UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Model order reduction of coupled thermo-hydro-mechanical processes in geo-environmental applications

Ygee Larion

ADVERTIMENT La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del repositori institucional UPCommons (http://upcommons.upc.edu/tesis) i el repositori cooperatiu TDX (http://www.tdx.cat/) ha estat autoritzada pels titulars dels drets de propietat intel·lectual **únicament per a usos privats** emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei UPCommons o TDX. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a UPCommons (*framing*). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del repositorio institucional UPCommons (http://upcommons.upc.edu/tesis) y el repositorio cooperativo TDR (http://www.tdx.cat/?locale-attribute=es) ha sido autorizada por los titulares de los derechos de propiedad intelectual **únicamente para usos privados enmarcados** en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio UPCommons No se autoriza la presentación de su contenido en una ventana o marco ajeno a UPCommons (*framing*). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the institutional repository UPCommons (http://upcommons.upc.edu/tesis) and the cooperative repository TDX (http://www.tdx.cat/?locale- attribute=en) has been authorized by the titular of the intellectual property rights **only for private uses** placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading nor availability from a site foreign to the UPCommons service. Introducing its content in a window or frame foreign to the UPCommons service is not authorized (framing). These rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Model order reduction of coupled thermo-hydro-mechanical processes in geo-environmental applications

by

Ygee Larion

A thesis submitted to the

Building, Architecture and Town Planning Department
École Polytechnique de Bruxelles
Université libre de Bruxelles

and

Departament d'Enginyeria Civil i Ambiental
Universitat Politècnica de Catalunya

in partial fulfilment of the requirements for the degree of
Doctor in Engineering Sciences

29 September 2021

Supervisors
Prof. Thierry J. Massart (Université libre de Bruxelles)
Prof. Sergio Zlotnik (Universitat Politècnica de Catalunya)
Prof. Pedro Díez (Universitat Politècnica de Catalunya)

Members of the Jury

Prof. Pierre Gérard(Université libre de Bruxelles)
Prof. Bertrand François (Université libre de Bruxelles)
Prof. Sebastia Olivella (Universitat Politècnica de Catalunya)
Prof. Gianluigi Rozza (International School for Advanced Studies, Trieste)
Prof. Fredrik Larsson (Chalmers University)

Abstract

In a large number of geo-environmental applications, it is essential to model coupled processes that depend on several design parameters such as material properties and geometrical features. Thermo-hydro-mechanical (THM) processes are, among others, key effects to consider in critical applications such as deep geological repository of hazardous waste. This thesis proposes novel model order reduction strategies to evaluate the thermo-hydro-mechanical response of the material, taking into account the complexities involved in the coupled processes for such applications.

To include variability of some design parameters, an a-posteriori model order reduction approach with reduced basis methods is applied to solve the high-dimensional parametric THM system. The reduction is based on an offline-online stage strategy. In the offline stage, reduced subspaces are constructed by a greedy adaptive procedure and in the online stage, multi-subspace projection is performed to quickly obtain the coupled THM response at any value of the design parameter. At the core of the greedy adaptive strategy is a goal-oriented error estimator that guides the selection of optimal design parameters where snapshots are evaluated. To tackle nonlinearity in the form of elasto-plastic material behaviour, the multi-subspace reduced basis method is combined with sub-structuring by domain decomposition.

The effectiveness of the model reduction strategies are demonstrated on inverse problems involving large-scale geomodels that depict the coupled response of host rocks in potential deep geological repository sites. Two types of scenarios are considered: (i) the host rock undergoing geomorphological process is investigated as glacier advances over it for a period lasting over thousands of years and (ii) the clay response of an underground research laboratory is modelled numerically to support and validate in-situ heating experiments.

Πᾶν ἐφήμερον, καὶ τὸ μνημονεῦον καὶ τὸ μνημονευόμενον

Marcus Aurelius

Contents

List of Figures

List of Tables

Table **Page**

List of Algorithms

1

Introduction

1 Background and Motivation

Characterizing the coupled processes in geological porous media plays a very important role for many geo-environmental applications including carbon dioxide sequestration [124, 132, 133], geothermal resource extraction [43, 47, 84], frictional heating in earthquake slip [2, 80, 117], freezing and thawing in soils [55, 88, 135] and deep geological repository for nuclear waste isolation [70, 93, 95, 120, 136]. The present challenges in dealing with time scales, spatial configuration and performance requirements of these engineered underground systems are becoming more and more demanding. In most cases, it is imperative that computational methods are employed in order to fully tackle the scope and complexity of these geo-environmental engineering problems.

Particularly for nuclear waste isolation projects where the deep geological repository is expected to confine radioactive waste over a long period of time, it is necessary to accurately predict the mechanical, thermal, hydrologic and chemical processes that influence the overall performance of the structure. These coupled processes are governed by changing conditions inside the disposal system – evolving from open-drift to saturation period and eventually, through the entire heating and cooling cycle of the decaying radioactive waste. Varying external conditions such as geomorphological processes (e.g. due to glaciation or erosion) to which the structure might be subjected will also contribute to the evolution of the coupled processes within the system. The structure which is typically designed as a multi-barrier system requires these basic functions: (i) the combination of engineered and natural barriers should structurally hold the radioactive waste canisters in place while preventing excavation collapse, (ii) the multi-barrier system should provide a very high level of water tightness to avoid radionuclide migration to the geosphere, (iii) during the operational phase, the high temperature and hydraulic pressure in the near field should not severely impact the the previous functions and (iv) the geological repository must remain stable while accounting for geomorphological processes over time scales of several thousands of years. Having little to no obligation passed onto future generations to actively maintain the facility, most countries including Argentina, Australia, Belgium, Canada, Czech Republic, Finland, France, Japan, the Netherlands, Republic of Korea, Russia, Spain, Sweden, Switzerland, the UK, and the USA prefer deep geological disposal to isolate nuclear waste from the biosphere [152]. With several countries studying the feasibility of operating a nuclear waste repository in the last decades, the repository site in the Olkiluoto bedrock (Finland) will be the first licensed deep geological disposal facility which is expected to start storing fuel-packed copper canisters in 2023 [68, 139].

Due to very long time scales involved in nuclear waste disposal, the safety and performance analysis of the repository must be carried out by combining experimental studies with numerical simulations. Experiments are performed in laboratory

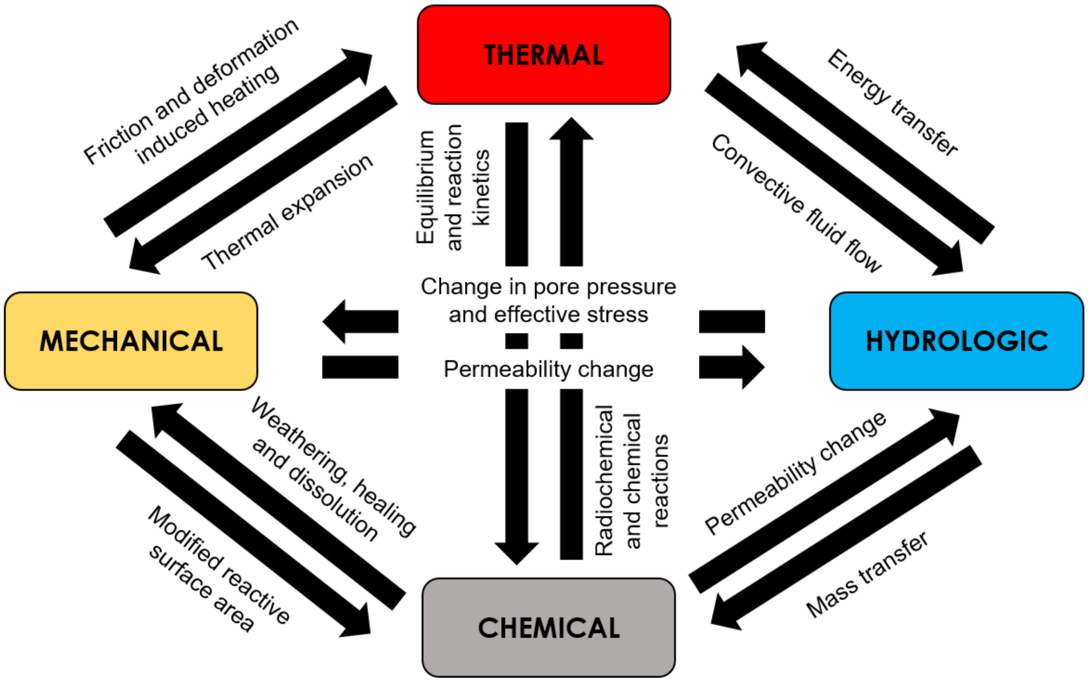


Fig. 1.1 Interaction of coupled thermo-hydro-mechanical-chemical (THMC) processes and its effects in rock mass behavior.

and in-situ to obtain information crucial to refining the numerical models used in performance assessment. There are numerous long-term laboratory tests and in-situ tests performed to investigate the interactions of coupled processes in near-field systems of different host rocks [6, 17, 19, 111]. To support these laboratory and in-situ tests validation, numerical models and computer codes have been developed (e..g ROCKMAS, FRACON, MOTIF, THAMES, RF/ RM, CAST3M, CHEF/HYDREF/VIPLEF, TOUGH2/JAS3D) [4, 89, 95, 120, 121]. Most of these modelling approaches are based on finite element (FE) analysis for which high-dimensionality, due to nonlinearity, complex discretization, numerous uncertain parameters and prolonged time evolution, becomes an unavoidable issue. Computational efficiency for the FE-based modelling of coupled systems has been typically addressed by different high performance computing strategies (e.g. using parallel solvers [99, 138, 144, 149], implementing object-oriented programming paradigm [148], devising algebraic multigrid methods [150, 151], applying staggered Newton schemes [99, 131]) to alleviate the computational burden.

In such applications where numerical simulations are combined with experimental studies, the existence of uncertainties in model parameters is a critical aspect of the problem. By treating the problem as a parametrized system, using robust tools for inverse modelling techniques can effectively identify uncertain parameters based on partial information and in-situ measurements. Describing the problem in

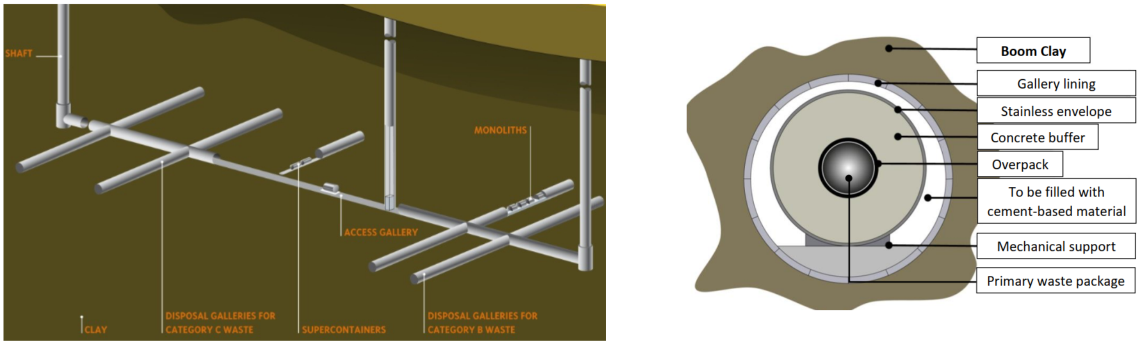


Fig. 1.2 Deep geological repository structure (left) and and schematic view of the multi-barrier system where radioactive waste is stored. Image taken from [13].

a parametric setting also allows for investigating the key consequences of variation in system parameters with respect to the coupled processes behavior in the deep geological repository. Within the framework of parametrized systems, a common requirement is the need for cheap and efficient simulation tools enabling fast and frequent solving of a set of problems that are similar to each other. For instance, the same problem needs to be solved many times by modifying one or several parameters in inverse modelling techniques. Having the set of problems share a common structure and form, reduced order modelling (ROM) techniques are crucial to efficiently solve the family of problems describing the parametrized system.

The purpose of ROM is to provide an accurate representation of generalized solutions using a so-called surrogate model. When compared to the full order approximation of the high-dimensional model, the reduced order model is obtained with a significant reduction of computational resources, associated with much fewer terms of the basis (number of degrees of freedom or modes) required to approximate the solution. The overall computational cost of ROM techniques depends on two phases:

(i) Offline Stage: An expensive procedure that involves construction of the reduced basis. This is done only once for a fixed range of parameters.

(ii) Online Stage: A much cheaper procedure than full order approximation where evaluation of the reduced model is done using the previously constructed reduced basis. This phase is performed every time a solution at a desired parameter value is evaluated. It is expected to produce ‘real-time’ response.

Model reduction techniques are typically classified based on how the reduced basis is constructed. Construction techniques of ROM can be classified either as a-posteriori or a-priori.

The a-posteriori approach is fundamentally based on preliminary solutions of the original problem evaluated at some selected points within the parametric space or the so-called 'snapshots'. Among a-posteriori approaches, the Proper Orthogonal Decomposition (POD) is one of the most widely used method [15, 27, 76, 85]. POD computes the optimal reduced basis using singular value decomposition (SVD) of the snapshots [42, 143]. The weakness of the method is related to the computational cost involved when dealing with large snapshot data. Other a-posteriori methods aim at reducing the number of snapshots or avoiding SVD computations which can result to loss of optimality in the reduced basis. A popular alternative to POD computation is applying greedy algorithm to smartly select the snapshots which refers to reduced basis (RB) method [97, 112]. The RB method performs the selection of optimal snapshots by implementing an adaptive strategy which is driven by a problem-dependent error estimator and greedy algorithm [11, 20, 61]. Other studies subsequently combined POD and greedy algorithm to build a more efficient reduced basis [86, 113].

The a-priori approach iteratively constructs the low-dimensional reduced basis that solves the high dimensional problem on the fly. Such methods do not involve using snapshots or SVD evaluations. The online stage for this approach does not need to solve any system which proves to be advantageous when dealing with real-time simulations as the solutions are always immediately available, regardless of the size of the reduced basis [31, 33, 98]. In a-posteriori approaches, the computational cost is pre-dominantly affected by the reduced basis dimension and the number of evaluations required – each evaluation entails projection of the original problem onto the reduced basis.

Despite being widely used in many engineering fields for several years, there are very few studies that explored model order reduction applications to coupled geomechanical problems (e.g. [22, 50, 87]). Due to the nature of model order reduction where applicability is problem-dependent such that an effective technique for one specific problem is not necessarily guaranteed to work for another problem, development of model reduction techniques that tackle time-dependent and nonlinear coupled systems is particularly challenging. Specifically for coupled processes in deep geological repository application, using model order reduction techniques is a fairly new research topic to be explored.

2 Scope and Objectives

In light of the challenges related to modelling of coupled geomechanical processes that are previously presented, the main objective pursued in the work presented in this thesis revolves around developing model order reduction strategies tailored to deep geological repository applications. To adequately tackle this objective, a two-fold approach needs to be considered. One aspect is focused on addressing the complexity of the physics that represents the coupled processes in the repository and the other aspect is focused on exploring model order reduction techniques. Each aspect belongs to a field of study which can be very broad to navigate hence, the scope of this thesis is limited to the following subject areas.

- The coupled processes are described by a thermo-hydro-mechanical (THM) system where uni-directional coupling is assumed with respect to the thermal process. Chemical, gas or biological impacts are considered insignificant in the types of applications presented in this study.
- The non-linearity investigated in the coupled THM system only occurs in the mechanical aspect – related to the elasto-plastic behavior of the porous material. The remaining aspects are described by a linear system.
- Snapshot-based (a-posteriori) reduced order modelling techniques are pursued in this study. To be more specific, Proper Orthogonal Decomposition and Reduced Basis methods are the main techniques explored for deep geological repository applications.
- The developed model reduction techniques address the complexity in the coupled system specific to the problem presented. Since ROM methods are known to be problem-dependent, it is not guaranteed that the techniques developed in this study will have the same effectiveness when applied to problems defined differently.

Furthermore, the direction of this thesis is guided by the topics stemming from the main objective and aiming to thoroughly answer the following research questions.

- Is it possible to carry out efficient (fast and accurate) computations for linear coupled THM systems using snapshot-based model reduction in direct query types of problem?
- How to extend and implement the previous approach to evaluate generalized solutions for multiple query problems that include variability of some model parameters?

- In the context of parametrized problems and a-posteriori model reduction, how can the offline stage be performed efficiently? In particular, how to optimally select the snapshots by minimizing the size of snapshot data that yields highly accurate reduced basis?
- How does the previously developed snapshot-based model reduction technique cope with non-linearity introduced in the coupled THM system by elasto-plastic constitutive laws?
- In real applications, how do the strategies developed in this study perform? Is the computational gain significant enough to solve inverse problems that are rather impractical in the context of full order finite element approximation? Is the ROM approximation accurate enough to support and validate real in-situ experiments?

3 Major Contributions

In order to accomplish the objectives defined previously, the following novelties are presented which constitute the main contributions of the thesis.

- A multi-subspace projection strategy is developed in the framework of snapshot-based model order reduction to tackle the coupled processes in the THM system. By doing so, the stability condition of the high dimensional reference model is inherited by the ROM.
- A goal-oriented a-posteriori error estimator is developed specifically for transient THM parametric systems. The error estimator involves solving an auxiliary adjoint problem that evolves backwards in time.
- Reduced bases for the coupled fields are constructed from a greedy adaptive procedure which is guided by the developed goal-oriented error estimator. The greedy adaptive procedure progressively constructs the reduced bases by selecting optimal snapshot points within the given parametric space. Driven by the goal-oriented error estimator, the adaptive procedure ensures that the resulting reduced basis is certified.
- A non-linear model reduction strategy that combines RB method and sub-structuring via domain decomposition is devised to tackle elasto-plasticity in the THM system.
- The developed framework is applied to large-scale geo-environmental problems describing the coupled processes in deep geological repository. Model order reduction is used to effectively investigate granite rock mass behavior undergoing geomorphological process (e.g. glacial advance) and to validate clay response subjected to in-situ heating based on real experiments.
- The above ROM features and careful implementation yield significant computational gain while maintaining highly accurate approximation of the solution. The work presented in this thesis illustrates how these aspects are extremely advantageous in fast and multiple query types of problems.

4 Structure of the Thesis

The main contributions of the thesis pursuing drastic reduction in computational cost for parametric THM systems are presented in three main parts. The methodology for solving linear THM systems by reduced basis model order reduction framework is introduced in Chapter 2. The multi-subspace projection, greedy-adaptive basis generation and goal-oriented error estimation are developed and implemented on a large-scale problem involving a 3D rock mass subjected to an advancing glacier. The ideas presented in the chapter are centred on developing a ROM that significantly yields computational speed up while accurately preserving the physics of the transient coupled system.

The methodology is further extended in Chapter 3, in order to handle more complex problems based on real in-situ heating experiments conducted at HADES undergound laboratory in Mol, Belgium. Some of the main challenges addressed are linearizing the system involving temperature-dependent hydro-mechanical parameters, dealing with 3D heterogeneous domain, including initial geostatic stresses, hydrostatic fluid pressure and temperature distribution, adapting the greedy adaptive strategy and error estimation technique to tackle the above complexities, and matching the numerical model as close as possible to experiment observations. The chapter focuses on effectively implementing the methodology for parametrized problems, specifically for inverse parameter identification where real sensor measurement data collected from in-situ experiments are used. The ROM is employed as a surrogate to replace the traditionally used FE model, which is proven impractical in optimization algorithms for large-scale parameter identification problems.

In Chapter 4, it is aimed to develop a model reduction strategy that deals with non-linear THM systems described by elato-plastic material behaviour. A new approach is developed where sub-structuring via domain composition and multi-subspace projection are combined to combat the expensive multi-level iterative procedure involved in solving the nonlinear system. The methodology is validated on parametrized problems where the effectiveness of the ROM is demonstrated in the framework of inverse modelling.

The closing chapter expounds the main findings and conclusions of the work presented in the thesis. It also discusses the drawbacks and challenges to overcome related to the current methodology. Moreover, the next directions are suggested for further improvement of the work presented in the thesis.

2

Certified Reduced Basis Approximation of Linear THM Systems

This chapter is based on
Larion, Y., Zlotnik, S., Massart, T. J., and Díez, P. **Building a certified reduced basis for coupled thermo-hydro-mechanical systems with goal-oriented error estimation.** Comput Mech, 66:559–573, 2020.

1 Introduction

Numerous studies have been conducted on coupled THM systems related to the behavior of geological formations for investigating the safety, performance and design optimization of underground nuclear waste repositories [10, 39, 40, 41, 70, 140, 141]. Over the years, some approaches started focusing on developing numerical models addressing these processes to support laboratory and in-situ tests validation [1, 26, 32, 34, 51, 100, 101, 142]. These large-scale numerical models face the challenge of dealing with systems of high dimensionality related to large and complex geometrical discretizations, nonlinearity, a number of uncertain field parameters and prolonged time evolution. This issue becomes even more critical for certain types of problems requiring fast and multiple queries, for instance inverse identification problems. The standard finite element based solvers are computationally demanding and consequently, different high performance computing strategies are used to alleviate the computational burden [3, 73, 99].

Despite being widely used in many engineering fields for several years, there are very few studies that explored model order reduction applications to coupled geomechanical problems [22, 50, 87]. Model reduction techniques aim at reducing the dimensionality of a system by projecting the reference full order model to a low-dimensional subspace while preserving key information up to an acceptable level of accuracy. In the present contribution, we focus our attention on the reduced basis (RB) method - an enrichment-based class of model order reduction that requires an adaptive strategy to construct a low-dimensional subspace spanning solutions of a system under parametric variation. Given a training sample that describes the parametric variation, it is desired to accurately approximate the reference solution at a much lower computational cost for any parameter inside the training set. The construction of this low-dimensional subspace mainly involves the assembly and post-processing of so-called ‘snapshots’ that refer to full order solutions evaluated at specific parameter values in the training sample.

An efficient a-posteriori error estimator is crucial to optimally select snapshots inside the parametric training sample. The efficiency of an error estimator is characterized by its rigor, sharpness and computational cost. Many of the error estimation techniques exploit the relationship between the residual and the error approximation of the system. This class of techniques is based on the post-processing of the residual, which can be done either explicitly by integration of residuals, or implicitly by solving local problems using residuals as source terms [71]. In recent years, several research efforts have been devoted to residual-based a-posteriori error estimators in the context of reduced basis applications for different types of PDE problems [62, 63, 65, 92, 112, 155].

In this contribution, we develop a goal-oriented a-posteriori error estimator for time-dependent coupled THM systems which assesses the error in specific features of the solution, termed as ‘Quantity of Interest’ (QoI). The time-dependence of the coupled THM system poses an additional complexity as it involves the computation of the auxiliary adjoint problem evolving backwards in time. Various implementations of goal-oriented error estimation for time-dependent problems have been investigated in [16, 44, 53, 90, 106, 130, 145, 154]. Similar to the methodology presented in [44, 90, 106], the error estimates for the coupled THM system are derived based on the explicitly-computed weak residual of the primal problem and on the implicitly-computed adjoint solution associated with the QoI. The error estimates are fed to the greedy-based adaptive strategy to optimally select the snapshots and certify the RB-projected solutions. For computational efficiency, the implicitly-computed adjoint is evaluated only once and reused to obtain the error estimates for the entire parametric training sample.

The rest of the chapter is organized as follows. In Section 2, we describe the model problem governing equations and present the notations for the discretized coupled THM system in full order and in reduced order form. In Section 3, a greedy adaptive strategy is proposed to optimally select the snapshot set used in the construction of the low-dimensional RB subspace. We derive the formulation of the backward-in-time evolving dual problem corresponding to the coupled THM system and develop the goal-oriented error estimation scheme in Section 4. Finally in Section 5, numerical examples for a regional model of glacier advance in 2D and 3D are presented in order to demonstrate the accuracy of the optimally built RB-approximation and the computational gain achieved with the proposed methodology.

2 Problem Definition: Linear Isotropic THM System

2.1 Governing Equations and Full Order Discretization

The coupled system for a fully-saturated porous geomaterial includes the three equations describing the deformation of the linearly-elastic body, the flow of the fluid phase and the heat conduction. The mechanical behavior is governed by poroelastic constitutive equations, while the hydraulic and thermal responses are governed by Darcy's law and Fourier's law, respectively. The classical conservation laws together with the constitutive relationships yield the governing equations for the time-dependent coupled THM system [93], resulting in an initial boundary value problem that reads: find temperature $T(\boldsymbol{x},t)$, displacement $\mathbf{u}(\boldsymbol{x},t)$ and pressure $p(\boldsymbol{x},t)$ with $\boldsymbol{x}\in\bar{\Omega}\subset\mathbb{R}^d$ and $t\in(0,t_f]$ such that

$$\left.\begin{aligned}
k_c\nabla^2 T-\rho c_p\dot{T}+Q&=0\\
G_D\nabla\cdot(\nabla\mathbf{u})+(G_D+\lambda)\nabla(\nabla\cdot\mathbf{u})&\\
-\alpha\nabla p-3K_D\alpha_s\nabla T&=0\\
-\frac{k}{\mu_f}\nabla\cdot(\nabla p)+\alpha\nabla\cdot\dot{\mathbf{u}}+\frac{1}{M}\dot{p}&\\
-[\phi 3\alpha_f+(\alpha-\phi)3\alpha_s]\dot{T}&=0
\end{aligned}\right\}\quad \text{in } \Omega\times(0,t_f] \tag{2.1}$$

with Neumann boundary conditions,

$$\left.\begin{aligned}
-k_c\nabla_{\mathrm{n}}T&=g_{\mathrm{N}}\\
\boldsymbol{\sigma}_{ij}\mathbf{n}_j&=\boldsymbol{\tau}_{\mathrm{N}_i}\\
\frac{k}{\mu_f}\nabla_{\mathrm{n}}p&=\upsilon_{\mathrm{N}}
\end{aligned}\right.\quad \text{on } \partial\Gamma_N\times(0,t_f] \tag{2.2}$$

Dirichlet boundary conditions,

$$\left.\begin{aligned}
T&=T_D\\
\mathbf{u}&=\mathbf{u}_D\\
p&=p_D
\end{aligned}\right.\quad \text{on } \partial\Gamma_D\times(0,t_f] \tag{2.3}$$

and initial conditions

$$\left.\begin{aligned}
T&=T_0\\
\mathbf{u}&=\mathbf{u}_0\\
p&=p_0
\end{aligned}\right.\quad \text{in } \Omega\times\{t=0\} \tag{2.4}$$

where the notation of physical parameters in the system of equations are listed in Table 2.1.

Table 2.1 Notation of physical parameters of the coupled THM system

Parameters	
K_D	bulk modulus of drained rock
G_D	shear modulus of the drained rock
λ	Lamé constants
α	Biot-Willis coefficient
M	constrained specific storage coefficient
μ_f	viscosity
k	permeability
ϕ	porosity
α_s	thermal expansion coefficient in the solid phase
α_f	thermal expansion coefficient in the fluid phase
c_p	specific heat
k_c	thermal conductivity
Q	heat source
ρ	solid density
ρ_f	fluid density
$\boldsymbol{\sigma}$	traction

We approximate the solution to the problem in (2.1)-(2.4) by the standard finite element (FE) method. The system is spatially discretized by associating the approximation to a mesh of characteristic size H in the functional space $\mathcal{Z}^H \subset \mathcal{H}_0^1(\Omega)$, $\mathcal{W}^H \subset \mathcal{H}_0^1(\Omega)$, $\mathcal{Q}^H \subset \mathcal{H}_0^1(\Omega)$:

$$\begin{aligned} T(\boldsymbol{x},t) &\approx T^H(\boldsymbol{x},t) = \sum_{i=1}^{\mathtt{n_{nodes}}} T_i(t) N_i(\boldsymbol{x}) = \mathbf{N}^\mathsf{T} \boldsymbol{T} \\ \mathbf{u}(\boldsymbol{x},t) &\approx \mathbf{u}^H(\boldsymbol{x},t) = \sum_{i=1}^{\mathtt{n_{nodes}}} \mathbf{u}_i(t) N_i(\boldsymbol{x}) = \tilde{\mathbf{N}}^\mathsf{T} \boldsymbol{U} \\ p(\boldsymbol{x},t) &\approx p^H(\boldsymbol{x},t) = \sum_{i=1}^{\mathtt{n_{nodes}}} p_i(t) N_i(\boldsymbol{x}) = \mathbf{N}^\mathsf{T} \boldsymbol{p} \end{aligned} \tag{2.5}$$

where $\boldsymbol{U} \in \mathbb{R}^{d \times \mathtt{n_{nodes}}}$, $\mathtt{n_{nodes}}$ is the number of nodes discretizing the field in Ω, while

$\mathbf{N}$ and $\tilde{\mathbf{N}}$ are compact support shape functions. In the case of a 2D domain, $\mathbf{N}$ and $\tilde{\mathbf{N}}$ are expressed as $\mathbf{N} := [N_1(\boldsymbol{x})\ N_2(\boldsymbol{x})\ ...N_{\mathtt{n_{nodes}}}(\boldsymbol{x})]^\mathsf{T}$ and

$$\tilde{\mathbf{N}} := \begin{bmatrix} N_1(\boldsymbol{x}) & 0 & N_2(\boldsymbol{x}) & 0 & ... & N_{\mathtt{n_{nodes}}}(\boldsymbol{x}) & 0 \\ 0 & N_1(\boldsymbol{x}) & 0 & N_2(\boldsymbol{x}) & ... & 0 & N_{\mathtt{n_{nodes}}}(\boldsymbol{x}) \end{bmatrix}.$$

For further use, we write the weak form of the problem (2.1) -(2.4) in a compact form as

$$\begin{aligned} B(T, \mathbf{u}, p; z, \mathbf{w}, q) &= L(z, \mathbf{w}, q) \\ &\qquad \forall z \in \mathcal{Z}^H, \mathbf{w} \in \mathcal{W}^H, q \in \mathcal{Q}^H \end{aligned} \tag{2.6}$$

where

$$\begin{aligned} B(T, \mathbf{u}, p; z, \mathbf{w}, q) = &\int_\Omega \frac{\partial z}{\partial x_j} k_c \frac{\partial T}{\partial x_j} d\Omega + \int_\Omega z \rho c_p \frac{\partial T}{\partial t} d\Omega \\ &+ \int_\Omega \frac{\partial \mathbf{w}}{\partial x_j} G \frac{\partial \mathbf{u}_i}{\partial x_j} d\Omega \\ &+ \int_\Omega \frac{\partial \mathbf{w}}{\partial x_i} (G + \lambda) \frac{\partial \mathbf{u}_j}{\partial x_j} d\Omega \\ &- \int_\Omega \mathbf{w} \alpha \frac{\partial p}{\partial x_i} d\Omega \\ &- 3 \int_\Omega \mathbf{w} K_D \alpha_s \frac{\partial T}{\partial x_i} d\Omega \\ &- \int_\Omega \frac{\partial q}{\partial x_j} \frac{k}{\mu_f} \frac{\partial p}{\partial x_j} d\Omega \\ &+ \int_\Omega q \frac{1}{M} \frac{\partial p}{\partial t} d\Omega \\ &+ \int_\Omega q \alpha \frac{\partial}{\partial t} (\frac{\partial u_i}{\partial x_i}) d\Omega \\ &- \int_\Omega q [\phi 3 \alpha_f + (\alpha - \phi) 3 \alpha_s] \frac{\partial T}{\partial t} d\Omega \end{aligned} \tag{2.7}$$

for $i, j = 1, 2, 3$ and

$$L(z, \mathbf{w}, q) = \int_\Gamma z g_\mathrm{N} d\Gamma_N + \int_\Gamma \mathbf{w} \boldsymbol{\tau}_{\mathrm{N}_i} d\Gamma_N - \int_\Gamma q \upsilon_\mathrm{N} d\Gamma_N. \tag{2.8}$$

The substitution of the approximation (2.5) in the weak form (2.6) leads to the discrete matrix form,

$$\mathbb{K}\,\mathbb{U} + \mathbb{C}\,\dot{\mathbb{U}} = \mathbb{F} \tag{2.9}$$

$$\underbrace{\begin{bmatrix} \mathbf{K_H} & 0 & 0 \\ \mathbf{K_T} & \mathbf{K_U} & \mathbf{C_P} \\ 0 & 0 & \mathbf{K_P} \end{bmatrix}}_{\mathbb{K}} \underbrace{\begin{bmatrix} \boldsymbol{T} \\ \boldsymbol{U} \\ \boldsymbol{p} \end{bmatrix}}_{\mathbb{U}} + \underbrace{\begin{bmatrix} \mathbf{C_H} & 0 & 0 \\ 0 & 0 & 0 \\ \mathbf{C_T} & \mathbf{C_P^{\intercal}} & \mathbf{C_M} \end{bmatrix}}_{\mathbb{C}} \underbrace{\begin{bmatrix} \dot{\boldsymbol{T}} \\ \dot{\boldsymbol{U}} \\ \dot{\boldsymbol{p}} \end{bmatrix}}_{\dot{\mathbb{U}}} = \underbrace{\begin{bmatrix} \mathbf{F_T} \\ \mathbf{F_u} \\ \mathbf{F_p} \end{bmatrix}}_{\mathbb{F}} \tag{2.10}$$

where $\mathbb{K}$ contains the stiffness and conductivity matrices, $\mathbb{C}$ contains the capacity (time-dependent) matrices and $\mathbb{F}$ contains the corresponding vectors to $L(\cdot)$.

For the time discretization, a time-discrete framework in the interval $I \equiv (0, t_f]$ is considered such that the subintervals are denoted as $\{0 = t^0 < t^1 < ... < t^l < .... < t^{\mathcal{L}} = t_f\}$ for $l = 1, 2, ..., \mathcal{L}$. For notation purposes, let Δt^l be the characteristic time step of the time-discrete framework, measured as $\Delta t^l = t^l - t^{l-1}$. To perform the time integration, an Euler-Backward scheme is used and the space-time discrete problem is written as

$$\begin{gathered} \Delta t^l \mathbb{K}\mathbb{U}^{\,l} + \mathbb{C}\,\mathbb{U}^{\,l} = \mathbb{C}\,\mathbb{U}^{\,l-1} + \Delta t^l \mathbb{F}^{\,l} \\ \forall l \in 1, 2, ..., \mathcal{L} \end{gathered} \tag{2.11}$$

Due to a lack of inf-sup condition, the coupled THM system can show oscillating solution patterns when temperature, displacement and pore pressure are approximated by the same set of basis functions [8, 12, 21]. We therefore use Taylor-Hood finite element spaces, assigning quadratic basis functions for displacement and linear basis functions for pore pressure and temperature, to avoid instability in the coupled THM system [83, 93, 157].

2.2 Reduced Order Model (ROM)

Considering the physical parameter constants of the coupled THM system appearing in (2.1), we assume that the material properties depend on a set of parameter $\boldsymbol{\mu} = (\mu_1, ..., \mu_{\mathtt{n_{par.}}}) \in \mathcal{P} \subset \mathbb{R}^P$, $\mathtt{n_{par.}} \geq 1$. The system unknowns are then denoted with an additional dependency on a set of material parameter $(T(\boldsymbol{x}, t, \boldsymbol{\mu})$, $\mathbf{u}(\boldsymbol{x}, t, \boldsymbol{\mu}), p(\boldsymbol{x}, t, \boldsymbol{\mu}))$. Note that in the next part, we simplify the field notations by writing them as $(T(\boldsymbol{\mu}), \mathbf{u}(\boldsymbol{\mu}),\, p(\boldsymbol{\mu}))$ instead of the full form. The problem in the parametric form is expressed as follows: find $(T(\boldsymbol{\mu}), \mathbf{u}(\boldsymbol{\mu}),\, p(\boldsymbol{\mu})) \in \mathcal{Z}^H \times \mathcal{W}^H \times \mathcal{Q}^H$ such that

$$B(T(\boldsymbol{\mu}),\mathbf{u}(\boldsymbol{\mu}),p(\boldsymbol{\mu});z,\mathbf{w},q)=L(z,\mathbf{w},q) \tag{2.12}$$

where

$$\begin{aligned}
B(\cdot\,;\cdot) = & \int_\Omega \frac{\partial z}{\partial x_j} k_c(\boldsymbol{\mu}) \frac{\partial T}{\partial x_j} d\Omega + \int_\Omega z\rho c_p(\boldsymbol{\mu}) \frac{\partial T}{\partial t} d\Omega \\
& + \int_\Omega \frac{\partial \mathbf{w}}{\partial x_j} G(\boldsymbol{\mu}) \frac{\partial \mathbf{u}_i}{\partial x_j} d\Omega + \int_\Omega \frac{\partial \mathbf{w}}{\partial x_i} (G(\boldsymbol{\mu}) \\
& + \lambda(\boldsymbol{\mu})) \frac{\partial \mathbf{u}_j}{\partial x_j} d\Omega - \int_\Omega \mathbf{w}\alpha(\boldsymbol{\mu}) \frac{\partial p}{\partial x_i} d\Omega \\
& - 3\int_\Omega \mathbf{w} K_D \alpha_s(\boldsymbol{\mu}) \frac{\partial T}{\partial x_i} d\Omega \\
& - \int_\Omega \frac{\partial q}{\partial x_j} \frac{k(\boldsymbol{\mu})}{\mu_(\boldsymbol{\mu})f} \frac{\partial p}{\partial x_j} d\Omega \\
& + \int_\Omega q \frac{1}{M(\boldsymbol{\mu})} \frac{\partial p}{\partial t} d\Omega \\
& + \int_\Omega q\alpha(\boldsymbol{\mu}) \frac{\partial}{\partial t} (\frac{\partial \mathbf{u}_i}{\partial x_i}) d\Omega \\
& - \int_\Omega q[\phi(\boldsymbol{\mu}) 3\alpha_f(\boldsymbol{\mu}) + (\alpha(\boldsymbol{\mu}) \\
& - \phi(\boldsymbol{\mu})) 3\alpha_s(\boldsymbol{\mu})] \frac{\partial T}{\partial t} d\Omega
\end{aligned} \tag{2.13}$$

for $i, j = 1, 2, 3$. The approximate solution of the ROM is based on a standard Galerkin projection of the full order FE model to the global approximation spaces generated from well-chosen solutions (snapshots) evaluated at specific parameter and time values. Given a set of parameter values $\mathcal{M} = \{\boldsymbol{\mu}^1, ..., \boldsymbol{\mu}^N\}$, the snapshots contain the corresponding solutions $S^N = \{T(\boldsymbol{\mu}^1), \mathbf{u}(\boldsymbol{\mu}^1), p(\boldsymbol{\mu}^1), \ldots, T(\boldsymbol{\mu}^N), \mathbf{u}(\boldsymbol{\mu}^N), p(\boldsymbol{\mu}^N)\}$ where N is much smaller than the characteristic size of the reference full order model.

It is critical that the ROM inherits the crucial properties of the reference full order model − particularly, maintaining the numerical stability of the model to guarantee the accuracy and convergence of the ROM. There are numerous works dedicated to the stability of ROM, proposing strategies to develop stability preserving models for specific types of problem [5, 75, 116, 119]. In general however, ensuring the numerical stability in model reduction is still an open issue. The numerical stability of the reduced coupled THM model is tackled similarly as in the FE context [8, 12, 21], by using different basis functions to approximate the displacement, pore pressure and temperature fields. Snapshot solutions are also orthonormalized to ensure algebraic stability for each basis.

We define the global approximation (RB) spaces separately for each of the field variable as follows:

$$\begin{aligned} \mathcal{Z}_N^H &= \mathrm{span}\{T(\boldsymbol{\mu}^n), \quad 1 \leq n \leq \mathtt{n}_{\mathtt{snapshots}}\} \quad \subset \mathcal{Z}^H \\ \mathcal{W}_N^H &= \mathrm{span}\{\mathbf{u}(\boldsymbol{\mu}^n), \quad 1 \leq n \leq \mathtt{n}_{\mathtt{snapshots}}\} \quad \subset \mathcal{W}^H \\ \mathcal{Q}_N^H &= \mathrm{span}\{p(\boldsymbol{\mu}^n), \quad 1 \leq n \leq \mathtt{n}_{\mathtt{snapshots}}\} \quad \subset \mathcal{Q}^H. \end{aligned} \tag{2.14}$$

To project the FE model to the set of RB spaces, the coupled system is 'decoupled' such that elements of the stiffness, conductivity and capacity matrices are extracted and grouped into blocks based on the field variable that it interacts with (2.15). Let $\mathbb{Z}, \mathbb{W}$ and $\mathbb{Q}$ denote the RB matrices for temperature, displacement and pressure, respectively. Each block is projected accordingly onto its corresponding basis, i.e., $\mathbf{K_T}$ interacts with temperature and displacement fields, therefore it is projected to the RB space from temperature ($\mathcal{Z}_N^H$) and displacement ($\mathcal{W}_N^H$) snapshots. Meanwhile, $\mathbf{K_P}$ that solely interacts with pore pressure is projected only to the RB space from pressure snapshots ($\mathcal{Q}_N^H$). These projected blocks are then reassembled in the same way as the FE coupled system as seen in (2.15) and (2.16), but the resulting blocks will have a much smaller characteristic size of order N.

$$\begin{aligned} &\begin{bmatrix} \mathbf{K}_{\mathbf{H}N} & 0 & 0 \\ \mathbf{K}_{\mathbf{T}N} & \mathbf{K}_{\mathbf{U}N} & \mathbf{C}_{\mathbf{P}N} \\ 0 & 0 & \mathbf{K}_{\mathbf{P}N} \end{bmatrix} \begin{bmatrix} \boldsymbol{T}_N \\ \boldsymbol{U}_N \\ \boldsymbol{p}_N \end{bmatrix} \\ &+ \begin{bmatrix} \mathbf{C}_{\mathbf{H}N} & 0 & 0 \\ 0 & 0 & 0 \\ \mathbf{C}_{\mathbf{T}N} & \mathbf{C}_{\mathbf{P}N}^{\mathsf{T}} & \mathbf{C}_{\mathbf{M}N} \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{T}}_N \\ \dot{\boldsymbol{U}}_N \\ \dot{\boldsymbol{p}}_N \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{\mathbf{T}N} \\ \mathbf{F}_{\mathbf{u}N} \\ \mathbf{F}_{\mathbf{p}N} \end{bmatrix} \end{aligned} \tag{2.15}$$

$$\begin{aligned} &\mathbf{K}_{\mathbf{H}N} = \mathbb{Z}^{\mathsf{T}}\, \mathbf{K_H}\, \mathbb{Z} && \mathbf{K}_{\mathbf{T}N} = \mathbb{W}^{\mathsf{T}}\, \mathbf{K_T}\, \mathbb{Z} \\ &\mathbf{K}_{\mathbf{U}N} = \mathbb{W}^{\mathsf{T}}\, \mathbf{K_U}\, \mathbb{W} && \mathbf{C}_{\mathbf{P}N} = \mathbb{W}^{\mathsf{T}}\, \mathbf{C_P}\, \mathbb{Q} \\ &\mathbf{K}_{\mathbf{P}N} = \mathbb{Q}^{\mathsf{T}}\, \mathbf{K_P}\, \mathbb{Q} && \mathbf{C}_{\mathbf{H}N} = \mathbb{Z}^{\mathsf{T}}\, \mathbf{C_H}\, \mathbb{Z} \\ &\mathbf{C}_{\mathbf{T}N} = \mathbb{Q}^{\mathsf{T}}\, \mathbf{C_T}\, \mathbb{Z} && \mathbf{C}_{\mathbf{P}N}^{\mathsf{T}} = \mathbb{Q}^{\mathsf{T}}\, \mathbf{C}_{\mathbf{P}}^{T}\, \mathbb{W} \\ &\mathbf{C}_{\mathbf{M}N} = \mathbb{Q}^{\mathsf{T}}\, \mathbf{C_M}\, \mathbb{Q} && \mathbf{F}_{\mathbf{T}N} = \mathbb{Z}^{\mathsf{T}}\, \mathbf{F_T} \\ &\mathbf{F}_{\mathbf{u}N} = \mathbb{W}^{\mathsf{T}}\, \mathbf{F_u} && \mathbf{F}_{\mathbf{p}N} = \mathbb{Q}^{\mathsf{T}}\, \mathbf{F_p} \\ &\boldsymbol{T}_N = \mathbb{Z}^{\mathsf{T}}\, \boldsymbol{T} && \boldsymbol{U}_N = \mathbb{W}^{\mathsf{T}}\, \boldsymbol{U} \\ &\boldsymbol{p}_N = \mathbb{Q}^{\mathsf{T}}\, \boldsymbol{p} \end{aligned} \tag{2.16}$$

Given $\boldsymbol{\mu} \in \mathcal{P}$, the approximated solutions belonging to the RB spaces ($T_N(\boldsymbol{\mu}) \in \mathcal{Z}_N^H$, $\mathbf{u}_N(\boldsymbol{\mu}) \in \mathcal{W}_N^H$, $p_N(\boldsymbol{\mu}) \in \mathcal{Q}_N^H$) are obtained by solving

$$\begin{gathered} B(T_N(\boldsymbol{\mu}), \mathbf{u}_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu}); z, \mathbf{w}, q) = L(z, \mathbf{w}, q) \\ \forall z \in \mathcal{Z}_N^H,\ \mathbf{w} \in \mathcal{W}_N^H,\ q \in \mathcal{Q}_N^H \end{gathered} \tag{2.17}$$

where the dimensionality of the system corresponding to the left hand side $B(\,\cdot\,)$ and right hand side $L(\,\cdot\,)$ terms, as described in (2.15) and (2.16), is much lower than the full order FE model.

