UNIVERSITY OF CALIFORNIA SAN DIEGO

An Optimal Sensor Design Framework for Structural Health Monitoring

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Structural Engineering

by

Yichao Yang

Committee in charge:

Professor Michael D. Todd, Chair
Professor Joel P. Conte
Professor Charles R. Farrar
Professor William S. Hodgkiss
Professor Zhen Hu
Professor Kenneth J. Loh

2021

The Dissertation of Yichao Yang is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

This dissertation is dedicated to my dear wife, Siyue Wu. During my PhD research journey, she took care of the entire family and provided a huge support both spiritually and physically to me. I also dedicate my work to my parents, Wei Yang and Yiqun Zhu, who gave me the opportunity to study abroad and always believed in me to achieve my academic goals.

EPIGRAPH

True optimization is the revolutionary contribution of modern research to decision processes.
—George B. Dantzig

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank many people who have supported me through my research career in this dissertation. First, I would like to acknowledge the chair of my committee, Prof. Michael Todd, for his critical advice and support academically and personally. Professor Todd provided me advice not only to conquer those difficulties in my research, but also, more importantly, to become a better person in my life. I really appreciate his guidance and proud to be his student. I would also like to acknowledge other committee members, Prof. Joel Conte, Prof. Charles Farrar, Prof. William Hodgkiss, Prof. Kenneth Loh, and Prof. Zhen Hu, for giving me suggestions in my research journey. Especially, I thank Professor Zhen Hu, who shared me his research techniques with personal experiences and contributed significantly to my research in last two years.

In addition, I would like to thank my colleague, Dr. Mayank Chadha, for helping and working together with me for my research publications. I also would like to extend my appreciation to many colleagues and collaborators: Dr. Vega Manuel, Dr. Wang Long, Zihan Wu, Guofeng Qian, Adrielly Hokama Razzini, and David Najera, for all the precious time we have spent together at UCSD.

The research presented in this dissertation was financially supported by the U.S. Army Corps of Engineers and the technical director Matthew Smith. All these sources of support are gratefully acknowledged.

Portions of this dissertation have been published or accepted for publication or are currently being considered for publication.

Chapter 2 is accepted for publication in Mechanical Systems and Signal Processing, the dissertation author was the primary investigator and author of this paper:

Y. Yang, M. Chadha, Z. Hu, and M. D. Todd, "An optimal sensor placement design framework for structural health monitoring using bayes risk," Mechanical Systems and Signal Processing.

Chapter 3 is composed of a first authored publication:

Y. Yang, M. Chadha, Z. Hu, M. A. Vega, M. D. Parno, and M. D. Todd, "A probabilistic optimal sensor design approach for structural health monitoring using risk-weighted f-divergence," Mechanical Systems and Signal Processing, vol. 161, p. 107920, 2021.

Chapter 4 is currently in preparation for publication, the dissertation author was the primary investigator and author of this paper:

Y. Yang, M. Chadha, Z. Hu, and M. D. Todd, "An optimal sensor design framework accounting for sensor reliability over the structural life cycle," Reliability Engineering & System Safety.

VITA

2015 Bachelor of Science, University of California San Diego

2018 Master of Science, University of California San Diego

2021 Doctor of Philosophy, University of California San Diego

PUBLICATIONS

Journal Papers:

Y. Yang, M. Chadha, Z. Hu, and M. D. Todd, "An optimal sensor placement design framework for structural health monitoring using bayes risk," Mechanical Systems and Signal Processing.

Y. Yang, M. Chadha, Z. Hu, M. A. Vega, M. D. Parno, and M. D. Todd, "A probabilistic optimal sensor design approach for structural health monitoring using risk-weighted f-divergence," Mechanical Systems and Signal Processing, vol. 161, p. 107920, 2021.

Y. Yang, M. Chadha, Z. Hu, and M. D. Todd, "An optimal sensor design framework accounting for sensor reliability over the structural life cycle," Reliability Engineering & System Safety (under review).

Vega, M.A., Hu, Z., **Yang, Y.**, Chadha, M. and Todd, M.D., 2022. Diagnosis, Prognosis, and Maintenance Decision Making for Civil Infrastructure: Bayesian Data Analytics and Machine Learning. In Structural Health Monitoring Based on Data Science Techniques (pp. 45-73). Springer, Cham.

Conference Papers:

Yang Y, Madarshahian R, Todd MD. Bayesian Damage Identification Using Strain Data from Lock Gates, Springer,; 2019, p. 47–54. https://doi.org/10.1007/978-3-030-12115-0_7.

ABSTRACT OF THE DISSERTATION

An Optimal Sensor Design Framework for Structural Health Monitoring

by

Yichao Yang

Doctor of Philosophy in Structural Engineering

University of California San Diego, 2021

Michael D. Todd, Chair

Over the life cycle of large civil infrastructure, there is often significant degradation in capability and design performance due to extend usage and, in many cases, deferred maintenance. The ultimate limit states in this case can result in unexpected failure, with potentially large economic or life-safety consequences. Structural health monitoring (SHM) is a framework for monitoring the structure over its life cycle, and this field has expanded rapidly in the past two to three decades due to the urgency of infrastructure aging beyond its design life. One of the critical challenges in a monitoring process is obtaining the most valuable information from the structure responses in the field under a continuous monitoring paradigm. This dissertation will present a

new optimal sensor design framework, using Machine Learning (ML) techniques, including Gaussian Process Regression (GPR), Bayesian inference, uncertainty quantification, and Bayesian optimization, that guides risk-informed SHM system design. In particular, this dissertation: (1) demonstrates a framework for optimal sensor design using Bayes risk as the objective function; (2) further explores the framework using risk-weighted *f*-divergence functions and implements it in a real miter gate problem as a case study; (3) investigates the effects of sensor reliability over a life-cycle monitoring of the miter gate for informing optimal sensor arrangement.

Chapter 1

INTRODUCTION

Structural Health Monitoring (SHM) generally makes assessments and maintenance decisions for the current structural state/performance, with implementation of damage identification strategies, by analyzing the obtained in-situ measured data from the field (Dawson 1976; Farrar and Worden 2007; Brownjohn 2007). In real-world engineering systems, many aging infrastructure systems require an efficient SHM system that can help to infer their current state before experiencing unexpected failures, which could lead to large economic or life safety impact (Cross 2013; Li and Ren 2016; Sony and Laventure 2019). In one example, the United States Army Corps of Engineers (USACE) maintains the country's inland waterway navigation corridor, corresponding to hundreds of billions of capital investment necessary to empower the transport of goods and services nationwide. The current maintenance strategy for these assets is periodic dewatering and inspection, which leads to downtime losses in commercial transport; unscheduled shutdowns due to unexpected failure leads to downtime losses several times higher (Foltz, 2017). The design and deployment of an SHM system promises to enable the minimization of these downtime costs by providing accurate estimates of the current asset which can inform maintenance and operations decisions.

In addition to this and other civil engineering applications, SHM system design is also widely being adopted in other engineering systems, such aerospace engineering, for life-safety justification (Sohn and Farrar 2003). Other motivations or benefits of a SHM system includes (but not limited to) predicting probabilities of failure and estimating the remain useful life in long term asset management, assuming the SHM diagnostics are coupled to a predictive model.

1.1. Structural Health Monitoring (SHM) System

The three fundamental components in a SHM system are: 1) in-situ measurements of structural performances, 2) data-to-information transformations, and 3) decision-making process based on the obtained information. A basic SHM system must be able to process from collecting the raw data all the way to making an operational decision regarding a target structure. To implement these components in a SHM system, a statistical pattern recognition process has evolved into the general modern paradigm of how SHM is done (Farrar 2000). The paradigm involves four main steps: operational evaluation, data acquisition, feature extractions, statistical modeling of the feature discrimination.

1.1.1 Operational Evaluation

Before designing and deploying an SHM system, it is critical to do a cost-benefit analysis of its value. In many civil engineering applications such as big waterway infrastructures, the maintenance costs have grown dramatically due to the degradation of the structural components over time under flat or reduced maintenance budgets (Foltz 2017; Modares and Waksmanski 2013). As with the inland waterway corridor, both unexpected shutdowns due to catastrophic structural failure and frequent unnecessary maintenance actions that induce downtime would have significant economic impact to nationwide transportation. One of the goals of the SHM system is to optimize the maintenance frequency and process based on a conditional basis rather than the current time-based maintenance (Vega, Hu, and Todd 2020). An active SHM system gathers effective information about the structural behavior so that unscheduled downtime (and thus costs) would be reduced while structural lifespan would be extended. Thus, SHM can be an important contributor to total cost reduction for an asset's life cycle.

Defining damage or failure modes of interest dictates what types of sensing and raw data might be most useful for monitoring the structural states (Pozzi and Der 2011). Defining the failure mode is highly application-dependent in the engineering fields. In much of the SHM literature, various damage modes (e.g. crack, corrosion, shear, and etc.) have been defined and estimated (Hayo and Frankenstein 2011; Boscato and Dal Cin 2014; Taha and Lucero 2005). On an aircraft, the most relevant failure mode might be corrosion or metal fatigue cracking, while for a concrete bridge over water, the most relevant failure mode might be pier scour or roller bearing wear (John 2016). In chapters 2 and 3 of this dissertation, a simple 2D beam-spring problem and a real-world Greenup miter gate problem will be presented as case studies. In the beam-spring problem, the damage state is defined as existences of springs connecting the cantilever beam to the wall, which leads to a binary damage state. In the miter gate problem, the most common damage mode is known as the "gap size" (or loss of contact in the quoin block) between the gate and the wall (Eric and Treece 2018). The "gap" produces stress re-distributed through the gate and induces failure. In both problems, strain gauges are installed as hardware equipment to measure strain responses of the structures under various loading scenarios.

Environmental conditions must be considered when installing sensors for the SHM system. For a miter gate partially submerged in water, the sensors must be waterproof and able to withstand harsh weather conditions. All sensors installed on a bridge for monitoring must be able to deal with large changes in temperature without distorting data. The response to the structure can be affected by water levels, changes in boundary conditions, or temperature effects inducing shrinking or swelling in the structure (Foltz 2017). A limitation of the SHM system could be the location of where the sensors are installed. For instance, sensors cannot be installed on the upstream side of the gate. Besides, sensor merged under the water level could have lower reliability

compared to those located above the water level. More discussions on sensor reliability will be illustrated in chapter 4.

1.1.2 Data Acquisition and Feature Extraction

Damage-sensitive features are quantities extracted from the measured response system data that could be used to indicate whether damage has been detected in the structure. If appropriate features are chosen, the pattern recognition and machine learning process can be easy, in which the presence of damage can be clearly inferred from changes in the features (Catbas and Gul 2008; Young Noh 2011). In contrast, if features chosen are not directly correlated with damage, it will be much more difficult to distinguish damage based on the changes in the features. When selecting features, ideally, one would select a feature that is sensitive to damage in the structure and insensitive to operational variability. However, often issues arise as features that are sensitive to damage in the structure are also sensitive to the environmental variables (such as, temperature, wind, moisture). For example, Farrar, in 1994, performed vibration tests on the I-40 Bridge and explored that the bridge's fundamental frequencies with respect to the four damage levels did not trend as expected due to large temperature gradient changes on the structure. As the damage level increased and the stiffness of the bridge decreased, the magnitudes of the bridge's natural frequencies were expected to reduce as well. In these situations, more complex pattern recognition and machine learning techniques can be used to help prevent these issues. Accelerometers and strain gauges are examples of sensors that could be used to gather data on features. Accelerometers can gather data to perform modal analysis, and strain gauges do the same to indicate excessive loading conditions. In the miter gate problem, strain responses from the miter gate are considered

as features for detection of gap length. The measured data is combined with observation noises that will be analyzed in chapter 3.

1.1.3 Damage Detection and Statistical Modeling

Damage detection requires data obtained from the structure responses. The supervised learning approach accrues data from the undamaged and damaged system, or among many damage levels if more than a binary choice is desired. This differs from the unsupervised learning approach, in which data is available only from the undamaged system (Pao and Sobajic 1992). In the majority cases in civil engineering, it would be difficult to obtain data from the damaged state. For this case, the undamaged state would be considered as the current baseline state of the structure. Damage on the structure would be accounted for when changes are observed in the structure with respect to this current baseline state (Rose and Croall 1991; Fugate and Sohn 2000). Unfortunately, using the unsupervised learning approach does not provide necessary information to determine what type and to what extent of damage is present. The type and extent of damage can only be determined through classification and regression analysis, both falling under the supervised learning approach (Lewis and Catlett 1994; Gathercole 1998).

To be able to perform supervised learning approach, Finite Element (FE) modeling is popular among researchers to generate the measurement data from various damaged states, particularly for civil infrastructure (Ren and Chen 2010; Liu and Edberg 2013). These physical-based FE models are generally complicated "forward" models that simulate the response measurement data from various input variables (e.g., damage level, loading conditions). These FE models are usually complex enough to be able to describe the real structure, subject to model validation or updating (Friswell and Mottershead 2001; Lin and Zhu 2006; Mottershead and Link

2011). On the other hand, these models are usually computationally expensive in damage detection problems because it becomes an "inverse" problem requiring estimates of the damage states from given measurements. Especially, in Bayesian calibration strategy, the model must run many samples in a random walk-based method for the mean value of the responses (Yang, Madarshahian, and Todd 2019). More discussions will be in Section 1.3.

1.2. Optimal Design Process in SHM

In the early days of SHM (before 1990's), no clear objective functions were really defined for designing an optimal SHM system (Kidder 1973; Berman 1979). A basic SHM system indicates damage based on comparing the prior undamaged or damaged data to a new observed data set, which uses quantitative models to provide a statistical estimate of the condition of a structure. A bad design of SHM could possibly mislead the maintenance processes which would create a tremendous cost in economic and even life-safety. Designing an SHM system, like any system, must invariably start with a well-defined objective function, and the optimal design should maximize the utility of the objective. Among the first well-defined objectives was formulated in 1991, when Kammer (Kammer 1991) implemented an optimal sensor placement problem for the purposes of modal identification by minimizing a Fisher information matrix (Behboodian 1972; Kleinman and Rao 1977). Other researchers eventually proposed using objective functions such as signal-to-noise ratio (Madu and Madu 1999; Bae and Flachsbart 2004; Pabari and Willson 2011), probability of detection (Peh and Liang 2007; Markmiller and Chang 2010; Bhattacharjee and Das 2015), estimation error minimization (Joshi and Boyd 2008).

None of them have taken the economic loss due to the detection decisions into consideration. A more advanced design technique (not originally intended for SHM but rather for design of statistical experiments), Bayesian Experimental Design (BED) (Chaloner and Verdinelli 1995), can be exploited to design a SHM system by minimizing the total risk or cost (Flynn and Todd 2010). The design implementation as applied for SHM usually includes four steps: i) evaluating the economic benefits, damage states, environmental conditions of the structure (essentially 'operational evaluation' from above), ii) obtaining useful features from the periodically sampled measurement data, iii) setting up the classification or binary decision hypothesis test, iv) transforming the decision statistics into making the most economical decisions (e.g., no action/maintenance/replacement) based on the minimizing the risk-based cost functions. Another advantage of Bayes risk-based strategy is that it accounts for the prior information of a damage state in the optimization process.

The optimal design of a SHM system has various perspectives; the optimization involves sensor acquisition parameters (Lukosi and Rust 2019; Li and Zhao 2019; Costas-Perez and Lago 2008), feature space (Iswandy and Koenig 2006; Gui and Pan 2017; Lin and Tsai 2015), and \ sensor locations (Padula and Kincaid 1999; Krause and Leskovec 2008; Ostachowicz and Soman 2019; Yi and Li 2011;), the latter of which have been the most studied as a design problem. This thesis is also going to focus on the optimal design of sensor locations based by minimizing multiple cost functions, both risk-based and information-based, in SHM. More explanations will be provided in Section 1.3 and chapters 2 and 3.

1.3. Optimal Sensor Design for SHM

Sensors essentially serve as a communication medium between structural response and human decisions regarding their behavior. Within the modern paradigm of SHM, sensors are the first step of the process, required to measure the in-situ responses of the structure under various loading conditions. In the early days (1990s) of the SHM field, the primary design strategy of traditional structural monitoring systems was to place as many sensors as possible on the structure to collect data without regard to the actionable information they collected, which was relatively inefficient and expensive; there was little early effort put into how to optimize such a design, according to some objective.

Optimal sensor placement is now widely studied and explored by researchers in many engineering systems including electrical engineering (Jung and Song 2015; Brandisky and Sankowski 2012), mechanical engineering (Jeon and Das 2021; Zabel and Kirchhof 2015), computer science engineering (Singh 2018), civil engineering (Liu and Yao 2005; Teng and Zhu 2011), aerospace engineering (Ye and Ling 2012; Khatami 2021). For instance, Tesla needs to optimally arrange its sensors around the vehicle to have the best or most accurate real-life response from the road in different situations. In civil engineering, infrastructure such as locks and dams for the inland waterway corridor requires an effective SHM system with to prevent unscheduled shutdowns, as mentioned above (Daniel 2000). Optimal sensor design becomes an important part in this cost minimization process, maximizing the value of any information gathered by the SHM system (Yang and Chadha 2021).

The key idea of optimal sensor placement is to design sensor locations over space in purpose to maximize the valuable information obtained from the structures with a minimum cost. When designing an optimal sensing system for a structure such as miter gate, the process integrates

feature extractions, damage diagnosis, risk-based cost functions, uncertainty quantifications, and optimization techniques. Figure 1.1 shows the steps of designing a risk-based SHM system.

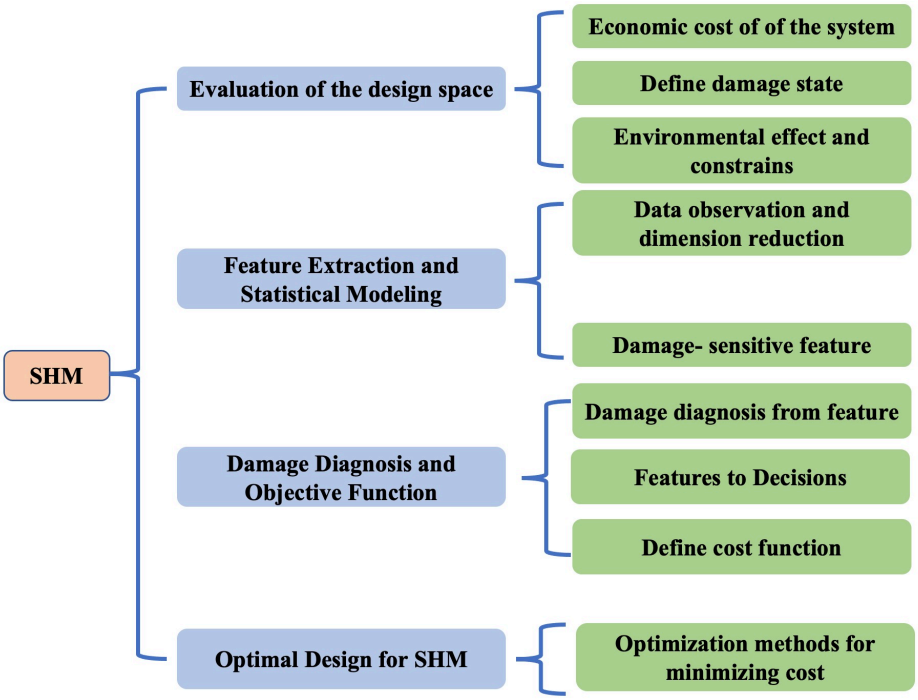


Figure 1.1: Risk-based SHM sensor design steps (process)

The framework involves, 1) surrogate modeling using Machine Learning (ML) strategies (Dupuis and Jouhaud 2018; Kim and Boukouvala 2020; Lu and Ricciuto 2019), 2) Bayesian inference for damage identification (Yan and Chronopoulos 2020; Yang, Madarshahian, and Todd 2019), 3) Risk-based objective functions (Chadha, Hu, and Todd 2021), and 4) Bayesian optimization techniques for optimal sensor placement (Wang and Zoghi 2013; Garnett and Osborne 2010).

1.3.1 Surrogate Modeling

During the calibration process, due to the high computational cost of the FE models, surrogate models (or meta-models) must be built. The goal of a surrogate model is to build a "run time" map from inputs and outputs. Machine learning techniques are widely deployed in building computationally efficient surrogate models. There are existing many different surrogate models that have been used in the literature such as support vector regression (SVR) (Xiang and Li 2017; Maolin and Liye 2020), neural networks (NN) (Tripathy and Bilionis 2018; Vega and Todd 2020), and Gaussian process regression (GPR) (Su and Peng 2017; Williams and Heng 2019). The data-driven surrogate model usually runs significantly faster than the physical-based FE model. For instance, in our optimal sensor design for the miter gate, the GPR method is used to build a forward surrogate model instead of frequently running the "ground truth" FE model to predict the strain gauge values from input parameters. The output of the surrogate model usually has a very large dimension. The single value decomposition (SVD) technique (Van Loan 1976) reduces the high-dimensional correlated output space to low-dimensional uncorrelated features to overcome the problem. The digital surrogate model computes 50000 times faster than the original FE model without sacrificing much accuracy. Figure 1.2 indicates the performance of surrogate model by plotting the predicted latent space value and true latent space value for each feature at each dimension.

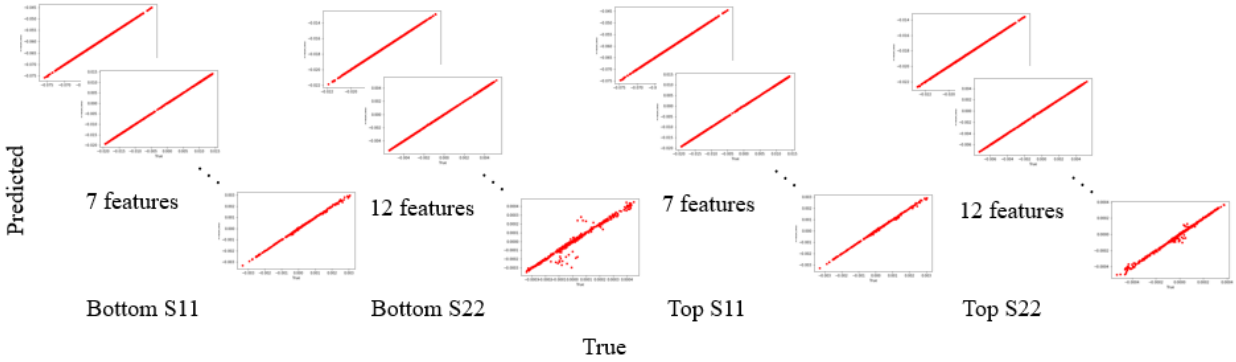


Figure 1.2: Accuracy test for surrogate model in Latent space

Figure 1.2 shows that the true values of the important features in latent space are almost same as the predicted values from our surrogate, as all the plots are almost diagonal for all four dimensions. Chapter 3 will explain the surrogate modeling in more detail.

1.3.2 Bayesian Inference

In damage detection process, Bayesian inference method is widely implemented for calibrating the damage state from the observed measurements (Box and Tiao 2011). The main advantage of Bayesian calibration is it offers a probabilistic estimation of damage by gathering prior information with obtained data. Chapter 2 will explain a deterministic solution for computing the likelihood in the inference process. Chapter 3 will illustrate the particles filter or sequential Monte Carlo (SMC) method (Lopes and Tsay 2011; Lee and Cohen 2002) to perform Bayesian inference for the continues variable gap length given observation data.

1.3.3 Bayes Risk

Bayes risk is an objective function that plays a critical role in the sensor optimal design framework. Generally, Bayes risk, in a design space, includes information from prior information,

posterior damage state prediction from observed data, and associated costs (or risks) with the decisions made from the estimated damage states. The primary goal is to find the best design over the entire design space that minimizes the total costs. The Bayes risk functions can be constructed in various ways depending on the type of objectives and decisions engineers want to achieve. Chapter 2 will show a linear binary decision-to-cost in the Bayes risk function design. The cost function includes intrinsic costs and extrinsic costs. Extrinsic costs include all consequences of the SHM system decisions, both correct and incorrect. Intrinsic costs include sensor design, installation, and maintenance. Chapter 3 will show a Bayes risk function mainly optimizes the information gain (*f*-divergences) of posterior estimation from prior distributions of the damage state.

1.3.4 Uncertainty Quantification

In the process of evaluating the Bayes risk, there are uncertainties from the input variables (damage level and loading position) and observation data (Smith 2013; Sullivan 2015). As part of formulating the risk function and extremizing it, we need to calculate its expected value. There are multiple methods of computing the expected costs. The first approach is sampling-based method (Gelfand 1996; Janson and Schmerling 2015). Basically, it generates inputs n_{mcs} number of times and compute posterior probability of damage and the corresponding costs n_{mcs} number of times. The expected cost is just the average value of these costs. The second approach is mean value approximation (or 0^{th} order Taylor expansion) (Huang and Du 2008). This method just evaluates the 0^{th} term of the Taylor expansion of the expected costs, which equals to evaluate the inputs variables at mean values. The third method is called univariate dimension reduction with Gaussian Hermite integration (Rahman and Xu 2004). This method neglects all terms with dimension two

and higher in the Taylor expansion and approximate the integral with Gaussian Hermite points. Chapter 2 compares and contrasts the tradeoff between computational efficiency and estimation accuracy for all three methods.

1.3.5 Optimal Sensor Placement

The basic idea of the proposed optimal sensor design is to select the best arrangement of sensors that produces minimum Bayes risk, as discussed in Section 1.3.3. The searching process for an optimal design is computationally expensive because it must be evaluated over the ranges of all the uncertain variables. As a result, machine learning techniques becomes a significant part of the optimization process. In the sensor optimization design, there are approaches such as gradient descent that can find an optimal point quickly (Nagarajan and Kolter 2017; Zhang 2019; Mustapha and Mohamed 2020). However, the main drawback of that is it can easily stack at a local optimum. To find the global optimal point it needs more computational efforts. The main motivation of using Bayesian optimization is locate the optimal solution by minimizing the sampling points in order to hasten the optimization process (Brochu and Cora 2010; Garnett and Osborne2010; Lisicki and Lubitz 2016). Bayesian optimization looks for the global optimum in minimum number of steps. It incorporates prior belief about objective function and updates the prior with samples drawn from the objective function to get a posterior that better approximates the function. One of the critical components of the Bayesian optimization is building a surrogate function using GPR.

1.4. Role of Sensor Reliability over the Life Cycle

Structures such as the miter gates employed herein as a case study generally degrade over time due to the aging of structural components. The sensors themselves can also degrade, also increasing their own probabilities of failure. In a life-cycle optimal sensor design framework, the time-dependent sensor reliability has critical impact on the optimization process. The acquired data can be biased if sensors are damaged, and then the inferred damage level (i.e., "gap" between the miter gate and the supporting wall) would be inaccurate. On top of the discussed optimal sensor design framework in Section 1.2, it is critical to account for changing of the sensor reliability when designing sensor locations multiple life cycle time-steps. More discussions of this study will be demonstrated in Chapter 4.

1.5. Research Objectives, Organizations of Thesis and Contributions:

The overarching objective of this thesis is to develop a risk-based optimal sensor design framework. Towards this goal, there are some challenges as:

1) Build the optimal sensor design framework based on evaluating Bayes risk in a manner that overcomes the heavy computational effort over multiple time steps (the life cycle).
2) Develop and implement the framework on the real miter gate problem with an informative cost function.
3) Extend the optimal sensor design framework in SHM to account for sensor reliability and to consider the structure state changes over time.

The remainder of the dissertation will be introduced by the order of the challenge tasks listed above. Chapter 2 will demonstrate a complete, well-constructed, and computationally fast framework for optimal sensor design in SHM using Bayes risk. A simple beam-spring problem is

used to show the power and effectiveness of the framework. Chapter 3 will show the framework in a real-world problem. The framework from Chapter 2 is further developed in terms of Bayes risk function, employing the information gain during Bayesian inference, and minimizing the computational expense in the optimization process. Chapter 4 will illustrate a life cycle sensor optimization design accounting for sensor reliably in SHM that integrates a structural degradation model.

The main contributions presented on the following chapters can be summarized as:

1) Developed a surrogate modeling approach with high prediction accuracy to overcome the lack of real-life data from various damage/loading conditions and heavy computations efforts of the physical-based FE model.
2) Implemented Bayesian inference in estimating the damage level probabilistically, including prior information.
3) Constructed two Bayes risk functions for the beam-spring and miter gate problems based on their different engineering demands.
4) Computed the uncertainties of the inputs and noise using a relatively accurate and efficient estimation method, called univariate dimension reduction with Gaussian Hermite integration.
5) Developed and deployed the Bayesian optimization strategy for sensor optimization, which includes surrogate modeling and expected improvement algorithms.
6) Designed and developed a reliability model for sensors.
7) Integrated sensor reliability with life cycle degradation in SHM system design process.

Chapter 2

AN OPTIMAL SENSOR PLACEMENT DESIGN FRAMEWORK FOR STRUCTURAL HEALTH MONITORING USING BABYES RISK

2.1. Abstract

This chapter presents a novel generalized framework for optimal sensor placement design for structural health monitoring (SHM) applications using Bayes risk as the objective function. Bayes risk considers the costs of consequences associated with making decisions and design selection (extrinsic cost) in the monitoring process, as well as intrinsic costs (e.g., sensor deployment and maintenance costs), which suggests that it is a natural choice for an SHM design objective function. The framework is intended to be sufficiently generalized to be applicable to any optimal sensor placement design used for SHM. To demonstrate the effectiveness and comprehensiveness of the proposed framework, it is applied to an example problem concerning the state detection of the boundary of a beam modeled by springs. I discuss in-depth the specific formulation of Bayes risk for this demonstration problem and detail multiple approaches to evaluate it. This work addresses the challenges encountered in optimal sensor design problem due to the computationally expensive physics-based model, and it considers various uncertainties through the investigation and integration of Bayesian inference methods, uncertainty quantification, and optimization strategies. The effect of the initial design assumption and the technique used to approximate the Bayes risk on the final optimal sensor design is discussed.

2.2. Introduction

Structural health monitoring (SHM) may be generally defined as the process of making an assessment, based on appropriate analyses of in-situ measured data, about the current ability of a structural component or system to perform its intended design function(s) successfully. When coupled with future predictive capabilities, a successful SHM strategy may enable significant ownership cost reduction through maintenance optimization, performance maximization during operation, and unscheduled downtime minimization, and/or enable significant life safety advantage through catastrophic failure mitigation. Such an SHM strategy inevitably must, for a sufficiently well-defined application, include in-situ data acquisition, feature extraction from the acquired data, statistical modeling of the features, and subsequent hypothesis-based synthesis of the feature probabilistic models to make informed decisions about what to do with the structural component or system. Clearly, an important underlying enabler for an SHM strategy is the design of the sensor system, as data acquisition is the initiator of this multi-part paradigm (Farrar Worden 2012; Farrar and Park 2010). As no widely accepted sensing strategy in SHM has been adopted for use, this chapter will propose and demonstrate the implementation of minimal Bayes risk as a natural target objective for SHM system design.

The application case that will be used in this chapter is taken from the inland waterway infrastructure. The locks and dams that comprise the inland waterways infrastructure require an effective SHM system to prevent their unexpected failure and continuous monitoring to prevent huge economic losses (Richardson 1964; Daniel 2000). The United States Army Corps of Engineers (USACE) spends billions of dollars in maintaining and operating this infrastructure, where unscheduled shutdown of these assets and dewatering for inspection or repair is very costly (Parno and Connor 2018; Schwieterman and Field 2010; Yang and Madarshahian 2020; Foltz

2017). The need for SHM to help facilitate maintenance and operations appears strong, but highly constrained budgets suggest SHM system allocation efforts must be optimized to meet risk-based goals. Within the navigation lock systems, miter gates are one of the most common locking gates used; their most common failure mechanisms include long-term corrosion and loss of load-transferring contact in the quoin block (boundary related damage) (Eric and Treece 2018). As many of these structures have been operational for over 50 years, many are presently potentially operating in a higher-risk profile without engineers knowing their real structural capability (Wilkins 1956); current practice involves engineering elicitation via inspection, followed by lock closures if the inspection so warrants. Since this process is based on the varied experience and interpretation of field engineers, it bears high uncertainties and variability (Vega and Todd 2020). The use of SHM could potentially reduce those uncertainties, but the value of information obtained depends upon its design (Thöns and Faber 2013). In general terms, the first step of the SHM system design is to decide what suitable sensors (e.g., strain-gauges, accelerators, etc.) provide measurements from which the extracted features are correlated to the type of damage or state to be inferred. The second step is then to obtain a sensor network design (e.g., number of sensors, location/placement, duty cycle, etc.) that provides the most valuable information at a minimal cost (Nath and Hu 2017; Padula and Kin 1999; Malings and Pozzi 2017).

Numerous seminal contributions have been made in optimal sensor placement design for a wide class of SHM applications (Chan 1997; Yi and Li 2011; Peddada and Tannous 2020). The overall goal of choosing the *best* sensor design is to let the monitoring system gather the most *effective* information from in-situ monitoring to detect the target state (Akbarzadeh and Lévesque 2014). During the optimization process, an optimality criterion or an objective function is used to evaluate the effectiveness of the design. The best sensor design for a considered application,

therefore, depends on the optimality criterion or objective function chosen. Thus, engineers from different fields may have different criteria for defining this to obtain the *best* design that leads to the most *effective* information use. In other words, the engineers look for the best objective function that is in line with the primary goal of the monitoring system, and it evaluates the value of that information in some way. Some classic such objective functions include the probability of detection (POD), and the probability of classification (Peh and Liang 2007). For instance, in the aviation sector, engineers maximize the probability of detection because the cost of life is assumed invaluable (Maul and Kopasakis 2008). Papadimitriou et al. (Papadimitriou and Beck 2000) have proposed sensor placement design by minimizing entropy focusing on the applications of structural model updating. Similarly, Udwadia (Unwadia 1994) and Basseville (Basseville and Benveniste 1987) have used the Fisher information matrix to maximize the parameter identification through SHM.

In many SHM systems used for large civil infrastructure such as the application area considered in this chapter, the primary goal of the SHM system is to minimize the long-term monitoring and maintenance costs (Parno and Connor 2018). In this context, optimal sensor design is tied to the rate of incorrect decisions (e.g., the probability of false alarms for a binary decision case) and the costs/consequences associated with those wrong decisions. Flynn and Todd (Flynn and Todd 2010; Todd and Flynn 2011) first introduced Bayesian experimental design (Chaloner and Verdinelli 1995) by minimizing expected loss or risk (also termed as Bayes risk) incurred as a consequence of making a decision (choosing optimal design in their case). They demonstrated it in an ultrasonic guided wave sensor design problem. Bayes risk is proposed as a suitable choice of the objective function because it considers the costs of consequences associated with making decisions (parameterized by the monitoring design variables), known as extrinsic costs, as well the

cost of sensors system design, deployment, and maintenance, known as intrinsic costs. The optimal sensor design essentially demands arriving at the sensor network design that minimizes the expected losses incurred as a consequence of making a decision, or equivalently, that minimizes the losses in an average sense (the idea adapted from Bayesian experiment design and Bayesian decision theory).

Because the monitoring process is subject to many sources of noise and variability, structural state determination is inherently stochastic. Thus, the goal is to arrive at a sensor network design that considers all the uncertainties and the consequences of inferring the structural state using the data gathered by the design of interest. The prediction of the structural state bears a cost/risk. For example, if the predicted structural state is not the same as the true state (unknown), there will be an associated penalty in the form of planned or unplanned maintenance costs, operational availability losses, or even structural failure costs. The design that leads to the least expected loss/risk due to making structural state decisions is the optimal design. Since I am operating in an uncertain domain, arriving at an optimal design that minimizes Bayes risk is the best one can do. In this chapter, Bayes risk will be used as an objective function in a strain-based measurement sensor optimization problem; however, I note that the framework proposed herein can also be applied to any SHM by formulating an appropriate form of Bayes risk constrained to that particular problem.

A common approach to function optimization includes iteratively evaluating the optimal value of the function locally guided by the steepest gradient descent. This approach has been used in machine-learning (Bottou 2010) and in developing an optimal sensor network (Ram and Nedic 2007). Akbarzadeh (Akbarzadeh and Lévesque 2014) used a gradient descent algorithm in sensor optimization by deriving derivatives at each step, which requires less computational effort.

However, in many problems, the exact analytical derivatives are not available. Agarwal (Agarwal and Ezra 2009) used the greedy algorithm to find a minimum number of sensors for covering a 2-dimensional space. The main shortcoming of the greedy algorithm is that it chooses the ``current best'' at each step so that it can easily converge to a local optimum instead of a global optimum. Heuristic algorithms are also widely used in the existing literature; for example, Jin (Jin and Zhou 2003) used a genetic algorithm to minimize the communication distance of sensors, while Yi et al. (Yi and Li 2011) utilized a genetic algorithm to obtain optimal sensor placement for a high-rise building monitoring system. However, the main drawback of these optimization strategies is that they must run many samples, and hence are computationally expensive to arrive at the global optimal value of the objective function. In complex large-scale civil structures SHM applications, the sensor design space is potentially colossal. This coupled with the fact that obtaining and evaluating Bayes risk is computationally expensive and I do not have its derivatives, Bayesian optimization is the most suitable technique to apply. Bayesian optimization can optimize objective functions parameterized by high-dimensional design spaces with relatively low computational effort (Jones and Schonlau 1998; Hu and Nannapaneni 2017; Frazier 2018). This chapter details a general Bayesian optimization framework for obtaining the optimal sensor network design for SHM applications by using Bayes risk as the objective function. I address three implementation-based challenges: (1) Bayesian calibration of the discrete parameters defining the damage state; (2) the expensive evaluation of Bayes risk; and (3) the global optimization of an extremely high-dimensional design space informing Bayes risk.

After laying the theoretical foundation of Bayes risk and Bayesian optimization, I detail the general framework. I believe that the best possible way to demonstrate our sensor optimization framework is through an example that by itself doesn't pose tremendous uninformative challenges,

is relatively simple to conceive, and has all the essential elements to utilize and showcase the optimization framework presented. To this end, I apply it to an example problem concerning the boundary condition detection state of a beam structure. This example was considered because it covers a broader spectrum of detection and inference-type problems that are common in SHM. One instance of a resembling but slightly different problem is that of contact loss detection between the quoin blocks of the miter gate. Moreover, the demonstration example is sufficiently complex to highlight the sensor optimization framework and the associated challenges while not inducing computational complexities and costs associated with more complex structural scenarios.

The rest of the chapter is arranged as follows. Section 2.3 briefs the concepts of the Bayes risk functional and explains the four steps of the general sensor optimization framework. Section 2.4 describes the demonstration problem and details the associated Bayes risk functional, followed by Section 2.5 that investigates three different approaches to evaluate the Bayes risk. Section 2.6 discusses the optimal sensor placement design using Bayesian optimization in detail and presents the algorithm used. After a general discussion on Bayesian optimization, the remaining part of Section 2.6 discusses the effect of the initial design assumption and the approaches used to evaluate the Bayes risk on the final optimal sensor design for the demonstration problem. I present three methods to evaluate the Bayes risk functional: a sampling-based method, mean-value approximation, and univariate dimensional reduction with Gauss-Hermite quadrature. The sampling-based method yields the most accurate Bayes risk if large sample size is considered. Consequently, the sampling-based method suffers from a high computational cost. This drawback makes the sampling-based method unsuitable for sensor placement optimization. Secondly, as is the inherent case with any Monte-Carlo based approach, the values of Bayes risk obtained from the sampling-based technique change as a different set of samples are chosen. The other two

methods overcome these challenges and disadvantages. However, the mean-value approximation of the Bayes risk does not yield accurate values. The univariate dimensional reduction with Gauss-Hermite quadrature is fairly accurate and has acceptable computational speed. Therefore, I use this third approach to evaluate Bayes risk, and then Bayesian optimization follows. Finally, Section 2.7 concludes the chapter and Section 2.8 provides a preview to the next chapter.

2.3. Bayes risk and general optimization framework

I first present some preliminary definitions and notations. The real number space is represented by $\mathbb{R}$. A random variable X is a real-valued function defined on a discrete or a continuous sample space S_X and the measurement space Ω_X such that $X: S_X \longrightarrow \Omega_X \subset \mathbb{R}$. Let x represent the realization of the random variable X, such that $x \in \Omega_X$. The probability density function and the cumulative density function is represented by $f_X(x \in \Omega_X)$ and $F_X(x \in \Omega_X)$. The expected value of a function $g(x)$ is denoted by $E_X[g(x)]$. Lastly, a random variable X following Gaussian distribution, with the mean μ_x and standard deviation σ_x is denoted by:

$$\begin{gathered} f_X(x) = \frac{1}{\sigma_x}\phi\left(\frac{x-\mu_x}{\sigma_x}\right) \\ F_X(x) = \Phi\left(\frac{x-\mu_x}{\sigma_x}\right) \\ X \sim N(\mu_x, \sigma_x) \end{gathered} \tag{2.1}$$

2.3.1 Bayes risk for decision-making

For a problem concerning Bayesian decision-making, the goal is to arrive at a decision that minimizes the expected risk (also referred to as Bayes risk in this chapter) or expected loss. The idea is that I have information about the system in the form of observable measured data. The goal

is to learn the behavior of the system from the data (called training) and then use the learned model to predict the *outcome*. Primarily, the outcomes can be categorized by detection, classification, and regression. For instance, detecting if the structure is *damaged* or *not damaged* given the measured strain gauge data is an example of detection; grouping the raw grades of the class into the letter grades is an example of classification; developing a *digital twin/surrogate* of a non-linear system is an example of regression. The goal is to make a decision that minimizes the expected loss or risk that arises associated with making a decision (every action/decision has a consequence). Therefore, the optimality criterion used in this chapter is the expected loss/risk, which is also referred to as Bayes risk functional and is a problem-dependent quantity. The strong similarity of Bayes risk with the action functional in variational structural mechanics is not surprising.

I focus on the classification type problem of which detection is a special case. Let Ω_X represents the measurement space, Ω_Y represents the true state (or outcome) space with M classes (for detection as defined above, $M = 2$), such that the feature/measurement/observable is $x \in \Omega_X$, true outcome (or decision) is $y \in \Omega_Y = \{y_0, y_1, \dots, y_{(M-1)}\}$, and the predicted outcome (or decision) is $g(x) \in \Omega_G = \{g_0, g_1, \dots, g_{(M-1)}\}$, where $i = \{0,1, \dots, (M-1)\}$. Let X, Y, and G represent the random variables corresponding to the uncertain measurement space, the true outcome, and the predicted outcome respectively. Note that $\Omega_G \equiv \Omega_Y$, and the two representations of outcome space is to distinguish between the true (but unknown) and the predicted states. In fact, $g(x)$ represents the trained model. For instance, in the case of a simple detection problem, y_0 denotes a true damaged state; and g_0 is a prediction of a damaged state. Bayes risk is designed such that it minimizes the effects of incorrect decisions. This is done by incorporating a loss function $L(g(x), y): \Omega_G \times \Omega_Y \longrightarrow \mathbb{R}$. It defines the consequence-cost of deciding the outcome to be $g(x)$ when y is the true outcome. Since our goal is to minimize losses, the Bayes risk (objective

functional) is defined as the expected loss, averaged over all possible (noisy) measurements and the true state y_i. Since the goal in this section is to estimate the state $g(x)$ when the true state is y using the measurement x, the Bayes risk Ψ_{state} is a function of the predicted outcome/state $g(x)$ which in turn is a function of newly acquired data $x \in \Omega_X$. The Bayes risk is then defined as:

$$\begin{aligned} \Psi_{\text{state}}(g(x)) &= E_{XY}[L(g(x), y)] = \sum_{i=0}^{M-1} \int_{\Omega_X} f_{XY}(x, y_i) L(g(x), y_i) \mathrm{d}x \\ &= \sum_{i=0}^{M-1} \int_{\Omega_X} L(g(x), y_i) P_{X|Y}(x \mid y_i) P_Y(y_i) \mathrm{d}x \end{aligned} \tag{2.2}$$

Bayes risk can also be written in terms of *conditional risk* $R_{\text{state}}(g(x))$, conditioned on measurement x, as:

$$\Psi_{\text{state}}(g(x)) = E_X[R_{\text{state}}(g(x))] = \int_{\Omega_X} f_X(x) R_{\text{state}}(g(x)) \mathrm{d}x, \text{ where,} \tag{2.3 a}$$

$$R_{\text{state}}(g(x)) = \sum_{i=0}^{M-1} L(g(x), y_i) f_{Y|X}(y_i \mid x) \tag{2.3 b}$$

The conditional risk is defined as the expected loss averaged over all possible true states and considering (or conditioned on) fixed measurement x. The optimal decision is the one that minimizes the expected loss, or,

$$\mathfrak{g}(x) = \arg\min_{g(x)} R_{\text{state}}(g(x)) \in \Omega_G \tag{2.4}$$

The Bayes risk $\Psi_{\text{state}}(g(x))$ defined in this section is an objective functional that is used to optimally predict the most likely state $g(x)$ given the measurement x (hence the subscript state in Ψ_{state} and R_{state}). However, among possible choices of an SHM system design, every design will

predict a unique state for a given set of measurements (obtained by Eq. (2.4)). Inversely, the predicted outcome is dependent on the sensor design. In the next section, I consider the problem of design selection that would warrant a different Bayes risk functional. The goal is to pick the design that leads to the least erroneous state estimation (the optimality criterion is defined in the next section). Unlike the problem of state-estimation, where Bayes risk was a function of the estimated state $g(x)$, the Bayes risk for the design selection, represented by $\Psi_{\text{design}}(e)$, will be a function of design e.

2.3.2 Bayes risk for design selection and optimal sensing framework

The primary goal of this chapter is to arrive at an optimal sensing design, and Bayes risk can accommodate this notion. Let Ω_E represent the design/experiment space, such that $e \in \Omega_E$ represents a design realization. Every design e yields different measurement data $x_e \in \Omega_X$, and corresponding likelihoods $f_{X_e|Y}(x_e \mid y)$. Here, Ω_{X_e} represents the measurement space for the design e, and X_e denotes the corresponding random variable. Let $\mathfrak{g}(x_e; e)$ represents the optimally estimated state obtained using Eq. (2.4) for the measurement x_e corresponding to the design $e \in \Omega_E$. Therefore, the decision $\mathfrak{g}(x_e; e)$ is also design-dependent. In other words, I now care about choosing the design with the least error/deviation in the decision $\mathfrak{g}(x_o; e)$ relative to the true value y. Equation Eq. (2.4) can be used to arrive at the optimal state $\mathfrak{g}(x_e; e)$ for a given design e; or equivalently, for each design e, a threshold (or a classifier) can be established using Eq. (2.4) in the measurement space Ω_{X_e} that helps classify each realization of measurement x_e into the optimal state $\mathfrak{g}(x_e; e)$. Therefore, Eq. (2.4) establishes a mapping between the continuous measurement parameter x_e and the decision (discrete in case of detection problem) $\mathfrak{g}(x_e; e)$. This allows me to

write the Bayes risk for each design e focusing on minimizing the deviation of the predicted outcome $\mathfrak{g}(x_\rho; e)$ relative to the true outcome y as:

$$\Psi_{\text{design}}(e) = E_{GY}[L(\mathfrak{g}(x_e; e), y)] \quad = \sum_{i,j=0}^{M-1} L(g_i, y_j) f_{G|Y}(\mathfrak{g}(x_e; e) = g_i \mid y_j) f_Y(y_j) \qquad (2.5\text{ a})$$

$$f_{G|Y}(\mathfrak{g}(x_e; e) \mid y) = \int_{\Omega_{X_e}} f_{G|X_e}(\mathfrak{g}(x_e; e) \mid x_e) f_{X_e|Y}(x_e \mid y) \mathrm{d}x_e. \qquad (2.5\text{ b})$$

For a design e, the true state y, and the observed measurement x_e, the estimated state $\mathfrak{g}(x_e; e)$ is one of the states in the set Ω_G. Equivalently, for a design e, the true state y, and the measurement x_e every state $g_i \in \Omega_G$ has a likelihood probability of $f_{G|Y}(\mathfrak{g}(x_e; e) = g_i \mid y)$ to be selected as the optimal estimated state $\mathfrak{g}(x_e; e)$. The Bayes risk functional $\Psi_{\text{design}}(e)$ defined in Eq. (2.5a) calculates the expected value of loss (or risk) considering all the possibilities of the estimated states $\mathfrak{g}(x_e; e) \in \Omega_G$ and considering all the possible true states $y \in \Omega_Y$. Minimizing this function yields a design that leads to the best prediction of the state. I will adapt the Bayes risk defined in Eq. (2.5a) focusing on a detection-type problems common in SHM. The following paragraphs detail the generalized step-by-step procedure for the proposed optimal sensor framework.

Step 1: Problem description

The first step involves a well-defined problem description. I need to decide our decision and the true space (Ω_G, Ω_Y), or what needs to be detected, and what is its true condition/state, respectively Both, the decision and the true space refer to the state of the structure defined accordingly. A discrete decision space in SHM answers the question "Is a structure critically damaged or not?", e.g., whether a bolted assembly is at design torque or not; a continuous decision

space, such as where regression may be utilized, might be to infer crack length. In this chapter, I focus on discrete decision spaces, but at the same time note that the framework can easily be extended to the continuous case. In theory, that would essentially replace the summation over the decision space in Eq. (2.5a) by an integral. In practice, the continuous decision space can be discretized by identifying the mutually exclusive and exhaustive subsets with a decision state. For instance, a corroded surface area $< 10\%$ of a bridge girder might be classified as not damaged, $10\% - 30\%$ can be identified as moderately damaged, and $> 30\%$ can be considered as severely damaged. Secondly, I need to define the measurement of the observable quantity using which the structural state is inferred. The features used to infer the structural state can be extracted from the measured quantity, although the measured data itself can be the feature. The measurement space Ω_{X_e} essentially is the space from which the decision is directly inferred; therefore, in current content, the measurement space is the feature space. Once know what needs to be measured (for example, strain values), the design space Ω_E follows (e.g., all the possible arrangements of a strain gauge network). Therefore, the problem description consists of defining the decision space, the true state space, the measurement (or feature) space, and the design space.

Step 2: Definition of the design dependent Bayes risk functional

For a simple classification problem, Eq. (2.5) represents the Bayes risk functional. However, as the complexity of the problem evolves, suitable adjustments to the Bayes risk should be made For instance, in our demonstration problem described in detail later, where I am focusing on the problem of multiple load path changes through boundary connections, the space of collective true states of the springs (denoted by Ω_A) becomes important. Second, in the case of collective decisionmaking problems, some decisions are more preferred or weighed than others.

To incorporate such situations, I can assign weights to each of these decisions. Third, unlike the Bayes risk expression in Eq. (2.5) that incorporates the cost of making a decision or extrinsic cost only, the intrinsic costs (like the sensor deployment and maintenance costs) must be included in SHM applications. All these considerations lead to Bayes risk $\Psi_{\text{design}}(e)$ to bear a form defined in Eq. (2.9), with the extrinsic cost defined in Eq. (2.10) of Section 2.4.2.

Step 3: Evaluation of the design-dependent Bayes risk functional

For a Bayes risk of a simple classification or detection type problems represented in Eq. (2.5a), the first challenge is to evaluate the three probabilities present in Eq. (2.5): $f_Y(y_j)$, $f_{X_e|Y}(x_e \mid y)$, and $f_{G|X_e}(\mathfrak{g}(x_e; e) \mid x_e)$. The quantity $f_Y(y_j)$ represents the prior probability of the true state, and in absence of any information can be assumed as 0.5 for a detection type problem. The likelihood $f_{X_e|Y}(x_e \mid y)$ is obtained using either a physics-based model or a digital twin. The posterior of the decision given the measurement $f_{G|X_e}(\mathfrak{g}(x_e; e) \mid x_e)$ is more involved to evaluate. For a binary detection problem it can be written using the law of total probability as:

$$f_{G|X_e}(g_i \mid x_e) = \sum_{j=1}^{2} f_{G|Y}(g_i \mid y_j) f_{Y|X_e}(y_j \mid x_e) \tag{2.6}$$

The probability of making a decision given the true state $f_{G|Y}(g_i \mid y_j)$ depends on the detection threshold evaluated for each design case using Eq. (2.4). The quantity $f_{Y|X_e}(y_j \mid x_e)$ is anti-causal and can be evaluated using Bayes theorem as:

$$f_{Y|X_e}(y_j \mid x_e) = \frac{f_{X_e|Y}(x_e \mid y_j) f_Y(y_j)}{f_{X_e}(x_e)} \tag{2.7}$$

The second difficulty in obtaining Bayes risk is to evaluate the integral in Eq. (2.5b). To approximate the integral, I first change the variable of the integral from the measurement space to the uncertain input space. For instance, in our demonstration problem, the load, its location, and the noise in the strain values are uncertain, causing randomness in the strain measurement. I realize that a unique value of the load, its location, and the noise in the strain gauge give a unique realization of the strain measurement. This allows me to change the variables of integration as defined in Eq. (2.17). The integral can then be numerically approximated. I discuss three different approaches to evaluate the integral in Section 2.5.4.

Step 4: Obtaining the optimal sensor design using Bayesian optimization

Once the problem is well defined (step 1) and the associated Bayes risk is obtainable (steps 2-3). the question that I intend to answer for optimal sensor design is: "Given $\Omega_G, \Omega_Y, \Omega_X$, and Ω_E and *given an assumed initial design* e_0*, what is the design* $e^* \in \Omega_E$ that minimizes the Bayes risk $\Psi_{\text{design}}(e)$?"

I very briefly detail the sensor optimization algorithm, which will be explained in great depth in section 2.6. I start with an initial design e_0 consisting of N_0 number of sensors. To obtain the optimal design e_1 with $(N_0 + 1)$ sensors, I search the entire design space for the $(N_0 + 1)^{\text{th}}$ sensor location. The $(N_0 + 1)^{\text{th}}$ sensor location that maximizes the acquisition function constitutes the next additional sensor. In this chapter, I use expected improvement (Mockus and Tiesis 1978; Jones and Schonlau 1998) as the acquisition function. Similarly, I repeat the optimization process to arrive at the optimal design $e_{n_{as}}$ consisting of $N_0 + n_{\text{as}}$ sensors (or n_{as} number of additional sensors relatively to the initially assumed design e_0). Finally, I pick $e^* = \arg\min_{e_{n_{\text{as}}}} \Psi_{\text{design}}\left(e_{n_{\text{as}}}\right)$ as the most optimal design, where $\Psi_{\text{design}}\left(e_{n_{\text{as}}}\right)$ represents the Bayes

risk associated with the design $e_{n_{\text{as}}}$. Figure 2.1 illustrates the pipeline of the proposed Bayesian optimization framework.

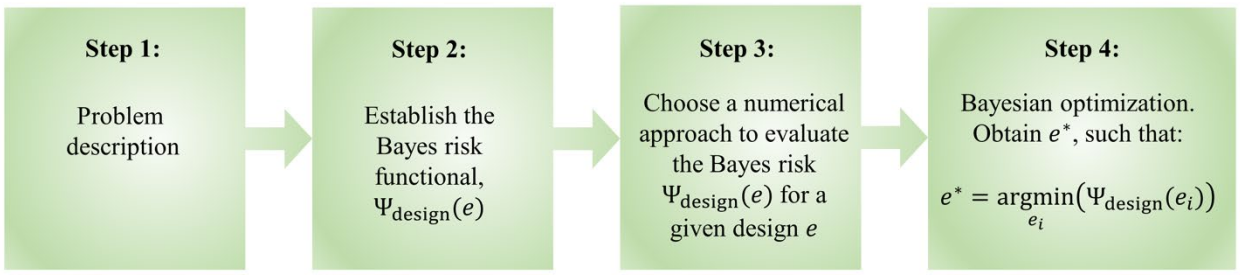


Figure 2.1: Bayesian optimization framework for optimal sensor network design

Section 2.4 deals with the description of a demonstration problem and derives the associated Bayes risk (steps 1 and 2). Section 2.5 discusses the approaches to evaluate the Bayes risk pertaining to the demonstration problem (step 3). Finally, Section 2.6 details the Bayesian optimization algorithm for optimal sensor placement 1 for general detection-type problems and discusses the results concerning the demonstration problem.

2.4. Demonstration problem description and the associated Bayes risk

As I mentioned in the introduction, the primary motivation for choosing the following example problem as a case study to demonstrate this framework is that it resembles in behavioral characteristics a typical detection-type problem in SHM that has a discrete decision space: the loss of contact in the quoin block of a miter gate. As discussed in section 2.3.2, even the continuous decision space can be reasonably broken down into a rather more convenient discrete decision space. Hence, the presented framework is also suitable for problems involving crack propagation, corrosion, weld defect growth, etc. To demonstrate the framework, I consider a beam modeled by 2D shell elements and focus on detecting the state of the boundary modeled using connecting

springs. This problem is complicated enough to highlight the Bayesian optimization framework for sensor placement and undemanding enough to implement the algorithms with a lower computational cost. The figure below shows similar types of problems in SHM where the presented sensor-design framework can be extended (although each specific problem would require its own carefully considered Bayes risk functional).

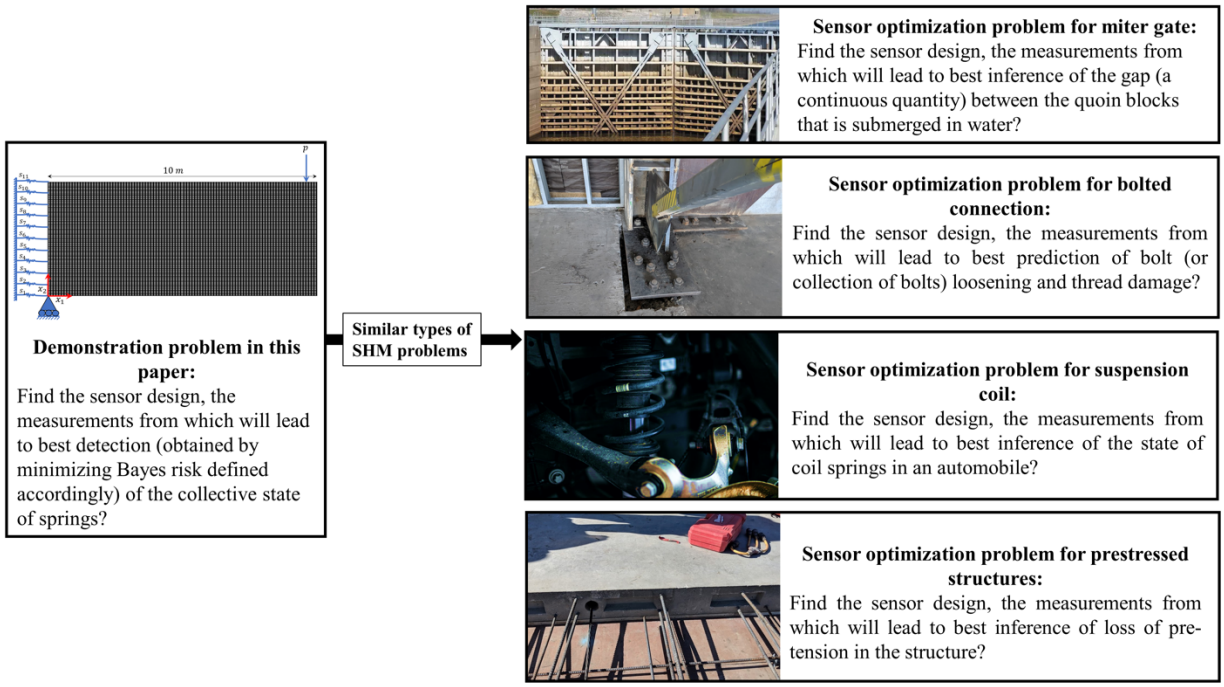


Figure 2.2: Design optimization problems in SHM similar to the demonstration example presented in this chapter

2.4.1 Demonstration problem description

The demonstration problem consists of a cantilever beam supported by a roller on the left end and a free boundary on the right end. The Young's modulus of the beam is $2.1 \times 10^9 \mathrm{Nm}^{-2}$. There exist 11 wall-to-beam springs with the stiffness $10^7 \mathrm{Nm}^{-1}$ connected to the left side of the 2D shell element providing structural stability as shown in Fig. 2.3. The finite element model for the beam was build in OpenSees (McKenna 2011) with quadrilateral meshing. The entire beam was meshed finely to 22500 elements to capture accurate strain responses, particularly at the left

edge of the beam, where the springs are attached. The horizontal axial strain of the element is considered to be the strain gauge measurement. Therefore, there are 22500 possible strain gauges (with horizontal orientation). In the general case, the strain gauge may be discrete or continuous and can have any orientation (Chadha and Todd 2019).

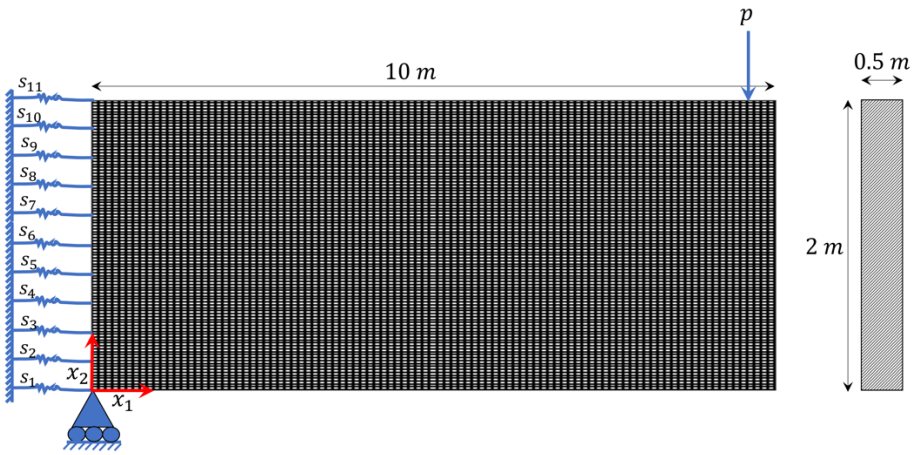


Figure 2.3: Schematic diagram of the 2D beam modeled by 2D shell elements

The problem statement is as follows: I aim to arrive at the best possible *sensor placement* design $e \in \Omega_E$, where Ω_E is the design space, such that the existence of the springs on the left of the beam can be most optimally predicted ("detected"), given that the magnitude of the load p and its location $p_{\text{loc}} \in [0,10]$ are uncertain. I also assume that the strain gauge readings are noisy. By sensor placement design, I mean the arrangement of the strain gauges (including the number used and their locations). To simplify the problem further, I fix the top six springs. Hence, I need not predict their existence. The goal is, therefore, to predict the existence of the remaining five springs s_1, s_2, s_3, s_4, and s_5.

