PARAMETER ESTIMATION VIA BAYESIAN INVERSION: THEORY, METHODS, AND APPLICATIONS

by

Ryan Michael Soncini

B.S. in Mechanical Engineering, University of Pittsburgh, 2012

Submitted to the Graduate Faculty of

Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

M.S in Mechanical Engineering

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This thesis was presented

by

Ryan Michael Soncini

It was defended on

November 21, 2013

and approved by

Anne M. Robertson, PhD, Associate Professor
Department of Mechanical Engineering and Materials Science

Giovanni P. Galdi, PhD, Associate Professor
Department of Mechanical Engineering and Materials Science

Thesis Advisor: Paolo Zunino, PhD, Assistant Professor
Department of Mechanical Engineering and Materials Science

PARAMETER ESTIMATION VIA BAYESIAN INVERSION: THEORY, METHODS, AND APPLICATIONS

Ryan Michael Soncini, M.S.

University of Pittsburgh, 2013

Uncertainty quantification is becoming an increasingly important area of investigation in the field of computational simulations. An understanding in the confidence of a simulation result requires information concerning the uncertainties associated with individual sub-models. The development of mathematical models for physical systems resides in the interpretation of experimental results. Inherent to physically interesting mathematical models is the occurrence of unobservable model parameters. The resolution of information concerning model parameters is typically performed through the use of least-squares regression analysis; however, least-squares analysis does not provide adequate information concerning the confidence which may be placed in the parameter estimates. Bayesian inversion provides quantifiable information concerning the confidence which may be placed in the parameter estimates allowing for overall simulation uncertainty quantification. Here, the application of Bayesian statistics to the general discrete inverse problem is presented. Following the presentation of the Bayesian formulation of the general discrete inverse problem, the procedure is applied to two scientifically interesting inverse problems: the reversible-reaction diffusion inverse problem and the Arrhenius inverse problem. The Arrhenius inverse problem is solved using a novel approach developed here. The novel approach is compared to other probabilistic and deterministic approaches to assess the validity of the method.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

1.1 MATHEMATICAL MODELS AND PARAMETER ESTIMATION

Mathematical models are tools which provide predictions of the behavior of physical systems. These may range from simple algebraic expressions to coupled systems of partial differential equations. The development of mathematical models is based in the physical interpretation of experimental observations and their application in improving some prior knowledge of a system's behavior. Measurements are collected, inspected, and parameterized into a mathematical expression in terms of observable and unobservable quantities. The presence of unobservable *model parameters* is intrinsic to the parameterization of any scientifically interesting physical system. The resolution of information concerning the value of model parameters falls under a field of study known as *parameter estimation*. The intended use of mathematical models is *forward modeling*: the making of predictions in observable quantities, provided some knowledge of the model parameters. It follows that the *inverse problem* may be described as the determination of model parameters, provided a set of experimental results. Estimation of model parameters from experimental observations is a task often performed through the use of *frequentist* regression techniques, an unconstrained optimization problem attempting to minimize some comparative metric between the forward model result and experimental observations. This deterministic approach to parameter estimation provides distinct

values of the model parameters along with simplistic metrics of estimate quality, e.g. coefficient of determination. A caveat to the frequentist approach is a lack of quantifiable parameter uncertainty. The eventual goal in the development of mathematical models is their application in the design of functional products and processes for industrial and consumer use. The advent of computational physics modeling has greatly accelerated the design process across many scientific and engineering disciplines. Computational simulations of physical systems may involve several sequential mathematical models, each requiring estimated parameters. The propagation of uncertainties due to parameter estimation may only be quantified if initial knowledge exists concerning the confidence in individual estimates; knowledge which is not conveyed in the result of a regression procedure. *Bayesian inversion* attempts to provide a more comprehensive state of information of the model parameters than frequentist techniques by expressing unobservable quantities in terms of a probability density. This probability density over the model parameters provides a means of quantifying parameter uncertainty. *Uncertainty quantification* (UQ) may be defined as: the process of quantifying uncertainties associated with forward modeling calculations, attempting to account for all potential sources of uncertainty and quantifying the contributions of each individual source [1]. The *probabilistic* understanding of parameter uncertainty provided by the Bayes' formulated inverse problem enables the application of forward modeling UQ techniques and provides a more informative state of knowledge regarding the results of computational simulations. The application of Bayesian statistics to inverse problems is far from a novel concept; however, the specific formulation of a Bayesian procedure for certain types of engineering problems is a new and interesting field of study.

1.2 LEAST-SQUARES PARAMETER ESTIMATION

1.2.1 General Least-Squares Formulation

The method of parameter estimation most typically employed is a regression technique referred to as least-squares analysis. Let $g(\boldsymbol{m}; x)$ be a mathematical model where $\boldsymbol{m}$ is a vector of the model parameters of interest and x is the control variable of the model. Taking $\{y_j\}_{j=1}^{m}$ be a set of experimentally determined responses at control values $\{x_j\}_{j=1}^{m}$, the method of least squares may be formulated as the unconstrained optimization problem given by:

$$\underset{\boldsymbol{m}}{minimize}\ f(\boldsymbol{m}) = \sum_{j=1}^{m} (g(\boldsymbol{m}; x_j) - y_j)^2$$

Here, the intent is to select values of the model parameters such that the sum of the squares of the absolute deviations between the experimentally determined values and the model results, referred to as the residuals, is minimized. This technique is referred to as a fixed-regressor method as it is assumed that the control variable, x, is known with high confidence and the response, y, is treated as a random variable. For non-linear mathematical models the optimal parameter vector is obtained through some search type method, e.g. Gauss-Newton, Levenberg-Marquardt, or Nelder-Mead Simplex, which works to refine some initial guess of the optimal solution [2]. The method returns some optimal values of the model parameters with no method of confidence quantification other than the value of the objective function at the optimal solution, providing little to no insight for forward modeling uncertainty quantification. Furthermore, the method of least squares heavily weights the solution toward measurements that significantly deviate from the model as the objective function involves the square of the residual. This results in a solution shifted toward possible erroneous experimental measurements.

1.2.2 Linear Least-Squares Formulation

The case where the mathematical model is a linear function of the control variable is particularly interesting in that a unique, closed-form solution to the least-squares problem may be formulated. Here, the objective function may be formulated as:

$$f(a,b) = \sum_{j=1}^{m} \left(y_j - ax_j - b\right)^2$$

The problem of estimating the values of the parameters is solved by differentiating the objective function with respect to each parameter and equating these to zero [3]. Define:

$$s_x = \frac{1}{m}\sum_{j=1}^{m} x_j \qquad s_{xx} = \frac{1}{m}\sum_{j=1}^{m} x_j^2$$

$$s_y = \frac{1}{m}\sum_{j=1}^{m} y_j \qquad s_{xy} = \frac{1}{m}\sum_{j=1}^{m} x_j y_j$$

These definitions allow for the calculation of the unique linear model parameters by:

$$a = \frac{s_{xy} - s_x s_y}{s_{xx} - (s_x)^2}$$

$$b = \frac{s_{xx} s_y - s_{xy} s_x}{s_{xx} - (s_x)^2}$$

A metric used in the assessment of the quality of the fit associated with this technique is referred to as the coefficient of determination, denoted by R^2, is given by:

$$SS_{tot} = \sum_{j=1}^{m} \left(y_j - s_y\right)^2 \qquad SS_{res} = \sum_{j=1}^{m} \left(g\left(a,b;x_j\right) - y_j\right)^2$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

Again, there is no means of accurately quantifying the forward modeling uncertainty.

1.3 BAYES' THEOREM AND THE DISCRETE INVERSE PROBLEM

Bayesian inversion intends to update some *a priori* information concerning the parameters of a given model using the results of experimental measurements. This problem involves the combination of information contained within probability densities over various mathematical manifolds. The Bayesian approach to statistical inference is a *subjective* interpretation of probability; that is, probability may be understood as a *degree of belief* concerning the true values of the model parameters. Bayes' theorem may be viewed as a systematic approach to the concept of learning, updating knowledge concerning some proposition, provided some relevant evidence [4]. Bayes' theorem may be developed through manipulation of Kolmogorov's axioms of probability. Let P be a probability measure over a probability space Ω and let $\boldsymbol{a}$, $\boldsymbol{b}$ be events contained within the probability space, see Figure 1.1

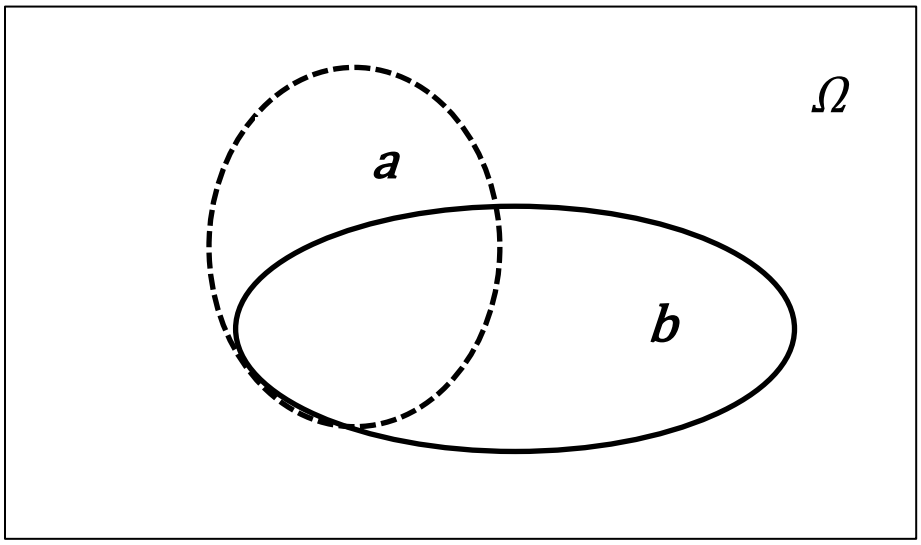

Figure 1.1: Diagram of Probability Space

The conditional probability of an event $\boldsymbol{a}$ given that event $\boldsymbol{b}$ has occurred, denoted by $P(\boldsymbol{a}|\boldsymbol{b})$, is defined as:

$$P(\boldsymbol{a}|\boldsymbol{b}) = \frac{P(\boldsymbol{a} \cap \boldsymbol{b})}{P(\boldsymbol{b})}$$

It follows from the previous definition and an understanding of set theory that:

$$P(\boldsymbol{b}|\boldsymbol{a}) = \frac{P(\boldsymbol{a} \cap \boldsymbol{b})}{P(\boldsymbol{a})}$$

Manipulation of the previous definitions yields the general form of *Bayes' theorem* [5]:

$$P(\boldsymbol{a}|\boldsymbol{b}) = \frac{P(\boldsymbol{b}|\boldsymbol{a})P(\boldsymbol{a})}{P(\boldsymbol{b})} \tag{1}$$

Considering event $\boldsymbol{a}$ to be some proposition and event $\boldsymbol{b}$ to be some evidence; Bayes' theorem provides a systematic approach to improve an *a priori* state of information, $P(\boldsymbol{a})$, utilizing the evidentiary support provided by $P(\boldsymbol{b}|\boldsymbol{a})/P(\boldsymbol{b})$, resulting in an *a posteriori* knowledge of the proposition, $P(\boldsymbol{a}|\boldsymbol{b})$. In the context of the Bayes' formulated inverse problem, the aforementioned evidence takes the form of experimental measurements, the proposition is a mathematical model with unobservable quantities, and the probability measure is a probability density [6-8].

1.3.1 The Model Space

Let $\mathbb{M}$ be a finite, *k*-dimensional, linear vector space containing the set of all conceivable models, referred to here as the *model space*. The dimensionally, *k*, of the model space corresponds to the number of unobservable quantities contained in the mathematical expression intended for inversion. Individual vectors of the model space are denoted by $\boldsymbol{m} = (m^1, m^2, \dots, m^k)$, where the coordinates of the *model vector* correspond to individual model parameters.

1.3.2 Bayesian Inversion Framework

1.3.2.1 *A Priori* Information

A priori information about the model parameters is expressed in the form of a probability density over the model space, referred to as the *prior probability density*, $\psi(\boldsymbol{m})$. This probability density contains the current state of information concerning the values of the model parameters, acquired before the interpretation of recent experimental results. The formulation of the prior probability may be performed through inspection of preceding experimental results or the application of accepted postulates governing the physical system of interest. Dated or simplistic prior experimental results may provide information as descriptive as a unimodal probability density or as ambiguous as upper or lower limits on parameter values. This information serves as a launch point for improved experimental techniques applied through the Bayesian approach to narrow the uncertainty of such states of information. Axioms of physics may allow for the exclusion of regions of the model space due to accepted non-physicality of certain model values. *A priori* information need not initially come in the form of a probability density over the entirety of the model space. Information concerning individual model parameters, termed *relative prior probability densities*, may be considered statistically independent. This allows the prior probability to be formed as a *joint probability density* from the product of the relative priors [6]. Equation (2) describes this method of prior combination.

$$\psi(\boldsymbol{m}) = \prod_i \psi^i(m^i) \tag{2}$$

The formulation of the relative priors in each model parameter is a difficult task as quantifiable prior information may not always be available. This concept delves further into the subjective

nature of Bayesian inversion as the selection of a prior probability from a qualitative understanding of the model parameters is a heuristic than a definite science [9].

1.3.2.2 The Likelihood Function

The updating of *a priori* information concerning model parameters is accomplished through an assessment of the compatibility of individual models and the experimental results. This assessment takes the form of a *conditional probability density* over the model space, termed the *likelihood*, $\lambda(\boldsymbol{d}|\boldsymbol{m})$. The likelihood probability density contains a state of information concerning model-data compatibility while accounting for uncertainties inherent to experimental measurements. Formulation of a likelihood expression involves knowledge concerning the confidence which may be placed in a specific measurement technique, e.g. uniform $\mp$2% of instrument response or Gaussian uncertainty with known variance. This information may from an uncertainty specification supplied by the manufacturer of the measurement instrument or from the result of a calibration procedure. The specific form of the likelihood function is based in the investigator's belief in the quality of the measurement technique. Discrete computation of the likelihood is accomplished through the evaluation of the forward problem for individual model vectors, followed by assessment of the agreement between individual models and their corresponding experimental results. Mathematical models typically attempt to predict the relationship between some *regressor variable* and some *response variable*. Experiments are conducted in a manner which attempts to provide the inverse problem with sufficient data to infer some knowledge of the model parameters. Inherent to this process is the collection of multiple measurements, each of which containing intrinsic uncertainty. Each measurement act may be considered statistically independent of the others, allowing the *relative likelihoods*

associated with individual measurements to be combined through products [6]. Equation (3) describes this method of relative likelihood combination.

$$\lambda(\boldsymbol{d}|\boldsymbol{m}) = \prod_i \lambda^i(d^i|\boldsymbol{m}) \tag{3}$$

1.3.2.3 Bayes' Theorem

The combination of the measurement and *a priori* states of information is performed through the application of Bayes' Theorem, Equation (4).

$$\eta(\boldsymbol{m}|\boldsymbol{d}) = \frac{\lambda(\boldsymbol{d}|\boldsymbol{m})\psi(\boldsymbol{m})}{\int_{\mathbb{M}} \lambda(\boldsymbol{d}|\boldsymbol{m})\psi(\boldsymbol{m})d\boldsymbol{m}} \tag{4}$$

The conditional probability of the model parameters given the experimental data is termed the *posterior probability density*. This conditional probability constitutes the solution to a Bayes' formulated inverse problem, providing information regarding the accuracy of the measurement technique as well as the previous understanding of the model parameters.

