
\mathbb{E}_t\left[\int_{t+\tau}^{T}\langle p_r,\varphi_r^2\rangle_{L^2(\Lambda)}+\langle q_r,\psi_r^2\rangle_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}r\right]=o\left(\tau+\|z\|_{L^2(\Lambda)}^2\right)
\tag{3.32}
$$

as $\tau \downarrow 0$*,* $z \to 0$*,* $z \in L^2(\Lambda)$*,* $\mathbb{P}$*–almost surely.*

Proof. Applying Itô's formula for variational solutions of SPDEs (see [Par21, Lemma 2.15] or [Kry13, Section 3]) to $\langle y_r^{\tau,z}, p_r\rangle_{L^2(\Lambda)}$ yields

$$
\begin{aligned}
&\langle y_T^{\tau,z}, h_x(\bar{x}_T)\rangle_{L^2(\Lambda)} \\
&= \langle y_{t+\tau}^{\tau,z}, p_{t+\tau}\rangle_{L^2(\Lambda)} + \int_{t+\tau}^T p_r \mathrm{d}y_r^{\tau,z} + \int_{t+\tau}^T y_r^{\tau,z}\mathrm{d}p_r + \int_{t+\tau}^T \mathrm{d}\langle y^{\tau,z}, p\rangle_r \\
&= \int_{t+\tau}^T \langle p_r, \Delta y_r^{\tau,z} + b_x(\bar{x}_r,\bar{u}_r)y_r^{\tau,z} + \frac{1}{2}b_{xx}(\bar{x}_r,\bar{u}_r)y_r^{\tau,z}y_r^{\tau,z} + \varphi_r^2\rangle_{L^2(\Lambda)}\mathrm{d}r \\
&\quad + \int_{t+\tau}^T \langle p_r, \sigma_x(\bar{x}_r,\bar{u}_r)y_r^{\tau,z} + \frac{1}{2}\sigma_{xx}(\bar{x}_r,\bar{u}_r)y_r^{\tau,z}y_r^{\tau,z} + \psi_r^2\mathrm{d}W_r\rangle_{L^2(\Lambda)} \\
&\quad - \int_{t+\tau}^T \langle y_r^{\tau,z}, \Delta p_r + b_x(\bar{x}_r,\bar{u}_r)p_r + \langle \sigma_x(\bar{x}_r,\bar{u}_r), q_r\rangle_{L_2(\Xi,\mathbb{R})} + l_x(\bar{x}_r,\bar{u}_r)\rangle_{L^2(\Lambda)}\mathrm{d}r \\
&\quad + \int_{t+\tau}^T \langle y_r^{\tau,z}, q_r\mathrm{d}W_r\rangle_{L^2(\Lambda)} \\
&\quad + \int_{t+\tau}^T \langle q_r, \sigma_x(\bar{x}_r,\bar{u}_r)y_r^{\tau,z} + \frac{1}{2}\sigma_{xx}(\bar{x}_r,\bar{u}_r)y_r^{\tau,z}y_r^{\tau,z} + \psi_r^2\rangle_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}r
\end{aligned}
\tag{3.33}
$$

Applying an integration by parts for the Laplace operator, canceling out matching terms with opposite sign, and taking the conditional expectation yields the claim (3.31).

For the remainder term estimate, we observe

$$
\begin{aligned}
&\mathbb{E}_t\left[\int_{t+\tau}^T \langle p_r, \varphi_r^2\rangle_{L^2(\Lambda)} + \langle q_r, \psi_r^2\rangle_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}r\right] \\
&\leq \mathbb{E}_t\left[\int_{t+\tau}^T \|p_r\|_{L^2(\Lambda)}\|\varphi_r^2\|_{L^2(\Lambda)} + \|q_r\|_{L_2(\Xi,L^2(\Lambda))}\|\psi_r^2\|_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}r\right] \\
&\leq \mathbb{E}_t\left[\int_{t+\tau}^T \|p_r\|_{L^2(\Lambda)}^2\mathrm{d}r\right]^{\frac{1}{2}} \mathbb{E}_t\left[\int_{t+\tau}^T \|\varphi_r^2\|_{L^2(\Lambda)}^2\mathrm{d}r\right]^{\frac{1}{2}} \\
&\qquad + \mathbb{E}_t\left[\int_{t+\tau}^T \|q_r\|_{L_2(\Xi,L^2(\Lambda))}^2\mathrm{d}r\right]^{\frac{1}{2}} \mathbb{E}_t\left[\int_{t+\tau}^T \|\psi_r^2\|_{L_2(\Xi,L^2(\Lambda))}^2\mathrm{d}r\right]^{\frac{1}{2}}
\end{aligned}
\tag{3.34}
$$

Since the first factor is finite in each case, the claim follows from the remainder estimates (3.22). □

Lemma 3.10. *Let* $y^{\tau,z}$ *be the process given by equation* (3.17)*, and let* P^η *be the* *mollified second order adjoint state. Then it holds for almost every $t \in [s,T]$,*

$$\begin{aligned}
&\mathbb{E}_t\left[\int_{t+\tau}^T \int_\Lambda (l_{xx}(\bar{x}_r(\lambda),\bar{u}_r) + b_{xx}(\bar{x}_r(\lambda),\bar{u}_r)p_r(\lambda))y_r^{\tau,z}(\lambda)y_r^{\tau,z}(\lambda)\mathrm{d}\lambda\mathrm{d}r\right]\\
&\quad + \mathbb{E}_t\left[\int_{t+\tau}^T \int_\Lambda \langle \sigma_{xx}(\bar{x}_r(\lambda),\bar{u}_r), q_r\rangle_{L_2(\Xi,\mathbb{R})} y_r^{\tau,z}(\lambda)y_r^{\tau,z}(\lambda)\mathrm{d}\lambda\mathrm{d}r\right]\\
&\quad + \mathbb{E}_t\left[\int_{\Lambda^2} h_{xx}^\eta(\lambda,\mu)y_T^{\tau,z}(\lambda)y_T^{\tau,z}(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&= \mathbb{E}_t\left[\int_{\Lambda^2} P_{t+\tau}^\eta(\lambda,\mu)y_{t+\tau}^{\tau,z}(\lambda)y_{t+\tau}^{\tau,z}(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\quad + \mathbb{E}_t\left[\int_{t+\tau}^T \langle P_r^\eta, \Phi_r^{\tau,z}\rangle_{L^2(\Lambda^2)} + \langle Q_r^\eta, \Psi_r^{\tau,z}\rangle_{L_2(\Xi,L^2(\Lambda^2))}\mathrm{d}r\right],
\end{aligned} \tag{3.35}$$

$\mathbb{P}$–almost surely, where

$$(\Phi^{\tau,z},\Psi^{\tau,z}) \in L^2([s,T]\times\Omega; L^2(\Lambda^2)) \times L^2([s,T]\times\Omega; L_2(\Xi, L^2(\Lambda^2))) \tag{3.36}$$

are given by

$$\begin{aligned}
\Phi_r^{\tau,z}(\lambda,\mu) :=& y_r^{\tau,z}(\lambda)\varphi_r^1(\mu) + y_r^{\tau,z}(\mu)\varphi_r^1(\lambda)\\
&+ \sigma_x(\bar{x}_r(\lambda),\bar{u}_r)y_r^{\tau,z}(\lambda)\psi_r^1(\mu) + \sigma_x(\bar{x}_r(\mu),\bar{u}_r)y_r^{\tau,z}(\mu)\psi_r^1(\lambda)\\
&+ \psi_r^1(\lambda)\psi_r^1(\mu),
\end{aligned} \tag{3.37}$$

and

$$\Psi_r^{\tau,z}(\lambda,\mu) := y_r^{\tau,z}(\lambda)\psi_r^1(\mu) + y_r^{\tau,z}(\mu)\psi_r^1(\lambda). \tag{3.38}$$

Furthermore, we have for almost every $t \in [s,T]$,

$$\mathbb{E}_t\left[\int_{t+\tau}^T \langle P_r^\eta, \Phi_r^{\tau,z}\rangle_{L^2(\Lambda^2)} + \langle Q_r^\eta, \Psi_r^{\tau,z}\rangle_{L_2(\Xi,L^2(\Lambda^2))}\mathrm{d}r\right] = o\left(\tau + \|z\|_{L^2(\Lambda)}^2\right) \tag{3.39}$$

as $\tau \downarrow 0$, $z \to 0$, $z \in L^2(\Lambda)$, $\mathbb{P}$–almost surely.

Proof. In order to invoke the second order adjoint state, we use the same idea as in the proof of the maximum principle in Chapter 2. We rewrite the quadratic terms in $y^{\tau,z}$ in the following way

$$\begin{aligned}
\langle p_r, b_{xx}(\bar{x}_r,\bar{u}_r)y_r^{\tau,z}y_r^{\tau,z}\rangle_{L^2(\Lambda)} &= \int_\Lambda p_r(\lambda)b_{xx}(\bar{x}_r(\lambda),\bar{u}_r)y_r^{\tau,z}(\lambda)y_r^{\tau,z}(\lambda)\mathrm{d}\lambda\\
&= \int_\Lambda p_r(\lambda)b_{xx}(\bar{x}_r(\lambda),\bar{u}_r)\delta(Y_r^{\tau,z})(\lambda)\mathrm{d}\lambda,
\end{aligned} \tag{3.40}$$

where $Y_r^{\tau,z}(\lambda,\mu) := y_r^{\tau,z}(\lambda)y_r^{\tau,z}(\mu)$ and δ is the operator introduced in Proposition 2.9.

Next, let us derive the equation for $Y^{\tau,z}$. Similar to the calculation in Chapter 2, we have

$$\begin{cases} \mathrm{d}Y_r^{\tau,z}(\lambda,\mu) = [\Delta Y_r^{\tau,z}(\lambda,\mu) + (b_x(\bar{x}_r(\lambda),\bar{u}_r) + b_x(\bar{x}_r(\mu),\bar{u}_r))Y_r^{\tau,z}(\lambda,\mu) \\ \qquad\qquad + \langle \sigma_x(\bar{x}_r(\lambda),\bar{u}_r), \sigma_x(\bar{x}_r(\mu),\bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})} Y_r^{\tau,z}(\lambda,\mu) + \Phi_r^{\tau,z}(\lambda,\mu)]\mathrm{d}r \\ \qquad\qquad + [(\sigma_x(\bar{x}_r(\lambda),\bar{u}_r) + \sigma_x(\bar{x}_r(\mu),\bar{u}_r))Y_r^{\tau,z}(\lambda,\mu) + \Psi_r^{\tau,z}(\lambda,\mu)]\mathrm{d}W_r \\ Y_{t+\tau}^{\tau,z} = (z + \bar{x}_t - \bar{x}_{t+\tau}) \otimes (z + \bar{x}_t - \bar{x}_{t+\tau}). \end{cases} \tag{3.41}$$

We want to apply Itô's formula to the product $\langle Y_r^{\tau,z}, P_r\rangle_{L^2(\Lambda^2)}$. Since P itself is not sufficiently regular, we take the mollified second order adjoint process given by equation (2.59), instead. Applying Itô's formula to $\langle Y_r^{\tau,z}, P_r^\eta\rangle_{L^2(\Lambda^2)}$ yields the duality relation (3.35).

For the first term in the remainder estimate, we observe

$$\mathbb{E}_t\left[\int_{t+\tau}^T \langle P_r^\eta, \Phi_r^{\tau,z}\rangle_{L^2(\Lambda^2)}\mathrm{d}r\right] \le \mathbb{E}_t\left[\int_{t+\tau}^T \|P_r^\eta\|^2_{L^2(\Lambda^2)}\mathrm{d}r\right]^{\frac{1}{2}} \mathbb{E}_t\left[\int_{t+\tau}^T \|\Phi_r^{\tau,z}\|^2_{L^2(\Lambda^2)}\mathrm{d}r\right]^{\frac{1}{2}}. \tag{3.42}$$

Since the first factor is finite, the claim follows from the a priori estimates in Lemma 3.6 and (3.21).

For the second term in the remainder estimate, we have

$$\begin{aligned} &\mathbb{E}_t\left[\int_{t+\tau}^T \langle Q_r^\eta, \Psi_r^{\tau,z}\rangle_{L_2(\Xi,L^2(\Lambda^2))}\mathrm{d}r\right] \\ &\le \mathbb{E}_t\left[\int_{t+\tau}^T \|Q_r^\eta\|^2_{L_2(\Xi,L^2(\Lambda^2))}\mathrm{d}r\right]^{\frac{1}{2}} \mathbb{E}_t\left[\int_{t+\tau}^T \|\Psi_r^{\tau,z}\|^2_{L_2(\Xi,L^2(\Lambda^2))}\mathrm{d}r\right]^{\frac{1}{2}}. \end{aligned} \tag{3.43}$$

Since

$$\begin{aligned} \|\Psi_r^{\tau,z}\|^2_{L_2(\Xi,L^2(\Lambda^2))} &= \|y_r^{\tau,z}\otimes\psi_r^1 + \psi_r^1\otimes y_r^{\tau,z}\|^2_{L_2(\Xi,L^2(\Lambda^2))} \\ &= \|y_r^{\tau,z}\|^2_{L^2(\Lambda)}\|\psi_r^1\|^2_{L_2(\Xi,L^2(\Lambda))} + \|y_r^{\tau,z}\|^2_{L^2(\Lambda)}\|\psi_r^1\|^2_{L_2(\Xi,L^2(\Lambda))}, \end{aligned} \tag{3.44}$$

the claim follows again from Lemma 3.6 and (3.21). ☐

3.2.3. Time-Increments

The next two lemmata address the time increment.

Lemma 3.11. *It holds for almost every $t \in [s,T]$,*

$$\begin{aligned} &\mathbb{E}_t\left[\langle p_{t+\tau}, \bar{x}_{t+\tau} - \bar{x}_t\rangle_{L^2(\Lambda)}\right] \\ &= \tau\mathbb{E}_t\left[\langle p_t, \Delta\bar{x}_t + b(\bar{x}_t,\bar{u}_t)\rangle_{H_0^1(\Lambda)\times H^{-1}(\Lambda)} + \langle q_t, \sigma(\bar{x}_t,\bar{u}_t)\rangle_{L_2(\Xi,L^2(\Lambda))}\right] + o(\tau), \end{aligned} \tag{3.45}$$

as $\tau \downarrow 0$, $\mathbb{P}$–almost surely.

Proof. Applying Itô's formula for variational solutions of SPDEs and taking the conditional expectation, we obtain

$$
\begin{aligned}
&\mathbb{E}_t\left[\langle p_{t+\tau}, \bar{x}_{t+\tau}-\bar{x}_t\rangle_{L^2(\Lambda)}\right] \\
&=\mathbb{E}_t\left[\int_t^{t+\tau}\langle \Delta p_r, \bar{x}_t\rangle_{H^{-1}(\Lambda)\times H_0^1(\Lambda)}+\langle p_r, b(\bar{x}_r,\bar{u}_r)\rangle_{L^2(\Lambda)}+\langle \sigma(\bar{x}_r,\bar{u}_r), q_r\rangle_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}r\right] \\
&\quad-\mathbb{E}_t\left[\int_t^{t+\tau}\langle \bar{x}_r-\bar{x}_t, b_x(\bar{x}_r,\bar{u}_r)p_r+\langle\sigma_x(\bar{x}_r,\bar{u}_r), q_r\rangle+l_x(\bar{x}_r,\bar{u}_x)\rangle_{L^2(\Lambda)}\,\mathrm{d}r\right].
\end{aligned}
\tag{3.46}
$$

Note that the stochastic integrals vanish under the expectation. For the first term, we have

$$
\begin{aligned}
&\left|\mathbb{E}_t\left[\frac{1}{\tau}\int_t^{t+\tau}\langle \Delta(p_r-p_t), \bar{x}_t\rangle_{H^{-1}(\Lambda)\times H_0^1(\Lambda)}\mathrm{d}r\right]\right| \\
&\leq \mathbb{E}_t\left[\frac{1}{\tau}\int_t^{t+\tau}\|p_r-p_t\|_{H_0^1(\Lambda)}\|\bar{x}_t\|_{H_0^1(\Lambda)}\mathrm{d}r\right] \\
&\leq \mathbb{E}_t\left[\frac{1}{\tau}\int_t^{t+\tau}\|p_r-p_t\|^2_{H_0^1(\Lambda)}\mathrm{d}r\right]^{\frac{1}{2}}\|\bar{x}_t\|_{H_0^1(\Lambda)}.
\end{aligned}
\tag{3.47}
$$

By Lebesgue's differentiation theorem, we have for almost every $t\in[s,T]$,

$$
\frac{1}{\tau}\int_t^{t+\tau}\mathbb{E}_t\left[\|p_r-p_t\|^2_{H_0^1(\Lambda)}\right]\mathrm{d}r\to 0 \tag{3.48}
$$

$\mathbb{P}$–almost surely. Hence, we obtain for almost every $t\in[s,T]$,

$$
\mathbb{E}_t\left[\frac{1}{\tau}\int_t^{t+\tau}\langle \Delta p_r, \bar{x}_t\rangle_{H^{-1}(\Lambda)\times H_0^1(\Lambda)}\mathrm{d}r\right]\to\langle \Delta p_t, \bar{x}_t\rangle_{H^{-1}(\Lambda)\times H_0^1(\Lambda)} \tag{3.49}
$$

$\mathbb{P}$–almost surely along some subsequence. Arguing similarly for the second and third term of equation (3.46) and noticing that the last line is of order $o(\tau)$ concludes the proof. □

Lemma 3.12. *It holds for almost every $t\in[s,T]$,*

$$
\begin{aligned}
&\mathbb{E}_t\left[\int_{\Lambda^2}P^\eta_{t+\tau}(\lambda,\mu)(\bar{x}_{t+\tau}-\bar{x}_t)(\lambda)(\bar{x}_{t+\tau}-\bar{x}_t)(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right] \\
&=\tau\mathbb{E}_t\left[\int_{\Lambda^2}P^\eta_t(\lambda,\mu)\langle\sigma(\bar{x}_t(\lambda),\bar{u}_t),\sigma(\bar{x}_t(\mu),\bar{u}_t)\rangle_{L_2(\Xi,\mathbb{R})}\mathrm{d}\lambda\mathrm{d}\mu\right]+o(\tau),
\end{aligned}
\tag{3.50}
$$

$\tau\downarrow 0$, $\mathbb{P}$–almost surely.

Proof. The equation for the tensor product $(\bar{x}_{t+\tau}-\bar{x}_t)\otimes(\bar{x}_{t+\tau}-\bar{x}_t)$ is

$$\begin{aligned}
&\mathrm{d}((\bar{x}_{t+\tau}-\bar{x}_t)(\lambda)(\bar{x}_{t+\tau}-\bar{x}_t)(\mu))\\
&=[(\bar{x}_{t+\tau}-\bar{x}_t)(\lambda)\Delta\bar{x}_{t+\tau}(\mu)+(\bar{x}_{t+\tau}-\bar{x}_t)(\mu)\Delta\bar{x}_{t+\tau}(\lambda)]\,\mathrm{d}\tau\\
&\quad+[(\bar{x}_{t+\tau}-\bar{x}_t)(\lambda)b(\bar{x}_{t+\tau}(\mu),\bar{u}_{t+\tau})+(\bar{x}_{t+\tau}-\bar{x}_t)(\mu)b(\bar{x}_{t+\tau}(\lambda),\bar{u}_{t+\tau})]\,\mathrm{d}\tau\\
&\quad+\langle\sigma(\bar{x}_{t+\tau}(\lambda),\bar{u}_{t+\tau}),\sigma(\bar{x}_{t+\tau}(\mu),\bar{u}_{t+\tau})\rangle_{L_2(\Xi,\mathbb{R})}\mathrm{d}\tau\\
&\quad+[(\bar{x}_{t+\tau}-\bar{x}_t)(\lambda)\sigma(\bar{x}_{t+\tau}(\mu),\bar{u}_{t+\tau})+(\bar{x}_{t+\tau}-\bar{x}_t)(\mu)\sigma(\bar{x}_{t+\tau}(\lambda),\bar{u}_{t+\tau})]\,\mathrm{d}W_\tau.
\end{aligned}\tag{3.51}$$

Again, applying Itô's formula for variational solutions of SPDEs and taking the conditional expectation yields

$$\begin{aligned}
&\mathbb{E}_t\left[\left\langle P^\eta_{t+\tau},(\bar{x}_{t+\tau}-\bar{x}_t)\otimes(\bar{x}_{t+\tau}-\bar{x}_t)\right\rangle\right]\\
&=\mathbb{E}_t\left[\int_t^{t+\tau}\langle P^\eta_r,\mathrm{d}((\bar{x}_r-\bar{x}_t)\otimes(\bar{x}_r-\bar{x}_t))\rangle+\int_t^{t+\tau}\langle(\bar{x}_r-\bar{x}_t)\otimes(\bar{x}_r-\bar{x}_t),\mathrm{d}P^\eta_r\rangle\right]\\
&\quad+\mathbb{E}_t\left[\langle P^\eta_\cdot,(\bar{x}_\cdot-\bar{x}_t)\otimes(\bar{x}_\cdot-\bar{x}_t)\rangle_{t+\tau}\right],
\end{aligned}\tag{3.52}$$

where

$$\begin{aligned}
&\mathbb{E}_t\left[\int_t^{t+\tau}\langle P^\eta_r,\mathrm{d}((\bar{x}_r-\bar{x}_t)\otimes(\bar{x}_r-\bar{x}_t))\rangle\right]\\
&=\mathbb{E}_t\left[\int_t^{t+\tau}\langle P^\eta_r,(\bar{x}_r-\bar{x}_t)(\lambda)\Delta\bar{x}_r(\mu)+(\bar{x}_r-\bar{x}_t)(\mu)\Delta\bar{x}_r(\lambda)\rangle\mathrm{d}r\right]\\
&\quad+\mathbb{E}_t\left[\int_t^{t+\tau}\langle P^\eta_r,(\bar{x}_r-\bar{x}_t)(\lambda)b(\bar{x}_r(\mu),\bar{u}_r)+(\bar{x}_r-\bar{x}_t)(\mu)b(\bar{x}_r(\lambda),\bar{u}_r)\rangle\mathrm{d}r\right]\\
&\quad+\mathbb{E}_t\left[\int_t^{t+\tau}\langle P^\eta_r,\langle\sigma(\bar{x}_r(\lambda),\bar{u}_r),\sigma(\bar{x}_r(\mu),\bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})}\rangle\mathrm{d}r\right],
\end{aligned}\tag{3.53}$$

and

$$
\begin{aligned}
&\mathbb{E}_t\left[\int_t^{t+\tau}\langle(\bar{x}_r-\bar{x}_t)\otimes(\bar{x}_r-\bar{x}_t),\mathrm{d}P_r^\eta\rangle\right]\\
&=\mathbb{E}_t\left[\int_t^{t+\tau}\langle(\bar{x}_r-\bar{x}_t)(\lambda)(\bar{x}_r-\bar{x}_t)(\mu),\Delta P_r^\eta(\lambda,\mu)\rangle\mathrm{d}r\right]\\
&\quad+\mathbb{E}_t\left[\int_t^{t+\tau}\langle(\bar{x}_r-\bar{x}_t)(\lambda)(\bar{x}_r-\bar{x}_t)(\mu),(b_x(\bar{x}_r(\lambda),\bar{u}_r)+b_x(\bar{x}_r(\mu),\bar{u}_r))P_r^\eta(\lambda,\mu)\rangle\mathrm{d}r\right]\\
&\quad+\mathbb{E}_t\left[\int_t^{t+\tau}\langle(\bar{x}_r-\bar{x}_t)(\lambda)(\bar{x}_r-\bar{x}_t)(\mu),\langle\sigma_x(\bar{x}_r(\lambda),\bar{u}_r),\sigma_x(\bar{x}_r(\mu),\bar{u}_r)\rangle P_r^\eta(\lambda,\mu)\rangle\mathrm{d}r\right]\\
&\quad+\mathbb{E}_t\left[\int_t^{t+\tau}\langle(\bar{x}_r-\bar{x}_t)(\lambda)(\bar{x}_r-\bar{x}_t)(\mu),\langle\sigma_x(\bar{x}_r(\lambda),\bar{u}_r)+\sigma_x(\bar{x}_r(\mu),\bar{u}_r),Q_r^\eta(\lambda,\mu)\rangle\rangle\mathrm{d}r\right]\\
&\quad+\mathbb{E}_t\left[\int_t^{t+\tau}\langle(\bar{x}_r-\bar{x}_t)(\lambda)(\bar{x}_r-\bar{x}_t)(\mu),\delta^*(l_{xx}(\bar{x}_r(\lambda),\bar{u}_r)+b_{xx}(\bar{x}_r(\lambda),\bar{u}_r)p_r(\lambda))\rangle\mathrm{d}r\right]\\
&\quad+\mathbb{E}_t\left[\int_t^{t+\tau}\langle(\bar{x}_r-\bar{x}_t)(\lambda)(\bar{x}_r-\bar{x}_t)(\mu),\delta^*(\langle\sigma_{xx}(\bar{x}_r(\lambda),\bar{u}_r),q_r\rangle_{L_2(\Xi,\mathbb{R})})\rangle\mathrm{d}r\right],
\end{aligned}
\tag{3.54}
$$

and

$$
\begin{aligned}
&\mathbb{E}_t\left[\langle P_\cdot^\eta,(\bar{x}_\cdot-\bar{x}_t)\otimes(\bar{x}_\cdot-\bar{x}_t)\rangle_{t+\tau}\right]\\
&=\mathbb{E}_t\left[\int_t^{t+\tau}\langle(\bar{x}_r-\bar{x}_t)\otimes\sigma(\bar{x}_r,\bar{u}_r)+\sigma(\bar{x}_r,\bar{u}_r)\otimes(\bar{x}_r-\bar{x}_t),Q_r^\eta\rangle_{L_2(\Xi,L^2(\Lambda^2))}\mathrm{d}r\right].
\end{aligned}
\tag{3.55}
$$

Except

$$
\mathbb{E}_t\left[\int_t^{t+\tau}\langle P_r^\eta,\langle\sigma(\bar{x}_r(\lambda),\bar{u}_r),\sigma(\bar{x}_r(\mu),\bar{u}_r)\rangle_{L_2(\Xi,\mathbb{R})}\rangle\mathrm{d}r\right],\tag{3.56}
$$

the integrand in each summand contains the term $\bar{x}_r-\bar{x}_t$. Therefore, arguing as in the proof of Lemma 3.11, when dividing by τ and taking the limit $\tau\downarrow 0$, this is the only remaining term. □

3.2.4. Mixed Time- and Space-Increments

Lemma 3.13. *It holds for almost every $t\in[s,T]$,*

$$
\mathbb{E}_t\left[\int_{\Lambda^2}P_{t+\tau}^\eta(\lambda,\mu)z(\lambda)(\bar{x}_{t+\tau}-\bar{x}_t)(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]=o\left(\tau+\|z\|_{L^2(\Lambda)}^2\right),\tag{3.57}
$$

as $\tau\downarrow 0$, $z\to 0$, $z\in L^2(\Lambda)$, $\mathbb{P}$–almost surely.

