MOX–Report No. 52/2013

Multilevel and weighted reduced basis method for stochastic optimal control problems constrained by Stokes equations

Chen, P.; Quarteroni, A.; Rozza, G.

MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox@mate.polimi.it http://mox.polimi.it

MULTILEVEL AND WEIGHTED REDUCED BASIS METHOD FOR STOCHASTIC OPTIMAL CONTROL PROBLEMS CONSTRAINED BY STOKES EQUATIONS

PENG CHEN *, ALFIO QUARTERONI * †, AND GIANLUIGI ROZZA ‡

Abstract. In this paper we develop and analyze a multilevel weighted reduced basis method for solving stochastic optimal control problems constrained by Stokes equations. Existence and uniqueness of the stochastic optimal solution is proved by establishing the equivalence between the constrained optimization problem and a stochastic saddle point problem. Analytic regularity of the optimal solution in the probability space is obtained under certain assumptions on the random input data. Finite element method and stochastic collocation method are employed for numerical approximation of the problem in deterministic space and probability space, respectively. A reduced basis method using a multilevel greedy algorithm based on isotropic and anisotropic sparse-grid techniques and weighted a posteriori error estimate is proposed in order to reduce the computational effort. A global error is obtained based on estimate results of error contribution from each method. Numerical experiments are performed with stochastic dimension ranging from 10 to 100, demonstrating that the proposed method is very efficient, especially for high dimensional and large-scale optimization problems.

Key words. uncertainty quantification, stochastic optimal control, stochastic collocation method, weighted reduced basis method, multilevel greedy algorithm, stochastic regularity, error estimate

AMS subject classifications. 35J20, 49N10, 65K10, 65C30, 60H15

1. Introduction. In computational science and engineering, often it is common practice to control selected variables of the underlying physical system modelled by partial differential equations (PDEs) in order to drive the simulation results as close as possible to some ideal data or experimental measurements. These problems can be generally formulated as optimal control of PDEs, where a so-called cost functional is minimized subject to some PDE model. In practical applications, uncertainties are inevitably encountered from various sources, for instance the PDE coefficients representing physical parameters, computational geometries, external loadings, boundary and initial conditions, etc.. This leads to the necessity of quantifying the uncertainties that play an important role in determining the optimal solutions. In the last few decades, deterministic optimal control problems without taking the uncertainties into account have been studied from both mathematical and computational perspectives [29, 19, 20, 45]. Instead, stochastic optimal control problems constrained by PDE models with random inputs have been considered only recently thanks to the development of efficient stochastic computational methods [23, 21, 38, 27, 44, 11, 46, 2, 34].

Several computational challenges arise from solving PDEs-constrained stochastic optimal control problems. Firstly, design of efficient and accurate numerical scheme for approximation of the optimal solution in the stochastic space has been a difficult task for most PDE models. Monte-Carlo method can be regarded as one of the most effective and simple schemes, however it is to be blamed for its low convergence rate, thus leading to heavy computational cost when a full deterministic

* Modelling and Scientific Computing, CMCS, Mathematics Institute of Computational Science and Engineering, MATHICSE, Ecole Polytechnique Fédérale de Lausanne, EPFL, Station 8, CH-1015 Lausanne, Switzerland. Corresponding to Peng Chen (peng.chen@epfl.ch)

†Modellistica e Calcolo Scientifico, MOX, Dipartimento di Matematica F. Brioschi, Politecnico di Milano, P.za Leonardo da Vinci 32, I-20133, Milano, Italy.

‡SISSA MathLab, International School for Advanced Studies, via Bonomea 265, 34136 Trieste, Italy.

optimal control problem has to be solved for every sample. Galerkin projection of the optimal solution onto some subspace (e.g. global polynomial space) of the stochastic space has been proven to converge exponentially fast for smooth problems [3, 23] but is not convenient to use in that the tensor-product projection scheme produces a large scale tensor system to be solved, which brings further computational difficulty. Another scheme known as stochastic collocation method based on multidimensional interpolation [2, 46] takes advantage of fast convergence of the Galerkin projection and non-intrusive property (thus easy implementation) of Monte-Carlo sampling. Secondly, to deal with high dimensional stochastic problems is commonly recognized as a computational challenge due to "curse-of-dimensionality". In order to harness the computational burden, sparse and adaptive algorithms have been well employed by making good use of sparse structure of numerical approximation and the different importance of each dimension, for instance (anisotropic) sparse-grid stochastic collocation method [46, 2, 34]. An additional computational challenge in solving PDEs-constrained stochastic optimal control problem comes from full solve of the optimality system obtained by variational approach [45], which is often ill-conditioned and naturally coupled. To undertake this difficulty, efficient preconditioning techniques have been developed to solve the optimality system by one "shot" approach [42, 37]; sequential quadratic programming [44] and trust-region algorithm [27] have also been applied as iterative approaches. However, when solving the full optimality system becomes very expensive, it is only affordable for tens or hundreds of full solve in practice, making the approaches introduced above impossible to be directly applied since the number of samples needed easily goes beyond that can be handled, especially for high dimensional problems. Since quantities of interest usually live in low dimensional manifold, model order reduction techniques may be applied using proper orthogonal decomposition or reduced basis method for parametrized optimal control problems and provide a computational opportunity, see [28, 30, 32, 31, 26].

In this paper, we study a stochastic optimal control problem constrained by Stokes equations with random inputs and distributed control function, which features all the aforementioned computational challenges, besides the additional difficulty in dealing with the saddle point structure of the underlying Stokes model [35, 6]. To tackle these challenges, a multilevel and weighted reduced basis method, using multilevel greedy algorithm and weighted a posteriori error estimate, is developed and applied to solve the stochastic optimal control problem. More in detail, (anisotropic) sparse grid stochastic collocation method is applied for stochastic approximation of the optimal solution in the probability space and finite element method with (optimal) preconditioning techniques is used for deterministic approximation in physical space, leading to a large number of finite element optimality systems to solve. To reduce the computational cost, we project the finite element optimality system into an adaptively constructed reduced basis space, leading to a reduced optimality system that can be solved with very cheap computational cost. For the construction of the reduced basis space, we design a multilevel greedy algorithm and propose a weighted a posteriori error bound, which produces quasi-optimal "snapshots" space that well approximate the low dimensional manifold of the quantities of interest. A global error analysis is carried out for the complete numerical approximation based on the regularity of the optimal solution, in particular the stochastic regularity obtained for the specific Stokes control problem. Numerical experiments with stochastic dimensions ranging from 10 to 100 are performed to verify the error convergence results and demonstrate the efficiency and accuracy of our computational method for large scale and high dimensional

PDEs-constrained optimization problems. The main contribution of this work is the development of efficient model order reduction techniques to solve stochastic optimal control problems with PDEs (Stokes equations) constraints. Another contribution is the detailed analysis of the stochastic regularity of the optimal solution with respect to input random variables and the associated error convergence analysis. Moreover, the numerical experiments demonstrate that the proposed method achieves considerable computational saving. For large-scale and "reducible" problems, it is definitely cheaper than both the stochastic collocation method [38] and Galerkin projection method [23] that have been recently developed for solving stochastic optimal control problems.

The paper is organized as follows: the stochastic optimal control problem with Stokes constraint is presented in section 2 with certainty assumptions on the random input data; section 3 is devoted to the well-posedness of the stochastic optimal solution, followed by section 4 for the study of stochastic regularity; detailed numerical approximation of the problem is presented in section 5, which provides the basis for the development of the multilevel and weighted reduced basis method in section 6; in section 7, global error estimates are carried out and verified by numerical experiments in section 8; concluding remarks are provided in the last section 9.

2. Problem statement. Let $(\Omega, \mathfrak{F}, P)$ denote a complete probability space, where Ω is a set of outcomes $\omega \in \Omega$, $\mathfrak{F}$ is a σ-algebra of events and $P : \mathfrak{F} \to [0,1]$ with $P(\Omega) = 1$ is a probability measure. A real-valued *random variable* is defined as a measurable function $Y : (\Omega, \mathfrak{F}) \to (\mathbb{R}, \mathfrak{B})$, being $\mathfrak{B}$ the Borel σ-algebra on $\mathbb{R}$. The distribution function of a random variable $Y : \Omega \to \Gamma \subset \mathbb{R}$, being Γ the image of Y, is defined as $F_Y : \Gamma \to [0,1]$ such that with $F_Y(y) = P(\omega \in \Omega : Y(\omega) \leq y)$ and its probability density function $\rho : \Gamma \to \mathbb{R}$ is given by $\rho(y)dy = dF_Y(y)$ if the random variable is continuous [15]. For any positive integer $k \in \mathbb{N}_+$, the k-th moment of Y is defined as

$$\mathbb{E}\left[Y^k\right] = \int_\Omega Y^k(\omega) dP(\omega) = \int_\Gamma y^k dF_Y(y) = \int_\Gamma y^k \rho(y) dy. \tag{2.1}$$

Let D be an open and bounded physical domain in $\mathbb{R}^d$ $(d = 2, 3)$ with Lipschitz continuous boundary ∂D. Let $v : D \times \Omega \to \mathbb{R}$ represent a real-valued *random field*, which is a real-valued random variable defined in Ω for each $x \in D$. We define the product Hilbert space $\mathcal{H}^s(D) := L^2(\Omega) \otimes H^s(D), s \in \mathbb{R}$ equipped with the norm

$$||v||_{\mathcal{H}^s(D)} := \left(\int_\Omega ||v(\cdot, \omega)||^2_{H^s(D)} dP(\omega) \right)^{1/2} < \infty, \tag{2.2}$$

where $H^s(D)$ is the Hilbert space of functions defined in the physical domain D [35, 36]. When $s = 0$, we denote $H^0(D) \equiv L^2(D)$, and thus $\mathcal{H}^0(D) \equiv \mathcal{L}^2(D)$ by convention. Moreover, the stochastic inner product is defined as

$$(w, v) = \int_\Omega \int_D wv dx dP(\omega) \quad \forall w, v \in \mathcal{L}^2(D). \tag{2.3}$$

For a random vector field $\mathbf{v} = (v_1, \ldots, v_d) : D \times \Omega \to \mathbb{R}^d$, we define the product Hilbert space $\mathcal{H}^{s,d}(D) := \left(L^2(\Omega) \otimes H^s(D)\right)^d$ $(= \mathcal{L}^{2,d}(D)$ for $s = 0)$ associated with the norm $||\mathbf{v}||_{\mathcal{H}^{s,d}(D)} = \sum_{i=1}^d ||v_i||_{\mathcal{H}^s(D)}$ and the inner product $(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^d (v_i, w_i)$.

2.1. Stochastic Stokes equations. We consider the following stochastic Stokes equations: given random variable $\nu : \Omega \to \mathbb{R}_+$, random vector fields $\mathbf{f} : D \times \Omega \to \mathbb{R}^d$ and $\mathbf{h} : \partial D_N \times \Omega \to \mathbb{R}^d$, find a solution $\{\mathbf{u}, p\} : D \times \Omega \to \mathbb{R}^d \times \mathbb{R}$ such that the following equations hold almost surely (for almost every $\omega \in \Omega$)

$$\begin{cases} -\nu(\omega)\triangle\mathbf{u}(\cdot,\omega) + \nabla p(\cdot,\omega) = \mathbf{f}(\cdot,\omega) & \text{in } D, \\ \nabla \cdot \mathbf{u}(\cdot,\omega) = 0 & \text{in } D, \\ \mathbf{u}(\cdot,\omega) = \mathbf{0} & \text{on } \partial D_D, \\ \nu(\omega)\nabla u(\cdot,\omega)\cdot\mathbf{n} - p(\cdot,\omega)\mathbf{n} = \mathbf{h}(\cdot,\omega) & \text{on } \partial D_N, \end{cases} \tag{2.4}$$

where ∂D_D and ∂D_N represent the Dirichlet and Neumann boundaries such that $\partial D_D \cup \partial D_N = \partial D$ and $\partial D_D \cap \partial D_N = \emptyset$. In particular, we consider a homogeneous Dirichlet boundary condition and a nonhomogeneous Neumann boundary condition.

At any realization $\omega \in \Omega$, the Stokes equations (2.4) is commonly used to quantify the velocity $\mathbf{u}$ and pressure p of fluid flow where advective inertial forces are negligible compared to viscous forces measured via the kinematic viscosity parameter ν. This occurs, e.g., for low speed channel flows, the flow of viscous polymers or microorganisms [1]. In practice, the viscosity ν may vary in a large extent rather than stay as a fixed constant for many fluids depending on the temperature, the multicomponent property of the fluid and some other factors [17]. Quantification of the body force $\mathbf{f}$ and boundary condition $\mathbf{h}$, for instance by experimental measurements, may also be faced with various noises or uncertainties. Incorporation of these different uncertainties leads to the study of stochastic Stokes equations.

In order to solve (2.4) in the distribution sense, we write its weak formulation as: find $\{\mathbf{u}, p\} \in \mathcal{V} \times \mathcal{Q}$ such that

$$\begin{cases} a(\mathbf{u},\mathbf{v}) + b(\mathbf{v},p) = (\mathbf{f},\mathbf{v}) + (\mathbf{h},\mathbf{v})_{\partial D_N} & \forall \mathbf{v} \in \mathcal{V}, \\ b(\mathbf{u},q) = 0 & \forall q \in \mathcal{Q}, \end{cases} \tag{2.5}$$

where $\mathcal{V} := \left\{\mathbf{v} \in \mathcal{H}^{1,d}(D) : \mathbf{v} = \mathbf{0} \text{ on } \partial D_D\right\}$, $\mathcal{Q} := L^2(\Omega) \otimes Q(D)$, with $Q(D)$ defined as

$$Q(D) := \left\{ q \in L^2(D) : \int_D q dx = 0 \right\}. \tag{2.6}$$

The bilinear form $a(\cdot,\cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ is defined as

$$a(\mathbf{w},\mathbf{v}) := \int_\Omega \int_D \nu \nabla\mathbf{w} \otimes \nabla\mathbf{v} dx dP(\omega) = \sum_{i,j=1}^{d} \int_\Omega \int_D \nu \frac{\partial w_i}{\partial x_j}\frac{\partial v_i}{\partial x_j} dx dP(\Omega) \tag{2.7}$$

and the bilinear form $b(\cdot,\cdot) : \mathcal{V} \times \mathcal{Q} \to \mathbb{R}$ reads

$$b(\mathbf{v},q) = -\int_\Omega \int_D \nabla\cdot\mathbf{v} q dx dP(\omega) = -\sum_{i=1}^{d} \int_\Omega \int_D \frac{\partial v_i}{\partial x_i} q dx dP(\omega). \tag{2.8}$$

The stochastic inner product $(\mathbf{f},\mathbf{v})$ and $(\mathbf{h},\mathbf{v})_{\partial D_N}$ are defined by the formula (2.3) on the domain D and Neumann boundary ∂D_N, respectively. To guarantee the well-posedness of saddle problem (2.5), we make the following assumption on the random variable ν and random vector fields $\mathbf{f}$ and $\mathbf{h}$.

Assumption 1. *The random viscosity ν is positive and uniformly bounded from below and from above, i.e. there exist two constants $0 < \nu_{\min} \le \nu_{max} < \infty$ such that*

$$P(\omega : \nu_{min} \le \nu(\omega) \le \nu_{max}) = 1. \tag{2.9}$$

The random force field $\mathbf{f}$ *and Neumann boundary field* $\mathbf{h}$ *satisfy*

$$||\mathbf{f}||_{\mathcal{L}} < \infty \text{ and } ||\mathbf{h}||_{\mathcal{H}} < \infty, \tag{2.10}$$

where we denote $\mathcal{L} = \mathcal{L}^{2,d}(D)$ *and* $\mathcal{H} = \mathcal{L}^{2,d}(\partial D_N)$ *for simplicity.*

The well-posedness of the stochastic Stokes problem (2.5) can be obtained by the following theorem, whose proof follows the same lines of the Brezzi theorem for the deterministic setting and will thus be omitted here, see [8, 35, 36] for details.

THEOREM 2.1. *Under Assumption 1, there exists a unique solution to the stochastic Stokes problem (2.5). Moreover, the following stability estimate holds*

$$||\mathbf{u}||_{\mathcal{V}} \leq \frac{1}{\alpha_a}\left(C_P||\mathbf{f}||_{\mathcal{L}} + \frac{\alpha_a+\gamma_a}{\beta_b}C_T||\mathbf{h}||_{\mathcal{H}}\right), \tag{2.11}$$

and

$$||p||_{\mathcal{Q}} \leq \frac{1}{\beta_b}\left(\left(1+\frac{\gamma_a}{\alpha_a}\right)C_P||\mathbf{f}||_{\mathcal{L}} + \frac{\gamma_a(\alpha_a+\gamma_a)}{\alpha_a\beta_b}C_T||\mathbf{h}||_{\mathcal{H}}\right), \tag{2.12}$$

where the positive constants $\alpha_a, \gamma_a, \beta_b, \gamma_b$ *are defined such that*

$$a(\mathbf{w},\mathbf{v}) \leq \gamma_a||\mathbf{w}||_{\mathcal{V}}||\mathbf{v}||_{\mathcal{V}} \quad \forall \mathbf{w},\mathbf{v} \in \mathcal{V} \tag{2.13}$$

and

$$a(\mathbf{v},\mathbf{v}) \geq \alpha_a||\mathbf{v}||_{\mathcal{V}}^2 \quad \forall \mathbf{v} \in \mathcal{V}_0, \tag{2.14}$$

being $\mathcal{V}_0 := \{\mathbf{v} \in \mathcal{V} : b(\mathbf{v},q) = 0, \forall q \in \mathcal{Q}\}$ *the kernel of* b*, and*

$$\inf_{q\in\mathcal{Q}}\sup_{\mathbf{v}\in\mathcal{V}}\frac{b(\mathbf{v},q)}{||\mathbf{v}||_{\mathcal{V}}||q||_{\mathcal{Q}}} \geq \beta_b, \tag{2.15}$$

where β_b *is called an inf-sup constant or compatibility constant, and the continuity*

$$b(\mathbf{v},q) \leq \gamma_b||\mathbf{v}||_{\mathcal{V}}||q||_{\mathcal{Q}} \quad \forall \mathbf{v} \in \mathcal{V}, \forall q \in \mathcal{Q}. \tag{2.16}$$

The constants C_P *and* C_T *are those of the Poincaré inequality and trace theorem [35],*

$$||\mathbf{v}||_{\mathcal{L}} \leq C_P||\mathbf{v}||_{\mathcal{V}} \text{ and } ||\mathbf{v}||_{\mathcal{H}} \leq C_T||\mathbf{v}||_{\mathcal{V}} \quad \forall \mathbf{v} \in \mathcal{V}. \tag{2.17}$$

2.2. Finite dimensional assumption. For the sake of numerical approximation of the Stokes equations (2.5) in stochastic space, we make the following finite dimensional assumption on the random input data.

Assumption 2. *The random data* ν*,* $\mathbf{f}$ *and* $\mathbf{h}$ *depend only on a finite number of random variables* $Y(\omega) = (Y_1(\omega), \ldots, Y_N(\omega)) : \Omega \to \Gamma = \Gamma_1 \times \cdots \times \Gamma_N \subset \mathbb{R}^N$ *with probability density function* $\rho = (\rho_1, \ldots, \rho_N) : \Gamma \to \mathbb{R}^N$*, i.e.* $\nu(\omega) = \nu(Y(\omega)) \in \mathbb{R}_+$*,* $\mathbf{f}(\cdot,\omega) = \mathbf{f}(\cdot, Y(\omega)) : D \to \mathbb{R}^d$ *and* $\mathbf{h}(\cdot,\omega) = \mathbf{h}(\cdot, Y(\omega)) : \partial D_N \to \mathbb{R}^d$ *almost surely.*

Remark 2.1. *The random variable* ν *and random vector fields* $\mathbf{f}$ *and* $\mathbf{h}$ *may not depend on the same random vector* Y *but on different ones* Y_ν, Y_f, Y_h*. For ease of notation, we still use a single random vector* $Y = (Y_\nu, Y_f, Y_h)$ *with total dimension* N*.*

Example 1. For a multicomponent fluid flow, the viscosity is propositional to the contribution of each component [25], which can be described by

$$\nu(Y(\omega)) = \sum_{n=1}^{N} \nu_n Y_n(\omega) + \nu_0 \left(1 - \sum_{n=1}^{N} Y_n(\omega)\right) = \nu_0 + \sum_{n=1}^{N} (\nu_n - \nu_0) Y_n(\omega), \quad (2.18)$$

where $Y_n, 1 \leq n \leq N$ are uniformly distributed in $[0, 1/N]$ and $\nu_n > 0, 0 \leq n \leq N$.