1.3.3 Point Estimation, the Covariance Matrix, and Marginalization

While the posterior probability density constitutes the solution to the Bayes' formulated inverse problem, it provides little utility without the definition of parameter value and uncertainty quantification techniques. Point estimation serves as a means of interpreting the posterior density to obtain distinct values of the model parameters which are indicative of the behavior of the posterior density. One simple method of interpretation of the posterior probability density is to determine the model vector where the posterior probability achieves a global maximum, termed the *maximum a posteriori* (MAP) estimator, Equation (5) [6].

$$\boldsymbol{m}_{MAP} = \underset{\boldsymbol{m} \in \mathbb{M}}{\arg\max}\, \eta(\boldsymbol{m}|\boldsymbol{d}) \tag{5}$$

Another form of point estimator is the first central moment of the posterior density, referred to as the *posterior mean* (EV), whose method of calculation is shown in Equation (6) [10].

$$E_i = \int_{\mathbb{M}} m^i \, \eta(\boldsymbol{m}|\boldsymbol{d}) d\boldsymbol{m} \tag{6}$$

While point estimation serves to produce a most probable value for individual model parameters, the covariance matrix provides information concerning parameter uncertainty and parameter interactions. The diagonal elements of the covariance matrix are the individual parameter variances and provide information about the confidence that may be placed in a single parameter estimate. The off-diagonal elements provide information about the correlations between model parameters. Equation (7) provides the method of covariance matrix calculation.

$$\Sigma_{ij} = \int_{\mathbb{M}} (m^i - E_i(\boldsymbol{m}))(m^j - E_j(\boldsymbol{m})) \, \eta(\boldsymbol{m}|\boldsymbol{d}) d\boldsymbol{m} \tag{7}$$

Lastly, when inspecting high dimensional probability densities it may be useful to eliminate the density's explicit dependence on certain parameters. This is accomplished by marginalizing the density over the parameter intended for elimination. Consider the bi-variate posterior density given by $\eta(m^1, m^2|\boldsymbol{d})$. Suppose that the parameter m^1 is of particular interest and the density's dependence on m^2 is of little concern. Parameter m^2 may be marginalized out by:

$$\zeta(m^1, |\boldsymbol{d}) = \int_{m^2} \eta(m^1, m^2|\boldsymbol{d}) d\boldsymbol{m} \tag{8}$$

2.0 THE REVERSIBLE REACTION-DIFFUSION INVERSE PROBLEM

2.1 REVERSIBLE REACTION-DIFFUSION MODEL

The Bayesian approach to the discrete inverse problem is best conveyed through the use of examples. This chapter presents a full computational example of Bayesian inversion as it applies to a complex system of partial differential equations; specifically, the reversible reaction diffusion problem. Consider the reversible reaction of two chemical species A and B.

$$A \underset{k_b}{\overset{k_f}{\leftrightarrow}} B \tag{9}$$

Taking the concentrations of A and B to be u and v respectively, the spatial concentration distribution of each species in time, with zero flux boundaries, may be described by the following system of partial differential equations:

$$\begin{aligned} \partial_t u - D_A \nabla^2 u + k_f u - k_b v &= 0 \ \ in\ \Omega \\ \partial_t v - D_B \nabla^2 v - k_f u + k_b v &= 0 \ \ in\ \Omega \\ \nabla u \cdot \boldsymbol{n} &= 0 \ \ on\ \partial\Omega \\ \nabla v \cdot \boldsymbol{n} &= 0 \ \ on\ \partial\Omega \end{aligned} \tag{10}$$

System (10) constitutes a mathematical model of the reversible reaction-diffusion system. There does not exist a measurement technique to directly transduce the mass diffusivities and kinetic rate constants of System (10), making these quantities unobservable model parameters. The

concentrations of each species; however, are observable through mass spectrum measurement techniques. This knowledge may be used to design an experiment to resolve information concerning the model parameters. The reversible reaction-diffusion inverse problem serves as an excellent exercise in the application of Bayesian parameter estimation to multi-parameter, differential systems. Here, experimental measurement data is numerically generated from a perturbed solution to the reversible reaction-diffusion mathematical model. Bayesian inference techniques are then applied to the data in an effort to recover the parameter values used in the data generation.

2.2 ARTIFICIATL REACTION-DIFFUSION EXPERIMENT

While no actual experiments were performed in the conduction of this study, a hypothetical experimental setup is proposed acquaint the reader with the method of artificial data generation. Consider a cylindrical vessel 10 cm in length and whose diameter is sufficiently small such that radial mass fluxes may be considered negligible. Five concentration sampling locations are mounted equidistantly along the length of the containment vessel and feed into a mass spectrometer. This mass spectrometer is capable of determining the concentration of each species at each sampling location simultaneously with Gaussian uncertainty of specified variance. Initially, the vessel is separated into three regions by negligibly thin splitter plates. Regions 1 and 3, see Figure 2.1, contain a mixture of 5 mol-cm^{-1} of species A and 20 mol-cm^{-1} of species B. Region 2 contains 30 mol-cm^{-1} of species A and 10 mol-cm^{-1} of species B. The splitter plates are then removed and the concentration of each species is recorded at each sampling location every 2 seconds for 30 seconds.

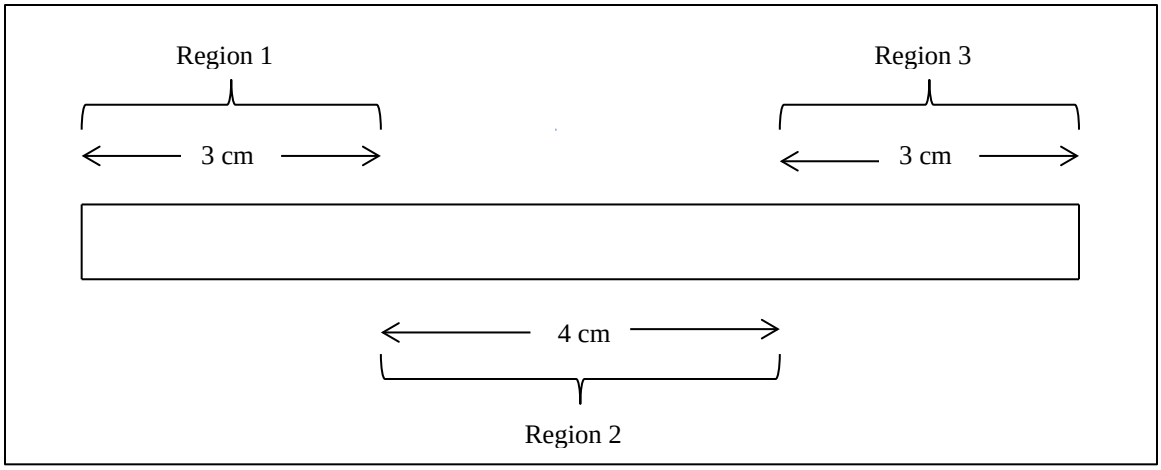

Figure 2.1: Schematic of Hypothetical Experiment

2.3 THE REACTION-DIFFUSION MODEL SOLUTION AND DATA GENERATION

2.3.1 Finite Difference Formulation

Bayesian inversion requires the existence of a method for forward problem evaluation to provide values of the observable quantities at a given model vector. There is no closed-form solution to the reaction-diffusion system, requiring the use of a numerical method for concentration evaluation. Here, the reaction-diffusion system is solved by the method of finite differences, employing a backward in time, central in space (BTCS) scheme [11]. The one-dimensional nature of the experiment allows for the specific formulation of System (10) as:

$$\begin{aligned} &\partial_t u - D_A \partial_{xx} u + k_f u - k_b v = 0 \quad x \in (a,b), t \in [0,T] \\ &\partial_t v - D_B \partial_{xx} v - k_f u + k_b v = 0 \quad x \in (a,b), t \in [0,T] \\ &\partial_x u = 0 \quad x = a,b \\ &\partial_x v = 0 \quad x = a,b \end{aligned} \tag{11}$$

To apply the method of finite differences, let $J, N \in \mathbb{N}$. The spatial and temporal step sizes are then defined by:

$$\Delta x = \frac{(b-a)}{(J-1)}, \Delta t = \frac{T}{(N-1)}$$

These definitions allow for the formulation of a space-time grid given by:

$$\{(x_j, t^n): 1 \leq j \leq J, 1 \leq n \leq N\}$$

The nodes of this grid are given by:

$$x_j = a + (j-1)\Delta x,\ t^n = (n-1)\Delta t$$

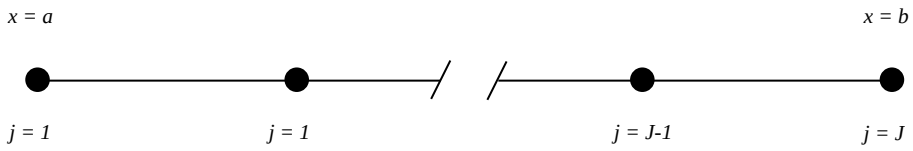

Figure 2.2: Schematic of Finite Difference Spatial Discretization

Use of Taylor's theorem provides second order approximations of the differential expressions present in System (11). Applying first order, backward and second order, central finite difference stencils to System (11), in time and space respectively, results in the discretized reaction diffusion system given by:

$$\begin{aligned} &\frac{u_j^{n+1} - u_j^n}{\Delta t} - D_A\left[\frac{u_{j-1}^{n+1} - 2u_j^{n+1} + u_{j+1}^{n+1}}{\Delta x^2}\right] + k_f u_j^{n+1} - k_b v_j^{n+1} = 0 \\ &\frac{v_j^{n+1} - v_j^n}{\Delta t} - D_B\left[\frac{v_{j-1}^{n+1} - 2v_j^{n+1} + v_{j+1}^{n+1}}{\Delta x^2}\right] - k_f u_j^{n+1} + k_b v_j^{n+1} = 0 \end{aligned} \tag{12}$$

System (12) may be simplified through algebraic manipulation and the definition of the *Fourier number*, $Fo_A = \frac{D_A \Delta t}{\Delta x^2}, Fo_B = \frac{D_B \Delta t}{\Delta x^2}$, in both transport species. The simplified discrete system is then given by:

$$\begin{aligned} -Fo_A u_{j-1}^{n+1} + (1 + 2Fo_A + \Delta t k_f) u_j^{n+1} - Fo_A u_{j+1}^{n+1} - \Delta t k_b v_j^{n+1} &= u_j^n \\ -Fo_B v_{j-1}^{n+1} + (1 + 2Fo_B + \Delta t k_b) v_j^{n+1} - Fo_B v_{j+1}^{n+1} - \Delta t k_f u_j^{n+1} &= v_j^n \end{aligned} \tag{13}$$

System (13) may be solved in a time marching fashion, solving a linear system at each time step. Letting $\vec{b} = (u_1^n, \dots, u_J^n, v_1^n, \dots, v_J^n)^T \in \mathbb{R}^{2J}$ and $\vec{w} = (u_1^{n+1}, \dots, u_J^{n+1}, v_1^{n+1}, \dots, v_J^{n+1})^T \in \mathbb{R}^{2J}$, a linear system of the form $C\vec{w} = \vec{b}$ may be formulated. The matrix C may be constructed in a block fashion through the following definitions:

$$\tilde{C}_u = tridiag\left(-Fo_A, \left[1 + 2Fo_A + \Delta t k_f\right], -Fo_A\right) \in \mathbb{R}^{J \times J}$$

$$\tilde{C}_v = tridiag(-Fo_B, [1 + 2Fo_B + \Delta t k_b], -Fo_B) \in \mathbb{R}^{J \times J}$$

$$Z_u = -\Delta t k_b I \in \mathbb{R}^{J \times J}$$

$$Z_v = -\Delta t k_f I \in \mathbb{R}^{J \times J}$$

$$\tilde{C} = \begin{bmatrix} \tilde{C}_u & Z_u \\ Z_v & \tilde{C}_v \end{bmatrix}) \in \mathbb{R}^{2J \times 2J}$$

This formulation of the stiffness matrix does not account for the Neumann boundary conditions. Boundary conditions are applied through the method of fictitious points by employing a second order approximation of the gradient about the boundary points and applying the effect to the stiffness matrix. The Neumann boundary condition asserts that:

$$\partial_x u|_{j=1} = 0, \ \partial_x u|_{j=J} = 0$$

$$\partial_x v|_{j=1} = 0, \partial_{\ x} v|_{j=J} = 0$$

Application of a second order, central approximation of the first derivative shows that applying $u_{j-1} = u_{j+1}$ at the boundaries for the fictitious points satisfies the Neumann condition. This is accomplished by:

$$-Fo_A u_0^{n+1} + (1 + 2Fo_A + \Delta t k_f) u_1^{n+1} - Fo_A u_2^{n+1} - \Delta t k_b v_1^{n+1} = u_1^n$$

= 1:

$$-Fo_B v_0^{n+1} + (1 + 2Fo_B + \Delta t k_b) v_1^{n+1} - Fo_B v_2^{n+1} - \Delta t k_f u_1^{n+1} = v_1^n$$

Application of $u_{j-1} = u_{j+1}$ and $v_{j-1} = v_{j+1}$ to the first spatial node yields:

$$-Fo_A u_2^{n+1} + (1 + 2Fo_A + \Delta t k_f) u_1^{n+1} - Fo_A u_2^{n+1} - \Delta t k_b v_1^{n+1} = u_1^n$$

= 1:

$$-Fo_B v_2^{n+1} + (1 + 2Fo_B + \Delta t k_b) v_1^{n+1} - Fo_B v_2^{n+1} - \Delta t k_f u_1^{n+1} = v_1^n$$

It follows that application of $u_{j-1} = u_{j+1}$ and $v_{j-1} = v_{j+1}$ to the last spatial node yields:

$$-Fo_A u_{J-1}^{n+1} + (1 + 2Fo_A + \Delta t k_f) u_J^{n+1} - Fo_A u_{J-1}^{n+1} - \Delta t k_b v_J^{n+1} = u_J^n$$

= J:

$$-Fo_B v_{J-1}^{n+1} + (1 + 2Fo_B + \Delta t k_b) v_J^{n+1} - Fo_B v_{J-1}^{n+1} - \Delta t k_f u_J^{n+1} = v_J^n$$

The resulting modifications to the stiffness matrix are:

$$C_u^{1,2} = \tilde{C}_u^{1,2} - Fo_A, C_u^{J,J-1} = \tilde{C}_u^{J,J-1} - Fo_A$$

$$C_v^{1,2} = \tilde{C}_v^{1,2} - Fo_B, C_v^{J,J-1} = \tilde{C}_v^{J,J-1} - Fo_B$$

$$C = \begin{bmatrix} C_u & Z_u \\ Z_v & C_v \end{bmatrix}) \in \mathbb{R}^{2J \times 2J}$$

The reaction-diffusion problem may then be solved in a time-marching fashion, solving the linear system $C\vec{w} = \vec{b}$ at each time step to resolve the space-time distribution of concentration over the $[a, b] \times [0, T]$ domain.

2.3.2 Computational Implementation

The solution to the reaction-diffusion forward problem is computationally resolved using the MATLAB programing environment in a modular design fashion. The model solver is coded as a function with inputs: model vector, initial spatial concentration distribution in species A, and initial spatial concentration distribution in species B. The model solver function returns two, two dimensional arrays representing the space-time concentration data in each chemical species. The columns of both arrays contain the spatial concentration distributions at each time node. This data storage structure may be depicted by:

$$U_M = \begin{bmatrix} u_1^1 & \cdots & u_1^N \\ \vdots & \ddots & \vdots \\ u_J^1 & \cdots & u_J^N \end{bmatrix} \in \mathbb{R}^{J\times N}, V_M = \begin{bmatrix} v_1^1 & \cdots & v_1^N \\ \vdots & \ddots & \vdots \\ v_J^1 & \cdots & v_J^N \end{bmatrix} \in \mathbb{R}^{J\times N}$$

The space time discretization used for the model solution is specified by a space range, a time range, the number of spatial nodes, and the number of temporal nodes. This information is stored in a call-only model grid function. The size of each array is initialized using the *ones* built-in MATLAB function. The pre-boundary stiffness matrices in each species are constructed according to the finite difference formulation using the *diag* MATLAB function. These matrices are modified to account for the Neumann boundaries and then pieced together to form the full stiffness matrix. The linear system of equations associated with the discretization is solved at each time step using the *mldivide*, "\", MATLAB function.