Proof. First note

$$
\begin{aligned}
&\mathbb{E}_t\left[\int_{\Lambda^2} P^\eta_{t+\tau}(\lambda,\mu)z(\lambda)(\bar{x}_{t+\tau}-\bar{x}_t)(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&=\mathbb{E}_t\left[\int_{\Lambda^2}(P^\eta_{t+\tau}-P^\eta_t)(\lambda,\mu)z(\lambda)(\bar{x}_{t+\tau}-\bar{x}_t)(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\quad+\mathbb{E}_t\left[\int_{\Lambda^2} P^\eta_t(\lambda,\mu)z(\lambda)(\bar{x}_{t+\tau}-\bar{x}_t)(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right].
\end{aligned}
\tag{3.58}
$$

For the first expectation, we have

$$
\begin{aligned}
&\mathbb{E}_t\left[\int_{\Lambda^2}(P^\eta_{t+\tau}-P^\eta_t)(\lambda,\mu)z(\lambda)(\bar{x}_{t+\tau}-\bar{x}_t)(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\le\mathbb{E}_t\left[\|P^\eta_{t+\tau}-P^\eta_t\|^2_{L^2(\Lambda^2)}\right]^{\frac{1}{2}}\|z\|_{L^2(\Lambda)}\mathbb{E}_t\left[\|\bar{x}_{t+\tau}-\bar{x}_t\|^2_{L^2(\Lambda)}\right]^{\frac{1}{2}}.
\end{aligned}
\tag{3.59}
$$

Since the second and third term are each of order $O\left(\sqrt{\tau+\|z\|^2_{L^2(\Lambda)}}\right)$, and P^η is continuous with values in $L^2(\Lambda^2)$ $\mathbb{P}$-almost surely, the whole expression is of order $o\left(\tau+\|z\|^2_{L^2(\Lambda)}\right)$.

For the second expectation in equation (3.58), we have

$$
\begin{aligned}
&\mathbb{E}_t\left[\int_{\Lambda^2} P^\eta_t(\lambda,\mu)z(\lambda)(\bar{x}_{t+\tau}-\bar{x}_t)(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\le\|P^\eta_t\|_{L^2(\Lambda^2)}\|z\|_{L^2(\Lambda)}\left\|\mathbb{E}_t\left[\bar{x}_{t+\tau}-\bar{x}_t\right]\right\|_{L^2(\Lambda)}.
\end{aligned}
\tag{3.60}
$$

Since

$$
\left\|\mathbb{E}_t\left[\bar{x}_{t+\tau}-\bar{x}_t\right]\right\|^2_{L^2(\Lambda)}=2\int_t^{t+\tau}\langle\mathbb{E}_t[\bar{x}_r-\bar{x}_t],\mathbb{E}_t[\Delta\bar{x}_r+b(\bar{x}_r,\bar{u}_r)]\rangle_{H^1_0(\Lambda)\times H^{-1}(\Lambda)}\mathrm{d}r,
\tag{3.61}
$$

by Lebesgue's differentiation theorem, we have for almost every $t\in[s,T]$,

$$
\left\|\mathbb{E}_t\left[\bar{x}_{t+\tau}-\bar{x}_t\right]\right\|_{L^2(\Lambda)}=o\left(\sqrt{\tau}\right)
\tag{3.62}
$$

$\mathbb{P}$-almost surely. Therefore, for almost every $t\in[s,T]$,

$$
\mathbb{E}_t\left[\int_{\Lambda^2} P^\eta_t(\lambda,\mu)z(\lambda)(\bar{x}_{t+\tau}-\bar{x}_t)(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]=o\left(\tau+\|z\|^2_{L^2(\Lambda)}\right),
\tag{3.63}
$$

$\mathbb{P}$–almost surely, which concludes the proof. ☐

Lemma 3.14. *It holds for almost every $t\in[s,T]$,*

$$
\mathbb{E}_t\left[\int_\Lambda(p_{t+\tau}(\lambda)-p_t(\lambda))z(\lambda)\mathrm{d}\lambda\right]=o\left(\tau+\|z\|^2_{L^2(\Lambda)}\right),
\tag{3.64}
$$

as $\tau \downarrow 0$, $z \to 0$, $z \in L^2(\Lambda)$, $\mathbb{P}$*–almost surely.*

Proof. As in the proof of Lemma 3.13, we have

$$\left\|\mathbb{E}_t\left[p_{t+\tau} - p_t\right]\right\|_{L^2(\Lambda)} = o\left(\sqrt{\tau}\right). \tag{3.65}$$

Therefore,

$$\begin{aligned}
&\mathbb{E}_t\left[\int_\Lambda (p_{t+\tau}(\lambda) - p_t(\lambda)) z(\lambda) \mathrm{d}\lambda\right] \\
&\leq \left\|\mathbb{E}_t\left[p_{t+\tau}(\lambda) - p_t(\lambda)\right]\right\|_{L^2(\Lambda)} \|z\|_{L^2(\Lambda)} \\
&= o\left(\tau + \|z\|^2_{L^2(\Lambda)}\right),
\end{aligned} \tag{3.66}$$

which concludes the proof. □

3.2.5. Proof of Theorem 3.4

Fix $t \in [s,T]$ such that all the preceding lemmata hold $\mathbb{P}$–almost surely, and let $\tau \in (0, T-t)$. Using the dynamic programming principle under the conditional expectation (see [YZ99, Chapter 4, Lemma 3.2 and Theorem 3.4] and [FGŚ17, Section 2.3.3]), we obtain for almost every $t \in [s,T]$,

$$\begin{aligned}
&V(t+\tau, \bar{x}_t + z) - V(t, \bar{x}_t) \\
&\leq \mathbb{E}_t\left[-\int_t^{t+\tau}\int_\Lambda l(\bar{x}_r(\lambda), \bar{u}_r)\mathrm{d}\lambda\mathrm{d}r + \int_{t+\tau}^T\int_\Lambda l(x_r^{\tau,z}(\lambda), \bar{u}_r) - l(\bar{x}_r(\lambda), \bar{u}_r)\mathrm{d}\lambda\mathrm{d}r\right] \\
&\quad + \mathbb{E}_t\left[\int_\Lambda h(x_T^{\tau,z}(\lambda)) - h(\bar{x}_T(\lambda))\mathrm{d}\lambda\right] \\
&= \mathbb{E}_t\left[-\int_t^{t+\tau}\int_\Lambda l(\bar{x}_r(\lambda), \bar{u}_r)\mathrm{d}\lambda\mathrm{d}r + \int_{t+\tau}^T\int_\Lambda l_x(\bar{x}_r(\lambda), \bar{u}_r) y_r^{\tau,z}(\lambda)\mathrm{d}\lambda\mathrm{d}r\right] \\
&\quad + \mathbb{E}_t\left[\int_\Lambda h_x(\bar{x}_T(\lambda)) y_T^{\tau,z}(\lambda)\mathrm{d}\lambda + \frac{1}{2}\int_{t+\tau}^T\int_\Lambda l_{xx}(\bar{x}_r(\lambda), \bar{u}_r) y_r^{\tau,z}(\lambda) y_r^{\tau,z}(\lambda)\mathrm{d}\lambda\mathrm{d}r\right] \\
&\quad + \mathbb{E}_t\left[\frac{1}{2}\int_\Lambda h_{xx}(\bar{x}_T(\lambda)) y_T^{\tau,z}(\lambda) y_T^{\tau,z}(\lambda)\mathrm{d}\lambda\right] + o\left(\tau + \|z\|^2_{L^2(\Lambda)}\right),
\end{aligned} \tag{3.67}$$

$\mathbb{P}$-almost surely, where the remainder terms of the Taylor expansion are of order $o\left(\tau + \|z\|^2_{L^2(\Lambda)}\right)$ for the same reason as in (3.21). Using the duality relations from Lemma 3.9 and Lemma 3.10, and the estimates for the remainder terms of the duality relations, we obtain

$$
\begin{aligned}
&V(t+\tau,\bar{x}_t+z)-V(t,\bar{x}_t)\\
&\le \mathbb{E}_t\left[-\int_t^{t+\tau}\int_\Lambda l(\bar{x}_r(\lambda),\bar{u}_r)\mathrm{d}\lambda\mathrm{d}r+\int_\Lambda p_{t+\tau}(\lambda)y^{\tau,z}_{t+\tau}(\lambda)\mathrm{d}\lambda\right]\\
&\quad+\mathbb{E}_t\left[\frac{1}{2}\int_{\Lambda^2}P^\eta_{t+\tau}(\lambda,\mu)y^{\tau,z}_{t+\tau}(\lambda)y^{\tau,z}_{t+\tau}(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\quad+\frac{1}{2}\mathbb{E}_t\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda))y^{\tau,z}_T(\lambda)y^{\tau,z}_T(\lambda)\mathrm{d}\lambda-\int_{\Lambda^2}h^\eta_{xx}(\lambda,\mu)y^{\tau,z}_T(\lambda)y^{\tau,z}_T(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\quad+o\left(\tau+\|z\|^2_{L^2(\Lambda)}\right).
\end{aligned}
\tag{3.68}
$$

Plugging in the initial condition

$$
\begin{aligned}
y^{\tau,z}_{t+\tau}&=z+\bar{x}_t-\bar{x}_{t+\tau}\\
&=z-\left(\int_t^{t+\tau}\Delta\bar{x}_r+b(\bar{x}_r,\bar{u}_r)\mathrm{d}r+\int_t^{t+\tau}\sigma(\bar{x}_r,\bar{u}_r)\mathrm{d}W_r\right),
\end{aligned}
\tag{3.69}
$$

we have

$$
\begin{aligned}
&V(t+\tau,\bar{x}_t+z)-V(t,\bar{x}_t)\\
&\le \mathbb{E}_t\left[-\int_t^{t+\tau}\int_\Lambda l(\bar{x}_r(\lambda),\bar{u}_r)\mathrm{d}\lambda\mathrm{d}r+\int_\Lambda p_{t+\tau}(\lambda)z(\lambda)\mathrm{d}\lambda\right]\\
&\quad-\mathbb{E}_t\left[\int_\Lambda p_{t+\tau}(\lambda)(\bar{x}_{t+\tau}-\bar{x}_t)(\lambda)\mathrm{d}\lambda\right]\\
&\quad+\frac{1}{2}\mathbb{E}_t\left[\int_{\Lambda^2}P^\eta_{t+\tau}(\lambda,\mu)z(\lambda)z(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\quad+\frac{1}{2}\mathbb{E}_t\left[\int_{\Lambda^2}P^\eta_{t+\tau}(\lambda,\mu)(\bar{x}_{t+\tau}-\bar{x}_t)(\lambda)(\bar{x}_{t+\tau}-\bar{x}_t)(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\quad-\mathbb{E}_t\left[\int_{\Lambda^2}P^\eta_{t+\tau}(\lambda,\mu)z(\lambda)(\bar{x}_{t+\tau}-\bar{x}_t)(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\quad+\frac{1}{2}\mathbb{E}_t\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda))y^{\tau,z}_T(\lambda)y^{\tau,z}_T(\lambda)\mathrm{d}\lambda-\int_{\Lambda^2}h^\eta_{xx}(\lambda,\mu)y^{\tau,z}_T(\lambda)y^{\tau,z}_T(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\quad+o\left(\tau+\|z\|^2_{L^2(\Lambda)}\right).
\end{aligned}
\tag{3.70}
$$

Now we apply Lemma 3.11, Lemma 3.12 and Lemma 3.13, and obtain

$$\begin{aligned}
&V(t+\tau,\bar{x}_t+z)-V(t,\bar{x}_t)\\
&\leq \tau\mathbb{E}_t\Bigg[-\langle p_t,\Delta\bar{x}_t+b(\bar{x}_t,\bar{u}_t)\rangle_{H_0^1(\Lambda)\times H^{-1}(\Lambda)}-\langle q_t,\sigma(\bar{x}_t,\bar{u}_t)\rangle_{L_2(\Xi,L^2(\Lambda))}\\
&\quad -\int_\Lambda l(\bar{x}_t(\lambda),\bar{u}_t)\mathrm{d}\lambda+\frac{1}{2}\int_{\Lambda^2}P_t^\eta(\lambda,\mu)\langle\sigma(\bar{x}_t(\lambda),\bar{u}_t),\sigma(\bar{x}_t(\mu),\bar{u}_t)\rangle_{L_2(\Xi,\mathbb{R})}\mathrm{d}\lambda\mathrm{d}\mu\Bigg]\\
&\quad +\mathbb{E}_t\left[\int_\Lambda p_{t+\tau}(\lambda)z(\lambda)\mathrm{d}\lambda+\frac{1}{2}\int_{\Lambda^2}P_{t+\tau}^\eta(\lambda,\mu)z(\lambda)z(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\quad +\frac{1}{2}\mathbb{E}_t\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda))y_T^{\tau,z}(\lambda)y_T^{\tau,z}(\lambda)\mathrm{d}\lambda-\int_{\Lambda^2}h_{xx}^\eta(\lambda,\mu)y_T^{\tau,z}(\lambda)y_T^{\tau,z}(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\quad +o\left(\tau+\|z\|_{L^2(\Lambda)}^2\right).
\end{aligned}\tag{3.71}$$

Adding a zero and rearranging terms yields

$$\begin{aligned}
&V(t+\tau,\bar{x}_t+z)-V(t,\bar{x}_t)-\left(-\langle\Delta\bar{x}_t,p_t\rangle_{H^{-1}(\Lambda)\times H_0^1(\Lambda)}-\mathcal{G}(t,\bar{x}_t,\bar{u}_t)\right)\tau\\
&-\int_\Lambda p_t(\lambda)z(\lambda)\mathrm{d}\lambda-\frac{1}{2}\int_{\Lambda^2}P_t(\lambda,\mu)z(\lambda)z(\mu)\mathrm{d}\lambda\mathrm{d}\mu\\
&\leq\frac{1}{2}\mathrm{tr}\left(\sigma(\bar{x}_t,\bar{u}_t)^*(P_t^\eta-P_t)\sigma(\bar{x}_t,\bar{u}_t)\right)\tau\\
&\quad +\mathbb{E}_t\left[\int_\Lambda(p_{t+\tau}(\lambda)-p_t(\lambda))z(\lambda)\mathrm{d}\lambda\right]\\
&\quad +\mathbb{E}_t\left[\frac{1}{2}\int_{\Lambda^2}(P_{t+\tau}^\eta(\lambda,\mu)-P_t(\lambda,\mu))z(\lambda)z(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\quad +\frac{1}{2}\mathbb{E}_t\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda))y_T^{\tau,z}(\lambda)y_T^{\tau,z}(\lambda)\mathrm{d}\lambda-\int_{\Lambda^2}h_{xx}^\eta(\lambda,\mu)y_T^{\tau,z}(\lambda)y_T^{\tau,z}(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\quad +o\left(\tau+\|z\|_{L^2(\Lambda)}^2\right).
\end{aligned}\tag{3.72}$$

Using Lemma 3.14 and elementary estimates for the right-hand side, we obtain

$$\begin{aligned}
&V(t+\tau,\bar{x}_t+z)-V(t,\bar{x}_t)-\left(-\langle\Delta\bar{x}_t,p_t\rangle_{H^{-1}(\Lambda)\times H_0^1(\Lambda)}-\mathcal{G}(t,\bar{x}_t,\bar{u}_t)\right)\tau\\
&-\int_\Lambda p_t(\lambda)z(\lambda)\mathrm{d}\lambda-\frac{1}{2}\int_{\Lambda^2}P_t(\lambda,\mu)z(\lambda)z(\mu)\mathrm{d}\lambda\mathrm{d}\mu\\
&\leq\frac{1}{2}\|P_t^\eta-P_t\|_{L^2(\Lambda^2)}\|\sigma(\bar{x}_t,\bar{u}_t)\|_{L_2(\Xi,L^2(\Lambda))}^2\tau+\frac{1}{2}\|P_{t+\tau}^\eta-P_t\|_{L^2(\Lambda^2)}\|z\|_{L^2(\Lambda)}^2\\
&\quad +\frac{1}{2}\mathbb{E}_t\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda))y_T^{\tau,z}(\lambda)y_T^{\tau,z}(\lambda)\mathrm{d}\lambda-\int_{\Lambda^2}h_{xx}^\eta(\lambda,\mu)y_T^{\tau,z}(\lambda)y_T^{\tau,z}(\mu)\mathrm{d}\lambda\mathrm{d}\mu\right]\\
&\quad +o\left(\tau+\|z\|_{L^2(\Lambda)}^2\right).
\end{aligned}\tag{3.73}$$

Now, let $(\tau_k, z_k) \to 0$, $\tau_k > 0$, be a sequence, which realizes the limit superior of the left-hand side divided by $\tau_k + \|z_k\|^2_{L^2(\Lambda)}$. By Lemma 3.8 and using the compact embedding $H_0^\gamma(\Lambda) \subset\subset L^2(\Lambda)$, $\gamma \in (0, 1/2)$ (see, e.g., [DD12, Theorem 4.54]), we can extract a subsequence – again denoted by (τ_k, z_k) – such that $y_T^{\tau_k,z_k}/\sqrt{\tau_k + \|z_k\|^2_{L^2(\Lambda)}}$ converges in $L^2(\Lambda)$ to some limit $\tilde{y}_T$. Therefore, dividing the inequality by $\tau_k + \|z_k\|^2_{L^2(\Lambda)}$ and sending (τ_k, z_k) to zero yields

$$\begin{aligned}
&\limsup_{\tau\downarrow 0, z\to 0} \frac{1}{\tau + \|z\|^2_{L^2(\Lambda)}} \Bigg\{ V(t+\tau, \bar{x}_t + z) - V(t, \bar{x}_t) - \big(-\langle \Delta \bar{x}_t, p_t\rangle_{L^2(\Lambda)} - \mathcal{G}(t, \bar{x}_t, \bar{u}_t)\big)\,\tau \\
&\quad - \int_\Lambda p_t(\lambda) z(\lambda) \mathrm{d}\lambda + \frac{1}{2}\int_{\Lambda^2} P_t(\lambda,\mu) z(\lambda) z(\mu) \mathrm{d}\lambda \mathrm{d}\mu \Bigg\} \\
&\le \frac{1}{2}\|P_t^\eta - P_t\|_{L^2(\Lambda^2)} \|\sigma(\bar{x}_t, \bar{u}_t)\|^2_{L_2(\Xi, L^2(\Lambda))} + \frac{1}{2}\|P_t^\eta - P_t\|_{L^2(\Lambda^2)} \\
&\quad + \frac{1}{2}\mathbb{E}_t\left[\int_\Lambda h_{xx}(\bar{x}_T(\lambda)) \tilde{y}_T(\lambda) \tilde{y}_T(\lambda) \mathrm{d}\lambda - \int_{\Lambda^2} h^\eta_{xx}(\lambda,\mu) \tilde{y}_T(\lambda) \tilde{y}_T(\mu) \mathrm{d}\lambda \mathrm{d}\mu\right].
\end{aligned} \tag{3.74}$$

Taking the limit $\eta \to 0$, the right-hand side vanishes, which concludes the proof of the first claim.

The second claim (3.10) follows along the same lines as in the finite-dimensional case with similar modifications as above. □

3.3. Space-Derivatives

In this section, we consider the case with differentials only in the spatial variable. To this end, we first recall the definition of viscosity super- and subdifferentials.

Definition 3.15. For $v \in C([s,T] \times L^2(\Lambda))$ the first order viscosity superdifferential in the space-variable of v at $(t,x) \in [s,T] \times L^2(\Lambda)$ is the set

$$D_x^{1,+} v(t,x) := \left\{ p \in L^2(\Lambda) \,\middle|\, \limsup_{z\to 0} \frac{v(t,x+z) - v(t,x) - \langle p, z\rangle_{L^2(\Lambda)}}{\|z\|_{L^2(\Lambda)}} \le 0 \right\}. \tag{3.75}$$

The first order viscosity subdifferential $D_x^{1,-} v$ is defined analogously with the lim sup replaced by lim inf and the $\le$ replaced by $\ge$.

Concerning the first order derivative, we obtain the following result.

Corollary 3.16. *It holds for almost every $t \in [s,T]$,*

$$D_x^{1,-} V(t, \bar{x}_t) \subset \{p_t\} \subset D_x^{1,+} V(t, \bar{x}_t) \tag{3.76}$$

$\mathbb{P}$–almost surely.

This follows from Theorem 3.4 by restricting the lim sup to $\tau = 0$ and estimating

$$|\langle z, P_t z\rangle_{L^2(\Lambda)}| \leq \|P_t\|_{L^2(\Lambda^2)} \|z\|^2_{L^2(\Lambda)}. \tag{3.77}$$

Next, we consider the second order viscosity differentials in the space-variable.

Definition 3.17. For $v \in C([s,T] \times L^2(\Lambda))$ the second order viscosity superdifferential in the space-variable of v at $(t,x) \in [s,T] \times L^2(\Lambda)$ is the set

$$D_x^{2,+} v(t,x) := \left\{ (p,P) \in L^2(\Lambda) \times \mathcal{S}(L^2(\Lambda)) \middle| \limsup_{z \to 0} \frac{v(t,x+z) - v(t,x) - \langle p, z\rangle_{L^2(\Lambda)} - \frac{1}{2}\langle z, Pz\rangle_{L^2(\Lambda)}}{\|z\|^2_{L^2(\Lambda)}} \leq 0 \right\}. \tag{3.78}$$

The second order viscosity subdifferential $D_x^{2,-} v$ is defined analogously with the lim sup replaced by lim inf and the $\leq$ replaced by $\geq$.

Corollary 3.18. *It holds for almost every $t \in [s,T]$,*

$$\{p_t\} \times \mathcal{S}_{\succeq P_t}(L^2(\Lambda)) \subset D_x^{2,+} V(t, \bar{x}_t), \tag{3.79}$$

$\mathbb{P}$–almost surely. Furthermore, it holds for almost every $t \in [s,T]$,

$$D_x^{2,-} V(t, \bar{x}_t) \subset \{p_t\} \times \mathcal{S}_{\preceq P_t}(L^2(\Lambda)). \tag{3.80}$$

$\mathbb{P}$–almost surely. Here $\mathcal{S}_{\succeq P_t}(L^2(\Lambda))$ and $\mathcal{S}_{\preceq P_t}(L^2(\Lambda))$ are defined as in Theorem 3.4.

The proof follows again from Theorem 3.4 by restricting the lim sup to $\tau = 0$.

3.4. Time-Derivatives

In this section, we consider the case with differentials only in the time-variable.

Definition 3.19. For $v \in C([s,T] \times L^2(\Lambda))$ the first order viscosity superdifferential in the time-variable of v at $(t,x) \in [s,T) \times L^2(\Lambda)$ is the set

$$D_{t+}^{1,+} v(t,x) := \left\{ G \in \mathbb{R} \middle| \limsup_{\tau \downarrow 0} \frac{v(t+\tau, x) - v(t,x) - G\tau}{\tau} \leq 0 \right\}. \tag{3.81}$$

The first order viscosity subdifferential $D_{t+}^{1,-} v$ is defined analogously with the lim sup replaced by lim inf and the $\leq$ replaced by $\geq$.