2.4.2 Bayes risk for the optimal sensor placement

As discussed before, for each spring, our predicted decision space Ω_G and the true outcome space Ω_Y consist of two possible outcomes, such that $\Omega_G = \{g_0, g_1\}$ and $\Omega_Y = \{y_0, y_1\}$, where

$$\begin{aligned} &g_0\text{: Prediction is that the spring exists;} \\ &g_1\text{: Prediction is that the spring does not exists;} \\ &y_0\text{: True state is that the spring exists;} \\ &y_1\text{: True state is that the spring does not exists.} \end{aligned} \tag{2.8}$$

Recall that the problem has 5 critical springs s_n, with $n \in \{1,2,3,4,5\}$. For n^{th} spring, I denote the predicted state by $g_i|_{n^{\text{th}}spring} = g_{ni}$, and the true state by $y_j|_{n^{\text{th}}spring} = y_{nj}$, with $i, j \in \{0,1\}$, such that $g_{ni} \in \Omega_G$ and $y_{nj} \in \Omega_Y$. Since there are five springs, with each of them existing in the either of two possible states $\{y_0, y_1\}$, there are 2^5 possibles states of the springs collectively. I define the *collective true state* of the five springs by a set of vectors Ω_A, such that $A_k = [\bar{y}_1, \bar{y}_2, \bar{y}_3, \bar{y}_4, \bar{y}_5] \in \Omega_A$ and $\bar{y}_n \in \{y_{n0}, y_{n1}\} = \Omega_Y$, with $k \in \{1,2,3,\dots,2^5\}$, and $f_A(A_k) = 2^{-5}, \forall k$. Similarly, I define the *collective prediction state* of the five springs by a set of vectors Ω_S, such that $S_k = [\bar{g}_1, \bar{g}_2, \bar{g}_3, \bar{g}_4, \bar{g}_5] \in \Omega_S$ and $\bar{g}_n \in \{g_{n0}, g_{n1}\} = \Omega_G$, with $k \in \{1,2,3,\dots,2^5\}$. I define A and S as the random variables corresponding to the space Ω_A and Ω_S respectively, such that, A_k and S_k represents the realizations of A and S respectively.

For the considered sensor placement design e, the Bayes risk specific to this problem consists of *intrinsic* and *extrinsic* costs. The intrinsic cost $\Psi_{\text{design-in}}(e)$ includes the expenses associated with the sensor installation and maintenance. On the other hand, the extrinsic cost $\Psi_{\text{design-ex}}(e)$ accounts for the cost of making a decision and the design selection. It resembles the form of Eq. (2.5). Therefore, the total Bayes risk is defined as:

$$\Psi_{\text{design}}(e) = \Psi_{\text{design-in}}(e) + \Psi_{\text{design-ex}}(e) \tag{2.9}$$

I now focus on constructing the extrinsic cost $\Psi_{\text{ex-design}}(e)$. I denote the cost $L(g_{ni}, A_k)$: $\Omega_G \times \Omega_A \longrightarrow \mathbb{R}$, defining the regret of making the decision g_i for the strain gauge n, when the true *collective state* is A_k. I further define the cost function: $C_{nij} = L\big(g_{ni}, A_k(n) = y_{nj}\big)$, and assume that it is independent of the selected design e. For the fixed spring n, the cost $C_{nij} = C_{ij}$ is defined as:

Table 2.1: Cost Function

		True State	
		y_0	y_1
Predicted State	g_0	C_{00}	C_{01}
	g_1	C_{10}	C_{11}

The cost values C_{ij} assign penalty/losses to each predicted state g_i when the true state is y_j. Consider the case where the spring exists in reality, i.e. the state y_0. If the prediction is correct, i.e. g_0, there is no loss since no action is warranted, or $C_{00} = 0$. On the contrary, if the prediction is g_1 (spring not existing), then the engineers would decide to perform unnecessary inspection and service leading to a loss of C_{10}. However, as a consequence of this incorrect decision, there would be no major failure since the spring exists in reality. Similarly, consider the case where the spring does not exist, represented by y_1 (the boundary is actually damaged). If I estimate (from strain gauge data) that spring does not exist (correct decision), then there will be cost (denoted by C_{11}) incurred to inspect and repair the spring (or the boundary). However, if the spring is predicted to exist, when it is non-existent (incorrect decision), it can lead to the most expensive mistake since the structure can potentially fail if appropriate actions are not taken. This leads to the maximum

cost of C_{01}. For simplicity, I assume $C_{01} = 200$ dollars and assume other costs to be fraction of C_{01}, such that $C_{10} = 0.1C_{01}$ and $C_{11} = 0.25C_{01}$. Individual costs C_{ij} are defined in Table 2.2.

Table 2.2: Assumed cost function values

Cost	Definition	Breakdown	Assumed Dollars
C_{00}	True positive cost	Zero cost, as no action is needed	0
C_{10}	False positive cost	Cost due to service and inspection	20
C_{01}	False negative cost	Cost due to service, failure, and replacement	200
C_{11}	True negative cost	Cost due to service and repair	50

In cases of problems involving multiple decisions, there may be instances where the consequences of making some decisions are more important or weighed for some cases than the others (like some springs being more important than the others). To incorporate these kinds of situations, I assume that the top two springs (s_4, s_5) are more important than the bottom three (s_1, s_2, s_3). I incorporate this assumption by assigning weights to each of these springs as $w = [1,1,1,1.5,1.5]$. The goal is to define the extrinsic Bayes risk functional, considering the importance of the consequence of decisions associated with each spring, as a quantity that minimizes itself with the most optimal sensor arrangement e. Along the similar lines of Eq. (2.5), the extrinsic Bayes Risk is defined as:

$$\begin{aligned}\Psi_{\text{design-ex}}(e) &= \sum_{n=1}^{5} w_n E_{GA}\big(L(g_{ni}, A_k)\big) \\ &= \sum_{n=1}^{5} w_n \sum_{k=1}^{32} \sum_{i=0}^{1} L(g_{ni}, A_k) f_{G|A}(g_{ni} \mid A_k) f_A(A_k)\end{aligned} \tag{2.10}$$

In the equation above, $f_A(A_k) = 2^{-5}$ is the prior probability of the collective state of springs being A_k. Secondly, $x_e \in \Omega_{X_e}$ represents the measured/observed data. For instance, x_e can be strain measurements for any design e. For the n^{th} spring, the quantity $f_{G|A}(g_{ni} \mid A_k)$ represents the probability of predicting the state g_i for the spring n, when the true collective state is A_k. This is a difficult bit to evaluate, and like Eq. (2.5b) can be broken down into more manageable pieces:

$$f_{G|A}(g_{ni} \mid A_k) = \int_{\Omega_{X_e}} f_{G|X_e}(g_{ni} \mid x_e) \cdot f_{X_e|A}(x_e \mid A_k) \mathrm{d}x_e \tag{2.11}$$

The likelihood $f_{G|X_e}(g_{ni} \mid x_e)$ depicts our belief of deciding the state of the spring n to be g_i for a given measurement $x_e \in \Omega_{X_e}$. I obtain the likelihood using Bayesian inference (detailed in the next section).

2.5. Evaluating Bayes risk for a fixed design

To perform Bayes optimization that yields the most optimal sensor placement design, I will have to start with a design that evolves/improves with every iteration of the optimization process At every iteration, for a suggested design e, I need to obtain the Bayes risk defined in Eq. (2.10). Therefore, in this Section, I detail on calculating the Bayes risk for a design e consisting of 30 sensors, the arrangement of which was obtained using Latin Hypercubic Sampling (LHS) technique (McKay and Beckman 2000). The first step of the process is to evaluate the likelihood of making a decision given the measurement, $f_{G|X_e}(g_{ni} \mid x_e)$.

2.5.1 Analytical formulation to obtain the likelihood

The goal is to obtain $f_{G|X_e}(g_{ni} \mid x_e)$. Recall that $S_k \in \Omega_S$ defines the collective prediction state of the springs. I can therefore write:

$$f_{G|X_e}(g_{ni} \mid x_e) = \sum_{k=1}^{32} f_{G|S}(g_{ni} \mid S_k) \cdot f_{S|X_e}(S_k \mid x_e) \tag{2.12}$$

Note that:

$$f_{G|S}(g_{ni} \mid S_k) = \begin{cases} 1 & \text{if } S_k(n) = g_{ni} \\ 0 & \text{otherwise} \end{cases} \tag{2.13}$$

To evaluate the distribution $f_{S|X_e}(S_k \mid x_e)$ in Eq. (2.12), I assume that to make a decision given the measurement data, I have a non-conflicting threshold or boundary to make a prediction of the spring state, such that $f_{Y|X_e}(y_{n0} \mid x_e)$ and $f_{Y|X_e}(y_{n1} \mid x_e)$ do not intersect. This also implies $f_{G|X}(g_{ni} \mid x_e) = f_{Y|X_e}(y_{ni} \mid x_e)$. With this assumption, I have $f_{S|X_e}(S_k \mid x_e) = f_{A|X_e}(A_k \mid x_e)$.

I note that the quantity $f_{A|X_e}(A_k \mid x_e)$ is anti-causal, as it is asking for the true state of the springs when the measurement x_e is given. I use Bayes theorem to write it in a more desirable and causal form:

$$f_{A|X_e}(A_k \mid x_e) = \frac{f_{X_e|A}(x_e \mid A_k) \cdot f_A(A_k)}{\sum_{l=1}^{32} f_{X_e|A}(x_e \mid A_l) \cdot f_A(A_l)} = \frac{f_{X_e|A}(x_e \mid A_k)}{\sum_{l=1}^{32} f_{X_e|A}(x_e \mid A_l)} \tag{2.14}$$

The likelihood $f_{X_e|A}(x_e \mid A_k)$ remains to be evaluated for all k. For a fixed collective spring state A_k, and a design e with $N_{\mathrm{sg}}(e)$ number of strain gauges, $f_{X_e|A}(x_e \mid A_k)$ is the joint distribution of the $N_{\mathrm{sg}}(e)$ strain measurements.

Although the true strain values of different strain-gauges are related due to the underlying physics, the noise in strain gauge measurements is taken to be statistically independent. I also assume (for modeling purposes) that the randomness in the strain gauge readings, primarily due to noise and uncertainties in loading, follows a Gaussian distribution. Let x_{en} represent the observed strain measurement in n-th strain-gauge of the design e, such that $x_e = \{x_{en}\}$ with $n \leq N_{\text{sg}}(e)$. For the selected spring state A_k and the design e, if x_{en}, μ_{en}, and σ_{en} represent the measurement of the strain gauge n (a random variable), its mean value, and the standard deviation respectively, I can write the following:

$$f_{X_e|A}(x_e \mid A_k) = \prod_{n=1}^{N_{\text{sg}}(e)} \frac{1}{\sqrt{2\pi\sigma_{en}^2}} \exp\left(-\frac{1}{2}\left(\frac{x_{en} - \mu_{en}}{\sigma_{en}}\right)^2\right) = \prod_{n=1}^{N_{\text{sg}}(e)} \frac{1}{\sigma_{en}} \phi\left(\frac{x_{en} - \mu_{en}}{\sigma_{en}}\right) \quad (2.15)$$

This gives all the pieces to obtain $f_{G|X_e}(g_{ni} \mid x_e)$. I obtain the measurement data x_e using *Finite Element Model* (FEM) developed using OpenSees (McKenna 2011) or using a surrogate model developed using Gaussian Process Regression (GPR) (Frazier 2018).

I note that obtaining the likelihood $f_{X_e|A}(x_e \mid A_k)$ and the posterior $f_{A|X_e}(A_k \mid x_e)$ is not complicated for the chosen demonstration problem. Since our emphasis is more on the optimization framework, for simplicity I have assumed strain values to be uncorrelated and evaluation of the posterior can be done analytically as the decision space is discrete. However, for more complicated problems with correlated measurement values and continuous decision space, evaluation of the likelihood and the posterior will be more involved. For instance, in such cases, I use numerical techniques like *Markov Chain Monte Carlo* (MCMC), Sequential Monte Carlo (SMC) (Ramancha and Astroza 2020), or other methods to evaluate posterior.

Remark 2.1: I note that the true strain values of different strain gauges are correlated (or functionally related) by the underlying physics of the problem. That is, each of the strain gauge readings embeds some information about the state of the structure. In this chapter, I use the Finite Element Model (FEM) as the ground truth (discussed more in the next section). This implies that the strain values obtained from the FEM are treated as the actual/true strain measurement that is impossible to be known since there will always be noise in observed strain readings. For a given load condition, I have a deterministic prediction of the mean value of the strain reading using the FEM model (or the respective digital surrogate constructed using strain data obtained from the FEM) which is considered to be the ground truth. However, the noise in the various strain gauge reading is statistically independent since the noise pertains to a given strain gauge itself. In this chapter, I have assumed a Gaussian structure to the noise.

2.5.2 Finite element and Surrogate model

Section 2.3 details the finite element model of the structure of interest built using shell elements. I consider that the loading in the beam is uncertain, such that, the concentrated load $p \in \Omega_P$ and its locations $p_{\text{loc}} \in \Omega_{P_{\text{loc}}}$ is represented by the random variables P and P_{loc} respectively. I run the FE model for 5000 samples of random input data consisting of seven quantities: the true state of the springs $A_k \in A$ (consisting of states of 5 springs), the magnitude of the load $P \sim N(\mu_p = 1000 \text{ newton}, \sigma_p = 100 \text{ newton})$, and the location of the load $P_{\text{loc}} \sim HN(\mu_{\text{loc}} = 10 \text{ m}, \sigma_{\text{loc}} = 1 \text{ m})$ Here, $HN(\cdot,\cdot)$ represents the half normal distribution. For each input sample, I obtain 22,500 strain responses. From here on Ω_Z represent the space of input sample, such that $z \in \Omega_Z$.

I would like to note that in many machine learning problems, physics-based models are unavailable, and the engineers must rely on fitting a numerical model using the data obtained from the experiments. In our case, I obtain the data from the finite element model, which I consider as "ground truth". Although I have the luxury of utilizing the finite element model, the computational cost is restrictive, and therefore, not the best option with which to carry out Bayesian optimization For Bayesian calibration, metamodels or surrogate models are preferable, e.g., Support Vector Regression (SVR) (Moustapha and Bourinet 2018), Gaussian Process Regression (GPR) (Moustapha and Bourinet 2018; Frazier 2018), Neural Network (Yu and Wang 2009), and Polynomial Chaos Expansion (PCE) (Capellari and Chatzi 2018). Models like PCE and SVR yield a point prediction of the output. Therefore, they are computationally cheaper than approaches like GPR that also predicts the uncertainties in the output. I use GPR to build our surrogate model which turns out to be 5000 times faster than the FEM model. The output of the surrogate model usually has a very large dimension. I overcome the issue of high-dimensional output space using the *Single Value Decomposition* (SVD) technique that reduces the high-dimensional correlated output space to low-dimensional uncorrelated features. I transform the strain response from 22500 dimensions to lower 28-dimensional latent space using SVD. These 28 important features cover 99.2% of the total information of the data. These 28 features can be inverted to obtain the complete strain gauge response. I have built the surrogate model for each of these 28 features using GPR. One-third of the 5000 data points were used for training the GPR, whereas the remaining was used for validation to verify the accuracy of the surrogate. Fig. 2.4 illustrates the discussion carried out so far.

The Figures 2.5a and 2.5b show the nearly identical strain field obtained for FEM and the surrogate model for a random input sample. Fig. 2.6a shows the error in the prediction of the strain

values using the FEM and GPR model for the bottom 100 strain gauge locations at the left boundary of the beam Overall, the absolute prediction error is of the order 10^{-7}, and the relative error is of the order 10^{-3} However, a relatively high prediction error is observed at the locations of springs. Similarly, Fig. 2.6b shows the distribution of the absolute prediction error across the beam for a random input sample Once again, relatively higher errors are observed at the spring locations and the location where the concentrated load acts.

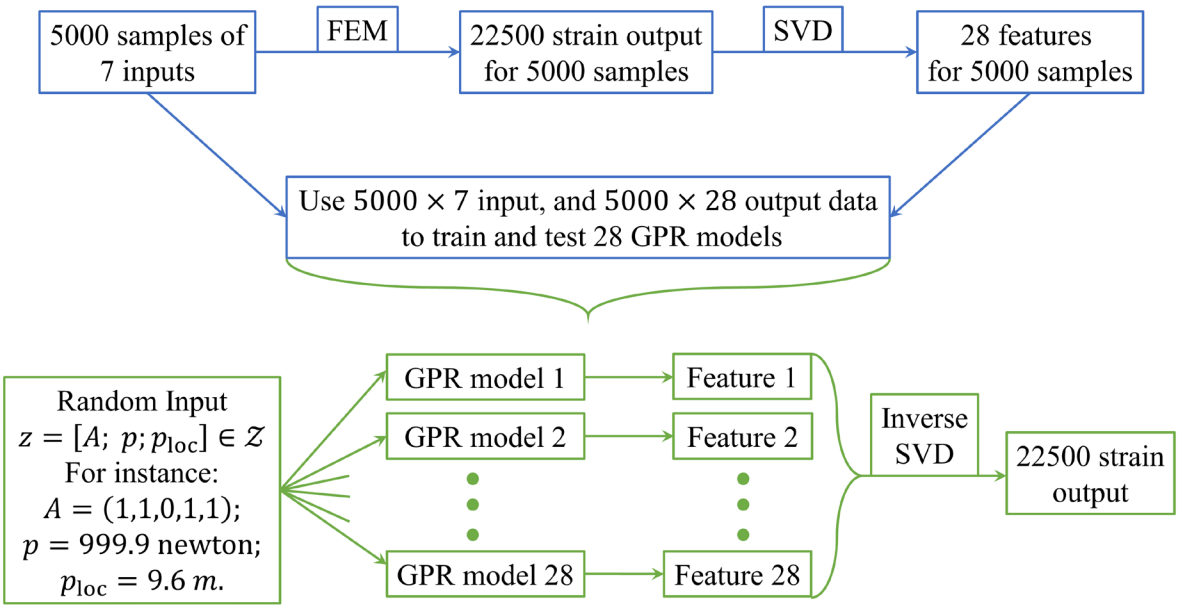
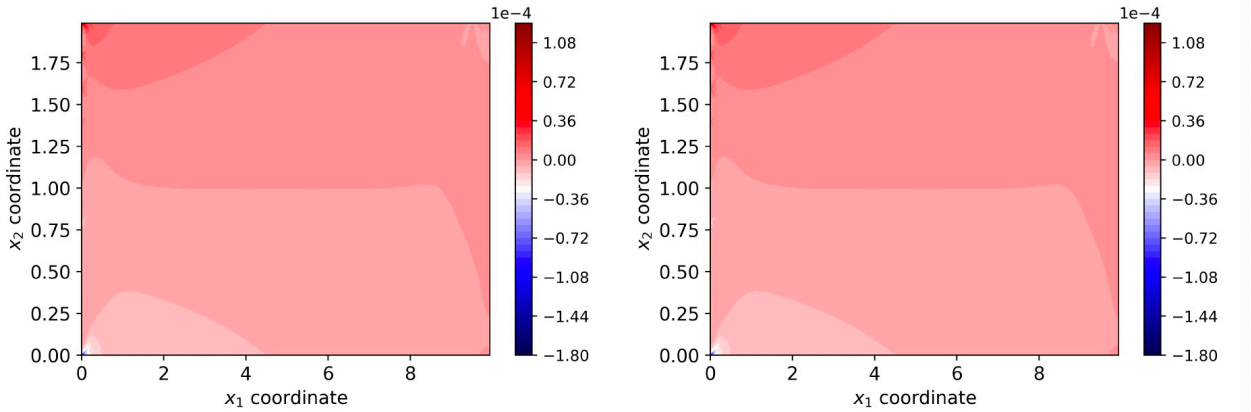


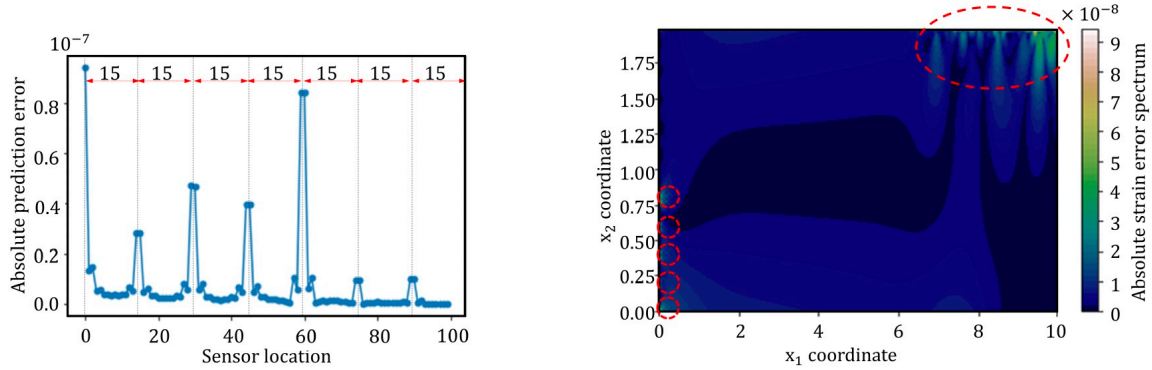
Figure 2.4: Flowchart describing strain data generation using FEM, and prediction using GPR surrogate model



(a) Strain field obtained using FEM

(b) Strain field obtained using GPR model

Figure 2.5: Comparison of the strain fields obtained using FEM and GPR model



(a) Error in the 100 strain values at the left boundary of the beam

(b) Error in the strain values for a random input sample

Figure 2.6: Error in the strain values obtained using FEM and GPR model

2.5.3 Revisiting Bayes risk

As seen in Section 2.5.2, the measurement $x_e \in \Omega_{X_e}$ depends on the load $p \in \Omega_P$, its location $p_{\text{loc}} \in \Omega_{P_{\text{loc}}}$. The randomness in the strain values (observations) $x_e \in \Omega_{X_e}$ are primarily due to the noise in strain gauge, uncertainties in the concentrated load, and its location. I assume a zero mean Gaussian noise structure $\zeta \sim N(\mu_\varepsilon = 0, \sigma_\varepsilon = 5 \times 10^{-7})$. Let ε represent the realization of noise and Ω_ζ represent the noise space, such that $\varepsilon \in \Omega_\zeta$.

Consider a design e with $N_{\text{sg}}(e)$ number of strain measurement locations. Let ζ_i represent the random variable for the noise in the i^{th} strain location. It is reasonable to assume that the

$\left(N_{\mathrm{sg}}(e)+2\right)$ random variables $P, P_{\mathrm{loc}}, \zeta_i, \cdots, \zeta_{N_{\mathrm{sg}}(e)}$ are statistically independent. I define a design dependent product space $\Omega_{\xi_e} = \Omega_P \times \Omega_{P_{\mathrm{loc}}} \times \Omega_{\zeta_1} \times \Omega_{\zeta_2} \times \cdots \times \Omega_{\zeta_{N_{\mathrm{sg}}(e)}}$. The random vector ξ_e consists of the realizations of the random variables $P, P_{\mathrm{loc}}, \zeta_i, \cdots, \zeta_{N_{\mathrm{sg}}(e)}$. The joint density function is then written as:

$$f_{\xi_e}(\beta) = f_P(p) \cdot f_{P_{\mathrm{loc}}}(p_{\mathrm{loc}}). \prod_{i=1}^{N_{\mathrm{sg}}(e)} f_{\zeta_i}(\varepsilon_i), \text{ where,}$$
$$\beta = \left(p \in \Omega_P, p_{\mathrm{loc}} \in \Omega_{P_{\mathrm{loc}}}, \varepsilon_1 \in \Omega_{\zeta_1}, \varepsilon_2 \in \Omega_{\zeta_2}, \cdots, \varepsilon_{N_{\mathrm{sg}}(e)} \in \Omega_{\zeta_{N_{\mathrm{sg}}(e)}}\right) \in \Omega_{\xi_e}. \tag{2.16}$$

Noting that the randomness in the measurement space Ω_{X_e} is by virtue of the uncertainty in Ω_{ξ_e} space, I rewrite Eq. (2.11) as follows,

$$f_{G|A}(g_{ni} \mid A_k) = \int_{\Omega_{\xi_e}} f_{G|\xi_e}(g_{ni} \mid \beta, A_k) \cdot f_{\xi_e}(\beta) \mathrm{d}\beta \text{ ,where,} \tag{2.17 a}$$

$$f_{G|\xi_e}(g_{ni} \mid \beta, A_k) = f_{G|X_e}(g_{ni} \mid x) \tag{2.17 b}$$

The second equation holds because a fixed input sample z, and noise value, yields a determinate and unique value of the measurement $x \in \Omega_{X_e}$. Substituting Eq. (2.17) into Eq. (2.10) yields:

$$\Psi_{\text{design-ex}}(e) = \int_{\Omega_{\xi_e}} \sum_{k=1}^{32} \mathcal{L}(\beta, A_k; e) f_A(A_k) f_{\xi_e}(\beta) d\beta, \text{ where,} \tag{2.18 a}$$

$$\mathcal{L}(\beta, A_k; e) = \sum_{n=1}^{5} \sum_{i=0}^{1} w_n L(g_{ni}, A_k) f_{G|\xi_e}(g_{ni} \mid \beta, A_k) \tag{2.18 b}$$

I note that these random variables constituting β can follow a generic distribution. I can always transform them to a standard normal random variable. Therefore, in an attempt to generalize. I transform the load P, its location P_{loc}, and the noise ζ_i into their respective standard normal forms. Since the load and the noise for the i^{th} strain gauge is Gaussian in our case, their standard normal forms can be written as $\mathcal{U}$ (standard normal counterpart of P), and $\mathcal{V}_i$ (standard normal counterpart of ζ_i), such that $p = u\sigma_p + \mu_p$, and $\varepsilon_i = v_i\sigma_\varepsilon + \mu_\varepsilon$, where u and v_i are the realizations of $\mathcal{U}$, and $\mathcal{V}_i$ respectively. I transform $f_{P_{\text{loc}}}(p_{\text{loc}})$ from Half Normal to a Standard Normal random variable , such that the cumulative density functions are equal: $F_{P_{\text{loc}}}(p_{\text{loc}}) = F_{\mathcal{U}_{\text{loc}}}(u_{\text{loc}})$, and $\mu_{p_{\text{loc}}} = F_{P_{loc}}^{-1}\left(F_{\mathcal{U}_{\text{loc}}}\left(\mu_{u_{\text{loc}}}\right)\right)$. This operation transforms ξ_e into a joint standard normal random variable $\mathcal{B}_e$ (with a realization $\mathscr{b}$, where $\mathscr{b} \in \Omega_{\mathcal{B}_e}$), such that

$$f_{\mathcal{B}_e}(\mathscr{b}) = f_{\mathcal{U}}(u) \cdot f_{\mathcal{U}_{\text{loc}}}(u_{\text{loc}}). \prod_{i=1}^{N_{\text{sg}}(e)} f_{\mathcal{V}_i}(v_i), \text{ where,}$$
$$\mathscr{b} = \left(u, u_{\text{loc}}, v_1, v_2, \cdots, v_{N_{\text{sg}}(e)}\right). \tag{2.19}$$

I can now rewrite Eq. (2.18) as:

$$\Psi_{\text{design-ex}}(e) = \int_{\Omega_{\mathcal{B}_e}} \sum_{k=1}^{32} \lambda(\mathscr{b}, A_k; e) f_A(A_k) f_{\mathcal{B}_e}(\mathscr{b}) \mathrm{d}\mathscr{b}, \text{ where,} \tag{2.20 a}$$

$$\lambda(\mathscr{b} = (u, u_{\text{loc}}, v_i), A_k; e)$$
$$= \mathcal{L}\left(\beta = \left(u\sigma_p + \mu_p, F_{P_{\text{loc}}}^{-1}\left(F_{\mathcal{U}_{\text{loc}}}(u_{\text{loc}})\right), v_i\sigma_\varepsilon + \mu_\varepsilon, A_k; e\right)\right) \tag{2.20 b}$$

Section 2.5.4 deals with evaluating the Bayes risk discussed in this section. To maintain generality, I present the formula for Bayes risk as an approximation of both, Eq. (2.18a), and Eq. (2.20a).

2.5.4 Evaluating the expected cost considering uncertainties in load and noise in the observed strains

2.5.4.1 Obtaining the cost $\mathcal{L}$ for a given input sample z and noise structure

Once having the GPR models, I can obtain $f_{X_e|A}(x_e \mid A_k)$, and hence evaluate $f_{G|X_e}(g_{ni} \mid x_e)$ using Eq. (2.14). To demonstrate a simple case of the evaluation of posterior probability of spring existence, I ignore the uncertainties due to load and its location by fixing the load as: $p \in \Omega_P$ and $p_{\text{loc}} \in \Omega_{P_{\text{loc}}}$. I consider a design with $N_{\text{sg}}(e)$ strain gauges, picked randomly using *Latin Hypercubic Sampling*. Assuming that the true state of the springs is A, I consider the input sample as $z = (A; p; p_{\text{loc}}) \in \Omega_Z$. For the chosen design e and the input sample z, I run multiple surrogate runs over different noise values in the Monte Carlo sense. The posterior can then be obtained using Eq. (2.15). Similarly, for the same fixed load and its location, the likelihood $f_{X_e|A}(x_e \mid A_k)$ for all possible spring states can be obtained, yielding $f_{A|X_e}(A_k \mid x_e)$ using Eq. (2.16). Finally, I can obtain $f_{G|X_e}(g_{ni} \mid x_e)$ using equations (2.12) and (2.13). Equations (2.17b) and (2.18b) yields $\mathcal{L}(\beta, A_k; e)$. Fig. 2.7 illustrates the discussion so far.

For this special example with a fixed load and its location, and that I have assumed a well-defined noise structure with zero mean $\zeta \sim N(0{,}5 \times 10^{-7})$, I can obtain the likelihood $f_{X_e|A}(x_e \mid A_k)$ without numerous surrogate runs. For the fixed input sample z, let $\bar{x}_{ei}$ represent the strain values at the n-th sensor locations obtained using either a forward FEM or a surrogate model. The closed form likelihood for such case can then be written as $f_{X_e|A}(x_e \mid A_k) = \prod_{n=1}^{N_{\text{sg}}(e)} \frac{1}{5\times10^{-7}} \phi\left(\frac{x_{en}-\bar{x}_{en}}{5\times10^{-7}}\right)$.

I need to incorporate the cumulative uncertainties due to all the aforementioned entities into evaluating the Bayes risk. Evaluating $\Psi_{\text{design-ex}}(e)$ and calculating the associated integral in

Eq. (2.18a) is computationally expensive and not so trivial. I do this by using three techniques discussed in the next section: a sampling-based approach, mean value approximation, and univariate dimension reduction with Gauss-Hermite quadrature.

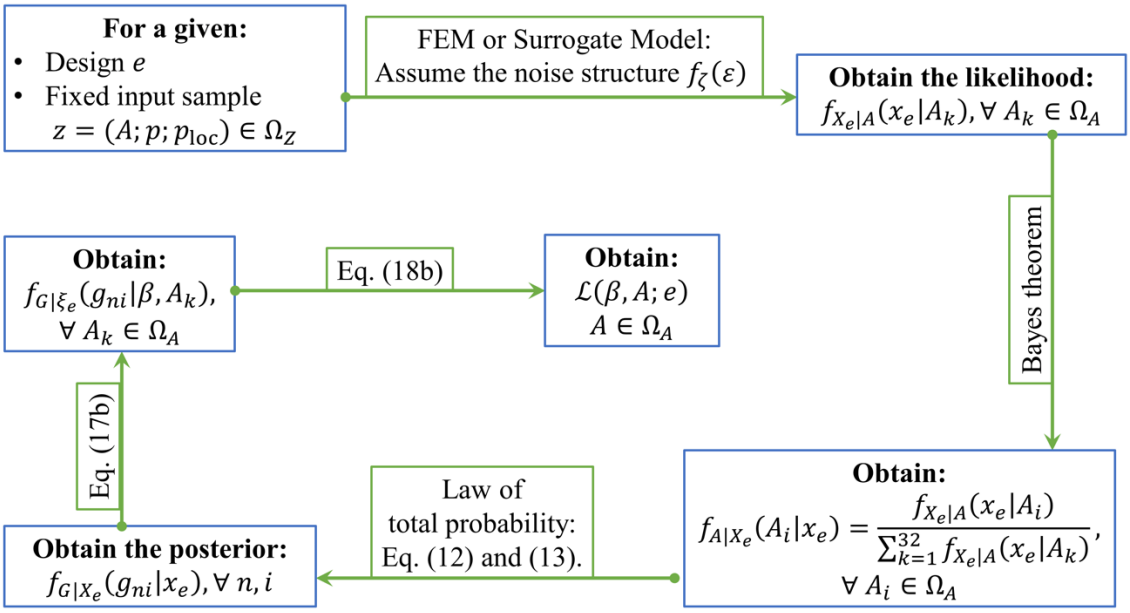


Figure 2.7: Flowchart describing the approach to obtain the cost $\mathcal{L}$ for a given design e and input sample z

2.5.4.2 Approach 1: Sampling-based method

This is a Monte Carlo based approach, where I generate large number of random samples of $A_i \in \Omega_A$, and $\beta_i \in \Omega_{\xi_e}$, with $i \in \{1,2,3,\cdots,N_{\text{mcs}}\}$. Here, N_{mcs} denotes the number of Monte Carlo samples. For the design e, I can obtain the cost $\mathcal{L}(\beta_i, A_i; e)$ for each A_i and β_i, as in procedure detailed in previous Section 2.5.4.1. The Bayes risk is then approximated as:

$$\Psi_{\text{design-ex}}(e) \approx \frac{1}{N_{\text{mcs}}} \sum_{i=1}^{N_{\text{mes}}} \mathcal{L}(\beta_i, A_i; e) \qquad (2.21\text{ a})$$

$$\Psi_{\text{design-ex}}(e) \approx \frac{1}{N_{\text{mcs}}} \sum_{i=1}^{N_{\text{mes}}} \lambda(\mathscr{b}_i, A_i; e) \qquad (2.21\text{ b})$$

The approximated Bayes risk depicted in the above equations is also called empirical risk. I recall here that Eq. (21a) and (21b) represents the approximated Bayes risk corresponding to Eq. (18), and Eq. (20) respectively. Fig. 2.8 illustrates a convergence plot for the sampling-based method obtained for a design e, with $N_{\text{sg}}(e) = 30$ strain gauges. The expected cost converges to 327.2 (showed by the red line) around $N_{\text{mcs}} = 10000$ samples with the average noise of 0.71. However, the evaluation of Bayes risk for 10000 samples takes 265 seconds. The optimization process demands an evaluation of the cost function around ten thousand times. Therefore, the sampling-based method is computationally expensive.

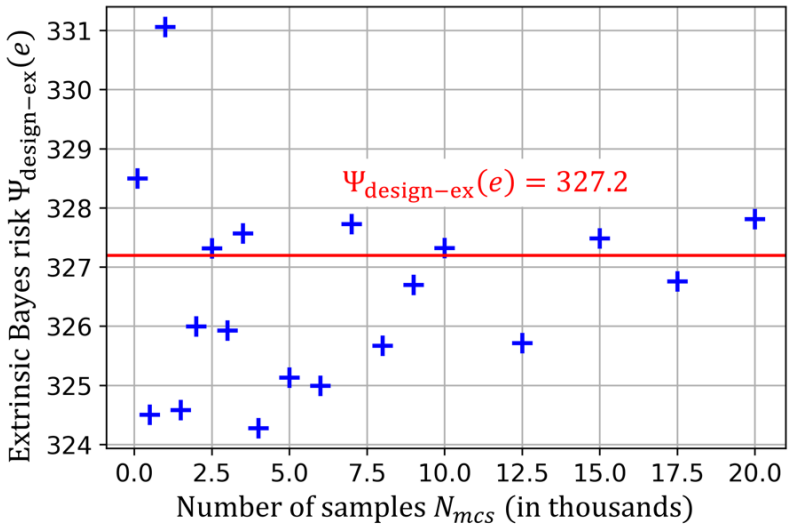


Figure 2.8: Convergence plot of Bayes risk obtained using sampling-based method

2.5.4.3 Approach 2: Mean value approximation

Here, I evaluate the Bayes risk as the cost $\mathcal{L}(\beta, A_k; e)$ evaluated for all the spring states A_k at the mean value of the load, its location, and noise, weighted over by the probability $f_A(A_k)$, such that:

$$\Psi_{\text{design-ex}}(e) \approx \sum_{k=1}^{32} \mathcal{L}\left(\beta = \left(\mu_p, \mu_{p_{\text{loc}}}, \mu_{\varepsilon_1}, \mu_{\varepsilon_2}, \cdots, \mu_{\varepsilon_{N_{\text{sg}}(e)}}\right), A_k; e\right) f_A(A_k) \tag{2.22 a}$$

$$\Psi_{\text{design-ex}}(e) \approx \sum_{k=1}^{32} \lambda\left(\mathscr{b} = \left(\mu_u, \mu_{p_{\text{loc}}}, \mu_{v_1}, \mu_{v_2}, \cdots, \mu_{v_{N_{\text{sg}}}}\right), A_k; e\right) f_A(A_k) \tag{2.22 b}$$

Since, $\mathcal{U}, \mathcal{U}_{\text{loc}}$, and $\mathcal{V}_i$ are standard normal random variables, I have $\mu_u = 0; \mu_{u_{loc}} = 0; \mu_{v_i} = 0$ $\forall i \leq N_{\text{sg}}(e)$.

2.5.4.4 Approach 3: Univariate dimensional reduction with Gauss-Hermite quadrature

Approach 1 is computationally expensive as it involves considering a large sample size, whereas approach 2 is feasible but not very accurate when there is large variability. I tackle these limitations using the current approach to evaluate Bayes risk.

I assume a design e, with $N_{\text{sg}}(e)$ number of strain gauges. I start by redefining Bayes risk in Eq. (2.20) as:

$$\Psi_{\text{design-ex}}(e) = \int_{\Omega_{\mathcal{B}_e}} h(\mathscr{b}; e) f_{\mathcal{B}_e}(\mathscr{b}) \mathrm{d}\mathscr{b}, \text{ where,} \tag{2.23 a}$$

$$h(\mathscr{b}; e) = \sum_{k=1}^{32} \lambda(\mathscr{b}, A_k; e) f_A(A_k) \tag{2.23 b}$$

Recall, that the vector $\mathscr{b} = \left(u, u_{\text{loc}}, v_1, v_2, \cdots, v_{N_{\text{sg}}(e)}\right)$ consist of $\left(N_{\text{sg}}(e) + 2\right)$ variables. I now define the following vectors consisting of $\left(N_{\text{sg}}(e) + 2\right)$ elements:

$$
\begin{aligned}
&b_0 = (0,0,0,0,\cdots,0)\\
&b_1 = (u,0,0,0,\cdots,0)\\
&b_2 = (0,u_{\mathrm{loc}},0,0,\cdots,0)\\
&b_3 = (0,0,v_1,0,\cdots,0)\\
&b_4 = (0,0,0,v_2,\cdots,0)\\
&\vdots\\
&b_{(N_{\mathrm{sg}}(e)+2)} = \left(0,0,0,0,\cdots,v_{N_{\mathrm{sg}}(e)}\right)
\end{aligned}
\tag{2.24}
$$

Using the definitions above and univariate dimensional reduction (Rahman and Xu 2004), I approximate the function $h(\mathscr{b};e)$ as:

$$
h(\mathscr{b};e) \approx -\left(N_{\mathrm{sg}}(e)+1\right)\lambda(b_0,A_k;e) + \sum_{i=1}^{(N_{\mathrm{sg}}(e)+2)} \lambda(b_i,A_k;e) \tag{2.25}
$$

Substituting Eq. (2.25) into Eq. (2.23),

$$
\begin{aligned}
\Psi_{\text{design-ex}}(e) \approx \sum_{k=1}^{32} \Bigg(&-\left(N_{\mathrm{sg}}(e)+1\right)\lambda(b_0,A_k;e)\\
&+ \sum_{i=1}^{(N_{\mathrm{sg}}(e)+2)} \int_{\Omega_{\mathcal{B}e}} \lambda(b_i,A_k;e) f_{\mathcal{B}_e}(\mathscr{b})\mathrm{d}\mathscr{b} \Bigg) f_A(A_k)
\end{aligned}
\tag{2.26}
$$

To simplify the expression above, firstly, I realize that $f_{\mathcal{B}_e}(\mathscr{b})$ is the joint probability density function of statistically independent standard normal random variables. Therefore,

$$
f_{\mathcal{B}_e}(\mathscr{b}) = \phi(u)\cdot\phi(u_{\mathrm{loc}}).\prod_{i=1}^{N_{\mathrm{sg}}(e)}\phi(v_i) = \prod_{i=1}^{N_{\mathrm{sg}}(e)+2}\phi(\mathscr{b}_i) = \prod_{i=1}^{N_{\mathrm{sg}}(e)+2}\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\mathscr{b}_i^2}\right) \tag{2.27}
$$

In the equation above, $\mathscr{b}_1 = u, \mathscr{b}_2 = u_{\text{loc}}$, and $\mathscr{b}_{j+2} = v_j$, for $j \in \left(1,2,\cdots,N_{\text{sg}}(e)\right)$. Secondly, I note that for any function of the form $g(x,y), E_{XY}(g(x,0)) = E_X(g(x,0))$, provided X and Y are statistically independent random variables. This allows me to simplify the integral in Eq. (2.26) as:

$$\int_{\Omega_{\mathcal{B}e}} \lambda(b_i, A_k; e) f_{\mathcal{B}_e}(\mathscr{b}) \mathrm{d}\mathscr{b} = \frac{1}{\sqrt{2\pi}} \int_{\tilde{\mathscr{b}}_i} \lambda(b_i, A_k; e) e^{-\frac{1}{2}\mathscr{b}_i^2} \, \mathrm{d}\mathscr{b}_i \tag{2.28}$$

Then, I realize that the Gauss-Hermite quadrature is a natural choice for approximating integral in the equation above. This is because Gauss-Hermite quadrature is meant to estimate integrals of form $\int g(x) e^{-x^2} \, \mathrm{d}x$, for any function g(x). Therefore, the integral above is approximated as:

$$\begin{gathered} \int_{\mathscr{b}} \lambda(b_i, A_k; e) f_{\mathcal{B}_e}(\mathscr{b}) \mathrm{d}\Omega_{\mathcal{B}_e} \approx \frac{1}{\sqrt{\pi}} \sum_n w_n \lambda\left(q_{i,n}, A_k; e\right); \\ q_{i,n}(j) = \begin{cases} b_i(j) = 0 & i \neq j; \\ \sqrt{2}\alpha_n & i = j. \end{cases} \end{gathered} \tag{2.29}$$

In the equation above, n represents quadrature order, w_n gives the weights, and α_n gives the point of evaluation of the function. For my calculations, I use $n = 2$, for which $w_n = 0.5\sqrt{\pi}$, and $\alpha_n = \pm\frac{1}{\sqrt{2}}$. The approximated Bayes risk can now be written as:

$$\Psi_{\text{design-ex}}(e) \approx \sum_{k=1}^{32} \left(-\big(N_{\text{sg}}(e)+1\big)\lambda(b_0, A_k; e) + \sum_{i=1}^{(N_{\text{sg}}(e)+2)} \sum_{n} w_n \lambda\big(q_{i,n}, A_k; e\big) \right) f_A(A_k) \tag{2.30}$$

The advantage of Bayes risk expressed in the form of Eq. (2.20) is clear from the discussion carried out so far. The expression of Bayes Risk in Eq. (2.30) can easily be extended to obtain Bayes risk in the form of Eq. (2.18).

The table below compares the value of Bayes risk and the run time for various approaches discussed in this section. The sampling-based method is the most accurate when a large sample size is considered. However, it is computationally expensive. Secondly, irrespective of the sample size, the Bayes risk approximated using approach 1 changes with the new sample even with the same sample size, hence, is random and non-unique. Approach 2 is the most feasible but not so accurate. Approach 3 enjoys acceptable accuracy and computational speed.

Table 2.3: Comparison of various approaches in evaluating Bayes risk for a design $\boldsymbol{e}$ with $\boldsymbol{N_{\text{sg}}(e) = 30}$

	Bayes risk $\Psi_{\text{design-ex}}(e)$	Run time in seconds
Approach 1 (10^4 samples)	327.32	229.97
Approach 2	309.81	0.82
Approach 3	321.35	46.53

The discussion in this chapter so far was about evaluating the Bayes risk for a given design e. The next section will focus on the problem of optimal sensor placement using Bayesian optimization. I use approach 3 to evaluate the extrinsic Bayes risk $\Psi_{\text{design-ex}}(e)$ and assume an intrinsic cost of unity per additional sensor.

2.6. Bayesian optimization: Optimal sensor placement design

2.6.1 Optimal sensor placement design algorithm

The primary objective is to obtain the optimal sensor placement design e^* that minimizes the Bayes risk functional discussed in the previous sections. Mathematically,

$$e^* = \arg\min_{e} \Psi_{\text{design}}(e) \in \Omega_E \tag{2.31}$$

In absolute terms, obtaining e^* involves looking at every possible design combination, and picking the one with the least Bayes risk. In our case, if $n = 22500$, this would be picking e^* from the $\sum_{r=1}^{n} \frac{n!}{r!(n-r)!} = (2^n - 1)$ possible combinations of sensor locations. Clearly, sampling the entire design space Ω_E, which consists of $(2^{22500} - 1) \approx 10^{6773}$ number of possible designs, is daunting even for this modest problem. The main motivation of using Bayesian optimization is to arrive at the optimal solution e^* by minimizing the sampling points to fasten the optimization process. Bayesian optimization looks for the global optimum in a minimum number of steps.

Unlike gradient-based optimization methods, Bayesian optimization is a global optimization technique that does not require the derivative of the objective function. Having a black-box model (like a surrogate function) of the objective function suffices to perform the

optimization. It involves two primary elements. The first element is developing surrogate function using *Gaussian process regression* (GPR) of the objective function using randomly evaluated sample. Consider, for example in our case, I assume an initial design $e_0 \in \Omega_E$, with 3 strain locations. To obtain the next optimal design with 4 strain gauges, I randomly sample, for instance, 20 locations to be the candidate for the 4th sensor. These locations yield 20 design samples $\tilde{e}(k), \forall k \leq 20$ each with four sensors. I obtain the exact cost $\Psi_{\text{design}}(\tilde{e}(k)), \forall k \leq 20$ using approach 3 discussed in previous section. Using the 20 set input data of the fourth sensor location $d = (x_1, x_2)$, and the output data of the exact cost, I train the surrogate function $\hat{\Psi}_{\text{design}}(d) \sim N(\mu_d, \sigma_d)$. It provides a posterior probability that describes possible values for the cost at a candidate fourth location d, with the mean value μ_d, and the standard deviation σ_d. The second component is the *acquisition function* that helps me locate the next most valuable candidate for the fourth location based on the current posterior over the cost. I use *Expected Improvement* EI as our acquisition function.

$$EI(d) = \left(\mu_d - \Psi^*_{\text{design}}\right)\Phi\left(\frac{\mu_d - \Psi^*_{\text{design}}}{\sigma_d}\right) + \sigma_d \phi\left(\frac{\mu_d - \Psi^*_{\text{design}}}{\sigma_d}\right) \tag{2.32}$$

Here, $\Psi^*_{\text{design}} = \min_{\tilde{e}(k)} \Psi_{\text{design}}(\tilde{e}(k))$ is the *current best values* of the objective function. For all the remaining $(22500 - 20 - 3) = 22477$ possible fourth location candidates, I evaluate $EI(d)$. The candidate with maximum EI is the next most valuable location. Once I locate the next most valuable fourth location candidate, I get 21st design sample. I re-train the GPR with 21 data points, and keep adding the next most valuable location until the maximum EI is less than a tolerance value ε. For detailed understanding of Bayesian optimization, readers are recommended to refer to (Jones and Schonlau 1998) and (Frazier 2018).

To generalize our optimization algorithm, I define the initial sensor design as $e_0 \in \Omega_E$ with $N_0 = N_{\text{sg}}(e_0)$ number of strain gauges. If $d^{(i)} = \left(x_1^{(i)}, x_2^{(i)}\right)$ represents the location of i^{th} strain gauge ($x_1^{(i)}$ and $x_2^{(i)}$ denote the horizontal and vertical coordinates of the sensor i), I have $e_0 = \left(d^{(1)}, d^{(2)}, \cdots, d^{(N_0)}\right)$. Let N_{as} represent number of additional sensors that will be added one by one to N_0 during the optimization process. Let $e_{n_{\text{as}}}$ represent the optimized sensor design with $(N_0 + n_{\text{as}})$ sensors, such that $n_{\text{as}} \leq N_{\text{as}}$, and $e^* = \arg\min_{e_{n_{\text{as}}}} \Psi_{\text{design}}\left(e_{n_{\text{as}}}\right)$. Finally, $N_{\text{total}} = 22500$ represents total number of strain gauge locations. Fig. 2.9 details the flowchart of the optimization algorithm 1 developed for obtaining optimal sensor placement. In Yang et al. (Yang and Chadha 2021), this algorithm was deployed to obtain a sensor placement design of a more complex real-world miter gate structure that had a different type of damage (unlike the detection type of problem here) and a different Bayes risk functional (quantifying the net relative gain in information). It shows the generality and applicability of the proposed algorithm.

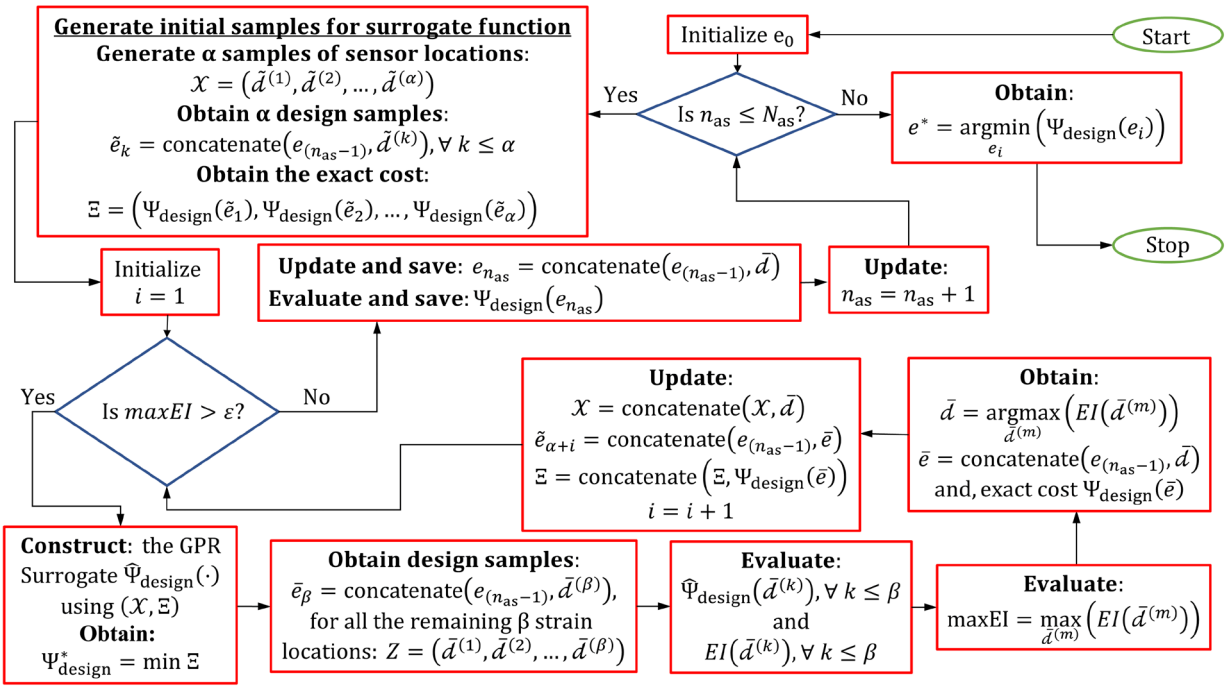


Figure 2.9: Flowchart of Bayesian optimization algorithm for optimal sensor placement design

Algorithm 1: Bayesian optimization for sensor placement

Initialize $e_0 = \left(\mathscr{d}^{(1)}, \mathscr{d}^{(2)}, \cdots, \mathscr{d}^{(N_0)}\right)$;

for $n_{as} = 1$ **to** N_{as} **do**

  Using LHS, randomly select α locations to be candidates for the $(N_0 + n_{\text{as}})$ sensor location, with coordinates $\mathscr{X} = \left(\tilde{\mathscr{d}}^{(1)}, \tilde{\mathscr{d}}^{(2)}, \cdots, \tilde{\mathscr{d}}^{(\alpha)}\right)$;

  Obtain α number of possible designs: $\tilde{e}_k = \text{concatenate}\left(e_{(n_{\text{as}}+1)}, \tilde{\mathscr{d}}^{(k)}\right)$, for all $k \leq \alpha$;

  Obtain the exact cost of all the α designs: $\Xi = \left(\Psi_{\text{design}}(\tilde{e}_1), \Psi_{\text{design}}(\tilde{e}_2), \cdots, \Psi_{\text{design}}(\tilde{e}_\alpha)\right)$;

  while $i = 1$ **or** $maxEI < \varepsilon$ **do**

    Construct the GPR model for $\hat{\Psi}_{\text{design}}(\cdot)$ trained using (X, Ξ);

    For all the remaining strain locations $Z = \left(\bar{\mathscr{d}}^{(1)}, \bar{\mathscr{d}}^{(2)}, \cdots, \bar{\mathscr{d}}^{(\beta)}\right)$, where $\beta = \left(N_{\text{total}} - \left(N_0 + n_{\text{as}} - 1\right) - \alpha\right)$, obtain β number of possible designs: $\bar{e}_m = \text{concatenate}\left(e_{(n_{\text{as}}+1)}, \bar{\mathscr{d}}^{(m)}\right)$, for all $m \leq \beta$;

    Obtain the cost $\hat{\Psi}_{\text{design}}\left(\bar{\mathscr{d}}^{(m)}\right)$ for all $m \leq \beta$ designs using GPR developed before;

    Obtain the currrent best $\Psi^*_{\text{design}} = \min \Xi$;

    Obtain the Expected Improvement for all the β designs using:

$$EI(\bar{\mathscr{d}}^{(m)}) = \left(\mu_{\bar{\mathscr{d}}^{(m)}} - \Psi^*_{\text{design}}\right) \Phi\left(\frac{\mu_{\bar{\mathscr{d}}^{(m)}} - \Psi^*_{\text{design}}}{\sigma_{\bar{\mathscr{d}}^{(m)}}}\right) + \sigma_{\bar{\mathscr{d}}^{(m)}} \phi\left(\frac{\mu_{\bar{\mathscr{d}}^{(m)}} - \Psi^*_{\text{design}}}{\sigma_{\bar{\mathscr{d}}^{(m)}}}\right)$$

    , where $m \leq \beta$;

    Obtain:

$$maxEI = \max_{\bar{\mathscr{d}}^{(m)}} \left(EI(\bar{\mathscr{d}}^{(m)})\right)$$

$$\bar{\mathscr{d}} = \arg\max_{\bar{\mathscr{d}}^{(m)}} \left(EI(\bar{\mathscr{d}}^{(m)})\right)$$

$$\bar{e} = \text{concatenate}\left(e_{(n_{\text{as}}+1)}, \bar{\mathscr{d}}\right)$$

    Evaluate the exact cost $\mathfrak{C}(\bar{e})$;

    Update:

$$\mathscr{X} = \text{concatenate}\left(\mathscr{X}, \bar{\mathscr{d}}\right)$$

$$\tilde{e}_{(\alpha+i)} = \bar{e}$$

$$\Xi = \text{concatenate}\left(\Xi, \Psi_{\text{design}}(\bar{e})\right)$$

    $i = i + 1$;

  end

  Update the sensor design: $e_{n_{\text{as}}} = \text{concatenate}(e_{n_{\text{as}}-1}, \bar{\mathscr{d}})$;

end

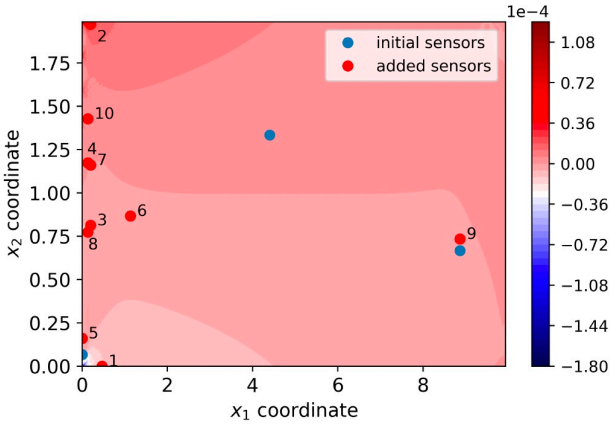
Obtain: $e^* = \arg\min_{e_k} \Psi_{\text{design}}(e_k)$, where, $k \leq N_{\text{as}}$;

2.6.2 Results and discussion

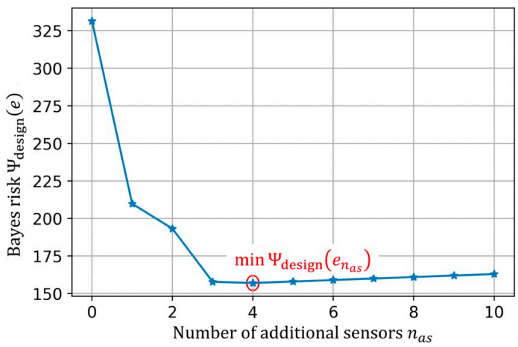
2.6.2.1 **Comparison of a Bayesian optimized sensor placement design with randomly chosen designs**

To numerically implement the optimization algorithm discussed in Section 2.6.1, I consider an initial design e_0 with $N_0 = 3$, with sensors picked randomly, and consider $N_{\text{as}} = 10$ additional sensors. I fix $\alpha = 20$. Fig. 2.10a below shows the sensors constituting e_0 by blue dots and the additional sensor location by red dots placed on the strain contour (for a random input sample) of the beam. For instance, the design e_1 consists of all the three initially considered sensors along with the fourth sensor in red (marked by number 4). Fig. 2.10a also shows the strain field for a realization of load and its location. Fig. 2.10 b illustrates Bayes risk for the designs $e_{n_{\text{as}}}$. I observe that the Bayes risk converges with 4 additional sensors, i.e., e_4 can be considered as the optimal design. I also observe that almost all these additional sensors are concentrated close to the boundary where the springs are present. I observe in all the following convergence plots (Figures 2.10 b, 2.13 b − 2.17 b) that the Bayes risk increases after the minimum value is attained because every additional sensor bears an intrinsic cost, which in this case was assumed to be unity per additional sensor.

To demonstrate the fact that Bayesian optimization produces the optimal sensor placement design, I consider a random design e_r, with $N_{\text{sg}}(e_r) = 13$. Figure 2.11 a shows the arrangement of the sensors for design e_r. Although design e_r has 6 more sensors than design e_4, the Bayes risk for e_r is much higher than that of the minimum Bayes risk for optimized design e_r. The reason is that the new information acquired by adding the 5-th sensor or more does not add to the value of decision-making as much as it leads to the increase in the intrinsic cost due to the addition of more sensors. This is clear from Fig. 2.11b.

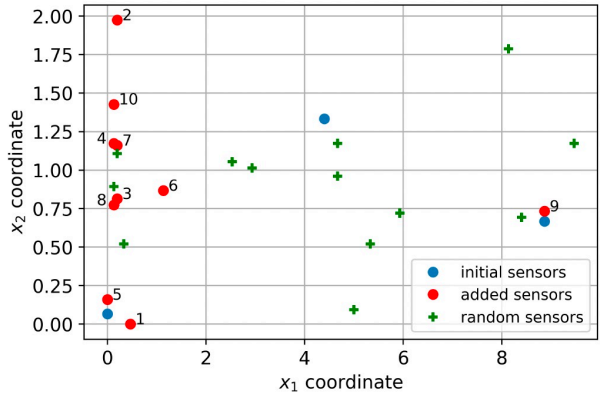


(a) Optimized sensor locations

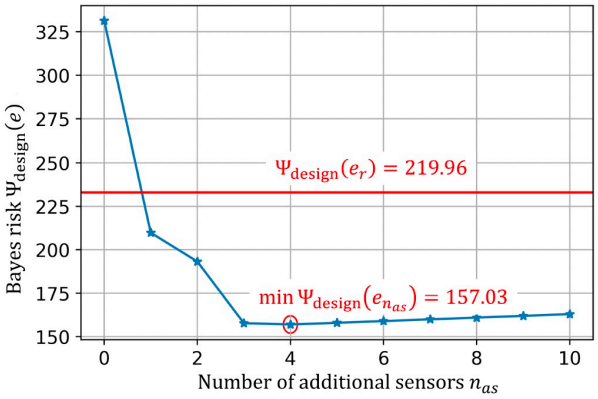


(b) Bayes risk

Figure 2.10: Optimized sensor placement and the associated Bayes risk obtained using approach 3



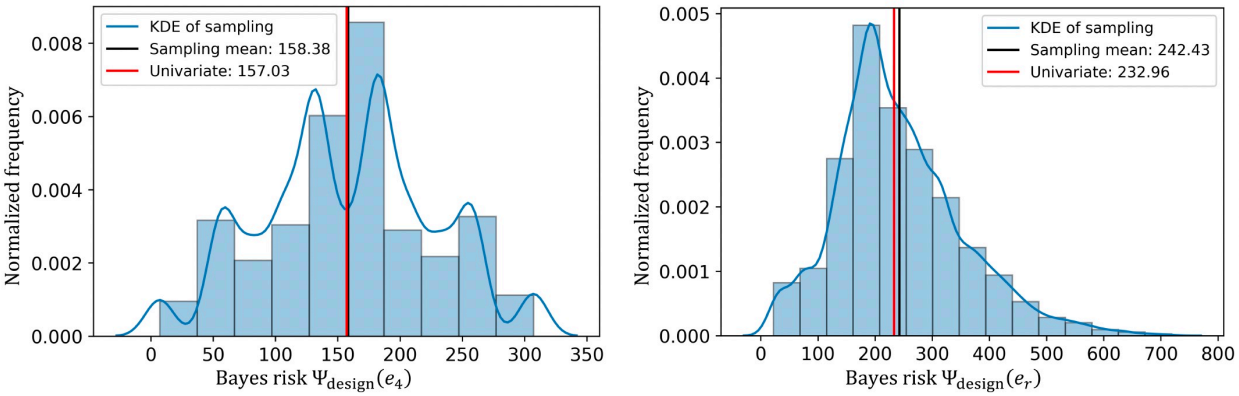
(a) Optimized sensor design $e_{n_{\text{as}}}$ and random design e_r



(b) Bayes risk $\Psi_{\text{design}}(e_{n_{\text{as}}})$ and $\Psi_{\text{design}}(e_r)$

Figure 2.11: Randomly selected sensor design e_r and the associated Bayes risk obtained using approach 3

I used approach 3 to evaluate Bayes risk while performing Bayesian optimization. Fig. 2.12a and 2.12 b, compare the expected cost obtained using the sampling-based method (approach 1 , with 10^4 samples) and approach 3 for designs e_4 and e_r respectively. As expected, the results obtained using approach 3 are very close to the sampling-based method (that can be assumed as ground truth) The plots also show the *Kernel Density Estimate* (KDE) for the sampling-based method. Finally, the deviation of Bayes risk in the case of random design e_r as compared to the optimal design e_4 is noteworthy.