2.3.3 Data Generator

The data generation function is designed to yield data of the form described in the artificial experiment. This involves the execution of the model solver using a known model vector. The

arrays generated by the model solver are of a higher resolution than the experimental procedure would produce, requiring the use of a data localization function which reduces the size of the data arrays in accordance with the experimental constraints. This procedure results in a data structure of the form:

$$U_D = \begin{bmatrix} \bar{u}_1^1 & \cdots & \bar{u}_1^{\tilde{N}} \\ \vdots & \ddots & \vdots \\ \bar{u}_{\tilde{J}}^1 & \cdots & \bar{u}_{\tilde{J}}^{\tilde{N}} \end{bmatrix} \in \mathbb{R}^{\tilde{J}\times\tilde{N}}, V_D = \begin{bmatrix} \bar{v}_1^1 & \cdots & \bar{v}_1^{\tilde{N}} \\ \vdots & \ddots & \vdots \\ \bar{v}_{\tilde{J}}^1 & \cdots & \bar{v}_{\tilde{J}}^{\tilde{N}} \end{bmatrix} \in \mathbb{R}^{\tilde{J}\times\tilde{N}}$$

The values of these arrays are then randomly perturbed in a manner consistent with the prescribed Gaussian uncertainty of the fictitious mass-spectrometer, treating the unperturbed model value as the mean and randomly sampling from the resulting distribution.

2.4 BAYES APPROACH TO THE REACTION-DIFFUSION INVERSE PROBLEM

2.4.1 *A Priori* Information

Inspection the physics occurring in the reaction-diffusion system informs that it would be unphysical for any of the model parameters to be less than zero. A review of current, in this case hypothetical, research literature concerning similar systems could provide an educated estimate of an upper bound to each of the four model parameters. Given no other information about the model parameters the only inference that may be made is that true value of each parameter resides somewhere between zero and a prescribed upper bound, with uniform probability within the bounded region. Each model parameter will therefore have a uniform prior density given by:

$$\psi^i(m^i) = \begin{cases} \frac{1}{m^{i,max}} & \forall\, m^i \in [0, m^{i,max}] \\ 0 & otherwise \end{cases} \qquad (14)$$

These, statistically independent, individual contributions to the prior information may be combined through products by:

$$\psi(\boldsymbol{m}) = \prod_i \psi^i(m^i) \tag{15}$$

Note that in the case of all uniform relative priors, the prior probability is a constant value over the entirety of the model space.

2.4.2 Measurement Uncertainty and the Likelihood Function

Inherent to the transduction of any observable quantity is random measurement uncertainty. The uncertainty of the mass spectrometer is taken to be Gaussian with known variance. Each individual concentration measurement will have an associated relative likelihood, described by Equations (16) and (17).

$$\varphi_j^n\left(u_j^n \middle| \boldsymbol{m}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{(u_j^n - \bar{u}_j^n)^2}{2\sigma^2}\right] \tag{16}$$

$$\theta_j^n\left(v_j^n \middle| \boldsymbol{m}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{(v_j^n - \bar{v}_j^n)^2}{2\sigma^2}\right] \tag{17}$$

Taking the measurement uncertainties to be statistically independent, the relative likelihoods may be combined through products to form a joint density which provides a metric of the cumulative compatibility of the data and each individual model vector.

$$\lambda(\boldsymbol{d}|\boldsymbol{m}) = \prod_{j,n} \varphi_j^n\left(u_j^n \middle| \boldsymbol{m}\right)\theta_j^n\left(v_j^n \middle| \boldsymbol{m}\right) \tag{18}$$

Here, the relative likelihoods are subjectively constructed based in the belief that the measurement device obeys Gaussian uncertainty.

2.4.3 Numerical Resolution of the Posterior Density

Numerical calculation of the posterior density is carried out by discretizing the 4-dimensional model space and computing the posterior probability at each discreet model. Due to the uniform nature of the relative priors, the model space is taken to be the region bounded by the prior density limits, that is:

$$\mathbb{M} = [0, D_A^{max}] \times [0, D_B^{max}] \times [0, k_f^{max}] \times [0, k_b^{max}]$$

To discretize the model space let $P, Q, R, S \in \mathbb{N}$, allowing for the definition of model parameter step sizes, given by:

$$\Delta D_A = \frac{D_A^{max}}{(P-1)}, \Delta D_B = \frac{D_B^{max}}{(Q-1)}, \Delta k_f = \frac{k_f^{max}}{(R-1)}, \Delta k_b = \frac{k_b^{max}}{(S-1)}$$

Individual coordinate grid points may then be described by:

$$D_A^p = (p-1)\Delta D_A, D_B^q = (q-1)\Delta D_B, k_f^r = (r-1)\Delta k_f, k_b^s = (s-1)\Delta k_b$$

These definitions allow for the statement of the model grid by:

$$\{(D_A^p, D_B^q, k_f^r, k_b^s) : 1 \le p \le P, 1 \le q \le Q, 1 \le r \le R, 1 \le s \le S\}$$

This grid may be computationally navigated in a systematic fashion through the use of four nested for loops, one over each model parameter index. Computation of the discrete posterior density is carried out using the MATLAB programming environment in a modular fashion. The prior information function reads in the lower and upper bounds on the individual model parameters and computes the value of the prior probability according to Equations (14) and (15). Due to the uniform nature of all of the relative priors, the prior probability is a constant value over the entirely of the model space, implying that it need not be recalculated at each grid loop recursion. The likelihood solver function is designed to read in the measurement data in each species and the model solution associated with a given model vector. Each model solution must

first be reduced to size compatible with the generated data, a task accomplished through the execution of the aforementioned data localization function. The measurement data and model solution are compared through Equations (16) and (17). The relative likelihoods associated with each measurement, in each species, φ and θ, are stored in two, two dimensional arrays similar in structure to the measurement data arrays, depicted by:

$$\Phi = \begin{bmatrix} \varphi_1^1 & \cdots & \varphi_1^{\tilde{N}} \\ \vdots & \ddots & \vdots \\ \varphi_{\tilde{J}}^1 & \cdots & \varphi_{\tilde{J}}^{\tilde{N}} \end{bmatrix} \in \mathbb{R}^{\tilde{J}\times\tilde{N}}, \Theta = \begin{bmatrix} \theta_1^1 & \cdots & \theta_1^{\tilde{N}} \\ \vdots & \ddots & \vdots \\ \theta_{\tilde{J}}^1 & \cdots & \theta_{\tilde{J}}^{\tilde{N}} \end{bmatrix} \in \mathbb{R}^{\tilde{J}\times\tilde{N}}$$

Because the data is held fixed, the likelihood is a function of the model space. The likelihood solver function must be executed at each recursion of the grid loop. At each for loop recursion the returned likelihood value is multiplied by the prior probability, resulting in a non-normalized value of the posterior density. These nodal, non-normalized posterior values are stored in a 4-dimensional array. Due to the successive nature of the method, the posterior probabilities may not be normalized into a probability density until all of the posterior probabilities have been computed. The nodal values of the non-normalized posterior probability are computed at each node of the discretized model space by:

$$\tilde{\eta}(\boldsymbol{m}^{p,q,r,s}|\boldsymbol{d}) = \lambda(\boldsymbol{d}|\boldsymbol{m}^{p,q,r,s})\psi(\boldsymbol{m}^{p,q,r,s}) \tag{19}$$

Upon solution of the non-normalized posterior probability density, the normalization constant is determined. The normalization constant, K, is given by:

$$K = \int_{\mathbb{M}} \tilde{\eta}(\boldsymbol{m}|\boldsymbol{d})d\boldsymbol{m}$$

In the case of the reversible reaction-diffusion problem, this integral may be expanded to:

$$K = \int_{\mathbb{M}} \tilde{\eta}\left(D_A, D_B, k_f, k_b\middle|\boldsymbol{d}\right)dm = dD_A dD_B dk_f dk_b$$

Fubini's theorem allows for this integral to be evaluated through the successive computation of one-dimensional integrals over each parameter coordinate by:

$$\tilde{\eta}'(D_A, D_B, k_f|\boldsymbol{d}) = \int_0^{k_b^{max}} \tilde{\eta}\big(D_A, D_B, k_f, k_b\big|\boldsymbol{d}\big)dk_b$$

$$\tilde{\eta}''(D_A, D_B|\boldsymbol{d}) = \int_0^{k_f^{max}} \tilde{\eta}'\big(D_A, D_B, k_f\big|\boldsymbol{d}\big)dk_f$$

$$\tilde{\eta}'''(D_A|\boldsymbol{d}) = \int_0^{D_B^{max}} \tilde{\eta}''(D_A, D_B|\boldsymbol{d})dD_B$$

$$K = \int_0^{D_A^{max}} \tilde{\eta}'''(D_A|\boldsymbol{d})dD_A$$

Numerically, the preceding succession of integrals is computed using trapezoidal quadrature. The composite trapezoidal quadratures for the computation of the normalization constant may be written as [12]:

$$\tilde{\eta}'(\boldsymbol{m}^{p,q,r}|\boldsymbol{d}) = \frac{\Delta k_b}{2}\sum_{s=1}^{S-1}(\tilde{\eta}(\boldsymbol{m}^{p,q,r,s}|\boldsymbol{d}) + \tilde{\eta}(\boldsymbol{m}^{p,q,r,s+1}|\boldsymbol{d}))$$

$$\tilde{\eta}''(\boldsymbol{m}^{p,q}|\boldsymbol{d}) = \frac{\Delta k_f}{2}\sum_{r=1}^{R-1}(\tilde{\eta}'(\boldsymbol{m}^{p,q,r}|\boldsymbol{d}) + \tilde{\eta}'(\boldsymbol{m}^{p,q,r+1}|\boldsymbol{d}))$$

$$\tilde{\eta}'''(\boldsymbol{m}^{p}|\boldsymbol{d}) = \frac{\Delta D_B}{2}\sum_{q=1}^{Q-1}(\tilde{\eta}''(\boldsymbol{m}^{p,q}|\boldsymbol{d}) + \tilde{\eta}''(\boldsymbol{m}^{p,q+1}|\boldsymbol{d}))$$

$$K = \frac{\Delta D_A}{2}\sum_{p=1}^{P-1}(\tilde{\eta}'''(\boldsymbol{m}^{p}|\boldsymbol{d}) + \tilde{\eta}'''(\boldsymbol{m}^{p+1}|\boldsymbol{d}))$$

Following the calculation of the normalization constant the raw posterior values are normalized to give a true probability density by:

$$\eta(\boldsymbol{m}|\boldsymbol{d}) = \frac{1}{K}\lambda(\boldsymbol{d}|\boldsymbol{m})\psi(\boldsymbol{m})$$

2.5 APPLICATIUON BAYESIAN INVERSION TO THE REVERSIBLE REACTION-DIFFUSION INVERSE PROBLEM

Experimental data of the type described by the artificial experiment is generated from the model using parameters listed in Table 2.1. The standard deviation of the measurement device is taken to be 0.015 mol-cm^{-1}. The model solver grid uses a spatial step-size of 0.25cm and a temporal step-size of 0.25s. Each model coordinate is discretized using 11 nodes.

Table 2.1: Parameter Information

Parameter	True Value	Lower Bound	Upper Bound
$\mathbf{D_A}$	0.8	0.0	1.0
$\mathbf{D_B}$	0.6	0.0	1.0
$\mathbf{k^f}$	0.4	0.0	1.0
$\mathbf{k^b}$	0.3	0.0	1.0

Figure 2.3 shows bivariate contours of the posterior density, holding the remaining two parameters at their true values. The dashed black line denotes the true value used in data generation. It can be seen from Figure 2.3 that the posterior density appears to be unimodal and is sharply centered near the true value model vector. The quantitative analysis shown in Table 2.2 shows that true values were in fact recovered with considerable confidence. This is not surprising as the experimental setup was designed to provide sufficient spatial and temporal information to accurately resolve the model parameters. These results serve as a means of algorithm verification. Now, suppose that the reactants used in these experiments were cost prohibitive and that this variety of experiment is to be conducted for many different species. It would be beneficial to reduce the experiment run time and reduce the quantity of each species used in each experiment.

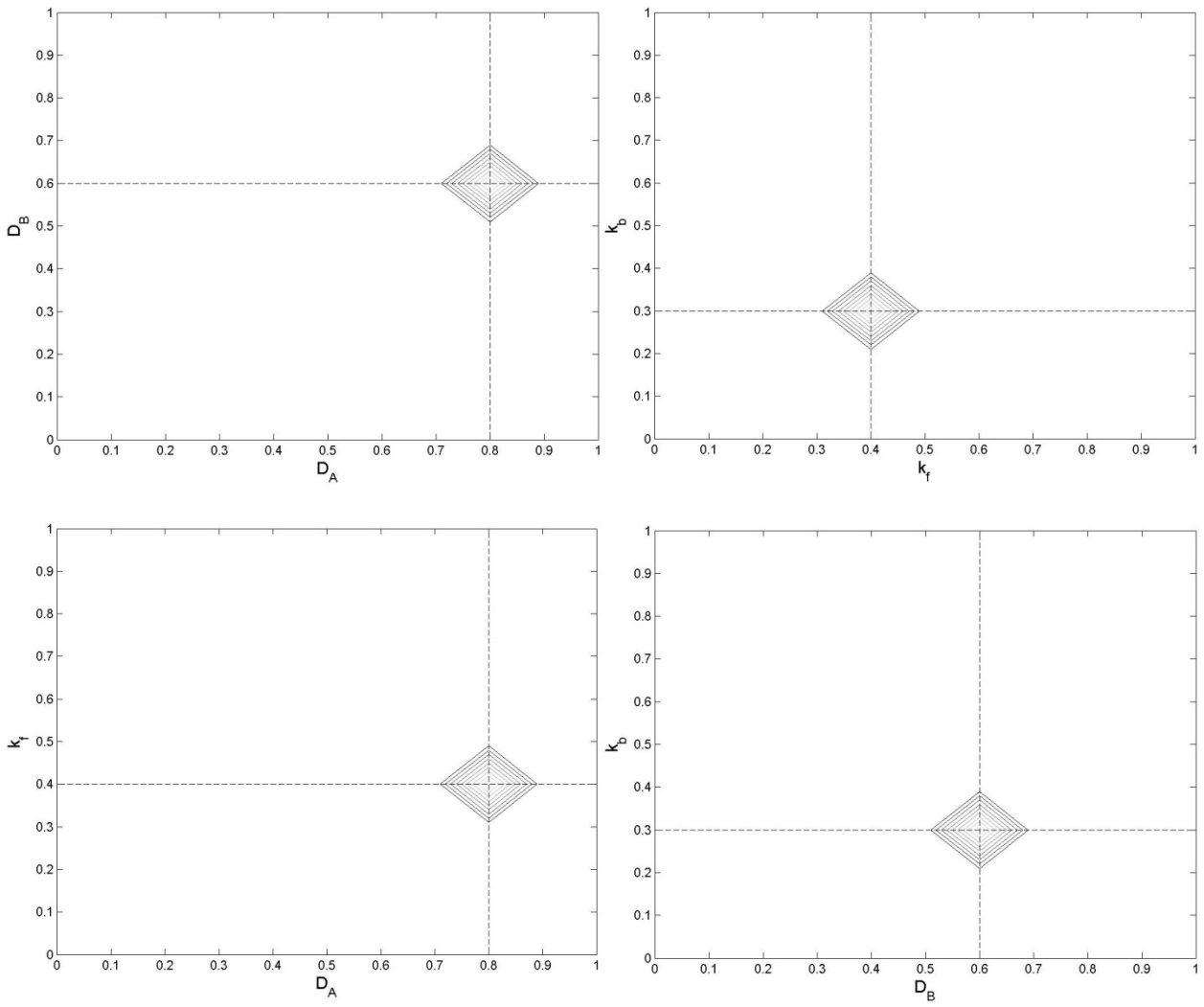

Figure 2.3: Bivariate Posterior Density Contours

Table 2.2: Statistical Analysis of Posterior Density

	Expeted Values	Covariance			
		D_A	D_B	k_f	k_b
D_A	0.8000	0.0000e+00	0.0000e+00	0.0000e+00	0.0000e+00
D_B	0.6000	0.0000e+00	1.2326e-32	0.0000e+00	0.0000e+00
k_f	0.4000	0.0000e+00	0.0000e+00	0.0000e+00	0.0000e+00
k_b	0.3000	0.0000e+00	0.0000e+00	0.0000e+00	0.0000e+00

