

A FULLY ADAPTIVE ALGORITHM FOR PDE-CONSTRAINED OPTIMAL CONTROL WITH AND WITHOUT UNCERTAINTY, PART 2: THE STOCHASTIC CASE AND LOW-RANK TENSOR METHODS*

SEBASTIAN GARREIS[†] AND MICHAEL ULBRICH[†]

Abstract. An adaptive algorithm for the numerical solution of a class of optimal control problems under uncertainty is presented. It is governed by a semilinear elliptic partial differential equation (PDEs) with uncertain coefficients. The existence of optimal solutions is shown and an adjoint-based representation of the reduced gradient is derived. Error estimates from the companion paper are generalized to the stochastic setting to apply an inexact trust-region algorithm to the problem. An adaptive solution technique with low-rank tensors for the arising PDEs is presented. It balances the error contributions stemming from the discretization of the physical and stochastic spaces as well as from the inexact solution of the discrete system. Numerical experiments complement the developed theory.

Key words. optimal control under uncertainty, adaptive algorithms, stochastic Galerkin discretization, a posteriori error estimation, low-rank tensor methods

AMS subject classifications. 15A69, 35J61, 35R60, 47H05, 47H30, 49J55, 65M70, 65N22, 90C15, 93E20

1. Introduction. We extend the results of [23], where we analyzed an adaptive inexact trust-region method for optimal control problems with partial differential equations (PDEs), and develop an inexact optimization algorithm with adaptive error control for the solution of optimal control problems *under uncertainty*. This introduces additional sources of inexactness that need to be controlled, in particular those stemming from the stochastic parameters. They increase the number of independent variables significantly in comparison with the spatial variables. For tackling this we work with low-rank tensors to represent the discretized state and other quantities. In addition to the sources of inexactness in the deterministic case, the stochastic discretization error and the algebraic error coming from low-rank tensor solvers need to be estimated and controlled. The state equation is a semilinear, elliptic PDE with stochastic coefficients and a smooth, monotone nonlinearity. It is formulated in a suitable Bochner space setting. In contrast to the linear case, which is covered in several papers [7, 16, 27, 39], it is not possible to choose the Hilbert space $L^2_{\mathbb{P}}(\Xi; H^1_0(\Omega))$ as state space, where $(\Xi, \mathcal{F}, \mathbb{P})$ is a probability space and $\Omega \subset \mathbb{R}^n$ is the spatial domain. Rather, we have to work in the Banach space $L^p_{\mathbb{P}}(\Xi; H^1_0(\Omega))$, $p \in (2, \infty)$, to make the semilinear nonlinearity well-defined and continuously differentiable. This makes the analysis of the equation more involved, in particular the proof of the existence of a unique solution. This PDE with random coefficients governs an optimal control problem with tracking type cost function and convex control constraints. The aim is to find a deterministic control prior to the observation of the uncertainty, cf. [22, 33, 34], whereas other works [2, 10] compute optimal controls for specific realizations of the random parameters. Due to the uncertainty in the state equation, the state of the system is a Banach space valued random variable. Hence, the optimal control shall minimize the objective *in expectation* in this paper, cf. [22, 32, 33]. Alternative approaches aiming for a *risk-averse* control additionally penalize the variance [1, 5] or minimize the conditional value-at-risk (tail expectation) of the objective [21, 34, 35]. Since our developments of an adaptive inexact optimization method with a rigorous and implementable

*Submitted to the editors DATE.

Funding: The work of the first author was partially supported by “TopMath”, which is a graduate program of the Elite Network of Bavaria, Germany, and a graduate center of TUM Graduate School, and by the DFG/FWF International Research training Group (IGDK Munich – Graz) “Optimization and Analysis for Partial Differential Equations with Nonsmooth Structures.”

[†]Technical University of Munich, Chair of Mathematical Optimization, Department of Mathematics, Boltzmannstr. 3, 85748 Garching b. München, Germany (garreis@ma.tum.de, mulbrich@ma.tum.de).

error estimation framework is quite complex in itself, we focus on the risk-neutral situation here where the expectation of the stochastic cost is minimized.

We show the existence of a solution to the optimal control problem and derive an adjoint based expression for the gradient of the reduced objective function. This is not as straightforward as in the deterministic case since the linearized state equation operator is not boundedly invertible as an operator mapping from $L^p_{\mathbb{P}}(\Xi; H_0^1(\Omega))$ to its dual. To apply the inexact trust-region algorithm presented in the companion paper [23], we derive bounds on the computed state and adjoint state which ensure the required accuracy of the computed objective function and gradient evaluation. As a result, it can be sufficient to control the state and adjoint state error in the $L^2_{\mathbb{P}}(\Xi; H_0^1(\Omega))$ norm although the integrability exponent of the state space is $p > 2$. This enables us to solve the equations adaptively using a stochastic Galerkin discretization and an a posteriori error estimate, cf. [7, 8, 16, 17]. The discretization is based on the tensor product structure of the parameter space $\Xi = \times_{i=1}^m \Xi_i$. We use a full tensor product polynomial basis for the discretization of $L^p_{\mathbb{P}}(\Xi)$ and finite elements (FE) for the deterministic spaces, such as $H_0^1(\Omega)$. The coefficients are tensors (multi-dimensional arrays), which quickly become infeasibly big if, e.g., the number of parameters m is large. To make the computations efficient and save memory, we represent them in a modern low-rank tensor format [29, 37] as in our previous work [22]. This comes with the drawback that only a limited set of arithmetic operations is available within these formats. Hence, the discretized operators and the a posteriori error indicators we propose are such that they can be evaluated efficiently with low-rank tensors. The developed error estimator is similar to [16, 18], but measures the error in a reference norm, makes a more rigorous use of a deterministic reference operator, cf. [7], is applicable to semilinear equations, and uses a polynomial basis which is more suitable for optimization if it comes to, e.g., state constraints.

Alternative approaches using low-rank tensors either use a fixed discretization [5, 6, 22] or solve only a linear PDE adaptively [18] instead of an optimal control problem governed by a semilinear PDE. Furthermore, alternative discretizations such as a multilevel Monte Carlo method [2] or adaptive sparse grids [9, 32, 33] with possible speed-ups by using a reduced basis method [11] or proper orthogonal decomposition [10] have been considered.

As the first part, this paper is based on and uses parts of the dissertation [20]. In section 2, the stochastic model problem is analyzed. Section 3 shows how the error estimates required in the inexact trust-region algorithm [23] can be fulfilled based on suitable error estimates in the state and adjoint state of the problem as well as the error in the inexact projection onto the feasible set. This resembles the strategy for the deterministic case and uses the corresponding results from [23]. In section 4 the a posteriori error estimator and adaptive solution technique for the considered PDEs with uncertain coefficients is presented. The numerical results in section 5 show how the algorithm adapts to the problem data by choosing suitable polynomial degrees and locally refined FE meshes.

2. Optimal control problem under uncertainty. Let Y and U be Hilbert spaces of deterministic functions and $\Xi \subset \mathbb{R}^m$ is equipped with the Borel σ -algebra and a probability measure \mathbb{P} . The pointwise formulation of the state equation is of the form

$$(2.1) \quad E[\xi](y(\xi), u) = A(\xi)y(\xi) + N(y(\xi)) - B(\xi)u - b(\xi) = 0 \in Y^*$$

and shall hold for almost every (a.e.) $\xi \in \Xi$. We have strongly measurable functions $A : \Xi \rightarrow \mathcal{L}(Y, Y^*)$, $B : \Xi \rightarrow \mathcal{L}(U, Y^*)$, and $b : \Xi \rightarrow Y^*$, and a continuous, monotone, possibly nonlinear operator $N : Y \rightarrow Y^*$. $A(\xi)$ shall be strongly monotone (in the sense of [45, Def. 25.2]) with constant $\underline{\kappa}$ for a.e. $\xi \in \Xi$. $y(\xi) \in Y$ is the parameter-dependent state and $u \in U$ the deterministic control. In this paper, we put the focus on the stochasticity of the problem and apply some results of its deterministic variant discussed in [23]. In particular,

we use Bochner spaces of $L_{\mathbb{P}}^p(\Xi)$ random variables with values in a Banach space, such as the state space $\mathbf{Y} := L_{\mathbb{P}}^p(\Xi; Y)$. To derive a suitable function space setting and apply a stochastic Galerkin discretization, we consider the weak formulation of (2.1) w.r.t. ξ , i.e.,

$$(2.2) \quad \mathbf{E}(\mathbf{y}, u) := \mathbf{A}\mathbf{y} + \mathbf{N}(\mathbf{y}) - \mathbf{B}u - \mathbf{b} = 0 \in \mathbf{Y}^* = L_{\mathbb{P}}^p(\Xi; \mathbf{Y}^*),$$

where $\frac{1}{p} + \frac{1}{p^*} = 1$. This equation is obtained by testing (2.1) and integrating w.r.t. ξ , i.e., we define

$$(2.3) \quad \begin{aligned} \mathbf{A} : \mathbf{Y} &\rightarrow \mathbf{Y}^*, \langle \mathbf{A}\mathbf{y}, \mathbf{v} \rangle_{\mathbf{Y}^*, \mathbf{Y}} := \int_{\Xi} \langle A(\xi)\mathbf{y}(\cdot, \xi), \mathbf{v}(\cdot, \xi) \rangle_{Y^*, Y} d\mathbb{P} \\ \mathbf{N} : \mathbf{Y} &\rightarrow \mathbf{Y}^*, \langle \mathbf{N}(\mathbf{y}), \mathbf{v} \rangle_{\mathbf{Y}^*, \mathbf{Y}} := \int_{\Xi} \langle N(\mathbf{y}(\cdot, \xi)), \mathbf{v}(\cdot, \xi) \rangle_{Y^*, Y} d\mathbb{P} \\ \mathbf{B} : U &\rightarrow \mathbf{Y}^*, \langle \mathbf{B}(u), \mathbf{v} \rangle_{\mathbf{Y}^*, \mathbf{Y}} := \int_{\Xi} \langle B(\xi)u, \mathbf{v}(\cdot, \xi) \rangle_{Y^*, Y} d\mathbb{P} \\ \mathbf{b} \in \mathbf{Y}^*, &\langle \mathbf{b}, \mathbf{v} \rangle_{\mathbf{Y}^*, \mathbf{Y}} := \int_{\Xi} \langle b(\xi), \mathbf{v}(\cdot, \xi) \rangle_{Y^*, Y} d\mathbb{P} \end{aligned}$$

The exponent $p \in [2, \infty)$ has to be chosen such that all operators are well-defined and twice continuously differentiable, in particular, the nonlinear operator \mathbf{N} . Usually, we identify $\mathbf{Y} \cong Y \otimes L_{\mathbb{P}}^p(\Xi)$ (possible for $p < \infty$) and view \mathbf{y} as a function of two variables, the first one corresponding to the independent variables of the function space Y and the second one being ξ . In Proposition 2.6 we show the equivalence of the formulations (2.1) and (2.2) for a concrete class of semilinear elliptic PDEs.

Let now H be a real Hilbert space, define $\mathbf{H} := L_{\mathbb{P}}^2(\Xi; H)$, and choose a desired state $\hat{\mathbf{q}} \in \mathbf{H}$ and a parametrized state-to-observation operator $Q(\xi)$.

Assumption 2.1. $\hat{\mathbf{q}} \in L_{\mathbb{P}}^{r_{\hat{\mathbf{q}}}}(\Xi, H) \subset \mathbf{H}$ is satisfied with some $r_{\hat{\mathbf{q}}} \in [2, \infty]$. Furthermore, we have $Q(\cdot) \in L_{\mathbb{P}}^{r_Q}(\Xi; \mathcal{L}(Y, H))$ with $r_Q \in [\frac{2p}{p-2}, \infty]$.

Under this assumption, the operator $Q(\cdot)$ can be identified with $\mathbf{Q} \in \mathcal{L}(\mathbf{Y}, \mathbf{H})$ by

$$(2.4) \quad \langle \mathbf{Q}\mathbf{y}, \mathbf{v} \rangle_{\mathbf{Y}^*, \mathbf{Y}} = \int_{\Xi} \langle Q(\xi)\mathbf{y}(\cdot, \xi), \mathbf{v}(\cdot, \xi) \rangle_{Y^*, Y} d\mathbb{P}.$$

By measurability and Hölder's inequality, the function

$$(2.5) \quad \xi \mapsto J[\xi](\mathbf{y}(\cdot, \xi), u) := \frac{1}{2} \|\mathbf{Q}\mathbf{y} - \hat{\mathbf{q}}(\cdot, \xi)\|_H^2 + \frac{\gamma}{2} \|u\|_U^2$$

with $\gamma > 0$ is an $L_{\mathbb{P}}^1(\Xi)$ random variable. $J[\cdot]$ is called the *random variable objective function* [35]. We want to minimize it prior to the observation of the uncertainty $\xi \in \Xi$ using a *deterministic* control u . To determine which probability distributions of the objective we prefer, we apply a *risk measure* $\mathcal{R} : L_{\mathbb{P}}^1(\Xi) \rightarrow \mathbb{R}$. In this paper, we discuss the *risk-neutral* [40, sec. 6.4] case $\mathcal{R} \equiv \mathbb{E}$. For recent works on risk-averse, PDE-constrained optimization we refer to [34, 35, 21]. With $U_{\text{ad}} \subset U$ nonempty, closed, and convex, the considered optimal control problem is

$$(2.6) \quad \min_{\mathbf{y} \in \mathbf{Y}, u \in U} \mathbf{J}(\mathbf{y}, u) \quad \text{s.t.} \quad \mathbf{E}(\mathbf{y}, u) = 0, \quad u \in U_{\text{ad}}$$

with $\mathbf{J}(\mathbf{y}, u) := \mathbb{E}[J[\cdot](\mathbf{y}(\cdot, \cdot), u)] = \frac{1}{2} \|\mathbf{Q}\mathbf{y} - \hat{\mathbf{q}}\|_H^2 + \frac{\gamma}{2} \|u\|_U^2$.

2.1. A class of semilinear, elliptic PDEs with stochastic coefficients. We consider the class of semilinear, elliptic PDEs from [23], now with uncertain problem data. Let $Y := H_0^1(\Omega)$ with an open, bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$ ($n \in \{2, 3\}$) and let the control space be $U := L^2(\Omega_u)$. We restrict the integrability exponent to $p \in (3, \infty)$ for $n = 2$ and even $p \in (3, 6]$ for $n = 3$ to prove differentiability of a superposition operator in [Proposition 2.3](#).

Assumption 2.2. The following data enters the considered class of PDEs and has the mentioned properties:

- $\kappa \in L^\infty(\Omega \times \Xi)$ fulfills $\underline{\kappa} \leq \kappa(x, \xi) \leq \bar{\kappa}$ for a.e. $(x, \xi) \in \Omega \times \Xi$ with $0 < \underline{\kappa} \leq \bar{\kappa} < \infty$.
- The possibly nonlinear function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is twice continuously differentiable, increasing, and satisfies

$$(2.7) \quad |\varphi''(t)| \leq a_\varphi'' + c_\varphi'' |t|^{p-3}$$

for all $t \in \mathbb{R}$ with constants $a_\varphi'', c_\varphi'' \geq 0$.

- $D \in \mathcal{L}(L^2(\Omega_u), L^2(\Omega))$ is linear and bounded.
- $\mathbf{f} \in L_{\mathbb{P}}^{r_f}(\Xi; L^2(\Omega))$ with $r_f \in [p, \infty]$.

The considered PDE has the form (2.2) with the definitions (2.3), where

$$(2.8) \quad \begin{aligned} \langle A(\xi)y, v \rangle_{Y^*, Y} &= (\kappa(\cdot, \xi) \nabla y, \nabla v)_{L^2(\Omega)^n}, & \langle N(y), v \rangle_{Y^*, Y} &= \int_{\Omega} \varphi(y) v \, dx, \\ \langle B(\xi)u, v \rangle_{Y^*, Y} &= (Du, v)_{L^2(\Omega)}, & \langle b(\xi), v \rangle_{Y^*, Y} &= (\mathbf{f}(\cdot, \xi), v)_{L^2(\Omega)}. \end{aligned}$$

The corresponding operators in (2.3) are well-defined due to measurability and identifications analogous to (2.4) with $A(\cdot) \in L_{\mathbb{P}}^\infty(\Xi; \mathcal{L}(Y, Y^*))$, $B(\cdot) \in L_{\mathbb{P}}^\infty(\Xi; \mathcal{L}(U, Y^*))$, $b(\cdot) \in L_{\mathbb{P}}^{r_f}(\Xi, Y^*)$ by [Assumption 2.2](#). For the nonlinear operator \mathbf{N} , the theory of superposition operators can be applied to show that it is well-defined and twice continuously differentiable. Due to the Sobolev embedding $H_0^1(\Omega) \hookrightarrow L^p(\Omega)$, we have

$$(2.9) \quad \mathbf{Y} = L_{\mathbb{P}}^p(\Xi; H_0^1(\Omega)) \hookrightarrow L_{\mathbb{P}}^p(\Xi, L^p(\Omega)) \cong L_{\lambda \otimes \mathbb{P}}^p(\Omega \times \Xi),$$

where λ is the Lebesgue measure on Ω . This gives the dual embedding $L_{\lambda \otimes \mathbb{P}}^{p^*}(\Omega \times \Xi) \hookrightarrow \mathbf{Y}^*$. Therefore, it is enough to show the desired properties of \mathbf{N} as a superposition operator from L^p to L^{p^*} .

PROPOSITION 2.3. *Let $\mathbf{N} : \mathbf{Y} \rightarrow \mathbf{Y}^*$ be defined in (2.3), (2.8) with the function $\varphi \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ satisfying (2.7). Then, \mathbf{N} is well-defined and twice continuously differentiable with the derivatives*

$$\begin{aligned} \langle \mathbf{N}'(\mathbf{y})\mathbf{v}, \tilde{\mathbf{v}} \rangle_{Y^*, Y} &= \int_{\Xi} \int_{\Omega} \varphi'(\mathbf{y}(x, \xi)) \mathbf{v}(x, \xi) \tilde{\mathbf{v}}(x, \xi) \, dx \, d\mathbb{P}, \\ \langle [\mathbf{N}''(\mathbf{y})\mathbf{v}]\mathbf{w}, \tilde{\mathbf{v}} \rangle_{Y^*, Y} &= \int_{\Xi} \int_{\Omega} \varphi''(\mathbf{y}(x, \xi)) \mathbf{v}(x, \xi) \mathbf{w}(x, \xi) \tilde{\mathbf{v}}(x, \xi) \, dx \, d\mathbb{P}. \end{aligned}$$

Proof. This can be shown in analogy to [43, sec. A.4] using the results from [3]. We sketch the main steps here. First note that (2.7) implies

$$(2.10) \quad \text{(a) } |\varphi'(t)| \leq a_\varphi' + c_\varphi' |t|^{p-2}, \quad \text{(b) } |\varphi(t)| \leq a_\varphi + c_\varphi |t|^{p-1}$$

with suitable constants $a_\varphi', c_\varphi', a_\varphi, c_\varphi \geq 0$ by the fundamental theorem of calculus. (2.10) (b) yields the well-definedness and continuity of the superposition operator \mathbf{N} as a map from $L_{\lambda \otimes \mathbb{P}}^p(\Omega \times \Xi)$ to $L_{\lambda \otimes \mathbb{P}}^{p^*}(\Omega \times \Xi)$ by [3, Thm. 3.1, Lem. 1.5, Thm. 1.1, Thm. 3.7]. Note that $\frac{p}{p^*} = p - 1$. Analogously, \mathbf{N}' is continuous as a mapping from $L_{\lambda \otimes \mathbb{P}}^p(\Omega \times \Xi)$ to $L_{\lambda \otimes \mathbb{P}}^{pp^*/(p-p^*)}(\Omega \times \Xi)$.

$\Xi \cong \mathcal{L}(L_{\lambda \otimes \mathbb{P}}^p(\Omega \times \Xi), L_{\lambda \otimes \mathbb{P}}^{p^*}(\Omega \times \Xi))$ by (2.10) (a). Here we require $p > p^*$ and thus $p > 2$. The Fréchet differentiability condition can now be shown as in [43, Prop. A.11]. The second derivative is handled analogously with $p > 3$. The embedding (2.9) yields the result. \square

Remark 2.4. A similar result can be found in [24], but only for the case of domains equipped with the Lebesgue measure. In particular, necessary conditions for continuity are more involved in the case of measures containing atomic parts, compare [3, Thm. 3.7] and [24, Thm. 6].

DEFINITION 2.5. For a.e. $\xi \in \Xi$ we define the parametrized control-to-state mapping $S[\xi] : U \rightarrow Y$ such that $S[\xi](u) = y(\xi) \in Y \cap \mathcal{C}(\bar{\Omega})$ is the unique solution of (2.1) with the operators defined in (2.8).

In [23] it is mentioned, that this mapping is well-defined. It satisfies the a priori bound and Lipschitz estimate

$$(2.11) \quad \|S[\xi](u)\|_Y \leq \frac{C_\Omega}{\kappa} \|Du + f(\xi) - \varphi(0)\|_{L^2(\Omega)},$$

$$(2.12) \quad \|S[\xi](w) - S[\xi](u)\|_Y \leq \frac{1}{\kappa} \iota D(w - u)\|_{Y^*},$$

where C_Ω is the constant from Poincaré's inequality and $\iota : L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$.

We now argue that the weak formulation (2.2) of the state equation has a unique solution, which can be constructed by solving the pointwise version (2.1) for a.e. $\xi \in \Xi$.

PROPOSITION 2.6. *Under Assumption 2.2, the control-to-state mapping*

$$(2.13) \quad \mathbf{S} : U \rightarrow \mathbf{Y}, \mathbf{S}(u)(\cdot, \xi) := S[\xi](u) \text{ for a.e. } \xi \in \Xi,$$

with $S[\xi](u)$ from Definition 2.5 is well-defined. It satisfies $\mathbf{S}(u) \in L_{\mathbb{P}}^{rf}(\Xi; Y) \subset \mathbf{Y}$ and $\mathbf{y} = \mathbf{S}(u)$ is the unique solution of (2.1) with the definitions (2.3) and (2.8).

Proof. The mapping $\xi \mapsto S[\xi](u)$ is measurable for every $u \in U$ because the solution of the pointwise state equation depends continuously on the data $A(\cdot)$, $B(\cdot)$, and $b(\cdot)$, which are strongly measurable. Furthermore, $\mathbf{S}(u) \in L_{\mathbb{P}}^{rf}(\Xi; Y) \subset \mathbf{Y}$ follows from (2.11) and the integrability assumption on \mathbf{f} in Assumption 2.2 by Hölder's inequality. Clearly, $\mathbf{S}(u)$ solves (2.2). The solution is unique on \mathbf{Y} because the nonlinear operator $\mathbf{A} + \mathbf{N}$ is strictly monotone, which follows from the uniform positivity of κ and the monotonicity of φ . \square

Remark 2.7. Since we have to require $p > 2$ for differentiability of \mathbf{N} , see Proposition 2.3, strong monotonicity of $\mathbf{A} + \mathbf{N}$ is not given on the state space \mathbf{Y} and existence of a solution cannot be derived this way. If κ meets the requirements in Assumption 2.2, the operator \mathbf{A} would be strongly monotone as a mapping from $L_{\mathbb{P}}^2(\Xi; Y)$ to $L_{\mathbb{P}}^2(\Xi; Y^*)$.