Corollary 3.20. *It holds for almost every $t \in [s,T]$,*

$$[-\langle \Delta \bar{x}_t, p_t\rangle_{H^{-1}(\Lambda) \times H_0^1(\Lambda)} - \mathcal{G}(t, \bar{x}_t, \bar{u}_t), \infty) \subset D_{t+}^{1,+} V(t, \bar{x}_t) \tag{3.82}$$

$\mathbb{P}$–almost surely.

The proof follows from Theorem 3.4 by restricting the lim sup to $z = 0$.

Remark 3.21. The results Theorem 3.4, Corollary 3.16, Corollary 3.18, and Corollary 3.20 extend the necessary condition in Peng's maximum principle by adjoint state inclusions.

3.5. Non-Positivity of the Correction Term

As another corollary of Theorem 3.4, we derive non-positivity of the correction term arising in non-smooth stochastic control problems.

Corollary 3.22. *Let*

$$\begin{cases} G \in L^2([s,T] \times \Omega; \mathbb{R}) \\ p \in L^2([s,T] \times \Omega; H_0^1(\Lambda)) \\ P \in L^2([s,T] \times \Omega; L_2(L^2(\Lambda))) \end{cases} \tag{3.83}$$

be adapted processes such that for almost every $t \in [s,T]$,

$$(G_t, p_t, P_t) \in D_{t+,x}^{1,2,+} V(t, \bar{x}_t) \tag{3.84}$$

$\mathbb{P}$*-almost surely. Then it holds for almost every $t \in [s,T]$,*

$$G_t + \langle \Delta \bar{x}_t, p_t \rangle_{H^{-1}(\Lambda) \times H_0^1(\Lambda)} + \mathcal{H}(\bar{x}_t, \bar{u}_t, p_t, P_t) \geq 0, \tag{3.85}$$

$\mathbb{P}$*–almost surely, where the Hamiltonian $\mathcal{H} : L^2(\Lambda) \times U \times L^2(\Lambda) \times L(L^2(\Lambda)) \to \mathbb{R}$ is given by*

$$\mathcal{H}(x, u, p, P) := \int_\Lambda l(x(\lambda), u) \mathrm{d}\lambda + \langle p, b(x,u) \rangle_{L^2(\Lambda)} + \frac{1}{2} tr\left(\sigma(x,u)^* P \sigma(x,u)\right). \tag{3.86}$$

Remark 3.23. The higher regularity assumptions on p and P in (3.83) are necessary due to the unbounded term in (3.85). Notice that the adjoint states given by (2.54) and (2.70), respectively, satisfy this higher regularity. In case p and P are the adjoint states and

$$G_t = -\langle \Delta \bar{x}_t, p_t \rangle_{H^{-1}(\Lambda) \times H_0^1(\Lambda)} - \mathcal{G}(t, \bar{x}_t, \bar{u}_t), \tag{3.87}$$

equation (3.85) is equivalent to

$$\mathcal{G}(t, \bar{x}_t, \bar{u}_t) \leq \mathcal{H}(\bar{x}_t, \bar{u}_t, p_t, P_t), \tag{3.88}$$

i.e.,

$$\mathrm{tr}(\sigma(x,u)^* [q_t - P_t \sigma(\bar{x}_t, \bar{u}_t)]) \leq 0. \tag{3.89}$$

The proof in the finite-dimensional case uses the following correspondence between test functions and points in the parabolic viscosity superdifferential, see [FS06, Chapter V, Lemma 4.1].

Proposition 3.24. *Let $v : (s,T) \times L^2(\Lambda) \to \mathbb{R}$ be upper semicontinuous. For $(t,x) \in (s,T) \times L^2(\Lambda)$, it holds $(G,p,P) \in D^{1,2,+}_{t+,x} v(t,x)$ if and only if there exists a function $\phi \in C^{1,2}((s,T) \times L^2(\Lambda))$ such that $v - \phi$ attains its strict global maximum over the set $[t,T) \times L^2(\Lambda)$ at the point (t,x), and*

$$(\phi(t,x), \partial_t \phi(t,x), D\phi(t,x), D^2\phi(t,x)) = (v(t,x), G, p, P). \tag{3.90}$$

However, due to the unbounded operator in the HJB equation, one has to restrict the class of admissible test functions and therefore loses the preceding equivalence. That is why we can't directly generalize the proof from the finite-dimensional case. Instead, in order to prove Corollary 3.22, we have to carry out the argument from the finite-dimensional case by hand. In addition to dealing with technical difficulties already arising in the proof of the verification theorem within the framework of viscosity solutions in finite dimensions (see [GŚZ05]), we have to perform a delicate regularity analysis due to the unbounded term in the state equation (2.1).

Now let us get to the proof of Corollary 3.22.

Proof. Fix $t \in [s,T]$ such that

$$(G_t, p_t, P_t) \in D^{1,2,+}_{t+,x} V(t, \bar{x}_t) \tag{3.91}$$

$\mathbb{P}$–almost surely. Following the idea from the finite-dimensional case (see [FS06, Chapter V, Lemma 4.1]), we define for $\beta > 0$,

$$g(\beta) := \sup \left\{ \frac{\left(V(t+\tau, \bar{x}_t + z) - V(t, \bar{x}_t) - G_t \tau - \langle p_t, z \rangle_{L^2(\Lambda)} - \frac{1}{2}\langle z, P_t z \rangle_{L^2(\Lambda)}\right)^+}{\left(\tau^2 + \|z\|^4_{L^2(\Lambda)}\right)^{\frac{1}{2}}} \,\middle|\, (t+\tau, z) \in (s,T) \times L^2(\Lambda), 0 < (\tau^2 + \|z\|^4_{L^2(\Lambda)})^{\frac{1}{2}} \leq \beta \right\}, \tag{3.92}$$

and set $g(0) := 0$. Since

$$\limsup_{\tau \downarrow 0, z \to 0} \frac{\tau + \|z\|^2_{L^2(\Lambda)}}{\left(\tau^2 + \|z\|^4_{L^2(\Lambda)}\right)^{\frac{1}{2}}} < \infty \tag{3.93}$$

and $(G_t, p_t, P_t) \in D^{1,2,+}_{t+,x} V(t, \bar{x}_t)$, g is continuous and non-decreasing on $[0,\infty)$. Using g, we define

$$\begin{aligned} a(\theta, x) &:= \left((\theta - t)^2 + \|x - \bar{x}_t\|^4_{L^2(\Lambda)}\right)^{\frac{1}{2}} \\ F(a) &:= \frac{2}{3a} \int_a^{2a} \int_\xi^{2\xi} g(\beta) \mathrm{d}\beta \mathrm{d}\xi, \quad F(0) = 0 \end{aligned} \tag{3.94}$$

and construct $\phi : (s, T) \times L^2(\Lambda) \to \mathbb{R}$ as

$$\begin{aligned} \phi(\theta, x) &:= F(a(\theta, x)) + V(t, \bar{x}_t) + G_t(\theta - t) \\ &\qquad + \langle p_t, x - \bar{x}_t \rangle_{L^2(\Lambda)} + \frac{1}{2} \langle x - \bar{x}_t, P_t(x - \bar{x}_t) \rangle_{L^2(\Lambda)}. \end{aligned} \tag{3.95}$$

In order to obtain higher regularity for the term $D\phi(r, \bar{x}_r)$, we replace the process P by an approximation. Let $(e_l)_{l \geq 1} \subset H_0^1(\Lambda)$ be an orthonormal basis of $L^2(\Lambda)$ and define $P^n x := \sum_{l=1}^n \langle Px, e_l \rangle_{L^2(\Lambda)} e_l$. Then we have for every $n \in \mathbb{N}$

- $P^n \in L^2([s, T] \times \Omega; L(L^2(\Lambda)))$, and $\|P_t^n\|_{L_2(L^2(\Lambda))} \leq \|P_t\|_{L_2(L^2(\Lambda))}$ $\mathrm{d}t \otimes \mathbb{P}$–almost everywhere;
- $P_t^n(H_0^1(\Lambda)) \subset H_0^1(\Lambda)$ $\mathrm{d}t \otimes \mathbb{P}$–almost everywhere;
- $P_t^n \to P_t$ in the uniform operator topology $\mathrm{d}t \otimes \mathbb{P}$–almost everywhere.

Note that these conditions also imply $P^n \to P$ in $L^2([s, T] \times \Omega; L(L^2(\Lambda)))$. We approximate ϕ by

$$\begin{aligned} \phi^n(\theta, x) &:= F(a(\theta, x)) + V(t, x_t^*) + G_t(\theta - t) \\ &\qquad + \langle p_t, x - x_t^* \rangle_{L^2(\Lambda)} + \frac{1}{2} \langle x - x_t^*, P_t^n(x - x_t^*) \rangle_{L^2(\Lambda)}. \end{aligned} \tag{3.96}$$

Since $V - \phi$ attains its maximum at $(t, \bar{x}_t)$ and by the dynamic programming principle, we have for every $\tau \geq 0$,

$$\begin{aligned} 0 &\geq \mathbb{E}_t \left[V(t + \tau, \bar{x}_{t+\tau}) - \phi(t + \tau, \bar{x}_{t+\tau}) - (V(t, \bar{x}_t) - \phi(t, \bar{x}_t)) \right] \\ &= \mathbb{E}_t \left[- \int_t^{t+\tau} \int_\Lambda l(\bar{x}_r(\lambda), \bar{u}_r) \mathrm{d}\lambda \mathrm{d}r - \phi^n(t + \tau, \bar{x}_{t+\tau}) + \phi^n(t, \bar{x}_t) \right] \\ &\quad + \mathbb{E}_t \left[\phi^n(t + \tau, \bar{x}_{t+\tau}) - \phi(t + \tau, \bar{x}_{t+\tau}) - \phi^n(t, \bar{x}_t) + \phi(t, \bar{x}_t) \right]. \end{aligned} \tag{3.97}$$

Now, we want to apply Itô's formula to ϕ^n. However, ϕ^n implicitly depends on ω via $\bar{x}_t$, G_t, p_t and P_t. Therefore, we fix an $\omega \in \Omega$ and switch to the probability space $(\Omega, \mathcal{F}^\nu, \mathbb{P}(\,\cdot\,|\mathcal{F}_{\nu,t}^s)(\omega))$, where $\mathbb{P}(\,\cdot\,|\mathcal{F}_{\nu,t}^s)(\cdot)$ denotes the regular conditional probability given $\mathcal{F}_{\nu,t}^s$. On this probability space, $\bar{x}_t$, G_t, p_t and P_t are almost surely constant, and are equal to $\bar{x}_t(\omega)$, $G_t(\omega)$, $p_t(\omega)$ and $P_t(\omega)$. See also [GŚZ05] for more details on this. In the following, we denote by $\mathbb{E}_t[\,\cdot\,](\omega)$ the expectation with respect to $\mathbb{P}(\,\cdot\,|\mathcal{F}_{\nu,t}^s)(\omega)$. Thus, we derive

$$\begin{aligned} &\mathbb{E}_t \left[\phi^n(t + \tau, \bar{x}_{t+\tau}) - \phi^n(t, \bar{x}_t) \right](\omega) \\ &= \mathbb{E}_t \Bigg[\int_t^{t+\tau} \partial_\theta \phi^n(r, \bar{x}_r) + \langle D\phi^n(r, \bar{x}_r), \Delta \bar{x}_r + b(\bar{x}_r, \bar{u}_r) \rangle_{H_0^1(\Lambda) \times H^{-1}(\Lambda)} \\ &\qquad + \frac{1}{2} \mathrm{tr} \left(\sigma(\bar{x}_r, \bar{u}_r)^* D^2 \phi^n(r, \bar{x}_r) \sigma(\bar{x}_r, \bar{u}_r) \right) \mathrm{d}r \Bigg](\omega). \end{aligned} \tag{3.98}$$

From Lemma 3.25 it follows

$$\begin{aligned}\mathbb{E}_t\Bigg[\frac{1}{\tau_k}\int_t^{t+\tau_k} &\partial_\theta\phi^n(r,\bar{x}_r)+\langle D\phi^n(r,\bar{x}_r),\Delta\bar{x}_r+b(\bar{x}_r,\bar{u}_r)\rangle_{H_0^1(\Lambda)\times H^{-1}(\Lambda)}\\&+\frac{1}{2}\mathrm{tr}\left(\sigma(\bar{x}_r,\bar{u}_r)^*D^2\phi^n(r,\bar{x}_r)\sigma(\bar{x}_r,\bar{u}_r)\right)\mathrm{d}r\Bigg](\omega)\\\to G_t(\omega)&+\langle p_t(\omega),\Delta\bar{x}_t(\omega)+b(\bar{x}_t(\omega),\bar{u}_t(\omega))\rangle\\&+\frac{1}{2}\mathrm{tr}(\sigma(\bar{x}_t(\omega),\bar{u}_t(\omega))^*P_t^n(\omega)\sigma(\bar{x}_t(\omega),\bar{u}_t(\omega))),\end{aligned}\tag{3.99}$$

$\mathrm{d}t\otimes\mathbb{P}$–almost everywhere.

For the last line in equation (3.97), we note that

$$\phi^n(\theta,x)-\phi(\theta,x)=\frac{1}{2}\langle x-\bar{x}_t,(P_t^n-P_t)(x-\bar{x}_t)\rangle_{L^2(\Lambda)}.\tag{3.100}$$

Therefore, $\mathbb{E}_t[\phi^n(t,\bar{x}_t)-\phi(t,\bar{x}_t)]$ vanishes, and

$$|\mathbb{E}_t\left[\phi^n(t+\tau_k,\bar{x}_{t+\tau_k})-\phi(t+\tau_k,\bar{x}_{t+\tau_k})\right]|\le\frac{1}{2}\|P_t^n-P_t\|^2_{L^2(\Lambda^2)}\mathbb{E}_t\left[\|\bar{x}_{t+\tau_k}-\bar{x}_t\|^2_{L^2(\Lambda)}\right].\tag{3.101}$$

Dividing by τ_k, taking the limit $k\to\infty$, and exploiting equation (3.16), we obtain

$$\lim_{k\to\infty}\frac{1}{\tau_k}|\mathbb{E}_t\left[\phi^n(t+\tau_k,\bar{x}_{t+\tau_k})-\phi(t+\tau_k,\bar{x}_{t+\tau_k})\right]|\le C\|P_t^n-P_t\|^2_{L^2(\Lambda^2)}.\tag{3.102}$$

Altogether, we derive from equation (3.97) for almost every $t\in[s,T]$,

$$\begin{aligned}0\ge&\lim_{k\to\infty}\frac{1}{\tau_k}\mathbb{E}_t\left[V(t+\tau_k,\bar{x}_{t+\tau_k})-\phi(t+\tau_k,\bar{x}_{t+\tau_k})-V(t,\bar{x}_t)+\phi(t,\bar{x}_t)\right]\\\ge&\mathrm{tr}\left(\sigma(\bar{x}_t,\bar{u}_t)^*\left[q_t-P_t^n\sigma(\bar{x}_t,\bar{u}_t)\right]\right)-C\|P_t^n-P_t\|^2_{L^2(\Lambda^2)}\end{aligned}\tag{3.103}$$

$\mathbb{P}$–almost surely. Taking the limit $n\to\infty$ concludes the proof. □

Lemma 3.25. *For almost every $t\in[s,T]$, it holds*

$$\begin{aligned}\mathbb{E}_t\Bigg[\frac{1}{\tau}\int_t^{t+\tau} &\partial_\theta\phi^n(r,\bar{x}_r)+\langle D\phi^n(r,\bar{x}_r),\Delta\bar{x}_r+b(\bar{x}_r,\bar{u}_r)\rangle_{H_0^1(\Lambda)\times H^{-1}(\Lambda)}\\&+\frac{1}{2}tr\left(\sigma(\bar{x}_r,\bar{u}_r)^*D^2\phi^n(r,\bar{x}_r)\sigma(\bar{x}_r,\bar{u}_r)\right)\mathrm{d}r\Bigg](\omega)\\\to G_t(\omega)&+\langle p_t(\omega),\Delta\bar{x}_t(\omega)+b(\bar{x}_t(\omega),\bar{u}_t(\omega))\rangle\\&+\frac{1}{2}tr(\sigma(\bar{x}_t(\omega),\bar{u}_t(\omega))^*P_t^n(\omega)\sigma(\bar{x}_t(\omega),\bar{u}_t(\omega))),\end{aligned}\tag{3.104}$$

as $\tau\to0$, $\mathbb{P}$–almost surely. Here, ϕ^n is given by (3.96).

Proof. We fix $t \in [s,T]$ such that (3.9) holds and the convergence in all the applications of Lebesgue's differentiation theorem below holds for t. Note that the set of such t is a set of full measure. Now, let us first discuss some properties of ϕ^n that can be proven similarly as in the finite-dimensional case, see [FS06, Chapter V, Lemma 4.1]. Recall the definition

$$\begin{aligned}\phi^n(\theta,x) := F(a(\theta,x)) + V(t,\bar{x}_t) + G_t(\theta-t)\\ + \langle p_t, x-\bar{x}_t\rangle_{L^2(\Lambda)} + \frac{1}{2}\langle x-\bar{x}_t, P_t^n(x-\bar{x}_t)\rangle_{L^2(\Lambda)}.\end{aligned} \tag{3.105}$$

where F and a are given by (3.94). The derivatives of F and a are given by

$$\begin{cases} F'(a) = \frac{4}{3a}\int_{2a}^{4a} g(\xi)\mathrm{d}\xi - \frac{2}{3a}\int_a^{2a} g(\xi)\mathrm{d}\xi - \frac{1}{a}F(a), \quad F'(0)=0 \\ F''(a) = \frac{2}{3a}(8g(4a)-6g(2a)+g(a)) - \frac{2}{a}F'(a), \quad F''(0)=0 \end{cases} \tag{3.106}$$

and

$$\begin{cases} \partial_\theta a(\theta,x) = \frac{\theta-t}{a(\theta,x)} \\ Da(\theta,x) = \frac{2\|x-\bar{x}_t\|^2_{L^2(\Lambda)}}{a(\theta,x)}(x-\bar{x}_t) \\ D^2a(\theta,x) = \left(\frac{4}{a(\theta,x)} - \frac{4\|x-\bar{x}_t\|^4_{L^2(\Lambda)}}{a(\theta,x)^3}\right)(x-\bar{x}_t)\otimes(x-\bar{x}_t) + \frac{2\|x-\bar{x}_t\|^2_{L^2(\Lambda)}}{a(\theta,x)}\langle\cdot,\cdot\rangle_{L^2(\Lambda)}. \end{cases} \tag{3.107}$$

The first and second derivative of ϕ^n are given by

$$\begin{cases} \partial_\theta\phi^n(\theta,x) = \frac{\theta-t}{a(\theta,x)}F'(a(\theta,x)) + G_t \\ D\phi^n(\theta,x) = F'(a(\theta,x))Da(\theta,x) + p_t + P_t^n(x-\bar{x}_t) \\ D^2\phi^n(\theta,x) = F''(a(\theta,x))Da(\theta,x)\otimes Da(\theta,x) + F'(a(\theta,x))D^2a(\theta,x) + P_t^n \end{cases} \tag{3.108}$$

Thus, $\phi^n \in C^{1,2}((s,T)\times L^2(\Lambda))$, and

$$(\phi^n(t,\bar{x}_t), \partial_\theta\phi^n(t,\bar{x}_t), D\phi^n(t,\bar{x}_t), D^2\phi^n(t,\bar{x}_t)) = (V(t,\bar{x}_t), G_t, p_t, P_t^n). \tag{3.109}$$

Furthermore, $F(a(\theta,x)) \in C^{1,2}([s,T]\times L^2(\Lambda))$ with vanishing time derivative and first and second order space derivative at $(\theta,x)=(t,\bar{x}_t)$. Furthermore, note that

$$|F'(a(\theta,x))| \le C\left(1+\|x\|^2_{L^2(\Lambda)}\right) \tag{3.110}$$

and

$$|F''(a(\theta,x))| \le C\left(1+\|x\|^2_{L^2(\Lambda)}\right). \tag{3.111}$$

Now let us start with the proof of (3.104). Let $(\tau_k)_{k\in\mathbb{N}}$ be an arbitrary sequence converging to zero. We fix $\omega\in\Omega$, and, as described in the discussion before (3.98), we switch to the probability space $(\Omega, \mathcal{F}^\nu, \mathbb{P}(\,\cdot\,|\mathcal{F}^s_{\nu,t})(\omega))$. First, we consider the term involving the time derivative of ϕ^n. We have

$$\partial_\theta \phi^n(r, \bar{x}_r) = \frac{r-t}{a(r,\bar{x}_r)} F'(a(r,\bar{x}_r)) + G_t. \tag{3.112}$$

By the almost sure continuity of

$$r \mapsto F'(a(r,\bar{x}_r)) \frac{r-t}{a(r,\bar{x}_r)}, \tag{3.113}$$

and the dominated convergence theorem, we have

$$\lim_{k\to\infty} \mathbb{E}_t \left[\left| \frac{1}{\tau_k} \int_t^{t+\tau_k} F'(a(r,\bar{x}_r)) \frac{r-t}{a(r,\bar{x}_r)} \mathrm{d}r \right| \right](\omega) = 0. \tag{3.114}$$

Now, let us consider the first space derivative of ϕ^n. We have

$$D\phi^n(r,\bar{x}_r) = F'(a(r,\bar{x}_r)) Da(r,\bar{x}_r) + p_t + P_t^n(\bar{x}_r - \bar{x}_t), \tag{3.115}$$

thus we have to show

$$\lim_{k\to\infty} \mathbb{E}_t \left[\left| \frac{1}{\tau_k} \int_t^{t+\tau_k} \langle F'(a(r,\bar{x}_r)) Da(r,\bar{x}_r), \Delta \bar{x}_r + b(\bar{x}_r, \bar{u}_r) \rangle_{H_0^1(\Lambda)\times H^{-1}(\Lambda)} \mathrm{d}r \right| \right](\omega) = 0, \tag{3.116}$$

and

$$\lim_{k\to\infty} \mathbb{E}_t \left[\left| \frac{1}{\tau_k} \int_t^{t+\tau_k} \langle p_t, \Delta(\bar{x}_r - \bar{x}_t) + b(\bar{x}_r, \bar{u}_r) - b(\bar{x}_t, \bar{u}_t) \rangle_{H_0^1(\Lambda)\times H^{-1}(\Lambda)} \mathrm{d}r \right| \right](\omega) = 0, \tag{3.117}$$

and

$$\lim_{k\to\infty} \mathbb{E}_t \left[\left| \frac{1}{\tau_k} \int_t^{t+\tau_k} \langle P_t^n(\bar{x}_r - \bar{x}_t), \Delta \bar{x}_r + b(\bar{x}_r, \bar{u}_r) \rangle_{H_0^1(\Lambda)\times H^{-1}(\Lambda)} \mathrm{d}r \right| \right](\omega) = 0. \tag{3.118}$$

We only consider the terms involving $\Delta\bar{x}_r$; the terms involving $b(\bar{x}_r, \bar{u}_r)$ can be handled similarly.

Let us start with (3.116). Using (3.110) and the bound

$$\|Da(r,\bar{x}_r)\|_{H_0^1(\Lambda)} \le 2\|\bar{x}_r - \bar{x}_t\|_{H_0^1(\Lambda)}, \tag{3.119}$$

we obtain

$$\begin{aligned} &\mathbb{E}_t \left[\left| \frac{1}{\tau_k} \int_t^{t+\tau_k} \langle F'(a(r,\bar{x}_r)) Da(r,\bar{x}_r), \Delta\bar{x}_r \rangle_{H_0^1(\Lambda)\times H^{-1}(\Lambda)} \mathrm{d}r \right| \right](\omega) \\ &\le \mathbb{E}_t \left[\frac{1}{\tau_k} \int_t^{t+\tau_k} C(1 + \|\bar{x}_r\|_{L^2(\Lambda)}^2) \|\bar{x}_r - \bar{x}_t\|_{H_0^1(\Lambda)} \|\bar{x}_r\|_{H_0^1(\Lambda)} \mathrm{d}r \right](\omega). \end{aligned} \tag{3.120}$$

Using Hölder's inequality, we obtain

$$\begin{aligned}&\mathbb{E}_t\left[\frac{1}{\tau_k}\int_t^{t+\tau_k}\|\bar{x}_r\|^2_{L^2(\Lambda)}\|\bar{x}_r-\bar{x}_t\|_{H^1_0(\Lambda)}\|\bar{x}_r\|_{H^1_0(\Lambda)}\mathrm{d}r\right](\omega)\\&\leq\left(\frac{1}{\tau_k}\int_t^{t+\tau_k}\mathbb{E}_t\left[\|\bar{x}_r-\bar{x}_t\|^2_{H^1_0(\Lambda)}\right](\omega)\mathrm{d}r\right)^{\frac{1}{2}}\left(\frac{1}{\tau_k}\int_t^{t+\tau_k}\mathbb{E}_t\left[\|\bar{x}_r\|^4_{L^2(\Lambda)}\|\bar{x}_r\|^2_{H^1_0(\Lambda)}\right](\omega)\mathrm{d}r\right)^{\frac{1}{2}}.\end{aligned}\tag{3.121}$$

Since $\bar{x}\in L^2([s,T]\times\Omega;H^1_0(\Lambda))$,

$$\lim_{k\to\infty}\frac{1}{\tau_k}\int_t^{t+\tau_k}\mathbb{E}\left[\|\bar{x}_r-\bar{x}_t\|^2_{H^1_0(\Lambda)}\right]\mathrm{d}r=0.\tag{3.122}$$

Thus, the first factor of (3.121) converges to zero $\mathbb{P}$-almost surely along some subsequence. Since

$$\|\bar{x}\|^4_{L^2(\Lambda)}\|\bar{x}\|^2_{H^1_0(\Lambda)}\in L^1([s,T]\times\Omega),\tag{3.123}$$

(see [LR15, Lemma 5.1.5]), the second factor is finite in the limit $k\to\infty$ along some subsequence.