(a) For the optimized design e_4 (b) For the random design e_r

Figure 2.12: Comparison of the Bayes risk evaluated using sample-based method (approach 1), and univariate dimensional reduction technique (approach 3)

Remark 2.2: I note that obtaining new information (for example: strain gauge data) is consequential in making a better decision (for example: detecting the existence of springs). However, acquiring information through a mechanism e bears cost, represented by $\Psi_{\text{design-in}}(e)$. Acquiring the new information is meaningful and economical if and only if the additional cost required to gather the information is outweighed by the reduction in the expected losses evaluated by considering the additional information (Chadha and Hu 2021). Recall the expression of the Bayes risk $\Psi_{\text{design}}(e)$ in Eq. (2.9). The Bayes risk is defined as the sum of intrinsic cost $\Psi_{\text{design-in}}(e)$ and extrinsic cost $\Psi_{\text{design-ex}}(e)$. Increasing the number of sensors has the following effects:

1 Every addition of the sensor increases the cost due to the intrinsic cost of the sensor, cost incurred to install and maintain the SHM system. Therefore, $\Psi_{\text{design-in}}(e)$ increases.

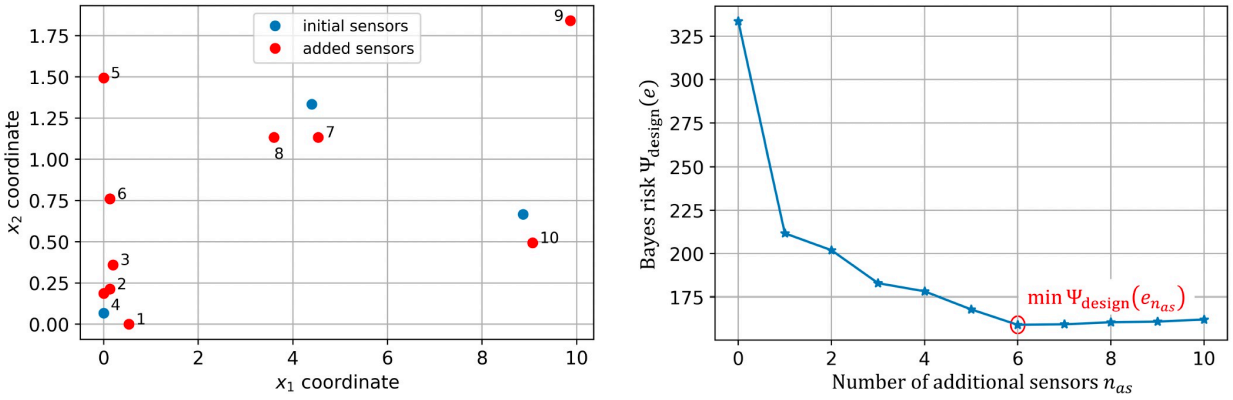
2 Every addition of the sensor also adds to the new information about the state of the structure leading to better decision making. Therefore, with the increase in sensor count, $\Psi_{\text{design-ex}}(e)$ decreases.

With the addition of a new sensor up to the optimal design e^*, the $\Psi_{\text{design-ex}}(e)$ decreases more than the increase in $\Psi_{\text{design-in}}(e)$, leading $\Psi_{\text{design}}(e)$ to decrease overall. However, beyond the optimal design, with any new addition of the sensors, $\Psi_{\text{design-ex}}(e)$ decreases less than the increase in $\Psi_{\text{design-in}}(e)$, leading $\Psi_{\text{design}}(e)$ to increase overall. In other words, there comes a time when the benefit of the additional information obtained by adding an additional sensor is dwarfed by the cost incurred due to a sensor addition. This effect is observed in all the convergence plots presented in this section.

2.6.2.2 **Comparison of a Bayesian optimized sensor placement design with Bayes risk evaluated using various approaches**

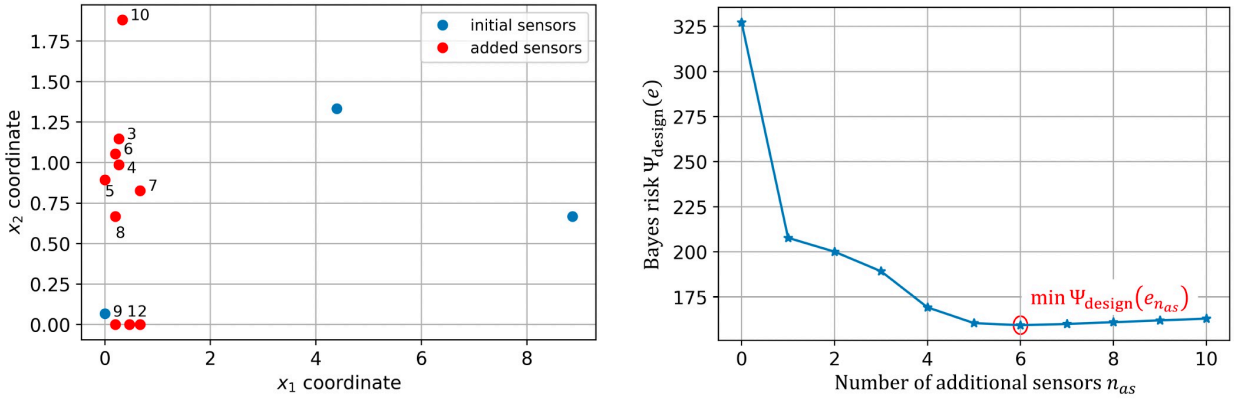
Now, I focus on the performance of various approaches detailed in Section 2.5.4 used in evaluating the Bayes risk while performing Bayesian optimization. Figures 2.10, 2.13, and 2.14 illustrate optimized sensor placement and the associated Bayes risk obtained using approaches 3, 1, and 2, respectively. It is not surprising that the convergence rate depends on the approach picked to evaluate the Bayes risk. Approach 2 (mean value approximation of Bayes risk) is not accurate, and the sampling-based method, like any Monte-Carlo based approach, bears uncertainties because it attempts to evaluate the integral in Bayes risk functional by sampling it. For a different sample with the same sample size, the Bayes risk evaluated using approach 1 is different. This randomness in the evaluation of Bayes risk using approach 1 leads the acquisition function to pick different sensor locations. Unlike these approaches, approach 3 attempts to evaluate the integral, and rather quickly and consistently (unlike sampling-based method), using Gaussian-Hermite quadrature These inherent advantages of approach 3 catalyze the optimization code to converge faster. It can be observed that approach 3 finds the first 4 sensors to be well spread in the vertical direction. The first four additional sensor locations obtained by using the Sampling-based technique are

concentrated to the bottom-left, and the code is forced to arrive at sensors 5 and 6 at the middle and the top left of the beam, respectively. I also note that there are instances where the sampling-based method converges faster than the other two approaches owing to the randomness in the prediction of Bayes risk by its very inherent nature. However, I note a commonality in the prediction by all three approaches. All the significant additional sensor locations (the first six additional sensors) are spread across the vertical direction near the left boundary of the beam, which is suitable for the spring detection problem.



(a) Optimized sensor locations (b) Bayes risk

Figure 2.13: Optimized sensor placement and the associated Bayes risk obtained using sampling-based method (approach 1)



(a) Optimized sensor locations (b) Bayes risk

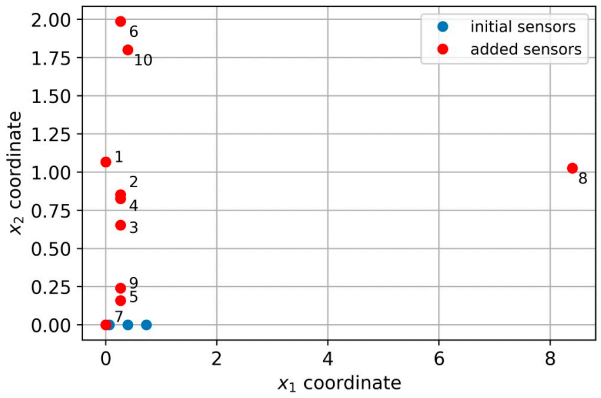
Figure 2.14: Optimized sensor placement and the associated Bayes risk obtained using mean value approximation (approach 2)

2.6.2.3 **Comparison of a Bayesian optimized sensor placement design with different initial designs**

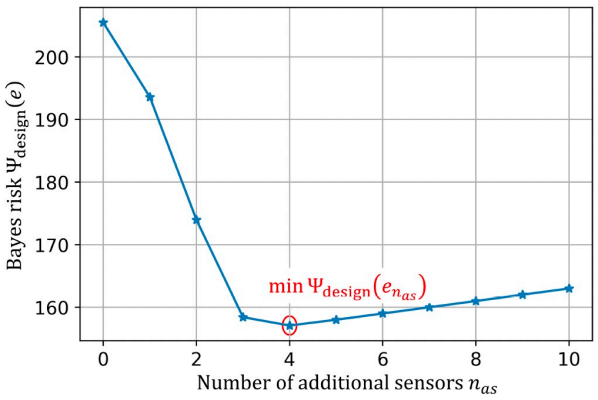
In this section, I compare the optimal sensor placement design evaluated using approach 3 for different initial designs e_0. Fig. 2.10 shows the sensor designs obtained when the initial sensor locations are well spread out across the beam. To demonstrate the effect of the choice of initial sensor designs, I then consider three extreme cases of e_0 with the sensors concentrated on the bottom-left, bottom-right, and top-right, as shown in figures 2.15, 2.16, and 2.17.

The *Expected Improvement* function defined in Eq. (2.32) guides the optimization algorithm to exploit and explore the design space to pick for the next sample. The algorithm exploits the strain locations at which the GP mean function is larger, and it explores the strain locations where the GP standard deviation is larger. For instance, in Fig. 2.15, the algorithm obtains the first 4 additional sensors by exploiting the strain locations with higher GP mean value, whereas in Fig. 2.16, with e_0 consisting of concentrated bottom-right sensors, the algorithm obtains the additional sensors mostly by exploring the region of high GP standard-deviation. Since it evaluates the additional sensors 2, 3, 4, and 5 concentrated at the top left, it is forced to obtain sensors 6, 7, and 8 in the middle of the left end, leading to late convergence. Like the previous section, I do observe that irrespective of the initial design, the algorithm arrives at a sensor design that consists of additional sensor locations spread out across the left end.

I note that the beam is so finely meshed that there exists a correlation between the strain values. Therefore, there are non-unique sensor locations that are sampled by the acquisition function, leading to non-unique sensor design depending on different initial design e_0.

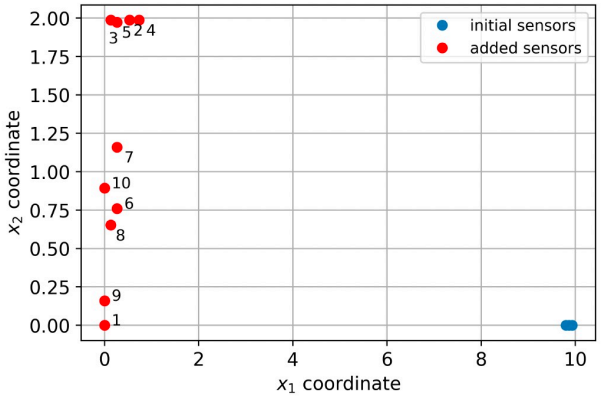


(a) Optimized sensor locations

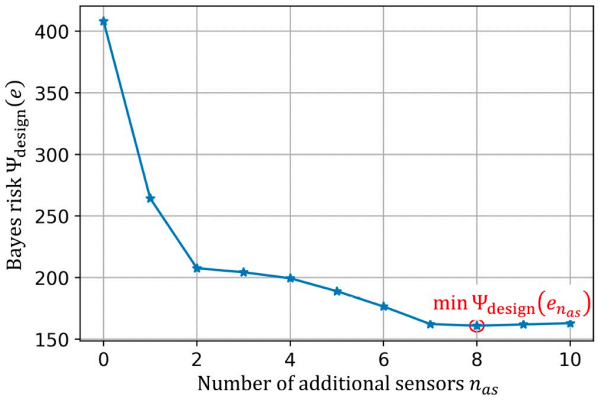


(b) Bayes risk

Figure 2.15: Optimized sensor placement and the associated Bayes risk obtained considering initial sensor design with three sensors concentrated at the bottom-left

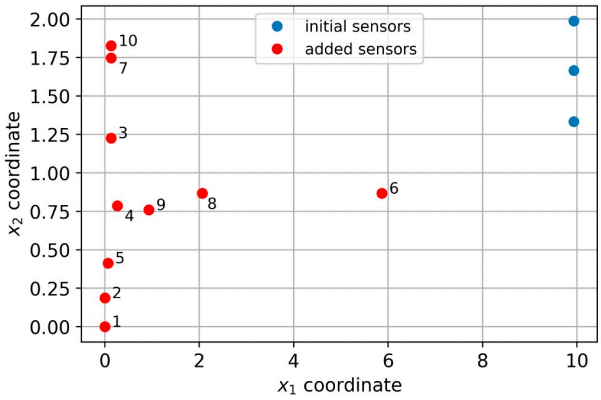


(a) Optimized sensor locations

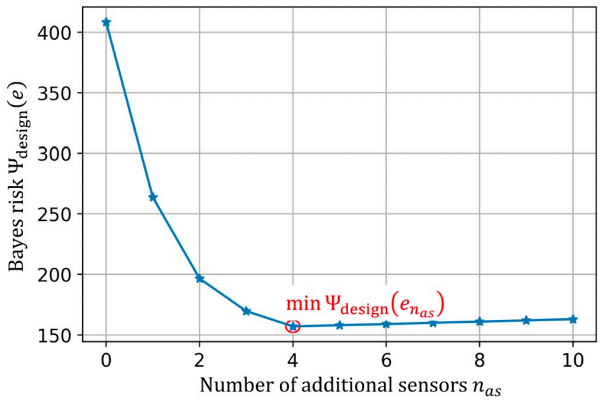


(b) Bayes risk

Figure 2.16: Optimized sensor placement and the associated Bayes risk obtained considering initial sensor design with three sensors concentrated at the bottom-right



(a) Optimized sensor locations



(b) Bayes risk

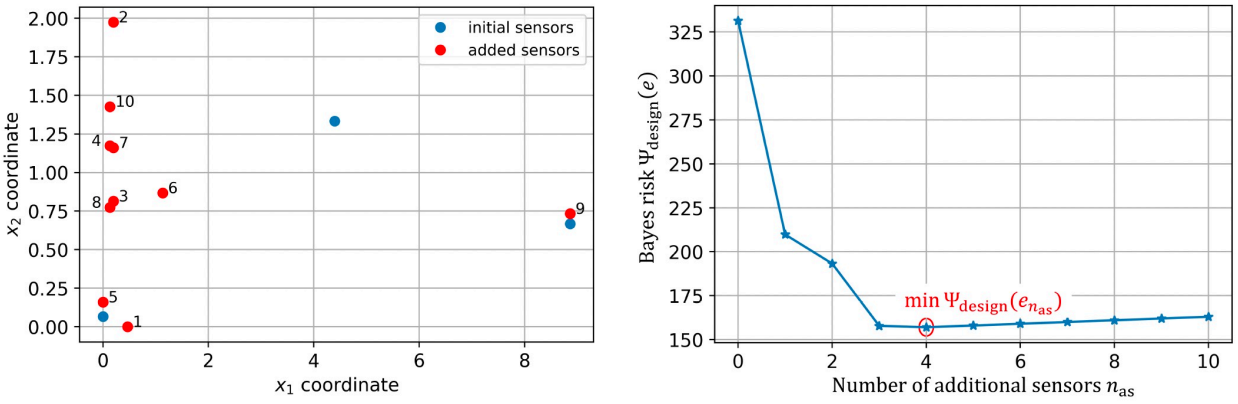
Figure 2.17: Optimized sensor placement and the associated Bayes risk obtained considering initial sensor design with three sensors concentrated at the top-right

2.6.2.4 **Comparison of a Bayesian optimized sensor placement design for different noise level in sensors**

In this section, I compare the optimal sensor placement design evaluated using approach 3 and considering initial design e_0 with $N_0 = 3$ for different noise levels as depicted in Table 4. Case 1 to 4 represents various noise levels (standard deviation in the strain measurements) in ascending order. It is observed that the number of sensors in the optimal sensor design increases with the increase in noise level in the acquired data. This is an expected result since a large amount of data is required to compensate for the increased uncertainty due to higher noise levels.

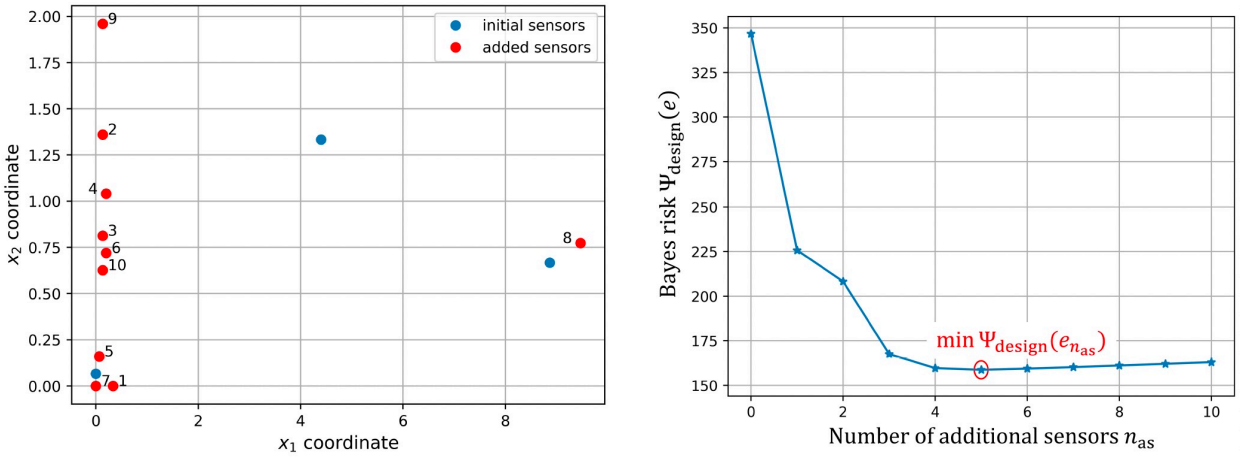
Table 2.4: Different cases of the noise level in strain gauges

Cases	Noise standard deviation σ_ε	Figure representing the resulting design	Number of sensors in optimal design
Case 1	5.0×10^{-7}	Fig. 2.18	4
Case 2	1.0×10^{-6}	Fig. 2.19	5
Case 3	2.5×10^{-6}	Fig. 2.20	9
Case 4	5.0×10^{-6}	Fig. 2.21	10



(a) Optimized sensor locations (b) Bayes risk

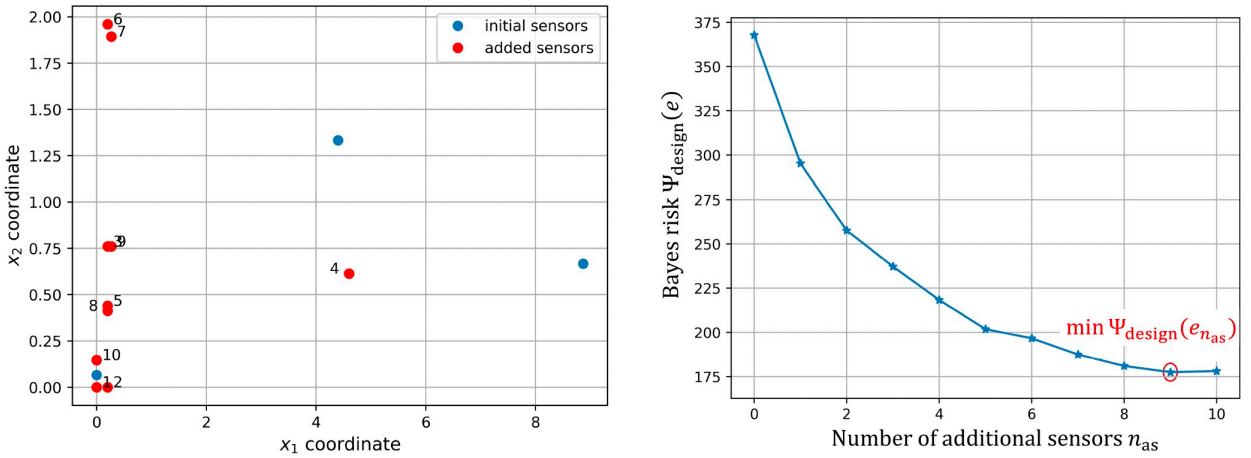
Figure 2.18: Optimized sensor placement and the associated Bayes risk obtained for case 1 of noise level



(a) Optimized sensor locations

(b) Bayes risk

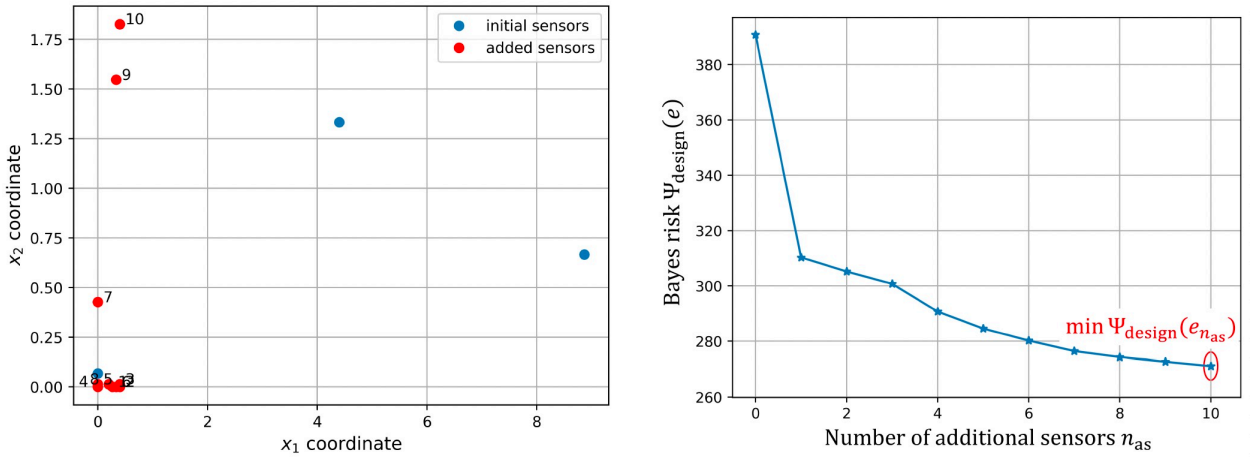
Figure 2.19: Optimized sensor placement and the associated Bayes risk obtained for case 2 of noise level



(a) Optimized sensor locations

(b) Bayes risk

Figure 2.20: Optimized sensor placement and the associated Bayes risk obtained for case 3 of noise level



(a) Optimized sensor locations

(b) Bayes risk

Figure 2.21: Optimized sensor placement and the associated Bayes risk obtained for case 4 of noise level

2.7. Summary and Conclusions

This chapter details an optimal sensor design framework for structural health monitoring applications where detection of a critical state is of prime importance. The primary contribution of the chapter is to present a sensor optimization framework and an algorithm that obtains the optimal sensor design yielding the least regrettable decision/inference of the state detection. The optimality criterion or the objective function used for optimization is the expected loss (arising as a consequence of decision making), the term also referred to as the *Bayes risk*. It is advantageous to use Bayes risk as it helps me incorporate the consequence-cost/regret of making a decision (extrinsic cost), as well as the intrinsic costs (e.g., sensor costs and their maintenance costs). A Bayes risk (or the expected loss/risk) minimized design leads to a prediction of the state that minimizes losses in an average sense.

The proposed optimal sensor placement design framework presented in this chapter can be summarized in four sequential steps as illustrated in Fig. 2.1. The optimization framework

proposed in this chapter is demonstrated on an example problem, where the existence of the boundary springs is in question (a binary decision problem). Noteworthy conclusions are: (1) Bayesian optimized sensor design is better than the random design since it leads to less expected loss/regret as a consequence of making a decision all the while using less number of sensors; (2) Generally, the Bayes risk functional has a non-linear integrand and is a high dimensional integral that demands sophisticated numerical approaches to evaluate it. Among the three approaches investigated, approach 3 (using univariate dimensional reduction with Gauss-Hermite quadrature) is the most desirable; (3) Irrespective of the initial sensor design, the proposed optimization algorithm arrives at a sensor design that is suitable for desirable decision making. In the proposed example problem, irrespective of the initial design, the optimal design consisted of additional sensor locations spread out across the left end close to the springs; (4) It is observed that the number of sensors in the optimal sensor design increases with the increase in noise level in the acquired data.

2.8. Preview to Chapter 3

This chapter has developed and explained in detail a Bayes risk-based sensor optimization framework. The framework has been implemented thoroughly to the simple 2D linear beam-spring case, which has shown the value of this design work compared to a randomly designed (or arbitrary) sensor arrangement system. However, the primary motivation of this study is to design a well-organized optimal sensor placement strategy that can be deployed directly to the SHM system in modern large-scale civil infrastructure, such as the Greenup miter gate in Mississippi. To make this framework more applicable to such infrastructure, there are still various challenges to be met. First, in this work, a discrete binary damage mode is considered for simplicity. For

most civil engineering applications, the structure of interest is nonlinear with much more complexity, and damage could occur in multiple ways. For instance, as mentioned in Section 1.1.1, a most common damage in the miter gate is the loss of contact between the gate and the wall as it is hydrostatically loaded. Instead of a discrete damage state, a continuous state damage should be considered with a more complicated calibration method. Second, to have an optimal sensor arrangement, multiple objective functions should be studied and deployed in comparison for the *best* design. Third, the miter gate has a much higher dimension in design space that could cause tremendous computational burden. More numerical strategies should be investigated to accelerate the sensor design process in a SHM system. The next chapter will present a more efficient optimal sensor design framework to overcome these challenges, and thus demonstrated through the miter gate problem in detail.

2.9. Remarks

Portions of this dissertation have been published or are currently being considered for publication. Chapter 2 is accepted for publication in Mechanical Systems and Signal Processing, the dissertation author was the primary investigator and author of this paper:

Y. Yang, M. Chadha, Z. Hu, and M. D. Todd, "An optimal sensor placement design framework for structural health monitoring using bayes risk," Mechanical Systems and Signal Processing.

Chapter 3

A PROBABILISTIC OPTIMAL SENSOR DESIGN APPROACH FOR STRUCTURAL HEALTH MONITORING USING RISK-WEIGHTED *f*-DIVERGENCE

3.1 Abstract

This chapter presents a new approach to optimal sensor design for structural health monitoring (SHM) applications using a modified f-divergence objective functional. One of the primary goals of SHM is to infer the unknown and uncertain damage state parameter(s) from the acquired data or features derived from the data. In this work, I consider the loss of boundary contact (a "gap") between a navigation lock miter gate and the supporting wall quoin block at the bottom of the gate to be the damage state parameter of concern. The design problem requires the optimal sensor placement of strain gages to obtain the best possible inference of the probability distribution of the gap length using the data from the multi-dimensional strain-gauge array. Using the notion of f-divergences (measures of difference between probability distributions), a risk-adjustment is made by using functions that weigh the importance of acquiring useful information for a given true value of the state-parameter and using Bayesian optimization. For this case study of miter gate monitoring, a computationally expensive high-fidelity finite element model and its digital surrogate is employed to provide efficient, previously validated data.

3.2 Introduction

Structural health monitoring (SHM) is a multi-part paradigm that aims at assessing the state of the structural system and its ability to perform the desired design functionality by analyzing in-

situ sensor measurement data. A well-designed SHM strategy enables the choice of optimal maintenance implementation, helps the structure achieve maximum performance, reduces ownership cost, minimizes unscheduled downtime, and potentially helps to avoid structural failures that can cause material or personal losses. Such an SHM system is desirable only if the benefits obtained from using the acquired information over the structure's life outweigh the cost of installing and maintaining that SHM system (Thöns and Faber 2013). Hence, the value of an SHM system essentially depends on its design; at the core of any well-designed SHM system is a data acquisition system that relies on (usually an array of) deployed sensors to initiate the information workflow from which ultimate decisions about operations, maintenance, and other life cycle actions will be made. Therefore, the optimal design of this sensor network-defined herein as the spatial arrangement of the sensor network-can significantly enhance the performance and life cycle value of the SHM system as a whole. Formulating and solving such an optimization problem is the central goal of this chapter. Of course, other design parameters (beyond spatial arrangements such as data acquisition rate or duty cycle) or constraints (such as power availability) will also play a role in any specific application, but such multi-objective don't fundamentally alter the ideas presented in this work.

Due to the many sources of variability and noises in any SHM system's observations, the SHM process contains inherent uncertainties that need to be considered. The optimal sensor placement problem, therefore, aims to find a sensor configuration that gathers the information most useful for detecting the target state(s) subject to uncertainty (Akbarzadeh and Lévesque 2014; Parno and Connor 2018). The key element to this is the optimality criterion or objective function that is used to evaluate design utility. However, there is not a universal objective for sensor design, as each application has a distinct goal for the use of a particular SHM system. Consequently,

different objective functions have been proposed in the past for optimal sensor placement design, starting with mode identification and correlation in some of the original works in this area, which were not necessarily SHM applications (Kammer 1991; Kammer 1996; Shi and Law 2000; Yao and Sethares 1993; Gomes and de Almeida 2019). For example, in Ref. (Kammer 1991; Gomes and de Almeida 2019), the optimal sensor arrangement minimized the condition number of the Fisher information matrix corresponding to the target modes of dynamical structures. Sun et al. (Sun and Büyüköztürk2015) proposed optimal sensor design by maximizing dynamic information of the structure using a limited number of sensors and proposed an artificial bee colony algorithm to solve the optimization problem. Austin et al. (Downey and Hu 2018) used objective function formulated to reduce the type I and II errors and used adaptive mutation-based genetic algorithm for the sensor design. Similarly, Papadimitriou et al. (Papadimitriou, Beck 2000) proposed minimizing entropy focusing on structural modal updating. In one of the first SHM-focused studies, Udwadia (Unwadia 1994) and Basseville (Basseville and Benveniste 1987) have also used the Fisher information matrix to maximize the performance of SHM for structural modal identification. For some other application domains such as the aviation sector, decision-makers are more concerned with detecting outlier states of the structure, since the cost of failure is catastrophic (Maul and Kopasakis 2008). Such maximization of outlier state detection has led to objective functions such as the probability of detection (POD), probability of classification (Peh and Liang 2007), and the Mahalanobis distance measure (Guratzsch and Mahadevan 2010). There are several other seminal contributions in optimal sensor placement design for a wide class of SHM applications found in Refs. (Chan 1997; Yi and Li 2011; Peddada and Tannous 2020).

Given that decision-makers are the typical curators of SHM utility, the objective function may also be defined from the perspective of decision theory that defines *loss as a consequence of*

decision making (or the associated risk) by considering various prior information and uncertainty sources in the decision-making process. The loss/risk is a subjective quantity and is defined according to the problem. Optimal sensor design therein requires finding the sensor network that minimizes the losses or risk expressed by an objective function in an average sense; such an objective function is defined as *Bayes risk*. This is a more general definition of traditional Bayes risk, and it expands its applicability from a pure monetary-based standpoint to a more general optimization problem in the sense that risk is no longer confined to the likelihood of losing money, but rather can be thought of as a regret of making an undesirable decision or predicting undesirable outcomes. One such Bayes risk objective function for sensor placement design was developed by Flynn and Todd (Flynn and Todd 2010; Todd and Flynn 2011) in which Bayesian experimental design (Chaloner and Verdinelli 1995) is used for optimal sensor placement design by minimizing an appropriate Bayes risk functional. It was demonstrated in an ultrasonic guided wave sensor design problem that Bayes risk can minimize the total presence of either type I or type II decision errors in SHM. The use of expected Kullback-Leibler (K-L) divergence or expected utility in sensor placement design (Li and Kiureghian 2016; Argyris and Chowdhury 2018; Hu and Ao 2017) can also be classified as a type of Bayes risk. While the idea of minimizing expected risk (or maximizing the utility of your desired outcome) using Bayes risk is powerful for optimizing sensor placement under uncertainty, its advantageously generic nature and currently unexploited benefits must be carefully considered. First, I note that the term risk is subjective and is somewhat open-ended to a desirable definition. The risk function or utility function can be formulated differently and will lead to different sensor placement designs, i.e., no utility function is generic to all problems. Second, current Bayes risk-type objective functions are incapable of incorporating human psychology or risk-perception of decision-makers in sensor network design

Using the monitoring of lock navigation gates as an application case, this chapter aims to address the two issues outlined above by expanding on the idea of Bayes risk-type objective functions to simultaneously evaluate the gain in information and consider the risk-perception of decision-makers in sensor network design. This is done by proposing a risk-weighted f-divergence functional for sensor placement design. Firstly, I investigate different types of f-divergence measures to evaluate the information gain of a particular sensor network design. Since the f-divergence gives a generic form to evaluate the distance (depicting gain information or information divergence) between two probability distribution functions, using different types of f-divergence helps to investigate and compare the effects of using different distance measures in the sensor placement optimization process. Secondly, the f-divergence is weighed with a risk-based weight function to incorporate a decision-maker's risk perception into sensor placement design. The f-divergence is modified using weight functions that weigh in the importance of acquiring good information for a given true value of the structural damage state. Thirdly, the f-divergence is also weighted by prior knowledge of the structural damage state. The proposed objective function in this chapter is, therefore, the integral of the weighted f-divergence of the posterior distribution relative to the prior distribution, weighted over the prior distribution, and risk-based weight function, integrated over all the physically possible values of the structural damage states. The goal is to obtain the sensor network that maximizes the objective function (or maximizing the gain in the additional information or minimizing the risk or regret of inferring meaningless information regarding the damage states). The chapter also proposes two different approaches to incorporate the risk weights into the Bayes risk functional. In the first approach, the risk-weights are included explicitly inside the integrand of the Bayes risk functional, whereas in the second approach, the

risk weights are used to modify the prior damage distribution. Although mathematically equivalent, numerical evaluation of risk-weighted Bayes risk yields slightly different results.

In addition to the objective function, another long-standing challenge in sensor placement design is how to effectively and efficiently solved optimization model. A common approach to optimization is iteratively searching for the optimal value guided by the steepest gradient descent This approach has been used in machine-learning (Bottou 2010) and in developing an optimal sensor network (Ram and Nedic 2007). For example, Akbarzadeh (Akbarzadeh and Lévesque 2014) used a gradient descent algorithm in sensor optimization by deriving derivatives at each step, which requires less computational effort. However, in many problems, the exact analytical derivatives are not available. Heuristic algorithms have also been widely used in the literature; for example, Jin (Jin, and Zhou 2003) used a genetic algorithm to minimize the communication distance of sensors, while Yi et al. (Yi and Li 2011) utilized a genetic algorithm to obtain optimal sensor placement for a high-rise building monitoring system. However, the main drawback of these optimization strategies is that they must evaluate many samples, yielding a computationally expensive path to the solution. In complex large-scale SHM applications such as the civil infrastructure problems considered in this chapter , the sensor design space itself is potentially prohibitively large, and this is coupled with the fact that obtaining and evaluating Bayes risk even once may require tens of thousands of runs of expensive simulations (such as if a finite element is used). This chapter also proposes a novel numerical framework that seamlessly synthesizes Gaussian process regression (Hu and Nannapaneni 2017), dimension reduction techniques (Nath and Hu 2017; Rahman and Xu 2004), Bayesian optimization (Jones and Schonlau 1998; Frazier 2018), and sequential Monte Carlo to break this computational challenge. In the proposed framework, once the desirable sensor measurements are obtained in one iteration of the

optimization, the predictive model need not be run in order to evaluate the observed sensor measurements at every optimization iteration. As shown in the result section, this significantly improves the computational efficiency of the sensor placement design optimization and allows for a reduction of computational time from years to hours. The review papers (Yi and Li 2012; Ostachowicz and Soman 2019; Tan and Zhang 2020) and the references therein serve as an excellent source of information on the optimal sensor network design and the computational methods to solve the optimization problems.

The proposed risk-weighted f-divergence functional and the efficient numerical framework to overcome the computational burden for sensor placement design will be demonstrated in a lock miter gate monitoring application. The United States Army Corps of Engineers (USACE) spends billions of dollars in maintaining and operating the USA's inland waterways navigation corridor, where the unscheduled shutdown of these assets and dewatering for inspection or repair is very costly (Parno and Connor 2018; Schwieterman and Field 2010; Foltz 2017) Within the navigation corridor, miter gates are one of the most common types of lock gates employed (Eick and Treece 2018). Many of these structures have been operational for over 50 years, and without knowledge of their actual structural residual strength capacity, they could potentially be operating with a higher risk of failure. Current practice involves engineering elicitation via inspection, followed by lock closures if the inspection so warrants. Since this process is based on the varied experience and interpretation of field inspectors, it bears high uncertainty and variability (Vega and Todd 2020), and USACE is investigating the use of SHM to potentially reduce those uncertainties. In general terms, the first step of the SHM system design is to decide what sensors are most suitable (e.g., discrete or continuous strain-gauges (Chadha and Todd 2019), accelerometers, etc.) to provide measurements that are most correlated to the type of damage or

state to be inferred. The second step is then to obtain a sensor network design (e.g., number of sensors, location/placement, duty cycle, etc.) that provides the most valuable information at a minimal cost (Padula and Kincaid 1999; Pozzi 2017), as broadly elucidated above. To this goal, in this chapter, an optimal sensor network will be designed for a miter gate using the proposed framework.

The primary contributions of this chapter may therefore be summarized as: (1) it proposes a risk-weighted f-divergence Bayes risk for sensor placement design and two different approaches to incorporate the risk weights into the Bayes risk functional; (2) it investigates and compares different types of f-divergence measures in the objective functional for sensor placement design; (3) it proposes a novel numerical framework that drastically reduces the required computational effort in sensor placement design by integrating Bayesian optimization, surrogate modeling, univariate dimensional reduction, and Sequential Monte Carlo; and (4) it demonstrates the proposed framework in a complex and practical miter gate monitoring application.

The rest of the chapter is arranged as follows. Section 3.3 details the background of the miter gate SHM application and briefly discusses the proposed sensor placement design optimization framework. Section 3.4 details the associated Bayes risk functional, followed by Section 3.5 that investigates univariate dimensional reduction with Gauss-Hermite quadrature approach to evaluate the Bayes risk. Section 3.6 discusses the optimal sensor placement design using Bayesian optimization in detail and presents the novel algorithm used to overcome the computational burden. After a general discussion on Bayesian optimization, the remaining part of Section 3.6 discusses numerical results. Finally, Section 3.7 concludes the chapter.

3.3 Problem Description

Some preliminary definitions and notations are first necessary. The real number space in d dimension is represented by $\mathbb{R}^d$, with $\mathbb{R}^1 \equiv \mathbb{R}$. A random variable X is a real-valued function defined on a discrete or a continuous sample space S_X and is assumed to take values in a measurement space $\Omega_X \in \mathbb{R}^d$, such that $X: S_X \longrightarrow \Omega_X \in \mathbb{R}^d$. Lower case letters x represent realizations of the random variable X, such that $x \in \Omega_X$. The probability density function and the cumulative density function are represented by $f_X(x)$ and $F_X(x)$. The expected value of a function $g(x)$ is denoted by $E_X[g(x)]$. Lastly, a random variable X following a Gaussian distribution, with the mean μ_x and standard deviation σ_x is denoted by:

$$\begin{gathered} f_X(x) = \frac{1}{\sigma_x}\phi\left(\frac{x-\mu_x}{\sigma_x}\right) \\ F_X(x) = \Phi\left(\frac{x-\mu_x}{\sigma_x}\right) \\ X \sim N(\mu_x, \sigma_x^2) \end{gathered} \tag{3.1}$$

No symbolic distinction is made for different dimensions d of the measurement space and the random variable. The vector-dimensionality of a random variable is contextual and is defined as needed.

Finally, let Ω_E represent the exhaustive sensor design space and $e \in \Omega_E$ represent a design realization. Let $\mathfrak{E}(e): \Omega_E \longrightarrow \mathbb{R}$ denote the Bayes risk functional. The goals of this chapter are: (1) to appropriately define the Bayes risk $\mathfrak{E}(e)$; (2) to devise a computationally-efficient approach to numerically evaluate the value of Bayes risk for a given design e; (3) to arrive at the most optimal design $e^* \in \Omega_E$, such that:

$$e^* = \arg\max_{e \in \Omega_E} \mathfrak{E}(e) \tag{3.2}$$

Evaluating the Bayes risk would require observable strain data under various damage scenarios. The sensor data is obtained using a validated finite element model. Therefore, the Bayes risk $\mathfrak{E}(e)$ and the design space Ω_E will both depend on a finite element model capable of estimating observable strains under various damage scenarios.

3.3.1 Miter gate: Finite element model

The Greenup miter gate, which is maintained and managed by USACE on the Ohio River in the USA, is considered for a case study. Fig. 3.1 shows the Greenup lock and the miter gate (image adapted from the USACE website and Eick et al. (Eick and Smith 2019)). Loss of contact in the quoin blocks is the most commonly observed damage mode in such systems (Foltz 2017; Vega and Todd 2020; Eick and Treece 2018). Loss of contact leads to a formation of a very thin gap between the gate and the wall quoin blocks at the bottom of the gate, which induces undesirable load redistribution in the system. The length of the of loss of contact at the bottom of the gate is referred to as gap length in this chapter; therefore, the gap length is considered as the continuous state-parameter $\theta \in \Omega_\Theta$ (refer to Fig. (3.2)), such that $\Omega_\Theta = \left[\theta_{\text{low}}, \theta_{\text{up}}\right]$. Here, θ_{low} is the lower bound of the gap length defining the existence of "damage", and θ_{up} is the upper bound of the gap length defining critical damage of failure. This value is suggested by the USACE engineers based on their experience, past inspection data, or numerical simulation. In most cases, data related to the failure of the structure may not be available because decision-makers are risk-averse and prevent the gap length from approaching failure levels. In such scenarios, a rigorous high-fidelity numerical simulation should be performed to estimate the $\theta_{\text{up.}}$. Based on feedback from the field-engineers (Eick and Treece 2018), the upper bound of the gap length can be considered as $\theta_{\text{up}} = 180$ inches for gates that have similar structural characteristics as the Greenup miter gate. If no value of θ_{low} is specified, it can be taken as 0 inch (indicating pristine state of the gate).

Figure 3.1: Greenup locks and miter gate

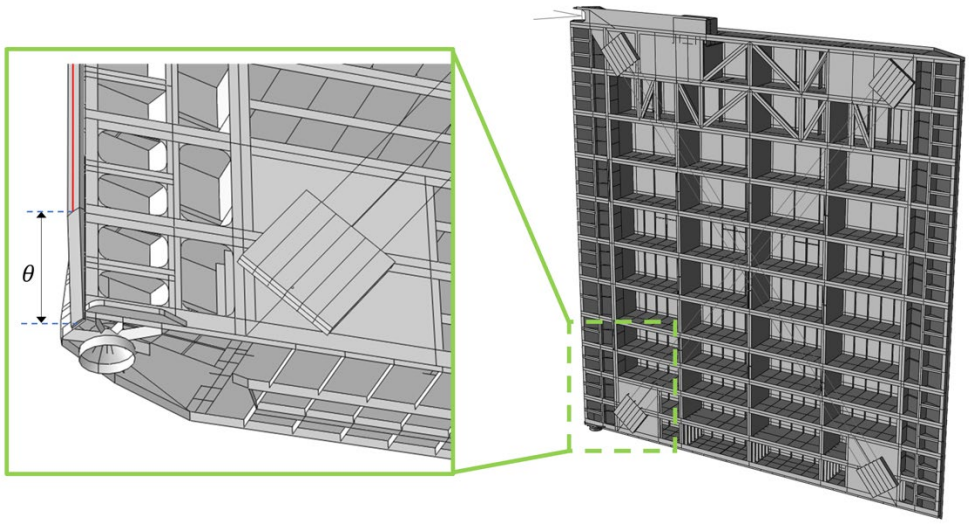


Figure 3.2: Physics-based model of miter gate and the bearing gap

The loss-of-contact part of the gate is always submerged in highly turbid water, and it consequently cannot be easily measured directly during normal operational conditions. Hence, gap length is an unknown parameter and must be inferred from indirect measurements. The Greenup miter gate is equipped with a strain gauge network that records the operational strain measurements in real-time and will be used to infer the gap length. The data acquisition process is simulated using a high-fidelity finite element model (FEM) of the Greenup miter gate previously validated in the

undamaged condition with the available strain sensor readings (Eick and Treece 2018). When the miter gate is new and pristine, the gap length could reasonably be presumed to be zero. As with any such model, its representative predictive value is only as good as its validation with regard to the real structure that it is modeling. In this case, the FEM was validated in the undamaged condition, but modeling of the damage itself could not be validated on actual data from the gate in a known damaged condition, so modeling bias in the damage state could exist. That does not invalidate the demonstration of the proposed approach or its utility but provides caution on interpreting the specific results for this case beyond the demonstration of the overall optimal sensor placement approach. With that caveat, the FEM will serve as the fundamental physical model for this study.

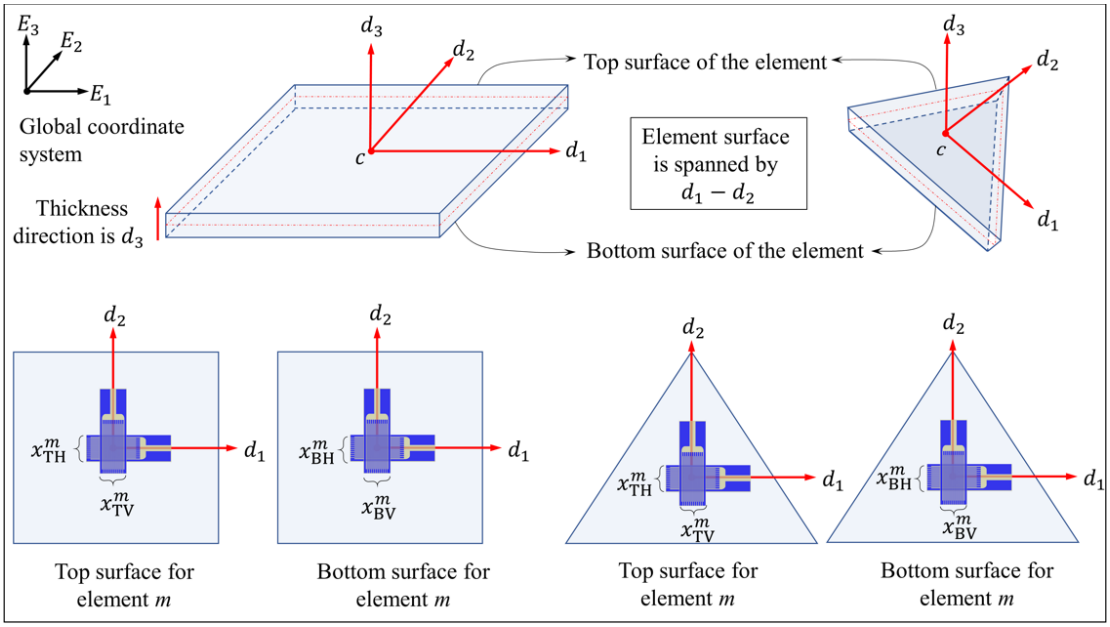


Figure 3.3: Orientation and the location of the strain gauge and different type of shell elements used in FEM.

To arrive at the optimal sensor placement design, I rely on the validated finite element model to obtain the observable strain values. A sensor-network can be designed by picking strain

gauges from a countable set of strain-locations where I have observable strain data for different damage scenarios. In my case, it is a set of 64919×4 strain locations as discussed later. Although there are infinite possible locations where strain gauges can be placed on a real miter gate, the finite element model discretely covers the possible sensor locations using a countable number of strain gauges The finite element modeling itself is constructed using 3D quadrilateral and triangular shell elements in ABAQUS and consists of a total of 64919 elements. Every element has a local coordinate system $\{d_i\}$ defined in the undeformed state, and a global coordinate system $\{E_i\}$. The thickness of the element is in the direction d_3, and the top and bottom surface of the element is spanned by the vectors $(d_1 - d_2)$ as shown in Fig. 3.3. The strain gauges are attached to the top and bottom surface of each element, measuring uniaxial strains along the direction d_1 and d_2. Each element is identified by its geometric centroid at the origin of the local coordinate system. Therefore, there are four possible arrangements of strain gauges on each element. These possibilities are identified using the following abbreviations:

TH: top element, horizontal orientation along d_1;
TV: top element, vertical orientation along d_2;
BH: bottom element, horizontal orientation along d_1;
BV : bottom element, vertical orientation along d_2. (3.3)

Based on the above abbreviations, for a typical element m, x_{TH}^m and x_{TV}^m represent the measurement of strain from gauges attached to the top surface and oriented along d_1 and d_2, respectively. Similarly, x_{RH}^m and x_{RV}^m represent the measurements of strain from gauges attached to the bottom surface and oriented along d_1 and d_2, respectively. Therefore, any element m has four candidate strain gauges attached to it, whose readings are represented by a four-dimensional

vector $x^m = (x^m_{\mathrm{TH}}, x^m_{\mathrm{TV}}, x^m_{\mathrm{BH}}, x^m_{\mathrm{BV}})$. Hence, there is a total of 64919×4 strain locations to be considered for optimal sensor design.

The gate is subjected to uncertain upstream and downstream hydrostatic loads quantified by the hydrostatic upstream and downstream heads; these are denoted by the random variables H_{up} and H_{down}, with realizations $h_{\mathrm{up}} \in \Omega_{h_{\mathrm{up}}}$ and $h_{\mathrm{down}} \in \Omega_{h_{\mathrm{down}}}$, respectively, where $\Omega_{H_{\mathrm{up}}}$ and $\Omega_{H_{\mathrm{down}}}$ represent the space of all possible values of upstream and downstream head, respectively. The water heads are modeled by a Gaussian distribution with their mean and variance reasonably assumed as

$$h_{\mathrm{up}} \sim N(552\mathrm{in}, 10^2\mathrm{in}^2) \tag{3.4a}$$

$$h_{\mathrm{down}} \sim N(168\mathrm{in}, 20^2\mathrm{in}^2) \tag{3.4b}$$

Independent zero-mean additive Gaussian noise, denoted by a random variable ζ_i with the realization ε_i, is assumed for each strain gauge,

$$\zeta_i \sim N\left(\mu_{\varepsilon_i} = 0, \sigma^2_{\varepsilon_i}\right) \tag{3.5}$$

The value ε_i represents the realization of noise, and Ω_{ζ_i} represents the noise space, such that $\varepsilon_i \in \Omega_{\zeta_i}$ The standard deviation of the noise is assigned to be $\sigma_{\varepsilon_1} = 5 \times 10^{-6}$ in accordance with reasonable commercial strain gauge performance.

The random nature of the water heads and strain gauge noise together make the observable strain values themselves random variables. Let $\Omega_X = \Omega_{X_{\mathrm{TH}}} \cup \Omega_{X_{\mathrm{TV}}} \cup \Omega_{X_{\mathrm{BH}}} \cup \Omega_{X_{\mathrm{BV}}}$ be the set of all the possible 64919×4 strain gauge locations. Here, $\Omega_{X_{\mathrm{TH}}}$ and $\Omega_{X_{\mathrm{TV}}}$ represents the space of all strain gauges attached to the top surface of element measuring strain in the direction d_1, and d_2 respectively. Similarly, $\Omega_{X_{\mathrm{BH}}}$ and $\Omega_{X_{\mathrm{BV}}}$ represents the space of all strain gauges attached to the

bottom surface of element measuring strain in the direction d_1, and d_2 respectively. Let X denote the random vectors consisting of all strain measurements corresponding to Ω_X space, such that $x \in \Omega_X$ represent realizations of the random vectors X (see Eq. (3.9) for the relationship between the observed strain realization x, the strain output of FEM, and the noise in strain gauge). Finally, I denote the true values of the gap length, and hydrostatic heads as: $\theta_{\text{true}} \in \Omega_\Theta, h_{\text{up-true}} \in \Omega_{H_{\text{up}}}$, $h_{\text{down-true}} \in \Omega_{H_{\text{down}}}$.

3.3.2 Miter gate: Surrogate model

Solving the optimization problem posed in Eq. (3.2) requires evaluating the Bayes risk (defined later in Section (3.4)) for various sensor network designs to arrive at an optimal design. As described below, each evaluation of the Bayes risk involves numerous FEM predictions to solve a Bayesian inference problem. This is computationally intractable, and I seek computationally efficient approximations to the FEM. For Bayesian calibration, metamodels or surrogate models are preferable, e.g., Support Vector Regression (SVR) (Moustapha and Bourinet 2018), Gaussian Process Regression (GPR) (Moustapha and Bourinet 2018; Frazier 2018), Neural Network (Yu and Wang 2009), and Polynomial Chaos Expansion (PCE) (Capellari and Chatzi 2018). Some such approaches like SVR or neural networks yield point estimates/prediction, while others like GPR also predict the uncertainties associated with an average estimate/prediction. GPR is used to build a surrogate model in this work, which turns out to be 50,000 times faster than the FEM model. The output of the surrogate model still has a very large dimension; this is addressed using principal components analysis, which can be efficiently computed using the *singular value decomposition* (SVD) that reduces the high-dimensional, highly-correlated output space to low-dimensional, uncorrelated features This is analogous to the "linear model of coregionalization" in the Gaussian process literature. Of the possible 64919×4 strain readings in Ω_X, 64919 strain

measurements corresponding to each of the spaces $\Omega_{X_{\mathrm{TH}}}, \Omega_{X_{\mathrm{TV}}}, \Omega_{X_{\mathrm{BH}}}$ and $\Omega_{X_{\mathrm{BV}}}$ are considered independently. The 64919-dimensional strain response corresponding to each of the spaces $\Omega_{X_{\mathrm{TH}}}, \Omega_{X_{\mathrm{TV}}}, \Omega_{X_{\mathrm{BH}}}$ and $\Omega_{X_{\mathrm{BV}}}$ are transformed to lower $7,12,7,12$ feature spaces, respectively. Equivalently, the 64919×4-dimensional strain response corresponding to the spaces Ω_X are transformed to a lower 38-dimensional feature space that covers 95% of the total information of the strain data. I realize that the vertically oriented strain measurements have a larger number of features (12 features for both top and bottom strain gauges) than the horizontally oriented strain measurements (7 features for both top and bottom strain gauges) One possible reason is that the vertical strain responses are more sensitive to the dynamic loading considered in this chapter (hydrostatic upstream and downstream loading) than their counterparts in the horizontal direction, and hence require a larger number of features. Inversely, the larger number of features required to represent vertical strain gauge measurements also implies that the vertically oriented strain measurements have higher complexity (that by itself is a subjective quantity as described in (Lloyd 2001)) than the horizontally oriented strain measurements. These 38 features can be inverted to obtain the complete strain gauge response. Four surrogate models for $7,12,7,12$ features corresponding to four strain measurement spaces $\Omega_{X_{\mathrm{TH}}}, \Omega_{X_{\mathrm{TV}}}, \Omega_{X_{\mathrm{BH}}}$ and $\Omega_{X_{\mathrm{BV}}}$ were built using GPR I used a squared exponential kernel and I evaluated the hyper-parameters using maximum likelihood estimation. Since the GP models for each of these features were trained independently, they have different hyper-parameters. One-third of the 1000 data points were randomly used for training the GPR, and the remaining two-thirds were used for validation to verify the accuracy of the surrogate. Fig. 3.4 illustrates the discussion carried out so far. Like the FEM, the GPR model yields the 64919×4 dimensional strain response, but at a much cheaper computational cost.

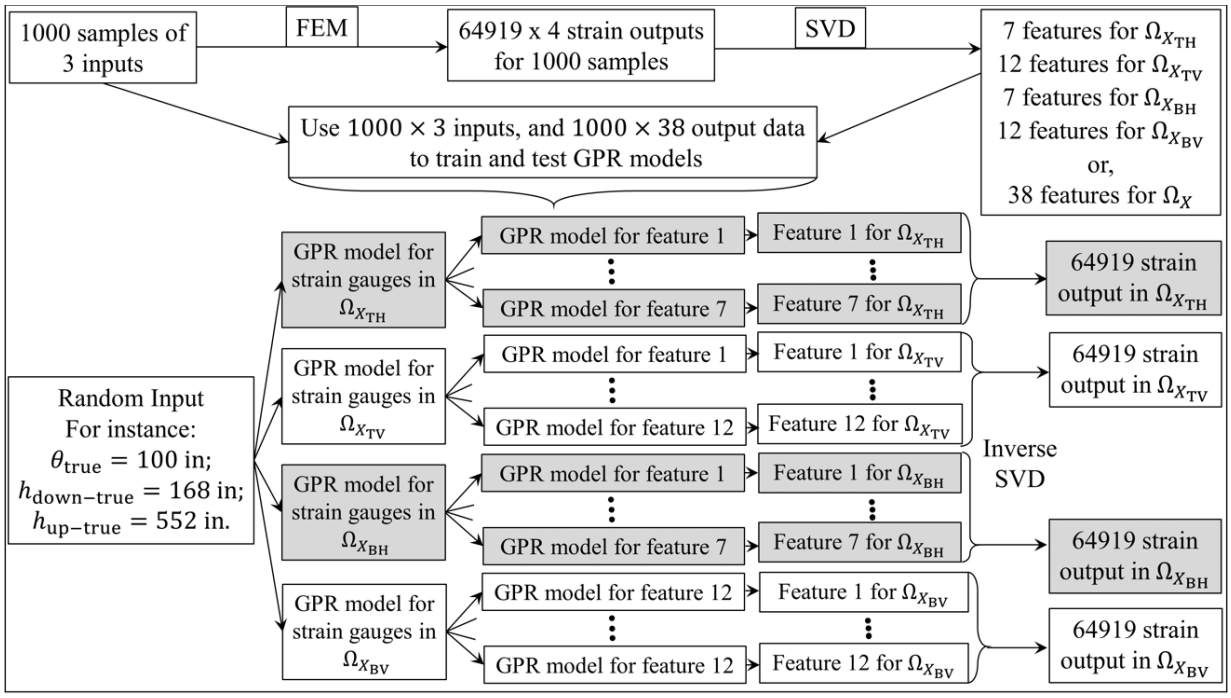


Figure 3.4: Flowchart describing strain data generation using FEM, and prediction using GPR surrogate model

3.3.3 Brief introduction to risk-weighted f-divergence based Bayesian optimization workflow for sensor placement design

Given the overall objective of an optimal sensing design, Ω_E represents the design space, such that $e \in \Omega_E$ represents a particular design realization. The design e consists of $N_{\mathrm{sg}}(e)$ number of strain gauge measurement locations. Every design e yields different measurement data $x_e \in \Omega_{X_e}$. Here, $\Omega_{X_a} \subset \Omega_X$ represents the measurement space for the design e, and X_e denotes the corresponding random variable (see Eq. (3.10)) that depicts how the observable strain realization x_e is obtained from the FEM and the strain gauge noise). Having defined the design space, four prominent steps are summarized below for the proposed risk-weighted f-divergence based Bayesian optimization framework for sensor network design.

Step 1: Problem description

This chapter's objective could be phrased as attempting to answer the question: "Given sources of uncertainty (noise in the sensors and the uncertain external conditions), which set of sensors should be chosen among the possible 64919×4 strain gauges measurements that yields the maximum relative gain in the information contained in the posterior distribution of the target damage (gap length) relative to the information contained in the prior distribution?"

Consider a sensor network design $e \in \Omega_E$ with the measurement space Ω_{X_e}. Before any new/additional information is available about the structure through the strain gauge measurements, the uncertainty in the gap length is described by its prior probability distribution $f_\Theta(\theta)$. When additional information or strain gauge measurements are observed, it informs the observer (or the engineer) about the current state of the structure. This new information translates into the further refinement of the understanding of the gap length, now described by its posterior distribution $f_{\Theta|X_e}(\theta \mid x_e)$. If the strain gauge readings are representative of the true state of the structure, the posterior distribution $f_{\Theta|X_e}(\theta \mid x_e)$ draws closer to an understanding of the true description of the gap length as compared to the prior distribution. Mathematically, then, the goal here is to obtain the sensor-design $e^* \in \Omega_E$ that yields the maximum relative gain in the information contained in the $f_{\Theta|X_e}(\theta \mid x_e)$ relative to the information contained in $f_\Theta(\theta)$.

I very briefly describe the remaining three steps of optimization next. The details are omitted in this brief description because each of these steps demands a complete section on its own.

Step 2: Definition of the design dependent Bayes risk functional

The next step of the Bayesian optimization is to define the optimality criterion or the objective functional, which is otherwise known as Bayes risk. The Bayes risk is a function of the design $e \in \Omega_E$ and is denoted by $\mathfrak{E}(e)$. Bayes risk is a problem-dependent functional. I aim to define Bayes risk such that:

1 The Bayes risk guides me to obtain a sensor-design that maximizes information gains on the gap length inferred from the sensor measurements. The gain in the information is quantified by the f-divergence that evaluates the similarity between two probability measures.

2 The Bayes risk incorporates the desire to obtain better information/description of the gap length when the structure approaches a higher degree of damage (an increased gap length approaching some critical size). This is accomplished by using a risk-based weight function.

3 The Bayes risk also takes into account prior knowledge of the gap length.

The Bayes risk for this chapter is defined as the integral of the weighted f-divergence of the posterior distribution relative to the prior distribution, weighted over the prior distribution, and risk-based weight function, integrated over all the physically possible values of the gap length. Section 3.4 is dedicated to detailing the Bayes risk functional.

Step 3: Evaluation of the design-dependent Bayes risk functional

Bayes risk is a non-linear functional. For a given design e, evaluating the Bayes risk requires one to obtain the posterior of the gap length and risk-based weight functions. Section 3.4.3

delineated a line of reasoning for incorporating risk-based weight function in the definition of the Bayes risk. Theoretically, the posterior distribution can be evaluated using Bayes' theorem:

$$f_{\Theta|X_e}(\theta \mid x_e) = \frac{f_{X_e|\Theta}(x_e \mid \theta) f_\Theta(\theta)}{f_{X_e}(x_e)} \tag{3.6}$$

The quantity $f_\Theta(\theta)$ represents the prior probability of the true state, and in absence of any information, it may be assumed to be a uniform (uninformed) distribution. The likelihood $f_{X_e|\Theta}(x_e \mid \theta)$ is obtained using either a physics-based model or a digital twin. I note that the posterior does not follow a canonical distribution and expectations with respect to the posterior in (6) cannot be obtained analytically. This is because the relationship between strain measurements x_ρ and the gap length θ is highly nonlinear. I deploy numerical approximation of the posterior distribution by using particle filters, or specifically a sequential Monte Carlo (SMC) approach.

The integrand in the Bayes risk expression is integrated over the gap length (defined later in Section 3.4). The second difficulty in obtaining Bayes risk is to evaluate this integral. To approximate the integral, I first change the variable of the integral from the measurement space to the uncertain input space. For instance, in my problem, the hydrostatic heads and the noise in the strain values are the primary random input variables. Since there is a unique one-to-one relationship between an input realization and an output realization, this allows me to change the variables of integration in the Bayes risk. The integral can then be numerically approximated using univariate dimensional reduction and Gauss-Hermite quadrature. Section 3.5 discusses the approach to evaluate the Bayes risk.