Example 2. Another example for the random vector field $\mathbf{h}$ is given by the truncated Karhunen-Loève expansion with $N+1$ terms as [43]

$$\mathbf{h}(x, Y(\omega)) = \mathbb{E}[\mathbf{h}](x) + \sum_{n=1}^{N} \sqrt{\lambda_n} \mathbf{h}_n(x) Y_n(\omega) \quad x \in \partial D_N, \quad (2.19)$$

where $(\lambda_n, \mathbf{h}_n)$ are the eigenpairs of the continuous and bounded covariance function $\mathbb{C}(x, x') = \mathbb{E}[(\mathbf{h}(x, Y) - \mathbb{E}[\mathbf{h}](x))(\mathbf{h}(x', Y) - \mathbb{E}[\mathbf{h}](x'))]$ and the random variables $Y_n, 1 \leq n \leq N$ are uncorrelated with zero mean and unit variance, given by [43]

$$Y_n(\omega) = \frac{1}{\sqrt{r_n}} \int_D \left(\mathbf{h}(x, Y(\omega)) - \mathbb{E}[\mathbf{h}](x)\right) \mathbf{h}_n(x) dx. \quad (2.20)$$

Under Assumption 2, the stochastic Stokes equations (2.4) can be viewed as a set of parameterized equations defined in a tensor product of the spatial domain and the parameter space $D \times \Gamma$. We remark that the Hilbert space $L^2(\Omega)$ is equivalent to $L^2_\rho(\Gamma)$ and we use the same notation $\mathcal{L}, \mathcal{H}, \mathcal{V}, \mathcal{Q}$ for the stochastic Hilbert spaces. Moreover, Theorem 2.1 holds under Assumption 2.

2.3. Constrained optimal control problem. We study a distributed optimal control problem constrained by the stochastic Stokes equations. Let us define a cost functional as follows

$$\begin{aligned} \mathcal{J}(\mathbf{u}, p, \mathbf{f}) &= \frac{1}{2}||\mathbf{u} - \mathbf{u}_d||^2_{\mathcal{L}} + \frac{1}{2}||p - p_d||^2_{\mathcal{L}^2(D)} + \frac{\alpha}{2}||\mathbf{f}||^2_{\mathcal{G}} \\ &= \mathbb{E}\left[\frac{1}{2}\int_D (\mathbf{u} - \mathbf{u}_d)^2 dx + \frac{1}{2}\int_D (p - p_d)^2 dx + \frac{\alpha}{2}\int_D \mathbf{f}^2 dx\right], \end{aligned} \quad (2.21)$$

where first two terms measure the discrepancy between the solution $\{\mathbf{u}, p\} \in \mathcal{V} \times \mathcal{Q}$ of the stochastic Stokes problem (2.5) and the observational data $\{\mathbf{u}_d, p_d\} \in L^{2,d}(D) \times Q(D)$ that represent the mean of measurements. The admissible control space $\mathcal{G}$ in the last term is a non-empty, closed, bounded and convex subset of $\mathcal{L}^{2,d}(D)$. This term is used to regularize in mathematical sense the control function $\mathbf{f}$ with a regularization parameter $\alpha > 0$, which can also be viewed as a penalization of the control energy. The optimal control problem constrained by the stochastic Stokes problem (2.5) can be formulated as: find an optimal solution $\{\mathbf{u}^*, p^*, \mathbf{f}^*\}$ such that

$$\mathcal{J}(\mathbf{u}^*, p^*, \mathbf{f}^*) = \min\left\{\mathcal{J}(\mathbf{u}, p, \mathbf{f}) : \{\mathbf{u}, p, \mathbf{f}\} \in \mathcal{V} \times \mathcal{Q} \times \mathcal{G} \text{ and solve (2.5)}\right\}. \quad (2.22)$$

It is easy to see that the cost functional $\mathcal{J}$ is weakly lower semicontinuous in $\mathcal{G}$, i.e.

$$\lim_{n\to\infty} \inf \mathcal{J}(\mathbf{f}_n) \geq \mathcal{J}(\mathbf{f}) \quad (2.23)$$

for any sequence $\{\mathbf{f}_n\}_{n=1}^{\infty} \in \mathcal{G}$ such that $\mathbf{f}_n \rightharpoonup \mathbf{f}$ as $n \to \infty$. Then, we have the following result by Lions' argument [29]:

THEOREM 2.2. *Under Assumption 1 and 2, there exists an optimal solution* $\{\mathbf{u}^*, p^*, \mathbf{f}^*\} \in \mathcal{V} \times \mathcal{Q} \times \mathcal{G}$ *to the stochastic optimal control problem (2.22).*

Remark 2.2. *In the cost functional, we have used the* $\mathcal{L}^{2,d}(D)$ *norm for measuring the discrepancy between the velocity field and its mean value of measurements. Extension to the case with* $\mathcal{V}$ *norm is straightforward by requiring that the data* $\{\mathbf{u}_d, p_d\}$ *possess higher regularity in the spatial domain. Another extension to stochastic data* $\{\mathbf{u}_d, p_d\}$ *can be handled in the same way as in this work provided they depend explicitly on a finite dimensional random vector, i.e.* $\{\mathbf{u}_d, p_d\}(\cdot, \omega) = \{\mathbf{u}_d, p_d\}(\cdot, Y(\omega))$.

Remark 2.3. *When the higher moments of the observational data* $\{\mathbf{u}_d, p_d\}$ *or the control function* $\mathbf{f}$*, e.g. variance, skewness, etc., or the probability distribution of* $\{\mathbf{u}_d, p_d\}$ *are incorporated into the cost functional in more general settings [44], we face essentially nonlinear and fully coupled problems, which will be addressed in [12].*

3. Saddle point formulation. In order to prove the uniqueness of the optimal solution of the constrained optimal control problem (2.22), we turn to a saddle point formulation and establish its equivalence to the optimality system obtained by Lagrangian variational approach in solving (2.22).

3.1. Optimality system. We first employ the variational approach [45] to derive an optimality system (known as Karush–Kuhn–Tucker (KKT) conditions) in solving the constrained optimal control problem (2.22). Define a compound bilinear form to represent the weak formulation of the stochastic Stokes equations (2.5) as

$$\mathcal{B}(\{\mathbf{u}, p, \mathbf{f}\}, \{\mathbf{v}, q\}) = a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) + b(\mathbf{u}, q) - (\mathbf{f}, \mathbf{v}). \tag{3.1}$$

Associated with this bilinear form, we define the Lagrangian functional

$$\mathcal{L}(\{\mathbf{u}, p, \mathbf{f}\}, \{\mathbf{u}^a, p^a\}) = \mathcal{J}(\mathbf{u}, p, \mathbf{f}) + \mathcal{B}(\{\mathbf{u}, p, \mathbf{f}\}, \{\mathbf{u}^a, p^a\}) - (\mathbf{h}, \mathbf{u}^a)_{\partial D_N}, \tag{3.2}$$

where $\{\mathbf{u}^a, p^a\} \in \mathcal{V} \times \mathcal{Q}$ are the adjoint (or dual) variables of the Stokes equations (2.5) corresponding to the state (or primal) variables $\{\mathbf{u}, p\}$. The Lagrangian functional (3.2) is Gâteaux differentiable with respect to $\{\mathbf{u}, p, \mathbf{f}, \mathbf{u}^a, p^a\}$ [45], so that we can take Gâteaux derivative of (3.2) with respect to the state variable $\{\mathbf{u}, p\}$ in test directions $\{\mathbf{v}^a, q^a\}$, control variable $\mathbf{f}$ in $\mathbf{g}$, and adjoint variable $\{\mathbf{u}^a, p^a\}$ in $\{\mathbf{v}, q\}$, respectively, obtaining the first order optimality system as

$$\begin{cases} (\{\mathbf{u}, p\}, \{\mathbf{v}^a, q^a\}) + \mathcal{B}(\{\mathbf{v}^a, q^a, \mathbf{0}\}, \{\mathbf{u}^a, p^a\}) & \\ \qquad = (\mathbf{u}_d, \mathbf{v}^a) + (p_d, p^a) & \forall \{\mathbf{v}^a, q^a\} \in \mathcal{V} \times \mathcal{Q}, \\ \alpha(\mathbf{f}, \mathbf{g}) - (\mathbf{u}^a, \mathbf{g}) = 0 & \forall \mathbf{g} \in \mathcal{G}, \\ \mathcal{B}(\{\mathbf{u}, p, \mathbf{f}\}, \{\mathbf{v}, q\}) = (\mathbf{h}, \mathbf{v})_{\partial D_N} & \forall \{\mathbf{v}, q\} \in \mathcal{V} \times \mathcal{Q}, \end{cases} \tag{3.3}$$

where we can identify that the third equation is in fact the state equation (2.5), the first one is the adjoint equation and the second one is the optimality equation. More explicitly, the optimality system can be rewritten as

$$\left\{\begin{array}{cccc:ccc} (\mathbf{u}, \mathbf{v}^a) & & & +a(\mathbf{u}^a, \mathbf{v}^a) & +b(\mathbf{v}^a, p^a) & = (\mathbf{u}_d, \mathbf{v}^a) & \forall \mathbf{v}^a \in \mathcal{V}, \\ & (p, q^a) & & +b(\mathbf{u}^a, \mathbf{q}^a) & & = (p_d, q^a) & \forall q^a \in \mathcal{Q}, \\ & & \alpha(\mathbf{f}, \mathbf{g}) & -(\mathbf{u}^a, \mathbf{g}) & & = 0 & \forall \mathbf{g} \in \mathcal{G}, \\ \hdashline a(\mathbf{u}, \mathbf{v}) & +b(\mathbf{v}, p) & -(\mathbf{f}, \mathbf{v}) & & & = (\mathbf{h}, \mathbf{v})_{\partial D_N} & \forall \mathbf{v} \in \mathcal{V}, \\ b(\mathbf{u}, q) & & & & & = 0 & \forall q \in \mathcal{Q}, \end{array}\right. \tag{3.4}$$

whose saddle point structure can be observed evidently. For the sake of numerical approximation, let us introduce the following operators corresponding to system (3.4):

$$\begin{array}{lllllll}
M_v: & \mathcal{V}\to\mathcal{V} & \text{such that} & (M_v\mathbf{u},\mathbf{v}) & = & (\mathbf{u},\mathbf{v}) & \forall \mathbf{u},\mathbf{v}\in\mathcal{V},\\
M_q: & \mathcal{Q}\to\mathcal{Q} & \text{such that} & (M_q p,q) & = & (p,q) & \forall p,q\in\mathcal{Q},\\
M_g: & \mathcal{G}\to\mathcal{G} & \text{such that} & (M_g\mathbf{f},\mathbf{g}) & = & (\mathbf{v},\mathbf{g}) & \forall \mathbf{f}\in\mathcal{G},\mathbf{g}\in\mathcal{G},\\
M_c: & \mathcal{G}\to\mathcal{V} & \text{such that} & (M_c\mathbf{g},\mathbf{v}) & = & (\mathbf{g},\mathbf{v}) & \forall \mathbf{g}\in\mathcal{G},\mathbf{v}\in\mathcal{V},\\
M_n: & \mathcal{H}\to\mathcal{V} & \text{such that} & (M_n\mathbf{h},\mathbf{v}) & = & (\mathbf{h},\mathbf{v})_{\partial D_N} & \forall \mathbf{h}\in\mathcal{H},\mathbf{v}\in\mathcal{V},\\
A: & \mathcal{V}\to\mathcal{V} & \text{such that} & a(\mathbf{u},\mathbf{v}) & = & (\nu\nabla\mathbf{u},\mathbf{v}) & \forall \mathbf{u},\mathbf{v}\in\mathcal{V},\\
B: & \mathcal{V}\to\mathcal{Q} & \text{such that} & b(\mathbf{v},q) & = & -(\nabla\cdot\mathbf{v},q) & \forall \mathbf{v}\in\mathcal{V},q\in\mathcal{Q},
\end{array} \tag{3.5}$$

from which we obtain the following saddle point linear optimality system as

$$\left(\begin{array}{ccc:cc}
M_v & 0 & 0 & A & B^T\\
0 & M_p & 0 & B & 0\\
0 & 0 & \alpha M_g & -M_c^T & 0\\
\hdashline
A & B^T & -M_c & 0 & 0\\
B & 0 & 0 & 0 & 0
\end{array}\right)
\left(\begin{array}{c}
\mathbf{u}\\ p\\ \mathbf{f}\\ \mathbf{u}^a\\ p^a
\end{array}\right)
=
\left(\begin{array}{c}
M_v\mathbf{u}_d\\ M_p p_d\\ 0\\ M_n\mathbf{h}\\ 0
\end{array}\right). \tag{3.6}$$

Remark 3.1. *The optimality system (3.6) is the first order necessary condition to guarantee that if there exists an an optimal solution to the optimal control problem (2.22), it should also satisfy the optimality system (3.6). However, the uniqueness of the optimal solution is not an immediate result.*

3.2. Saddle point formulation. In order to obtain the uniqueness and study the stochastic regularity (Sec. 4) of the optimal solution, we introduce a compound saddle point formulation of the constrained optimal control problem (2.22).

Let $\mathcal{A}:(\mathcal{V}\times\mathcal{Q}\times\mathcal{G})\times(\mathcal{V}\times\mathcal{Q}\times\mathcal{G})\to\mathbb{R}$ be a compound bilinear form defined as

$$\mathcal{A}(\{\mathbf{u},p,\mathbf{f}\},\{\mathbf{v},q,\mathbf{g}\}) = (\mathbf{u},\mathbf{v})+(p,q)+\alpha(\mathbf{f},\mathbf{g}), \tag{3.7}$$

then we have that the cost functional (2.21) can be expressed as

$$\mathcal{J}(\mathbf{u},p,\mathbf{f}) = \frac{1}{2}\mathcal{A}(\{\mathbf{u},p,\mathbf{f}\},\{\mathbf{u},p,\mathbf{f}\}) - \mathcal{A}(\{\mathbf{u},p,\mathbf{f}\},\{\mathbf{u}_d,p_d,\mathbf{0}\}) + C, \tag{3.8}$$

where C is a constant given by $C=\mathcal{A}(\{\mathbf{u}_d,p_d,\mathbf{0}\},\{\mathbf{u}_d,p_d,\mathbf{0}\})/2$. We recall that $\mathcal{B}:(\mathcal{V}\times\mathcal{Q}\times\mathcal{G})\times(\mathcal{V}\times\mathcal{Q})\to\mathbb{R}$ is a compound bilinear form defined in (3.1). Then the following proposition establishes the equivalence between the constrained optimal control problem (2.22) and the saddle point problem (3.14), whose proof follows the one in the deterministic setting, see [8, 7] for details.

PROPOSITION 3.1. *Suppose that the bilinear form $\mathcal{A}$ is symmetric, non-negative and continuous, i.e. there exists a constant $\gamma>0$ such that $\forall\{\mathbf{u},p,\mathbf{f}\},\{\mathbf{v},q,\mathbf{g}\}\in\mathcal{V}\times\mathcal{Q}\times\mathcal{G}$, we have*

$$\mathcal{A}(\{\mathbf{u},p,\mathbf{f}\},\{\mathbf{v},q,\mathbf{g}\}) \le \gamma||\{\mathbf{u},p,\mathbf{f}\}||_{\mathcal{V}\times\mathcal{Q}\times\mathcal{G}}||\{\mathbf{v},q,\mathbf{g}\}||_{\mathcal{V}\times\mathcal{Q}\times\mathcal{G}}. \tag{3.9}$$

Moreover, suppose that $\mathcal{A}$ is strongly coercive in the kernel space of $\mathcal{B}$, defined as

$$\mathcal{K} := \big\{\{\mathbf{u},p,\mathbf{f}\}\in\mathcal{V}\times\mathcal{Q}\times\mathcal{G}: \quad \mathcal{B}(\{\mathbf{u},p,\mathbf{f}\},\{\mathbf{v},q\})=0 \quad \forall\{\mathbf{v},q\}\in\mathcal{V}\times\mathcal{Q}\big\}, \tag{3.10}$$

i.e. there exists a constant $\epsilon>0$ such that $\forall\{\mathbf{v},q,\mathbf{g}\}\in\mathcal{V}\times\mathcal{Q}\times\mathcal{G}$, we have

$$\mathcal{A}(\{\mathbf{v},q,\mathbf{g}\},\{\mathbf{v},q,\mathbf{g}\}) \ge \epsilon||\{\mathbf{v},q,\mathbf{g}\}||^2_{\mathcal{V}\times\mathcal{Q}\times\mathcal{G}}. \tag{3.11}$$

Suppose that $\mathcal{B}$ is continuous, i.e. there exists a constant $\delta > 0$ such that $\forall \{\mathbf{u}, p, \mathbf{f}\} \in \mathcal{V} \times \mathcal{Q} \times \mathcal{G}, \{\mathbf{v}, q\} \in \mathcal{V} \times \mathcal{Q}$, we have

$$\mathcal{B}(\{\mathbf{u}, p, \mathbf{f}\}, \{\mathbf{v}, q\}) \leq \delta ||\{\mathbf{u}, p, \mathbf{f}\}||_{\mathcal{V}\times\mathcal{Q}\times\mathcal{G}} ||\{\mathbf{v}, q\}||_{\mathcal{V}\times\mathcal{Q}}. \tag{3.12}$$

Furthermore, suppose that $\mathcal{B}$ satisfies the inf-sup condition, i.e. there exists a constant $\beta > 0$ such that

$$\inf_{\{\mathbf{v},q\}\in\mathcal{V}\times\mathcal{Q}} \sup_{\{\mathbf{u},p,\mathbf{f}\}\in\mathcal{V}\times\mathcal{Q}\times\mathcal{G}} \frac{\mathcal{B}(\{\mathbf{u}, p, \mathbf{f}\}, \{\mathbf{v}, q\})}{||\{\mathbf{u}, p, \mathbf{f}\}||_{\mathcal{V}\times\mathcal{Q}\times\mathcal{G}} ||\{\mathbf{v}, q\}||_{\mathcal{V}\times\mathcal{Q}}} \geq \beta. \tag{3.13}$$

Then the constrained optimal control problem (2.22) is equivalent to the following saddle point problem: find $\{\mathbf{u}, p, \mathbf{f}\} \in \mathcal{V} \times \mathcal{Q} \times \mathcal{G}$ and $\{\mathbf{u}^a, p^a\} \in \mathcal{V} \times \mathcal{Q}$ such that

$$\begin{cases} \mathcal{A}(\{\mathbf{u}, p, \mathbf{f}\}, \{\mathbf{v}^a, q^a, \mathbf{g}\}) + \mathcal{B}(\{\mathbf{v}^a, q^a, \mathbf{g}\}, \{\mathbf{u}^a, p^a\}) \\ = (\{\mathbf{u}_d, p_d, \mathbf{0}\}, \{\mathbf{v}^a, q^a, \mathbf{g}\}) \quad \forall \{\mathbf{v}^a, q^a, \mathbf{g}\} \in \mathcal{V} \times \mathcal{Q} \times \mathcal{G}, \\ \mathcal{B}(\{\mathbf{u}, p, \mathbf{f}\}, \{\mathbf{v}, q\}) = (\mathbf{h}, \mathbf{v})_{\partial D_N} \quad \forall \{\mathbf{v}, q\} \in \mathcal{V} \times \mathcal{Q}. \end{cases} \tag{3.14}$$