Now, let us consider (3.117). We have

$$\begin{aligned}&\mathbb{E}_t\left[\left|\frac{1}{\tau_k}\int_t^{t+\tau_k}\langle p_t,\Delta(\bar{x}_r-\bar{x}_t)\rangle_{H^1_0(\Lambda)\times H^{-1}(\Lambda)}\mathrm{d}r\right|\right](\omega)\\&\leq\|p_t\|_{H^1_0(\Lambda)}\left(\frac{1}{\tau_k}\int_t^{t+\tau_k}\mathbb{E}_t\left[\|\bar{x}_r-\bar{x}_t\|^2_{H^1_0(\Lambda)}\right](\omega)\mathrm{d}r\right)^{\frac{1}{2}}.\end{aligned}\tag{3.124}$$

The first factor is finite for almost every t, and the second factor again converges to zero along some subsequence.

Finally, for (3.118), we obtain using Hölder's inequality

$$\begin{aligned}&\mathbb{E}_t\left[\left|\frac{1}{\tau_k}\int_t^{t+\tau_k}\langle P^n_t(\bar{x}_r-\bar{x}_t),\Delta\bar{x}_r\rangle_{H^1_0(\Lambda)\times H^{-1}(\Lambda)}\mathrm{d}r\right|\right](\omega)\\&\leq\|P^n_t\|_{H^1_0(\Lambda^2)}\left(\frac{1}{\tau_k}\int_t^{t+\tau_k}\mathbb{E}_t\left[\|\bar{x}_r-\bar{x}_t\|^2_{L^2(\Lambda)}\right](\omega)\mathrm{d}r\right)^{\frac{1}{2}}\left(\frac{1}{\tau_k}\int_t^{t+\tau_k}\mathbb{E}_t\left[\|\bar{x}_r\|^2_{H^1_0(\Lambda)}\right](\omega)\mathrm{d}r\right)^{\frac{1}{2}}.\end{aligned}\tag{3.125}$$

The second factor converges to zero and the third factor is finite along some subsequence, which shows that the right-hand side of equation (3.125) converges to zero.

Now, let us consider the second space derivative of ϕ^n. We have

$$D^2\phi^n(r,\bar{x}_r)=F''(a(r,\bar{x}_r))Da(r,\bar{x}_r)\otimes Da(r,\bar{x}_r)+F'(a(r,\bar{x}_r))D^2a(r,\bar{x}_r)+P^n_t,\tag{3.126}$$

thus we have to show

$$\lim_{k\to\infty} \mathbb{E}_t \left[\left| \frac{1}{\tau_k} \int_t^{t+\tau_k} \operatorname{tr}\left(\sigma(\bar{x}_r, \bar{u}_r)^* F''(a(r, \bar{x}_r)) Da(r, \bar{x}_r) \otimes Da(r, \bar{x}_r) \sigma(\bar{x}_r, \bar{u}_r)\right) \mathrm{d}r \right| \right] (\omega) = 0 \tag{3.127}$$

and

$$\lim_{k\to\infty} \mathbb{E}_t \left[\left| \frac{1}{\tau_k} \int_t^{t+\tau_k} \operatorname{tr}\left(\sigma(\bar{x}_r, \bar{u}_r)^* F'(a(r, \bar{x}_r)) D^2 a(r, \bar{x}_r) \sigma(\bar{x}_r, \bar{u}_r)\right) \mathrm{d}r \right| \right] (\omega) = 0 \tag{3.128}$$

and

$$\lim_{k\to\infty} \mathbb{E}_t \left[\left| \frac{1}{\tau_k} \int_t^{t+\tau_k} \operatorname{tr}\left((\sigma(\bar{x}_r, \bar{u}_r)\sigma(\bar{x}_r, \bar{u}_r)^* - \sigma(\bar{x}_t, \bar{u}_t)\sigma(\bar{x}_t, \bar{u}_t)^*) P_t^n\right) \mathrm{d}r \right| \right] (\omega) = 0. \tag{3.129}$$

For the first term, we have

$$\begin{aligned}
&\mathbb{E}_t \left[\left| \frac{1}{\tau_k} \int_t^{t+\tau_k} \operatorname{tr}\left(\sigma(\bar{x}_r, \bar{u}_r)^* F''(a(r, \bar{x}_r)) Da(r, \bar{x}_r) \otimes Da(r, \bar{x}_r) \sigma(\bar{x}_r, \bar{u}_r)\right) \mathrm{d}r \right| \right] (\omega) \\
&\le \left(\frac{1}{\tau_k} \int_t^{t+\tau_k} \mathbb{E}_t \left[\|\sigma(\bar{x}_r, \bar{u}_r)\|^2_{L_2(\Xi, L^2(\Lambda))} \right] (\omega) \mathrm{d}r \right)^{\frac{1}{2}} \\
&\qquad \left(\frac{1}{\tau_k} \int_t^{t+\tau_k} \mathbb{E}_t \left[\|F''(a(r, \bar{x}_r)) Da(r, \bar{x}_r) \otimes Da(r, \bar{x}_r)\|^2_{L(L^2(\Lambda))} \right] (\omega) \mathrm{d}r \right)^{\frac{1}{2}}.
\end{aligned} \tag{3.130}$$

Since $\sigma(\bar{x}, \bar{u}) \in L^2([s, T] \times \Omega; L_2(\Xi, L^2(\Lambda)))$, the first factor is finite along some subsequence, and using continuity of the second derivative of F and the first derivative of a, we obtain

$$\lim_{k\to\infty} \frac{1}{\tau_k} \int_t^{t+\tau_k} \mathbb{E}_t \left[\|F''(a(r, \bar{x}_r)) Da(r, \bar{x}_r) \otimes Da(r, \bar{x}_r)\|^2_{L(L^2(\Lambda))} \right] (\omega) \mathrm{d}r = 0. \tag{3.131}$$

Similar arguments can be employed to prove (3.128). Finally, for (3.129), we have

$$\begin{aligned}
&\mathbb{E}_t \left[\left| \frac{1}{\tau_k} \int_t^{t+\tau_k} \operatorname{tr}\left((\sigma(\bar{x}_r, \bar{u}_r)\sigma(\bar{x}_r, \bar{u}_r)^* - \sigma(\bar{x}_t, \bar{u}_t)\sigma(\bar{x}_t, \bar{u}_t)^*) P_t^n\right) \mathrm{d}r \right| \right] (\omega) \\
&\le \|P_t^n\|_{L(L^2(\Lambda))} \frac{1}{\tau_k} \int_t^{t+\tau_k} \mathbb{E}_t \left[\|\sigma(\bar{x}_r, \bar{u}_r)\sigma(\bar{x}_r, \bar{u}_r)^* - \sigma(\bar{x}_t, \bar{u}_t)\sigma(\bar{x}_t, \bar{u}_t)^*\|_{L_1(L^2(\Lambda))} \right] (\omega) \mathrm{d}r.
\end{aligned} \tag{3.132}$$

The first factor is finite for almost every t, and the second factor converges to zero along some subsequence since $\sigma(x, u) \in L^2([s, T] \times \Omega; L_2(\Xi, L^2(\Lambda)))$. □

Part II.

Sufficient Optimality Conditions

4. Verification Theorem

Part II of this dissertation is devoted to sufficient optimality conditions. Our results do not rely on the coefficients being of Nemytskii-type. Therefore, we introduce a more general setting. First, we prove a sufficient condition in terms of the value function. Afterwards, we combine this result with a well-known result that identifies the value function as the unique B-continuous viscosity solution of the HJB equation yielding a stochastic verification theorem for controlled semilinear SPDEs in the framework of viscosity solutions. This chapter is based on [SW22].

4.1. Introduction

Consider the following control problem: Minimize

$$J(s,x;u) := \mathbb{E}\left[\int_s^T L(x_t^u, u_t)\mathrm{d}t + H(x_T^u)\right] \tag{4.1}$$

over $u \in \mathcal{U}_s$ subject to

$$\begin{cases} \mathrm{d}x_t^u = [\Delta x_t^u + B(x_t^u, u_t)]\mathrm{d}t + \Sigma(x_t^u, u_t)\mathrm{d}W_t, \quad t \in [s,T] \\ x_s^u = x \in L^2(\Lambda). \end{cases} \tag{4.2}$$

Here, Assumption 3.1 is still in place and we impose the following assumptions on the coefficients of the control problem.

Assumption 4.1.(B1)$''$ *Let $L : L^2(\Lambda) \times U \to \mathbb{R}$ be continuous and satisfy*

$$|L(x,u)| \leq C\left(1 + \|x\|^2_{L^2(\Lambda)}\right) \tag{4.3}$$

for all $x \in L^2(\Lambda)$ and all $u \in U$.

(B2)$''$ *Let $H : L^2(\Lambda) \to \mathbb{R}$ be continuous and satisfy*

$$|H(x)| \leq C\left(1 + \|x\|^2_{L^2(\Lambda)}\right) \tag{4.4}$$

for all $x \in L^2(\Lambda)$.

(B3)$''$ *Let $B : L^2(\Lambda) \times U \to L^2(\Lambda)$ satisfy*

$$\begin{cases} \|B(x,u) - B(y,u)\|_{L^2(\Lambda)} \leq C\|x-y\|_{L^2(\Lambda)} \\ \|B(x,u)\|_{L^2(\Lambda)} \leq C\left(1 + \|x\|_{L^2(\Lambda)}\right) \end{cases} \tag{4.5}$$

for all $x, y \in L^2(\Lambda)$ *and all* $u \in U$.

(B4)″ *Let* $\Sigma : L^2(\Lambda) \times U \to L_2(\Xi, L^2(\Lambda))$ *satisfy*

$$\begin{cases} \|\Sigma(x,u) - \Sigma(y,u)\|_{L_2(\Xi,L^2(\Lambda))} \leq C\|x-y\|_{L^2(\Lambda)} \\ \|\Sigma(x,u)\|_{L_2(\Xi,L^2(\Lambda))} \leq C\left(1 + \|x\|_{L^2(\Lambda)}\right), \end{cases} \tag{4.6}$$

for all $x, y \in L^2(\Lambda)$ *and all* $u \in U$, *as well as*

$$\|\Sigma(x,u)\|_{L_2(\Xi,H_0^1(\Lambda))} \leq C\left(1 + \|x\|_{H_0^1(\Lambda)}\right) \tag{4.7}$$

for all $x \in H_0^1(\Lambda)$ *and all* $u \in U$.

In this setting, the Hamiltonian $\mathcal{H} : L^2(\Lambda) \times U \times L^2(\Lambda) \times L(L^2(\Lambda)) \to \mathbb{R}$ is given by

$$\mathcal{H}(x,u,p,P) := L(x,u) + \langle p, B(x,u)\rangle_{L^2(\Lambda)} + \frac{1}{2}\mathrm{tr}\left(\Sigma(x,u)^* P \Sigma(x,u)\right). \tag{4.8}$$

Remark 4.2. In this chapter, u^* denotes an arbitrary admissible control and x^* denotes the corresponding controlled state.

4.2. Verification Theorem

Theorem 4.3. *Assume there exists a constant* $C > 0$ *such that for every* $t \in [s,T)$, $\tau \in [0, T-t]$, *and* $x \in H_0^1(\Lambda)$,

$$V(t+\tau, x) - V(t,x) \leq C\left(1 + \|x\|^2_{H_0^1(\Lambda)}\right)\tau, \tag{4.9}$$

and let V *be semiconcave uniformly in* t, *i.e., there exists a constant* $C \geq 0$ *such that for every* $t \in (s,T]$ *it holds*

$$V(t,\cdot) - C\|\cdot\|^2_{L^2(\Lambda)} \tag{4.10}$$

is concave on $L^2(\Lambda)$. *Suppose further that there are adapted processes*

$$\begin{cases} G \in L^2([s,T]\times\Omega;\mathbb{R}) \\ p \in L^2([s,T]\times\Omega;H_0^1(\Lambda)) \\ P \in L^2([s,T]\times\Omega;L_2(L^2(\Lambda))) \end{cases} \tag{4.11}$$

such that for almost every $t \in [s,T]$,

$$(G_t, p_t, P_t) \in D^{1,2,+}_{t+,x} V(t, x_t^*) \tag{4.12}$$

$\mathbb{P}$-*almost surely, and*

$$\mathbb{E}\left[\int_s^T G_t + \langle \Delta x_t^*, p_t\rangle_{H^{-1}(\Lambda)\times H_0^1(\Lambda)} + \mathcal{H}(x_t^*, u_t^*, p_t, P_t)\mathrm{d}t\right] \leq 0. \tag{4.13}$$

Then (x^*, u^*) *is an optimal pair.*

Remark 4.4. (i) Together with Corollary 3.22, this result implies

$$G_t + \langle \Delta x_t^*, p_t \rangle_{H^{-1}(\Lambda) \times H_0^1(\Lambda)} + \mathcal{H}(x_t^*, u_t^*, p_t, P_t) = 0 \tag{4.14}$$

$\mathrm{d}t \otimes \mathbb{P}$-almost everywhere.

(ii) Conditions under which the growth condition (4.9) and the semiconcavity (4.10) hold are given in the subsequent Proposition 4.6 and Proposition 4.7.

In the proof of Theorem 4.3, we need the following lemma.

Lemma 4.5. *Let* V *satisfy* (4.9) *and* (4.10)*. Then, for almost every* $t \in [s, T)$*, there is a function* $\rho_1 \in L^1(\Omega)$ *such that for every* $\tau \in (0, T - t]$*,*

$$\frac{1}{\tau} \mathbb{E}_t \left[V(t + \tau, x_{t+\tau}^*) - V(t, x_t^*) \right] \leq \rho_1(\omega) \tag{4.15}$$

$\mathbb{P}$*-almost surely. Furthermore, there is a function* $\rho_2 \in L^1(s, T)$ *such that for almost every* $t \in [s, T]$ *and for every* $\tau \in (0, T - t]$*,*

$$\frac{1}{\tau} \mathbb{E} \left[V(t + \tau, x_{t+\tau}^*) - V(t, x_t^*) \right] \leq \rho_2(t). \tag{4.16}$$

Proof. The proof follows along the same lines as in the finite-dimensional case, see [GŚZ10]. In order to obtain Lipschitz continuity of the state trajectories in $H^{-1}(\Lambda)$ we rely on the analyticity of the heat semigroup.

First, we split up the increment into separate space and time increments:

$$V(t+\tau, x_{t+\tau}^*) - V(t, x_t^*) = V(t, x_{t+\tau}^*) - V(t, x_t^*) + V(t+\tau, x_{t+\tau}^*) - V(t, x_{t+\tau}^*). \tag{4.17}$$

For the space increment, we observe that by the semiconcavity of $V(t, \cdot)$, we have for $(G_t, p_t, P_t) \in D_{t+,x}^{1,2,+} V(t, x_t^*)$,

$$V(t, x_{t+\tau}^*) - V(t, x_t^*) \leq \langle p_t, x_{t+\tau}^* - x_t^* \rangle_{L^2(\Lambda)} + C \| x_{t+\tau}^* - x_t^* \|_{L^2(\Lambda)}^2. \tag{4.18}$$

For the time increment, we apply the growth condition (4.9). Therefore, we obtain altogether

$$\begin{aligned} &V(t + \tau, x_{t+\tau}^*) - V(t, x_t^*) \\ &\leq \langle p_t, x_{t+\tau}^* - x_t^* \rangle_{L^2(\Lambda)} + C \| x_{t+\tau}^* - x_t^* \|_{L^2(\Lambda)}^2 + C \left(1 + \| x_{t+\tau}^* \|_{H_0^1(\Lambda)}^2 \right) \tau. \end{aligned} \tag{4.19}$$

For the first term, we have

$$\mathbb{E} \left[\langle p_t, x_{t+\tau}^* - x_t^* \rangle_{L^2(\Lambda)} \right] \leq \mathbb{E} \left[\| p_t \|_{H_0^1(\Lambda)}^2 \right]^{\frac{1}{2}} \mathbb{E} \left[\left\| \mathbb{E}_t \left[x_{t+\tau}^* - x_t^* \right] \right\|_{H^{-1}(\Lambda)}^2 \right]^{\frac{1}{2}}. \tag{4.20}$$

For the second factor, we first note

$$\mathbb{E}_t\left[x^*_{t+\tau} - x^*_t\right] = (S_\tau - I)x^*_t + \mathbb{E}_t\left[\int_t^{t+\tau} S_{t+\tau-r}B(x^*_r, u^*_r)\mathrm{d}r\right], \tag{4.21}$$

where $(S_r)_{r\geq 0}$ denotes the heat semigroup. Since the heat semigroup is analytic and 0 is in the resolvent set of the Laplace operator with Dirichlet boundary conditions, using [Paz83, Chapter 2, Theorem 6.13] we obtain for the first term

$$\|(S_\tau - I)x^*_t\|_{H^{-1}(\Lambda)} = \|(S_\tau - I)\Delta^{-\frac{1}{2}}x^*_t\|_{L^2(\Lambda)} \leq C\tau\|x^*_t\|_{H^1_0(\Lambda)}. \tag{4.22}$$

For the second term in (4.21), we have by the linear growth assumption (4.5) on B

$$\begin{aligned}\mathbb{E}_t\left[\int_t^{t+\tau} \|S_{t+\tau-r}B(x^*_r, u^*_r)\|_{H^{-1}(\Lambda)}\,\mathrm{d}r\right] &\leq C\tau \sup_{r\in[t,t+\tau]} \mathbb{E}_t\left[\|B(x^*_r, u^*_r)\|_{L^2(\Lambda)}\right] \\ &\leq C\left(1 + \|x^*_t\|_{L^2(\Lambda)}\right)\tau.\end{aligned} \tag{4.23}$$

Thus, we obtain

$$\left\|\mathbb{E}_t\left[x^*_{t+\tau} - x^*_t\right]\right\|_{L^2(\Lambda)} \leq C\left(1 + \|x^*_t\|_{H^1_0(\Lambda)}\right)\tau, \tag{4.24}$$

and therefore together with (4.20),

$$\mathbb{E}\left[\langle p_t, x^*_{t+\tau} - x^*_t\rangle_{L^2(\Lambda)}\right] \leq C\left(1 + \mathbb{E}\left[\|x^*_t\|^2_{H^1_0(\Lambda)} + \|p_t\|^2_{H^1_0(\Lambda)}\right]\right)\tau. \tag{4.25}$$

For the second term in (4.19), we obatin using standard regularity arguments for solutions of SPDEs

$$\mathbb{E}\left[\|x^*_{t+\tau} - x^*_t\|^2_{L^2(\Lambda)}\right] \leq C\tau. \tag{4.26}$$

For the last term in (4.19), we first observe

$$\begin{aligned}\|x^*_{t+\tau}\|^2_{H^1_0(\Lambda)} &= \|x^*_t\|^2_{H^1_0(\Lambda)} + \int_t^{t+\tau} \langle \Delta x^*_r + B(x^*_r, u^*_r), x^*_r\rangle_{H^1_0(\Lambda)}\mathrm{d}r \\ &\quad + \int_t^{t+\tau} \|\Sigma(x^*_r, u^*_r)\|^2_{L_2(\Xi, H^1_0(\Lambda))}\mathrm{d}r + \int_t^{t+\tau} \langle x^*_r, \Sigma(x^*_r, u^*_r)\mathrm{d}W_r\rangle_{H^1_0(\Lambda)}.\end{aligned} \tag{4.27}$$

Therefore,

$$\begin{aligned}&\mathbb{E}\left[\|x^*_{t+\tau}\|^2_{H^1_0(\Lambda)}\right] \\ &\leq \mathbb{E}\left[\|x^*_t\|^2_{H^1_0(\Lambda)}\right] + \int_t^{t+\tau} \mathbb{E}\left[\|B(x^*_r, u^*_r)\|^2_{L^2(\Lambda)} + \|\Sigma(x^*_r, u^*_r)\|^2_{L_2(\Xi, H^1_0(\Lambda))}\right]\mathrm{d}r.\end{aligned} \tag{4.28}$$

Using the growth condition on B and Σ and applying Grönwall's inequality, we obtain for the last term in (4.19)

$$\mathbb{E}\left[C\left(1+\|x_{t+\tau}^*\|_{L^2(\Lambda)}^2\right)\tau\right] \leq C\left(1+\mathbb{E}\left[\|x_t^*\|_{H_0^1(\Lambda)}^2\right]\right)\tau. \tag{4.29}$$

Thus, we have the bound

$$\frac{1}{\tau}\mathbb{E}\left[V(t+\tau,x_{t+\tau}^*)-V(t,x_t^*)\right] \leq C\left(1+\mathbb{E}\left[\|x_t^*\|_{H_0^1(\Lambda)}^2+\|p_t\|_{H_0^1(\Lambda)}^2\right]\right), \tag{4.30}$$

where the right-hand side is in $L^1(s,T)$. This proves (4.16). The proof of (4.15) is similar. □

Now, let us prove Theorem 4.3.

Proof. Using Proposition 3.24, we obtain for given $(t,\omega) \in [s,T]\times\Omega$ a function $\phi \in C^{1,2}((s,T)\times L^2(\Lambda))$ such that $V-\phi$ attains its strict global maximum over $[t,T)\times L^2(\Lambda)$ at the point $(t,x_t^*(\omega))$, and

$$(\phi(t,x_t^*),\partial_t\phi(t,x_t^*),D\phi(t,x_t^*),D^2\phi(t,x_t^*)) = (V(t,x_t^*),G_t,p_t,P_t). \tag{4.31}$$

For ϕ and ϕ^n, we use the same construction as in the proof of Corollary 3.22.