Step 4: Obtaining the optimal sensor design using Bayesian optimization

With the problem well-defined (step 1) and the associated Bayes risk optimality criterion formulated (steps 2-3), the problem becomes: "Given $\Omega_\Theta, \Omega_{X_e}, \Omega_E, \Omega_\zeta, \Omega_{h_{\text{up}}}, \Omega_{h_{\text{down}}}$, given an assumed uncertainty structure (as in Eq.(3.4) and (3.5)), what is the design $e^* \in \Omega_E$ that maximizes the Bayes risk objective functional $\mathfrak{E}(e)$? " I carefully note that intuitively, "risk" must be minimized However, in this chapter , Bayes risk represents relative gain in information, and therefore, must be maximized. I could have called the objective function "Bayes utility", but as noted in the introduction, I take advantage of the fact that "*risk*" is a subjective quantity.

I very briefly detail the sensor optimization algorithm, which will be explained in great depth in Section 3.6. I start with an initial design e_0 consisting of N_0 number of sensors. To obtain the optimal design e_1 with $(N_0 + 1)$ sensors, I search the entire design space for the $(N_0 + 1)^{\text{th}}$ sensor location. The $(N_0 + 1)^{\text{th}}$ sensor location that maximizes the acquisition function constitutes the next additional sensor. In this chapter, I use expected improvement (Mockus and Tiesis 1978, Jones, and Schonlau 1998) as the acquisition function. Similarly, I repeat the optimization process to arrive at the optimal design $e_{n_{as}}$ consisting of $N_0 + n_{\text{as}}$ sensors (or n_{as} number of additional sensors relatively to the initially assumed design e_0). Finally, I pick $e^* = \arg\max_{e_{n_{0s}}} \mathfrak{E}(e_{n_{\text{as}}})$ as the most optimal design, where $\mathfrak{E}(e_{n_{\text{as}}})$ represents the Bayes risk associated with the design $e_{n_{as}}$. Section 3.6 details the Bayesian optimization algorithm for optimal sensor placement.

3.3.4 Bayesian inference of gap length for a given sensor design

As discussed in the previous section, the state parameter is the gap length θ, and for a given sensor-design e, the measurement vector $x_e \in \Omega_{X_e}$ is the strain recorded at $N_{\text{sg}}(e)$ number of

strain gauge locations. Therefore, X_e is also a random vector. The measurements obtained from the strain gauges are used to infer the gap length θ using Eq. (3.6). In the context of inferring θ, the evidence $f_{X_e}(x_e)$ is just a normalizing constant. Therefore, Eq. (3.6) may be written as:

$$f_{\Theta|X_e}(\theta \mid x_e) \propto f_{X_e|\Theta}(x_e \mid \theta) f_{\Theta}(\theta) \tag{3.7}$$

The distribution $f_{\Theta}(\theta)$ reflects the prior knowledge about the parameter θ before any new information/measurements are obtained. Assuming only basic geometrical constraints on the gap length, I assume the prior to be a uniform distribution spanning over $\Omega_{\Theta} = [\theta_{\text{low}}, \theta_{\text{up}}]$, such that

$$f_{\Theta}(\theta) = \begin{cases} (\theta_{\text{up}} - \theta_{\text{low}})^{-1}, & \theta \in \Omega_{\Theta} \\ 0, & \text{otherwise} \end{cases} \tag{3.8}$$

Evaluating the posterior using Eq. (3.7) requires me to obtain the likelihood $f_{X_e|\Theta}(x_e \mid \theta)$. Constructing the likelihood $f_{X_e|\Theta}(x_e \mid \theta)$ requires a model of the measurement process. I note that the design e consists of the selected sensors chosen from a total of 64919×4 possibilities, or $\Omega_{X_e} \subset \Omega_X$. For a given gap length and the hydrostatic heads, the FEM or GPR yields strain values for all of the 64919×4 sensors. Therefore, to detail the measurement model, I consider the total measurement space Ω_X. In this chapter, I use the following measurement model

$$x = g(\theta, h_{\text{up}}, h_{\text{down}}) + \varepsilon \tag{3.9}$$

In the equation above, $x = (x_1, x_2, \cdots, x_{64919\times4}) \in \Omega_X$ is a realization of the random vector X consisting of 64919×4 strain measurements, where x_i represents the strain value corresponding to the i^{th} strain gauge. For a given gap length θ, the digital surrogate g yields $g(\theta, h_{\text{up}}, h_{\text{down}}) = \left(g_1(\theta, h_{\text{up}}, h_{\text{down}}), g_2(\theta, h_{\text{up}}, g_{\text{down}}), \cdots, g_{64919\times4}(\theta, h_{\text{up}}, h_{\text{down}})\right)$ at 64919×4 location of the strain gauges, where $g_i(\theta, h_{\text{up}}, h_{\text{down}})$ represents the strain response of the i^{th} strain gauge

obtained by the GPR surrogate model g. In the equation above $\varepsilon \in \Omega_\zeta$ is the realization of the random vector ζ defining the noise in 64919×4 sensors.

Equation (3.9) defines the measurement model considering all the sensor locations in Ω_X. However, a design e consists of only $N_{\text{sg}}(e)$ sensors with the measurement space Ω_{X_e}. Utilizing the fact that $\Omega_{X_e} \subset \Omega_X$, let g_e define the true strain response for the sensors included in design e obtained by the GPR model, such that $g_e(\theta, h_{\text{up}}, h_{\text{down}}) = \left(g_{e_1}(\theta, h_{\text{up}}, h_{\text{down}}), \cdots, g_{e_{N_{\text{sg}(e)}}}(\theta, h_{\text{up}}, h_{\text{down}})\right)$ Similarly, let $x_e = \left(x_{e_1}, \cdots, x_{e_{N_{sg(e)}}}\right) \in \Omega_{X_e}$ denotes the observed/measured strain response. Here, $g_{e_i}(\theta, h_{\text{up}}, h_{\text{down}})$ and x_{e_i} represents the true (obtained by the GPR model) and the observed strain response of the i^{th} strain gauge in the sensor design e, respectively. The measurement model for the strain gauges included in the design e is given by

$$x_e = g_e(\theta, h_{\text{up}}, h_{\text{down}}) + \varepsilon_e \tag{3.10}$$

In the equation above, x_e is one of the realizations of the random vector X_e. The vector ε_e is the realization of the random vector ζ_e with $\varepsilon_e = \left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{N_{\text{sg}}(e)}\right)$. It represents the measurement noise/error vector for the design e, where ε_i denotes the error between the measurement output and GPR predicted response (assumed to be the true response) corresponding to the i^{th} strain gauge in the design e as defined in Eq. (3.5). I assume that ε_o follows a zero-mean Gaussian distribution with independent components, i.e., the noise/error terms of all $N_{\text{sg}}(e)$ strain gauges are assumed to be statistically independent. In addition, I assume that each strain gauge has same standard deviation σ_{ε_i}, such that

$$f_{\zeta_e}\left(\varepsilon_e = \left(\varepsilon_1, \cdots, \varepsilon_{N_{\mathrm{sg}}(e)}\right)\right) = \prod_{i=1}^{N_{\mathrm{sg}}(e)} f_{\zeta_i}(\varepsilon_i) = \prod_{i=1}^{N_{\mathrm{sg}}(e)} \frac{1}{\sigma_{\varepsilon_i}} \phi\left(\frac{\varepsilon_i}{\sigma_{\varepsilon_i}}\right) \tag{3.11}$$

Using the measurement model defined in Eq. (3.10), and the description of noise in Eq. (3.11), the likelihood of observing the strain measurement $x_e \in \Omega_{X_e}$ for the gap length θ can be written as

$$f_{X_e|\Theta}(x_e \mid \theta) = \prod_{i=1}^{N_{\mathrm{sg}}(e)} \frac{1}{\sigma_{\varepsilon_i}} \phi\left(\frac{x_{e_i} - g_{e_i}\left(\theta, h_{\mathrm{up}}, h_{\mathrm{down}}\right)}{\sigma_{\varepsilon_i}}\right) \tag{3.12}$$

Having defined the prior and the likelihood in Eq. (3.8) and (3.12), I note that the posterior cannot be obtained analytically using Eq. (3.6). This is because the relationship between strain measurements x_e and the gap length θ is highly nonlinear. One can rely on numerical approximation techniques like Markov chain Monte Carlo (MCMC) methods, particle filter, and sequential Monte Carlo (SMC) approach in recursive mode to solve the inference problem. As mentioned, because the evaluation of the likelihood $f_{X\wedge\Theta}(x_e \mid \theta)$ at numerous values of θ using the full finite element model was too expensive, this was achieved by running instead the the GPR model g_e. Furthermore, I employ the particle filter method (or sequential Monte Carlo (SMC) in recursive mode) to obtain the posterior.

I also simulate the measurement data numerically. For simulating such data, I obtain the response of the digital surrogate $g\left(\theta_{\text{true}}, h_{\text{up-true}}, h_{\text{down-true}}\right)$ parameterized by a chosen/fixed value of true gap length θ_{true} subjected to chosen/fixed input loading $\left(h_{\text{up-true}}, h_{\text{down-true}}\right)$. This strain gauge response is now corrupted by Gaussian noise of standard deviation σ_{ε_i} to mimic the real-

world measurement noise. This corrupted strain response is now used as the measurement/observed data $x_e \in \Omega_{X_e}$

Obtaining the posterior numerically using particle filtering requires evaluating the likelihood at numerous values of the gap length, called the particles. Usually, particle filtering is used for sequential updating of the posterior distribution for a dynamic system, i.e., the case where new information on the system is available as time evolves. However, in this case, I just have one set of data, and I aim at obtaining the posterior in a single step. The following summarizes the process of obtaining the posterior distribution of the gap length:

(a) For the assumed true gap length θ_{true} and the chosen/fixed input loading ($h_{\text{up-true}}$, $h_{\text{down-true}}$), simulate the observed/measurement strain data $x_e \in \Omega_{X_e}$.

(b) I choose $N_{\text{particle}} = 5000$ discrete values of the gap length, called particles. It is at these 5000 particles for the likelihood is to be evaluated.

(c) At each of these 5000 particles (or the gap length), and for a given loads ($h_{\text{up-true}}$, $h_{\text{down-true}}$), the GPR model yields the true strain value at the i^{th} strain location for the gap length particle θ_j is denoted by $g_{e_i}(\theta_j, h_{\text{up-true}}, h_{\text{down-true}})$, where $i \leq N_{\text{sg}}(e)$ and $j \leq 5000$. I obtain the numerical value of the likelihood of the measurement given the gap length θ_j at the observed strain $x_e \in \Omega_{X_e}$ using Eq. (3.12) as:

$$f_{X_e|\Theta}(x_e \mid \theta_j) = \prod_{i=1}^{N_{\text{sg}}(e)} \frac{1}{\sigma_{\varepsilon_i}} \phi\left(\frac{x_{e_i} - g_{e_i}(\theta_j, h_{\text{up}}, h_{\text{down}})}{\sigma_{\varepsilon_i}}\right) \tag{3.13}$$

In the equation above, x_{e_i} represents the observed strain value at the i^{th} strain gauge.

(d) Evaluate the weight w_j for each particle as:

$$w_j = \frac{f_{X_e|\Theta}(x_e \mid \theta_j)}{\sum_{k=1}^{N_{\text{paricle}}} f_{X_e|\Theta}(x_e \mid \theta_k)} \tag{3.14}$$

(e) Calculate the cumulative weights to observe big jumps. Resample the weighted particles to obtain unweighted samples of the posterior distribution over gap length.

It is evident from discussion carried above that evaluating the posterior distribution of the gap length θ for a given sensor measurements X_e requires obtaining the likelihood $f_{X_e|\Theta}(x_e \mid \theta_j)$ for $j \leq N_{\text{particles}}$ This requires running the GPR model $N_{\text{particles}}$ times.

3.4 The objective functional, Bayes risk

3.4.1 Bayes risk: Expected utility function

Recall Step 2 of the Bayesian optimization framework discussed in section 3.3.3. There are three primary goals that I aim to achieve: (1) Maximize the relative gain in the information; (2) Obtain better information/description of the gap length when the true value of gap length is larger, or the state of the structure approaches a higher degree of damage; (3) Include the prior knowledge of the gap length.

3.4.1.1 Relative gain of information: f-divergence

We start with the first goal. As discussed before, the prior distribution of gap length $f_\Theta(\theta)$ quantifies my understanding of the gap length when no additional/new information on the system (through the strain gauge measurements) is available. When the new information or the sensor measurements $x_e \in \Omega_{X_e}$ are available, Bayesian inference allows me to refine my understanding of the gap length, now quantified by the posterior $f_{\Theta|X_e}(\theta \mid x_e)$. In this section, for brevity, I denote

the posterior distribution as $f_{\Theta|X_e}(\theta \mid x_e) = g_\Theta(\theta)$. Traditionally, the information divergence (similarity or dissimilarity) between two distributions, for example, deviation of the posterior $g_\Theta(\theta)$ from the prior $f_\Theta(\theta)$, is quantified by the relative entropy or Kullback-Leibler (KL) divergence or i-divergence (i for information), defined by:

$$D_{KL}(g_\Theta(\theta) \parallel f_\Theta(\theta)) = \int_{\Omega_\Theta} g_\Theta(\theta) \log\left(\frac{g_\Theta(\theta)}{f_\Theta(\theta)}\right) \mathrm{d}\theta. \tag{3.15}$$

Since KL divergence is not symmetric, I prefer this form of information divergence because it quantifies the information gain in the posterior distribution of the gap length as compared to the prior distribution (refer to (Lindley 1956, Huan and Marzouk 2013)). Although KL divergence measures the distance between two probability distributions, it does not qualify as a statistical metric of spread because it violates the symmetric property and triangular inequality. However, KL divergence does satisfy the other two properties of a metric: non-negativity and the identity of indiscernible. Therefore, KL divergence of probability distributions may be loosely interpreted as a nonsymmetric analog of squared Euclidean distance. Like KL divergence, there are many other divergences used to evaluate the similarity and dissimilarity between probability distributions. Many of these divergences can be unified under the generic framework of f-divergence (Csiszár and Shields 2004). Therefore, i-divergence is a special case of f-divergence. For a convex function $\ell(t)$ defined for $t > 0$, with $\ell(1) = 0$, the f-divergence of the posterior $g_\Theta(\theta)$ from the prior $f_\Theta(\theta)$ is defined by:

$$D_\ell(g_\Theta(\theta) \parallel f_\Theta(\theta)) = \int_{\Omega_\Theta} f_\Theta(\theta) \ell\left(\frac{g_\Theta(\theta)}{f_\Theta(\theta)}\right) \mathrm{d}\theta \tag{3.16}$$

Note that the constraint $\ell(1) = 0$ implies that all the f-divergences satisfy the identity of indiscernible. Table 3.1 lists some of the important and commonly used f-divergences; more

information may be found in (Csiszár and Shields 2004; Verdu 2016; Nielsen and Nock 2013; Osterreicher and Vajda 2003). Among all the f-divergences listed in Table 1 (see appendix 7), only the total variance satisfies all the properties of a metric: non-negativity, symmetry, the identity of indiscernible, and triangular inequality (Khosravifard and Fooladivanda 2007).

Table 3.1: Common f-divergences

Types of f-divergences	Denoted by	Function $f(t)$
Kullback-Leibler	$D_{KL}\left(g_\Theta(\theta)\|\|f_\Theta(\theta)\right)$	$t\log t$
Reverse Kullback-Leibler	$D_{KL}\left(f_\Theta(\theta)\|\|g_\Theta(\theta)\right)$	$-\log t$
Pearson Chi Square	$D_{\chi_P^2}\left(g_\Theta(\theta)\|\|f_\Theta(\theta)\right)$	$(t-1)^2$
Neyman Chi Square	$D_{\chi_N^2}\left(g_\Theta(\theta)\|\|f_\Theta(\theta)\right)$	$\frac{(1-t)^2}{t}$
Pearson-Vajda	$D_{\chi_P^k}\left(g_\Theta(\theta)\|\|f_\Theta(\theta)\right)$	$(t-1)^k$
Squared Hellinger	$D_{H^2}\left(g_\Theta(\theta)\|\|f_\Theta(\theta)\right)$	$(\sqrt{t}-1)^2$
Total variation	$D_\delta\left(g_\Theta(\theta)\|\|f_\Theta(\theta)\right)$	$\frac{1}{2}\|t-1\|$
K-divergence	$D_K\left(g_\Theta(\theta)\|\|f_\Theta(\theta)\right)$	$t\log\left(\frac{2t}{t+1}\right)$
Skewed K-divergence	$D_{K_\alpha}\left(g_\Theta(\theta)\|\|f_\Theta(\theta)\right)$	$t\log\left(\frac{t}{1+\alpha(1-t)}\right)$
Jensen-Shannon	$D_{JS}\left(g_\Theta(\theta)\|\|f_\Theta(\theta)\right)$	$\frac{1}{2}\left(t\log t-(t+1)\log\left(\frac{t+1}{2}\right)\right)$
α Jensen-Shannon	$D_{JS_\alpha}\left(g_\Theta(\theta)\|\|f_\Theta(\theta)\right)$	$\frac{1}{2}\left(t\log t-(t+1)\log\left(1+\alpha(t-1)\right)\right)$
α-divergence	$D_\alpha\left(g_\Theta(\theta)\|\|f_\Theta(\theta)\right)$	$\begin{cases}\frac{4}{1-\alpha^2}\left(1-t^{\frac{1+\alpha}{2}}\right) & \alpha\neq\pm1;\\ t\log t & \alpha=1;\\ -\log t & \alpha=-1.\end{cases}$
f_β-divergence	$D_{f_\beta}\left(g_\Theta(\theta)\|\|f_\Theta(\theta)\right)$	$\begin{cases}\frac{\beta}{\beta-1}\left((1+t^\beta)^{\frac{1}{\beta}}-2^{\left(\frac{1}{\beta}-1\right)}(1+t)\right) & \beta\in\mathbb{R}\backslash\{1\};\\ t\log t-(1+t)\log\frac{1+t}{2} & \beta=1;\\ \frac{1}{2}\|t-1\| & \beta=\infty.\end{cases}$

In this chapter, the state parameter (the gap length) is a single-dimensional quantity. However, in many problems, the state parameter is a multi-dimensional vector. In such scenarios, evaluating the f divergence becomes computationally expensive. Many approximation techniques for f-divergence have been proposed, like using higher-order Chi distances (Nielsen and Nock 2013), penalized convex risk minimization (Nguyen and Wainwright 2008), and random mixture estimator (Rubenstein and Bousquet2019). For completion's sake and for ensuring generality, I

briefly present approximating f-divergence using Taylor series expansion and higher-order Chi distances (Nielsen and Nock 2013). The function $\ell(t)$ can be expanded about the point t_0 using the Taylor series as

$$\ell(t) = \sum_{i=0}^{\infty} \frac{1}{i!}(t - t_0)^i \cdot \frac{\partial^i \ell(t)}{\partial t^i}\bigg|_{t=t_0} \tag{3.17}$$

The f-divergence defined in Eq. (3.16) can be written as:

$$\begin{aligned} D_\ell(g_\Theta(\theta) \parallel f_\Theta(\theta)) \quad &= \int_{\Omega_\Theta} f_\Theta(\theta) \sum_{i=0}^{\infty} \frac{1}{i!}\left(\frac{g_\Theta(\theta)}{f_\Theta(\theta)} - t_0\right)^i \cdot \frac{\partial^i \ell(t)}{\partial t^i}\bigg|_{t=t_0} \mathrm{d}\theta \\ &= \sum_{i=0}^{\infty} \frac{1}{i!}\frac{\partial^i \ell(t)}{\partial t^i}\bigg|_{t=t_0} \left(\int_{\Omega_\Theta} \frac{(g_\Theta(\theta) - t_0 f_\Theta(\theta))^i}{f_\Theta(\theta)^{i-1}} \mathrm{d}\theta\right) \\ &= \sum_{i=0}^{\infty} \frac{1}{i!}\frac{\partial^i \ell(t)}{\partial t^i}\bigg|_{t=t_0} . D_{\chi^i_{P,t_0}}(g_\Theta(\theta) \parallel f_\Theta(\theta)) \end{aligned} \tag{3.18}$$

Here, $D_{\chi^i_{P\ell_0}}(g_\Theta(\theta) \parallel f_\Theta(\theta))$ is the generalization of the i^{th} order Pearson-Vajda f-divergence. The equation above allows me to write any f-divergence as the weighted sum of the generalized i^{th} order Pearson-Vajda f-divergence, which in turn can be approximated by the restricted class of exponential families that are easy to evaluate (Nielsen and Nock 2013).

3.4.1.2 Implicit and explicit inclusion of the risk weights into Bayes risk

The space of all the uncertainties in the current problem is defined as $\Omega_{\xi_e} = \Omega_{H_{\text{up}}} \times \Omega_{H_{\text{down}}} \times \Omega_{\zeta_1} \times \Omega_{\zeta_2} \times \cdots \times \Omega_{\zeta_{N_{\text{sg}(e)}}}$, such that the random variable ξ_e represents all the uncertainty sources considered to affect the design e. Let $\beta_e = \left(h_{\text{up}}, h_{\text{down}}, \varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{N_{\text{ca}\,(e)}}\right) \in$

Ω_{ε_e} represents a realization of the random variable ξ_e. Secondly, since the strain measurements $x_e \in \Omega_{X_e}$ are representative of the physics of the miter gate, its value also depends on the gap length value $\theta \in \Omega_\Theta$. This fact is mathematically denoted by redefining the random variable X_e to be a function of the uncertainties and the true gap length, $x_e = X_e(\theta_{\text{true}}, \beta_e)$. If there is no external noise and if the true value of gap length θ_{true} exactly known, then x_e represents true value of the strain measurements. However, since the true gap length can't be obtained under all the inevitable uncertainty, the best one can do is to define the Bayes risk as the expected value of risk-weighted f-divergence averaged over the entire space Ω_Θ and Ω_{ξ_e}, i.e., by considering the entire range of possible true values of the gap length and taking into account the uncertainties in strain gauge readings and external loads. I reasonably assume that the random variables $\Theta, H_{\text{up}}, H_{\text{down}}$, and ζ_i are statistically independent. With all the necessary pieces defined, I first state the Bayes risk functional without including any risk weights as

$$\mathfrak{E}(e) = \int_{\Omega_\Theta} \int_{\Omega_{\xi_e}} f_{\xi_e}(\beta_e) f_\Theta(\theta) D_\ell\big(f_{\Theta|X_e(\theta_{\text{true}}=\theta,\beta_e)}(\varphi \mid x_e) \parallel f_\Theta(\varphi)\big) \mathrm{d}\beta_e \, \mathrm{d}\theta \tag{3.19}$$

In the equation above, the variable φ gets integrated out in the expression of f-divergence. The f-divergence is a function of $(\theta_{\text{true}} = \theta, \beta_e)$, in the sense that

$$\begin{aligned} D_\ell\big(f_{\Theta|X_e(\theta_{\text{true}}=\theta,\beta_e)}(\varphi \mid x_e) || f_\Theta(\varphi)\big) &= \int_{\Omega_\theta} f_\Theta(\varphi) \ell\left(\frac{f_{\Theta|X_e(\theta_{\text{true}}=\theta,\beta_e)}(\varphi \mid x_e)}{f_\Theta(\varphi)}\right) \mathrm{d}\varphi \\ &= \mathcal{D}_\ell(\theta_{\text{true}} = \theta, \beta_e) \end{aligned} \tag{3.20}$$

In the Eq. (3.19) and (3.20), the f-divergence $\mathcal{D}_\ell(\theta_{\text{true}} = \theta, \beta_e)$ measures the divergence in the posterior distribution in the gap length relative to its prior distribution.

I consider two approaches to incorporate the risk weights in the Bayes risk. In the first approach, I explicitly weigh the integrand of the Bayes risk defined in Eq. (3.19) with the risk weights, such that

$$\mathfrak{E}_{\text{explicit-risk}}(e) = \int_{\Omega_\theta}\int_{\Omega_{\xi_e}} f_{\xi_e}(\beta_e) f_\Theta(\theta) r(\theta_{\text{true}} = \theta)\mathcal{D}_\ell(\theta_{\text{true}} = \theta, \beta_e)\mathrm{d}\beta_e\ \mathrm{d}\theta \tag{3.21}$$

The quantity $r(\theta_{\text{true}} = \theta)$ weighs the risk-based importance factor for all the possible value of true gap length, i.e. $\forall\theta_{\text{true}} = \theta \in \Omega_\Theta$. The prior $f_\Theta(\theta)$ accounts for the prior knowledge of the gap length, and the distribution $f_{\xi_e}(\beta_e)$ accounts for all the uncertainties. Finally, I define the utility function $\mathcal{L}(\theta_{\text{true}} = \theta, \beta_e)$ as the risk-weighted f-divergence

$$\mathcal{L}(\theta_{\text{true}} = \theta, \beta_e) = r(\theta_{\text{true}} = \theta)\mathcal{D}_\ell(\theta_{\text{true}} = \theta, \beta_e) \tag{3.22}$$

I understand that in the definition of the Bayes risk, I consider all the possible values of the true gap length. From here on, I omit writing $\theta_{\text{true}} = \theta$ in the argument of utility or the weight function, such that Eq. (3.22) becomes

$$\mathcal{L}(\theta, \beta_e) = r(\theta)\mathcal{D}_\ell(\theta, \beta_e) \tag{3.23}$$

In the equation above, the argument θ represents one possibility of true gap length.

With the simplified notation of f-divergence and definition of the utility function, $\mathcal{L}(\theta, \beta_e)$, and the Bayes risk explicitly considering the risk weights is compactly written as

$$\mathfrak{E}_{\text{explicit-risk}}(e) = E_{\Theta\xi_e}[\mathcal{L}(\theta, \beta)] = \int_{\Omega_\Theta}\int_{\Omega_{\xi_e}} f_\Theta(\theta) f_{\xi_e}(\beta_e) r(\theta)\mathcal{D}_\ell(\theta, \beta_e)\mathrm{d}\beta_e\ \mathrm{d}\theta \tag{3.24}$$

Another mathematically equivalent approach to consider risk weighing is by modifying the prior distribution to

$$f_{\hat{\Theta}}(\hat{\theta} = \theta) = \frac{f_{\Theta}(\theta)r(\theta)}{\int_{\theta_{\text{low}}}^{\theta_{\text{up}}} f_{\Theta}(\theta)r(\theta)\mathrm{d}\theta} \tag{3.25}$$

such that $f_{\hat{\Theta}}(\hat{\theta})$ is transformed prior probability distribution function of the random variable $\Theta \longrightarrow \hat{\Theta}$ with the realization $\hat{\theta} \in \Omega_{\hat{\Theta}}$, such that $\Omega_{\Theta} = \Omega_{\hat{\Theta}}$. The Bayes risk that implicitly incorporates the risk weight in the form of modified prior distribution is defined as

$$\mathfrak{E}_{\text{implicit-risk}}(e) = E_{\hat{\Theta}\xi_e}[\mathcal{D}_{\ell}(\hat{\theta}, \beta_e)] = \int_{\Omega_{\hat{\theta}}} \int_{\Omega_{\xi_e}} f_{\hat{\Theta}}(\hat{\theta}) f_{\xi_e}(\beta_e) \mathcal{D}_{\ell}(\hat{\theta}, \beta_e) \mathrm{d}\beta_e \, \mathrm{d}\hat{\theta} \tag{3.26}$$

The Bayes risk functional $\mathfrak{E}_{\text{implicit-risk}}(e)$ implicitly considers the risk-weight $r(\theta)$. I note that the implicit Bayes risk $\mathfrak{E}_{\text{implicit-risk}}(e)$ and the explicit Bayes risk $\mathfrak{E}_{\text{explicit-risk}}(e)$ are proportional to each other, with a positive constant of proportionality, i.e., $\left(\int_{\theta_{\text{low}}}^{\theta_{\text{up}}} f_{\Theta}(\theta)r(\theta)\mathrm{d}\theta\right)$:

$$\mathfrak{E}_{\text{implicit-risk}}(e) \propto \mathfrak{E}_{\text{explicit-risk}}(e) \tag{3.27}$$

Therefore, either explicit or implicit Bayes risk can be used in the optimization problem. Finally, I note that the implicit and explicit Bayes risk converge when I assign a constant risk-weight (importance factor) to all the true gap length values, such that

$$\mathfrak{E}(e) = \mathfrak{E}_{\text{implicit-risk}}(e)\big|_{r(\theta)=1} = \mathfrak{E}_{\text{explicit-risk}}(e)\big|_{r(\theta)=1} \tag{3.28}$$

As shown in Eq. (3.24) Bayes risk is defined as the expected value of the utility function. Ideally, the goal is to maximize the utility, but due to the uncertainties in the system quantified by $\beta_e \in \Omega_{\xi_e}$ and my inability to know the true value of the gap length, the best I can do is to pick a sensor design that maximizes the expected value of the utility averaged over all the possible values of the true gap length and the uncertainties. An optimal sensor design that maximizes the expected utility is the most optimal. The next Section 3.4.2 discusses the quantity $r(\theta_{\text{true}})$ in detail.

Evaluating the f-divergence $\mathcal{D}_\ell(\theta, \beta_e)$ for a given true gap length $\theta_{\text{true}} = \theta$ and the external uncertainties β_e (consisting of hydrostatic heads and the noise in strain gauge readings) requires me to obtain the posterior distribution $f_{\Theta|X_e(\theta,\beta_e)}(\varphi \mid x_e)$ of the gap length for a given true-variables (θ, β_e). Therefore, as was mentioned above, for a given measurement $X_e(\theta, \beta_e)$, obtaining the posterior using particle filter requires running the GPR or digital surrogate model $N_{\text{particles}}$ times. However, since I am simulating the measurement data $X_e(\theta, \beta_e)$, I need to run GPR model once as mentioned in Eq. (3.10). Therefore, evaluating the f-divergence $\mathcal{D}_\ell(\theta, \beta_e)$ (or the utility $\mathcal{L}(\theta, \beta_e)$) requires running the GPR model ($N_{\text{particles}} + 1$) times.

For the prior $f_\Theta(\varphi)$ and the posterior distribution $f_{\Theta|X_e(\theta,\beta_e)}(\varphi \mid x_e)$, the f-divergence is numerically evaluated by approximating Eq. (3.20) as:

$$\mathcal{D}_\ell(\theta, \beta_e) \approx \frac{1}{N}\sum_{i=1}^{N} r(\varphi_i)\ell\left(\frac{f_{\Theta|X_e(\theta,\beta_e)}(\varphi_i \mid x_e)}{f_\Theta(\varphi_i)}\right) \tag{3.29}$$

Fig. 3.5 illustrates a function or a module called "Evaluate the Utility $\mathcal{L}(\theta, \beta_e)$ "or" Evaluating the f-divergence $\mathcal{D}_\ell(\theta, \beta_e)$ " that obtains the f-divergence $\mathcal{D}_\ell(\theta, \beta_e)$ and the utility $\mathcal{L}(\theta, \beta_e)$ for a given design e and the input variables (θ, β_e). Note that obtaining utility function is just a step away from the f-divergence. It does so in a three-step process that requires running the GPR model ($N_{\text{particles}} + 1$) times. The first step is to simulate observed/measurement data x_e for a design e by assuming a true gap length θ, hydraulic heads $(h_{\text{up}}, h_{\text{down}})$, and a noise structure $f_{\zeta_i}(\varepsilon_i)$. The second step is to obtain the posterior distribution of the gap length given the measurement x_e obtained in the first step using particle filter. Finally, the third step is to evaluate f-divergence of the posterior (obtained in step 2) relative to the prior distribution of the gap length that ultimately yields the utility function.

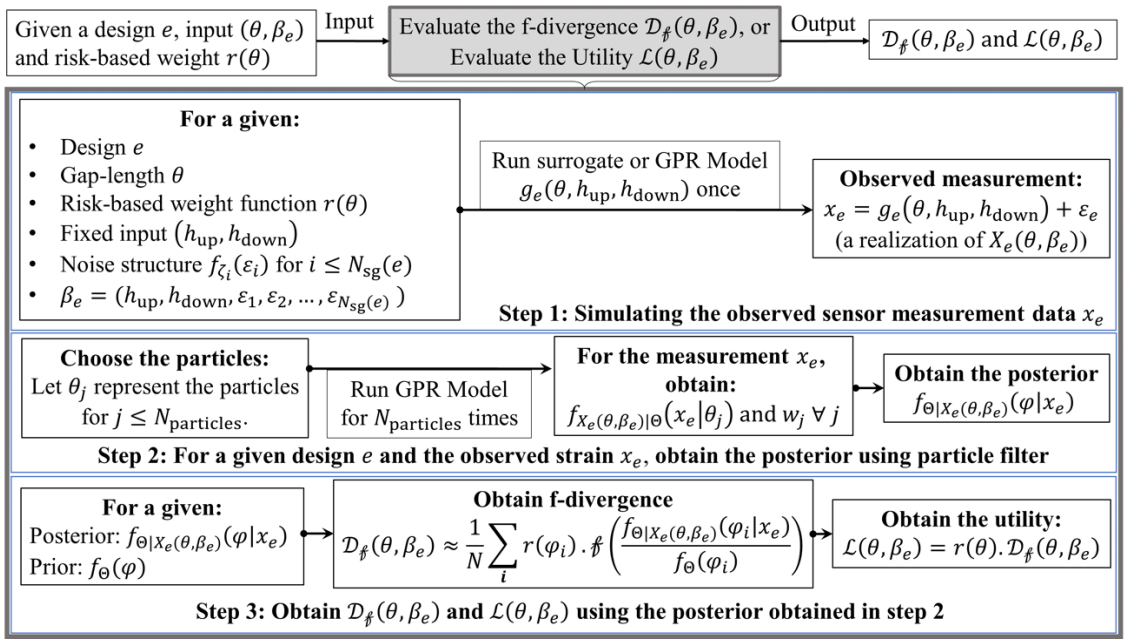


Figure 3.5: Evaluating the f-divergence $\mathcal{D}_{\mathbb{f}}(\theta, \beta_e)$ and the utility $\mathcal{L}(\theta, \beta_e)$ for a given case of true gap length θ, uncertaintiesβ_e, and the risk-based weight function $r(\theta)$.

3.4.2 Risk-based weight function

In some applications, there is a well-defined limit state that defines damage criticality, and a given structural owner/stakeholder might want an SHM system to determine the proximity to that limit state. In the current application, there is no well-defined limit state, and the risk weight function $r(\theta_{\text{trme}})$ can serve as a surrogate to that proximity by assigning relative importance to the values of true gap length θ_{true} in terms of the degree of damage. For instance, $r(\theta_{\text{true}} = \theta_2) > r(\theta_{\text{true}} = \theta_1)$ implies that the structural owner believes a gap length of θ_2 is more concerning with regard to criticality than when the true gap length is θ_1. This approach is inspired by the fact that different decision-makers mentally assign a different importance factor (or in economic terms: utility or risk-intensity) to the seriousness/urgency to take necessary actions with the increasing intensity of structural damage. For instance, given that the true state of the structure is moderately damaged, a decision-maker who is fearful of making any mistake leading to heavy losses, a risk-

averter, might suggest major repair and continuous inspection. On the other hand, another decision-maker might suggest only minor repairs, a risk-seeker. Commensurate with the notion of subjective risk perception, I suggest risk weights $r(\theta_{\text{true}})$ that follow the following two properties:

1 The risk weights $r(\theta_{\text{true}})$ must have a zero or positive slope. This is because, physically, an increase in gap length reflects higher damage to the structural state of the miter gate. Therefore. to satisfy this physical constraint, I can either assign constant or monotonically increasing risk weights for all the gap lengths.

2 The risk weight need not be unique and can be selected based on a desirable optimization criterion. For example, if I decide to equally weigh all the values of true gap length, then the risk weight can be taken as a constant $r(\theta_{\text{true}}) = 1$ (zero slope). On the other hand, if I desire to make a better prediction of the state of the miter gate at a higher gap length value that implies higher damage, I may pick a monotonically increasing risk weight.

Since the state estimation depends on the probabilistic description of the gap length, obtaining a better estimate of the miter gate damage intensity demands a better estimate of the probability distribution of the gap length. I aim at assigning an increasing importance factor to damage estimation as the value of the gap length increases. Therefore, for the sake of the optimization problem considered in this chapter, I consider a particular case of monotonically increasing risk-weight of the following form:

$$r(\theta_{\text{true}}) = e^{-\left(\frac{\theta_{\text{ruve}} - \theta_{\text{cricical}}}{b}\right)^2}, \text{ for } \theta_{\text{low}} < \theta_{\text{true}} < \theta_{\text{up}} = \theta_{\text{critical}} \tag{3.30}$$

In the equation above, θ_{critical} represents the critical value of gap length such that as the true gap length θ_{true} approaches this critical value θ_{critical}, the risk-weight increases. I consider $\theta_{\text{critical}} =$

θ_{up} The factor b controls how quickly the risk-weight decays as θ_{true} deviates from θ_{critical}. Now with the Bayes risk functional fully defined, the next section deals with evaluating the Bayes risk.

3.5 Evaluating the Bayes risk for a fixed sensor design

3.5.1 Revisiting Bayes risk

Section 3.4.1 defined the utility function $\mathcal{L}(\theta, \beta_e)$, and the explicit Bayes risk was defined as the expected utility, i.e., averaged over all the values of the uncertainties $\beta_e \in \Omega_{\xi_e}$, and the possible true values of gap length $\theta \in \Omega_\Theta$. I note that these random variables constituting β_e and θ (or $\hat{\theta}$ when Bayes risk considers the risk-weights implicitly) can follow a generic continuous distribution I can always transform them to a standard normal random variable. Therefore, in an attempt to generalize, I transform the true gap length θ (or $\hat{\theta}$), hydrostatic heads h_{up} , h_{down} , and the noise ζ_i into their respective standard normal forms denoted by a tilde ($\tilde{\cdot}$) over the respective quantity. Since the hydrostatic heads and noise for the i^{th} strain gauge is Gaussian in my case, their standard normal forms can be written as $\tilde{h}_{\text{up}}$, and $\tilde{h}_{\text{down}}$, such that $h_{\text{up}} = \tilde{h}_{\text{up}}\, \sigma_{h_{\text{up}}} + \mu_{h_{\text{upp}}}$, $h_{\text{down}} = \tilde{h}_{\text{down}}\, \sigma_{h_{\text{down}}} + \mu_{h_{\text{down}}}$, and $\varepsilon_i = \tilde{\varepsilon}_i \sigma_{\varepsilon_i} + \mu_{\varepsilon_j}$, where $\tilde{h}_{\text{up}}, \tilde{h}_{\text{down}}$ and $\tilde{\varepsilon}_i$ are the realizations of standard normal random variables $\tilde{H}_{\text{up}}, \tilde{H}_{\text{down}}$, and $\tilde{\zeta}_i$ respectively. I transform the prior $f_\Theta(\theta)$ and the modified prior $f_{\hat{\Theta}}(\hat{\theta})$ to a standard normal random variable $\tilde{\Theta}$ and $\tilde{\hat{\Theta}}$ respectively, such that the cumulative density functions are equal: $F_\Theta(\theta) = F_{\tilde{\Theta}}(\tilde{\theta})$, and $F_{\hat{\Theta}}(\hat{\theta}) = F_{\tilde{\hat{\theta}}}(\tilde{\hat{\theta}})$. This transforms ξ_e into a joint standard normal random variable $\tilde{\xi}_e$ (with a realization $\tilde{\beta}_e$), such that

$$f_{\tilde{\xi}_e}(\tilde{\beta}_e) = f_{\tilde{H}_{\text{up}}}(\tilde{h}_{\text{up}}) \cdot f_{\tilde{H}_{\text{down}}}(\tilde{h}_{\text{down}}) \cdot \prod_{i=1}^{N_{\text{sg}}(e)} f_{\tilde{\xi}_i}(\tilde{\varepsilon}_i), \text{ where,}$$
$$\tilde{\beta}_e = \left(\tilde{h}_{\text{up}}, \tilde{h}_{\text{down}}, \tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \cdots, \tilde{\varepsilon}_{N_{\text{sg}}(e)}\right) \tag{3.31}$$

We can now rewrite Bayes risk in Eq. (3.24) as:

$$\mathfrak{E}_{\text{explicit-risk}}(e) = E_{\tilde{\Theta}\tilde{\xi}_e}\left[\tilde{\mathcal{L}}(\tilde{\theta}, \tilde{\beta}_e)\right], \text{ where, } \tilde{\mathcal{L}}(\tilde{\theta}, \tilde{\beta}_e) = \mathcal{L}(\theta, \beta_e) \tag{3.32a}$$

$$\mathfrak{E}_{\text{implicit-risk}}(e) = E_{\tilde{\hat{\theta}}\tilde{\xi}_e}\left[\tilde{\mathcal{D}}_\ell\left(\tilde{\hat{\theta}}, \tilde{\beta}_e\right)\right], \text{ where, } \tilde{\mathcal{D}}_\ell\left(\tilde{\hat{\theta}}, \tilde{\beta}_e\right) = \mathcal{D}_\ell(\hat{\theta}, \beta_e) \tag{3.32b}$$

The next section approximates the Bayes risk defined in Eq. (3.32) by using univariate dimensional reduction and Gauss-Hermite quadrature to carry out the integration.

3.5.2 Univariate dimensional reduction and Gauss-Hermite quadrature

To obtain the optimal sensor placement design, I may either optimize $\mathfrak{E}_{\text{explicit-risk}}(e)$ or $\mathfrak{E}_{\text{implicit-risk}}(e)$. Since these are both integrals, I will use Gauss-Hermite quadrature to approximate the Bayes risk. In this section, I in parallel detail the numerical approximation of $\mathfrak{E}_{\text{explicit-risk}}(e)$ or $\mathfrak{E}_{\text{implicit-risk}}(e)$. Recall that the vector $\tilde{\beta}_e$ consists of $\left(N_{\text{sg}}(e) + 2\right)$ variables. To catalyze the derivation to estimate the Bayes risk using univariate dimensional reduction and Gauss-Hermite quadrature, I define the following spaces

$$\Omega_{\tilde{\Psi}_e} = \Omega_{\tilde{\Theta}} \times \Omega_{\tilde{\xi}_e}, \text{ such that } \tilde{\psi}_e = (\tilde{\theta}, \tilde{\beta}_e) = \left(\tilde{\theta}, \tilde{h}_{\text{up}}, \tilde{h}_{\text{down}}, \tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \cdots, \tilde{\varepsilon}_{N_{\text{sg}}(e)}\right) \in \Omega_{\tilde{\Psi}_e} \tag{3.33a}$$

$$\Omega_{\tilde{\hat{\Psi}}_e} = \Omega_{\tilde{\hat{\Theta}}} \times \Omega_{\tilde{\xi}_e}, \text{ such that } \tilde{\hat{\psi}}_e = \left(\tilde{\hat{\theta}}, \tilde{\beta}_e\right) = \left(\tilde{\hat{\theta}}, \tilde{h}_{\text{up}}, \tilde{h}_{\text{down}}, \tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \cdots, \tilde{\varepsilon}_{N_{\text{sg}}(e)}\right) \in \Omega_{\tilde{\hat{\Psi}}_e} \tag{3.33b}$$

To distinguish between a variable with or without the hat ($\hat{*}$), refer to Eq. (3.25) for the definition of transformed gap length $\hat{\theta}$ used in the expression of implicit-risk Bayes risk. Equation set (33) allows me to write the Bayes risk in a more desirable form

$$\mathfrak{E}_{\text{explicit-risk}}(e) = E_{\tilde{\Psi}_e}\left[\tilde{\mathcal{L}}(\tilde{\psi}_e)\right] = \int_{\Omega_{\tilde{\psi}_e}} f_{\tilde{\Psi}_e}(\tilde{\psi}_e)\tilde{\mathcal{L}}(\tilde{\psi}_e)\mathrm{d}\tilde{\psi}_e, \text{ where, } \tilde{\mathcal{L}}(\tilde{\psi}_e) = \tilde{\mathcal{L}}(\tilde{\theta}, \tilde{\beta}_e) \quad (3.34a)$$

$$\begin{aligned}\mathfrak{E}_{\text{implicit-risk}}(e) = E_{\tilde{\tilde{\Psi}}_e}\left[\tilde{\mathcal{D}}_\ell\left(\tilde{\tilde{\psi}}_e\right)\right] &= \int_{\Omega_{\tilde{\tilde{\Psi}}_e}} f_{\tilde{\tilde{\Psi}}_e}\left(\tilde{\tilde{\psi}}_e\right)\tilde{\mathcal{D}}_\ell\left(\tilde{\tilde{\psi}}_e\right)\mathrm{d}\tilde{\tilde{\psi}}_e, \text{ where, } \tilde{\mathcal{D}}_\ell\left(\tilde{\tilde{\psi}}_e\right) \\ &= \tilde{\mathcal{D}}_\ell\left(\tilde{\tilde{\theta}}, \tilde{\beta}_e\right)\end{aligned} \quad (3.34b)$$

Like Fig. (3.5), Fig. (3.6) illustrates a function or a module called "Evaluate the Utility $\tilde{\mathcal{L}}(\tilde{\theta}, \tilde{\beta}_e)$" and

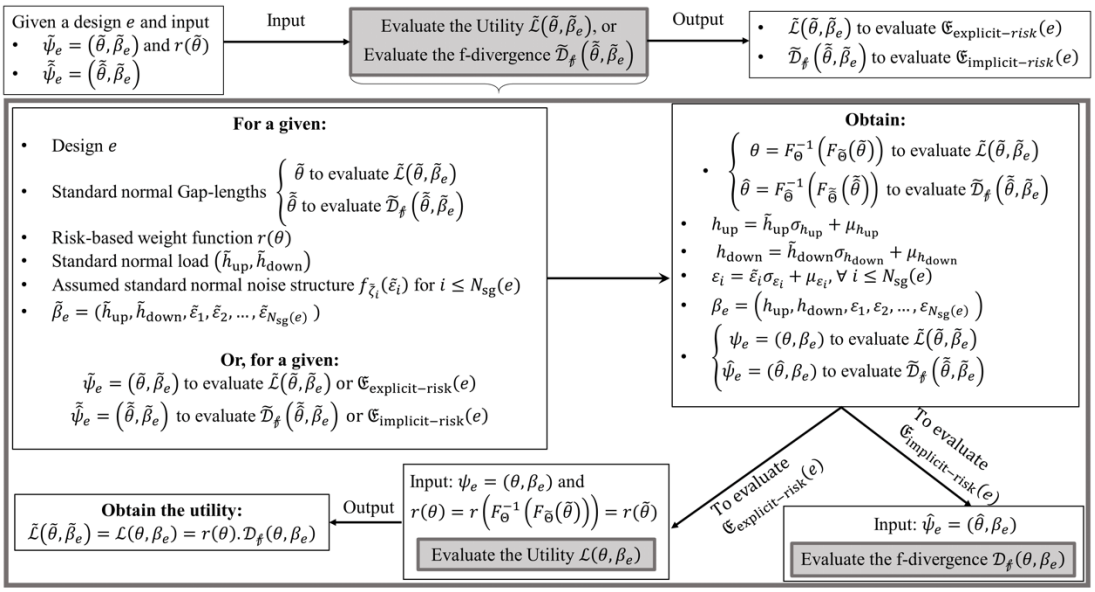


Figure 3.6: Evaluating the utility $\tilde{\mathcal{L}}(\tilde{\psi}_e) = \tilde{\mathcal{L}}(\tilde{\theta}, \tilde{\beta}_e)$ or the f-divergence $\tilde{\mathcal{D}}_\ell\left(\tilde{\tilde{\psi}}_e\right) = \tilde{\mathcal{D}}_\ell\left(\tilde{\tilde{\theta}}, \tilde{\beta}_e\right)$ for a given $(\tilde{\theta}, \tilde{\beta}_e)$, or $\left(\tilde{\tilde{\theta}}, \tilde{\beta}_e\right)$, respectively

"Evaluate the f-divergence $\tilde{\mathcal{D}}_\ell\left(\tilde{\tilde{\theta}}, \tilde{\beta}_e\right)$ " that obtains the utility $\tilde{\mathcal{L}}(\tilde{\psi}_e)$ or the f-divergence $\tilde{\mathcal{D}}_\ell\left(\tilde{\tilde{\psi}}_e\right)$ for a given design e and the input variables $(\tilde{\theta}, \tilde{\beta}_e)$ or $\left(\tilde{\tilde{\theta}}, \tilde{\beta}_e\right)$ respectively. The module "Evaluate the Utility $\tilde{\mathcal{L}}(\tilde{\theta}, \tilde{\beta}_e)$ " is required to evaluate $\mathfrak{E}_{\text{explicit-risk}}(e)$, and the module "Evaluate the f-

divergence $\tilde{D}_{\ell}\left(\tilde{\tilde{\psi}}_e\right)$ " is required to evaluate $\mathfrak{E}_{\text{implicit-risk}}(e)$. It does so by transforming the standard-normal variables back to their original form, i.e., $\left(\tilde{\theta}, \tilde{\beta}_e\right) \longrightarrow (\theta, \beta_e)$ or $\left(\tilde{\tilde{\theta}}, \tilde{\beta}_e\right) \longrightarrow \left(\hat{\theta}, \beta_e\right)$, and then using the module "Evaluate the Utility $\mathcal{L}(\theta, \beta_e)$ "or" Evaluate the f-divergence $\mathcal{D}_{\ell}\left(\hat{\theta}, \beta_e\right)$ " (illustrated in Fig. (3.5)) to obtain the respective quantities.

The integrals in Equations (3.34a) and (3.34a) are high dimensional expectations in $\left(N_{\text{sg}}(e) + 3\right)$ dimensional spaces, making classic multivariate quadrature rules (e.g., quasi Monte Carlo or Smolyak sparse grids) prohibitively expensive. Monte Carlo approximations converge slowly and require a large number of samples to approximate the expectations. This is problematic because an expensive Bayesian inference problem needs to be solved to evaluate the integrands in Equations (3.34a) and (3.34a). To overcome these issues, I employ an approximation to the integrals in Equations (3.34a) and (3.34a) based on univariate dimension reduction. To do so, I define the following vectors, each consisting of $\left(N_{\text{sg}}(e) + 3\right)$ components:

$$
\begin{aligned}
&\tilde{\psi}_0 = (0,0,0,0,0,\cdots,0)\\
&\tilde{\psi}_1 = (\tilde{\theta},0,0,0,0,\cdots,0)\\
&\tilde{\psi}_2 = \left(0,\tilde{h}_{\text{up}},0,0,0,\cdots,0\right)\\
&\tilde{\psi}_3 = \left(0,0,\tilde{h}_{\text{down}},0,0,\cdots,0\right)\\
&\tilde{\psi}_4 = (0,0,0,\tilde{\varepsilon}_1,0,\cdots,0)\\
&\tilde{\psi}_5 = (0,0,0,0,\tilde{\varepsilon}_2,\cdots,0)\\
&\vdots\\
&\tilde{\psi}_{(N_{\text{sg}}(e)+3)} = \left(0,0,0,0,0,\cdots,\tilde{\varepsilon}_{N_{\text{sg}}(e)}\right)
\end{aligned}
\tag{3.35}
$$

$$
\begin{aligned}
\tilde{\tilde{\psi}}_0 &= (0,0,0,0,0,\cdots,0) \\
\tilde{\tilde{\psi}}_1 &= (\tilde{\tilde{\theta}},0,0,0,0,\cdots,0) \\
\tilde{\tilde{\psi}}_2 &= \left(0,\tilde{h}_{\mathrm{up}},0,0,0,\cdots,0\right) \\
\tilde{\tilde{\psi}}_3 &= \left(0,0,\tilde{h}_{\mathrm{down}},0,0,\cdots,0\right) \\
\tilde{\tilde{\psi}}_4 &= (0,0,0,\tilde{\varepsilon}_1,0,\cdots,0); \\
\tilde{\tilde{\psi}}_5 &= (0,0,0,0,\tilde{\varepsilon}_2,\cdots,0) \\
&\vdots \\
\tilde{\tilde{\psi}}_{(N_{\mathrm{sg}}(e)+3)} &= \left(0,0,0,0,0,\cdots,\tilde{\varepsilon}_{N_{\mathrm{sg}}(e)}\right)
\end{aligned}
\tag{3.36}
$$

Note that $\tilde{\psi}_i$ are exactly same as $\tilde{\tilde{\psi}}_i$ except for $i = 1$. Using the definitions above and univariate dimensional reduction (refer to (Rahman and Xu 2004)), I approximate the utility function $\tilde{\mathcal{L}}(\tilde{\psi}_e)$ or the f-divergence $\tilde{\mathcal{D}}_\ell\left(\tilde{\tilde{\psi}}_e\right)$ as:

$$
\tilde{\mathcal{L}}(\tilde{\psi}_e) \approx -\left(N_{\mathrm{sg}}(e)+2\right)\tilde{\mathcal{L}}(\tilde{\psi}_0) + \sum_{i=1}^{(N_{\mathrm{sg}}(e)+3)} \tilde{\mathcal{L}}(\tilde{\psi}_i) \tag{3.37a}
$$

$$
\mathcal{D}_\ell\left(\tilde{\tilde{\psi}}_e\right) \approx -\left(N_{\mathrm{sg}}(e)+2\right)\tilde{\mathcal{D}}_\ell\left(\tilde{\tilde{\psi}}_0\right) + \sum_{i=1}^{(N_{\mathrm{sg}}(e)+3)} \tilde{\mathcal{D}}_\ell\left(\tilde{\tilde{\psi}}_i\right) \tag{3.37b}
$$

Substituting Eq. (3.37a) into Eq. (3.32a), I get,

$$
\begin{aligned}
\mathfrak{E}_{\text{explicit-risk}}(e) &\approx -\left(N_{\mathrm{sg}}(e)+2\right)E_{\tilde{\Psi}_e}\left[\tilde{\mathcal{L}}(\tilde{\psi}_0)\right] + \sum_{i=1}^{(N_{\mathrm{sg}}(e)+3)} E_{\tilde{\Psi}_e}\left[\tilde{\mathcal{L}}(\tilde{\psi}_i)\right] \\
&= -\left(N_{\mathrm{sg}}(e)+2\right)\int_{\Omega_{\tilde{\psi}_e}} f_{\tilde{\Psi}_e}(\tilde{\psi}_e)\tilde{\mathcal{L}}(\tilde{\psi}_0)\mathrm{d}\tilde{\psi}_e + \sum_{i=1}^{(N_{\mathrm{sg}}(e)+3)} \int_{\Omega_{\tilde{\psi}_e}} f_{\tilde{\Psi}_e}(\tilde{\psi}_e)\tilde{\mathcal{L}}(\tilde{\psi}_i)\mathrm{d}\tilde{\psi}_e \\
&= -\left(N_{\mathrm{sg}}(e)+2\right)\tilde{\mathcal{L}}(\tilde{\psi}_0) + \sum_{i=1}^{(N_{\mathrm{sg}}(e)+3)} \int_{\Omega_{\tilde{\psi}_e}} f_{\tilde{\psi}_e}(\tilde{\psi}_e)\tilde{\mathcal{L}}(\tilde{\psi}_i)\mathrm{d}\tilde{\psi}_e
\end{aligned}
$$

(3.38)

Similarly, substituting Eq. (3.37b) into Eq. (3.32b), I get,

$$\mathfrak{E}_{\text{implicit-risk}}(e) \approx -\big(N_{\text{sg}}(e)+2\big)\tilde{\mathcal{D}}_{\ell}\left(\tilde{\tilde{\psi}}_0\right) + \sum_{i=1}^{\left(N_{\text{sg}}(e)+3\right)} \int_{\Omega_{\tilde{\tilde{\Psi}}_e}} f_{\tilde{\Psi}_e}\left(\tilde{\tilde{\psi}}_e\right) \mathcal{D}_{\ell}\left(\tilde{\tilde{\psi}}_i\right) \mathrm{d}\tilde{\tilde{\psi}}_e \tag{3.39}$$

To simplify the expression above, firstly, I realize that $f_{\tilde{\psi}}\big(\tilde{\psi}_e\big)$ and $f_{\tilde{\tilde{\Psi}}_e}\left(\tilde{\tilde{\psi}}_e\right)$ are the joint probability density function of statistically independent standard normal random variables. Therefore,

$$f_{\tilde{\Psi}_e}\big(\tilde{\psi}_e\big) = f_{\tilde{\Theta}}(\tilde{\theta}) \cdot f_{\tilde{\xi}_e}\big(\tilde{\beta}_e\big) = f_{\tilde{\Theta}}(\tilde{\theta}) \cdot f_{\tilde{H}_{\text{up}}}\big(\tilde{h}_{\text{up}}\big) \cdot f_{\tilde{H}_{\text{down}}}\big(\tilde{h}_{\text{down}}\big) \cdot \prod_{i=1}^{N_{\text{sg}}(e)} f_{\tilde{\zeta}_i}(\tilde{\varepsilon}_i) \tag{3.40a}$$

$$f_{\tilde{\tilde{\Psi}}_e}\left(\tilde{\tilde{\psi}}_e\right) = f_{\tilde{\tilde{\Theta}}}(\tilde{\tilde{\theta}}) \cdot f_{\tilde{\xi}_e}\big(\tilde{\beta}_e\big) = f_{\tilde{\tilde{\Theta}}}(\tilde{\tilde{\theta}}) \cdot f_{\tilde{H}_{\text{up}}}\big(\tilde{h}_{\text{up}}\big) \cdot f_{\tilde{H}_{\text{down}}}\big(\tilde{h}_{\text{down}}\big) \cdot \prod_{i=1}^{N_{\text{sg}}(e)} f_{\tilde{\xi}_i}(\tilde{\varepsilon}_i) \tag{3.40b}$$

Since all these random variables are standard normal, using the notation defined in Eq. (3.1) I can re-write Eq. (3.40) in a more desirable form as:

$$\begin{aligned} f_{\tilde{\psi}_e}\big(\tilde{\psi}_e\big) &= \phi(\tilde{\theta}) \cdot \phi\big(\tilde{h}_{\text{up}}\big) \cdot \phi\big(\tilde{h}_{\text{down}}\big). \prod_{i=1}^{N_{\text{sg}}(e)} \phi(\tilde{\varepsilon}_i) = \prod_{i=1}^{N_{\text{sg}}(e)+3} \phi\big(\tilde{b}_i\big) \\ &= \prod_{i=1}^{N_{\text{sg}}(e)+3} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\tilde{b}_i^2}\right) \end{aligned} \tag{3.41a}$$

$$f_{\tilde{\Phi}_e}\left(\tilde{\hat{\psi}}_e\right) = \phi(\tilde{\hat{\theta}}) \cdot \phi\left(\tilde{h}_{\text{up}}\right) \cdot \phi\left(\tilde{h}_{\text{down}}\right). \prod_{i=1}^{N_{\text{sg}}(e)} \phi(\tilde{\varepsilon}_i) = \prod_{i=1}^{N_{\text{sg}}(e)+3} \phi\left(\tilde{\hat{b}}_i\right)$$
$$= \prod_{i=1}^{N_{\text{sg}}(e)+3} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\tilde{b}_i^2}\right) \tag{3.41b}$$

In the equation above, $\tilde{b}_1 = \tilde{\theta}, \tilde{b}_2 = \tilde{h}_{\text{up}}, \tilde{b}_3 = \tilde{h}_{\text{down}}$ and $\tilde{b}_{j+3} = \tilde{\varepsilon}_j$, for $j \in \left(1,2,\cdots,N_{\text{sg}}(e)\right)$ with $\Omega_{\tilde{b}_i}$ representing the respective space (for example: $\Omega_{\tilde{b}_1} = \Omega_{\tilde{\Theta}}$). Similarly, $\tilde{\hat{b}}_1 = \tilde{\hat{\theta}}, \tilde{\hat{b}}_2 = \tilde{h}_{\text{up}}$ $\tilde{\hat{b}}_3 = \tilde{h}_{\text{down}}$ and $\tilde{\hat{b}}_{j+3} = \tilde{\varepsilon}_j$, for $j \in \left(1,2,\cdots,N_{\text{sg}}(e)\right)$, with $\Omega_{\tilde{\hat{b}}_i}$ representing the respective space (for example: $\Omega_{\tilde{b}_1} = \Omega_{\tilde{\theta}}$). Secondly, I note that for any function of the form $g(x \in X, y = 0 \in Y)$, $E_{XY}(g(x,0)) = E_X(g(x,0))$, provided X and Y are statistically independent random variables. This allows me to simplify the integral in Eq. (3.38) and (3.39) as:

$$\int_{\Omega_{\tilde{\psi}_e}} f_{\tilde{\Psi}_e}(\tilde{\psi}_e)\tilde{\mathcal{L}}(\tilde{\psi}_i)\mathrm{d}\tilde{\psi}_e = \frac{1}{\sqrt{2\pi}}\int_{\Omega_{\tilde{b}_i}} \tilde{\mathcal{L}}(\tilde{\psi}_i) e^{-\frac{1}{2}\tilde{b}_i^2}\, \mathrm{d}\tilde{b}_i \tag{3.42a}$$

$$\int_{\Omega_{\tilde{\Phi}_e}} f_{\tilde{\Phi}_e}\left(\tilde{\hat{\psi}}_e\right)\tilde{\mathcal{D}}_\ell\left(\tilde{\hat{\psi}}_i\right)\mathrm{d}\tilde{\hat{\psi}}_e = \frac{1}{\sqrt{2\pi}}\int_{\Omega_{\tilde{b}_i}} \tilde{\mathcal{D}}_\ell\left(\tilde{\hat{\psi}}_i\right) e^{-\frac{1}{2}\tilde{\hat{b}}_i^2}\, \mathrm{d}\tilde{\hat{b}}_i \tag{3.42b}$$

I realize that the Gauss-Hermite quadrature is a natural choice for approximating the integrals in the equation above. This is because Gauss-Hermite quadrature is meant to estimate integrals of the form $\int_x g(x)e^{-x^2}\,\mathrm{d}x$, for any function $g(x)$. Therefore, the approximations are

$$\int_{\Omega_{\tilde{\psi}_e}} f_{\tilde{\Psi}_e}(\tilde{\psi}_e)\tilde{\mathcal{L}}(\tilde{\psi}_i)\mathrm{d}\tilde{\psi}_e \approx \frac{1}{\sqrt{\pi}}\sum_{n=1}^{r} w_n \tilde{\mathcal{L}}(\tilde{q}_{i,n}), \text{ where } \tilde{q}_{i,n}(j) = \begin{cases} \tilde{\psi}_i(j) = 0 & i \neq j \\ \alpha_n & i = j \end{cases} \tag{3.43a}$$

$$\int_{\Omega_{\tilde{\Phi}_e}} f_{\tilde{\Phi}_e}\left(\tilde{\hat{\psi}}_e\right)\tilde{\mathcal{D}}_\ell\left(\tilde{\hat{\psi}}_i\right)\mathrm{d}\tilde{\hat{\psi}}_e \approx \frac{1}{\sqrt{\pi}}\sum_{n=1}^{r} w_n \mathcal{D}_\ell\left(\tilde{\hat{q}}_{i,n}\right), \text{ where } \tilde{\hat{q}}_{i,n}(j)$$

$$= \begin{cases} \tilde{\hat{\psi}}_i(j) = 0 & i \neq j \\ \alpha_n & i = j \end{cases} \tag{3.43b}$$

In the equations above, $\tilde{q}_{i,n}(j)$ (or $\tilde{\hat{\psi}}_i(j)$) represents the j^{th} component of the vector $\tilde{q}_{i,n}$ (or $\tilde{\hat{\psi}}_i$); r represents quadrature order; w_n gives the weights; and α_n gives the point of evaluation of the function for $n \leq r$. For my calculations, I use $r = 3$, for which $w_1 = \frac{2}{3}\sqrt{\pi}, w_2 = \frac{1}{6}\sqrt{\pi}$, $w_3 = -\frac{1}{6}\sqrt{\pi}, \alpha_1 = 0, \alpha_2 = \frac{\sqrt{6}}{2}$, and $\alpha_3 = -\frac{\sqrt{6}}{2}$. This choice of the quadrature order satisfies the computational accuracy that this problem demands and at the same time leads to a computationally efficient numerical estimation of Bayes risk. The approximated Bayes risk functions can now be written as

$$\mathfrak{E}_{\text{explicit-risk}}(e) \approx -\left(N_{\text{sg}}(e)+2\right)\tilde{\mathcal{L}}\left(\tilde{\psi}_0\right) + \sum_{i=1}^{\left(N_{\text{sg}}(e)+3\right)}\sum_{n=1}^{r=3} w_n \tilde{\mathcal{L}}\left(\tilde{q}_{i,n}\right) \tag{3.44a}$$

$$\mathfrak{E}_{\text{implicit-risk}}(e) \approx -\left(N_{\text{sg}}(e)+2\right)\tilde{\mathcal{D}}_\ell\left(\tilde{\hat{\psi}}_0\right) + \sum_{i=1}^{\left(N_{\text{sg}}(e)+3\right)}\sum_{n=1}^{r=3} w_n \tilde{\mathcal{D}}_\ell\left(\tilde{\hat{q}}_{i,n}\right) \tag{3.44b}$$

Figure 3.7 illustrates the algorithmic flowchart to obtain the explicit and implicit form of Bayes risk functional defined in Eq. (3.44).