Remark 3.2. *By establishing the equivalence between the optimality system (3.3) and the saddle point system (3.14), it can be shown that the variables $\{\mathbf{u}^a, p^a\}$ (and $\{\mathbf{v}^a, q^a, \mathbf{g}\}$) used, with slight abuse of notation, in the saddle point formulation (3.14) are coincident with the adjoint variables (and test variables) as introduced in the Lagrangian functional (3.2). Moreover, we highlight that some mathematical properties such as stochastic regularity (Sec. 4) of the two systems hold the same.*

3.3. Equivalence, uniqueness and stability estimates.

LEMMA 3.2. *The constrained optimal control problem (2.22), the saddle point problem (3.14) and the first order optimality system (3.3) are equivalent problems.*

Proof. To prove the equivalence between the first two problems, we only need to verify the assumptions in Proposition 3.1. By the definition (3.7), it is easy to check that $\mathcal{A}$ is symmetric and non-negative; $\mathcal{A}$ is also continuous

$$\begin{aligned} |\mathcal{A}(\{\mathbf{u}, p, \mathbf{f}\},\{\mathbf{v}, q, \mathbf{g}\})| &\leq ||\mathbf{u}||_{\mathcal{V}} ||\mathbf{v}||_{\mathcal{V}} + ||p||_{\mathcal{Q}} ||q||_{\mathcal{Q}} + \alpha ||\mathbf{f}||_{\mathcal{G}} ||\mathbf{g}||_{\mathcal{G}} \\ &\leq \gamma ||\{\mathbf{u}, p, \mathbf{f}\}||_{\mathcal{V}\times\mathcal{Q}\times\mathcal{G}} ||\{\mathbf{v}, q, \mathbf{g}\}||_{\mathcal{V}\times\mathcal{Q}\times\mathcal{G}}, \forall \{\mathbf{v}, q, \mathbf{g}\} \in \mathcal{V} \times \mathcal{Q} \times \mathcal{G}, \end{aligned} \tag{3.15}$$

where the continuity constant $\gamma = 1$ and $||\{\mathbf{v}, q, \mathbf{g}\}||_{\mathcal{V}\times\mathcal{Q}\times\mathcal{G}} := ||\mathbf{v}||_{\mathcal{V}} + ||q||_{\mathcal{Q}} + \sqrt{\alpha}||\mathbf{g}||_{\mathcal{G}}$. For any $\{\mathbf{v}, q, \mathbf{g}\} \in \mathcal{K}$, the kernel of $\mathcal{B}$ defined in (3.10), we have by Theorem 2.1 that $||\mathbf{v}||_{\mathcal{V}} \leq C_P ||\mathbf{g}||_{\mathcal{G}}/\alpha_a$, which yields

$$\begin{aligned} \mathcal{A}(\{\mathbf{v}, q, \mathbf{g}\}, \{\mathbf{v}, q, \mathbf{g}\}) &= ||\mathbf{v}||^2_{\mathcal{L}} + ||q||^2_{\mathcal{Q}} + \alpha ||\mathbf{g}||^2_{\mathcal{G}} \\ &\geq \frac{\alpha_a^2 \alpha}{2C_P^2} ||\mathbf{v}||^2_{\mathcal{V}} + ||q||^2_{\mathcal{Q}} + \frac{\alpha}{2} ||\mathbf{g}||^2_{\mathcal{G}} \\ &\geq \frac{1}{3} \min\left\{ \frac{\alpha_a^2 \alpha}{2C_P^2}, \frac{1}{2} \right\} ||\{\mathbf{v}, q, \mathbf{g}\}||^2_{\mathcal{V}\times\mathcal{Q}\times\mathcal{G}}, \end{aligned} \tag{3.16}$$

from which we can infer that $\mathcal{A}$ is coercive on $\mathcal{K}$ with a coercivity constant $\epsilon = (1/3)\min\{\alpha_a^2\alpha/(2C_P^2), 1/2\}$. As for the continuity of the bilinear form $\mathcal{B}$ defined in (3.1), by Assumption 1 and Theorem 2.1 we have for any $\{\{\mathbf{u}, p, \mathbf{f}\}, \{\mathbf{v}, q\}\} \in (\mathcal{V} \times \mathcal{Q} \times \mathcal{G}) \times (\mathcal{V} \times \mathcal{Q})$,

$$\begin{aligned} |\mathcal{B}(\{\mathbf{u}, p, \mathbf{f}\}, \{\mathbf{v}, q\})| &\leq \nu_{max} ||\mathbf{u}||_{\mathcal{V}} ||\mathbf{v}||_{\mathcal{V}} + \gamma_b ||\mathbf{v}||_{\mathcal{V}} ||p||_{\mathcal{Q}} + \gamma_b ||\mathbf{u}||_{\mathcal{V}} ||q||_{\mathcal{Q}} + ||\mathbf{f}||_{\mathcal{G}} ||\mathbf{v}||_{\mathcal{V}} \\ &\leq \max\{\nu_{max}, \gamma_b, 1/\sqrt{\alpha}\} ||\{\mathbf{u}, p, \mathbf{f}\}||_{\mathcal{V}\times\mathcal{Q}\times\mathcal{G}} ||\{\mathbf{v}, q\}||_{\mathcal{V}\times\mathcal{Q}}, \end{aligned} \tag{3.17}$$

the continuity constant being $\delta = \max\{\nu_{max}, \gamma_b, 1/\sqrt{\alpha}\}$. Finally, $\mathcal{B}$ satisfies the inf-sup condition, as

$$\begin{aligned} &\inf_{\{\mathbf{v},q\}\in\mathcal{V}\times\mathcal{Q}} \sup_{\{\mathbf{u},p,\mathbf{f}\}\in\mathcal{V}\times\mathcal{Q}\times\mathcal{G}} \frac{\mathcal{B}(\{\mathbf{u},p,\mathbf{f}\},\{\mathbf{v},q\})}{||\{\mathbf{u},p,\mathbf{f}\}||_{\mathcal{V}\times\mathcal{Q}\times\mathcal{G}}||\{\mathbf{v},q\}||_{\mathcal{V}\times\mathcal{Q}}} \\ \geq &\inf_{\{\mathbf{v},q\}\in\mathcal{V}\times\mathcal{Q}} \sup_{\{\mathbf{u},p,\mathbf{0}\}\in\mathcal{V}\times\mathcal{Q}\times\mathcal{G}} \frac{\mathcal{B}(\{\mathbf{u},p,\mathbf{0}\},\{\mathbf{v},q\})}{||\{\mathbf{u},p,\mathbf{0}\}||_{\mathcal{V}\times\mathcal{Q}\times\mathcal{G}}||\{\mathbf{v},q\}||_{\mathcal{V}\times\mathcal{Q}}} \geq \beta. \end{aligned} \tag{3.18}$$

The inf-sup constant $\beta > 0$ depends on $\alpha_a, \gamma_a, \beta_b$ as follows (see [47]):

$$\beta = \frac{1}{k_{12} + \max\{k_{11}, k_{22}\}}, \tag{3.19}$$

where

$$k_{11} = \alpha_a^{-2}(1 + \beta_b^{-2}\gamma_a^2), \quad k_{22} = \beta_b^{-2}\gamma_a^2 k_{11} + \beta_b^{-2} \text{ and } k_{12} = \beta_b^{-1}\gamma_a k_{11}. \tag{3.20}$$

We conclude that solving the constrained optimal control problem (2.22) is equivalent to solve the saddle point problem (3.14), thanks to Proposition 3.1. The equivalence between the optimality system (3.3) and the saddle point system (3.14) can be observed by noticing that by adding the second equation (optimal equation) of (3.3) to its first one (adjoint equation), we obtain the first equation of (3.14). □

Thanks to Lemma 3.2 and using Theorem 2.1, we can conclude that the optimal solution is unique and satisfies a priori (boundedness) estimate:

THEOREM 3.3. *There exists a unique optimal solution to the constrained optimal control problem (2.22). Moreover, the optimal solution* $\{\mathbf{u}, p, \mathbf{f}\}$ *and the adjoint variables* $\{\mathbf{u}^a, p^a\}$ *satisfy the following stability estimates:*

$$||\{\mathbf{u}, p, \mathbf{f}\}||_{\mathcal{V}\times\mathcal{Q}\times\mathcal{G}} \leq \alpha_1 ||\{\mathbf{u}_d, p_d\}||_{\mathcal{L}\times\mathcal{Q}} + \beta_1 ||\mathbf{h}||_{\mathcal{H}} \tag{3.21}$$

and

$$||\{\mathbf{u}^a, p^a\}||_{\mathcal{V}\times\mathcal{Q}} \leq \alpha_2 ||\{\mathbf{u}_d, p_d\}||_{\mathcal{L}\times\mathcal{Q}} + \beta_2 ||\mathbf{h}||_{\mathcal{H}} \tag{3.22}$$

where the constants $\alpha_1, \beta_1, \alpha_2, \beta_2$ *are defined as*

$$\alpha_1 = \frac{1}{\epsilon}\max\{C_P, 1\}, \quad \beta_1 = \frac{1}{\epsilon}\frac{\epsilon+\gamma}{\beta}C_T, \tag{3.23}$$

and

$$\alpha_2 = \frac{1}{\beta}\left(1 + \frac{\gamma}{\epsilon}\right)\max\{C_P, 1\}, \quad \beta_2 = \frac{1}{\beta}\frac{\gamma(\epsilon+\gamma)}{\epsilon\beta}C_T. \tag{3.24}$$

4. Stochastic regularity. In this section, we show that under suitable assumptions for the regularity of the viscosity $\nu : \Gamma \to \mathbb{R}_+$ and boundary data $\mathbf{h} : \Gamma \to H$ in the stochastic space Γ, the solution $\{\mathbf{u}, p, \mathbf{f}, \mathbf{u}^a, p^a\} : \Gamma \to V \times Q \times G \times V \times Q$ can be analytically extended to a complex region that covers the stochastic space Γ. (Here and in the following, we denote L, V, Q, G, H as the deterministic Hilbert space corresponding to their stochastic counterparts $\mathcal{L}, \mathcal{V}, \mathcal{Q}, \mathcal{G}, \mathcal{H}$, e.g. $H = L^{2,d}(\partial D_N)$.)

Let $\mathbf{k} = (k_1, \ldots, k_N) \in \mathbb{N}_0^N$ be a N-dimensional multi-index of non-negative integers, with $\mathbf{k}! = \prod_{i_1}^{k_1} i_1 \cdots \prod_{i_N}^{k_N} i_N$, $|\mathbf{k}| = \sum_{n=1}^N k_n$, and $|\mathbf{k}|! = \prod_{i=1}^{|\mathbf{k}|} i$; let $\partial_y^{\mathbf{k}}\{\cdot\} =$

$\partial_{y_1}^{k_1}\partial_{y_2}^{k_2}\cdots\partial_{y_N}^{k_N}\{\cdot\}$ represent the $\mathbf{k}$-th order partial derivative with respect to the parameter $y=(y_1,\ldots,y_N)$. Let us also define the following constants for ease of notation

$$C_\alpha=\alpha_1+\alpha_2, C_\beta=\beta_1+\beta_2, C_{\alpha,\beta}=\max\{\alpha_1+\alpha_2,\beta_1+\beta_2\}, \tag{4.1}$$

where $\alpha_1,\alpha_2,\beta_1,\beta_2$ are the stability constants defined in the Brezzi theorem (3.3).

We make the following assumption of stochastic regularity on the input data:

Assumption 3. *For every $y\in\Gamma$, there exists a N-dimensional positive rate vector $\mathbf{r}=(r_1,\ldots,r_N)\in\mathbb{R}_+^N$ such that the $\mathbf{k}$-th order derivative of the viscosity $\nu:\Gamma\to\mathbb{R}_+$ and the boundary condition $\mathbf{h}:\Gamma\to H$ satisfy*

$$C_{\alpha,\beta}\frac{|\partial_y^{\mathbf{k}}\nu(y)|}{\nu(\bar{y})}\leq\mathbf{r}^{\mathbf{k}} \text{ and } \frac{C_\beta||\partial_y^{\mathbf{k}}\mathbf{h}(y)||_H}{C_\alpha||\{\mathbf{u}_d,p_d\}||_{L\times Q}+C_\beta||\mathbf{h}(y)||_H}\leq|\mathbf{k}|!\mathbf{r}^{\mathbf{k}}. \tag{4.2}$$

THEOREM 4.1. *Under assumption 3, we have the following a priori estimate for the $\mathbf{k}$-th order derivative of the solution $\{\mathbf{u},p,\mathbf{f},\mathbf{u}^a,p^a\}:\Gamma\to V\times Q\times G\times V\times Q$*

$$\begin{aligned}&||\partial_y^{\mathbf{k}}\{\mathbf{u}(y),p(y),\mathbf{f}(y)\}||_{V\times Q\times G}+||\partial_y^{\mathbf{k}}\{\mathbf{u}^a(y),p^a(y)\}||_{V\times Q}\\&\leq C(C_\alpha||\{\mathbf{u}_d,p_d\}||_{L\times Q}+C_\beta||\mathbf{h}(y)||_H)|\mathbf{k}|!(r\mathbf{r})^{\mathbf{k}},\end{aligned} \tag{4.3}$$

where $r\mathbf{r}=(rr_1,rr_2,\ldots,rr_N)$ with the constant rate $r>1/\log(2)$, and the constant C is independent of $\mathbf{k}$, which will be provided explicitly in the proof.

Proof. The semi-weak formulation of the saddle point problem (3.14) reads: find $\{\mathbf{u}(y),p(y),\mathbf{f}(y)\}\in V\times Q\times G$ and $\{\mathbf{u}^a(y),p^a(y)\}\in V\times Q$ such that

$$\begin{cases}\mathcal{A}\left(\{\mathbf{u}(y),p(y),\mathbf{f}(y)\},\{\mathbf{v}^a,q^a,\mathbf{g}\}\right)+\mathcal{B}(\{\mathbf{v}^a,q^a,\mathbf{g}\},\{\mathbf{u}^a(y),p^a(y)\};y)\\\qquad=(\mathbf{u}_d,\mathbf{v}^a)+(p_d,q^a)\quad\forall\{\mathbf{v}^a,q^a,\mathbf{g}\}\in V\times Q\times G,\\\mathcal{B}(\{\mathbf{u}(y),p(y),\mathbf{g}(y)\},\{\mathbf{v},q\};y)=(\mathbf{h}(y),\mathbf{v})_{\partial D_N}\quad\forall\{\mathbf{v},q\}\in V\times Q,\end{cases} \tag{4.4}$$

where we have used the same bilinear forms $\mathcal{A}$ and $\mathcal{B}$ for ease of notation, which can be identified in the semi-weak sense by their explicit dependence on the parameter y. Taking $\mathbf{k}$-th ($|\mathbf{k}|>0$) order partial derivative of problem (4.4) with respect to the parameter y, we obtain the following problem thanks to the general Leibniz rule: find $\partial_y^{\mathbf{k}}\{\mathbf{u}(y),p(y),\mathbf{f}(y)\}\in V\times Q\times G$ and $\partial_y^{\mathbf{k}}\{\mathbf{u}^a(y),p^a(y)\}\in V\times Q$ such that

$$\begin{cases}\mathcal{A}\left(\partial_y^{\mathbf{k}}\{\mathbf{u}(y),p(y),\mathbf{f}(y)\},\{\mathbf{v}^a,q^a,\mathbf{g}\}\right)+\mathcal{B}(\{\mathbf{v}^a,q^a,\mathbf{g}\},\partial_y^{\mathbf{k}}\{\mathbf{u}^a(y),p^a(y)\};y)\\\qquad=-\sum_{\mathbf{k}'\in\Lambda(\mathbf{k})}(\partial_y^{\mathbf{k}-\mathbf{k}'}\nu(y)\nabla\partial_y^{\mathbf{k}'}\mathbf{u}^a(y),\nabla\mathbf{v}^a)\quad\forall\{\mathbf{v}^a,q^a,\mathbf{g}\}\in V\times Q\times G,\\\mathcal{B}(\partial_y^{\mathbf{k}}\{\mathbf{u}(y),p(y),\mathbf{g}(y)\},\{\mathbf{v},q\};y)=(\partial_y^{\mathbf{k}}\mathbf{h}(y),\mathbf{v})_{\partial D_N}\\\qquad-\sum_{\mathbf{k}'\in\Lambda(\mathbf{k})}(\partial_y^{\mathbf{k}-\mathbf{k}'}\nu(y)\nabla\partial_y^{\mathbf{k}'}\mathbf{u}(y),\nabla\mathbf{v})\quad\forall\{\mathbf{v},q\}\in V\times Q,\end{cases} \tag{4.5}$$

where the multivariate index set $\Lambda(\mathbf{k})$ is defined as

$$\Lambda(\mathbf{k})=\left\{\mathbf{k}'\in\mathbb{N}_0^N:k_n'\leq k_n,\forall 1\leq n\leq N,\text{ and }\mathbf{k}'\neq\mathbf{k}\right\}. \tag{4.6}$$

By the Brezzi theorem 3.3, the solution of problem (4.5) admits the following estimate

$$\begin{cases} ||\partial_y^{\mathbf{k}}\{\mathbf{u}(y),p(y),\mathbf{f}(y)\}||_{V\times Q\times G} \leq \alpha_1 \displaystyle\sum_{\mathbf{k}'\in\Lambda(\mathbf{k})} \frac{|\partial_y^{\mathbf{k}-\mathbf{k}'}\nu(y)|}{\nu(\bar{y})}||\partial_y^{\mathbf{k}'}\mathbf{u}^a(y)||_V \\ \qquad\qquad +\beta_1\left(||\partial_y^{\mathbf{k}}\mathbf{h}(y)||_H + \displaystyle\sum_{\mathbf{k}'\in\Lambda(\mathbf{k})} \frac{|\partial_y^{\mathbf{k}-\mathbf{k}'}\nu(y)|}{\nu(\bar{y})}||\partial_y^{\mathbf{k}'}\mathbf{u}(y)||_V\right), \\ ||\partial_y^{\mathbf{k}}\{\mathbf{u}^a(y),p^a(y)\}||_{V\times Q} \leq \alpha_2 \displaystyle\sum_{\mathbf{k}'\in\Lambda(\mathbf{k})} \frac{|\partial_y^{\mathbf{k}-\mathbf{k}'}\nu(y)|}{\nu(\bar{y})}||\partial_y^{\mathbf{k}'}\mathbf{u}^a(y)||_V \\ \qquad\qquad +\beta_2\left(||\partial_y^{\mathbf{k}}\mathbf{h}(y)||_H + \displaystyle\sum_{\mathbf{k}'\in\Lambda(\mathbf{k})} \frac{|\partial_y^{\mathbf{k}-\mathbf{k}'}\nu(y)|}{\nu(\bar{y})}||\partial_y^{\mathbf{k}'}\mathbf{u}(y)||_V\right), \end{cases} \tag{4.7}$$

where the parameters $\alpha_1, \alpha_2, \beta_1, \beta_2$ are given in (3.23) and (3.24). Adding the second inequality of (4.3) to the first one and noting that $\forall \mathbf{k}' \in \Lambda(\mathbf{k})$,

$$||\partial_y^{\mathbf{k}'}\mathbf{u}^a(y)||_V \leq ||\partial_y^{\mathbf{k}'}\{\mathbf{u}^a(y),p^a(y)\}||_{V\times Q} \tag{4.8}$$

and

$$||\partial_y^{\mathbf{k}'}\mathbf{u}(y)||_V \leq ||\partial_y^{\mathbf{k}'}\{\mathbf{u}(y),p(y),\mathbf{f}(y)\}||_{V\times Q\times G}, \tag{4.9}$$

thus

$$\begin{aligned} &||\partial_y^{\mathbf{k}}\{\mathbf{u}(y),p(y),\mathbf{f}(y)\}||_{V\times Q\times G} + ||\partial_y^{\mathbf{k}}\{\mathbf{u}^a(y),p^a(y)\}||_{V\times Q} \leq C_\beta||\partial_y^{\mathbf{k}}\mathbf{h}(y)||_H + C_{\alpha,\beta} \\ &\sum_{\mathbf{k}'\in\Lambda(\mathbf{k})} \frac{|\partial_y^{\mathbf{k}-\mathbf{k}'}\nu(y)|}{\nu(\bar{y})}\left(||\partial_y^{\mathbf{k}'}\{\mathbf{u}(y),p(y),\mathbf{f}(y)\}||_{V\times Q\times G} + ||\partial_y^{\mathbf{k}'}\{\mathbf{u}^a(y),p^a(y)\}||_{V\times Q}\right), \end{aligned} \tag{4.10}$$

where the constants C_β and $C_{\alpha,\beta}$ are defined in (4.1).