For $t>T$, we set $V(t,x_t^*) := V(T,x_T^*)$. Then we have

$$\begin{aligned}
&\mathbb{E}\left[V(T,x_T^*)\right]-\mathbb{E}\left[V(s,x_s^*)\right] \\
&= \lim_{\tau\downarrow 0}\frac{1}{\tau}\left(\int_T^{T+\tau}\mathbb{E}\left[V(t,x_t^*)\right]\mathrm{d}t-\int_s^{s+\tau}\mathbb{E}\left[V(t,x_t^*)\right]\mathrm{d}t\right) \\
&= \lim_{\tau\downarrow 0}\frac{1}{\tau}\left(\int_{s+\tau}^{T+\tau}\mathbb{E}\left[V(t,x_t^*)\right]\mathrm{d}t-\int_s^{T}\mathbb{E}\left[V(t,x_t^*)\right]\mathrm{d}t\right) \\
&= \lim_{\tau\downarrow 0}\int_s^T\frac{1}{\tau}\mathbb{E}\left[V(t+\tau,x_{t+\tau}^*)-V(t,x_t^*)\right]\mathrm{d}t.
\end{aligned} \tag{4.32}$$

By Lemma 4.5, we can apply Fatou's lemma to obtain

$$\mathbb{E}\left[V(T,x_T^*)\right]-\mathbb{E}\left[V(s,x_s^*)\right] \leq \int_s^T\limsup_{\tau\downarrow 0}\left(\frac{1}{\tau}\mathbb{E}\left[V(t+\tau,x_{t+\tau}^*)-V(t,x_t^*)\right]\right)\mathrm{d}t. \tag{4.33}$$

Since $V-\phi$ attains its maximum at (t,x_t^*), we have

$$\begin{aligned}
&\mathbb{E}_t\left[V(t+\tau,x_{t+\tau}^*)-V(t,x_t^*)\right] \\
&\leq \mathbb{E}_t\left[\phi^n(t+\tau,x_{t+\tau}^*)-\phi^n(t,x_t^*)\right] \\
&\quad+\mathbb{E}_t\left[\phi(t+\tau,x_{t+\tau}^*)-\phi^n(t+\tau,x_{t+\tau}^*)+\phi^n(t,x_t^*)-\phi(t,x_t^*)\right].
\end{aligned} \tag{4.34}$$

For the last line, we first observe

$$\phi^n(\theta,x)-\phi(\theta,x) = \frac{1}{2}\langle x-x_t^*,(P_t^n-P_t)(x-x_t^*)\rangle_{L^2(\Lambda)}, \tag{4.35}$$

hence $\phi^n(t,x_t^*) - \phi(t,x_t^*)$ vanishes. Furthermore

$$\begin{aligned}
&\left|\mathbb{E}_t\left[\phi(t+\tau,x_{t+\tau}^*) - \phi^n(t+\tau,x_{t+\tau}^*)\right]\right| \\
&\le \frac{1}{2}\mathbb{E}_t\left[|\langle x_{t+\tau}^* - x_t^*, (P_t - P_t^n)(x_{t+\tau}^* - x_t^*)\rangle_{L^2(\Lambda)}|\right] \\
&\le \frac{1}{2}\mathbb{E}_t\left[\|x_{t+\tau}^* - x_t^*\|_{L^2(\Lambda)}^4\right]^{\frac{1}{2}} \|P_t - P_t^n\|_{L(L^2(\Lambda))} \\
&\le C\tau\|P_t - P_t^n\|_{L(L^2(\Lambda))}.
\end{aligned} \tag{4.36}$$

As in the proof of Lemma 3.25, we fix $\omega \in \Omega$ and apply Itô's formula to ϕ^n on the probability space $(\Omega, \mathcal{F}^\nu, \mathbb{P}(\,\cdot\,|\mathcal{F}_{\nu,t}^s)(\omega))$. This yields for every $t \in (s,T)$ and every $\tau \ge 0$ such that $t+\tau \le T$,

$$\begin{aligned}
&\mathbb{E}_t\left[\phi^n(t+\tau,x_{t+\tau}^*) - \phi^n(t,x_t^*)\right](\omega) \\
&= \mathbb{E}_t\left[\int_t^{t+\tau} \partial_\theta\phi^n(r,x_r^*) + \langle D\phi^n(r,x_r^*), \Delta x_r^* + B(x_r^*,u_r^*)\rangle \mathrm{d}r\right](\omega) \\
&\quad + \mathbb{E}_t\left[\int_t^{t+\tau} \frac{1}{2}\mathrm{tr}(\Sigma(x_r^*,u_r^*)^* D^2\phi^n(r,x_r^*)\Sigma(x_r^*,u_r^*))\mathrm{d}r\right](\omega).
\end{aligned} \tag{4.37}$$

Now, we take the expectation in (4.34), divide by τ and take the limit superior $\tau \to 0$. By Lemma 4.5, we can again apply Fatou's lemma which yields

$$\begin{aligned}
&\limsup_{\tau\downarrow 0} \frac{1}{\tau}\mathbb{E}\left[V(t+\tau,x_{t+\tau}^*) - V(t,x_t^*)\right] \\
&\le \mathbb{E}\left[\limsup_{\tau\downarrow 0} \frac{1}{\tau}\mathbb{E}_t\left[V(t+\tau,x_{t+\tau}^*) - V(t,x_t^*)\right](\omega)\right] \\
&\le \mathbb{E}\Bigg[\limsup_{\tau\downarrow 0} \frac{1}{\tau}\mathbb{E}_t\Bigg[\int_t^{t+\tau} \partial_\theta\phi^n(r,x_r^*) + \langle D\phi^n(r,x_r^*), \Delta x_r^* + B(x_r^*,u_r^*)\rangle \mathrm{d}r \\
&\qquad + \int_t^{t+\tau} \frac{1}{2}\mathrm{tr}(\Sigma(x_r^*,u_r^*)^* D^2\phi^n(r,x_r^*)\Sigma(x_r^*,u_r^*))\mathrm{d}r\Bigg](\omega) + C\|P_t - P_t^n\|_{L(L^2(\Lambda))}\Bigg].
\end{aligned} \tag{4.38}$$

By Lemma 3.25, this implies

$$\begin{aligned}
&\limsup_{\tau\downarrow 0} \frac{1}{\tau}\mathbb{E}\left[V(t+\tau,x_{t+\tau}^*) - V(t,x_t^*)\right] \\
&\le \mathbb{E}\left[G_t + \langle p_t, \Delta\bar{x}_t + b(\bar{x}_t,\bar{u}_t)\rangle + \frac{1}{2}\mathrm{tr}(\sigma(\bar{x}_t,\bar{u}_t)^* P_t^n \sigma(\bar{x}_t,\bar{u}_t)) + C\|P_t - P_t^n\|_{L(L^2(\Lambda))}\right].
\end{aligned} \tag{4.39}$$

Together with (4.33), after taking the limit $n \to \infty$ and plugging in the terminal condition for V, this yields

$$\mathbb{E}\left[\int_s^T L(x_t^*, u_t^*)\mathrm{d}t + H(x_T^*)\right] \leq V(s,x), \tag{4.40}$$

which concludes the proof. □

Next, we present conditions under which the value function satisfies the growth condition (4.9) and the semiconcavity (4.10).

Proposition 4.6. *Let $H : L^2(\Lambda) \to \mathbb{R}$ be Lipschitz continuous with respect to the $H^{-1}(\Lambda)$-norm, and twice Fréchet differentiable with bounded second derivative. Then there exists a constant C such that for every $t \in [s,T)$, $\tau \in (0, T-t]$, and $x \in H_0^1(\Lambda)$,*

$$V(t+\tau, x) - V(t,x) \leq C\left(1 + \|x\|^2_{H_0^1(\Lambda)}\right)\tau. \tag{4.41}$$

Proof. Again, the proof is similar as in the finite-dimensional case, see [FS06, Section IV.8], and the Lipschitz continuity of the state trajectories in $H^{-1}(\Lambda)$ relies on the analyticity of the heat semigroup.

Let u be an admissible control defined on $[t,T]$, and let x^u denote the associated state with initial condition $x_t^u = x$. We introduce a time shifted control $\tilde{u}$ defined on $[t+\tau, T]$ given by

$$\tilde{u}(r) := u(r-\tau), \tag{4.42}$$

and denote by $\tilde{x}$ the associated state with initial condition $\tilde{x}_{t+\tau} = x$. Then, we obtain

$$\begin{aligned} J(t+\tau, x; \tilde{u}) &= \mathbb{E}\left[\int_{t+\tau}^T L(\tilde{x}_r, \tilde{u}_r)\mathrm{d}r + H(\tilde{x}_T)\right] \\ &= \mathbb{E}\left[\int_t^{T-\tau} L(x_r^u, u_r)\mathrm{d}r + H(x^u_{T-\tau})\right]. \end{aligned} \tag{4.43}$$

Hence,

$$J(t+\tau, x; \tilde{u}) - J(t,x,u) = \mathbb{E}\left[-\int_{T-\tau}^T L(x_r^u, u_r)\mathrm{d}r + H(x^u_{T-\tau}) - H(x^u_T)\right]. \tag{4.44}$$

For the running costs, we have by the quadratic growth assumption (4.3) on L and standard estimates for solutions of SPDEs

$$\mathbb{E}\left[-\int_{T-\tau}^T L(x_r^u, u_r)\mathrm{d}r\right] \leq C\left(1 + \|x\|^2_{L^2(\Lambda)}\right)\tau. \tag{4.45}$$

For the terminal costs, using the boundedness of H_{xx} and a Taylor expansion, we obtain

$$\begin{aligned} &\mathbb{E}\left[H(x^u_{T-\tau}) - H(x^u_T)\right] \\ &\leq \mathbb{E}\left[\langle H_x(x^u_{T-\tau}), x^u_{T-\tau} - x^u_T\rangle_{L^2(\Lambda)}\right] + C\mathbb{E}\left[\|x^u_{T-\tau} - x^u_T\|^2_{L^2(\Lambda)}\right]. \end{aligned} \tag{4.46}$$

For the first summand, we have

$$\begin{aligned}&\mathbb{E}\left[\langle H_x(x^u_{T-\tau}), x^u_{T-\tau} - x^u_T\rangle_{L^2(\Lambda)}\right]\\&\leq \mathbb{E}\left[\|H_x(x^u_{T-\tau})\|^2_{H^1_0(\Lambda)}\right]^{\frac{1}{2}} \mathbb{E}\left[\left\|\mathbb{E}_{T-\tau}[x^u_T - x^u_{T-\tau}]\right\|^2_{H^{-1}(\Lambda)}\right]^{\frac{1}{2}}.\end{aligned} \tag{4.47}$$

Since H is Lipschitz continuous with respect to the $H^{-1}(\Lambda)$-norm, the Fréchet derivative of H maps to $H^1_0(\Lambda)$ and is bounded. For the second factor, we obtain analogous to (4.24)

$$\mathbb{E}\left[\left\|\mathbb{E}_{T-\tau}[x^u_T - x^u_{T-\tau}]\right\|^2_{H^{-1}(\Lambda)}\right]^{\frac{1}{2}} \leq C\left(1 + \|x\|^2_{H^1_0(\Lambda)}\right)\tau. \tag{4.48}$$

For the second summand in (4.46) we have

$$\mathbb{E}\left[\|x^u_{T-\tau} - x^u_T\|^2_{L^2(\Lambda)}\right] \leq C\left(1 + \|x\|^2_{L^2(\Lambda)}\right)\tau \tag{4.49}$$

by standard regularity results for solutions of SPDEs. Hence, we obtain together with (4.45)

$$J(t+\tau, x; \tilde{u}) - J(t, x, u) \leq C\left(1 + \|x\|^2_{H^1_0(\Lambda)}\right)\tau. \tag{4.50}$$

Now, for given $\varepsilon > 0$, let u^ε be a control such that

$$J(t, x; u^\varepsilon) \leq V(t, x) + \varepsilon. \tag{4.51}$$

Then we have

$$\begin{aligned}V(t+\tau, x) - V(t, x) &\leq J(t+\tau, x; \tilde{u}^\varepsilon) - J(t, x, u^\varepsilon) + \varepsilon\\&\leq C\left(1 + \|x\|^2_{H^1_0(\Lambda)}\right)\tau + \varepsilon.\end{aligned} \tag{4.52}$$

Since this holds for an arbitrary ε, this concludes the proof. □

Proposition 4.7. *Let L and H be Lipschitz continuous and semiconcave in x uniformly in u. Let B and Σ be Fréchet differentiable in x with Lipschitz continuous derivative uniformly in u. Then the value function V of the control problem* (4.1) *and* (4.2) *is semiconcave uniformly in $t \in [s, T]$, i.e., there exists a $C \geq 0$ such that for every $t \in (s, T]$,*

$$V(t, \cdot) - C\|\cdot\|^2_{L^2(\Lambda)} \tag{4.53}$$

is concave.

Proof. The corresponding result in the finite-dimensional case can be found in [YZ99, Chapter 4, Proposition 4.5]. The proof in the infinite-dimensional case is exactly the same upon replacing the finite-dimensional derivatives by Fréchet derivatives. □

It is well-known that under additional regularity assumptions the value function can be characterized as the unique B-continuous viscosity solution of the Hamilton-

Jacobi-Bellman equation

$$\begin{cases} V_s + \langle \Delta x, DV \rangle_{L^2(\Lambda)} + \inf_{u \in U} \mathcal{H}(x, u, DV, D^2 V) = 0 \\ V(T, x) = H(x), \end{cases} \tag{4.54}$$

in the sense of [FGŚ17, Definition 3.35]. Thus, in conjunction with the previous result we obtain the following verification theorem in terms of a B-continuous viscosity subsolution of (4.54).

Theorem 4.8 (Verification Theorem)**.** *Suppose there exists a constant $C > 0$ such that*

$$\|\Sigma(x, u) - \Sigma(y, u)\|_{L_2(\Xi, L^2(\Lambda))} \leq C \|x - y\|_{H^{-1}(\Lambda)}, \tag{4.55}$$

for every $x, y \in L^2(\Lambda)$ and $u \in U$, and let L and H be locally uniformly continuous in $x \in L^2(\Lambda)$, uniformly in $u \in U$. Let $\mathcal{V}$ be a B-continuous viscosity subsolution (see [FGŚ17, Definition 3.35]) of the HJB equation (4.54) *satisfying the growth condition* (4.9) *and the semiconcavity* (4.10)*, as well as*

$$\mathcal{V}(T, x) = H(x) \tag{4.56}$$

for all $x \in L^2(\Lambda)$,

$$|\mathcal{V}(t, x)| \leq C \left(1 + \|x\|^2_{L^2(\Lambda)}\right), \tag{4.57}$$

and

$$\lim_{t \to T} \left(\mathcal{V}(t, x) - H(S_{T-t} x)\right)^+ = 0 \tag{4.58}$$

uniformly on bounded subsets of $L^2(\Lambda)$, where $(S_r)_{r \geq 0}$ denotes the heat semigroup. Then we have:

(i) $\mathcal{V}(s, x) \leq V(s, x) \leq J(s, x; u)$ for any $(s, x) \in (0, T] \times L^2(\Lambda)$ and any admissible control u.

(ii) Suppose there are adapted processes

$$\begin{cases} G \in L^2([s, T] \times \Omega; \mathbb{R}) \\ p \in L^2([s, T] \times \Omega; H^1_0(\Lambda)) \\ P \in L^2([s, T] \times \Omega; L_2(L^2(\Lambda))) \end{cases} \tag{4.59}$$

such that for almost every $t \in [s, T]$,

$$(G_t, p_t, P_t) \in D^{1,2,+}_{t+,x} \mathcal{V}(t, x^*_t) \tag{4.60}$$

$\mathbb{P}$–almost surely, and

$$\mathbb{E}\left[\int_s^T G_t + \langle \Delta x^*_t, p_t \rangle_{H^{-1}(\Lambda) \times H^1_0(\Lambda)} + \mathcal{H}(x^*_t, u^*_t, p_t, P_t) \mathrm{d}t\right] \leq 0. \tag{4.61}$$

Then (x^*, u^*) is an optimal pair.

Proof. Under the additional assumption (4.55), the value function is the unique B-continuous viscosity solution of the HJB equation, see [FGŚ17, Theorem 3.67] for details. Thus, (i) follows from the comparison result [FGŚ17, Theorem 3.54].

The proof of (ii) follows by using exactly the same arguments as in the proof of Theorem 4.3 up until inequality (4.40), which we now have for $\mathcal{V}$ instead of the value function. Applying part (i) concludes the proof. □

Remark 4.9. (i) In the case of the Nemytskii operator σ as defined in Assumption 2.2, the assumption (4.55) cannot be satisfied in general. However, one can approximate the Nemytskii operator $\sigma(x(\lambda), u)(\xi)$ for $x \in L^2(\Lambda)$, $\lambda \in \Lambda$ and $\xi \in \Xi$, with smooth noise coefficients of the following type

$$\Sigma_\varepsilon(x, u)(\xi)(\lambda) = \sigma\left(\int_\Lambda f_\varepsilon(\lambda - \mu)x(\mu)\mathrm{d}\mu, u\right)(\xi) \tag{4.62}$$

where $(f_\varepsilon)_{\varepsilon>0} \subset C_c^\infty(\mathbb{R})$ is a mollifier, i.e., $\int_\Lambda f_\varepsilon(\lambda - \mu)x(\mu)\mathrm{d}\mu$ converges to x in $L^2(\Lambda)$ as $\varepsilon \to 0$. Under the assumptions imposed on σ in Assumption 2.2, Σ_ε satisfies the additional regularity condition (4.55).

(ii) All our results in this chapter generalize to the case of uniformly elliptic operators in divergence form with Dirichlet boundary conditions formally given by

$$Ax(\lambda) = \partial_\lambda(a\partial_\lambda x)(\lambda) \tag{4.63}$$

for some $a \in L^\infty(\Lambda)$ with $a \geq a_0 > 0$. Indeed, the variational setting relies on the monotonicity and the coercivity of the unbounded operator which also holds for the operator A, see also Remark 2.13 for more details. Furthermore, A is the generator of an analytic semigroup, hence Lemma 3.8 and Proposition 4.7 still hold. Finally, the strong B-condition needed for the proof of Theorem 4.3 part (i) is shown in [FGŚ17, Example 3.16].

Part III.

Applications

5. The Stochastic Nagumo Model

Part III of this dissertation is devoted to the analysis of the optimal control of the stochastic Nagumo model with a view towards efficient numerical implementations. First, we prove existence of optimal controls, analyze the regularity of the control-to-state operator, and derive a necessary optimality condition. In the case of additive noise, we characterize the adjoint state as the solution of a random backward PDE and develop a gradient descent method for the approximation of optimal controls. Finally, we present numerical examples. This chapter is based on [SW21a].

5.1. Introduction

The Hodgkin-Huxley model proposed in [HH52] is one of the most famous models for the propagation of action potentials in neurons. However, due to the high complexity of the model, simulations of large networks of neurons quickly become unfeasible. Therefore, one often restrains to one of the simplified models proposed by FitzHugh, see [Fit61], or Nagumo et al., see [NAY62]. For a neurobiological derivation of these models, see [ET10, Mur02, Mur03].

More specifically, the stochastic Nagumo model, also known as the Schlögl model due to [Sch72], is given by

$$\mathrm{d}x_t = [\Delta x_t + \gamma x_t(x_t - 1)(a - x_t)]\,\mathrm{d}t + \Sigma(t, x_t)\mathrm{d}W_t, \tag{5.1}$$

for some $\gamma > 0$ and $a \in (0, 1)$. The main feature of this model is the cubic nonlinearity which, in particular, does not fit into the framework discussed in Part I and Part II of this dissertation. In order to capture this kind of nonlinearity, we consider the controlled SPDE

$$\begin{cases} \mathrm{d}x_t^u = [\Delta x_t^u + b(x_t^u) + d(t)u(t)]\,\mathrm{d}t + \Sigma(t, x_t^u)\mathrm{d}W_t, \quad t \in [0, T] \\ x_0^u = x \in L^6(\Omega, \mathcal{F}_0, \mathbb{P}; L^2(\Lambda)) \end{cases} \tag{5.2}$$

with homogeneous Neumann boundary conditions, where $\Lambda \subset \mathbb{R}$ is a bounded interval, $T > 0$ is fixed, $(W_t)_{t\in[0,T]}$ is a cylindrical Wiener process taking values in some real, separable Hilbert space Ξ and defined on a given filtered probability space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\big)$, where $(\mathcal{F}_t)_{t\in[0,T]}$ is the filtration generated by (W_t) and augmented by all $\mathbb{P}$-null sets, $d \in L^\infty([0, T] \times \Lambda)$, $\Sigma : [0, T] \times L^2(\Lambda) \to L_2(\Xi, L^2(\Lambda))$ is Fréchet differentiable for every fixed $t \in [0, T]$ and satisfies for all $t \in [0, T]$ and

$x, y \in L^2(\Lambda)$,

$$
\begin{cases}
\|(\Sigma(t,x) - \Sigma(t,y))\|_{L_2(\Xi, L^2(\Lambda))} \leq C\|x - y\|_{L^2(\Lambda)}, \\
\|\Sigma(t,x)\|_{L_2(\Xi, L^2(\Lambda))} \leq C\left(1 + \|x\|_{L^2(\Lambda)}\right), \\
\|\Sigma_x(t,x)y\|_{L_2(\Xi, L^2(\Lambda))} \leq C\|y\|_{L^2(\Lambda)},
\end{cases}
\tag{5.3}
$$

for some constant $C \in \mathbb{R}$, where $\|\cdot\|_{L_2(\Xi, L^2(\Lambda))}$ denotes the Hilbert-Schmidt norm on the space of all Hilbert-Schmidt operators on $L^2(\Lambda)$. Furthermore, $b : \mathbb{R} \to \mathbb{R}$ is continuously differentiable satisfying $b(0) = 0$,

$$
\sup_{x \in \mathbb{R}} b'(x) < \infty, \tag{5.4}
$$

and for all $x \in \mathbb{R}$,

$$
|b'(x)| < C(1 + |x|^2), \tag{5.5}
$$

for some constant $C \in \mathbb{R}$.

Remark 5.1. (i) Notice that the upper bound of the derivative implies a one-sided Lipschitz condition, i.e., there exists a constant $\widetilde{\mathrm{Lip}}_b \in \mathbb{R}$ such that

$$
(b(x) - b(y))(x - y) \leq \widetilde{\mathrm{Lip}}_b (x - y)^2, \tag{5.6}
$$

for all $x, y \in \mathbb{R}$.

(ii) The nonlinearity in the Nagumo equation satisfies these conditions since the leading coefficient of the polynomial is negative and the derivative is a polynomial of degree 2.

(iii) A possible choice for the diffusion coefficient Σ would be the Nemytskii operator associated with $\sigma : \mathbb{R} \to \mathbb{R}$,

$$
\sigma(x) := \bar{\sigma} \min\{x(x-1), M\},
$$

for some constants $\bar{\sigma}, M > 0$. In the stochastic Nagumo model, this choice imposes noise in particular on the wave front of the resulting traveling wave. For a more detailed discussion, see [LPS14, Example 10.2].

Let $H^1(\Lambda)$ denote the Sobolev space of order 1 with Neumann boundary conditions and consider the Gelfand triple

$$
H^1(\Lambda) \subset L^6(\Lambda) \subset \left(H^1(\Lambda)\right)^*. \tag{5.7}
$$

Under appropriate assumptions on the control u (to be specified later), the existence of a variational solution of equation (5.2) in the space

$$
E := L^2([0,T] \times \Omega, \mathrm{d}t \otimes \mathbb{P}; H^1(\Lambda)) \cap L^6(\Omega, \mathbb{P}; C([0,T]; L^2(\Lambda))) \tag{5.8}
$$

is assured (see e.g. [LR15, Example 5.1.8]).

Our objective is to study the optimal control problem associated with the state equation (5.2). Let $I_1 : L^6(\Omega; C([0,T]; L^2(\Lambda))) \to \mathbb{R}$ be given by

$$I_1(x) := \mathbb{E}\left[\frac{c_{\overline{\Lambda}}}{2}\int_0^T\int_\Lambda \left(x_t(\lambda) - x_{\overline{\Lambda}}(t,\lambda)\right)^2 \mathrm{d}\lambda\mathrm{d}t + \frac{c_T}{2}\int_\Lambda \left(x_T(\lambda) - x^T(\lambda)\right)^2 \mathrm{d}\lambda\right] \tag{5.9}$$

and $I_2 : L^2([0,T]\times\Lambda) \to \mathbb{R}$ be given by

$$I_2(u) := \frac{\nu}{2}\int_0^T\int_\Lambda u^2(t,\lambda)\mathrm{d}\lambda\mathrm{d}t, \tag{5.10}$$

where $c_{\overline{\Lambda}}$, c_T, $\nu \geq 0$, and $x_{\overline{\Lambda}} \in L^2([0,T]\times\Lambda)$ and $x^T \in L^2(\Lambda)$ are given running and terminal reference profiles, respectively. We want to minimize the cost functional

$$J(u) := I_1(x^u) + I_2(u), \tag{5.11}$$

subject to the state equation (5.2), where

$$u \in U_{\mathrm{ad}} := \left\{u \in L^6([0,T]\times\Lambda) \,|\, \|u\|_{L^6([0,T]\times\Lambda)} \leq \kappa\right\}, \tag{5.12}$$

for given $\kappa \geq 0$.

Remark 5.2. The proof of the Gateaux-differentiability of $u \mapsto x^u$ (see Proposition 5.8 below), requires a moment bound of the solution in $L^6(\Omega\times[0,T]\times\Lambda)$ due to the upper bound (5.5) on the derivative b' of the nonlinearity. Therefore the minimal requirement for an admissible control is $u \in L^6([0,T]\times\Lambda)$.

In the work by Buchholz et al. [BEKT13a] on the deterministic case, the set of admissible controls

$$\tilde{U}_{\mathrm{ad}} := \{u \in L^\infty([0,T]\times\Lambda) \,|\, u_a \leq u(t,x) \leq u_b \text{ for a.a. } (t,\lambda) \in [0,T]\times\Lambda\}, \tag{5.13}$$

for some $u_a < u_b$ is considered. We could use the same set in our analysis as well.

5.2. Well-Posedness of the Optimal Control Problem

First we want to show that the control problem is well-posed. In order to do so, we need the following a priori bound for solutions of the state equation (5.2).

Proposition 5.3. *There is a constant $C = C(b,d,\Sigma,T,x)$ such that for every solution $x^u \in E$ of the state equation (5.2) associated with $u \in U_{ad}$ on the right hand side we have*

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|x_t^u\|^6_{L^2(\Lambda)} + \left(\int_0^T \|x_t^u\|^2_{H^1(\Lambda)}\,\mathrm{d}t\right)^3\right] \leq C\left(1 + \int_0^T \|u(t)\|^6_{L^2(\Lambda)}\,\mathrm{d}t\right). \tag{5.14}$$

Proof. By the Itô formula from [LR15, Theorem 4.2.5], we have

$$\begin{aligned}
&\|x_t^u\|_{L^2(\Lambda)}^2 \\
&= \|x\|_{L^2(\Lambda)}^2 + 2\int_0^t {}_{(H^1(\Lambda))^*}\langle \Delta x_s^u, x_s^u\rangle_{H^1(\Lambda)}\,\mathrm{d}s + 2\int_0^t \langle b\left(x_s^u\right), x_s^u\rangle_{L^2(\Lambda)}\,\mathrm{d}s \\
&\quad + 2\int_0^t \langle d(s)u(s), x_s^u\rangle_{L^2(\Lambda)}\,\mathrm{d}s + \int_0^t \|\sigma(s,x_s^u)\|_{L_2(\Xi,L^2(\Lambda))}^2 \mathrm{d}s \\
&\quad + 2\int_0^t \langle x_s^u, \sigma(s,x_s^u)dW_s\rangle_{L^2(\Lambda)}.
\end{aligned} \tag{5.15}$$

Using the one-sided Lipschitz continuity of b (5.6) and the fact that $b(0) = 0$, we obtain

$$\langle b\left(x_s^u\right), x_s^u\rangle_{L^2(\Lambda)} \le \widetilde{\mathrm{Lip}}_b \|x_s^u\|_{L^2(\Lambda)}^2. \tag{5.16}$$

The rest of the proof is analogous to the proof of Lemma 2.5. □

As a consequence, the finiteness of all of the integrals appearing in the cost functional J is assured. Furthermore, we have the following result.