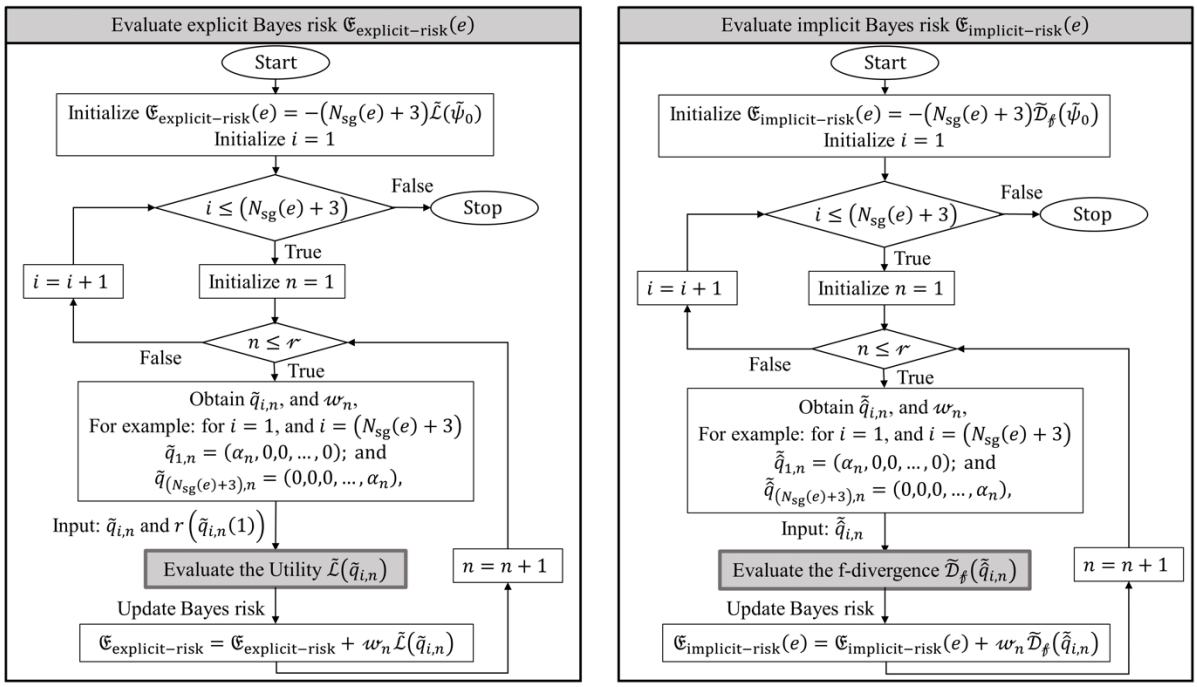


Figure 3.7: Algorithm to evaluate both the explicit and implicit Bayes risk.

As was noted in Eq. (3.27) that, mathematically, optimization using explicit and implicit Bayes risk functional should yield the same result. However, since I numerically estimate $\mathfrak{E}_{\text{explicit-risk}}(e)$ and $\mathfrak{E}_{\text{implicit-risk}}(e)$ using Gauss-Hermite quadrature in conjuncture with univariate dimensional reduction, optimization using these functional leads to a different sensor design. This is because evaluating $\mathfrak{E}_{\text{explicit-risk}}(e)$ requires using observed strain measurements corresponding to the gap length values $\theta = F_{\Theta}^{-1}\left(F_{\tilde{\Theta}}(\tilde{\theta} = \alpha_n)\right)$, whereas evaluating $\mathfrak{E}_{\text{implicit-risk}}(e)$ requires using a different set of observed strain measured corresponding to the true gap length values $\tilde{\theta} = F_{\hat{\Theta}}^{-1}\left(F_{\tilde{\tilde{\Theta}}}\left(\tilde{\tilde{\theta}} = \alpha_n\right)\right)$. In other words, since the cumulative distribution functions $F_{\tilde{\Theta}}(\tilde{\theta})$ (obtained from the prior distribution of gap length) and $F_{\tilde{\tilde{\Theta}}}\left(\tilde{\tilde{\theta}}\right)$ (obtained from the modified prior distribution of gap length) are different, the Gauss point α_n maps to a different values of true gap lengths θ and $\tilde{\theta}$, and hence requires using different simulated strain data. I use implicit Bayes risk

as the objective functional because it shows the effect of risk-weights more prominently than obtained using the explicit Bayes risk.

It is natural to address the question "How many times do I need to run the GPR model to *obtain the Bayes risk* $\mathfrak{E}_{\text{explicit-risk}}(e)$ or $\mathfrak{E}_{\text{implicit-risk}}(e)$ *defined by* Eq. (3.44) for a single design consideration *e*?" I show the calculation for the explicit Bayes risk and note that the number of GPR runs are the same for explicit or implicit Bayes risk functional evaluation. To start with, I realize that there are two sums in the expression of $\mathfrak{E}_{\text{explicit-risk}}(e)$ in Eq. (3.44). Therefore, the cost function $\tilde{\mathcal{L}}(\tilde{q}_{i,n})$ has to be obtained $r \cdot (N_{\text{sg}}(e) + 3)$ times for all the possible $\tilde{q}_{i,n}$, where $i \leq (N_{\text{sg}}(e) + 3)$ and $n \leq \mu$. For a fixed value of the index i and n, the vector $\tilde{q}_{i,n}$ consist of some realization of the standard-normal quantities $(\tilde{\theta}, \tilde{\beta}_o)$ obtained by Eq. (3.43). From Remark 3.3, obtaining the utility function $\tilde{\mathcal{L}}(\tilde{q}_{i,n})$ for a fixed $\tilde{q}_{i,n}$ requires running the GPR model $(N_{\text{particles}} + 1)$ times. It can also be seen in the Fig. 3.7 that the utility $\tilde{\mathcal{L}}(\tilde{q}_{i,n})$ is evaluated for every $\tilde{q}_{i,n}$ by calling a function "Expected the Utility $\tilde{\mathcal{L}}(\tilde{q}_{i,n})$ " that requires running GPR model ($N_{\text{particles}} + 1$) times at every instance the function is called (refer to Fig. 3.5). Therefore, evaluating Bayes risk using Eq. (3.44) for a given design e requires running the GPR model $N_{\text{GPR 1}}$ times, such that:

$$N_{\text{GPR1}} = r \times (N_{\text{particles}} + 1) \times (N_{\text{sg}}(e) + 3) \tag{3.45}$$

These GPR model runs make evaluating Bayes risk computationally expensive.

Bayesian optimization aimed at obtaining the optimal sensor network design consists of evaluating many such designs, denoted by e_i, consisting of $N_{\text{sg}}(e_i)$ number of sensors. To obtain an optimized sensor placement design, I start with an initially assumed design, denoted by e_0 that

consists of $N_{\mathrm{sg}}(e_0)$ number of sensors. Starting with e_0, the subsequent sensor design e_i with $N_{\mathrm{sg}}(e_i)$ sensors is obtained by picking the most optimal sensor location from the available sensors and adding that sensor location to the previous design e_{i-1} with $N_{\mathrm{sg}}(e_{i-1}) = N_{\mathrm{sg}}(e_i) - 1$ sensors. Picking the additional sensor required to update the design e_{i-1} to the design e_i requires $N_{\text{iter}}(e_i)$ number of iterations. Since Bayes risk is the optimiality criteria, it needs to be evaluated at every iteration for the design e_i. Let e_I, with $i = I$, represent the final optimal sensor network design. The total number of GPR runs to arrive at e_I (starting from e_0) is denoted by N_{GPR2}, such that:

$$N_{\mathrm{GPR2}} = \sum_{i=1}^{I} \mu \times \left(N_{\text{particles}} + 1\right) \times \left(N_{\mathrm{sg}}(e_i) + 3\right) \times N_{\text{iter}}(e_i) \tag{3.46}$$

So far, I have taken two major steps to reduce the computational cost. First, I have used a digital surrogate (GPR) of the finite element model. Secondly, I have used SVD to reduce the dimension of the GPR model's output. In the next section, I propose a novel and innovative approach to further minimize the computational cost for evaluating the Bayes risk by minimizing the number of times I run the GPR model to evaluate the Bayes risk.

3.5.3 An efficient computational approach to evaluate the Bayes risk

In this section, I highlight the disadvantage of using the algorithm illustrated in Fig. 3.7 to obtain the explicit and implicit Bayes risk and propose an alternative novel approach that significantly reduces the computational cost of optimization. For the sake of discussion, I consider explicit Bayes risk functional. An approach to evaluate Bayes risk $\mathfrak{E}_{\text{explicit-risk}}(e)$ for a given design e as illustrated in Fig. 3.7 involves many repeated evaluation of GPR model $g_e\left(\theta, h_{\text{up}}, h_{\text{down}}\right)$ for same input arguments $\left(\theta, h_{\text{up}}, h_{\text{down}}\right)$. The computational cost can be significantly reduced by realizing that in the entire process of evaluating the Bayes risk, there are

only $7\left(N_{\text{particles}}+1\right)$ unique runs of GPR model. This follows from the following line of reasoning. The vector $\tilde{q}_{i,n}$ is a special case of the vector $\tilde{\psi}_i$ as defined in Eq. (3.35). The first three components of the vector $\tilde{q}_{i,n}$ constitute a sub-vector $\tilde{q}_{i,n}(1:3)=\left(\tilde{\theta},\tilde{h}_{\text{up}},\tilde{h}_{\text{down}}\right)$, the inverse standard-normal transformation of which are the argument of the GPR model $g_e\left(\theta,h_{\text{up}},h_{\text{down}}\right)=\tilde{g}_e\left(\tilde{\theta},\tilde{h}_{\text{up}},\tilde{h}_{\text{down}}\right)$. The remaining $N_{\text{sg}}(e)$ components constitute a vector $\tilde{q}_{i,n}\left(4:N_{\text{sg}}(e)+3\right)$ representing external noise. For any vector $\tilde{q}_{i,n}$, the sub-vector $\tilde{q}_{i,n}(1:3)=\left(\tilde{\theta},\tilde{h}_{\text{up}},\tilde{h}_{\text{down}}\right)$ bears α_n as numerical value of one of the components and zero for others Therefore, I have a set of 7 unique sub-vectors $\left(\tilde{\theta},\tilde{h}_{\text{up}},\tilde{h}_{\text{down}}\right)$ of interest to us. From remark 3.1, Fig 3.5 and Fig 3.6, obtaining the utility $\tilde{\mathcal{L}}\left(\tilde{q}_{i,n}\right)$ for each $\tilde{q}_{i,n}$ requires $\left(N_{\text{particles}}+1\right)$ GPR runs. Therefore, considering all the 7 unique arguments of GPR model, I essentially need to run GPR model only N_{GPR3} times, such that:

$$N_{\text{GPR3}}=7\times\left(N_{\text{particles}}+1\right) \tag{3.47}$$

For each of these stand-alone GPR runs, I store the strain values in all the 64919 sensors constituting data of size $7\times\left(N_{\text{particles}}+1\right)\times 64919$ in a matrix called "strain-data" and pick the strain measurements of the sensors constituting a design e. Therefore, even while carrying out Bayesian optimization that may consider many designs, the number of GPR runs remains $7\times\left(N_{\text{particles}}+1\right)$ cutting computational cost intensively. Figure 3.8 illustrates this process of storing the strain data. Once the matrix straindata is obtained independently, the utility $\tilde{\mathcal{L}}\left(\tilde{q}_{i,n}\right)$ can be evaluated by extracting the relevant sensor readings from the matrix strain-data as demonstrated in Fig. 3.9. Replacing θ with $\hat{\theta}$, and $\tilde{q}_{i,n}$ with $\tilde{\hat{q}}_{i,n}$ in the flowchart 8 gives the strain-data matrix

required to obtain the intrinsic Bayes risk and replacing them in in the flowchart 9 yields the f-divergence $\tilde{D}_{\ell}\left(\tilde{\hat{\theta}}, \beta_e\right)$ required to obtain the intrinsic Bayes risk.

The modified algorithm to evaluate the Bayes risk $\mathfrak{E}_{\text{explicit-risk}}(e)$ is illustrated in Fig. 3.10. The most important difference between the algorithm in Fig. 3.7 and the one in Fig. 3.10 is that in the modified algorithm the GPR models are not run at every iteration step.

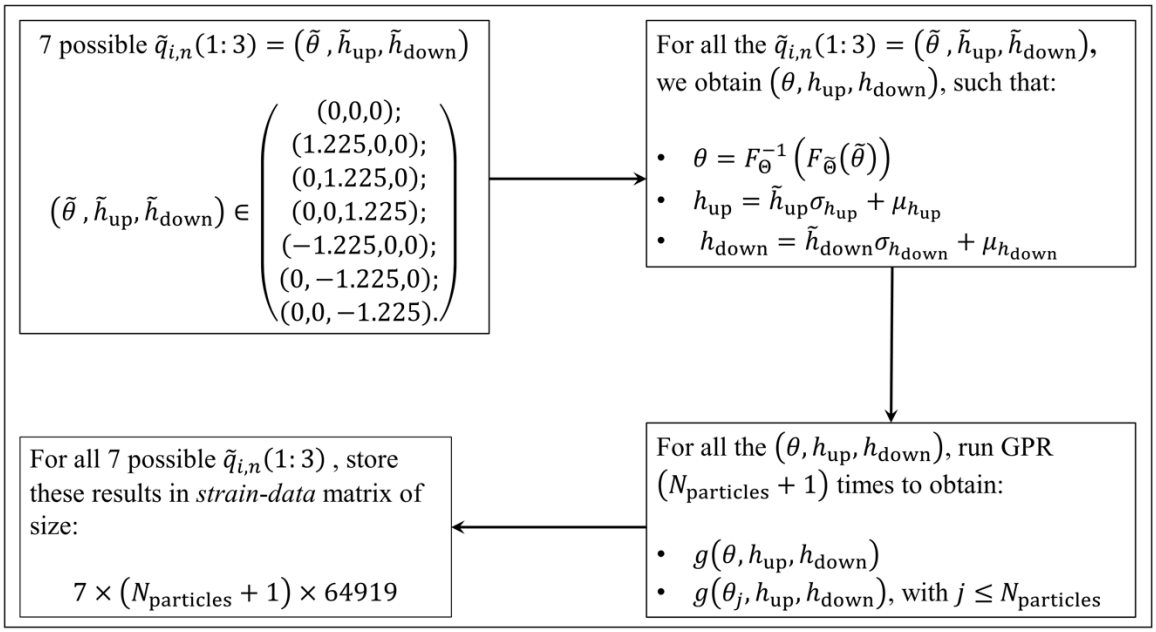


Figure 3.8: Obtaining and storing the strain-data required to obtain Bayes risk $\mathfrak{E}_{\text{extrinsic-risk}}(e)$

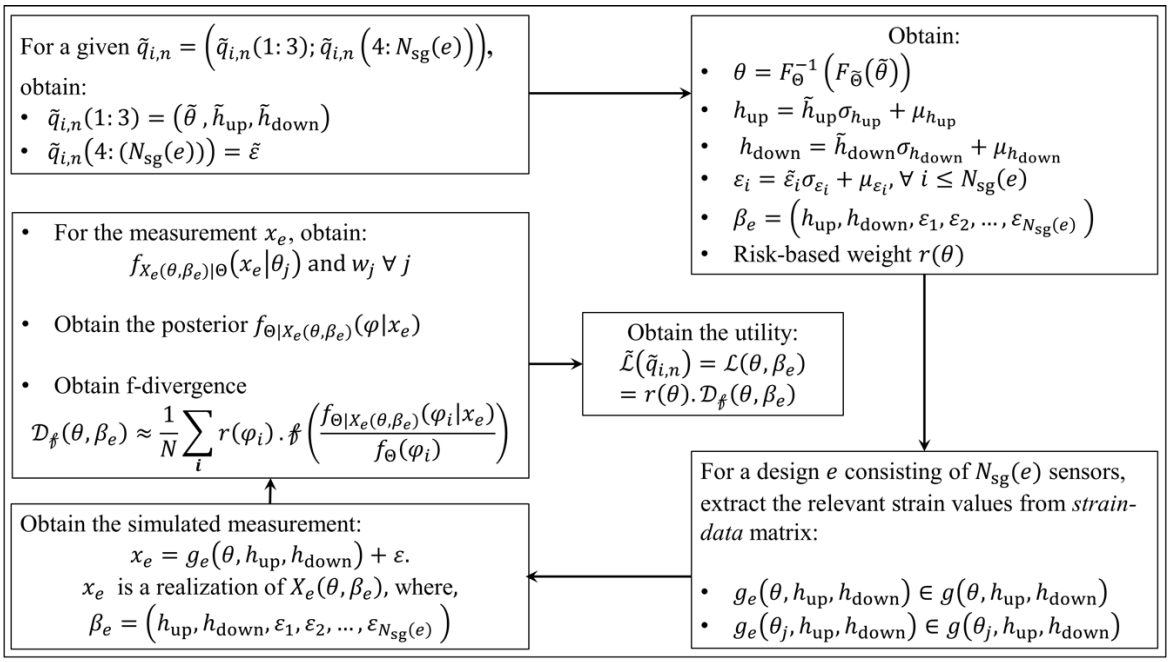


Figure 3.9: Obtaining the utility $\tilde{\mathcal{L}}(\tilde{q}_{i,n})$

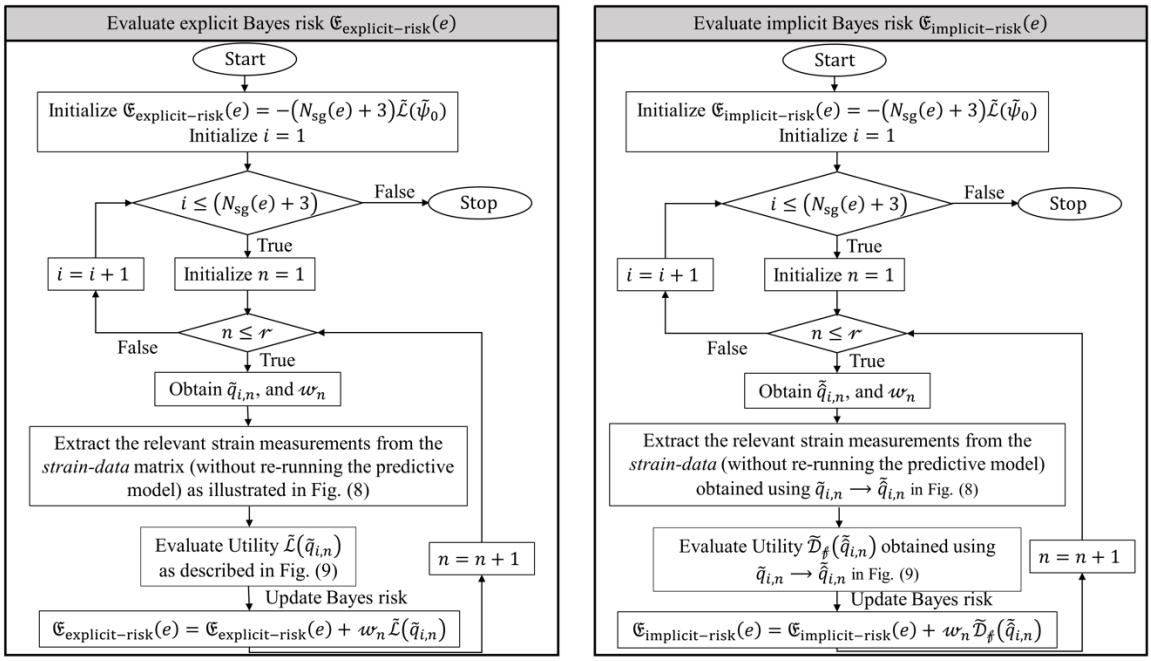


Figure 3.10: Computationally efficient algorithm to evaluate the explicit and implicit Bayes risk.

The Fig. 3.11 compares the number of GPR runs required for Bayesian optimization by considering two approaches used to evaluate Bayes risk as defined in Fig. 3.7 and 3.10. Carrying

out Bayesian optimization using first approach to evaluate Bayes risk (either extrinsic or intrinsic) Fig. (3.7) needs N_{GPR2} runs of GPR model that depends on the design e_i considered as defined in Eq (3.46). I start by assuming $N_{\mathrm{sg}}(e_0) = 3$ number of initial sensors and update the design using upto 10 additional sensors. For simplicity, I assume a constant average number of iterations for each design $N_{\mathrm{iter}}\,(e_i) = 20$. I assume $N_{\mathrm{particles}} = 5000$. I observe that number of GPR runs using modified algorithm illustrated in Fig. 3.10 is $N_{\mathrm{GPR2}} = 35007$, which is order of magnitudes smaller as compared to N_{GPR2}. For instance, $N_{\mathrm{GPR2}} = 2100420$ for the first additional sensor, and $N_{GPR2} = 34506900$ for ten additional sensors.

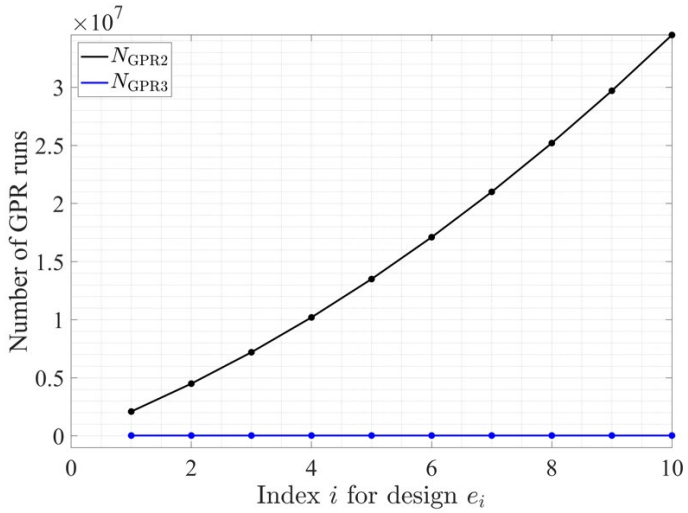


Figure 3.11: Number of GPR runs for Bayesian optimization

3.6 Bayesian optimization: optimal sensor placement design

3.6.1 Optimal sensor placement design algorithm

Given that my objective is to find the design e^* that maximizes a Bayes risk functional, I will use the implicit Bayes risk form to solve

$$e^* = \arg\max_{e \in \Omega_E} \mathfrak{E}_{\text{implicit-risk}}(e) \tag{3.48}$$

In general, obtaining e^* involves looking at every possible design combination and choosing the one with the maximum implicit Bayes risk. In my case, where the total number of sensor locations is $n = 64919 \times 4$, this would be choosing e^* from the $\sum_{r=1}^{n} \frac{n!}{r!(n-r)!} = (2^n - 1)$ possible combinations of sensor locations. Therefore, the exhaustive design space Ω_E consists of $(2^{259676} - 1) \approx 10^{78170}$ possible designs, which is approximately 78090 orders of magnitude more than the number of estimated atoms in the universe. Given the intractable nature of exhaustive search, Bayesian optimization is used to look for a global optimum in a minimum number of steps, thus minimizing the sampling points to rapidly speed up the optimization process. Unlike gradient-based optimization methods, Bayesian optimization does not require the derivative of the objective function; having a black-box model (like a surrogate function) of the objective function suffices to perform the optimization. It involves two primary elements. The first element is developing surrogate function using another GPR of the objective function using randomly evaluated samples. The second component is the acquisition function that helps me locate the next most valuable candidate to update the design (Jones and Schonlau1998, Frazier 2018).

The process begins by choosing an initial design $e_0 = \left(d^{(1)}, d^{(2)}, \cdots, d^{\left(N_{\text{sg}}(e_0)\right)}\right) \in \Omega_E$ consisting of $N_{\text{sg}}(e_0) = N_0 \geq 0$ sensors. Here, $d^{(l)}$ represents the location of l^{th} strain gauge in the design e_0. The next step is to obtain an updated design e_1 by adding an additional sensor to e_0, such that $N_{\text{sg}}(e_1) = N_0 + 1$. To obtain the optimal e_1, I randomly sample α sensor locations using Latin Hypercube Sampling (LHS), subjected to a space filling property, to be the candidate for the additional sensor from the unused sensors constituting the measurement space Ω_X. These locations

yield α number of design samples $\tilde{e}_k, \forall k \leq \alpha$ each with $N_{\text{sg}}(e_1)$ sensors. I obtain the exact cost $\mathfrak{E}_{\text{implicit-risk}}(\tilde{e}_k), \forall k \leq \alpha$ using approach discussed in the previous section. Using the set of α additional sensor locations as input data, denoted by $\tilde{d}$, and the exact cost as output data, I train my surrogate function $\hat{\mathfrak{E}}(\tilde{d}) \sim N(\mu_{\tilde{d}}, \sigma_{\tilde{d}}^2)$. This surrogate can be used to quickly estimate a posterior probability that describes possible values for the Bayes risk at a remaining candidate location $\bar{d}$ spanning the entire design space, with mean value $\mu_{d\bar{d}}$ and standard deviation $\sigma_{d\bar{d}}$. I use Expected Improvement EI as my acquisition function that helps me locate the next most valuable candidate for the next sensor location based on the current posterior over the Bayes risk,

$$EI(\bar{d}) = (\mu_{\bar{d}} - \mathfrak{E}^*)\Phi\left(\frac{\mu_{\bar{d}} - \mathfrak{E}^*}{\sigma_{\bar{d}}}\right) + \sigma_{\bar{d}}\phi\left(\frac{\mu_{\bar{d}} - \mathfrak{E}^*}{\sigma_{\bar{d}}}\right) \tag{3.49}$$

Here, $\mathfrak{E}^* = \max_{\tilde{e}_k} \mathfrak{E}_{\text{implicitrisk}}(\tilde{e}_k)$ is the current best value of the objective function. For all the remaining possible additional sensor location candidates, I evaluate $EI(\bar{d})$. The candidate with maximum EI is the next most valuable location. Once I locate the next most valuable sensor location candidate, I get $(\alpha + 1)^{\text{th}}$ design samples. I retrain the GPR with $(\alpha + 1)$ data points and keep adding the next most valuable location from the set of strain locations constituting Ω_X until the maximum EI is less than a tolerance value ε.

Note that the aforementioned details update an initial design e_0 to e_1 by adding one additional sensor. I keep updating the designs by adding one sensor at a time until one of the following two conditions is reached:

1 Bayes risk converges to a constant value, i.e., the design $e_I = e^*$(with $i = I$) can be considered as the most optimal design if $\mathfrak{E}(e_I) \approx \mathfrak{E}(e_{(I-1)})$. Given an updated design $e_i =$

$\left(d^{(1)}, d^{(2)}, \cdots, d^{\left(N_{\mathrm{sg}}(e_i)\right)}\right)$ with $N_{\mathrm{sg}}(e_i)$ number of sensors, the aforementioned steps can be generalized to obtain the updated design $e_{(i+1)}$

2 The total number of sensors in the design reaches the maximum number of sensors limited/constrained by the decision-maker or other factors.

Given the design e_i, the updated design $e_{(i+1)}$ can be obtained following similar exercise as described above. Let $N_{\text{total}} = 64919 \times 4$ represent the total number of strain-gauges attached in the structure (at top and bottom of each 64919 elements in both vertical and horizontal direction). Let $e_{n\ \text{wepresent}}$ the optimized sensor design with $(N_0 + n_{\mathrm{as}})$ sensors, such that $n_{\mathrm{as}} \leq N_{\mathrm{as}}$. Here, N_{as} represents the maximum additional sensors considered over the initially assumed number of sensors N_0. The number of sensors int final design shall then be $\leq (N_0 + N_{as}$).The optimal design e^* is then given by:

$$e^* = \arg\max_{e_{n_{as}}} \mathfrak{E}_{\text{implicit-risk}}\left(e_{n_{as}}\right) \tag{3.50}$$

The following algorithm 1 demonstrates the Bayesian optimization procedure to evaluate the design e^*

Algorithm 1: Bayesian optimization for sensor placement

Initialize $e_0 = \left(\mathcal{d}^{(1)}, \mathcal{d}^{(2)}, \cdots, \mathcal{d}^{(N_0)}\right)$;

for $n_{as} = 1$ **to** N_{as} **do**

  Using LHS, randomly select α locations to be candidates for the $(N_0 + n_{\text{as}})$ sensor location, with coordinates $\mathcal{X} = \left(\tilde{\mathcal{d}}^{(1)}, \tilde{\mathcal{d}}^{(2)}, \cdots, \tilde{\mathcal{d}}^{(\alpha)}\right)$;

  Obtain α number of possible designs: $\tilde{e}_k = \text{concatenate}\left(e_{(n_{\text{as}}+1)}, \tilde{\mathcal{d}}^{(k)}\right)$, for all $k \leq \alpha$;

  Obtain the exact cost of all the α designs:
  $\Xi = \left(\mathfrak{E}_{\text{implicit-risk}}(\tilde{e}_1), \mathfrak{E}_{\text{implicit-risk}}(\tilde{e}_2), \cdots, \mathfrak{E}_{\text{implicit-risk}}(\tilde{e}_\alpha)\right)$;

  while $i = 1$ **or** $maxEI < \varepsilon$ **do**

    Construct the GPR model for $\hat{\mathfrak{E}}(\cdot)$ trained using $(\mathcal{X}, \Xi)$;

    For all the remaining strain locations $Z = \left(\bar{\mathcal{d}}^{(1)}, \bar{\mathcal{d}}^{(2)}, \cdots, \bar{\mathcal{d}}^{(\beta)}\right)$, where $\beta = \left(N_{\text{total}} - \left(N_0 + n_{\text{as}} - 1\right) - \alpha\right)$, obtain β number of possible designs: $\bar{e}_m = \text{concatenate}\left(e_{(n_{\text{as}}+1)}, \bar{\mathcal{d}}^{(m)}\right)$, for all $m \leq \beta$;

    Obtain the cost $\hat{\mathfrak{E}}\left(\bar{\mathcal{d}}^{(m)}\right)$ for all $m \leq \beta$ designs using GPR developed before;

    Obtain the currrent best $\mathfrak{E}^* = \min \Xi$;

    Obtain the Expected Improvement for all the β designs using:

$$EI(\bar{\mathcal{d}}^{(m)}) = \left(\mu_{\bar{\mathcal{d}}^{(m)}} - \mathfrak{E}^*\right) \Phi\left(\frac{\mu_{\bar{\mathcal{d}}^{(m)}} - \mathfrak{E}^*}{\sigma_{\bar{\mathcal{d}}^{(m)}}}\right) + \sigma_{\bar{\mathcal{d}}^{(m)}} \phi\left(\frac{\mu_{\bar{\mathcal{d}}^{(m)}} - \mathfrak{E}^*}{\sigma_{\bar{\mathcal{d}}^{(m)}}}\right)$$

    , where $m \leq \beta$;

    Obtain:

$$maxEI = \max_{\bar{x}^{(m)}} \left(EI(\bar{\mathcal{d}}^{(m)})\right)$$

$$\bar{\mathcal{d}} = \arg\max_{\bar{x}^{(m)}} \left(EI(\bar{\mathcal{d}}^{(m)})\right)$$

$$\bar{e} = \text{concatenate}\left(e_{(n_{\text{as}}+1)}, \bar{\mathcal{d}}\right)$$

    Evaluate the exact cost $\mathfrak{E}_{\text{implicit-risk}}(\bar{e})$;

    Update:

$$\mathcal{X} = \text{concatenate}\left(\mathcal{X}, \bar{\mathcal{d}}\right)$$

$$\tilde{e}_{(\alpha+i)} = \bar{e}$$

$$\Xi = \text{concatenate}\left(\Xi, \mathfrak{E}_{\text{implicit-risk}}(\bar{e})\right)$$

    $i = i + 1$;

  end

  Update the sensor design: $e_{n_{\text{as}}} = \text{concatenate}(e_{n_{\text{as}}-1}, \bar{\mathcal{d}})$;

end

Obtain: $e^* = \arg\max_{e_k} \mathfrak{E}_{\text{implicit-risk}}(e_k)$, where, $k \leq N_{\text{as}}$;

3.6.2 Results and discussion

The initial design e_0 consisting of $N_0 \geq 0$ sensors may be randomly obtained via LHS, predefined based on judgment/experience, or taken as 0 if no pre-defined design is available and if one does not want to assume a random initial design. I note in my case study that the miter gate finite element model is so finely meshed that there exists a spatial correlation between the strain values. Therefore, there are non-unique sensor locations that are sampled by the acquisition function, leading to non-unique and slightly different sensor designs depending on the different initial design e_0. In this chapter, I numerically implement the optimization algorithm by fixing the initial design e_0 with $N_0 = 0$. Secondly, I consider $\theta_{\text{low}} = 70$ inches, and $\theta_{\text{up}} = 180$ inches. I reasonably assume the gap values below 70 inches do not represent significant damage to the gate, and the gap value of 180 inches represents the upper limit of the possible gap value beyond which the gate is considered critically damaged, based on discussions with USACE engineers. I perform the following studies:

1 In Section 3.6.2.1, I investigate and compare the capability of predicting the posterior distribution of the gap length using a random design consisting of 10 sensors (obtained using LHS) vs. optimal sensor design obtained using Bayes risk $\mathfrak{E}(e)$ that ignores the risk weight and considers KL divergence as the choice of f-divergence in the Bayes risk functional.

2 In Section 3.6.2.2, I investigate and compare the capability of a sensor design in predicting the posterior distribution of the gap length obtained using Bayesian optimization of Bayes risk functional constructed using various f-divergences, with and without the risk

weighting I consider 5 types of f-divergences in constructing the Bayes risk and compare their effectiveness in arriving at the optimal design.

3.6.2.1 Comparison of an optimal sensor placement design based on KL divergence (no risk weight) vs. randomly chosen design

Figure 3.13 illustrates the randomly chosen sensor design (left image), and the one arrived at by using KL divergence without risk weights in the Bayes risk functional (right image). I observe that all these sensors constituting the design obtained using KL divergence-based functional are concentrated close to the boundary of the quoin block and the gate. This location is desirable to capture the change in the strain values due to loss of contact between the quoin block and the gate (or for obtaining a better inference of the gap length), reflecting an unquestionable advantage of using Bayesian optimized sensor design over the randomly chosen design. It is also seen in Fig. 3.12, from the strain field plot obtained for a fixed set of loading parameters, that there is a stress concentration near the gap length that intuitively justifies the fact that optimal sensor design should contain at least a few sensors near the gap. Secondly, from Fig. 3.14, I observe that sensor placement optimization significantly increases the effectiveness of Bayesian inference. This leads to a significant reduction in the uncertainty associated with the posterior distribution of the gap length for different realizations of the true gap values. Thirdly, as seen in Fig. 3.14b, the uncertainty in the posterior distribution is not equally/uniformly reduced for different realizations of true gap length. This is because the obtained optimal sensor design can reduce more uncertainty for certain true gap lengths and less for the other ones. Despite this local non-uniformity, the sensor design obtained using Bayesian optimization is optimal in a global sense.

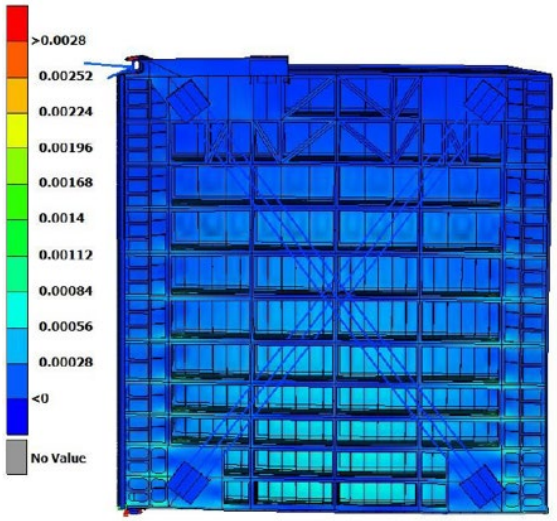


Figure 3.12: Stress field plot of the miter gate structure obtained for a fixed set of load parameters

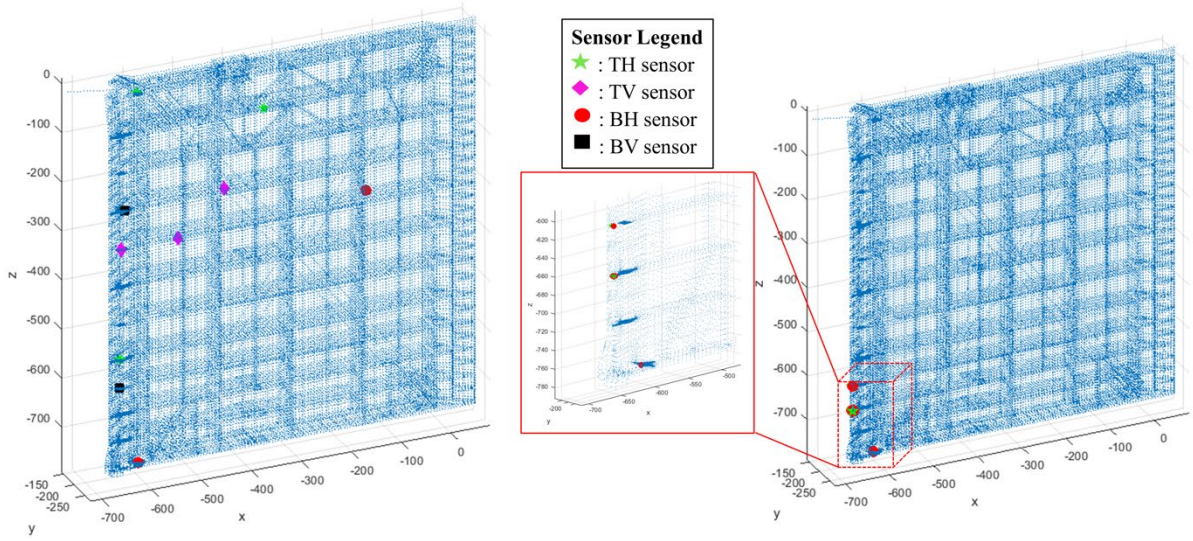
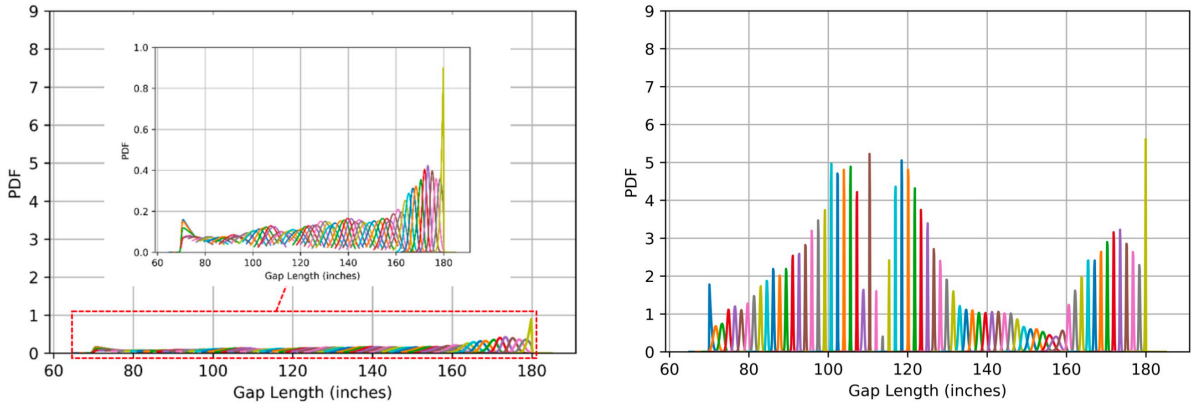


Figure 3.13: Randomly chosen 10 sensor design (left) vs. KL divergence optimized (no risk weight) 10 sensor design (right)



(a) A random design (b) Optimized design using KL divergence

Figure 3.14: Posterior distributions obtained using a randomly chosen 10 sensor design vs. KL divergence (no risk weight) optimized 10 sensor design

3.6.2.2 Comparison of an optimal sensor placement design obtained using f-divergence with risk weights vs. without risk weights

In this section, I compare the effectiveness of sensor design in Bayesian inference obtained using different kinds of f-divergences with and without the risk weight. I consider the 5 different f-divergences: KL, Hellinger, total variation, Pearson, and Jensen divergence. The following subsections illustrate a set of three plots for each of the f-divergence considered:

1 Figures 3.15, 3.18, 3.21, 3.24 and 3.27, illustrates the sensor design arrived by ignoring the risk weights (left figure), and by including the risk weights (right figure) in Bayes risk functional.

2 Figures 3.16, 2.19, 3.22, 3.25 and 3.28, illustrates the posterior distribution of the gap length arrived for different realizations of the true gap length values by ignoring the risk weight (left figure) and by including the risk weights (right figure).

3 Figures 3.17, 3.20, 3.23, 3.26, and 3.29, illustrates the ratio of the maximum value of the posterior distribution of the gap length with and without the risk weights (left figure), and the ratio of the standard deviation of the posterior distribution of the gap length with and without the risk weights (right figure) for different realizations of the true gap length values.

First, I recall that one of the criteria for the Bayes risk functional was to incorporate my desire to obtain better information/description of the gap length when the true value of gap length is larger or when the state of the structure approaches a higher degree of damage. As is seen in Figures 3.16,3.19, 3.22, 3.25 and 3.28, the optimization using risk-weighted (implicit) Bayes risk functional allows me to have higher confidence in the inference results for larger value of true gap length (or for higher degree of damage). However, accomplishing better inference at a higher value of the true gap length leads to sacrifice in the performance of the Bayesian optimization at a lower value of the true gap length. This fact is reflected in Figures 3.17a, 3.20a, 3.23a, 3.26a and 3.29a, such that the ratio of the maximum value of the posterior with or without risk weight is higher at larger value of true gap length. Similarly, as seen in Figures 3.17b, 3.20b, 3.23b, 3.26b and 3.29b, the ratio of the standard deviation of the posterior with or without risk weight is lower at larger value of true gap length. Second, most of the sensors identified are in the horizontal direction (TH and BH), and only a few are in the vertical direction (TV and BV). Third, the optimization results are dependent on the choice of f-divergence.

As observed in Fig. 3.15 for the risk-weighted KL divergence case (right figure) that one sensor (highlighted by a red circle) is far away from the gap (unlike the other sensors that are close to the gap). This is counter-intuitive and deserves an explanation. The algorithm searches the global domain for the next possible candidate. There are two possibilities of such selection:

1 The outlier (for example: BV sensor circled in red in Fig. 3.15) could have been selected in the initial iteration steps. It would have been an incorrect choice that the algorithm would self-correct by picking appropriate sensors in the next iterations.

2 The outlier (for example: BV sensor circled in red in Fig. 3.15) was selected in the later iteration step (closer to the converging point). In that case, the information provided by the sensors selected in the preceding iteration steps was enough to capture the required information and the BV sensor did not add much value to the design. This was the case in selecting the BV sensor in Fig. 3.15 highlighted by red circle.

Since the algorithm searches the entire space (global search), it may have been duped by the resembling and related strain information at different coordinates. However, the self-correcting nature of the algorithm would eventually select a combination of sensors (in the final design) that would capture the necessary information.

For the case of optimization where risk-weights are ignored, it is observed from Figures 3.16a, 3.19a, 3.22a, 3.25a, and 3.28a, that at a true gap length about 110 inches, the posterior distribution of the gap-length has higher variability than the distributions at other gap values slightly higher or lower to 110 inches. It can be seen with a closer look that there are some sensors near the 110 inch gap value in the final design (see the green circle highlighted portion of the left part of Fig. 3.15). These sensors may be relatively more sensitive to the gap value lower or higher than 110 inches partly because of their location and the component of the gate they are attached to (a gate is a complex structure consisting of many elements welded together). Although these sensors may not be as sensitive to 110 inches gap degradation, they certainly are sensitive to other gap values Since the optimization framework presented in this chapter chooses the optimal design

in an average sense, the relative advantage (optimal sensitivity to the overall damaged state) of these sensors is one possible reason for them to be picked by the optimization algorithm in the first place.

(a) Comparison of an optimal sensor placement design obtained using KL divergence with vs. without risk weight

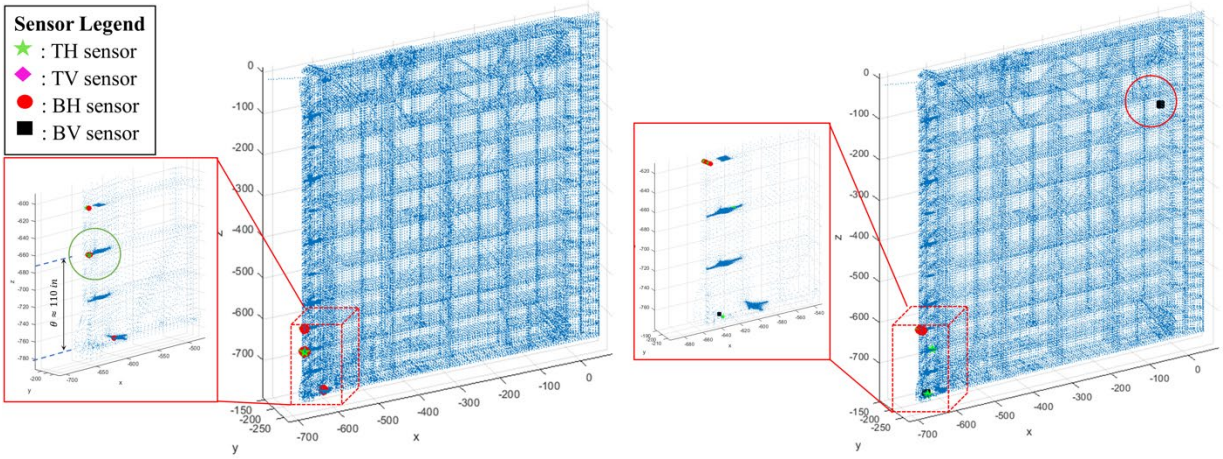
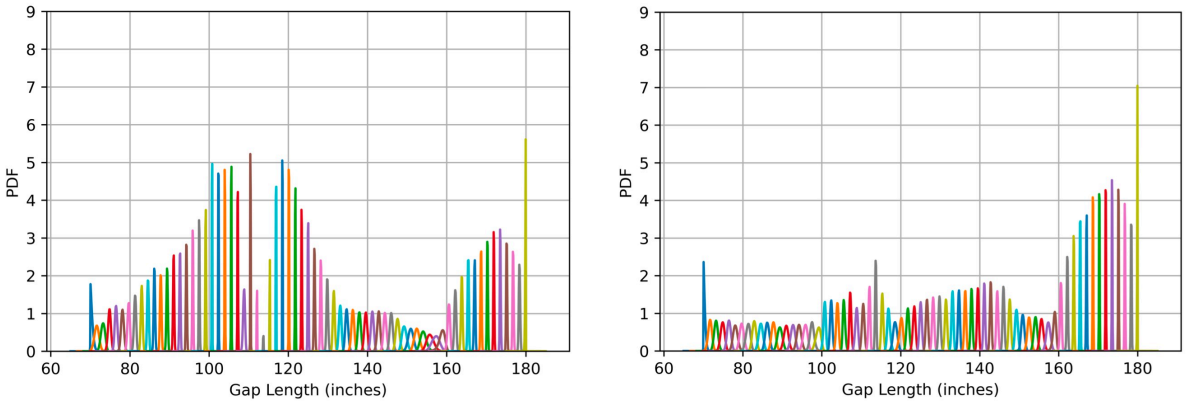
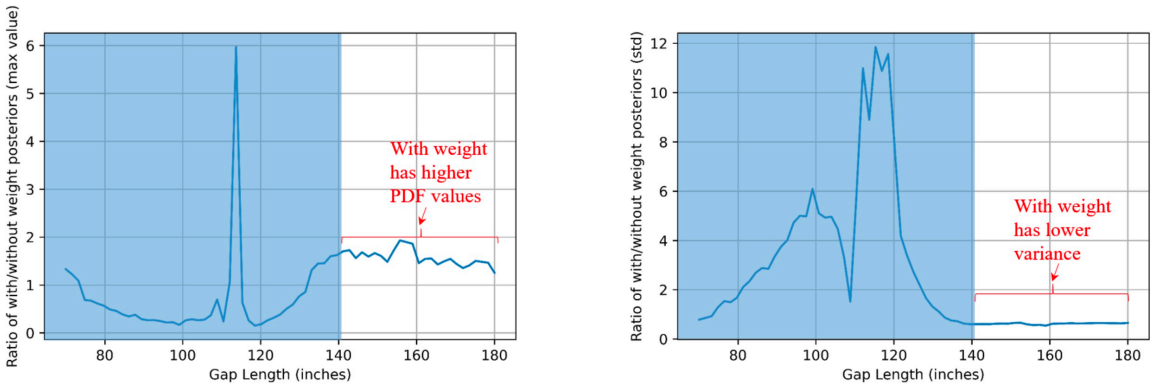


Figure 3.15: Sensor placement design using KL divergence: without (left) and with (right) the risk weight



(a) Without risk weight (b) With risk weight

Figure 3.16: Posterior distributions obtained using KL divergence



(a) Ratio of the maximum value of the posterior (b) Ratio of the standard deviation of the posterior

Figure 3.17: Ratio of the maximum value and the standard deviation of the posterior obtained using KL divergence with and without risk weight

(b) Comparison of an optimal sensor placement design obtained using Hellinger divergence with vs. without risk weight

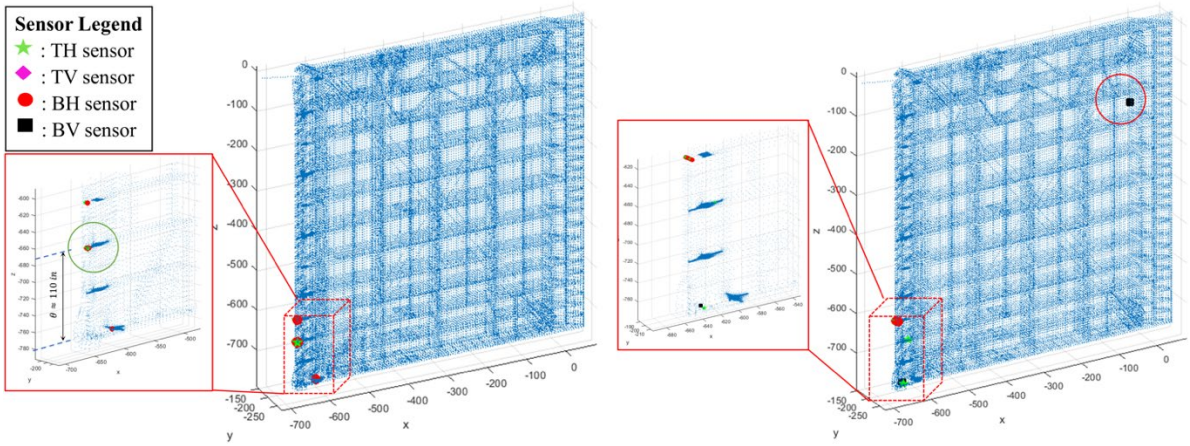


Figure 3.18: Sensor placement design using Hellinger divergence: without (left) and with (right) the risk weight

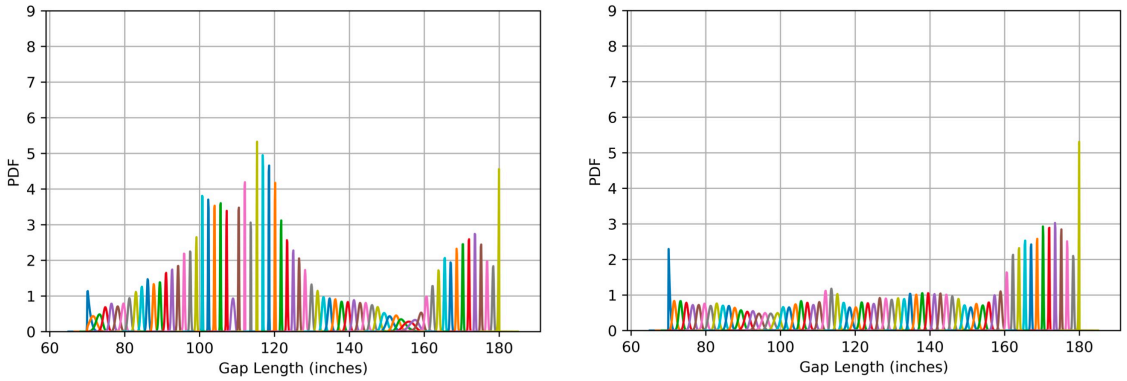


Figure 3.19: Posterior distributions obtained using Hellinger divergence

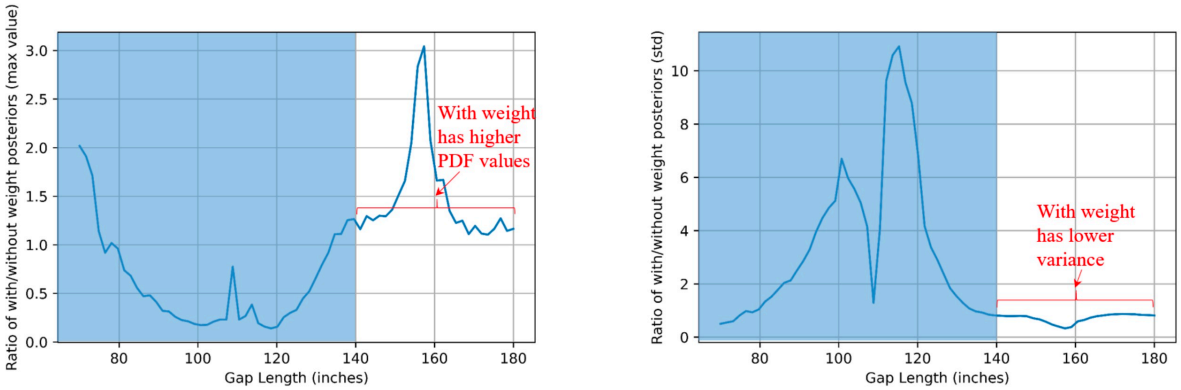


Figure 3.20: Ratio of the maximum value and the standard deviation of the posterior obtained using Hellinger divergence with and without risk weight

(c) Comparison of an optimal sensor placement design obtained using total variation f-divergence with vs. without risk weight

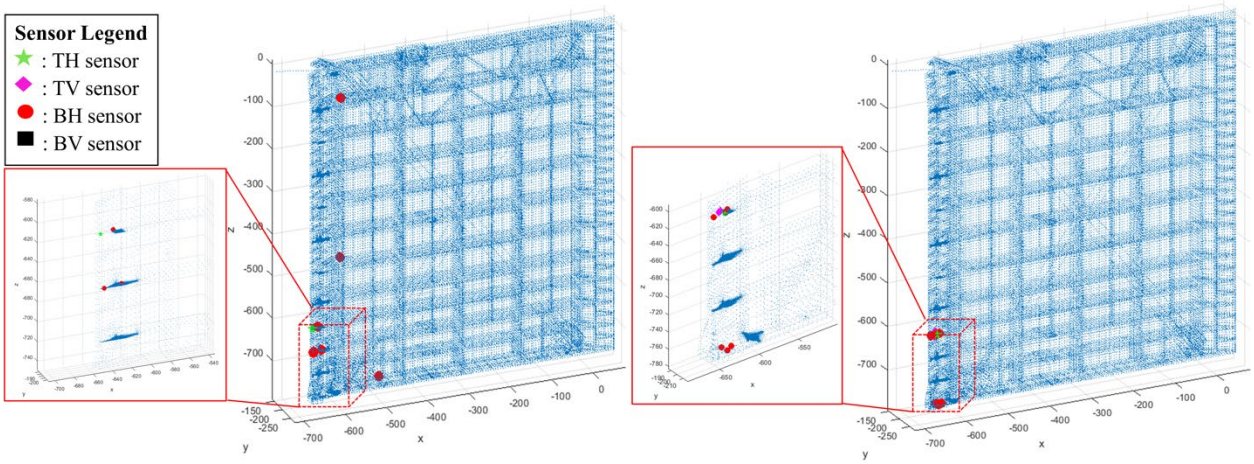


Figure 3.21: Sensor placement design using total variation f-divergence: without (left) and with (right) the risk weight

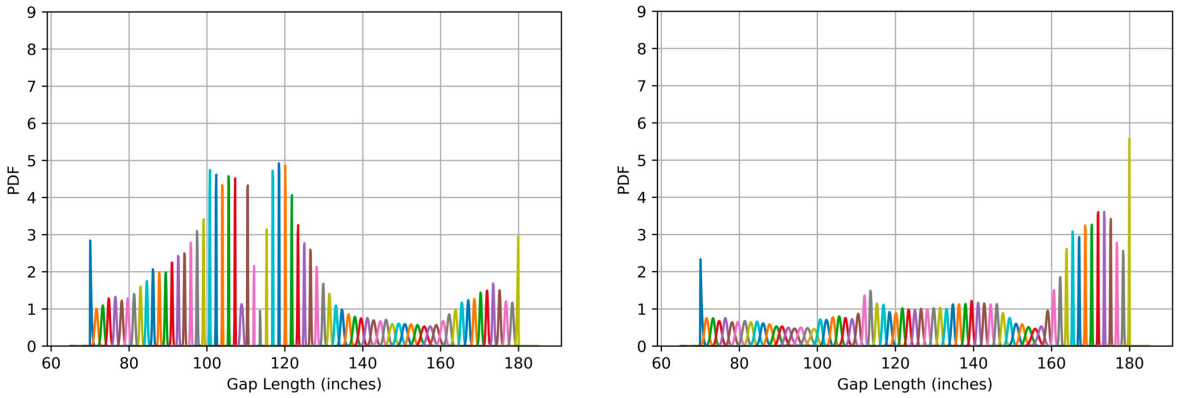
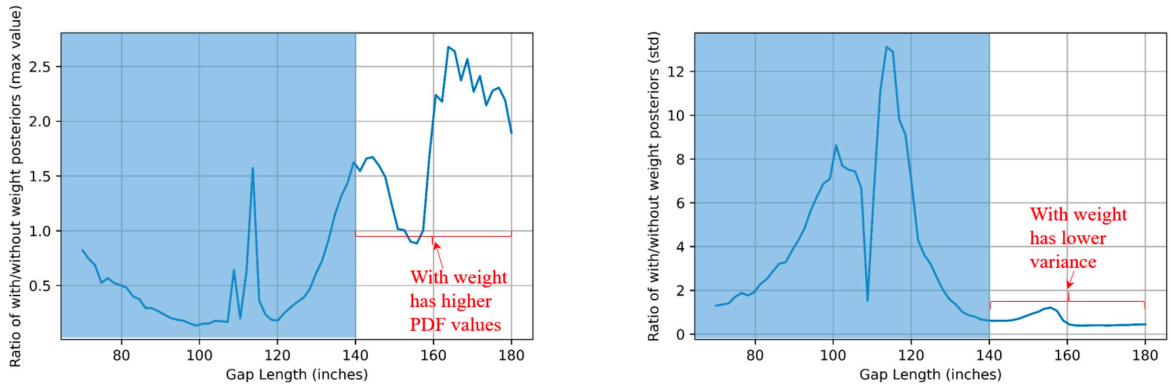


Figure 3.22: Posterior distributions obtained using total variation f-divergence



(a) Ratio of the maximum value of the posterior (b) Ratio of the standard deviation of the posterior

Figure 3.23: Ratio of the maximum value and the standard deviation of the posterior obtained using total variation f-divergence with weight and without weight

(d) Comparison of an optimal sensor placement design obtained using Pearson f-divergence with vs. without risk weight