To prove the estimate (4.3) for a general $\mathbf{k} \in \mathbb{N}_0^N$, we adopt an induction argument based on the recursive result (4.10). To start, we consider the case when $|\mathbf{k}| = 0$. Applying the Brezzi theorem to the semi-weak problem (4.4), we have

$$\begin{cases} ||\{\mathbf{u}(y),p(y),\mathbf{f}(y)\}||_{V\times Q\times G} \leq \alpha_1||\{\mathbf{u}_d,p_d\}||_{L\times Q} + \beta_1||\mathbf{h}(y)||_H, \\ ||\{\mathbf{u}^a(y),p^a(y)\}||_{V\times Q} \leq \alpha_2||\{\mathbf{u}_d,p_d\}||_{L\times Q} + \beta_2||\mathbf{h}(y)||_H. \end{cases} \tag{4.11}$$

Adding the second inequality of (4.11) to the first one, we find

$$\begin{aligned} &||\{\mathbf{u}(y),p(y),\mathbf{f}(y)\}||_{V\times Q\times G} + ||\{\mathbf{u}^a(y),p^a(y)\}||_{V\times Q} \\ &\leq C_\alpha||\{\mathbf{u}_d,p_d\}||_{L\times Q} + C_\beta||\mathbf{h}(y)||_H \\ &= (C_\alpha||\{\mathbf{u}_d,p_d\}||_{L\times Q} + C_\beta||\mathbf{h}(y)||_H)|\mathbf{k}|!\mathbf{r}^{\mathbf{k}}, \end{aligned} \tag{4.12}$$

which verifies the estimate (4.3) for $|\mathbf{k}| = 0$ by noting that $r > 1$ and $C = 1$.

When $|\mathbf{k}| = 1$, i.e. there exists $n, 1 \leq n \leq N$ such that $k_n = 1$ and $k_{n^*} = 0$ for all

$n^* \neq n, 1 \leq n^* \leq N$, we have by the estimates (4.10) and (4.12) and Assumption 3

$$\begin{aligned}
&||\partial_y^{\mathbf{k}}\{\mathbf{u}(y), p(y), \mathbf{f}(y)\}||_{V\times Q\times G} + ||\partial_y^{\mathbf{k}}\{\mathbf{u}^a(y), p^a(y)\}||_{V\times Q} \\
&= ||\partial_y^{k_n}\{\mathbf{u}(y), p(y), \mathbf{f}(y)\}||_{V\times Q\times G} + ||\partial_y^{k_n}\{\mathbf{u}^a(y), p^a(y)\}||_{V\times Q} \\
&\leq C_\beta||\partial_y^{k_n}\mathbf{h}(y)||_H + C_{\alpha,\beta}\frac{|\partial_y^{k_n}\nu(y)|}{\nu(\bar{y})}\left(C_\alpha||\{\mathbf{u}_d, p_d\}||_{L\times Q} + C_\beta||\mathbf{h}(y)||_H\right) \\
&\leq 2\left(C_\alpha||\{\mathbf{u}_d, p_d\}||_{L\times Q} + C_\beta||\mathbf{h}(y)||_H\right) r_n \\
&= 2\left(C_\alpha||\{\mathbf{u}_d, p_d\}||_{L\times Q} + C_\beta||\mathbf{h}(y)||_H\right) |\mathbf{k}|!\mathbf{r}^{\mathbf{k}},
\end{aligned} \tag{4.13}$$

which yields the estimate (4.3) for $|\mathbf{k}| = 1$ by noting that $r > 1$ and $C = 2$.

As for more general $\mathbf{k}$ with $|\mathbf{k}| > 1$, we first prove the following auxiliary estimate

$$\begin{aligned}
&||\partial_y^{\mathbf{k}}\{\mathbf{u}(y), p(y), \mathbf{f}(y)\}||_{V\times Q\times G} + ||\partial_y^{\mathbf{k}}\{\mathbf{u}^a(y), p^a(y)\}||_{V\times Q} \\
&\leq \left(C_\alpha||\{\mathbf{u}_d, p_d\}||_{L\times Q} + C_\beta||\mathbf{h}(y)||_H\right) s(|\mathbf{k}|)\mathbf{r}^{\mathbf{k}},
\end{aligned} \tag{4.14}$$

where $s(|\mathbf{k}|)$ depends only on $|\mathbf{k}|$ according to the following recursive formula,

$$s(0) = 1, s(1) = 2, s(|\mathbf{k}|) = 1 + \sum_{|\mathbf{k}'|=0}^{|\mathbf{k}|-1} \left(\begin{array}{c} |\mathbf{k}| \\ |\mathbf{k}'| \end{array} \right) s(|\mathbf{k}'|). \tag{4.15}$$

In fact, (4.14) holds for $|\mathbf{k}| = 0$ and $|\mathbf{k}| = 1$ due to (4.12) and (4.13). By induction, we assume that the stability estimate (4.14) holds for every $\mathbf{k}' \in \Lambda(\mathbf{k})$, so that (4.10) implies

$$\begin{aligned}
&||\partial_y^{\mathbf{k}}\{\mathbf{u}(y), p(y), \mathbf{f}(y)\}||_{V\times Q\times G} + ||\partial_y^{\mathbf{k}}\{\mathbf{u}^a(y), p^a(y)\}||_{V\times Q} \leq C_\beta||\partial_y^{\mathbf{k}}\mathbf{h}(y)||_H \\
&+ C_{\alpha,\beta}\sum_{\mathbf{k}'\in\Lambda(\mathbf{k})}\frac{|\partial_y^{\mathbf{k}-\mathbf{k}'}\nu(y)|}{\nu(\bar{y})}\left(C_\alpha||\{\mathbf{u}_d, p_d\}||_{L\times Q} + C_\beta||\mathbf{h}(y)||_H\right) s(|\mathbf{k}'|)\mathbf{r}^{\mathbf{k}'} \\
&\leq \left(C_\alpha||\{\mathbf{u}_d, p_d\}||_{L\times Q} + C_\beta||\mathbf{h}(y)||_H\right)\left(|\mathbf{k}|!\mathbf{r}^{\mathbf{k}} + \sum_{\mathbf{k}'\in\Lambda(\mathbf{k})}\mathbf{r}^{\mathbf{k}-\mathbf{k}'}s(|\mathbf{k}|')\mathbf{r}^{\mathbf{k}'}\right) \\
&= \left(C_\alpha||\{\mathbf{u}_d, p_d\}||_{L\times Q} + C_\beta||\mathbf{h}(y)||_H\right)\left(|\mathbf{k}|! + \sum_{|\mathbf{k}'|=0}^{|\mathbf{k}|-1}\left(\begin{array}{c} |\mathbf{k}| \\ |\mathbf{k}'| \end{array} \right) s(|\mathbf{k}'|)\right)\mathbf{r}^{\mathbf{k}} \\
&= \left(C_\alpha||\{\mathbf{u}_d, p_d\}||_{L\times Q} + C_\beta||\mathbf{h}(y)||_H\right) s(|\mathbf{k}|)\mathbf{r}^{\mathbf{k}},
\end{aligned} \tag{4.16}$$

where we have used the assumption 3 for the second inequality, the fact that $\mathbf{r}^{\mathbf{k}} = \mathbf{r}^{\mathbf{k}-\mathbf{k}'}\mathbf{r}^{\mathbf{k}'}$ for any $\mathbf{k}' \in \Lambda(\mathbf{k})$, and the following relation by summation reordering

$$\sum_{\mathbf{k}'\in\Lambda(\mathbf{k})} s(|\mathbf{k}|') = \sum_{|\mathbf{k}'|=0}^{|\mathbf{k}|-1}\left(\begin{array}{c} |\mathbf{k}| \\ |\mathbf{k}'| \end{array} \right) s(|\mathbf{k}'|), \tag{4.17}$$

thanks to the definition of $\Lambda(\mathbf{k})$ in (4.6). By this end, it is left to establish a suitable bound for $s(|\mathbf{k}|)$ in order to prove the estimate (4.3) from the estimate (4.14). Let us denote $k = |\mathbf{k}|$, $k' = |\mathbf{k}'|$, and define $t(k) = s(k)/k!$, so that from (4.15) we have

$$t(k) = \frac{1}{k!}\left(k! + \sum_{k'=0}^{k-1}\frac{k!}{(k-k')!}\frac{s(k')}{k'!}\right) = 1 + \sum_{k'=0}^{k-1}\frac{t(k')}{(k-k')!}. \tag{4.18}$$

Suppose that for all k, $t(k) \leq c_r r^k$ for some positive constants c_r, r to be determined, so that (4.18) yields

$$t(k) - 1 = \sum_{k'=0}^{k-1} \frac{t(k')}{(k-k')!} = \sum_{k'=1}^{k} \frac{t(k-k')}{k'!} \leq c_r r^k \sum_{k'=1}^{k} \frac{r^{-k'}}{k'!} \leq c_r r^k \left(e^{\frac{1}{r}} - 1\right). \quad (4.19)$$

On the other hand, $t(k) - 1 \leq c_r r^k - 1$ from our assumption. Hence, we only require that $c_r r^k \left(e^{1/r} - 1\right) \leq c_r r^k - 1$, which can be satisfied when $r > 1/\log(2)$ and $c_r \geq 1/(2 - e^{1/r})$. Therefore, $s(k) = t(k)k! \leq c_r r^k k!$, implying that

$$s(|\mathbf{k}|) \leq c_r r^{|\mathbf{k}|} |\mathbf{k}|! = c_r \mathbf{r}_r^{\mathbf{k}} |\mathbf{k}|!, \quad (4.20)$$

where the N-dimensional constant rate vector $\mathbf{r}_r = (r, \ldots, r)$. The proof is concluded by substituting (4.20) into (4.14), noting $\mathbf{r}_r^{\mathbf{k}} \mathbf{r}^{\mathbf{k}} = (r\mathbf{r})^{\mathbf{k}}$, and setting $C = c_r$ in (4.3). ∎

Let us define a complex region associated with the stability estimate (4.3) as

$$\Sigma := \left\{ z \in \mathbb{C} : \exists y \in \Gamma \text{ such that } \sum_{n=1}^{N} r r_n |z_n - y_n| < 1 \right\}. \quad (4.21)$$

Then we have that the solution does not only have bounded partial derivative but can be analytically extended to the complex region Σ, as stated in the following theorem:

THEOREM 4.2. *Under assumption 3, the solution of the semi-weak saddle point problem (4.4) admits an analytical extension to the region Σ defined in (4.21).*

Proof. Given any $y \in \Gamma$, the Taylor expansion of the solution of problem (4.4) $\{\mathbf{u}, p, \mathbf{f}\} : \Gamma \to V \times Q \times G$ and $\{\mathbf{u}^a, p^a\} : \Gamma \to V \times Q$ about y reads

$$\{\mathbf{u}(z), p(z), \mathbf{f}(z)\} = \sum_{\mathbf{k} \in \mathbb{N}_0^N} \frac{\partial_y^{\mathbf{k}} \{\mathbf{u}(y), p(y), \mathbf{f}(y)\}}{\mathbf{k}!} (z - y)^{\mathbf{k}} \quad (4.22)$$

and

$$\{\mathbf{u}^a(z), p^a(z)\} = \sum_{\mathbf{k} \in \mathbb{N}_0^N} \frac{\partial_y^{\mathbf{k}} \{\mathbf{u}^a(y), p^a(y)\}}{\mathbf{k}!} (z - y)^{\mathbf{k}}, \quad (4.23)$$

where $(z-y)^{\mathbf{k}} = \prod_{n=1}^{N} (z_n - y_n)^{k_n}$. By Theorem 4.1, we have

$$\begin{aligned} &||\{\mathbf{u}(z), p(z), \mathbf{f}(z)\}||_{V \times Q \times G} + ||\{\mathbf{u}^a(z), p^a(z)\}||_{V \times Q} \\ &\leq \sum_{\mathbf{k} \in \mathbb{N}_0^N} \left(||\partial_y^{\mathbf{k}} \{\mathbf{u}(y), p(y), \mathbf{f}(y)\}||_{V \times Q \times G} + ||\partial_y^{\mathbf{k}} \{\mathbf{u}^a(y), p^a(y)\}||_{V \times Q} \right) \frac{|z-y|^{\mathbf{k}}}{\mathbf{k}!} \\ &\leq C(C_\alpha ||\{\mathbf{u}_d, p_d\}||_{L \times Q} + C_\beta ||\mathbf{h}(y)||_H) \sum_{\mathbf{k} \in \mathbb{N}_0^N} |\mathbf{k}|! (r\mathbf{r})^{\mathbf{k}} \frac{|z-y|^{\mathbf{k}}}{\mathbf{k}!}. \end{aligned} \quad (4.24)$$

Upon reordering, we have

$$\sum_{\mathbf{k} \in \mathbb{N}_0^N} |\mathbf{k}|! (r\mathbf{r})^{\mathbf{k}} \frac{|z-y|^{\mathbf{k}}}{\mathbf{k}!} = \sum_{k=0}^{\infty} \sum_{|\mathbf{k}|=k} \frac{k!}{\mathbf{k}!} \prod_{n=1}^{N} (r r_n |z_n - y_n|)^{k_n}. \quad (4.25)$$

By Newton's generalized binomial formula, we have

$$\sum_{k=0}^{\infty}\sum_{|\mathbf{k}|=k}\frac{k!}{\mathbf{k}!}\prod_{n=1}^{N}(rr_n|z_n-y_n|)^{k_n}=\sum_{k=0}^{\infty}\left(\sum_{n=1}^{N}rr_n|z_n-y_n|\right)^k, \tag{4.26}$$

which converges when $\sum_{n=1}^{N} rr_n|z_n-y_n|<1$. This concludes our proof. □

5. Numerical approximation. In order to solve the constrained optimization problem (2.22), we need to solve the equivalent saddle point problem (3.14) (or, equivalently (3.6)) thanks to Lemma 3.2. Hereafter, we introduce a numerical approximation of system (3.6) in the probability space Γ by the stochastic collocation method and in the physical domain D by the finite element method.

5.1. Stochastic collocation method. For stochastic problems with smooth solution in the probability space, the stochastic collocation method based on sparse grid techniques [46, 2, 34, 33] features both fast convergence of stochastic Galerkin method and the non-intrusive structure of Monte Carlo method. This makes it an efficient method in solving stochastic optimal control problems [44, 27, 11].

Let X denote a general Hilbert space defined in the physical domain D, e.g. $H^1(D)$. Let $C(\Gamma;X)$ be the space of continuous functions with values in X, i.e.

$$C(\Gamma;X):=\left\{v:\Gamma\to X|v \text{ is continuously measurable and } \max_{y\in\Gamma}||v(y)||_X<\infty\right\}. \tag{5.1}$$

Let $\mathcal{P}_m(\Gamma)$ be a space of polynomials with degree less than or equal to m in each coordinate $y_n, 1\le n\le N$. Let $\mathcal{U}^{i_n}: C(\Gamma;X)\to\mathcal{P}_{m(i_n)-1}(\Gamma_n)\otimes X$ denote a one-dimensional Lagrangian interpolation operator based on the set of collocation nodes $\Theta_n^{i_n}=\{y_n^1,\dots,y_n^{m(i_n)}\}$, $1\le n\le N$, defined as

$$\mathcal{U}^{i_n}v(y_n)=\sum_{j_n=1}^{m(i_n)}v(y_n^{j_n})l_n^{j_n}(y_n),\ \text{with } l_n^{j_n}(y_n)=\prod_{1\le k\le m(i_n):k\ne j_n}\frac{y_n-y_n^k}{y_n^{j_n}-y_n^k}, \tag{5.2}$$

where $m(k)$ is a function of k depending on the choice of collocation nodes, e.g. $m(k)=1$ when $k=1$ and $m(k)=2^{k-1}+1, 1\le n\le N$ when $k>1$ [34]. We define the sparse grid Smolyak formula $\mathcal{S}_q: C(\Gamma;X)\to\mathcal{P}_{m(q-N+1)-1}(\Gamma)\otimes X$ as [34]

$$\mathcal{S}_q v(y)=\sum_{q-N+1\le|\mathbf{i}|\le q}(-1)^{q-|\mathbf{i}|}\begin{pmatrix}N-1\\ q-|\mathbf{i}|\end{pmatrix}\mathcal{I}_{\mathbf{i}}v(y),\quad q=N,N+1,\dots \tag{5.3}$$

where $\mathbf{i}=(i_1,\dots,i_N)\in\mathbb{N}_+^N$ with $|\mathbf{i}|=i_1+\cdots+i_N$ is a multi-index. The tensor-product interpolation operator $\mathcal{I}_{\mathbf{i}}: C(\Gamma;X)\to\mathcal{P}_{m(\mathbf{i})-1}(\Gamma)\otimes X$ is defined on the set of collocation nodes $\Theta^{\mathbf{i}}=\Theta_1^{i_1}\times\cdots\times\Theta_N^{i_N}$ as

$$\mathcal{I}_{\mathbf{i}}v(y)=(\mathcal{U}^{i_1}\otimes\cdots\otimes\mathcal{U}^{i_N})v(y)=\sum_{j_1=1}^{m(i_1)}\cdots\sum_{j_N=1}^{m(i_N)}v(y_1^{j_1},\dots,y_N^{j_N})\bigotimes_{n=1}^{N}l_n^{j_n}(y_n). \tag{5.4}$$

With the definition of Lagrangian interpolation operator (5.3) and (5.4), we can approximate statistics of interest, e.g. expectation, by

$$\mathbb{E}[\mathcal{S}_q v]=\sum_{q-N+1\le|\mathbf{i}|\le q}(-1)^{q-|\mathbf{i}|}\begin{pmatrix}N-1\\ q-|\mathbf{i}|\end{pmatrix}\mathbb{E}[\mathcal{I}_{\mathbf{i}}v], \tag{5.5}$$

being $\mathbb{E}[\mathcal{I}_{\mathbf{i}}v]$ defined as

$$\mathbb{E}[\mathcal{I}_{\mathbf{i}}v] = \sum_{j_1=1}^{m(i_1)} \cdots \sum_{j_N=1}^{m(i_N)} v(y_1^{j_1}, \ldots, y_N^{j_N}) \prod_{n=1}^{N} w_n^{j_n}, \tag{5.6}$$

where the quadrature weights are given by

$$w_n^{j_n} = \int_{\Gamma_n} l_n^{j_n}(y_n)\rho(y_n)dy_n \quad 0 \le j_n \le m(i_n), 1 \le n \le N. \tag{5.7}$$