THEOREM 2.8. *Under Assumptions 2.1 and 2.2, problem (2.6) with the state equation defined by (2.2), (2.3), (2.8) has a solution.*

Proof. The statement is proven using ideas from [43, Lem. 9.4]. Let $(\mathbf{y}_k, u_k)_{k \in \mathbb{N}}$ be a feasible, infimizing sequence, where $\mathbf{y}_k = \mathbf{S}(u_k)$. Because of the regularization term in the objective function with $\gamma > 0$ and the non-negative tracking term, the sequence (u_k) is bounded in $L^2(\Omega_u) = U$. Since U_{ad} is convex and closed, we can extract a subsequence, again denoted by (u_k) , converging weakly to some limit $u \in U_{\text{ad}}$. Therefore, $Du_k + f(\xi) \rightharpoonup Du + f(\xi)$ weakly in $L^2(\Omega)$ and strongly in $H^{-1}(\Omega)$. By (2.12), the sequences of corresponding pointwise states $(\mathbf{y}_k(\cdot, \xi))$ converge strongly to some $\mathbf{y}(\cdot, \xi) \in H_0^1(\Omega)$. The continuity of the superposition operator induced by φ , which holds in analogy to Proposition 2.3, yields that $\varphi(\mathbf{y}_k(\cdot, \xi)) \rightarrow \varphi(\mathbf{y}(\cdot, \xi))$ strongly in $H^{-1}(\Omega)$. It follows that

$$A(\xi)\mathbf{y}_k(\cdot, \xi) + \varphi(\mathbf{y}_k(\cdot, \xi)) \rightarrow A(\xi)\mathbf{y}(\cdot, \xi) + \varphi(\mathbf{y}(\cdot, \xi)) \text{ in } H^{-1}(\Omega)$$

so that $\mathbf{y}(\cdot, \xi)$ solves the pointwise counterpart of (2.2) for a.e. ξ . Thus, we have that $\mathbf{y} = \mathbf{S}(u) \in L^{r_f}(\Xi; Y)$ giving that (\mathbf{y}, u) is feasible for (2.6).

The regularization term $u \mapsto \frac{\gamma}{2} \|u\|_U^2$ is convex and continuous under strong convergence in $L^2(\Omega_u)$, and thus weakly lower semicontinuous. Furthermore, $\|Q(\xi)\mathbf{y}_k(\cdot, \xi) - \hat{q}(\xi)\|_H^2 \rightarrow \|Q(\xi)\mathbf{y}(\cdot, \xi) - \hat{q}(\xi)\|_H^2$ for a.e. ξ as $k \rightarrow \infty$ due to strong continuity. Because of the boundedness of $\|u_k\|_U$ by some constant $C_u > 0$ and (2.11), we can bound

$$(2.14) \quad \begin{aligned} & \|Q(\xi)\mathbf{y}_k(\cdot, \xi) - \hat{q}(\xi)\|_H^2 \\ & \leq \left(\|Q(\xi)\|_{\mathcal{L}(Y, H)} \frac{C_\Omega}{\underline{\kappa}} (\|D\|_{\mathcal{L}(U, L^2(\Omega))} C_u + \|f(\xi) - \varphi(0)\|_{L^2(\Omega)}) + \|\hat{q}(\xi)\|_H \right)^2 \end{aligned}$$

for every k . The integrability properties of $Q(\cdot)$, $f(\cdot)$ and $\hat{q}(\cdot)$ are chosen exactly such that the estimate on the right-hand side is integrable w.r.t. \mathbb{P} . Therefore, the dominated convergence theorem can be applied to conclude

$$\int_{\Xi} \|Q(\xi)\mathbf{y}_k(\cdot, \xi) - \hat{q}(\xi)\|_H^2 d\mathbb{P} \longrightarrow \int_{\Xi} \|Q(\xi)\mathbf{y}(\cdot, \xi) - \hat{q}(\xi)\|_H^2 d\mathbb{P} \quad \text{as } k \rightarrow \infty.$$

It follows from the derived continuity properties of \mathbf{J} that (\mathbf{y}, u) solves (2.6). \square

2.2. The reduced objective function. Using the control-to-state mapping (2.13), problem (2.6) can be reduced to the control:

$$(2.15) \quad \min_{u \in U} \hat{\mathbf{J}}(u) := \mathbf{J}(\mathbf{S}(u), u) = \mathbb{E}[J[\cdot](\mathbf{S}(u)(\cdot, \cdot), u)] \quad \text{s.t.} \quad u \in U_{\text{ad}}.$$

With the reduced random variable objective function $\hat{J}[\cdot](u) := J[\cdot](\mathbf{S}(u)(\cdot, \cdot), u)$ we can also write $\hat{\mathbf{J}}(u) = \mathbb{E}[\hat{J}[\cdot](u)]$. By Proposition 2.6, we have $\mathbf{S}(u) \in L_{\mathbb{P}}^{r_f}(\Xi; Y)$ with $r_f \in [p, \infty]$, and, by Assumption 2.1, $Q(\cdot) \in L_{\mathbb{P}}^{r_Q}(\Xi; \mathcal{L}(Y, H))$, $\hat{q}(\cdot) \in L_{\mathbb{P}}^{r_{\hat{q}}}(\Xi; H)$ with $r_Q \in [\frac{2p}{p-2}, \infty]$, and $r_{\hat{q}} \in [2, \infty]$. Therefore, by Hölder's inequality,

$$Q(\cdot)\mathbf{S}(u)(\cdot, \cdot) - \hat{q}(\cdot) \in L_{\mathbb{P}}^{\hat{r}}(\Xi; H) \quad \text{with } \hat{r} = \min\{\check{r}, r_{\hat{q}}\} \geq 2, \frac{1}{\hat{r}} = \frac{1}{r_Q} + \frac{1}{r_f}.$$

The tracking term $U \ni u \mapsto \|Q(\cdot)\mathbf{S}(u)(\cdot, \cdot) - \hat{q}(\cdot)\|_H^2 \in L_{\mathbb{P}}^{p_J}(\Xi)$, $p_J = \frac{\hat{r}}{2}$, admits weak-strong continuity: If $u^k \rightharpoonup u \in U$, then $\mathbf{S}(u^k)(\cdot, \xi) \rightarrow \mathbf{S}(u)(\cdot, \xi)$ in Y for almost every ξ by (2.12). Thus, the tracking term converges almost surely. By the dominated convergence theorem, $L_{\mathbb{P}}^{p_J}(\Xi)$ convergence follows from (2.14) in the case $p_J < \infty$. For $p_J = \infty$, one can use the uniform Lipschitz continuity (2.12) of the PDE solution operator, which yields uniform, i.e., $L_{\mathbb{P}}^{\infty}(\Xi)$ convergence of the tracking term. This property is important when deriving existence results with more general, strongly continuous risk measures, see [35].

We want to use the adjoint approach [30, sec. 1.6] for the computation of the gradient of the reduced objective function $\hat{\mathbf{J}}$. This is not straight-forward since the operator $\mathbf{A} + \mathbf{N}'(\mathbf{y}) : \mathbf{Y} \rightarrow \mathbf{Y}^*$ is not boundedly invertible. Still, as seen in [23], we can compute the derivative

$$(2.16) \quad \hat{J}[\xi]'(u) = -B(\xi)^* T[\xi](u) + \gamma(u, \cdot)_U$$

of the parametric reduced objective function for a.e. $\xi \in \Xi$ with the adjoint state $T[\xi](u)$:

DEFINITION 2.9. For a.e. $\xi \in \Xi$, the parametrized control-to-adjoint-state mapping $T[\xi] : U \rightarrow Y$ is defined such that $T[\xi](u) = z(\xi) \in Y$ is the unique solution of the adjoint equation

$$(2.17) \quad A(\xi)z(\xi) + N'(S[\xi](u))z(\xi) = -Q(\xi)^*(Q(\xi)S[\xi](u) - \hat{q}(\xi)) \text{ for a.e. } \xi \in \Xi$$

with the operators defined in (2.8) and the state $S[\xi](u)$ defined in Definition 2.5.

Function and space	Exponents	Estimation 1	Estimation 2	Example
$\mathbf{y} \in L_{\mathbb{P}}^{r_f}(\Xi; Y)$	$r_f \geq p \in (3, \infty)$	$r_f \in [p, \infty)$	$r_f = \infty$	$r_f \geq p$
$\mathbf{Q} \in L_{\mathbb{P}}^{r_Q}(\Xi; \mathcal{L}(Y, H))$	$r_Q \geq \frac{2p}{p-2}$	$r_Q \geq \frac{2pr_f}{r_f-p}$	$r_Q \geq 2p$	$r_Q = \infty$
$\hat{\mathbf{q}} \in L_{\mathbb{P}}^{r_{\hat{q}}}(\Xi; H)$	$r_{\hat{q}} \geq 2$	$r_{\hat{q}} \geq \frac{2pr_f}{r_f+p}$	$r_{\hat{q}} \geq r_Q$	$r_{\hat{q}} \geq r_f$
$\mathbf{Q}\mathbf{y} \in L_{\mathbb{P}}^{\check{r}}(\Xi; H)$	$\frac{1}{\check{r}} = \frac{1}{r_Q} + \frac{1}{r_f}, \check{r} \geq 2$	$\check{r} \geq \frac{2pr_f}{r_f+p}$	$\check{r} = r_Q$	$\check{r} = r_f$
$\mathbf{Q}\mathbf{y} - \hat{\mathbf{q}} \in L_{\mathbb{P}}^{\hat{r}}(\Xi; H)$	$\hat{r} = \min\{\check{r}, r_{\hat{q}}\} \geq 2$	$\hat{r} \geq \frac{2pr_f}{r_f+p}$	$\hat{r} = r_Q$	$\hat{r} = r_f$
$\mathbf{z} \in L_{\mathbb{P}}^{r_z}(\Xi; Y)$	$\frac{1}{r_z} = \frac{1}{\check{r}} + \frac{1}{r_Q}, r_z \geq \frac{p}{p-1}$	$r_z \geq p$	$r_z = \frac{r_Q}{2} \geq p$	$r_z = r_f$

TABLE 1

Overview of the integrability inheritance

In analogy to (2.11), the adjoint state satisfies the a priori estimate

$$(2.18) \quad \|T[\xi](u)\|_Y \leq \frac{1}{\underline{\kappa}} \|Q(\xi)\|_{\mathcal{L}(Y, H)} \|Q(\xi)S[\xi](u) - \hat{q}(\xi)\|_H.$$

The weak formulation of (2.17) is given by

$$(2.19) \quad \mathbf{A}\mathbf{z} + \mathbf{N}'(\mathbf{S}(u))\mathbf{z} = -\mathbf{Q}^*(\mathbf{Q}\mathbf{S}(u) - \hat{\mathbf{q}}).$$

Again, we want to discuss whether it has a unique solution.

PROPOSITION 2.10. *Under Assumptions 2.1 and 2.2, the control-to-adjoint-state mapping*

$$(2.20) \quad \mathbf{T} : U \rightarrow L_{\mathbb{P}}^{p/(p-1)}(\Xi; Y), \quad \mathbf{T}(u)(x, \xi) := T[\xi](u)(x) \text{ for a.e. } \xi \in \Xi.$$

with $T[\xi](u)$ from Definition 2.9 is well-defined. It satisfies $\mathbf{T}(u) \in L_{\mathbb{P}}^{r_z}(\Xi; Y) \subset \mathbf{Y}$ with r_z from the column ‘‘Exponents’’ in Table 1.

Proof. To argue that the mapping \mathbf{T} is well-defined we firstly see that $\xi \mapsto T[\xi](u)$ is measurable because the solution $T[\xi](u)$ of (2.17) depends continuously on the operator $A(\xi) + N'(y(\xi))$ and the right-hand side $-Q(\xi)^*(Q(\xi)y(\xi) - \hat{q}(\xi))$, which are both measurable w.r.t. ξ . Secondly, since $(A(\xi) + N'(y(\xi)))^{-1} \in L_{\mathbb{P}}^{\infty}(\Xi; \mathcal{L}(Y^*, Y))$, the adjoint state has the same integrability w.r.t. ξ as $Q^*(\cdot)(Q(\cdot)S[\cdot](u) - \hat{q}(\cdot))$, which belongs to $L_{\mathbb{P}}^{p/(p-1)}(\Xi; Y^*)$ by Assumption 2.1 and Proposition 2.6. The integrability exponents in the column ‘‘Exponents’’ of Table 1 can be concluded by Hölder’s inequality. \square

PROPOSITION 2.11. *Let Assumptions 2.1 and 2.2 hold. Then, the adjoint equation (2.19) is well-defined. If the exponents r_Q and $r_{\hat{q}}$ fulfill the r_f -dependent bounds given in the ‘‘Estimation’’-columns of Table 1, it has the unique solution $\mathbf{z} = \mathbf{T}(u)$ defined in (2.20).*

Proof. We have that $Q^*(\mathbf{Q}\mathbf{S}(u) - \hat{\mathbf{q}}) \in L_{\mathbb{P}}^{p^*}(\Xi; Y^*) = \mathbf{Y}^*$ as well as $\mathbf{A}\mathbf{z} + \mathbf{N}'(\mathbf{S}(u))\mathbf{z} \in \mathbf{Y}^*$ for $\mathbf{z} \in \mathbf{Y}$ by Hölder’s inequality, i.e., the equation is well-posed on \mathbf{Y} . Since $\mathbf{A} + \mathbf{N}'(\mathbf{S}(u))$ is strictly monotone, it has at most one solution. As listed in Table 1, the adjoint state $\mathbf{z} = \mathbf{T}(u)$ belongs to $\mathbf{Y} = L_{\mathbb{P}}^p(\Xi; Y)$ if r_Q and $r_{\hat{q}}$ fulfill the mentioned bounds. It clearly solves (2.19). \square

In the case $r_Q = \infty$ and $r_{\hat{q}} \geq r_f$ (column ‘‘Example’’ of Table 1), we even get $\mathbf{T}(u) \in L_{\mathbb{P}}^{r_f}(\Xi; Y)$, i.e., the adjoint state has the same ξ -integrability as the state itself. We will focus on this case when deriving error estimates in section 3.

To derive an adjoint-based expression of the gradient $\nabla \hat{J}(u)$, we observe that $\xi \mapsto \hat{J}[\xi]'(u)$ is measurable as a composition of measurable functions by (2.16). It admits the same ξ -integrability as the adjoint state because $B(\cdot) \in L_{\mathbb{P}}^{\infty}(\Xi; \mathcal{L}(U, Y^*))$. Hence, $\|\hat{J}[\cdot]'(u)\|_{U^*} \in$

$L_{\mathbb{P}}^{p/(p-1)}(\Xi) \subset L_{\mathbb{P}}^1(\Xi)$. Furthermore, the function $\xi \mapsto \hat{J}[\xi](u)$ belongs to $L_{\mathbb{P}}^{\hat{r}/2}(\Xi) \subset L_{\mathbb{P}}^1(\Xi)$. Therefore, we can apply the chain rule and get

$$(2.21) \quad \frac{d}{du} \hat{J}(u) = \frac{d}{du} \int_{\Xi} \hat{J}[\xi](u) d\mathbb{P} = \int_{\Xi} \hat{J}[\xi]'(u) d\mathbb{P} = -\check{\mathbf{B}}^* \mathbf{T}(u) + \gamma(u, \cdot)_U$$

since $\mathbb{E} : L_{\mathbb{P}}^1(\Xi) \rightarrow \mathbb{R}$ is linear and bounded. The operator $\check{\mathbf{B}} \in \mathcal{L}(U, L_{\mathbb{P}}^{p/(p-1)}(\Xi; Y^*))$ is defined as \mathbf{B} in (2.3). The pointwise adjoint state $\mathbf{T}(u)$ can thus be used for the computation of the gradient of the reduced objective function \hat{J} . Similar arguments can be applied to derive an expression of the Hessian $\nabla^2 \hat{J}(u)$, see [20].

3. Error estimation procedure in the stochastic case. We apply the inexact trust-region algorithm presented in the first part [23] to the example from section 2 with the concrete definitions and assumptions given in subsection 2.1, in particular (2.8). As derived, we have to control the errors in the objective function and gradient evaluation as well as the inexact projection to ensure global convergence and to implement a projected linesearch. We require error estimates of the form

$$(3.1) \quad \begin{aligned} (a) \quad & |\hat{J}_k(u) - \hat{J}(u)| \leq c_o \varepsilon_o, & (b) \quad & \|\nabla m_k(0) - \nabla \hat{J}(u^k)\|_U \leq c_g \varepsilon_g, \\ (c) \quad & \|P_{U_{\text{ad}}}(w^k(t)) - \hat{P}_{U_{\text{ad}}}(w^k(t))\|_U \leq c_p \varepsilon_p \end{aligned}$$

with arbitrarily small, computable tolerances $\varepsilon_o, \varepsilon_g, \varepsilon_p > 0$ and fixed, but possibly unknown constants $c_o, c_g, c_p > 0$. The objective function is $J : U \rightarrow \mathbb{R}$ with $U_{\text{ad}} \subset \tilde{U} \subset U$, $\hat{J}_k : U_{\text{ad}} \rightarrow \mathbb{R}$ is the inexact evaluation of it in iteration k of the algorithm. The current iterate and step are $u^k \in U_{\text{ad}}$ and $s^k \in U$, respectively, and $u \in \{u^k, u^k + s^k\}$. The function $m_k : U \rightarrow \mathbb{R}$ is the trust-region model, $t > 0$ is a gradient stepsize, and $P_{U_{\text{ad}}}, \hat{P}_{U_{\text{ad}}} : U \rightarrow U_{\text{ad}}$ are the exact and the inexact projection, respectively.

In [23], it is shown how to control the projection error (3.1) (c). The estimates from [23] ensuring (3.1) (a) and (b) in the deterministic case are generalized to the stochastic setting in the following. For this purpose, we require stronger integrability assumptions on the problem data, namely Assumption 2.1 with $r_Q = \infty$ and $r_{\hat{q}} = r_f$, whereas Assumption 2.2 remains unchanged. This ensures $\mathbf{y}, \mathbf{z} \in L_{\mathbb{P}}^{r_f}(\Xi; Y)$, see the ‘‘Example’’ column of Table 1. The index k denoting the iteration number is skipped for readability purposes in this section.

Model gradient error. The model gradient is computed as described in subsection 2.2, but with inexact solutions of the respective state and adjoint equations, i.e., we have the exact state $\mathbf{y} = \mathbf{S}(u)$ and adjoint state $\mathbf{z} = \mathbf{T}(u)$ as well as inexact versions $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{z}}$. Quantities like $\|\mathbf{B}^*(\mathbf{z} - \tilde{\mathbf{z}})\|_{U^*}$ can be estimated as follows.

LEMMA 3.1. *Let $\mathbf{B} : U \rightarrow Y^*$ be defined as in (2.3) with $B(\cdot) \in L_{\mathbb{P}}^{\infty}(\Xi; \mathcal{L}(U, Y^*))$ and let $\mathbf{z}, \tilde{\mathbf{z}} \in Y$ be given. Then, $\|\mathbf{B}^*(\mathbf{z} - \tilde{\mathbf{z}})\|_{U^*} \leq \|B(\cdot)\|_{L_{\mathbb{P}}^{\infty}(\Xi; \mathcal{L}(U, Y^*))} \|\tilde{\mathbf{z}} - \mathbf{z}\|_{L_{\mathbb{P}}^1(\Xi; Y)}$ holds.*

Proof. The result follows from Hölder’s inequality noting that

$$\begin{aligned} \|\mathbf{B}^*(\mathbf{z} - \tilde{\mathbf{z}})\|_{U^*} &= \sup_{w \in U, \|w\|_U \leq 1} \langle \mathbf{B}w, \mathbf{z} - \tilde{\mathbf{z}} \rangle_{Y^*, Y} \\ &\leq \sup_{w \in U, \|w\|_U \leq 1} \int_{\Xi} \|B(\xi)w\|_{Y^*} \|(\tilde{\mathbf{z}}(\cdot, \xi) - \mathbf{z}(\cdot, \xi))\|_Y d\mathbb{P}. \quad \square \end{aligned}$$

If $\langle B(\xi)u, v \rangle_{Y^*, Y} = (Du, v)_{L^2(\Omega)}$ as in (2.8), we can use $\|B(\cdot)\|_{L_{\mathbb{P}}^{\infty}(\Xi; \mathcal{L}(U, Y^*))} \leq C_{\Omega} \|D\|_{\mathcal{L}(U, L^2(\Omega))}$ in the derived estimate. We introduce the exact solution $\hat{\mathbf{z}}$ of the perturbed adjoint equation

$$(3.2) \quad \mathbf{A}\mathbf{z} + \mathbf{N}'(\tilde{\mathbf{y}})\mathbf{z} = -\mathbf{Q}^*(\mathbf{Q}\tilde{\mathbf{y}} - \hat{\mathbf{q}})_{\mathbf{H}}$$

involving the inexact state $\tilde{\mathbf{y}}$ to derive the following error estimate for the model gradient.