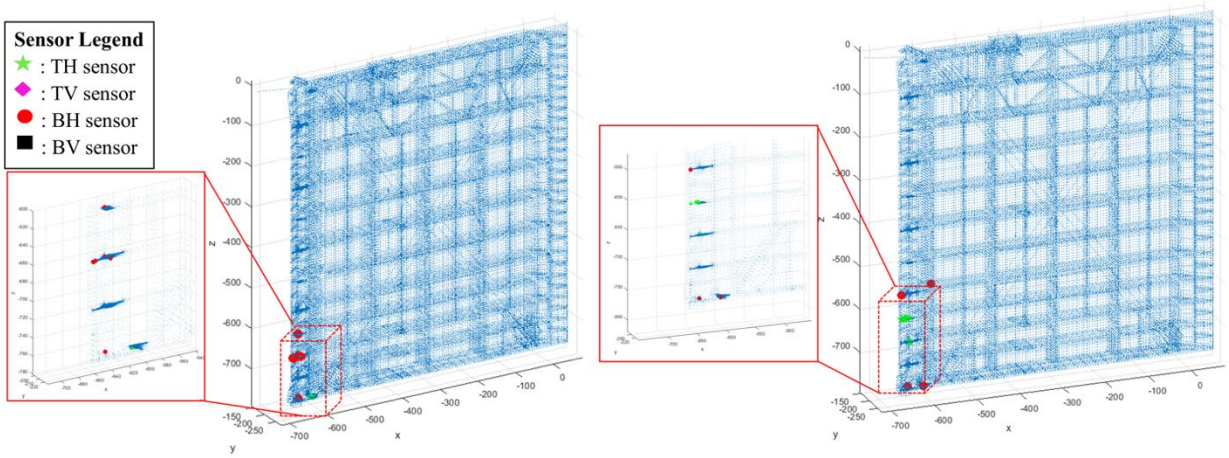
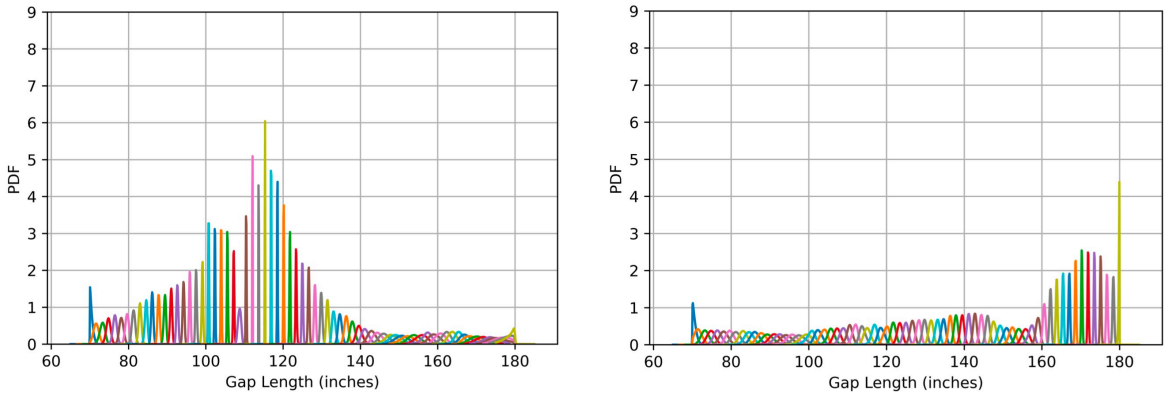
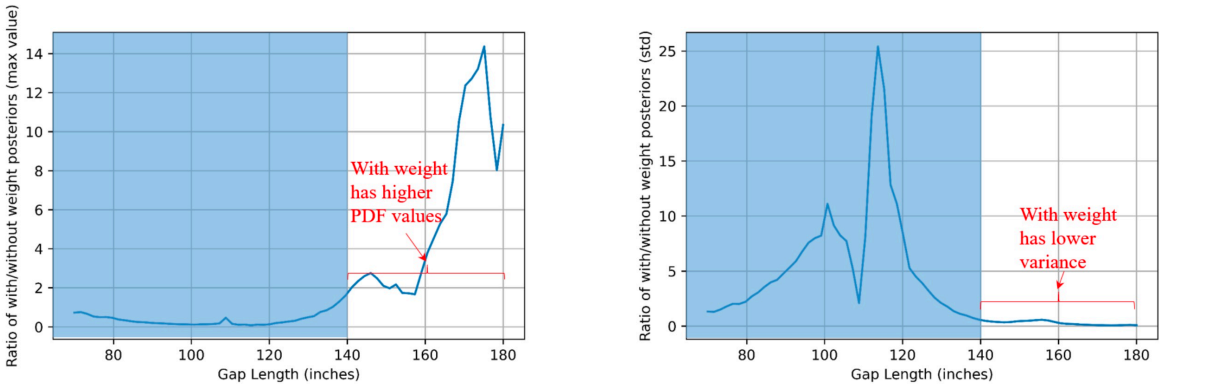


Figure 3.24: Sensor placement design using Pearson f-divergence: without (left) and with (right) the risk weight



(a) Without risk weight (b) With risk weight

Figure 3.25: Posterior distributions obtained using Pearson f-divergence



(a) Ratio of the maximum value of the posterior (b) Ratio of the standard deviation of the posterior

Figure 3.26: Ratio of the maximum value and the standard deviation of the posterior obtained using Pearson f-divergence with and without risk weight

(e) Comparison of an optimal sensor placement design obtained using Jensen f-divergence with vs. without risk weight

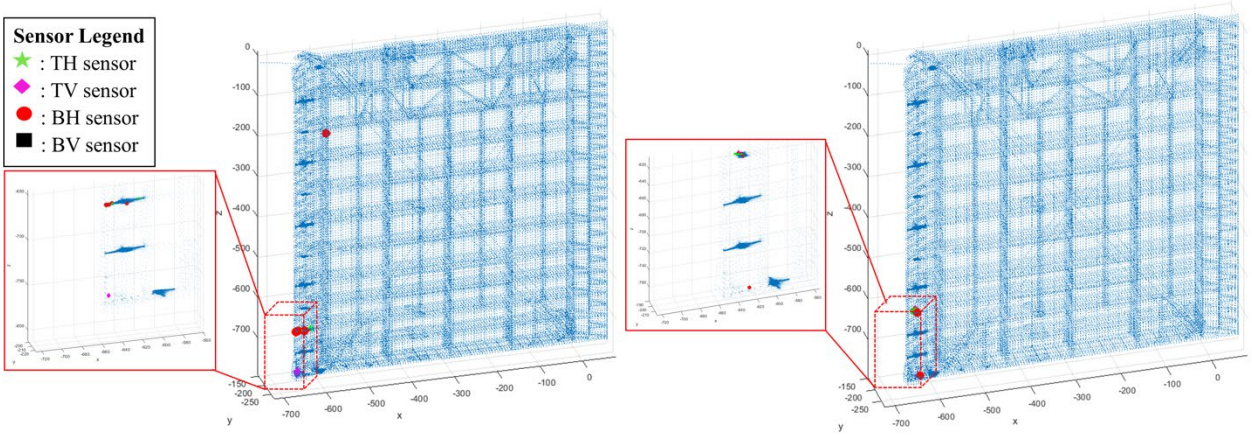
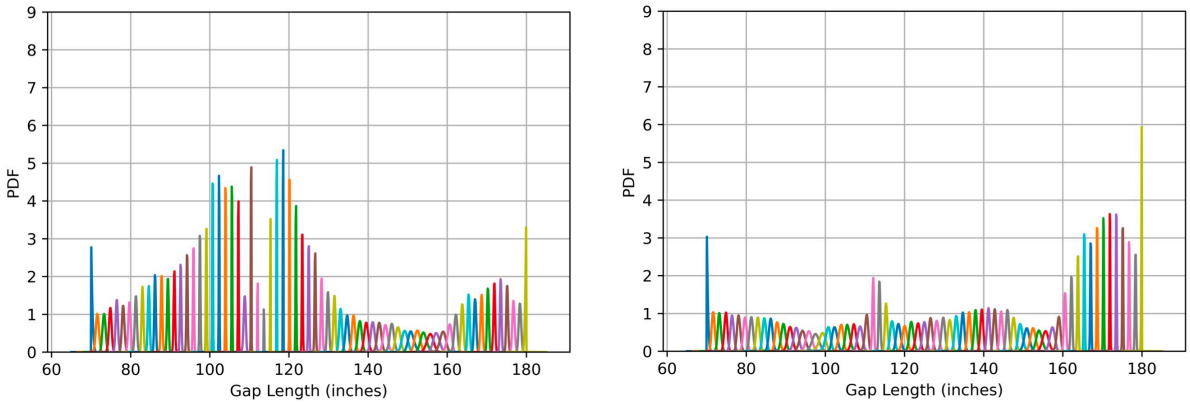
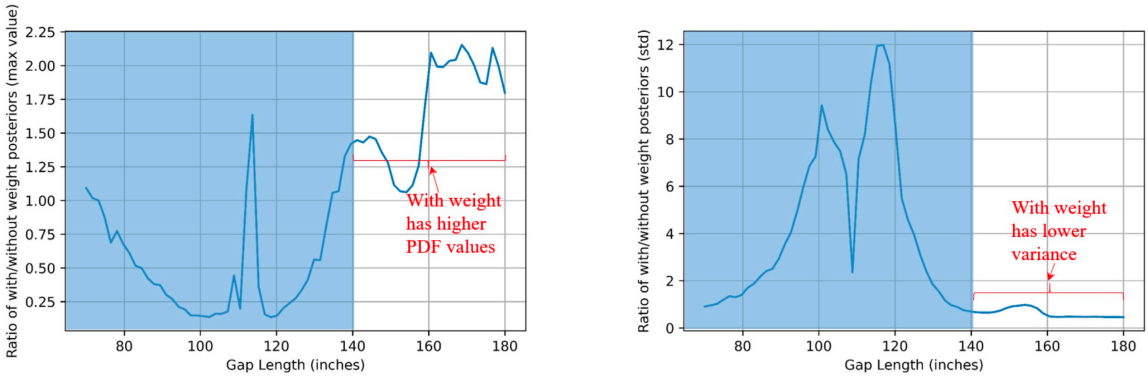


Figure 3.27: Sensor placement design using Jensen f-divergence: without (left) and with (right) the risk weight



(a) Without risk weight (b) With risk weight

Figure 3.28: Posterior distributions obtained using Jensen f-divergence



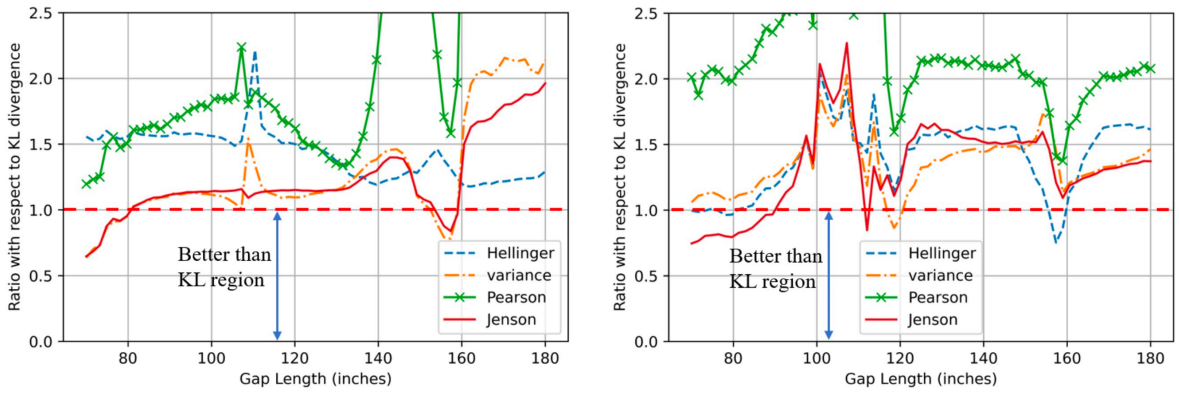
(a) Ratio of the maximum value of the posterior (b) Ratio of the standard deviation of the posterior

Figure 3.29: Ratio of the maximum value and the standard deviation of the posterior obtained using Jensen f-divergence with and without weight

(f) Comparison of an optimal sensor placement design obtained using different f-divergences relative to KL divergence with and without risk weight

Figure 3.30a illustrates the ratio of the standard deviation of the posterior distribution of the gap length obtained by various f-divergences without risk weight with respect to the standard deviation obtained using KL divergence without risk weight. Similarly, Fig. 3.30b represents the same ratio when risk weights are considered in the Bayes risk functional.

I observe that among the f-divergences studied for the application of optimal sensor design; Bayes risk functional considering KL divergence leads to the best sensor placement design for both with or without risk weights.



(a) Without risk weight

(b) With risk weight

Figure 3.30: Ratio of the standard deviation of the posterior obtained using various f-divergence relative to the standard-deviation obtained using KL divergence

3.7 Summary and Conclusions

This chapter details an optimal sensor design framework for structural health monitoring applications. It was demonstrated on a miter gate case study with the primary goal of arriving at the optimal strain gauge network design used to infer the posterior distribution of the loss-in-

contact gap length (the damage state parameter). I arrived at such a design by maximizing an objective functional referred to as Bayes risk. The Bayes risk is designed to accommodate three crucial elements: first, it aims to obtain a design that maximizes the gains in the information on the gap length inferred from the strain-gauge measurements. This gain in the information is quantified by f-divergence that evaluates the similarity or dissimilarity between two probability measures by evaluating the distance (relative gain of information) between two distributions. Second, the Bayes risk incorporates my desire to obtain better information/description of the gap length when the true value of gap length is larger, or the state of the structure approaches a higher degree of damage. It is crucial for deciding the maintenance strategies and appropriate action preventing significant losses. It is accomplished by using a risk-weight. Third, the Bayes risk also takes into account my prior knowledge of the gap length.

Arriving at the optimal sensing design was accomplished by minimizing risk or equivalently maximizing the utility (defined as the risk-weighted gain of information) in an average sense. In this regard, evaluation of the Bayes risk for a given sensor network design demands considering all the possible degrees of damage (indicated by the true gap length value). This requires obtaining a large set of simulated observation data and a quick Bayesian inference of the posterior distribution of gap-length for many different realizations of the true gap length values. Given a true gap length value and the loading parameters, the simulated strain gauge readings can be obtained using a high-fidelity finite element model (FEM). The randomness in the strain measurements is primarily due to uncertainties in the hydrostatic load parameters, and the noise in strain gauges. Since the high-fidelity finite element model is computationally expensive, I built a digital surrogate using Gaussian Process Regression (GPR), which is around 50000 times faster than FEM. I deploy numerical approximation of the posterior distribution by using particle

filters, or specifically the sequential Monte Carlo (SMC) approach. I define two approaches, intrinsic and extrinsic, to incorporate risk weights into Bayes risk functional. Although both, the intrinsic and extrinsic, definitions of Bayes risk are mathematically equivalent, the numerical evaluation of the intrinsic and extrinsic Bayes risk functional yields slightly different results. I use implicit Bayes risk because it shows the effect of risk weights more prominently than the result obtained using the explicit form of Bayes risk. Numerically evaluating Bayes risk expression involves evaluating a non-linear, multi-dimensional integral. I use univariate dimension reduction in conjuncture with the Gauss-Hermite quadrature. Apart from reducing the computational cost by using GPR, I also proposed a novel and innovative approach to further minimize the computational cost for evaluating the Bayes risk by minimizing the number of times I run the GPR model to evaluate the Bayes risk

I observe that as compared to random sensor design, the optimal sensor design significantly increases the effectiveness of Bayesian inference and reduces the uncertainty in the posterior distribution of the gap-length value. Inclusion of the risk-weight in Bayes risk allows me to have higher confidence in the inference results for a higher degree of damage (as was intended). However, accomplishing better inference at a higher value of the true gap length leads to sacrifice in the performance of the Bayesian optimization at a lower value of the true gap length. Amongst the chosen f-divergences, I conclude that KL divergence is the most suitable choice for this particular class of problems. A future possible work can be to compare the efficiency and results obtained by the presented Bayesian optimization with other algorithms in the literature for the given problem The computational speed, as well as the final design, arrived using various optimization algorithms may depend on how the algorithm is engineered, and the constraints on the problem at hand, and it remains to be investigated.

3.8 Preview to Chapter 4

Up through Chapter 3, I have successfully built up a complete efficient framework for optimal sensor design for a large civil infrastructure system, i.e., the miter gate. The challenges addressed in Section 2.8 have been solved thoroughly in this chapter. A new Bayes risk function is addressed with a combination of maximizing the information gain during the gap length inference process and creating a risk-weight related to the structural damage level. Various information gain functions are evaluated for comparing their performances in selecting the *best* sensor arrangement. Besides, a numerical strategy is introduced to reduce the computational expanse when calculating the expected values for the objective function in optimization process.

Although the updated framework is capable of designing an optimal sensor placement, the current framework has not yet considered the time evolution of the structural performance or the reliability of the monitoring system itself (the sensors' reliability). As discussed in Section 1.4, real systems degrade under loading over time and possible degradation in the sensor system itself can induce error into the data-to-decision SHM workflow. Sensors are more likely to be damaged or biased when they are merged under the water level (a harsh environment). However, sometimes sensors are installed in the water because they could be closer to the damage area, which could provide more information, implying a tradeoff. Thus, on top of the current framework, sensor reliability is an important factor to be considered in a sensor placement design. The next chapter will account for the sensor reliability in the optimal sensor design framework and extend to a structural life cycle analysis.

3.9 Remarks

Portions of this dissertation have been published or are currently being considered for publication. Chapter 3 is composed of a first authored publication:

Y. Yang, M. Chadha, Z. Hu, M. A. Vega, M. D. Parno, and M. D. Todd, "A probabilistic optimal sensor design approach for structural health monitoring using risk-weighted f-divergence," Mechanical Systems and Signal Processing, vol. 161, p. 107920, 2021.

Chapter 4

AN OPTIMAL SENSOR DESIGN FRAMEWORK ACCOUNTING FOR SENSOR RELIABILITY OVER THE STRUCTURAL LIFE CYCLE

4.1 Abstract

One of the most prominent goals of a structural health monitoring system (SHM) is to infer the state of the structure to inform appropriate maintenance actions that affect performance and safety over the lifecycle of the structure. An SHM system infers the structural state by acquiring measurement data that is expected to contain information about the structural state. This inference may be fundamentally flawed, however, if the sensing system that initiates the SHM workflow is unreliable. The operational and environmental load state that these sensors can face, in addition to normal manufacturing defects, result in varying functionality at different monitoring locations, at different times. Therefore, it becomes imperative to account for sensor reliability in the optimal sensor design process for the SHM system at the outset. This chapter details an optimal sensor design framework with a target of probabilistically inferring the structural state accounting for the reliability of sensors over the lifecycle of the structure. The generic framework is laid out and is finally demonstrated on a complex real-world miter-gate structure.

4.2 Introduction

At a global level, the primary goal of an SHM system is to infer the structural state by gathering in-situ sensor measurement data and converting it to actionable information. The benefits of a well-designed structural health monitoring (SHM) system are potentially meaningful,

including (but not limited to) reduction in ownership cost through data-driven optimal maintenance decision-making, performance maximization during operation, and mitigation of catastrophic failure to provide life safety margins. Since an SHM system monitors the structure by exploiting in-situ measurement data, data acquisition is one of the most important components of the SHM paradigm. Therefore, the optimal design of the sensor network–defined herein as the spatial arrangement of the sensor network in order to limit the scope appropriately–may significantly enhance the performance and life cycle value of the SHM system as a whole.

In recent work (Yang and Chadha 2021), I investigated and proposed an optimal sensor network design framework using Bayesian optimization with the target of maximizing the gain in the information of the structural state when measurement data is available. The information gain was quantified by risk-weighted f-divergence of the posterior distribution of the state parameter (conditioned on newly acquired measurement data) relative to the prior knowledge of the structural state when new measurements are unavailable. It was assumed that the sensors were fully functional, and their measurements were unbiased and reliable. However, this is seldom the case when an SHM system is used to monitor the structure that is subjected to harsh and variable environmental conditions over its lifecycle. In the real world, parts of a structure may be exposed to different environmental conditions and loads. As a result, sensors have varying reliability and malfunctioning chances at different monitoring locations and at different times. Since the overall value of an SHM system depends on its design (Chadha and Hu 2021; Thöns and Faber 2013; Malings and Pozzi 2016), it is crucial to investigate and account for the reliability of sensors along with all the uncertainties (for example, varying load conditions and noise in the sensors) in an optimal sensor network design framework. This chapter focuses on extending my optimal sensor

design framework to a more realistic scenario that accounts for the reliability of sensors over the lifecycle of the structure.

As mentioned in Yang et al. (Yang and Chadha 2021), there is not a universal objective for sensor design, as each application has a distinct goal for the use of a particular SHM system. Given that decision-makers are the typical curators of SHM utility, the objective function may also be defined from the perspective of decision theory that defines *loss as a consequence of decision-making* (or the associated risk) by considering various prior information and uncertainty sources in the decision-making process. The loss/risk is a subjective quantity and is defined according to the problem. Optimal sensor design therein requires finding the sensor network that minimizes the losses or risk expressed by an objective function in an average sense; such an objective function is defined as *Bayes risk*. Readers are referred to research papers cited in (Yang and Chadha 2021) that use different objective functions suitable for various applications. When the sensors malfunction and the measurement data is not reliable (not representative of the true structural state), it runs the *risk* of incorrectly inferring the state parameter, i.e., there is a shift in the properties (such as the mean value) of the posterior distribution of the state parameter (conditioned upon newly acquired sensor data) farther away from the true state parameter. Since the primary target of the optimal sensor design process delineated in this chapter is to account for the reliability of sensors over the structure's life cycle, the proposed objective function in this chapter is, therefore, the expected value of the aggregate absolute deviation of the mean value of the inferred posterior distribution of the state parameter from the true state parameter over the lifecycle of the structure, *averaged* over all the uncertainties (an example of *Bayes risk*).

In such complex sensor optimization types of problems, the objective function is usually a high-dimensional non-linear integral. Minimizing such a function is challenging and requires advanced numerical methods and computational techniques. Optimal sensor design has been a topic of significant interest in the SHM community, especially when it comes to implementing SHM for damage detection, classification, or inference in large and complex infrastructure (Flynn and Todd 2010; Meo and Zumpano 2004; Yi and Li 2012). Since each detection, inference, or classification problem is distinct and focuses on a unique target, different objective functions and various solution strategies have been proposed in a long list of seminal literature available in this research area (Friswell and Castro-Triguero 2015; Entezami and Sarmadi 2020; Li and Zhang 2016; Yang and Liang 2019). Yang et al. (Yang and Chadha 2021; Yang and Chadha 2021) list numerous research work that describes various optimization approaches and lay the foundation of Bayesian optimization numerical algorithm that will be extended and used in this chapter. Bayesian optimization is the most suitable technique to apply complex large-scale civil structures SHM applications because the global sensor design space is huge, and therefore evaluating Bayes risk for each of these designs is computationally impossible (thus requires an iterative optimization approach) and I do not have the derivatives of the objective function that prevents me from using gradient decent types of approaches (Akbarzadeh and Lévesque 2014).

The previous work (Yang and Chadha 2021; Yang and Chadha 2021) leading up to this chapter established a Bayesian optimization framework that targeted different optimization objectives. Yang et al. (Yang and Chadha 2021) aimed at obtaining a sensor design that maximizes gain in information, and (Yang and Chadha 2021) aimed at detecting discrete boundary conditions. Although the objectives were different, both of my previous work considered that the sensors were fully reliable and worked just as warranted. Mathematically, this behavior was modeled by

assuming a zero-mean Gaussian noise structure for sensor measurements. Although such construct allowed me to converge my focus on the optimization framework itself, it completely ignored the possibility of sensors malfunctioning. Inclusion of sensor reliability as a sensor design optimization criterion has been investigated for various applications, like fault and failure detection (Li and Upadhyaya 2011; Duan and Lin 2018;), communication reliability in wireless networks (Damaso and Rosa 2014; Zonouz and Xing 2014), reliability in corrosion rate sensor data (Marsh and Frangopol 2008), and on system reliability (Li and Sadoughi 2020; Kang and Song 2008). Coit et al. (Coit and Zio 2019) categorizes and discusses evolution of system reliability optimization in three era and also delineates potential challenges encountered in system reliability optimization. Needless to say, given that a structure is exposed to wide varieties of environmental conditions over its lifecycle, sensor reliability is crucial to consider while analyzing lifecycle performance of an SHM system. This chapter does exactly that.

The chapter will first describe the sensor placement optimization framework accounting for sensor reliability, followed by application to a demonstration case study. The case study is a navigation lock miter gate that is managed and operated by USACE (Richardson 1964; Vega and Hu 2020; Vega and Hu 2021). This structure serves as an excellent demonstration example for two reasons. The first is that the damage parameter is the loss of boundary contact (a "gap") between the miter gate and its supporting wall quoin block at the bottom of the gate. It is reasonably assumed that the damage state is characterized by a one-dimensional continuous state parameter that is probabilistically inferred from the observable sensor. Although miter-gate is a significantly complex real-world structure, it is suitable to visualize, interpret, and evaluate the optimization results since it allows us to consider a single-dimensional state parameter. The second reason is that in terms of the reliability aspect, at any point in time, the miter gate may be conveniently

divided into two zones. The first zone is the part of the gate above the water level (called splash zone) and the second zone is the part of the gate that is submerged in the water (called the submerged zone). The sensors that are in the submerged zone for a majority of the time have a higher probability of malfunctioning than the sensors that are installed in the splash zone. Therefore, the miter gate case study captures the generic case of different sensor reliability zones in a structure, and hence, it is an excellent example that can be used to reveal and comprehensively understand the full potential of the proposed optimization algorithm that accounts for sensor reliability.

The primary highlights and novelty of this chapter may therefore be summarized as: (1) it develops a framework to investigate the impact of sensor reliability on Bayesian inference-based damage detection of a structure. To the best of my knowledge, this is the first time that a study of such sort has been conducted; (2) it proposes a sensor placement optimization framework with the consideration of spatial and time-dependent sensor reliability. The consideration of the spatial- and time-dependent sensor reliability allows me to design a reliable sensor network such that it is robust and reliable to the hazards over space and during the sensor lifecycle. The objective function proposed and used in this chapter caters to the need of accounting for the reliability of the sensors over the lifecycle of the structure. The objective function used in this chapter is defined as the expected value of the aggregate absolute deviation of the mean value of the inferred posterior distribution of the state parameter from the true state parameter over the lifecycle of the structure (referred to as expected absolute mean deviation from hereon); (3) it incorporates the degradation of damage over time into the design of a reliable sensor network from the damage detection perspective. The designed sensor network is reliable for the whole lifecycle of the miter gate; (4) it extends the novel and efficient numerical framework established in my previous work (Yang

and Chadha 2021) to solve the computational challenges in the design of a reliable sensor network by integrating Bayesian optimization, surrogate modeling, univariate dimensional reduction, and Sequential Monte Carlo simulation. For the same reason, certain results are borrowed from (Yang and Chadha 2021); (5) finally, this work demonstrates the efficacy of the proposed novel sensor placement design framework on a complex and practical miter gate monitoring application.

The rest of the chapter is arranged as follows. Section 4.3 details the generic Bayesian optimization framework focusing on accounting for sensor reliability over the lifecycle of the structure. Section 4.4 and 4.5 discusses the application of the proposed approach in a real-world miter gate structure. Finally, Section 4.6 concludes the chapter

4.3 The sensor placement optimization framework accounting for sensor reliability

4.3.1 A brief discussion on sensor reliability

Sensor readings are inherently uncertain and possibly unreliable due to technical factors, practical limitations, and noise. Some causes of deviation in sensor measurements relative to the ground truth are primarily due to meteorological conditions, uncertainty in loading, noise in measurements, possible malfunctioning, unforeseen installation issues, improper or no periodic maintenance, or other stochastic influences, to name a few. Broadly, in the ascending order of sensor functionality, there are three scenarios to consider while accounting for uncertainty in sensor readings:

Sensor reliability scenario 1–The sensor is perfectly functional (hence yields reliable readings) but has measurement noise: In this case, the measurement noise is modeled by a zero-mean Gaussian white noise. The previous work (Yang and Chadha 2021; Yang and Chadha 2021)

falls in this category. It is assumed that the mean value of the distribution of *observed* sensor measurement is equal to the *true* sensor reading.

Sensor reliability scenario 2–The sensor is partially malfunction and has measurement noise: The uncertainty in sensor reading now consist of two parts–*measurement noise ignoring sensor reliability* and *reliability bias*–collectively called as *observation noise*. As a consequence of sensor malfunction, the mean value of the distribution of *observed* sensor measurement can no longer be equated to the *true* sensor reading. This bias in measurement (i.e., a shift/deviation in the mean sensor reading relative to the ground truth) must be accounted for in such a case. Additionally, the probability of a sensor malfunction over time varies from one part of the structure to another. This is because parts of the structure may be subjected to different environmental or loading conditions. There are numerous ways to define the reliability of the sensor over time (Ebeling 2004). I adapt an exponentially decaying reliability model, generically defined as:

$$R(t) = e^{-\int_0^t \lambda(t')\mathrm{d}t'} \tag{4.1}$$

In the expression above, $\lambda(t')$ denotes the failure rate at time t'. The variable $R(t)$ could be interpreted as the probability of a sensor being reliable at time t. An exponential decaying reliability model defined in Eq. (4.1) allows me to model a scenario where a sensor tends to malfunction at an accelerating pace as time passes.

Sensor reliability scenario 3-The sensor has completely failed: when a sensor completely fails. there is an omission of a reading. In some cases, this can be easily recognized and rectified by an on-site repair/replacement. Completely missing readings can compromise information quality or

robustness, but the concept of completely missing data will not be considered within the scope of this paper, as it is a topic that has been considered by other researchers in SHM, e.g., (Tang and Bao 2020).

As mentioned above, I focus on the second scenario of accounting for sensor reliability over the lifecycle of the structure in an optimal sensor placement framework. The goals of this section are as follows:

1 To clearly describe the generic problem statement.

2 To generically delineate the Bayes risk functional that accounts for sensor reliability over structure's lifecycle.

3 To describe the optimization framework and all the related numerical pieces required to computationally solve the optimization problem.

For a better understanding of the proposed framework and related concepts, I break down the framework into pieces, each described by a single subsection in the following write-up. I start with the mathematical description of the structural state and its evolution over time.

Remark 1: To proceed further, I lay down common nomenclature used in this chapter. Let Y denote a random variable. Lower case letters y represent realizations of the random variable Y, such that $y \in \Omega_Y$. Here, Ω_Y denotes the measurement space. The probability density function and the cumulative density function are represented by $f_Y(y)$ and $F_Y(y)$. The expected value of a function

$g(y)$ is denoted by $E_Y[g(y)]$. A random variable Y following a Gaussian distribution, with the mean μ_y and standard deviation σ_y is denoted by:

$$f_Y(y) = \frac{1}{\sigma_y}\phi\left(\frac{y-\mu_y}{\sigma_y}\right)$$
$$F_Y(y) = \Phi\left(\frac{y-\mu_y}{\sigma_y}\right) \tag{4.2}$$

No symbolic distinction is made for different dimensions of the measurement space and the random variable. The vector-dimensionality of a random variable is contextual and is defined as needed.

4.3.2 The structural state and the prior damage evolution model

Over the lifespan of the structure as quantified by the time-space $\Omega_T = [0, t_{\max}]$, the structural state evolves, in the absence of any external actions such as maintenance/repair, from *as-built* to approaching a *limit state*, which very generally could be defined as the state in which the structure may no longer perform its intended design functions successfully. Let $\Theta(t)$ denote a random variable representing the state-parameter vector at any time $t \in \Omega_T$. Let $\Theta_{\text{true}}(t)$ denote the true/actual state-parameter vector at any time t. It is assumed that the true state of structure must be inferred probabilistically due to the various inherent uncertainties and stochastic processes. Since our goal is to consider the reliability of the sensors over the life cycle of a structure, we require a damage evolution model. Since the time evolution of structural state is also not precisely known, it is modeled probabilistically, such that, in the absence of any new sensor data, the damage evolution is defined by $f_{\Theta(t)}(\theta(t))$, which denotes the prior distribution of the state parameter vector at time t. Availability of acquired sensor data from an SHM system allows an updated inference of the structural health; since entire life cycles of sensor readings must be

obtained in order to evaluate such a framework, the observed sensor readings will be computationally simulated via a previously validated high-resolution finite element model. The following three subsections deal with modeling sensor measurement data and the Bayesian inference technique used for model-updating.

4.3.3 The sensor design space and the ground truth

Let Ω_E represent the exhaustive sensor design space. Let $e \in \Omega_E$ denote a sensor network design realization consisting of $N_{\text{sg}}(e)$ number of sensors with the measurement space $\Omega_{X_e(t)}$, such that $x_e(t) \in \Omega_{X_e(t)}$ denotes a realization of an observed sensor measurement vector at time t. For SHM related problems, realizations of observed measurement data can be simulated using a validated high-fidelity finite element model (FEM), denoted by g_e, or by a digital surrogate built using FEMgenerated data. Over its lifecycle, a structure is subjected to uncertain loading, generically denoted by the random vector $H(t)$, with a realization $h(t) \in \Omega_{H(t)}$. At any time t, for the given true state $\theta_{\text{true}}(t)$ and the load conditions $h(t)$, the *true/exact/ground-truth* value of the sensor measurements constituting a design e is obtained from the FEM model $g_e(\theta_{\text{true}}(t), h(t); t)$. The *observed* sensor readings are not same as the *true/exact* of the sensor measurements. The next subsection discusses the model adopted to simulate observed data.

4.3.4 The observed sensor reading model, and observation noise model

As mentioned in Section 4.3.1, an *observed* sensor readings are corrupted by *noise and reliability bias* as a consequence of possible sensor partial malfunction, collectively called as the *observation noise*. Sensor reliability is directly affected by the functional state of a sensor which is probabilistic and depends on the zone/location where the sensor is attached. Usually, a structure is divided into various reliability-zones depending on parts of structures exposed to different

environmental and loading conditions. Environmental conditions (like part of structure submerged in water, or subjected to different temperatures), are modeled by their respective load terms (like hydrostatic load or thermal stresses). Therefore, I assume that a structure is zoned (or sub divided) based on various parts of structures being subjected to different loading conditions quantified by the load vector $h(t)$ (or a subset of the load vector). Let $\Omega_S = \{s_{\text{unrel}}, s_{\text{rel}}\}$ denote the set of functional states of a sensor, such that s_{unrel} denote that a sensor is malfunctional and unreliable (reliability scenario 2 mentioned in Section 4.3.1), and s_{rel} represents a *fully-functional* and *reliable* functional state of a sensor (reliability scenario 1 mentioned in Section 4.3.1). The probability mass function $P_{S_{ei}(t)|H(t)}(S_{\text{unrel}} \mid h(t))$ and $P_{S_{ei}(t)|H(t)}(S_{\text{rel}} \mid h(t)) = 1 - P_{S_{ei}(t)|H(t)}(S_{\text{unrel}} \mid h(t))$, denotes the probabilities of i-th sensor being *malfunctional* or *fully-functional* at time t conditioned upon the loading-zone (defined in turn by the load vector and the sensor location of the i-th vector in design e) respectively. Here, $S_{ei}(t)$ denotes the random variable modeling the functional state of the i-th sensor in design e at time t. Since the functional state of a sensor is dependent on the loading-zone the sensor is installed, the probabilities $P_{S_{ei}(t)|H(t)}(s_{\text{unrel}} \mid h(t))$ and $P_{S_{ei}(t)|H(t)}(s_{\text{rel}} \mid h(t))$ are conditioned on the load vector $h(t)$. The approach for evaluating these probabilities may vary from one problem to another. In this chapter, the probability of a sensor being fully functional at time t takes the exponential decay form as described in Eq. (4.1).

Let the random vector $\zeta_e(t)$ represent *observation noise* in sensor readings at time t, such that $\varepsilon_e(t) = \left(\varepsilon_{e1}(t), \varepsilon_{e2}(t), \cdots, \varepsilon_{eN_{\text{sg}}(e)}(t)\right) \in \Omega_{\zeta_e(t)}$ is a realization of observation noise vector for the design e at time t (where $\varepsilon_{ei}(t)$ denotes observation noise in the i-th sensor). The observation noise in the i-th sensor in design e also depends on the loading-zone where it is

attached. Let $f_{\zeta_{ei}(t)|H(t)}(\varepsilon_{ei}(t) \mid h(t))$ denote the distribution of observation noise in i-th sensor conditioned upon the loading zone at which the sensor is installed (which is assumed to be defined in terms of the various loading situations), and $f_{\zeta_e(t)|H(t)}(\varepsilon_e(t) \mid h(t))$ denote the joint distribution of observation noise in all the sensors constituting the design e. For the purpose of Bayesian inference, I need to model the conditional distribution of observation noise $f_{\zeta_e(t)|H(t)}(\varepsilon_e(t) \mid h(t))$. I assume the following:

1 The observation noise in sensors is assumed to be statistically independent of each other Although the true sensor measurements are related by virtue of the physics of the problem, noise in these readings can be reasonably assumed to be statistically independent. That is,

$$f_{\zeta_e(t)|H(t)}(\varepsilon_e(t) \mid h(t)) = \prod_{i=1}^{N_{\mathrm{sg}}(e)} f_{\zeta_{ei}(t)|H(t)}(\varepsilon_{ei}(t) \mid h(t)) \tag{4.3}$$

2 I assume that if sensors are fully reliable and functional, the observation noise follows a zero-mean Gaussian distribution (white noise). That is, when sensors are *fully functional* (scenario 1 mentioned in Section 4.3.1), let the random vector $\bar{\zeta}_e(t)$ represent noise in sensor readings at time t, such that $\bar{\varepsilon}_e(t) = \left(\bar{\varepsilon}_{e1}(t), \bar{\varepsilon}_{e2}(t), \cdots, \bar{\varepsilon}_{eN_{sg}(e)}(t)\right) \in \Omega_{\bar{\zeta}_e(t)}$ is a realization of noise vector for the design e at time t (where $\bar{\varepsilon}_{ei}(t)$ denotes observation noise in the i-th sensor). That is:

$$f_{\bar{\zeta}_{ei}(t)}(\bar{\varepsilon}_{ei}(t)) = \frac{1}{\sigma_{\bar{\varepsilon}_{ei}(t)}} \phi\left(\frac{\bar{\varepsilon}_{ei}(t)}{\sigma_{\bar{\varepsilon}_{ei(t)}}}\right) \tag{4.4}$$

Here, $\sigma_{\bar{\varepsilon}_{ei}(t)}$ denotes the standard-deviation in reading of the i-th sensor in design e at time t. I note that $f_{\bar{\varepsilon}_e(t)}(\bar{\varepsilon}_{ei}(t))$ is not conditioned upon the loading-zones in which a structure

is divided since sensors are fully functional irrespective of the loading-zone they are attached to.

3 I now consider reliability scenario 2 mentioned in Section 4.3.1. I model the reliability aspect of sensors by incorporating *reliability bias* in observation noise. The consequence of sensors malfunctioning is that the probability distribution of observation noise has a non-zero mean. The generic form of the probability distribution of observation noise $f_{\zeta_e(t)|H(t)}(\varepsilon_e(t) \mid h(t))$ is obtained as a weighted sum of the zero-mean probability distribution defined in Eq. (4.4) and the probability distribution of sensor reliability bias. The sensor reliability bias in the i-th sensor of design e at time t is modeled by sensor bias factor, denoted by a random variable $\eta_{ei}(t)$. Let $n_{ei}(t)$ denote a realization of $\eta_{ei}(t)$. I model the sensor bias factor by a non-zero mean $\mu_{n_{ei}(t)}$ Gaussian distribution with the standard-deviation $\sigma_{n_{ei}(t)}$. Therefore, a realization of observed sensor noise can be written as:

$$\varepsilon_{ei}(t) = \begin{cases} \bar{\varepsilon}_{ei}(t) & \text{for sensor functional state } s_{\text{rel}} \\ \bar{\varepsilon}_{ei}(t) + n_{ei}(t) & \text{for sensor functional state } s_{\text{unrel}} \end{cases} \tag{4.5}$$

With all the pieces at hand, I obtain the distribution of the observation noise as:

$$f_{\zeta_{ei}(t)|H(t)}(\varepsilon_{ei}(t)|h(t)) = P_{S_{ei}(t)|H(t)}(s_{\text{rel}}|h(t)) \overbrace{\frac{1}{\sigma_{\bar{\varepsilon}_{ei}(t)}}\phi\left(\frac{\varepsilon_{ei}(t)}{\sigma_{\bar{\varepsilon}_{ei}(t)}}\right)}^{\substack{\text{Noise PDF for} \\ \text{reliability scenario 1}}} + P_{S_{ei}(t)|H(t)}(s_{\text{unrel}}|h(t)) \overbrace{\frac{1}{\sqrt{\sigma^2_{\bar{\varepsilon}_{ei}(t)} + \sigma^2_{n_{ei}(t)}}}\phi\left(\frac{\varepsilon_{ei}(t) - \mu_{n_{ei}(t)}}{\sqrt{\sigma^2_{\bar{\varepsilon}_{ei}(t)} + \sigma^2_{n_{ei}(t)}}}\right)}^{\substack{\text{Noise PDF for} \\ \text{reliability scenario 2}}}. \quad (4.6)$$

In the expression above, $P_{S_{ei}(t)|H(t)}(S_{\text{rel}} \mid h(t))$ and $P_{S_{ei}(t)|H(t)}(s_{\text{unrel}} \mid h(t))$ can be interpreted as weights assigned to observation noise weighing the possibility of the i-th sensor of design e being *reliable* or *unreliable* at time t respectively. It is these sensor functional state probabilities that makes the distribution $f_{\zeta_{ei}(t)|H(t)}(\varepsilon_{ei}(t) \mid h(t))$ conditioned on $H(t)$.

I obtain the observed sensor readings at time t using the following measurement model:

$$x_e(t) = g_e(\theta_{\text{true}}(t), h(t); t) + \varepsilon_e(t) \quad (4.7)$$

Having decided on the observation noise model and observed sensor readings, I now investigate the impact of sensor reliability on Bayesian inference that will be later used to obtain posterior damage evolution model. This is discussed in the following subsection.

4.3.5 Bayesian inference and the posterior damage evolution model

Bayesian inference performed in my previous works (Yang and Chadha 2021, Yang and Chadha 2021, Chadha and Hu 2021) focused on *reliability* scenario 1 mentioned in Section 4.3.1. In this section, I expand the Bayesian inference to a more generic *reliability scenario* 2. As a result

of accounting for sensor reliability, it is not only vital to consider the observed measurements $x_e(t)$ obtained from sensors constituting the design e at time t, but I must also consider the information about which loading-zone (or reliability-zone) each sensor in the design e are attached to at time t for updating the posterior distribution of state parameter. Let $f_{\Theta(t)|X_e(t),H(t)}(\theta(t) \mid x_e(t), h(t))$ denote the posterior distribution of the state parameter conditioned on observed measurements $x_e(t)$ and the loading-zone decided by $h(t)$. In other words, $f_{\Theta(t)X_{(t)}H(t)}(\theta(t) \mid x_e(t), h(t))$ defines the posterior damage evolution model. To elaborate on the Bayesian inference of state parameter in this section, I fix time t and omit explicitly mentioning the time dependence (for example, $\theta(t)$ is simply written as θ). The Bayes theorem in current context may be written as:

$$f_{\Theta|X_e,H}(\theta \mid x_e, h) = \frac{f_{X_e|\Theta,H}(x_e \mid \theta, h) f_{\Theta|H}(\theta \mid h)}{f_{X_e|H}(x_e \mid h)} \tag{4.8}$$

I note the following two points:

1 In the context of inferring state parameter, the evidence $f_{X_e|H}(x_e \mid h)$ is just a normalizing constant.

2 In practice, for time-evolving systems, the prior distribution used in the Bayesian inference should be obtained by using the prior damage evolution model $f_{\Theta(t)}(\theta(t))$, observations from all the previous time steps, and the historic information about loading zone that the sensors in design e were exposed to. That would essentially be model updating that incorporates all the historic data. However, Bayesian inference is a computationally expensive process, and performing historic data updating would exponentially increase the computational cost and complexity. As a consequence of this, I simply use the prior

distribution as the distribution obtained from the prior damage evolution model $f_{\Theta(t)}(\theta(t))$ at time t.

Therefore, the Bayes theorem now becomes,

$$f_{\Theta|X_e,H}(\theta \mid x_e, h) \propto f_{X_e|\Theta}(x_e \mid \theta) f_{\Theta}(\theta) \tag{4.9}$$

Since the sensor noise is conditioned on the loading-zone (which in turn is decided by the loading-vector $h(t)$), the likelihood of obtaining a realization x_e of sensor readings is conditioned on the load vector $h(t)$. This in turn is the reason for conditional dependence of the posterior distribution of state parameter on the load vector $h(t)$. The likelihood $f_{X_e|\Theta,H}(x_e \mid \theta, h)$ can be obtained by using the observation noise model described in Eq. (4.6) and the measurement model described in Eq. (4.7), such that:

$$\begin{aligned} f_{X_e|\Theta,H}(x_e \mid \theta, h) \quad = \prod_{i=1}^{N_{\mathrm{sg}}(e)} \Bigg(& P_{S_{ei}|H}(s_{\mathrm{rel}} \mid h) \frac{1}{\sigma_{\bar{\varepsilon}_{ei}}} \phi\left(\frac{x_{ei} - g_{ei}}{\sigma_{\bar{\varepsilon}_{ei}}}\right) \\ & + P_{S_{ei}|H}(s_{\mathrm{unrel}} \mid h) \frac{1}{\sqrt{\sigma_{\bar{\varepsilon}_{ei}}^2 + \sigma_{n_{ei}}^2}} \phi\left(\frac{(x_{ei} - g_{ei}) - \mu_{n_{ei}}}{\sqrt{\sigma_{\bar{\varepsilon}_{ei}}^2 + \sigma_{n_{ei}}^2}}\right) \Bigg) \end{aligned} \tag{4.10}$$

Since the relationship between the sensor measurement x_e and the state parameter θ is highly nonlinear, I must rely on numerical approximation techniques like Markov chain Monte Carlo (MCMC) methods, particle filter, and sequential Monte Carlo (SMC) approach in recursive mode to solve the inference problem. I use the particle filtering technique in this chapter. Obtaining the posterior numerically using particle filtering requires evaluating the likelihood at numerous values of state parameter, called the particles. Usually, particle filtering is used for sequential updating of the posterior distribution for a dynamic system, i.e., the case where new information on the system

is available as time evolves. However, in this case, I just have one set of data, and I aim at obtaining the posterior in a single step. The following summarizes the process of obtaining the posterior distribution of state parameter:

(a) For the assumed true state parameter θ_{true} and the chosen/fixed input loading h_{true}, simulate an observed/measurement strain data $x_e \in \Omega_{X_e}$.

(b) I choose $N_{\text{particles}}$ number of discrete values of the state parameter, called particles.

(c) At each of these $N_{\text{particles}}$ particles, and for a given loads h_{true}, the digital twin or the FEM yields the true strain value at the i-th strain location for the state parameter particle θ_j and is denoted by $g_{e_i}(\theta_j, h_{\text{true}})$, where $i \leq N_{\text{sg}}(e)$ and $j \leq N_{\text{particles}}$. Using Eq. (4.10), I obtain numerical value of likelihood of the measurement $x_e \in \Omega_{X_e}$ given the state parameter θ_j.

(d) Evaluate the weight w_j for each particle as:

$$w_j = \frac{f_{X_e|\Theta,H}(x_e \mid \theta_j, h_{\text{true}})}{\sum_{k=1}^{N_{\text{paricles}}} f_{X_e|\Theta,H}(x_e \mid \theta_k, h_{\text{true}})} \tag{4.11}$$

(e) Calculate the cumulative weights to observe big jumps. Resample the weighted particles to obtain unweighted samples of the posterior distribution over state parameter.

I now have all the pieces required to define the Bayes risk functional.

4.3.6 Bayes risk functional accounting for reliability of sensors

The proposed Bayes risk targets the following: (1) quantify the reliability of a sensor design over the lifecycle; (2) consider all the uncertainties in the sensor data, loading, and the prior

understanding of the state over the structure's lifecycle. Intending to eventually capture the reliability of a sensor design e over the lifecycle Ω_T, I start with quantifying the reliability of a sensor design at time t. I define *sensor reliability* risk at time t, denoted by $\mathcal{L}(\theta_{\text{true}}(t), h(t), \varepsilon_e(t); t)$, as the absolute deviation between the true structural state $\theta_{\text{true}}(t)$ at time t and the mean of the posterior distribution of structural state, denoted by $\mu_{\Theta(t)|X(t)H(t)}$, such that:

$$\mathcal{L}(\theta_{\text{true}}(t), h(t), \varepsilon_e(t); t) = \left|\mu_{\Theta(t)|X_e(t),H(t)} - \theta_{\text{true}}(t)\right| \tag{4.12}$$

Fig. (4.1) illustrates a function or a module called "*Evaluate the risk* $\mathcal{L}(\theta(t), h(t), \varepsilon_e(t); t)$ " that obtains the risk $\mathcal{L}\big(\theta(t), h(t), \varepsilon_\rho(t); t\big)$ for a given design e at time t and the input variables $\big(\theta(t), h(t), \varepsilon_\rho(t)\big)$ The module does so in a three step process that requires running the FEM model or digital twin $(N_{\text{Particles}} + 1)$ times. The first step is to simulate observed/measurement data $x_e(t)$ for a design e by assuming a true state parameter $\theta_{\text{true}}(t)$ (simply denoted here as $\theta(t)$), load vector $h(t)$, and a noise structure $f_{\zeta_{ei}(t)|H(t)}(\varepsilon_{ei}(t) \mid h(t))$. The second step is to obtain the posterior distribution of state parameter given the measurement $x_e(t)$ obtained in the first step using particle filter. Finally, the third step is to evaluate the absolute deviation of the posterior's mean from the assumed true state parameter $\theta(t)$.

The expected sensor reliability risk at time t, denoted by $\mathfrak{E}(e; t)$: $\Omega_E \times \Omega_T \longrightarrow \mathbb{R}$, is then defined as follows:

$$\begin{aligned}\mathfrak{E}(e; t) = \iiint_{\Omega_\Theta(t)\times\Omega_{\zeta_e(t)}\times\Omega_{H(t)}} & f_{\Theta(t)}(\theta(t)) f_{H(t)}(h(t)) f_{\zeta_e(t)|H(t)}(\varepsilon_e(t) \\ & \mid h(t)) \mathcal{L}(\theta(t), h(t), \varepsilon_e(t); t) \mathrm{d}\theta(t) \mathrm{d}h(t) \mathrm{d}\varepsilon_e(t)\end{aligned} \tag{4.13}$$

Finally, the *aggregate expected sensor reliability* risk over the lifecycle, or simply the Bayes risk, denoted by $\mathfrak{E}_{\mathrm{LC}}(e)\colon \Omega_E \longrightarrow \mathbb{R}$, is defined as follows:

$$\mathfrak{E}_{\mathrm{LC}}(e) = \int_{\Omega_T} \mathfrak{E}(e;t)\mathrm{d}t \approx \sum_{k=1}^{t_k \in \Omega_T} \mathfrak{E}(e;t_k) \tag{4.14}$$

Here, the subscript LC in the notation of Bayes risk $\mathfrak{E}_{\mathrm{LC}}(e)$ represents *life cycle*.

Remark 2: The figure 4.1 shows that to obtain the risk $\mathcal{L}(\theta_{\text{true}}, h(t), \varepsilon_e(t); t)$ for a given instance of the design, time, and input parameters requires running the digital twin or the FEM a total of $\left(N_{\text{particles}} + 1\right)$ times.

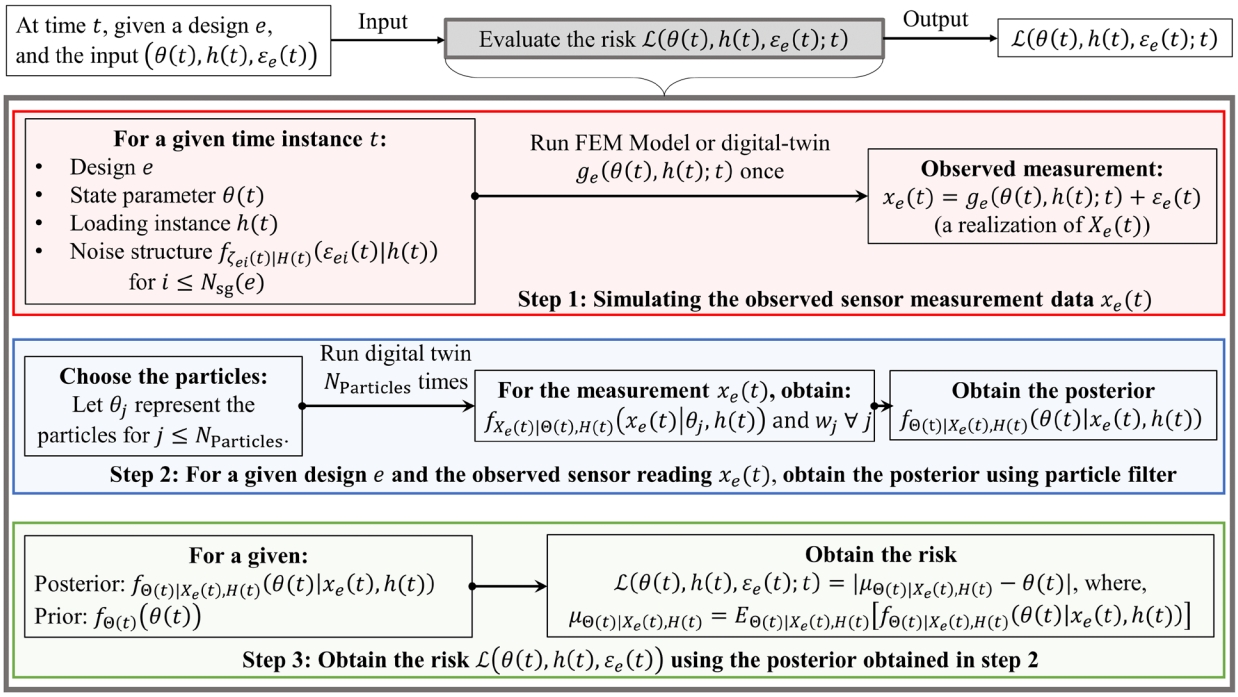


Figure 4.1: A three step process for evaluating the risk $\mathcal{L}(\theta_{\text{true}}, \varepsilon_e(t), h(t); t)$

4.3.7 Evaluating sensor reliability risk

The Bayes risk, the sensor observation noise model, and by extension the Bayesian inference approach adopted in this chapter is different from the respective entities used in Yang et al. (Yang and Chadha 2020) primarily because the target optimization objectives of these papers are different. However, the numerical techniques used to approximate the risk function $\mathfrak{E}(e;t)$ fundamentally stay the same. I briefly delineate a more generic form of the approximation procedure that was initially proposed in Yang et al. (Yang and Chadha 2020).

Its clear from expression Eq. (4.14) that calculating $\mathfrak{E}_{\mathrm{LC}}(e)$ requires evaluating $\mathfrak{F}(e;t)$ over different time instances spanning Ω_T. Since $\mathfrak{E}(e;t)$ (as defined in Eq. (4.12)) is a high-dimensional integral with non-linear integrand, numerical techniques are indispensable to approximate $\mathfrak{E}(e;t)$. This section deals with exploiting univariate dimensional reduction used in tandem with Gasuss-Hermite quadrature to evaluate $\mathfrak{E}(e;t)$. The overall process can be summarized in three steps:

1 Transform all the probability density functions in Eq. (4.13) into their standard normal form.

2 Obtain an approximated form of the risk function $\mathcal{L}(\theta_{\text{true}}, h(t), \varepsilon_e(t); t)$ by using univariate dimensional reduction approximation.

3 Obtain the approximated form of the integral in Eq. (4.13) using Gauss-Hermite quadrature.

4.3.7.1 Revisiting sensor reliability risk

I redefine the expression of $\mathfrak{E}(e;t)$ that is desirable for further simplifying it using univariate dimensional reduction and Gauss-Hermite quadrature. To use Gauss-Hermite quadrature approximation, I transforms the random variables $\Theta(t), \zeta_e(t)$, and $H(t)$, into their standard normal counterparts denoted by a tilde ($\tilde{.}$) over the respective quantity. Therefore, let $\tilde{\Theta}(t), \tilde{\zeta}_e(t)$, and $\tilde{H}(t)$ denote the standard normal random variables that are counterparts of $\Theta(t), \zeta_e(t)$, and $H(t)$, such that their realizations are denoted as $\tilde{\theta}(t), \tilde{\varepsilon}_e(t)$, and $\tilde{h}(t)$. I can redefine the sensor reliability risk as:

$$\mathfrak{E}(e;t) = E_{\tilde{\Theta}(t)\tilde{\zeta}_e(t)\tilde{H}(t)}\left[\tilde{\mathcal{L}}\left(\tilde{\theta}(t), \tilde{h}(t), \tilde{\varepsilon}_e(t); t\right)\right] \tag{4.15}$$

In the equation above, I have:

$$\tilde{\mathcal{L}}\left(\tilde{\theta}(t), \tilde{h}(t), \tilde{\varepsilon}_e(t); t\right) = \mathcal{L}(\theta(t), h(t), \varepsilon_e(t); t) \tag{4.16}$$

The current form of $\mathfrak{E}(e;t)$, as expressed in Eq. (15), is suitable for its numerical approximation as detailed in the next subsection.