Remark 5.1. *The accuracy of stochastic collocation approximation depends on the choice of the collocation nodes. Among the most popular, we mention Clenshaw-Curtis abscissas, Gauss abscissas of certain orthogonal polynomials corresponding to the joint probability density function ρ, e.g. Gauss-Jacobi abscissas for beta density function, Gauss-Hermite abscissas for normal density function, see [34, 9] for more details.*

By using the difference operator $\triangle^{i_n} = \mathcal{U}^{i_n} - \mathcal{U}^{i_n-1}$, with $\mathcal{U}^0 = 0$, we have an alternative representation of the sparse grid Smolyak formula (5.3) as follows

$$\mathcal{S}_q v(y) = \sum_{\mathbf{i} \in X(q,N)} (\triangle^{i_1} \otimes \cdots \otimes \triangle^{i_N}) v(y) \tag{5.8}$$

with the multivariate index set defined as

$$X(q,N) := \left\{ \mathbf{i} \in \mathbb{N}_+^N : \sum_{n=1}^{N} i_n \le q \right\}, \quad q = N, N+1, \ldots. \tag{5.9}$$

Let $H(q,N) := \{\Theta^{\mathbf{i}}, \mathbf{i} \in X(q,N)\}$ denote the set of collocation nodes associated to the index set $X(q,N)$. The cardinality of $H(q,N)$ grows exponentially with respect to the dimension of the problem [46, 34]. In tackling high dimensional problems, each dimension may be given appropriate relevance by applying anisotropic sparse grid interpolation formula written as [33]

$$\mathcal{S}_q^{\boldsymbol{\alpha}} v(y) = \sum_{\mathbf{i} \in X_{\boldsymbol{\alpha}}(q,N)} (\triangle^{i_1} \otimes \cdots \otimes \triangle^{i_N}) v(y), \tag{5.10}$$

where the anisotropic multivariate index set $X_{\boldsymbol{\alpha}}(q,N)$ is defined as

$$X_{\boldsymbol{\alpha}}(q,N) := \left\{ \mathbf{i} \in \mathbb{N}_+^N : \sum_{n=1}^{N} \alpha_n i_n \le \min_{1 \le n \le N} \alpha_n q \right\}, \; q = N, N+1, \ldots. \tag{5.11}$$

Here, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)$ is a positive multivariate weight, which can be obtained by a priori or a posteriori estimate [33], or by a suitable dimension adaptive algorithm [18]. Similarly, we define the set of collocation nodes $H_\alpha(q,N) := \{\Theta^{\mathbf{i}}, \mathbf{i} \in X_{\boldsymbol{\alpha}}(q,N)\}$. Note that the isotropic sparse grid interpolation (5.8) is a special case corresponding to $\boldsymbol{\alpha} = \mathbf{1}$. Evaluation of statistics based on the anisotropic sparse grid stochastic collocation method, e.g. expectation, is straightforward by the following approximation

$$\mathbb{E}[\mathcal{S}_q^{\boldsymbol{\alpha}} v] = \sum_{\mathbf{i} \in X_{\boldsymbol{\alpha}}(q,N)} \mathbb{E}\left[(\triangle^{i_1} \otimes \cdots \otimes \triangle^{i_N}) v \right]. \tag{5.12}$$

5.2. Finite element method. Given a regular triangulation $\mathcal{T}_h$ of the physical domain $\bar{D} \subset \mathbb{R}^d$ with mesh size h, we define the finite element space [36, 35]

$$X_h^k := \{v_h \in C^0(\bar{D}) : v_h|_K \in \mathbb{P}_k \quad \forall K \in \mathcal{T}_h\}, \quad k \geq 1, \tag{5.13}$$

where $C^0(\bar{D})$ is the space of continuous functions in $\bar{D}$, $\mathbb{P}_k, k \geq 1$, is the space of polynomials of degree less than or equal to k in the variables $x_1, \ldots, x_d$. In particular, we define $V_h^k := (X_h^k)^d \cap V$, $Q_h^m := X_h^m \cap Q$, and $G_h^l := (X_h^l)^d \cap G$ with $k, m, l \geq 1$ as finite element approximation spaces corresponding to the Hilbert spaces V, Q and G, respectively, defined in Section 4. The semi-weak finite element approximation of the saddle point problem (3.14) reads: for any $y \in \Gamma$, find $\{\mathbf{u}_h(y), p_h(y), \mathbf{f}_h(y)\} \in V_h^k \times Q_h^m \times G_h^l$ and $\{\mathbf{u}_h^a(y), p_h^a(y)\} \in V_h^k \times Q_h^m$ such that

$$\begin{cases} \mathcal{A}\left(\{\mathbf{u}_h(y), p_h(y), \mathbf{f}_h(y)\}, \{\mathbf{v}_h^a, q_h^a, \mathbf{g}_h\}\right) + \mathcal{B}(\{\mathbf{v}_h^a, q_h^a, \mathbf{g}_h\}, \{\mathbf{u}_h^a(y), p_h^a(y)\}; y) \\ \qquad = (\mathbf{u}_d, \mathbf{v}_h^a) + (p_d, q_h^a) \quad \forall \{\mathbf{v}_h^a, q_h^a, \mathbf{g}_h\} \in V_h^k \times Q_h^m \times G_h^l, \\ \mathcal{B}(\{\mathbf{u}_h(y), p_h(y), \mathbf{g}_h(y)\}, \{\mathbf{v}_h, q_h\}; y) = (\mathbf{h}(y), \mathbf{v}_h)_{\partial D_N} \quad \forall \{\mathbf{v}_h, q_h\} \in V_h^k \times Q_h^m. \end{cases} \tag{5.14}$$

The well-posedness of problem (5.14) can be guaranteed by fulfilling the conditions in Proposition 3.1 in finite element spaces. In particular, the compatibility condition (3.13) needs to be satisfied in the finite element spaces V_h^k, Q_h^m, G_h^l. In fact, it is a consequence (as can be observed from the proof (3.18)) of the compatibility condition (2.15) in V_h^k, Q_h^m, for which we may use, e.g. the Taylor-Hood elements ($m = k-1, k \geq 2$) among many feasible choices [35], leading to stable finite element approximation featuring optimal convergence rate. We set $l = k$ for the control function space G_h^l.

Let the finite element solution of the saddle point problem (5.14) be written as

$$\mathbf{u}_h(y) = \sum_{n=1}^{N_v} u_n(y)\boldsymbol{\psi}_n, p_h(y) = \sum_{n=1}^{N_p} p_n(y)\varphi_n, \mathbf{f}_h(y) = \sum_{n=1}^{N_v} f_n(y)\boldsymbol{\psi}_n, \tag{5.15}$$

and

$$\mathbf{u}_h^a(y) = \sum_{n=1}^{N_v} u_n^a(y)\boldsymbol{\psi}_n, p_h^a(y) = \sum_{n=1}^{N_p} p_n^a(y)\varphi_n, \tag{5.16}$$

where $\boldsymbol{\psi}_n, 1 \leq n \leq N_v$ and $\varphi_n, 1 \leq n \leq N_p$ are the bases of the finite element spaces V_h^k and Q_h^k, respectively. Note that the bases of V_h^k and G_h^l are the same when $k = l$. Corresponding to the matrix operators defined in (3.5), the finite element mass matrices $M_{v,h}$ (note that $M_{g,h} = M_{c,h} = M_{v,h}$ when $k = l$) and $M_{p,h}$ are obtained as

$$(M_{v,h})_{mn} = (\boldsymbol{\psi}_n, \boldsymbol{\psi}_m), 1 \leq m, n \leq N_v; (M_{p,h})_{mn} = (\varphi_n, \varphi_m), 1 \leq m, n \leq N_p, \tag{5.17}$$

and the mass matrix for Neumann boundary condition is given by

$$(M_{n,h})_{mn} = (\boldsymbol{\psi}_m, \boldsymbol{\psi}_n)_{\partial D_N}, 1 \leq m, n \leq N_v. \tag{5.18}$$

The stiffness matrix A_h^y is obtained as

$$(A_h^y)_{mn} = a(\boldsymbol{\psi}_n, \boldsymbol{\psi}_m; y), 1 \leq m, n \leq N_v, \tag{5.19}$$

and the matrix B_h^y corresponding to the compatibility condition is written as

$$(B_h)_{mn} = b(\boldsymbol{\psi}_m, \varphi_n), 1 \leq m \leq N_v, 1 \leq n \leq N_p. \tag{5.20}$$

Let $U_h(y) = (u_1(y), \ldots, u_{N_v}(y))^T$ represent the coefficient vector for the finite element function $\mathbf{u}_h(y)$, and $P_h(y), F_h(y), U_h^a(y), P_h^a(y)$ the coefficient vectors for the functions $p_h(y), \mathbf{f}_h(y), \mathbf{u}_h^a(y)$, and $U_{d,h}, P_{d,h}, H_h(y)$ the values of $\mathbf{u}_d, p_d, \mathbf{h}(y)$ at the finite element nodes. To this end, the algebraic formulation of problem (5.14) can be written via the optimality operator system (3.6) as

$$\left(\begin{array}{ccc:cc} M_{v,h} & 0 & 0 & A_h^y & B_h^T \\ 0 & M_{p,h} & 0 & B_h & 0 \\ 0 & 0 & \alpha M_{g,h} & -M_{c,h}^T & 0 \\ \hdashline A_h^y & B_h^T & -M_{c,h} & 0 & 0 \\ B_h & 0 & 0 & 0 & 0 \end{array}\right) \left(\begin{array}{c} U_h(y) \\ P_h(y) \\ F_h(y) \\ U_h^a(y) \\ P_h^a(y) \end{array}\right) = \left(\begin{array}{c} M_{v,h}U_{d,h} \\ M_{p,h}P_{d,h} \\ 0 \\ M_{n,h}H_h(y) \\ 0 \end{array}\right). \quad (5.21)$$

The matrix of the linear system (5.21) becomes ill-conditioned with large condition number when h or α is very small, which makes it unsuitable for direct solve. Alternatively, we seek the solution by MINRES iteration with the help of the following block diagonal preconditioner [42, 37],

$$P(y) = \left(\begin{array}{ccc} \hat{M}_{s,h} & 0 & 0 \\ 0 & \alpha\hat{M}_{g,h} & 0 \\ 0 & 0 & \hat{K}_{s,h}^y M_{s,h}^{-1} (\hat{K}_{s,h}^y)^T \end{array}\right). \quad (5.22)$$

The mass matrix $M_{s,h}$ and the saddle point matrix $K_{s,h}^y$ corresponding to the Stokes problem (2.5) in deterministic setting are defined as

$$M_{s,h} = \left(\begin{array}{cc} M_{v,h} & 0 \\ 0 & M_{p,h} \end{array}\right) \text{ and } K_{s,h}^y = \left(\begin{array}{cc} A_h^y & B_h^T \\ B_h & 0 \end{array}\right), \quad (5.23)$$

where the matrices $\hat{M}_{s,h}$, $\hat{M}_{g,h}$ and $\hat{K}_{s,h}^y$ can be regarded as convenient approximations of $M_{s,h}$, $M_{g,h}$ and $K_{s,h}^y$ obtained by using suitable iteration methods [32, 37], e.g. symmetric Gauss-Seidel iteration for $\hat{M}_{s,h}$ and $\hat{M}_{g,h}$, and inexact Uzawa iteration for $\hat{K}_{s,h}^y$.

6. Multilevel and weighted reduced basis method. To solve a full system (5.21) at one sample $y \in \Gamma$ is very expensive when the number of degrees of freedom of the finite element approximation is large. The task becomes prohibitive when the dimension of the probability space Γ is so high that a large number of samples are necessary to be used in order to obtain accurate statistics of interest. To circumvent this computational obstacle, we propose a reduced basis method [10, 14, 16] featuring multilevel greedy algorithm and weighted a posteriori error bound. The crucial consideration is that the optimal solution of the constrained optimization problem (2.22) lies in a low dimensional manifold, despite the fact that the random inputs live in high dimensional probability space.

6.1. Reduced basis approximation. The idea behind reduced basis approximation is to take "snapshots" - that is high fidelity solutions of the underlying PDE model - as bases and then approximate the solution at a new sample by Galerkin projection on the pre-selected snapshots [40, 14]. Specific to the finite element problem (5.21), the associated reduced basis problem can be formulated as: for any $y \in \Gamma$, find

$\{\mathbf{u}_r(y), p_r(y), \mathbf{f}_r(y)\} \in V_{N_r} \times Q_{N_r} \times G_{N_r}$ and $\{\mathbf{u}_r^a(y), p_r^a(y)\} \in V_{N_r} \times Q_{N_r}$ such that

$$\begin{cases} \mathcal{A}\left(\{\mathbf{u}_r(y), p_r(y), \mathbf{f}_r(y)\}, \{\mathbf{v}_r^a, q_r^a, \mathbf{g}_r\}\right) + \mathcal{B}(\{\mathbf{v}_r^a, q_r^a, \mathbf{g}_r\}, \{\mathbf{u}_r^a(y), p_r^a(y)\}; y) \\ \qquad = (\mathbf{u}_d, \mathbf{v}_r^a) + (p_d, q_r^a) \quad \forall \{\mathbf{v}_r^a, q_r^a, \mathbf{g}_r\} \in V_{N_r} \times Q_{N_r} \times G_{N_r}, \\ \mathcal{B}(\{\mathbf{u}_r(y), p_r(y), \mathbf{g}_r(y)\}, \{\mathbf{v}_r, q_r\}; y) = (\mathbf{h}(y), \mathbf{v}_r)_{\partial D_N} \quad \forall \{\mathbf{v}_r, q_r\} \in V_{N_r} \times Q_{N_r}, \end{cases} \tag{6.1}$$

where $V_{N_r}, Q_{N_r}, G_{N_r}$ are reduced basis spaces constructed from the snapshots at the pre-selected samples $y^1, \ldots, y^{N_r}$. More in detail, G_{N_r} is constructed by

$$G_{N_r} = \text{span}\{\mathbf{f}_h(y^n), 1 \le n \le N_r\}. \tag{6.2}$$

As for Q_{N_r}, we take the union of the state and adjoint snapshots of pressure in order to guarantee the approximate stability in the reduced basis space [32], written as

$$Q_{N_r} = Q_{N_r}^s \cup Q_{N_r}^a = \text{span}\{p_h(y^n), p_h^a(y^n), 1 \le n \le N_r\}. \tag{6.3}$$

As for V_{N_r}, a simple union of the state and adjoint snapshots of velocity is not sufficient to satisfy the compatibility condition (3.13) in the reduced basis spaces. To overcome this difficulty, it is sufficient to enrich the reduced basis velocity space. With this aim, we introduce the supremizer operator $T : Q_h^m \to V_h^k$ [41, 39],

$$(Tq_h, \mathbf{v}_h)_A = b(\mathbf{v}_h, q_h) \quad \forall \mathbf{v} \in V_h^k, \tag{6.4}$$

where the A-scalar product is defined as

$$(\mathbf{u}, \mathbf{v})_A = a(\mathbf{u}, \mathbf{v}; \bar{y}) \quad \forall \mathbf{u}, \mathbf{v} \in V, \tag{6.5}$$

being $\bar{y} \in \Gamma$ a reference value, for instance, the center of Γ. Then, we construct the reduced basis velocity space V_{N_r} as the union of state and adjoint velocity snapshots enriched by pressure supremizers

$$V_{N_r} = V_{N_r}^s \cup V_{N_r}^a = \text{span}\{\mathbf{u}_h(y^n), Tp_h(y^n), \mathbf{u}_h^a(y^n), Tp_h^a(y^n), 1 \le n \le N_r\}. \tag{6.6}$$

It can be proven [39] that the compatibility condition (2.15) is satisfied in $V_{N_r}^s$ and $V_{N_r}^a$ with $\beta_b^{N_r} \ge \beta_b^h$, being $\beta_b^{N_r}$ and β_b^h the compatibility constants of the bilinear form b of (2.15) in the reduced basis space and finite element space, respectively. Consequently, the compatibility condition (3.13) is satisfied in V_{N_r} following the proof of (3.18), with the compatibility constants $\beta^{N_r} \ge \beta^h$ corresponding to that in (3.19). Following the argument in the proof of Lemma 3.2, it is straightforward to check that the other conditions in Proposition 3.1 are also satisfied in the reduced basis space $V_{N_r} \times Q_{N_r} \times G_{N_r}$. Hence, there exists a unique reduced basis solution to problem (6.1).

For the sake of algebraic stability, we perform Gram-Schmidt orthonormalization [39] to the reduced basis spaces V_{N_r}, Q_{N_r} and G_{N_r}, obtaining the orthonormal bases such that $V_{N_r} = \text{span}\{\boldsymbol{\zeta}_n^v, 1 \le n \le 4N_r\}$, $Q_{N_r} = \text{span}\{\zeta_n^p, 1 \le n \le 2N_r\}$ and $G_{N_r} = \text{span}\{\boldsymbol{\zeta}_n^g, 1 \le n \le N_r\}$. Finally, at any $y \in \Gamma$, we project the finite element solution $\{\mathbf{u}_h(y), p_h(y), \mathbf{f}_h(y)\} \in V_h^k \times Q_h^m \times G_h^l$ into the reduced basis space $V_{N_r} \times Q_{N_r} \times G_{N_r}$ as

$$\mathbf{u}_h(y) = \sum_{n=1}^{4N_r} u_n(y)\boldsymbol{\zeta}_n^v, p_h(y) = \sum_{n=1}^{2N_r} p_n(y)\zeta_n^p, \mathbf{f}_h(y) = \sum_{n=1}^{N_r} f_n(y)\boldsymbol{\zeta}_n^g, \tag{6.7}$$

and the adjoint variables $\{\mathbf{u}_h^a(y), p_h^a(y)\} \in V_h^k \times Q_h^m$ into $V_{N_r} \times Q_{N_r}$ as

$$\mathbf{u}_h^a(y) = \sum_{n=1}^{4N_r} u_n^a(y)\boldsymbol{\zeta}_n^v, p_h^a(y) = \sum_{n=1}^{2N_r} p_n^a(y)\zeta_n^p. \tag{6.8}$$

Let $U_r(y) = (u_1(y), \ldots, u_{4N_r}(y))$ denote the coefficient vector of the reduced basis approximation, and define $P_r(y), F_r(y), U_r^a(y)$ and $P_r^a(y)$ similarly, corresponding to those of the finite element approximation. Let $\mathcal{Z}_{N_r}^v = (\boldsymbol{\zeta}_1^v, \ldots, \boldsymbol{\zeta}_{4N_r}^v)^T$, $\mathcal{Z}_{N_r}^p = (\zeta_1^p, \ldots, \zeta_{2N_r}^p)^T$ and $\mathcal{Z}_{N_r}^g = (\boldsymbol{\zeta}_1^g, \ldots, \boldsymbol{\zeta}_{N_r}^g)^T$, by which we define the reduced basis mass matrices as follows: $M_{v,r} = (\mathcal{Z}_{N_r}^v)^T M_{v,h} \mathcal{Z}_{N_r}^v$, $M_{p,r} = (\mathcal{Z}_{N_r}^p)^T M_{p,h} \mathcal{Z}_{N_r}^p$, $M_{g,r} = (\mathcal{Z}_{N_r}^g)^T M_{g,h} \mathcal{Z}_{N_r}^g$, $M_{c,r} = (\mathcal{Z}_{N_r}^v)^T M_{c,h} \mathcal{Z}_{N_r}^g$, $M_{n,r} = (\mathcal{Z}_{N_r}^v)^T M_{n,h} \mathcal{Z}_{N_r}^v$, and the Stokes matrices A_r^y and B_r as $A_r^y = (\mathcal{Z}_{N_r}^v)^T A_h^y \mathcal{Z}_{N_r}^v$, and $B_r = (\mathcal{Z}_{N_r}^p)^T B_h \mathcal{Z}_{N_r}^v$. The reduced basis data vector $U_{d,r}, P_{d,r}, H_r(y)$ are defined as $U_{d,r} = (\mathcal{Z}_{N_r}^v)^T U_{d,h}, P_{d,r} = (\mathcal{Z}_{N_r}^p)^T P_{d,h}, H_r(y) = (\mathcal{Z}_{N_r}^v)^T H_h(y)$. By projecting the finite element system (5.21) into the reduced basis spaces, we obtain the algebraic formulation of the reduced basis problem corresponding to the finite element algebraic system (5.21) as

$$\left(\begin{array}{ccc:cc} M_{v,r} & 0 & 0 & A_r^y & B_r^T \\ 0 & M_{p,r} & 0 & B_r & 0 \\ 0 & 0 & \alpha M_{g,r} & -M_{c,r}^T & 0 \\ \hdashline A_r^y & B_r^T & -M_{c,r} & 0 & 0 \\ B_r & 0 & 0 & 0 & 0 \end{array}\right) \left(\begin{array}{c} U_r(y) \\ P_r(y) \\ F_r(y) \\ U_r^a(y) \\ P_r^a(y) \end{array}\right) = \left(\begin{array}{c} M_{v,r}U_{d,r} \\ M_{p,r}P_{d,r} \\ 0 \\ M_{n,r}H_r(y) \\ 0 \end{array}\right), \tag{6.9}$$

which is a $13N_r \times 13N_r$ linear system, whose numerical solution costs far less computational effort than solving the finite element system (5.21) thanks to the fact that N_r is much smaller than the number of degrees of freedom of the finite element discretization.