An alternative experimental procedure, called reduced experiment α, is proposed that records the concentration at each of the five sampling location each second for only five seconds using an initial condition where each concentration of the previous case reduced by a factor of five. The results of such a procedure are shown in Figure 2.4 and Table 2.3. Figure 2.4 and Table 2.3 show that the confidence in the mass diffusivity of species B has been reduced and that the expected value has shifted away from the true value; however, the cost of the experiment may have been significantly reduced. To gain an understanding of the complexities of the method another experimental procedure is proposed; this one termed reduced experiment β. Here, the reduced concentrations of each species are regionally interchanged, i.e. Region 1 and 3: 4 mol-cm^{-1} of species A and 1 mol-cm^{-1} of species B, Region 2: 2 mol-cm^{-1} of species A and 6 mol-cm^{-1} of species B. Figure 2.5 and Table 2.4 convey the results of this procedure. Here, it can be seen that the confidence in the mass diffusivity of species B has been reduced and its expected value has shifted away from the true value. The results of the two reduced experiments suggest that quality of parameter estimation in the mass diffusivities is sensitive to the initial concentration. Now, suppose an severely ill-advised experimental procedure is proposed where the reaction vessel size is reduced to 1 cm, the concentration at five sampling locations is recorded every 0.5 milliseconds for 10 milliseconds, and an initial condition of: Region 1 and 3: 0.25 mol-cm^{-1} of species A and 1 mol-cm^{-1} of species B, Region 2: 1 mol-cm^{-1} of species A and 0.75 mol-cm^{-1} of species B. The partitions are placed at 0.3 cm from the ends of the vessel. The results of the ill-advised experiment may be seen in Figure 6 and Table 5. Inspection of both the Figure 2.6 and Table 2.5 reveal a significant correlation between the forward and reverse reaction rates. It can also be seen that the confidence in all of the parameters is relatively low; however, the diffusivity means are close to their true values.

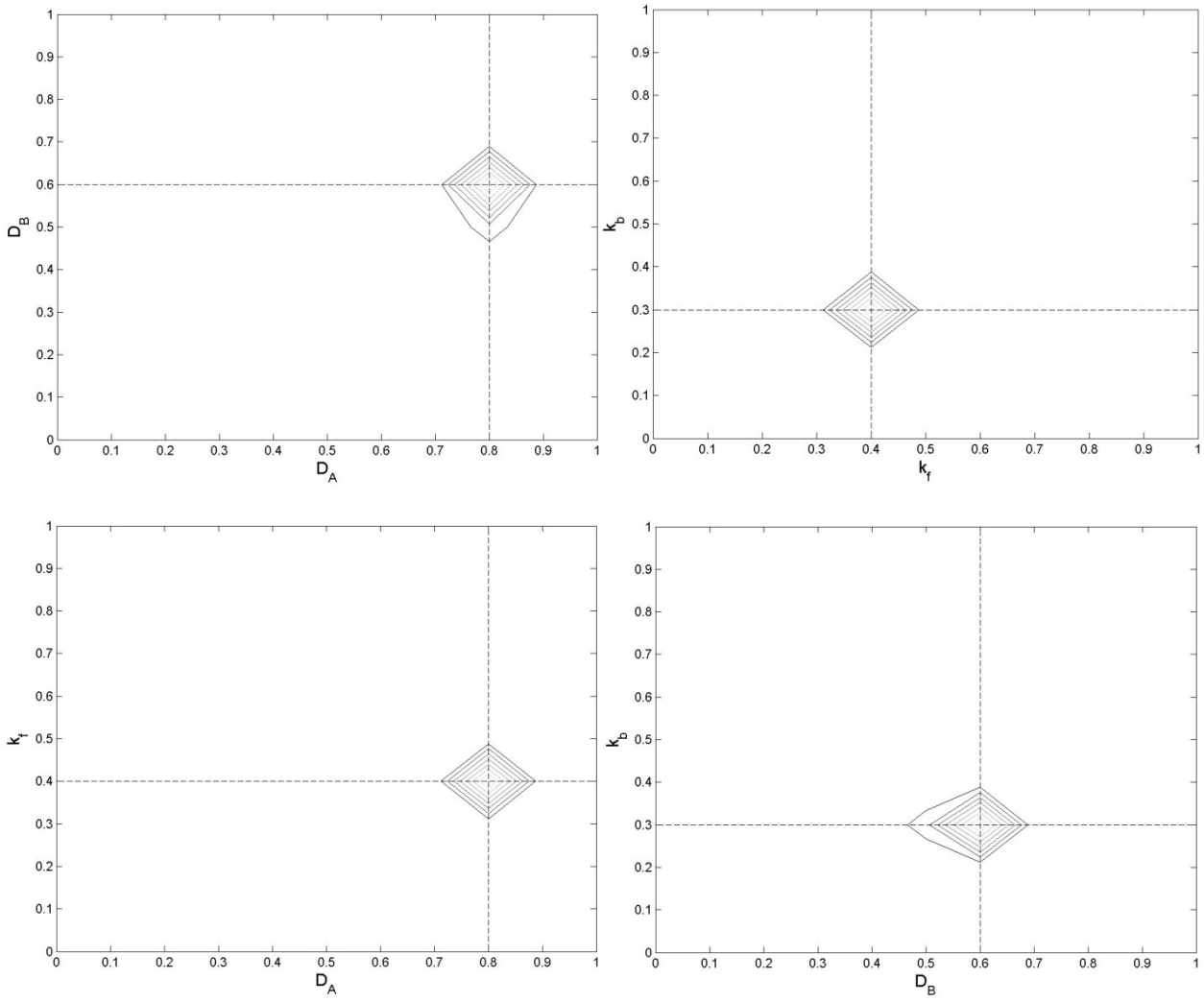

Figure 2.4: Bivariate Posterior Density Contours of Reduced Experiment α

Table 2.3: Statistical Analysis of Posterior Density for Reduced Experiment α

	Expected Values	Covariance			
		D_A	D_B	k_f	k_b
D_A	0.8000	7.4244e-39	-1.0410e-38	-4.0243e-54	0.0000e+00
D_B	0.5861	-1.0410e-38	1.4172e-03	-7.0231e-33	0.0000e+00
k_f	0.4000	-4.0243e-54	-7.0231e-33	3.0815e-33	0.0000e+00
k_b	0.3000	0.0000e+00	0.0000e+00	0.0000e+00	0.0000e+00

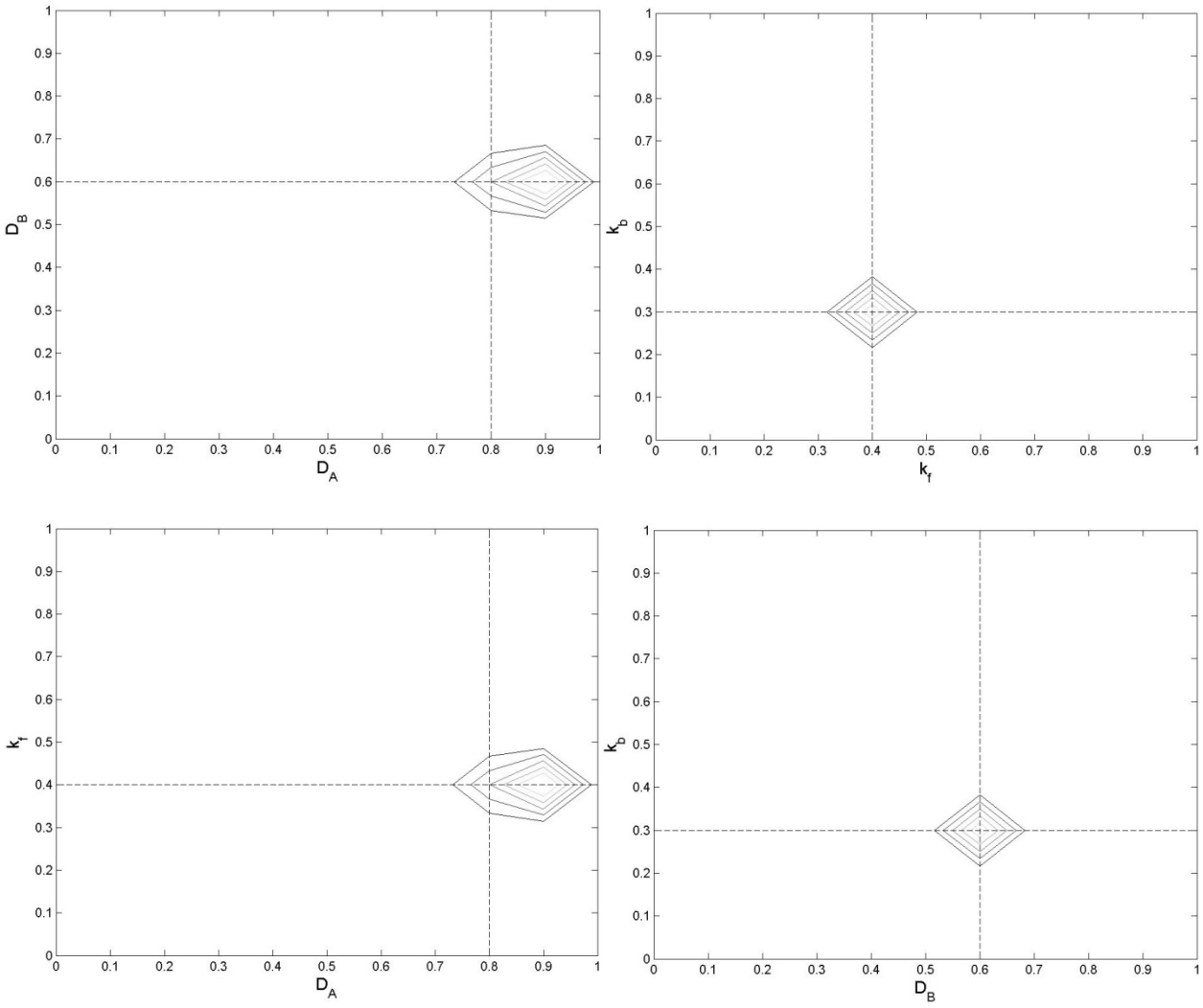

Figure 2.5: Bivariate Posterior Density Contours of Reduced Experiment β

Table 2.4: Statistical Analysis for Reduced Experiment β

	Expected Values	Covariance			
		D_A	D_B	k^f	k^b
D_A	0.8710	2.3422e-03	1.6383e-65	0.0000e+00	0.0000e+00
D_B	0.6000	1.6383e-65	1.9549e-65	0.0000e+00	0.0000e+00
k^f	0.4000	0.0000e+00	0.0000e+00	0.0000e+00	0.0000e+00
k^b	0.3000	0.0000e+00	0.0000e+00	0.0000e+00	0.0000e+00

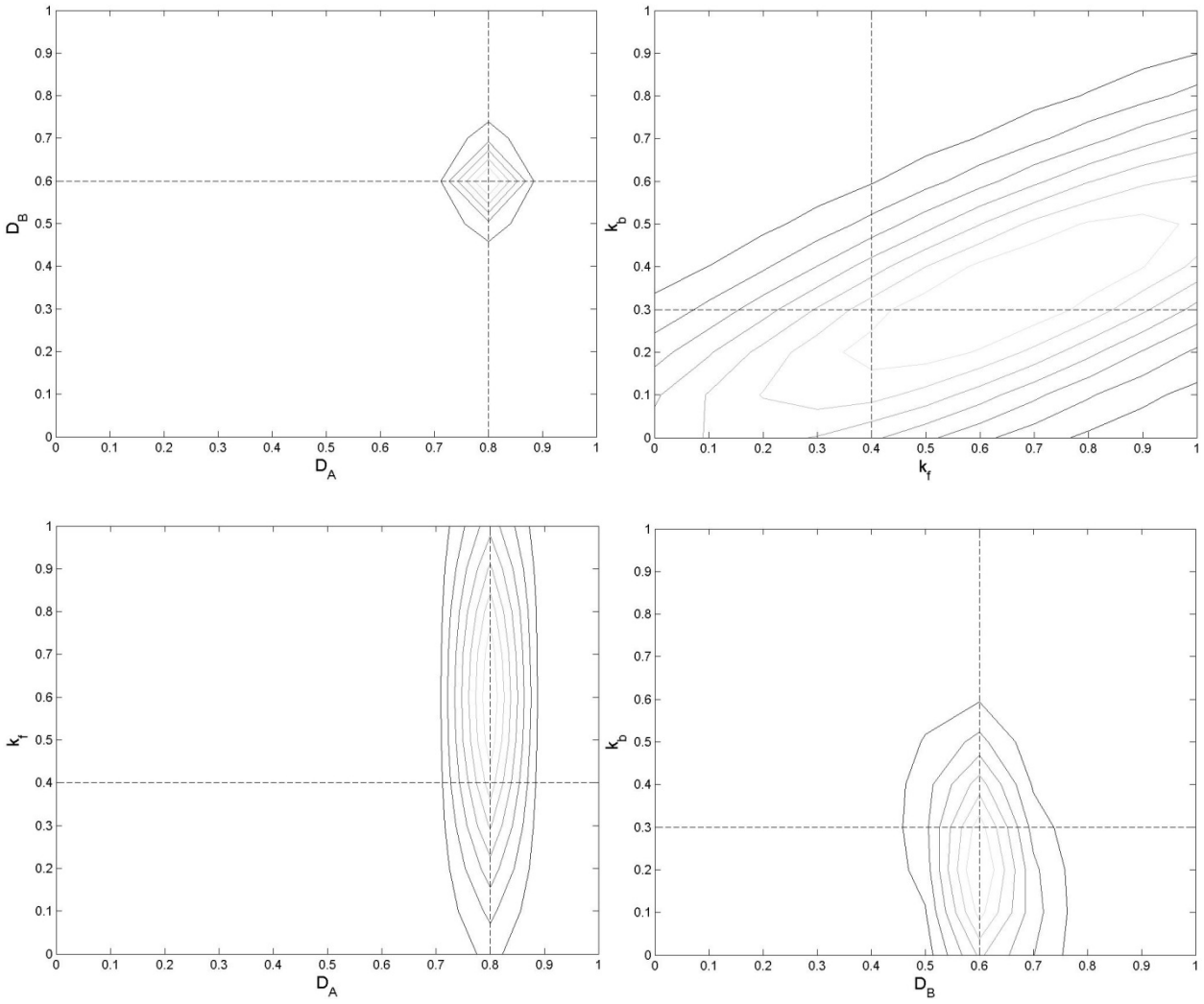

Figure 2.6: Bivariate Posterior Density Contours of Ill-Advised Experiment

Table 2.5: Statistical Analysis of Posterior Density for Ill-Advised Experiment

	Expected Values	Covariance			
		D_A	D_B	k^f	k^b
D_A	0.7921	7.2814e-04	1.5943e-04	-1.0349e-03	-1.8908e-03
D_B	0.5981	1.5943e-04	3.8386e-03	-1.2699e-03	-3.6357e-03
k^f	0.5849	-1.0349e-03	-1.2699e-03	6.9132e-02	2.9621e-02
k^b	0.3594	-1.8908e-03	-3.6357e-03	2.9621e-02	4.6111e-02

This exercise has shown the effect that experimental design has on the estimation of model parameters. The confidence in the estimation of the mass diffusivities is related to the initial concentration distribution in the containment vessel; specifically, this occurrence is related to the initial concentration gradients in each species. If the initial regional gradient is low, the quality of the estimate in mass diffusivity suffers. The ill-advised experiment shows the relation between the estimation of the reaction rates and the temporal spacing of the data collection.