THEOREM 3.2. *Given a control $u \in U$, let $\mathbf{y} = \mathbf{S}(u) \in L_{\mathbb{P}}^{r_f}(\Xi; Y)$ be the exact state and let $\tilde{\mathbf{y}} \in L_{\mathbb{P}}^{r_f}(\Xi; Y)$. Furthermore, let $\mathbf{z} = \mathbf{T}(u) \in L_{\mathbb{P}}^{r_f}(\Xi; Y)$ be the exact adjoint state, let $\hat{\mathbf{z}} \in L_{\mathbb{P}}^{r_f}(\Xi; Y)$ be the unique solution of (3.2), and let $\tilde{\mathbf{z}} \in L_{\mathbb{P}}^{r_f}(\Xi; Y)$. Using $\mathbf{J} : \mathbf{Y} \times U \rightarrow \mathbb{R}$ from (2.6), set $\mathbf{m}'(0) := -\mathbf{B}^* \tilde{\mathbf{z}} + \mathbf{J}_u(\tilde{\mathbf{y}}, u)$. Suppose that $\tilde{r} \in (2, \infty)$ (for $n = 2$) or even $\tilde{r} \in (2, 6]$ (for $n = 3$) and let $c_{\tilde{r}}$ be the Sobolev constant such that $\|y\|_{L^{\tilde{r}}(\Omega)} \leq c_{\tilde{r}} \|y\|_Y$ holds for every $y \in Y$. Let $p_1 \in [1, \frac{r_f}{p-2}]$, $p_2 \in [1, r_f]$ such that $\frac{1}{p_1} + \frac{1}{p_2} = 1$. Then the following estimate holds true with $\hat{\mathbf{J}} : U \rightarrow \mathbb{R}$ from (2.15):*

$$(3.3) \quad \begin{aligned} & \|\mathbf{m}'(0) - \hat{\mathbf{J}}'(u)\|_{U^*} \leq \\ & \|\mathbf{B}^*(\tilde{\mathbf{z}} - \hat{\mathbf{z}})\|_{U^*} + \frac{1}{\kappa} \|B(\cdot)\|_{L_{\mathbb{P}}^{\infty}(\Xi; \mathcal{L}(U, Y^*))} \|Q(\cdot)\|_{L_{\mathbb{P}}^{\infty}(\Xi; \mathcal{L}(Y, H))} \left(\|Q(\mathbf{y} - \tilde{\mathbf{y}})\|_{L_{\mathbb{P}}^1(\Xi, H)} \right. \\ & \left. + \frac{1}{\kappa} c_{\tilde{r}}^2 \|\phi'(\tilde{\mathbf{y}}) - \phi'(\mathbf{y})\|_{L_{\mathbb{P}}^{p_1}(\Xi, L^{\tilde{r}/(\tilde{r}-2)}(\Omega))} (\|Q\tilde{\mathbf{y}} - \hat{q}\|_{L_{\mathbb{P}}^{p_2}(\Xi, H)} + \|Q(\mathbf{y} - \tilde{\mathbf{y}})\|_{L_{\mathbb{P}}^{p_2}(\Xi, H)}) \right). \end{aligned}$$

Proof. Analogously to [23, Eq. (4.4)], we have that

$$\|\mathbf{m}'(0) - \hat{\mathbf{J}}'(u)\|_{U^*} \leq \|\mathbf{B}^*(\hat{\mathbf{z}} - \tilde{\mathbf{z}})\|_{U^*} + \|\mathbf{B}^*(\mathbf{z} - \hat{\mathbf{z}})\|_{U^*}.$$

The second summand is estimated by Lemma 3.1 and it remains to bound $\|\mathbf{z} - \hat{\mathbf{z}}\|_{L_{\mathbb{P}}^1(\Xi; Y)}$. Since \mathbf{z} and $\hat{\mathbf{z}}$ are defined pointwise (almost everywhere), the estimate [23, Eq. (4.9)] yields

$$(3.4) \quad \begin{aligned} \|\mathbf{z} - \hat{\mathbf{z}}\|_{L_{\mathbb{P}}^1(\Xi; Y)} & \leq \frac{1}{\kappa} \|Q^* Q(\mathbf{y} - \tilde{\mathbf{y}})\|_{L_{\mathbb{P}}^1(\Xi; Y^*)} + \frac{1}{\kappa^2} \|N'(\tilde{\mathbf{y}}) - N'(\mathbf{y})\|_{L_{\mathbb{P}}^{p_1}(\Xi; \mathcal{L}(Y, Y^*))} \\ & \cdot \left(\|Q^*(Q\tilde{\mathbf{y}} - \hat{q})\|_{L_{\mathbb{P}}^{p_2}(\Xi; Y^*)} + \|Q^* Q(\mathbf{y} - \tilde{\mathbf{y}})\|_{L_{\mathbb{P}}^{p_2}(\Xi; Y^*)} \right). \end{aligned}$$

with $p_1, p_2 \in [1, \infty]$, $\frac{1}{p_1} + \frac{1}{p_2} = 1$. Using $\|N'(\mathbf{y}) - N'(\tilde{\mathbf{y}})\|_{\mathcal{L}(Y, Y^*)} \leq c_{\tilde{r}}^2 \|\phi'(\mathbf{y}) - \phi'(\tilde{\mathbf{y}})\|_{L^{\tilde{r}/(\tilde{r}-2)}(\Omega)}$ results in (3.3). The admissible values of p_1 and p_2 ensure together with the integrability of $\mathbf{y}, \tilde{\mathbf{y}}$ and the growth of ϕ' that every appearing quantity in (3.3) is finite. \square

Remark 3.3. The parameters p_1 and p_2 in Theorem 3.2 make different estimates involving $\|Q(\mathbf{y} - \tilde{\mathbf{y}})\|_{L_{\mathbb{P}}^{p_2}(\Xi; H)}$ with $2 \leq p_2 \leq r_f$ possible. For larger p_2 , a weaker norm w.r.t. ξ can be used to estimate the error in $N'(\tilde{\mathbf{y}})$.

To bound the error in the computed adjoint state, we identify $U^* = U$ and $L^2(\Omega)^* = L^2(\Omega)$ and estimate

$$\|\mathbf{B}^*(\mathbf{z} - \tilde{\mathbf{z}})\|_{U^*} \leq \|D\|_{\mathcal{L}(U, L^2(\Omega))} \|\tilde{\mathbf{z}} - \mathbf{z}\|_{L_{\mathbb{P}}^1(\Xi; L^2(\Omega))} \leq C_{\Omega} \|D\|_{\mathcal{L}(U, L^2(\Omega))} \|\tilde{\mathbf{z}} - \mathbf{z}\|_{L_{\mathbb{P}}^2(\Xi; Y)},$$

with the concrete definition (2.8) of $B(\cdot)$ and using that $\|\cdot\|_{L_{\mathbb{P}}^1(\Xi)} \leq \|\cdot\|_{L_{\mathbb{P}}^2(\Xi)}$ since \mathbb{P} is a probability measure. We see that it would be sufficient to control the $L_{\mathbb{P}}^1(\Xi; L^2(\Omega))$ -error in the adjoint state. This is no longer true if we consider, e.g., a boundary control problem. Thus, we will use the $L_{\mathbb{P}}^2(\Xi; Y)$ -error to keep the algorithm flexible.

LEMMA 3.4. *Let $\phi : \mathbb{R} \rightarrow \mathbb{R}$ be as in Assumption 2.2 and $\mathbf{N} : \mathbf{Y} \rightarrow \mathbf{Y}^*$ as in (2.3), (2.8) and let $\mathbf{y}, \tilde{\mathbf{y}} \in \mathbf{Y}$ be given. Let p_1 be as in Theorem 3.2 and let $p_3, p_4 \in [1, \infty]$ such that $\frac{1}{p_3} + \frac{1}{p_4} = \frac{1}{p_1}$. Then it holds that*

$$\begin{aligned} & \|\phi'(\tilde{\mathbf{y}}) - \phi'(\mathbf{y})\|_{L_{\mathbb{P}}^{p_1}(\Xi; L^{p/(p-2)}(\Omega))} \leq \\ & c_p \|\tilde{\mathbf{y}} - \mathbf{y}\|_{L_{\mathbb{P}}^{p_3}(\Xi; Y)} \left(a_{\phi}'' \lambda(\Omega)^{(p-3)/p} + c_{\phi}'' c_p^{p-3} \max\{\|\mathbf{y}\|_{L_{\mathbb{P}}^{p_4(p-3)}(\Xi; Y)}, \|\tilde{\mathbf{y}}\|_{L_{\mathbb{P}}^{p_4(p-3)}(\Xi; Y)}\}^{p-3} \right). \end{aligned}$$

Proof. As in the proof of [23, Lem. 4.3] we estimate

$$\begin{aligned} & \|\varphi'(\tilde{\mathbf{y}}) - \varphi'(\mathbf{y})\|_{L_{\mathbb{P}}^{p_1}(\Xi; L^{\tilde{r}/(\tilde{r}-2)}(\Omega))} = \left\| \int_0^1 \varphi''(\mathbf{y} + \tau(\tilde{\mathbf{y}} - \mathbf{y}))(\tilde{\mathbf{y}} - \mathbf{y}) \, d\tau \right\|_{L_{\mathbb{P}}^{p_1}(\Xi; L^{\tilde{r}/(\tilde{r}-2)}(\Omega))} \\ & \leq \|\tilde{\mathbf{y}} - \mathbf{y}\|_{L_{\mathbb{P}}^{p_3}(\Xi; L^{\tilde{r}}(\Omega))} \cdot \sup_{\tau \in [0,1]} \|\varphi''(\mathbf{y} + \tau(\tilde{\mathbf{y}} - \mathbf{y}))\|_{L_{\mathbb{P}}^{p_4}(\Xi; L^{\tilde{r}/(\tilde{r}-3)}(\Omega))} \end{aligned}$$

for some $\tilde{r} \in (3, \infty)$, using $\frac{1}{p_3} + \frac{1}{p_4} = \frac{1}{p_1}$. The second factor can be estimated as

$$\begin{aligned} & \sup_{\tau \in [0,1]} \|\varphi''(\mathbf{y} + \tau(\tilde{\mathbf{y}} - \mathbf{y}))\|_{L_{\mathbb{P}}^{p_4}(\Xi; L^{\tilde{r}/(\tilde{r}-3)}(\Omega))} \\ & \leq \sup_{\tau \in [0,1]} \|a''_{\varphi} + c''_{\varphi}|\mathbf{y} + \tau(\tilde{\mathbf{y}} - \mathbf{y})|^{p-3}\|_{L_{\mathbb{P}}^{p_4}(\Xi; L^{\tilde{r}/(\tilde{r}-3)}(\Omega))} \\ & = a''_{\varphi} \cdot \lambda(\Omega)^{(\tilde{r}-3)/\tilde{r}} + c''_{\varphi} \cdot \max\left\{\|\mathbf{y}\|_{L_{\mathbb{P}}^{p_4(p-3)}(\Xi; L^{r_5}(\Omega))}, \|\tilde{\mathbf{y}}\|_{L_{\mathbb{P}}^{p_4(p-3)}(\Xi; L^{r_5}(\Omega))}\right\}^{p-3} \end{aligned}$$

with $r_5 = \frac{\tilde{r}(p-3)}{\tilde{r}-3}$ by $|\varphi''(t)| \leq a''_{\varphi} + c''_{\varphi}|t|^{p-3}$ with $c''_{\varphi} \geq 0$ and $p > 3$. If we choose $\tilde{r} = p$ and use $H_0^1(\Omega) \hookrightarrow L^p(\Omega)$, we obtain the result. \square

Combining [Theorem 3.2](#) and [Lemma 3.4](#) and choosing $p_2 = p_3$, we observe that we have to estimate the error $\|\tilde{\mathbf{y}} - \mathbf{y}\|_{L_{\mathbb{P}}^{p_2}(\Xi; Y)}$ in the inexact state, compute the norm $\|\tilde{\mathbf{y}}\|_{L_{\mathbb{P}}^{p_4(p-3)}(\Xi; Y)}$ or bound it from above, and bound the norm $\|\mathbf{y}\|_{L_{\mathbb{P}}^{p_4(p-3)}(\Xi; Y)}$, e.g., by the a priori estimate (2.11).

For $r_f = \infty$ we can choose $p_2 = p_3 = 2$ and $p_4 = \infty$. Then it is enough to estimate the $L_{\mathbb{P}}^2(\Xi; Y)$ state error as long as we can compute or bound $\|\tilde{\mathbf{y}}\|_{L_{\mathbb{P}}^{\infty}(\Xi; Y)}$ and $\|\mathbf{y}\|_{L_{\mathbb{P}}^{\infty}(\Xi; Y)}$. This enables us to use the fact that the operator \mathbf{A} is strongly monotone with constant $\underline{\kappa}$ on $L_{\mathbb{P}}^2(\Xi; Y)$, but not strongly monotone on $L_{\mathbb{P}}^p(\Xi; Y)$ for $p > 2$.

For $r_f < \infty$, we have to choose $p_4 \in [\frac{r_f}{r_f-2}, \frac{r_f}{p-3}]$ and $p_2 = p_3 = \frac{2p_4}{p_4-1} \in [\frac{2r_f}{r_f-p+3}, r_f]$.¹ Observe that $p_1 = \frac{2p_4}{p_4+1} \leq \frac{2r_f}{r_f+p-3} \leq \frac{r_f}{p-\frac{3}{2}}$ follows then. On the one hand, for $r_f = p$ we can take $p_4 = \frac{p}{p-2}$. Then we have $p_2 = p_3 = p$ and the error in the computed state has to be estimated in the Y -norm. On the other hand, for increasing r_f and using $p_4 = \frac{r_f}{p-3}$, the exponents p_2 and p_3 get close to 2. Then it is enough to use a weaker norm than the Y -norm for the state error, but bounds on the exact and the inexact state have to be computed in a stronger norm.

Objective function evaluation error.

PROPOSITION 3.5. *Let $\mathbf{y}, \tilde{\mathbf{y}} \in Y$ and $u \in U$ be given. With $\mathbf{J} : Y \times U \rightarrow \mathbb{R}$ from (2.6) it then holds that*

$$(3.5) \quad |\mathbf{J}(\mathbf{y}, u) - \mathbf{J}(\tilde{\mathbf{y}}, u)| \leq \frac{1}{2} \|\mathbf{Q}(\mathbf{y} - \tilde{\mathbf{y}})\|_{L_{\mathbb{P}}^2(\Xi; H)}^2 + \|\mathbf{Q}\tilde{\mathbf{y}} - \hat{\mathbf{q}}\|_{L_{\mathbb{P}}^2(\Xi; H)} \|\mathbf{Q}(\mathbf{y} - \tilde{\mathbf{y}})\|_{L_{\mathbb{P}}^2(\Xi; H)}.$$

Proof. The estimate from [23, Prop. 4.3] and Hölder's inequality yield

$$\begin{aligned} & |\mathbf{J}(\mathbf{y}, u) - \mathbf{J}(\tilde{\mathbf{y}}, u)| = \left| \int_{\Xi} J[\xi](\mathbf{y}(\xi), u) - J[\xi](\tilde{\mathbf{y}}(\xi), u) \, d\mathbb{P} \right| \\ & \leq \int_{\Xi} \frac{1}{2} \|\mathbf{Q}(\xi)(\mathbf{y}(\xi) - \tilde{\mathbf{y}}(\xi))\|_H^2 + \|\mathbf{Q}(\xi)\tilde{\mathbf{y}}(\xi) - \hat{\mathbf{q}}(\xi)\|_H \|\mathbf{Q}(\xi)(\mathbf{y}(\xi) - \tilde{\mathbf{y}}(\xi))\|_H \, d\mathbb{P} \\ & \leq \frac{1}{2} \|\mathbf{Q}(\mathbf{y} - \tilde{\mathbf{y}})\|_{L_{\mathbb{P}}^2(\Xi; H)}^2 + \|\mathbf{Q}\tilde{\mathbf{y}} - \hat{\mathbf{q}}\|_{L_{\mathbb{P}}^2(\Xi; H)} \|\mathbf{Q}(\mathbf{y} - \tilde{\mathbf{y}})\|_{L_{\mathbb{P}}^2(\Xi; H)}. \quad \square \end{aligned}$$

¹Note that $1 < \frac{r_f}{r_f-2} \leq \frac{r_f}{r_f-3} \leq \frac{r_f}{p-3}$ and $2 < \frac{2r_f}{r_f-p+3} \leq \frac{2}{3}r_f$ hold for $p \in (3, \infty)$, $r_f \in [p, \infty)$.

We see that we have to compute or bound the $L_{\mathbb{P}}^2(\Xi; H)$ -norm of $\mathbf{Q}\tilde{\mathbf{y}} - \hat{\mathbf{q}}$ and have to estimate $\|\mathbf{Q}(\mathbf{y} - \tilde{\mathbf{y}})\|_{L_{\mathbb{P}}^2(\Xi; H)}$. If, e.g., $Q(\xi) \equiv \iota : H_0^1(\Omega) \hookrightarrow L^2(\Omega) = H$, it is enough to estimate the $L_{\mathbb{P}}^2(\Xi; L^2(\Omega))$ -norm of $\mathbf{y} - \tilde{\mathbf{y}}$. This is not true anymore if we consider a problem with boundary observation. Again, to have a flexible algorithm and to use strong monotonicity of \mathbf{A} , we estimate the $L_{\mathbb{P}}^2(\Xi; Y)$ -error and use that $\|\mathbf{Q}(\mathbf{y} - \tilde{\mathbf{y}})\|_{L_{\mathbb{P}}^2(\Xi; H)} \leq \|Q\|_{L_{\mathbb{P}}^\infty(\Xi; \mathcal{L}(Y, H))} \|\mathbf{y} - \tilde{\mathbf{y}}\|_{L_{\mathbb{P}}^2(\Xi; Y)}$, cf. [Lemma 3.1](#).

How the $L_{\mathbb{P}}^2(\Xi; Y)$ error in the computed state and adjoint state can be measured up to fixed, but possibly unknown constant factors, is discussed in the following.

4. Adaptive solution of the PDEs with uncertain coefficients. From now on, we consider the case that (Ξ, \mathbb{P}) has tensor product form, i.e., $\Xi := \times_{i=1}^m \Xi_i$ with $\Xi_i \subset \mathbb{R}$ and $\mathbb{P} := \otimes_{i=1}^m \mathbb{P}_i$, where \mathbb{P}_i are probability measures on Ξ_i . This means that the entries ξ_i of the vector $\xi \in \Xi$ are independent and—using $p < \infty$ —we can identify $L_{\mathbb{P}}^p(\Xi) \cong \otimes_{i=1}^m L_{\mathbb{P}_i}^p(\Xi_i)$, giving $\mathbf{Y} \cong Y \otimes L_{\mathbb{P}_1}^p(\Xi_1) \otimes \cdots \otimes L_{\mathbb{P}_m}^p(\Xi_m)$. For more details on tensor products of Banach spaces we refer to [\[28, chaps. 3, 4\]](#), [\[13, chap. I, sec. 7\]](#), and [\[20, sec. 2.2\]](#). In this section, we present an adaptive solution technique for the stochastic state equation [\(2.2\)](#) with the operators [\(2.3\)](#), [\(2.8\)](#) and the adjoint equation [\(2.19\)](#) based on a stochastic Galerkin discretization. The tensor product structure of Ξ makes a discretization by polynomial chaos with a tensor product basis possible. The deterministic spaces are discretized by finite elements as in [\[23\]](#). The coefficients in tensor form are represented in a low-rank format, cf. [\[22\]](#). We refer to this paper for an introduction to tensor notation, operations and modern low-rank formats and for the discretization of a very similar problem. We present an a posteriori error estimator giving separate error contributions from the finite element and the polynomial chaos discretization as well as the algebraic error caused by the iterative low-rank tensor solver. Based on that, the PDEs can be solved adaptively using efficient low-rank tensor computations.

4.1. Stochastic discretization: polynomial chaos. We proceed with the discretization of the spaces $L_{\mathbb{P}}^p(\Xi)$ and $L_{\mathbb{P}}^p(\Xi; Y)$.

Assumption 4.1. For the discretization of these spaces we assume the following:

- The sets $\Xi_i \subset \mathbb{R}$ are open and bounded² intervals for all $i \in [m]$.
- Each probability measure \mathbb{P}_i does not consist of finitely many atoms such that discretization is really necessary.
- The assumptions for the discretization of $Y = H_0^1(\Omega)$ given in [\[23, sec. 5.1\]](#) are still valid. In particular, $\Omega \subset \mathbb{R}^2$ is polygonal.

For $i \in \{1, \dots, m\}$ and a fixed $d \in \mathbb{N}^m$, the spaces $L_{\mathbb{P}_i}^p(\Xi_i)$ are discretized by polynomials of degree at most $d_i - 1 \in \mathbb{N}_0$. The sets of these polynomials on Ξ_i are denoted by $\mathcal{P}_{d_i-1}(\Xi_i) \subset L_{\mathbb{P}_i}^p(\Xi_i)$. By [Assumption 4.1](#), all polynomials defined on Ξ_i are \mathbb{P}_i -integrable and the space of polynomials of arbitrary degree is dense in $L_{\mathbb{P}_i}^p(\Xi_i)$, see [\[41, chap. 8\]](#). Furthermore, there exist sets $\{\beta_{k_i}^{(i)}\}_{k_i=1}^\infty \subset L_{\mathbb{P}_i}^2(\Xi_i)$ of orthonormal polynomials w.r.t. the $L_{\mathbb{P}_i}^2(\Xi_i)$ -inner product, where $\beta_{k_i}^{(i)}$ has degree $k_i - 1$ by [\[41, Thm. 8.5\]](#). These sets are Hilbert bases of $L_{\mathbb{P}_i}^2(\Xi_i)$, respectively, and can be constructed by applying the Grad-Schmidt process to the monomial basis $\{1, \xi_i, \xi_i^2, \dots\}$ for example. We want to mention that some papers [\[18, 4\]](#) dealing with uncertainty quantification restrict the discussion to certain probability distributions, for which the “classical” orthonormal polynomials of increasing degree such as Legendre or Hermite polynomials are well-known. In order to have a more general setting, we only assume that the orthonormal polynomials can be constructed and evaluated, e.g., by the three-term recurrence

²Boundedness of the intervals is assumed because then the set of polynomials of arbitrary degree is a dense subset of $L_{\mathbb{P}_i}^p(\Xi_i)$. This condition can possibly be relaxed, cf. [\[41, chap. 8\]](#).

relation [41, sec. 8.2]. In particular, we do not assume a purely continuous [19, 42, 11, 32] or a symmetric [16] distribution.

Let $\{a_{l_i}^{(i)}\}_{l_i=1}^{d_i} \subset \Xi_i$ be the d_i pairwise distinct roots of the polynomial $\beta_{d_i+1}^{(i)}$ in ascending order, respectively. They exist and have the mentioned properties due to [41, Thm. 8.16] and are known as Gaussian quadrature nodes. Let $\{w_{l_i}^{(i)}\}_{l_i=1}^{d_i}$ be the positive Gaussian quadrature weights associated to the nodes $\{a_{l_i}^{(i)}\}_{l_i=1}^{d_i}$ for $i \in [m]$ and $l_i \in [d_i]$, see [41, Def. 9.2]. We use the Gaussian quadrature nodes to define weighted Lagrange polynomials $\{\theta_{k_i}^{(i)}\}_{k_i=1}^{d_i}$, which fulfill $\theta_{k_i}^{(i)}(a_{l_i}^{(i)}) = \delta_{k_i l_i} \cdot \omega_{k_i}^{(i)}$ for some weights $\omega_{k_i}^{(i)} > 0$. These polynomials are used as bases of the polynomial subspaces $\mathcal{P}_{d_i-1}(\Xi_i)$ instead of the typical orthonormal polynomials $\{\beta_{k_i}^{(i)}\}_{k_i=1}^{d_i}$ of increasing degree. This has some useful consequences:

- These Lagrange polynomials are $L_{\mathbb{P}_i}^2(\Xi_i)$ -orthogonal. If we choose $\omega_{k_i}^{(i)} = (w_{k_i}^{(i)})^{-1/2}$ for $k_i \in \{1, \dots, d_i\}$, it follows from the exactness of Gaussian quadrature, that they are even orthonormal:

$$\int_{\Xi_i} \theta_{k_i}^{(i)} \theta_{l_i}^{(i)} d\mathbb{P}_i = \sum_{\hat{l}_i=1}^{d_i} w_{\hat{l}_i}^{(i)} \theta_{k_i}^{(i)}(a_{\hat{l}_i}^{(i)}) \theta_{l_i}^{(i)}(a_{\hat{l}_i}^{(i)}) = \sum_{\hat{l}_i=1}^{d_i} w_{\hat{l}_i}^{(i)} \delta_{k_i \hat{l}_i} (w_{k_i}^{(i)})^{-\frac{1}{2}} \delta_{l_i \hat{l}_i} (w_{l_i}^{(i)})^{-\frac{1}{2}} = \delta_{k_i l_i}.$$

The product of two Lagrange polynomials has degree $2d_i - 2$ and is integrated exactly by Gaussian quadrature, which is exact up to degree $2d_i - 1$ [41, Thm. 9.9].