4.3.7.2 Univariate dimensional reduction and Gauss-Hermite quadrature

Let N_H, N_Θ, and $N_{\text{sg}}(e)$ denote the size of the load vector $h(t)$ (or $\tilde{h}(t)$), the state parameter vector $\theta(t)$ (or $\tilde{\theta}(t)$), and the observation noise vector $\varepsilon_e(t)$ (or $\tilde{\varepsilon}_e(t)$) respectively, such that:

$$
\begin{aligned}
h(t) &= \left(h_1(t), h_2(t), \cdots, h_{N_H}(t)\right) \\
\tilde{h}(t) &= \left(\tilde{h}_1(t), \tilde{h}_2(t), \cdots, \tilde{h}_{N_H}(t)\right) \\
\theta(t) &= \left(\theta_1(t), \theta_2(t), \cdots, \theta_{N_\theta}(t)\right) \\
\tilde{\theta}(t) &= \left(\tilde{\theta}_1(t), \tilde{\theta}_2(t), \cdots, \tilde{\theta}_{N_\theta}(t)\right); \\
\varepsilon_e(t) &= \left(\varepsilon_{e1}(t), \varepsilon_{e2}(t), \cdots, \varepsilon_{eN_{\mathrm{sg}(e)}}(t)\right) \\
\tilde{\varepsilon}_e(t) &= \left(\tilde{\varepsilon}_{e1}(t), \tilde{\varepsilon}_{e2}(t), \cdots, \tilde{\varepsilon}_{eN_{\mathrm{sg}(e)}}(t)\right)
\end{aligned}
\tag{4.17}
$$

The realizations of the random vector and their standard normal counterparts are related by the Rosenblatt transformation as follows (see Eq. (4.26) of (Rahman and Xu 2004)):

$$
\begin{aligned}
h(t) &= F_{H(t)}^{-1}\left(F_{\tilde{H}(t)}(\tilde{h}(t))\right) \\
\theta(t) &= F_{\Theta(t)}^{-1}\left(F_{\tilde{\Theta}(t)}(\tilde{\theta}(t))\right) \\
\varepsilon_e(t) &= F_{\zeta_e(t)|H(t)}^{-1}\left(F_{\tilde{\zeta}_e(t)|\tilde{H}(t)}(\tilde{\zeta}(t) \mid \tilde{h}(t))\right)
\end{aligned}
\tag{4.18}
$$

Here, $F_A^{-1}(\cdot)$ is an inverse cumulative distribution function oa a generic random variable A. Let $\Theta_i(t), H_i(t)$, and $\zeta_{ei}(t)$ denote the random variables corresponding to the i-th component of the random vector $\Theta(t), H(t)$, and $\zeta_e(t)$ respectively. It is assumed that the dimensionality of state parameter and load vector does not change with time. As a result of the Rosenblatt transformation transformation, collective components of the standard normal random vectors $\tilde{\Theta}(t), \tilde{H}(t)$, and $\tilde{\zeta}_e(t)$ are statistically independent. I obtain a further compacted form of $\mathfrak{E}(e;t)$ by defining the following space,

$$
\begin{aligned}
\Omega_{\tilde{\Psi}_e(t)} &= \Omega_{\tilde{\Theta}(t)} \times \Omega_{\tilde{H}(t)} \times \Omega_{\tilde{\zeta}_e(t)} \\
&= \left(\Omega_{\tilde{\Theta}_1(t)} \times \cdots \times \Omega_{\tilde{\Theta}_{N_\Theta}(t)}\right) \times \left(\Omega_{\tilde{H}_1(t)} \times \cdots \times \Omega_{\tilde{H}_{N_H}(t)}\right) \times \left(\Omega_{\tilde{\zeta}_{e1}(t)} \times \cdots \times \Omega_{\tilde{\zeta}_{eN_{\mathrm{sg}(e)}}(t)}\right)
\end{aligned}
\tag{4.19}
$$

I note that $\Omega_{\tilde{\Psi}_e(t)}$ is $\left(N_\Theta + N_H + N_{\mathrm{sg}}(e)\right)$ dimensional vector space. Let $\tilde{\psi}_e(t)$ denote a realization of $\tilde{\Psi}_e(t)$, such that:

$$\begin{aligned}\tilde{\psi}_e(t) \quad &= \left(\tilde{\theta}(t); \tilde{h}(t); \tilde{\varepsilon}_e(t)\right) \\ &= (\overbrace{\tilde{\theta}_1(t), \tilde{\theta}_2(t), \cdots, \tilde{\theta}_{N_\Theta}(t); \tilde{h}_1(t), \tilde{h}_2(t), \cdots, \tilde{h}_{N_H}(t); \tilde{\varepsilon}_{e1}(t), \tilde{\varepsilon}_{e2}(t), \cdots, \tilde{\varepsilon}_{eN_{\mathrm{sg}}(e)}(t)}^{\text{Total Length of } N_\Theta + N_H + N_{\mathrm{sg}}(e)}\Big)) \\ &= \left(\tilde{b}_1(t), \cdots, \tilde{b}_{N_\Theta}(t); \tilde{b}_{N_\Theta+1}(t), \cdots, \tilde{b}_{N_\Theta+N_H}(t); \tilde{b}_{N_\Theta+N_H+1}(t), \cdots, \tilde{b}_{N_\Theta+N_H+N_{\mathrm{sg}}(e)}(t)\right)\end{aligned} \tag{4.20}$$

In the equation set above, semicolon is used to distinguish between sub-vectors corresponding to the random vectors $\Theta(t), H(t)$, and $\zeta_e(t)$ (consisting of N_Θ, N_H, and $N_{\mathrm{sg}}(e)$ random variables respectively). The variable $\tilde{b}_i(t)$ represents the i-th component of the realization vector $\tilde{\psi}_e(t)$. Let $\Omega_{\tilde{B}_i(t)}$ denote the space to which $\tilde{b}_i(t)$ belongs. Then,

$$\Omega_{\tilde{\Psi}_e(t)} = \Omega_{\tilde{B}_1(t)} \times \cdots \times \Omega_{\tilde{B}_{(N_\Theta+N_H+N_{\mathrm{sg}}(e))}(t)}. \tag{4.21}$$

The sensor reliability risk can now be written as:

$$\mathfrak{E}(e;t) = E_{\tilde{\Psi}_e(t)}\left[\tilde{\mathcal{L}}\left(\tilde{\Psi}_e(t); t\right)\right] \tag{4.22}$$

As such, the integral in Eq. (4.22) is high-dimensional expectation in $\left(N_\Theta + N_H + N_{\mathrm{sg}}(e)\right)$ dimensional space, making classical multivariate quadrature rules (e.g., quasi Monte Carlo or Smolyak sparse grids) prohibitively expensive. I conquer this challenge by using univariate dimensional reduction and Gauss-Hermite quadrature to approximate the integral in Eq. (4.22). I start by defining the following vectors, each containing $\left(N_\Theta + N_H + N_{\mathrm{sg}}(e)\right)$ elements:

$$
\begin{aligned}
&\tilde{\psi}_0(t) = \overbrace{(0,0,\cdots,0;\, 0,0,\cdots,0;\, 0,0,\cdots,0)}^{\text{Length } N_\Theta+N_H+N_{\mathrm{sg}}(e)} \\
&\tilde{\psi}_1(t) = \left(\tilde{\theta}_1(t),0,\cdots,0;\, 0,0,\cdots,0;\, 0,0,\cdots,0\right) \\
&\tilde{\psi}_2(t) = \left(0,\tilde{\theta}_2(t),\cdots,0;\, 0,0,\cdots,0;\, 0,0,\cdots,0\right); \\
&\quad\vdots \\
&\tilde{\psi}_{N_\Theta}(t) = \left(0,0,\cdots,\tilde{\theta}_{N_\Theta}(t);\, 0,0,\cdots,0;\, 0,0,\cdots,0\right) \\
&\tilde{\psi}_{(N_\theta+1)}(t) = \left(0,0,\cdots,0;\, \tilde{h}_1(t),0,\cdots,0;\, 0,0,\cdots,0\right) \\
&\tilde{\psi}_{(N_\theta+2)}(t) = \left(0,0,\cdots,0;\, 0,\tilde{h}_2(t),\cdots,0;\, 0,0,\cdots,0\right) \\
&\quad\vdots \\
&\tilde{\psi}_{(N_\theta+N_H)}(t) = \left(0,0,\cdots,0;\, 0,0,\cdots,\tilde{h}_{N_H}(t);\, 0,0,\cdots,0\right) \\
&\tilde{\psi}_{(N_\theta+N_H+1)}(t) = (0,0,\cdots,0;\, 0,0,\cdots,0;\, \tilde{\varepsilon}_{e1}(t),0,\cdots,0) \\
&\tilde{\psi}_{(N_\theta+N_H+2)}(t) = (0,0,\cdots,0;\, 0,0,\cdots,0;\, 0,\tilde{\varepsilon}_{e2}(t),\cdots,0) \\
&\quad\vdots \\
&\tilde{\psi}_{(N_\theta+N_H+N_{\mathrm{sg}}(e))}(t) = \left(0,0,\cdots,0;\, 0,0,\cdots,0;\, 0,0,\cdots,\tilde{\varepsilon}_{eN_{\mathrm{sg}}(e)}(t)\right)
\end{aligned}
\tag{4.23}
$$

The sensor reliability risk $\tilde{\mathcal{L}}\left(\tilde{\psi}_e(t);t\right)$ can now be approximated using the definitions in Eq. (4.23) and univariate dimensional reduction (Rahman and Xu 2004), such that:

$$
\tilde{\mathcal{L}}\left(\tilde{\psi}_e(t);t\right) \approx -\left(N_\Theta+N_H+N_{\mathrm{sg}}(e)-1\right)\tilde{\mathcal{L}}\left(\tilde{\psi}_0(t);t\right) + \sum_{i=1}^{\left(N_\Theta+N_H+N_{\mathrm{sg}}(e)\right)} \tilde{\mathcal{L}}\left(\tilde{\psi}_i(t);t\right) \tag{4.24}
$$

Substituting Eq. (4.24) into Eq. (4.22), I get,

$$
\begin{aligned}
\mathfrak{E}(e;t) \quad &\approx -\left(N_\Theta+N_H+N_{\mathrm{sg}}(e)-1\right)E_{\tilde{\Psi}_e(t)}\left[\tilde{\mathcal{L}}\left(\tilde{\psi}_0(t);t\right)\right] + \sum_{i=1}^{\left(N_\Theta+N_H+N_{\mathrm{sg}}(e)\right)} E_{\tilde{\Psi}_e(t)}\left[\tilde{\mathcal{L}}\left(\tilde{\psi}_i(t);t\right)\right] \\
&= -\left(N_\Theta+N_H+N_{\mathrm{sg}}(e)-1\right)\tilde{\mathcal{L}}\left(\tilde{\psi}_0(t);t\right) + \sum_{i=1}^{\left(N_\Theta+N_H+N_{\mathrm{sg}}(e)\right)} E_{\tilde{\Psi}_e(t)}\left[\tilde{\mathcal{L}}\left(\tilde{\psi}_i(t);t\right)\right]
\end{aligned}
\tag{4.25}
$$

In the equation above,

$$E_{\tilde{\Psi}_{e(t)}}\left[\tilde{\mathcal{L}}\left(\tilde{\ }_i(t);t\right)\right]=\int_{\Omega_{\tilde{\psi}_{e(t)}}} f_{\tilde{\Psi}_{e(t)}}\left(\tilde{\psi}_e(t)\right)\tilde{\mathcal{L}}\left(\tilde{\psi}_i(t);t\right)\mathrm{d}\tilde{\psi}_e(t) \tag{4.26}$$

To further simplify $\mathfrak{E}(e;t)$, I focus on the integral in Eq. (4.26). Since $\tilde{\Theta}(t), \tilde{\zeta}_e(t)$, and $\tilde{H}(t)$ are statistically independent random variables, the probability density function $f_{\tilde{\Psi}_{e(t)}}\left(\tilde{\psi}_e(t)\right)$ can be written as:

$$\begin{aligned} f_{\tilde{\Psi}_e}\left(\tilde{\psi}_e\right) &= \prod_{i=1}^{N_\ominus} f_{\tilde{\Theta}_{i(t)}}\left(\tilde{\theta}_i(t)\right)\cdot\prod_{j=1}^{N_H} f_{\tilde{H}_{j(t)}}\left(\tilde{h}_j(t)\right)\cdot\prod_{k=1}^{N_{\mathrm{sg}}(e)} f_{\tilde{\xi}_{ek(t)}}\left(\tilde{\varepsilon}_{ek}(t)\right) \\ &= \prod_{i=1}^{N_\ominus}\phi\left(\tilde{\theta}_i(t)\right)\cdot\prod_{j=1}^{N_H}\phi\left(\tilde{h}_j(t)\right)\cdot\prod_{k=1}^{N_{\mathrm{sg}}(e)}\phi\left(\tilde{\varepsilon}_{ek}(t)\right) \\ &= \prod_{m=1}^{\left(N_\Theta+N_H+N_{\mathrm{sg}}(e)\right)}\phi\left(\tilde{b}_m(t)\right)=\prod_{m=1}^{\left(N_\Theta+N_H+N_{\mathrm{sg}}(e)\right)}\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\tilde{b}_m^2(t)}\right) \end{aligned} \tag{4.27}$$

In the equation above, $\tilde{b}_m(t)$ represents the m-th component of the realization vector $\tilde{\psi}_e(t)$. I note that for any function of the form $g(x\in X, y=0\in Y), E_{XY}(g(x,0))=E_X(g(x,0))$, provided X and Y are statistically independent random variables. This allows me to simplify the integral in Eq. (4.26) as:

$$E_{\tilde{\Psi}_{e(t)}}\left[\tilde{\mathcal{L}}\left(\tilde{\Psi}_i(t);t\right)\right]=\frac{1}{\sqrt{2\pi}}\int_{\Omega_{\tilde{B}_i(t)}}\tilde{\mathcal{L}}\left(\tilde{\psi}_i(t);t\right)e^{-\frac{1}{2}\tilde{b}_i^2(t)}\mathrm{d}\tilde{b}_i(t) \tag{4.28}$$

By its very design, Gauss-Hermite quadrature is suitable to discretely estimate an integral of the form $\left(\int_x g(x)e^{-x^2}\right)$, for any function $g(x)$. Therefore, using Gauss-Hermite quadrature, I approximate the integral in Eq. (4.28) as follows,

$$E_{\tilde{\Psi}_{e(t)}}\left[\tilde{\mathcal{L}}\left(\tilde{\psi}_i(t);t\right)\right]\approx\frac{1}{\sqrt{\pi}}\sum_{n=1}^{r} w_n\tilde{\mathcal{L}}\left(\tilde{q}_{i,n};t\right), \text{ where } \tilde{q}_{ij,n}=\begin{cases}\tilde{\psi}_{ij}(t)=0 & i\neq j\\ \alpha_n & i=j\end{cases} \tag{4.29}$$

In the equation above, $\tilde{q}_{ij,n}$ and $\tilde{\psi}_{ij}(t)$ represents the j-th component of the vectors $\tilde{q}_{i,n}$ and $\tilde{\psi}_i(t)$ respectively; $\boldsymbol{r}$ denote the quadrature order; w_n gives the quadrature weights; and α_n gives the point of evaluation of the function for $n \leq r$. In this chapter, I use $r = 3$, for which $w_1 = \frac{2}{3}\sqrt{\pi}$ $w_2 = \frac{1}{6}\sqrt{\pi}, w_3 = -\frac{1}{6}\sqrt{\pi}, \alpha_1 = 0, \alpha_1 = \frac{\sqrt{6}}{2}, \alpha_1 = -\frac{\sqrt{6}}{2}$. The quadrature order of $r = 3$ satisfies the computational accuracy that this problem demands and at the same time leads to a computationally efficient numerical estimation of Bayes risk. Substituting Eq. (4.29) into Eq. (4.25), I get

$$\begin{aligned}\mathfrak{E}(e;t) \approx & -\big(N_\Theta + N_H + N_{\mathrm{sg}}(e) - 1\big)\tilde{\mathcal{L}}\big(\tilde{\psi}_0(t);t\big) \\ & + \frac{1}{\sqrt{\pi}} \sum_{i=1}^{(N_\Theta + N_H + N_{\mathrm{sg}}(e))} \sum_{n=1}^{r} w_n \tilde{\mathcal{L}}\big(\tilde{q}_{i,n};t\big)\end{aligned} \tag{4.30}$$

Finally, the Bayes risk functional over the lifecycle is approximated as:

$$\begin{aligned}\mathfrak{E}_{\mathrm{LC}}(e) \approx \sum_{k=1}^{N_T} \Bigg(& -\big(N_\Theta + N_H + N_{\mathrm{sg}}(e) - 1\big)\tilde{\mathcal{L}}\big(\tilde{\Psi}_0(t_k);t_k\big) \\ & + \frac{1}{\sqrt{\pi}} \sum_{i=1}^{(N_\Theta + N_H + N_{\mathrm{sg}}(e))} \sum_{n=1}^{r} w_n \tilde{\mathcal{L}}\big(\tilde{q}_{i,n};t_k\big) \Bigg)\end{aligned} \tag{4.31}$$

For sake of evaluating the sum in Eq. (4.31), I discretize Ω_T into N_T time steps. The figure below shows an algorithmic module "*Evaluate the Bayes ri*sk $\mathfrak{E}_{\mathrm{LC}}(e)$" that obtain the Bayes risk functional.

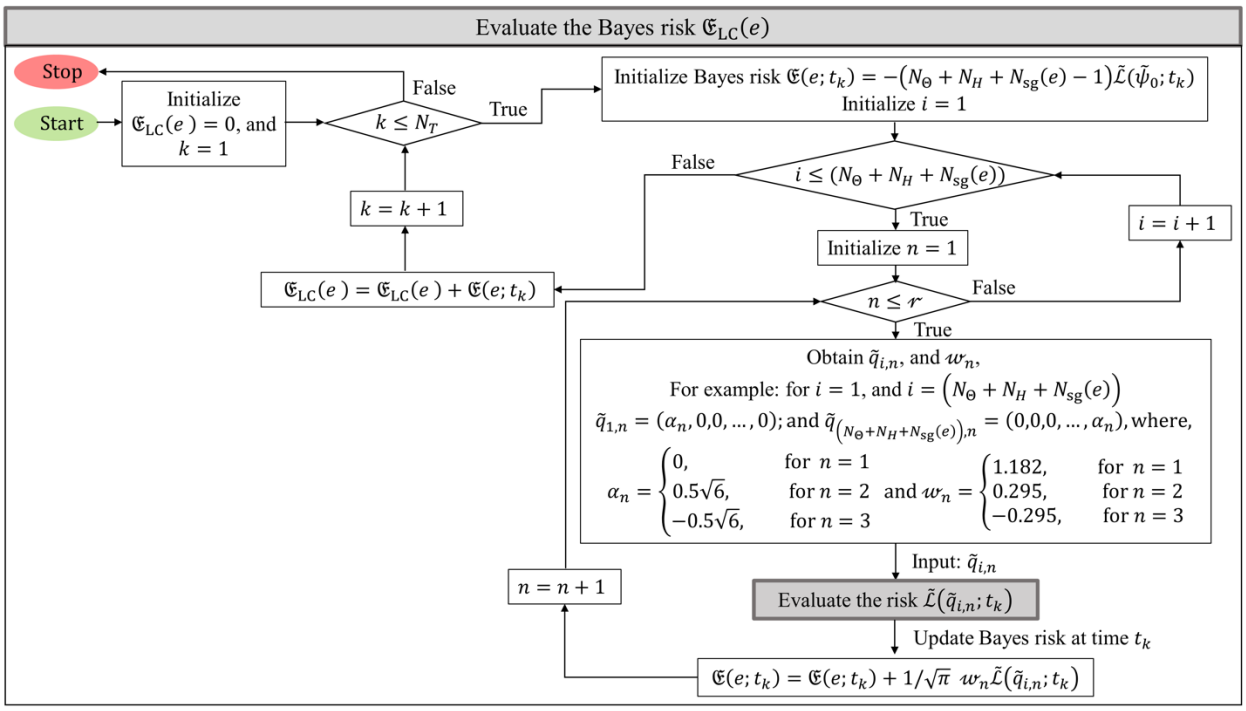


Figure 4.2: Algorithm to evaluate Bayes risk for design e over the lifecycle of structure.

4.3.8 A note on computational challenge and a novel solution

4.3.8.1 Computational challenge in evaluating Bayes risk

Using the algorithm described in Fig. 4.2 to evaluate $\mathfrak{E}(e; t)$ (and by extension $\mathfrak{E}_{\mathrm{LC}}(e)$) for a given design e requires immense computational power. I note some limitations of the algorithm described above. I observe that there are three sums in Eq. (4.31). This implies that the risk $\mathcal{L}(\tilde{q}_{i,n}; t_k)$ is to be evaluated for all $1 \le k \le N_T$, $1 \le i \le (N_\Theta + N_H + N_{\mathrm{sg}}(e))$, and $1 \le n \le \eta$ Recall from remark 1 that obtaining one instance of the risk requires running the FEM or the digital twin a total of $(N_{\text{Particles}} + 1)$ times. Therefore, evaluating $\mathfrak{E}_{\mathrm{LC}}(e)$ using Eq. (4.31) for a given design e requires running the FEM or the digital twin for a total of ($N_{\text{Runs-Per-Design}}$) times, such that:

$$N_{\text{Runs-Per-Design}} = (N_{\text{Particles}} + 1) \times N_T \times (N_\Theta + N_H + N_{\mathrm{sg}}(e)) \times r \tag{4.32}$$

Bayesian optimization aimed at obtaining the optimal sensor network design consists of evaluating many such designs, denoted by e_i, consisting of $N_{\mathrm{sg}}(e_i)$ number of sensors. To obtain an optimized sensor placement design, I start with an initially assumed design, denoted by e_0, that consists of $N_{\mathrm{sg}}(e_0)$ number of sensors. Starting with e_0, the subsequent sensor design e_i with $N_{\mathrm{sg}}(e_i)$ sensors is obtained by picking the most optimal sensor location from the available sensors and adding that sensor location to the previous design e_{i-1} with $N_{\mathrm{sg}}(e_{i-1}) = \left(N_{\mathrm{sg}}(e_i) - 1\right)$ sensors. Picking the additional sensor required to update the design e_{i-1} to the design e_i requires $N_{\text{iter}}\,(e_i)$ number of iterations. Since Bayes risk is the optimiality criteria, it needs to be evaluated at every iteration for the design e_i. Let e_I, with $i = I$, represent the final optimal sensor network design. The total number of GPR runs to arrive at e_I (starting from e_0) is denoted by $N_{\text{Total-Runs}}$, such that:

$$N_{\text{Total-Runs}} = \sum_{i=1}^{I} \left(N_{\text{particles}} + 1\right) \times N_T \times \left(N_\Theta + N_H + N_{\mathrm{sg}}(e)\right) \times r \times N_{\text{iter}}\,(e_i) \qquad (4.33)$$

4.3.8.2 Solution adopted for significantly reducing the number of digital-twin runs

I have adopted three major steps to reduce the computational cost at both, data simulation level, and at optimization algorithm level. These undertakings are outlined below:

Approximate FEM by a digital surrogate: Simulating the true sensor reading $N_{\text{Total-Runs}}$ times using a high-fidelity FEM is computationally intractable and I seek computationally efficient approximations to the FEM. For Bayesian calibration, metamodels or surrogate models are preferable, e.g., Support Vector Regression (SVR) (Moustapha and Bourinet 2018), Gaussian Process Regression (GPR) (Moustapha and Bourinet 2018, Frazier 2018), Neural Network (Yu

and Wang 2009), and Polynomial Chaos Expansion (PCE) (Capellari and Chatzi 2018). Some such approaches like SVR or neural networks yield point estimates/prediction, while others like GPR also predict the uncertainties associated with an average estimate/prediction. In this chapter, I have used the GPR model to approximate the FEM. It turns out to be faster than the FEM model by an order of 10^4 to 10^5 times

Further optimize digital surrogate: Let $N_{\text{Total-Sensors}}$ denote the total number of sensors available to choose from for a design e, such that $N_{\text{sg}}(e) \leq N_{\text{Total-Sensors}}$. Running a FEM for a given instance of state-parameter, and load will yield a total of $N_{\text{Total-Sensors}}$ dimensional output (true sensor readings). Usually, $N_{\text{Total-Sensors}}$ is a reasonably large number and the sensor readings are correlated to each other. I address this using principal components analysis, which can be efficiently computed using the singular value decomposition (SVD) that reduces the high-dimensional, highly correlated output space to low-dimensional, uncorrelated feature space. Let N_{Features} denote the total number of features obtained as a result of the SVD analysis. Instead of training digital twin to yield $N_{\text{Total-Sensors}}$ dimensional output, I train them to yield N_{Features} features as their output. I can then retrieve $N_{\text{Total-Sensors}}$ sensor readings from N_{Features} features by using inverse SVD. I build a digital surrogate for each feature. Figure 4.3 below illustrates schematic flowchart of developing and utilizing digital-twin to obtain sensor measurements.

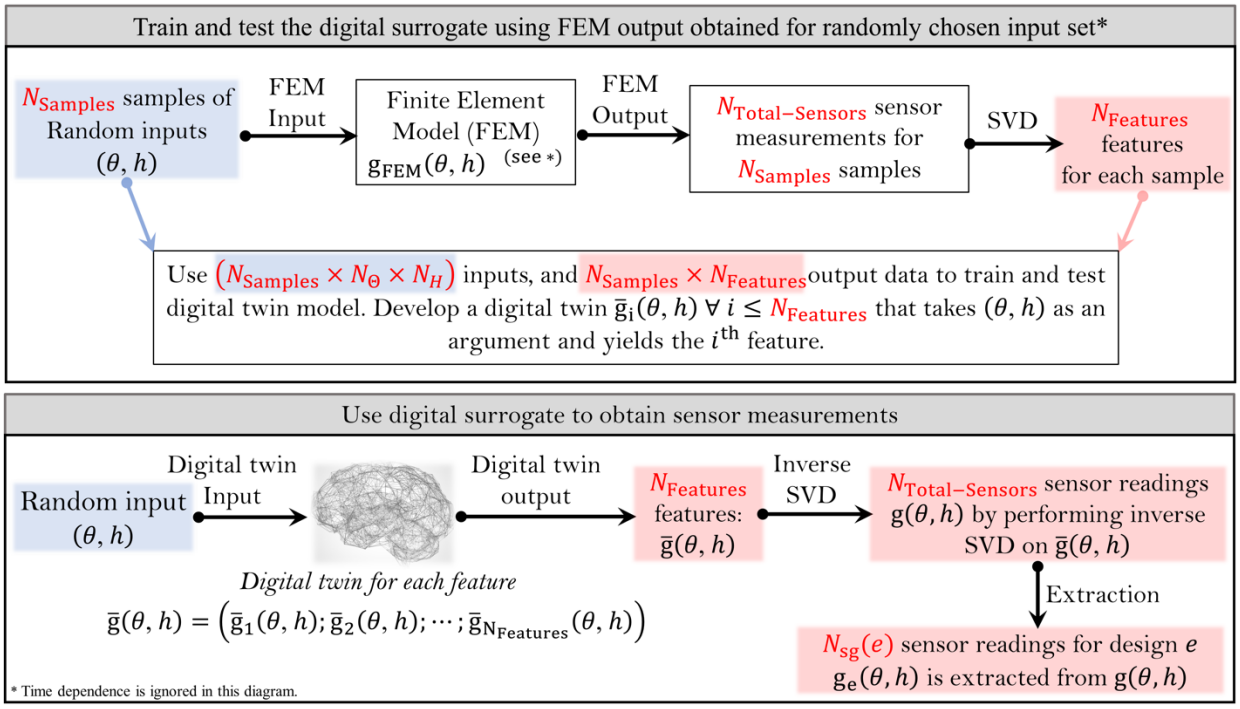


Figure 4.3: Schematic flowchart describing sensor data generation using a digital twin trained using data obtained through FEM.

An efficient algorithm for evaluating the Bayes risk: The computational cost can be further reduced by realizing that running the digital-twin $N_{\text{Total-Runs}}$ times includes numerous non-unique runs for many repetitive arguments (or input). Consider the following line of reasoning as an extension to the explanation provided in Yang et al. (Yang and Chadha 2021). At a given time instance t, the vector $\tilde{q}_{i,n}$ is a special case of the vector $\tilde{\psi}_i(t)$ as defined in Eq. (4.23). The first $(N_\Theta + N_H)$ components of the vector $\tilde{q}_{i,n}$ constitute a sub-vector $\tilde{q}_{i,n}(1{:}\,N_\Theta + N_H) = (\tilde{\theta}(t), \tilde{h}(t)) = \left(\tilde{\theta}_1(t), \tilde{\theta}_2(t), \cdots, \tilde{\theta}_{N_\Theta}(t), \tilde{h}_1(t), \tilde{h}_2(t), \cdots, \tilde{h}_{N_H}(t)\right)$ the inverse standard-normal transformation of which are the argument of the digital-twin model $g_e(\theta(t), h(t); t)$. The remaining $N_{\text{sg}}(e)$ components constitute a vector $\tilde{q}_{i,n}(N_\Theta + N_H{:}\,N_\Theta + N_H + N_{s\,g}(e))$ representing external noise. For any vector $\tilde{q}_{i,n}$, the sub-vector $\tilde{q}_{i,n}(1{:}\,N_\Theta + N_H)$ bears α_n as numerical value

of one of the components and zero for others. Therefore, I have a set of $(2 \times (N_\Theta + N_H) + 1)$ unique sub-vectors $(\tilde{\theta}(t), \tilde{h}(t))$ of interest, such that:

$$\tilde{q}_{i,n}(1 : N_\Theta + N_H) = \left(\tilde{\theta}(t), \tilde{h}(t)\right) \in \left\{ \begin{array}{l} \underbrace{(0, 0, \cdots, 0;}_{\text{Length } N_\Theta} \underbrace{0, 0, \cdots, 0)}_{\text{Length } N_H}; \Big\} \text{the zero vector} \\ \left.\begin{array}{l} (0.5\sqrt{6}, 0, \cdots, 0; 0, 0, \cdots, 0); \\ (0, 0.5\sqrt{6}, \cdots, 0; 0, 0, \cdots, 0); \\ \vdots \\ (0, 0, \cdots, 0.5\sqrt{6}; 0, 0, \cdots, 0); \\ (0, 0, \cdots, 0; 0.5\sqrt{6}, 0, \cdots, 0); \\ (0, 0, \cdots, 0; 0, 0.5\sqrt{6}, \cdots, 0); \\ \vdots \\ (0, 0, \cdots, 0; 0, 0, \cdots, 0.5\sqrt{6}); \end{array}\right\} (N_H + N_\Theta) \text{ number of vectors} \\ \left.\begin{array}{l} (-0.5\sqrt{6}, 0, \cdots, 0; 0, 0, \cdots, 0); \\ (0, -0.5\sqrt{6}, \cdots, 0; 0, 0, \cdots, 0); \\ \vdots \\ (0, 0, \cdots, -0.5\sqrt{6}; 0, 0, \cdots, 0); \\ (0, 0, \cdots, 0; -0.5\sqrt{6}, 0, \cdots, 0); \\ (0, 0, \cdots, 0; 0, -0.5\sqrt{6}, \cdots, 0); \\ \vdots \\ (0, 0, \cdots, 0; 0, 0, \cdots, -0.5\sqrt{6}); \end{array}\right\} (N_H + N_\Theta) \text{ number of vectors} \end{array} \right. \quad (4.34)$$

Since obtaining the utility $\tilde{\mathcal{L}}(\tilde{q}_{i,n}; t)$ for each $\tilde{q}_{i,n}$ at a fixed time instance requires $(N_{\text{Particles}} + 1)$ digital-twin runs, and there are N_T discrete time periods, considering all the unique arguments of digital-twin model, I essentially need to run GPR model only $N_{\text{Reduced-Total-Runs}}$ times, such that:

$$N_{\text{Reduced-Total-Runs}} = \left(1 + 2 \times (N_\Theta + N_H)\right) \times \left(N_{\text{particles}} + 1\right) \times N_T \quad (4.35)$$

The sensor readings for all $N_{\text{Total-Sensors}}$ sensors are obtained for these $N_{\text{Reduced-Total-Runs}}$ runs and is stored in a "true-sensor-data" matrix. The sensor measurements for design e can be extracted

from "true-sensor-data" matrix. The risk $\tilde{\mathcal{L}}(\tilde{q}_{i,n}; t)$ can be evaluated by extracting the relevant sensor readings from "true-sensor-data" matrix as illustrated in Fig. 4.4 An efficient algorithm to evaluate the Bayes risk $\mathfrak{E}_{\mathrm{LC}}(e)$ that requires only $N_{\text{Reduced-Total-Runs}}$ digital-twin runs is illustrated in Fig. 4.5.

In summary, the computational speed is further enhanced by setting me the Bayesian optimization code in a way that once the desirable sensor measurements are obtained in one iteration of the optimization, the predictive model need not be run in order to evaluate the observed sensor measurements at every optimization iteration, significantly improving the computational efficiency of the sensor placement design optimization and allows for a reduction of computational time from years to hours.

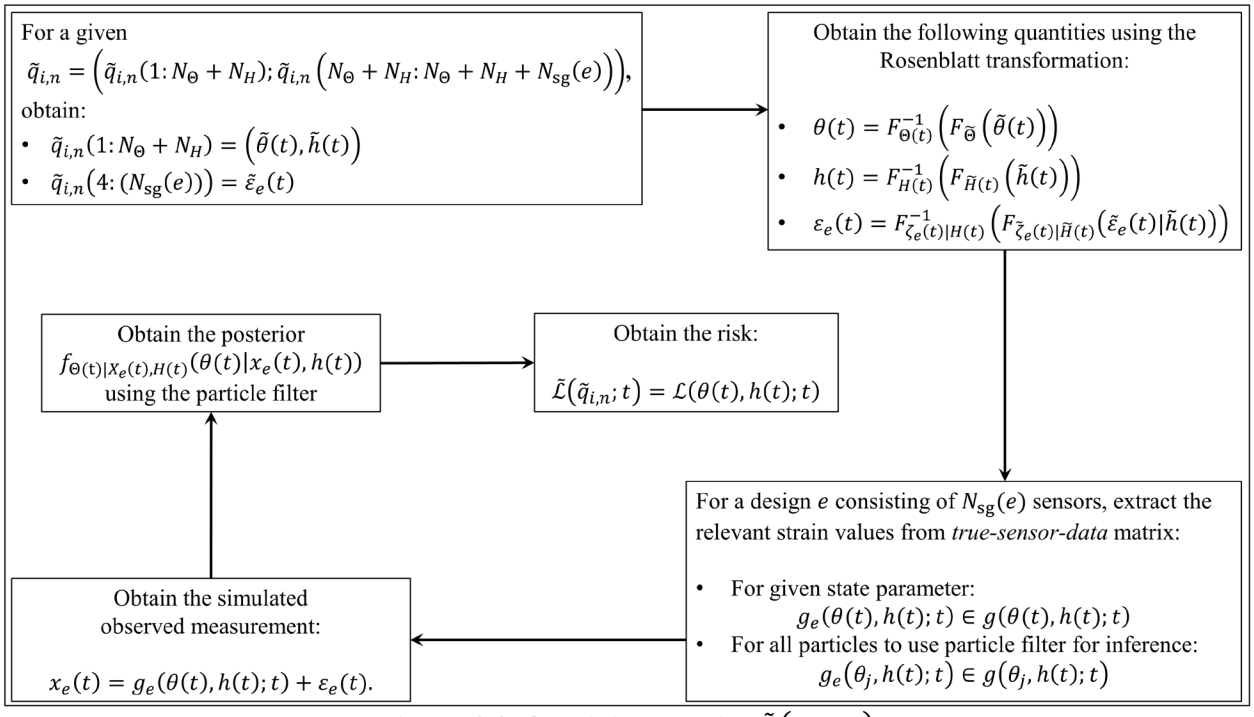


Figure 4.4: Obtaining the risk $\tilde{\mathcal{L}}(\tilde{q}_{i,n}; t)$

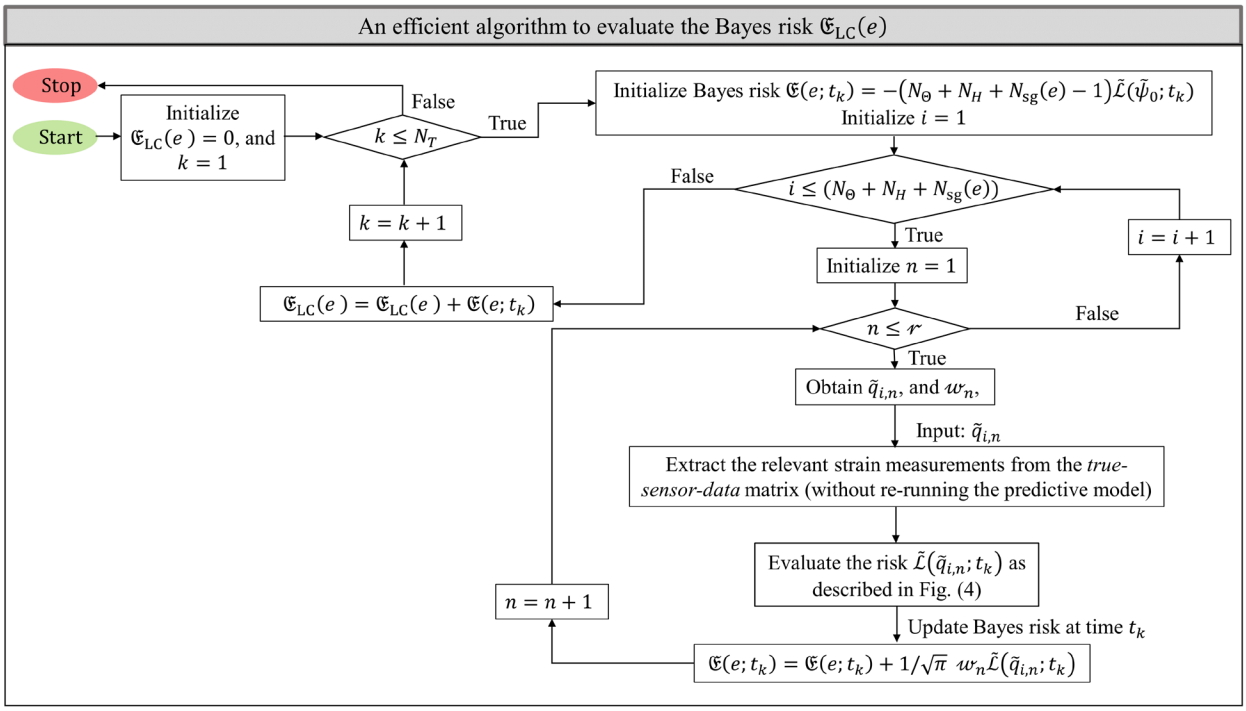


Figure 4.5: An efficient algorithm to evaluate the Bayes risk for design e over the lifecycle of structure.

4.3.9 Optimal sensor design in a nutshell

The optimal design $e^* \in \Omega_E$ is the design that minimizes the expected Bayes risk. In my case, it is the design that leads to the most reliable observed sensor readings that have a minimum *expected* deviation in the inferred state (quantified by mean of the posterior distribution of the state parameter conditioned on the acquired information) relative to the true structural state considering the design's *overall* performance throughout the lifecycle of the structure. Mathematically,

$$e^* = \underset{e \in \Omega_E}{\arg\min}\ \mathfrak{E}_{\mathrm{LC}}(e) \tag{4.36}$$

I realize that in a real complex structure, the total number of sensor locations of interest, denoted by $N_{\text{Total-Sensors}}$, , is usually a large number spanning anywhere from $10^3 - 10^6$ or so. An increase in $N_{\text{Total-Sensors}}$ increases the size of the exhaustive sensor design space Ω_E exponentially. The size

of Ω_E can be obtained as the number of all the possible combinations of $N_{\text{Total-Sensors}}$ sensor locations, such that:

$$\text{size}(\Omega_E) = \sum_{r=1}^{r \leq N_{\text{Total-Sensors}}} \frac{N_{\text{Total-Sensors}}\,!}{r!\,(N_{\text{Total-Sensors}} - r)!} = 2^{N_{\text{Tocal-Sensors}}} - 1 \tag{4.37}$$

In theory, the design $e^* \in \Omega_E$ can be picked simply by evaluating the Bayes risk for all the possible designs. However, evaluating the objective function for all possible designs is not only difficult but computationally impossible even for a lower value of $N_{\text{Total-Sensors.}}$. Given the intractable nature of the exhaustive search, I use an iterative Bayesian optimization to look for a global optimum in a minimum number of steps, thus minimizing the sampling points to rapidly speed up the optimization process. The algorithmic details of carrying Bayesian optimization detailed exhaustively in Yang et al. (Yang and Chadha 2021) stays the same. To avoid repetition, the readers are referred to algorithm 1 in Section 4.6 of Yang et al. (Yang and Chadha 2021) for algorithmic details of carrying out Bayesian optimization.

The discussion so far completes the description of the generic Bayesian optimization framework considering the overall reliability of a sensor design over the lifecycle of a structure. In the following section, I focus on applying the sensor optimization framework to a miter gate structure. I expect this demonstration example to serve as an excellent case study to get acquainted with applications of these ideas to a practical problem of interest to engineers.

4.4 A case study: a miter gate

4.4.1 Why miter gate?

I consider the same demonstration example as was used in chapter 3 for the following reasons:

1 The state of the miter gate can be modeled by a scalar one-dimensional continuous state parameter. Therefore, the miter gate is a complex and yet a suitable case study that is appropriate for visualizing, interpreting, evaluating, and comparing the optimization results.

2 In terms of the reliability aspect, at any point in time, the miter gate can be conveniently divided into two zones. The first zone is the part of the gate above the water level (called splash zone) and the second zone is the part of the gate that is submerged in the water (called the submerged zone). The sensors that are in the submerged zone for a majority of the time have a higher probability of malfunctioning than the sensors that spends higher average time in the splash zone. Therefore, the miter gate case study captures the generic case of different sensor reliability zones in a structure, and hence, it is an excellent example that can be used to unfold and comprehensively understand the full potential of the proposed optimization algorithm that accounts for sensor reliability.

3 Adopting the same demonstration examples as used in Yang et al. (Yang and Chadha 2021) allows for a direct comparison in optimal sensor design obtained for different target objective functional. As described in the upcoming Section 4.5, the numerical results demonstrate that for a more realistic sensor network design, not only maximizing the gain in information (the target objective in (Yang and Chadha 2021)) is crucial but considering sensor reliability in sensor network design (the target objective in this chapter) is even more important.

4 Finally, I had used a high-fidelity finite element model in Yang et al. (Yang and Chadha 2021) to simulate true sensor measurements of strain gauges installed in the Greenup miter

gate. Building a highfidelity FEM usually takes a tremendous amount of tedious modeling work that is not the focus of this chapter. Adopting the same demonstration example allows me to conveniently use the same FEM that was used in (Yang and Chadha 2021) for an application case of the proposed framework.

4.4.2 Description of the miter gate structure

In the last chapter, I have briefly described the miter gate structure and the damage parameter selected.

4.4.2.1 Design space:

As discussed in chapter 3, in total, there are 64919×4 sensor locations of choice. However, for practical reasons, the USACE installs the sensors only on the downstream side of the miter gate. This reduces my available sensor locations by approximately half, that is, $N_{\text{Total-Sensors}} \approx 0.5 \times 64919 \times 4$

4.4.2.2 The load vector:

The gate is subjected to uncertain upstream and downstream hydrostatic loads quantified by the hydrostatic upstream and downstream heads, such that the load vector $h(t)$ consists of two components, that is $h(t) = \left(h_{\text{up}}(t), h_{\text{down}}(t)\right)$ and $\Omega_{H(t)} = \Omega_{H_{\text{m}}(t)} \times \Omega_{H_{\text{domen}}(t)}$. To have less discrepancy to the real model. In this chapter, instead of normal distribution, the water heads are modeled by employing a three-parameter-based Weibull distribution. It is the distribution of choice because it provides flexibility to model a complex variable like water heads and has a positive support (i.e., it naturally constraints water heads to be non-negative). I have:

$$f_{H_{\text{up}}(t)}(h_{\text{up}}(t)) = c_1\left(\frac{h_{\text{up}}(t) - c_{\text{up}}}{c_2}\right)^{(c_1-1)} e^{-\left(\frac{h_{\text{up}}(t)-c_{\text{up}}}{c_2}\right)^{c_1}};$$
$$f_{H_{\text{down}}(t)}(h_{\text{down}}(t)) = c_1\left(\frac{h_{\text{down}}(t) - c_{\text{down}}}{c_2}\right)^{(c_1-1)} e^{-\left(\frac{h_{\text{down}}(t)-c_{\text{down}}}{c_2}\right)^{c_1}} \quad (4.39)$$

In the expression above, the parameters $(c_1 = 2, c_2 = 50, c_{\text{up}} = 516$ inches $)$ models the distribution of the upstream water load, and $(c_1 = 2, c_2 = 50, c_{\text{down}} = 132$ inches $)$ models the distribution of the downstream water load. These parameters were selected based on consultation with USACE research team. Figure 4.7 illustrates the two distributions.

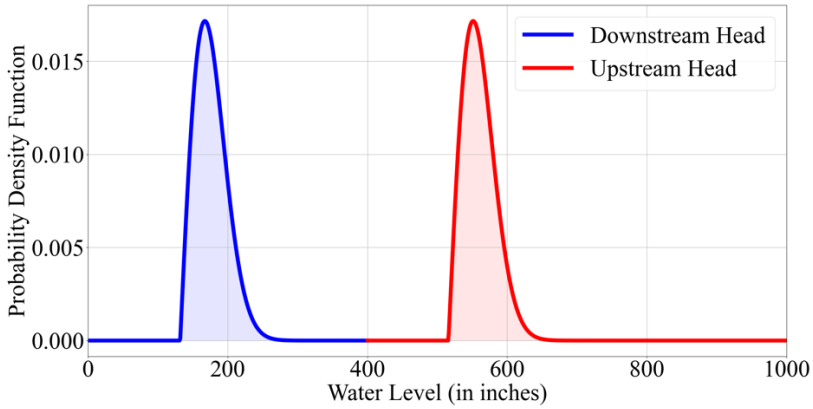


Figure 4.6: Assumed probability distributions of the upstream and the downstream water heads.

The next heading discusses various zones of reliability in the miter gate structure.

4.4.2.3 Various zones of reliability, and modeling observation noise of a sensor

To model the noise in the i-th strain gauge constituting the design e at time t as per Eq. (4.6), I assume that $\sigma_{\varepsilon_{ei}}(t) = 5 \times 10^{-6}$. This assumption is in accordance with reasonable commercial strain gauge performance. Secondly, the sensor reliability bias factor $\eta_{ei}(t)$ is modeled by a Gaussian distribution with the mean $\mu_{n_{ei}(t)}$, and the standard deviation $\sigma_{n_{ei}}(t)$. I consider two cases of reliability bias factor in the numerical simulation discussed in Section 4.5.

The miter gate is divided into two reliability-zones decided by the downstream water-head H_{down} (since strain gauges are installed on the downstream side of the gate only). The first zone is part of the gate above the downstream water level (called splash zone and is denoted by z_{splash}) and the second zone is part of the gate that is submerged in water (called the submerged zone and is denoted by z_{sub}). Let the set of zones be defined as $\Omega_Z = \left(z_{\text{splash}}, z_{\text{sub}}\right)$. Sensors that are in the submerged zone for a majority of the time have a higher probability of malfunctioning than the sensors that are installed in the splash zone. Let $Z_{ei}(t)$ denotes the random variable modeling the functional state of the i-th sensor in design e at time t. The probability mass function $P_{Z_{ei}(t)|H_{\text{down}}(t)}\left(z_{\text{splash}} \mid h_{\text{down}}(t)\right)$ and $P_{Z_{ei}(t)|H_{\text{down}}(t)}(z_{\text{sub}} \mid h_{\text{down}}(t)) = 1 - P_{Z_{ei}(t)|H_{\text{down}}(t)}\left(z_{\text{splash}} \mid h_{\text{down}}(t)\right)$, denotes the probabilities of i-th sensor of design e exposed to the *splash zone* or *submerged* zone at time t conditioned upon the downstream water head respectively. I define the conditional probability of the i-th sensor reading constituting the design e being unreliable or reliable conditioned on their location in either the splash zone or the submerged zone at any time t as:

$$\begin{aligned} P_{S_{ei}(t)|Z_{ei}(t)}\left(S_{\text{unrel}} \mid z_{\text{splash}}\right) &= 1 - e^{-\lambda_{\text{splash}}\, t} \\ P_{S_{ei}(t)|Z_{ei}(t)}\left(s_{\text{unrel}} \mid z_{\text{sub}}\right) &= 1 - e^{-\lambda_{\text{sub}}\, t} \\ P_{S_{ei}(t)|Z_{ei}(t)}\left(s_{\text{rel}} \mid z_{\text{splash}}\right) &= e^{-\lambda_{\text{splash}}\, t} \\ P_{S_{ei}(t)|Z_{ei}(t)}\left(s_{\text{rel}} \mid z_{\text{sub}}\right) &= e^{-\lambda_{\text{sub}}\, t} \end{aligned} \tag{4.40}$$

I note two properties of the probabilities assigned above: (1) it is widely accepted that the submerged sensors fail more often than the sensors in the splash zone. I assume $\lambda_{\text{sub}} > \lambda_{\text{splash}}$. As a consequence of this, the probability of a sensor malfunctioning as defined in Eq. (4.40) is higher in the submerged zone than in the splash zone. For simulation, I assume that $\lambda_{\text{sub}} = 0.2$ and $\lambda_{\text{splash}} = 0.002$; (2) Given that a sensor's zone is unchanged, the probability of a sensor

malfunctioning increases with time. Finally, the probability of the i-th sensor constituting the design e being unreliable or reliable at any time t is obtained using total probability theorem as:

$$P_{S_{ei}(t)|H_{\text{down}}(t)}(s_{\text{unrel}} \mid h_{\text{down}}(t)) = P_{S_{ei}(t)|Z_{ei}(t)}(s_{\text{unrel}} \mid z_{\text{splash}})P_{Z_{ei}(t)|H_{\text{down}}(t)}(z_{\text{splash}} \mid h_{\text{down}}(t)) + P_{S_{ei}(t)|Z_{ei}(t)}(s_{\text{unrel}} \mid z_{\text{sub}})P_{Z_{ei}(t)|H_{\text{down}}(t)}(z_{\text{sub}} \mid h_{\text{down}}(t))$$

$$P_{S_{ei}(t)|H_{\text{down}}(t)}(s_{\text{rel}} \mid h_{\text{down}}(t)) = \left(1 - P_{S_{ei}(t)|H_{\text{down}}(t)}(s_{\text{unrel}} \mid h_{\text{down}}(t))\right)$$

(4.41)

Here, $P_{Z_{ei}(t)|H_{\text{down}}(t)}(z_{\text{splash}} \mid h_{\text{down}}(t))$ and $P_{Z_{ei}(t)|H_{\text{down}}(t)}(z_{\text{sub}} \mid h_{\text{down}}(t))$ denotes the probability of the i-th sensor in design e being in the splash zone or the submerged zone at time t. Since the sensors are attached to the downstream side, given a realization of downstream water head at time t, I have the following:

The i-th sensor in design e is in the spash zone, or above $h_{\text{down}}(t)$: $\begin{Bmatrix} P_{Z_{ei}(t)|H_{\text{down}}(t)}(z_{\text{splash}} \mid h_{\text{down}}(t)) = 1 \\ P_{Z_{ei}(t)|H_{\text{down}}(t)}(z_{\text{sub}} \mid h_{\text{down}}(t)) = 0 \end{Bmatrix}$

The i-th sensor in design e is in the submerged zone, or below $h_{\text{down}}(t)$: $\begin{Bmatrix} P_{Z_{ei}(t)|H_{\text{down}}(t)}(z_{\text{splash}} \mid h_{\text{down}}(t)) = 0 \\ P_{Z_{ei}(t)|H_{\text{down}}(t)}(z_{\text{sub}} \mid h_{\text{down}}(t)) = 1 \end{Bmatrix}$

(4.42)

4.4.2.4 The digital twin:

I have all the information required to simulate observed strain gauge readings by running the FEM for a given input $(\theta(t), h(t), \varepsilon_e(t))$ as per Eq. (4.7). To simulate the true strain gauge data, I rely on a Gaussian Process Regression (GPR) based model trained using simulated true strain values obtained from the validated FEM as discussed in Section 4.3.8.2. The SVD analysis reduces the output space of FEM from $N_{\text{Total-Sensors}} = 0.5 \times 64919 \times 4$ sensors, to $N_{\text{Features}} = 38$

dimensional space. Therefore, for a given input $\big(\theta(t), h(t), \varepsilon_\rho(t)\big)$, the GPR model yields 38 features which could be used to retrieve $N_{\text{Total-Sensors}}$ strain readings through inverse SVD analysis.

4.4.2.5 The prior gap-degradation model

Since the time evolution of gap is not precisely known, I model it probabilistically (as shown in Fig. 4.8), such that $f_{\Theta(t)}(\theta(t))$ denotes the prior distribution of gap length at time t. The gap evolution over time is described by a piecewise multi-stage degradation model as follows:

$$\theta(t_{k+1}) = \theta(t_k) + \theta(t_k)^{w(t_{k+1})} \cdot Q(t_{k+1}) \cdot \exp\big(\sigma(t_{k+1}) \cdot U(t_{k+1})\big) \tag{4.43}$$

In the equation above, $\theta(t_{k+1})$ denotes the gap length at time step t_k; N_t is the total number of time steps; $U(t_{k+1})$ is a stationary Gaussian stochastic process; $\sigma_{t_{k+1}}$, $Q(t_{k+1})$, and $w(t_{k+1})$ are degradation state-dependent model parameters, which are given as follows

$$\begin{aligned} \sigma(t_{k+1}) &= \sigma_j \\ Q(t_{k+1}) &= Q_j \\ w(t_{k+1}) &= w_j \end{aligned} \tag{4.44}$$

where, the index j represents the degradation state, such that

$$j = \hbar_s\big(\theta(t_k)\big) \tag{4.45}$$

The function $\hbar_s(\cdot)$ maps the gap length to the degradation state j, such that

$$j = \hbar_s(\theta(t)) = \begin{cases} 1, & \text{if } \theta(t) \in [0, \varphi_1] \\ 2, & \text{if } \theta(t) \in [\varphi_1, \varphi_2] \\ & \vdots \\ N_d, & \text{if } \theta(t) \in \big[\varphi_{N_d-1}, \infty\big] \end{cases} \tag{4.46}$$

where, φ_i for $i \in \{1,2,\cdots,(N_d-1)\}$ are the switching points that govern the transition between different degradation stages and N_d is the number of degradation stages. I assumed $N_d = 3$ for the

current study. Since the switching points φ_i in Eq. (4.46) are uncertain in nature, they are modeled by the Gaussian distribution as shown below

$$\varphi_i \sim N\left(\mu_{\varphi_i}, \sigma^2_{\varphi_i}\right) \forall\, i \in \{1,2,\cdots, N_d - 1\} \tag{4.47}$$

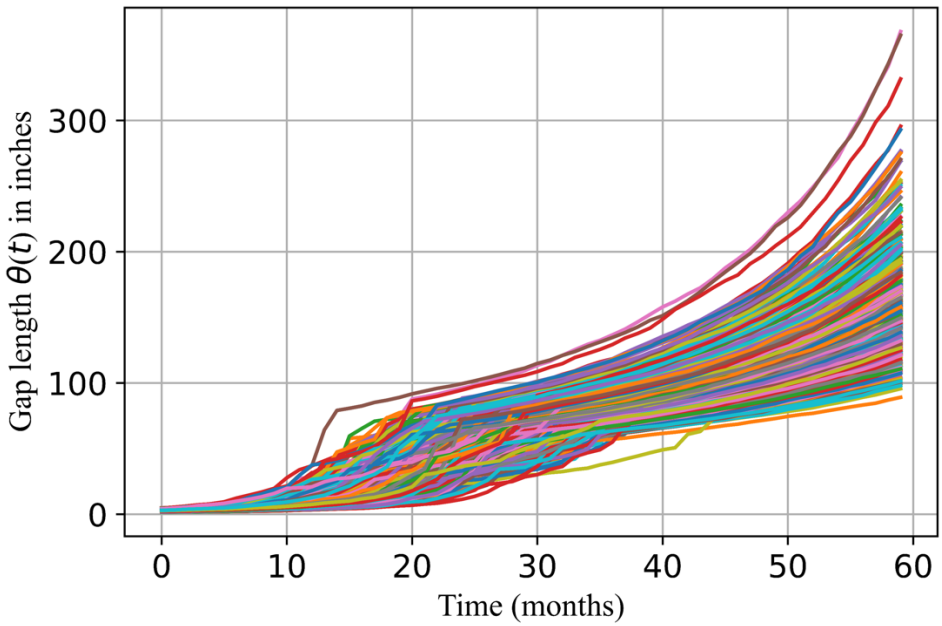


Figure 4.7: The time-evolution of gap length

For a given design e, the posterior gap-degradation model over the lifecycle of the structure can be obtained by using Bayesian inference as described in Section 4.3.5. In words, the sensor reliability risk $\mathcal{L}(\theta_{\text{true}}, h(t), \varepsilon_e(t))$ is defined as absolute deviation of mean of the posterior distribution of the gap length (conditioned upon observed strain readings) from the true value of the gap length at time t. Mathematically, the definition of the sensor reliability risk in Eq. (4.12), and the Bayes risk in Eq. (4.14) stays intact. I use the framework described in Section 4.3.7 in tandem with Section 4.3.8 to obtain the value of the Bayes risk $\mathfrak{E}_{\text{LC}}(e)$. Finally, optimal sensor design is obtained using Bayesian optimization as discussed in Section 4.3.9. In the next section, I discuss numerical examples of the proposed framework applied to the miter gate.

4.5 Numerical results and discussion

4.5.1 Various optimal designs for comparison

For the purposes of investigation, I consider three optimal designs:

1 Design e_{KL}: Acquire strain measurements using Bayesian optimized strain-gauge network containing $N_{\mathrm{sg}}(e_{\mathrm{KL}}) = 10$ sensors as detailed in Yang et al. (Yang and Chadha 2021) (using KL divergence without risk weights as an objective functional).

2 Design e_{R_1} : Acquire strain measurements using Bayesian optimized strain-gauge network containing $N_{\mathrm{sg}}\left(e_{\mathrm{R}_1}\right) = 10$ sensors obtained using reliability-based Bayes risk defined in Eq. (4.14) with the following reliability bias factor statistics: $\mu_{n_{ei}} = 10^{-4}$ and $\sigma_{n_{ei}} = 2 \times 10^{-6}$.

3 Design e_{R_2} : Acquire strain measurements using Bayesian optimized strain-gauge network containing $N_{\mathrm{sg}}\left(e_{\mathrm{R}_2}\right) = 10$ sensors obtained using reliability-based Bayes risk defined in Eq.(4.14) with the following reliability bias factor statistics: $\mu_{n_{ei}} = 5 \times 10^{-5}$ and $\sigma_{n_{ei}} = 5 \times 10^{-6}$.

4.5.1.1 The KL divergence optimized design $\boldsymbol{e_{\mathrm{KL}}}$

To illustrate the importance of accounting for the reliability of sensors in the optimal design framework, I first obtain the design $\boldsymbol{e_{\mathrm{KL}}}$. Figures 4.9 and 4.10 illustrate the front and side rendering of the design $\boldsymbol{e_{\mathrm{KL}}}$ respectively.

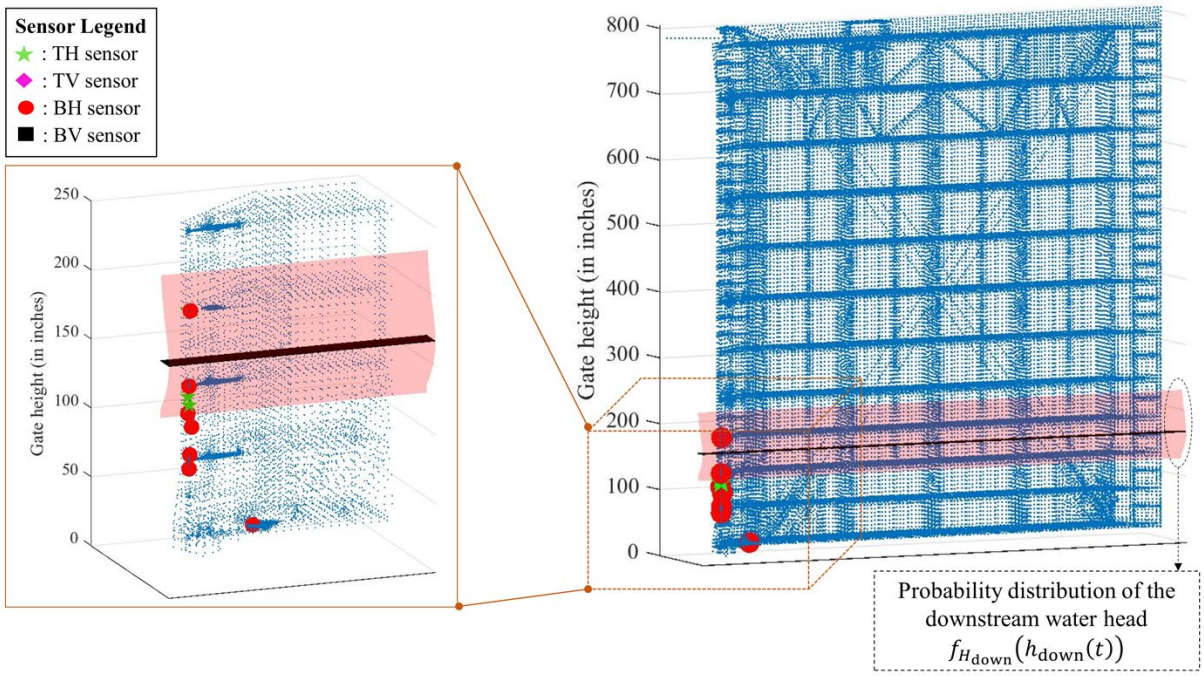


Figure 4.8: Design e_{KL}: Rendered front view of the gate

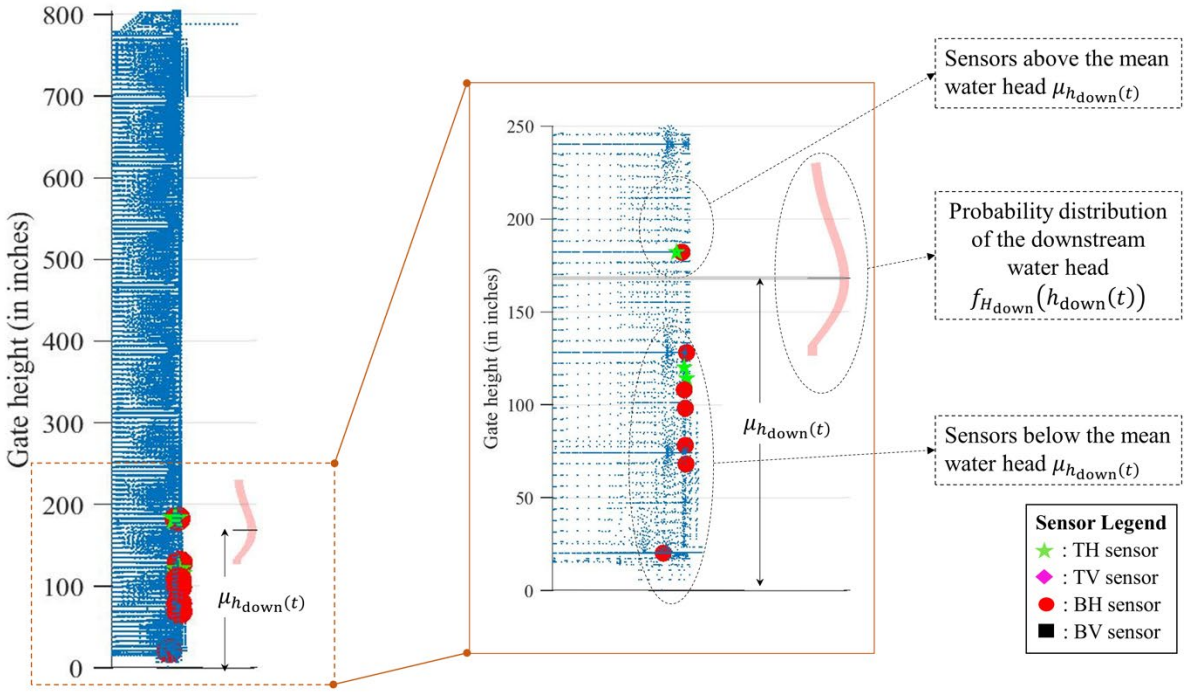


Figure 4.9: Design e_{KL}: Rendered side view of the gate

The design e_{KL} targets maximizing the information gain and does not account for sensor reliability. Naturally, I expect the following:

1 I expect that the sensors constituting the design e_{KL} will be closer to the gap because the information about the gap can be best captured when sensors are in its vicinity. It can be seen from Fig. 4.10 that 8 out of 10 sensors are in the submerged zone.

2 Since the design e_{KL} does not account for the sensor reliability (sensors have a higher likelihood of malfunctioning when submerged), I expect that the inference of gap length obtained using the strain gauge data (constituting the design e_{KL}) will be unreliable and skewed from reality when considering reliability bias in the strain gauge measurements along with the usual observation noise. The following sections will illustrate that this is indeed the case by comparing the inference results obtained from the design e_{KL} relative to the reliability-focused designs e_{R_1} and e_{R_2} respectively.

4.5.1.2 The optimal design $e_{\mathbf{R_1}}$

Bayesian optimization process requires obtaining the Bayes risk numerous times. To evaluate the sum in time domain in Bayes risk defined in Eq. (4.14), I divide a 60 months time interval ($\Omega_T = [0,60]$ months) into $N_T = 13$ discrete time steps. Figures 4.11 and 4.12 illustrates the front and side rendering of the design e_{R_1} respectively.

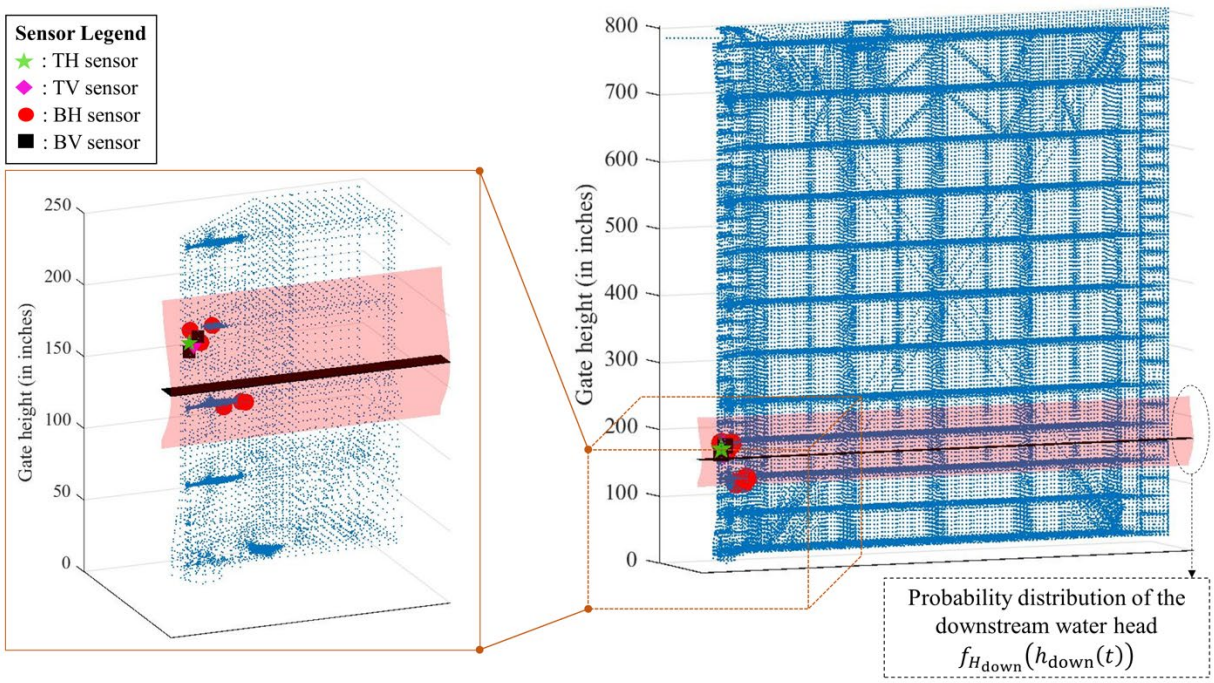


Figure 4.10: Design e_{R_1}: Rendered front view of the gate

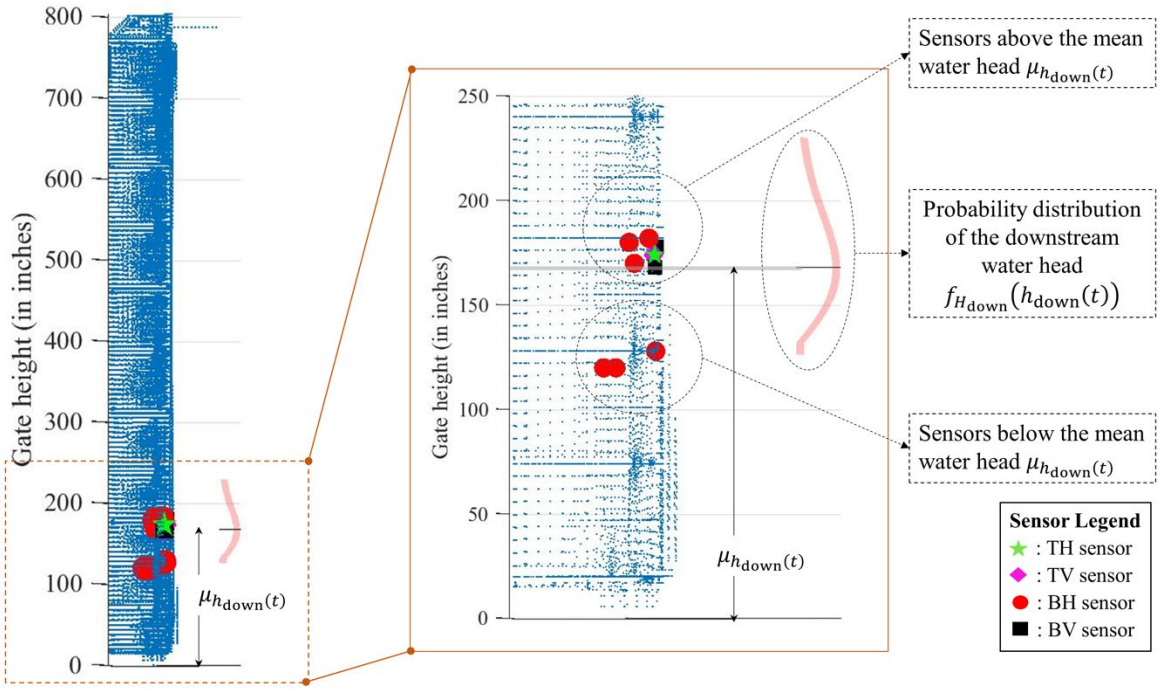


Figure 4.11: Design e_{R_1}: Rendered side view of the gate

4.5.1.3 The optimal design e_{R_2}

Similar to the design e_{R_1}, I use $N_T = 13$ discrete time steps to obtain the optimal design e_{R_2}. Figures 4.13 and 4.14 illustrates the front and side rendering of the design e_{R_2} respectively.