3 Generation of the Reduced Basis

The computational procedure involved in generating the reduced basis is explained in this section. The RB is built from an assembly of snapshot solutions of the full order model evaluated within a parametric training sample. Optimally selecting the parameters where such snapshots are evaluated is evidently critical for efficiently building the RB. It is desired that the selected snapshots give the most accurate approximation while keeping the number of snapshots to a minimum. In order to satisfy this requirement, we pursue an adaptive strategy in which the criterion to select the snapshots is driven by an a-posteriori error estimator. In turn, this error estimator also provides certification to the RB-approximated solution. The error estimator developed for the coupled THM system is discussed thoroughly in Section 4.

The adaptive strategy is implemented in a greedy-based sampling framework to build hierarchical RB approximation spaces [62, 63, 87, 107], in which the basis is enriched by appending new snapshot solutions at every enrichment step. The greedy adaptive strategy is explained by the schematic procedure below.

Given a parametric training sampling space discretization $\Xi_{train} \subset \mathtt{n}_{\mathtt{sample}}$, with $\Xi_{train} = \text{span}\ \{\boldsymbol{\mu}^1,\ \boldsymbol{\mu}^2 \ldots,\ \boldsymbol{\mu}^{\mathtt{n}_{\mathtt{sample}}}\}$, a stopping tolerance for estimated error ϵ, and a maximum number of enrichment steps N_{max}, initiate the greedy strategy at $N = 0$ with empty snapshot matrices $S_T^N, S_{\mathbf{u}}^N, S_p^N$ and a pre-selected initial parameter value $\boldsymbol{\mu}^* = \boldsymbol{\mu}^n$

Algorithm 2.1 Greedy basis generation

```
while $N < N_{max}$ and $\delta_N > \epsilon$
    $N \leftarrow N + 1$
    $[T, \mathbf{u}, p] = \mathbf{FE}(\boldsymbol{\mu}^*)$
    $S_T^N = [S_T^{N-1};\ T],\ S_{\mathbf{u}}^N = [S_{\mathbf{u}}^{N-1};\ \mathbf{u}],\ S_p^N = [S_p^{N-1};\ p]$
    $\mathbb{V}_T = \mathbf{POD}(S_T^N),\ \mathbb{V}_{\mathbf{u}} = \mathbf{POD}(S_{\mathbf{u}}^N),\ \mathbb{V}_p = \mathbf{POD}(S_p^N)$
    for $\boldsymbol{\mu} \in \Xi_{train}$
        $\Delta(\boldsymbol{\mu}) = \mathbf{ERROR\,ESTIMATE}(\mathbb{V}_T, \mathbb{V}_{\mathbf{u}}, \mathbb{V}_p, \boldsymbol{\mu})$
    end for
    $\boldsymbol{\mu}^* = \arg \max_{\boldsymbol{\mu} \subset \Xi_{train}} \Delta(\boldsymbol{\mu})$
    $\delta_N = \max_{\boldsymbol{\mu} \subset \Xi_{train}} \Delta(\boldsymbol{\mu})$
end while
```

In the first enrichment step, a full order space-time FE solution is evaluated at a pre-selected parameter value and denoted as a snapshot. To construct the basis for each field, the snapshot is separated $(S_T^N, S_{\mathbf{u}}^N, S_p^N)$ and orthonormalized with the Proper Orthogonal Decomposition (POD) technique. The modes are truncated by prescribing a tolerance that corresponds to their energy contribution relative to the contribution of the first mode. The same tolerance level is prescribed for temperature, displacement and pore pressure fields, which results in a number of truncated modes that can vary accordingly, reflecting the different features of each snapshot field (see Figure 2.6).Using the truncated and orthonormalized matrices as projection bases, the error estimator selects the point in the training sample which gives the worst (least accurate) approximation in the current RB space. The next step is to generate snapshot solutions at this worst point and append it to the RB space to improve its approximation. The enrichment procedure is repeated until the stopping criteria are satisfied.

The most critical ingredient in the RB generation strategy is to ensure that the error estimator gives a good approximation of the exact error, while being computationally inexpensive. Particularly for cases where the training sample is very large, estimating the error at every set of parameter value is an exhaustive procedure.

4 Goal-Oriented Error Estimation

4.1 Error-Residual Representation

Residual-based a-posteriori error estimators are typically employed to certify reduced order models built from projection-based techniques. In goal-oriented error estimation, we focus on specific features of the solution (QoI) and assess the error in these features. The error in the QoI is estimated by relating the primal residual to an appropriate adjoint solution of a dual problem [16, 44, 145, 154]. The ideas introduced in [44, 90, 106], in which an implicitly-solved adjoint is directly injected into the weak primal residual of a parabolic PDE, are followed to develop an a-posteriori goal-oriented error estimator for the coupled THM system.

We represent the error in the the temperature, displacement and pore pressure fields between the reference FE and RB-projected solution as

$$e = (e_T, e_{\mathbf{u}}, e_p) = (T - T_N, \mathbf{u} - \mathbf{u}_N, p - p_N). \tag{2.18}$$

Recalling the weak form of the model problem (2.6) we define the residual as

$$\begin{aligned} R(z, \mathbf{w}, q) &= B(e_T, e_{\mathbf{u}}, e_p; z, \mathbf{w}, q) \\ &= L(z, \mathbf{w}, q) - B(T_N, \mathbf{u}_N, p_N; z, \mathbf{w}, q) \\ &\qquad \forall z \in \mathcal{Z}^H, \mathbf{w} \in \mathcal{W}^H, q \in \mathcal{Q}^H. \end{aligned} \tag{2.19}$$

Denoting the QoI as $J(\cdot)$, we introduce a corresponding dual problem. Let $T_{adj} \in \mathcal{Z}^H$, $\mathbf{u}_{adj} \in \mathcal{W}^H$, $p_{adj} \in \mathcal{Q}^H$ be the set of adjoint solutions to the dual problem such that

$$B(\,\cdot\,; T_{adj}, \mathbf{u}_{adj}, p_{adj}) = J(\,\cdot\,). \tag{2.20}$$

It is aimed to evaluate the QoI, $J(T, \mathbf{u}, p)$, without having to explicitly compute the primal solution $(T, \mathbf{u}, p)$. The QoI is estimated by relating the source term of the primal problem to the adjoint solution of the dual problem.

$$J(T, \mathbf{u}, p) = L(T_{adj}, \mathbf{u}_{adj}, p_{adj}) \tag{2.21}$$

and consequently, we can write the error in the QoI as an adjoint-residual representation

$$\begin{aligned} J(e_T, e_{\mathbf{u}}, e_p) &= R(T_{adj}, \mathbf{u}_{adj}, p_{adj}) \\ &= L(T_{adj}, \mathbf{u}_{adj}, p_{adj}) \\ &\quad - B(T_N, \mathbf{u}_N, p_N; T_{adj}, \mathbf{u}_{adj}, p_{adj}). \end{aligned} \tag{2.22}$$

The implementation of adjoint-residual error representation is simple once the adjoint solution and weak residuals are computed – the adjoint solution is directly plugged in to the residual at every time step, adapting the space-time grid. Particularly for weak residuals $R(\cdot)$ which are explicitly evaluated, the computational cost is minimal even if solved for each parameter value in the training sample. In contrast, computing the backward-in-time-evolving adjoint solution of the dual problem is more involved and computationally expensive.

4.2 Backward-in-Time Evolving Adjoint Solution of a Dual Problem

For the QoI, we consider a general functional output that accounts for the behavior of the solution evolving over time

$$J(T,\mathbf{u},p)=\int_0^{t_f}\int_\Omega \bar{f}_T(\boldsymbol{x},t)T d\Omega\, dt+\int_0^{t_f}\int_\Omega \bar{\mathbf{f}}_{\mathbf{u}}(\boldsymbol{x},t)\mathbf{u} d\Omega\, dt+\int_0^{t_f}\int_\Omega \bar{f}_p(\boldsymbol{x},t)p d\Omega\, dt \tag{2.23}$$

For the sake of simplicity, we normalize the terms to have consistent units with contributions from the temperature, displacement and pore pressure fields resulting to a scalar QoI bearing the unit of the chosen field. By choosing the QoI to have the unit of displacement field, we introduce the following characteristic factors derived from norms of a previously solved primal problem

$$\hat{\gamma}=\|\mathbf{u}\|/\|T\|; \qquad \hat{T}=\hat{\gamma}T \tag{2.24}$$

$$\hat{\delta}=\|\bar{\mathbf{f}}_{\mathbf{u}}\|/\|\bar{f}_T\|; \quad \hat{f}_T=\hat{\delta}\bar{f}_T \tag{2.25}$$

$$\hat{\alpha}=\|\mathbf{u}\|/\|p\|; \qquad \hat{p}=\hat{\alpha}p \tag{2.26}$$

$$\hat{\beta}=\|\bar{\mathbf{f}}_{\mathbf{u}}\|/\|\bar{f}_p\|; \quad \hat{f}_p=\hat{\beta}\bar{f}_p. \tag{2.27}$$

Making use of arbitrary characteristic factors, the QoI in (2.23) is expressed as

$$\begin{aligned}\hat{J}(T,\mathbf{u},p)=\hat{\gamma}\int_0^{t_f}\int_\Omega \bar{f}_T(\boldsymbol{x},t)T\; d\Omega\; dt+\int_0^{t_f}\int_\Omega \bar{\mathbf{f}}_{\mathbf{u}}(\boldsymbol{x},t)\mathbf{u}\; d\Omega\; dt\\ +\hat{\alpha}\int_0^{t_f}\int_\Omega \bar{f}_p(\boldsymbol{x},t)p\; d\Omega\; dt.\end{aligned} \tag{2.28}$$

Recalling the dual problem associated with the coupled THM system,

$$\begin{aligned}B(z,\mathbf{w},q;T_{adj},\mathbf{u}_{adj},p_{adj})=J(z,\mathbf{w},q)\\ \forall z\in\mathcal{Z}^H,\mathbf{w}\in\mathcal{W}^H,q\in\mathcal{Q}^H\end{aligned} \tag{2.29}$$

it is observed that due to the coupling terms in the system of equations, the ensemble of units in the left hand side $B(\,\cdot\,)$ is only compatible with the right hand side of the primal problem $L(\,\cdot\,)$, and not with $J(z, \mathbf{w}, q)$. It becomes critical to express the left hand side of the dual problem (2.29) with normalized units in order to satisfy compatibility with the right hand side vector of the QoI. The dual problem bearing compatible units is therefore written as:

$$\hat{B}(z, \mathbf{w}, q; T_{adj}, \mathbf{u}_{adj}, p_{adj}) = J(z, \mathbf{w}, q) \tag{2.30}$$

where

$$\begin{aligned}
\hat{B}(\cdot\,;\cdot) = {} & \frac{\hat{\delta}}{\hat{\gamma}} \int_0^T \int_\Omega \frac{\partial T_{adj}}{\partial x} k_c \frac{\partial z}{\partial x} d\Omega dt + \int_0^T \int_\Omega \frac{\partial \mathbf{u}_{adj}}{\partial x} G \frac{\partial \mathbf{w}}{\partial x} d\Omega dt \\
& + \int_0^T \int_\Omega \frac{\partial \mathbf{u}_{adj}}{\partial x} (G + \lambda) \frac{\partial \mathbf{w}}{\partial x} d\Omega dt - \frac{1}{\hat{\alpha}} \int_0^T \int_\Omega \frac{\partial p_{adj}}{\partial x} \alpha \mathbf{w} d\Omega dt \\
& - \frac{1}{\hat{\gamma}} \int_0^T \int_\Omega T_{adj}\; 3K_D \alpha_s \frac{\partial \mathbf{w}}{\partial x} d\Omega dt - \frac{\hat{\beta}}{\hat{\alpha}} \int_0^T \int_\Omega \frac{\partial p_{adj}}{\partial x} \kappa \frac{\partial q}{\partial x} d\Omega dt \\
& + \frac{\hat{\delta}}{\hat{\gamma}} \int_\Omega T_{adj}(t = T) \rho c_p z d\Omega - \frac{\hat{\delta}}{\hat{\gamma}} \int_0^T \int_\Omega z \rho c_p \frac{\partial T_{adj}}{\partial t} d\Omega dt \\
& + \frac{\hat{\beta}}{\hat{\alpha}} \int_\Omega p_{adj}(t = T) \frac{1}{M} q d\Omega - \frac{\hat{\beta}}{\hat{\alpha}} \int_0^T \int_\Omega q \frac{1}{M} \frac{\partial p_{adj}}{\partial t} d\Omega dt \\
& + \hat{\beta} \int_\Omega \mathbf{u}_{adj}(t = T) \alpha \frac{\partial q}{\partial x} d\Omega - \hat{\beta} \int_0^T \int_\Omega \frac{\partial q}{\partial x} \alpha \frac{\partial \mathbf{u}_{adj}}{\partial t} d\Omega dt \\
& + \frac{\hat{\beta}}{\hat{\gamma}} \int_\Omega T_{adj}(t = T) [\phi 3\alpha_f + (\alpha - \phi) 3\alpha_s] q d\Omega \\
& - \frac{\hat{\beta}}{\hat{\gamma}} \int_0^T \int_\Omega q [\phi 3\alpha_f + (\alpha - \phi) 3\alpha_s] \frac{\partial T_{adj}}{\partial t} d\Omega dt.
\end{aligned} \tag{2.31}$$

In strong compact form, the dual problem is thus written as:

$$\begin{aligned}
\frac{\hat{\delta}}{\hat{\gamma}} k_c \nabla^2 T_{adj} + \frac{\hat{\delta}}{\hat{\gamma}} \rho c_p \dot{T}_{adj} &= \bar{f}_T \\
G_D \nabla \cdot (\nabla \mathbf{u}_{adj}) + (G_D + \lambda) \nabla (\nabla \cdot \mathbf{u}_{adj}) & \\
- \frac{1}{\hat{\alpha}} \alpha \nabla p_{adj} - \frac{1}{\hat{\gamma}} 3K_D \alpha_s \nabla T_{adj} &= \bar{\mathbf{f}}_{\mathbf{u}} \\
- \frac{\hat{\beta}}{\hat{\alpha}} \frac{k}{\mu_f} \nabla \cdot (\nabla p_{adj}) - \hat{\beta} \alpha \nabla \cdot \dot{\mathbf{u}}_{adj} & \\
- \frac{\hat{\beta}}{\hat{\alpha}} \frac{1}{M} \dot{p}_{adj} + \frac{\hat{\beta}}{\hat{\gamma}} [\phi 3\alpha_f + (\alpha - \phi) 3\alpha_s] \dot{T}_{adj} &= \bar{f}_p \\
\text{in } \Omega \times (0, t_f] &
\end{aligned} \tag{2.32}$$

with Neumann boundary conditions,

$$\begin{aligned} -k_c \nabla_{\mathrm{n}} T_{adj} &= 0 \\ \boldsymbol{\sigma}_{ij,adj} \mathbf{n}_j &= \mathbf{0} \qquad \text{on } \partial\Gamma_N \times (0, t_f] \\ \frac{k}{\mu_f} \nabla_{\mathrm{n}} p_{adj} &= 0 \end{aligned} \tag{2.33}$$

Dirichlet boundary conditions,

$$\begin{aligned} T_{adj} &= 0 \\ \mathbf{u}_{adj} &= 0 \qquad \text{on } \partial\Gamma_D \times (0, t_f] \\ p_{adj} &= 0 \end{aligned} \tag{2.34}$$

and initial conditions

$$\begin{aligned} T_{adj} &= 0 \\ \mathbf{u}_{adj} &= 0 \qquad \text{in } \Omega \times \{t = t^f\}. \\ p_{adj} &= 0 \end{aligned} \tag{2.35}$$

Note that the adjoint solution of each field $(T_{adj}, \mathbf{u}_{adj}, p_{adj})$ bears the same units arising from the characteristic factors $(\hat{\gamma}, \hat{\delta}, \hat{\alpha}, \hat{\beta})$ introduced in the coupled system of equations. This allows for the adjoint to be injected into the residual in a straightforward procedure that yields the error estimate in the QoI,

$$\hat{J}(e_T, e_{\mathbf{u}}, e_p) = \hat{R}(T_{adj}, \mathbf{u}_{adj}, p_{adj}) \tag{2.36}$$

where $\hat{R}$ contains the residuals from the temperature, displacement and pore pressure fields normalized to have consistent units.

The strong form of the dual problem is very similar to the primal problem differing only by the time-dependent terms that have opposite signs − reflecting the backward in time propagation of the adjoint solution. Evidently, the computational cost of solving the dual problem is the same as solving the primal problem.

In this methodology, note that only one adjoint solution is needed to represent the entire parametric training sample. Further details on this assumption will be explained in section 5.1.4.

5 Numerical Examples

5.1 2D Advancing Glacier Problem

As a first illustrative application, we consider a homogeneous region for a fully-saturated rock mass in a 2D domain, see Figure 2.1. A 22000m by 1700m rock mass is subjected to a glacier with non-uniform thickness advancing along the upper surface Γ_{TOP}. The thickness variation of the glacier is described by the function $H(x,t)$ as inspired by reference [136],

$$H(x,t) = 2^{1/8} H_{max} \left(\frac{L - X + v\,t}{L} \right)^{1/2}. \tag{2.37}$$

where $H_{max} = 3200m$ is the maximum height of the glacier, $L = 1329870m$ is the distance from the center to the glacier front, X is the distance from the center of the glacier (the datum point) which is related to the coordinate system of the model domain by $X = x + 1333029m$, v is velocity of the glacier specified as 1.27 m per year. At $t = 0$, the glacier thickness function $H(X)$ satisfies $H(0) = H_{max}$ and $H(L) = 0$. For an observation period of 6000 years, the glacier thus moves 7620m on top of the upper surface Γ_{TOP} of the rock mass. The loading effects of glacier are treated as boundary conditions for the governing equations of the system. These effects are applied as traction ($\boldsymbol{\sigma}_N$) in the mechanical system, and as prescribed pore pressure (p_D) and temperature (T_D) on the upper surface. Additional Neumann and Dirichlet boundary conditions are imposed on the lower surface and edges of the rock mass,

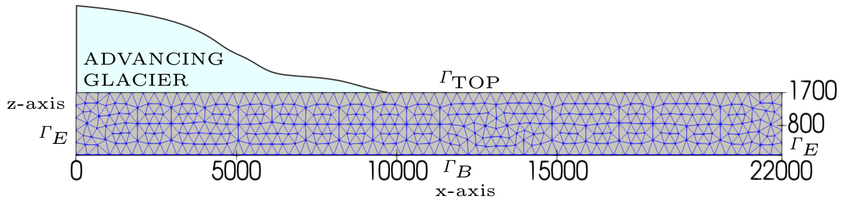


Fig. 2.1 Homogeneous rock mass subjected to glacier loading

leading to

$$\begin{aligned}\boldsymbol{\sigma}_{zz}(x,t) &= \boldsymbol{\sigma}_N = -\rho_f g H(x,t) \\ T(x,t) &= T_D = T_{\max}\,\frac{H(x,t)}{H_{\max}} \quad \text{on} \quad \partial\Gamma_{\mathrm{TOP}} \\ p(x.t) &= p_D = \rho_f g H(x,t)\end{aligned} \tag{2.38}$$

$$\mathbf{u}_z = 0; \quad \frac{\partial p}{\partial z} = 0; \quad \frac{\partial T}{\partial z} = 0 \qquad \text{on} \quad \partial\Gamma_B \tag{2.39}$$

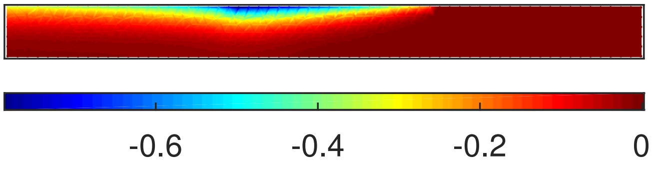
$$\mathbf{u}_n = 0; \quad \frac{\partial p}{\partial n} = 0; \quad \frac{\partial T}{\partial n} = 0 \qquad \text{on} \quad \partial\Gamma_E. \tag{2.40}$$

In this illustrative application, only changes in the rock mass response due to small glacier advances are investigated. Furthermore, the problem is assumed to be linear and that the initial fields at the onset of advancing glacier (i.e. stresses due to the weight of the glacier at its initial position, geostatic stresses, hydrostatic fluid pressure, temperature distribution from geo-thermal heat flux, etc.) can be ignored. Zero initial conditions are therefore imposed on the system.

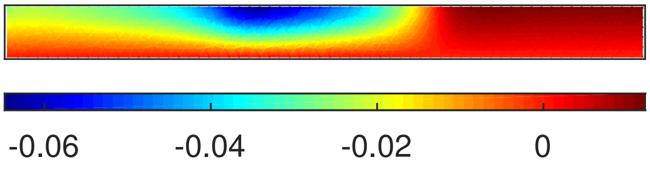
5.1.1 Primal Problem

The FE solution to the coupled THM problem (2.1-2.4) subjected to the glacier loading with the physical properties described in Table 2.2 is obtained from solving the discrete problem in (2.11). For spatial discretization, Taylor-Hood $P_2 - P_1$ elements on triangles are used - quadratic interpolation for displacements and linear interpolation for pore pressure and temperature. There are 4398 degrees of freedom (DOF) in total and the time grid is spaced by 100 logarithmically increasing timesteps.

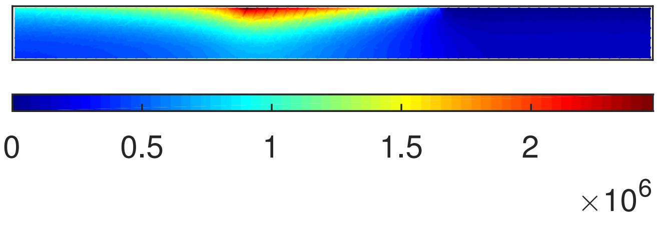
The behavior of the coupled THM system at $t = t_f$, as obtained by a direct FE simulation, is described in Figure 2.2. The pore pressure distribution is maximal closer to the glacier front's initial location as expected. The vertical displacement of the rock mass is negative under the glacier (compression) but exhibits positive values outside the glacier, indicating heaving which is typically observed with glacial loading [50]. The non-uniform temperature distribution under the glacier indicates that steady-state condition is not yet reached after 6000 years.



(a) FE Temperature, T (°C)

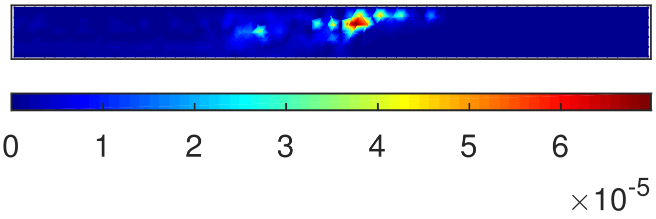


(b) FE z-Displacement, u_z (m)

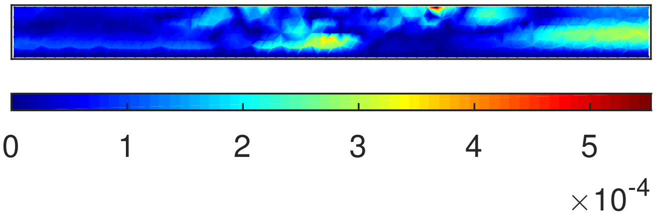


(c) FE Pore Pressure, p (Pa)

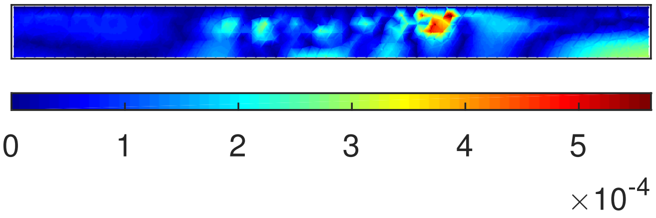
Fig. 2.2 Evolution of thermo-hydro-mechanical fields after 6000 years



(a) $\frac{|T-T_N|}{\|T\|_{\max}}$



(b) $\frac{|u_z-u_{zN}|}{\|u_z\|_{\max}}$



(c) $\frac{|p-p_N|}{\|p\|_{\max}}$

Fig. 2.3 Error map of FE and RB-projected coupled THM solution. Error is expressed relative to the maximum field value of the FE solution.

Table 2.2 Physical parameters describing the homogeneous rock mass

K_D	2×10^{10} Pa
G_D	1.2×10^{10} Pa
α	0.70
M	3.08×10^{10} Pa
μ_f	1×10^{-3} Pa·s
k	1.55×10^{-19} m^2
ϕ	0.05
α_s	8.3×10^{-6} °C^{-1}
α_f	6.9×10^{-7} °C^{-1}
c_p	1.83×10^6 N m$^2 \cdot$ K
k_c	3.66 W/ m $\cdot$ K

5.1.2 Dual Problem

For the setup of the dual problem, we consider the QoI as the average pressure in the domain Ω combined with the average vertical displacement in the upper surface Γ_{TOP} over time $t = (0, t_f]$,

$$\begin{aligned} \hat{J}(T, \mathbf{u}, p) = & \frac{1}{|\Gamma_{\text{TOP}}|} \int_0^{t_f} \int_{\Gamma_T} \mathbf{u} \; d\Gamma_T \; dt \\ & + \hat{\alpha} \frac{1}{|\Omega|} \int_0^{t_f} \int_{\Omega} p \; d\Omega \; dt. \end{aligned} \tag{2.41}$$

Applying the formulation for the backward-in-time evolving dual problem (2.32-2.35), we obtain the adjoint solution corresponding to the temperature, displacement and pressure fields. The dual problem is solved with the same spatial and time discretization as the primal problem. Figure 2.4 shows the backward propagation of the adjoint for the vertical displacement in the upper surface with zero initial condition imposed at the final time.

The adjoint solution is verified by showing that the QoI estimated by relating the source term and adjoint (2.21) holds. Figure 2.5 shows the accuracy of the source term-adjoint estimate with respect to the spatial and time discretization refinement. It is observed that the time discretization of the dual problem affects the QoI estimate convergence more dominantly than the spatial discretization refinement in this particular case where the QoI is chosen as a functional of field solutions in the entire time span $t : (0, t_f]$.

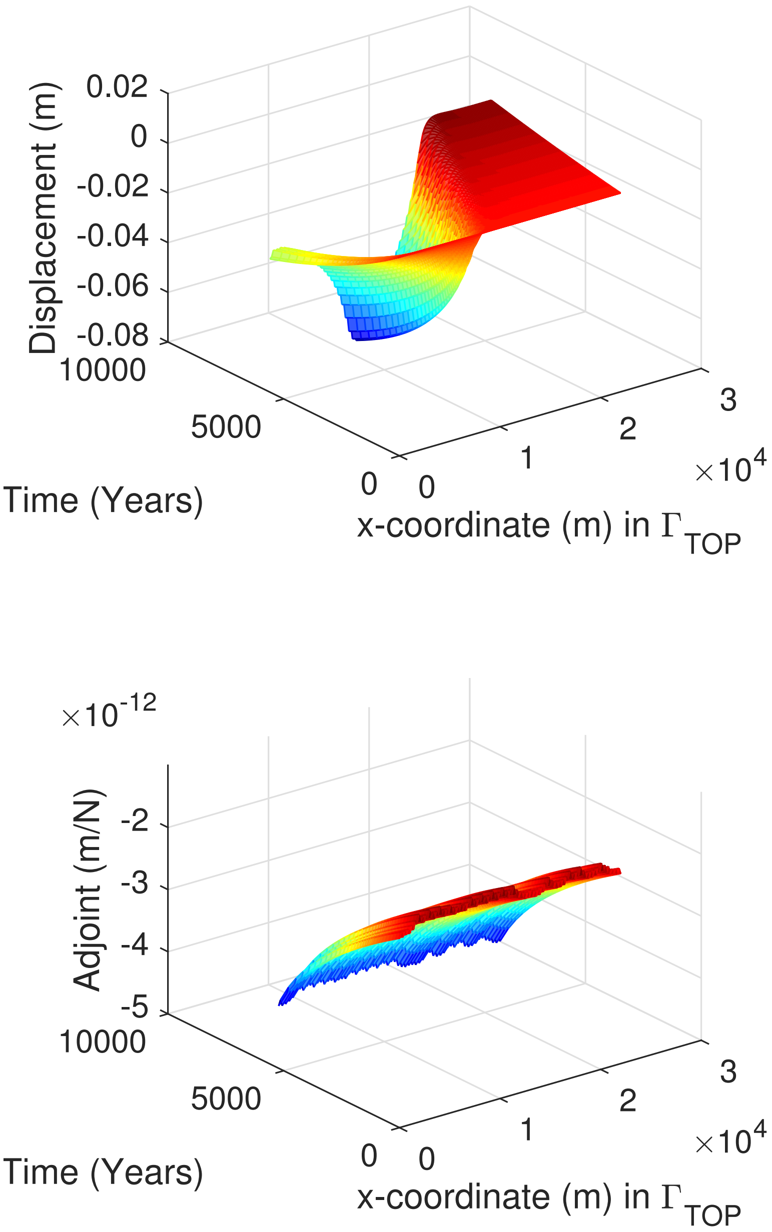


Fig. 2.4 Vertical displacement of the primal problem (top) and the corresponding adjoint solution (bottom) in the upper surface of the domain Γ_{TOP}

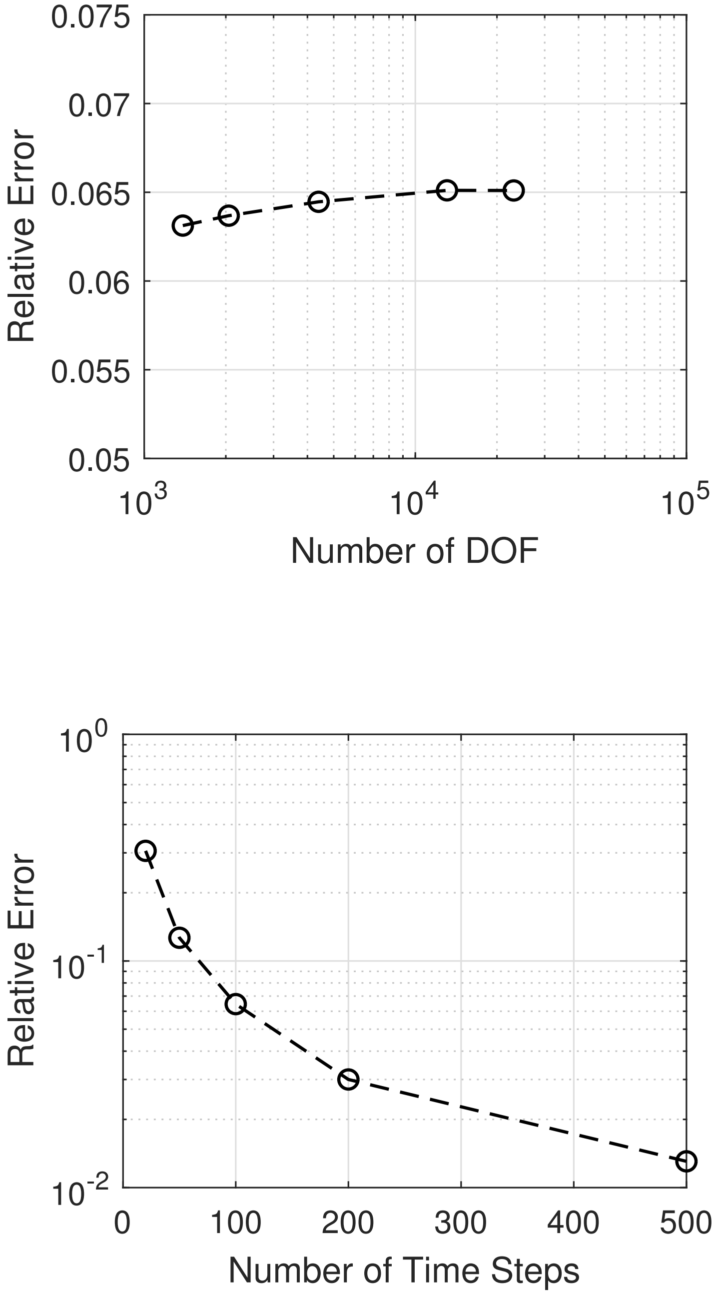


Fig. 2.5 Error plot verifying $J(T, u, p) = L(T_{adj}, u_{adj}, p_{adj})$ with respect to mesh refinement (top) and time step refinement (bottom)

5.1.3 Reduced Basis Generation

This 2D coupled THM problem is parametrized by varying some of the parameters related to material properties. The elastic modulus and permeability corresponding to the mechanical and hydraulic properties are selected to vary within a range of realistic values for a granitic type of rock; elastic modulus $\mu_1 : [1.5 \times 10^{10},\ 4.4 \times 10^{10}]$ Pa and permeability $\mu_2 : [1 \times 10^{-24},\ 1 \times 10^{-18}]$ m^2. The parametric training sample is discretized with 1050 parameter sets ($n_{\mu_1} \times n_{\mu_2} = 35 \times 30$), where each set contains the 2 material properties $\boldsymbol{\mu}_n = (\mu_1, \mu_2)$.

The formulation of the error in the QoI at a certain parameter set is equivalent to the product of the residual and the adjoint solution evaluated in that parameter set. Recalling that the computational cost involved in the solution of the dual problem is the same as in the primal problem, if the adjoint solution is evaluated for each parameter set in the training sample, $(T_{adj}(\boldsymbol{\mu}_n), u_{adj}(\boldsymbol{\mu}_n), p_{adj}(\boldsymbol{\mu}_n))$, it is equivalent to solving the full order (FE) model 1050 times at each enrichment step.

$$\begin{aligned} \hat{J}(e_T(\boldsymbol{\mu}), e_{\mathbf{u}}(\boldsymbol{\mu}), e_p(\boldsymbol{\mu})) : &= \hat{R}(T_{adj}(\boldsymbol{\mu}), \mathbf{u}_{adj}(\boldsymbol{\mu}), p_{adj}(\boldsymbol{\mu})) \\ &= \Delta(\boldsymbol{\mu}) \end{aligned} \tag{2.42}$$

To avoid this extremely expensive procedure, we impose a condition based on the assumption that the adjoint solution is not very sensitive with respect to parametric variations. We therefore solve the dual problem only once in a selected parameter set ($\boldsymbol{\mu}^0$); and this adjoint solution is assumed to be sufficient to be reused for evaluating the error estimate in the entire training sample. The enrichment procedure is stopped once the normalized estimated error in QoI is below the prescribed tolerance level of 10^{-4}.

In the present case, the greedy adaptive procedure required 9 enrichment steps to reach the error level below 10^{-4}. The corresponding number of POD modes at a truncation tolerance of 10^{-5} on their relative amplitude after each enrichment step is shown in Figure 2.6.

The accuracy of the projected RB solution with respect to the FE model is qualitatively assessed in Figure 2.3. The evolution of the temperature, displacement and pore pressure fields after 6000 years is accurately depicted as shown in the error map distribution of the THM fields in the domain. The accuracy is quantitatively assessed in the entire parametric training sample by looking at the global error norm in the domain in Table 2.3. The error norm is computed at every parameter set in the training sample and plotted in Figure 2.7. With a maximum error on the order of 10^{-7} for temperature, and 10^{-3} for displacement and pore pressure, the projected RB solution is in good agreement with the reference FE solution.

Table 2.3 Maximum error level in the entire parametric training sample

$\max\left[\frac{e_T(\boldsymbol{\mu})}{T(\boldsymbol{\mu})}\right]$	$\max\left[\frac{e_{\mathbf{u}}(\boldsymbol{\mu})}{\mathbf{u}(\boldsymbol{\mu})}\right]$	$\max\left[\frac{e_p(\boldsymbol{\mu})}{p(\boldsymbol{\mu})}\right]$
1.8×10^{-7}	1.1×10^{-3}	7.8×10^{-3}

5.1.4 Efficiency of the Goal Oriented Error Estimator

To assess the efficiency of the goal-oriented error estimator, the effectivity index describing the rigor and sharpness of the estimate is measured. The effectivity index is computed as

$$\eta(\boldsymbol{\mu}) = \frac{J(e(\boldsymbol{\mu}))}{\Delta(\boldsymbol{\mu})} \tag{2.43}$$

where $\Delta(\boldsymbol{\mu})$ is the estimated error and $J(e(\boldsymbol{\mu}))$ is the exact error evaluated in the QoI. Ideally, it is desired to have $\eta \equiv 1$ to obtain a sharp bound for the error.

By first evaluating the effectivity index in the parameter, $\boldsymbol{\mu}^0$, where the adjoint is computed (Figure 2.8), it is shown that the error estimator indeed yields a sharp bound with an index value of 1.05 for the given spatial and time discretization (Table 2.4).

To investigate the effect of solving the adjoint only in one parameter set and reusing it for the entire training sample, the sensitivity of the adjoint is evaluated by computing the effectivity indices at different points in the training sample (Figure 2.9). The corresponding effectivity indices at these points are reported in Table 2.4. The error estimator performs well with the maximum underestimation index less than 5 ($\eta(\boldsymbol{\mu}^f) = 3.83$) for the parameter sets tested. As expected, the effectivity index gets worse when the parameter set is further from the reference adjoint parameter $\boldsymbol{\mu}^0$. There is no conclusive observation when an overestimation or an underestimation occurs (i.e. the error is overestimated in both cases where the permeability is lower ($\boldsymbol{\mu}^{\boldsymbol{g}}$) and higher ($\boldsymbol{\mu}^{\boldsymbol{b}}$) than the adjoint's permeability).

5.2 3D Advancing Glacier Problem

To illustrate more quantitatively the computational gain that can be expected from the proposed methodology, the 2D homogeneous rock mass subjected to an advancing glacier model is now extended to a more realistic 3D problem.

Table 2.4 Effectivity index of error estimator within the parameteric training sample

Point	$J(e(\boldsymbol{\mu}))$	$\Delta(\boldsymbol{\mu})$	$\eta(\boldsymbol{\mu})$
$\boldsymbol{\mu}^0$	3.4×10^{-3}	3.2×10^{-3}	1.05
$\boldsymbol{\mu}^a$	1.37×10^{-2}	1.17×10^{-1}	0.12
$\boldsymbol{\mu}^b$	1.35×10^{-2}	1.16×10^{-1}	0.12
$\boldsymbol{\mu}^c$	1.34×10^{-2}	1.15×10^{-1}	0.12
$\boldsymbol{\mu}^d$	3.2×10^{-3}	3.7×10^{-3}	0.87
$\boldsymbol{\mu}^e$	3.4×10^{-3}	3.0×10^{-3}	1.13
$\boldsymbol{\mu}^f$	6.44×10^{-5}	1.68×10^{-5}	3.83
$\boldsymbol{\mu}^g$	2.48×10^{-6}	3.28×10^{-6}	0.75
$\boldsymbol{\mu}^h$	1.75×10^{-5}	2.08×10^{-5}	0.84

The geometry of the 3D homogeneous rock mass (Figure 2.10) is taken from the topography of a valley located in southern Andes. For the FE model, 27-node hexahedral elements are used to interpolate the solution with quadratic shape functions for displacement and linear shape functions for pore pressure and temperature fields to ensure the stability of the coupled system. The 3D domain is a 22000m by 10200m region with non-uniform thickness characterizing the valley with a maximum elevation of +1341m and a minimum elevation of -3985m. The discretized domain consists of 3774 hexahedral elements, 34125 nodes and 115833 DOF. The time discretization is kept the same as in the 2D case with 100 time steps for an observation period of 6000 years.

The QoI is taken as the average of the pressure in the whole domain and the average of the vertical displacement along the upper surface. The glacier advance on the upper surface is described by the same function as in the 2D problem (2.37), imposed at zero-elevation. The elevation profile of the glacier is extruded along the lateral direction and enforced to conform to the rock mass elevation (i.e., when the effective glacier thickness is lower than the rock mass elevation at a specific point, no loading is applied).

To build the reduced basis, the coupled system is parametrized as in the 2D case. The ranges of values for the elastic modulus μ_1 and the permeability μ_2 are varied and discretized to 70 sets ($n_{\mu_1} \times n_{\mu_2} = 10 \times 7$), defining the parametric training sample. Implementing the greedy adaptive strategy, it took 8 enrichment steps to reach the tolerance level below 10^{-4}. The reduced basis after enrichment contains 306 POD modes in total, resulting in RB spaces with characteristic size of 48, 79 and 179 modes for temperature, displacement and pore pressure, respectively.

The dimensionality of the coupled THM system is thus significantly reduced –

having only 306 DOF instead of 115833 DOF to solve the primal problem. Comparing the computational costs, the FE model solution with 115833 DOF requires 17683 seconds (4.6 hours) in MATLAB implementation while the RB-projected system takes 35 seconds to run which yields a computational gain of 505x speed up, see Table 2.5. Note that with even more realistic (i.e. larger) problems and involving more parameters, the expected speed up is even higher.

Table 2.5 Computational gain with RB model in 3D application

	FE Model	RB Model	Gain
DOF	115 833	306	**378:1**
CPU Time	17 683s	35s	**505x***speed up*

The accuracy of the RB-projected solution is assessed by comparing it to the FE model solution evaluated for a reference parameter set value presented in [136] $(3 \times 10^{10}\ \mathrm{Pa}, 1.55 \times 10^{-19}\ \mathrm{m}^2)$, which is not used as a snapshot point to build the RB. Figure 2.11 and Figure 2.12 show that the temperature, displacement and pore pressure field distributions in the domain after 6000 years are recovered well by the RB model. The relative error norms in the QoI at 6000 years are 4.3×10^{-4} for pore pressure and 2.6×10^{-5} for displacement which gives a good level of accuracy for the RB approximation. Further comparison of RB-projected solutions to FE model solutions (Table **??**) is performed for different parameter set values chosen by varying parameter μ_1 or μ_2 with respect to the reference parameter set which is useful for sensitivity analysis of the parametric system.