6.2. A multilevel greedy algorithm. The efficiency of the reduced basis approximation depends critically on the choice of reduced bases, and thus on the samples $y^1, \ldots, y^{N_r}$ selected in the construction of the reduced basis spaces $V_{N_r}, Q_{N_r}, G_{N_r}$. In order to choose the most representative samples, we propose a multilevel greedy algorithm based on the sparse grid construction for stochastic collocation method and reduce the computational cost of the construction of the reduced basis spaces.

To begin, we choose the first sample from the zeroth level of the sparse grid, i.e. $y^1 \in H(q, N)$ (or $H_{\boldsymbol{\alpha}}(q, N)$ for anisotropic sparse grid) with $q - N = 0$, where only one collocation node is available. We solve the finite element problem (5.21) at y^1 and construct the reduced basis space V_1, Q_1, G_1 according to (6.2), (6.3) and (6.6).

Let $\mathcal{E}_r$ denote the reduced basis approximation error defined as

$$\mathcal{E}_r(y) := ||\mathrm{u}_h - \mathrm{u}_r||_{\mathrm{V}}, \tag{6.10}$$

where we denote the Hilbert space $\mathrm{V} = V \times Q \times G \times V \times Q$, the solution $\mathrm{u}(y) := \{\mathbf{u}(y), p(y), \mathbf{f}(y), \mathbf{u}^a(y), p^a(y)\}$ with finite element approximation u_h and reduced basis approximation u_r. At each of the level $q - N = l, l = 1, 2, \ldots, L$ with prescribed $L \in \mathbb{N}_+$, we first construct the set of collocation nodes $H(q, N)$ of the sparse grid and then choose the "most representative" sample y^{N_r+1} by minimizing $\mathcal{E}_r(y)$ over the new collocation nodes in the current level of the sparse grid, i.e.

$$y^{N_r+1} = \arg \max_{y \in H(q,N) \setminus H(q-1,N)} \mathcal{E}_r(y). \tag{6.11}$$

Note that in the hierarchical sparse grid with nested collocation nodes, we have $H(q-1,N) \subset H(q,N), q \geq N+1$, which provides further computational efficiency since there is no need to evaluate the error at the collocation nodes in the previous level. After updating the reduced basis spaces V_{N_r}, Q_{N_r} and G_{N_r} by the finite element solution of problem (5.21) at y^{N_r+1}, we set $N_r+1 \to N_r$ and proceed to choose the next sample until the error $\mathcal{E}_r(y^{N_r+1})$ is smaller than a prescribed tolerance ϵ_{tol}. Then we move to the next level $q-N=l+1$. However, in order to compute the reduced basis approximation error (6.10), we have to solve the full finite element system (5.21), which is out of reach. Instead of computing an exact reduced basis approximation error $\mathcal{E}_r(y)$, we seek to evaluate a cheap, sharp and reliable error bound $\triangle_r(y)$ depending on $\{\mathbf{u}_r(y), p_r(y), \mathbf{f}_r(y), \mathbf{u}_r^a(y), p_r^a(y)\}$ at $y \in \Gamma$ such that

$$c\triangle_r(y) \leq \mathcal{E}_r(y) \leq \triangle_r(y) \tag{6.12}$$

with the constant c as close to 1 as possible. The multilevel greedy algorithm for construction of the reduced basis space can be summarized as follows:

Algorithm 1 A multilevel greedy algorithm

procedure INITIALIZATION
  Set maximum sparse grid level L, tolerance ϵ_{tol}, $q=N$, take $y^1 \in H(q,N)$;
  Solve (5.21), construct the initial reduced basis spaces V_1, Q_1, G_1, set $N_r=1$.
end procedure

procedure CONSTRUCTION
  for $q=N+1,\ldots,N+L$ **do**
    Construct the set of collocation nodes $H(q,N)$, take $H(q,N)\setminus H(q-1,N)$;
    Solve (6.9) to obtain $y^{N_r+1}=\arg\max_{y\in H(q,N)\setminus H(q-1,N)} \triangle_r(y)$;
    while $\triangle_r(y^{N_r+1}) \geq \epsilon_{tol}$ **do**
      Set $N_r \leftarrow N_r+1$;
      Solve (5.21) at y^{N_r}, update the reduced basis spaces $V_{N_r}, Q_{N_r}, G_{N_r}$;
      Solve (6.9) to obtain $y^{N_r+1}=\arg\max_{y\in H(q,N)\setminus H(q-1,N)} \triangle_r(y)$.
    end while
  end for
end procedure

6.3. A weighted a posteriori error bound. In order to efficiently evaluate a sharp and reliable bound for the reduced basis approximation error, we carry out a residual-based a posteriori error estimate. At first, we reformulate the semi-weak saddle point problem (4.4) as an elliptic problem: for any $y \in \Gamma$, find $\mathrm{u}(y) \in \mathrm{V}$

$$\mathrm{B}(\mathrm{u}(y), \mathrm{v}; y) = \mathrm{F}(\mathrm{v}; y) \quad \forall \mathrm{v} \in \mathrm{V}, \tag{6.13}$$

where the bilinear form $\mathrm{B}(\cdot,\cdot;y) : \mathrm{V}\times\mathrm{V} \to \mathbb{R}$ is given by

$$\begin{aligned} &\mathrm{B}(\mathrm{u}(y), \mathrm{v}; y) = \mathcal{A}\left(\{\mathbf{u}(y), p(y), \mathbf{f}(y)\}, \{\mathbf{v}^a, q^a, \mathbf{g}\}\right) \\ &+ \mathcal{B}(\{\mathbf{v}^a, q^a, \mathbf{g}\}, \{\mathbf{u}^a(y), p^a(y)\}; y) + \mathcal{B}(\{\mathbf{u}(y), p(y), \mathbf{g}(y)\}, \{\mathbf{v}, q\}; y), \end{aligned} \tag{6.14}$$

and the linear functional

$$\mathrm{F}(\mathrm{v}; y) = (\mathbf{u}_d, \mathbf{v}^a) + (p_d, q^a) + (\mathbf{h}(y), \mathbf{v})_{\partial D_N}. \tag{6.15}$$

The bilinear form $\mathrm{B}(\cdot,\cdot;y):\mathrm{V}\times\mathrm{V}\to\mathbb{R}$ can be proven to be continuous and weakly coercive [47] since the bilinear forms $\mathcal{A},\mathcal{B}$ satisfy the conditions in Proposition 3.1, yielding the continuous and weak coercivity constants $\gamma_c(y)$ and $\beta_c(y)$ defined as

$$\gamma_c(y):=\frac{\mathrm{B}(\mathrm{u},\mathrm{v};y)}{||\mathrm{u}||_{\mathrm{V}}||\mathrm{v}||_{\mathrm{V}}}<\infty \text{ and } \beta_c(y):=\inf_{\mathrm{v}\in\mathrm{V}}\sup_{\mathrm{u}\in\mathrm{V}}\frac{\mathrm{B}(\mathrm{u},\mathrm{v};y)}{||\mathrm{u}||_{\mathrm{V}}||\mathrm{v}||_{\mathrm{V}}}>0. \tag{6.16}$$

Therefore, by Babuška theorem [47] we have the following stability estimate

$$||\mathrm{u}(y)||_{\mathrm{V}}\leq\frac{||\mathrm{F}(y)||_{\mathrm{V}'}}{\beta_c(y)}, \tag{6.17}$$

where V' is the dual space of V. By the construction of the finite element approximation in section 5.2 and the reduced basis approximation in section 6.1, we have that the relation (6.16) holds in both finite element space $\mathrm{V}_h=V_h\times Q_h\times G_h\times V_h\times Q_h$ and reduced basis space $\mathrm{V}_r=V_r\times Q_r\times G_r\times V_r\times Q_r$ with constants $\gamma_c^{N_r}(y)\leq\gamma_c^h(y)\leq\gamma_c(y)$ and $\beta_c^{N_r}(y)\geq\beta_c^h(y)\geq\beta_c(y)$. Moreover, the stability estimate (6.17) holds for the finite element solution and the reduced basis solution with the constant $\beta_c^h(y)$ and $\beta_c^{N_r}(y)$, respectively. Let the reduced basis approximation error be defined as $\mathrm{e}(y)=\mathrm{u}_h(y)-\mathrm{u}_r(y)$. To seek an error bound for $\mathrm{e}(y)$, we consider the residual

$$\mathrm{R}(\mathrm{v}_h;y):=\mathrm{F}(\mathrm{v}_h;y)-\mathrm{B}(\mathrm{u}_r(y),\mathrm{v}_h;y)\quad \mathrm{v}_h\in\mathrm{V}_h. \tag{6.18}$$

Noting that $\mathrm{F}(\mathrm{v}_h;y)=\mathrm{B}(\mathrm{u}_h,\mathrm{v}_h;y),\forall\mathrm{v}_h\in\mathrm{V}_h$, we have from (6.18)

$$\mathrm{B}(\mathrm{e}(y),\mathrm{v}_h;y)=\mathrm{R}(\mathrm{v}_h;y)\quad \mathrm{v}_h\in\mathrm{V}_h. \tag{6.19}$$

By the stability estimate (6.17) in the finite element space, we obtain

$$||\mathrm{e}(y)||_{\mathrm{V}_h}\leq\frac{||\mathrm{R}(y)||_{\mathrm{V}_h'}}{\beta_c^h(y)}=:\triangle_r(y). \tag{6.20}$$

Taking the probability density function $\rho:\Gamma\to\mathbb{R}_+$ into account, we replace $\mathcal{E}_r(y)$ in (6.11) by a weighted a posteriori error bound [14] $\triangle_r^\rho(y)=\sqrt{\rho(y)}\triangle_r(y)$. The error bound $\triangle_r^\rho(y)$ assigns high importance at the sample with big probability density, leading to more efficient (using less bases to achieve the same accuracy) evaluation of statistical moments of interest, see [14] for proof and illustrative examples. In order to evaluate the error bound (6.20), we need to compute both the constant $\beta_c^h(y)$ and the norm of the residual $||\mathrm{R}||_{\mathrm{V}_h'}$. For the former, we may apply successive constraint method [24] to compute a lower bound $\beta_c^{LB}(y)\leq\beta_c^h(y)$ with cheap computational cost, or simply use a uniform lower bound $\beta_c^{LB}\leq\beta_c^h(y)$ evaluated at the minimum random viscosity ν_{min} provided that the random coefficient $\nu(y)$ varies in a relatively small range. As for the latter, we turn to an offline-online decomposition procedure in order to reduce computational effort in the many-query context.

6.4. Offline-online decomposition. The offline-online decomposition takes advantage of the affine structure of the data, as given in examples (2.18) and (2.19). If the data are provided in a non-affine structure, e.g. log-normal Karhunen-Loève expansion [34], we may apply a weighted empirical interpolation method to obtain an affine decomposition of the data function at first, see [13] for details and error analysis. Let us assume that the random viscosity and the Neumann boundary condition undergoes, after possibly performing empirical interpolation [4, 13], the following

affine structure

$$\nu(y) = \sum_{n=1}^{N_\nu} \nu_n \theta_n^\nu(y) \text{ and } \mathbf{h}(x,y) = \sum_{n=1}^{N_h} \mathbf{h}_n(x)\theta_n^h(y) \quad \forall (x,y) \in \partial D_N \times \Gamma, \tag{6.21}$$

where $\theta_n^\nu, 1 \le n \le N_\nu$ and $\theta_n^h, 1 \le n \le N_h$ are functions of the random vector $y \in \Gamma$. Let the matrix A_r^y and vector $H_r(y)$ in (6.9) be assembled as

$$A_r^y = \sum_{n=1}^{N_\nu} A_r^n \theta_n^\nu(y) \text{ and } H_r(y) = \sum_{n=1}^{N_h} H_r^n \theta_n^h(y), \tag{6.22}$$

where the deterministic reduced basis matrices $A_r^n, 1 \le n \le N_\nu$ are defined as

$$A_r^n = (\mathcal{Z}_{N_r}^v)^T A_h^n \mathcal{Z}_{N_r}^v \text{ with } (A_h^n)_{ij} = (\nu_n \nabla \boldsymbol{\psi}_i, \nabla \boldsymbol{\psi}_j), 1 \le i,j \le N_v, \tag{6.23}$$

and the deterministic reduced basis vectors $H_r^n, 1 \le n \le N_h$ are defined as

$$H_r^n = (\mathcal{Z}_{N_r}^v)^T H_h^n \text{ with } (H_h^n)_i = (\mathbf{h}_n, \boldsymbol{\psi}_i)_{\partial D_N}, 1 \le i \le N_v. \tag{6.24}$$

Accordingly, we decompose the global matrix of the linear system (6.9) as

$$\mathrm{B}_r^0 = \left(\begin{array}{ccc:cc} M_{v,r} & 0 & 0 & 0 & B_r^T \\ 0 & M_{p,r} & 0 & B_r & 0 \\ 0 & 0 & \alpha M_{g,r} & -M_{c,r}^T & 0 \\ \hdashline 0 & B_r^T & -M_{c,r} & 0 & 0 \\ B_r & 0 & 0 & 0 & 0 \end{array}\right) \tag{6.25}$$

and $\mathrm{B}_r^n, 1 \le n \le N_\nu$ with only the blocks $(4,1),(1,4)$ as A_r^n the other blocks zero. Similarly, we decompose the vector on the right hand side of the linear system (6.9) as $\mathrm{F}_r^0 = (M_{v,r}U_{d,r}, M_{p,r}P_{d,r}, 0, 0, 0)^T$ and $\mathrm{F}_r^n = (0,0,0, M_{n,r}H_r^n, 0)^T, 1 \le n \le N_h$. Thus, the algebraic formulation of the problem (6.13) can be written as: for any $y \in \Gamma$, find $\mathrm{U}_r(y) := (U_r(y), P_r(y), F_r(y), U_r^a(y), P_r^a(y))^T \in \mathbb{R}^{13N_r}$ such that

$$\left(\sum_{n=0}^{N_\nu} \theta_n^\nu(y)\mathrm{B}_r^n\right)\mathrm{U}_r(y) = \sum_{n=0}^{N_h} \theta_n^h(y)\mathrm{F}_r^n. \tag{6.26}$$

Since $\mathrm{B}_r^n, 1 \le n \le N_\nu$ and $\mathrm{F}_r^n, 1 \le n \le N_h$ are independent of y, we can assemble them in offline stage. Given any $y \in \Gamma$, the reduced basis solution can be obtained by solving the linear system (6.26) with at most $O(N_v + N_h)$ operations for assembling and $O((13N_r)^3)$ operations for solve.

As for the evaluation of the residual norm $||\mathrm{R}(y)||_{\mathrm{V}_h'}$, we first seek the Riesz representation [36] of $\mathrm{R}(y)$ as $\hat{\mathrm{e}}(y) \in \mathrm{V}_h$ such that

$$(\hat{\mathrm{e}}(y), \mathrm{v}_h)_{\mathrm{V}_h} = \mathrm{R}(\mathrm{v}_h; y) \quad \forall \mathrm{v}_h \in \mathrm{V}_h, \tag{6.27}$$

so that we have $||\mathrm{R}(y)||_{\mathrm{V}_h'} = ||\hat{\mathrm{e}}(y)||_{\mathrm{V}_h}$. Let $\mathrm{B}_n : \mathrm{V}_h \times \mathrm{V}_h \to \mathbb{R}$ denote the bilinear form defined in the finite element space corresponding to the matrix $\mathrm{B}_r^n, 0 \le n \le N_\nu$ and $\mathrm{F}_n : \mathrm{V}_h \to \mathbb{R}$ the linear functional corresponding to the vector $\mathrm{F}_r^n, 0 \le n \le N_h$, then the residual defined in (6.18) can be decomposed as

$$\mathrm{R}(\mathrm{v}_h; y) = \sum_{n=0}^{N_h} \theta_n^h(y)\mathrm{F}_n(\mathrm{v}_h) - \sum_{n=0}^{N_\nu} \theta_n^\nu(y)\mathrm{B}_n(\mathrm{u}_r, \mathrm{v}_h) \quad \forall \mathrm{v}_h \in \mathrm{V}_h. \tag{6.28}$$

By Riesz representation theorem, we have that there exist $\mathrm{f}_n \in \mathrm{V}_h, 0 \le n \le N_h$ and $\mathrm{b}_n^k \in \mathrm{V}_h, 0 \le n \le N_\nu, 1 \le k \le 13N_r$ such that

$$(\mathrm{f}_n, \mathrm{v}_h)_{\mathrm{V}_h} = \mathrm{F}_n(\mathrm{v}_h) \text{ and } (\mathrm{b}_n^k, \mathrm{v}_h)_{\mathrm{V}_h} = -\mathrm{B}_n(\mathrm{u}_h^k, \mathrm{v}_h) \quad \forall \mathrm{v}_h \in \mathrm{V}_h, \tag{6.29}$$

where we have set the reduced basis solution as $\mathrm{u}_h^k = (\boldsymbol{\psi}_k^v, 0, 0, 0, 0), 1 \le k \le 4N_r$, $\mathrm{u}_h^k = (0, \varphi_{k-4N_r}^p, 0, 0, 0), 4N_r < k \le 6N_r$, $\mathrm{u}_h^k = (0, 0, \boldsymbol{\psi}_{k-6N_r}^g, 0, 0), 6N_r < k \le 7N_r$, $\mathrm{u}_h^k = (0, 0, 0, \boldsymbol{\psi}_{k-7N_r}^v, 0), 7N_r < k \le 11N_r$, $\mathrm{u}_h^k = (0, 0, 0, 0, \varphi_{k-11N_r}^p), 11N_r < k \le 13N_r$, being 0 the vector with length N_v, N_p, N_v, N_v, N_p at the first to fifth argument. Finally, we obtain the norm $||\hat{e}(y)||_{\mathrm{V}_h}$ as

$$\begin{aligned} ||\hat{e}(y)||_{\mathrm{V}_h}^2 = &\sum_{n=1}^{N_h}\sum_{n'=1}^{N_h} \theta_n^h(y)(\mathrm{f}_n, \mathrm{f}_{n'})_{\mathrm{V}_h}\theta_{n'}^h(y) \\ &+ 2\sum_{n=1}^{N_h}\sum_{n'=1}^{N_\nu}\sum_{k=1}^{13N_r} \theta_n^h(y)(\mathrm{f}_n, \mathrm{b}_{n'}^k)_{\mathrm{V}_h}(\mathrm{u}_r)_k\theta_{n'}^\nu(y) \\ &+ \sum_{n=1}^{N_h}\sum_{n'=1}^{N_\nu}\sum_{k=1}^{13N_r}\sum_{k'=1}^{13N_r} \theta_n^\nu(y)(\mathrm{u}_r)_k(\mathrm{b}_n^k, \mathrm{b}_{n'}^{k'})_{\mathrm{V}_h}(\mathrm{u}_r)_{k'}\theta_{n'}^\nu(y), \end{aligned} \tag{6.30}$$

where $(\mathrm{f}_n, \mathrm{f}_{n'})_{\mathrm{V}_h}, 1 \le n, n' \le N_h$, $(\mathrm{f}_n, \mathrm{b}_{n'}^k)_{\mathrm{V}_h}, 1 \le n \le N_h, 1 \le n' \le N_\nu, 1 \le k \le 13N_r$ and $(\mathrm{b}_n^k, \mathrm{b}_{n'}^{k'})_{\mathrm{V}_h}, 1 \le n, n' \le N_\nu, 1 \le k, k' \le 13N_r$ are independent of y and can be computed and stored in the offline stage, while in the online stage, we only need to assemble the formula (6.30) by $O((N_h + 13N_rN_\nu)^2)$ operations. Recall that N_h and N_ν are the number of affine terms of the random Neumann boundary condition and the viscosity, and N_r is the number of selected samples in the construction of reduced basis space, leading to fast evaluation of the error bound as they are small.