Corollary 5.4. *Let E be defined as in* (5.8). *Every solution $x^u \in E$ of the state equation* (5.2) *associated with $u \in U_{ad}$ on the right hand side is in $L^6(\Omega \times [0,T] \times \Lambda)$.*

Proof. We apply the Gagliardo-Nirenberg interpolation inequality which can be found in [Rou13]. This yields for almost all $(t,\omega) \in [0,T] \times \Omega$,

$$\|x_t^u\|_{L^6(\Lambda)}^6 \le C\,\|x_t^u\|_{H^1(\Lambda)}^2\,\|x_t^u\|_{L^2(\Lambda)}^4. \tag{5.17}$$

Integrating over $[0,T] \times \Omega$ yields

$$\begin{aligned}
\mathbb{E}\left[\int_0^T \|x_t^u\|_{L^6(\Lambda)}^6\,\mathrm{d}t\right] &\le \mathbb{E}\left[\int_0^T \|x_t^u\|_{H^1(\Lambda)}^2\,\|x_t^u\|_{L^2(\Lambda)}^4\,\mathrm{d}t\right] \\
&\le \mathbb{E}\left[\sup_{t\in[0,T]} \|x_t^u\|_{L^2(\Lambda)}^4 \int_0^T \|x_t^u\|_{H^1(\Lambda)}^2\,\mathrm{d}t\right] \\
&\le \mathbb{E}\left[\sup_{t\in[0,T]} \|x_t^u\|_{L^2(\Lambda)}^6\right]^{\frac{2}{3}} \mathbb{E}\left[\left(\int_0^T \|x_t^u\|_{H^1(\Lambda)}^2\,\mathrm{d}t\right)^3\right]^{\frac{1}{3}} \\
&< \infty,
\end{aligned} \tag{5.18}$$

where we used Hölder's inequality and Proposition 5.3. □

Next, we show that the control-to-state operator of the state equation (5.2) is globally Lipschitz continuous.

Proposition 5.5. *Let E be defined as in* (5.8). *For the control-to-state operator*

$$\begin{aligned}
L^2\left([0,T]\times\Lambda\right) &\to E \\
u &\mapsto x^u,
\end{aligned} \tag{5.19}$$

there exists a constant $C = C(b, d, \Sigma, \Lambda, T) \in \mathbb{R}$ *such that*

$$\|x_t^u - x_t^v\|_E^2 \leq C \int_0^T \|u - v\|_{L^2(\Lambda)}^2 \,\mathrm{d}s, \tag{5.20}$$

i.e., the control-to-state operator is Lipschitz continuous from $L^2([0,T] \times \Lambda)$ *to* E.

Proof. By Itô's formula (see [LR15, Theorem 4.2.5]), we have almost surely

$$\begin{aligned}
\|x_t^u - x_t^v\|_{L^2(\Lambda)}^2 = 2 \int_0^t & {}_{(H^1(\Lambda))^*} \langle \Delta\left(x_s^u - x_s^v\right), x_s^u - x_s^v \rangle_{H^1(\Lambda)} \,\mathrm{d}s \\
&+ 2 \int_0^t \langle b\left(x_s^u\right) - b\left(x_s^v\right), x_s^u - x_s^v \rangle_{L^2(\Lambda)} \,\mathrm{d}s \\
&+ 2 \int_0^t \langle d(s)\left(u(s) - v(s)\right), x_s^u - x_s^v \rangle_{L^2(\Lambda)} \,\mathrm{d}s \\
&+ \int_0^t \|\Sigma(s, x_s^u) - \Sigma(s, x_s^v)\|_{L_2(\Xi, L^2(\Lambda))}^2 \mathrm{d}s \\
&+ 2 \int_0^t \langle x_s^u - x_s^v, \Sigma(s, x_s^u) - \Sigma(s, x_s^v) \mathrm{d}W_s \rangle.
\end{aligned} \tag{5.21}$$

Using the Lipschitz condition (5.3), the one-sided Lipschitz continuity of b (5.6), and similar arguments as in the proof of Proposition 5.3 yields the claim. $\square$

Now we want to prove the existence of an optimal control:

Theorem 5.6. *There is at least one optimal control* $\bar{u} \in U_{ad}$ *such that*

$$J(\bar{u}) = \inf_{u \in U_{ad}} J(u). \tag{5.22}$$

Proof. First, we notice that J is nonnegative and hence bounded from below. Let $(u_n)_{n\in\mathbb{N}} \subset U_{\mathrm{ad}}$ be a minimizing sequence, i.e.,

$$\lim_{n\to\infty} J(u_n) = \inf_{u \in U_{\mathrm{ad}}} J(u), \tag{5.23}$$

Since $(u_n)_{n\in\mathbb{N}} \subset U_{\mathrm{ad}}$, $(u_n)_{n\in\mathbb{N}}$ is in particular bounded in $L^2([0,T] \times \Lambda)$. Hence, we can extract a weakly convergent subsequence – again denoted by u_n – such that $u_n \rightharpoonup \bar{u}$ in $L^2([0,T] \times \Lambda)$ for some $\bar{u} \in L^2([0,T] \times \Lambda)$. The point is now to show that $\bar{u} \in U_{\mathrm{ad}}$, and $\bar{u}$ minimizes J in U_{ad}.

Since U_{ad} is convex and strongly closed, it follows that U_{ad} is also weakly closed, hence $\bar{u} \in U_{\mathrm{ad}}$.

Next, we show that $\bar{u}$ minimizes J. Let $x^n, \bar{x} \in E$ denote the unique solution of the state equation (5.2) associated with u_n and $\bar{u}$ on the right hand side, respectively. We first show that x^n converges strongly to $\bar{x}$. In the deterministic case, the a priori bound in Proposition 5.3 holds pathwise and we can apply a compact embedding theorem in order to show strong convergence of the solutions. Since we only have the a priori bound under the expectation, we cannot use the same technique. Instead we apply the so-called compactness method introduced in [FG95]. Let us sketch this technique here:

From the bound

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left[\sup_{t\in[0,T]} \|x_t^n\|_{L^2(\Lambda)}^2 + \int_0^T \|x_s^n\|_{H^1(\Lambda)}^2 \, \mathrm{d}s\right] < \infty \tag{5.24}$$

we can conclude tightness of the measures $\mathbb{P}^n := \mathbb{P}\circ(x^n)^{-1}$ on $L^2([0,T]\times\Lambda)$. Therefore, $(\mathbb{P}^n)_{n\in\mathbb{N}}$ is relatively compact and we can extract a converging subsequence $\mathbb{P}^n \to \bar{\mathbb{P}}$. It remains to identify the limit $\bar{\mathbb{P}}$. By the Skorohod representation theorem, see [Sko56], there exists a probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ and a sequence of random variables $(X^n)_{n\in\mathbb{N}}$ and $\bar{X}$ defined on $\hat{\Omega}$ with the same law as $(x^n)_{n\in\mathbb{N}}$ and $\bar{x}$, respectively, such that $X^n \to \bar{X}$ strongly in $L^2([0,T]\times\Lambda)$ $\hat{\mathbb{P}}$-almost surely. Therefore, using the martingale representation theorem, we can identify $\bar{X}$ as a solution of our state equation associated with $\bar{u}$ on the right hand side, see [DPZ14, Section 8.4] for details.

Now, we split the cost functional into one part that depends on x^u and into one part that depends on u. For the first part, I_1, we have

$$\begin{aligned}
&\lim_{n\to\infty} I_1(x^n)\\
&= \lim_{n\to\infty} \mathbb{E}\left[\frac{c_{\overline{\Lambda}}}{2}\int_0^T\int_\Lambda \left(x_t^n(\lambda) - x_{\overline{\Lambda}}(t,\lambda)\right)^2 \mathrm{d}\lambda\mathrm{d}t + \frac{c_T}{2}\int_\Lambda \left(x_T^n(\lambda) - x^T(\lambda)\right)^2 \mathrm{d}\lambda\right]\\
&= \lim_{n\to\infty} \hat{\mathbb{E}}\left[\frac{c_{\overline{\Lambda}}}{2}\int_0^T\int_\Lambda \left(X_t^n(\lambda) - x_{\overline{\Lambda}}(t,\lambda)\right)^2 \mathrm{d}\lambda\mathrm{d}t + \frac{c_T}{2}\int_\Lambda \left(X_T^n(\lambda) - x^T(\lambda)\right)^2 \mathrm{d}\lambda\right]\\
&\geq \hat{\mathbb{E}}\left[\liminf_{n\to\infty}\left(\frac{c_{\overline{\Lambda}}}{2}\int_0^T\int_\Lambda \left(X_t^n(\lambda) - x_{\overline{\Lambda}}(t,\lambda)\right)^2 \mathrm{d}\lambda\mathrm{d}t + \frac{c_T}{2}\int_\Lambda \left(X_T^n(\lambda) - x^T(\lambda)\right)^2 \mathrm{d}\lambda\right)\right]\\
&= \hat{\mathbb{E}}\left[\frac{c_{\overline{\Lambda}}}{2}\int_0^T\int_\Lambda \left(\bar{X}_t(\lambda) - x_{\overline{\Lambda}}(t,\lambda)\right)^2 \mathrm{d}\lambda\mathrm{d}t + \frac{c_T}{2}\int_\Lambda \left(\bar{X}_T(\lambda) - x^T(\lambda)\right)^2 \mathrm{d}\lambda\right]\\
&= \mathbb{E}\left[\frac{c_{\overline{\Lambda}}}{2}\int_0^T\int_\Lambda \left(\bar{x}_t(\lambda) - x_{\overline{\Lambda}}(t,\lambda)\right)^2 \mathrm{d}\lambda\mathrm{d}t + \frac{c_T}{2}\int_\Lambda \left(\bar{x}_T(\lambda) - x^T(\lambda)\right)^2 \mathrm{d}\lambda\right]\\
&= I_1(\bar{x}),
\end{aligned} \tag{5.25}$$

where we used Fatou's Lemma, and exploited that uniqueness in law holds for the state equation (5.2) and we have a solution in the space E (see equation (5.8) for the definition of E).

Furthermore, since I_2 is continuous and convex, it is also weakly lower semicontinuous, i.e.,

$$u_n \rightharpoonup \bar{u} \qquad \Longrightarrow \qquad \liminf_{n\to\infty} I_2(u_n) \geq I_2(\bar{u}). \tag{5.26}$$

Therefore, we have

$$\inf_{u\in U_{\mathrm{ad}}} J(u) = \lim_{n\to\infty} J(u_n) \geq \lim_{n\to\infty} I_1(x^n) + \liminf_{n\to\infty} I_2(u_n) \geq I_1(\bar{x}) + I_2(\bar{u}) = J(\bar{u}), \tag{5.27}$$

which completes the proof. □

Remark 5.7. This proof does not rely on the explicit form of our cost functional. The crucial point is, that the cost functional is sequentially weakly lower semicontinuous.

5.3. First Order Condition for Critical Points

In this section, we are first going to derive the Gâteaux derivative of the control-to-state operator and the cost functional, and then prove a necessary condition for a control to be locally optimal.

Proposition 5.8. *Let $b : \mathbb{R} \to \mathbb{R}$ satisfy the assumptions of Section 5.2 and $u \in L^6([0,T]\times\Lambda)$ be fixed. Then, for every $v \in L^6([0,T]\times\Lambda)$, the Gâteaux derivative of the control-to-state operator $u \mapsto x^u$, $L^6([0,T]\times\Lambda) \to E$ in direction v is given by the solution of the linear SPDE*

$$\begin{cases} \mathrm{d}y_t^v = [\Delta y_t^v + b'(x_t^u)y_t^v + d(t)v(t)]\mathrm{d}t + \Sigma_x(t,x_t^u)y_t^v\mathrm{d}W_t, \quad t\in[0,T] \\ y_0^v = 0 \in L^2(\Lambda). \end{cases} \tag{5.28}$$

Proof. Let y^v denote the solution of equation (5.28) associated with v on the right hand side. Set

$$z_\varepsilon(t) := \frac{x_t^{u+\varepsilon v} - x_t^u}{\varepsilon} - y_t^v. \tag{5.29}$$

We want to show that $z_\varepsilon \to 0$ in $L^2(\Omega\times[0,T];H^1(\Lambda)) \cap L^2(\Omega;C([0,T];L^2(\Lambda)))$ as $\varepsilon \to 0$. First notice

$$\begin{aligned} z_\varepsilon(t) = &\int_0^t \Delta z_\varepsilon(s) + \frac{1}{\varepsilon}\left(b(x_s^{u+\varepsilon v}) - b(x_s^u)\right) - b'(x_s^u)y_s^v\mathrm{d}s \\ &+ \int_0^t \frac{1}{\varepsilon}\left(\Sigma(s,x_s^{u+\varepsilon v}) - \Sigma(s,x_s^u)\right) - \Sigma_x(s,x_s^u)y_s^v\mathrm{d}W_s. \end{aligned} \tag{5.30}$$

Note that

$$\begin{aligned} &\frac{1}{\varepsilon}\left(b(x_s^{u+\varepsilon v}) - b(x_s^u)\right) - b'(x_s^u)y_s^v \\ &= \underbrace{\frac{1}{\varepsilon}\left(b(x_s^u + \varepsilon y_s^v) - b(x_s^u)\right) - b'(x_s^u)y_s^v}_{=:\mathcal{R}_\varepsilon^1(s)} + \underbrace{\frac{1}{\varepsilon}\left(b(x_s^{u+\varepsilon v}) - b(x_s^u + \varepsilon y_s^v)\right)}_{=:\mathcal{R}_\varepsilon^2(s)}. \end{aligned} \tag{5.31}$$

and similarly

$$\begin{aligned}
&\frac{1}{\varepsilon}\left(\Sigma(s,x_s^{u+\varepsilon v})-\Sigma(s,x_s^u)\right)-\Sigma_x(s,x_s^u)y_s^v \\
&=\underbrace{\frac{1}{\varepsilon}\left(\Sigma(s,x_s^u+\varepsilon y_s^v)-\Sigma(s,x_s^u)\right)-\Sigma_x(s,x_s^u)y_s^v}_{=:\mathcal{S}_\varepsilon^1(s)}+\underbrace{\frac{1}{\varepsilon}\left(\Sigma(s,x_s^{u+\varepsilon v})-\Sigma(s,x_s^u+\varepsilon y_s^v)\right)}_{=:\mathcal{S}_\varepsilon^2(s)}.
\end{aligned} \tag{5.32}$$

Together with equation (5.30), Itô's formula yields

$$\begin{aligned}
\frac{1}{2}\|z_\varepsilon(t)\|_{L^2(\Lambda)}^2 =& \int_0^t {}_{(H^1(\Lambda))^*}\langle\Delta z_\varepsilon(s),z_\varepsilon(s)\rangle_{H^1(\Lambda)}\,\mathrm{d}s+\int_0^t\left\langle\mathcal{R}_\varepsilon^1(s),z_\varepsilon(s)\right\rangle_{L^2(\Lambda)}\mathrm{d}s \\
&+\int_0^t\left\langle\mathcal{R}_\varepsilon^2(s),z_\varepsilon(s)\right\rangle_{L^2(\Lambda)}\mathrm{d}s+\int_0^t\left\langle z_\varepsilon(s),\mathcal{S}_\varepsilon^1(s)\mathrm{d}W_s\right\rangle_{L^2(\Lambda)} \\
&+\int_0^t\left\langle z_\varepsilon(s),\mathcal{S}_\varepsilon^2(s)\mathrm{d}W_s\right\rangle_{L^2(\Lambda)} \\
&+\frac{1}{2}\int_0^t\|\mathcal{S}_\varepsilon^1(s)+\mathcal{S}_\varepsilon^2(s)\|_{L_2(\Xi,L^2(\Lambda))}^2\mathrm{d}s.
\end{aligned} \tag{5.33}$$

First notice that

$$\int_0^t {}_{(H^1(\Lambda))^*}\langle\Delta z_\varepsilon(s),z_\varepsilon(s)\rangle_{H^1(\Lambda)}\mathrm{d}s=-\int_0^t\|\nabla z_\varepsilon(s)\|_{L^2(\Lambda)}^2\mathrm{d}s. \tag{5.34}$$

Furthermore, we have $\langle\mathcal{R}_\varepsilon^1(s),z_\varepsilon(s)\rangle_{L^2(\Lambda)}\le(\|\mathcal{R}_\varepsilon^1(s)\|_{L^2(\Lambda)}^2+\|z_\varepsilon(s)\|_{L^2(\Lambda)}^2)/2$, and, since b is one-sided Lipschitz continuous, we have

$$\begin{aligned}
\left\langle\mathcal{R}_\varepsilon^2(s),z_\varepsilon(s)\right\rangle_{L^2(\Lambda)} &= \frac{1}{\varepsilon^2}\left\langle b(x_s^{u+\varepsilon v})-b(x_s^u+\varepsilon y_s^v),x_s^{u+\varepsilon v}-(x_s^u+\varepsilon y_s^v)\right\rangle_{L^2(\Lambda)} \\
&\le\widetilde{\mathrm{Lip}}_b\|z_\varepsilon(s)\|_{L^2(\Lambda)}^2.
\end{aligned} \tag{5.35}$$

For the last term in equation (5.33), we have

$$\begin{aligned}
&\frac{1}{2}\int_0^T\|\mathcal{S}_\varepsilon^1(s)+\mathcal{S}_\varepsilon^2(s)\|_{L_2(\Xi,L^2(\Lambda))}^2\mathrm{d}s \\
&\le\int_0^T\|\mathcal{S}_\varepsilon^1(s)\|_{L_2(\Xi,L^2(\Lambda))}^2\mathrm{d}s+\int_0^T\|\mathcal{S}_\varepsilon^2(s)\|_{L_2(\Xi,L^2(\Lambda))}^2\mathrm{d}s,
\end{aligned} \tag{5.36}$$

where, by the Lipschitz condition (5.3) on Σ,

$$\begin{aligned}\left\|\mathcal{S}_\varepsilon^2(s)\right\|_{L_2(\Xi,L^2(\Lambda))}^2 &= \left\|\frac{1}{\varepsilon}\left(\Sigma(s,x_s^{u+\varepsilon v}) - \Sigma(s,x_s^u+\varepsilon y_s^v)\right)\right\|_{L_2(\Xi,L^2(\Lambda))}^2 \\ &\leq C\left\|\frac{1}{\varepsilon}\left(x_s^{u+\varepsilon v}-x_s^u\right)-y_s^v\right\|_{L^2(\Lambda)}^2 \\ &= C\|z_\varepsilon(s)\|_{L^2(\Lambda)}^2.\end{aligned} \tag{5.37}$$

Therefore, taking the supremum with respect to $t\in[0,T]$ in equation (5.33) and taking expectations, it follows

$$\begin{aligned}&\mathbb{E}\left[\sup_{t\in[0,T]}\|z_\varepsilon(t)\|_{L^2(\Lambda)}^2+\int_0^T\|\nabla z_\varepsilon(t)\|_{L^2(\Lambda)}^2\mathrm{d}t\right] \\ &\leq C\Bigg\{\mathbb{E}\left[\int_0^T\left\|\mathcal{R}_\varepsilon^1(t)\right\|_{L^2(\Lambda)}^2\mathrm{d}t\right]+\int_0^T\mathbb{E}\left[\sup_{s\in[0,t]}\|z_\varepsilon(s)\|_{L^2(\Lambda)}^2\right]\mathrm{d}t \\ &\qquad+\mathbb{E}\left[\sup_{t\in[0,T]}\int_0^t\left\langle z_\varepsilon(t),\mathcal{S}_\varepsilon^1(t)\mathrm{d}W_t\right\rangle_{L^2(\Lambda)}\right] \\ &\qquad+\mathbb{E}\left[\sup_{t\in[0,T]}\int_0^t\left\langle z_\varepsilon(s),\mathcal{S}_\varepsilon^2(s)\mathrm{d}W_s\right\rangle_{L^2(\Lambda)}\right]\Bigg\}.\end{aligned} \tag{5.38}$$

Using Burkholder-Davis-Gundy inequality, we have

$$\mathbb{E}\left[\sup_{t\in[0,T]}\int_0^t\left\langle z_\varepsilon(s),\mathcal{S}_\varepsilon^2(s)\mathrm{d}W_s\right\rangle_{L^2(\Lambda)}\right]\leq C\mathbb{E}\left[\left\langle\int_0^\cdot\left\langle z_\varepsilon(s),\mathcal{S}_\varepsilon^2(s)\mathrm{d}W_s\right\rangle_{L^2(\Lambda)}\right\rangle_T^{\frac{1}{2}}\right]. \tag{5.39}$$

Now we compute the quadratic variation. To this end, let $(e_k)_{k\geq 1}$ be an orthonormal basis of $L^2(\Lambda)$. Then we have

$$\begin{aligned}&\left\langle\int_0^\cdot\left\langle z_\varepsilon(s),\mathcal{S}_\varepsilon^2(s)\mathrm{d}W_s\right\rangle_{L^2(\Lambda)}\right\rangle_T^{\frac{1}{2}} \\ &=\left(\int_0^T\sum_{k=1}^\infty\left|\left\langle z_\varepsilon(s),\mathcal{S}_\varepsilon^2(s)e_k\right\rangle_{L^2(\Lambda)}\right|^2\mathrm{d}s\right)^{\frac{1}{2}} \\ &\leq\frac{\alpha}{2}\sup_{t\in[0,T]}\|z_\varepsilon(t)\|_{L^2(\Lambda)}^2+\frac{1}{2\alpha}\int_0^T\left\|\mathcal{S}_\varepsilon^2(s)\right\|_{L_2(\Xi,L^2(\Lambda))}^2\mathrm{d}s,\end{aligned} \tag{5.40}$$

for arbitrary $\alpha>0$. With the same estimates as above for $\|\mathcal{S}_\varepsilon^2(s)\|_{L_2(\Xi,L^2(\Lambda))}^2$ and with inequality (5.39) this yields

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}\int_0^t \left\langle z_\varepsilon(s), \mathcal{S}_\varepsilon^2(s)\mathrm{d}W_s\right\rangle_{L^2(\Lambda)}\right] \\
&\leq C\alpha\mathbb{E}\left[\sup_{t\in[0,T]}\|z_\varepsilon(t)\|^2_{L^2(\Lambda)}\right] + \frac{C}{\alpha}\mathbb{E}\left[\int_0^T \|z_\varepsilon(s)\|^2_{L^2(\Lambda)}\mathrm{d}s\right].
\end{aligned}
\tag{5.41}
$$

Furthermore, with similar calculations as above, we obtain

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}\int_0^t \left\langle z_\varepsilon(s), \mathcal{S}_\varepsilon^1(s)\mathrm{d}W_s\right\rangle_{L^2(\Lambda)}\right] \\
&\leq C\alpha\,\mathbb{E}\left[\int_0^T \|z_\varepsilon(s)\|^2_{L^2(\Lambda)}\mathrm{d}s\right] + \frac{C}{\alpha}\,\mathbb{E}\left[\int_0^T \left\|\mathcal{S}_\varepsilon^1(s)\right\|^2_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}s\right],
\end{aligned}
\tag{5.42}
$$

for arbitrary $\alpha > 0$. Choosing $\alpha > 0$ in (5.41) and (5.42) sufficiently small, we derive from (5.38)

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}\|z_\varepsilon(t)\|^2_{L^2(\Lambda)} + \int_0^T \|\nabla z_\varepsilon(s)\|^2_{L^2(\Lambda)}\mathrm{d}s\right] \\
&\leq C\Bigg\{\int_0^T \mathbb{E}\left[\sup_{s\in[0,t]}\|z_\varepsilon(s)\|^2_{L^2(\Lambda)}\right]\mathrm{d}t + \mathbb{E}\left[\int_0^T \left\|\mathcal{R}_\varepsilon^1(s)\right\|^2_{L^2(\Lambda)}\mathrm{d}s\right] \\
&\quad + \mathbb{E}\left[\int_0^T \left\|\mathcal{S}_\varepsilon^1(s)\right\|^2_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}s\right]\Bigg\}.
\end{aligned}
\tag{5.43}
$$

By Grönwall's inequality, this yields

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[0,T]}\|z_\varepsilon(s)\|^2_{L^2(\Lambda)} + \int_0^T \|\nabla z_\varepsilon(s)\|^2_{L^2(\Lambda)}\mathrm{d}s\right] \\
&\leq C\mathbb{E}\left[\int_0^T \|\mathcal{R}_\varepsilon^1(s)\|^2_{L^2(\Lambda)} + \left\|\mathcal{S}_\varepsilon^1(s)\right\|^2_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}s\right].
\end{aligned}
\tag{5.44}
$$

Since $\mathcal{R}_\varepsilon^1 \to 0$ as $\varepsilon \to 0$ for almost all $(\omega, t, \lambda) \in \Omega \times [0,T] \times \Lambda$, we obtain using the dominated convergence theorem

$$
\lim_{\varepsilon\to 0}\mathbb{E}\left[\int_0^T \|\mathcal{R}_\varepsilon^1(t)\|^2_{L^2(\Lambda)}\mathrm{d}t\right] = 0. \tag{5.45}
$$

Here, we used that $\mathcal{R}_\varepsilon^1$ is dominated in the following way: By assumption (5.5), Taylor's formula and elementary estimates, we have

$$
|\mathcal{R}_\varepsilon^1| \leq C\left(1 + |x^u|^3 + |y^v|^3\right). \tag{5.46}
$$

The boundedness of the right hand side in $L^2(\Omega \times [0,T] \times \Lambda)$ follows immediately from Corollary 5.4 (notice that we obtain the boundedness of y^v in $L^6(\Omega \times [0,T] \times \Lambda)$ by the same arguments as for x^u). Furthermore, we have

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\int_0^T \left\|\mathcal{S}_\varepsilon^1(s)\right\|^2_{L_2(\Xi, L^2(\Lambda))} \mathrm{d}s\right] = 0 \tag{5.47}$$

since by the Lipschitz condition on Σ and the bound on the Fréchet derivative of Σ (see (5.3)), we have the following bound:

$$\begin{aligned}
&\left\|\mathcal{S}_\varepsilon^1(s)\right\|^2_{L_2(\Xi, L^2(\Lambda))} \\
&= \left\|\frac{1}{\varepsilon}\left(\Sigma(s, x_s^u + \varepsilon y_s^v) - \Sigma(s, x_s^u)\right) - \Sigma_x(s, x_s^u) y_s^v\right\|^2_{L_2(\Xi, L^2(\Lambda))} \\
&\leq 2\left\|\frac{1}{\varepsilon}\left(\Sigma(s, x_s^u + \varepsilon y_s^v) - \Sigma(s, x_s^u)\right)\right\|^2_{L_2(\Xi, L^2(\Lambda))} + 2\left\|\Sigma_x(s, x_s^u) y_s^v\right\|^2_{L_2(\Xi, L^2(\Lambda))} \\
&\leq C\left(1 + \|y_s^v\|^2_{L^2(\Lambda)}\right).
\end{aligned} \tag{5.48}$$

This completes the proof that z_ε converges to 0 in $L^2([0,T] \times \Omega; H^1(\Lambda))$ and in $L^2(\Omega; C([0,T]; L^2(\Lambda)))$. From the definition of $v \mapsto y^v$, it follows immediately that this is linear. Thus, for the Gâteaux differentiability it remains to show that $v \mapsto y^v$ is continuous. But this follows with the same arguments as in Proposition 5.5. □

As a corollary we obtain the following representation for the Gâteaux derivative of the cost functional.