Table 2.6 Relative error in the pore pressure and displacement fields evaluated in the QoI domain (temperature field is unaffected by μ_1 and μ_2)

μ_1(Pa)	μ_2(m^2)	$\frac{\|e_p^*\|}{\|p^*\|}$	$\frac{\|e_{\mathbf{u}}^*\|}{\|\mathbf{u}^*\|}$
3.0×10^{10}	1.55×10^{-19}	4.3×10^{-4}	2.6×10^{-5}
3.0×10^{10}	6.00×10^{-18}	9.3×10^{-3}	4.4×10^{-4}
2.4×10^{10}	1.55×10^{-19}	3.3×10^{-4}	1.6×10^{-5}
3.5×10^{10}	1.55×10^{-19}	4.1×10^{-4}	2.5×10^{-5}
3.0×10^{10}	1.00×10^{-21}	3.8×10^{-4}	1.5×10^{-5}

* evaluated in the QoI, at time=6000 years

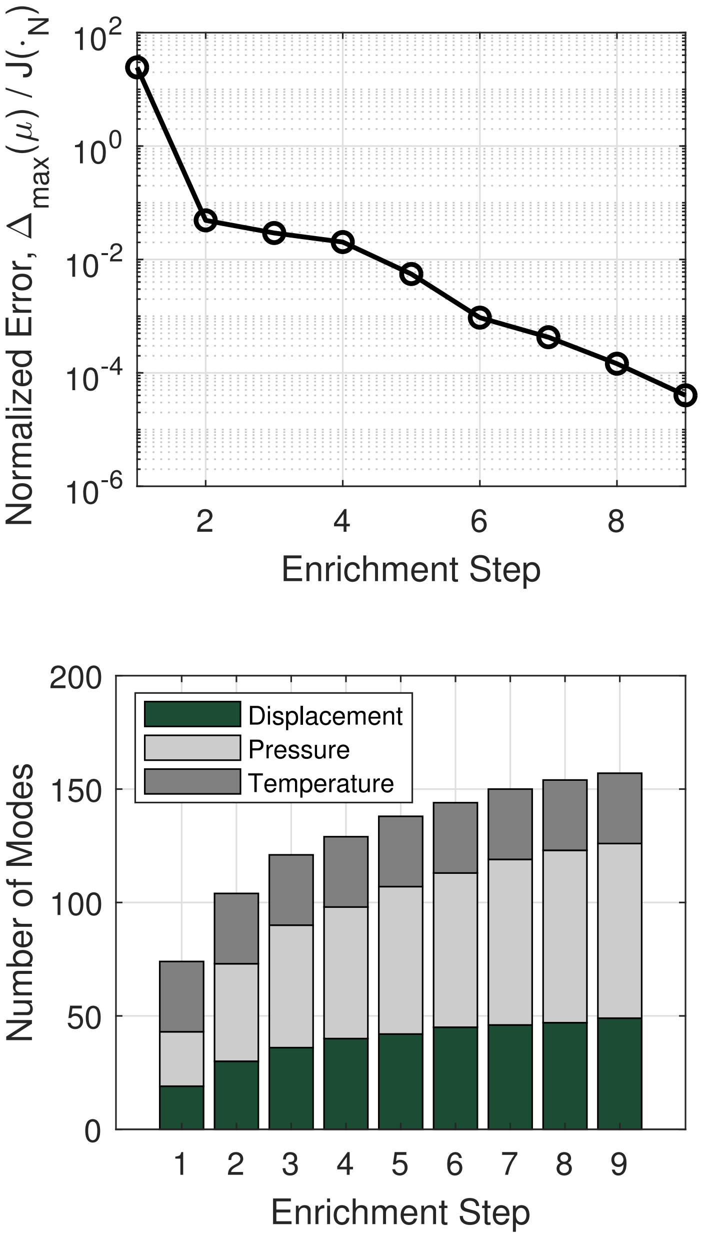
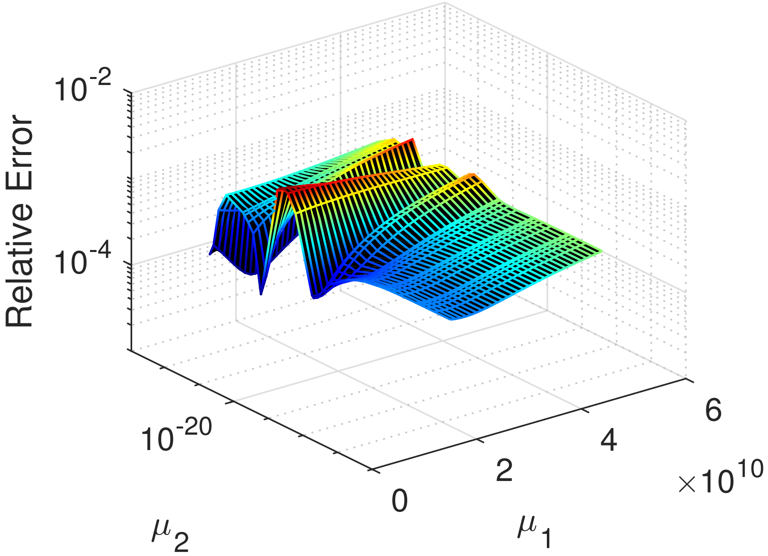


Fig. 2.6 Convergence of greedy-based strategy to the prescribed tolerance level, $\frac{\Delta_{max}(\mu)}{\hat{J}(\cdot_N)}$ (top) and number of POD modes at each enrichment step (bottom)



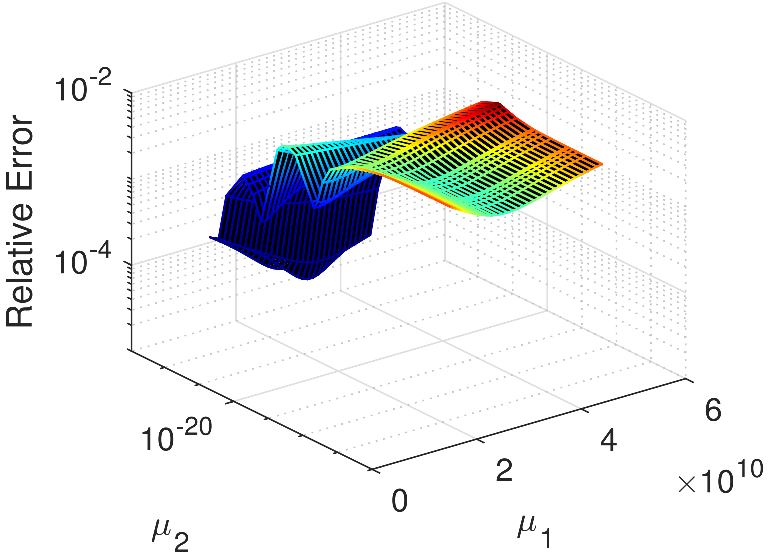


Fig. 2.7 Error in the entire parametric training sample: displacement $\| \frac{e_{\mathbf{u}}}{\mathbf{u}} \|$ (top) and pressure $\| \frac{e_p}{p} \|$ (bottom)

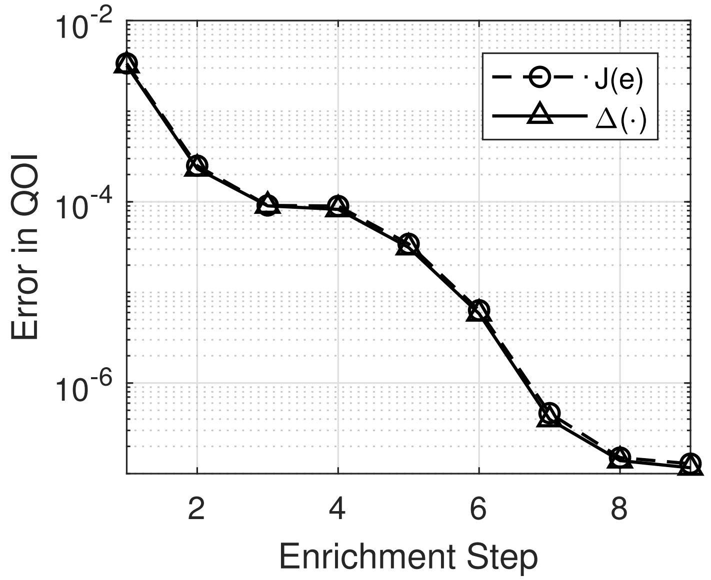


Fig. 2.8 Efficiency of goal-oriented error estimator with respect to basis enrichment - assessed at parameter set where the adjoint is evaluated ($\boldsymbol{\mu}^0$)

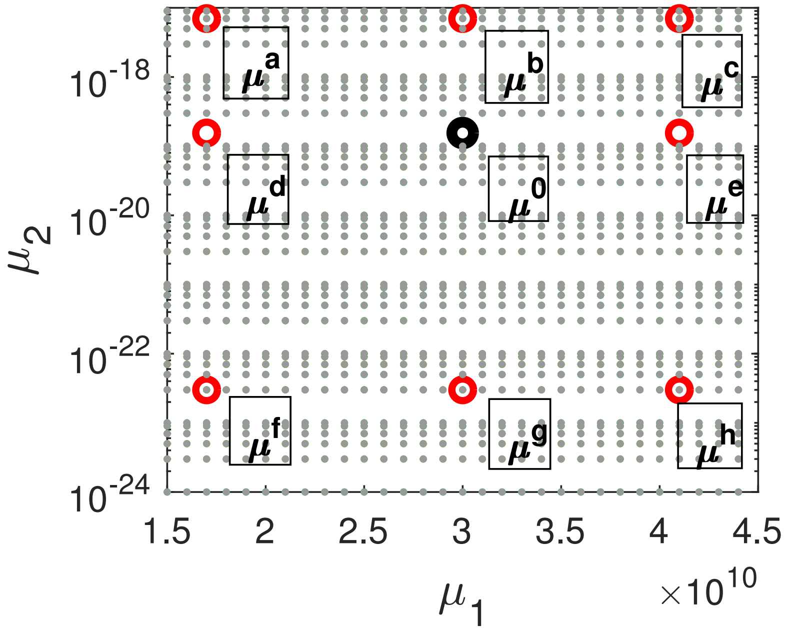


Fig. 2.9 Selected parameters where effectivity index is evaluated

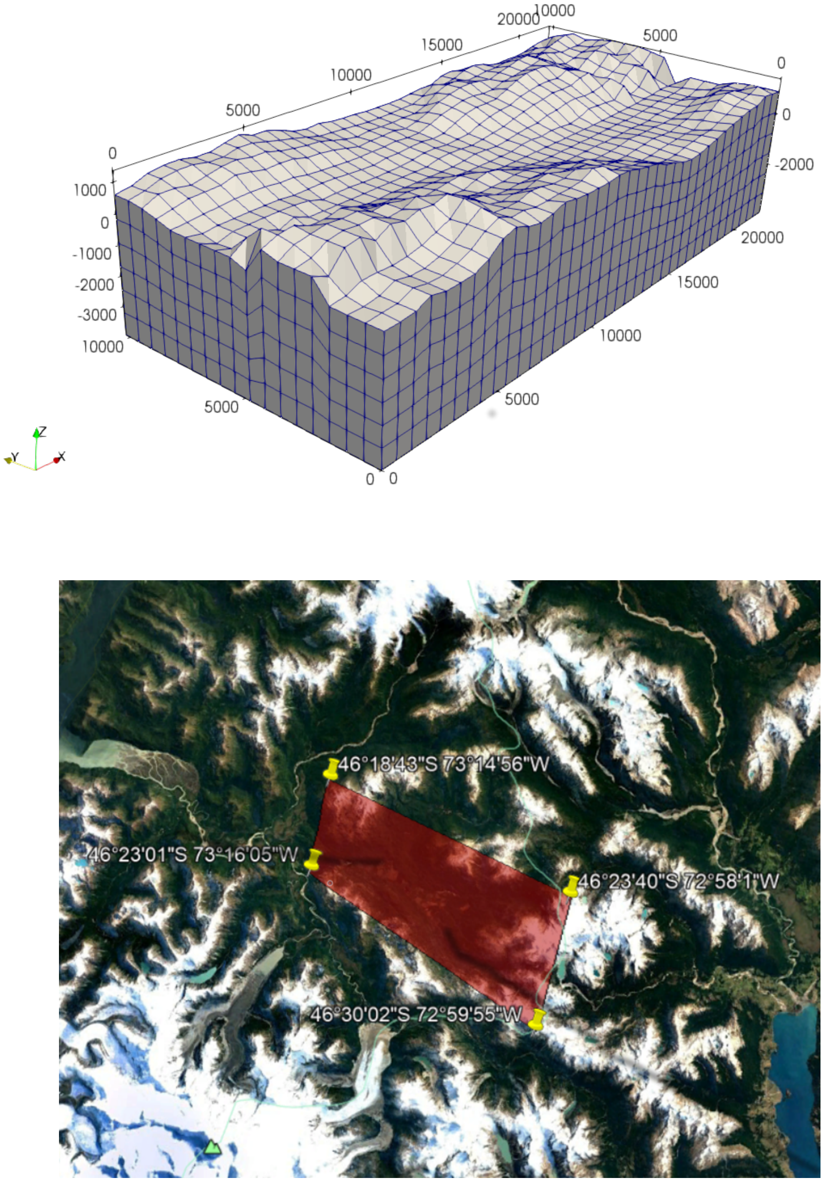
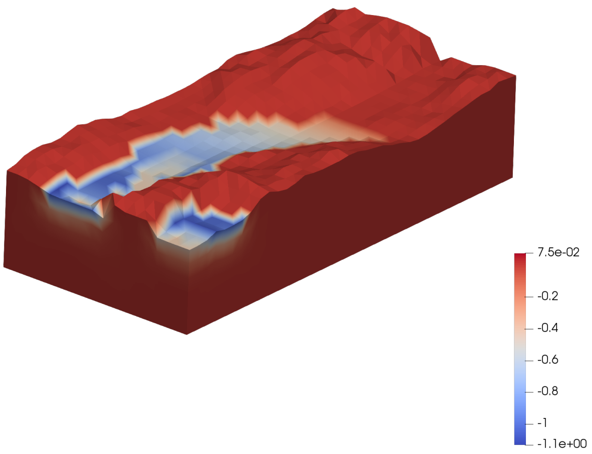
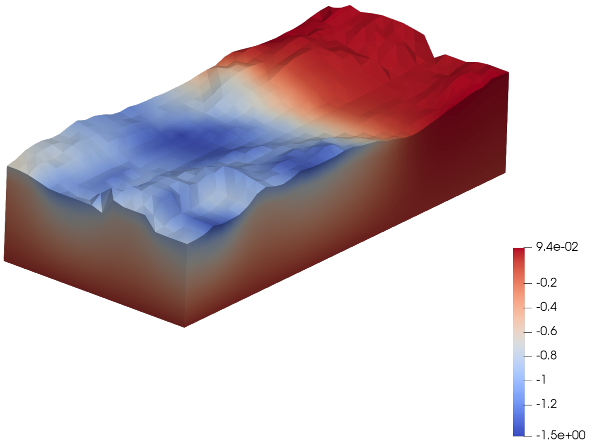


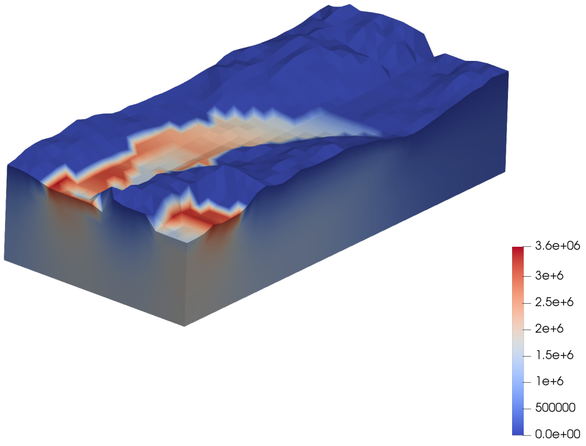
Fig. 2.10 Homegeneous rock mass valley with non-uniform top elevation taken from a region in southern Andes



(a) FE Temperature, T (°C)

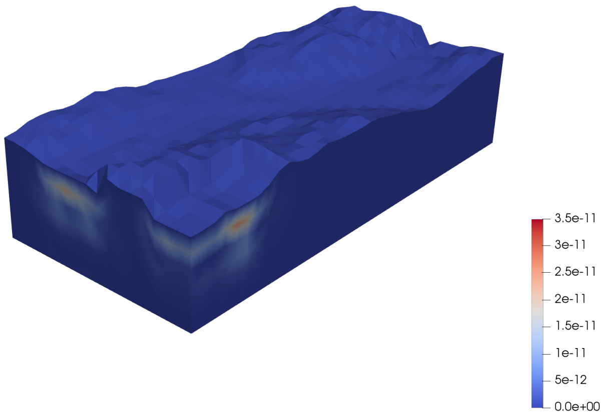


(b) FE z-Displacement, u_z (m)

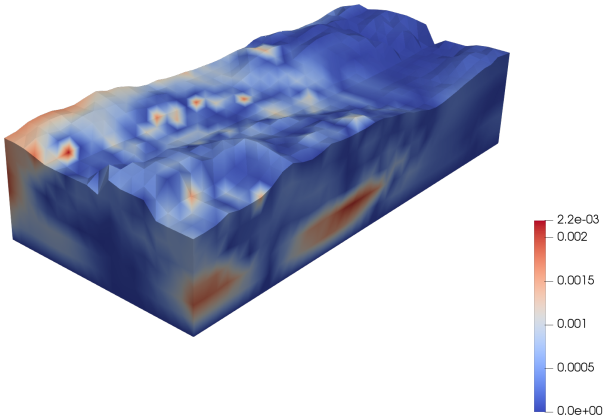


(c) FE Pore Pressure, p (Pa)

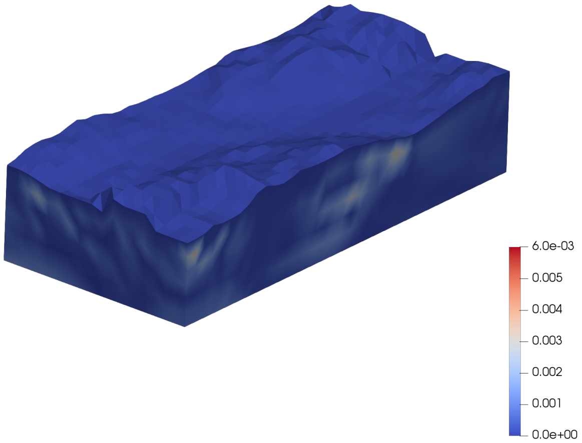
Fig. 2.11 3D Evolution of thermo-hydro-mechanical properties after 6000 years



(a) $\frac{|T-T_N|}{\|T\|_{\max}}$



(b) $\frac{|u_z-u_{zN}|}{\|u_z\|_{\max}}$



(c) $\frac{|p-p_N|}{\|p\|_{\max}}$

Fig. 2.12 Error map of FE and RB-projected coupled THM solution. Error is expressed relative to the maximum field value of the FE solution.

6 Summary and Findings

We presented a strategy to solve transient coupled THM systems in the framework of reduced basis approximation. To preserve the stability properties of the reference FE model, the same interpolation principle is applied in the RB projection by using three separate bases - each basis corresponding respectively to temperature, displacement and pore pressure fields. The greedy adaptive strategy, employed to optimally enrich the bases, strongly relies on the error estimator that provides the criterion to select the next snapshots and certifies the RB-projected solutions.

We developed an efficient a-posteriori error estimator for the coupled THM system which evaluates the error in specific quantities of interest. The strength of this goal-oriented scheme lies in the combination of implicit and explicit error assessment approaches - first, assessing the error implicitly in the dual problem and then injecting it to the explicitly evaluated primal residual. Particularly for a time-dependent system, this error assessment enables a simplified implementation that adapts the space-time grid such that at every time step, the adjoint is simply plugged in to the weak residual. For computational efficiency, it was proven for the application examples that solving the adjoint only once and reusing it for error evaluation in the entire training sample is effective as evident in the sharp error estimates.

The numerical examples demonstrated the ability of the RB strategy to accurately approximate the reference FE solutions of the coupled THM system with a significant reduction in computational costs. Furthermore, higher computational gain is foreseen for more complex and realistic problems that feature non-linearity, heterogeneity and higher-dimensional parametric space. For problems that deal with much larger sets of uncertain parameters, it is worth investigating in future works whether the scheme of solving the dual problem only at one parameter set can provide sufficiently sharp error estimates. Potential strategies (e.g., interpolation and iterative optimization procedures) that tackle such issue are to be explored.

3

Parametrized THM Problems: RB Surrogate Modelling and Inverse Identification for ATLAS III Experiment

This chapter is based on
Larion, Y., Chen, G.J., Zlotnik,S., Díez, P., Seetharam S. , and, Massart, T. J.
Using reduced basis approximation for efficient surrogate-based inverse identification of thermo-hydro-mechanical parameters from an in-situ heating test. *In preparation.*

1 Introduction

In Belgium, an underground research facility (URF) named HADES (High Activity Disposal Experimental Site) was constructed in the Boom Clay layer to investigate the safety and feasibility of deep geological disposal for high level radioactive wastes (HLW), Figure 3.1. The thermo-hydro-mechanical (THM) perturbation of the Boom Clay in response to the heat-emitting HLW has been studied in HADES URF through in-situ heating tests at different scales including a small-scale ATLAS heater test that ran in four phases (ATLAS I, II, III, and IV) between 1993 and 2012. These tests are performed to confirm and refine the THM material properties obtained from previous laboratory experiments.

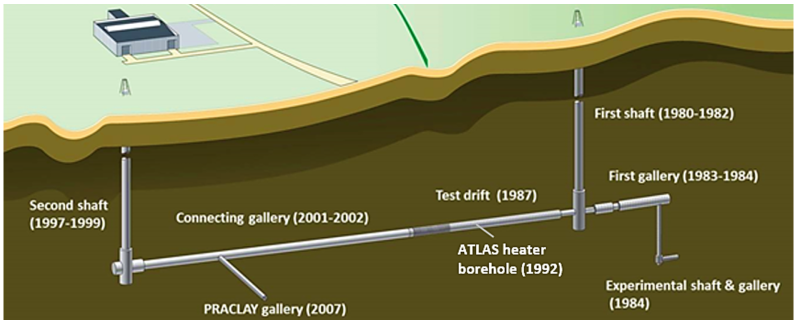


Fig. 3.1 HADES URF in Mol, Belgium

For deep geological disposal, performance assessment of the host formation focused on isolating the HLW from the biosphere requires that in-situ experiments are complemented by numerical simulations to allow predictions of THM behaviour for extended periods of time [7, 29, 37, 51, 134], or to allow investigating the effect of host rock degradation due to excavation [96, 104, 109, 122]. Mathematical models for such coupled geomechanical systems are often characterized by high dimensionality in which the associated computational cost becomes critical when dealing with inverse problems. In previous studies related to this application, the identification of THM parameters in large-scale models was most often restricted to a trial and error approach by manually calibrating the model until experimentally measured data are reproduced [18, 29, 67, 94, 120, 123].

Most computationally intensive inverse identification procedures rely on using surrogates as a substitute to high fidelity forward models with the aim to implement optimization procedures more efficiently. Surrogate models are approximation models built based on repeated forward model simulations probed in a parametric space. The parameters or inputs (e.g. thermal conductivity, hydraulic permeability) configure a particular behaviour of solutions or outputs (e.g. heat flux, flow rates) in the constitutive model. Surrogate modelling techniques can be classified to two different types: physics-based and non-physics based models [48]. Data-fit models are constructed by non-physics-based approximation through interpolation or regression of input-output data resulting in a black box modelling approach. On the other hand, physics-based hierarchical models are constructed from either defining coarser discretization, simplifying physical assumptions or relaxing convergence tolerances. Reduced order models can be regarded as a class of hierarchical surrogates as evident from its physics-dependent projection framework. Aimed at reducing the dimensionality of the system, the high-fidelity forward model is projected to a low dimensional subspace. The subspace is defined by basis functions that are typically constructed using a set of so-called 'snapshots' referring to high-fidelity solutions for specific values of the physical parameters. Numerous studies have implemented reduced order modelling in the framework of inverse problems for different types of applications [14, 52, 102, 110, 147].

We focus our attention on the reduced basis (RB) method - an enrichment-based class of model order reduction that requires an adaptive strategy to construct a low dimensional subspace spanning solutions of a system under parametric variation. At the core of this enrichment procedure, an 'a posteriori error indicator' drives the optimal selection of snapshots used to build the low dimensional subspace underlying the surrogate model. A similar goal-oriented error estimation technique was developed for time-dependent coupled THM systems in [81]. The error indicator, derived based on the explicitly-solved weak residual of the primal problem and on the implicitly-solved adjoint of the dual problem, is extended for the case where some hydro-mechanical parameters are temperature-dependent.

In the present work, a reduced basis method is used to generate surrogate models for solving inverse problems related to coupled processes in ATLAS III Heater test performed in HADES URF [29]. A two-step parameter identification procedure is performed by first identifying the unknown thermal parameters which are subsequently used as inputs to the coupled THM model, followed by identifying the unknown hydro-mechanical parameters in the second stage. The effectiveness of the proposed methodology is demonstrated through a parameter identification procedure using synthetically manufactured data. Furthermore, the feasibility of this approach in real applications is demonstrated by using real sensor measurement data from an in-situ experiment conducted in an underground laboratory.

The chapter is organized as follows. In Section 2, we describe the governing equations for the time-dependent coupled THM system and their discretized forms in both full order and reduced order. In Section 3, we outline the greedy-based adaptive strategy and error estimation used to optimally select the snapshot set during the construction of the low dimensional RB subspace. In Section 4, the parameter identification strategy and the underlying optimization techniques are detailed. Application of the reduced basis generation strategy and inverse problems related to the in-situ experiment are presented in Section 5.

2 Problem Definition: Linear Anisotropic THM System

2.1 Governing Equations and Full Order Discretization

We consider a time-dependent coupled THM system involving a fully-saturated linear poroelastic material. The equations governing the THM system take the form,

$$\left.\begin{aligned} \boldsymbol{k}_c \nabla^2 T - \rho c_p \dot{T} + Q = 0 \\ \nabla \cdot (\boldsymbol{C} : \nabla \mathbf{u}) - \alpha \nabla p - \boldsymbol{\Theta} \nabla T = 0 \\ -\frac{\boldsymbol{k}}{\mu_f(T)} \nabla \cdot (\nabla p) + \alpha \nabla \cdot \dot{\mathbf{u}} + \frac{1}{M} \dot{p} - [\phi 3 \alpha_f + (\alpha - \phi) 3 \alpha_s] \dot{T} = 0 \end{aligned}\right\} \text{in } \Omega \times (0, t_f] \tag{3.1}$$

with Neumann boundary conditions

$$\left.\begin{aligned} \psi \cdot \mathbf{n} = g_N \\ \boldsymbol{\sigma} \cdot \mathbf{n} = \boldsymbol{\sigma}_N \\ \Upsilon \cdot \mathbf{n} = p_N \end{aligned}\right\} \qquad \text{on } \partial\Gamma_N \times (0, t_f], \tag{3.2}$$

Dirichlet boundary conditions

$$\left.\begin{aligned} T = T_D \\ \mathbf{u} = \mathbf{u}_D \\ p = p_D \end{aligned}\right\} \qquad \text{on } \partial\Gamma_D \times (0, t_f], \tag{3.3}$$

and initial condition

$$\left.\begin{aligned} T = T_0 \\ \mathbf{u} = \mathbf{u}_0 \\ p = p_0 \end{aligned}\right\} \qquad \text{in } \Omega \times \{t = 0\}, \tag{3.4}$$

where $T, \mathbf{u}$ and p are the field variables corresponding to temperature, displacement and pore water pressure. The liquid dynamic viscosity is temperature-dependent, obeying the empirical function [56, 146],

$$\mu_f(T) = 2.1 \times 10^{-12} \exp\left(\frac{1808.5}{273.15 + T}\right) \tag{3.5}$$

where μ_f is in MPa · s and T is in Celsius. Assuming cross-anisotropic condition with the horizontal bedding plane and its perpendicular vertical plane denoted by

(h) and (v) subscripts respectively, the fourth order elasticity tensor (3.6) contains five independent coefficients [24, 60]. The physical parameters used in the system of equations are listed in Table 3.1.

$$\boldsymbol{C} = \begin{bmatrix} \frac{E_h(1-\nu_{hh})}{(1+\nu_{hh})(1-2\nu_{hh})} & \frac{E_h\nu_{hh}}{(1+\nu_{hh})(1-2\nu_{hh})} & \frac{E_h\nu_{vh}}{(1+\nu_{hh})(1-2\nu_{hh})} & 0 & 0 & 0 \\ \frac{E_h\nu_{hh}}{(1+\nu_{hh})(1-2\nu_{hh})} & \frac{E_h(1-\nu_{hh})}{(1+\nu_{hh})(1-2\nu_{hh})} & \frac{E_h\nu_{vh}}{(1+\nu_{hh})(1-2\nu_{hh})} & 0 & 0 & 0 \\ \frac{E_h\nu_{vh}}{(1+\nu_{hh})(1-2\nu_{hh})} & \frac{E_h\nu_{vh}}{(1+\nu_{hh})(1-2\nu_{hh})} & \frac{E_v(1-\nu_{hh})}{(1+\nu_{hh})(1-2\nu_{hh})} & 0 & 0 & 0 \\ 0 & 0 & 0 & G_{vh} & 0 & 0 \\ 0 & 0 & 0 & 0 & G_{vh} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{E_h}{2(1+\nu_{hh})} \end{bmatrix} \tag{3.6}$$

Table 3.1 Physical parameters of the THM system

Parameters	
G	shear modulus
E	Young's modulus
ν	Poisson's ratio
α	Biot-Willis coefficient
M	constrained specific storage coefficient
μ_f	viscosity
k	permeability
ϕ	porosity
Θ	thermal expansion factor
α_s	thermal expansion coefficient in the solid phase
α_f	thermal expansion coefficient in the fluid phase
c_p	bulk specific heat
$c_{p,s}$	solid specific heat
k_c	thermal conductivity
Q	heat source
ψ	thermal flux
σ	traction
Υ	fluid velocity

The solution to (3.1) -(3.4) is approximated with the finite element (FE) method by discretizing the domain Ω using a mesh of characteristic size H in the functional space $\mathcal{Z}^H \subset \mathcal{H}_0^1(\Omega)$, $\boldsymbol{\mathcal{W}}^H \subset [\mathcal{H}_0^1(\Omega)]^2$, $\mathcal{Q}^H \subset \mathcal{H}_0^1(\Omega)$:

$$\begin{aligned}
T(\boldsymbol{x},t) \approx T^H(\boldsymbol{x},t) &= \sum_{i=1}^{\mathtt{n_{nodes}}} T_i(t) N_i(\boldsymbol{x}) = \mathbf{N}^\mathsf{T} \boldsymbol{T} \\
\mathbf{u}(\boldsymbol{x},t) \approx \mathbf{u}^H(\boldsymbol{x},t) &= \sum_{i=1}^{\mathtt{n_{nodes}}} \mathbf{u}_i(t) \tilde{N}_i(\boldsymbol{x}) = \tilde{\mathbf{N}}^\mathsf{T} \boldsymbol{U} \\
p(\boldsymbol{x},t) \approx p^H(\boldsymbol{x},t) &= \sum_{i=1}^{\mathtt{n_{nodes}}} p_i(t) N_i(\boldsymbol{x}) = \mathbf{N}^\mathsf{T} \boldsymbol{p}
\end{aligned} \tag{3.7}$$

where $\boldsymbol{U} \in \mathbb{R}^{d \times \mathtt{n_{nodes}}}$, $\mathtt{n_{nodes}}$ is the number of nodes discretizing the field in Ω, while $\mathbf{N}$ and $\tilde{\mathbf{N}}$ are compact support shape functions. The weak form is written in a compact format as

$$B(T,\mathbf{u},p;z,\mathbf{w},q) = L(z,\mathbf{w},q) \quad \forall z \in \mathcal{Z}^H, \mathbf{w} \in \boldsymbol{\mathcal{W}}^H, q \in \mathcal{Q}^H \tag{3.8}$$

where

$$\begin{aligned}
B(T,\mathbf{u},p;z,\mathbf{w},q) = &\int_\Omega \boldsymbol{k}_c \frac{\partial z}{\partial x_j}\frac{\partial T}{\partial x_j} d\Omega + \int_\Omega z \rho c_p \frac{\partial T}{\partial t} d\Omega \\
&+ \int_\Omega \boldsymbol{C}_{ijkl} \frac{\partial \mathbf{w}_i}{\partial x_j}\frac{\partial \mathbf{u}_k}{\partial x_l} d\Omega - \int_\Omega \mathbf{w}_i \alpha \frac{\partial p}{\partial x_i} d\Omega \\
&- \int_\Omega \mathbf{w}_i \boldsymbol{\Theta} \frac{\partial T}{\partial x_i} d\Omega \\
- \int_\Omega \frac{\boldsymbol{k}}{\mu_f}\frac{\partial q}{\partial x_j}\frac{\partial p}{\partial x_j} d\Omega &+ \int_\Omega q \frac{1}{M}\frac{\partial p}{\partial t} d\Omega + \int_\Omega q\alpha \frac{\partial}{\partial t}(\frac{\partial \mathbf{u}_i}{\partial x_i}) d\Omega \\
&- \int_\Omega q[\phi 3\alpha_f + (\alpha - \phi) 3\alpha_s] \frac{\partial T}{\partial t} d\Omega
\end{aligned} \tag{3.9}$$

for $i,j,k,l = 1,2,3$ and

$$L(z,\mathbf{w},q) = \int_\Gamma z g_N d\Gamma + \int_\Gamma \mathbf{w}\boldsymbol{\sigma}_N d\Gamma - \int_\Gamma q p_N d\Gamma. \tag{3.10}$$

The substitution of the approximation (3.7) in the weak form (3.8) leads to the discrete matrix form,

$$\mathbb{K}\,\mathbb{U} + \mathbb{C}\,\dot{\mathbb{U}} = \mathbb{F} \tag{3.11}$$

$$\underbrace{\begin{bmatrix}\mathbf{K_H} & 0 & 0\\ \mathbf{K_T} & \mathbf{K_U} & \mathbf{C_P}\\ 0 & 0 & \mathbf{K_P}\end{bmatrix}}_{\mathbb{K}}\underbrace{\begin{bmatrix}\boldsymbol{T}\\ \boldsymbol{U}\\ \boldsymbol{p}\end{bmatrix}}_{\mathbb{U}}+\underbrace{\begin{bmatrix}\mathbf{C_H} & 0 & 0\\ 0 & 0 & 0\\ \mathbf{C_T} & \mathbf{C_P^\intercal} & \mathbf{C_M}\end{bmatrix}}_{\mathbb{C}}\underbrace{\begin{bmatrix}\dot{\boldsymbol{T}}\\ \dot{\boldsymbol{U}}\\ \dot{\boldsymbol{p}}\end{bmatrix}}_{\dot{\mathbb{U}}}=\underbrace{\begin{bmatrix}\mathbf{F_T}\\ \mathbf{F_u}\\ \mathbf{F_p}\end{bmatrix}}_{\mathbb{F}} \tag{3.12}$$

where $\mathbb{K}$ contains the stiffness and conductivity matrices, $\mathbb{C}$ contains the capacity (time-dependent) matrices and $\mathbb{F}$ contains the vectors corresponding to $L(\cdot)$.

For the time discretization, one considers a time-discrete framework in the interval $I \equiv (0, t_f]$ such that the subintervals are denoted as $\{0 = t^0 < t^1 < ... < t^l < .... < t^{\mathcal{L}} = t_f\}$ for $l = 1, 2, ..., \mathcal{L}$. For notation purposes, let Δt^l be the characteristic time step of the time-discrete framework, measured as $\Delta t^l = t^l - t^{l-1}$. To perform the time integration, an Euler-Backward scheme is used and the space-time discrete problem is written as

$$\Delta t^l \mathbb{K}\mathbb{U}^{\,l} + \mathbb{C}\,\mathbb{U}^{\,l} = \mathbb{C}\,\mathbb{U}^{\,l-1} + \Delta t^l \mathbb{F}^{\,l} \qquad \forall l \in 1, 2, ..., \mathcal{L} \tag{3.13}$$

In this study, the effect of the hydro-mechanical processes on the thermal aspect can be neglected due to the low permeability and small deformation of the Boom Clay. This unidirectional coupling permits separation of the thermal balance equation from the monolithic system (3.11). In the case of temperature-dependent viscosity, this separation allows preserving the linearity in the system of equations since temperature will be obtained independently of the hydro-mechanical system. The non-monolithic system that involves a two-step solving procedure: first, solving the thermal balance equation and then using the temperature solution as input variable for the hydro-mechanical system, is expressed as

$$\mathbf{K_H}\ \boldsymbol{T} + \mathbf{C_H}\ \dot{\boldsymbol{T}} = \mathbf{F_T} \tag{3.14}$$

$$\begin{bmatrix}\mathbf{K_U} & \mathbf{C_P}\\ 0 & \mathbf{K_P}\end{bmatrix}\begin{bmatrix}\boldsymbol{U}\\ \boldsymbol{p}\end{bmatrix}+\begin{bmatrix}0 & 0\\ \mathbf{C_P^\intercal} & \mathbf{C_M}\end{bmatrix}\begin{bmatrix}\dot{\boldsymbol{U}}\\ \dot{\boldsymbol{p}}\end{bmatrix}=\begin{bmatrix}\mathbf{F_u} - \mathbf{K_T}\boldsymbol{T}\\ \mathbf{F_p} - \mathbf{C_T}\dot{\boldsymbol{T}}\end{bmatrix}. \tag{3.15}$$

To avoid instability in the system, Taylor-Hood finite element spaces are defined by assigning quadratic basis functions for displacement and linear basis functions for pore pressure and temperature [83, 93, 157]. The ‘high fidelity full order model’ referred to in the next sections is described by the finite element approximation in (3.14 - 3.15).

2.2 Reduced Order Model Form

For a set of parameter values $\mathcal{M} = \{\boldsymbol{\mu}^1, ..., \boldsymbol{\mu}^N\}$, the snapshots are denoted as the corresponding solutions $S^N = \{(T(\boldsymbol{\mu}^1), \mathbf{u}(\boldsymbol{\mu}^1), p(\boldsymbol{\mu}^1)), ..., (T(\boldsymbol{\mu}^N), \mathbf{u}(\boldsymbol{\mu}^N),\ p(\boldsymbol{\mu}^N))\}$ where N is typically much smaller than the characteristic size of the reference full order model. To inherit the stability condition of the reference full order model, the global approximation spaces are defined separately for each field variable:

$$\begin{aligned} \mathcal{Z}_N^H &= \operatorname{span}\{T(\boldsymbol{\mu}^n), \quad 1 \leq n \leq \mathtt{n}_{\mathtt{snapshots}}\} \subset \mathcal{Z}^H \\ \boldsymbol{\mathcal{W}}_N^H &= \operatorname{span}\{\mathbf{u}(\boldsymbol{\mu}^n), \quad 1 \leq n \leq \mathtt{n}_{\mathtt{snapshots}}\} \subset \boldsymbol{\mathcal{W}}^H \\ \mathcal{Q}_N^H &= \operatorname{span}\{p(\boldsymbol{\mu}^n), \quad 1 \leq n \leq \mathtt{n}_{\mathtt{snapshots}}\} \subset \mathcal{Q}^H \end{aligned} \tag{3.16}$$

The Galerkin projection of the reference model to global approximation (RB) spaces starts by decoupling the system. In the discrete form, elements of stiffness, conductivity and capacity matrices are extracted and grouped into blocks based on the field variable it interacts with (3.17). With $\mathbb{Z}, \mathbb{W}$ and $\mathbb{Q}$ denoting the RB matrices for temperature, displacement and pore pressure respectively, each block is projected to its corresponding bases and assembled similar to the reference FE matrix system. The resulting blocks are described by a much smaller characteristic size of order N.

$$\begin{bmatrix} \mathbf{K}_{\mathbf{H}N} & 0 & 0 \\ \mathbf{K}_{\mathbf{T}N} & \mathbf{K}_{\mathbf{U}N} & \mathbf{C}_{\mathbf{P}N} \\ 0 & 0 & \mathbf{K}_{\mathbf{P}N} \end{bmatrix} \begin{bmatrix} \boldsymbol{T}_N \\ \boldsymbol{U}_N \\ \boldsymbol{p}_N \end{bmatrix} + \begin{bmatrix} \mathbf{C}_{\mathbf{H}N} & 0 & 0 \\ 0 & 0 & 0 \\ \mathbf{C}_{\mathbf{T}N} & \mathbf{C}_{\mathbf{P}N}^{\mathsf{T}} & \mathbf{C}_{\mathbf{M}N} \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{T}}_N \\ \dot{\boldsymbol{U}}_N \\ \dot{\boldsymbol{p}}_N \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{\mathbf{T}N} \\ \mathbf{F}_{\mathbf{u}N} \\ \mathbf{F}_{\mathbf{p}N} \end{bmatrix} \tag{3.17}$$

$$\begin{array}{ccc} \mathbf{K}_{\mathbf{H}N} = \mathbb{Z}^{\mathsf{T}}\ \mathbf{K}_{\mathbf{H}}\ \mathbb{Z} & \mathbf{K}_{\mathbf{T}N} = \mathbb{W}^{\mathsf{T}}\ \mathbf{K}_{\mathbf{T}}\ \mathbb{Z} & \mathbf{K}_{\mathbf{U}N} = \mathbb{W}^{\mathsf{T}}\ \mathbf{K}_{\mathbf{U}}\ \mathbb{W} \\ \mathbf{C}_{\mathbf{P}N} = \mathbb{W}^{\mathsf{T}}\ \mathbf{C}_{\mathbf{P}}\ \mathbb{Q} & \mathbf{K}_{\mathbf{P}N} = \mathbb{Q}^{\mathsf{T}}\ \mathbf{K}_{\mathbf{P}}\ \mathbb{Q} & \mathbf{C}_{\mathbf{H}N} = \mathbb{Z}^{\mathsf{T}}\ \mathbf{C}_{\mathbf{H}}\ \mathbb{Z} \\ \mathbf{C}_{\mathbf{T}N} = \mathbb{Q}^{\mathsf{T}}\ \mathbf{C}_{\mathbf{T}}\ \mathbb{Z} & \mathbf{C}_{\mathbf{P}N}^{\mathsf{T}} = \mathbb{Q}^{\mathsf{T}}\ \mathbf{C}_{\mathbf{P}}^{T}\ \mathbb{W} & \mathbf{C}_{\mathbf{M}N} = \mathbb{Q}^{\mathsf{T}}\ \mathbf{C}_{\mathbf{M}}\ \mathbb{Q} \\ \mathbf{F}_{\mathbf{T}N} = \mathbb{Z}^{\mathsf{T}}\ \mathbf{F}_{\mathbf{T}} & \mathbf{F}_{\mathbf{u}N} = \mathbb{W}^{\mathsf{T}}\ \mathbf{F}_{\mathbf{u}} & \mathbf{F}_{\mathbf{p}N} = \mathbb{Q}^{\mathsf{T}}\ \mathbf{F}_{\mathbf{p}} \\ \boldsymbol{T}_N = \mathbb{Z}^{\mathsf{T}}\ \boldsymbol{T} & \boldsymbol{U}_N = \mathbb{W}^{\mathsf{T}}\ \boldsymbol{U} & \boldsymbol{p}_N = \mathbb{Q}^{\mathsf{T}}\ \boldsymbol{p} \end{array} \tag{3.18}$$

For a specific value of each, physical parameter, $\boldsymbol{\mu} \in \mathcal{P}$, the approximated solution belonging to the RB spaces $(T_N(\boldsymbol{\mu}) \in \mathcal{Z}_N^H,\ \mathbf{u}_N(\boldsymbol{\mu}) \in \mathcal{W}_N^H,\ p_N(\boldsymbol{\mu}) \in \mathcal{Q}_N^H)$ is obtained by solving

$$B(T_N(\boldsymbol{\mu}), \mathbf{u}_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu}); z, \mathbf{w}, q) = L(z, \mathbf{w}, q) \quad \forall z \in \mathcal{Z}_N^H,\ \mathbf{w} \in \mathcal{W}_N^H,\ q \in \mathcal{Q}_N^H \tag{3.19}$$

where the dimensionality of the system corresponding to the left hand side $B(\,\cdot\,)$ and right hand side $L(\,\cdot\,)$ terms, as described in (3.9) and (3.10), is much lower than the reference full order model. In the subsequent sections, we will denote the reduced variables (3.17) with the subscript N.