2.5.1 Quantification of Computational Cost

At each discrete model vector the forward model must be evaluation for calculation of the nodal likelihood. In the case of the reaction-diffusion problem, the forward model comes in the form of a time-marching numerical procedure which requires a reasonably small temporal step-size to ensure negligible model error. Furthermore, this numerical procedure requires the solution to a linear system at each time step. The totality of the computational expense associated with forward model evaluation limits the discretization of the model space. Here, a modest grid of 11 nodes in each parameter results in 14641 evaluations of the BTCS algorithm. The spatial discretization used in the BTCS algorithm contains 41 nodes. The temporal discretization for the BTCS procedure contains 121 nodes, implying that an 82×82 linear system is solved 120 times for each of 14641 model vectors. The dimensionality of the problem indicates that increases in the resolution of the model space result in exponential increases in the number of computations. What can be seen from this exercise is the method of application associated with numerical Bayesian inversion as well as the correlation between experimental design and quality of parameter estimation.

3.0 THE ARRHENIUS INVERSE PROBLEM

3.1 MOTIVATION

Chemical reactor engineering is a fundamental facet in the planning and construction of chemical and energy production facilities. This area of study has application in industries such as commercial power generation, pharmaceutical manufacture, and petroleum refining. The use of computational physics modeling has greatly accelerated the reactor design process for throughput optimization and process safety; however, the results provided by such methods are only as reliable as the experimental inputs used in their generation. Inherent to these computational simulations is the use of experimentally determined parameters of individual simulation sub-models. Uncertainty quantification techniques provide a means for understanding the confidence that may be placed in the results of a simulation and may be used to modify experimental design for improved model parameter estimation. One such area of chemical reactor design is the assessment of rate law expressions and the determination of Arrhenius equation parameters as chemical kinetics play a pivotal role in process planning for operation efficiency and design safety. Design facets such as process throughput, temperature control, as well as pressure vessel mechanics are all related in some way to the rate of production and consumption of various chemical species, making kinetics modeling and uncertainty quantification a key area of study.

3.2 THE ARRHENIUS EQUATION

The Arrhenius equation, Equation (19), is a mathematical expression intended to model the temperature dependence of reaction rates.

$$k(T) = k_0 e^{\left(\frac{-E_A}{\tilde{R}T}\right)} \tag{20}$$

The solution to the Arrhenius inverse problem is a state of information pertaining to activation energy, and the pre-exponential factor, obtained from observations concerning specific reaction rate and temperature. The Arrhenius inverse problem presents a challenge in that the specific rate of reaction is not a directly observable quantity. Furthermore, the concept of the specific rate of reaction is trivial unless placed in the context of a phenomenologically developed rate law expression as the specific rate constant serves as a factor of proportionality for the rate law mathematical model. This means that the Arrhenius inverse problem requires data in the form of solutions to rate law expression inverse problems. Here, a sequential method for the Bayesian formulation of the Arrhenius inverse problem treatment is presented.

3.3 ARRHENIUS INVERSE PROBLEM FOR AN ELEMENTARY REACTION

3.3.1 Development of the First-Order Integrated Rate Law Expression

Consider the elementary chemical reaction:

$$A \rightarrow B \tag{21}$$

It is believed that this chemical reaction obeys first-order chemical kinetics, allowing the rate of production of B to be modeled using the power rate law expression:

$$r_B = kC_A \tag{22}$$

Accounting for the stoichiometry of Reaction (20), the rate of production of B may be written as a concentration differential with respect to time as:

$$r_B = -\frac{dC_A}{dt} \tag{23}$$

Substitution of Equation (21) into (22) yields:

$$-\frac{dC_A}{dt} = kC_A \tag{24}$$

Equation (23) is a first-order initial value problem with particular solution:

$$C_A(t) = C_{A,0}e^{-kt} \tag{25}$$

Equation (24) is a mathematical model attempting to predict the concentration of A over time at isothermal conditions in terms of the initial concentration and specific rate of reaction. This expression is referred to as the Integrated Rate Law (IRL).

The sequential approach presented here involves the application of Bayesian inversion to the Integrated Rate Law model using isothermal concentration-time data. This will result in a state of information concerning the specific rate of reaction associated with each temperature level. These states of information will then be interpreted as data in the subsequent application of Bayesian inversion to the Arrhenius equation. The specific of the utilization of IRL posterior densities as data for the Arrhenius inverse problem will be discussed in detail in section 3.3.3.

3.3.2 Bayesian Inversion of Integrated Rate Law Expression

In the context of the Bayesian inverse problem, specific reaction rate and initial concentration are taken to be model parameters while concentration vs. time measurements are treated as the data. The initial concentration is treated as a model parameter in order to capture uncertainty in its

transduction while maintaining the ability to evaluate the forward problem, a necessity in discrete computation. Uncertainty concerning initial concentration is incorporated in the formulation of the prior probability density. For the development of this procedure the prior probability densities in both initial concentration and specific rate of reaction are taken to be uniform described by:

$$\psi_{IRL}^{i}(m^i) = \begin{cases} \dfrac{1}{m^{i,max} - m^{i,min}} & \forall\, m^i \in [m^{i,min}, m^{i,max}] \\ 0 & otherwise \end{cases} \tag{26}$$

These relative priors are taken to be statistically independent, allowing for their combination through products by:

$$\psi_{IRL}(\boldsymbol{m}) = \prod_i \psi_{IRL}^{i}(m^i) \tag{27}$$

Suppose that the mass-spectrometer used for concentration measurement is believed to obey Gaussian uncertainty with known variance. The concentration is sampled at a single location over some period of time. The relative likelihoods according to this experimental procedure are given by:

$$\lambda_{IRL}^{n}(C_A^n|\boldsymbol{m}) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{(C_A^n - \bar{C}_A^n)^2}{2\sigma^2}\right] \tag{28}$$

These statistically independent relative likelihoods may be combined through products by:

$$\lambda_{IRL}^{n}(\boldsymbol{d}|\boldsymbol{m}) = \prod_n \lambda_{IRL}^{n}(C_A^n|\boldsymbol{m}) \tag{29}$$

Following the determination of the prior density and likelihood over the model space, the posterior probability density may be constructed through the use of Bayes' theorem by:

$$\eta_{IRL}(\boldsymbol{m}|\boldsymbol{d}) = \frac{\lambda_{IRL}(\boldsymbol{d}|\boldsymbol{m})\psi_{IRL}(\boldsymbol{m})}{\int_{\mathbb{M}_{IRL}} \lambda_{IRL}(\boldsymbol{d}|\boldsymbol{m})\psi_{IRL}(\boldsymbol{m})d\boldsymbol{m}} \tag{30}$$

3.3.3 Bayesian Inversion of the Arrhenius Equation

The Arrhenius inverse problem involves the determination of the activation energy and pre-exponential factor from specific rate constant and temperature data; however, the specific rate constant is not a directly transducible quantity. This presents a challenge in the formulation of a likelihood expression for the Arrhenius inverse problem as there is no measurement device specification or calibration to provide information concerning the measurement uncertainty in the isothermal specific rate constants. However, the discrete posterior probability densities associated with individual isothermal IRL inverse problems capture this uncertainty. It follows that these isothermal posterior probability densities may be treated as relative likelihoods in the formulation of the Arrhenius inverse problem. Let $\{\eta_{IRL}^{j}(k, C_{A,0})\}_{j=1}^{J}$ be a set of discrete, isothermal posterior probability densities resulting from concentration-time experiments conducted at J distinct temperature levels. Marginalizing these densities over initial concentration results in a set of discrete, univariate posterior probability densities $\{\zeta^{j}(k)\}_{j=1}^{J}$. These marginalized IRL posteriors capture the uncertainty in the experimental measurement technique while providing a probabilistic representation of the compatibility of the concentration-time data and individual Arrhenius model vectors. This comparison accomplished by treating the result of the forward Arrhenius model at isothermal model vectors as the corresponding posterior density's argument. Each $\zeta^{j}(k)$ contains information concerning the confidence which may be placed in the experimental process; however, instead of being a continuous function obtained from some instrument specification or calibration, the information comes in the form of a discrete probability density. To elaborate, consider the relative likelihood expression associated with the IRL problem. The relative likelihoods are expressed as a

continuous, Gaussian probability density functions with IRL model concentration as their arguments and means specified by the experimental data. For the IRL inverse problem, the data come in the form of discrete concentration values. In the case of the Arrhenius inverse problem, there is no means to obtain discrete data values of the specific rate of reaction, only a probabilistic interpretation; however, what Equation (28) accomplishes for the IRL inverse problem, the set of discrete posterior densities, $\{\zeta^j(k)\}_{j=1}^{J}$, accomplishes for the Arrhenius inverse problem. The primary complicating factor is that each $\zeta^j(k)$ is not a continuous function. Each isothermal relative likelihood, $\zeta^j(k)$, is only known for distinct values of k due to the discrete nature of the method of computation, i.e. the discretization of the specific rate constant model space coordinate. The forward Arrhenius model may return values of k not used as nodes in the IRL inverse problem, requiring the use of an interpolation technique to allow for the computation of relative likelihoods for any value of k the forward Arrhenius model may produce. Here a mid-point step function approximation is used for this interpolation. Figure 3.1 displays this technique.

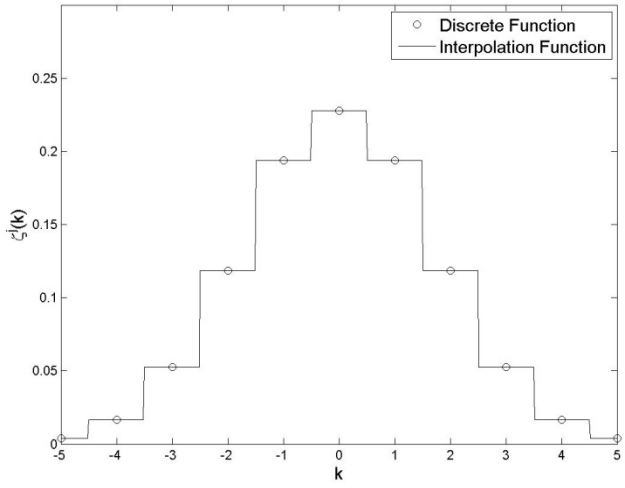

Figure 3.1: Arrhenius Likelihood Interpolation Technique

Each isothermal experiment used in the generation of each $\zeta^j(k)$ is statistically independent allowing the combination of the relative likelihoods by:

$$\lambda_{AR}(\boldsymbol{d}|\boldsymbol{m}) = \prod_j \zeta^j(k) \tag{31}$$

This likelihood expression coupled with appropriate prior information may be combined via Bayes' theorem by:

$$\eta_{AR}(\boldsymbol{m}|\boldsymbol{d}) = \frac{\lambda_{AR}(\boldsymbol{d}|\boldsymbol{m})\psi_{AR}(\boldsymbol{m})}{\int_{\mathbb{M}_{AR}} \lambda_{AR}(\boldsymbol{d}|\boldsymbol{m})\psi_{AR}(\boldsymbol{m})d\boldsymbol{m}} \tag{32}$$

3.3.4 Sequential Inverse Problem Numerical Implementation

3.3.4.1 Integrated Rate Law Inverse Problem

Suppose that J isothermal concentration vs. time experiments are conducted resulting in J concentration-time data sets. For simplicity of this procedural development, the relative priors for each of these J inverse problems are taken to be uniform in both the specific rate constant and initial concentration coordinates. Due to the uniform nature of the relative priors, the model spaces associated with the individual isothermal concentration time inverse problems may be given by:

$$\mathbb{M}_{IRL}^j = \left[k^{j,min}, k^{j,max}\right] \times \left[C_{A,0}^{j,min}, C_{A,0}^{j,max}\right]$$

Let $P^j, Q^j \in \mathbb{N}$ denote the number of nodes in each model coordinate. Note that the discretization for each concentration-time inverse problem need not be compatible with the others, allowing for greater control of the resolution of each individual inverse problem. It is to be understood that J distinct discretizations must be constructed. The parameter step sizes are defined by:

$$\Delta k^j = \frac{(k^{j,max} - k^{j,min})}{(P^j - 1)}, \Delta C_{A,0}^j = \frac{(C_{A,0}^{j,max} - C_{A,0}^{jmin})}{(Q^j - 1)}$$

Individual grid points may then be described by:

$$k^{j,p} = (p-1)\Delta k^j, C_{A,0}^{j,q} = (q-1)\Delta C_{A,0}^j$$

These definitions allow for the statement of the model grid by:

$$\{(k^{j,p}, C_{A,0}^{j,q}) : 1 \leq p \leq P^j, 1 \leq q \leq Q^j\}$$

This grid may be computationally navigated in a systematic fashion through the use of two nested for loops. In the case of uniform priors, the prior probability is a constant value over the entirely of the model grid, implying that it need only be calculated once. The likelihood must be computed at each recursion of the nested loop structure. The relative likelihoods are computed according to Equation (27) and are combined according to Equation (28). Due to the successive nature of discrete computation, the posterior probabilities may not be normalized until each model vector of the grid has been accessed. The non-normalized posteriors are determined by:

$$\tilde{\eta}_{IRL}^j(k^{j,p}, C_{A,0}^{j,q}|\boldsymbol{d}) = \lambda_{IRL}^j(\boldsymbol{d}|k^{j,p}, C_{A,0}^{j,q})\psi_{IRL}^j(k^{j,p}, C_{A,0}^{j,q})$$

Upon solution of the non-normalized posterior probability the distribution is normalized by the normalization constant K, given by:

$$K^j = \int_{\mathbb{M}_{IRL}^j} \tilde{\eta}_{IRL}^j(\boldsymbol{m}|\boldsymbol{d}) d\boldsymbol{m}$$

This integral may be expanded to:

$$K^j = \int_{C_{A,0}} \int_{k} \tilde{\eta}_{IRL}^j(k, C_{A,0}|\boldsymbol{d}) dk dC_{A,0}$$

Fubini's theorem allows for this integral to be evaluated through the successive computation of one-dimensional integrals over each parameter coordinate by:

$$\tilde{\eta}'^{j}_{IRL}(k|\boldsymbol{d}) = \int_{C_{A,0}^{j,min}}^{C_{A,0}^{j,max}} \tilde{\eta}^{j}_{IRL}\left(k, C_{A,0}\middle|\boldsymbol{d}\right) dC_{A,0}$$

$$K^{j} = \int_{k^{j,min}}^{k^{j,max}} \widetilde{\eta}'^{j}_{IRL}(k|\boldsymbol{d}) dk$$