Corollary 5.9. *For every $v \in L^6([0,T] \times \Lambda)$, the cost functional $J : L^6([0,T] \times \Lambda) \to \mathbb{R}$ is Gâteaux differentiable in the direction v with Gâteaux derivative*

$$\begin{aligned}
\frac{\partial J(u)}{\partial v} = \mathbb{E}\Bigg[&c_{\overline{\Lambda}} \int_0^T \int_\Lambda y_t^v(\lambda)\left(x_t^u(\lambda) - x_{\overline{\Lambda}}(t,\lambda)\right) \mathrm{d}\lambda \mathrm{d}t \\
&+ c_T \int_\Lambda y_T^v(\lambda)\left(x_T^u(\lambda) - x^T(\lambda)\right) \mathrm{d}\lambda + \nu \int_0^T \int_\Lambda u(t,\lambda) v(t,\lambda) \mathrm{d}\lambda \mathrm{d}t\Bigg],
\end{aligned} \tag{5.49}$$

where y^v denotes the variational solution of the SPDE (5.28).

Proof. Recall that the cost functional is given by

$$J(u) := I_1(x^u) + I_2(u), \tag{5.50}$$

where

$$I_1(x) := \mathbb{E}\left[\frac{c_{\overline{\Lambda}}}{2}\int_0^T\int_\Lambda \left(x_t(\lambda)-x_{\overline{\Lambda}}(t,\lambda)\right)^2 \mathrm{d}\lambda\mathrm{d}t + \frac{c_T}{2}\int_\Lambda \left(x_T(\lambda)-x^T(\lambda)\right)^2 \mathrm{d}\lambda\right] \tag{5.51}$$

and

$$I_2(u) := \frac{\nu}{2}\int_0^T\int_\Lambda u^2(t,\lambda)\mathrm{d}\lambda\mathrm{d}t. \tag{5.52}$$

Hence

$$\frac{\partial J(u)}{\partial v} = \frac{\partial I_1\left(x^u\right)}{\partial v} + \frac{\partial I_2(u)}{\partial v}. \tag{5.53}$$

Let $u \in L^6\left([0,T]\times\Lambda\right)$ be fixed. For $v \in L^6([0,T]\times\Lambda)$, we have for the Gâteaux derivative of I_2

$$\frac{\partial I_2(u)}{\partial v} = \lambda\int_0^T\int_\Lambda u(t,\lambda)v(t,\lambda)\mathrm{d}\lambda\mathrm{d}t. \tag{5.54}$$

On the other hand we have for the Gâteaux derivative of I_1

$$\frac{\partial I_1(x)}{\partial y} = \mathbb{E}\left[c_{\overline{\Lambda}}\int_0^T\int_\Lambda y\left(x-u_{\overline{\Lambda}}\right)\mathrm{d}\lambda\mathrm{d}t + c_T\int_\Lambda y\left(x-x^T\right)\mathrm{d}\lambda\right]. \tag{5.55}$$

Hence, by the chain rule, we obtain

$$\frac{\partial I_1\left(x^u\right)}{\partial v} = \mathbb{E}\left[c_{\overline{\Lambda}}\int_0^T\int_\Lambda \frac{\partial x^u}{\partial v}\left(x^u-x_{\overline{\Lambda}}\right)\mathrm{d}\lambda\mathrm{d}t + c_T\int_\Lambda \frac{\partial x^u}{\partial v}\left(x^u-x^T\right)\mathrm{d}\lambda\right], \tag{5.56}$$

which, together with equation (5.54) and Proposition 5.8, completes the proof. □

Now we can state a necessary condition for J to attain a minimum.

Theorem 5.10. *Let J attain a (local) minimum at $\bar{u} \in U_{ad}$. Then, for every $v \in U_{ad}$ we have*

$$\begin{aligned}\mathbb{E}\Bigg[&c_{\overline{\Lambda}}\int_0^T\int_\Lambda y_t^{v-\bar{u}}(\lambda)\left(\bar{x}_t(\lambda)-x_{\overline{\Lambda}}(t,\lambda)\right)\mathrm{d}\lambda\mathrm{d}t\\ &+c_T\int_\Lambda y_T^{v-\bar{u}}(\lambda)\left(\bar{x}_T(\lambda)-x^T(\lambda)\right)\mathrm{d}\lambda + \nu\int_0^T\int_\Lambda \bar{u}(t,\lambda)(v-\bar{u})(t,\lambda)\mathrm{d}\lambda\mathrm{d}t\Bigg] \geq 0.\end{aligned} \tag{5.57}$$

Proof. Let $v \in U_{\mathrm{ad}}$, and observe for $\varepsilon \in [0,1]$ that $\bar{u}+\varepsilon(v-\bar{u}) \in U_{\mathrm{ad}}$. Since $\bar{u}$ is a local minimizer, there exists a $1 \geq \varepsilon_0 > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have

$$J(\bar{u}) \leq J(\bar{u}+\varepsilon(v-\bar{u})). \tag{5.58}$$

which implies

$$\frac{1}{\varepsilon}\left(J(\bar{u}+\varepsilon(v-\bar{u}))-J(\bar{u})\right) \geq 0. \tag{5.59}$$

Letting ε tend to zero and plugging in the representation from Corollary 5.9 yields the claim. □

5.4. The Gradient of the Cost Functional

In this section, we are going to derive a representation for the gradient of the cost functional via adjoint calculus. Recall the state equation

$$\begin{cases} \mathrm{d}x_t^u = \left[\Delta x_t^u + b\left(x_t^u\right) + d(t)u(t)\right]\mathrm{d}t + \Sigma(t, x_t^u)\mathrm{d}W_t, & t \in [0,T] \\ x_0^u = x \in L^2(\Lambda). \end{cases} \tag{5.60}$$

In Section 5.3, we proved the following representation

$$\begin{aligned} \frac{\partial J(u)}{\partial v} = \mathbb{E}\Bigg[& c_{\overline{\Lambda}} \int_0^T \int_\Lambda y_t^v(\lambda)\left(x_t^u(\lambda) - u_{\overline{\Lambda}}(t,\lambda)\right)\mathrm{d}\lambda\mathrm{d}t \\ & + c_T \int_\Lambda y_T^v(\lambda)\left(x_T^u(\lambda) - x^T(\lambda)\right)\mathrm{d}\lambda + \nu \int_0^T \int_\Lambda u(t,\lambda)v(t,\lambda)\mathrm{d}\lambda\mathrm{d}t\Bigg], \end{aligned} \tag{5.61}$$

where y^v is the variational solution of

$$\begin{cases} \mathrm{d}y_t^v = [\Delta y_t^v + b'(x_t^u)y_t^v + d(t)v(t)]\mathrm{d}t + \Sigma_x(t, x_t^u)y_t^v\mathrm{d}W_t, & t \in [0,T] \\ y_0^v = 0 \in L^2(\Lambda). \end{cases} \tag{5.62}$$

The adjoint equation that is used in the existing literature is

$$\begin{cases} -\mathrm{d}p_t = \left[\Delta p_t + b'(x_t^u)p_t + c_{\overline{\Lambda}}\left(x_t^u - x_{\overline{\Lambda}}(t,\cdot)\right) + \partial_x\Sigma(t, x_t^u)^* q_t\right]\mathrm{d}t - q_t\mathrm{d}W_t \\ p_T = c_T\left(x_T^u - x^T\right) \in L^2(\Lambda), \end{cases} \tag{5.63}$$

for some processes $p \in L^2(\Omega \times [0,T]; H^1(\Lambda))$ and $q \in L^2(\Omega \times [0,T]; L_2(\Xi, L^2(\Lambda)))$. The derivation of the Stochastic Minimum Principle with this adjoint equation works in our setting as well. However, the numerical approximation of the solution of this adjoint equation is extremely costly. Therefore, we restrict our analysis to the case of additive noise in the state equation.

$$\begin{cases} \mathrm{d}x_t^u = \left[\Delta x_t^u + b\left(x_t^u\right) + d(t)u(t)\right]\mathrm{d}t + \sigma\mathrm{d}W_t, & t \in [0,T] \\ x_0^u = x \in L^2(\Lambda), \end{cases} \tag{5.64}$$

for some $\sigma \in L_2(\Xi, L^2(\Lambda))$. In this case, the linearized equation becomes

$$\begin{cases} \partial_t y_t^v = \Delta y_t^v + b'(x_t^u)y_t^v + d(t)v(t), & t \in [0,T] \\ y_0^v = 0 \in L^2(\Lambda), \end{cases} \tag{5.65}$$

which is a random partial differential equation (the coefficient $b'(x_t^u)$ is random). Now, we introduce the following random backward PDE for the adjoint state:

$$\begin{cases} -\partial_t p_t = \Delta p_t + b'(x^u) p_t + c_{\overline{\Lambda}} \left(x_t^u - x_{\overline{\Lambda}}(t, \cdot) \right), & t \in [0, T] \\ p_T = c_T \left(x_T^u - x^T \right) \in L^2(\Lambda). \end{cases} \tag{5.66}$$

One crucial point for our algorithm is the fact that the adjoint equation is a random backward PDE. The following adjoint state property is the main ingredient in the derivation of the gradient of the cost functional.

Proposition 5.11. *Let p be the solution of the adjoint equation* (5.66) *and let y^v be the solution of equation* (5.65) *associated with x^u. Then we have almost surely for every $v \in L^6([0, T] \times \Lambda)$,*

$$\int_0^T \int_\Lambda d(t) v(t) p_t \mathrm{d}\lambda \mathrm{d}t = \int_0^T \int_\Lambda c_{\overline{\Lambda}} (x_t^u - x_{\overline{\Lambda}}(t, \cdot)) y_t^v \mathrm{d}\lambda \mathrm{d}t + \int_\Lambda c_T (x_T^u - x^T) y_T^v \mathrm{d}\lambda. \tag{5.67}$$

Proof. By the deterministic integration by parts formula, it holds pathwise

$$y_T^v p_T - y_0^v p_0 = \int_0^T y_t^v \mathrm{d}p_t + \int_0^T p_t \mathrm{d}y_t^v. \tag{5.68}$$

Plugging in equations (5.65) and (5.66), respectively, this yields

$$\begin{aligned} y_T^v c_T (x_T^u - x^T) = &- \int_0^T y_t^v \left(\Delta p_t + b'(x_t^u) p_t + c_{\overline{\Lambda}} (x_t^u - x_{\overline{\Lambda}}(t, \cdot)) \right) \mathrm{d}t \\ &+ \int_0^T p_t \left(\Delta y_t^v + b'(x_t^u) y_t^v + d(t) v(t) \right) \mathrm{d}t. \end{aligned} \tag{5.69}$$

Integrating over Λ, and integrating the Laplace operator by parts, we obtain

$$\int_\Lambda y_T^v c_T (x_T^u - x^T) \mathrm{d}\lambda = \int_0^T \int_\Lambda d(t) v(t) p_t - c_{\overline{\Lambda}} (x_t^u - x_{\overline{\Lambda}}(t, \cdot)) y_t^v \mathrm{d}\lambda \mathrm{d}t, \tag{5.70}$$

which is the claimed result. □

As a corollary, we obtain the following representation for the gradient of the cost functional.

Theorem 5.12. *The gradient of the cost functional is given by*

$$\nabla J(u)(t, \lambda) = \mathbb{E} \left[d(t) p_t(\lambda) + \nu u(t, \lambda) \right], \tag{5.71}$$

where p is the solution of the adjoint equation

$$\begin{cases} -\partial_t p_t = \Delta p_t + b'(x_t^u) p_t + c_{\overline{\Lambda}} \left(x_t^u - x_{\overline{\Lambda}}(t, \cdot) \right), & t \in [0, T] \\ p_T = c_T \left(x_T^u - x^T \right) \in L^2(\Lambda). \end{cases} \tag{5.72}$$

Proof. By Corollary 5.9, we have

$$\frac{\partial J(u)}{\partial v} = \mathbb{E}\Bigg[c_{\overline{\Lambda}} \int_0^T \int_\Lambda y_t^v(\lambda) \left(x_t^u(\lambda) - x_{\overline{\Lambda}}(t,\lambda)\right) \mathrm{d}\lambda \mathrm{d}t + c_T \int_\Lambda y_T^v(\lambda) \left(x_T^u(\lambda) - x^T(\lambda)\right) \mathrm{d}\lambda + \nu \int_0^T \int_\Lambda u(t,\lambda)v(t,\lambda)\mathrm{d}\lambda \mathrm{d}t\Bigg], \quad (5.73)$$

where y^v denotes the variational solution of the random PDE (5.28). Now, by Proposition 5.11, this yields

$$\frac{\partial J(u)}{\partial v} = \mathbb{E}\left[\int_0^T \int_\Lambda d(t)v(t)p_t \mathrm{d}\lambda \mathrm{d}t + \nu \int_0^T \int_\Lambda u(t)v(t)\mathrm{d}\lambda \mathrm{d}t\right], \quad (5.74)$$

which completes the proof. □

Furthermore, by plugging this representation into the necessary condition derived in Theorem 5.10, we obtain the Stochastic Minimum Principle.

Theorem 5.13. *Let J attain a (local) minimum at $\bar{u} \in U_{ad}$. Then, for every $v \in U_{ad}$ we have*

$$\mathbb{E}\left[\int_0^T \int_\Lambda (d(t)p_t(\lambda) + \nu\bar{u}(t,\lambda))(v(t,\lambda) - \bar{u}(t,\lambda))\mathrm{d}\lambda \mathrm{d}t\right] \geq 0. \quad (5.75)$$

5.5. Nonlinear Conjugate Gradient Descent

Now that we have identified a representation for the gradient, we can apply a probabilistic nonlinear conjugate gradient descent method in order to approximate the optimal control. We are going to briefly sketch our algorithm here. For a survey of nonlinear conjugate gradient descent methods see [HZ06].

Let the initial control $u_0 \in L^6([0,T] \times \Lambda)$ be given and fix an initial step size $s_0 > 0$, as well as a stopping criterion $\eta > 0$. Then, the next control can be found as follows.

1. Solve the state equation

$$\begin{cases} \mathrm{d}x_t^n = \left[\Delta x_t^n + b\left(x_t^n\right) + d(t)u_n(t)\right] \mathrm{d}t + \sigma \mathrm{d}W_t, & t \in [0,T] \\ x_0^n = x \in L^2(\Lambda) \end{cases} \quad (5.76)$$

   for one realization of the noise.

2. Solve the adjoint equation

$$\begin{cases} -\partial_t p_t^n = \Delta p_t^n + b'(x_t^n)p_t^n + c_{\overline{\Lambda}}\left(x_t^n - x_{\overline{\Lambda}}(t,\cdot)\right), & t \in [0,T] \\ p_T^n = c_T\left(x_T^n - x^T\right) \in L^2(\Lambda) \end{cases}$$

   with the data given by the sample of the solution of the state equation that was calculated in Step 1.

3. Repeat Step 1 and Step 2 to approximate

$$\nabla J(u_n)(t, \lambda) = \mathbb{E}\left[d(t)p_t^n(\lambda) + \nu u_n(t, \lambda)\right]$$

via a Monte Carlo method.

4. The direction of descent is given by $D_n = -\nabla J(u_n) + \beta_n D_{n-1}$, where $\beta_n = \frac{\|\nabla J(u_n)\|}{\|\nabla J(u_{n-1})\|}$. (In the first step, $\beta_1 = 0$.)

5. Compute the new control via $u_{n+1} = u_n + s_n D_n$.

6. Accept or deny the new control: Again using a Monte Carlo method, we compare the costs under the new control with the costs under the old control. If the new control decreases the costs, we accept the new control and go back to Step 1. Otherwise, we decrease the step size $s_n = s_n/2$ and then go back to Step 5. (In our simulations, it has proven useful to accept the new control even if the costs are non-decreasing, once the step size gets too small, e.g. $s_n < 10^{-4}$.)

7. Stop if $\|\nabla J(u_n)\| < \eta$, otherwise reset the step size $s_n = s_0$ and go to step 1.

5.6. Numerical Experiments

In this section we want to present the application of the algorithm that was introduced in Section 5.5 to the stochastic Nagumo model. We are going to investigate two examples. The first one is to control the speed and the direction of travel of the wave front developing in the Nagumo model with additive noise; the second one is an example, where the optimal control of the deterministic system differs from the optimal control of the stochastic system. Corresponding results for the deterministic model can be found in the work by Buchholz et al., see [BEKT13a].

5.6.1. Steering of a Wave Front

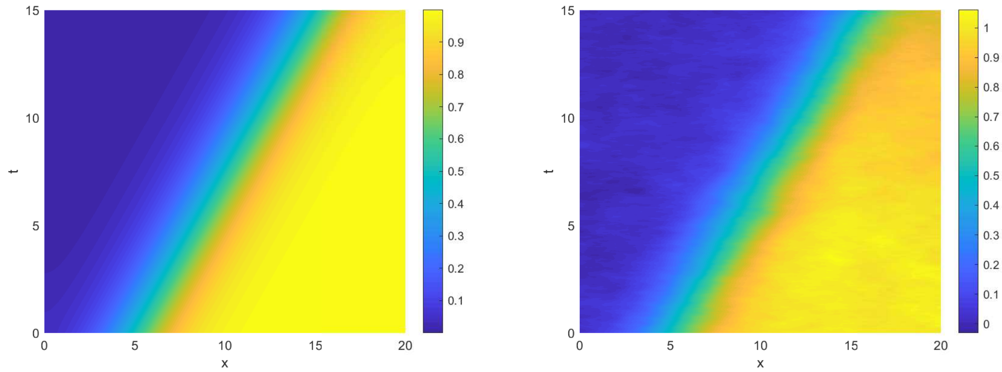
Let us first recall the stochastic Nagumo model. We consider the state equation

$$\begin{cases} \mathrm{d}x_t^u = \left[\Delta x_t^u + b\left(x_t^u\right) + d(t)u(t)\right]\mathrm{d}t + \sigma \mathrm{d}W_t, \quad t \in [0, T] \\ x_0^u = x \in L^2(\Lambda) \end{cases} \tag{5.77}$$

with homogeneous Neumann boundary conditions, where $d \equiv 1$, and the nonlinearity is of the form $b(x) = \gamma x(x-1)(a-x)$ for some $\gamma > 0$, $a \in (0, 1)$, i.e., the state equation takes the form

$$\begin{cases} \mathrm{d}x_t^u = \left[\Delta x_t^u + \gamma x_t^u(x_t^u - 1)(a - x_t^u) + u(t)\right]\mathrm{d}t + \sigma \mathrm{d}W_t, \quad t \in [0, T] \\ x_0^u = x \in L^2(\Lambda). \end{cases} \tag{5.78}$$

We choose $\Xi = L^2(\Lambda)$. In this case, the noise coefficient $\sigma \in L_2(L^2(\Lambda))$ can be associated with an intergral kernel $k \in L^2(\Lambda^2)$. We choose the correlation length of



(a) Deterministic Case, $\bar{\sigma} = 0$ (b) Stochastic Case, $\bar{\sigma} = 0.5$

Figure 5.1.: Uncontrolled Nagumo Equation

this integral kernel shorter than our chosen space discretization for the simulations. Therefore, we use independent Brownian motions with intensity $\bar{\sigma} \in \mathbb{R}$ on every node.

In our example, we choose the time-horizon $[0, 15]$, the space interval $\Lambda = [0, 20]$, $\gamma = 1$, and $a = 39/40$. As the initial condition we choose the wave profile

$$x(\lambda) = \left(1 + \exp\left(-\frac{\sqrt{2}}{2}(\lambda - 5)\right)\right)^{-1}. \tag{5.79}$$

In the deterministic case on an unbounded domain, these parameters lead to a traveling wave of the form $x(\lambda + ct)$, where $c = \sqrt{2}(a - \frac{1}{2}) \approx 0.672$, see [CG92]. Even though we are on a bounded domain, due to the homogeneous Neumann boundary conditions and the flatness of the wave profile close to the boundary we expect a similar behavior. Figure 5.1a displays a simulation in the deterministic case. Figure 5.1b shows one realization of the solution in the stochastic case with $\bar{\sigma} = 0.5$.

We can see that the traveling wave slowly travels to the right. Our objective is now to first speed up the wave and then change the direction of travel. To this end, we consider the cost functional given by

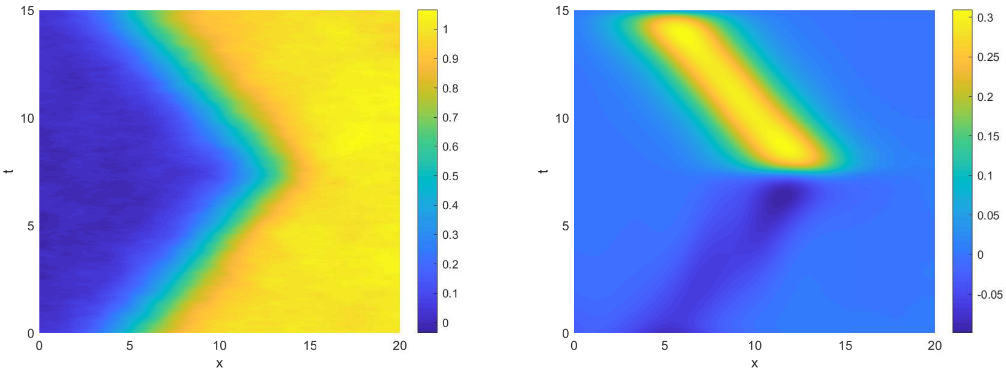
$$J(u) = \mathbb{E}\left[\frac{c_{\overline{\Lambda}}}{2}\int_0^T \int_\Lambda \left(x_t^u(\lambda) - x_{\overline{\Lambda}}(t,\lambda)\right)^2 \mathrm{d}\lambda \mathrm{d}t\right] \tag{5.80}$$

where $c_{\overline{\Lambda}} = 1$, and the reference profile $x_{\overline{\Lambda}}$ is given by

$$x_{\overline{\Lambda}}(t,\lambda) = \begin{cases} \left(1 + \exp\left(-\frac{\sqrt{2}}{2}(\lambda - t - 5)\right)\right)^{-1}, & t \leq \frac{T}{2} \\ \left(1 + \exp\left(-\frac{\sqrt{2}}{2}(\lambda - (T - t) - 5)\right)\right)^{-1}, & t > \frac{T}{2} \end{cases}, \tag{5.81}$$

for $(t, \lambda) \in [0, T] \times \Lambda$.