- A connection to stochastic collocation methods can be established, see [19, 22].
- Nonlinear dependence on the parameters and nonlinear operators can be approximated nicely. This was already done in [22, Example 3.6] and is used here for the discretization of the nonlinearity.

Defining $\theta_{k_i}^{(i)} := \beta_{k_i}^{(i)}$ for $k_i \geq d_i + 1$, we get that $\{\theta_{k_i}^{(i)}\}_{k_i=1}^{\infty}$ is also a Hilbert basis of $L_{\mathbb{P}_i}^2(\Xi_i)$.

For the discretization of $L_{\mathbb{P}}^p(\Xi)$ we take the full tensor product

$$\mathcal{P}_{d-1}(\Xi) := \bigotimes_{i=1}^m \mathcal{P}_{d_i-1}(\Xi_i) \subset \bigotimes_{i=1}^m L_{\mathbb{P}_i}^p(\Xi_i) = L_{\mathbb{P}}^p(\Xi),$$

of polynomials of *coordinate degree* $d - 1 \in \mathbb{N}_0^m$ with $\mathbb{1} = (1, \dots, 1)^\top \in \mathbb{N}^m$. Defining $\beta_k(\xi) := \prod_{i=1}^m \beta_{k_i}^{(i)}(\xi_i)$ we obtain a set $\{\beta_k\}_{k \in \mathbb{N}^m}$ of orthonormal polynomials which form a Hilbert basis of $L_{\mathbb{P}}^2(\Xi)$, see [44, Thm. 3.12(b)]. Note that k is an index vector. Analogously, the Hilbert basis $\{\theta_k\}_{k \in \mathbb{N}^m}$ with $\theta_k(\xi) := \prod_{i=1}^m \theta_{k_i}^{(i)}(\xi_i)$ can be defined. Writing $a_l := (a_{l_1}^{(1)}, \dots, a_{l_m}^{(m)})^\top$ and $\omega_l := \prod_{i=1}^m \omega_{l_i}^{(i)}$, it holds that $\theta_k(a_l) = \delta_{kl} \omega_l$ for $k \leq d$ componentwise. We will use the set $\{\theta_k\}_{\mathbb{1} \leq k \leq d}$ as basis of $\mathcal{P}_{d-1}(\Xi)$. A polynomial $\mathbf{p} \in \mathcal{P}_{d-1}(\Xi)$ is represented by a tensor $\mathbf{p} \in \mathbb{R}^{d_1 \times \dots \times d_m}$ via

$$\mathbf{p}(\xi) = \mathbf{p}(\xi_1, \dots, \xi_m) = \sum_{k_1=1}^{d_1} \dots \sum_{k_m=1}^{d_m} \mathbf{p}(k_1, \dots, k_m) \theta_{k_1}^{(1)}(\xi_1) \dots \theta_{k_m}^{(m)}(\xi_m) = \sum_{\mathbb{1} \leq k \leq d} \mathbf{p}(k) \theta_k(\xi).$$

By inserting $\xi = a_l$ for $\mathbb{1} \leq l \leq d$ we obtain that the entries of this tensor are given by $\mathbf{p}(k) = \mathbf{p}(a_l) \omega(k)^{-1}$. The $L_{\mathbb{P}}^2(\Xi)$ -inner product of two functions $\mathbf{p}, \tilde{\mathbf{p}} \in \mathcal{P}_{d-1}(\Xi)$ is discretized by applying Gaussian quadrature:

$$(4.1) \quad (\mathbf{p}, \tilde{\mathbf{p}})_{L_{\mathbb{P}}^2(\Xi)} = \sum_{\mathbb{1} \leq k \leq d} \mathbf{w}(k) \mathbf{p}(a_k) \tilde{\mathbf{p}}(a_k) = \langle \mathbf{w} \odot \omega \odot \mathbf{p}, \omega \odot \tilde{\mathbf{p}} \rangle =: \langle \mathbf{p}, \tilde{\mathbf{p}} \rangle_{\mathbf{w} \odot \omega},$$

where “ \odot ” denotes the Hadamard (componentwise) product and “ $\langle \cdot, \cdot \rangle$ ” the Frobenius inner product of tensors. The weight tensors $\boldsymbol{\omega}, \mathbf{w} \in \mathbb{R}^{d_1 \times \dots \times d_m}$ are defined by $\boldsymbol{\omega}(k) := \prod_{i=1}^m \omega_{k_i}^{(i)}$ and $\mathbf{w}(k) := \prod_{i=1}^m w_{k_i}^{(i)}$ and have rank 1. We see that for the special choice of orthonormal polynomials with $\boldsymbol{\omega} = \mathbf{w}^{-1/2}$ (componentwise exponentiation) we get $\langle \mathbf{p}, \tilde{\mathbf{p}} \rangle_{L^2_{\mathbb{P}}(\Xi)} = \langle \mathbf{p}, \tilde{\mathbf{p}} \rangle$, a simple Frobenius inner product. A special case is the expectation of a function \mathbf{p} , computed as an inner product with the constant function $\mathbb{1}$, which is represented by the tensor $\boldsymbol{\omega}^{-1}$.

Choosing the linear FE basis $\{\phi_{k_0}\}_{k_0=1}^{d_0}$ for the FE space $\mathbf{Y} \subset Y$ as in [23], the state space $\mathbf{Y} = L^p_{\mathbb{P}}(\Xi; Y) \cong Y \otimes L^p_{\mathbb{P}_1}(\Xi_1) \otimes \dots \otimes L^p_{\mathbb{P}_m}(\Xi_m)$ is discretized by the full tensor product of the respective finite-dimensional subspaces with the basis

$$\{\phi_{k_0} \otimes \theta_{k_1}^{(1)} \otimes \dots \otimes \theta_{k_m}^{(m)}, k_i \in \{1, \dots, d_i\}, i \in \{0, \dots, m\}\}.$$

This is a basis due to [28, Lem. 3.11]. A function $\mathbf{y} \in \mathbf{Y}$ belonging to the finite-dimensional space is represented by a coefficient tensor $\mathbf{y} \in \mathbb{R}^{d_0 \times \dots \times d_m}$ corresponding to weighted values of the function \mathbf{y} since nodal FE ansatz functions and Lagrange polynomials are used. This means that

$$(4.2) \quad \mathbf{y}(x, \xi_1, \dots, \xi_m) = \sum_{k_0=1}^{d_0} \sum_{k_1=1}^{d_1} \dots \sum_{k_m=1}^{d_m} \mathbf{y}(k_0, k_1, \dots, k_m) \phi_{k_0}(x) \theta_{k_1}^{(1)}(\xi_1) \dots \theta_{k_m}^{(m)}(\xi_m)$$

and in particular $\mathbf{y}(x, \mathbf{a}_{k_1}^{(1)}, \dots, \mathbf{a}_{k_m}^{(m)}) = \sum_{k_0=1}^{d_0} \mathbf{y}(\mathring{k}) \phi_{k_0}(x) \boldsymbol{\omega}(k)$, abbreviating $\mathring{k} = (k_0, k_1, \dots, k_m)$ as well as $\mathring{l} = (l_0, l_1, \dots, l_m)$ and $(\mathring{k}, \mathring{l}) = (k_0, k_1, \dots, k_m, l_0, l_1, \dots, l_m)$ etc. here and in the following in contrast to $l = (l_1, \dots, l_m)$ etc.

We now discretize the operators defined in (2.3), (2.8) using the discretized deterministic operators

$$(4.3) \quad \begin{aligned} \mathbf{A}(\xi) &\in \mathbb{R}^{d_0 \times d_0}, & \mathbf{A}(\xi)_{k_0 l_0} &= (\kappa(\cdot, \xi) \nabla \phi_{l_0}, \nabla \phi_{k_0})_{L^2(\Omega)^n}, \\ \mathbf{B} &\in \mathbb{R}^{d_0 \times d_u}, & \mathbf{B}_{k_0 l_u} &= (D\psi_{l_u}, \phi_{k_0})_{L^2(\Omega)}, \\ \mathbf{b}(\xi) &\in \mathbb{R}^{d_0}, & \mathbf{b}(\xi)_{k_0} &= (\mathbf{f}(\cdot, \xi), \phi_{k_0})_{L^2(\Omega)}. \end{aligned}$$

from [23], where $\{\psi_\ell\}_{\ell=1}^{d_u} \subset U \subset U$ is the basis of the control subspace.

Let the control $\mathbf{u} \in U$ be represented by $\mathbf{u} \in \mathbb{R}^{d_u}$. Testing with $\mathbf{v} \in \mathbf{Y}$ represented by $\mathbf{v} \in \mathbb{R}^{d_0 \times \dots \times d_m}$ we have

$$\begin{aligned} \langle \mathbf{B}\mathbf{u}, \mathbf{v} \rangle_{\mathbf{Y}^*, \mathbf{Y}} &= \int_{\Xi} (D\mathbf{u}, \mathbf{v}(\cdot, \xi))_{L^2(\Omega)} d\mathbb{P} = \sum_{\mathbf{1} \leq \mathring{k} \leq \mathring{d}} \mathbf{v}(\mathring{k}) (D\mathbf{u}, \phi_{k_0})_{L^2(\Omega)} \int_{\Xi} \prod_{i=1}^m \theta_{k_i}^{(i)}(\xi_i) d\mathbb{P} \\ &= \langle (\mathbf{B}\mathbf{u}) \otimes (\mathbf{w} \odot \boldsymbol{\omega}), \mathbf{v} \rangle =: \langle \mathbf{B}\mathbf{u}, \mathbf{v} \rangle_{\mathbf{1} \otimes (\mathbf{w} \odot \boldsymbol{\omega}, 2)} \end{aligned}$$

with $\mathbf{B}\mathbf{u} = (\mathbf{B}\mathbf{u}) \otimes \boldsymbol{\omega}^{-1}$ denoting by “ \otimes ” the Kronecker (outer) product of tensors. Furthermore, we get

$$\langle \mathbf{b}, \mathbf{v} \rangle_{\mathbf{Y}^*, \mathbf{Y}} = \sum_{\mathbf{1} \leq \mathring{k} \leq \mathring{d}} \mathbf{v}(\mathring{k}) \int_{\Xi} \mathbf{b}_{k_0}(\xi) \prod_{i=1}^m \theta_{k_i}^{(i)}(\xi_i) d\mathbb{P} =: \langle \mathbf{b}, \mathbf{v} \rangle_{\mathbf{1} \otimes (\mathbf{w} \odot \boldsymbol{\omega}, 2)}$$

with $\mathbf{b}(\mathring{k}) = \frac{1}{\prod_{i=1}^m w_{k_i}^{(i)} (\omega_{k_i}^{(i)})^2} \left(\int_{\Xi} \mathbf{b}_{k_0}(\xi) \prod_{i=1}^m \theta_{k_i}^{(i)}(\xi_i) d\mathbb{P} \right)$. We assume that this tensor can be constructed and has sufficiently small rank. In particular, if $\mathbf{b}_{k_0}(\cdot)$ are polynomials of total

degree at most d , Gaussian quadrature, which is exact up to coordinate degree $2d - 1$, can be applied to compute

$$(4.4) \quad \int_{\Xi} \mathbf{b}_{k_0}(\xi) \prod_{i=1}^m \theta_{k_i}^{(i)}(\xi_i) d\mathbb{P} = \sum_{1 \leq l \leq d} \mathbf{w}(l) \mathbf{b}_{k_0}(\mathbf{a}_l) \prod_{i=1}^m \theta_{k_i}^{(i)}(\mathbf{a}_{l_i}^{(i)}) = \mathbf{w}(k) \boldsymbol{\omega}(k) \mathbf{b}_{k_0}(\mathbf{a}_k)$$

so that $\mathbf{b}(\cdot, k) = \boldsymbol{\omega}(k)^{-1} \mathbf{b}(\mathbf{a}_k)$ holds. Moreover,

$$(4.5) \quad \begin{aligned} \langle \mathbf{A}\mathbf{y}, \mathbf{v} \rangle_{\mathbf{Y}^*, \mathbf{Y}} &= \int_{\Xi} (\kappa(\cdot, \xi) \nabla_x \mathbf{y}(\cdot, \xi), \nabla_x \mathbf{v}(\cdot, \xi))_{L^2(\Omega)^n} d\mathbb{P} \\ &= \sum_{1 \leq \mathring{k} \leq \mathring{d}} \sum_{1 \leq \mathring{l} \leq \mathring{d}} \mathbf{y}(\mathring{l}) \mathbf{v}(\mathring{k}) \int_{\Xi} (\mathbf{A}(\xi))_{k_0 l_0} \prod_{i=1}^m \theta_{l_i}^{(i)}(\xi_i) \theta_{k_i}^{(i)}(\xi_i) d\mathbb{P} \\ &= \langle \langle \tilde{\mathbf{a}}, \mathbf{y} \rangle_{(m+2, \dots, 2m+2), (1, \dots, m+1)}, \mathbf{v} \rangle_{\mathbb{1} \otimes (\mathbf{w} \odot \boldsymbol{\omega}, 2)} =: \langle \mathbf{A}\mathbf{y}, \mathbf{v} \rangle_{\mathbb{1} \otimes (\mathbf{w} \odot \boldsymbol{\omega}, 2)} \end{aligned}$$

holds with the tensor $\tilde{\mathbf{a}} \in \mathbb{R}^{d_0 \times d_1 \times \dots \times d_m \times d_0 \times d_1 \times \dots \times d_m}$ defined by

$$(4.6) \quad \tilde{\mathbf{a}}(\mathring{k}, \mathring{l}) := \frac{1}{\prod_{i=1}^m \mathbf{w}_{k_i}^{(i)}(\boldsymbol{\omega}_{k_i}^{(i)})^2} \left(\int_{\Xi} (\mathbf{A}(\xi))_{k_0 l_0} \prod_{i=1}^m \theta_{l_i}^{(i)}(\xi_i) \theta_{k_i}^{(i)}(\xi_i) d\mathbb{P} \right).$$

The notation $\langle \cdot, \cdot \rangle_{s,t}$ stands for a contraction of tensors along the mode tuples s and t . Here, we would obtain $\langle \tilde{\mathbf{a}}, \mathbf{y} \rangle_{(m+2, \dots, 2m+2), (1, \dots, m+1)} \in \mathbb{R}^{d_0 \times \dots \times d_m}$ by reshaping the result of the matrix-vector product of the matricization $\text{mat}(\tilde{\mathbf{a}}) \in \mathbb{R}^{d_0 \dots d_m \times d_0 \dots d_m}$ and the vectorization $\text{vec}(\mathbf{y}) \in \mathbb{R}^{d_0 \dots d_m}$. The nonlinear part of the equation is approximated by Gaussian quadrature/interpolation to simplify the implementation.

$$(4.7) \quad \begin{aligned} \langle \mathbf{N}(\mathbf{y}), \mathbf{v} \rangle_{\mathbf{Y}^*, \mathbf{Y}} &= \int_{\Xi} \int_{\Omega} \varphi(\mathbf{y}(x, \xi)) \mathbf{v}(x, \xi) dx d\mathbb{P} \\ &= \sum_{1 \leq \mathring{k} \leq \mathring{d}} \mathbf{v}(\mathring{k}) \int_{\Xi} \int_{\Omega} \varphi(\mathbf{y}(x, \xi)) \phi_{k_0}(x) dx \prod_{i=1}^m \theta_{k_i}^{(i)}(\xi_i) d\mathbb{P} \\ &\approx \sum_{1 \leq \mathring{k} \leq \mathring{d}} \mathbf{v}(\mathring{k}) \sum_{1 \leq l \leq d} \mathbf{w}(l) \int_{\Omega} \varphi(\mathbf{y}(x, \mathbf{a}_l)) \phi_{k_0}(x) dx \prod_{i=1}^m \theta_{k_i}^{(i)}(\mathbf{a}_{l_i}^{(i)}) \\ &= \sum_{1 \leq \mathring{k} \leq \mathring{d}} \mathbf{v}(\mathring{k}) \mathbf{w}(k) \boldsymbol{\omega}(k) \int_{\Omega} \varphi \left(\sum_{l_0=1}^{d_0} \mathbf{y}(l_0, k_1, \dots, k_m) \phi_{l_0}(x) \boldsymbol{\omega}(k) \right) \phi_{k_0}(x) dx \\ &\approx \sum_{1 \leq \mathring{k} \leq \mathring{d}} \mathbf{v}(\mathring{k}) \mathbf{w}(k) \boldsymbol{\omega}(k) \varphi(\mathbf{y}(\mathring{k}) \boldsymbol{\omega}(k)) (\mathbf{M}_L)_{k_0 k_0} \\ &= \langle (\mathbb{1} \otimes \boldsymbol{\omega}^{-1}) \odot (\mathbf{M}_L \circ_1 \varphi(\mathbb{1} \otimes \boldsymbol{\omega}) \odot \mathbf{y}), \mathbf{v} \rangle_{\mathbb{1} \otimes (\mathbf{w} \odot \boldsymbol{\omega}, 2)} \\ &= \langle \hat{\boldsymbol{\omega}}^{-1} \odot (\mathbf{M}_L \circ_1 \varphi(\hat{\boldsymbol{\omega}} \odot \mathbf{y})), \mathbf{v} \rangle_{\mathbb{1} \otimes (\mathbf{w} \odot \boldsymbol{\omega}, 2)} =: \langle \mathbf{N}(\mathbf{y}), \mathbf{v} \rangle_{\mathbb{1} \otimes (\mathbf{w} \odot \boldsymbol{\omega}, 2)}, \end{aligned}$$

where the first approximate equality is due to Gaussian quadrature in the parameter space and the second one is due to an FE nodes based quadrature formula, which is related to mass lumping, see [23]. By $\mathbf{M}_L \in \mathbb{R}^{d_0 \times d_0}$ we denote the lumped mass matrix corresponding to the deterministic, discrete state space \mathbf{Y} . It is applied to the first mode of the tensor $\varphi(\hat{\boldsymbol{\omega}} \odot \mathbf{y}) \in \mathbb{R}^{d_0 \times \dots \times d_m}$, i.e., we obtain $\mathbf{M}_L \circ_1 \varphi(\hat{\boldsymbol{\omega}} \odot \mathbf{y}) \in \mathbb{R}^{d_0 \times \dots \times d_m}$ by tensorizing the result of the matrix product $\mathbf{M}_L \text{mat}(\varphi(\hat{\boldsymbol{\omega}} \odot \mathbf{y})) \in \mathbb{R}^{d_0 \times d_1 \dots d_m}$ with $\text{mat}(\varphi(\hat{\boldsymbol{\omega}} \odot \mathbf{y})) \in \mathbb{R}^{d_0 \times d_1 \dots d_m}$. We define the rank-1-tensor $\hat{\boldsymbol{\omega}} := \mathbb{1} \otimes \boldsymbol{\omega}$ for readability purposes. Here we see the importance of

using mass lumping and weighted Lagrange polynomials: It yields that the nonlinear function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ can be applied componentwise to a tensor in the discrete setting (4.7). In fact, the tensor $\hat{\omega} \odot \mathbf{y}$ contains exactly the function values of the state at the FE nodes and Gaussian quadrature grid points. Therefore, pointwise operations, such as the application of φ' , carry over to componentwise operations on tensors in the discrete setting. Multiplying a tensor of functions values by $\hat{\omega}^{-1}$ again transforms it back to the representation with weighted Lagrange polynomials.

Overall, the discrete state equation reads

$$(4.8) \quad \mathbf{A}\mathbf{y} + \mathbf{N}(\mathbf{y}) = \mathbf{B}\mathbf{u} + \mathbf{b}.$$

As in [23], the discrete subspace H of H is isomorphic to \mathbb{R}^{d_H} ($d_H \in \mathbb{N}$) equipped with the inner product induced by $M_H \in \mathbb{R}^{d_H \times d_H}$. The discrete version of $Q(\xi)$ and $\hat{q}(\xi)$ are $Q(\xi) \in \mathbb{R}^{d_H \times d_0}$ and $\hat{q}(\xi) \in \mathbb{R}^{d_H}$, respectively. The discrete, parameter-dependent objective function is given by

$$(4.9) \quad J[\xi](y, u) = \frac{1}{2} \|Q(\xi)y - \hat{q}(\xi)\|_{M_H}^2 + \frac{\gamma}{2} u^\top \tilde{M}u,$$

see [23, Eq. (5.3)].

Assumption 4.2. In order to be able to evaluate the objective function from (2.6) on the finite-dimensional subspace exactly in a simple way, we assume that the operator $Q(\xi) = Q$ is constant and the desired state $\hat{q}(\cdot)$ is a polynomial of coordinate degree at most $d - 1$.