3 Reduced Basis Surrogate Model

3.1 Reduced Basis Generation

The reduced basis is built from an assembly of optimally selected snapshots evaluated within a given parametric training sample defined by specific values of the physical properties in Table 3.1. It is desired that the selected snapshots yield the most accurate reduced basis approximation, while keeping the number of snapshots minimal. An adaptive strategy driven by an a-posteriori error estimator is used to efficiently carry out the snapshot selection procedure. In turn, the error estimator provides certification to the RB-approximated solutions.

For the adaptive strategy, hierarchical RB approximation spaces are constructed from a greedy-based sampling framework [62, 63, 87, 107]. An enrichment procedure is implemented in which newly-selected snapshots are appended to enrich the reduced basis at every step (Algorithm 2.1).

At every enrichment step, a full order solution is evaluated at the optimally selected parameter value $\boldsymbol{\mu}^*$ and denoted as the optimal snapshot. Three bases are constructed for each field variable by decoupling $(S_T^M, S_{\mathbf{u}}^M, S_p^M)$ and orthonormalizing the optimal snapshot with the Proper Orthogonal Decomposition (POD) technique. POD modes are truncated based on a prescribed tolerance value corresponding to the desired level of accuracy. Using the truncated and orthonormalized modes as projection bases, the worst point in the training sample is located by finding the parameter values in which the largest error estimate occurs for the current RB space. A new snapshot solution is then generated at this worst point and appended to the current RB space to improve its approximation. The enrichment procedure is repeated until the stopping criteria are satisfied.

As the main driver of the greedy adaptive strategy, it is critical to ensure that the error estimates give a good approximation of the exact error while being computationally inexpensive. For cases where the training sample is very large indeed, estimating the error at every set of parameters in the sampling space is an exhaustive and expensive procedure.

3.2 Goal-Oriented Error Estimate

Goal-oriented error estimation provides a very effective tool for numerical simulations where one is particularly interested in specific features of the solution as it selectively gives information about a more relevant error in the quantity of interest (QoI) instead of error estimates derived from energy norms. The error in the QoI is estimated by a dual weighted residual which relates the primal residual to an appropriate adjoint solution of a dual problem. Following the methodologies developed in [81], we obtain the error estimates for the coupled THM system by directly injecting the adjoint to the weak primal residual

$$\hat{J}(e_T, e_{\mathbf{u}}, e_p) = \hat{R}(T_{adj}, \mathbf{u}_{adj}, p_{adj}) \tag{3.20}$$

$$(e_T, e_{\mathbf{u}}, e_p) = (T - T_N, \mathbf{u} - \mathbf{u}_N, p - p_N) \tag{3.21}$$

where $(T_{adj}, \mathbf{u}_{adj}, p_{adj})$ are adjoint solutions, $J(\cdot)$ is a linear functional describing the QoI, $R(\cdot)$ contains the residuals from the temperature, displacement and pore pressure fields and the $[\hat{\ }]$ superscript denotes vectors normalized to obtain consistent units. Recalling the weak form of the THM model, the residual is defined as

$$\begin{aligned} R(z, \mathbf{w}, q) &= B(e_T, e_{\mathbf{u}}, e_p; z, \mathbf{w}, q) \\ &= L(z, \mathbf{w}, q) - B(T_N, \mathbf{u}_N, p_N; z, \mathbf{w}, q) \quad \forall z \in \mathcal{Z}^H, \mathbf{w} \in \mathcal{W}^H, q \in \mathcal{Q}^H. \end{aligned} \tag{3.22}$$

Consider a quantity of interest described by the functional $(\bar{\mathrm{f}}_T, \bar{\mathbf{f}}_{\mathbf{u}}, \bar{\mathrm{f}}_p)$ for temperature, displacement and pore pressure, the dual problem for the coupled THM system is derived in strong compact form as:

$$\left.\begin{aligned} \frac{\hat{\delta}}{\hat{\gamma}} \boldsymbol{k}_c \nabla^2 T_{adj} + \frac{\hat{\delta}}{\hat{\gamma}} \rho c_p \dot{T}_{adj} &= \bar{\mathrm{f}}_T \\ \nabla \cdot (\boldsymbol{C} : \nabla \mathbf{u}_{adj}) - \frac{1}{\hat{\alpha}} \alpha \nabla p_{adj} - \frac{1}{\hat{\gamma}} \boldsymbol{\Theta} \nabla T_{adj} &= \bar{\mathbf{f}}_{\mathbf{u}} \\ -\frac{\hat{\beta}}{\hat{\alpha}} \frac{\boldsymbol{k}}{\mu_f(T)} \nabla \cdot (\nabla p_{adj}) - \hat{\beta} \alpha \nabla \cdot \dot{\mathbf{u}}_{adj} - \frac{\hat{\beta}}{\hat{\alpha}} \frac{1}{M} \dot{p}_{adj} & \\ + \frac{\hat{\beta}}{\hat{\gamma}} [\phi 3 \alpha_f + (\alpha - \phi) 3 \alpha_s] \dot{T}_{adj} &= \bar{\mathrm{f}}_p \end{aligned} \right\} \text{ in } \Omega \times (0, t_f] \tag{3.23}$$

with zero flux Neumann boundary conditions in $\partial\Gamma_N$, zero valued Dirichlet boundary conditions in $\partial\Gamma_D$ and zero final conditions imposed at $t = t^f$. The dual problem formulation is very similar to the primal problem differing only by the time-dependent terms bearing opposite signs which reflects the backward in time propagation of the adjoint solution. Characteristic factors $(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$ derived from norms of previously solved primal problem are also introduced to ensure compatibility of variables within the dual problem. Assuming that the adjoint sensitivity with respect to variation of material properties in the THM system is low, only one adjoint solution is needed to be computed to represent the entire parametric training sample [81].

4 Inverse Analysis: Two-Step Parameter Identification

Exploiting the unidirectional coupling of the thermal system, it is possible to separate the THM system into two systems (thermal, hydro-mechanical) and solve them sequentially(3.14-3.15). In the context of parameter identification where one is interested in identifying material parameters in both separated systems, this entails devising a two-step parameter identification procedure. In the first step, thermal parameters are sought to match a given measurement data. The temperature evolution generated from the identified parameters will be used as input to compute the necessary temperature-dependent material properties to build the hydro-mechanical model. In the second step, material parameters related to hydro-mechanical system are identified based on another set of measurement data.

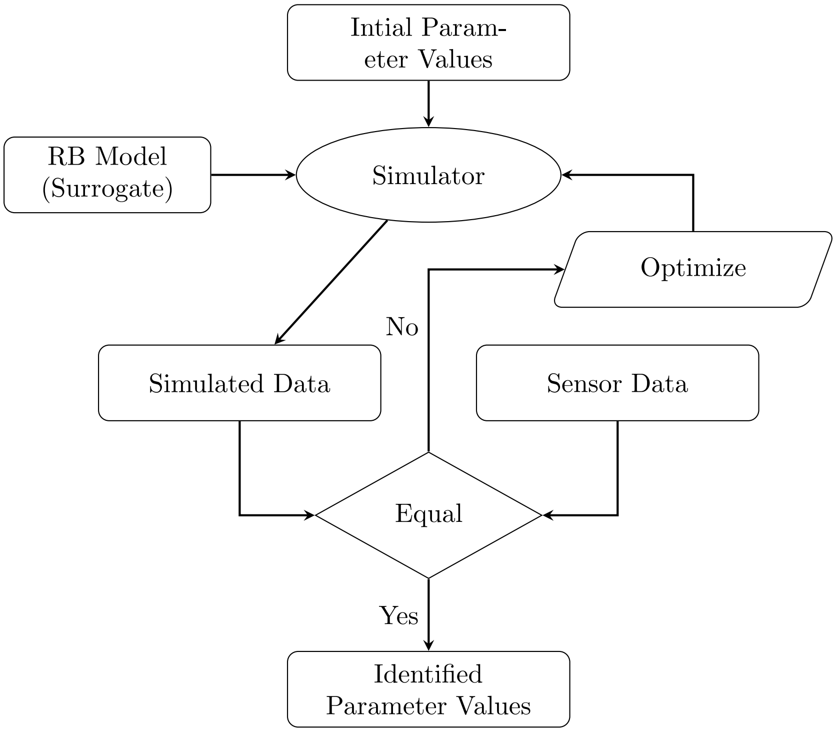


Fig. 3.2 Parameter identification procedure

The identification procedure is further explained by the flowchart depicted in Figure 3.2. Given an initial guess of parameter values, a forward model is run on a 'simulator' to generate simulated field data. The simulated data is then compared to sensor measurement data such that if the difference between the two is above a certain threshold, it detects that the previous guess of the parameter values is incorrect and requires to be adjusted. The goal of the identification procedure is to find the correct parameter values that will match the simulated and sensor measurement data. The set up can be thought of as a function minimization problem where one must converge to the point where the difference between the simulated and sensor measurement data is minimum. To perform the function minimization efficiently, a suitable optimization technique needs to be employed. In the methodology proposed here, the 'simulator' is replaced by a reduced basis surrogate model built in Section 3.

In this work, we implemented two optimization algorithms in the parameter identification procedure. The Nelder-Mead simplex algorithm [91], one of the most widely used direct search methods designed to solve classical unconstrained optimization, is employed when tackling simple cases where the search space is not very large. It is advantageous in the context of multiple parameter realizations as it typically requires only one or two function evaluations at every iteration. To deal with more complex optimization problems, a stochastic global search method is employed. Based on natural evolutionary processes, the genetic algorithm (GA) identifies the optimal parameter values by submitting a group of individuals (population) to the principle of natural selection. By creating randomly sampled individuals and subjecting them to fitness tests for which only outstanding individuals are allowed to carry on to the next generation, this process leads to the evolution of individuals better suited to the studied problem than the initially created individuals [59, 82, 103, 105, 115, 129]. Relative to Nelder-Mead simplex algorithm, GA requires more function evaluation throughout the entire evolutionary process.

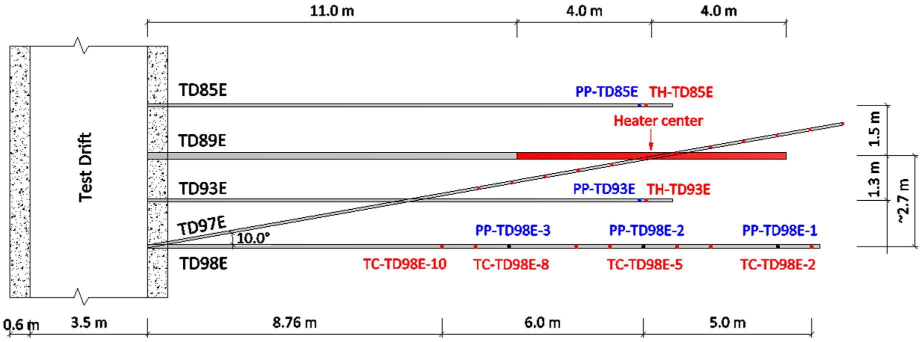
5 Application to ATLAS III Heater Test

In this section, we consider a problem describing the coupled THM processes in an in-situ ATLAS III Heater test conducted in HADES URF. The heating test was performed to characterize the thermo-hydro-mechanical behaviour of the Boom Clay when subjected to transient thermal loads.

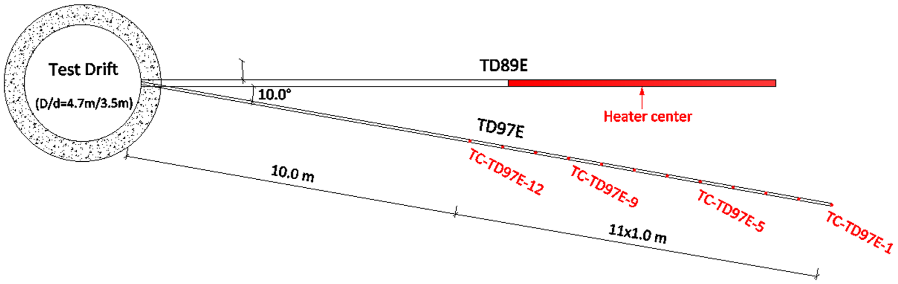
The schematic view of ATLAS III Heater test is presented in Figure 3.3. The original test set-up for ATLAS Heater test was established in 1992 by SCK CEN within the framework of the European project Interclay II (1990-1994) [79]. The ATLAS Heater test had run in three phases between 1993 and 2008. During the first two phases (1993-1997), the test set-up consisted of one heater borehole (TD89E) and two horizontal observation boreholes (TD85E and TD93E). The set-up was extended in 2006 by drilling one inclined downward observation borehole TD97E and one horizontal borehole TD98E. The heater was switched on from April 2007 to April 2008 with a stepwise power increase followed by an instantaneous shut down. This phase is called ATLAS III Heater test.

The central heater borehole TD89E has a drilling depth of 19 m and a drilling diameter of 230 mm. The heated section depth runs from 11 m to 19 m. The borehole is cased with a stainless tube having an external diameter of 190 mm and an internal diameter of 160 mm. The initial gap between the tube and the surrounding Boom Clay closed within a few months due to soil convergence. The two observation boreholes were installed in both sides of the central heater borehole – with TD85E 1.5 m to the left and TD93E 1.3 m to the right. At the end of both boreholes, sensors were installed in the same depth ($\sim$15 m) as the centre of the heater to measure the temperature and pore pressure. Borehole TD97E has a depth of 21 m with a downward inclination of 10° and a horizontal deviation (to the left) of 10°. TD97E is equipped with 12 temperature sensors that are evenly distributed along the depth of 10 to 21 m. There are no pore pressure sensors in borehole TD97E. Borehole TD98E is located about 2.7 m away from the central borehole, with a length of 20 m. TD98E is equipped with 10 temperature sensors distributed along the depth of 9 to 20 m. It is instrumented by three pore pressure sensors with respective depth of 11, 15 and 19 m.

For ATLAS III, the heater was re-activated on April 2, 2007 with the power being increased to 400 W within 4 days. The heating power increased from 400 W to 900 W on May 25, 2007 within 5 days, and finally increased from 900 W to 1400 W on August 3, 2007 within 5 days. On April 17, 2008, the heater was shut off. Figure 3.4 sketches the stepwise increase of the total heater power applied. The measurement data covering a 381-day heating phase and a 519-day cooling phase are used for the inverse analysis.



(a) Top View



(b) Front View

Fig. 3.3 ATLAS III in-situ test set-up and instrumentation

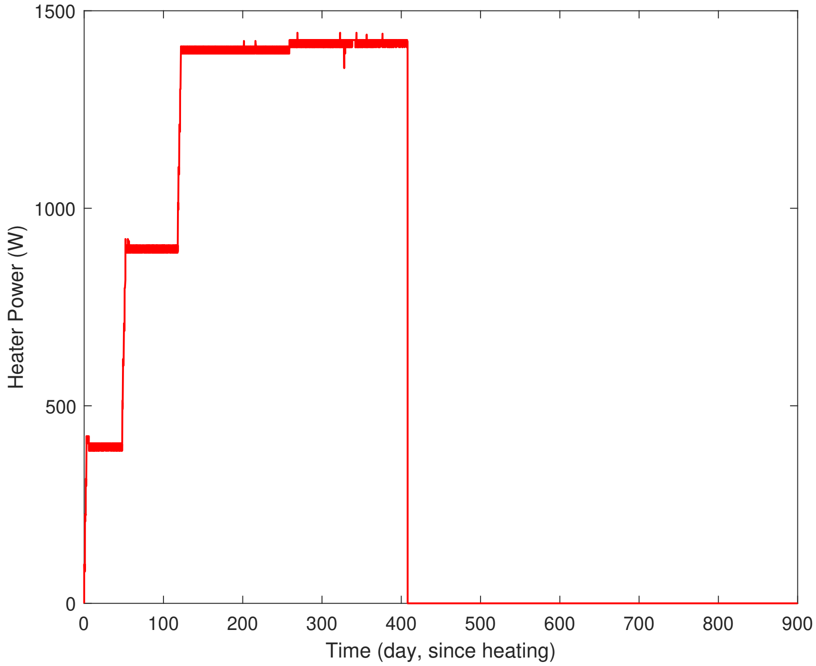


Fig. 3.4 ATLAS III heating power

The coupled THM processes in the ATLAS III Heater test is modelled numerically assuming full saturation conditions. Analysis in [29] indicates that the plasticity is very limited in the Boom Clay and elasticity plays a dominant role in the test responses. Therefore, the Boom Clay is assumed to be linearly poroelastic. Cross-anisotropy is considered for the properties of all three physics involved - thermal conductivity, permeability and elasticity. The THM behaviour is isotropic in the sub-horizontal bedding (x-y plane) and anisotropic with respect to the perpendicular vertical plane (along z-axis). The governing equations presented in (3.1)-(3.4) are applied to cross-anisotropic cases and the Boom Clay material properties are presented in Table 3.3 [17, 29, 37, 70, 153]. The steel casing material properties are presented in Table 3.2. Note that some of these material property values will be modified in the case where they will be selected as the parameters to be identified in the inverse analysis.

Table 3.2 Material parameters describing the steel casing

E	210 GPa
v	0.20
G	88 GPa
c_p	460 J/ (kg $\cdot$ K)
k_c	15 W/ (m $\cdot$ K)

Table 3.3 Material parameters describing the Boom Clay

E_h	1400 MPa
E_v	700 MPa
v_{hh}	0.25
v_{vh}	0.125
G_{vh}	280 MPa
α	1
M	5698 MPa
k_h	4×10^{-19} m^2
k_v	2×10^{-19} m^2
ϕ	0.39
α_s	1×10^{-5} °C^{-1}
$3\alpha_f$	2.8×10^{-4} °C^{-1}
$c_{p,s}$	740 J/ (kg $\cdot$ K)
k_{ch}	1.65 W/ (m $\cdot$ K)
k_{cv}	1.31 W/ (m $\cdot$ K)

h and v subscripts refer to the orientation of cross-anisotropic variables

5.1 Thermal Analysis

This section deals with the first step of the THM analysis for ATLAS III Heater test which is a purely thermal analysis.

5.1.1 Full Order Model (FOM)

Figure 3.5 illustrates the 3D geometry and the corresponding mesh developed for the thermal model, covering a computational domain of 100 m-radius in the x-z plane and 119 m-length in the direction parallel to y-axis. The domain consists of two regions: a 15 mm-thick, 19 m-long steel casing and the Boom Clay for the rest of the domain. The resulting mesh discretization contains 36 489 nodes and 33 024 linearly interpolated 8-noded hexahedral elements. To model the thermal loading, a homogeneously distributed heat flux is prescribed on the intrados of the steel casing which runs from 11 m to 19 m along the y axis. The remaining boundaries are defined as adiabatic. An initial temperature of 16.5°C is prescribed in the whole domain.

In the modelling, the influence of the Test Drift boundary conditions on the Boom Clay THM responses is considered negligible. This assumption primarily takes into account the Test Drift excavation (in 1987) which happened 20 years before ATLAS III started. Moreover, the ATLAS III heater was installed 11 m from the Test Drift and previously conducted sensitivity analysis indicated that heating-induced disturbance does not reach the Test Drift during a one-year heating phase. In 1992, the central heater borehole AT89E was drilled and cased with a steel tube resulting in a 20 mm thick overdrill. Further 15 years had passed and quasi-equilibrium conditions prevailed which makes is acceptable to assume a homogeneous Boom Clay in the test domain.

After implementing the FE model with the assumptions described, the temperature evolution corresponding to reference material properties (Table 3.2 and 3.3) at sensor locations are shown in Figure 3.7.

5.1.2 Reduced Order Model (ROM)

The thermal system is parametrized by varying several material properties within a wide range of values well beyond expected for the Boom Clay [54]. The cross-anisotropic thermal conductivities (k_{ch}, k_{cv}) are selected to vary between $[1, 3]$W/(m·K), while the solid phase specific heat coefficient ($c_{p,s}$) varies between $[500, 1200]$ J/(kg·K). The parametric training sample is discretized by uniformly sampling k_{ch}, k_{cv} and $c_{p,s}$ with 12 points resulting in 1728 ($12 \times 12 \times 12$) parameter sets where each set

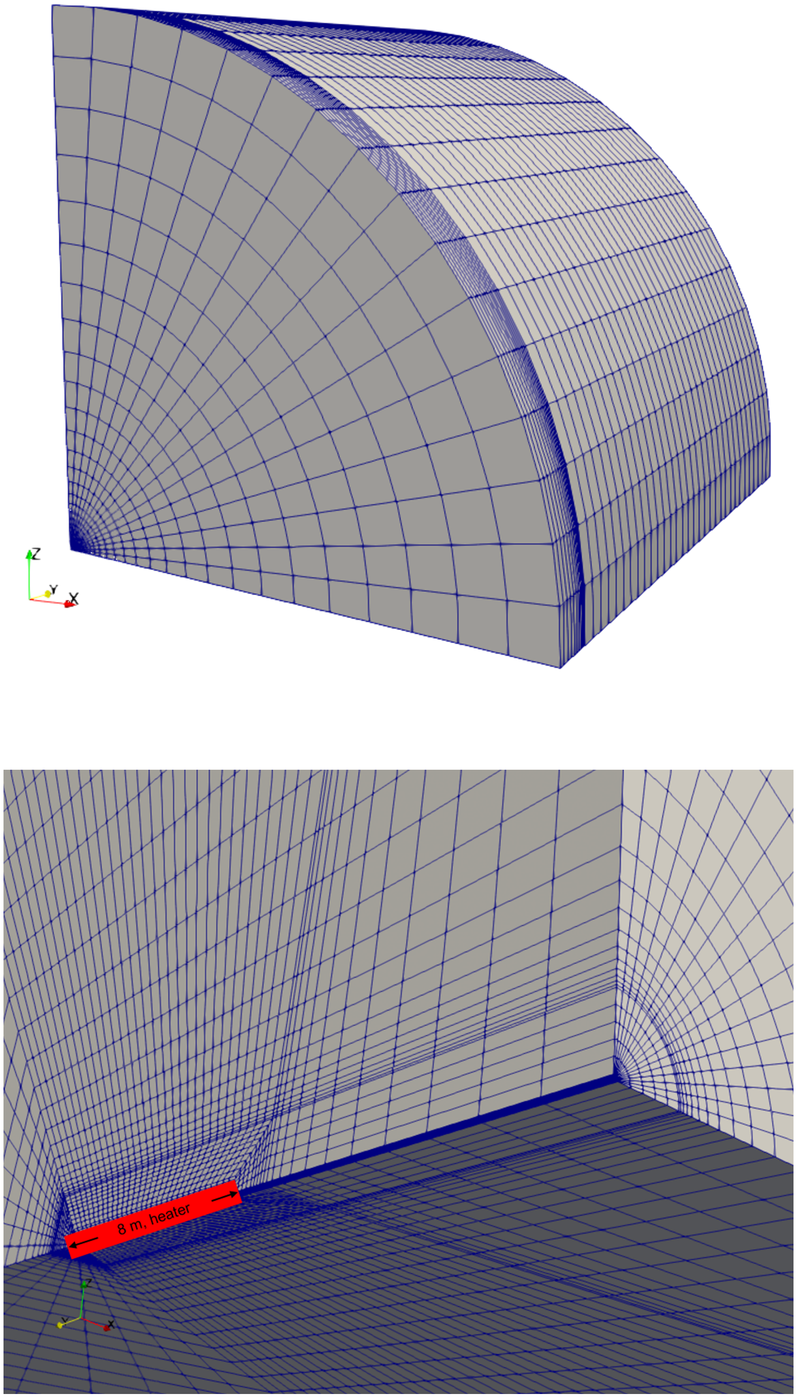


Fig. 3.5 3D mesh of the thermal model

contains 3 thermal property values, $\boldsymbol{\mu} = (k_{ch}, k_{cv}, c_{p,s})$.

We define the quantity of interest (QoI) as the average temperature over time in regions surrounding the thermocouple sensors. The error in the QoI is estimated by relating the weak residual to the adjoint solution (3.20). The parameter set defined in the reference FOM is used to formulate the dual problem which will represent the adjoint for the entire parametric training sample. With these assumptions, the enrichment procedure described in Section 3 and detailed in [81] is performed and stopped once the normalized estimated error in the QoI is below the 10^{-4} tolerance.

It took five enrichment steps to build the reduced basis for the parametrized system. In the final enrichment step, the modes are truncated to 19 modes which correspond to the 10^{-5} tolerance prescribed for the POD technique in the greedy basis generation algorithm. A significant size reduction relative to the reference FOM (36 489 free degrees of freedom) is observed in the system, yielding a reduction ratio of 1:1920.

To validate the accuracy of the reduced basis approximation, the temperature evolution at different sensor locations obtained by the full order FE model are compared against the RB-projected solutions (Figure 3.7) for the material properties given in Table 3.3. In addition to this validation on the reference FOM, random parameter values within the training sample are also used as comparison points. A quantitative measure of accuracy is formulated as $\bar{e} = \frac{\|\text{FOM}-\text{ROM}\|_2}{\|\text{FOM}\|_2}$, taking the norms of temperature discrepancies between the FE and RB-projected solutions (Table 3.4).

5.1.3 Thermal Properties Identification

As previously mentioned, specific heat and thermal conductivity are chosen as the unknown thermal properties to be identified. The objective function describing the relative difference between measured and simulated temperature data is evaluated as:

$$\chi = \frac{1}{\mathtt{n}_{\mathtt{sensor}}} \sum_{i=1}^{\mathtt{n}_{\mathtt{sensor}}} \int \frac{\| T_{sensor} - T_{surr} \|^2}{\| T_{sensor} \|^2} dt \tag{3.24}$$

where T_{sensor} is the sensor measurement data, T_{surr} is the simulated data from the surrogate model and $\mathtt{n}_{\mathtt{sensor}}$ is the total number of sensors considered. The integral over time imposes a weight effect to avoid biasing and favouring periods where more frequent measurements are taken. The resulting relative norms are averaged over the total number of sensors.

With only three unknown thermal properties to identify and 24 sensor measurement

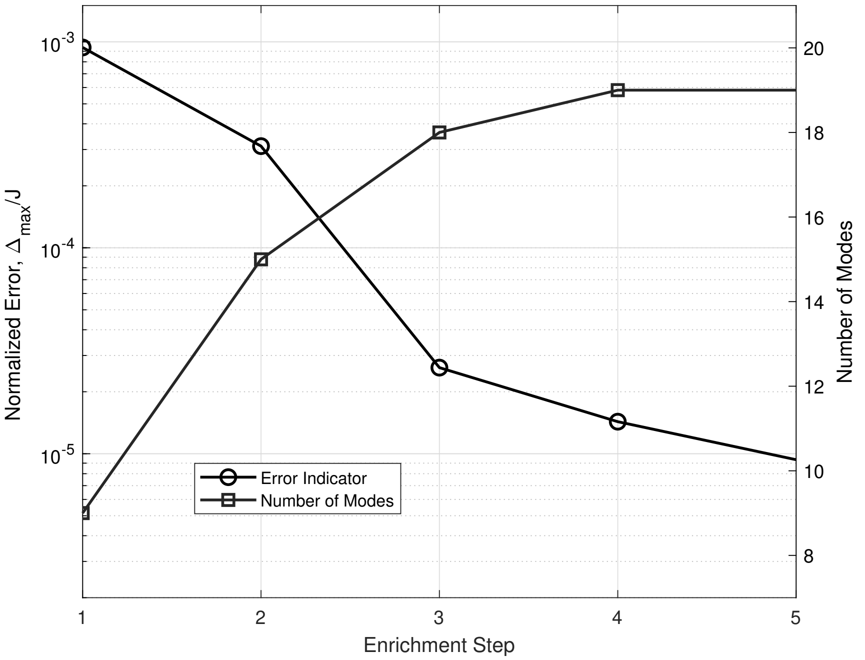


Fig. 3.6 Convergence of the greedy-based procedure to the prescribed error level and the number of corresponding POD modes at each enrichment step - thermal analysis

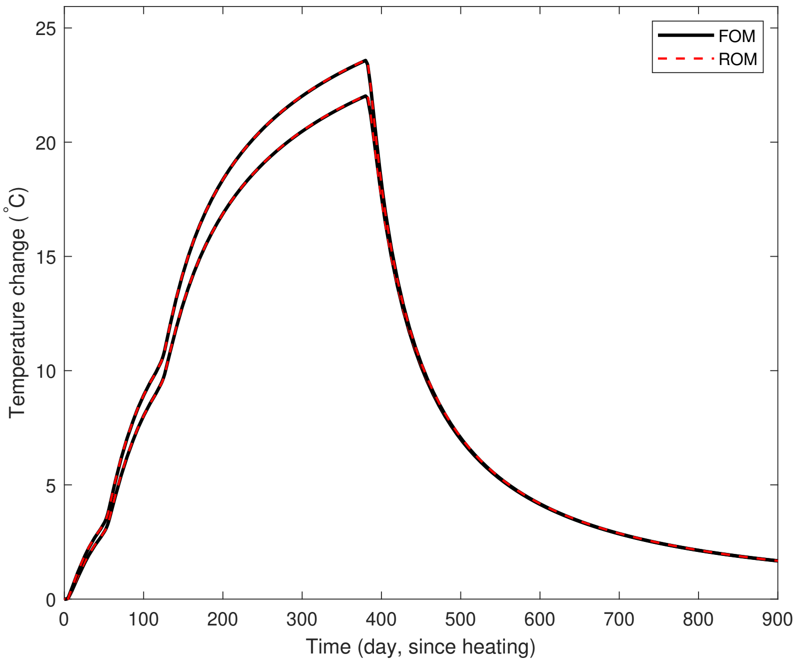
data, a local search method, Nelder-Mead simplex algorithm, is chosen to perform the optimization. The entire measurement taken from day 0 to 900 is considered for the time period in the objective function.

Table 3.4 Relative error norm ($\bar{e}$) of thermal RB-projected solutions at random parameter values ($k_{ch}, k_{cv}, c_{p,s}$)

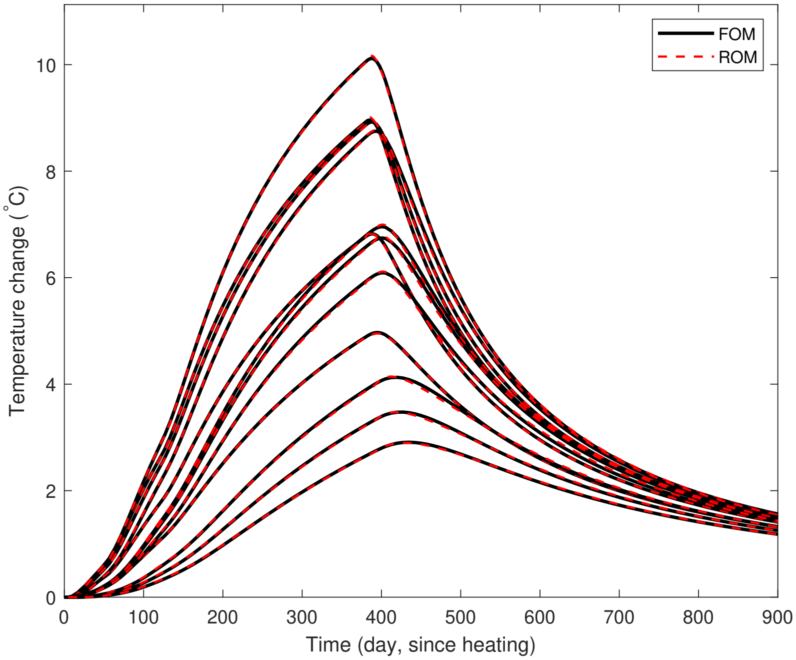
Parameter Values	TH-TD85E TH-TD93E	TC-TD97E	TC-TD98E
1.65, 1.31, 740	8.3×10^{-5}	5.2×10^{-5}	3.5×10^{-5}
2.00, 2.00, 850	3.3×10^{-5}	3.2×10^{-5}	2.3×10^{-5}
1.00, 3.00, 1200	3.7×10^{-5}	3.1×10^{-5}	1.8×10^{-5}
1.15, 2.55 1100	4.1×10^{-5}	4.8×10^{-5}	3.5×10^{-5}
2.50, 2.80 600	2.6×10^{-5}	2.1×10^{-5}	1.5×10^{-5}

Synthetic Data

To verify the ability of the ROM to replace FOM acting as surrogate in the inverse analysis, we first consider an identification problem using synthetically manufactured data. This eliminates possible sensor data-related errors such as measurement

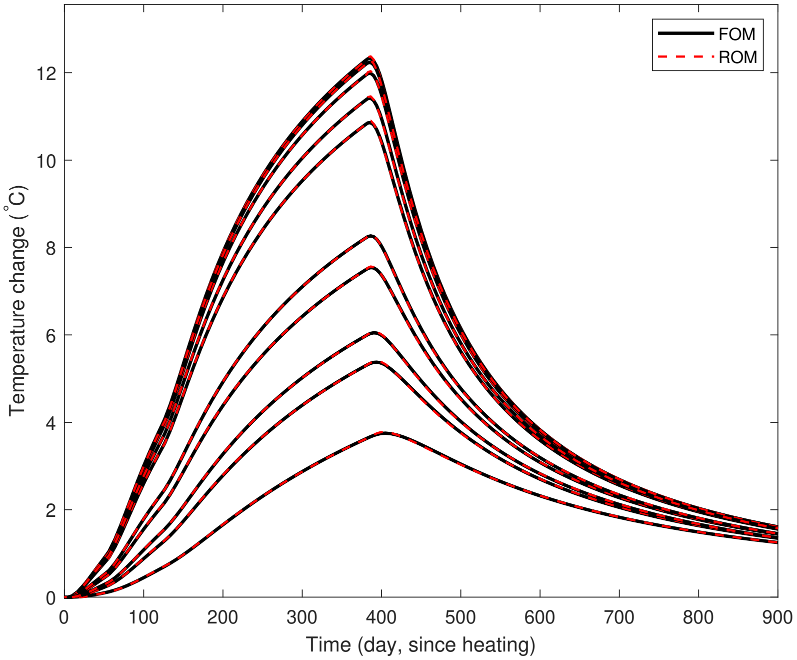


(a) One sensor in TD85E and one sensor in TH-TD93E



(b) 12 sensors in TH-TD97E

Fig. 3.7 Evolution of temperature in sensor locations: reduced order model versus full order model solutions at $\boldsymbol{\mu}^{\text{ref}} = (1.65, 1.31, 740)$



(c) 10 sensors in TH-TD98E

Fig. 3.7 Evolution of temperature in sensor locations: reduced order model versus full order model solutions at $\boldsymbol{\mu}^{\mathrm{ref}} = (1.65, 1.31, 740)$

noise effects. The synthetic data used is obtained from the reference FOM where temperature solutions at 24 sensor locations from day 0 to 900 are taken.

Convergence point in the optimization is attained with $\chi = 2.2 \times 10^{-9}$ after 132 iterations and 265 objective functions evaluated. The identified parameters $\boldsymbol{\mu^{opt}} = (1.648, 1.312, 738)$ are very close to the target parameter values $\boldsymbol{\mu^{ref}} = (1.65, 1.31, 740)$ used in the reference FOM, see Figure 3.8. It is expected that there is a shift in the minimum of the optimization search space since the surrogate model is not exactly the same as the reference FOM that generated the synthetic data.

For comparison purposes, we also performed the same identification procedure while using a FOM simulator (direct FE approach) in place of the surrogate model. The identified parameters converged to the exact values as expected. However, a significantly longer run time is expected where it took 29 hours to finish the identification procedure compared to less than 500 sec. using a surrogate model despite having almost the same number of iterations (Figure 3.9).

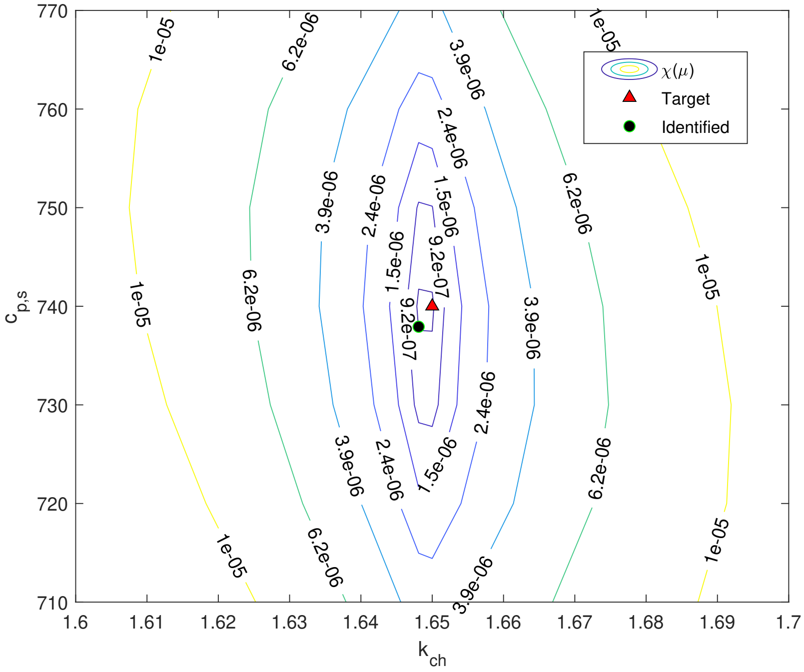


Fig. 3.8 Optimization search space minimum using surrogate RB model (synthetic data)

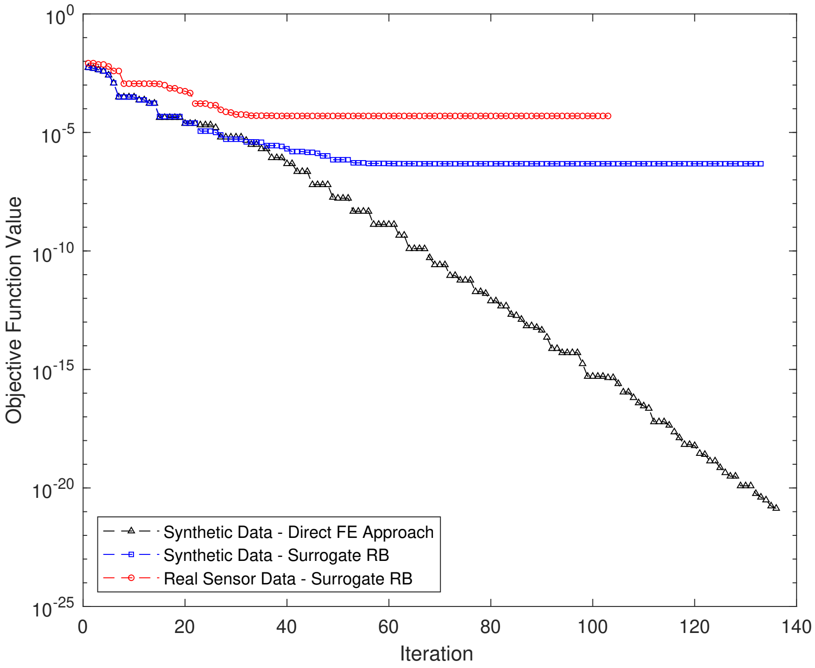


Fig. 3.9 Minimization of objective function during optimization (Nelder-Mead simplex algorithm) using synthetic data and real sensor data

Real Measurement Data

The in-situ measured data from 24 thermocouple sensors are simplified to some extent for this inverse problem application.

Using a similar configuration with the synthetic case, we now define T_{sensor} as the real measurement from sensors. The objective function is slightly modified by introducing additional weight factors $w_i = \frac{|d_{max}-d_i|}{\sum |d_{max}-d_i|}$ where d pertains to the distance of the sensor from the heater. If sensors are installed far the from the heater, it is expected that less thermal effect is experienced in such regions. Sensors closer to the heater are favoured by applying higher weights as their temperature profiles are more significant in describing the temperature evolution of the system.

$$\chi = \frac{1}{\mathtt{n_{sensor}}} \sum_{i=1}^{\mathtt{n_{sensor}}} w_i \int \frac{\parallel T_{sensor} - T_{surr} \parallel^2}{\parallel T_{sensor} \parallel^2} dt \tag{3.25}$$

Using Nelder-Mead simplex algorithm, the unknown parameters are identified after evaluating 224 objective functions in 102 iterations. Contrary to the synthetic case, no target values are available to compare with the identified values since we intend to derive unknown thermal properties. To measure the accuracy of the identified values, their corresponding simulated temperatures are fitted over the sensor data (Figure 3.10). It is observed that majority of the sensor data fit well to the numerical model solution using the identified parameters. However there are some sensors (TC-TD97E1-6) wherein discrepancies are apparent. Some experimental factors may affect the parameter identification, such as the sensor location uncertainty, heater power uncertainty, temperature measurement precision or theoretical assumptions used in the numerical model development. It is important to emphasize that the main objective of this study is not to interpret the test results but to demonstrate the efficiency and accuracy of the presented ROM in the in-situ test inverse analysis Therefore, further investigation on these factors is not explored here. The identified thermal parameters, $\boldsymbol{\mu}^{opt} = (1.80, 1.17, 835)$, are used to build the coupled THM system described in the subsequent sections.

5.2 Thermo-Hydro-Mechanical (THM) Analysis

This section deals with the second step of the THM analysis for the ATLAS III Heater test, in which the independently solved temperature is mapped to the HM system.