7. Error estimates. The global error of the numerical approximation presented in sections 5 and 6 comprises three components: the stochastic collocation approximation error [2, 34, 33], the finite element approximation error [36, 35], and the weighted reduced basis approximation error [5, 10, 14], which have been analyzed individually in different contexts. In the following, we provide individual error estimate as well as a global error estimate in the context of the constrained optimization problem (2.22).

7.1. Stochastic collocation approximation error. The error of stochastic collocation approximation of the optimal solution depends on the stochastic regularity of the latter. We consider the case that Γ is bounded, however similar results can be obtained in the same way for unbounded Γ as in [2]. Let the complex region $\Sigma(\Gamma; \tau)$ be defined as

$$\Sigma(\Gamma; \tau) := \{z \in \Sigma : \exists\, y \in \Gamma \text{ such that } |z_n - y_n| \le \tau_n, 1 \le n \le N\}, \tag{7.1}$$

where Σ has been defined in (4.21); $\tau = (\tau_1, \ldots, \tau_N)$ with each element taking the largest possible value ($\tau_n = 1/(rr_n), 1 \le n \le N$). Thanks to the analytic regularity obtained in Theorem 4.2, we have the following a priori error estimate for tensor-product stochastic collocation approximation of the optimal solution $\mathrm{u} : \Gamma \to \mathrm{V}$ (recall that $\mathrm{u} = \{\mathbf{u}, p, \mathbf{f}, \mathbf{u}^a, p^a\}$ and $\mathrm{V} = V \times Q \times G \times V \times Q$)

$$\mathcal{E}_s := ||\mathrm{u} - \mathrm{u}_s||_{C(\Gamma;\mathrm{V})} \le \sum_{n=1}^{N} C_n^{\mathbf{i}} \exp(-(m(i_n) - 1)r_n), \tag{7.2}$$

where we denote $\mathrm{u}_s = \mathcal{I}_{\mathbf{i}}\mathrm{u}$; the constants $C_n^{\mathbf{i}}, 1 \le n \le N$ are bounded by [2, 10]

$$C_n^{\mathbf{i}} \le (1 + \Lambda(m(i_n)))C_n, \text{ being } C_n := \frac{2}{e^{r_n} - 1}\left(\max_{z\in\Sigma(\Gamma;\tau)} ||\mathrm{u}(z)||_{\mathrm{V}}\right) \tag{7.3}$$

with Lebesgue constant $\Lambda(m) \le 1 + (2/\pi)\log(m+1)$, and convergence rate

$$r_n = \log\left(\frac{2\tau_n}{|\Gamma_n|} + \sqrt{1 + \frac{4\tau_n^2}{|\Gamma_n|^2}}\right) > 1, \quad 1 \le n \le N. \tag{7.4}$$

Remark 7.1. *In the case of unbounded Γ, the convergence rate has been obtained as $r_n = \tau_n/\delta_n, 1 \le n \le N$ with δ_n depending on the decay of the probability density function at infinity, e.g. $\delta_n = 1$ for normal density function, see details in [2].*

As for the error of isotropic sparse grid Smolyak interpolation (5.3) with Gauss-abscissas, the following error estimate can be proved via (7.2) [34] (denote $\mathrm{u}_s = \mathcal{S}_q\mathrm{u}$)

$$\mathcal{E}_s := ||\mathrm{u} - \mathrm{u}_s||_{C(\Gamma;\mathrm{V})} \le C_s N_q^{-r}, \tag{7.5}$$

where N_q represents the number of collocation nodes, C_s depends on Lebesgue constant but not on N_q (see [34, 10] for more explicit expression), r is such that (see [10])

$$r \ge \frac{e\log(2)\min\{r_n, 1 \le n \le N\}}{3 + \log(N)}. \tag{7.6}$$

As for the error of anisotropic sparse grid Smolyak interpolation (5.10) based on Gauss-abscissas, we have the following estimate [33] (denote $\mathrm{u}_s = \mathcal{S}_q^{\boldsymbol{\alpha}}\mathrm{u}$)

$$\mathcal{E}_s^{\boldsymbol{\alpha}} := ||\mathrm{u} - \mathrm{u}_s||_{C(\Gamma;\mathrm{V})} \le C_s^{\boldsymbol{\alpha}} N_q^{-r(\boldsymbol{\alpha})}, \tag{7.7}$$

where $C_s^{\boldsymbol{\alpha}}$ depends on Lebesgue constant but not on N_q and the algebraic convergence rate $r(\boldsymbol{\alpha})$ is defined as

$$r(\boldsymbol{\alpha}) = \frac{e\log(2)\alpha_{min}}{2\log(2) + \sum_{n=1}^{N}\frac{\alpha_{min}}{\alpha_n}}, \tag{7.8}$$

being $\alpha_{min} = \min_{1\le n\le N}\alpha_n$ with the choice $\alpha_n = r_n/2, 1 \le n \le N$, with r_n defined in (7.4). Moreover, the error of the expectation of the stochastic optimal solution evaluated by isotropic or anisotropic sparse grid Smolyak formula is bounded by [34, 10]

$$\mathcal{E}_s^e := ||\mathbb{E}[\mathrm{u}] - \mathbb{E}[\mathrm{u}_s]||_{\mathrm{V}} \le ||\mathrm{u} - \mathrm{u}_s||_{L^2_\rho(\Gamma;\mathrm{V})} \le C_s^e N_q^{-r(\boldsymbol{\alpha})}, \tag{7.9}$$

where C_s^e is a constant independent of both Lebesgue constant and N_q, see [10].

7.2. Finite element approximation error. Recall that the bilinear forms $\mathcal{A}$ and $\mathcal{B}$ of the finite element problem (5.14) satisfy the conditions of Proposition 3.1 in the finite element space V_h^k, Q_h^m, G_h^l with the choice of Taylor-Hood elements. More explicitly, the finite element constants corresponding to those stated in the conditions of Proposition 3.1 can be bounded by

$$\gamma_h \le 1, \epsilon_h \ge \frac{1}{3}\min\left\{\frac{\alpha_a^2\alpha}{2C_P^2}, \frac{1}{2}\right\}, \delta_h \le \max\{\nu_{max}, \gamma_b, 1/\sqrt{\alpha}\}, \beta_h \ge \beta, \tag{7.10}$$

being the constants $\alpha_a, \alpha, C_P, \nu_{max}, \gamma_b, \beta$ presented in Lemma 3.2. Therefore, by Brezzi theorem [36, 35], we have the following estimate for the error $\mathcal{E}_h$ of the finite element approximation to solution of the semi-weak saddle point problem (4.4):

$$\begin{aligned}
\mathcal{E}_h(y) &:= ||\mathrm{u}(y) - \mathrm{u}_h(y)||_{\mathrm{V}} \\
&\leq C_1^h \inf_{\{\mathbf{v}_h, q_h, \mathbf{g}_h\} \in V_h^k \times Q_h^m \times G_h^l} ||\{\mathbf{u}(y), p(y), \mathbf{f}(y)\} - \{\mathbf{v}_h, q_h, \mathbf{g}_h\}||_{V \times Q \times G} \\
&+ C_2^h \inf_{\{\mathbf{v}_h^a, q_h^a\} \in V_h^k \times Q_h^m} ||\{\mathbf{u}^a(y), p^a(y)\} - \{\mathbf{v}_h^a, q_h^a\}||_{V \times Q} \qquad (7.11) \\
&= O(h^k) \left(C_1^h (||\mathbf{u}(y)||_{k+1} + ||p(y)||_k + \sqrt{\alpha}||\mathbf{f}(y)||_{k+1}) \right) \\
&+ O(h^k) \left(C_2^h (||\mathbf{u}^a(y)||_{k+1} + ||p^a(y)||_k) \right),
\end{aligned}$$

where we have chosen $m = k - 1$ and $l = k$; the constants C_1^h and C_2^h are given by

$$C_1^h = \left(1 + \frac{\gamma_h}{\epsilon_h}\right)\left(1 + \frac{\gamma_h}{\beta_h}\right)\left(1 + \frac{\delta_h}{\beta_h}\right) \text{ and } C_2^h = 1 + \frac{\delta_h}{\epsilon_h} + \frac{\delta_h}{\beta_h} + \frac{\gamma_h \delta_h}{\epsilon_h \beta_h}. \qquad (7.12)$$

Remark 7.2. *Equivalently, we may formulate the semi-weak saddle point finite element problem (5.14) as a weakly coercive elliptic problem and apply Babuška theorem to obtain similar finite element error estimate.*

7.3. Reduced basis approximation error. In addition to the a posteriori error bound $\triangle_r$ obtained in section 6.3, we present some results about a priori error estimate for reduced basis approximation following those obtained in [14, 5].

Thanks to the analytic regularity in Theorem 4.2, we have a priori error estimate for reduced basis solution of (6.1) when $\Gamma \subset \mathbb{R}$ [14]

$$\mathcal{E}_r := ||\mathrm{u}_h - \mathrm{u}_r||_{C(\Gamma;\mathrm{V})} \leq C_r \exp(-rN_r), \qquad (7.13)$$

where r is defined as in (7.4) for a single dimension, the constant C_r is bounded by

$$C_r \leq C \left(\max_{z \in \Sigma(\Gamma;\tau)} ||\mathrm{u}(z)||_{\mathrm{V}} \right), \qquad (7.14)$$

being C a constant independent of the number of reduced bases N_r, see [14] for details.

As in multidimensional case, the error estimate has been obtained via Kolmogorov N-width defined in an abstract Hilbert space X as [5]

$$d_N(\Gamma; X) := \inf_{X_N \subset X} \sup_{y \in \Gamma} \inf_{w_N \in X_N} ||v(y) - w_N||_X, \qquad (7.15)$$

where X_N is a N-dimensional subspace of X. We have the following result for $\mathcal{E}_r$ [5]: suppose that there exists $M > 0$ such that $d_0(\Gamma) \leq M$; moreover, suppose that there exist two positive constants $c_1 > 0, c_2 > 0$, such that

$$\text{if } d_{N_r}(\Gamma; \mathrm{V}_h) \leq M \exp(-c_1 N_r^{c_2}) \text{ then } \mathcal{E}_r \leq c_5 M \exp(-c_3 N_r^{c_4}), \qquad (7.16)$$

where $c_4 = c_2/(c_2 + 1)$, $c_3 > 0, c_5 > 0$ depend only on c_1, c_2 and $c_6 > 0$, which measures the sharpness of the reduced basis error bound in (6.20), i.e.

$$c_6 \triangle_r(y) \leq ||\mathrm{u}_h(y) - \mathrm{u}_r(y)||_{\mathrm{V}}. \qquad (7.17)$$

Remark 7.3. *The result (7.16) implies that whenever the error of the best possible approximation decays exponentially, the reduced basis error also enjoys an exponential decay with rate depending on the sharpness of the greedy algorithm (7.17).*

7.4. Global error estimate. With the individual error estimate presented above, we obtain the global error estimate in the following theorem.

THEOREM 7.1. *Under Assumption 1, 2 and 3, for any given $y \in \Gamma$, by finite element approximation and reduced basis approximation we have*

$$||\mathrm{u}(y) - \mathrm{u}_r(y)||_{\mathrm{V}} \leq \mathcal{E}_h(y) + \mathcal{E}_r(y). \tag{7.18}$$

Moreover, the error for evaluation of the expectation using stochastic collocation method, finite element method and weighted reduced basis method can be bounded by

$$||\mathbb{E}[\mathrm{u}] - \mathbb{E}[\mathrm{u}_r]||_{\mathrm{V}} \leq \mathcal{E}_s^e + \max_{y \in H_{\boldsymbol{\alpha}}(q,N)} \mathcal{E}_h(y) + \max_{y \in H_{\boldsymbol{\alpha}}(q,N)} \mathcal{E}_r(y), \tag{7.19}$$

where $\boldsymbol{\alpha} = \mathbf{1}$ when using the isotropic sparse grid stochastic collocation method.

Proof. The proof is straightforward by applying triangular inequality as follows:

$$||\mathrm{u}(y) - \mathrm{u}_r(y)||_{\mathrm{V}} \leq ||\mathrm{u}(y) - \mathrm{u}_h(y)||_{\mathrm{V}} + ||\mathrm{u}_h(y) - \mathrm{u}_r(y)||_{\mathrm{V}} \leq \mathcal{E}_h(y) + \mathcal{E}_r(y). \tag{7.20}$$

Similarly, we have the error estimate for the expectation of the optimal solution as

$$\begin{aligned} ||\mathbb{E}[\mathrm{u}] - \mathbb{E}[\mathrm{u}_r]||_{\mathrm{V}} &\leq ||\mathrm{u} - \mathrm{u}_r||_{L^2_\rho(\Gamma;\mathrm{V})} \\ &\leq ||\mathrm{u} - \mathrm{u}_s||_{L^2_\rho(\Gamma;\mathrm{V})} + ||\mathrm{u}_s - \mathrm{u}_h||_{L^2_\rho(\Gamma;\mathrm{V})} + ||\mathrm{u}_h - \mathrm{u}_r||_{L^2_\rho(\Gamma;\mathrm{V})} \\ &\leq \mathcal{E}_s^e + \max_{y \in H_{\boldsymbol{\alpha}}(q,N)} \mathcal{E}_h(y) + \max_{y \in H_{\boldsymbol{\alpha}}(q,N)} \mathcal{E}_r(y). \end{aligned} \tag{7.21}$$

We remark that $\mathcal{E}_r(y)$ is bounded by $\triangle_r(y)$, explicitly computed at $y \in H_{\boldsymbol{\alpha}}(q,N)$. ∎

8. Numerical experiments. In this section, we perform two numerical experiments in testing reduced basis approximation error and stochastic collocation approximation error with sparse grid techniques in isotropic and anisotropic settings. The aim is to demonstrate the efficiency of the proposed reduced basis method in solving constrained optimization problem (2.22). Numerical examples for verifying finite element approximation error in a similar context can be found in [11].

We consider a two dimensional physical domain $D = (0,1)^2$. The observation data is set as in [22], $\mathbf{u}_d = (u_{d1}, u_{d2})$ and $p_d = 0$, where $u_{d1}(x) = \partial_{x_2}(\phi(x_1)\phi(x_2))/10$ and $u_{d2}(x) = -\partial_{x_1}(\phi(x_1)\phi(x_2))/10$ with $\phi(\xi) = (1 - \cos(0.8\pi\xi))(1-\xi)^2, \xi \in [0,1]$. The random viscosity ν is given as in (2.18) which can be transformed as

$$\nu(y^\nu) = \frac{1}{2}\sum_{n=0}^{N_\nu} \nu_n + \frac{1}{2N_\nu}\sum_{n=1}^{N_\nu} (\nu_n - \nu_0) y_n^\nu, \tag{8.1}$$

where $y^\nu \in \Gamma_\nu = [-1,1]^{N_\nu}$ corresponding to N_ν uniformly distributed random variables. We set $\nu_0 = 0.01$, $\nu_n = \nu_0/2^n$ and use $N_\nu = 3$ for both the isotropic and anisotropic tests without loss of generality. Homogeneous Dirichlet boundary condition is imposed on the upper, lower and left edge. Random Neumann boundary condition is imposed on the right edge as given in (2.19) on the Neumann boundary, more explicitly, we set $\mathbf{h}(x, y^h) = (h_1(x_2, y^h), 0)$ with

$$h_1(x_2, y^h) = \frac{1}{10}\left(\left(\frac{\sqrt{\pi}L}{2}\right)^{1/2} y_1^h + \sum_{n=1}^{N_h} \sqrt{\lambda_n}\left(\sin(n\pi x_2) y_{2n}^h + \cos(n\pi x_2) y_{2n+1}^h\right)\right), \tag{8.2}$$

which comes from truncation of Karhunen-Loève expansion of a Gauss covariance field with correlation length $L = 1/16$ [34]; the eigenvalues $\lambda_n, 1 \leq n \leq N_h$ are given by

$$\lambda_n = \sqrt{\pi} L \exp\left(-(n\pi L)^2/4\right); \tag{8.3}$$

$y_n^h, 1 \leq n \leq 2N_h + 1$ are uncorrelated with zero mean and unit variance, which are independent of y^ν. Therefore, the random inputs are $y = (y^\nu, y^h)$, living in $N = N_\nu + 2N_h + 1$ dimensional probability space. As for the specification of the finite element approximation, we use P1 element for pressure space and P2 element for velocity and control space with 1342 elements in total.

8.1. Isotropic case. In the first experiment, we set $y_n^h, 1 \leq n \leq 2N_h + 1$ with $N_h = 3$ as independent standard normal random variables (thus the total stochastic dimension $N = 10$), and apply isotropic sparse grid stochastic collocation method with Gauss-Legendre abscissa for the collocation of y^ν and Gauss-Hermite abscissa for the collocation of y^h. In the multilevel greedy algorithm 1, we set the tolerance $\epsilon_{tol} = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}$, and the interpolation level $q - N = 0, 1, 2, 3$ in the isotropic sparse grid Smolyak formula (5.3). A uniformly lower bound of the inf-sup constant $\beta_c^{LB} = 0.1436$ is used since the fluctuation or variance of ν is small compared to its mean value. The results for reduced basis construction is reported in Table 8.1. The number of collocation nodes in each level is shown in the second row; the number of selected samples as new bases in each level and the samples whose weighted error bound $\triangle_r^\rho$ is larger than the tolerance ϵ_{tol}, thus potential as new bases are shown in the 3rd-6th lines, from which we can see that the number of reduced bases is much less than that of collocation nodes. For example, with the smallest tolerance $\epsilon_{tol} = 10^{-5}$, we only need $1, 10, 22, 14$ new bases in each level, respectively, resulting in 47 bases in total out of 1581 collocation nodes. Since the number of samples as potential bases is also small (216 in total), the computational cost for sample selection in the construction of reduced basis space is negligible compared to the full solve of the finite element problem (5.21), especially for large scale problems featuring a small mesh size h.