Using (4.9), the discrete version of the objective function is

$$(4.10) \quad \begin{aligned} \mathbf{J}(\mathbf{y}, u) &:= \int_{\Xi} J[\xi] \left(\sum_{1 \leq k \leq d} \mathbf{y}(\cdot, k_1, \dots, k_m) \theta_{k_1}^{(1)}(\xi_1) \cdots \theta_{k_m}^{(m)}(\xi_m), u \right) d\mathbb{P} \\ &= \int_{\Xi} \frac{1}{2} \left\| \sum_{1 \leq k \leq d} Q(\xi) \mathbf{y}(\cdot, k) \theta_{k_1}^{(1)}(\xi_1) \cdots \theta_{k_m}^{(m)}(\xi_m) - \hat{q}(\xi) \right\|_{M_H}^2 d\mathbb{P} + \frac{\gamma}{2} u^\top \tilde{M}u, \end{aligned}$$

where $\mathbf{y}(\cdot, k) \in \mathbb{R}^{d_0}$ is obtained from the tensor \mathbf{y} by fixing all indices except the first one. By *Assumption 4.2*, the integrand in (4.10) has degree at most $2d - 2 \cdot 1$ and the integral can be evaluated exactly by Gaussian quadrature. This gives

$$(4.11) \quad \begin{aligned} \mathbf{J}(\mathbf{y}, u) &= \sum_{1 \leq l \leq d} \mathbf{w}(l) \frac{1}{2} \left\| \omega(l) (Q \mathbf{y}(\cdot, l)) - \hat{q}(a_l) \right\|_{M_H}^2 + \frac{\gamma}{2} u^\top \tilde{M}u \\ &=: \langle \mathbf{M}_H (\mathbf{Q}\mathbf{y} - \hat{\mathbf{q}}), \mathbf{Q}\mathbf{y} - \hat{\mathbf{q}} \rangle + \frac{\gamma}{2} u^\top \tilde{M}u \end{aligned}$$

with $\mathbf{Q}\mathbf{y} = Q \circ_1 \mathbf{y}$, $\hat{\mathbf{q}} \in \mathbb{R}^{d_H \times d_1 \times \cdots \times d_m}$ defined by $\hat{\mathbf{q}}(\cdot, l) = \hat{q}(a_l) \omega(l)^{-1}$, and

$$(4.12) \quad \mathbf{M}_H : \mathbb{R}^{d_H \times d_1 \times \cdots \times d_m} \rightarrow \mathbb{R}^{d_H \times d_1 \times \cdots \times d_m}, \quad \mathbf{M}_H \hat{\mathbf{q}} := (\mathbb{1} \otimes (\mathbf{w} \odot \omega^2)) \odot (M_H \circ_1 \hat{\mathbf{q}}).$$

This evaluation is exact on the discrete subspace and the objective function evaluation error can be estimated using *Proposition 3.5*.

For the computation of the gradient of the reduced objective function, it remains to discretize the adjoint equation (2.19). The adjoint state \mathbf{z} is represented by the tensor \mathbf{z} analogously to the state. We have

$$\begin{aligned} \langle \mathbf{N}'(\mathbf{y})\mathbf{z}, \mathbf{v} \rangle_{\mathbf{Y}^*, \mathbf{Y}} &= \int_{\Xi} \int_{\Omega} \varphi'(\mathbf{y}(x, \xi)) \mathbf{z}(x, \xi) \mathbf{v}(x, \xi) dx d\mathbb{P} \\ &\approx \sum_{1 \leq k \leq d} \mathbf{v}(k) \sum_{1 \leq l \leq d} \mathbf{w}(l) \int_{\Omega} \varphi'(\mathbf{y}(x, a_l)) \mathbf{z}(x, a_l) \phi_{k_0}(x) dx \prod_{i=1}^m \theta_{k_i}^{(i)}(a_{l_i}^{(i)}) \end{aligned}$$

$$\begin{aligned}
&\approx \sum_{\substack{\circ \\ \mathbb{1} \leq k \leq d}} \mathbf{v}(k) \mathbf{w}(k) \boldsymbol{\omega}(k) \boldsymbol{\varphi}'(\mathbf{y}(k) \boldsymbol{\omega}(k)) \mathbf{z}(k) \boldsymbol{\omega}(k) (\mathbf{M}_L)_{k_0 k_0} \\
&= \langle \hat{\boldsymbol{\omega}}^{-1} \circ (\mathbf{M}_L \circ_1 (\boldsymbol{\varphi}'(\hat{\boldsymbol{\omega}} \circ \mathbf{y}) \circ (\hat{\boldsymbol{\omega}} \circ \mathbf{z}))), \mathbf{v} \rangle_{\mathbb{1} \otimes (\mathbf{w} \circ \boldsymbol{\omega}, 2)} \\
&= \langle \mathbf{N}'(\mathbf{y}) \mathbf{z}, \mathbf{v} \rangle_{\mathbb{1} \otimes (\mathbf{w} \circ \boldsymbol{\omega}, 2)},
\end{aligned}$$

cf. (4.7), where \mathbf{N}' is exactly the derivative of the discretized operator \mathbf{N} . Furthermore, the right-hand side is

$$\begin{aligned}
&\langle -\boldsymbol{Q}^* (\boldsymbol{Q} \mathbf{y} - \hat{\mathbf{q}}), \mathbf{v} \rangle_{\mathbf{y}^*, \mathbf{y}} \\
&= \sum_{\mathbb{1} \leq k \leq d} \mathbf{v}(\cdot, k)^\top \int_{\Xi} \left(-\mathbf{Q}^\top \mathbf{M}_H \left(\sum_{l_1, \dots, l_m=1}^{d_1, \dots, d_m} \mathbf{Q} \mathbf{y}(\cdot, l) \prod_{i=1}^m \theta_i^{(i)}(\xi_i) - \hat{\mathbf{q}}(\xi) \right) \right) \prod_{i=1}^m \theta_{k_i}^{(i)}(\xi_i) d\mathbb{P} \\
&= \sum_{\mathbb{1} \leq k \leq d} \mathbf{v}(\cdot, k)^\top \mathbf{w}(k) \boldsymbol{\omega}(k) \left(-\mathbf{Q}^\top \mathbf{M}_H \left(\mathbf{Q} \mathbf{y}(\cdot, k) \boldsymbol{\omega}(k) - \hat{\mathbf{q}}(\mathbf{a}_k) \right) \right) \\
(4.13) \quad &= \langle -(\mathbf{Q}^\top \mathbf{M}_H) \circ_1 (\boldsymbol{Q} \mathbf{y} - \hat{\mathbf{q}}), \mathbf{v} \rangle_{\mathbb{1} \otimes (\mathbf{w} \circ \boldsymbol{\omega}, 2)}.
\end{aligned}$$

Given the solution \mathbf{z} of the discretized adjoint equation

$$(4.14) \quad \mathbf{A} \mathbf{z} + \mathbf{N}'(\mathbf{y}) \mathbf{z} = -(\mathbf{Q}^\top \mathbf{M}_H) \circ_1 (\boldsymbol{Q} \mathbf{y} - \hat{\mathbf{q}}),$$

the gradient of the reduced, discretized objective function reads

$$(4.15) \quad \nabla \hat{\mathbf{J}}(\mathbf{u}) = -(\tilde{\mathbf{M}}^{-1} \mathbf{B}^\top) \langle \mathbf{z}, \mathbf{w} \circ \boldsymbol{\omega} \rangle_{(2, \dots, m+1), (1, \dots, m)} + \gamma \mathbf{u}.$$

Due to $\mathbf{w} \circ \boldsymbol{\omega} = (\mathbf{w} \circ \boldsymbol{\omega}, 2) \circ \boldsymbol{\omega}, 1$, the term $\langle \mathbf{z}, \mathbf{w} \circ \boldsymbol{\omega} \rangle_{(2, \dots, m+1), (1, \dots, m)}$ corresponds to computing the $L_{\mathbb{P}}^2(\Xi)$ -inner product (induced by $\mathbf{w} \circ \boldsymbol{\omega}, 2$) of \mathbf{z} and the function $\mathbb{1}$ (represented by $\boldsymbol{\omega}, 1$), i.e., computing the expectation of \mathbf{z} , see also (4.1) and the paragraph below this equation. As in [23], the matrix $(\tilde{\mathbf{M}}^{-1} \mathbf{B}^\top)$ can often be applied without inverting the mass matrix $\tilde{\mathbf{M}}$ so that [Theorem 3.2](#) can be applied to estimate the gradient error.

Choice of the trust-region model. As in [23], let $\mathbf{u}^k \in U$ be the current iterate of the trust-region algorithm, represented by the vector $\mathbf{u}^k \in \mathbb{R}^{d_u}$. The tensors $\tilde{\mathbf{y}}^k$ and $\tilde{\mathbf{z}}^k$ are (approximate) solutions of (4.8) and (4.14) with $\mathbf{u} = \mathbf{u}^k$ and $\mathbf{y} = \tilde{\mathbf{y}}^k$, respectively. Then, we choose the quadratic trust-region model $m_k(\mathbf{s}) = m_k(\mathbf{s}) := \nabla m_k(0)^\top \tilde{\mathbf{M}} \mathbf{s} + \frac{1}{2} \mathbf{s}^\top \tilde{\mathbf{M}} \nabla^2 m_k(0) \mathbf{s}$, where $\tilde{\mathbf{M}} \in \mathbb{R}^{d_u \times d_u}$ induces the U inner product and

$$(4.16) \quad \nabla m_k(0) = -(\tilde{\mathbf{M}}^{-1} \mathbf{B}^\top) \langle \tilde{\mathbf{z}}^k, \mathbf{w} \circ \boldsymbol{\omega} \rangle_{(2, \dots, m+1), (1, \dots, m)} + \gamma \mathbf{u}^k,$$

cf. (4.15). As model Hessian $\nabla^2 m_k(0)$, we choose a deterministic reference Hessian, which is computed using the reference operator $\mathbf{A}_{\text{ref}} = \mathbf{A}(\tilde{\xi})$ with the expected value $\tilde{\xi} := \int_{\Xi} \xi d\mathbb{P}$ and the expected computed state and adjoint state $\langle \tilde{\mathbf{y}}^k, \mathbf{w} \circ \boldsymbol{\omega} \rangle_{(2, \dots, m+1), (1, \dots, m)}$ and $\langle \tilde{\mathbf{z}}^k, \mathbf{w} \circ \boldsymbol{\omega} \rangle_{(2, \dots, m+1), (1, \dots, m)}$, respectively. Doing that, we can reuse the code from the deterministic Hessian computation and save many costly solves of PDEs with uncertain coefficients. Furthermore, this approximation has turned out to work well in practice [22].

Application of the operators to low-rank tensors. To make the solution of the state and the adjoint equation and further computations efficient, we represent all tensors, such as $\mathbf{y} \in \mathbb{R}^{d_0 \times \dots \times d_m}$ and $\mathbf{z} \in \mathbb{R}^{d_0 \times \dots \times d_m}$, in a modern low-rank format, namely the tensor train (TT) [37] or the hierarchical Tucker (HT) [29] format. As already noted in [18], this is related to reduced basis methods [12, 11]. The aim of this procedure is to overcome the curse of dimensionality because the tensor \mathbf{y} has $\prod_{i=0}^m d_i$ entries so that its storage requirement easily exceeds

any available memory if, e.g., $d_i \geq 2$ and m is large. For a short introduction to low-rank tensors and the notation used here we refer to [22], for overviews to [28, 26]. The essential idea of these formats is to generalize the notion of low-rank matrices and low-rank approximations by truncated singular value decompositions (SVDs) to higher dimensions. Like a low-rank matrix, which can be represented as product of two smaller matrices, a low-rank tensor is represented as a contraction of smaller tensors. In particular, the required storage of the mentioned formats scales linearly in the tensor order $m + 1$ instead of exponentially. An issue when dealing with these formats is that full tensors should never be formed explicitly. Instead, one aims to stay in the low-rank format, which comes with the disadvantage that only a limited set of operations is available within those. For example, the tensor contractions mentioned earlier, which include outer and inner products as well as applications of matrices to a tensor mode, can be implemented efficiently. Furthermore, componentwise addition and multiplication are available, but come with the drawback that the respective tensor ranks sum or multiply as well, respectively. To avoid infeasible rank growth—the storage requirements of TT and HT tensors scale quadratically and cubically in the ranks, respectively—it is therefore necessary to truncate the tensors to lower rank, i.e., approximate them with, e.g., the TT-SVD [37] or the HSVD [25], which are algorithms using truncated SVDs. If desired, the approximation error can be prescribed during this procedure so that required ranks are chosen automatically. Further componentwise operations such as the application of nonlinear functions and even computing the elementwise reciprocal requires more work, e.g., the application of iterative methods or sampling some entries of the result and using low-rank tensor completion [38, 31, 36].

For the solution of linear systems in tensor space, one challenge is guessing the required tensor rank for obtaining a sufficiently accurate solution. We will solve equations (4.8) and (4.14) using AMEn [14], which operates on the core tensors of the TT format and chooses the tensor rank adaptively. In order to be able to apply any low-rank tensor solver, it is important that all operators are efficiently applicable to low-rank tensors. This shall be discussed in the following. We see that the nonlinear function \mathbf{N} and its derivative \mathbf{N}' can be applied to low-rank tensors as long as the elementwise application of the functions φ and φ' can be implemented efficiently because the rest of the computations consists of multiplications with rank-1-tensors, i -mode matrix multiplications, and componentwise multiplications to apply $\mathbf{N}'(\mathbf{y})$. In particular, the componentwise multiplication by a rank-1-tensor can be written as i -mode multiplications with diagonal matrices. In our case, it holds that

$$(4.17) \quad \mathbf{N}(\mathbf{y}) = \mathbf{M}_L \circ_1 \text{diag}(\omega^{(1)}) \circ_2 \dots \text{diag}(\omega^{(m)}) \circ_{m+1} \varphi(\text{diag}(\omega^{(1)}) \circ_2 \dots \text{diag}(\omega^{(m)}) \circ_{m+1} \mathbf{y}).$$

Analogously, the operator \mathbf{M}_H defined in (4.12) can be implemented efficiently using i -mode matrix products only. More i -mode matrix products, outer and inner products appear in form of the operator \mathbf{Q} used in (4.11) and (4.14), the operator \mathbf{B} , and in (4.13), (4.15). There, the contraction with a rank-1-tensor can be carried out by

$$\langle \mathbf{z}, \mathbf{w} \odot \boldsymbol{\omega} \rangle_{(2, \dots, m+1), (1, \dots, m)} = ((\mathbf{w}^{(1)} \odot \boldsymbol{\omega}^{(1)})^\top \circ_2 \dots (\mathbf{w}^{(m)} \odot \boldsymbol{\omega}^{(m)})^\top \circ_{m+1} \mathbf{z})(\bullet, 1, \dots, 1),$$

where the indexing at the end is done to remove the tensor modes of dimension one. It only remains to investigate the application of the operator \mathbf{A} to low-rank tensors. Then all operations in (4.8), (4.11), (4.14), and (4.15) can be realized in an efficient way.

By (4.5), the application of the operator \mathbf{A} can be performed as contraction with the tensor $\tilde{\mathbf{a}}$ defined in (4.6). If this tensor can be represented or approximated in the used low-rank format, this contraction can be computed easily. Another favorable option is having the operator \mathbf{A} of small canonical rank, i.e., as a sum of a few rank-1-operators. We can achieve this by using additional structure.

EXAMPLE 4.3. We consider the specific, affine form $\kappa(x, \xi) = \kappa_0(x) + \sum_{i=1}^m \xi_i \kappa_i(x)$ of the coefficient function since the a posteriori error estimator developed in subsection 4.2 depends on this structure. This form can originate in a truncated Karhunen-Loève expansion. For $i \in \{0, \dots, m\}$ we define the operators $A_i \in \mathcal{L}(Y, Y^*)$ by $\langle A_i \mathbf{y}, \mathbf{v} \rangle_{Y^*, Y} := (\kappa_i \nabla \mathbf{y}, \nabla \mathbf{v})_{L^2(\Omega)^n}$. Their discrete counterparts are

$$A_i \in \mathbb{R}^{d_0 \times d_0}, \quad (A_i)_{k_0 l_0} := (\kappa_i \nabla \phi_{l_0}, \nabla \phi_{k_0})_{L^2(\Omega)^n}$$

and it holds that $A(\xi) = A_0 + \sum_{i=1}^m \xi_i A_i$. Then, as pointed out in [22, Example 3.6], we obtain $\mathbf{A} \mathbf{y} = A_0 \circ_1 \mathbf{y} + \sum_{i=1}^m A_i \circ_1 \tilde{S}_i \circ_{i+1} \mathbf{y}$ with $\tilde{S}_i = \text{diag}(\mathbf{a}^{(i)})$, where $\mathbf{a}^{(i)} \in \mathbb{R}^{d_i}$ is the vector of Gaussian quadrature nodes on Ξ_i . We see that the operator \mathbf{A} admits CP rank $m+1$ and can be implemented using simple i -mode matrix multiplications and summation.

4.2. A posteriori error estimation. The a posteriori error estimator for elliptic, semi-linear PDEs with uncertain coefficients, which is presented in the following, is based on [7, 8, 16, 17, 18] and the estimator for the deterministic equation discussed in [23, sec. 5.2]. Assumption 4.1 shall still be valid, i.e., we work with a vector ξ of independent random variables distributed on $\Xi := \times_{i=1}^m \Xi_i$ with the probability measure $\mathbb{P} := \otimes_{i=1}^m \mathbb{P}_i$. We combine the deterministic FE a posteriori error estimator from [23] with one for the polynomial discretization of $L_{\mathbb{P}}^2(\Xi)$ in tensor form presented in subsection 4.1 to obtain an overall error estimator for the solution of (2.2) and (2.19).

To unify the discussion of the state and adjoint equation, cf. [23, Remark 5.3], we define the following operators in analogy to (2.3) and (2.8), but with slight differences:

$$\begin{aligned} \langle \hat{\mathbf{A}} \mathbf{y}, \mathbf{v} \rangle_{Y^*, Y} &:= \int_{\Xi} \int_{\Omega} \hat{\kappa}(x, \xi) \nabla_x \mathbf{y}(x, \xi) \cdot \nabla_x \mathbf{v}(x, \xi) + \hat{\chi}(x, \xi) \mathbf{y}(x, \xi) \mathbf{v}(x, \xi) \, dx \, d\mathbb{P}, \\ (4.18) \quad \langle \hat{\mathbf{N}}(\mathbf{y}), \mathbf{v} \rangle_{Y^*, Y} &:= \int_{\Xi} \int_{\Omega} \hat{\phi}(\mathbf{y}(x, \xi)) \mathbf{v}(x, \xi) \, dx \, d\mathbb{P}, \quad \langle \hat{\mathbf{b}}, \mathbf{v} \rangle_{Y^*, Y} := \int_{\Xi} \int_{\Omega} \hat{\mathbf{f}} \mathbf{v} \, dx \, d\mathbb{P}, \\ \langle \hat{\mathbf{A}}_{\text{ref}} \mathbf{y}, \mathbf{v} \rangle_{Y^*, Y} &:= \int_{\Xi} \int_{\Omega} \hat{\kappa}_{\text{ref}}(x) \nabla_x \mathbf{y}(x, \xi) \cdot \nabla_x \mathbf{v}(x, \xi) + \hat{\chi}_{\text{ref}}(x) \mathbf{y}(x, \xi) \mathbf{v}(x, \xi) \, dx \, d\mathbb{P} \end{aligned}$$

with $\hat{\kappa}, \hat{\chi} \in L^\infty(\Omega \times \Xi)$ such that $\hat{\chi}(x, \xi) \geq 0$ and $\underline{\kappa} \leq \hat{\kappa}(x, \xi) \leq \bar{\kappa}$ for a.e. $(x, \xi) \in \Omega \times \Xi$ with $0 < \underline{\kappa} \leq \bar{\kappa} < \infty$. Furthermore, we have the deterministic reference coefficients $\hat{\kappa}_{\text{ref}} \in L^\infty(\Omega)$ (uniformly positive) and $\hat{\chi}_{\text{ref}} \in L^\infty(\Omega)$ (nonnegative). As before, $\hat{\mathbf{f}} \in L_{\mathbb{P}}^{\hat{r}_f}(\Xi; L^2(\Omega))$, $\hat{r}_f \in [p, \infty]$, and the function $\hat{\phi} \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ shall be increasing and fulfill (2.7).

We consider the operator equation

$$(4.19) \quad \langle \hat{\mathbf{A}} \mathbf{y} + \hat{\mathbf{N}}(\mathbf{y}), \mathbf{v} \rangle_{Y^*, Y} = \langle \hat{\mathbf{b}}, \mathbf{v} \rangle_{Y^*, Y} \quad \text{for all } \mathbf{v} \in Y.$$

The exact solution \mathbf{y} of (4.19) is contained in $Y = L_{\mathbb{P}}^p(\Xi; Y)$ with $p \in (3, \infty)$, which we require for the weak formulation to be well-defined and for the nonlinear operator $\hat{\mathbf{N}}$ to be twice continuously differentiable, but we will provide error estimation only in $\check{Y} := L_{\mathbb{P}}^2(\Xi; Y)$, because we want to make use of (Galerkin) orthogonality. For this purpose, we use the reference operator $\hat{\mathbf{A}}_{\text{ref}}$ and the inner product induced by its version $\check{\mathbf{A}}_{\text{ref}} \in \mathcal{L}(\check{Y}, \check{Y}^*)$ on \check{Y} with $\check{Y}^* = L_{\mathbb{P}}^2(\Xi; Y^*)$, i.e., the operator $\check{\mathbf{A}}_{\text{ref}}$ is defined exactly as $\hat{\mathbf{A}}_{\text{ref}}$ in (4.18), but on the larger space \check{Y} . It is well-defined and has a bounded inverse $\check{\mathbf{A}}_{\text{ref}}^{-1} \in \mathcal{L}(\check{Y}^*, \check{Y})$ due to the mentioned properties of the coefficients $\hat{\kappa}_{\text{ref}}$ and $\hat{\chi}_{\text{ref}}$. Analogously, the version $\check{\mathbf{A}} \in \mathcal{L}(\check{Y}, \check{Y}^*)$ and the inverse operators $\check{\mathbf{A}}^{-1}, \check{\mathbf{A}}_{\text{ref}}^{-1}$ induce inner products and norms on \check{Y} and \check{Y}^* , respectively. These norms are equivalent to the standard norms on \check{Y} and \check{Y}^* , respectively. Hence, as in [23,

Eqs. (5.10) and (5.11)], we have the equivalence estimates

$$(4.20) \quad \lambda \langle \check{\mathbf{A}} \mathbf{v}, \mathbf{v} \rangle_{\check{\mathbf{Y}}^*, \check{\mathbf{Y}}} \leq \langle \check{\mathbf{A}}_{\text{ref}} \mathbf{v}, \mathbf{v} \rangle_{\check{\mathbf{Y}}^*, \check{\mathbf{Y}}} = (\mathbf{v}, \mathbf{v})_{\check{\mathbf{A}}} \leq \Lambda \langle \check{\mathbf{A}} \mathbf{v}, \mathbf{v} \rangle_{\check{\mathbf{Y}}^*, \check{\mathbf{Y}}},$$

$$(4.21) \quad \frac{1}{\Lambda} \|\mathbf{b}\|_{\check{\mathbf{A}}^{-1}}^2 \leq \langle \mathbf{b}, \check{\mathbf{A}}_{\text{ref}}^{-1} \mathbf{b} \rangle_{\mathbf{Y}^*, \mathbf{Y}} = \|\mathbf{b}\|_{\check{\mathbf{A}}_{\text{ref}}^{-1}}^2 \leq \frac{1}{\lambda} \|\mathbf{b}\|_{\check{\mathbf{A}}^{-1}}^2$$

for every $\mathbf{v} \in \check{\mathbf{Y}}$, $\mathbf{b} \in \check{\mathbf{Y}}^*$ with $0 < \lambda \leq \Lambda < \infty$.