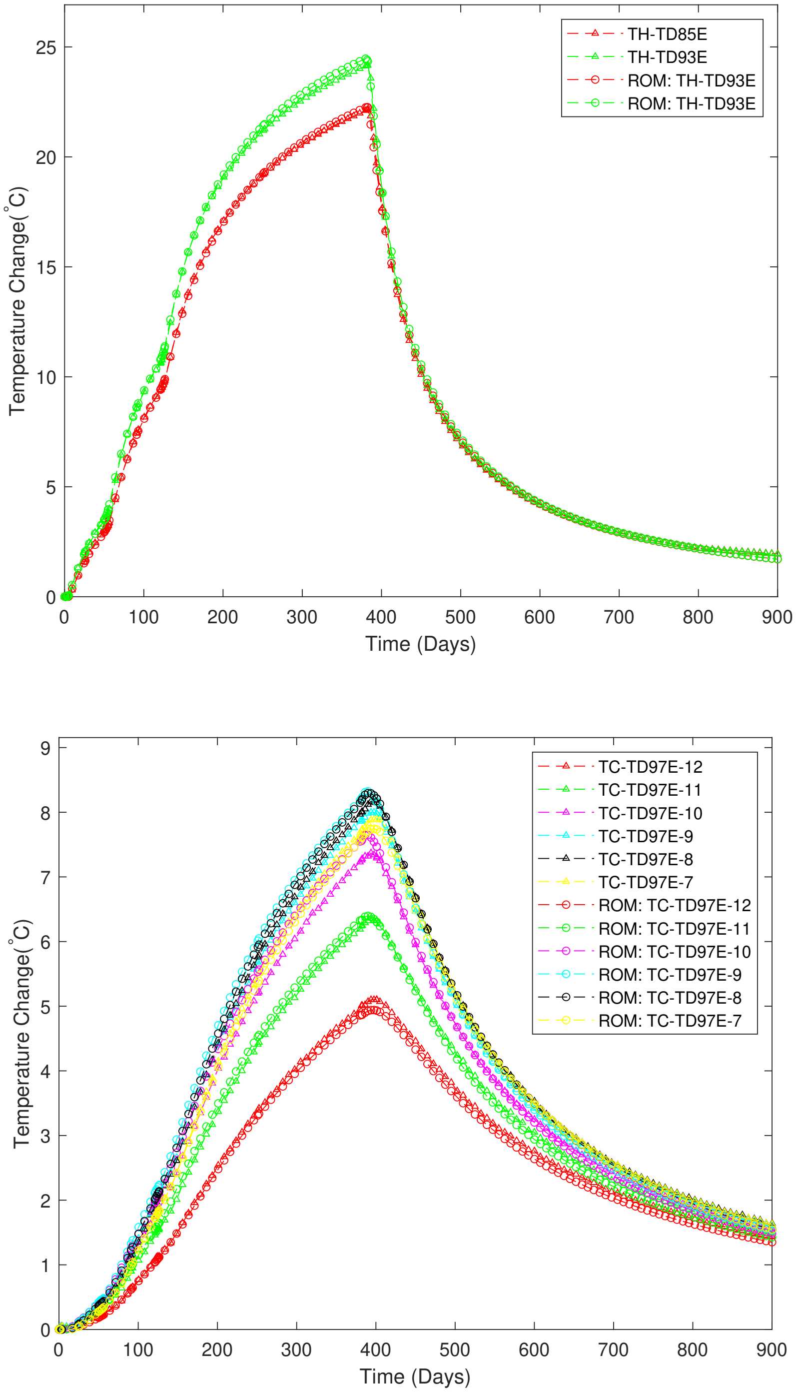


Fig. 3.10 Modelled temperature using identified parameters fitted over sensor measurement data

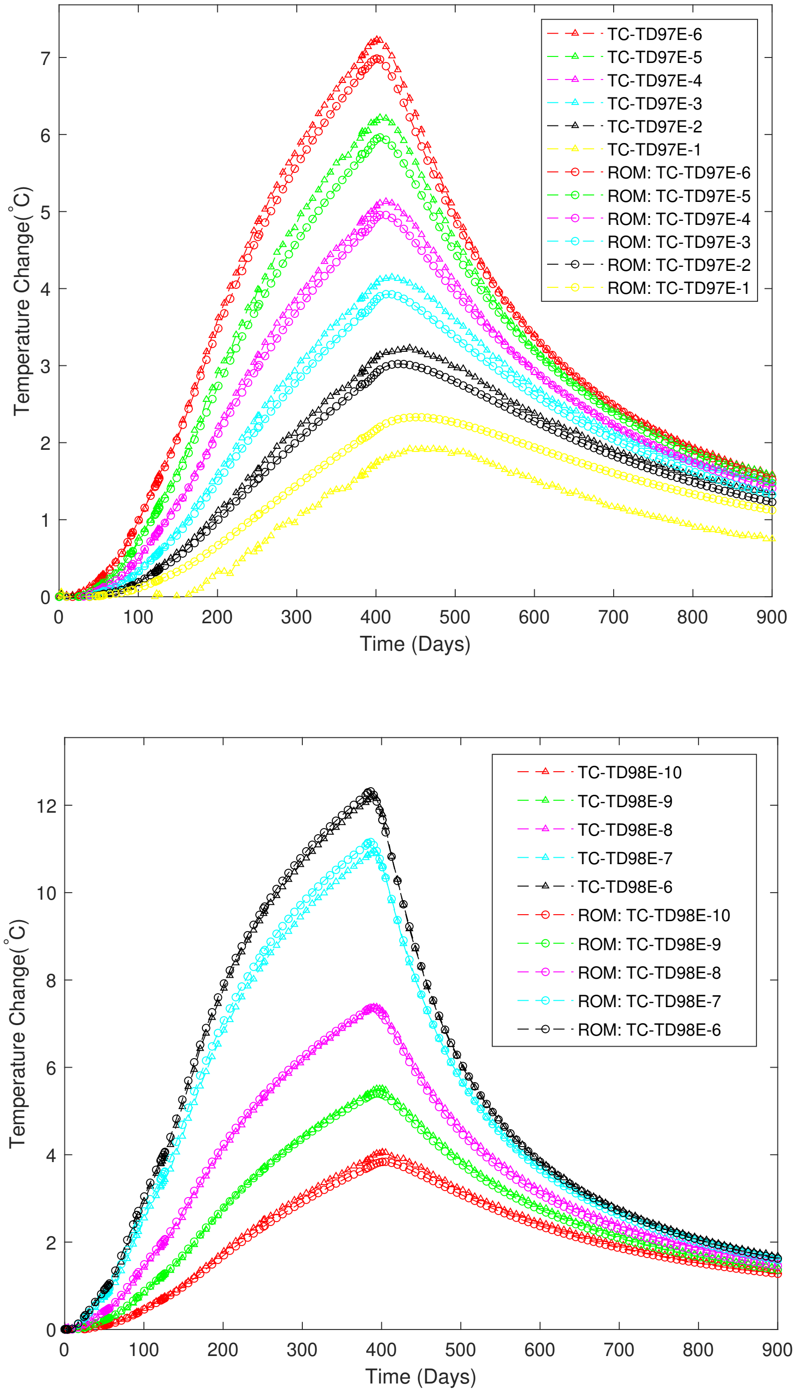


Fig. 3.10 Modelled temperature using identified parameters fitted over sensor measurement data

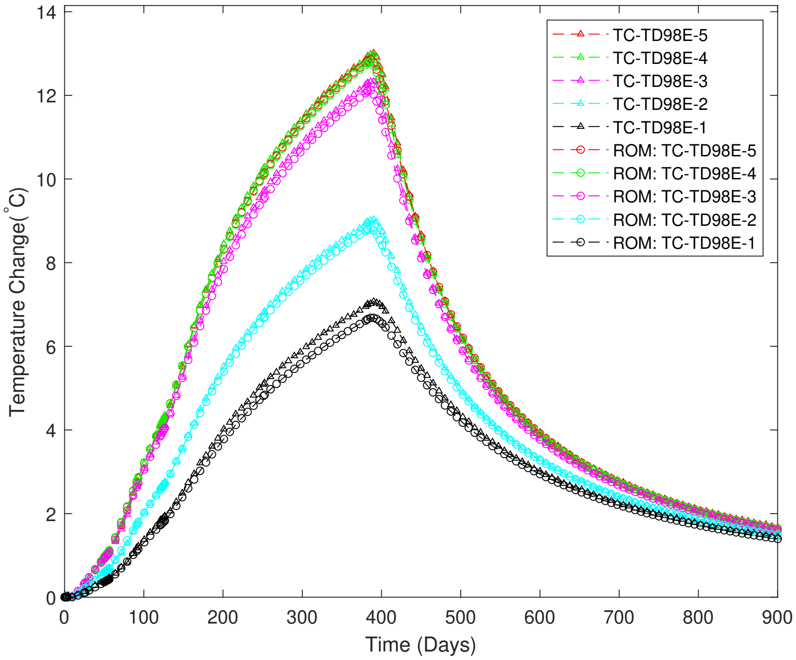


Fig. 3.10 Modelled temperature using identified parameters fitted over sensor measurement data

5.2.1 Full Order Model

A 3D geometry similar to the thermal model is considered for the HM model, consisting of a heterogeneous domain but with a smaller Boom Clay region covering only 30 m of radius in the x-z plane and 49 m of length in the y-axis. Comparing to the thermal model, a smaller geometrical domain is considered as hydro-mechanical effects far from the heat source are found to be insignificant in characterizing the coupled processes. The discretized mesh contains 60 509 nodes and 13 760 quadratic (20-noded hexahedral) elements.

The hydraulic boundary conditions include a 2.25 MPa pore water pressure imposed on ABFE and BCF surfaces while the remaining surfaces are impervious (Figure 3.11). For the mechanical boundary conditions, displacements are fixed in the normal direction for surface ABCD, CDEF and ADE. Normal boundary stress is applied on surface BCF (3.825 MPa) and ABFE (4.5 MPa in z-direction, 3.825 MPa in x-direction). The HM model is described with 2.25 MPa initial pore water pressure and initial stress conditions of 4.5 MPa and 3.825 MPa in the vertical and horizontal direction, respectively. To establish the initial conditions, an offset period before the start of heating is introduced in the model for which very high permeability values are temporarily assigned to the system in order to reach the desired equilibrium state. The aforementioned HM boundary conditions are imposed while no thermal load is applied during the offset period.

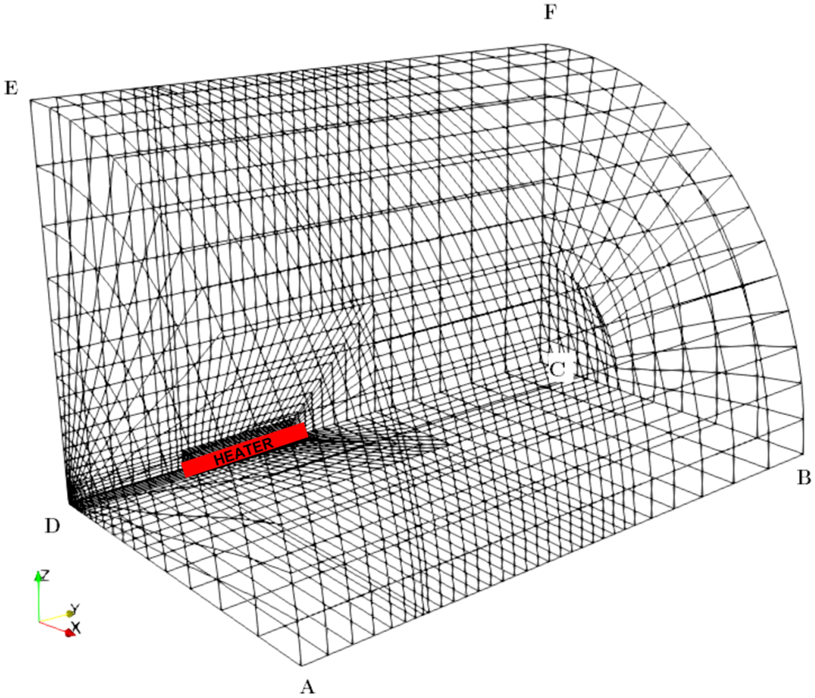


Fig. 3.11 3D mesh of the coupled THM model

The temperature change is the main loading that affects the hydro-mechanical processes. The varying temperature in the domain is primarily reflected on the temperature-dependent viscosity defined in (3.5). For this numerical model, the volumetric thermal expansion coefficient is assumed to be independent of temperature with a constant value of 2.8×10^{-4} °C^{-1} corresponding to the average coefficient value between 16°C and 40°C based on the in-situ environment.

5.2.2 Reduced Order Model

Five material properties corresponding to cross-anisotropic Young's modulus (E_h, E_v), cross-anisotropic permeability (k_h, k_v) and shear modulus (G_{vh}) are varied to define the parametrized coupled system. The solution of the parametrized system belongs to a 5-dimensional space corresponding to the five material parameters: $E_h \in (200, 2000)$, $E_v \in (100, 1600)$, $k_h \in (2{\times}10^{-19}, 8{\times}10^{-19})$, $k_v \in (1{\times}10^{-19}, 5{\times}10^{-19})$, $G_{vh} \in (50, 1000)$. The parametric training sample is discretized to 1425 parameter sets $\boldsymbol{\mu} = (E_h, E_v, k_h, k_v, G_{vh})$ – each material parameter is uniformly sampled with 5 discrete points and sets containing $E_h < E_v$ and $k_h < k_v$ are excluded from the training sample.

We define the QoI as the average pore water pressure over time in regions surrounding the piezometer filters. Similar to the thermal system, the adjoint solution

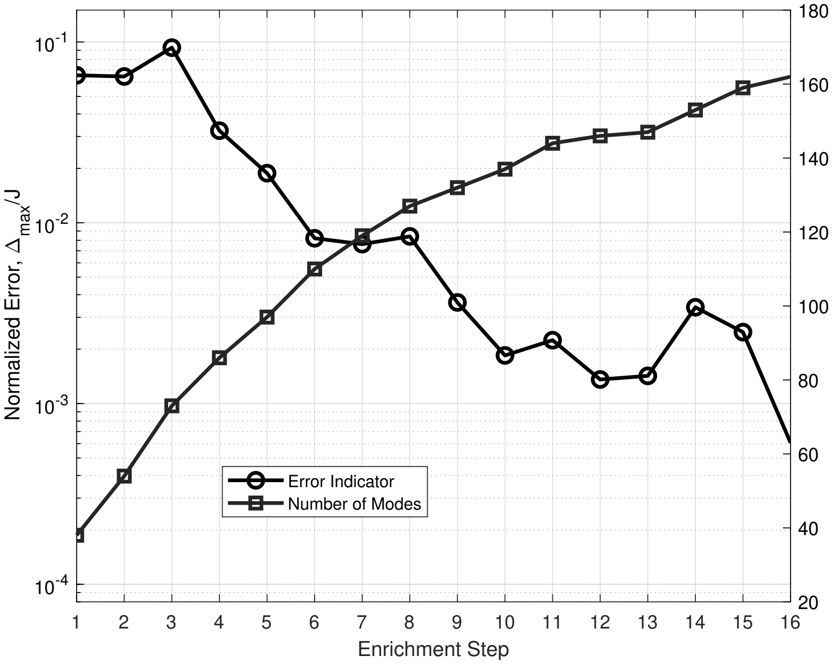


Fig. 3.12 Convergence of the greedy-based procedure to the prescribed error level and the number of corresponding POD modes at each enrichment step - THM system

to the dual problem of the HM system (**??**) is weighted to the weak residual to obtain error estimates in the QoI. The offset period introduced to establish the initial conditions is disregarded in the residuals since we are only interested in the solutions after the initial conditions are reached. Moreover, the offset period does not affect the dual problem formulation and the adjoint solution.

The adaptive greedy procedure in Section 3 is followed to generate an enriched reduced basis. In comparison with the full order HM model containing 183 387 free degrees of freedom, the constructed RB contains only 162 modes (with 10^{-5} truncation tolerance) after a total of 16 enrichment steps. A reduction ratio of 1:1130 is achieved. To validate the accuracy of the reduced basis approximation, full order pore water pressure solutions at different sensor locations are compared to RB-projected pore water pressure solutions for the material properties reported in Table 3.3. Random parameter values within the training sample are also used as additional comparison points. The relative norms of pore water pressure discrepancies between the full order and RB-projected solutions are less than 5.6×10^{-3} as shown in Table 3.5).

Figure 3.13 compares the ROM pore pressure changes at the five sensor locations with those by FOM using material parameter set $\boldsymbol{\mu} = (1000, 500, 3.5 \times 10^{-19}, 3 \times 10^{-19}, 600)$, from which an excellent agreement can be observed. It is remarkable that the reduced order model was able to capture the temporary pore pressure decrease (or increase) after power increase (or cooling) due to mechanical anisotropy [29].

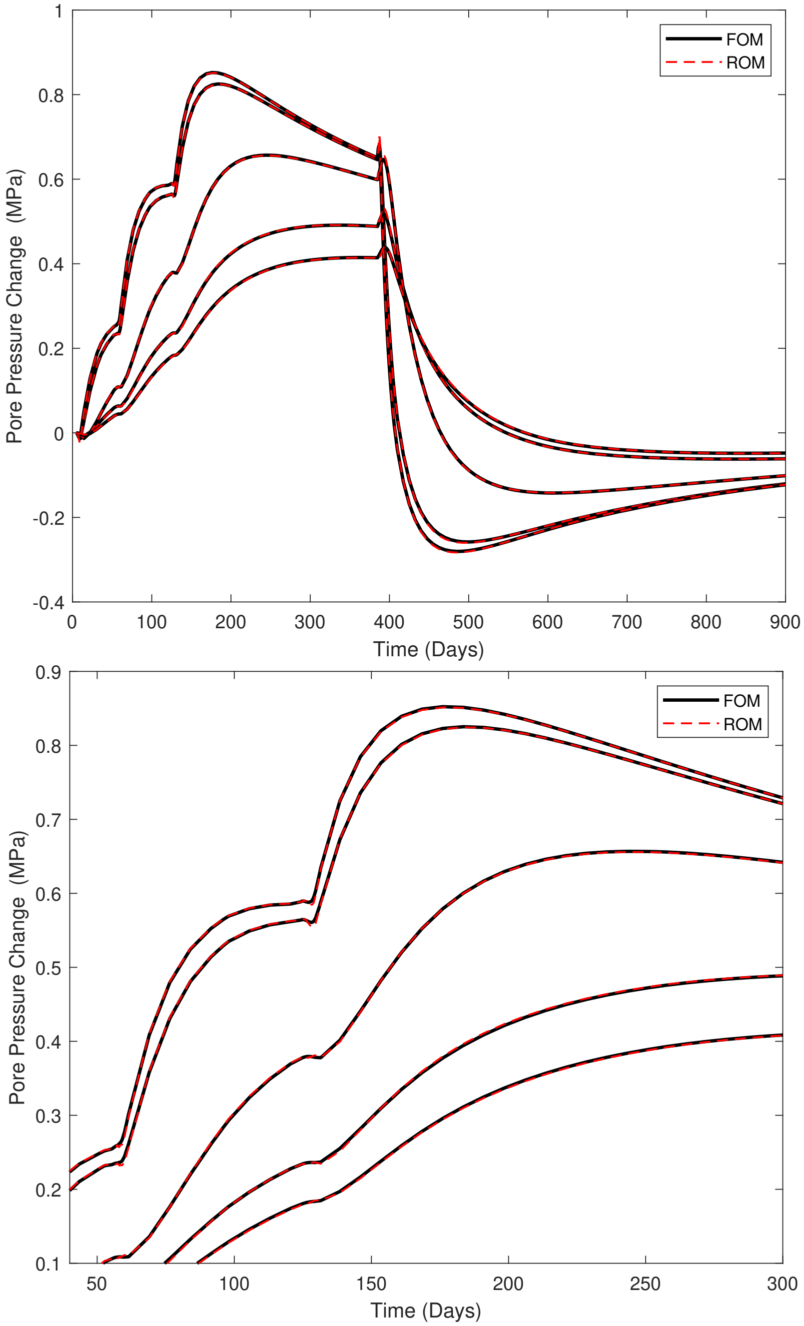


Fig. 3.13 Evolution of pore pressure change at 5 sensor locations (PP-TD85E, PP-TD93E, PP-TD98E1, PP-TD98E2, PP-TD98E3): reduced order model versus full order model solutions at $\boldsymbol{\mu} = (1000, 500, 3.5 \times 10^{-19},\ 3 \times 10^{-19}, 600)$

Table 3.5 Relative error norm ($\bar{e}$) of RB-projected pore water pressure at random parameter values ($E_h, E_v, k_h, k_v, G_{vh}$)

Parameter Values	PP-TD 85E	PP-TD 93E	PP-TD 98E1	PP-TD 98E2	PP-TD 98E3
$1400, 700,$ $4 \times 10^{-19},$ $2 \times 10^{-19},$ 280	$1.6{\times}10^{-3}$	$1.6{\times}10^{-3}$	$1.3{\times}10^{-3}$	$1.1{\times}10^{-3}$	$6.6{\times}10^{-4}$
$1700, 850,$ $5 \times 10^{-19},$ $1 \times 10^{-19},$ 350	$1.4{\times}10^{-3}$	$1.3{\times}10^{-3}$	$1.1{\times}10^{-3}$	$9.5{\times}10^{-4}$	$7.5{\times}10^{-4}$
$1000, 500,$ $3.5 \times 10^{-19},$ $3 \times 10^{-19},$ 600	$1.6{\times}10^{-3}$	$1.6{\times}10^{-3}$	$1.3{\times}10^{-3}$	$1.1{\times}10^{-3}$	$6.6{\times}10^{-4}$
$650, 450,$ $4 \times 10^{-19},$ $4 \times 10^{-19},$ 200	$1.8{\times}10^{-3}$	$1.9{\times}10^{-3}$	$1.1{\times}10^{-3}$	$1.4{\times}10^{-3}$	$5.6{\times}10^{-3}$
$400, 400,$ $3 \times 10^{-19},$ $3 \times 10^{-19},$ 150	$2.3{\times}10^{-3}$	$2.5{\times}10^{-3}$	$1.4{\times}10^{-3}$	$1.8{\times}10^{-3}$	$1.3{\times}10^{-3}$
$800, 600,$ $6 \times 10^{-19},$ $5 \times 10^{-19},$ 400	$1.3{\times}10^{-3}$	$1.1{\times}10^{-3}$	$1.0{\times}10^{-3}$	$9.1{\times}10^{-4}$	$3.9{\times}10^{-4}$

5.2.3 Hydro-mechanical Properties Identification

Pore water pressure measurements taken at five sensor locations (PP-TD85E, PP-TD93E, PP-TD98E1, PP-TD98E2, PP-TDT98E3) for a duration of 900 days are used in the inverse analysis. The hydro-mechanical properties to be identified are selected to be the same five independent parameters ($E_h, E_v, k_h, k_v, G_{vh}$) varied when building the reduced basis.

With five material properties to identify and having limited measurement data to fully describe the coupled processes, using a local search method (Nelder-Mead simplex algorithm) did not yield desirable results as it often converged to local minima instead of identifying the target values. A global search method (genetic algorithm), which is known to perform better in tackling more complex optimization problems,

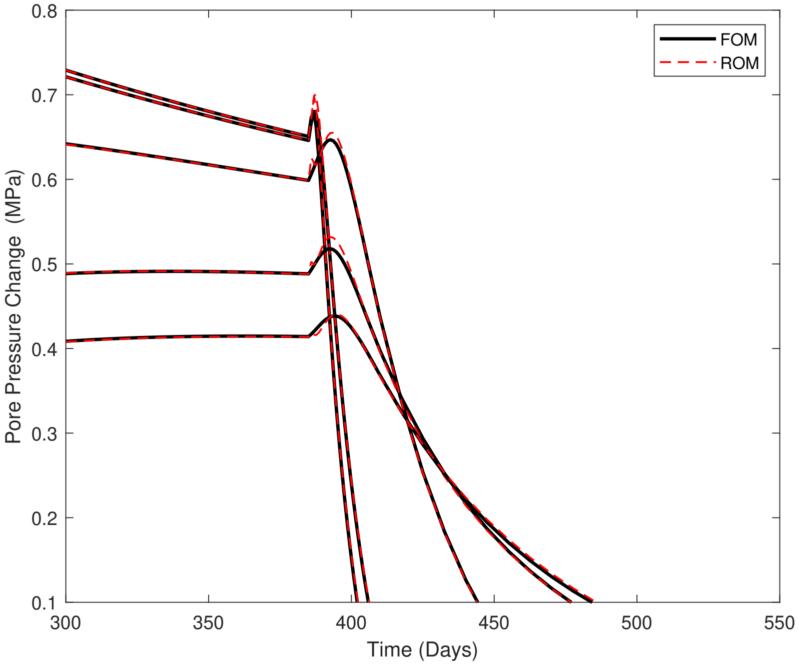


Fig. 3.13 Evolution of pore pressure change at 5 sensor locations (PP-TD85E, PP-TD93E, PP-TD98E1, PP-TD98E2, PP-TD98E3): reduced order model versus full order model solutions at $\boldsymbol{\mu} = (1000, 500, 3.5 \times 10^{-19},\, 3 \times 10^{-19}, 600)$

is therefore employed to perform the following identification procedure. The fitness function describing the discrepancy between in-situ measured and simulated pore water pressure data is formulated as:

$$\chi = \frac{1}{\mathtt{n}_{\mathtt{sensor}}} \sum_{i=1}^{\mathtt{n}_{\mathtt{sensor}}} \int \frac{\| p_{sensor} - p_{surr} \|^2}{\| p_{sensor} \|^2} dt \tag{3.26}$$

where p_{sensor} is the measured sensor data, p_{surr} is the simulated data from the surrogate model and $\mathtt{n}_{\mathtt{sensor}}$ is the total number of sensors considered. The resulting relative norms are averaged over the total number of sensors. Since the five sensors are located close to one another, a weight factor reflecting their distances from the heater is deemed unnecessary. The optimization is constrained by imposing that elastic moduli and permeabilities are always higher in the horizontal direction. We limit the number of population per generation to 50 individuals as parameter realizations which require evaluation of the surrogate model to compute the fitness value dominate the genetic algorithm computational cost.

Synthetic Data To verify the surrogate-based inverse analysis method described above, we start with the identification of hydro-mechanical properties using syntheti-

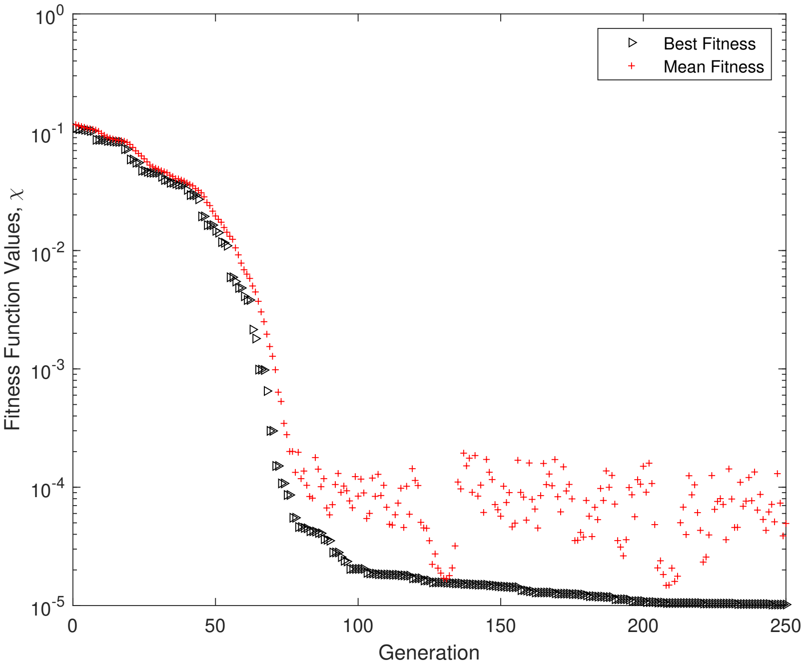
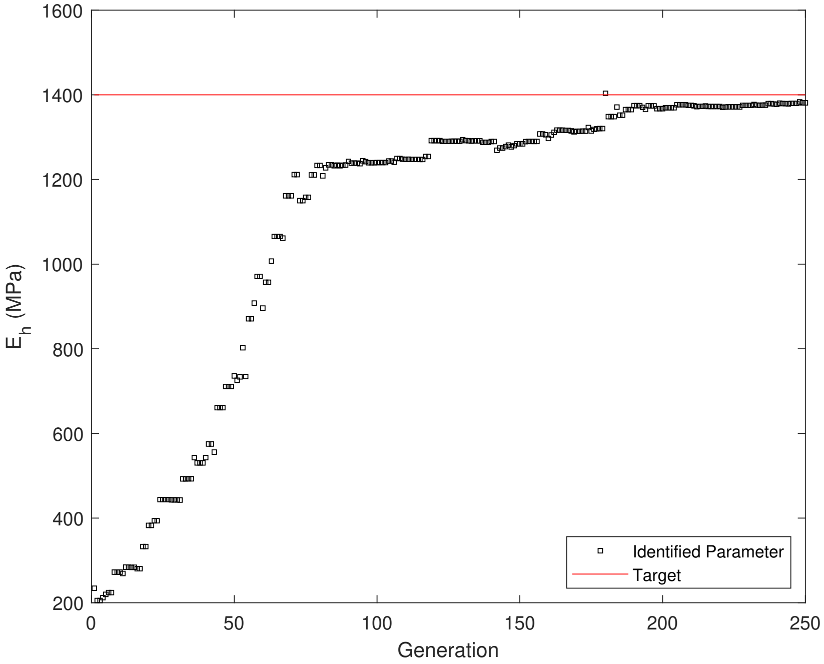


Fig. 3.14 Genetic algorithm fitness function convergence (synthetic data)

cally generated data from the reference FOM. To converge to the identified values, a total of 12500 individuals were subjected to fitness test over 250 generations (Figure 3.14). The identified parameters $\boldsymbol{\mu}^{opt} = (1381, 707, 4 \times 10^{-19}, 2 \times 10^{-19}, 269)$ are very close to the target parameter values $\boldsymbol{\mu}^{ref} = (1400, 700, 4 \times 10^{-19}, 2 \times 10^{-19}, 280)$ with a 4% maximum discrepancy occurring in the shear modulus G_{vh}, see Figure 3.15. As discussed in the thermal case, it is expected that there is a shift in the minimum of the optimization search space, resulting from the synthetic data generated from FOM instead of ROM, which explains the optimization's inability to converge perfectly to the target parameter values.

Real Measurement Data

In place of synthetically generated data, we use real pore water pressure measurements at PD-TD85E, PD-TD93E, PD-TD98E1, PD-TD98E2, PD-TD98E3 and set up the inverse problem using the same genetic algorithm configuration. There are 261 generations and 13 050 fitness tests evaluated before convergence is achieved. The linear constraints are defined more rigidly for elastic moduli $(1.8 < \frac{E_h}{E_v} < 2.2)$, $(\frac{G_{vh}}{E_h} < 0.4)$ and permeabilities $(1.8 < \frac{k_h}{k_v} < 2.2)$. To investigate the quality of the identified parameter values, $\boldsymbol{\mu}^{opt} = (1061, 483, 3.3 \times 10^{-19}, 1.8 \times 10^{-19}, 422)$, its corresponding RB-projected solution is fitted over the measured pore water pressure sensor data. Figure 3.16 illustrates that the surrogate model reproduces the measured pressure val-



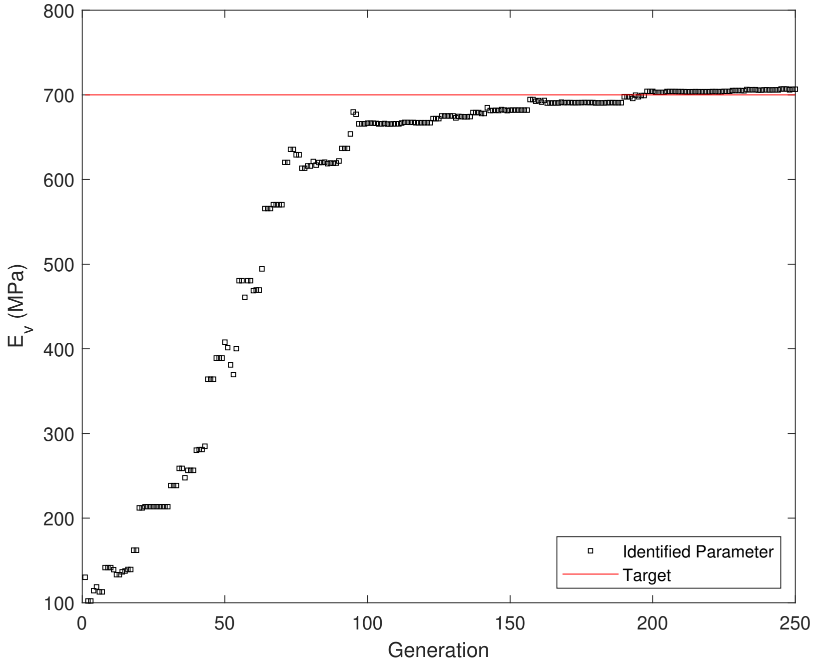


Fig. 3.15 Unknown HM properties convergence to identified values (synthetic data)

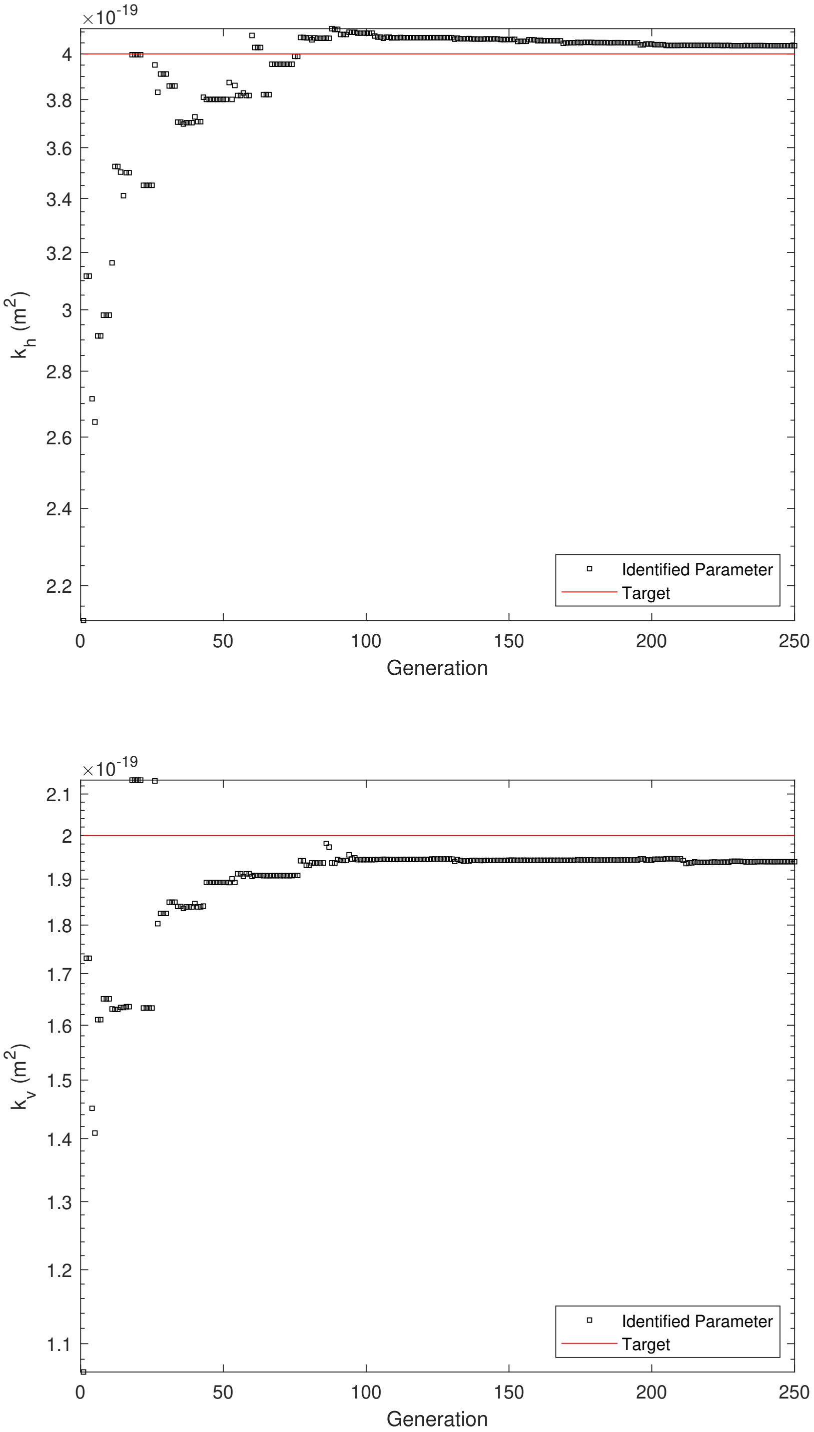


Fig. 3.15 Unknown HM properties convergence to identified values (synthetic data)

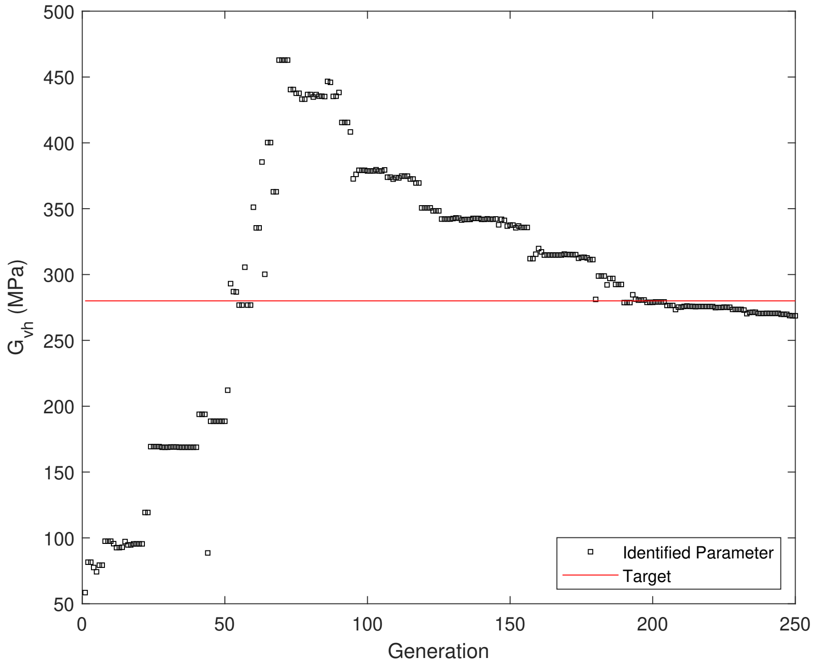


Fig. 3.15 Unknown HM properties convergence to identified values (synthetic data)

ues reasonably well, except for sensor PP-TD98E3 in which the discrepancy between simulated and measured values is significantly larger than in other sensors starting from time, t = 300 days.

As previously mentioned, the aim of this study is not to further interpret the in-situ experiment results. We focus our attention instead on the computational gain demonstrated in the solution of the inverse problem. Each of the 13 050 fitness tests is carried out more efficiently due to the significantly reduced system size of the surrogate model with a 1:1130 degrees of freedom reduction. The RB-projected solution for one parameter realization is obtained in less than 40 sec., which is several orders of magnitude faster than using a standard FE code (approximately 11 hours). Thus for a single inverse problem with thousands of parameter realizations, it is still much cheaper to use a RB surrogate model even if the offline RB generation cost is added. It would take approximately 6000 days using the direct FE approach to evaluate 13 050 parameter realizations compared to 22 days (16 days offline RB generation, 6 days online parameter realizations) using the methodology presented here.

The actual inverse analysis simulations were executed on 10 cores of CPU with 3.2 GHz processor that allowed parallel fitness test evaluations for individuals belonging to the same generation. This reduced the total run time of the parameter identification procedure to less than 18 hours. It is worth noting that for the direct FE approach, even if parallel computing is implemented, solving an identification problem of

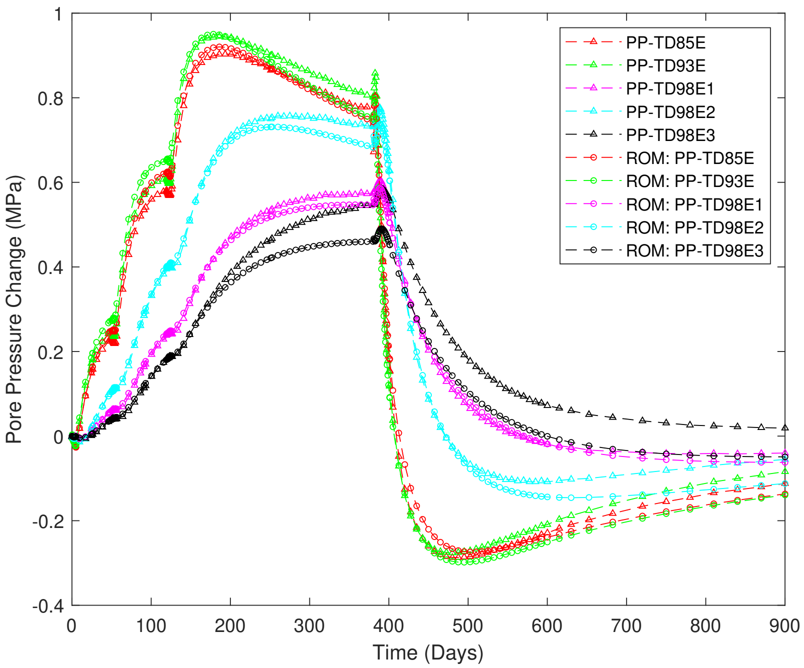


Fig. 3.16 Comparison of measured pore water pressure with modelled pore water pressure using identified parameters

Table 3.6 Computational time of inverse analysis using real sensor data

	ROM	FOM
OFFLINE	16 days	–
ONLINE: 1-Core CPU	6 days	6000 days*
ONLINE: 10-Core CPU	18 hrs	600 days*

*expected run time based on extrapolation of FE simulation time

this scale for hundreds of days becomes impractical. Table 3.6 summarizes the computational time of inverse analysis using real measurement data.

6 Summary and Findings

In this work, we address a coupled THM problem described as a parametrized system where variation of material properties is of special interest. We tackle the typical difficulties that come with high dimensional systems by building a reduced basis certified with goal-oriented error estimators.

The illustrative THM problem for waste repository applications demonstrated the capabilities of carefully constructed reduced order models to act as surrogates to replace full order models in inverse analysis. The accuracy of reduced models is validated by comparing its generated solutions against full order solutions. Synthetic sensor data are used to show that unknown THM properties can be identified very close to the target values using the proposed methodology. With real sensor data, good agreement between simulated solutions from identified parameters and in-situ measurements is observed. Both of the reduced models built for thermal and coupled THM system have greater than 1:1000 degrees of freedom reduction ratio which allowed reaching practical identification costs. Such computational gain is critical when multiple parameter realizations are involved where it becomes impractical for direct FE simulations to perform inverse analysis. Aside from parameter identification, the methodology presented here can be used to analyze the sensitivity of the coupled processes and investigate predictive modelling scenarios without the expensive computational costs typically attached to full order model simulations.

4

Model Order Reduction of Nonlinear Elasto-plastic THM Systems

This chapter is based on
Larion, Y., Massart, T. J., Díez, P., Chen, G.J., Seetharam S., and Zlotnik, S.**Nonlinear model order reduction of coupled THM systems by means of elastic and plastic domain sub-structuring.** *In preparation.*

1 Introduction

The modelling of coupled thermo-hydro-mechanical (THM) systems is extensively required in various areas of geomechanics due to the multiphase nature of geomaterials. For many geo-environmental applications such as nuclear waste repository, geo-thermal energy extraction and reservoir engineering, the coupled THM model that describes the geological material response is observed to exhibit a highly nonlinear behaviour particularly in the mechanical aspect. Studying the numerical difficulties arising from the typically large and complex coupled systems and developing modeling solutions have been the focus of several research efforts. Performance of standard finite element (FE) modelling schemes are often improved either by using parallel solvers [99, 138, 144, 149], implementing object-oriented programming paradigms [148], devising algebraic multi-grid methods [150, 151] or applying staggered Newton schemes [99, 131]. So far, very few research studies have however focused on developing model order reduction techniques specifically tailored to THM systems in a geomechanical context [22, 81].

When dealing with nonlinear systems, the computational gain achieved from model order reduction techniques are often limited. Reduced order modelling schemes are typically formulated with offline and online stage procedures, wherein expensive system matrices are pre-computed and stored during the offline stage, resulting in lower computation times during the online stage. This holds valid for linear systems in which matrices can be computed independently of intermediate solutions, with online calculation costs that can be liberated from discretization dependency. For systems involving path-dependent constitutive models, the assembly of the reduced problem has to be embodied online. As new system matrices are updated in the nonlinear iteration, the constitutive law needs to be evaluated at each integration point regardless of the reduced subspace dimension. Consequently, many of the established model order reduction techniques fail to reduce the computational cost of nonlinear systems effectively [28].