With the algorithm from Section 5.5 we can approximate the optimal control. Let



(a) Solution with Optimal Control

(b) Optimal Control

Figure 5.2.: Controlled Stochastic Nagumo Equation, $\bar{\sigma} = 0.5$

us apply the algorithm to the stochastic case with $\bar{\sigma} = 0.5$, the stopping criterion $\eta = 0.05$ and 100 Monte Carlo simulations for the approximation of the gradient. One realization of the solution with applied optimal control is displayed in Figure 5.2a. Figure 5.2b shows the corresponding optimal control.

5.6.2. Comparison with the Control of the Deterministic System

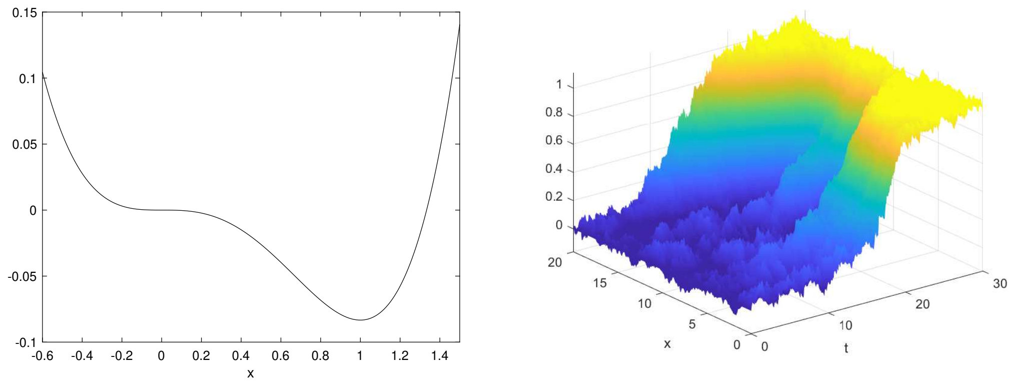
Simulations suggest that the optimal control for the deterministic system in the preceding example does not differ qualitatively from the optimal control for the stochastic system. This is because the fixed points 0 and 1 are stable. The situation changes, however, if one of the fixed points becomes unstable from one side, as the following example shows. Consider the state equation

$$\begin{cases} \mathrm{d}x_t^u = \left[\Delta x_t^u - (x_t^u)^3 + (x_t^u)^2 + u(t)\right] \mathrm{d}t + \sigma \mathrm{d}W_t, \quad t \in [0, T] \\ x_0^u = x \in L^2(\Lambda), \end{cases} \tag{5.82}$$

where $\Lambda = [0, 20]$, $T = 30$ and $\sigma \in L_2(\Xi, L^2(\Lambda))$. These choices lead to only one stable steady state, $x = 1$ and one unstable steady state $x = 0$, as illustrated by the potential of the nonlinearity in Figure 5.3a. Now, as initial condition, we choose $x_0^u = 0$, and consider the cost functional

$$J(u) = \mathbb{E}\left[\frac{1}{2}\int_\Lambda (x_T^u(\lambda))^2 \,\mathrm{d}\lambda\right], \tag{5.83}$$

i.e., we want the final state to be unchanged, in the unstable steady state 0. In the deterministic case, the optimal control is clearly $\bar{u} \equiv 0$, since we start in the steady state $x = 0$ and without any forcing, we stay in this state and accomplish the minimal possible costs $J(\bar{u}) = 0$. In the stochastic case, however, the noise term pushes the state out of the unstable steady state. Whenever the noise pushes the state above



(a) Potential (b) Solution without Control

Figure 5.3.: Uncontrolled Stochastic Nagumo Equation, One Unstable Steady State, $\bar{\sigma} = 1$

0, the dynamics of the state equation force the state towards the stable steady state $x = 1$. As an illustration of this effect, Figure 5.3b shows one realization in the stochastic case without a control function. When we introduce a control, the control tries to counteract this effect by keeping the state below 0 for times $t < T$. This effect can be seen in the simulations, as well. As the stopping criterion we used $\eta = 0.002$. Figure 5.4 displays the optimal control in the stochastic case with the same noise coefficient as in Section 5.6.1 and $\bar{\sigma} = 0.5$, and one realization of the corresponding state, respectively.

5.6.3. Mathematical Analysis in a Simplified Setting

Since we are not able to prove the previous result in that setting rigorously, we consider a simpler similar example in which the optimal control in the deterministic case and the optimal control in the stochastic case differ.

Let us consider the SDE

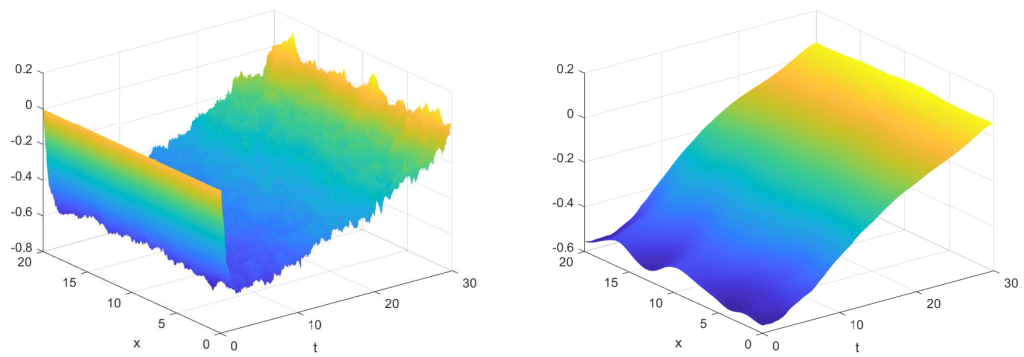
$$\begin{cases} \mathrm{d}x_t^u = [-\mathcal{P}'(x_t^u) + u(t)]\,\mathrm{d}t + \bar{\sigma}\mathrm{d}B_t, & t \in [0,T] \\ x_0^u = 0, \end{cases} \tag{5.84}$$

where $(B_t)_{t\geq 0}$ is a Brownian motion on $\mathbb{R}$, $\bar{\sigma} \in \mathbb{R}$, the potential $\mathcal{P} : \mathbb{R} \to \mathbb{R}$ is given by

$$\mathcal{P}(x) = \begin{cases} \frac{1}{2}(\arctan(x) - x), & \text{for } x \geq 0 \\ 0, & \text{for } x < 0, \end{cases} \tag{5.85}$$

and hence $-\mathcal{P}'$ is given by

$$-\mathcal{P}'(x) = \begin{cases} \frac{x^2}{2(1+x^2)}, & \text{for } x \geq 0 \\ 0, & \text{for } x < 0. \end{cases} \tag{5.86}$$



(a) Solution with Optimal Control

(b) Optimal Control

Figure 5.4.: Controlled Stochastic Nagumo Equation, One Unstable Steady State, $\bar{\sigma} = 0.5$

Notice that this potential qualitatively resembles the potential used in the previous example in the interval $[0, 1]$. That is why we observe a similar effect in this example. We consider the cost functional

$$J(u) := \mathbb{E}\left[\frac{1}{2}(x_T^u)^2\right]. \tag{5.87}$$

As in the previous example, the initial condition and the desired final state are both the unstable steady state $x = 0$. Hence, in the deterministic case ($\bar{\sigma} = 0$), the optimal control is given by $\bar{u} \equiv 0$, since the constant function $x \equiv 0$ solves the deterministic equation without control and the associated costs are zero.

Now, we are going to show that the optimal control in the stochastic case ($\bar{\sigma} > 0$), however, is not equal to zero. First, notice that the adjoint equation associated with our control problem is given by

$$\begin{cases} -\partial_t p_t = -\mathcal{P}''(x_t^u)p_t, & t \in [0, T] \\ p_T = x_T^u, \end{cases} \tag{5.88}$$

where $-\mathcal{P}''$ is given by

$$-\mathcal{P}''(x) = \begin{cases} \frac{x}{(1+x^2)^2}, & \text{for } x \geq 0 \\ 0, & \text{for } x < 0. \end{cases} \tag{5.89}$$

Hence, the solution of the adjoint equation is given explicitly by

$$p_t = x_T^u \exp\left(\int_t^T -\mathcal{P}''(x_s^u)\mathrm{d}s\right), \tag{5.90}$$

and the gradient of the cost functional is given by

$$\nabla J(u)(t) = \mathbb{E}[p_t] = \mathbb{E}\left[x_T^u \exp\left(\int_t^T -\mathcal{P}''(x_s^u)\mathrm{d}s\right)\right]. \tag{5.91}$$

Now, we are going to show that the gradient for $u \equiv 0$ is not equal to zero and hence, $u \equiv 0$ is not an optimal control. To this end, consider

$$\partial_t(\nabla J(u))(t) = \mathbb{E}[\partial_t p_t] = \mathbb{E}\left[\mathcal{P}''(x_t^u)x_T^u \exp\left(\int_t^T -\mathcal{P}''(x_s^u)\mathrm{d}s\right)\right]. \tag{5.92}$$

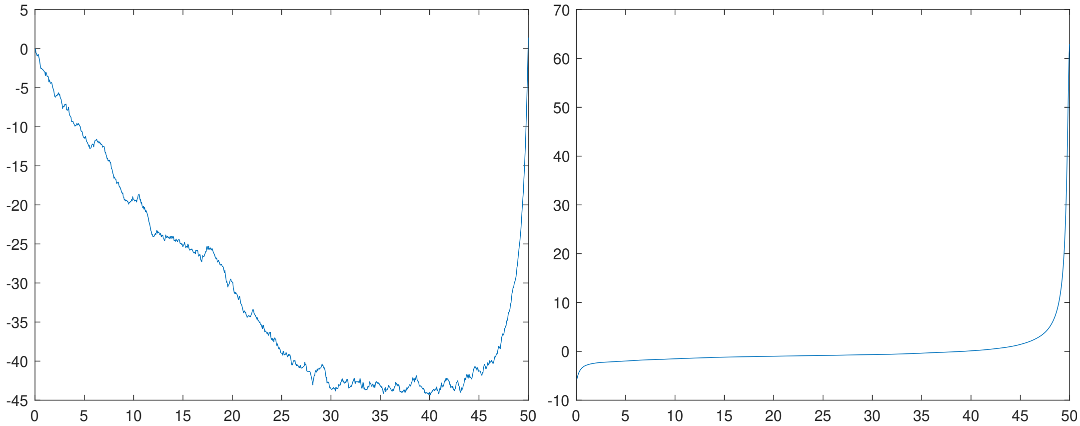
This yields

$$\begin{aligned}
&\liminf_{t\to T}\{-\partial_t(\nabla J(u))(t)\} \\
&= \liminf_{t\to T}\mathbb{E}\left[-\mathcal{P}''(x_t^u)x_T^u \exp\left(\int_t^T -\mathcal{P}''(x_s^u)\mathrm{d}s\right)\right] \\
&\geq \mathbb{E}\left[\liminf_{t\to T}\left\{-\mathcal{P}''(x_t^u)x_T^u \exp\left(\int_t^T -\mathcal{P}''(x_s^u)\mathrm{d}s\right)\right\}\right] \\
&= \mathbb{E}\left[-\mathcal{P}''(x_T^u)x_T^u\right] \\
&= \mathbb{E}\left[\frac{(x_T^u)^2}{\left(1+\left(x_T^u\right)^2\right)^2}1_{\{x_T^u>0\}}\right] > 0,
\end{aligned} \tag{5.93}$$

where the last part is strictly positive since x_T has a strictly positive density with respect to the Lebesgue measure. Therefore, the gradient is not equal to zero and thus, $u \equiv 0$ is not an optimal control.

Remark 5.14. Notice that we did not use that $u \equiv 0$ in this proof. This shows, that the optimal control in the stochastic case is unbounded.

Figure 5.5 illustrates our results in case of the SDE (5.84) as the constraint and the cost functional (5.87).



(a) Solution with Optimal Control (b) Optimal Control

Figure 5.5.: Controlled SDE, One Unstable Steady State, $\bar{\sigma} = 1$

Bibliography

[AF03] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, Elsevier, 2003.

[BEKT13a] R. Buchholz, H. Engel, E. Kammann, and F. Tröltzsch, *On the optimal control of the Schlögl-model*, Comput. Optim. Appl. **56** (2013), 153–185.

[BEKT13b] ———, *Erratum to: On the optimal control of the Schlögl-model*, Comput. Optim. Appl. **56** (2013), 187–188.

[Bel57] R. Bellman, *Dynamic programming*, Princeton University Press, 1957.

[Ben83] A. Bensoussan, *Stochastic maximum principle for distributed parameter systems*, J. Franklin Inst. **315** (1983), 387–406.

[Bis73] J. M. Bismut, *Conjugate convex functions in optimal stochastic control*, J. Math. Anal. Appl. **44** (1973), 384–404.

[Bis78] ———, *An introductory approach to duality in optimal stochastic control*, SIAM Rev. **20** (1978), 62–78.

[CDP18] F. Cordoni and L. Di Persio, *Optimal control for the stochastic FitzHugh-Nagumo model with recovery variable*, Evol. Eq. Control Theory **7** (2018), 571–585.

[CDP21] ———, *Optimal control of the FitzHugh-Nagumo stochastic model with nonlinear diffusion*, Appl. Math. Optim. **84** (2021), 2947–2968.

[CF91] P. Cannarsa and H. Frankowska, *Some characterizations of optimal trajectories in control theory*, SIAM J. Control Optim. **29** (1991), 1322–1347.

[CF92] ———, *Value function and optimality conditions for semilinear control problems*, Appl. Math. Optim. **26** (1992), 139–169.

[CG92] Z. X. Chen and B. Y. Guo, *Analytic solutions of the Nagumo equation*, IMA J. Appl. Math. **48** (1992), 107–115.

[CKSY22] J. Chessari, R. Kawai, Y. Shinozaki, and T. Yamada, *Numerical methods for backward stochastic differential equations: A survey*, preprint, https://arxiv.org/abs/2101.08936, 2022.

[CL83] M. G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc. **277** (1983), 1–42.

[CL90] ______, *Viscosity solutions of Hamilton-Jacobi equations in infinite dimensions. IV. Hamiltonians with unbounded linear terms*, J. Funct. Anal. **90** (1990), 237–283.

[CL91] ______, *Viscosity solutions of Hamilton-Jacobi equations in infinite dimensions. V. Unbounded linear terms and B-continuous solutions*, J. Funct. Anal. **97** (1991), 417–465.

[DD12] F. Demengel and G. Demengel, *Functional spaces for the theory of elliptic partial differential equations*, Springer, 2012.

[DFT07] A. Debussche, M. Fuhrman, and G. Tessitore, *Optimal control of a stochastic heat equation with boundary-noise and boundary-control*, ESAIM Control Optim. Calc. Var. **13** (2007), 178–205.

[DM13] K. Du and Q. Meng, *A maximum principle for optimal control of stochastic evolution equations*, SIAM J. Control Optim. **51** (2013), 4343–4362.

[DMPV19] T. Dunst, A. K. Majee, A. Prohl, and G. Vallet, *On stochastic optimal control in ferromagnetism*, Arch. Ration. Mech. Anal. **233** (2019), 1383–1440.

[DP16] T. Dunst and A. Prohl, *The forward-backward stochastic heat equation: Numerical analysis and simulation*, SIAM J. Sci. Comput. **38** (2016), A2725–A2755.

[DPZ14] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge University Press, 2014.

[ET10] G. B. Ermentrout and D. H. Terman, *Mathematical foundations of neuroscience*, Springer, 2010.

[Fab06] G. Fabbri, *First order HJB equations in Hilbert spaces and applications*, Doctoral thesis, Università degli Studi di Roma – La Sapienza, Rome, Italy, 2006.

[FG95] F. Flandoli and D. Gatarek, *Martingale and stationary solutions for stochastic Navier-Stokes equations*, Probab. Theory Related Fields **102** (1995), 367–391.

[FG18] S. Federico and F. Gozzi, *Verification theorems for stochastic optimal control problems in Hilbert spaces by means of a generalized Dynkin formula*, Ann. Appl. Probab. **28** (2018), 3558–3599.

[FGG11] S. Federico, B. Goldys, and F. Gozzi, *HJB equations for the optimal control of differential equations with delays and state constraints, II: Verification and optimal feedbacks*, SIAM J. Control Optim. **49** (2011), 2378–2414.

[FGŚ10] G. Fabbri, F. Gozzi, and A. Święch, *Verification theorem and construction of ϵ-optimal controls for control of abstract evolution equations*, J. Convex Anal. **17** (2010), 611–642.

[FGŚ17] ———, *Stochastic optimal control in infinite dimension*, Springer, 2017.

[FHT12] M. Fuhrman, Y. Hu, and G. Tessitore, *Stochastic maximum principle for optimal control of SPDEs*, C. R. Math. Acad. Sci. Paris **350** (2012), 683–688.

[FHT13] ———, *Stochastic maximum principle for optimal control of SPDEs*, Appl. Math. Optim. **68** (2013), 181–217.

[FHT18] ———, *Stochastic maximum principle for optimal control of partial differential equations driven by white noise*, Stoch. Partial Differ. Equ. Anal. Comput. **6** (2018), 255–285.

[Fit61] R. FitzHugh, *Impulses and physiological states in theoretical models of nerve membrane*, Biophys. J. **1** (1961), 445–466.

[FO16] M. Fuhrman and C. Orrieri, *Stochastic maximum principle for optimal control of a class of nonlinear SPDEs with dissipative drift*, SIAM J. Control Optim. **54** (2016), 341–371.

[FR17] G. Fabbri and R. Russo, *HJB equations in infinite dimension and optimal control of stochastic evolution equations via generalized Fukushima decomposition*, SIAM J. Control Optim. **55** (2017), 4072–4091.

[FS06] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, Springer, 2006.

[FT02] M. Fuhrman and G. Tessitore, *Nonlinear Kolmogorov equations in infinite dimensional spaces: The backward stochastic differential equations approach and applications to optimal control*, Ann. Probab. **30** (2002), 1397–1465.

[FT04] ———, *Infinite horizon backward stochastic differential equations and elliptic equations in Hilbert spaces*, Ann. Probab. **32** (2004), 607–660.

[FT05] ———, *Generalized directional gradients, backward stochastic differential equations and mild solutions of semilinear parabolic equations*, Appl. Math. Optim. **51** (2005), 279–332.

[FZ20] H. Frankowska and X. Zhang, *Necessary conditions for stochastic optimal control problems in infinite dimensions*, Stochastic Process. Appl. **130** (2020), 4081–4103.

[GŚZ05] F. Gozzi, A. Święch, and X. Y. Zhou, *A corrected proof of the stochastic verification theorem within the framework of viscosity solutions*, SIAM J. Control Optim. **43** (2005), 2009–2019.

[GŚZ10] ______, *Erratum: "a corrected proof of the stochastic verification theorem within the framework of viscosity solutions"*, SIAM J. Control Optim. **48** (2010), 4177–4179.

[GT14] G. Guatteri and G. Tessitore, *Well posedness of operator valued backward stochastic Riccati equations in infinite dimensional spaces*, SIAM J. Control Optim. **52** (2014), 3776–3806.

[Gua11] G. Guatteri, *Stochastic maximum principle for SPDEs with noise and control on the boundary*, Systems Control Lett. **60** (2011), 198–204.

[Hai09] M. Hairer, *An introduction to stochastic PDEs*, preprint, https://arxiv.org/abs/0907.4178, 2009.

[HH52] A. L. Hodgkin and A. F. Huxley, *A quantitative description of membrane current and its application to conduction and excitation in nerve*, J Physiol. **117** (1952), 500–544.

[HP90] Y. Hu and S. Peng, *Maximum principle for semilinear stochastic evolution control systems*, Stochastics Stochastics Rep. **33** (1990), 159–180.

[Hu19] Y. Hu, *Stochastic maximum principle*, in Encyclopedia of Systems and Control (J. Baillieul and T. Samad, eds.), Springer, 2019, pp. 1–4.

[HZ06] W. W. Hager and H. Zhang, *A survey of nonlinear conjugate gradient methods*, Pac. J. Optim. **2** (2006), 35–58.

[Kry13] N. V. Krylov, *A relatively short proof of Itô's formula for SPDEs and its applications*, Stoch. Partial Differ. Equ. Anal. Comput. **1** (2013), 152–174.

[KS91] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, Springer, 1991.

[Lis02] H. Lisei, *Existence of optimal and ε-optimal controls for the stochastic Navier-Stokes equation*, Nonlinear Anal. **51** (2002), 95–118.

[LPS14] G. J. Lord, C. E. Powell, and T. Shardlow, *An introduction to computational stochastic PDEs*, Cambridge University Press, 2014.

[LR15] W. Liu and M. Röckner, *Stochastic partial differential equations: An introduction*, Springer, 2015.

[LZ14] Q. Lü and X. Zhang, *General Pontryagin-type stochastic maximum principle and backward stochastic evolution equations in infinite dimensions*, Springer, 2014.

[LZ15] ———, *Transposition method for backward stochastic evolution equations revisited, and its application*, Math. Control Relat. Fields **5** (2015), 529–555.

[LZ18] ———, *Operator-valued backward stochastic Lyapunov equations in infinite dimensions, and its application*, Math. Control Relat. Fields **8** (2018), 337–381.

[Mur02] J. D. Murray, *Mathematical biology. I*, Springer, 2002.

[Mur03] ———, *Mathematical biology. II*, Springer, 2003.

[NAY62] J. Nagumo, S. Arimoto, and S. Yoshizawa, *An active pulse transmission line simulating nerve axon*, Proceedings of the IRE **50** (1962), 2061–2070.

[Par21] E. Pardoux, *Stochastic partial differential equations: An introduction*, Springer, 2021.

[Paz83] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Springer, 1983.

[PBGM62] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The mathematical theory of optimal processes*, Interscience Publishers John Wiley & Sons, Inc., 1962.

[Pen90] S. Peng, *A general stochastic maximum principle for optimal control problems*, SIAM J. Control Optim. **28** (1990), 966–979.

[RLM$^+$16] C. Ryll, J. Löber, S. Martens, H. Engel, and F. Tröltzsch, *Analytical, optimal, and sparse optimal control of traveling wave solutions to reaction-diffusion systems*, Control of self-organizing nonlinear systems (E. Schöll, S. H. L. Klapp, and P. Hövel, eds.), Springer, 2016, pp. 189–210.

[Rou13] T. Roubíček, *Nonlinear partial differential equations with applications*, Birkhäuser/Springer Basel AG, Basel, 2013.

[RS80] M. Reed and B. Simon, *Methods of modern mathematical physics. I*, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], 1980.

[Ryl17] C. Ryll, *Optimal control of patterns in some reaction-diffusion-systems*, Doctoral thesis, Technische Universität Berlin, Berlin, Germany, 2017.

[Sch72] F. Schlögl, *Chemical reaction models for non-equilibrium phase transitions*, Z. Physik **253** (1972), 147–161.

[Sko56] A. V. Skorohod, *Limit theorems for stochastic processes*, Teor. Veroyatnost. i Primenen. **1** (1956), 289–319.

[SW21a] W. Stannat and L. Wessels, *Deterministic control of stochastic reaction-diffusion equations*, Evol. Equ. Control Theory **10** (2021), 701–722.

[SW21b] ______, *Peng's maximum principle for stochastic partial differential equations*, SIAM J. Control Optim. **59** (2021), 3552–3573.

[SW22] ______, *Necessary and sufficient conditions for optimal control of semilinear stochastic partial differential equations*, preprint, https://arxiv.org/abs/2112.09639, 2022.

[Świ94] A. Święch, *"Unbounded" second order partial differential equations in infinite-dimensional Hilbert spaces*, Comm. Partial Differential Equations **19** (1994), 1999–2036.

[TL94] S. J. Tang and X. J. Li, *Maximum principle for optimal control of distributed parameter stochastic systems with random jumps*, in Differential equations: Dynamical Systems, and Control Science (K. D. Elworthy, W. N. Everitt, and E. B. Lee, eds.), Dekker, 1994, pp. 867–890.

[YZ99] J. Yong and X. Y. Zhou, *Stochastic controls: Hamiltonian systems and HJB equations*, Springer, 1999.

[Zei86] E. Zeidler, *Nonlinear functional analysis and its applications. I*, Springer, 1986.

[Zho91a] X. Y. Zhou, *Remarks on optimal controls of stochastic partial differential equations*, Systems Control Lett. **16** (1991), 465–472.

[Zho91b] ______, *A unified treatment of maximum principle and dynamic programming in stochastic controls*, Stochastics Stochastics Rep. **36** (1991), 137–161.

[ZYL97] X. Y. Zhou, J. Yong, and X. Li, *Stochastic verification theorems within the framework of viscosity solutions*, SIAM J. Control Optim. **35** (1997), 243–253.