Numerically, the preceding succession of integrals is computed using trapezoidal quadrature. The composite trapezoidal quadratures for the computation of the normalization constant may be written as:

$$\tilde{\eta}'^{j}_{IRL}(k^{j,p}|\boldsymbol{d}) = \frac{\Delta C_{A,0}^{j}}{2} \sum_{q=1}^{Q-1} (\tilde{\eta}^{j}_{IRL}(k^{j,p}, C_{A,0}^{j,q}|\boldsymbol{d}) + \tilde{\eta}^{j}_{IRL}(k^{j,p}, C_{A,0}^{j,q+1}|\boldsymbol{d}))$$

$$K^{j} = \frac{\Delta k^{j}}{2} \sum_{p=1}^{P-1} (\tilde{\eta}'^{j}_{IRL}(k^{j,p}|\boldsymbol{d}) + \tilde{\eta}'^{j}_{IRL}(k^{j,p+1}|\boldsymbol{d}))$$

These expressions are numerically evaluated through the use of nested for loops for the computation of the sum operator. Following the calculation of the normalization constant the raw posterior values are normalized to give a discrete probability density by:

$$\eta^{j}_{IRL}\left(k, C_{A,0}\middle|\boldsymbol{d}\right) = \frac{1}{K^{j}} \lambda^{j}_{IRL}\left(\boldsymbol{d}\middle|k, C_{A,0}\right) \psi^{j}_{IRL}\left(k, C_{A,0}\right)$$

Each of the isothermal IRL posteriors must now be marginalized for use as the relative likelihoods in the subsequent Arrhenius inverse problem. The marginal density is found using:

$$\zeta^{j}(k) = \int_{C_0} \eta^{j}_{IRL}(k, C_{A,0}) dC_{A,0} \tag{33}$$

Computationally this task may be accomplished through use of trapezoidal quadrature

$$\zeta^{j}\left(k^{j,p}\right) = \frac{\Delta C_{A,0}^{j}}{2} \sum_{q=1}^{Q-1} \eta^{j}_{IRL}(k^{j,p}, C_{A,0}^{j,q}) + \eta^{j}_{IRL}(k^{j,p}, C_{A,0}^{j,q+1})$$

3.3.4.2 Arrhenius Inverse Problem

With each of the discrete marginalized IRL posteriors generated, the Arrhenius inverse problem may now be numerically formulated. As before, for simplicity the prior information in activation energy and pre-exponential factor are taken to be uniform, allowing the Arrhenius inverse problem model space to be defined by:

$$\mathbb{M}_{AR} = \left[E_A^{min}, E_A^{max}\right] \times \left[k_0^{min}, k_0^{max}\right]$$

Let $R, S \in \mathbb{N}$ denote the number of nodes in each Arrhenius parameter coordinate, allowing for the definition of parameter step sizes by:

$$\Delta E_A = \frac{(E_A^{max} - E_A^{min})}{(R-1)}, \Delta k_0 = \frac{(k_0^{max} - k_0^{min})}{(S-1)}$$

Individual coordinate grid points may then be described by:

$$E_A^r = (r-1)\Delta E_A,\ k_0^s = (s-1)\Delta k_0$$

These definitions allow for the statement of the model grid by:

$$\{(E_A^r, k_0^s) : 1 \leq r \leq R, 1 \leq s \leq S\}$$

This grid may be computationally navigated in a manner similar to that in the concentration time inverse problem. The interesting point made here is the use of $\zeta^j(k)$ as the relative likelihoods. The model will be evaluated at each nodal model vector, at each temperature level. This will result in a value of specific rate constant for each model vector at each temperature level. Inherent to this approach is the assumption that the uncertainty in temperature transduction is negligible, an experimentally valid assumption considering the accuracy of modern thermocouples. The likelihood is taken to be the value of $\zeta^j(k)$ corresponding to each model value of the specific rate constant. The discrete nature of $\zeta^j(k)$ does not allow for the determination of a likelihood for any value of specific rate constant produced by the model, requiring the use of the aforementioned midpoint interpolation technique.

3.3.5 Sequential Versus Direct Arrhenius Inverse Problem Formulation

The sequential method presented here is not the only mathematically sound method of Bayesian inversion that may be applied to the Arrhenius inverse problem. The problem may be directly formulated by substituting the IRL expression into the Arrhenius equation resulting in:

$$C_A = C_{A,0}\exp\left[-tk_0\left(\frac{-E_A}{\tilde{R}T}\right)\right] \tag{34}$$

The selection of the sequential formulation is made out of a desire for computational tractability. Consider first the sequential method, where $P, Q, R, S \in \mathbb{N}$ denote the number of nodes used in the discretization of the specific rate constant, initial concentration, activation energy, and pre-exponential factor coordinates respectively. Here, the number of posterior computations is $P \times Q$ for each temperature level resulting in $J(P \times Q)$ posterior computations for the IRL inverse problems. The Arrhenius inverse problem requires $R \times S$ posterior computations bringing the total number of posterior computations to $J(P \times Q) + R \times S$. In the direct formulation of the Arrhenius inverse problem each isothermal initial concentration adds a dimension to the model space. This coupled with the dimensions from the activation energy and pre-exponential factor, while taking the discretisations in each initial concentration to be equal, results in $Q^J \times R \times S$ posterior computations for the direct formulation. Inspection of these number of posterior computation expressions shows that the number of posterior computations in the direct formulation grows exponentially with the number of temperatures while the number of posterior computations in the sequential formulation grows multiplicatively with the number of temperature levels. Furthermore, the relationship between the sequential problems is additive while the relationship is multiplicative for the direct formulation. The stark contrast in the number of posterior complications is best conveyed through use of an example. Given a modest

discretization of the model spaces, taking $P = 101, Q = 101, R = 101$, and $S = 101$, with a reasonable number of temperature levels, $J = 5$, results in 61206 posterior computations for the sequential case and 1.0721e+14 posterior computations for the direct case. The direct formulation of the Arrhenius inverse problem quickly becomes computationally large. Increases in the resolution of the sequential formulation increase the number of posterior computations at a significantly more modest pace. While each formulation will vary in actual number of floating point operations in a way not entirely described by these formulations due to differences in specific implementation, these estimates provide sufficient information to clearly show that the direct formulation will drastically increase the total number of computational operations.

3.4 CHEMICAL KINETICS OF BENZENE DIAZONIUM CHLORIDE DECOMPOSTION

3.4.1 The Decomposition Reaction and Artificial Experiment

An aqueous solution of benzene diazonium chloride (BDC) will decompose yeilding aqueous benzene chloride and nitrogen gas in accordance with a first order power law expression. Suppose that isothermal concentration-time experiments are conducted in a stirred batch reactor; sized such that concentration and temperature gradients may be considered negligible. On start-up, the reactor is filled with a 0.1 M aqueous solution of BDC. The reactor is then heated. Upon equilibrating at the desired temperature level, the concentration of BDC is recorded over time using a mass spectrometer whose uncertainty is believed to be Gaussian with known variance. The concentration is recorded every minute for a total time of ten minutes. This procedure is performed for five temperature levels: 313 K, 319 K, 323 K, 328 K, and 333 K.

3.4.2 Numerical Generation of Concentration vs. Time Data

Concentration-time data is generated from sequentially evaluation of the Arrhenius and IRL forward problems, followed by a random, normal perturbation of the concentration values in accordance with the uncertainty of the mass spectrometer. The true values of activation energy and pre-exponential factor are taken from an example in *Fogler*: E_a= 116.5 kJ-mol^{-1}, k_0=7.20e17 min^{-1} [13]. The Arrhenius parameters, along with the aforementioned temperate levels are used to generate five values of the specific rate constant. The IRL forward problem requires a specific rate constant and initial concentration. It is expected that the 0.1M solution will decompose during the heat up phase of the experiment. The initial concentration for the 323 K experiment was taken from an example in *Hill* [14]. The others were selected around this value in accordance with expected decomposition associated with equilibration time. The values of initial concentration and their corresponding temperatures are shown in Table 3.1

Table 3.1: True Values of Initial Concentration

T	**313 K**	**319 K**	**323 K**	**328 K**	**333 K**
$\mathbf{C_{BDC,0}}$	0.0750M	0.0700 M	0.0650 M	0.0600 M	0.0550 M

The specific rate constants from the forward Arrhenius problem, the initial concentration values and the time interval described in the experimental procedure allow for the generation of concentration-time data sets. These data sets are then randomly perturbed, including the initial concentration value, in accordance with the mass spectrometer uncertainty.

3.4.3 The Integrated Rate Law Inverse Problem

3.4.3.1 *A Priori* Information and the Likelihood Function

Because the initial concentration is recorded using a measurement device believed to obey Gaussian uncertainty with known variance, the prior probability in initial concentration may be described by a normal distribution of the form:

$$\psi^j\left(C_{BDC,0}\middle|\boldsymbol{m}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{(C^j_{BDC,0} - \bar{C}^j_{BDC,0})^2}{2\sigma^2}\right] \qquad (35)$$

For this numerical example the relative prior in specific rate constant is taken to be uniform, arbitrarily bounded by $\left[k^{j,true} - 0.25k^{j,true}, k^{j,true} + 0.15k^{j,true}\right]$ resulting in a prior probability in this interval given by:

$$\psi^j(k) = \frac{1}{(k^{j,true} + 0.15k^{j,true}) - (k^{j,true} - 0.25k^{j,true})} \qquad (36)$$

Relative likelihoods are described by Equation (27) and may be combined through Equation (28)

3.4.3.2 Numerical Resolution of Posterior Density

In the previous examples the relative prior probabilities have been taken to be uniform densities making the selection of the model space obvious. However, in this case the initial concentration is recorded using the same normal uncertainty measurement device as the experimental data, providing a prior in the form of a normal distribution of known variance and mean of the measured initial concentration. To capture an appropriate amount of the initial concentration prior, the model space is bounded four standard deviations, of the measurement Gaussian density, to the left and right of the measured initial concentration, allowing the formulation of a model space by:

$$\mathbb{M}_{IRL}^{j} = \left[k^{j,true} - 0.25k^{j,true}, k^{j,true} + 0.15k^{j,true}\right] \times \left[\bar{C}_{BDC,0}^{j} - 4\sigma, \bar{C}_{BDC,0}^{j} + 4\sigma\right]$$

Let $P^j, Q^j \in \mathbb{N}$ denote the number of nodes in each model coordinate. The parameter step sizes are defined by:

$$\Delta k^j = \frac{(k^{max} - k^{min})}{(P-1)}, \Delta C_{BDC,0}^{j} = \frac{\left[\left(\bar{C}_{BDC,0}^{j} + 4\sigma\right) - \left(\bar{C}_{BDC,0}^{j} - 4\sigma\right)\right]}{(Q-1)}$$

Individual grid points may then be described by:

$$k^{j,p} = (p-1)\Delta k^j, C_{BDC,0}^{j,q} = (q-1)\Delta C_{BDC,0}^{j}$$

These definitions allow for the statement of the model grid by:

$$\{\left(k^{j,p}, C_{BDC,0}^{j,q}\right): 1 \leq p \leq P, 1 \leq q \leq Q\}$$

This grid may be computationally navigated in a systematic fashion through the use of two nested for loops. Computation of the discrete posterior density is carried out using the MATLAB programming environment in a modular fashion. The prior information function reads in the measured initial concentration, the standard deviation of the measurement device, the true value of specific rate constant and the current nodal model vector. Because the relative prior in initial concentration is Gaussian, the prior probability function must be executed at each loop recursion. The likelihood function is also evaluated at each loop recursion in accordance with Equations (27) and (28). The remainder of the procedure for computation of the posterior probability density of the integrated rate law inverse problem may be found in sub-section 3.3.4.

3.4.3.3 Application and Results

The IRL inverse problem was solved at each temperature level first using un-perturbed data to validate the computational process and then using the randomly perturbed data. The variance of the concentration measurement device is taken to be 0.0005 M. The model space was discretized

using 1001 nodes in each coordinate to ensure sufficient resolution of the posterior for the subsequent application of the marginalized IRL posterior as the Arrhenius likelihood. Figure 3.2 depicts contours of the isothermal posteriors for the un-perturbed data with true value denoted by the black lines. Table 3.2 shows a comparison of the maximum a posteriori and expected value point estimators to the true value for the IRL posterior densities. It can be seen from Figure 3.2 that each of the isothermal posteriors is centered on the true value of each parameter. The program is quantifiably verified upon inspection of Table 3.2 as the parameters have been approximately recovered by the both point estimation techniques with some slight deviation in the expected value which may be attributed to the method of numerical computation. Figure 3.3 depicts the marginalized densities with the vertical line denoting the true value. Table 3.3 shows the results of the point estimation techniques as applied to the marginalized IRL posteriors. It can be seen that from comparison of Tables 3.2 and 3.3 that the same point estimates of specific rate constant are recovered from both the full IRL posterior and the marginalized IRL posterior.

Table 3.2: IRL Posterior Point Estimate Comparison (Un-Perturbed Case)

	T = 313 K		**T = 319 K**		**T = 323 K**		**T = 328 K**		**T = 333 K**	
	k	C_0	k	C_0	k	C_0	k	C_0	k	C_0
True	0.0287	0.0750	0.0665	0.0700	0.1146	0.0650	0.2219	0.0600	0.4215	0.0550
MAP	0.0287	0.0750	0.0665	0.0700	0.1146	0.0650	0.2219	0.0600	0.4215	0.0550
EV	0.0287	0.0750	0.0665	0.0700	0.1146	0.0650	0.2220	0.0600	0.4217	0.0550

Table 3.3: Marginalized IRL Posterior Point Estimate Comparison (Un-Perturbed Case)

	T = 313 K	**T = 319 K**	**T = 323 K**	**T = 328 K**	**T = 333 K**
	k	k	k	k	k
True	0.0287	0.0665	0.1146	0.2219	0.4215
MAP	0.0287	0.0665	0.1146	0.2219	0.4215
EV	0.0287	0.0665	0.1146	0.2220	0.4217