Let $\mathbf{y} \in \mathbf{Y}$ be the exact solution of (4.19), and let $\tilde{\mathbf{y}} \in \mathbf{Y} \subset \mathbf{Y}$ be an inexact solution from the discrete subspace $\mathbf{Y} = \mathbf{Y} \otimes \mathcal{P}_{d-1}(\Xi)$, defining the residual $\mathbf{r} := \check{\mathbf{A}} \tilde{\mathbf{y}} + \check{\mathbf{N}}(\tilde{\mathbf{y}}) - \hat{\mathbf{b}}$. Then, by [23, Lem. 5.1], we have the estimate $\|\tilde{\mathbf{y}} - \mathbf{y}\|_{\check{\mathbf{A}}_{\text{ref}}} \leq \Lambda \|\mathbf{r}\|_{\check{\mathbf{A}}_{\text{ref}}^{-1}}$, provided $\mathbf{r} \in L_{\mathbb{P}}^2(\Xi; \mathbf{Y}^*)$. The required integrability of the residual can be concluded using the integrability properties of $\tilde{\mathbf{y}}$, cf. section 2. In fact, if we have $\tilde{\mathbf{y}} \in L_{\mathbb{P}}^{\hat{r}_f}(\Xi; \mathbf{Y})$, it follows that $\hat{\mathbf{A}} \tilde{\mathbf{y}}, \hat{\mathbf{b}} \in L_{\mathbb{P}}^{\hat{r}_f}(\Xi; \mathbf{Y}^*)$ and $\hat{\mathbf{N}}(\tilde{\mathbf{y}}) \in L_{\mathbb{P}}^{\frac{\hat{r}_f}{p-1}}(\Xi; \mathbf{Y}^*)$ so that $\hat{r}_f \geq 2p - 2$ is sufficient for $\mathbf{r} \in L_{\mathbb{P}}^2(\Xi; \mathbf{Y}^*)$.

Defining $\mathbf{w} := \check{\mathbf{A}}_{\text{ref}}^{-1} \mathbf{r} \in L_{\mathbb{P}}^2(\Xi; \mathbf{Y})$, we have $\|\mathbf{r}\|_{\check{\mathbf{A}}_{\text{ref}}^{-1}}^2 = \|\mathbf{w}\|_{\check{\mathbf{A}}_{\text{ref}}}^2$. We compute a discrete version $\mathbf{w} \in \mathbf{Y}$ fulfilling

$$(4.22) \quad \langle \check{\mathbf{A}}_{\text{ref}} \mathbf{w}, \mathbf{v} \rangle_{\check{\mathbf{Y}}^*, \check{\mathbf{Y}}} = \langle \mathbf{r}, \mathbf{v} \rangle_{\check{\mathbf{Y}}^*, \check{\mathbf{Y}}} \text{ for all } \mathbf{v} \in \mathbf{Y}.$$

Having in mind that $\check{\mathbf{A}}_{\text{ref}}$ comes from deterministic reference coefficients, we assume that this equation is solved exactly. If this is not the case, the algebraic error caused by the inexact solution of the discrete equation has to be incorporated additionally. As in [23, Lem. 5.2] it follows that

$$(4.23) \quad \|\tilde{\mathbf{y}} - \mathbf{y}\|_{\check{\mathbf{A}}_{\text{ref}}}^2 \leq \Lambda^2 (\|\mathbf{w}\|_{\check{\mathbf{A}}_{\text{ref}}}^2 + \|\mathbf{w} - \mathbf{w}\|_{\check{\mathbf{A}}_{\text{ref}}}^2)$$

due to Galerkin orthogonality. The first summand in (4.23) is the algebraic error, which shall be minimized by a low-rank tensor solver applied to a discretized version of (4.19). The second summand will be estimated by applying a posteriori error estimates to (4.22).

Assumption 4.4. We need the following prerequisites in addition to Assumption 4.1:

- The space $Y = H_0^1(\Omega)$ is discretized by linear finite elements (continuous on $\bar{\Omega}$) on a triangulation \mathcal{T} as in [23].
- The coefficient function $\hat{\kappa}$ is affine in ξ , i.e., $\hat{\kappa}(x, \xi) = \hat{\kappa}_0(x) + \sum_{i=1}^m \xi_i \hat{\kappa}_i(x)$, cf. Example 4.3.
- The functions $\hat{\kappa}_i$ ($i \in \{0, \dots, m\}$) and $\hat{\kappa}_{\text{ref}}$ are assumed to be piecewise constant, i.e., constant on each triangle $T \in \mathcal{T}$.
- Let $\tilde{\mathbf{y}} \in \mathbf{Y} \otimes \mathcal{P}_{d-1}(\Xi)$. We assume that the nonlinear part $\hat{\phi}(\tilde{\mathbf{y}})$ of the residual can be approximated sufficiently well by a function $\tilde{\phi} \in L^2(\Omega) \otimes \mathcal{P}_{d-1}(\Xi)$, e.g., by interpolation. Moreover, we assume that the function $\hat{\chi} \cdot \tilde{\mathbf{y}}$ can be approximated sufficiently well by a function $\tilde{\chi} \in L^2(\Omega) \otimes \mathcal{P}_{d-1}(\Xi)$.
- The right-hand side $\hat{\mathbf{b}}$ stems from a function $\hat{\mathbf{f}} \in L^2(\Omega) \otimes \mathcal{P}_{d-1}(\Xi)$, cf. (4.18).

To separate the interpolation error from the remaining error contributions, we compute:

$$(4.24) \quad \begin{aligned} \langle \mathbf{r} - \check{\mathbf{A}}_{\text{ref}} \mathbf{w}, \mathbf{v} \rangle_{\check{\mathbf{Y}}^*, \check{\mathbf{Y}}} &= \int_{\Xi} \int_{\Omega} \hat{\kappa} \nabla_x \tilde{\mathbf{y}} \cdot \nabla_x \mathbf{v} + \hat{\chi} \tilde{\mathbf{y}} \mathbf{v} + \hat{\phi}(\tilde{\mathbf{y}}) \mathbf{v} - \hat{\mathbf{f}} \mathbf{v} - \hat{\kappa}_{\text{ref}} \nabla_x \mathbf{w} \cdot \nabla_x \mathbf{v} - \hat{\chi}_{\text{ref}} \mathbf{w} \mathbf{v} \, dx \, d\mathbb{P} \\ &= \int_{\Xi} \int_{\Omega} \hat{\kappa} \nabla_x \tilde{\mathbf{y}} \cdot \nabla_x \mathbf{v} + \tilde{\chi} \mathbf{v} + \tilde{\phi} \mathbf{v} - \hat{\mathbf{f}} \mathbf{v} - \hat{\kappa}_{\text{ref}} \nabla_x \mathbf{w} \cdot \nabla_x \mathbf{v} - \hat{\chi}_{\text{ref}} \mathbf{w} \mathbf{v} \, dx \, d\mathbb{P} \\ &\quad + \int_{\Xi} \int_{\Omega} (\hat{\chi} \tilde{\mathbf{y}} - \tilde{\chi} + \hat{\phi}(\tilde{\mathbf{y}}) - \tilde{\phi}) \mathbf{v} \, dx \, d\mathbb{P}. \end{aligned}$$

In the following, we will neglect $\hat{\chi} \cdot \tilde{\mathbf{y}} - \tilde{\chi}$ and $\hat{\boldsymbol{\phi}}(\tilde{\mathbf{y}}) - \tilde{\boldsymbol{\phi}}$ for the ease of presentation and implementation. Moreover, we allow for different orthonormal polynomial bases of $L^2_{\mathbb{P}}(\Xi)$ in the derivation of the error estimates. More concretely, we use the tensor product orthonormal Hilbert bases $\{\beta_k\}_{k \in \mathbb{N}^m}$ constructed from polynomials of increasing degree and $\{\theta_k\}_{k \in \mathbb{N}^m}$, where adequately weighted Lagrange polynomials w.r.t. the Gaussian quadrature nodes appear for $k_i \leq d_i$, see [subsection 4.1](#). Moreover, we introduce alternative, still to be specified orthonormal bases $\{\vartheta_{k_i}^{(i)}\}_{k_i=1}^{\infty}$ of $L^2_{\mathbb{P}_i}(\Xi_i)$ with $\vartheta_{k_i}^{(i)} \equiv \theta_{k_i}^{(i)} \equiv \beta_{k_i}^{(i)}$ for $k_i \geq d_i + 1$ and the corresponding tensor product Hilbert basis $\{\vartheta_k\}_{k \in \mathbb{N}^m}$ of $L^2_{\mathbb{P}}(\Xi)$. Since only the first d_i basis polynomials differ, it is possible to convert between the coefficients w.r.t. the different bases by applying suitable orthonormal transformations.

For a simple, interpolation-based approximation of the nonlinearity, it is beneficial to use the Lagrange basis $\{\theta_k\}_{k \in \mathbb{N}_0^m}$, as pointed out in [subsection 4.1](#). Hence, the polynomial basis representations of $\tilde{\mathbf{y}}$ and \mathbf{w} shall be $\tilde{\mathbf{y}}(x, \xi) = \sum_{l \leq d} \tilde{y}_l(x) \theta_l(\xi)$ and $\mathbf{w}(x, \xi) = \sum_{l \leq d} w_l(x) \theta_l(\xi)$ with $\tilde{y}_l, w_l \in Y$. Be aware of the fact that these sums are actually large since l is an index vector. Analogously, we write $\hat{\mathbf{f}}, \hat{\boldsymbol{\phi}},$ and $\tilde{\chi}$ with the deterministic functions $\hat{f}_l, \hat{\phi}_l, \tilde{\chi}_l \in L^2(\Omega)$. Moreover, we will use a transformed representation $\tilde{\mathbf{y}}(x, \xi) = \sum_{l \leq d} \tilde{y}_l^*(x) \vartheta_l(\xi)$ with the alternative basis $\{\vartheta_k\}_{k \in \mathbb{N}^m}$ and $\tilde{y}_l^* \in Y$. These coefficients can be obtained from \tilde{y}_l by the mentioned orthonormal transformations.

For an arbitrary test function \mathbf{v} we use a mixed representation of the form

$$\mathbf{v}(x, \xi) = \sum_{k \in \mathbb{N}^m} v_k^*(x) \vartheta_k(\xi) = \sum_{k \leq d} v_k(x) \theta_k(\xi) + \sum_{k \not\leq d} v_k^*(x) \vartheta_k(\xi)$$

with unique coefficients $v_k, v_k^* \in Y$ for all k . The series converges in the $L^2_{\mathbb{P}}(\Xi; Y)$ sense. With arbitrary $v_k \in Y$, Galerkin orthogonality yields

$$(4.25) \quad \langle \check{\mathbf{A}}_{\text{ref}}(\mathbf{w} - \mathbf{w}), \mathbf{v} \rangle_{\tilde{\mathbf{y}}^*, \tilde{\mathbf{y}}} = \sum_{k \leq d} \langle \mathbf{r} - \check{\mathbf{A}}_{\text{ref}} \mathbf{w}, (v_k - v_k) \theta_k \rangle_{\tilde{\mathbf{y}}^*, \tilde{\mathbf{y}}} + \sum_{k \not\leq d} \langle \mathbf{r} - \check{\mathbf{A}}_{\text{ref}} \mathbf{w}, v_k^* \vartheta_k \rangle_{\tilde{\mathbf{y}}^*, \tilde{\mathbf{y}}}.$$

For $k \leq d$, we obtain from (4.24):

$$\begin{aligned} & \langle \mathbf{r} - \check{\mathbf{A}}_{\text{ref}} \mathbf{w}, v_k \theta_k \rangle_{L^2_{\mathbb{P}}(\Xi; Y^*), L^2_{\mathbb{P}}(\Xi; Y)} \\ & \approx \sum_{l \leq d} \int_{\Xi} \left(\int_{\Omega} (\hat{\mathbf{k}}_0 + \xi_1 \hat{\mathbf{k}}_1 + \dots + \xi_m \hat{\mathbf{k}}_m) \nabla \tilde{y}_l \cdot \nabla v_k + \tilde{\chi}_l v_k + \hat{\phi}_l v_k \right. \\ & \quad \left. - \hat{f}_l v_k - \hat{\mathbf{k}}_{\text{ref}} \nabla \mathbf{w}_l \cdot \nabla v_k - \hat{\chi}_{\text{ref}} \mathbf{w}_l v_k \, dx \right) \theta_l(\xi) \theta_k(\xi) \, d\mathbb{P} \\ & = \sum_{l \leq d} \left(\int_{\Omega} (\hat{\mathbf{k}}_0 \nabla \tilde{y}_l - \hat{\mathbf{k}}_{\text{ref}} \nabla \mathbf{w}_l) \cdot \nabla v_k + (\tilde{\chi}_l + \hat{\phi}_l - \hat{f}_l - \hat{\chi}_{\text{ref}} \mathbf{w}_l) v_k \, dx \right) \left(\int_{\Xi} \theta_l(\xi) \theta_k(\xi) \, d\mathbb{P} \right) \\ & \quad + \sum_{l \leq d} \sum_{i=1}^m \left(\int_{\Omega} \hat{\mathbf{k}}_i \nabla \tilde{y}_l \cdot \nabla v_k \, dx \right) \left(\int_{\Xi} \xi_i \theta_l(\xi) \theta_k(\xi) \, d\mathbb{P} \right) \\ (4.26) \quad & = \int_{\Omega} \left((\hat{\mathbf{k}}_0 + \mathbf{a}_{k_1}^{(1)} \hat{\mathbf{k}}_1 + \dots + \mathbf{a}_{k_m}^{(m)} \hat{\mathbf{k}}_m) \nabla \tilde{y}_k - \hat{\mathbf{k}}_{\text{ref}} \nabla \mathbf{w}_k \right) \cdot \nabla v_k + (\tilde{\chi}_k + \hat{\phi}_k - \hat{f}_k - \hat{\chi}_{\text{ref}} \mathbf{w}_k) v_k \, dx. \end{aligned}$$

This is due to orthogonality of the polynomials $\{\theta_k\}_{k \in \mathbb{N}^m}$ and the fact that

$$\int_{\Xi} \xi_i \theta_l(\xi) \theta_k(\xi) \, d\mathbb{P} = \int_{\Xi_i} \xi_i \theta_{l_i}^{(i)}(\xi_i) \theta_{k_i}^{(i)}(\xi_i) \, d\mathbb{P}_i \prod_{j=1, j \neq i}^m \delta_{l_j k_j} = \begin{cases} \mathbf{a}_{k_i}^{(i)} & \text{if } k = l, \\ 0 & \text{otherwise.} \end{cases}$$

Recall that $\{\mathbf{a}_{k_i}^{(i)}\}_{k_i=1}^{d_i} \subset \Xi_i$ are the Gaussian quadrature nodes w.r.t. \mathbb{P}_i . For a different polynomial basis we would obtain different values here and possibly no decoupling of deterministic terms.

For $k \not\leq d$ (yielding $k \neq l$) we get

$$(4.27) \quad \begin{aligned} & \langle \mathbf{r} - \check{\mathbf{A}}_{\text{ref}} \mathbf{w}, v_k^* \vartheta_k \rangle_{L^2_{\Xi}(\Xi; Y^*), L^2_{\Xi}(\Xi; Y)} \approx \sum_{l \leq d} \sum_{i=1}^m \left(\int_{\Omega} \hat{\mathbf{k}}_i \nabla \tilde{y}_l^* \cdot \nabla v_k^* \, dx \right) \left(\int_{\Xi} \xi_i \vartheta_l(\xi) \vartheta_k(\xi) \, d\mathbb{P} \right) \\ & = \begin{cases} \sum_{l_i=1}^{d_i} c_{l_i}^{(i)} \int_{\Omega} \hat{\mathbf{k}}_i \nabla \tilde{y}_{k_1, \dots, l_i, \dots, k_m}^* \cdot \nabla v_k^* \, dx & \text{if } \begin{cases} k_i = d_i + 1 \text{ for an } i \in [m], \\ k_j \leq d_j \text{ for } j \in [m] \setminus \{i\}, \end{cases} \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

In this computation we have used the following property of any orthonormal polynomial basis $\{\vartheta_l\}_{l \in \mathbb{N}^m}$ in tensor product form, where $\vartheta_{l_i}^{(i)} \perp \mathcal{P}_{l_i-2}(\Xi_i)$ for all $l_i \geq d_i + 1$:

$$\int_{\Xi} \xi_i \vartheta_l(\xi) \vartheta_k(\xi) \, d\mathbb{P} = \begin{cases} c_{l_i}^{(i)} & \text{if } k_i = d_i + 1 \text{ and } k_j = l_j \text{ for all } j \in [m] \setminus \{i\} \\ 0 & \text{otherwise} \end{cases}$$

with $c_{l_i}^{(i)} := \int_{\Xi_i} \xi_i \vartheta_{l_i}^{(i)}(\xi_i) \vartheta_{d_i+1}^{(i)}(\xi_i) \, d\mathbb{P}_i$ holds for any $l \leq d$ and $k \not\leq d$.

From (4.25), (4.26), and (4.27) we get

$$(4.28) \quad \begin{aligned} & \langle \check{\mathbf{A}}_{\text{ref}}(\mathbf{w} - \mathbf{w}), \mathbf{v} \rangle_{\check{\mathbf{Y}}^*, \check{\mathbf{Y}}} \\ & \approx \sum_{k \leq d} \int_{\Omega} (\hat{\mathbf{k}}(\cdot, \mathbf{a}_k) \nabla \tilde{y}_k - \hat{\mathbf{k}}_{\text{ref}} \nabla \mathbf{w}_k) \cdot \nabla (v_k - \mathbf{v}_k) + (\tilde{\chi}_k + \tilde{\varphi}_k - \hat{f}_k - \hat{\chi}_{\text{ref}} \mathbf{w}_k) (v_k - \mathbf{v}_k) \, dx \\ & \quad + \sum_{i=1}^m \sum_{k \leq d} c_{k_i}^{(i)} \int_{\Omega} \hat{\mathbf{k}}_i \nabla \tilde{y}_k^* \cdot \nabla v_{k_1, \dots, d_i+1, \dots, d_m}^* \, dx. \end{aligned}$$

Recall that we write $\mathbf{a}_k := (\mathbf{a}_{k_1}^{(1)}, \dots, \mathbf{a}_{k_m}^{(m)})^\top$.

We see that the first term at the end of (4.28) can be treated as in [23, sec. 5.2] and constitutes the deterministic part of the error. The second term can be viewed as the stochastic part of the error, where the i -th summand corresponds to the discretization error of $L^2_{\mathbb{P}_i}(\Xi_i)$.

The first term in (4.28) is estimated as in [23, Eq. (5.21)] and is bounded as in [23, Eq. (5.22)] using the Clément interpolant \mathbf{v}_k of v_k :

$$(4.29) \quad \begin{aligned} & \sum_{k \leq d} \int_{\Omega} (\hat{\mathbf{k}}(\cdot, \mathbf{a}_k) \nabla \tilde{y}_k - \hat{\mathbf{k}}_{\text{ref}} \nabla \mathbf{w}_k) \cdot \nabla (v_k - \mathbf{v}_k) + (\tilde{\chi}_k + \tilde{\varphi}_k - \hat{f}_k - \hat{\chi}_{\text{ref}} \mathbf{w}_k) (v_k - \mathbf{v}_k) \, dx \\ & \leq \sum_{T \in \mathcal{T}} \sum_{k \leq d} \|\tilde{\chi}_k + \tilde{\varphi}_k - \hat{f}_k - \hat{\chi}_{\text{ref}} \mathbf{w}_k\|_{L^2(T)} \|v_k - \mathbf{v}_k\|_{L^2(T)} \\ & \quad + \sum_{E \in \mathcal{E}^0} \sum_{k \leq d} \|[(\hat{\mathbf{k}}(\cdot, \mathbf{a}_k) \nabla \tilde{y}_k - \hat{\mathbf{k}}_{\text{ref}} \nabla \mathbf{w}_k) \cdot \mathbf{n}_E]\|_{L^2(E)} \|v_k - \mathbf{v}_k\|_{L^2(E)} \\ & \leq c_{\mathcal{T}} c_{\hat{\mathbf{A}}_{\text{ref}}} \left(\sum_{T \in \mathcal{T}} \boldsymbol{\eta}_T(\tilde{\mathbf{y}})^2 + \sum_{E \in \mathcal{E}^0} \boldsymbol{\eta}_E(\tilde{\mathbf{y}})^2 \right)^{\frac{1}{2}} \left(\sum_{k \leq d} \|v_k\|_{\hat{\mathbf{A}}_{\text{ref}}}^2 \right)^{\frac{1}{2}}, \text{ where} \end{aligned}$$

$$(4.30) \quad \boldsymbol{\eta}_T(\tilde{\mathbf{y}}) := h_T \left(\sum_{k \leq d} \|\tilde{\chi}_k + \tilde{\varphi}_k - \hat{f}_k - \hat{\chi}_{\text{ref}} \mathbf{w}_k\|_{L^2(T)}^2 \right)^{\frac{1}{2}},$$

$$(4.31) \quad \boldsymbol{\eta}_E(\tilde{\mathbf{y}}) := h_E^{1/2} \left(\sum_{k \leq d} \|[(\hat{\mathbf{k}}(\cdot, \mathbf{a}_k) \nabla \tilde{y}_k - \hat{\mathbf{k}}_{\text{ref}} \nabla \mathbf{w}_k) \cdot \mathbf{n}_E]\|_{L^2(E)}^2 \right)^{\frac{1}{2}}.$$

Here, \mathcal{E}^0 denotes the set of interior edges of the triangulation \mathcal{T} , \mathbf{n}_E is the unit normal of the edge E , and $[[\cdot]]$ stands for the jump of the function over the edge E , see [23, sec. 5.2]. The constant $c_{\mathcal{T}} > 0$ depends only on the smallest angle in the triangulation \mathcal{T} , the constant $c_{\hat{A}_{\text{ref}}} > 0$ on the coefficients $\hat{\kappa}_{\text{ref}}$ and $\hat{\chi}_{\text{ref}}$. The diameters of the elements T and the edges E are denoted by h_T and h_E , respectively. Note that $\tilde{\mathbf{y}}$ enters into the definitions (4.30), (4.31) via $\tilde{\mathbf{y}}_k$, \mathbf{w}_k , $\tilde{\chi}_k$, and $\tilde{\varphi}_k$.