A class of techniques called ‘hyper-reduction’ was therefore developed to accelerate the evaluation of the nonlinearity in the field of model reduction. One of the most commonly used techniques is discrete empirical interpolation method (DEIM), which uses linear combinations of collateral basis functions to approximate the nonlinear contributions. The coefficients are obtained from interpolation based on nonlinear functions sampled at a much smaller number of points [28, 61]. Gappy Proper Orthogonal Decomposition (POD) is implemented using a similar approach to DEIM and differs mainly by replacing the interpolation procedure with linear regression [49].

A modified version of Gappy POD that uses Petrov-Galerkin projection to improve the stability of reduced order models [23] is referred to as Gauss-Newton Approximated Tensors (GNAT) method. Reduced Integration Domain (RID) [125, 126] is a hyper-reduction technique that defines test functions with support confined to the reduced domain. The resulting hyper-reduced equations constitute a Petrov-Galerkin formulation of the balance equations which allows for model reduction while keeping the constitutive model equations unchanged.

We now shift the attention to model order reduction of nonlinear systems in which nonlinearity more particularly arises from plastic behaviour of materials. Several intrusive approaches have been proposed in which modifications of the high fidelity or full order (e.g. finite element) model code are required. The study presented in [77] combined RID hyper-reduction with POD and augmented Newton-Krylov algorithms to solve highly nonlinear mechanical problems. The local dependency in elasto-plastic models was tackled in [127, 128] through multi-level RID hyper-reduction, and [155] extended the technique to handle thermo-elasto-plastic models. In [69], Empirical Cubature hyper-reduction method was used for the model reduction of elasto-plastic composite plates. Such hyper-reduction methods are however found to be susceptible to robustness deficiencies, in addition to being highly problem-dependent which limits their effectiveness for complex nonlinear problems. Conversely, more versatile non-intrusive methods have also been proposed which involve a more general approach in modelling nonlinear constitutive laws through a ‘black-box’ model. A study tackled elasto-plasticity in tunnel engineering problems using a multi-grid strategy [25], involving the computation of less expensive and less accurate FE approximations during the online stage and an improvement of those solutions using a reduced basis-based rectification method. In [64], Gaussian process regression was used to ensure a full decoupling between offline and online stages: projection coefficients are generated with supervised learning in the offline stage, and are subsequently combined with scalar parameter values to rapidly predict the reduced basis solution during the online stage. Similarly, a purely data-driven approach using forward feed neural networks training was proposed in [118] for laminated composites, while [58] proposed a recurrent neural network surrogate to simulate history-dependent material models for quasi-brittle failure analysis. Black-box models are known to have limited extrapolation capabilities which poses a risk when using them away from their training sets. Furthermore, the physical information of the problem is not directly preserved in the black box model, which can lead to high errors and nonsensical predictions if inadequate data are trained.

In this contribution, we propose an intrusive approach that combines traditional Galerkin reduced basis (RB) projection and sub-structuring by domain decomposition to effectively reduce the computational cost involved in solving nonlinear elasto-plastic THM systems. The sub-structuring strategy stems from the early works on component mode synthesis [36, 72] which were then further extended for various elasto-plastic problems [35, 114, 156]. The occurrence of plasticity in specific zones of the models (e.g. around cavities and galleries in underground storage facilities) is then exploited by defining sub-domains corresponding to elastic, plastic and interface regions and by constructing localized reduced subspaces to perform multi-level reduced basis projection [66, 108]. We focus our study on extending these strategies for coupled THM systems in geo-environmental applications with the nonlinearity originating from the mechanical response. Therefore, the paper is organized as follows: in Section 2, we define the formulation of the coupled THM system along with the considered nonlinear elasto-plastic constitutive laws in the full order and reduced order model forms. The sub-structuring methodology combined with model order reduction is explained in Section 3. Numerical examples describing a model for an in-situ heating test for nuclear waste repository applications are presented in Section 4. An inverse identification procedure is also implemented to further illustrate the expected computational speed up that can be attained with the proposed methodology in a multi-query problem setting. Conclusions are finally drawn in Section 5.

2 Problem Definition: Nonlinear Elasto-plastic THM System

2.1 Governing Equations and Full Order Discretization

The coupled THM system for a fully-saturated porous geomaterial is described by a set of three equations characterizing the deformation of the elasto-plastic body, the flow of the fluid phase and the heat conduction. The mechanical behaviour is governed by an elasto-plastic constitutive law while the hydraulic and thermal responses are governed by Darcy's law and Fourier's law, respectively. The time-dependent coupled THM system takes the strong form,

$$\left.\begin{aligned} k_c\nabla^2 T - \rho c_p \dot{T} + Q = 0 \\ \nabla\cdot(\boldsymbol{C} : (\epsilon(\mathbf{u}) - \epsilon_p(\mathbf{u}))) - \alpha\nabla p - \Theta\nabla T = 0 \\ -\frac{k}{\mu_f(T)}\nabla\cdot(\nabla p) + \alpha\nabla\cdot\dot{\mathbf{u}} + \frac{1}{M}\dot{p} - [\phi 3\alpha_f + (\alpha-\phi)3\alpha_s]\dot{T} = 0 \end{aligned}\right\} \text{in } \Omega\times(0,t_f] \tag{4.1}$$

with Neumann boundary conditions

$$\left.\begin{aligned} \psi\cdot\mathbf{n} = g_N \\ \boldsymbol{\sigma}\cdot\mathbf{n} = \boldsymbol{\sigma}_N \\ \Upsilon\cdot\mathbf{n} = p_N \end{aligned}\right\} \qquad \text{on } \partial\Gamma_N\times(0,t_f], \tag{4.2}$$

Dirichlet boundary conditions

$$\left.\begin{aligned} T = T_D \\ \mathbf{u} = \mathbf{u}_D \\ p = p_D \end{aligned}\right\} \qquad \text{on } \partial\Gamma_D\times(0,t_f], \tag{4.3}$$

and initial conditions

$$\left.\begin{aligned} T = T_0 \\ \mathbf{u} = \mathbf{u}_0 \\ p = p_0 \end{aligned}\right\} \qquad \text{in } \Omega\times\{t=0\}, \tag{4.4}$$

where $T, \mathbf{u}$ and p are the field variables corresponding to temperature, displacement and pore water pressure. The thermal expansion coefficient in the fluid phase (Figure

4.1) is assumed to be temperature-dependent [9, 38, 57]. The liquid dynamic viscosity is also temperature-dependent, obeying the empirical function [56, 146]

$$\mu_f(T) = 2.1 \times 10^{-12} \exp\left(\frac{1808.5}{273.15 + T}\right) \tag{4.5}$$

where μ_f is in MPa · s and T is expressed in Celsius. Assuming isotropic conditions for a 2D plane-strain analysis, the elasticity tensor contains two independent coefficients as described in (4.6). The notation for the physical parameters used in the system of equations is given in Table 4.1.

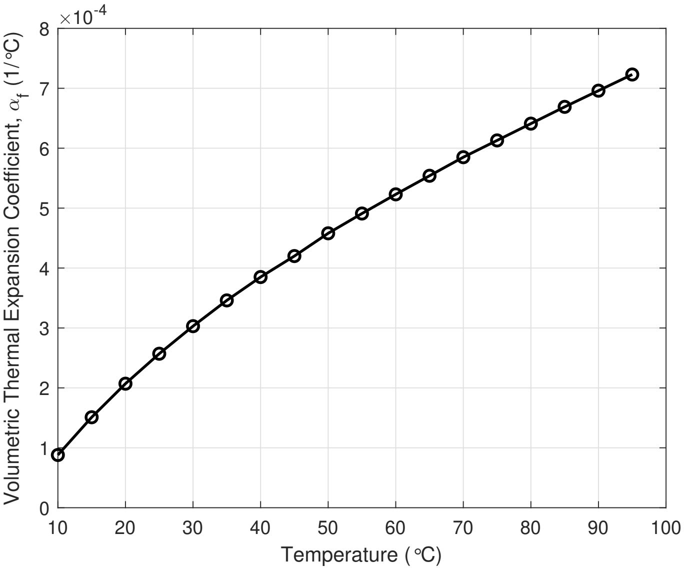


Fig. 4.1 Temperature-dependent thermal expansion coefficient in the fluid phase

$$\boldsymbol{C} = \begin{bmatrix} \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} & \frac{E\nu}{(1+\nu)(1-2\nu)} & 0 \\ \frac{E\nu}{(1+\nu)(1-2\nu)} & \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} & 0 \\ 0 & 0 & \frac{E}{2(1+\nu)} \end{bmatrix} \tag{4.6}$$

The solution of the problems defined by (4.1)-(4.4) is approximated with the finite element (FE) method by discretizing the domain Ω using a mesh of characteristic

Table 4.1 Physical parameters of the THM system

Parameters	
E	Young's modulus
ν	Poisson's ratio
α	Biot-Willis coefficient
M	constrained specific storage coefficient
μ_f	viscosity
k	permeability
ϕ	porosity
Θ	thermal expansion factor
α_s	thermal expansion coefficient in the solid phase
α_f	thermal expansion coefficient in the fluid phase
c_p	bulk specific heat
$c_{p,s}$	solid specific heat
k_c	thermal conductivity
Q	heat source
ψ	thermal flux
σ	traction
Υ	fluid velocity

size H in the functional space $\mathscr{Z}^H \subset \mathcal{H}_0^1(\Omega)$, $\boldsymbol{\mathcal{W}}^H \subset [\mathcal{H}_0^1(\Omega)]^2$, $\mathscr{Q}^H \subset \mathcal{H}_0^1(\Omega)$:

$$\begin{aligned} T(\boldsymbol{x},t) \approx T^H(\boldsymbol{x},t) &= \sum_{i=1}^{\mathtt{n_{nodes}}} T_i(t)N_i(\boldsymbol{x}) = \mathbf{N}^\mathsf{T}\boldsymbol{T} \\ \mathbf{u}(\boldsymbol{x},t) \approx \mathbf{u}^H(\boldsymbol{x},t) &= \sum_{i=1}^{\mathtt{n_{nodes}}} \mathbf{u}_i(t)\tilde{N}_i(\boldsymbol{x}) = \tilde{\mathbf{N}}^\mathsf{T}\boldsymbol{U} \\ p(\boldsymbol{x},t) \approx p^H(\boldsymbol{x},t) &= \sum_{i=1}^{\mathtt{n_{nodes}}} p_i(t)N_i(\boldsymbol{x}) = \mathbf{N}^\mathsf{T}\boldsymbol{p} \end{aligned} \tag{4.7}$$

where $\boldsymbol{U} \in \mathbb{R}^{d\times\mathtt{n_{nodes}}}$, with $\mathtt{n_{nodes}}$ the number of nodes discretizing the field in Ω, while $\mathbf{N}$ and $\tilde{\mathbf{N}}$ are compact support shape functions. For simplicity, a time-discrete framework is considered directly for the interval $I \equiv (0, t_f]$ such that the subintervals are denoted as $\{t^0 < t^1 < ... < t^l < .... < t^{\mathcal{L}} = t_f\}$ for $l = 1, 2, ..., \mathcal{L}$. For notation purposes, let Δt^l be the characteristic time step of the time-discrete framework, measured as $\Delta t^l = t^l - t^{l-1}$.

Considering the elasto-plastic system, the weak form is written in a compact format by separating the nonlinear term (A_p) containing the plastic strain (ϵ_p) as follows:

$$\begin{aligned} &\Delta t^l A\ (T(\mu,t^l),\mathbf{u}(\mu,t^l),p(\mu,t^l);z,\mathbf{w},q) - \Delta t^l A_p\ (\epsilon_p(\mathbf{u}(\mu,t^l);\mathbf{w}) \\ &\qquad + M\ (T(\mu,t^l),\mathbf{u}(\mu,t^l),p(\mu,t^l);z,\mathbf{w},q) \\ &= M\ (T(\mu,t^{l-1}),\mathbf{u}(\mu,t^{l-1}),p(\mu,t^{l-1});z,\mathbf{w},q) + \Delta t^l L\ (z,w,q;t^l) \end{aligned} \tag{4.8}$$

where

$$\begin{aligned} A\ (T(\cdot),\mathbf{u}(\cdot),p(\cdot);z,\mathbf{w},q) &= \int_\Omega k_c \frac{\partial z}{\partial x_j}\frac{\partial T}{\partial x_j} d\Omega + \int_\Omega \epsilon(\mathbf{u}) : \boldsymbol{C} : \epsilon(\mathbf{w}) d\Omega \\ &- \int_\Omega \mathbf{w}_i \alpha \frac{\partial p}{\partial x_i} d\Omega - \int_\Omega \mathbf{w}_i \Theta \frac{\partial T}{\partial x_i} d\Omega - \int_\Omega \frac{k}{\mu_f}\frac{\partial q}{\partial x_j}\frac{\partial p}{\partial x_j} d\Omega \end{aligned} \tag{4.9}$$

$$A_p\ (\epsilon_p(\mathbf{u}(\cdot);\mathbf{w}) = \int_\Omega \epsilon_p(\mathbf{u}) : \boldsymbol{C} : \epsilon(\mathbf{w}) d\Omega \tag{4.10}$$

$$\begin{aligned} M\ (T(\cdot),\mathbf{u}(\cdot),p(\cdot);z,\mathbf{w},q) &= \int_\Omega z\rho c_p T d\Omega + \int_\Omega q\frac{1}{M} p d\Omega \\ &+ \int_\Omega q\alpha(\frac{\partial \mathbf{u}_i}{\partial x_i}) d\Omega - \int_\Omega q[\phi 3\alpha_f + (\alpha-\phi)3\alpha_s] T d\Omega \end{aligned} \tag{4.11}$$

for $i,j = 1,2$ and

$$L(z,\mathbf{w},q) = \int_\Gamma z g_N d\Gamma + \int_\Gamma \mathbf{w}\boldsymbol{\sigma}_N d\Gamma - \int_\Gamma q p_N d\Gamma. \tag{4.12}$$

The substitution of the approximation (4.7) in the weak form (4.8) leads to the discrete matrix form,

$$\mathbb{K}\ \mathbb{U} + \mathbb{C}\ \dot{\mathbb{U}} = \mathbb{F} \tag{4.13}$$

$$\underbrace{\begin{bmatrix} \mathbf{K_H} & 0 & 0 \\ \mathbf{K_T} & \mathbf{K_U} & \mathbf{C_P} \\ 0 & 0 & \mathbf{K_P} \end{bmatrix}}_{\mathbb{K}} \underbrace{\begin{bmatrix} \boldsymbol{T} \\ \boldsymbol{U} \\ \boldsymbol{p} \end{bmatrix}}_{\mathbb{U}} + \underbrace{\begin{bmatrix} \mathbf{C_H} & 0 & 0 \\ 0 & 0 & 0 \\ \mathbf{C_T} & \mathbf{C_P^\intercal} & \mathbf{C_M} \end{bmatrix}}_{\mathbb{C}} \underbrace{\begin{bmatrix} \dot{\boldsymbol{T}} \\ \dot{\boldsymbol{U}} \\ \dot{\boldsymbol{p}} \end{bmatrix}}_{\dot{\mathbb{U}}} = \underbrace{\begin{bmatrix} \mathbf{F_T} \\ \mathbf{F_u} \\ \mathbf{F_p} \end{bmatrix}}_{\mathbb{F}} \tag{4.14}$$

where $\mathbb{K}$ contains the stiffness and conductivity matrices, $\mathbb{C}$ contains the capacity (time-dependent) matrices and $\mathbb{F}$ contains the load vectors corresponding to $L(\cdot)$. For the case in which uni-directional coupling is observed due to hydro-mechanical processes having no direct effect on the thermal process, the thermal balance equation can be separated from the monolithic system (4.13). Such non-monolithic system permits linear solving approach even with the presence of temperature-dependent parameters (μ_f, α_f) in the hydro-mechanical system as the thermal system is separately solved prior to the HM system.

$$\mathbf{K_H}\ \boldsymbol{T} + \mathbf{C_H}\ \dot{\boldsymbol{T}} = \mathbf{F_T} \tag{4.15}$$

$$\begin{bmatrix} \mathbf{K_U} & \mathbf{C_P} \\ 0 & \mathbf{K_P} \end{bmatrix} \begin{bmatrix} \boldsymbol{U} \\ \boldsymbol{p} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \mathbf{C}_\mathbf{P}^{\mathsf{T}} & \mathbf{C_M} \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{U}} \\ \dot{\boldsymbol{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{F_u} - \mathbf{K_T}\boldsymbol{T} \\ \mathbf{F_p} - \mathbf{C_T}\dot{\boldsymbol{T}} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{F}}_\mathbf{u} \\ \bar{\mathbf{F}}_\mathbf{p} \end{bmatrix} \tag{4.16}$$

In general, the nonlinear components in equation (4.16) reflecting the elasto-plastic constitutive law is solved by Newton-Raphson algorithm which results in the linearized form

$$\bar{\mathbf{K}}_\mathbf{U}(\boldsymbol{U}_l^k)\ \Delta\boldsymbol{U}_l = -\mathbf{G}(\boldsymbol{U}_l^k) \tag{4.17}$$

at each time step l and iteration step k. The residual vector $\mathbf{G}$ denotes the difference between the internal and external forces. The tangential stiffness matrix is defined as $\bar{\mathbf{K}}_\mathbf{U}(\boldsymbol{U}_l^k) = \frac{\partial \mathbf{G}(\boldsymbol{U}_l}{\partial \boldsymbol{U}_l}|_{\boldsymbol{U}_l^k}$ while the displacement vector is updated with $\boldsymbol{U}_l^{k+1} = \boldsymbol{U}_l^k + \Delta\boldsymbol{U}_l$.

2.2 Reduced Order Discrete System

We now consider a parametrized system defined by a set of parameter values $\mathcal{M} = \{\boldsymbol{\mu}^1, ..., \boldsymbol{\mu}^N\}$ with the corresponding solutions $S^N = \{(T(\boldsymbol{\mu}^1), \mathbf{u}(\boldsymbol{\mu}^1), p(\boldsymbol{\mu}^1)), ...,$ $(T(\boldsymbol{\mu}^N), \mathbf{u}(\boldsymbol{\mu}^N),\ p(\boldsymbol{\mu}^N))\}$ denoted as snapshots, where N is typically much smaller than the characteristic size of the reference full order finite element model. To inherit the stability condition of the reference full order model, the global approximation spaces are defined separately for each field variable:

$$\begin{aligned} \mathcal{Z}_N^H &= \mathrm{span}\{T(\boldsymbol{\mu}^n), \quad 1 \le n \le \mathtt{n_{snapshots}}\} \subset \mathcal{Z}^H \\ \boldsymbol{\mathcal{W}}_N^H &= \mathrm{span}\{\mathbf{u}(\boldsymbol{\mu}^n), \quad 1 \le n \le \mathtt{n_{snapshots}}\} \subset \boldsymbol{\mathcal{W}}^H \\ \mathcal{Q}_N^H &= \mathrm{span}\{p(\boldsymbol{\mu}^n), \quad 1 \le n \le \mathtt{n_{snapshots}}\} \subset \mathcal{Q}^H \end{aligned} \tag{4.18}$$

The RB projection matrices are obtained by subjecting each snapshots separately to a Proper Orthogonal Decomposition (POD) technique [27]. POD modes are truncated based on a prescribed tolerance value corresponding to the desired level of accuracy. The orthonormalized and truncated modes denote the RB projection matrices ($\mathbb{Z}, \mathbb{W}$ and $\mathbb{Q}$) for temperature, displacement and pore pressure respectively. To project the reference model onto global approximation RB spaces, the full system is decoupled and grouped into blocks based on the field variable it interacts with. Each block is projected onto its corresponding bases (4.21) and assembled similar to the non-monolithic FE matrix system (4.16). The resulting reduced order blocks (4.20) are described by a much smaller characteristic size of order N.

$$\mathbf{K}_{\mathbf{H}N}\ \boldsymbol{T}_N + \mathbf{C}_{\mathbf{H}N}\ \dot{\boldsymbol{T}}_N = \mathbf{F}_{\mathbf{T}N} \tag{4.19}$$

$$\begin{bmatrix} \mathbf{K}_{\mathbf{U}N} & \mathbf{C}_{\mathbf{P}N} \\ 0 & \mathbf{K}_{\mathbf{P}N} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_N \\ \boldsymbol{p}_N \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \mathbf{C}^{\mathsf{T}}_{\mathbf{P}N} & \mathbf{C}_{\mathbf{M}N} \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{U}}_N \\ \dot{\boldsymbol{p}}_N \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{\mathbf{u}N} \\ \mathbf{F}_{\mathbf{p}N} \end{bmatrix} \tag{4.20}$$

$$\begin{array}{ccc} \mathbf{K}_{\mathbf{H}N} = \mathbb{Z}^{\mathsf{T}}\ \mathbf{K}_{\mathbf{H}}\ \mathbb{Z} & \mathbf{C}_{\mathbf{H}N} = \mathbb{Z}^{\mathsf{T}}\ \mathbf{C}_{\mathbf{H}}\ \mathbb{Z} & \mathbf{F}_{\mathbf{T}N} = \mathbb{Z}^{\mathsf{T}}\ \mathbf{F}_{\mathbf{T}} \\ \mathbf{K}_{\mathbf{U}N} = \mathbb{W}^{\mathsf{T}}\ \mathbf{K}_{\mathbf{U}}\ \mathbb{W} & \mathbf{C}_{\mathbf{P}N} = \mathbb{W}^{\mathsf{T}}\ \mathbf{C}_{\mathbf{P}}\ \mathbb{Q} & \mathbf{K}_{\mathbf{P}N} = \mathbb{Q}^{\mathsf{T}}\ \mathbf{K}_{\mathbf{P}}\ \mathbb{Q} \\ \mathbf{C}^{\mathsf{T}}_{\mathbf{P}N} = \mathbb{Q}^{\mathsf{T}}\ \mathbf{C}^{T}_{\mathbf{P}}\ \mathbb{W} & \mathbf{C}_{\mathbf{M}N} = \mathbb{Q}^{\mathsf{T}}\ \mathbf{C}_{\mathbf{M}}\ \mathbb{Q} & \mathbf{F}_{\mathbf{u}N} = \mathbb{W}^{\mathsf{T}}\ \bar{\mathbf{F}}_{\mathbf{u}} \\ \mathbf{F}_{\mathbf{p}N} = \mathbb{Q}^{\mathsf{T}}\ \bar{\mathbf{F}}_{\mathbf{p}} & \boldsymbol{U}_N = \mathbb{W}^{\mathsf{T}}\ \boldsymbol{U} & \boldsymbol{p}_N = \mathbb{Q}^{\mathsf{T}}\ \boldsymbol{p} \\ \boldsymbol{T}_N = \mathbb{Z}^{\mathsf{T}}\ \boldsymbol{T} & \dot{\boldsymbol{T}}_N = \mathbb{Z}^{\mathsf{T}}\ \dot{\boldsymbol{T}} & \dot{\boldsymbol{U}}_N = \mathbb{W}^{\mathsf{T}}\ \dot{\boldsymbol{U}} \end{array} \tag{4.21}$$

The re-assembled discrete matrix system is solved the same way as the full order finite element model. In the subsequent sections, we will denote all reduced field variables with the subscript N.

3 Nonlinear Model Order Reduction Strategy

The nonlinear system described in (4.8) is classically solved in finite element analysis by the iterative procedure outlined in Algorithm 4.1, giving extra attention to the mechanical aspect containing the nonlinear contribution of the elasto-plastic material behaviour. Note that some notations are further simplified for better readability, i.e. $(u(\mu, t^l) \rightarrow u^l)$. The constitutive model describing the mechanical system is governed by a non-associated elasto-plastic law. In particular, we consider the Drucker-Prager criterion [46], with a yield surface expressed as,

$$f(\sigma) \equiv {J'_2}^{\frac{1}{2}} - \frac{2\sin\varphi}{\sqrt{3}(3-\sin\varphi)}(J_1 + \frac{3c}{\tan\varphi}) - \kappa(\lambda) = 0 \tag{4.22}$$

where J'_2 is the second invariant of the deviatoric stress tensor, J_1 is the first stress invariant, φ is the friction angle in compression and c is the cohesion of the soil. A non-associated flow is considered by using a slightly different plastic potential function defined as

$$g(\sigma) \equiv {J'_2}^{\frac{1}{2}} - \frac{2\sin\psi}{\sqrt{3}(3-\sin\psi)}(J_1 + \frac{3c}{\tan\varphi}) - \kappa(\lambda) = 0 \tag{4.23}$$

where ψ is the angle of dilatancy. Isotropic hardening is considered for the yield surface as defined by

$$\kappa(\lambda) \equiv H\lambda^{\mathsf{n_x}} \tag{4.24}$$

where λ is the plastic multiplier while H and $\mathsf{n_x}$ are generic hardening function coefficients.

In general, the incremental iterative procedure used to deal with the material nonlinearity is organized in two levels: global and local iterations. The global level iteration checks equilibrium (i.e. such that balance of forces is achieved by comparing internal forces with external loads). A local level iterative procedure is used to solve the constitutive model equations, where the stresses and internal variables are iteratively sought based on a total strain state furnished by the equilibrium iteration until convergence is reached. In Algorithm 4.1, the global iteration is implemented inside the **while** loop (line 4-22). The local iteration involved in finding the plastic state variables is implemented inside a nested loop (line 13-15). Although the local level constitutive model integration could also be performed explicitly, an implicit technique is used in this study. This is achieved using a ‘return mapping’ algorithm, which first integrates the elastic equations with total strain increments in the prediction stage and later relaxes the elastically predicted stresses onto the yield surface by iteratively correcting the plastic strain increments [137].

Algorithm 4.1 Elasto-plastic THM system iterative solver

```
Initialization $l = 0,\ t^l = 0,\ \epsilon_p^0 = 0,\ \sigma^0 = 0$
for $l = 1$ to $\mathcal{L} \rightarrow t^l$
    $L(z, \mathbf{w}, q) \leftarrow L(z, \mathbf{w}, q; t^l)$
    while $\mathsf{k} < \mathsf{Niter}_{\mathsf{max}}$ and $\mathsf{error} > \mathsf{tol}$
        $\mathsf{k} \leftarrow \mathsf{k} + 1$
        Compute $(T^{l,\mathsf{k}},\ \mathbf{u}^{l,\mathsf{k}}\ p^{l,\mathsf{k}})$ such that
        $\Delta t^{l,\mathsf{k}} A(T^{l,\mathsf{k}}, \mathbf{u}^{l,\mathsf{k}}, p^{l,\mathsf{k}}; z, \mathbf{w}, q) \ +\ M(T^l, \mathbf{u}^{l,\mathsf{k}}, p^{l,\mathsf{k}}; z, \mathbf{w}, q)$
                 $- \Delta t^l A_p(\epsilon_p(\mathbf{u}^{l,\mathsf{k}-1}; \mathbf{w}) = M\ (T^{l-1}, \mathbf{u}^{l-1}, p^{l-1}; z, \mathbf{w}, q)$
                 $+ \Delta t^l L\ (z, w, q)$
        Compute $\hat{\sigma}^{\mathsf{k}} = \boldsymbol{C} : \epsilon(\mathbf{u}^{l,\mathsf{k}})$
        for all elements, $\mathsf{x}_\Omega = 1 : \mathsf{N}_{\mathsf{elem}}$
            for all integration points, $\mathsf{x_p} = 1 : \mathsf{n}_{\mathsf{intp}}$
                if $f(\hat{\sigma}^{\mathsf{k}}(\mathsf{x_p})) > 0$ then
                Plastic State: $\lambda_{\mathsf{k}} = \frac{f(\hat{\sigma}^{\mathsf{k}})}{\frac{\partial f(\hat{\sigma}^{\mathsf{k}})}{\partial \sigma} : \boldsymbol{C} : \frac{\partial g(\hat{\sigma}^{\mathsf{k}})}{\partial \sigma}}$
                        $\epsilon_p(\mathbf{u}^{l,\mathsf{k}}) = \lambda_{\mathsf{k}} \frac{\partial f}{\partial \sigma}(\sigma^{\mathsf{k}-1})$
                        $\sigma^{\mathsf{k}} = \hat{\sigma}^{\mathsf{k}} - \boldsymbol{C} : \epsilon_p^{\mathsf{k}}$
                else
                Elastic State: $\sigma^{\mathsf{k}} = \hat{\sigma}^{\mathsf{k}}$ and $\epsilon_p(\mathbf{u}^{l,\mathsf{k}}) = \epsilon_p(\mathbf{u}^{l,\mathsf{k}-1})$
            end for
        end for
        $\mathsf{error} = \Delta t^l A(T^{l,\mathsf{k}}, \mathbf{u}^{l,\mathsf{k}}, p^{l,\mathsf{k}}; z, \mathbf{w}, q) + M(T^{l,\mathsf{k}}, \mathbf{u}^{l,\mathsf{k}}, p^{l,\mathsf{k}}; z, \mathbf{w}, q)$
                 $-\ \Delta t^l A_p(\epsilon_p(\mathbf{u}^{l,\mathsf{k}}; \mathbf{w}) - M\ (T^{l-1}, \mathbf{u}^{l-1}, p^{l-1}; z, \mathbf{w}, q)$
                 $-\ \Delta t^l L\ (z, w, q)$
    end while
    $\mathbf{u}^l \leftarrow \mathbf{u}^{l,\mathsf{k}},\ \sigma^l \leftarrow \sigma^{\mathsf{k}},\ \epsilon_p^l \leftarrow \epsilon_p^{l,\mathsf{k}}$
    $T^l \leftarrow T^{l,\mathsf{k}},\ p^l \leftarrow p^{l,\mathsf{k}}$
end for
```

In the context of model order reduction, efficiency mainly relies on liberating the computational cost from the discretization dependency, and on smartly implementing the offline-online stage solving procedure. For the reduced basis (RB) method, the low dimensional subspace is built from snapshots [27] and the system of equations is affinely decomposed such that parameter-independent matrices can be pre-computed and stored during the offline stage. The pre-computed matrices are reused during the online stage when system matrices are assembled prior to projection onto the low dimensional subspace [61, 107]. In the case of elasto-plastic materials, the affine decomposition procedure is not fully applicable since path-dependent constitutive models do not allow the extraction of parameter-independent matrices during the offline stage. Thus, the system matrices need to be generated and assembled at every iteration which can be extremely expensive to implement. Furthermore, it is worth noting that the constitutive model describing the nonlinear elasto-plastic behaviour cannot be liberated from the discretized space as the plastic constitutive law has to be treated locally at the integration points level.

The strategy proposed here takes advantage of the typical scenario for coupled THM problems in geomechanics, where the occurrence of plasticity is restricted to specific regions of small size with respect to the model size [29, 30, 45]. We divide the full domain into two sub-domains: an elastic region governed by linear THM equations and a plastic region governed by the nonlinear elasto-plastic constitutive law. We exploit these sub-domains by confining the time-consuming procedures in the plastic region while exploiting the full advantage of model order reduction for linear systems in the elastic region. The following sections further explain the three step strategy – decomposition, reduction and re-assembly to form a reduced order model representing the original full order model in the best possible way.

3.1 Domain Decomposition

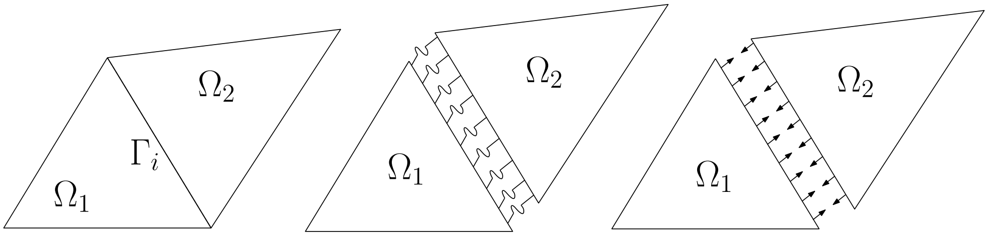


Fig. 4.2 Decomposed sub-domains connected by a fictitious interface

The central idea of the model reduction strategy for coupled THM models with locally increasing plasticification revolves around decomposing the entire domain (Ω) into plastic (Ω_1) and elastic(Ω_2) sub-domains. Figure 4.2 shows an illustration of the decomposition where the sub-domains are connected by fictitious interface. The hypothesis of displacement continuity between these identified decomposed sub-domains is enforced with a linear elastic relationship between tractions and displacement jump [35, 114]. In terms of nodal solutions, the displacement is separated into three parts,

$$\boldsymbol{U} = \begin{bmatrix} \boldsymbol{U}_{\Omega_1} \\ \boldsymbol{U}_{\Gamma_i} \\ \boldsymbol{U}_{\Omega_2} \end{bmatrix} \tag{4.25}$$

in which boundary nodes lying on the interface are separated from their previously containing sub-domain. A Boolean matrix is used to identify the pairing of the degrees of freedom (DoF) at the interface. The separated parts constitute the degrees of freedom of the plastic($\boldsymbol{U}_{\Omega_1}$), interface($\boldsymbol{U}_{\Gamma_i}$) and elastic($\boldsymbol{U}_{\Omega_2}$) domains. Consequently, the relevant matrices to the elasto-plastic constitutive law (4.16) are sub-structured in the form of

$$\begin{aligned} &\begin{bmatrix} \mathbf{K}_{\mathbf{U}\Omega_1} & \mathbf{K}_{\mathbf{U}\Omega_1,\Gamma_i} & 0 & \mathbf{C}_{\mathbf{P}\Omega_1} \\ \mathbf{K}_{\mathbf{U}\Gamma_i,\Omega_1} & \mathbf{K}_{\mathbf{U}\Gamma_i} & \mathbf{K}_{\mathbf{U}\Gamma_i,\Omega_2} & \mathbf{C}_{\mathbf{P}\Gamma_i} \\ 0 & \mathbf{K}_{\mathbf{U}\Omega_2,\Gamma_i} & \mathbf{K}_{\mathbf{U}\Omega_2} & \mathbf{C}_{\mathbf{P}\Omega_2} \\ 0 & 0 & 0 & \mathbf{K}_{\mathbf{P}} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_{\Omega_1} \\ \boldsymbol{U}_{\Gamma_i} \\ \boldsymbol{U}_{\Omega_2} \\ \boldsymbol{p} \end{bmatrix} \\ &+ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \mathbf{C}^{\mathsf{T}}_{\mathbf{P}\Omega_1} & \mathbf{C}^{\mathsf{T}}_{\mathbf{P}\Gamma_i} & \mathbf{C}^{\mathsf{T}}_{\mathbf{P}\Omega_2} & \mathbf{C}_{\mathbf{M}} \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{U}}_{\Omega_1} \\ \dot{\boldsymbol{U}}_{\Gamma_i} \\ \dot{\boldsymbol{U}}_{\Omega_2} \\ \dot{\boldsymbol{p}} \end{bmatrix} \\ &= \begin{bmatrix} \bar{\mathbf{F}}_{\mathbf{u}\Omega_1} \\ \bar{\mathbf{F}}_{\mathbf{u}\Gamma_i} \\ \bar{\mathbf{F}}_{\mathbf{u}\Omega_2} \\ \bar{\mathbf{F}}_{\mathbf{P}} \end{bmatrix}. \end{aligned} \tag{4.26}$$

With this sub-structuring approach, it is possible to affinely decompose the elastic sub-domain ($\mathbf{K}_{\mathbf{U}\Omega_2}$) and store a set of parameter-independent matrices during the offline stage. In this case, only the plastic sub-domain ($\mathbf{K}_{\mathbf{U}\Omega_1}$) needs to be explicitly computed at every global iteration in the online stage. The local return mapping iterative procedure, performed locally at every integration point to integrate the elasto-plastic constitutive law, is still fully evaluated but only in the plastic sub-domain Ω_1.

3.2 Order Reduction and Assembly

The reduced basis subspace established in previous sections is altered to conform with the proposed sub-structuring strategy. The displacement snapshot $S_{\mathbf{u}} = \{(\mathbf{u}(\boldsymbol{\mu}^1), \mathbf{u}(\boldsymbol{\mu}^2), ..., \mathbf{u}(\boldsymbol{\mu}^N)\}$ is decomposed to three parts that correspond to the degrees of freedom of the plastic, interface and elastic sub-domains. Consequently, the RB matrices for the displacement subspace are now denoted as $\mathbb{W}_{\Omega_1}, \mathbb{W}_{\Gamma_i}$ and $\mathbb{W}_{\Omega_2}$ after post-processing the decomposed snaphsots to obtain the orthonormalized bases in Algorithm 4.2. On the other hand, the reduced basis space defined for the pore pressure $\mathbb{Q}$ in (4.18) remains unchanged.

$$\boldsymbol{U}_N = \begin{bmatrix} \boldsymbol{U}_{\Omega_1 N} \\ \boldsymbol{U}_{\Gamma_i N} \\ \boldsymbol{U}_{\Omega_2 N} \end{bmatrix} = \begin{bmatrix} \mathbb{W}_{\Omega_1}^{\mathsf{T}} \boldsymbol{U}_{\Omega_1} \\ \mathbb{W}_{\Gamma_i}^{\mathsf{T}} \boldsymbol{U}_{\Gamma_i} \\ \mathbb{W}_{\Omega_2}^{\mathsf{T}} \boldsymbol{U}_{\Omega_2} \end{bmatrix} \tag{4.27}$$

To reduce the order of the sub-structured model, we perform a multi-level Galerkin projection of the decomposed domain onto their respective RB subspaces separately. In comparison to earlier studies on model reduction in plasticity where only a portion of the subdomain is reduced [114] we also perform model order reduction on the interface degrees of freedom[74, 78]. The projection procedure performed in (4.21) is extended for the decomposed displacement degrees of freedom as shown in (4.28).

Algorithm 4.2 OFFLINE Stage: Sub-structured RB construction by POD algorithm

```
    Input:   Training sample $\mathcal{M}$, Tolerance $\varepsilon_{\mathsf{POD}}$,
             Boolean matrix sub-domain DOF extractor $\mathcal{X}_{\Omega_1}, \mathcal{X}_{\Gamma_i}, \mathcal{X}_{\Omega_2}$
    Output: $\mathbb{W}_{\Omega_1}, \mathbb{W}_{\Gamma_i}, \mathbb{W}_{\Omega_2}, \mathbb{Q}$
for $\boldsymbol{\mu} \in \mathcal{M}$
    $[\mathbf{u}, p] = \mathsf{SolveFE}(\boldsymbol{\mu})$
    $S_{\mathbf{u}} \leftarrow [S_{\mathbf{u}},\ \mathbf{u}],\ S_p \leftarrow [S_p,\ p]$
end for
$S_{\mathbf{u}\Omega_1} = \mathcal{X}_{\Omega_1} S_{\mathbf{u}},\ S_{\mathbf{u}\Gamma_i} = \mathcal{X}_{\Gamma_i} S_{\mathbf{u}},\ S_{\mathbf{u}\Omega_2} = \mathcal{X}_{\Omega_2} S_{\mathbf{u}}$
$\mathbb{W}_{\Omega_1} = \mathbf{POD}(S_{\mathbf{u}\Omega_1}, \varepsilon_{\mathsf{POD}}), \mathbb{W}_{\Gamma_i} = \mathbf{POD}(S_{\mathbf{u}\Gamma_i}, \varepsilon_{\mathsf{POD}}), \mathbb{W}_{\Omega_2} = \mathbf{POD}(S_{\mathbf{u}\Omega_2}, \varepsilon_{\mathsf{POD}})$
$\mathbb{Q} = \mathbf{POD}(S_p, \varepsilon_{\mathsf{POD}})$
```

The characteristic size of the reduced system is described by the total sum of sub-structured RB modes ($N = \mathtt{n}_{\mathtt{modes},\mathbb{W}_{\Omega_1}} + \mathtt{n}_{\mathtt{modes},\mathbb{W}_{\Gamma_i}} + \mathtt{n}_{\mathtt{modes},\mathbb{W}_{\Omega_2}} + \mathtt{n}_{\mathtt{modes},\mathbb{Q}}$). The nonlinear discrete matrix system is solved following the iterative procedure explained in Algorithm 4.1. For the proposed model order reduction strategy, the computational time is significantly reduced due to two contributions of the reduction: (1) the faster assembly with the pre-computed system matrices in the elastic sub-domain, and (2) the much smaller system solved as a result of projection onto low dimensional RB spaces for all subdomains.

$$
\begin{bmatrix}
\mathbb{W}_{\Omega_1}^{\mathsf{T}}\mathbf{K}_{\mathbf{U}\Omega_1}\mathbb{W}_{\Omega_1} & \mathbb{W}_{\Omega_1}^{\mathsf{T}}\mathbf{K}_{\mathbf{U}\Omega_1,\Gamma_i}\mathbb{W}_{\Gamma_i} & 0 & \mathbb{W}_{\Omega_1}^{\mathsf{T}}\mathbf{C}_{\mathbf{P}\Omega_1}\mathbb{Q} \\
\mathbb{W}_{\Gamma_i}^{\mathsf{T}}\mathbf{K}_{\mathbf{U}\Gamma_i,\Omega_1}\mathbb{W}_{\Omega_1} & \mathbb{W}_{\Gamma_i}^{\mathsf{T}}\mathbf{K}_{\mathbf{U}\Gamma_i}\mathbb{W}_{\Gamma_i} & \mathbb{W}_{\Gamma_i}^{\mathsf{T}}\mathbf{K}_{\mathbf{U}\Gamma_i,\Omega_2}\mathbb{W}_{\Omega_2} & \mathbb{W}_{\Gamma_i}^{\mathsf{T}}\mathbf{C}_{\mathbf{P}\Gamma_i}\mathbb{Q} \\
0 & \mathbb{W}_{\Omega_2}^{\mathsf{T}}\mathbf{K}_{\mathbf{U}\Omega_2,\Gamma_i}\mathbb{W}_{\Gamma_i} & \mathbb{W}_{\Omega_2}^{\mathsf{T}}\mathbf{K}_{\mathbf{U}\Omega_2}\mathbb{W}_{\Omega_2} & \mathbb{W}_{\Omega_2}^{\mathsf{T}}\mathbf{C}_{\mathbf{P}\Omega_2}\mathbb{Q} \\
0 & 0 & 0 & \mathbb{Q}\mathbf{K}_{\mathbf{P}}\mathbb{Q}
\end{bmatrix}
\begin{bmatrix}
\boldsymbol{U}_{\Omega_1 N} \\
\boldsymbol{U}_{\Gamma_i N} \\
\boldsymbol{U}_{\Omega_2 N} \\
\boldsymbol{p}_N
\end{bmatrix}
+
\begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\mathbb{Q}^{\mathsf{T}}\mathbf{C}_{\mathbf{P}\Omega_1}^{\mathsf{T}}\mathbb{W}_{\Omega_1} & \mathbb{Q}^{\mathsf{T}}\mathbf{C}_{\mathbf{P}\Gamma_i}^{\mathsf{T}}\mathbb{W}_{\Gamma_i} & \mathbb{Q}^{\mathsf{T}}\mathbf{C}_{\mathbf{P}\Omega_2}^{\mathsf{T}}\mathbb{W}_{\Omega_2} & \mathbb{Q}^{\mathsf{T}}\mathbf{C}_{\mathbf{M}}\mathbb{Q}
\end{bmatrix}
\begin{bmatrix}
\dot{\boldsymbol{U}}_{\Omega_1 N} \\
\dot{\boldsymbol{U}}_{\Gamma_i N} \\
\dot{\boldsymbol{U}}_{\Omega_2 N} \\
\dot{\boldsymbol{p}}_N
\end{bmatrix}
=
\begin{bmatrix}
\mathbb{W}_{\Omega_1}^{\mathsf{T}}\bar{\mathbf{F}}_{\mathbf{u}\Omega_1} \\
\mathbb{W}_{\Gamma_i}^{\mathsf{T}}\bar{\mathbf{F}}_{\mathbf{u}\Gamma_i} \\
\mathbb{W}_{\Omega_2}^{\mathsf{T}}\bar{\mathbf{F}}_{\mathbf{u}\Omega_2} \\
\mathbb{Q}^{\mathsf{T}}\bar{\mathbf{F}}_{\mathbf{p}}
\end{bmatrix}
\tag{4.28}
$$

4 Numerical Examples

4.1 Boom Clay In-situ Heating Test

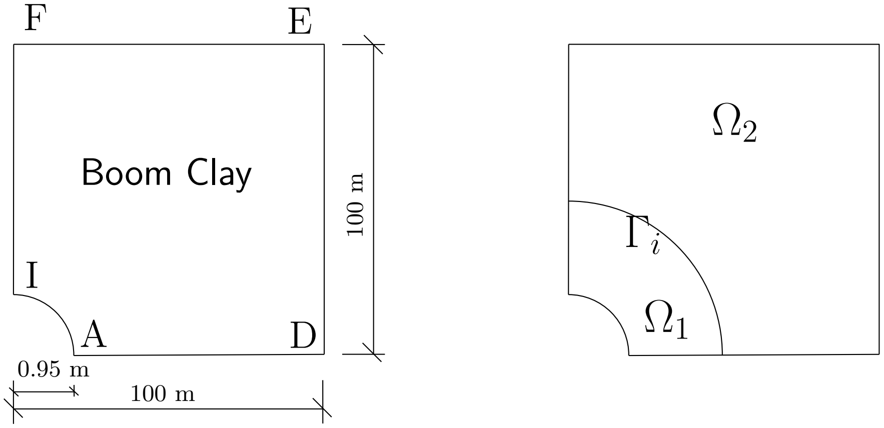


Fig. 4.3 2D plane-strain model domain and decomposed sub-domains

We now turn to a particular numerical example on coupled THM processes in a simplified in-situ heating test model relevant to nuclear waste repository applications. A 2D plane-strain model characterizing the properties of Boom clay material is shown in Figure 4.3. The coupled processes in the domain $\Omega \equiv \{[0, 100] \times [0, 100]\} \backslash \{(\mathsf{r} = 0.95)\}$ are described by the underlying partial differential equations in (4.1–4.4). By assuming uni-directional coupling, the thermal process is modelled independently from the hydro-mechanical process by prescribing uniformly an increasing temperature from 16°C to 120°C on boundary IA over the course of 300 days. The hydro-mechanical process is started from an initial 2.25 MPa pore pressure and an initial 4.5MPa triaxial stress conditions. It is subsequently driven by the temperature change resulting from the thermal process. Fixed normal displacements and impervious conditions are assumed on boundaries IF, AD and IA while the outer boundaries ED and EF are prescribed with the corresponding initial hydro-mechanical conditions. The physical values of the thermo-hydro-mechanical parameters used in the reference model are given in Table 4.2. It must be noted that the values of some material parameters may change in the subsequent sections, specifically when they are selected as varying properties in the parametrized problem.