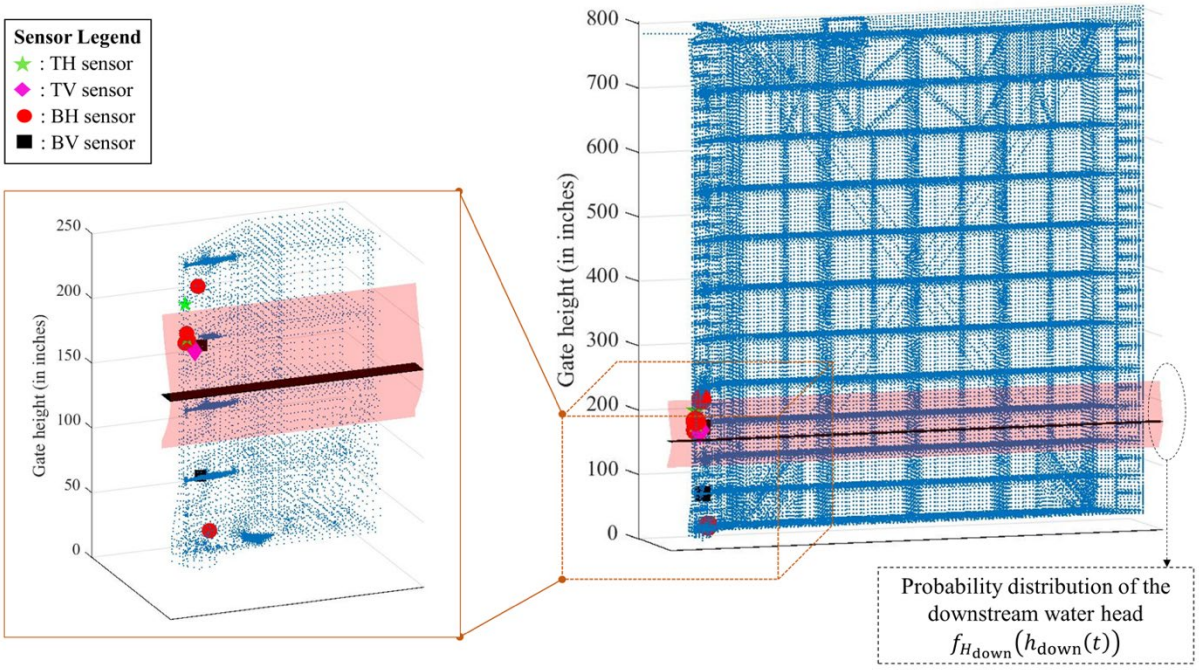


Figure 4.12: Design e_{R_2} : Rendered front view of the gate

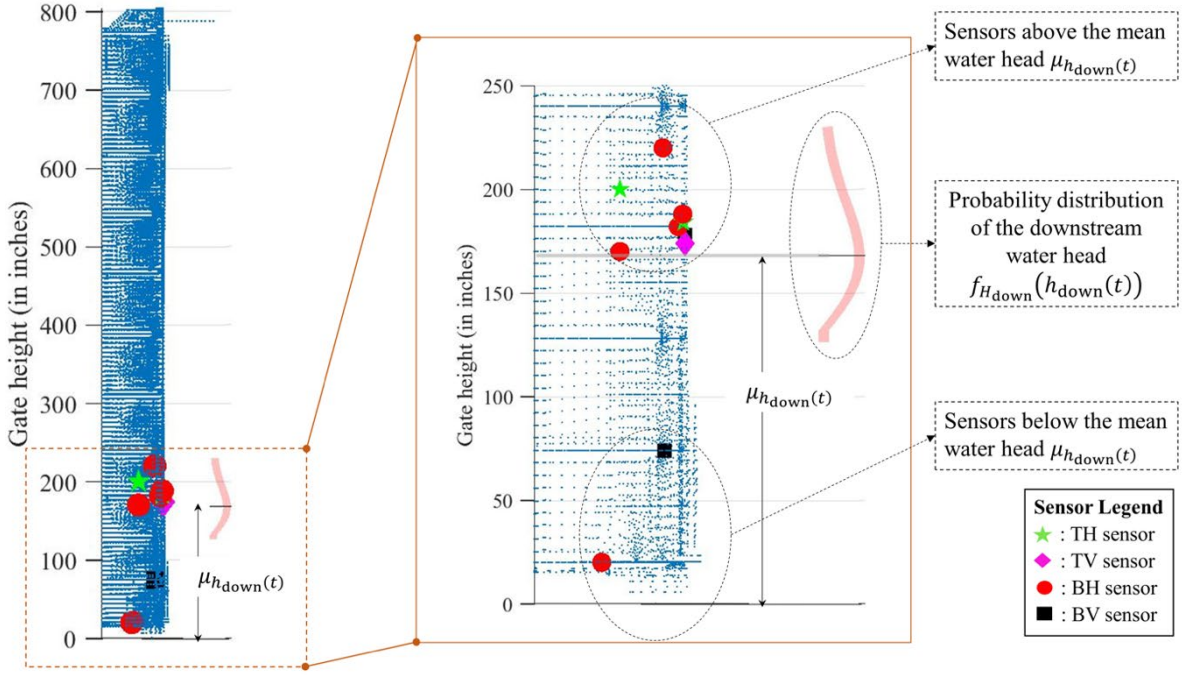


Figure 4.13: Design e_{R_2} : Rendered side view of the gate

One noteworthy observation is that unlike in the design e_{KI}, in both the reliability-focused designs e_{R_1} and e_R, there are more sensors above the mean downstream water head than the number of sensors below it. As discussed in Section 4.4.2.4, the probability of sensors malfunctioning is higher when they are located below the mean downstream water head (higher likelihood of being in the submerged zone) than when they are installed above the mean water head (higher likelihood of being in the splash zone). The design e_{KL} does not take this into account. The designs e_{R_1} and e_{R_2} are not only emboldened with strength of the design e_{KL} (carrying excellent inference when the sensors are fully reliable and functional), it also overcomes the limitations of the design e_{KL} (accounting for the sensor reliability). In the designs e_{R_1} and e_{R_2}, the sensors are strategically placed in the gap's neighborhood allowing for a realistic inference of the gap length (just like the design e_{KL}). At the same time, collectively, sensors spend higher average time in the splash zone over the lifespan of the structure, such that if submerged sensors malfunction, the sensors in the splash zone can carry the burden of performing acceptable inference over the lifecycle of the miter gate (unlike the design e_{KL}). This construct invariably makes the newly proposed sensor optimization framework a superior choice for state inference. In the next section, I compare the inference results for all three designs.

4.5.2 Comparison of various designs

To compare the reliability-focused optimal designs e_{R_1} and e_{R}, with the design e_{KL} that does not account for sensor reliability, I consider the following:

1. There must be a unique true/real gap evolution over time. However, while in the design phase, the uncertain nature of the problem does not permit knowing deterministically how the gap would evolve with time. Therefore, I must consider different realizations of true

gap evolution over time (each $\theta(t)$ curve in Fig. 4.8 is a realization of a possible true gap evolution $\theta_{\text{true}}(t)$) and evaluate the robustness and performance of the proposed optimization framework for various possible realities of system evolution. For illustration purposes, I show inference results for two such realizations of $\theta_{\text{true}}(t)$.

2 I consider $N_T = 13$ discrete time instances and consider 5 realizations of load vector satisfying Eq. (4.7) for each of these discrete time steps. For all the input variables $\left(\theta_{\text{true}}(t_i), h_{\text{up}}(t_i), h_{\text{down}}(t_i)\right)$, where $1 \leq i \leq N_T$, I obtain the inferred probability distribution of the gap-length conditioned upon strain data obtained for the designs $e_{\text{KL}}, e_{\text{R}_1}$, and e_{R_2}. For a given design and for each time step, I obtain 5 posterior distributions of gap-length obtained for 5 instances of water heads. I use the average value of the mean and standard deviation of these 5 distributions to obtain piecewise linearly interpolated posterior gap-degradation curves. These inferred gap evolution curves are used to compare various designs.

The Fig. 4.15 and 4.16 shows piecewise linearly interpolated posterior gap-degradation curves for the first and the second true gap length evolution curve for designs e_{KL} and e_{R_1}, respectively. Similarly, Fig. 4.17 and 4.18 illustrates the posterior gap-degradation curves for the first and the second true gap length evolution curve for designs e_{KL} and e_{R_2}, respectively. All four figures also zoom in on two instances of true gap length (or the respective time) and illustrate 5 cases of inferred posterior distributions of the gap length corresponding to 5 realizations of loading.

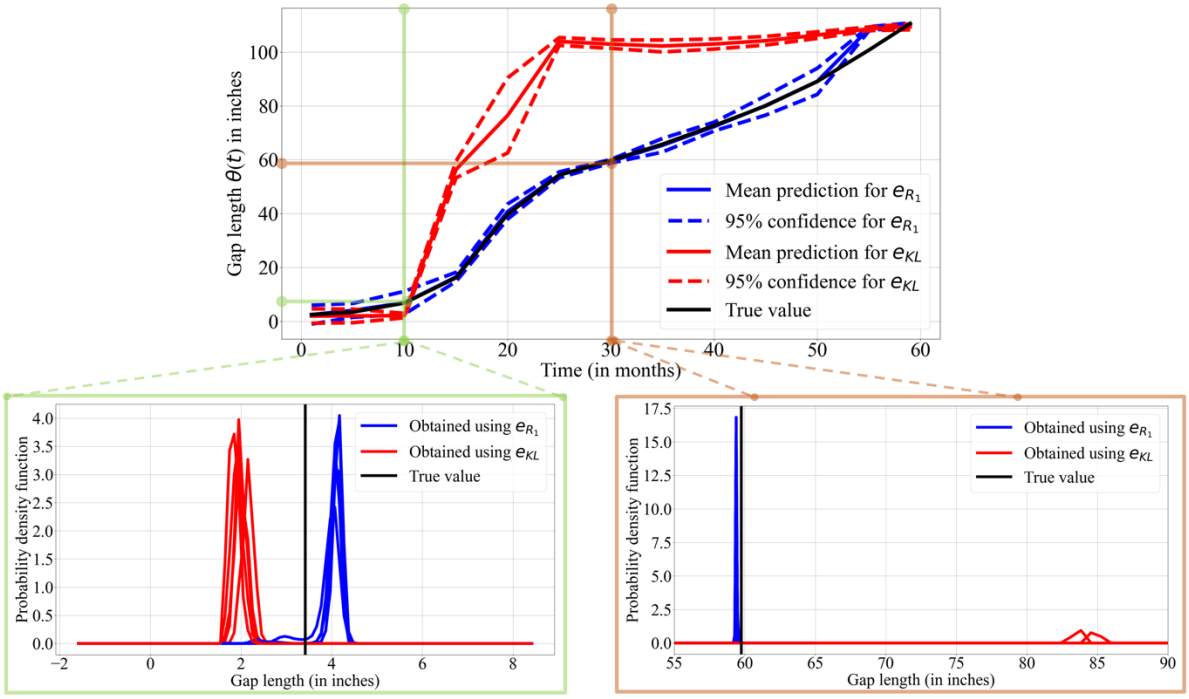


Figure 4.14: Design e_{KL} vs. e_{R_1}: Inference results for the first realization of the true gap length evolution curve

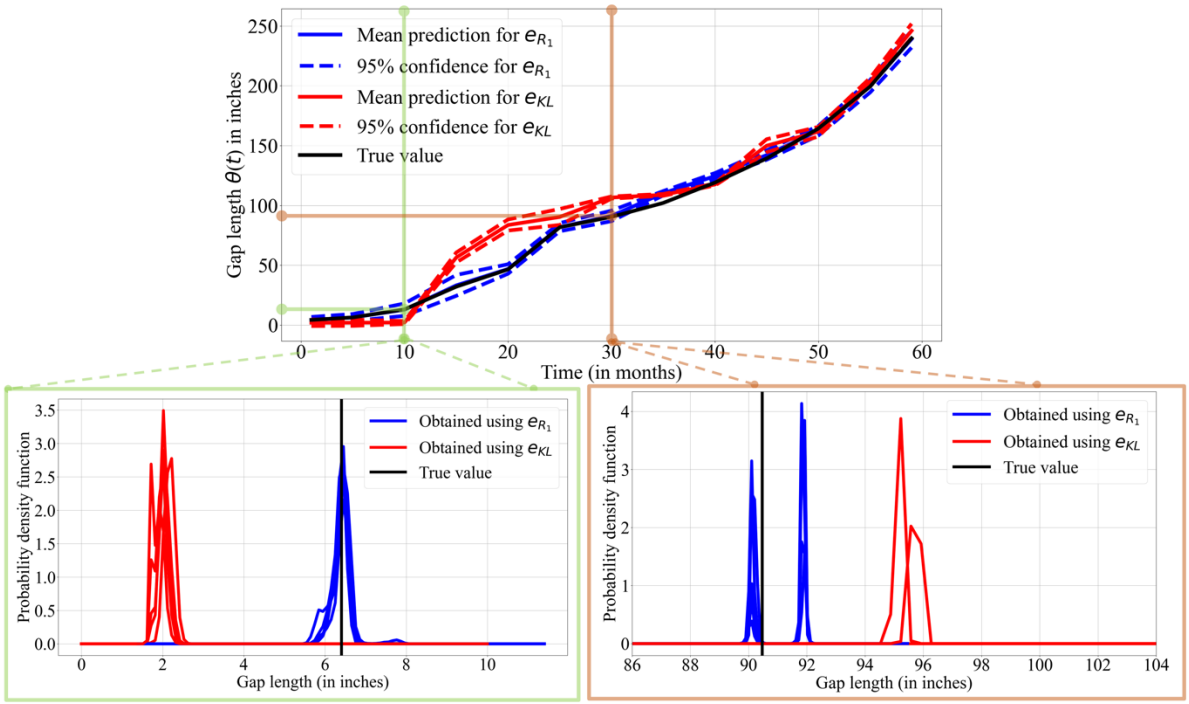


Figure 4.15: Design e_{KL} vs. e_{R_1} : Inference results for the second realization of the true gap length evolution curve

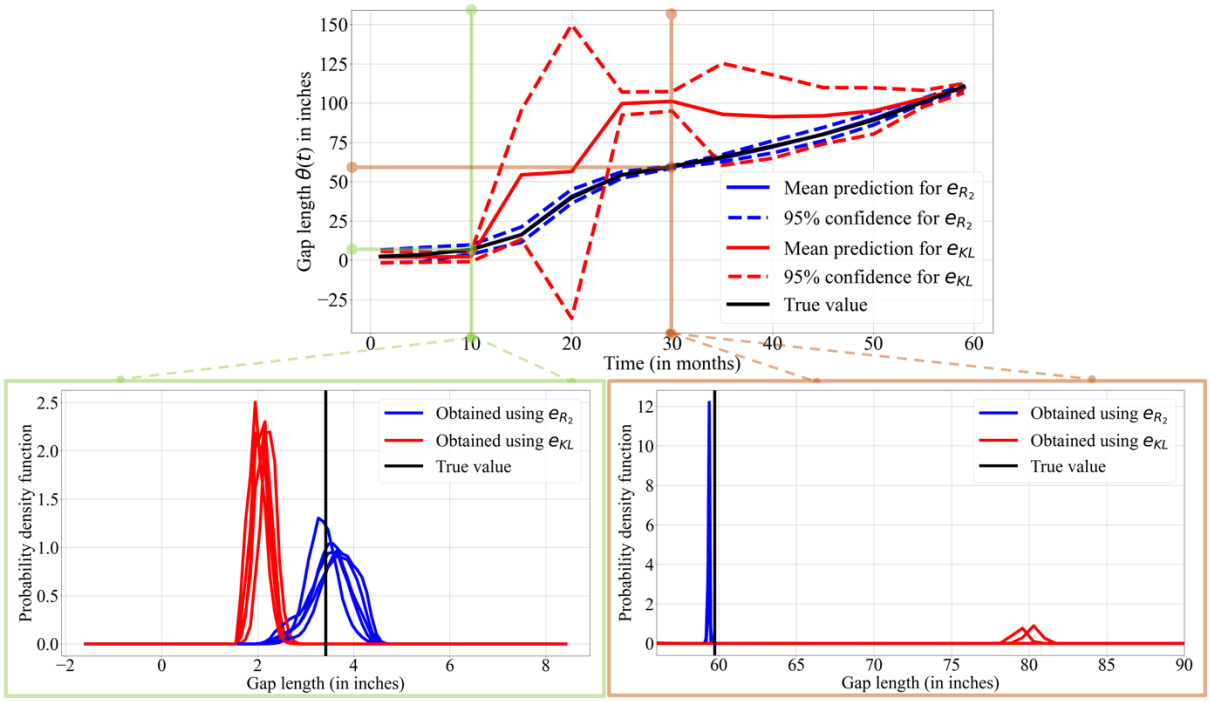


Figure 4.16: Design e_{KL} vs. e_{R_2} : Inference results for the first realization of the true gap length evolution curve

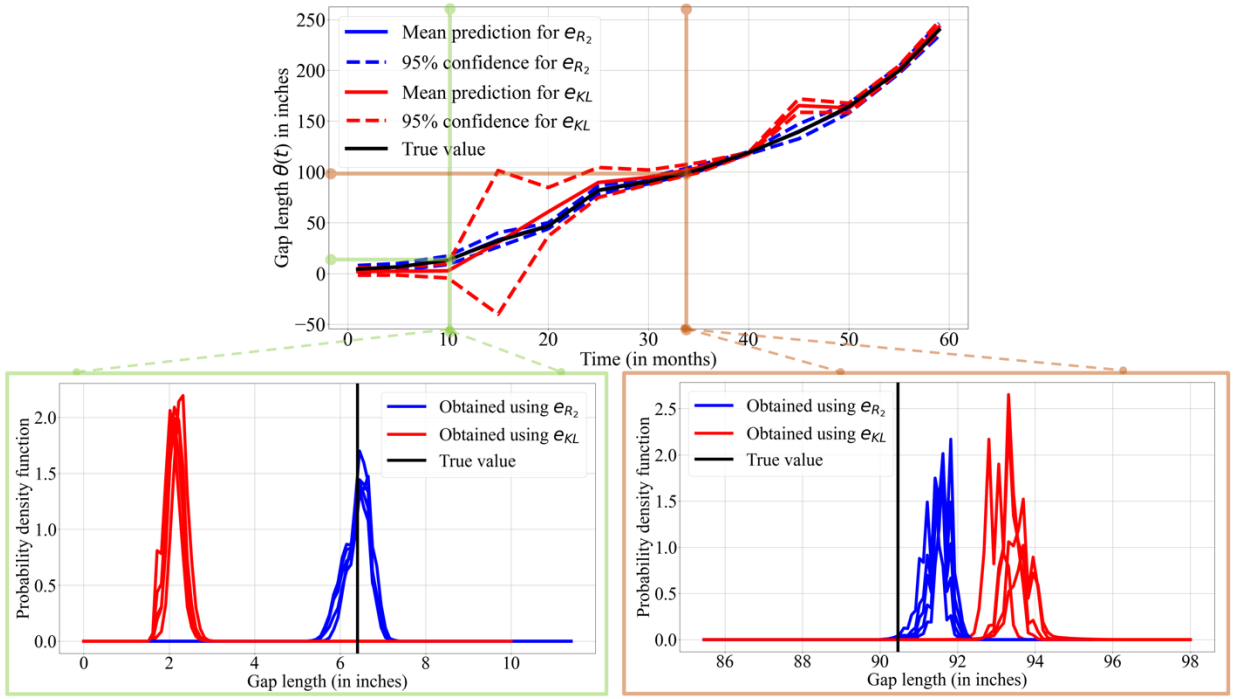


Figure 4.17: Design e_{KL} vs. e_{R2} : Inference results for the second realization of the true gap length evolution curve

As expected, I observe in Fig. 4.15-4.18 that reliability-focused designs lead to better overall inference over the lifecycle (mean of the 5 posterior distributions are closely following the true gap evolution curve). The inference obtained using the design e_{KL} are unreliable and frequently skewed away from reality (true gap length curve). Unlike the results obtained using the design e_{KL}, the inference results obtained by using the designs e_{R_1} and e_{R_2} are overall consistent for different realizations of true gap length evolution curves and different reliability bias factor.

4.6 Summary and Conclusions

This chapter proposes an optimal sensor design framework using Bayesian optimization with a dual target: (1) the design framework should obtain a sensor design that can lead to probabilistically inferring the structural state to an acceptable degree of accuracy; (2) the design framework must consider all the uncertainties that the system is subjected to and account for the sensor reliability over the structure's life cycle. This goal is achieved by using the aggregate expected sensor reliability *risk over the life cycle* as the Bayes risk functional.

The sensor placement optimization process demands evaluating the Bayes risk functional numerous times while iteratively searching the global sensor space to obtain the optimal sensor network that minimizes the Bayes risk. Doing so, I observe two major challenges: (1) the proposed Bayes risk is a high-dimensional integral with a non-linear integrand; (2) since my goal is to consider the uncertainties in load and sensor noise over the life cycle of the structure, it is computationally expensive to obtain simulated sensor data for numerous instances of true structural state parameter (observed over the structure's lifespan), loading scenario, and sensor noise structure. I overcome this numerical challenge by adopting these innovative solutions: (1) I

employ univariate dimensional reduction in conjuncture with Gauss-Hermite quadrature rule that approximates the Bayes risk in an easy-to-evaluate form; (2) I use the FEM generated simulated sensor data in tandem with singular value decomposition (SVD) technique to train a digital twin that could be deployed to obtain simulated sensor data at much lower computational cost; (3) I realized that the Bayesian optimization process requires a unique and small subset of exhaustive sensor data over the structure's lifespan. This allowed me to first store the unique set of sensor measurements that are needed for Bayesian inference and extract it as and when required in the optimization process increasing the computational speed by many folds.

The proposed framework is applied to a real-world miter gate structure. A comparison of approximated inferred posterior gap evolution over time obtained using three sensor designs (the first obtained by maximizing information gain, and the other two obtained by accounting for sensor reliability over the structure's lifecycle) indicates the importance of accounting for sensor reliability in the design framework. The reliability-focused designs lead to inference results that are overall reliable, consistent, and representative of true gap evolution over time.

4.7 Remarks

Portions of this dissertation have been published or accepted for publication or are currently being considered for publication. Chapter 4 is currently in preparation for publication, the dissertation author was the primary investigator and author of this paper:

Y. Yang, M. Chadha, Z. Hu, and M. D. Todd, "An optimal sensor design framework accounting for sensor reliability over the structural life cycle," Reliability Engineering & System Safety.

Chapter 5

CONCLUSIONS AND FUTURE RESEARCH

SHM system design with considering objective functions started in early 1990's. Over the several decades, the researchers in engineering systems have further developed and deployed various techniques for designing an optimal SHM system. Among all, placing sensors in suitable location that provides most valuable information about the structure has been widely studied for application in civil engineering. The primary goal of this research is to propose a practical and efficient optimal sensor design framework that can be implemented to large infrastructures such as miter gates. The old school human visualization technique in SHM should be replaced by a life cycle digital twin model. This dissertation demonstrates that an optimal sensor design framework that creates large benefits in SHM due to the detection accuracy, minimum designed risk, and practical to applications. The framework involves feature extraction, damage diagnosis, Bayes risk-based objective functions, uncertainty quantification, and sensor optimization.

The main contributions in this dissertation can be summarized as: 1) overcome the lack of field data under various inputs (e.g. damage states, loading vector) scenarios by building an efficient and accurate surrogate model with integrating machine learning techniques such as *GPR*; 2) estimate the damage state using *Bayesian inference* method that updates the initial belief with observed measurements; 3) constructed *Bayes risk* functions, with considering the consequences of making decision, as the objective in optimization; 4) reduced the uncertainties in the optimization process by computationally efficient *univariate dimension reduction with Gaussian Hermite integration* method; 5) Implement *Bayesian optimization* strategy in selecting best sensor placement; 6) create a sensor reliability model and 7) apply to a life cycle degradation structure.

Chapter 2 presents basic ideas and fundamental approaches for designing an optimal sensor placement. This part considers a binary and linear decision-making process in a simple Bayes risk function that is constructed the used for optimization. Three different statistical methods are introduced and compared in their performances in computing the expected value for Bayes risk. There is a tradeoff between estimation accuracy and model running times. The final optimally designed sensors result shows value of the sensor design framework against to a randomized sensor arrangement.

Chapter 3 overcomes multiple obstacles presented in realistic scenarios, including complexity of big infrastructures and high computational cost, and then further develops the framework for the miter gate problem. This chapter changes the objective function to the equations that measures information gain in damage estimation. In addition, a human knowledge-based risk weight function is added to account for the decision making. Various machine learning techniques and efficient computational strategies have been introduced to the framework for reducing the computational efforts. The framework is applied to the miter gate. A critical design of sensor locations has been achieved by maximizing the information gain.

Chapter 4 further develops the framework to account for the sensor reliability in life cycle monitoring. It is noted that the probabilities of sensor failure are highly related to their environmental variants (such as, moisture and temperature). Sensors located in a harsh position would have higher probability of failure that other sensors. In the miter gate, many sensors are installed under water to get more information from the damage position. This chapter introduces a sensor reliability model that considers the huge bias caused by sensor failures. The framework is also expanded to provide a design for the life cycle miter gate.

In addition, the current framework is a static sensor design that does not evolves over time. In real-world cases, the original design of optimal sensor network could become more and more ineffective and inaccurate as the structure damage state changes over time. Updating of the sensor network design would be critical in this case. In the future research, it would be valuable to design a time-dependent optimal sensor design framework for SHM. It would be interesting to study when and where to add new sensors to the structure based on the information observed from the current sensor design network. Further, in the future, it is also worth to integrate the optimal sensor design into a big picture of decision-making digital-twin model in SHM.

REFERENCES

A. Damaso, N. Rosa, and P. Maciel, “Reliability of wireless sensor networks,” Sensors, vol. 14, no. 9, pp. 15760–15785, 2014.

A. Downey, C. Hu, and S. Laflamme, “Optimal sensor placement within a hybrid dense sensor network using an adaptive genetic algorithm with learning gene pool,” Structural Health Monitoring, vol. 17, no. 3, pp. 450–460, 2018.

A. E. Zonouz, L. Xing, V. M. Vokkarane, and Y. L. Sun, “Reliability-oriented single-path routing protocols in wireless sensor networks,” IEEE Sensors Journal, vol. 14, no. 11, pp. 4059–4068, 2014.

A. Entezami, H. Sarmadi, and B. S. Razavi, “An innovative hybrid strategy for structural health monitoring by modal flexibility and clustering methods,” Journal of Civil Structural Health Monitoring, vol. 10, no. 5, pp. 845–859, 2020.

B. A. Eick, M. D. Smith, and T. B. Fillmore, “Feasibility of retrofitting existing miter-type lock gates with discontinuous contact blocks,” Journal of Structural Integrity and Maintenance, vol. 4, no. 4, pp. 179–194, 2019.

B. Li and A. Der Kiureghian, “Robust optimal sensor placement for operational modal analysis based on maximum expected utility,” Mechanical Systems and Signal Processing, vol. 75, pp. 155–175, 2016.

Bae, B., Flachsbart, B.R., Park, K. and Shannon, M.A., 2004. Design optimization of a piezoresistive pressure sensor considering the output signal-to-noise ratio. Journal of Micromechanics and Microengineering, 14(12), p.1597.

Behboodian, J., 1972. Information matrix for a mixture of two normal distributions. Journal of statistical computation and simulation, 1(4), pp.295-314.

Berman, A., 1979. Mass matrix correction using an incomplete set of measured modes. AIAA journal, 17(10), pp.1147-1148.

Bhattacharjee, S., Das, P., Mandal, S. and Sardar, B., 2015, July. Optimization of probability of false alarm and probability of detection in cognitive radio networks using GA. In 2015 IEEE 2nd International Conference on Recent Trends in Information Systems (ReTIS) (pp. 53-57). IEEE.

Boscato, G., Dal Cin, A., Russo, S. and Sciarretta, F., 2014. SHM of historic damaged churches. In Advanced materials research (Vol. 838, pp. 2071-2078). Trans Tech Publications Ltd.

Box, G.E. and Tiao, G.C., 2011. Bayesian inference in statistical analysis (Vol. 40). John Wiley & Sons.

Brandisky, K., Sankowski, D., Banasiak, R. and Dolapchiev, I., 2012. ECT sensor optimization based on RSM and GA. COMPEL-The international journal for computation and

mathematics in electrical and electronic engineering.

Brochu, E., Cora, V.M. and De Freitas, N., 2010. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv preprint arXiv:1012.2599.

Brownjohn, J.M., 2007. Structural health monitoring of civil infrastructure. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365(1851), pp.589-622.

C. Argyris, S. Chowdhury, V. Zabel, and C. Papadimitriou, "Bayesian optimal sensor placement for crack identification in structures using strain measurements," Structural Control and Health Monitoring, vol. 25, no. 5, p. e2137, 2018.

C. E. Ebeling, An introduction to reliability and maintainability engineering. Tata McGraw-Hill Education, 2004.

C. Farrar, G. Park, K. Farinholt, and M. Todd, "Integrated solutions to shm problems: an overview of shm research at the lanl/ucsd engineering institute," tech. rep., Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2010.

C. Malings and M. Pozzi, "Optimal sensor placement and scheduling with value of information for spatio-temporal infrastructure system management," in Proc., 12th Int. Conf. on Structural Safety and Reliability, pp. 3320–3330, 2017.

C. Malings and M. Pozzi, "Value of information for spatially distributed systems: Application to sensor placement," Reliability Engineering & System Safety, vol. 154, pp. 219–233, 2016.

C. Papadimitriou, J. L. Beck, and S.-K. Au, "Entropy-based optimal sensor location for structural model updating," Journal of Vibration and Control, vol. 6, no. 5, pp. 781–800, 2000.

C. R. Farrar and K. Worden, Structural Health Monitoring: A Machine Learning Perspective. Wiley, 2012.

C. Yang, K. Liang, X. Zhang, and X. Geng, "Sensor placement algorithm for structural health monitoring with redundancy elimination model based on sub-clustering strategy," Mechanical Systems and Signal Processing, vol. 124, pp. 369–387, 2019.

Capellari, G., Chatzi, E. and Mariani, S., 2018. Structural health monitoring sensor network optimization through Bayesian experimental design. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, 4(2), p.04018016.

Catbas, F.N., Gul, M. and Burkett, J.L., 2008. Conceptual damage-sensitive features for structural health monitoring: laboratory and field demonstrations. Mechanical systems and signal processing, 22(7), pp.1650-1669.

Chadha, M., Hu, Z. and Todd, M.D., 2021. An alternative quantification of the value of information in structural health monitoring. Structural Health Monitoring,

p.14759217211028439.

Costas-Perez, L., Lago, D., Farina, J. and Rodriguez-Andina, J.J., 2008. Optimization of an industrial sensor and data acquisition laboratory through time sharing and remote access. IEEE Transactions on Industrial Electronics, 55(6), pp.2397-2404.

Cross, E.J., Worden, K. and Farrar, C.R., 2013. Structural health monitoring for civil infrastructure. In Health Assessment of Engineered Structures: Bridges, Buildings and Other Infrastructures (pp. 1-31).

D. C. Kammer, "Optimal sensor placement for modal identification using system-realization methods," Journal of Guidance, Control, and Dynamics, vol. 19, no. 3, pp. 729–731, 1996.

D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," Journal of Global optimization, vol. 13, no. 4, pp. 455–492, 1998.

D. V. Lindley, "On a measure of the information provided by an experiment," The Annals of Mathematical Statistics, pp. 986–1005, 1956.

D. W. Coit and E. Zio, "The evolution of system reliability optimization," Reliability Engineering & System Safety, vol. 192, p. 106259, 2019.

Dawson, B., 1976. Vibration condition monitoring techniques for rotating machinery. The shock and vibration digest, 8(12), p.3.

Dupuis, R., Jouhaud, J.C. and Sagaut, P., 2018. Surrogate modeling of aerodynamic simulations for multiple operating conditions using machine learning. Aiaa Journal, 56(9), pp.3622-3635.

E. Chan, "Optimal design of building structures using genetic algorithms," California Institute of Technology, Report No. EERL 97-06, 1997.

E. Wilkins, "Cumulative damage in fatigue," in Colloquium on Fatigue/Colloque de Fatigue/Kolloquium über Ermüdungsfestigkeit, pp. 321–332, Springer, 1956.

Eick, B.A., Treece, Z.R., Spencer, B.J.F., Smith, M.D., Sweeney, S.C., Alexander, Q.G. and Foltz, S.D., 2018. Miter gate gap detection using principal component analysis. ERDC CHAMPAIGN United States.

F. Chaloner and I. Verdinelli, "Bayesian experimental design: A review," Statistical Science, vol. 10, no. 3, pp. 273–304, 1995.

F. Li and B. R.Upadhyaya, "Design of sensor placement for an integral pressurizedwater reactor using fault diagnostic observability and reliability criteria," Nuclear Technology, vol. 173, no. 1, pp. 17–25, 2011.

F. McKenna, "Opensees: a framework for earthquake engineering simulation," Computing in Science & Engineering, vol. 13, no. 4, pp. 58–66, 2011.

F. Nielsen and R. Nock, “On the chi square and higher-order chi distances for approximating f-divergences,” IEEE Signal Processing Letters, vol. 21, no. 1, pp. 10–13, 2013.

F. Osterreicher and I. Vajda, “A new class of metric divergences on probability spaces and its applicability in statistics,” Annals of the Institute of Statistical Mathematics, vol. 55, no. 3, pp. 639–653, 2003.

F. Unwadia, “Methodology for optimal sensor locations for parameter identification in dynamic system,” Journal of Engineering Mechanics, vol. 120, pp. 368–390, 1994.

Farrar, C. R., Baker, W. E., Bell, T. M., Cone, K. M., Darling, T. W., Duffey, T. A., Eklund, A. & Migliori, A. 1994 Dynamic characterization and damage detection in the I-40 bridge over the Rio Grande. Los Alamos National Laboratory Report: LA-12767-MS.

Farrar, C.R. and Worden, K., 2007. An introduction to structural health monitoring. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365(1851), pp.303-315.

Farrar, C.R., Duffey, T.A., Doebling, S.W. and Nix, D.A., 1999. A statistical pattern recognition paradigm for vibration-based structural health monitoring. Structural Health Monitoring, 2000, pp.764-773.

Flynn, E.B. and Todd, M.D., 2010, July. An active sensor placement optimization strategy using data-driven Bayesian experimental design. In Bridge Maintenance, Safety, Management and Life-Cycle Optimization: Proceedings of the Fifth International IABMAS Conference, Philadelphia, USA, 11-15 July 2010 (p. 87). CRC Press.

Flynn, E.B. and Todd, M.D., 2010. A Bayesian approach to optimal sensor placement for structural health monitoring with application to active sensing. Mechanical Systems and Signal Processing, 24(4), pp.891-903.

Flynn, E.B., 2010. A Bayesian experimental design approach to structural health monitoring with application to ultrasonic guided waves. University of California, San Diego.

Foltz, SD. Investigation of mechanical breakdowns leading to lock closures, 2017.

Friswell, M.I., Mottershead, J.E. and Ahmadian, H., 2001. Finite–element model updating using experimental test data: parametrization and regularization. Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 359(1778), pp.169-186.

Fugate, M.L., Sohn, H. and Farrar, C.R., 2000, February. Unsupervised learning methods for vibration-based damage detection. In Proceedings of 18th International Modal Analysis Conference–IMAC (p. 18).

G. C. Richardson, “Navigation locks: navigation lock gates and valves,” Journal of the Waterways and Harbors Division, vol. 90, no. 1, pp. 79–102, 1964.

G. Capellari, E. Chatzi, and S. Mariani, "Cost–benefit optimization of structural health monitoring sensor networks," Sensors, vol. 18, no. 7, p. 2174, 2018.

G. F. Gomes, F. A. de Almeida, P. d. S. L. Alexandrino, S. S. da Cunha, B. S. de Sousa, and A. C. Ancelotti, "A multiobjective sensor placement optimization for shm systems considering fisher information matrix and mode shape interpolation," Engineering with Computers, vol. 35, no. 2, pp. 519–535, 2019.

Garnett, R., Osborne, M.A. and Roberts, S.J., 2010, April. Bayesian optimization for sensor set selection. In Proceedings of the 9th ACM/IEEE international conference on information processing in sensor networks (pp. 209-219).

Gathercole, C., 1998. An investigation of supervised learning in genetic programming.

Gelfand, A.E., 1996. Model determination using sampling-based methods. Markov chain Monte Carlo in practice, pp.145-161.

Gui, G., Pan, H., Lin, Z., Li, Y. and Yuan, Z., 2017. Data-driven support vector machine with optimization techniques for structural health monitoring and damage detection. KSCE Journal of Civil Engineering, 21(2), pp.523-534.

H. Sun and O. Büyüköztürk, "Optimal sensor placement in structural health monitoring using discrete optimization," Smart Materials and Structures, vol. 24, no. 12, p. 125034, 2015.

Hajela, P. and Soeiro, F.J., 1990. Structural damage detection based on static and modal analysis. AIAA journal, 28(6), pp.1110-1115.

Hayo, T., Frankenstein, B., Boller, C. and Bockenheimer, C., 2011. Approach to the Technical Qualification of a SHM System in terms of Damage Detection in Aerospace Industry. In Proc. Intl Workshop Smart Materials, Structures & NDT in Aerospace, Montreal (pp. 1-9).

Huang, B. and Du, X., 2008. Probabilistic uncertainty analysis by mean-value first order saddlepoint approximation. Reliability Engineering & System Safety, 93(2), pp.325-336.

I. Csiszár and P. C. Shields, Information theory and statistics: A tutorial. Now Publishers Inc, 2004.

I. Sason and S. Verdu, "f-divergence inequalities," IEEE Transactions on Information Theory, vol. 62, no. 11, pp. 5973–6006, 2016.

Iswandy, K. and Koenig, A., 2006. Feature selection with acquisition cost for optimizing sensor system design. Advances in Radio Science, 4(C. 1), pp.135-141.

J. Mockus, V. Tiesis, and A. Zilinskas, "The application of bayesian methods for seeking the extremum," Towards Global Optimization, vol. 2, no. 2, pp. 117–129, 1978.

J. P. Schwieterman, S. Field, L. Fischer, and A. Pizzano, "An analysis of the economic effects of terminating operations at the chicago river controlling works and o'brien locks on the chicago

area waterway system," DePaul University, Chicago, IL, 2010.

Janson, L., Schmerling, E., Clark, A. and Pavone, M., 2015. Fast marching tree: A fast marching sampling-based method for optimal motion planning in many dimensions. The International journal of robotics research, 34(7), pp.883-921.

Jeon, T., Das, P.T., Kim, M., Jeon, C., Lim, B., Soldatov, I. and Kim, C., 2021. Operational parameters for sub-nano tesla field resolution of phmr sensors in harsh environments. Sensors, 21(20), p.6891.

Jin, S., Zhou, M. and Wu, A.S., 2003, July. Sensor network optimization using a genetic algorithm. In Proceedings of the 7th world multiconference on systemics, cybernetics and informatics (pp. 109-116).

John, Weeks A. "I-20 Mississippi River Crossing At Vicksburg." Vicksburg Bridge. N.p., n.d. Web. 5 Apr. 2016.

Joshi, S. and Boyd, S., 2008. Sensor selection via convex optimization. IEEE Transactions on Signal Processing, 57(2), pp.451-462.

Jung, H.J., Song, Y., Hong, S.K., Yang, C.H., Hwang, S.J., Jeong, S.Y. and Sung, T.H., 2015. Design and optimization of piezoelectric impact-based micro wind energy harvester for wireless sensor network. Sensors and Actuators A: Physical, 222, pp.314-321.

Kammer, D.C., 1991. Sensor placement for on-orbit modal identification and correlation of large space structures. Journal of Guidance, Control, and Dynamics, 14(2), pp.251-259.

Kang, F., Xu, Q. and Li, J., 2016. Slope reliability analysis using surrogate models via new support vector machines with swarm intelligence. Applied Mathematical Modelling, 40(11-12), pp.6105-6120.

Khatami, I. and Abdollahzadeh Jamalabadi, M.Y., 2021. Optimal design of microphone array in a planar circular configuration by genetic algorithm enhanced beamforming. Journal of Thermal Analysis and Calorimetry, 145(4), pp.1817-1825.

Kidder, R.L., 1973. Reduction of structural frequency equations. AIAA journal, 11(6), pp.892-892.

Kim, S.H. and Boukouvala, F., 2020. Machine learning-based surrogate modeling for data-driven optimization: a comparison of subset selection for regression techniques. Optimization Letters, 14(4), pp.989-1010.

Kleinman, D.L. and Rao, P.K., 1977, December. An information matrix approach for aircraft parameter-insensitive control. In 1977 IEEE Conference on Decision and Control including the 16th Symposium on Adaptive Processes and A Special Symposium on Fuzzy Set Theory and Applications (pp. 316-325). IEEE.

Krause, A., Leskovec, J., Guestrin, C., VanBriesen, J. and Faloutsos, C., 2008. Efficient sensor

placement optimization for securing large water distribution networks. Journal of Water Resources Planning and Management, 134(6), pp.516-526.

L. Bottou, “Large-scale machine learning with stochastic gradient descent,” in Proceedings of COMPSTAT’2010, pp. 177–186, Springer, 2010.

L. Yao,W. A. Sethares, and D. C. Kammer, “Sensor placement for on-orbit modal identification via a genetic algorithm,” AIAA journal, vol. 31, no. 10, pp. 1922–1928, 1993.

L. Yu, S. Wang, and K. K. Lai, “A neural-network-based nonlinear metamodeling approach to financial time series forecasting,” Applied Soft Computing, vol. 9, no. 2, pp. 563–574, 2009.

Lee, M.W., Cohen, I. and Jung, S.K., 2002, December. Particle filter with analytical inference for human body tracking. In Workshop on Motion and Video Computing, 2002. Proceedings. (pp. 159-165). IEEE.

Lewis, D.D. and Catlett, J., 1994. Heterogeneous uncertainty sampling for supervised learning. In Machine learning proceedings 1994 (pp. 148-156). Morgan Kaufmann.

Li, H., Zhao, G., Qin, L. and Yang, Y., 2019. Design and optimization of a hybrid sensor network for traffic information acquisition. IEEE Sensors Journal, 20(4), pp.2132-2144.

Li, H.N., Ren, L., Jia, Z.G., Yi, T.H. and Li, D.S., 2016. State-of-the-art in structural health monitoring of large and complex civil infrastructures. Journal of Civil Structural Health Monitoring, 6(1), pp.3-16.

Lin, R.M. and Zhu, J., 2006. Model updating of damped structures using FRF data. Mechanical systems and signal processing, 20(8), pp.2200-2218.

Lin, Y.C., Tsai, P.A., Lin, C.L., Chang, K.T. and Tsai, M.D., 2015, July. Feature space optimization of multispectral imagery and LiDAR waveform data. In 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 589-592). IEEE.

Lisicki, M., Lubitz, W. and Taylor, G.W., 2016. Optimal design and operation of Archimedes screw turbines using Bayesian optimization. Applied Energy, 183, pp.1404-1417.

LIU, B., YAO, Y.D. and YE, G.R., 2005. STUDY OF OPTIMAL SENSOR PLACEMENT FOR CABLE-STAYED BRIDGE [J]. Engineering Mechanics, 5.

Liu, J., Edberg, J., Tan, M.J., Lindgren, L.E., Castagne, S. and Jarfors, A.E., 2013. Finite element modelling of superplastic-like forming using a dislocation density-based model for AA5083. Modelling and Simulation in Materials Science and Engineering, 21(2), p.025006.

Lopes, H.F. and Tsay, R.S., 2011. Particle filters and Bayesian inference in financial econometrics. Journal of Forecasting, 30(1), pp.168-209.

Lu, D. and Ricciuto, D., 2019. Efficient surrogate modeling methods for large-scale Earth system models based on machine-learning techniques. Geoscientific Model Development, 12(5),

pp.1791-1807.

Lukosi, E., Rust, M., Stand, L. and Melcher, C.L., 2019. Optimization of data acquisition parameters for single gamma line sensing: KSr2I5: Eu2+ and NaI (Tl) compared. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 938, pp.36-40.

M. A. Vega, Z. Hu, T. B. Fillmore, M. D. Smith, and M. D. Todd, "A novel framework for integration of abstracted inspection data and structural health monitoring for damage prognosis of miter gates," Reliability Engineering & System Safety, vol. 211, p. 107561, 2021.

M. Basseville, A. Benveniste, G. Moustakides, and A. Rougee, "Optimal sensor location for detecting changes in dynamical behavior," IEEE Transactions on Automatic control, vol. 32, no. 12, pp. 1067–1075, 1987.

M. Chadha and M. D. Todd, "A comprehensive kinematic model of single-manifold cosserat beam structures with application to a finite strain measurement model for strain gauges," International Journal of Solids and Structures, vol. 159, pp. 58–76, 2019.

M. Chadha, Z. Hu, and M. D. Todd, "An alternative quantification of the value of information in structural health monitoring," Structural Health Monitoring, p. 14759217211028439, 2021.

M. D. McKay, R. J. Beckman, andW. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," Technometrics, vol. 42, no. 1, pp. 55–61, 2000.

M. I. Friswell and R. Castro-Triguero, "Clustering of sensor locations using the effective independence method," AIAA journal, vol. 53, no. 5, pp. 1388–1391, 2015.

M. Khosravifard, D. Fooladivanda, and T. A. Gulliver, "Confliction of the convexity and metric properties in f-divergences," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol. 90, no. 9, pp. 1848–1853, 2007.

M. Meo and G. Zumpano, "Optimal sensor placement on a large-scale civil structure," in Health monitoring and smart nondestructive evaluation of structural and biological systems III, vol. 5394, pp. 108–117, International Society for Optics and Photonics, 2004.

M. Moustapha, J.-M. Bourinet, B. Guillaume, and B. Sudret, "Comparative study of kriging and support vector regression for structural engineering applications," ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, vol. 4, no. 2, p. 04018005, 2018.

M. Parno, D. O'Connor, and M. Smith, "High dimensional inference for the structural health monitoring of lock gates," arXiv preprint arXiv:1812.05529, 2018.

M. Ramancha, R. Astroza, J. P. Conte, J. I. Restrepo, and M. D. Todd, "Bayesian nonlinear finite element model updating of a full-scale bridge-column using sequential monte carlo," in

Proceedings of the 38th IMAC, A Conference and Exposition on Structural Dynamics 2020, Springer, 2021.

Madu, I.E. and Madu, C.N., 1999. Design optimization using signal-to-noise ratio. Simulation Practice and Theory, 7(4), pp.349-372.

Maolin, S., Liye, L., Sun, W. and Xueguan, S., 2020. A multi-fidelity surrogate model based on support vector regression. Structural and Multidisciplinary Optimization, 61(6), pp.2363-2375.

Markmiller, J.F. and Chang, F.K., 2010. Sensor network optimization for a passive sensing impact detection technique. Structural Health Monitoring, 9(1), pp.25-39.

Modares, M. and Waksmanski, N., 2013. Overview of structural health monitoring for steel bridges. Practice Periodical on Structural Design and Construction, 18(3), pp.187-191.

Mottershead, J.E., Link, M. and Friswell, M.I., 2011. The sensitivity method in finite element model updating: A tutorial. Mechanical systems and signal processing, 25(7), pp.2275-2296.

Mustapha, A., Mohamed, L. and Ali, K., 2020, June. An overview of gradient descent algorithm optimization in machine learning: application in the ophthalmology field. In International Conference on Smart Applications and Data Analysis (pp. 349-359). Springer, Cham.

Nagarajan, V. and Kolter, J.Z., 2017. Gradient descent GAN optimization is locally stable. arXiv preprint arXiv:1706.04156.

Ostachowicz, W., Soman, R. and Malinowski, P., 2019. Optimization of sensor placement for structural health monitoring: A review. Structural Health Monitoring, 18(3), pp.963-988.

P. I. Frazier, "A tutorial on bayesian optimization," arXiv preprint arXiv:1807.02811, 2018.

P. K. Agarwal, E. Ezra, and S. K. Ganjugunte, "Efficient sensor placement for surveillance problems," in International Conference on Distributed Computing in Sensor Systems, pp. 301– 314, Springer, 2009.

P. Nath, Z. Hu, and S. Mahadevan, "Sensor placement for calibration of spatially varying model parameters," Journal of Computational Physics, vol. 343, pp. 150–169, 2017.

P. Rubenstein, O. Bousquet, J. Djolonga, C. Riquelme, and I. O. Tolstikhin, "Practical and consistent estimation of f-divergences," in Advances in Neural Information Processing Systems, pp. 4072–4082, 2019.

P. S. Marsh and D. M. Frangopol, "Reinforced concrete bridge deck reliability model incorporating temporal and spatial variations of probabilistic corrosion rate sensor data," Reliability Engineering & System Safety, vol. 93, no. 3, pp. 394–409, 2008.

Pabari, P.A., Willson, K., Stegemann, B., van Geldorp, I.E., Kyriacou, A., Moraldo, M., Mayet, J., Hughes, A.D. and Francis, D.P., 2011. When is an optimization not an optimization?

Evaluation of clinical implications of information content (signal-to-noise ratio) in optimization of cardiac resynchronization therapy, and how to measure and maximize it. Heart failure reviews, 16(3), pp.277-290.

Padula, S.L. and Kincaid, R.K., 1999. Optimization strategies for sensor and actuator placement.

Pao, Y.H. and Sobajic, D.J., 1992. Combined use of unsupervised and supervised learning for dynamic security assessment. IEEE Transactions on Power Systems, 7(2), pp.878-884.

Peh, E. and Liang, Y.C., 2007, March. Optimization for cooperative sensing in cognitive radio networks. In 2007 IEEE wireless communications and networking conference (pp. 27-32). IEEE.

Pozzi, M. and Der Kiureghian, A., 2011, April. Assessing the value of information for long-term structural health monitoring. In Health monitoring of structural and biological systems 2011 (Vol. 7984, p. 79842W). International Society for Optics and Photonics.

R. A. Daniel, "Miter gates in some recent lock projects in the netherlands (stemmtore in einigenneuen schleusenanlagen in den niederlanden)," Stahlbau, vol. 69, no. 12, pp. 952–964, 2000.

R. Duan, Y. Lin, and T. Feng, "Optimal sensor placement based on system reliability criterion under epistemic uncertainty," IEEE Access, vol. 6, pp. 57061–57072, 2018.

R. F. Guratzsch and S. Mahadevan, "Structural health monitoring sensor placement optimization under uncertainty," AIAA journal, vol. 48, no. 7, pp. 1281–1289, 2010.

Rahman, S. and Xu, H., 2004. A univariate dimension-reduction method for multi-dimensional integration in stochastic mechanics. Probabilistic Engineering Mechanics, 19(4), pp.393-408.

Ren, W.X. and Chen, H.B., 2010. Finite element model updating in structural dynamics by using the response surface method. Engineering structures, 32(8), pp.2455-2465.

Retout, S., Duffull, S. and Mentré, F., 2001. Development and implementation of the population Fisher information matrix for the evaluation of population pharmacokinetic designs. Computer methods and Programs in Biomedicine, 65(2), pp.141-151.

Rose, V.S., Croall, I.F. and Macfie, H.J., 1991. An Application of Unsupervised Neural Network Methodology Kohonen Topology‐Preserving Mapping to QSAR Analysis. Quantitative Structure‐Activity Relationships, 10(1), pp.6-15.

S. Li, H. Zhang, S. Liu, and Z. Zhang, "Optimal sensor placement using frfs-based clustering method," Journal of sound and vibration, vol. 385, pp. 69–80, 2016.

S. Lloyd, "Measures of complexity: a nonexhaustive list," IEEE Control Systems Magazine, vol. 21, no. 4, pp. 7–8, 2001.

S. R. Peddada, P. J. Tannous, A. G. Alleyne, and J. T. Allison, "Optimal sensor placement methods

in active high power density electronic systems with experimental validation," Journal of Mechanical Design, vol. 142, no. 2, 2020.

S. Rahman and H. Xu, "A univariate dimension-reduction method for multi-dimensional integration in stochastic mechanics," Probabilistic Engineering Mechanics, vol. 19, no. 4, pp. 393–408, 2004.

S. S. Ram, A. Nedic, and V. V. Veeravalli, "Stochastic incremental gradient descent for estimation in sensor networks," in 2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers, pp. 582–586, IEEE, 2007.

S. Thöns and M. H. Faber, "Assessing the value of structural health monitoring," ICOSSAR'13 - 11th International conference on structural safety and reliability - Safety, reliability, riskand life-cycle performance of structures and infrastructures (Proceedings), 2013.

Sakai, J. and Kitsunezuka, M., 2019, March. Optimization of sensor disposition for accurate localization in a practical environment. In 2019 URSI Asia-Pacific Radio Science Conference (AP-RASC) (pp. 1-4). IEEE.

Singh, K., Singh, K. and Aziz, A., 2018. Congestion control in wireless sensor networks by hybrid multi-objective optimization algorithm. Computer Networks, 138, pp.90-107.

Smith, R.C., 2013. Uncertainty quantification: theory, implementation, and applications (Vol. 12). Siam.

Sohn, H., Farrar, C.R., Hemez, F.M., Shunk, D.D., Stinemates, D.W., Nadler, B.R. and Czarnecki, J.J., 2003. A review of structural health monitoring literature: 1996–2001. Los Alamos National Laboratory, USA, 1.

Sony, S., Laventure, S. and Sadhu, A., 2019. A literature review of next‐generation smart sensing technology in structural health monitoring. Structural Control and Health Monitoring, 26(3), p.e2321.

Su, G., Peng, L. and Hu, L., 2017. A Gaussian process-based dynamic surrogate model for complex engineering structural reliability analysis. Structural Safety, 68, pp.97-109.

Sullivan, T.J., 2015. Introduction to uncertainty quantification (Vol. 63). Springer.

T.-H. Yi and H.-N. Li, "Methodology developments in sensor placement for health monitoring of civil infrastructures," International Journal of Distributed Sensor Networks, vol. 8, no. 8, p. 612726, 2012.

T.-H. Yi, H.-N. Li, and M. Gu, "A new method for optimal selection of sensor location on a highrise building using simplified finite element model," Structural Engineering and Mechanics, vol. 37, no. 6, pp. 671–684, 2011.

T.-H. Yi, H.-N. Li, and M. Gu, "Optimal sensor placement for health monitoring of high-rise structure based on genetic algorithm," Mathematical Problems in Engineering, vol. 2011,

2011.

Taha, M.R. and Lucero, J., 2005. Damage identification for structural health monitoring using fuzzy pattern recognition. Engineering Structures, 27(12), pp.1774-1783.

Teng, J. and Zhu, Y.H., 2011. Optimal sensor placement for modal parameters test of large span spatial steel structural. Engineering Mechanics, 28(3), pp.150-156.

Todd, M.D. and Flynn, E.B., 2011, March. A Bayesian experimental design approach for structural health monitoring. In Proceedings of the XIV International Symposium on Dynamic Problems of Mechanics (DINAME 2011), Brazil.

Tripathy, R.K. and Bilionis, I., 2018. Deep UQ: Learning deep neural network surrogate models for high dimensional uncertainty quantification. Journal of computational physics, 375, pp.565-588.

V. Akbarzadeh, J.-C. Lévesque, C. Gagné, and M. Parizeau, "Efficient sensor placement optimization using gradient descent and probabilistic coverage," Sensors, vol. 14, no. 8, pp. 15525–15552, 2014.

Van Loan, C.F., 1976. Generalizing the singular value decomposition. SIAM Journal on numerical Analysis, 13(1), pp.76-83.

Vega, M.A. and Todd, M.D., 2020. A variational Bayesian neural network for structural health monitoring and cost-informed decision-making in miter gates. Structural Health Monitoring, p.1475921720904543.

Vega, M.A., Hu, Z. and Todd, M.D., 2020. Optimal maintenance decisions for deteriorating quoin blocks in miter gates subject to uncertainty in the condition rating protocol. Reliability Engineering & System Safety, 204, p.107147.

W. A. Maul, G. Kopasakis, L. M. Santi, T. S. Sowers, and A. Chicatelli, "Sensor selection and optimization for health assessment of aerospace systems," Journal of Aerospace Computing, Information, and Communication, vol. 5, no. 1, pp. 16–34, 2008.

W.-H. Kang, J. Song, and P. Gardoni, "Matrix-based system reliability method and applications to bridge networks," Reliability Engineering & System Safety, vol. 93, no. 11, pp. 1584–1593, 2008.

M. Li, M. Sadoughi, Z. Hu, and C. Hu, "A hybrid gaussian process model for system reliability analysis," Reliability Engineering & System Safety, vol. 197, p. 106816, 2020.

Wang, Z., Zoghi, M., Hutter, F., Matheson, D. and De Freitas, N., 2013, June. Bayesian optimization in high dimensions via random embeddings. In Twenty-Third international joint conference on artificial intelligence.

Williams, D., Heng, I.S., Gair, J., Clark, J.A. and Khamesra, B., 2019. A precessing numerical relativity waveform surrogate model for binary black holes: a Gaussian process regression

approach. arXiv preprint arXiv:1903.09204.

X. Huan and Y. M. Marzouk, “Simulation-based optimal bayesian experimental design for nonlinear systems,” Journal of Computational Physics, vol. 232, no. 1, pp. 288–317, 2013.

X. Nguyen, M. J. Wainwright, and M. I. Jordan, “Estimating divergence functionals and the likelihood ratio by penalized convex risk minimization,” in Advances in neural information processing systems, pp. 1089–1096, 2008.

Xiang, H., Li, Y., Liao, H. and Li, C., 2017. An adaptive surrogate model based on support vector regression and its application to the optimization of railway wind barriers. Structural and Multidisciplinary Optimization, 55(2), pp.701-713.

Y. Tan and L. Zhang, “Computational methodologies for optimal sensor placement in structural health monitoring: A review,” Structural Health Monitoring, vol. 19, no. 4, pp. 1287–1308, 2020.

Y. Tang, Zhiyi aand Bao, , and H. Li, “Group sparsity-aware convolutional neural network for continuous missing data recovery of structural health monitoring,” Structural Health Monitoring, vol. 20, no. 4, pp. 1738–1759, 2020.

Y. Yang, M. Chadha, Z. Hu, and M. D. Todd, “An optimal sensor placement design framework for structural health monitoring using bayes risk,” Mechanical Systems and Signal Processing, vol. 168, 2022, p.108618, ISSN 0888-3270,https://doi.org/10.1016/j.ymssp.2021.108618.

Y. Yang, M. Chadha, Z. Hu, M. A. Vega, M. D. Parno, and M. D. Todd, “A probabilistic optimal sensor design approach for structural health monitoring using risk-weighted f-divergence,” Mechanical Systems and Signal Processing, vol. 161, p. 107920, 2021.

Y. Yang, R. Madarshahian, and M. D. Todd, “Bayesian damage identification using strain data from lock gates,” in Dynamics of Civil Structures, Volume 2, pp. 47–54, Springer, 2020.

Yan, W.J., Chronopoulos, D., Papadimitriou, C., Cantero-Chinchilla, S. and Zhu, G.S., 2020. Bayesian inference for damage identification based on analytical probabilistic model of scattering coefficient estimators and ultrafast wave scattering simulation scheme. Journal of Sound and Vibration, 468, p.115083.

Ye, Z.H.A.O., Ling, L. and Huabao, Y., 2012. The optimal design of bracket structure for the second mirror of a large-aperture space optical remote sensor. Mechanical Science and Technology for Aerospace Engineering, 31(3), pp.373-378.

Yi, T.H., Li, H.N. and Gu, M., 2011. Optimal sensor placement for structural health monitoring based on multiple optimization strategies. The Structural Design of Tall and Special Buildings, 20(7), pp.881-900.

Young Noh, H., Krishnan Nair, K., Lignos, D.G. and Kiremidjian, A.S., 2011. Use of wavelet-based damage-sensitive features for structural damage diagnosis using strong motion data.

Journal of Structural Engineering, 137(10), pp.1215-1228.

Z. Hu, D. Ao, and S. Mahadevan, “Calibration experimental design considering field response and model uncertainty,” Computer Methods in Applied Mechanics and Engineering, vol. 318, pp. 92–119, 2017.

Z. Hu, S. Nannapaneni, and S. Mahadevan, “Efficient kriging surrogate modeling approach for system reliability analysis,” AI EDAM, vol. 31, no. 2, pp. 143–160, 2017.

Z. Shi, S. Law, and L. Zhang, “Optimum sensor placement for structuraldamage detection,” Journal of Engineering Mechanics, vol. 126, no. 11, pp. 1173–1179, 2000.

Zabel, S., Kirchhof, C., Yarar, E., Meyners, D., Quandt, E. and Faupel, F., 2015. Phase modulated magnetoelectric delta-E effect sensor for sub-nano tesla magnetic fields. Applied Physics Letters, 107(15), p.152402.

Zhang, J., 2019. Gradient descent based optimization algorithms for deep learning models training. arXiv preprint arXiv:1903.03614.