Defining $\langle \hat{A}_{\text{ref}} \mathbf{v}, \mathbf{v} \rangle_{Y^*, Y} := \int_{\Omega} \hat{\kappa}_{\text{ref}} \nabla \mathbf{y} \cdot \nabla \mathbf{v} \, dx$ as well as $\langle \hat{A}_i \mathbf{y}, \mathbf{v} \rangle_{Y^*, Y} := \int_{\Omega} \hat{\kappa}_i \nabla \mathbf{y} \cdot \nabla \mathbf{v} \, dx$ for $i \in \{0, \dots, m\}$, the second term in (4.28) can be treated as follows:

$$\begin{aligned}
(4.32) \quad & \sum_{i=1}^m \sum_{k \leq d} c_{k_i}^{(i)} \int_{\Omega} \hat{\kappa}_i \nabla \tilde{\mathbf{y}}_k^* \cdot \nabla \mathbf{v}_{k_1, \dots, d_i+1, \dots, k_m}^* \, dx = \sum_{i=1}^m \sum_{k \leq d} c_{k_i}^{(i)} \langle \hat{A}_i \tilde{\mathbf{y}}_k^*, \mathbf{v}_{k_1, \dots, d_i+1, \dots, k_m}^* \rangle_{Y^*, Y} \\
& \leq \sum_{i=1}^m \sum_{k_j \leq d_j, j \neq i} \left\| \sum_{k_i \leq d_i} (c_{k_i}^{(i)} \cdot \hat{A}_i \tilde{\mathbf{y}}_k^*) \right\|_{\hat{A}_{\text{ref}}^{-1}} \|\mathbf{v}_{k_1, \dots, d_i+1, \dots, k_m}^*\|_{\hat{A}_{\text{ref}}} \\
& \leq \left(\sum_{i=1}^m \sum_{k_j \leq d_j, j \neq i} \left\| \hat{A}_i \sum_{k_i \leq d_i} (c_{k_i}^{(i)} \tilde{\mathbf{y}}_k^*) \right\|_{\hat{A}_{\text{ref}}^{-1}}^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \sum_{k_j \leq d_j, j \neq i} \|\mathbf{v}_{k_1, \dots, d_i+1, \dots, k_m}^*\|_{\hat{A}_{\text{ref}}}^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

The term

$$\tilde{\mathbf{y}}_{k,i}^* := \sum_{k_i \leq d_i} (c_{k_i}^{(i)} \tilde{\mathbf{y}}_k^*) = \int_{\Xi_i} \xi_i \beta_{d_i+1}^{(i)}(\xi_i) \sum_{k_i \leq d_i} \vartheta_{k_i}^{(i)}(\xi_i) \tilde{\mathbf{y}}_k^* \, d\mathbb{P}_i$$

is independent of the concrete choice of the basis $\{\vartheta_{k_i}^{(i)}\}_{k_i \in \mathbb{N}}$.

To estimate $\|\hat{A}_i \tilde{\mathbf{y}}_{k,i}^*\|_{\hat{A}_{\text{ref}}^{-1}}$, we define $\hat{\mathbf{y}}_{k,i}^* := \hat{A}_{\text{ref}}^{-1} \hat{A}_i \tilde{\mathbf{y}}_{k,i}^*$, i.e.,

$$\int_{\Omega} \hat{\kappa}_{\text{ref}} \nabla \hat{\mathbf{y}}_{k,i}^* \cdot \nabla \mathbf{v} + \hat{\chi}_{\text{ref}} \hat{\mathbf{y}}_{k,i}^* \mathbf{v} \, dx = \int_{\Omega} \hat{\kappa}_i \nabla \tilde{\mathbf{y}}_{k,i}^* \cdot \nabla \mathbf{v} \, dx$$

holds for all $\mathbf{v} \in Y$. Since we want to avoid the cost of computing a discrete version $\hat{\mathbf{y}}_{k,i}^*$ of $\tilde{\mathbf{y}}_{k,i}^*$, we estimate

$$\begin{aligned}
\|\hat{A}_i \tilde{\mathbf{y}}_{k,i}^*\|_{\hat{A}_{\text{ref}}^{-1}}^2 &= \langle \hat{A}_{\text{ref}}^{-1} \hat{A}_i \tilde{\mathbf{y}}_{k,i}^*, \hat{A}_i \tilde{\mathbf{y}}_{k,i}^* \rangle_{Y, Y^*} = \langle \hat{\mathbf{y}}_{k,i}^*, \hat{A}_i \tilde{\mathbf{y}}_{k,i}^* \rangle_{Y, Y^*} = \int_{\Omega} \hat{\kappa}_i \nabla \tilde{\mathbf{y}}_{k,i}^* \cdot \nabla \hat{\mathbf{y}}_{k,i}^* \, dx \\
&\leq \left(\int_{\Omega} |\hat{\kappa}_i| \nabla \tilde{\mathbf{y}}_{k,i}^* \cdot \nabla \tilde{\mathbf{y}}_{k,i}^* \, dx \right)^{\frac{1}{2}} \left(\int_{\Omega} |\hat{\kappa}_i| \nabla \hat{\mathbf{y}}_{k,i}^* \cdot \nabla \hat{\mathbf{y}}_{k,i}^* \, dx \right)^{\frac{1}{2}} \\
&\leq \left(\int_{\Omega} |\hat{\kappa}_i| \nabla \tilde{\mathbf{y}}_{k,i}^* \cdot \nabla \tilde{\mathbf{y}}_{k,i}^* \, dx \right)^{\frac{1}{2}} \|\frac{\hat{\kappa}_i}{\hat{\kappa}_{\text{ref}}}\|_{L^\infty(\Omega)}^{\frac{1}{2}} \left(\int_{\Omega} \hat{\kappa}_{\text{ref}} \nabla \hat{\mathbf{y}}_{k,i}^* \cdot \nabla \hat{\mathbf{y}}_{k,i}^* + \chi_{\text{ref}} (\hat{\mathbf{y}}_{k,i}^*)^2 \, dx \right)^{\frac{1}{2}} \\
&= \left(\int_{\Omega} |\hat{\kappa}_i| \nabla \tilde{\mathbf{y}}_{k,i}^* \cdot \nabla \tilde{\mathbf{y}}_{k,i}^* \, dx \right)^{\frac{1}{2}} \|\frac{\hat{\kappa}_i}{\hat{\kappa}_{\text{ref}}}\|_{L^\infty(\Omega)}^{\frac{1}{2}} \left(\int_{\Omega} \hat{\kappa}_i \nabla \tilde{\mathbf{y}}_{k,i}^* \cdot \nabla \hat{\mathbf{y}}_{k,i}^* \, dx \right)^{\frac{1}{2}}.
\end{aligned}$$

It follows that

$$(4.33) \quad \|\hat{A}_i \tilde{\mathbf{y}}_{k,i}^*\|_{\hat{A}_{\text{ref}}^{-1}}^2 = \int_{\Omega} \hat{\kappa}_i \nabla \tilde{\mathbf{y}}_{k,i}^* \cdot \nabla \hat{\mathbf{y}}_{k,i}^* \, dx \leq \|\frac{\hat{\kappa}_i}{\hat{\kappa}_{\text{ref}}}\|_{L^\infty(\Omega)} \left(\int_{\Omega} |\hat{\kappa}_i| \nabla \tilde{\mathbf{y}}_{k,i}^* \cdot \nabla \tilde{\mathbf{y}}_{k,i}^* \, dx \right).$$

Thus, we define

$$(4.34) \quad \boldsymbol{\zeta}_i(\tilde{\mathbf{y}}) := \|\frac{\hat{\kappa}_i}{\hat{\kappa}_{\text{ref}}}\|_{L^\infty(\Omega)}^{1/2} \left(\sum_{k_j \leq d_j, j \neq i} \left(\int_{\Omega} |\hat{\kappa}_i| \left(\sum_{k_i \leq d_i} c_{k_i}^{(i)} \nabla \tilde{\mathbf{y}}_k^* \right) \cdot \left(\sum_{k_i \leq d_i} c_{k_i}^{(i)} \nabla \tilde{\mathbf{y}}_k^* \right) \, dx \right) \right)^{\frac{1}{2}}$$

in order to formulate the following error estimation theorem.

THEOREM 4.5. *Let $\tilde{\mathbf{y}} \in Y \otimes \mathcal{P}_d(\Xi)$ be given such that $\mathbf{r} = \hat{\mathbf{A}}\tilde{\mathbf{y}} + \hat{\mathbf{N}}(\tilde{\mathbf{y}}) - \hat{\mathbf{b}} \in L^2_{\mathbb{P}}(\Xi; Y^*)$. Suppose that [Assumptions 4.1](#) and [4.4](#) are satisfied. Then,*

$$(4.35) \quad \begin{aligned} \|\tilde{\mathbf{y}} - \mathbf{y}\|_{\hat{\mathbf{A}}_{\text{ref}}}^2 &\leq \Lambda^2 \left(c_{\mathcal{T}} c_{\hat{\mathbf{A}}_{\text{ref}}} \left(\sum_{T \in \mathcal{T}} \boldsymbol{\eta}_T(\tilde{\mathbf{y}})^2 + \sum_{E \in \mathcal{E}^0} \boldsymbol{\eta}_E(\tilde{\mathbf{y}})^2 \right)^{\frac{1}{2}} + \left(\sum_{i=1}^m \boldsymbol{\zeta}_i(\tilde{\mathbf{y}})^2 \right)^{\frac{1}{2}} \right. \\ &\quad \left. + \left\| (\hat{\chi} \cdot \tilde{\mathbf{y}} - \tilde{\boldsymbol{\chi}} + \hat{\boldsymbol{\phi}}(\tilde{\mathbf{y}}) - \tilde{\boldsymbol{\phi}}, \cdot) \right\|_{L^2_{\mathbb{P}}(\Xi; L^2(\Omega))} \right\|_{\hat{\mathbf{A}}_{\text{ref}}^{-1}} \right)^2 + \Lambda^2 \|\mathbf{w}\|_{\hat{\mathbf{A}}_{\text{ref}}}^2, \end{aligned}$$

where \mathbf{w} is defined by [\(4.22\)](#), $\boldsymbol{\eta}_T$, $\boldsymbol{\eta}_E$ are defined by [\(4.30\)](#), [\(4.31\)](#), and $\boldsymbol{\zeta}_i$ is defined in [\(4.34\)](#).

Proof. For an arbitrary $\mathbf{v} \in \check{Y}$, $\mathbf{v}(x, \xi) = \sum_{k \in \mathbb{N}^m} v_k(x) \boldsymbol{\theta}_k(\xi) = \sum_{k \in \mathbb{N}^m} v_k^*(x) \boldsymbol{\vartheta}_k(\xi)$ with $v_k, v_k^* \in Y$ for all k , we have that $\|\mathbf{v}\|_{\hat{\mathbf{A}}_{\text{ref}}}^2 = \sum_{k \in \mathbb{N}^m} \|v_k\|_{\hat{\mathbf{A}}_{\text{ref}}}^2 = \sum_{k \in \mathbb{N}^m} \|v_k^*\|_{\hat{\mathbf{A}}_{\text{ref}}}^2$ because the polynomials $\{\boldsymbol{\theta}_k\}_k$, $\{\boldsymbol{\vartheta}_k\}_k$ are orthonormal. Combining [\(4.24\)](#), [\(4.28\)](#), [\(4.29\)](#), [\(4.32\)](#), [\(4.33\)](#), and the fact that

$$\langle \hat{\mathbf{A}}_{\text{ref}}(\mathbf{w} - \mathbf{w}), \mathbf{v} \rangle_{\check{Y}^*, \check{Y}} = (\mathbf{w} - \mathbf{w}, \mathbf{v})_{\hat{\mathbf{A}}_{\text{ref}}} \leq c \boldsymbol{\eta}(\mathbf{w}) \|\mathbf{v}\|_{\hat{\mathbf{A}}_{\text{ref}}}$$
 for all $\mathbf{v} \in \check{Y}$

yields $\|\mathbf{w} - \mathbf{w}\|_{\hat{\mathbf{A}}_{\text{ref}}} \leq c \boldsymbol{\eta}(\mathbf{w})$, gives

$$(4.36) \quad \begin{aligned} \|\mathbf{w} - \mathbf{w}\|_{\hat{\mathbf{A}}_{\text{ref}}} &\leq c_{\mathcal{T}} c_{\hat{\mathbf{A}}_{\text{ref}}} \left(\sum_{T \in \mathcal{T}} \boldsymbol{\eta}_T(\tilde{\mathbf{y}})^2 + \sum_{E \in \mathcal{E}^0} \boldsymbol{\eta}_E(\tilde{\mathbf{y}})^2 \right)^{\frac{1}{2}} + \left(\sum_{i=1}^m \boldsymbol{\zeta}_i(\tilde{\mathbf{y}})^2 \right)^{\frac{1}{2}} \\ &\quad + \left\| (\hat{\chi} \cdot \tilde{\mathbf{y}} - \tilde{\boldsymbol{\chi}} + \hat{\boldsymbol{\phi}}(\tilde{\mathbf{y}}) - \tilde{\boldsymbol{\phi}}, \cdot) \right\|_{L^2_{\mathbb{P}}(\Xi; L^2(\Omega))} \right\|_{\hat{\mathbf{A}}_{\text{ref}}^{-1}}, \end{aligned}$$

which yields the result together with [\(4.23\)](#). \square

[Theorem 4.5](#) provides a split of the error into different contributions. The last term in [\(4.35\)](#) is exactly the algebraic error due to the inexact solution of the discrete system by, e.g., a low-rank tensor solver. The first term consists of

- a term related to the FE discretization error, which can itself be split into error contributions for each element $T \in \mathcal{T}$,
- the error contribution due to the discretization of the stochastic space, which consists of terms for each stochastic parameter ξ_i , $i \in \{1, \dots, m\}$, and
- the interpolation error coming from the approximation of the nonlinear terms in the equation.

Based on these error contributions, the respective PDEs are solved adaptively by either running more iterations of the low-rank tensor solver if the algebraic error dominates, refining the triangles contributing a certain amount of the FE discretization error if this one dominates, or choosing a higher polynomial degree for the parameter with the largest contribution if the stochastic error is large.

The error estimate given in [Theorem 4.5](#) uses a deterministic reference operator as the one discussed in [\[7, 8\]](#). The results in these papers rely on a saturation assumption, which we do not make here, but derive an error estimate similar to the one presented in [\[16\]](#) and used in [\[17, 18\]](#). All mentioned papers present error estimates for linear equations whereas we consider a class of semilinear equations with a monotone nonlinearity, but have used a linear reference equation to derive the error estimate. The convergence of the adaptive solution technique should be investigated. Since it is already discussed in the linear case for a very similar estimator in [\[17\]](#), we skip this topic at this point.

Realization with low-rank tensors. In order to solve [\(4.19\)](#) adaptively based on the estimate [\(4.35\)](#) and using low-rank tensors, we apply Newton's method to the discretized

system $\hat{\mathbf{A}}\mathbf{y} + \hat{\mathbf{N}}(\mathbf{y}) - \hat{\mathbf{b}} = 0$. Each Newton step is computed with the iterative low-rank tensor solver AMEn [14]. We have extended it such that it can also handle componentwise multiplication operators [22] to be able to apply the operator $\mathbf{N}'(\mathbf{y})$. By default, AMEn and other low-rank tensor solvers such as ALS [31] aim for minimizing the squared Frobenius norm $\|\hat{\mathbf{A}}\mathbf{y} - \hat{\mathbf{b}}\|_{\mathbb{F}}^2$ of the residual, setting $\hat{\mathbf{N}} \equiv 0$ for simplicity. Since we are indeed interested in minimizing the norm induced by $\check{\mathbf{A}}_{\text{ref}}^{-1}$, we have extended AMEn such that it solves the preconditioned, symmetric system $\check{\mathbf{R}}_{\text{ref}}^{-*} \check{\mathbf{A}} \check{\mathbf{R}}_{\text{ref}}^{-1} \hat{\mathbf{y}} = \check{\mathbf{R}}_{\text{ref}}^{-*} \hat{\mathbf{b}}$, where the rank-1-operator $\check{\mathbf{A}}_{\text{ref}}$ is decomposed as $\check{\mathbf{A}}_{\text{ref}} = \hat{\mathbf{A}}_{\text{ref}} \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1} = (\hat{\mathbf{R}}_{\text{ref}}^{\top} \hat{\mathbf{R}}_{\text{ref}}) \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1} = \check{\mathbf{R}}_{\text{ref}}^* \check{\mathbf{R}}_{\text{ref}}$ by a sparse Cholesky decomposition, cf. Example 4.3 and [22]. The solution $\hat{\mathbf{y}}$ is then transformed to the solution $\mathbf{y} = \check{\mathbf{R}}_{\text{ref}}^{-1} \hat{\mathbf{y}}$ of the original system $\hat{\mathbf{A}}\mathbf{y} = \hat{\mathbf{b}}$.

In addition, we have to implement the evaluation of the error indicators $\boldsymbol{\eta}_T(\tilde{\mathbf{y}})$, $\boldsymbol{\eta}_E(\tilde{\mathbf{y}})$, and $\boldsymbol{\zeta}_i(\tilde{\mathbf{y}})$ from the low-rank solution in an efficient manner. If, e.g., $\tilde{\mathbf{y}}(x, \xi) = \sum_{k \leq d} \tilde{y}_k(x) \boldsymbol{\theta}_k(\xi)$ is represented by the tensor $\tilde{\mathbf{y}}$ analogously to (4.2), we obtain that $\tilde{y}_k(x) = \sum_{k_0=1}^{d_0} \tilde{\mathbf{y}}(k_0, k) \phi_{k_0}(x)$ holds, i.e., the function \tilde{y}_k is represented by the vector $\tilde{\mathbf{y}}(\cdot, k)$.

Knowing that, we establish the evaluation of the triangle error contribution (4.30). We approximate each $\tilde{\chi}_k + \tilde{\varphi}_k - f_k - \hat{\chi}_{\text{ref}} w_k \approx \tilde{f}_k$ by interpolation, where \tilde{f}_k are linear finite element functions for all k . These functions are represented altogether by a single tensor $\tilde{\mathbf{f}} \in \mathbb{R}^{\tilde{d} \times d_1 \times \cdots \times d_m}$ of values at the FE nodes. Note that $\tilde{d} > d_0$ is the number of all FE nodes whereas d_0 counts only the interior nodes.

Let now $x^1, x^2, x^3 \in \Omega$ be the three vertices of the triangle T and let $k_0^{(1)}, k_0^{(2)}, k_0^{(3)} \in [\tilde{d}]$ be the indices corresponding to the vertices x^1, x^2, x^3 , respectively, so that $\tilde{\mathbf{f}}(k_0^{(j)}, k) = \tilde{f}_k(x^j)$ for all j and all k . Then, we obtain

$$\begin{aligned} \boldsymbol{\eta}_T(\tilde{\mathbf{y}})^2 &\approx h_T^2 \sum_{k \leq d} \|\tilde{f}_k\|_{L^2(T)}^2 = h_T^2 \sum_{k \leq d} \left(\frac{a_T}{6} \sum_{j=1}^3 \sum_{l=j}^3 \tilde{f}_k(x^j) \tilde{f}_k(x^l) \right) \\ &= h_T^2 \frac{a_T}{6} \sum_{k \leq d} \left(\sum_{j=1}^3 \sum_{l=j}^3 \tilde{\mathbf{f}}(k_0^{(j)}, k) \tilde{\mathbf{f}}(k_0^{(l)}, k) \right) = h_T^2 \frac{a_T}{6} \sum_{j=1}^3 \sum_{l=j}^3 \langle \mathbb{1}, \tilde{\mathbf{f}}(k_0^{(j)}, \cdot) \odot \tilde{\mathbf{f}}(k_0^{(l)}, \cdot) \rangle. \end{aligned}$$

where a_T is the area of T . The advantage of this formulation is that it can be vectorized using the tensors of evaluations at all first, second, and third triangle vertices, respectively, to evaluate the error indicator for all triangles T simultaneously. Then, only componentwise multiplication and contraction with the rank-1-tensor $\mathbb{1}$ are needed.

To compute the edge error contribution (4.31), one first computes the values of the partial derivatives $\partial_{x_1} \tilde{y}_k$ and $\partial_{x_2} \tilde{y}_k$ on the elements. These values are contained in the tensors $\mathbf{G}_1 \circ_1 \tilde{\mathbf{y}}$ and $\mathbf{G}_2 \circ_1 \tilde{\mathbf{y}}$, where $\mathbf{G}_1, \mathbf{G}_2 \in \mathbb{R}^{|\mathcal{T}| \times d_0}$ are sparse matrices mapping a vector of function values on the FE nodes to a vector of derivative values on the triangles numbered from 1 to $|\mathcal{T}|$. Analogously, we obtain the tensors $\mathbf{G}_1 \circ_1 \mathbf{w}$ and $\mathbf{G}_2 \circ_1 \mathbf{w}$ containing the values of $\partial_{x_1} w_k$ and $\partial_{x_2} w_k$. Furthermore, we create the tensors $\hat{\mathbf{k}}_{\text{ref}} = \hat{\mathbf{k}}_{\text{ref}} \otimes \mathbb{1} \in \mathbb{R}^{|\mathcal{T}| \times d_1 \times \cdots \times d_m}$, where $\hat{\mathbf{k}}_{\text{ref}} \in \mathbb{R}^{|\mathcal{T}|}$ contains the values of $\hat{\mathbf{k}}_{\text{ref}}$ on the triangles, and $\hat{\mathbf{k}} \in \mathbb{R}^{|\mathcal{T}| \times d_1 \times \cdots \times d_m}$ such that $\hat{\mathbf{k}}(\cdot, k)$ contains the respective values of $\hat{\mathbf{k}}(\cdot, a_k)$. Let now $\boldsymbol{\nu}^{j,1} \in \mathbb{R}^{|\mathcal{T}|}$ contain the first components of the outwards-pointing, normal vectors corresponding to the j -th edge of the triangles T , and let $\boldsymbol{\nu}^{j,2} \in \mathbb{R}^{|\mathcal{T}|}$ contain the second components for $j \in \{1, 2, 3\}$. This means that if T is the ℓ -th triangle, the outer normal vector $\mathbf{n}_{T,j}$ corresponding to its j -th edge is given by $\mathbf{n}_{T,j} = (\boldsymbol{\nu}_\ell^{j,1} \quad \boldsymbol{\nu}_\ell^{j,2})^\top$. Then, the values of $\hat{\mathbf{k}}(\cdot, a_k) \nabla \tilde{y}_k \cdot \mathbf{n}_{T,j}$ for all T are given by the single tensor

$$\hat{\mathbf{k}} \odot (\text{diag}(\boldsymbol{\nu}^{j,1}) \circ_1 \mathbf{G}_1 \circ_1 \tilde{\mathbf{y}} + \text{diag}(\boldsymbol{\nu}^{j,2}) \circ_1 \mathbf{G}_2 \circ_1 \tilde{\mathbf{y}}) = \hat{\mathbf{k}} \odot ((\text{diag}(\boldsymbol{\nu}^{j,1}) \mathbf{G}_1 + \text{diag}(\boldsymbol{\nu}^{j,2}) \mathbf{G}_2) \circ_1 \tilde{\mathbf{y}}).$$

With an analogous consideration we get that the values of $(\hat{\kappa}(\cdot, \mathbf{a}_k) \nabla \tilde{\mathbf{y}}_k - \hat{\kappa}_{\text{ref}} \nabla \mathbf{w}_k) \cdot \mathbf{n}_{T,j}$ are contained in the tensor

$$\tilde{\mathbf{g}}_j := \hat{\kappa} \odot ((\text{diag}(\nu^{j,1}) \mathbf{G}_1 + \text{diag}(\nu^{j,2}) \mathbf{G}_2) \circ_1 \tilde{\mathbf{y}}) - \hat{\kappa}_{\text{ref}} \odot ((\text{diag}(\nu^{j,1}) \mathbf{G}_1 + \text{diag}(\nu^{j,2}) \mathbf{G}_2) \circ_1 \mathbf{w}),$$

which can be computed using standard low-rank tensor arithmetics, namely i -mode matrix products as well as componentwise multiplication and subtraction. In a next step, we number the interior edges $E \in \mathcal{E}_0$ from 1 to $|\mathcal{E}_0|$ and create sparse matrices $\mathbf{H}_j \in \{0, 1\}^{|\mathcal{E}_0| \times |\mathcal{T}|}$ ($j \in \{1, 2, 3\}$) mapping vectors of values on the triangles to vectors of values on the respective j -th triangle edges in the correct order, provided they are interior edges. Then the tensor of jumps over the interior edges is given by $\tilde{\mathbf{h}} := \sum_{j=1}^3 \mathbf{H}_j \circ_1 \tilde{\mathbf{g}}_j \in \mathbb{R}^{|\mathcal{E}_0| \times d_1 \times \dots \times d_m}$. It is correct to sum here because the outer unit normals of two neighboring triangles point in the opposite direction. Using the presented definitions, the edge error indicator is given by

$$\boldsymbol{\eta}_E(\tilde{\mathbf{y}})^2 = h_E \sum_{k \leq d} \|\llbracket (\hat{\kappa}(\cdot, \mathbf{a}_k) \nabla \tilde{\mathbf{y}}_k - \hat{\kappa}_{\text{ref}} \nabla \mathbf{w}_k) \cdot \mathbf{n}_E \rrbracket_E\|_{L^2(E)}^2 = h_E^2 \langle \mathbb{1}, \tilde{\mathbf{h}}(l, \cdot)^2 \rangle,$$

where $l \in \{1, \dots, |\mathcal{E}_0|\}$ is the number of the edge E . As before, this procedure can be vectorized to compute the error contributions of all interior edges simultaneously. Finally, we assign half of the error to each of the two neighboring triangles. Based on that, we mark all triangles with the largest error contributions which constitute a certain amount $\vartheta_{\boldsymbol{\eta}} \in (0, 1)$ of the total error, see [16, sec. 7.1], a so-called Dörfler strategy [15]. These triangles are refined regularly, i.e., divided into four triangles of the same shape. To avoid hanging nodes, additional triangles have to be divided into two new ones possibly.