As explained in the previous section, the computational gain from the model order reduction technique presented here heavily relies on the extent of the plastic region. In this context, the number of elements visited during evaluation of elasto-plastic constitutive law (in Ω_1) with respect to the total number of elements in the entire domain (Ω) plays an important role in achieving a significant computational speed up in the reduced order model. To further illustrate this, we consider two different

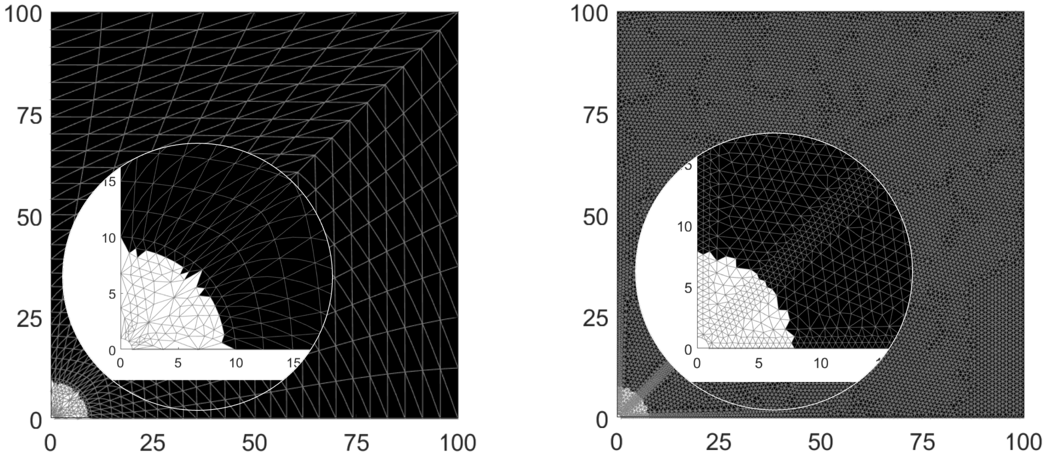
Table 4.2 Material parameters describing the Boom Clay

E	1700 MPa	$c_{p,s}$	740 J/ (kg $\cdot$ K)
k	2×10^{-19} m^2	k_c	1.65 W/ (m $\cdot$ K)
v	0.15	c	0.3 MPa
α	1	φ	10°
M	5698 MPa	ψ	0°
ϕ	0.39	H	0.30 MPa
α_s	1×10^{-5} °C^{-1}	n_x	0.80

discretizations in the domain. Figure 4.4a depicts a coarse discretized mesh with finer concentration of elements closer to IA where the input heat originates. Figure 4.4b on the other hand depicts the domain discretized finely with a more uniform mesh distribution throughout the entire domain. The concentrated coarse mesh consisting of 994 elements results to a high plastic element ratio $\Omega_1 : \Omega \cong 1 : 5$ as compared to $1 : 75$ for the uniform fine mesh consisting of 23 672 elements. The region defining the plastic sub-domain (Ω_1) is determined by running a finite element (FE) model simulation for the fixed values of the parameters given in Table 4.2 and observing at which integration points and elements plasticity occurred. In both discretizations, the plastic region locally occurred within 8m to 10m radially from the origin.

The decomposition-reduction-assembly procedure is followed to implement the model reduction and sub-structuring strategy. For the coarse mesh, this involved decomposing the displacement field degrees of freedom (DoF) into three parts with ($\mathtt{n}_{\mathtt{DoF},\Omega_1} = 738, \mathtt{n}_{\mathtt{DoF},\Gamma_i} = 32, \mathtt{n}_{\mathtt{DoF},\Omega_2} = 3200$). For the fine mesh, it corresponds to a decomposition of the displacement field to ($\mathtt{n}_{\mathtt{DoF},\Omega_1} = 1260, \mathtt{n}_{\mathtt{DoF},\Gamma_i} = 82, \mathtt{n}_{\mathtt{DoF},\Omega_2} = 94260$). In the reduction phase, we considered constructing the reduced basis (Algorithm 4.2) using only one snapshot solution $[\mathbf{u}, p] = \mathsf{SolveFE}(\boldsymbol{\mu}_{ref})$ at this stage. The mode numbers in the RB subspace are ($\mathtt{n}_{\mathtt{modes},\mathbb{W}_{\Omega_1}} = 14, \mathtt{n}_{\mathtt{modes},\mathbb{W}_{\Gamma_i}} = 13, \mathtt{n}_{\mathtt{modes},\mathbb{W}_{\Omega_2}} = 10$, $\mathtt{n}_{\mathtt{modes},\mathbb{Q}} = 17$) for the coarse mesh and ($\mathtt{n}_{\mathtt{modes},\mathbb{W}_{\Omega_1}} = 23, \mathtt{n}_{\mathtt{modes},\mathbb{W}_{\Gamma_i}} = 20, \mathtt{n}_{\mathtt{modes},\mathbb{W}_{\Omega_2}} = 20$, $\mathtt{n}_{\mathtt{modes},\mathbb{Q}} = 24$) for the fine mesh. Once the RB spaces are constructed, the sub-structured system matrices are projected to their appropriate bases. These matrices are then assembled similar to the coupled THM FE system, only differing by the significantly smaller size of the projected matrices. Such DoF reduction, in combination with the optimized offline-online stage calculation that isolates the expensive procedure to the plastic sub-domain (Ω_1), yield significant computational savings. Figure 4.5 and Figure 4.6 illustrate the pore pressure distribution and deformation (horizontal and vertical displacements) in the vicinity of the heat source (boundary IA) after 300 days, along with the error map with respect to the RB-approximated solution. It must be noted that the figures are zoomed up to the region $\Omega \equiv \{[0, 20] \times [0, 20]\}$ where THM responses are prominent, instead of illustrating the full domain $\Omega \equiv \{[0, 100] \times [0, 100]\}$ as THM responses are very small farther from the heat source at IA .

To further investigate the computational gain of the proposed strategy, it is compared with the traditional model reduction approach involving Galerkin projection but without sub-structuring of the domain. In this case, system matrices are updated and assembled at every iteration in solving the nonlinear system and the only gain stems from the reduced size of the system to be solved at each iteration. The RB is constructed by subjecting the snapshot solution $(\mathbf{u}, p)$ to POD directly, which resulted to 31 modes for the coarse mesh and 46 modes for the fine mesh in total. It can be observed that the computation time compared to the FE model simulation was not reduced significantly using the traditional RB-projection approach with only 1.2x faster computation for the coarse and fine mesh. With the proposed model reduction and sub-structuring strategy, a speed up of 3.7x is observed in the coarse mesh, while a significantly faster 20x speed up was achieved in the fine mesh. There was no trade off in the accuracy of the approximation as evident in the similar magnitudes of error between the traditional RB and the proposed strategy as presented in Table 4.3.

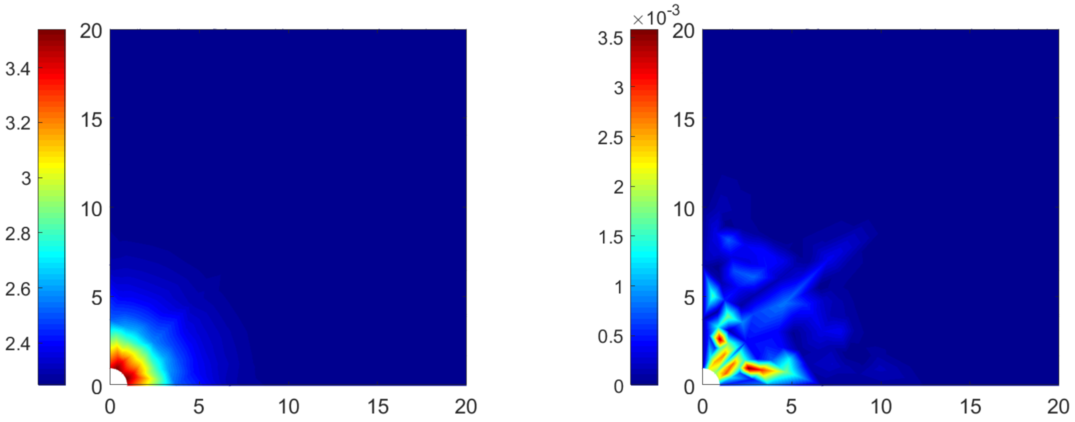


(a) Coarse mesh **(b)** Fine mesh

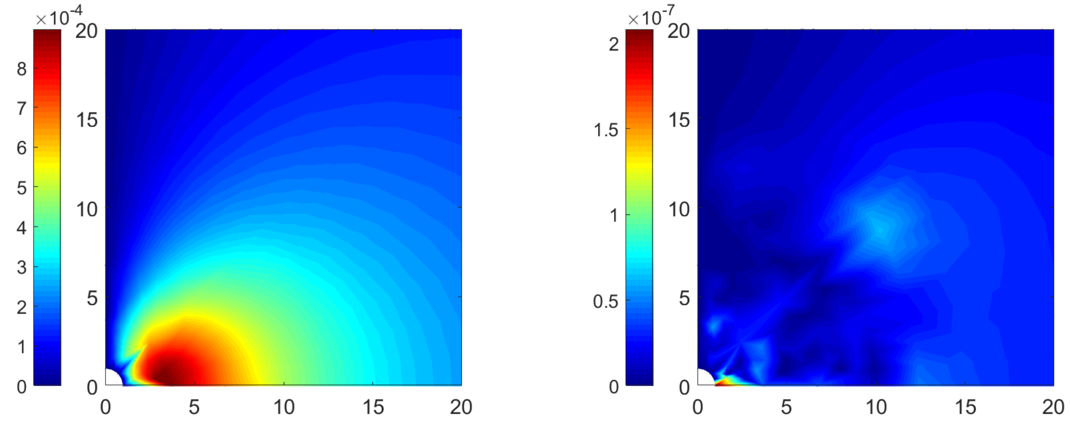
Fig. 4.4 Domain discretization and plastic zone at time = 300 days, corresponding to material parameters in Table 4.2

4.2 Model Order Reduction of a Parametrized System

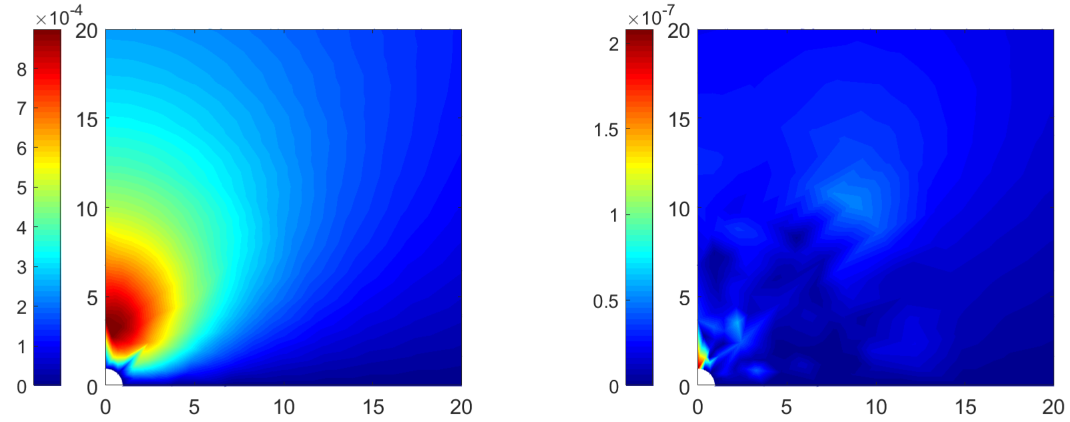
The following section deals with the same coupled THM processes in the in-situ heating test model problem and we proceed using the fine mesh discretization to further illustrate the advantages of the proposed model reduction strategy in a parametric problem setting. We aim to construct a reduced basis that can approximate solutions of the THM system at any point within a pre-defined space under parametric variation. The material properties relating to the mechanical (elastic



(a) FE Pore Pressure (MPa) **(b)** |RB-projected – FE Pore Pressure| (MPa)

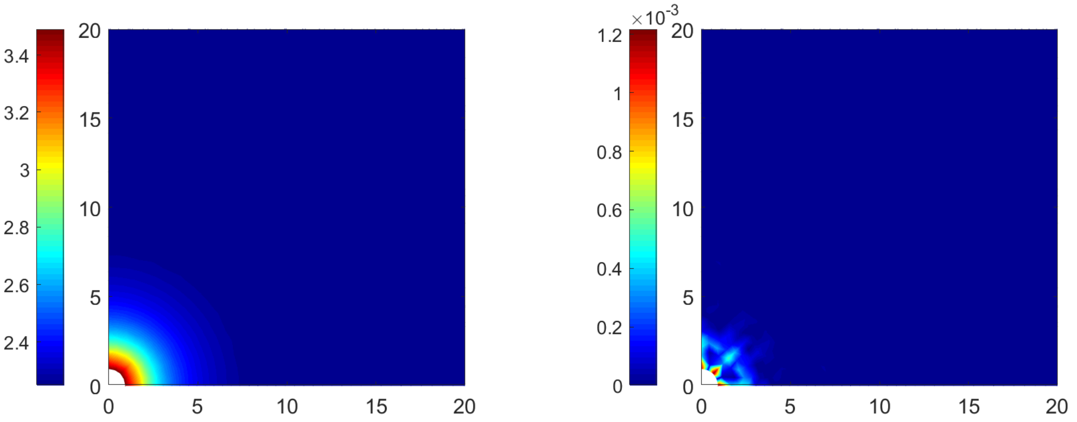


(c) FE Horizontal Displacement (m) **(d)** |RB-projected – FE Hor. Displacement| (m)

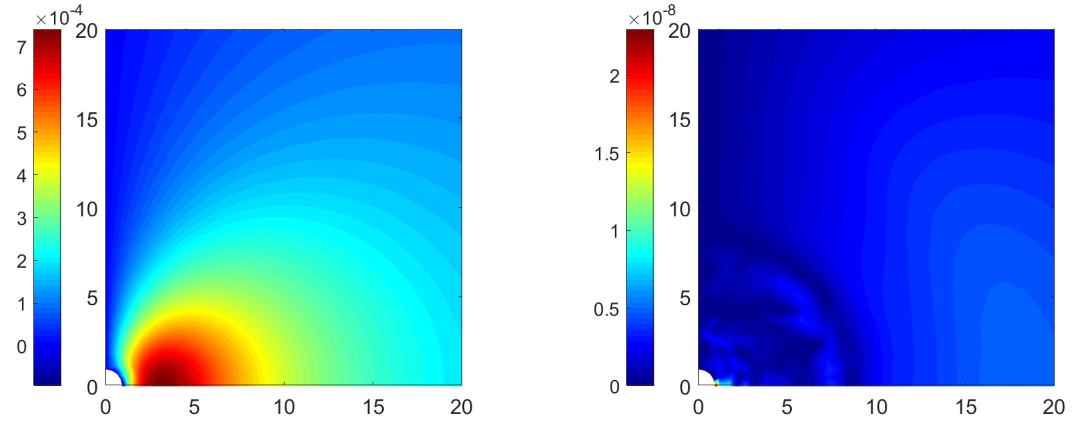


(e) FE Vertical Displacement (m) **(f)** |RB-projected – FE Ver. Displacement| (m)

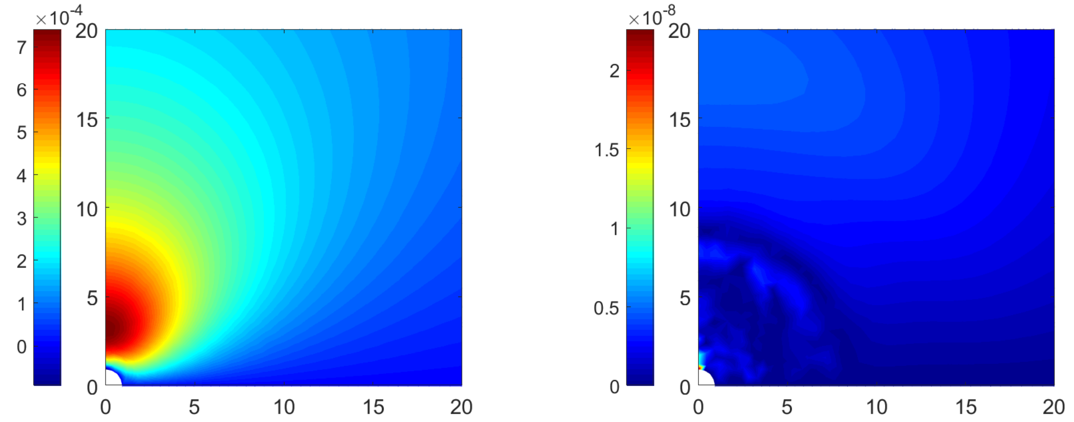
Fig. 4.5 Hydro-mechanical solution at t=300 days, coarse mesh



(a) FE Pore Pressure (MPa) **(b)** |RB-projected – FE Pore Pressure| (MPa)



(c) FE Horizontal Displacement (m) **(d)** |RB-projected – FE Hor. Displacement| (m)



(e) FE Vertical Displacement (m) **(f)** |RB-projected – FE Ver. Displacement| (m)

Fig. 4.6 Hydro-mechanical solution at t=300 days, fine mesh

Table 4.3 RB-approximation of the reference in-situ heating test model

	RB-Projection Coarse Mesh	RB-Projection Fine Mesh	Sub-structured RB-projection Coarse Mesh	Sub-structured RB-projection Fine Mesh
Relative error, $\mathbf{u}$	$2. \times 10^{-4}$	3.9×10^{-6}	1.9×10^{-4}	1.6×10^{-5}
Relative error, p	8.4×10^{-4}	1.2×10^{-5}	8.4×10^{-4}	1.3×10^{-5}
DoF reduction	1:145	1:2315	1:85	1:1225
RB Run time	60 s	2310 s	19 s	137 s
FE Run Time	70 s	2835 s	70 s	2835 s

modulus, E) and hydraulic system (permeability, k) are now considered as varying parameters. The training sample is defined by the space depicted in Figure 4.7 which contains parametric combinations in the sets $E = [300, 500, 700, 900, 1100]$ MPa and $k = [2 \times 10^{-19}, 3.5 \times 10^{-19}, 5 \times 10^{-19}, 7 \times 10^{-19}]$ m^2. In total, the snapshots collected are from 20 different parametric combinations. Subjecting the snapshot solutions to the procedure in Algorithm 4.2 leads to 1:410 DoF reduction containing ($\mathtt{n}_{\mathtt{modes},\mathbb{W}_{\Omega_1}} = 128, \mathtt{n}_{\mathtt{modes},\mathbb{W}_{\Gamma_i}} = 27, \mathtt{n}_{\mathtt{modes},\mathbb{W}_{\Omega_2}} = 23, \mathtt{n}_{\mathtt{modes},\mathbb{Q}} = 96$) in the sub-structured domains. The same sub-structured domains illustrated in Figure 4.4b are pre-selected for the parametric problem. It must be noted that for all parametric combinations within the training space, it has to be ensured that the plastic sub-domain will always be inside the pre-selected Ω_1. For this reason, the elastic modulus in the training sample is selected to be much lower than the reference material property value in Table 4.2. If this condition is violated, the reduced model can not fully capture the plastic behaviour in the appropriate regions and could result in convergence issues.

In Table 4.4, the performance of the RB subspace constructed from the proposed model reduction and sub-structuring strategy is reported. The RB subspace is tested at selected parameter values used to generate the snapshots and at random non-snapshot parameter values to check its ability to approximate any solution within the defined parametric space. All tested samples resulted to low approximation error below the order of 10^{-5} (Table 4.4). The simulation time varied depending on the tested parameter values which reflects the extent of the plastic region where the elasto-plastic constitutive law is evaluated in each case. For instance, a low elastic modulus generally guarantees a smaller plastic region and faster simulation as it entails less time spent on performing the local-iterative 'return mapping' procedure. Overall, the constructed RB for the coupled THM system has shown accurate RB-approximation within the defined parametric space at significantly reduced costs, ranging from 15x to 18x faster than standard FE model simulations. In the next section, the effectiveness of this RB is further put into test by using it to solve a multiple-query type of problem.

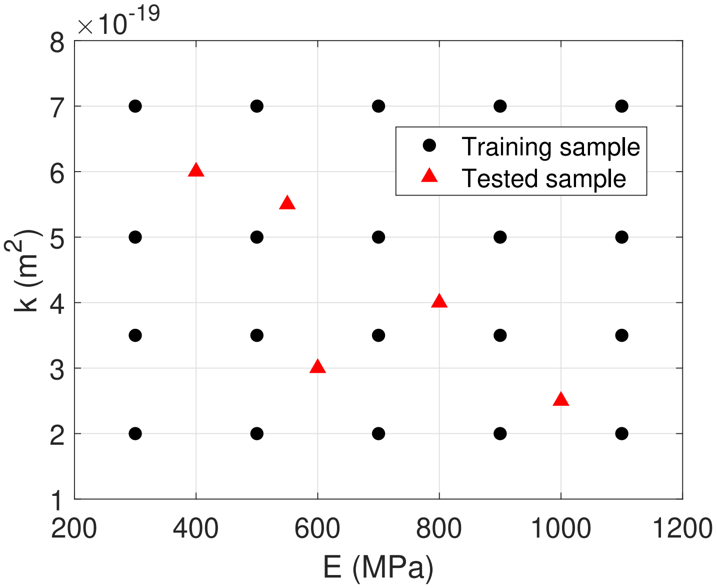


Fig. 4.7 Reduced basis parametric space and training samples

4.3 Application to an Inverse Parameter Identification Problem

To have a more realistic view of the computational gain that can be achieved from the proposed model reduction strategy, we set up a multiple query problem in the form of inverse parameter identification. In this particular problem, we are interested in identifying two independent parameters relating to material properties of the model, elastic modulus (E) and permeability (k). These are the same parameters selected to vary when constructing the reduced basis in the previous section. To carry out the inverse parameter identification, an optimization algorithm minimizing a suitable objective function is implemented. We define the objective function to describe the relative difference between measured and simulated pore pressure at specific measurement locations,

$$\beta = \frac{1}{\mathtt{n}_{\mathtt{sensor}}} \sum_{i=1}^{\mathtt{n}_{\mathtt{sensor}}} \int_0^{t_f} \frac{\| p_m - p_s \|_2}{\| p_s \|_2} \, dt \tag{4.29}$$

where p_m corresponds to the observed measurement data, p_s is the simulated data coming from the numerical model and $\mathtt{n}_{\mathtt{sensor}}$ is the total number of points considered as measurement locations in the domain. The measurement data is synthetically generated by solving a standard FE model evaluated at the parameter values to be identified. The measurement locations are selected to be within the sub-domain where plasticity is expected to occur (Figure 4.8). For the sake of simplicity, the sensor locations are randomly selected and should not necessarily be expected to represent

Table 4.4 Accuracy of the reduced model with respect to varying parameter values within the sampling space

Elastic Modulus, E	Permeability, k	Relative error, $\mathbf{u}$	Relative error, p	Run time
300 MPa	5.0×10^{-19} m^2	2.9×10^{-7}	2.2×10^{-8}	118 s
500 MPa	3.5×10^{-19} m^2	1.2×10^{-7}	1.0×10^{-8}	122 s
700 MPa	3.5×10^{-19} m^2	5.5×10^{-8}	1.2×10^{-8}	124 s
900 MPa	7.0×10^{-19} m^2	1.0×10^{-7}	1.1×10^{-8}	135 s
1100 MPa	2.0×10^{-19} m^2	5.1×10^{-8}	1.3×10^{-8}	137 s
400 MPa	6.0×10^{-19} m^2	1.7×10^{-7}	7.1×10^{-8}	122 s
600 MPa	3.0×10^{-19} m^2	5.4×10^{-6}	5.7×10^{-7}	128 s
800 MPa	4.0×10^{-19} m^2	1.7×10^{-5}	1.2×10^{-6}	128 s
1000 MPa	2.5×10^{-19} m^2	1.3×10^{-5}	1.6×10^{-6}	130 s
550 MPa	5.5×10^{-19} m^2	3.8×10^{-6}	7.7×10^{-7}	120 s

real experiments. The main focus in this numerical example is to demonstrate the ability of the reduced model to adequately substitute the FE model in general inverse problems. Having only two unknown parameters to identify and ample synthetic data to characterize the inverse problem, a local search method, Nelder-Mead simplex algorithm, is found to be appropriate to perform an unconstrained optimization procedure.

The unknown parameters to be targeted are selected as $\boldsymbol{\mu}_{\mathsf{target}} : [E = 800 \text{ MPa}, k = 4{\times}10^{-19} \text{ m}^2]$. The synthetic measurement data include the pore pressure evolution $p(\boldsymbol{\mu}_{\mathsf{target}}; x_{\Omega 1}, t)$ from day 0 to 300 in the plastic sub-domain. When the inverse problem is solved using the standard FE model simulation tool, $p_s = \mathsf{SolveFE}(\boldsymbol{\mu})$, the identified parameters converged to the target values $\mu : [800 \text{ MPa}, 4.0 \times 10^{-19} \text{ m}^2]$ at $\beta = 2.6 \times 10^{-13}$ after 83 iterations with 156 functions evaluated. It took 60 hours to complete the optimization procedure (Figure 4.9a). In comparison, the RB-approximation driven inverse problem was also successful in identifying the unknown parameters converging at $\boldsymbol{\mu} : [799.9998 \text{ MPa}, 3.9997 \times 10^{-19} \text{ m}^2]$, $\beta = 4.3 \times 10^{-6}$ after 77 iterations and 151 functions evaluated (Figure 4.9b). The run time is 4.8 hours which is significantly shorter than the FE-driven optimization. A small shift in the minimum is expected for the RB-approximation driven optimization since the synthetically generated measurement data is obtained from the FE model which does not yield exactly the same solution as that obtained from the surrogate reduced order model. It is observed that for this particular inverse identification problem, 12x faster computation is achieved by using the surrogate model built from the proposed model reduction and sub-structuring strategy.

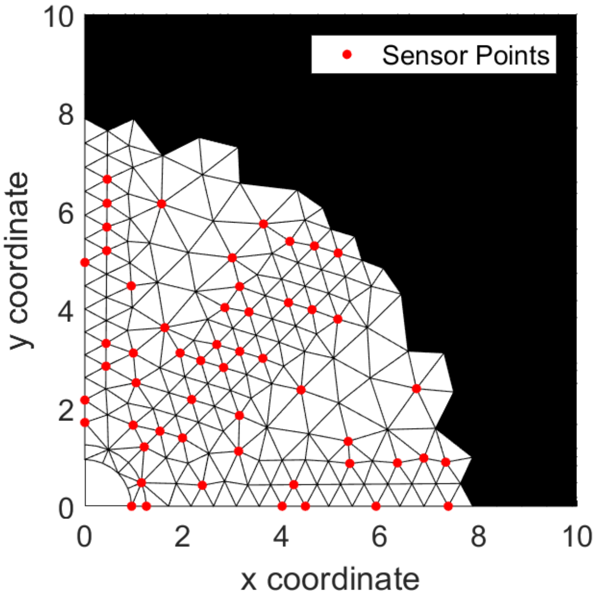
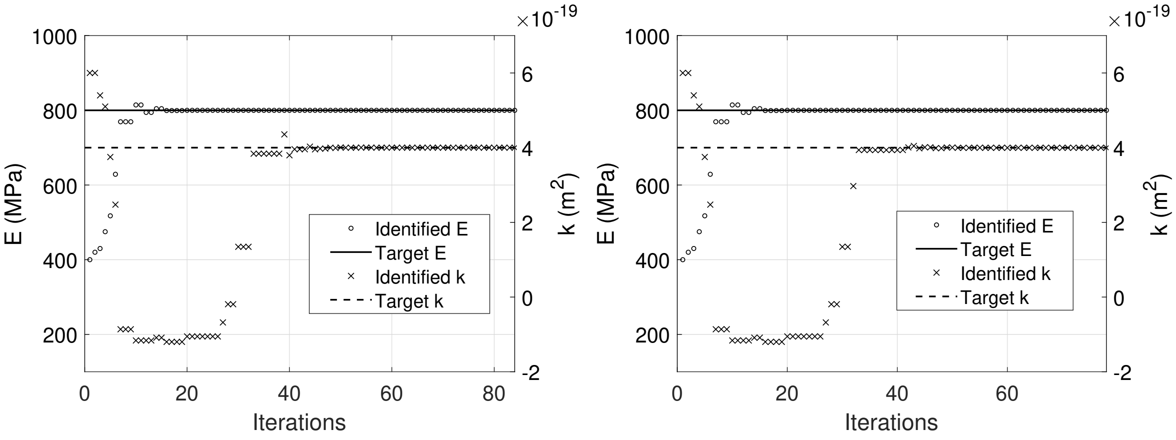


Fig. 4.8 Location of sensor points considered in the manufactured measurement data



(a) FE model driven optimization **(b)** Reduced model driven optimization

Fig. 4.9 Parameter convergence plots (E and k) using Nelder-Mead simplex algorithm

5 Summary and Findings

In this work, we presented a model order reduction strategy that combines the traditional reduced basis projection and sub-structuring by domain decomposition to tackle nonlinear elasto-plasticity in coupled THM systems for possible nuclear waste repository applications. The proposed strategy takes advantage of the occurrence of plasticity in restricted zones of coupled THM models, which allows defining separated elastic and plastic sub-domains easily. This method offers simplicity in implementation where expensive computational efforts (e.g. nonlinear iterative procedure) during the online stage are only evaluated in the plastic sub-domain which constitutes a small region in the entire domain. The main limitation, on the other hand, relates to the efficiency level of the model reduction depending on how small or large the plastic sub-domain is. For scenarios in which the entire domain is in the plastic state, no additional speed up can be expected more than the traditional fully projected RB-approximation method. It would be recommended to explore other techniques that treat the model reduction globally in the entire domain such as hyper-reduction, if the plastic region is not restricted in small zones over the observed period of interest.

For an in-situ heating test, the plastic region is typically expected to cover only a small region in the vicinity of the heat source. Such problem set-up provided a good demonstration of the computational gain that can be achieved in the proposed model reduction strategy. The accuracy of the reduced model is shown to be in line with the traditional reduced basis projection technique while obtaining a significantly faster speed up in computation time (up to 20x faster). The effectiveness of the reduced order model is further validated in the context of parametrized systems that culminated in a manufactured inverse problem which successfully identified unknown parameters in the coupled THM system at a much faster rate (12x faster than FE-driven optimization).

The parametric problem presented here is rather idealized, having a pre-defined space that can be uniformly sampled easily. For future work, it is suggested to formulate a smarter snapshot sampling strategy to construct the reduced basis in cases where the parametric space to consider is very large. This issue becomes highly relevant in more realistic parameter identification problems in coupled THM systems where a handful of uncertain parameters are to be identified. Another area worth exploring in future works is the idea of further reducing the computational effort spent in the plastic sub-domain by applying an additional nonlinear model reduction technique to remove the discretization dependency of the elasto-plastic constitutive law.

5

Conclusions and Outlook

1 Summary of Results

This thesis deals with developing a-posteriori model order reduction strategies tailored to parametric THM systems. It is aimed to build accurate ROMs that are extremely fast to solve and use them as surrogates replacing the traditionally expensive FE models in the framework of multi-query problems. Within the scope and objectives defined, the research questions stated in Chapter 1 are sufficiently answered as demonstrated in numerical examples pertaining to deep geological repository applications. A synthesis that highlights the main results are elaborated in the discussion below.

To tackle the coupling aspect of the transient THM system, a multi-subspace projection approach is pursued such that the thermal, hydrologic and mechanical reduced bases are individually constructed from separated snapshot fields (e.g. temperature, pore pressure, displacements). To efficiently construct the reduced basis, an error estimator guiding the greedy adaptive strategy is devised which in turn, provides certification that the constructed bases always approximate solutions at a certain level of accuracy. An a-posteriori goal-oriented error estimator is formulated particularly for the parabolic PDE describing the transient coupled system. The formulation combines implicit and explicit methods – easily adapting to the space time grid by injecting the implicitly solved adjoint to the explicitly evaluated weak residual. Its capability to estimate the true error value is rigorously investigated and shown to be highly effective despite using only one adjoint solution to represent the dual problem in the entire parametric space. In a large-scale model describing a rock mass under glacial advance, up to 500x faster computational speed up is observed.

The model order reduction strategy further solved a higher dimensional parametric system with up to five material parameters variation to support an in-situ heating experiment. There are several uncertain parameters (THM material properties) to be identified based on partial information and in-situe measurements. With real site measurements to match, the numerical model assumptions are defined as close as possible to the observed geological conditions of the underground laboratory. The thermal aspect is decoupled from the THM system to accommodate temperature-dependent hydro-mechanical material properties while preserving linearity in the system. In the context of inverse parameter identification, this entails a two-step identification. Uncertain thermal parameters are first identified by constructing a RB-based surrogate model of the thermal system and implementing optimization algorithms to automatically identify material property values that correspond to the observed measurement data. The identified thermal parameters are then used to build the hydro-mechanical model in which another RB-based surrogate is constructed to efficiently carry out the inverse hydro-mechanical parameter identification. The surrogate model used in the optimization procedure is demonstrated to be very efficient with the ability to successfully identify the material parameters up to 800x faster than the FE-driven procedure. In contrast, the time it would take to solve such multi-query problems is known to be impractical using FE models which could reach up to several hundreds of days.

To include the complexity arising from nonlinear systems, a novel strategy is developed to deal with the elasto-plastic behaviour of geomaterials. The nonlinearity originating from the history-dependent constitutive law requires a two-level iteration procedure which is very costly to implement in FE models. The presented model order reduction strategy alleviates the involved computational burden by sub-structuring the domain and confining the expensive procedures in the plastic region. This strategy does not free the computation fully from online procedures as the system matrices must still be updated online at every iteration in the plastic region. When the reduced model is used as a surrogate in a multi-query problem, up to 12x faster computation time is observed compared to FE-driven simulations.

With the numerical examples depicting different scenarios in deep geological repository, the proposed strategies demonstrated significant advantages to replace the traditionally expensive FE model of coupled THM systems in parametrized problems. The need for reliable error estimation is crucial for snapshot-based model order reduction – particularly, when dealing with high dimensional systems as the parametric space grows exponentially with every additional parameter of interest considered. In linear THM systems, the computational gain primarily comes from the matrix system size reduction which drastically offloads the system solver (i.e. from solving systems of 10^5 DoF to 10^2 DoF). The linearity also permits affinely decomposing the system and therefore allowing seamless implementation of the offline-online stage procedures. In the case of nonlinear systems where affine decomposition is not fully applicable, the computational gain from reduced subspace projection alone becomes insignificant as the cost of generating system matrices at every iteration dominates the simulation. The strategy presented takes advantage of the plasticity occurring in a small confined region which permits implementing linear model order reduction in the elastic region.

Overall, it is observed that the initial cost related to offline stage procedures becomes marginal compared to the computational savings and flexibility that ROM offers. For instance, ROM can be used to accurately substitute FE models in performing inverse analysis, real-time simulations, generating parametric maps and sensitivity analysis at a much faster rate. The strategies presented here are rather intrusive which may not be easily adaptable for all types of problems. However, using intrusive approaches is favourable in preserving the physics presented in the original problem which is critical for model order reduction of coupled multi-physics systems.

2 Future Developments

The work presented in this thesis successfully fulfilled the research objectives within the defined scope of study. In future works, it is recommended to further investigate and explore the following issues and topics that are either not previously attempted or arose from findings of the current research work.

- **Physical laws.** Considering more complex physical laws for numerical modelling of deep geological repository represents an interesting challenge. For instance, material nonlinearity and multi-phase flow of partially saturated soils are vital in particular studies. The effects of chemical processes (e.g. radionuclide transport) can also be significant in characterizing the behaviour of the geological host in specific repository sites.

- **Complex geometrical domain.** In many subsurface applications, large-scale geomodels are characterized by complex networks of fractures that alter groundwater flow, affect mass transport and modify the pore pressure generation of the geological mass. Modelling of fracture zones in the framework of model order reduction could be an interesting topic to explore.

- **Pre-operational phase of deep geological repository.** In this work, the time scales investigated for the coupled THM behaviour is limited to the operational phase of the repository. It is also critical to study the processes during the construction phase of the multi-barrier concrete system. Building a reduced order model that captures the pre-operational phase could be useful in safety assessment of the structure.

- **Snapshot selection with respect to time.** In the methodology presented, each snapshot contains the solution evaluated in the entire time period. Instead of only parametrizing the snapshot with respect to the material parameter, it can also be parametrized with respect to time. This strategy reduces the size of snapshots and possibly the number of POD modes by taking fewer solutions at selected time steps. In effect, an additional optimality selection criterion needs to be formulated for the greedy adaptive procedure.

- **Representative adjoint solution.** With the assumption that a single adjoint solution is sufficient to represent the entire parametric space, it is worth exploring in much higher dimensional parametric space if this still holds. For instance, it might be of interest to evaluate more adjoint solutions and devise a smart selection procedure that creates a rigorous representation of the dual problem in the parametric space.

- **Greedy adaptive procedure and nonlinear elasto-plasticity.** Following the greedy procedure for linear THM systems, it is useful to develop a similar framework for the nonlinear elasto-plastic system in order to optimally select snapshots that construct the reduced basis. It is particularly challenging to devise an efficient error estimator that incorporates the history-dependent constitutive law of the material.

- **Discretization-dependency of elasto-plastic constitutive laws.** One of the hurdles in nonlinear model order reduction which was not addressed in this work is freeing the plastic computation from spatial discretization. The plastic constitutive law is still evaluated at every integration point in the plastic domain at every iteration. Some techniques such as hyper-reduction and data-driven non-intrusive methods can be explored to address this condition.

- **Using ROM in other practical investigations.** In further assessment of potential site of deep geological repository, other types of parametrized problems might be of strong interest. Aside from parameter identification, the reduced model can be used as a surrogate for sensitivity analysis, generating parametric maps and fast direct-query simulations.

Appendices

A POD in Parametrized Systems

Proper orthogonal decomposition (POD) is a data analysis technique that aims at obtaining low-dimensional approximations of high-dimensional processes. Also known as Principal Component Analysis or Karhunen-Loéve Decomposition, the POD method transforms the original variables into a new set of uncorrelated variables (e.g. POD modes or principal components) where most of the energy or dominant model features is retained in the first few modes. Data analysis using POD often extracts basis functions from collected data, either by experiments or high-dimensional system simulations, for a Galerkin projection procedure that yields a representative low-dimensional system.