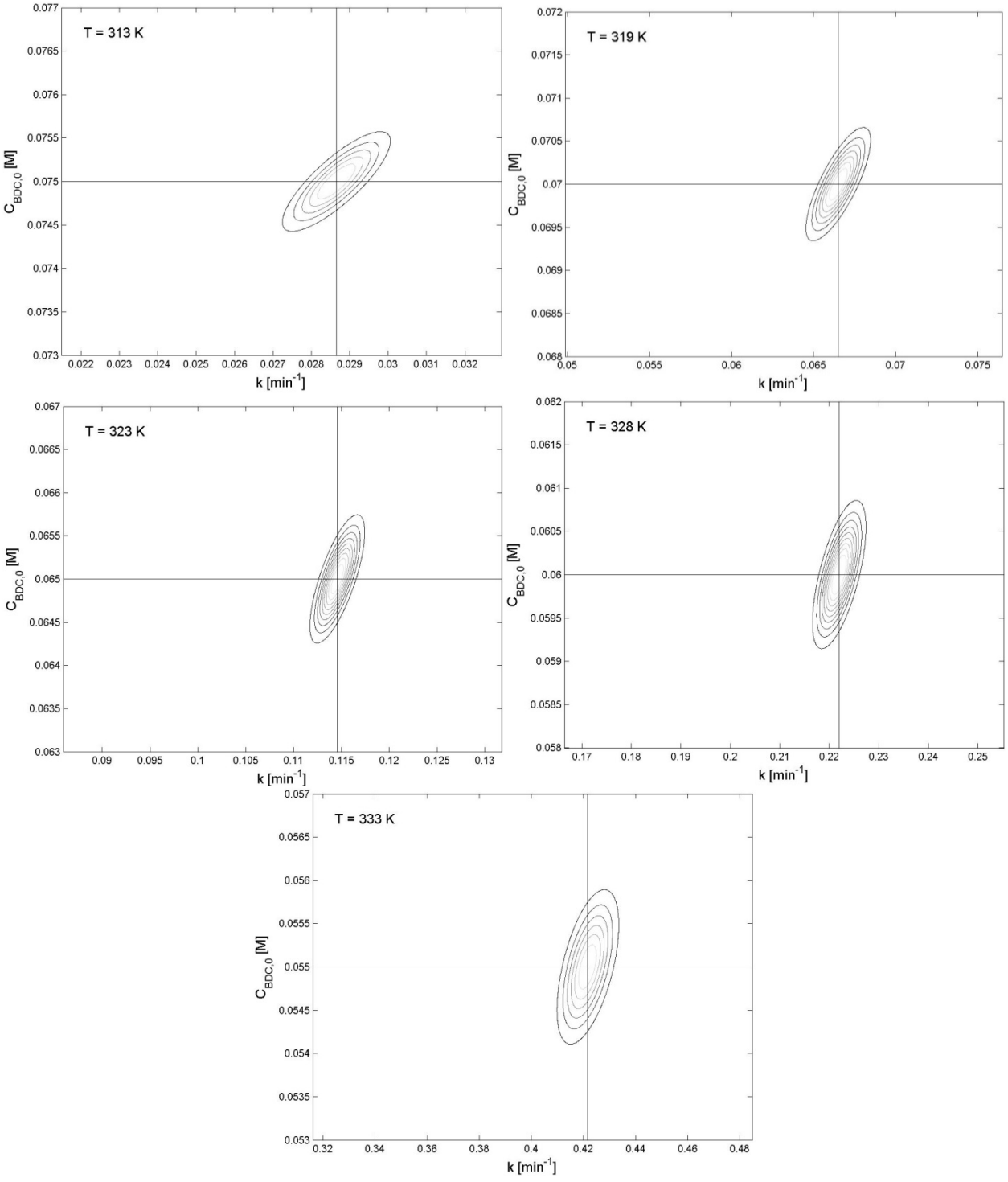

Figure 3.2: IRL Posterior Densities (Un-Perturbed Case)

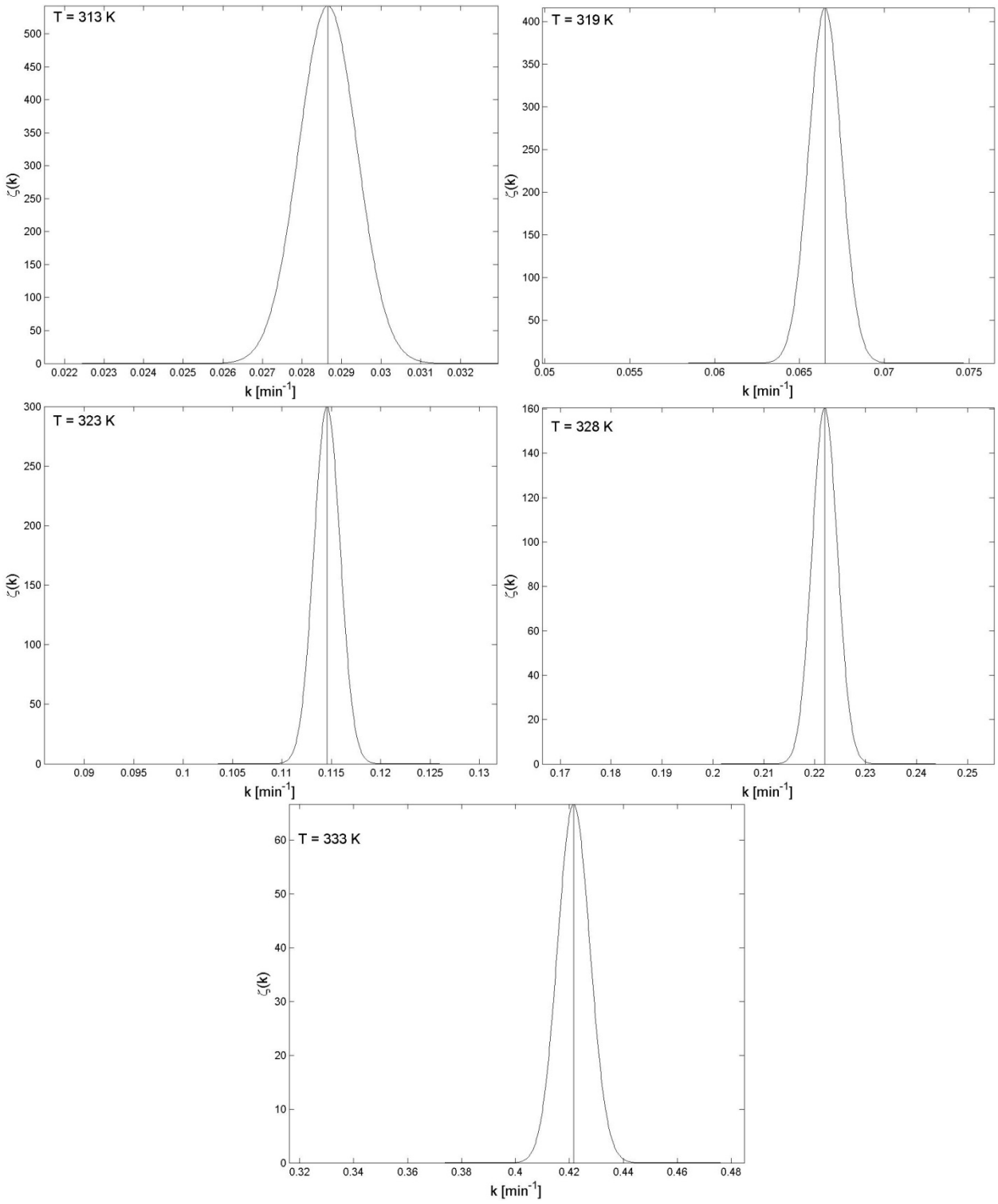

Figure 3.3: Marginalized IRL Posteriors (Un-Perturbed Case)

Figure 3.4 depicts contours of the IRL posteriors for the randomly perturbed data. It can be seen that the perturbation shifts the mean and maximum a posteriori of each distribution away from the true value of each parameter. Table 3.4 shows a comparison of the maximum a posteriori and expected value point estimators to the true value for the IRL posterior densities. The results presented in Table 3.4 quantify this shift away from the true values. Figure 3.5 depicts the marginalized densities for the perturbed IRL posteriors. Table 3.5 shows a comparison of the maximum a posteriori and expected value point estimators to the true value for the marginalized IRL posteriors. It can be seen that the mean and maximum a posteriori of the marginalized density is shifted away from the true value. These shifts are not surprising as the perturbation is expected to affect the quality of the estimates. Inspection Figures 3.4 shows that the true values still reside in a probable region of the IRL posterior.

Table 3.4: IRL Posterior Point Estimate Comparison (Perturbed Case)

	T = 313 K		**T = 319 K**		**T = 323 K**		**T = 328 K**		**T = 333 K**	
	k	C_0	k	C_0	k	C_0	k	C_0	k	C_0
True	0.0287	0.0750	0.0665	0.0700	0.1146	0.0650	0.2219	0.0600	0.4215	0.0550
MAP	0.0283	0.0751	0.0666	0.0698	0.1138	0.0651	0.2214	0.0598	0.4264	0.0543
EV	0.0283	0.0751	0.0666	0.0698	0.1138	0.0651	0.2215	0.0598	0.4266	0.0543

MAP: maximum a posteriori, EV: expected value

Table 3.5: Marginalized IRL Posterior Point Estimate Comparison (Perturbed Case)

	T = 313 K	**T = 319 K**	**T = 323 K**	**T = 328 K**	**T = 333 K**
	k	k	k	k	k
True	0.0287	0.0665	0.1146	0.2219	0.4215
MAP	0.0283	0.0666	0.1138	0.2214	0.4264
EV	0.0283	0.0666	0.1138	0.2215	0.4266

MAP: maximum a posteriori, EV: expected value

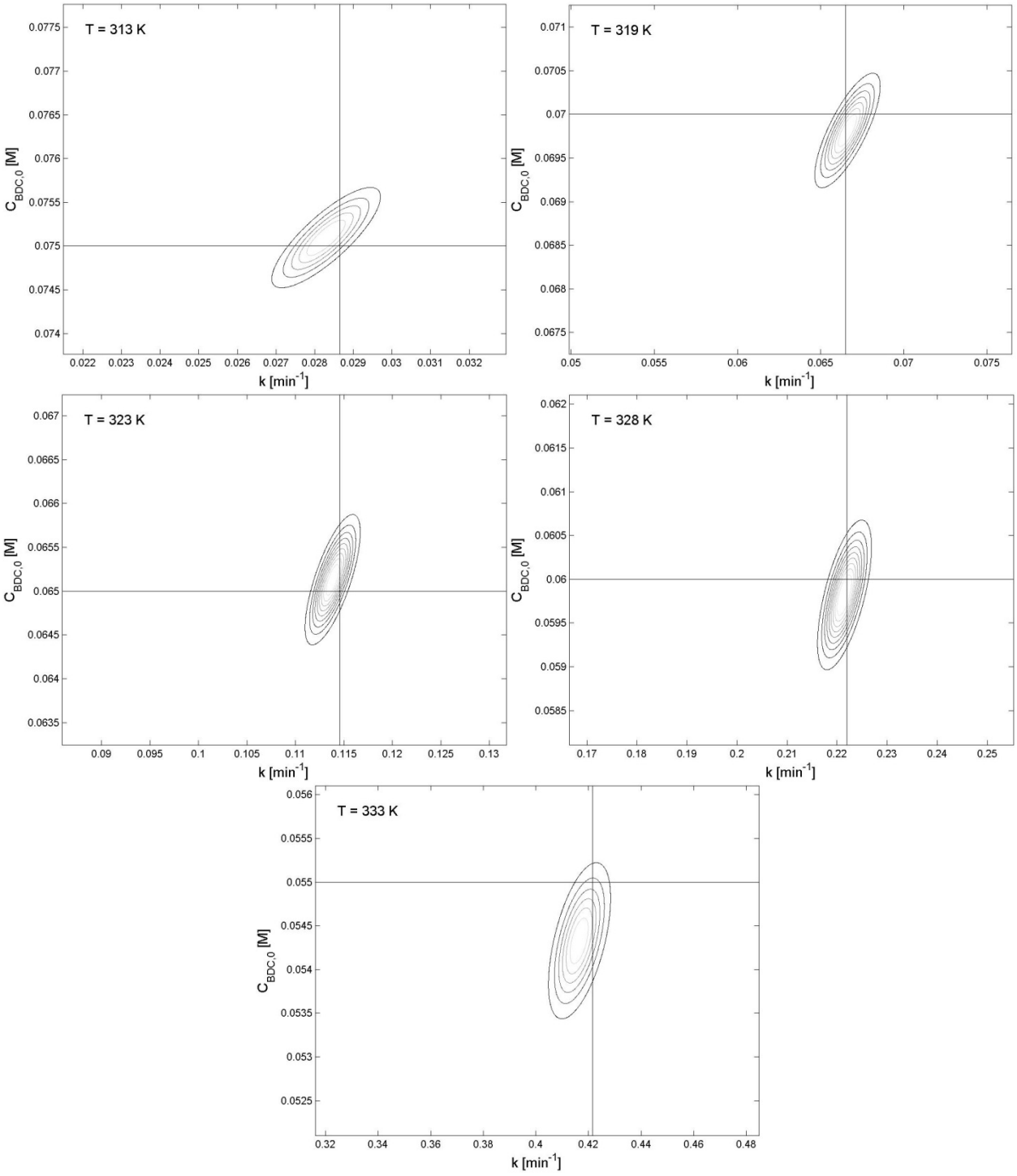

Figure 3.4: IRL Posterior Densities (Perturbed Case)

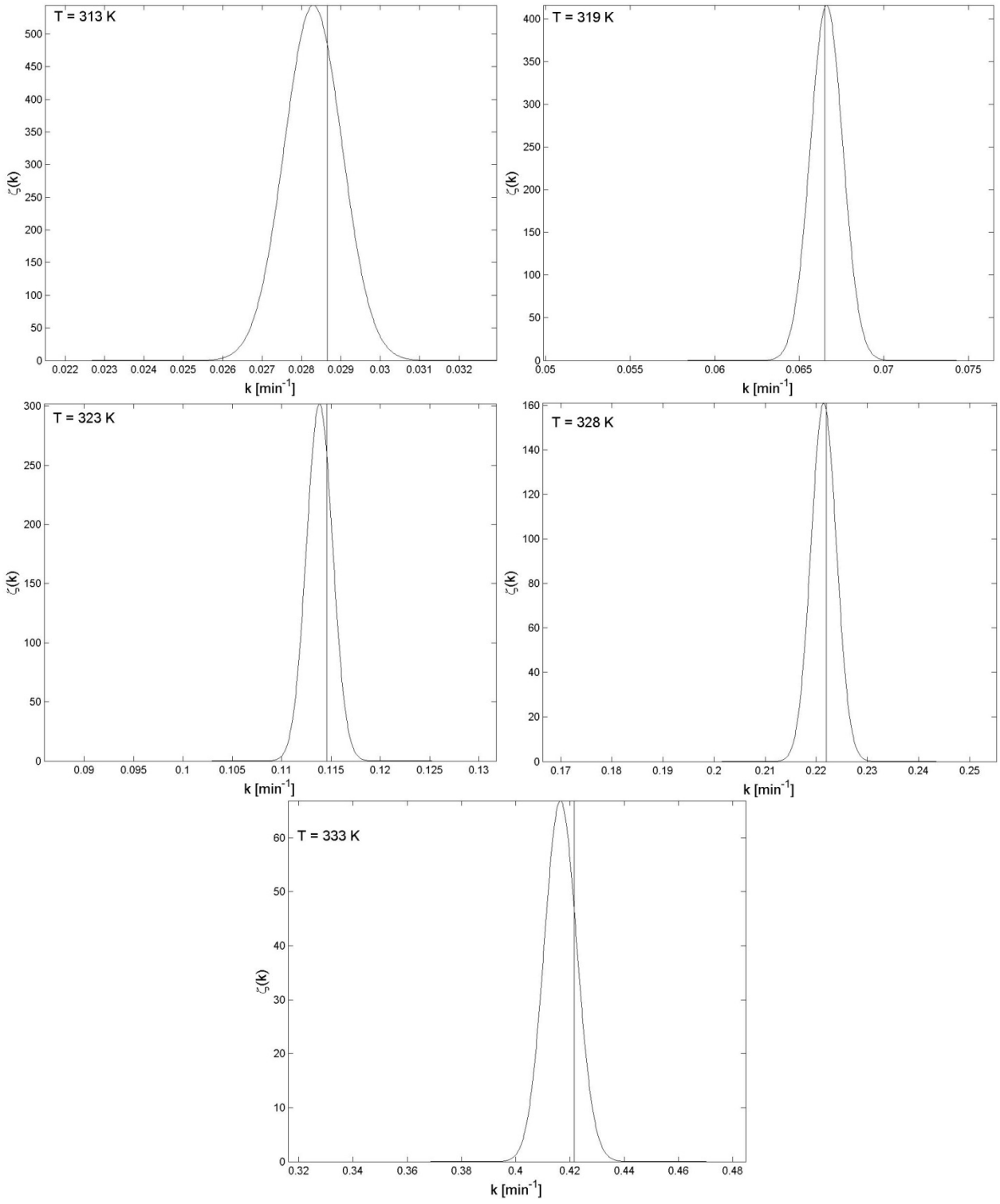

Figure 3.5: Marginalized IRL Posteriors (Perturbed Case)