To evaluate $\zeta_i(\tilde{\mathbf{y}})$ (see (4.34)) based on the tensor $\tilde{\mathbf{y}}$ representing the function $\tilde{\mathbf{y}}$, one first computes the values $c_{l_i}^{(i)} = \int_{\Xi_i} \xi_i \vartheta_{l_i}^{(i)}(\xi_i) \vartheta_{d_{i+1}}^{(i)}(\xi_i) d\mathbb{P}_i$ by a quadrature formula of high enough order, e.g., by Gaussian quadrature with $d_i + 1$ nodes if $\vartheta_{l_i}^{(i)}$ has degree $d_i - 1$ and $\vartheta_{d_{i+1}}^{(i)}$ has degree d_i , or analytically and writes them as one vector $c^{(i)} \in \mathbb{R}^{d_i}$. Furthermore, the matrix $\bar{\mathbf{A}}_i \in \mathbb{R}^{d_0 \times d_0}$ given by $(\bar{\mathbf{A}}_i)_{k_0 l_0} := (|\hat{\kappa}_i| \nabla \phi_{l_0}, \nabla \phi_{k_0})_{L^2(\Omega)^n}$, cf. Example 4.3, is assembled. Then we have

$$\zeta_i(\tilde{\mathbf{y}})^2 = \|\frac{\hat{\kappa}_i}{\hat{\kappa}_{\text{ref}}}\|_{L^\infty(\Omega)} \langle c^{(i)\top} \circ_{i+1} \tilde{\mathbf{y}}, \bar{\mathbf{A}}_i \circ_1 c^{(i)} \circ_{i+1} \tilde{\mathbf{y}} \rangle.$$

If this is the largest contribution to the total error (4.35), we increase the respective polynomial degree $d_i - 1$ by 1.

For the interpolation to finer FE or polynomial spaces, we create the respective matrices mapping the coefficients of the coarser subspaces to the new ones and apply them to the respective mode of the tensor.

5. Implementation and numerical results. To test the approach numerically, we consider the same setup as in [23], i.e., a distributed control problem on the L-shaped domain Ω with $\mathbf{f} \equiv 0$ and the desired state $\hat{\mathbf{q}} \equiv 1$. The control and the observation space are $U = H = L^2(\Omega)$ and we choose $Q(\cdot) \equiv \iota : H_0^1(\Omega) \hookrightarrow L^2(\Omega)$, $D \equiv I : L^2(\Omega) \rightarrow L^2(\Omega)$, $U_{\text{ad}} = \{u \in L^2(\Omega) : u(x) \leq 14 \text{ for a.e. } x \in \Omega\}$, and $\gamma = 10^{-3}$. The parameters are distributed uniformly on $\Xi_i = (-1, 1)$. The uncertain coefficient function is defined by $\kappa(x, \xi) = \kappa_0(x) + \sum_{i=1}^m \xi_i \kappa_i(x)$, $m = 6$, where $\kappa_0 \equiv 1$ and $\kappa_i(x) = \sigma_i 1_{\Omega_i}(x)$ are weighted indicator functions of six triangles Ω_i ($i \in \{1, \dots, m\}$) of the same area covering the domain Ω in clockwise order. This means, that the weight $\sigma_i \in [0, 1)$ determines the impact of the parameter ξ_i on the coefficient function. Since the nonlinearity $\varphi(t) = t^3$ and its derivatives are monomial, we can evaluate them using componentwise multiplication of low-rank tensors. Due to the mentioned choices Assumptions 2.1 and 2.2 are satisfied with $p = 4$, $r_Q = r_{\hat{q}} = r_f = \infty$, and $\underline{\kappa} = 1 - \max_{i \in \{1, \dots, m\}} \sigma_i$, $\bar{\kappa} = 1 + \max_{i \in \{1, \dots, m\}} \sigma_i$, $d''_{\varphi} = 0$, and $c''_{\varphi} = 6$. This yields that the

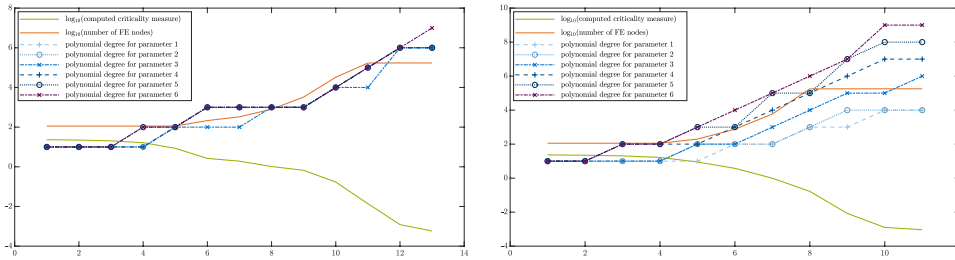


FIG. 1. Convergence and refinement plots for $\sigma = (0.25, 0.25, 0.25, 0.25, 0.25, 0.25)^\top$ (top) and $\sigma = (0.05, 0.1, 0.2, 0.3, 0.4, 0.45)^\top$ (bottom).

state and the adjoint state both belong to $L_{\mathbb{P}}^\infty(\Xi; Y)$, see [Propositions 2.6](#) and [2.10](#). Hence, it is enough to estimate the $L_{\mathbb{P}}^2(\Xi; Y)$ -error in the state and adjoint state according to [section 3](#). To bound $\|\mathbf{y}\|_{L_{\mathbb{P}}^\infty(\Xi; Y)}$, we use the a priori estimate [\(2.11\)](#).

The a posteriori error estimator from [subsection 4.2](#) is applied for the adaptive solution of the state and adjoint equation. The reference coefficients are $\hat{\kappa}_{\text{ref}} \equiv 1$ and $\hat{\chi}_{\text{ref}} \equiv 0$. Hence, we have $\lambda = \underline{\kappa}$ and $\Lambda = \bar{\kappa}$ in [\(4.20\)](#), [\(4.21\)](#). In the error estimate [\(4.35\)](#) we neglect the interpolation error term for simplicity. In contrast to [\[23\]](#) we now have to care about the constants balancing the different error contributions and choose $c_{\mathcal{T}} c_{\hat{A}_{\text{ref}}} = 10^{-3}$. The algorithmic parameters are chosen as in [\[23\]](#) with the only differences that $\tilde{c}_{\text{c}} = 0.05$, $\tilde{c}_{\text{g}} = 0.05$, and $\tilde{c}_{\text{o}} = 10^3$, which reflects the different error estimator, and we stop the algorithm if $\chi_k(0) < 10^{-3}$.

Results. We consider two different setups $\sigma = (0.25, 0.25, 0.25, 0.25, 0.25, 0.25)^\top$ and $\sigma = (0.05, 0.1, 0.2, 0.3, 0.4, 0.45)^\top$, i.e., in the first setup all parameters influence the coefficient function κ in the same manner whereas their influence is increasing in the second setup. [Figure 1](#) shows the convergence and adaptivity over the iterations. In particular, the used polynomial degrees are approximately equal in the first setup, but increasing in the second, i.e., larger for the parameters with higher influence on the coefficient function. The total computing time for the first setup is about 42 hours. Due to expensive high-fidelity PDE solves, about 37 hours are spent for the last two iterations, during which the computed criticality measure is decreased from $1.39 \cdot 10^{-2}$ to $5.82 \cdot 10^{-4}$. The obtained controls and grids are similar to the ones from [\[23\]](#), but the size of the active sets change significantly and the obtained meshes reflect the boundaries of these sets. The differences between the deterministic and the robustified controls are depicted in [Figure 2](#). Here, the subdomains Ω_i can be recognized together with the clockwise increasing uncertainty yielding an increased difference in the second setup.

REFERENCES

- [1] A. ALEXANDERIAN, N. PETRA, G. STADLER, AND O. GHATTAS, *Mean-Variance Risk-Averse Optimal Control of Systems Governed by PDEs with Random Parameter Fields Using Quadratic Approximations*, SIAM/ASA J. Uncertain. Quantif., 5 (2017), pp. 1166–1192, <https://doi.org/10.1137/16M106306X>.
- [2] A. ALI, E. ULLMANN, AND M. HINZE, *Multilevel Monte Carlo Analysis for Optimal Control of Elliptic PDEs with Random Coefficients*, SIAM/ASA J. Uncertain. Quantif., 5 (2017), pp. 466–492, <https://doi.org/10.1137/16M109870X>.
- [3] J. APPELL AND P. ZABREĀKO, *Nonlinear Superposition Operators*, vol. 95 of Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge, 1990, <https://doi.org/10.1017/CBO9780511897450>.
- [4] M. BACHMAYR, A. COHEN, AND W. DAHMEN, *Parametric PDEs: sparse or low-rank approximations?*, IMA J. Numer. Anal., (2017), <https://doi.org/10.1093/imanum/drx052>.
- [5] P. BENNER, S. DOLGOV, A. ONWUNTA, AND M. STOLL, *Low-rank solvers for unsteady Stokes–Brinkman optimal control problem with random data*, Comput. Methods Appl. Mech. Engrg., 304 (2016), pp. 26–54,

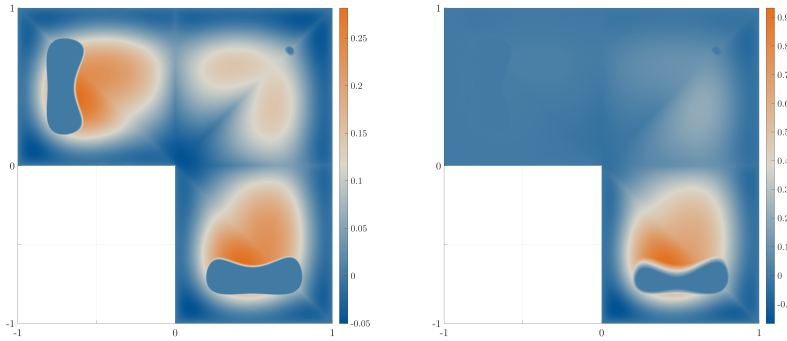


FIG. 2. Difference between the deterministic and the robustified control for $\sigma = (0.25, 0.25, 0.25, 0.25, 0.25, 0.25)^\top$ (left) and $\sigma = (0.05, 0.1, 0.2, 0.3, 0.4, 0.45)^\top$ (right).

- <https://doi.org/10.1016/j.cma.2016.02.004>.
- [6] P. BENNER, A. ONWUNTA, AND M. STOLL, *Block-Diagonal Preconditioning for Optimal Control Problems Constrained by PDEs with Uncertain Inputs*, SIAM J. Matrix Anal. Appl., 37 (2013), pp. 491–518, <https://doi.org/10.1137/15M1018502>.
- [7] A. BESPALOV, C. E. POWELL, AND D. SILVESTER, *Energy norm a posteriori error estimation for parametric operator equations*, SIAM J. Sci. Comput., 36 (2014), pp. A339 – A363, <https://doi.org/10.1137/130916849>.
- [8] A. BESPALOV AND D. SILVESTER, *Efficient adaptive stochastic Galerkin methods for parametric operator equations*, SIAM J. Sci. Comput., 38 (2016), pp. A2118 – A2140, <https://doi.org/10.1137/15M1027048>.
- [9] A. BORZÌ, V. SCHULZ, C. SCHILLINGS, AND G. VON WINCKEL, *On the treatment of distributed uncertainties in PDE-constrained optimization*, GAMM-Mitt., 33 (2010), pp. 230–246, <https://doi.org/10.1002/gamm.201010017>.
- [10] A. BORZÌ AND G. VON WINCKEL, *A POD framework to determine robust controls in PDE optimization*, Comput. Vis. Sci., 14 (2011), pp. 91–103, <https://doi.org/10.1007/s00791-011-0165-5>.
- [11] P. CHEN AND A. QUARTERONI, *Weighted Reduced Basis Method for Stochastic Optimal Control Problems with Elliptic PDE Constraint*, SIAM/ASA Journal on Uncertainty Quantification, 2 (2014), pp. 364–396, <https://doi.org/10.1137/130940517>.
- [12] P. CHEN, A. QUARTERONI, AND G. ROZZA, *A Weighted Reduced Basis Method for Elliptic Partial Differential Equations with Random Input Data*, SIAM J. Numer. Anal., 51 (2013), pp. 3163–3185, <https://doi.org/10.1137/130905253>.
- [13] A. DEFANT AND K. FLORET, *Tensor Norms and Operator Ideals*, Elsevier Science Publishers, Amsterdam, 1993.
- [14] S. V. DOLGOV AND D. V. SAVOSTYANOV, *Alternating Minimal Energy Methods for linear systems in higher dimensions*, SIAM J. Sci. Comput., 36 (2014), pp. A2248–A2271, <https://doi.org/10.1137/140953289>.
- [15] W. DÖRFLER, *A Convergent Adaptive Algorithm for Poisson’s Equation*, SIAM J. Numer. Anal., 33 (1996), pp. 1106–1124, <https://doi.org/10.1137/0733054>.
- [16] M. EIGEL, C. J. GITTELSON, C. SCHWAB, AND E. ZANDER, *Adaptive stochastic Galerkin FEM*, Comput. Methods Appl. Mech. Engrg., 270 (2014), pp. 247–269, <https://doi.org/10.1016/j.cma.2013.11.015>.
- [17] M. EIGEL, C. J. GITTELSON, C. SCHWAB, AND E. ZANDER, *A convergent adaptive stochastic Galerkin finite element method with quasi-optimal spatial meshes*, ESAIM Math. Model. Numer. Anal., 49 (2015), pp. 1367–1398, <https://doi.org/10.1051/m2an/2015017>.
- [18] M. EIGEL, M. PFEFFER, AND R. SCHNEIDER, *Adaptive stochastic Galerkin FEM with hierarchical tensor representations*, Numer. Math., 136 (2017), pp. 765–803, <https://doi.org/10.1007/s00211-016-0850-x>.
- [19] R. FORSTER AND R. KORNHUBER, *A polynomial chaos approach to stochastic variational inequalities*, J. Numer. Math., 18 (2010), pp. 235–255, <https://doi.org/10.1515/jnum.2010.012>.
- [20] S. GARREIS, *Optimal Control under Uncertainty: Theory and Numerical Solution with Low-Rank Tensors*, PhD thesis, Technische Universität München, 2019.
- [21] S. GARREIS, T. M. SUROWIEC, AND M. ULBRICH, *An Interior-Point Approach for Solving Risk-Averse PDE-Constrained Optimization Problems with Coherent Risk Measures*, Preprint, submitted, Technical University of Munich, 2019.
- [22] S. GARREIS AND M. ULBRICH, *Constrained Optimization with Low-Rank Tensors and Applications to Parametric Problems with PDEs*, SIAM J. Sci. Comput., 39 (2017), pp. A25–A54, <https://doi.org/10.1137/16M1057607>.
- [23] S. GARREIS AND M. ULBRICH, *A Fully Adaptive Algorithm for PDE-Constrained Optimal Control with and without Uncertainty, Part 1: An Inexact Trust-Region Algorithm and the Deterministic Case*, Preprint,

- submitted, Technical University of Munich, 2019.
- [24] H. GOLDBERG, W. KAMPOWSKY, AND F. TRÖLTZSCH, *On NEMYTSKII Operators in l_p -Spaces of Abstract Functions*, *Math. Nachr.*, 155 (1992), pp. 127–140, <https://doi.org/10.1002/mana.19921550110>.
 - [25] L. GRASEDYCK, *Hierarchical singular value decomposition of tensors*, *SIAM J. Matrix Anal. Appl.*, 31 (2010), pp. 2029–2054, <https://doi.org/10.1137/090764189>.
 - [26] L. GRASEDYCK, D. KRESSNER, AND C. TOBLER, *A literature survey of low-rank tensor approximation techniques*, *GAMM-Mitt.*, 36 (2013), pp. 53–78, <https://doi.org/10.1002/gamm.201310004>.
 - [27] M. GUNZBURGER, C. WEBSTER, AND G. ZHANG, *Stochastic finite element methods for partial differential equations with random input data*, *Acta Numer.*, 23 (2014), pp. 521–650, <https://doi.org/10.1017/S0962492914000075>.
 - [28] W. HACKBUSCH, *Tensor Spaces and Numerical Tensor Calculus*, Springer, Heidelberg, 2012, <https://doi.org/10.1007/978-3-642-28027-6>.
 - [29] W. HACKBUSCH AND S. KÜHN, *A new scheme for the tensor representation*, *J. Fourier Anal. Appl.*, 15 (2009), pp. 706–722, <https://doi.org/10.1007/s00041-009-9094-9>.
 - [30] M. HINZE, R. PINNAU, M. ULBRICH, AND S. ULBRICH, *Optimization with PDE Constraints*, Springer, New York, 2009, <https://doi.org/10.1007/978-1-4020-8839-1>.
 - [31] S. HOLTZ, T. ROHWEDDER, AND R. SCHNEIDER, *The alternating linear scheme for tensor optimization in the tensor train format*, *SIAM J. Sci. Comput.*, 34 (2012), pp. A683–A713, <https://doi.org/10.1137/100818893>.
 - [32] D. P. KOURI, *A Multilevel Stochastic Collocation Algorithm for Optimization of PDEs with Uncertain Coefficients*, *SIAM/ASA J. Uncertain. Quantif.*, 2 (2014), pp. 55–81, <https://doi.org/10.1137/130915960>.
 - [33] D. P. KOURI, M. HEINKENSCHLOSS, D. RIDZAL, AND B. G. VAN BLOEMEN WAANDERS, *A Trust-Region Algorithm with Adaptive Stochastic Collocation for PDE Optimization under Uncertainty*, *SIAM J. Sci. Comput.*, 35 (2013), pp. A1847–A1879, <https://doi.org/10.1137/120892362>.
 - [34] D. P. KOURI AND T. M. SUROWIEC, *Risk-Averse PDE-Constrained Optimization Using the Conditional Value-At-Risk*, *SIAM J. Optim.*, 26 (2016), pp. 365–396, <https://doi.org/10.1137/140954556>.
 - [35] D. P. KOURI AND T. M. SUROWIEC, *Existence and Optimality Conditions for Risk-Averse PDE-Constrained Optimization*, *SIAM/ASA J. Uncertain. Quantif.*, 6 (2018), pp. 787–815, <https://doi.org/10.1137/16M1086613>.
 - [36] D. KRESSNER, M. STEINLECHNER, AND B. VANDEREYCKEN, *Low-rank tensor completion by Riemannian optimization*, *BIT*, 54 (2014), pp. 447–468, <https://doi.org/10.1007/s10543-013-0455-z>.
 - [37] I. V. OSELEDETS, *Tensor-train decomposition*, *SIAM J. Sci. Comput.*, 33 (2011), pp. 2295–2317, <https://doi.org/10.1137/090752286>.
 - [38] I. V. OSELEDETS AND E. E. TYRTYSHNIKOV, *TT-cross approximation for multidimensional arrays*, *Linear Algebra Appl.*, 432 (2010), pp. 70–88, <https://doi.org/10.1016/j.laa.2009.07.024>.
 - [39] C. SCHWAB AND C. J. GITTELSON, *Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs*, *Acta Numer.*, 20 (2011), pp. 291–467, <https://doi.org/10.1017/S0962492911000055>.
 - [40] A. SHAPIRO, D. DENTCHEVA, AND A. RUSZCZYŃSKI, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, Philadelphia, 2009, <https://doi.org/10.1137/1.9780898718751>.
 - [41] T. J. SULLIVAN, *Introduction to Uncertainty Quantification*, Springer, Cham, 2015, <https://doi.org/10.1007/978-3-319-23395-6>.
 - [42] H. TIESLER, R. M. KIRBY, D. XIU, AND T. PREUSSER, *Stochastic Collocation for Optimal Control Problems with Stochastic PDE Constraints*, *SIAM J. Control Optim.*, 50 (2012), pp. 2659–2682, <https://doi.org/10.1137/110835438>.
 - [43] M. ULBRICH, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, vol. 11 of MOS-SIAM Ser. Optim., SIAM, Philadelphia, 2011, <https://doi.org/10.1137/1.9781611970692>.
 - [44] J. WEIDMANN, *Linear Operators in Hilbert Spaces*, Springer, New York, 1980, <https://doi.org/10.1007/978-1-4612-6027-1>.
 - [45] E. ZEIDLER, *Nonlinear Functional Analysis and its Applications*, vol. II/B: Nonlinear Monotone Operators, Springer, New York, 1990, <https://doi.org/10.1007/978-1-4612-0981-2>.