In the context of model order reduction of parametrized systems, we consider a set of parameter samples $\Xi_{train} = \text{span}\{\mu^1, \mu^2, ..., \mu^{\mathtt{n_{sample}}}\}$ and the corresponding set of snapshots (e.g. solutions of the high-dimensional system). The snapshot matrix $\mathbb{U} \in \mathbb{R}^{H \times \mathtt{n_{sample}}}$ as

$$\mathbb{U} = [u(\mu^1)\, u(\mu^2)\, ...\, u(\mu^{\mathtt{n_{sample}}})] \tag{A.1}$$

From singular value decomposition of the snapshot matrix,

$$\mathbb{U} = \mathbb{X}\Sigma\mathbb{Y}^{\mathsf{T}} \tag{A.2}$$

where $\mathbb{X} = [\zeta_1\ \zeta_2 ...\ \zeta_H] \in \mathbb{R}^{H \times H}$ and $\mathbb{Y} = [\Psi_1\ \Psi_2\ ...\Psi_{\mathtt{n_{sample}}}] \in \mathbb{R}^{\mathtt{n_{sample}} \times \mathtt{n_{sample}}}$ are orthogonal matrices, and $\Sigma = \mathsf{diag}(\sigma_1,\ \sigma_2,\ ...\ , \sigma_r)$, with $r \leq \mathsf{min}(H, \mathtt{n_{sample}})$ denoting the rank of $\mathbb{U}$. Equivalently, the following eigenvalue relationships are written as:

$$\mathbb{U}\Psi_i = \sigma_i\zeta_i \quad \text{and} \quad \mathbb{U}^{\mathsf{T}}\Psi_i = \sigma_i\Psi_i, \quad i = 1, ..., r \tag{A.3}$$

$$\mathbb{U}^\mathsf{T}\mathbb{U}\Psi_i = \sigma_i^2\Psi_i \quad \text{and} \quad \mathbb{U}\mathbb{U}^\mathsf{T}\Psi_i = \sigma_i^2\zeta_i, \quad i = 1, ..., r \tag{A.4}$$

The orthonormal POD basis $\mathbb{V}$ of dimension $\mathtt{n_{modes}}$ defining the first $\mathtt{n_{modes}}$ left singular vectors $[\zeta_1\ \zeta_2\ ...\ \zeta_{\mathtt{n_{modes}}}]$ of $\mathbb{U}$, minimizes the least squares error of snapshot reconstruction

$$\min_{\mathbb{V}\in\mathbb{R}^{H\times \mathtt{n_{modes}}}} \parallel \mathbb{U} - \mathbb{V}\mathbb{V}^\mathsf{T}\mathbb{U} \parallel_F^2 = \min_{\mathbb{V}\in\mathbb{R}^{H\times \mathtt{n_{modes}}}} \sum_{i=1}^{\mathtt{n_{sample}}} \parallel u(\mu^i) - \mathbb{V}\mathbb{V}^\mathsf{T}u(\mu^i) \parallel_2^2 = \sum_{i=r+1}^{\mathtt{n_{sample}}} \sigma_i^2 \tag{A.5}$$

It follows that the error in the POD basis is equal to the sum of the squares of the singular values corresponding to the neglected POD modes. A suitable criterion can be deduced to select the minimal POD dimension $\mathtt{n_{modes}} \leq r$ such that the projection error is smaller than the desired tolerance (A.6).

$$\mathsf{E}(\mathtt{n_{modes}}) = \frac{\sum_{i=1}^{\mathtt{n_{modes}}} \sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2} \geq 1 - \mathsf{tol}^2 \tag{A.6}$$

B Affine Decomposition: Anisotropic Coupled THM System

In the framework of offline–online stage procedures, the reduced basis approximation is implemented starting from the weak formulation (3.8-3.10). For a parametric system that depends on variation of material parameters μ, the bilinear form $B(T, \mathbf{u}, p; z, \mathbf{w}, q)$ $= L(z, \mathbf{w}, q)$ admits an affine expansion. Following the non-monolithic approach, for the cross-anisotropic system with heterogeneous domains describing the steel casing (Ω_1) and Boom Clay (Ω_2) material,

$$B(T, \mathbf{u}, p; z, \mathbf{w}, q) = \sum_{\pi=1}^{\Pi_a} \theta_a^\pi(\mu) a_\pi(T, \mathbf{u}, p; z, \mathbf{w}, q) + \sum_{\pi=1}^{\Pi_m} \theta_m^\pi(\mu) m_\pi(T, \mathbf{u}, p; z, \mathbf{w}, q) \tag{B.1}$$

$$L(z, \mathbf{w}, q) = \sum_{\pi=1}^{\Pi_f} \theta_f^\pi(\mu) f_\pi(z, \mathbf{w}, q) \tag{B.2}$$

where $\theta_a^\pi(\mu)$, $\theta_m^\pi(\mu)$ and $\theta_f^\pi(\mu)$ are the varying parameters of interest, a_π, m_π and f_π are the decomposed parameter-independent terms. The terms m relates to the time dependent terms of the bilinear term in the left hand side of the weak form while f denotes the right hand side term. Taking into account all the possible parameter variation in the defined THM system, $\Pi_a = 11$, $\Pi_m = 2$, $\Pi_f = 3$, ×2 domains (Ω_1, Ω_2). The number of affine terms $(\Pi = \Pi_a + \Pi_m + \Pi_f)$ can be reduced depending on the number of material parameter variation considered in the problem (i.e. a system varying permeability (k) and elastic modulus (E) involves only two affine terms).

For the cross-anisotropic material properties in the mechanical system, we introduce arbitrary variables $(\omega_{hh}, \omega_{hv}, \omega_{vh}, \lambda_h, \lambda_v)$ to express the elasticity tensor in separable form

$$\boldsymbol{C} = \begin{bmatrix} \lambda_h - \omega_{hh} & \omega_{hh} & \omega_{hv} & 0 & 0 & 0 \\ \omega_{hh} & \lambda_h - \omega_{hh} & \omega_{hv} & 0 & 0 & 0 \\ \omega_{hv} & \omega_{hv} & \lambda_v - \omega_{vh} & 0 & 0 & 0 \\ 0 & 0 & 0 & G_{vh} & 0 & 0 \\ 0 & 0 & 0 & 0 & G_{vh} & 0 \\ 0 & 0 & 0 & 0 & 0 & G_{hh} \end{bmatrix} \tag{B.3}$$

where $\omega_{hh} = \frac{E_h v_{hh}}{(1+v_{hh})(1-2v_{hh})}$, $\omega_{hv} = \frac{E_h v_{hv}}{(1+v_{hh})(1-2v_{hh})}$, $\omega_{vh} = \frac{E_v v_{hh}}{(1+v_{hh})(1-2v_{hh})}$, $\lambda_h = \frac{E_h}{(1+v_{hh})(1-2v_{hh})}$ and $\lambda_v = \frac{E_v}{(1+v_{hh})(1-2v_{hh})}$. With these arbitrary variables introduced and shear moduli (G_{vh}, G_{hh}), the elastic tensor can be expressed in 7 components independent with respect to each variables.

$$\boldsymbol{C} = \boldsymbol{C_1} + \boldsymbol{C_2} + \boldsymbol{C_3} + \boldsymbol{C_4} + \boldsymbol{C_5} + \boldsymbol{C_6} + \boldsymbol{C_7} \tag{B.4}$$

$$\boldsymbol{C_1} = \omega_{hh} \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{C_2} = \omega_{hv} \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{C_3} = \omega_{vh} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{C_4} = \lambda_h \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{C_5} = \lambda_v \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{C_6} = G_{vh} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{C_7} = G_{hh} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The parameter-independent terms are then defined as:

$$a_1 = \int_\Omega \frac{\partial \mathbf{w}_i}{\partial x_j} \boldsymbol{C_1} \frac{\partial \mathbf{u}_k}{\partial x_l} d\Omega \tag{B.5}$$

$$a_2 = \int_\Omega \frac{\partial \mathbf{w}_i}{\partial x_j} \boldsymbol{C_2} \frac{\partial \mathbf{u}_k}{\partial x_l} d\Omega \tag{B.6}$$

$$a_3 = \int_\Omega \frac{\partial \mathbf{w}_i}{\partial x_j} \boldsymbol{C_3} \frac{\partial \mathbf{u}_k}{\partial x_l} d\Omega \tag{B.7}$$

$$a_4 = \int_\Omega \frac{\partial \mathbf{w}_i}{\partial x_j} \boldsymbol{C_4} \frac{\partial \mathbf{u}_k}{\partial x_l} d\Omega \tag{B.8}$$

$$a_5 = \int_\Omega \frac{\partial \mathbf{w}_i}{\partial x_j} \boldsymbol{C_5} \frac{\partial \mathbf{u}_k}{\partial x_l} d\Omega \tag{B.9}$$

$$a_6 = \int_\Omega \frac{\partial \mathbf{w}_i}{\partial x_j} \boldsymbol{C_6} \frac{\partial \mathbf{u}_k}{\partial x_l} d\Omega \tag{B.10}$$

$$a_7 = \int_\Omega \frac{\partial \mathbf{w}_i}{\partial x_j} \boldsymbol{C_7} \frac{\partial \mathbf{u}_k}{\partial x_l} d\Omega \tag{B.11}$$

$$a_8 = \int_\Omega \frac{1}{\mu_f(T)} \frac{\partial q}{\partial x_1} \frac{\partial p}{\partial x_1} d\Omega \tag{B.12}$$

$$a_9 = \int_\Omega \frac{1}{\mu_f(T)} \frac{\partial q}{\partial x_2} \frac{\partial p}{\partial x_2} d\Omega \tag{B.13}$$

$$a_{10} = \int_\Omega \frac{1}{\mu_f(T)} \frac{\partial q}{\partial x_3} \frac{\partial p}{\partial x_3} d\Omega \tag{B.14}$$

$$a_{11} = \int_\Omega \mathbf{w}_i \frac{\partial p}{\partial x_i} d\Omega \tag{B.15}$$

$$m_1 = \int_\Omega q \frac{\partial p}{\partial t} d\Omega \tag{B.16}$$

$$m_2 = \int_\Omega q \frac{\partial}{\partial t}(\frac{\partial \mathbf{u}_i}{\partial x_i}) d\Omega \tag{B.17}$$

$$f_1 = \int_\Omega \mathbf{w}_i \boldsymbol{\Theta} \frac{\partial T}{\partial x_i} d\Omega \tag{B.18}$$

$$f_2 = \int_\Omega q \frac{\partial T}{\partial t} d\Omega \tag{B.19}$$

$$f_3 = \int_\Omega 3\alpha_f(T) q \frac{\partial T}{\partial t} d\Omega \tag{B.20}$$

with their corresponding parameters:

$$\theta_a^1(\mu) = \omega_{hh} \tag{B.21}$$

$$\theta_a^2(\mu) = \omega_{hv} \tag{B.22}$$

$$\theta_a^3(\mu) = \omega_{vh} \tag{B.23}$$

$$\theta_a^4(\mu) = \lambda_h \tag{B.24}$$

$$\theta_a^5(\mu) = \lambda_v \tag{B.25}$$

$$\theta_a^6(\mu) = G_{vh} \tag{B.26}$$

$$\theta_a^7(\mu) = G_{hh} \tag{B.27}$$

$$\theta_a^8(\mu) = k_h \tag{B.28}$$

$$\theta_a^9(\mu) = k_h \tag{B.29}$$

$$\theta_a^1 0(\mu) = k_v \tag{B.30}$$

$$\theta_a^1 1(\mu) = \alpha \tag{B.31}$$

$$\theta_m^1(\mu) = \frac{1}{M} \tag{B.32}$$

$$\theta_m^2(\mu) = \alpha \tag{B.33}$$

$$\theta_f^1(\mu) = \Theta \tag{B.34}$$

$$\theta_f^2(\mu) = 3\alpha_s(\alpha - \phi) \tag{B.35}$$

$$\theta_f^3(\mu) = \phi \tag{B.36}$$

Bibliography

[1] FEBEX project: Full-scale engineered barriers experiment for a deep geological repository for high-level radioactive waste in crystalline host rock, Final report. Technical report, ENRESSA, 2010.

[2] Aagaard, B. T., Heaton, T. H., and Hall, J. F. Dynamic earthquake ruptures in the presence of lithostatic normal stresses: implications for friction models and heat production. *Bull. Seismol. Soc. Am*, 91:1765–1796, 2001.

[3] Aagaard, B. T., Knepley, M. G., and Williams, C. A. A domain decomposition approach to implementing fault slip in finite-element models of quasi-static and dynamic crustal deformation. *J. Geophys. Res. Solid Earth*, 118:3059–3079, 2013.

[4] Alonso, E. E., Alcoverro, J., and Coste, F. The FEBEX benchmark test: case definitionand comparison of modeling approaches. *Int J Rock Mech Mining Sci*, 42(5–6):611–638, 2005.

[5] Amsallem, D. and Farhat, C. Stabilization of projection based reduced-order models. *Internat. J. Numer. Methods Eng.*, 91:358–377, 2012.

[6] Andra. Dossier 2005 argile: evaluation of the feasibility of a geological repository in an argillaceous formation. Technical report, National Agency for Radioactive Waste Management, France, 2005.

[7] Armand, G., Bumbieler, F., Conil, N., de la Vaissière, R., Bosgiraud, J. M., and Vu, M. N. Main outcomes from in situ thermo-hydro-mechanical experiments programme to demonstrate feasibility of radioactive high-level waste disposal in the Callovo-Oxfordian claystone. *Journal of Rock Mechanics and Geotechnical Engineering*, 9(3):415–427, 2017.

[8] Babuska, I. The finite element method with lagrangian multipliers. *Numerische Mathematik*, 20(3):179–192, 1973.

[9] Baldi, G., Hueckel, T., and Pellegrini, R. Thermal volume changes of the mineral-water system in low-porosity clay soils. *Can Geotech J* , 25:807–825, 1988.

[10] Baldi, G., Hueckel, T., Peano, A., and Pellegrini, R. Developments in modelling of thermo-hydro-geomechanical behaviour of boom clay and clay-based buffer materials. Technical Report EUR-13365/2, Commission of the European Community, 1991.

[11] Barrault, M., Maday, Y., Nguyen, N. C., and Patera, A. T. An 'empirical interpolation' method: Application to efficient reduced-basis discretization of partial differential equations. *C. R. Math.*, 339(9):667–672, 2004.

[12] Bathe, K. J. The inf-sup condition and its evaluation for mixed finite element methods. *Computers and Structures*, 79(2):243–252, 2001.

[13] Berckmans, A., Boulanger, D., Brassinnes, S., Capouet, M., Depaus, C., and Lopez, E. D. ONDRAF/NIRAS research development and demonstration plan–for the geological disposal of high-level and/or long-lived radioactive waste including irradiated fuel if considered as waste: State-of-the-art report as of December 2012. Technical Report NIROND-TR 2013-12 E, Belgian Agency for Radioactive Waste and Enriched Fissile Materials, 2013.

[14] Berger, J., Helcio, R., Orlande, B., and Mendes, N. Proper Generalized Decomposition model reduction in the Bayesian framework for solving inverse heat transfer problems. *Inverse Problems in Science and Engineering*, 25(2): 260–278, 2017.

[15] Berkooz, G., Holmes, P., and Lumle, J. L. The proper orthogonal decomposition in the analysis of turbulent flows. *Annu. Rev. Fluid Mech*, 25(1):539–575, 1993.

[16] Bermejo, R. and Carpio, J. A space-time adaptive finite element algorithm based on dual weighted residual methodology for parabolic equations. *SIAM J Sci Comput*, 31:3324–3355, 2009.

[17] Bernier, F., Li, X. L., and Bastiaens, W. Twenty-five years geotechnical observation and testing in the tertiary boom clay formation. *Geotechnique*, 57 (2):229–237, 2007.

[18] Blessent, D., Therrien, R., and Gable, C. Large-scale numerical simulation of groundwater flow and solute transport in discretely-fractured crystalline bedrock. *Advances in Water Resources*, 34(12):1539–1552, 2011.

[19] Bossart, P. and Thury, M. Mont terri rock laboratory – project, programme 1996 to 2007 and results. Technical Report 3, Reports of the Swiss Geological Survey, 2008.

[20] Boyaval, S., Bris, C. L., Maday, Y., Nguyen, C. N., and Patera, A. T. A reduced basis approach for variational problems with stochastic parameters: Application to heat conduction with variable robin coefficient. *Comput. Methods Appl. Mech. Eng.*, 198(41–44):3187–3206, 2009.

[21] Brezzi, F., Douglas, J., and Marini, L. D. Two families of mixed finite elements for second order elliptic problems. *Numerische Mathematik*, 47:217–235, 1985.

[22] Carey, J., Harp, D., Lewis, K., Chu, S., and Kelkar, S. A Reduced-Order Model for Wellbore Permeability Induced by Geomechanical Damage. *NRAP-TRS-III-006- 2017; NRAP Technical Report Series; U.S. Department of Energy, National Energy Technology Laboratory*, 2017.

[23] Carlberg, K., Farhat, C., Cortial, J., and Amsallem, D. The GNAT method for nonlinear model reduction: Effective implementation and application to computational fluid dynamics and turbulent flows. *J. Comput. Phys*, 242: 623–647, 2013.

[24] Castro, J. and Sivasithamparam, N. A constitutive model for soft clays incorporating elastic and plastic cross-anisotropy. *Materials*, 10(6), 2017.

[25] Chakir, R. and Hammond, J. K. A non-intrusive reduced basis method for elastoplasticity problems in geotechnics . *Journal of Computational and Applied Mathematics,Elsevier*, 337:1–17, 2018.

[26] Chandler, N. A., Wan, A. W. L., and Roach, P. J. The buffer/container experiment design and construction report. Technical Report AECL-11792, COG-97-186-I, AECL, 1998.

[27] Chatterjee, A. An introduction to the proper orthogonal decomposition. *Curr Sci* , 78(7):808–817, 2000.

[28] Chaturantabut, S. and Sorensen, D. C. Nonlinear model reduction via Discrete Empirical Interpolation. *SIAM J. Sci. Comput*, 32:2737–2764, 2010.

[29] Chen, G. J., Sillen, X., Verstricht, J., and Li, X. L. ATLAS III in situ heating test in boom clay: Field data, observation and interpretation. *Computers and Geotechnics*, 38(5):683–696, 2011.

[30] Chen, G. J., Dizier, A., Li, X. L., Verstricht, J., Sillen, X., and Levasseur, S. Numerical prediction of the large-scale in situ PRACLAY Heater test in the Boom Clay. *Rock Mechanics and Rock Engineering.*, 54(5):2197–2218, 2021.

[31] Chevreuil, M. and Nouy, A. Model order reduction based on proper generalized decomposition for the propagation of uncertainties in structural dynamics. *Int. J. Numer. Methods Eng.*, 89(2):241–268, 2012.

[32] Chijimatsu, M., Fujita, T., Sugita, Y., Amemiya, K., and Kobayashi, A. Field experiment, results and thm behavior in the kamaishi mine experiment. *International Journal of Rock Mechanics and Mining Science*, 38(1):67–78, 2001.

[33] Chinesta, F., Ladeveze, P., and Cueto, E. A short review on model order reduction based on proper generalized decomposition. *Arch. Comput. Methods Eng.*, 18:395–404, 2011.

[34] Collin, F., Li, X. L., Radu, J. P., and Charlier, R. Thermo-hydro-mechanical coupling in clay barriers. *Engineering Geology*, 64(2–3):179–193, 2002.

[35] Corigliano, A., Dossi, M., and Mariani, S. Model order reduction and domain decomposition strategies for the solution of the dynamic elastic–plastic structural problem. *Comput. Methods Appl. Mech. Eng.*, 290:127–155, 2015.

[36] Craig, R. and Bampton, M. C. C. Coupling of substructures for dynamic analysis. *American Institute of Aeronautics and Astronautics Journal*, pages 1313–1319, 1968.

[37] De-Bruyn, D. and Labat, S. The second phase of ATLAS: the continuation of a running THM test in the HADES underground research facility at Mol. *Engineering Geology*, 64(2–3):309–316, 2002.

[38] Delage, P., Sultan, and Cui, Y. J. On the thermal consolidation of Boom Clay. *Can Geotech J*, 37:343–354, 2000.

[39] Deng, Y. F., Tang, A. M., Cui, Y. J., and Li, X. L. Study on the hydraulic conductivity of boom clay. *Canadian Geotechnical Journal*, 48(10):1461–1470, 2011.

[40] Deng, Y. F., Tang, A. M., Cui, Y. J., Nguyen, X. P., Li, X. L., and Wouters, L. Laboratory Hydromechanical Characterisation of Boom Clay at Essen and Mol. *Physics and Chemistry of the Earth, Parts A,B,C, Clays in Natural and Engineered Barriers for Radioactive Waste Confinement*, 36(17–18):1878–1890, 2011.

[41] Deng, Y. F., Cui, Y. J., Tang, A. M., Li, X. L., and Sillen, X. An experimental study on the secondary deformation of boom clay. *Applied Clay Science*, 59–60: 19–25, 2012.

[42] De Lathauwer, L., Moor, B. D., and Vandewalle, J. A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl*, 21(4):1253–1278, 2000.

[43] Dickson, M. and Fanelli, M. *Geothermal energy*, volume 61 of *Lecture Notes in Computational Science and Engineering*. Springer, 2008.

[44] Diez, P. and Calderon, G. Goal-oriented error estimation for transient parabolic problems. *Computational Mechanics*, 39:631–646, 2007.

[45] Dizier, A., Chen, G. J., Verstricht, J., Li, X. L., Sillen, X., and Levasseur, S. The large-scale in situ PRACLAY heater test: main observations on the thermo-hydro-mechanical behaviour of the Boom Clay. *Intern J Rock Mech Mining Sci*, 137:3619–3631, 2021.

[46] Drucker, D. C. and Prager, W. Soil mechanics and plasticity analysis of limit design. *Quarterly Applied Mathematics*, 10(2):157–165, 1952.

[47] Duffield, W. and Sass, J. Geothermal energy-clean power from the earth's heat. Technical Report C1249, US Geological Survey, 2003.

[48] Eldred, M., Giunta, S., and Collis, S. Second-order corrections for surrogate-based optimization with model hierarchies. *In Proceedings of the 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, 2004.

[49] Everson, R. and Sirovich, L. Karhunen-Loeve procedure for gappy data. *EJ. Opt. Soc. Am. A*, 12:1657–1664, 1995.

[50] Florez, H. Applications of model-order reduction to thermo-poroelasticity. In *Proceedings of the 51st US Rock Mechanics / Geomechanics Symposium*, 2014.

[51] Francois, B., Laloui, L., and Laurent, C. Thermo-hydromechanical simulation of ATLAS in situ large scale test in boom clay. *Computers and Geotechnics*, 36 (4):626–640, 2009.

[52] Frangos, M., Marzouk, Y., Willcox, K., and van Bloemen Waanders, B. *Surrogate and Reduced-Order Modeling: A Comparison of Approaches for Large-Scale Statistical Inverse Problems.* John Wiley and Sons, Ltd, 2010.

[53] Fuentes, D., Littlefield, D., Oden, J. T., and Prudhomme, S. Extensions of goal-oriented error estimation methods to simulations of highly-nonlinear response of shock-loaded elastomer-reinforced structures. *Comput. Methods Appl. Mech. Eng.*, 195:4659–4680, 2006.

[54] Garitte, B., Gens, A., Vaunat, J., and Armand, G. Thermal conductivity of argillaceous rocks: determination methodology using in situ heating tests. *Rock Mechanics and Rock Engineering*, 47(1):111–129, 2014.

[55] Gens, A. Soil–environment interaction in geotechnical engineering. *Géotechnique*, 60(1):3–74, 2010.

[56] Gens, A., Garcia-Molina, A. J., Olivella, S., Alonso, E. E., and Huertas, F. Analysis of a full scale in situ test simulating repository conditions. *International Journal for Numerical and Analytical methods in Geomechanics*, 22:515–548, 1998.

[57] Ghabezloo, S. and Sulem, J. Stress dependent thermal pressurization of a fluid-saturated rock. *Rock Mech Rock Eng*, 42:1–24, 2009.

[58] Ghavamian, F. and Simone, A. Accelerating multiscale finite element simulations of history-dependent materials using a recurrent neural network. *Comput. Methods Appl. Math.Eng.*, 357:23, 2019.

[59] Goldberg, D. E. and Holland, J. H. Genetic algorithms and machine learning. *Machine Learning*, 3(2):95–99, 1988.

[60] Graham, J. and Houlsby, G. T. Anisotropic elasticity of natural clay. *Géotechnique*, 33:165–180, 1983.

[61] Grepl, M., Maday, Y., Nguyen, N., and Patera, A. Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 41(3):575–605, 2007.

[62] Grepl, M. A. Certified reduced basis methods for nonaffine linear time-varying and nonlinear parabolic partial differential equations. *Mathematical Models and Methods in Applied Sciences*, 22(03):1150015, 2012.

[63] Grepl, M. A. and Patera, A. T. A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations. *M2AN Math. Model. Numer. Anal.*, 39:157–181, 2005.

[64] Guo, M. and Hesthaven, J. S. Reduced order modeling for nonlinear structural analysis using Gaussian process regression. *Computer Methods in Applied Mechanics and Engineering*, 341:807–826, 2018.

[65] Haasdonk, B. and Ohlberger, M. Reduced basis method for finite volume approximations of parametrized linear evolution equations. *M2AN Math. Model. Numer. Anal.*, 42:277–302, 2008.

[66] Haasdonk, B., Dihlmann, M., and Ohlberger, M. A training set and multiple basis generation approach for parametrized model reduction based on adaptive grids in parameter space. *Math. Comput. Model. Dyn. Syst.*, 17:423–442, 2011.

[67] Hakami, H. Rock characterisation facility (RCF) shaftsinking—numerical computations using FLAC. *Int J Rock MechMin Sci*, 38(1):59–65, 2001.

[68] Hanninen, H., Forsstrom, A., and Yagodsinzkyy, Y. Copper behavior in geological nuclear waste disposal. *European Federation of Corrosion (EFC) Series, Nuclear Corrosion, Woodhead Publishin*, pages 391–402, 2020.

[69] Hernandez, J. A., Caicedo, M. A., and Ferrer, A. Dimensional hyper-reduction of nonlinear finite element models via empirical cubature. *Computer Methods in Applied Mechanics and Engineering*, 313:687–722, 2017.

[70] Horseman, S. T., Winter, M. G., and Entwistle, D. C. Geotechnical characterization of boom clay in relation to the disposal of radioactive waste. Technical report, Commission of the European Community, 1987.

[71] Huerta, A. and Diez, P. *Implicit residual type error estimators*, pages 19–32. Springer, 2016.

[72] Hurty, W. C. Dynamic analysis of structural systems using component modes. *American Institute of Aeronautics and Astronautics Journal*, 3(4):678–685, 1965.

[73] Jha, B. and Juanes, R. Coupled multiphase flow and poromechanics: a computational model of pore pressure effects on fault slip and earthquake triggering. *Water Resources*, 50(5):3776–3808, 2014.

[74] Junge, M., Brunner, D., Becker, J., and Gaul, L. Interface-reduction for the Craig–Bampton and Rubin method applied to FE–BE coupling with a large fluid-structure interface. *Int. J. Numer. Methods Eng*, 77(12):1731–1752, 2009.

[75] Kalashnikova, I., Bloemen, B. V., Arunajatesan, S., and Barone, M. Stabilization of projection-based reduced order models for linear time-invariant systems via optimization-based eigenvalue reassignment. *Computer Methods in Applied Mechanics and Engineering*, 272:251–270, 2014.

[76] Karhunen, K. Uber lineare methoden in der wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fenn*, 37:1–79, 1946.

[77] Kerfriden, P., Gosselet, P., Adhikari, S., and Bordas, S. Bridging Proper Orthogonal Decomposition methods and augmented Newton-Krylov algorithms: an adaptive model order reduction for highly nonlinear mechanical problems. *Comput Methods Appl Mech Eng* , 200(5–8):850–866, 2011.

[78] Kim, J. G. and Markovic, D. High-fidelity flexibility-based Component Mode Synthesis method with interface degrees of freedom reduction. *American Institute of Aeronautics and Astronautics Journal*, 54:3619–3631, 2016.

[79] Knowles, N. C., Jeffries, R. M., and Come, B. Benchmarking the thermomechanical behaviour of clays a progress report the CEC interclay project. *Engineering Geology*, 41(1-4):65–71, 1996.

[80] Lachenburch, A. Friction heating, fluid pressure and the resistance to fault motion. *J. Geophys. Res*, 85:6097 – 6112, 1980.

[81] Larion, Y., Zlotnik, S., Massart, T. J., and Díez, P. Building a certified reduced basis for coupled thermo-hydro-mechanical systems with goal-oriented error estimation. *Comput Mech*, 66:559–573, 2020.

[82] Levasseur, S., Malécot, Y., Boulon, M., and Flavigny, E. Soil parameter identification using a genetic algorithm. *Int. J. Numer. Anal. Meth. Geomech.*, 32:189–213, 2008.

[83] Lewis, R. B. and Schrefler, B. A. *The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media.* 2. Wiley, 1999.

[84] Li, S., Fan, C., Han, J., Luo, M., Yang, Z., and Bi, H. A fully coupled thermal-hydraulic-mechanical model with two-phase flow for coalbed methane extraction. *J. Nat. Gas Sci. Eng*, 33:324 –336, 2016.

[85] Loeve, M. *Probability Theory.* Third Edition. Van Nostrand Princeton, 1963.

[86] Manzoni, A., Quarteroni, A., and Rozza, G. Computational reduction for parametrized pdes: Strategies and applications. *Milan Journal of Mathematics*, 80:293–309, 2012.

[87] Martini, I., Rozza, G., and Haasdonk, B. Reduced basis approximation and a-posteriori error estimation for the coupled stokes-darcy system. *Advances in Computational Mathematics*, 41(5):1131–1157, 2015.

[88] Michalowski, R. L. and Zhu, M. Modelling of freezing in frost-susceptible soils. *Comput. Assist. Mech. Eng. Sci.*, 13:613–625, 2006.

[89] Milliard, A., Rejeb, A., and Chijimatsu, M. Numerical study of the THM effects on the near-field safety of a hypothetical nuclear waste repository—BMT1 of the DECOVALEX-III project. Part 1: Conceptualization and characterization of the problems and summary of results. *Int J Rock Mech Mining Sci*, 42(5–6): 720–730, 2005.

[90] N Pares and P Diez and A Huerta, t. . B. *Comput. Methods Appl. Mech. Engrg.*, 197:1661–1679, 2008.

[91] Nelder, J. A. and Mead, R. A simplex method for function minimization. *Comput. J.*, 7:308–313, 1965.

[92] Nguyen, N. C., Rozza, G., and Patera, A. T. Reduced basis approximation and a posteriori error estimation for the time-dependent viscous Burgers' equation. *Calcolo*, 46:157–185, 2009.

[93] Nguyen, T. S. and Selvadurai, A. P. S. Coupled thermal-mechanical-hydrological behaviour of sparsely fractured rock: Implications for nuclear fuel waste disposal. *Int. J. Rock Mech. Min. Sci. Geomech.*, 32:465– 479, 1995.

[94] Nguyen, T. S., Börgesson, L., Chijimatsu, M., Rutqvist, J., Fujita, T., Hernelind, J., Kobayashi, A., Ohnishi, Y., Tanaka, M., and Jing, L. Hydro-mechanical response of a fractured granitic rock mass toexcavation of a test pit—the

Kamaishi Mine experiment in Japan. *Int J Rock Mech Min Sci*, 38(1):79–94, 2001.

[95] Nguyen, T. S., Selvadurai, A. P. S., and Armand, G. Modeling the FEBEX THM experiment using a state surface approach. *Int J Rock Mech Mining Sci*, 42(5–6):639–651, 2005.

[96] Nguyen, T. S., Börgesson, L., and Chijimatsu, M. A case study on the influence of THM coupling on the near field safety of a spent fuel repository in sparsely fractured granite. *Environ Geol*, 57:1239–1254, 2009.

[97] Noor, A. K. and Peters, J. M. Reduced basis technique for nonlinear analysis of structures. *American Institute of Aeronautics and Astronautics Journal*, 18 (4):455–462, 1980.

[98] Nouy, A. Proper generalized decompositions and separated representations for the numerical solution of high dimensional stochastic problems. *Arch. Comput. Methods Engrg*, 17:403–434, 2010.

[99] Nowak, T., Kunz, H., Dixon, D., Wang, W., Goerke, U., and Kolditz, O. Coupled 3D THM analysis of geotechnical in situ tests. *Int. J. of Rock Mech. Min. Sci.*, 48(1):1–15, 2011.

[100] Olivella, S. and Gens, A. Double structure thm analyses of a heating test in a fractured tuff incorporating intrinsic permeability variations. *International Journal of Rock Mechanics and Mining Sciences*, 42(5–6):667–679, 2005.

[101] Onofrei, C. and Gray, M. Modelling hygro-thermomechanical behaviour of engineered clay barriers: validation phase. *Engineering Geology*, 41(1):301–318, 1996.

[102] Ortega-Gelabert, O., Zlotnik, S., Afonso, J. C., and Díez, P. Fast stokes flow simulations for geophysical-geodynamic inverse problems and sensitivity analyses based on reduced order modeling. *Journal of Geophysical Research: Solid Earth*, 125, 2020.

[103] Pal, S., Wathugala, G., and Kundu, S. Calibration of constitutive model using genetic algorithms. *Comput Geotech*, 19(4):325–348, 1996.

[104] Pan, P. Z., Feng, X. T., and Huang, X. H. Coupled THM processes in EDZ of crystalline rocks using an elasto-plastic cellular automaton. *Environ Geol*, 57 (1299), 2009.

[105] Papon, A., Riou, Y., Dano, C., and Hicher, P. Y. Single-and multi-objective genetic algorithm optimization for identifying soil parameters. *International Journal for Numerical and Analytical Methods in Geomechanics*, 36(5):597–618, 2011.

[106] Pares, N., Diez, P., and Huerta, A. Bounds of functional outputs for parabolic problems. Part I: exact bounds of the discontinuous galerkin time discretization. *Comput. Methods Appl. Mech. Engrg.*, 197:1641–1660, 2008.

[107] Patera, A. T. and Rozza, G. *Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations.* MIT-Pappalardo Graduate Monographs in Mechanical Engineering, Preface and Part I. Cambridge, MA, US, (C) MIT, Massachusetts Institute of Technology, 2007.

[108] Peherstorfer, B., Butnaru, D., Willcox, K., and Bungartz, H. J. Localized discrete empirical interpolation method. *SIAM J. Sci. Comput.* , 36(1):168–192, 2014.

[109] Perras, M. and Diederichs, M. Predicting excavation damage zone depths in brittle rocks. *Journal of Rock Mechanics and Geotechnical Engineering*, 8(1): 60–74, 2016.

[110] Pfaller, M. R., Varona, M. C., Lang, J., Bertoglio, C., and Wall, W. A. Using parametric model order reduction for inverse analysis of large nonlinear cardiac simulations. *Int J Numer Meth Biomed Engng.*, 36, 2020.

[111] Posiva, O. Nuclear Waste Management of the Olkiluoto and Loviisa Power Plants: Programme for Research, Development and Technical Design for 2004-2006. Technical report, POSIVA-TKS–2003. Finland, 2003.

[112] Prudhomme, C., Rovas, D., Veroy, K., Maday, Y., Patera, A. T., and Turinici, G. Reliable real-time solution of parametrized partial differential equations: Reduced basis output bounds methods. *J. Fluids Eng*, 124(1):70–80, 2002.

[113] Quarteroni, A., Rozza, G., and Manzoni, A. Certified reduced basis approximation for parametrized partial differential equations in industrial applications. *J. Math. Ind.*, 1(3), 2011.

[114] Radermacher, A. and Reese, S. Model reduction in elastoplasticity: proper orthogonal decomposition combined with adaptive sub-structuring. *Comput Mech*, 54(3):677–687, 2014.

[115] Rechea, C., Levasseur, S., and Finno, R. Inverse analysis techniques for parameter identification in simulation of excavation support systems. *Computers and Geotechnics*, 35(3):331–345, 2008.

[116] Reis, T. and Stykel, T. Stability analysis and model order reduction of coupled systems. *Math. Comput. Model.Dyn. Syst*, 13:413–436, 2007.

[117] Rice, J. R. Heating and weakening of faults during earthquake slip. *J. Geophys. Res.*, 111:1–29, 2006.

[118] Rocha, I., Kerfriden, P., and van der Meer, F. P. Micromechanics-based surrogate models for the response of composites: A critical comparison between a classical mesoscale constitutive model, hyper-reduction and neural networks. *European Journal of Mechanics - A/Solids*, 82, 2020.

[119] Rozza, G. and Veroy, K. On the stability of the reduced basis method for stokes equations in parametrized domains. *Comput. Methods Appl. Mech. Engrg*, 196: 1244–1260, 2007.

[120] Rutqvist, J., Borgesson, L., Kobayashi, A., Jing, L., Nguyen, T., JNoorishad, and Tsang, C. F. Thermo-hydro-mechanics of partially saturated geological media: Governing equations and formulation of four finite element codes. *Int. J. Geomater. Mech. Min. Sci.*, 36:105–127, 2001.

[121] Rutqvist, J., Barr, D., and Datta, R. Coupled thermal–hydrological–mechanical analyses of the yucca mountain drift scale test – comparison of field measurements to predictions of four different numerical models. *Int J Rock Mech Mining Sci*, 42(5–6):680–697, 2005.

[122] Rutqvist, J., Bäckström, A., and Chijimatsu, M. A multiple-code simulation study of the long-term EDZ evolution of geological nuclear waste repositories. *Environ Geol*, 57:1313–1324, 2009.

[123] Rutqvist, J., Börgesson, L., and Chijimatsu, M. Modeling of damage, permeability changes and pressure responses during excavation of the TSX tunnel in granitic rock at URL, Canada. *Environ Geol*, 57:1263–1274, 2009.

[124] Rutqvist, J., Vasco, D., and Meyer, L. Coupled reservoir-geomechanical analysis of co2 injection and ground deformations at in salah, algeria. *Int. J. Greenhouse Gas Control*, 4(12):225–230, 2010.

[125] Ryckelynck, D. A priori hyperreduction method: An adaptive approach. *J. Comput. Phys*, 202:346–366, 2005.

[126] Ryckelynck, D. Hyper Reduction of finite strain elasto-plastic models. *International Journal of Material Forming*, 2(1):567–571, 2009.

[127] Ryckelynck, D. and Benziane, D. M. Multi-level A Priori Hyper-Reduction of mechanical models involving internal variables . *Comput. Methods Appl. Mech. Eng.*, 199:1134–1142, 2010.

[128] Ryckelynck, D., Vincent, F., and Cantournet, S. Multidimensional a priori hyperreduction of mechanical models involving internal variables. *Computer Methods in Applied Mechanics and Engineering*, 28(43):225–228, 2012.

[129] Samarajiva, P., Macari, E., and Wathugala, W. Genetic algorithms for the calibration of constitutive models of soils. *Int J Geomech*, 5(3):206–217, 2005.

[130] Schmich, M. and Vexler, B. Adaptivity with dynamic meshes for space-time finite element discretization of parabolic equations. *SIAM J Sci Comput*, 30: 369–393, 2008.

[131] Schrefler, B. A., Simoni, L., and Turska, E. Standard staggered and staggered Newton schemes in thermo-hydro-mechanical problems. *Computer Methods in Applied Mechanics and Engineering*, 144(1–2):93–109, 1997.

[132] Selvadurai, A. P. S. Heave of a surficial rock layer due to pressures generated by injected fluids. *Geophys. Res. Lett.*, 36(14), 2009.

[133] Selvadurai, A. P. S. *Caprock breach: a potential threat to secure geologic sequestration of CO2*, pages 75–93. John Wiley and Sons, Hoboken,NJ, 2013.

[134] Selvadurai, A. P. S. and Nguyen, T. S. Scoping analyses of the coupled thermal-hydrological-mechanical behaviour of the rock mass around a nuclear fuel waste repository. *Engineering Geology*, 47(4):379–400, 1997.

[135] Selvadurai, A. P. S., Hu, J., and Konuk, I. Computational modelling of frost heave induced soilpipeline interaction. ii. modelling of experiments at the caen test facility. *Cold Regions Sci. Technol.*, 29:229–257, 1999.

[136] Selvadurai, A. P. S., Surorov, A. P., and Selvadurai, P. A. Thermo-hydro-mechanical processes in fractured rock formations during a glacial advance. *Geosci. Model Dev.*, 8:2167–2185, 2015.

[137] Simo, J. C. and Taylor, R. L. Consistent tangent operators for rate-independent elastoplasticity. *Computer Methods in Applied Mechanics and Engineering*, 48 (1):101–118, 1985.

[138] Simoni, L., Secchi, S., and Schrefler, B. Numerical difficulties and computational procedures for thermo-hydro-mechanical coupled problems of saturated porous media. *Computational Mechanics*, 43(1):179–189, 2008.

[139] Solomon, B. D. Solving a wicked problem in deep time: Nuclear waste disposal. *Sci Eng Ethics*, 27(21), 2021.

[140] Sultan, N., Delage, P., and Cui, Y. J. Temperature effects on the volume change behaviour of boom clay. *Engineering Geology*, 64(2–3):135–145, 2002.

[141] Sultan, N., Cui, Y. J., and Delage, P. Yielding and plastic behaviour of boom clay. *Geotechnique*, 60(9):657–666, 2010.

[142] Thomas, H. R. and He, Y. Modelling the behaviour of unsaturated soil using an elastoplastic constitutive relationship. *1998*, 48(5):589–603, 1998.

[143] Tucker, L. R. *The extension of factor analysis to three-dimensional matrices*, pages 110–127. Holt, Rinehart and Winston, New York, 1964.

[144] Vardon, P., Banicescu, I., Cleall, P., Thomas, H., and Philp, R. Coupled thermo-hydro-mechanical modelling: A new parallel approach. *IEEE International Symposium on Parallel and Distributed Processing*, 1(5):2535–2542, 2009.

[145] Verdugo, F., Pares, N., and Diez, P. Goal-oriented spacetime adaptivity for transient dynamics using a modal description of the adjoint solution. *Comput Mech*, 54(2):331–352, 2014.

[146] Viswanath, D. S., Ghosh, T., Prasadand, D. H. L., Dutt, N. V. K., and Rani, K. Y. *Viscosity of Liquids: Theory, Estimation, Experiment, and Data*. Springer, 2007.

[147] Wang, J. and Zabaras, N. Using Bayesian statistics in the estimation of heat source in radiation. *International Journal of Heat and Mass Transfer*, 48:15–29, 2005.

[148] Wang, W. and Kolditz, O. Object-oriented finite element analysis of thermo-hydromechanical (THM) problems in porous media. *International Journal for Numerical Methods in Engineering*, 69(1):162–201, 2007.

[149] Wang, W., Kosakowski, G., and Kolditz, O. A parallel finite element scheme for thermo-hydro-mechanical (THM) coupled problems in porous media. *Computers and Geosciences*, 35(8):1631–1641, 2009.

[150] Wang, X. and Schrefler, B. Fully coupled thermo-hydro-mechanical analysis by an algebraic multigrid method. *Engineering Computations*, 20(1–2):211–229, 2003.

[151] Wang, X., Ge, Z., and Wu, H. An algebraic multigrid method for coupled thermohydro-mechanical problems. *Applied Mathematics and Mechanics*, 23 (12):1464–1471, 2002.

[152] World Nuclear Association. Storage and disposal of radioactive waste, May 2021.

[153] Yu, L., Rogiers, B., Gedeon, M., Marivoet, J., Craen, M. D., and Mallants, D. A critical review of laboratory and in-situ hydraulic conductivity measurements for the Boom Clay in Belgium. *Applied Clay Science*, 75-76:1–12, 2013.

[154] Zee, K. G., Oden, J. T., Prudhomme, S., and Hawkins-Daarud, A. Goal-oriented error estimation for cahn-hilliard models of binary phase transition. *Numerical Methods for Partial Differential Equations*, 27(1):160–196, 2010.

[155] Zhang, Y., Feng, L., Li, S., and Benner, P. An efficient output error estimation for model order reduction of parametrized evolution equations. *SIAM Journal on Scientific Computing*, 37:910–936, 2015.

[156] Zhou, L., Simon, J. W., and Reese, S. Proper orthogonal decomposition for substructures in nonlinear finite element analysis: coupling by means of tied contact. *Arch Appl Mech* , 88:1975–2001, 2018.

[157] Zienkiewicz, O. C., Chan, A., Pastor, M., Schrefler, B., and Shiomi, T. *Computational Geomechanics with Special Reference to Earthquake Engineering*. Wiley, 1999.

Acknowledgements

Firstly, I would like to thank the financial support granted by the European Commission Education, Audiovisual and Culture Executive Agency (EACEA) under Erasmus Mundus Joint Doctorate Simulation in Engineering and Entrepreneurship Development (SEED) program during the first three years of this PhD. The financial support of the Belgian nuclear research centre (SCK-CEN) in the final year is also greatly acknowledged.

I would like to express my deepest gratitude to my thesis supervisors, Thierry J. Massart, Pedro Díez and Sergio Zlotnik for their continuous support and guidance during this PhD journey. I am forever grateful for their generosity, insights, flexibility, encouragement, and trust. I am also thankful to my industry mentors from SCK-CEN, Guangjing Chen and Suresh Seetharam, for their invaluable contributions to this thesis and their advices outside the academic work.

I would also like to thank my current and former colleagues from ULB and UPC for enriching this PhD experience through our meaningful discussions within and beyond the scope of research – Anqi, Arash G., Arash M., Batoma, Christina, Fabiola, Jonatha, Karim E., Karim G., Li, Liesbeth, Peter, Priyam, Rohith, Roland and Zoltan, I feel very fortunate to have had your company both in and out of the office. I must also acknowledge these wonderful people I have met in Brussels and Barcelona in the last four years – Alex, Anna, Ausi, Babass, Bethan, Caroline, Grajo, Guillaume, Jess, Julie, Matthieu, Rahel, Raul, Roger, Salim and Van, I am very thankful for the friendship and moments we shared together.

Lastly, I dedicate this thesis to my family for their unconditional support in all my endeavours – to my parents, Lydia & Leonardo, and to my sisters, Yvette, Yvonne & Yzelle. Thank you for always believing in me.