TABLE 8.1
The number of samples selected by multilevel greedy algorithm 1 with different tolerance ϵ_{tol} in each of the sparse grid level; the value in $(\cdot)$ reports the number of samples potential as new bases.

tolerance \ level	$q - N = 0$	$q - N = 1$	$q - N = 2$	$q - N = 3$	in total
# nodes	1	21	221	1581	1581
$\epsilon_{tol} = 10^{-1}$	1 (1)	6 (14)	1 (21)	0 (0)	8 (36)
$\epsilon_{tol} = 10^{-2}$	1 (1)	8 (20)	7 (80)	4 (28)	20 (129)
$\epsilon_{tol} = 10^{-3}$	1 (1)	9 (20)	13 (86)	5 (62)	28 (169)
$\epsilon_{tol} = 10^{-4}$	1 (1)	9 (20)	18 (90)	9 (67)	37 (178)
$\epsilon_{tol} = 10^{-5}$	1 (1)	10 (20)	22 (90)	14 (105)	47 (216)

Fig. 8.1 (left) displays the weighted error bound $\triangle_r^\rho$ and the true error of the reduced basis approximation in each level of the construction, from which we can see that the error bound is accurate and relatively sharp, providing good estimate of the true error with cheap computation. On the right of Fig. 8.1 we plot the expectation error (in $L_\rho^2(\Gamma; \mathrm{V})$ norm) of the reduced basis approximation using quadrature formula based on sparse grid of different levels, where the expectation error is defined as

$$\text{exp. error} = |||\mathrm{u}||_{L_\rho^2(\Gamma;\mathrm{V})} - ||\mathrm{u}_{s,r}||_{L_\rho^2(\Gamma;\mathrm{V})}| = |(\mathbb{E}[||\mathrm{u}||_{\mathrm{V}}^2])^{1/2} - (\mathbb{E}[||\mathrm{u}_{s,r}||_{\mathrm{V}}^2])^{1/2}|. \tag{8.4}$$

Note that the "true" value of $||\mathrm{u}||_{L^2_\rho(\Gamma;\mathrm{V})}$ is approximated by the finite element solution u_h computed at the deepest level $q-N=3$. From this figure, different accuracy with different ϵ_{tol} can be observed, implying that decreasing tolerance for the construction of the reduced basis space results in more accurate evaluation of statistics of the solution. How to balance the reduced basis approximation error (by choice of ϵ_{tol}) and the sparse grid quadrature error (by choice of $q-N$) is subject to further investigation.

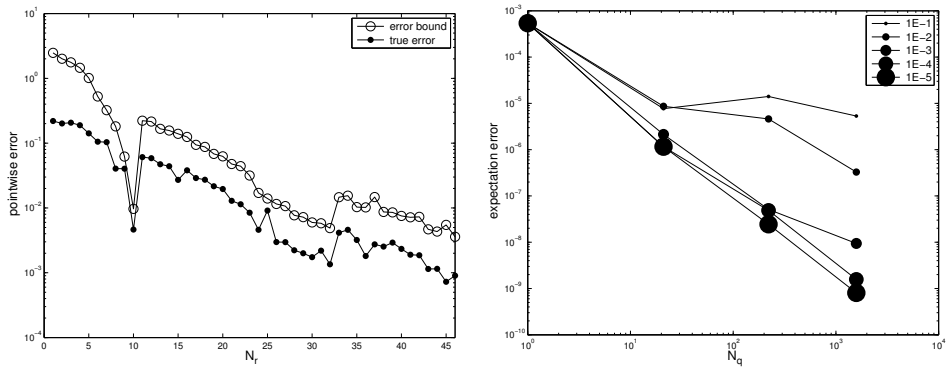

FIG. 8.1. *Left, weighted error bound $\triangle_r^\rho$ and true error of the reduced basis approximation at the selected samples; right, expectation error at different levels with different tolerance ϵ_{tol}.*

8.2. Anisotropic case. In the second experiment, we solve the constrained optimization problem (2.22) in high dimensional probability space by combination of the anisotropic sparse grid techniques and the multilevel weighted reduced basis method. We set $y_n^h, 1 \le n \le N_h$ in (8.2) with $N_h = 3, 8, 13, 18, 48$ as uniformly distributed random variables, thus leading to $N = 10, 20, 30, 40, 100$ stochastic dimensions in total. The weight parameter $\boldsymbol{\alpha}$ is chosen a priori according to [33] in the following conservative way

$$\alpha_n = \frac{1}{2}\log\left(1+\frac{2\tau_n}{|\Gamma_n|}\right), \text{ with } \tau_n = \frac{1}{4\sqrt{\lambda_n}}, \quad 1 \le n \le N_h. \tag{8.5}$$

We remark that for a more general random field where $\boldsymbol{\alpha}$ is difficult to be obtained from a priori estimate, we may use a posteriori estimate by fitting a empirical convergence rate in each dimension [33], or use dimension-adaptive approach which determines the weight automatically [18]. The sparse grid level is chosen as $q-N = 0, 1, 2, 3, 4$. As for the tolerance for the construction of the reduced basis space, we use $\epsilon_{tol} = 10^{-5}$. The results for the construction of the reduced basis space with different dimension N and different sparse grid level $q-N$ (results for $q-N=0$ are the same as in Table 8.1, thus omitted here) are presented in Table 8.2. Similar conclusion as for results in the isotropic case in Table 8.1 can be drawn for those in the anisotropic case in Table 8.2. For example, when $N=40$, only 97 samples out of 40479 are used for the construction of the reduced basis space, thus resulting in only 97 full solve the finite element problem (5.21) instead of 40479, which considerably reduces the total computational cost. This observation holds even in the 100 dimensional case. Moreover, the number of nodes of sparse grid and the number of reduced bases increase as the dimension increase when N is small, see the change from 10 to 40. However, they stay almost the same when N becomes large, see the change from

40 to 100, which indicates that out of 100 random variables, the first 40 play the most important role on the impact of the stochastic optimal solution when we set sparse grid level at $q - N = 4$.

TABLE 8.2
The number of samples selected by multilevel greedy algorithm 1 in each of the sparse grid level with different dimensions; the value in $(\cdot)$ reports the number of samples potential as new bases.

dimension \ level	$q - N = 1$	$q - N = 2$	$q - N = 3$	$q - N = 4$	in total
$N = 10$	5 (10)	13 (40)	19 (85)	10 (100)	48 (236)
# nodes	11	71	401	2141	2141
$N = 20$	5 (10)	21 (60)	36 (205)	15 (204)	78 (480)
# nodes	11	91	1021	12121	12121
$N = 40$	5 (10)	25 (92)	47 (397)	19 (432)	97 (932)
# nodes	11	123	2381	40769	40769
$N = 100$	5 (10)	25 (92)	47 (397)	19 (436)	97 (936)
# nodes	11	123	2393	41349	41349

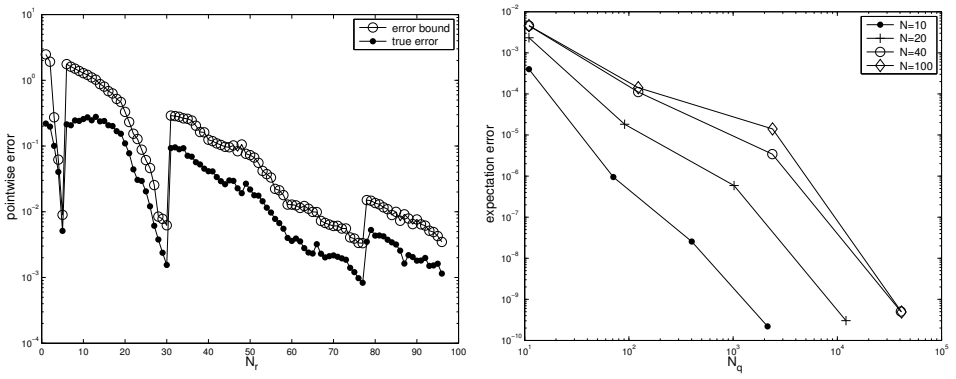

FIG. 8.2. *Weighted error bound $\triangle_r^\rho$ and true error of the reduced basis approximation at the selected samples in the case of stochastic dimension $N = 10$ (left) and $N = 100$ (right).*

On the left of Fig. 8.2, we plot the weighted a posteriori error bound $\triangle_r^\rho$ and the true error of the reduced basis approximation at each sparse grid level with stochastic dimension $N = 100$. We can observe that the error bound is indeed accurate and sharp for the high dimensional case, especially when the reduced basis space become large. The right of Fig. 8.2 depicts the expectation error at different sparse grid level. We show the expectation error with the "true" expectation for each stochastic dimension computed the same as in the isotropic sparse grid case, from which we can see that the expectation error converges with an algebraic rate that verifies the error estimate in section 7. Moreover, the error becomes very small at around 4×10^4 nodes for the 100 dimensional problem by anisotropic sparse grid technique, which would need around 7×10^7 nodes for isotropic sparse grid technique at the same sparse grid level $q - N = 4$. Furthermore, we can observe that no "plateau" (flattening) of expectation error appears as in Fig. 8.1, demonstrating that the multilevel reduced basis method is very efficient in producing the accurate statistics of the stochastic optimal solution even when the number of the reduced bases shown in Table 8.2 remains critically small (around 97 for high dimensions).

9. Concluding remarks. In this paper we studied the mathematical properties of an optimal control problem constrained by stochastic Stokes equations and developed a computational strategy by using sparse grid techniques and the model order reduction approach. The existence and uniqueness of the stochastic optimal solution was proved by establishing the equivalence between the constrained optimization problem and an stochastic saddle point problem. Moreover, we obtained some stochastic regularity results of the optimal solution in the probability space under some mild assumptions on the random input data. In the fully discretized problem, we used finite element approximation in the deterministic space and stochastic collocation approximation in the probability space, and proposed a multilevel and weighted reduced basis method in order to reduce the computational effort in the many-query context, for which a global error estimate was carried out. This computational approach was proven to be very efficient by two numerical experiments, especially for high dimensional and large-scale problems requiring a large number of samples and heavy computational cost for a full solve of the optimization problem at each sample. Further study on more general statistical cost functional, adaptive scheme to balance various computational errors and applications to practical flow control problems are ongoing [12].

Acknowledgement: We acknowledge the use of the Matlab packages *MLife* previously developed by Prof. Fausto Saleri from MOX, Politecnico di Milano and *rbMIT* developed by Prof. Anthony Patera and his coworkers from Massachusetts Institute of Technology. This work is partially supported by FNS 200021_141034. G. Rozza acknowledges NOFYSAS exellence grant of SISSA.

REFERENCES

[1] M. Arroyo, L. Heltai, D. Millán, and A. DeSimone. Reverse engineering the euglenoid movement. *Proceedings of the National Academy of Sciences*, 109(44):17874–17879, 2012.

[2] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 45(3):1005–1034, 2007.

[3] I. Babuška, R. Tempone, and G.E. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM Journal on Numerical Analysis*, 42(2):800–825, 2005.

[4] M. Barrault, Y. Maday, N.C. Nguyen, and A.T. Patera. An empirical interpolation method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathematique, Analyse Numérique*, 339(9):667–672, 2004.

[5] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Convergence rates for greedy algorithms in reduced basis methods. *SIAM Journal of Mathematical Analysis*, 43(3):1457–1472, 2011.

[6] P.B. Bochev and M.D. Gunzburger. Least-squares finite-element methods for optimization and control problems for the Stokes equations. *Computers & Mathematics with Applications*, 48(7):1035–1057, 2004.

[7] P.B. Bochev and M.D. Gunzburger. *Least-squares finite element methods*, volume 166. Springer, 2009.

[8] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from lagrangian multipliers. *RAIRO. Modélisation Mathématique et Analyse Numérique*, 8(2):129–151, 1974.

[9] C. Canuto, M.Y. Hussaini, A. Quarteroni, and T.A. Zang. *Spectral Methods: Fundamentals in Single Domains*. Springer, 2006.

[10] P. Chen, A. Quarteroni, and G. Rozza. Comparison of reduced basis method and collocation method for stochastic elliptic problems. *accepted in Journal of Scientific Computing, EPFL, MATHICSE Report 34*, 2012.

[11] P. Chen, A. Quarteroni, and G. Rozza. Stochastic optimal Robin boundary control problems of advection-dominated elliptic equations. *accepted in SIAM Journal on Numerical Analysis,*

EPFL, MATHICSE Report 23, 2012.

[12] P. Chen, A. Quarteroni, and G. Rozza. Fast solver for optimal control problems constrained by stochastic PDE with general statistical observation data. *in preparation*, 2013.

[13] P. Chen, A. Quarteroni, and G. Rozza. A weighted empirical interpolation method: A priori convergence analysis and applications. *EPFL, MATHICSE Report 05, submitted*, 2013.

[14] P. Chen, A. Quarteroni, and G. Rozza. A weighted reduced basis method for elliptic partial differential equations with random input data. *EPFL, MATHICSE Report 04, submitted*, 2013.

[15] R. Durrett. *Probability: theory and examples*. Cambridge University Press, 2010.

[16] H. Elman and Q. Liao. Reduced basis collocation methods for partial differential equations with random coefficients. *SIAM/ASA Journal on Uncertainty Quantification*, 1(1):192–217, 2013.

[17] L. Formaggia, A. Quarteroni, and A. Veneziani. *Cardiovascular Mathematics: Modeling and simulation of the circulatory system*, volume 1. MS&A, Springer, 2009.

[18] T. Gerstner and M. Griebel. Dimension–adaptive tensor–product quadrature. *Computing*, 71(1):65–87, 2003.

[19] R. Glowinski and J.L. Lions. *Exact and approximate controllability for distributed parameter systems*. Cambridge University Press, 1996.

[20] M.D. Gunzburger. *Perspectives in flow control and optimization*, volume 5. SIAM, 2003.

[21] M.D. Gunzburger, H.C. Lee, and J. Lee. Error estimates of stochastic optimal neumann boundary control problems. *SIAM Journal on Numerical Analysis*, 49:1532–1552, 2011.

[22] M.D. Gunzburger and S. Manservisi. Analysis and approximation of the velocity tracking problem for Navier–Stokes flows with distributed control. *SIAM Journal on Numerical Analysis*, 37(5):1481–1512, 2000.

[23] L.S. Hou, J. Lee, and H. Manouzi. Finite element approximations of stochastic optimal control problems constrained by stochastic elliptic PDEs. *Journal of Mathematical Analysis and Applications*, 384(1):87–103, 2011.

[24] D.B.P. Huynh, D.J. Knezevic, Y. Chen, J.S. Hesthaven, and A.T. Patera. A natural-norm successive constraint method for inf-sup lower bounds. *Computer Methods in Applied Mechanics and Engineering*, 199(29):1963–1975, 2010.

[25] Junseok K. Phase-field models for multi-component fluid flows. *Communications in Computational Physics*, 12(3):613–661, 2012.

[26] M. Kärcher and M. Grepl. A certified reduced basis method for parametrized elliptic optimal control problems. *accepted in ESAIM: Control, Optimisation and Calculus of Variations*, 2013.

[27] D.P. Kouri, D. Heinkenschloos, M. Ridzal, and B.G. Van Bloemen Waanders. A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty. *Preprint ANL/MCS-P3035-0912*, 2012.

[28] Karl Kunisch and Stefan Volkwein. Proper orthogonal decomposition for optimality systems. *ESAIM: Mathematical Modelling and Numerical Analysis*, 42(1):1, 2008.

[29] J.L. Lions. *Optimal control of systems governed by partial differential equations*. Springer, 1971.

[30] A. Manzoni. *Reduced models for optimal control, shape optimization and inverse problems in haemodynamics*. PhD thesis, EPFL, 2012.

[31] F. Negri, A. Quarteroni, G. Rozza, and A. Manzoni. Reduced basis method for parametrized saddle point problems. *in preparation*, 2013.

[32] F. Negri, G. Rozza, A. Manzoni, and A. Quarteroni. Reduced basis method for parametrized elliptic optimal control problems. *accepted in SIAM Journal on Scientific Computing, EPFL, MATHICES Report 40*, 2012.

[33] F. Nobile, R. Tempone, and C.G. Webster. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2411–2442, 2008.

[34] F. Nobile, R. Tempone, and C.G. Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2309–2345, 2008.

[35] A. Quarteroni. *Numerical Models for Differential Problems*. Springer, MS & A, vol 2, 2009.

[36] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*. Springer, 1994.

[37] T. Rees and A.J. Wathen. Preconditioning iterative methods for the optimal control of the Stokes equations. *SIAM Journal on Scientific Computing*, 33(5):2903–2926, 2011.

[38] E. Rosseel and G.N. Wells. Optimal control with stochastic PDE constraints and uncertain controls. *Computer Methods in Applied Mechanics and Engineering*, 213C216(0):152 –

167, 2012.

[39] G. Rozza, D.B.P. Huynh, and A. Manzoni. Reduced basis approximation and a posteriori error estimation for Stokes flows in parametrized geometries: roles of the inf-sup stability constants. *Numerische Mathematik*, pages 1–38, 2013.

[40] G. Rozza, D.B.P. Huynh, and A.T. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Archives of Computational Methods in Engineering*, 15(3):229–275, 2008.

[41] G. Rozza and K. Veroy. On the stability of the reduced basis method for stokes equations in parametrized domains. *Computer methods in applied mechanics and engineering*, 196(7):1244–1260, 2007.

[42] J. Schöberl and W. Zulehner. Symmetric indefinite preconditioners for saddle point problems with applications to PDE-constrained optimization problems. *SIAM Journal on Matrix Analysis and Applications*, 29(3):752–773, 2007.

[43] C. Schwab and R. A. Todor. Karhunen-Loève approximation of random fields by generalized fast multipole methods. *Journal of Computational Physics*, 217(1):100–122, 2006.

[44] H. Tiesler, R.M. Kirby, D. Xiu, and T. Preusser. Stochastic collocation for optimal control problems with stochastic PDE constraints. *SIAM Journal on Control and Optimization*, 50(5):2659–2682, 2012.

[45] F. Tröltzsch. *Optimal control of partial differential equations: theory, methods, and applications*, volume 112. American Mathematical Society, 2010.

[46] D. Xiu and J.S. Hesthaven. High-order collocation methods for differential equations with random inputs. *SIAM Journal on Scientific Computing*, 27(3):1118–1139, 2005.

[47] J. Xu and L. Zikatanov. Some observations on Babuška and Brezzi theories. *Numerische Mathematik*, 94(1):195–202, 2003.

MOX Technical Reports, last issues

Dipartimento di Matematica "F. Brioschi",
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

51/2013 Chen, P.; Quarteroni, A.
Weighted reduced basis method for stochastic optimal control problems with elliptic PDE constraint

52/2013 Chen, P.; Quarteroni, A.; Rozza, G.
Multilevel and weighted reduced basis method for stochastic optimal control problems constrained by Stokes equations

47/2013 Chkifa, A.; Cohen, A.; Migliorati, G.; Nobile, F.; Tempone, R.
Discrete least squares polynomial approximation with random evaluations - application to parametric and stochastic elliptic PDEs

50/2013 Antonietti, P.F.; Verani, M.; Zikatanov, L.
A two-level method for Mimetic Finite Difference discretizations of elliptic problems

49/2013 Micheletti, S.
Fast simulations in Matlab for Scientific Computing

48/2013 Simone Palamara, Christian Vergara, Elena Faggiano, Fabio Nobile
An effective algorithm for the generation of patient-specific Purkinje networks in computational electrocardiology

45/2013 Sangalli, L.M.; Secchi, P.; Vantini, S.
Analysis of AneuRisk65 data: K-mean Alignment

46/2013 Marron, J.S.; Ramsay, J.O.; Sangalli, L.M.; Srivastava, A.
Statistics of Time Warpings and Phase Variations

44/2013 Sangalli, L.M.; Secchi, P.; Vantini, S.
AneuRisk65: a dataset of three-dimensional cerebral vascular geometries

43/2013 Patriarca, M.; Sangalli, L.M.; Secchi, P.; Vantini, S.
Analysis of Spike Train Data: an Application of K-mean Alignment