3.4.4 The Arrhenius Inverse Problem

3.4.4.1 *A Priori* Information and the Discrete Likelihood

For this numerical example the relative priors in both activation energy and pre-exponential factor will be taken to be uniform with bounds selected by some percentage of the true value. These will also provide the bounds of the model space. The set of discrete integrated rate law posteriors will serve as the likelihood functions as stated in subsection 3.3.3

3.4.4.2 Numerical Resolution of the Posterior Density

The uniform nature of the relative priors allows the model space to be constructed by:

$$\mathbb{M}_{AR} = \left[E_A^{min}, E_A^{max}\right] \times \left[k_0^{min}, k_0^{max}\right]$$

Let $R, S \in \mathbb{N}$ denote the number of nodes in each model coordinate. The parameter step sizes are defined by:

$$\Delta E_A = \frac{E_A^{max} - E_A^{min}}{(R-1)}, \Delta k_0 = \frac{k_0^{max} - k_0^{min}}{(S-1)}$$

Individual grid points may then be described by:

$$E_A^r = (r-1)\Delta E_A, k_0^s = (s-1)\Delta k_0$$

These definitions allow for the statement of the model grid by:

$$\{(E_A^r, k_0^s) : 1 \leq r \leq R, 1 \leq s \leq S\}$$

This grid may be computationally navigated through the use of two nested for loops. Computation of the posterior density is carried out using the MATLAB programming environment. Because the relative prior probabilities are uniform the relative prior probability density is described by a constant value and need not be recalculated at each loop recursion. At each loop recursion the forward Arrhenius model must be evaluated five times with the same

model vector, one for each temperature level. The results of the forward model are then applied as the argument for its corresponding marginalized IRL posterior. This is accomplished by the aforementioned mid-point step function interpolation technique with may numerically be described by the following algorithm:

for j = 1:Number of temperatures
if $k^{min} \leq k^{model} \leq k^{max}$
if $k^{min} \leq k^{model} < k^{min} + \frac{\Delta k}{2}$
$\lambda_{AR}^{j}(k^{model}) = \zeta^{j}(k^{min})$
elseif $k^{max} - \frac{\Delta k}{2} \leq k^{model} \leq k^{max}$
$\lambda_{AR}^{j}(k^{model}) = \zeta^{j}(k^{max})$
else
for p = 1:Number of k nodes in IRL discretization
if $k^{p} - \frac{\Delta k}{2} \leq k^{model} < k^{p} + \frac{\Delta k}{2}$
$\lambda_{AR}^{j}(k^{model}) = \zeta^{j}(k^{p})$
end
end
end
else
$\lambda_{AR}^{j}(k^{model}) = 0$
end
end

These five relative likelihoods are then combined through products. The non-normalized value of the Arrhenius posterior is then calculated by taking the product of the nodal likelihood and prior. This process is performed for each nodal model vector. Upon completion the posterior is normalized. The computations associated with the determination of the Arrhenius posterior normalization constant are similar to those associated with the determination of the IRL posterior normalization constant.

3.4.4.3 Application and Results

The upper and lower bounds of the Arrhenius parameter priors were taken to be ∓10% of their true value. These ranges were also used in the construction of the model space. The model space was discretized using 101 nodes in each parameter coordinate. Figure 3.6 displays a contour plot of the posterior density generated from the un-perturbed data. Table 3.6 shows a comparison of the point estimates for this density.

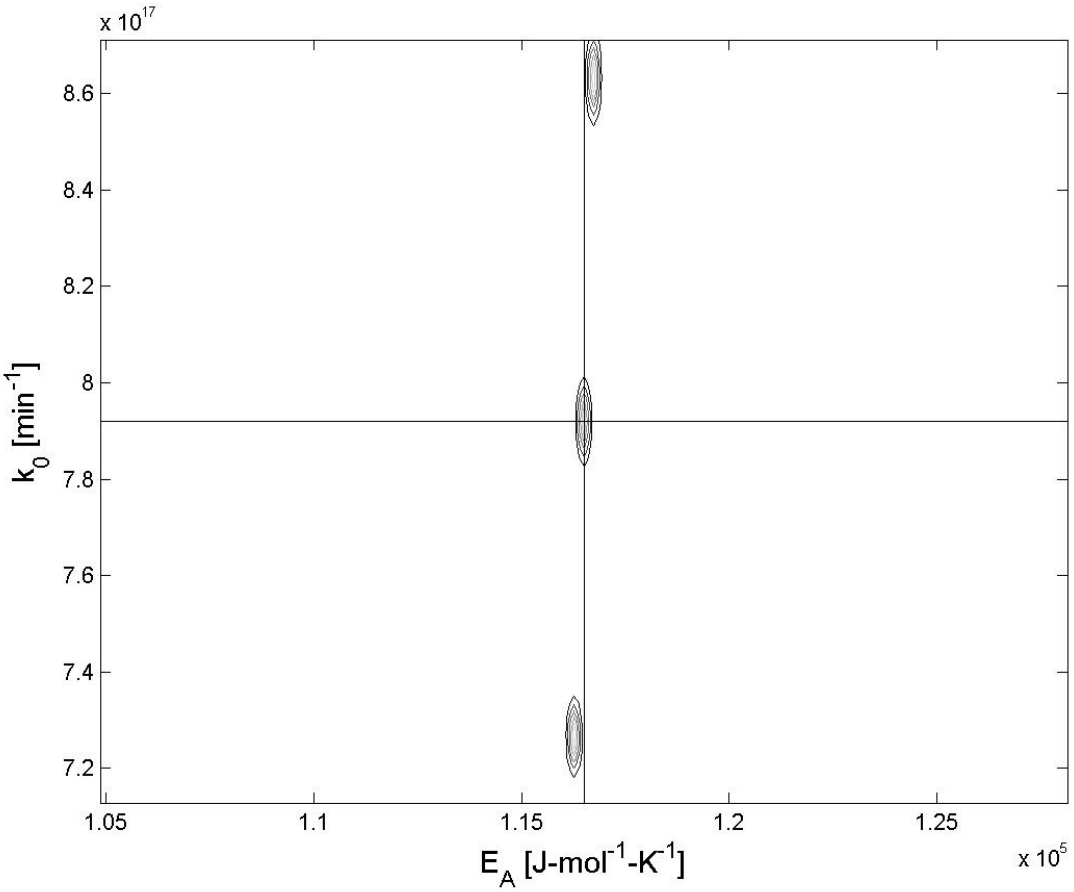

Figure 3.6: Arrhenius Posterior Density (Un-Perturbed Case)

Table 3.6: Arrhenius Posterior Density Point Estimate Comparison (Un-Perturbed Case)

	E_a [kJ-mol^{-1}]	k_0 [min^{-1}]
True	1.165000e+05	7.920000e+17
MAP	1.165000e+05	7.920000e+17
EV	1.165076e+05	7.959566e+17

MAP: maximum a posteriori, EV: expected value

The multi-modal nature of the Arrhenius posterior is an interesting and unexpected phenomenon. Note that the apex mode of the posterior density is sharply centered at the true value. Figure 3.7 shows the side view of a surface plot of the Arrhenius posterior density, depicting the multi-modal nature and the amplitude of the modes.

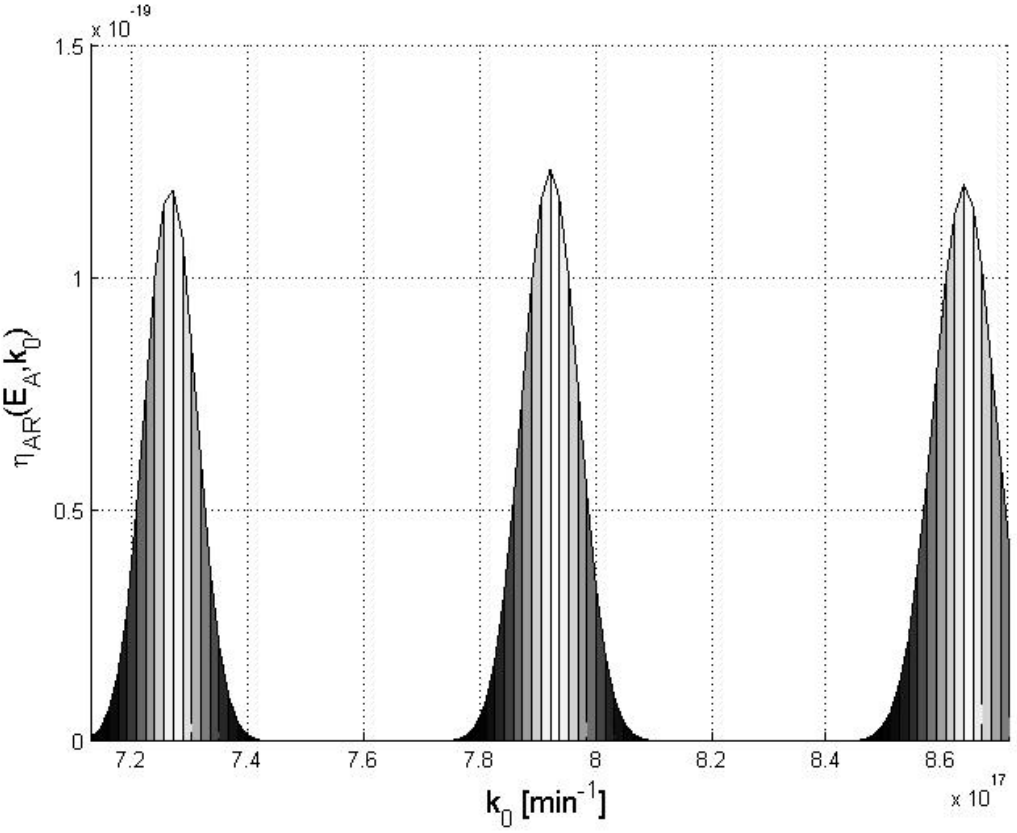

Figure 3.7: Side View of Arrhenius Posterior Surface Plot (Un-Perturbed Case)

It can be seen from Figures 3.6 and 3.7, as well as from Table 3.6, that the true value of the Arrhenius parameters resides at the apex of the central mode. This implies that there are multiple probable model vectors for the Arrhenius inverse problem; however, one is more probable than the others. Since the relative priors in this case are uniform, and therefore do little more than truncate the model space, this occurrence is solely due to the likelihood formulation. Figure 3.8 displays the marginalized IRL posteriors with the specific rate constants returned from a MAP estimate of each posterior mode peak. Inspection of Figure 3.8 corroborates this notion that certain false Arrhenius model vectors result in quality estimates of specific reaction rate. It can be seen that the false modes of the Arrhenius posterior result in specific rates of reaction which lie to the left and right of the marginalized IRL mean while the true mode results in values of the specific rate constant which lies precisely at the marginalized IRL mean. This shows that these false modes produce probable values of the specific rate constant; however, the true mode produces the most probable value. This multi-modal phenomenon reduces the credibility of the expected value point estimator as the false value modes skew the integral away from the true value vector. Table 3.7 shows the values of the Arrhenius parameters at the apex of each posterior mode.

Table 3.7: Maximum A Posteriori Point Estimate Peak Comparison (Un-Perturbed Case)

	E_a [kJ-mol^{-1}]	k_0 [min^{-1}]
Peak 1	1.162670e+05	7.270560e+17
Peak 2	1.165000e+05	7.920000e+17
Peak 3	1.167330e+05	8.632800e+17

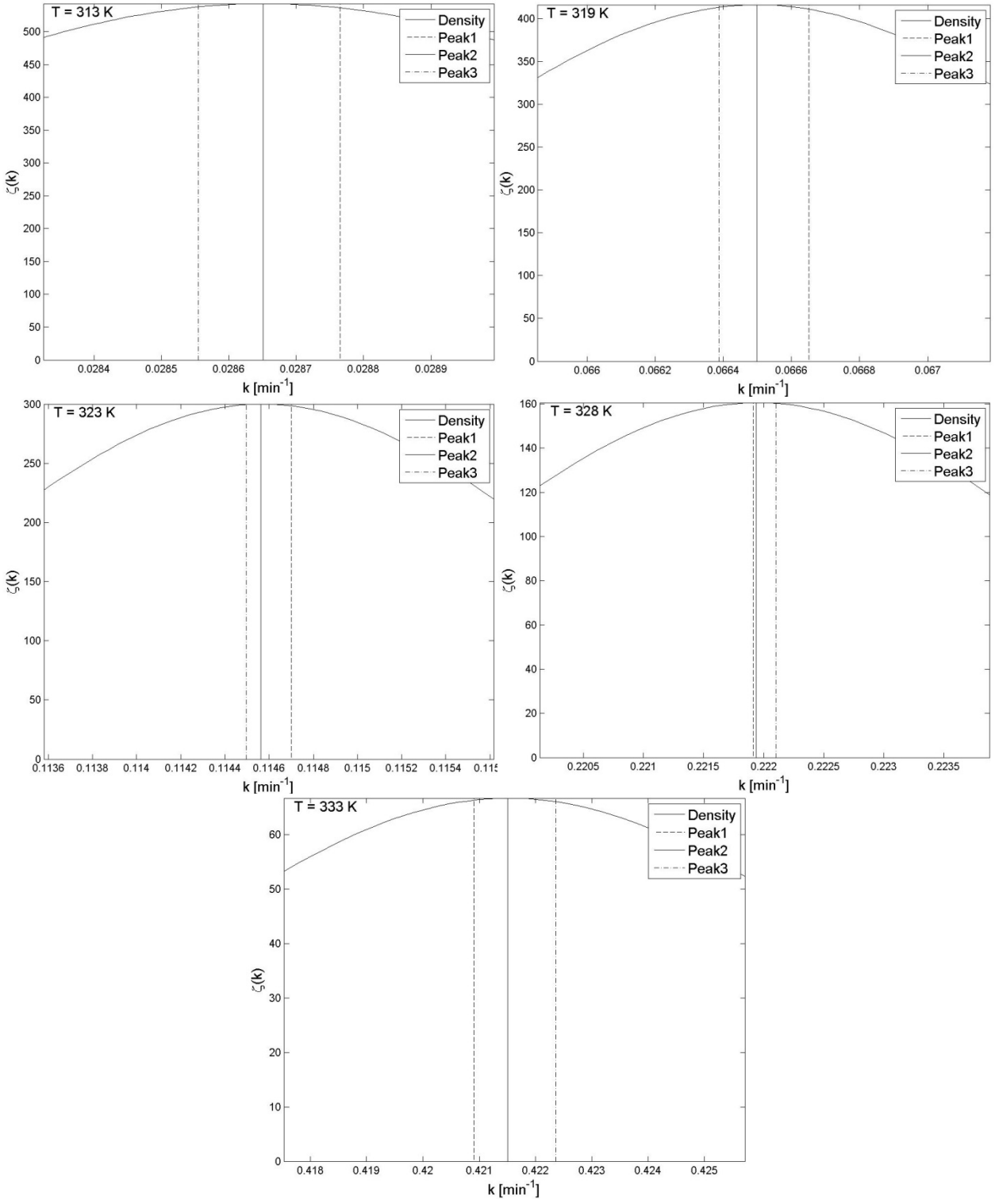

Figure 3.8: Marginalized IRL Posteriors with Peak Probabilities (Un-Perturbed Case)

Further numerical experimentation shows that the multi-modal nature of the posterior is related to both the discretization of the Arrhenius model space as well as likelihood expression used in the IRL inverse problem. If the space is discretized using 201 nodes in each Arrhenius parameter coordinate the resulting posterior contains five probable modes, as shown in Figure 3.9.

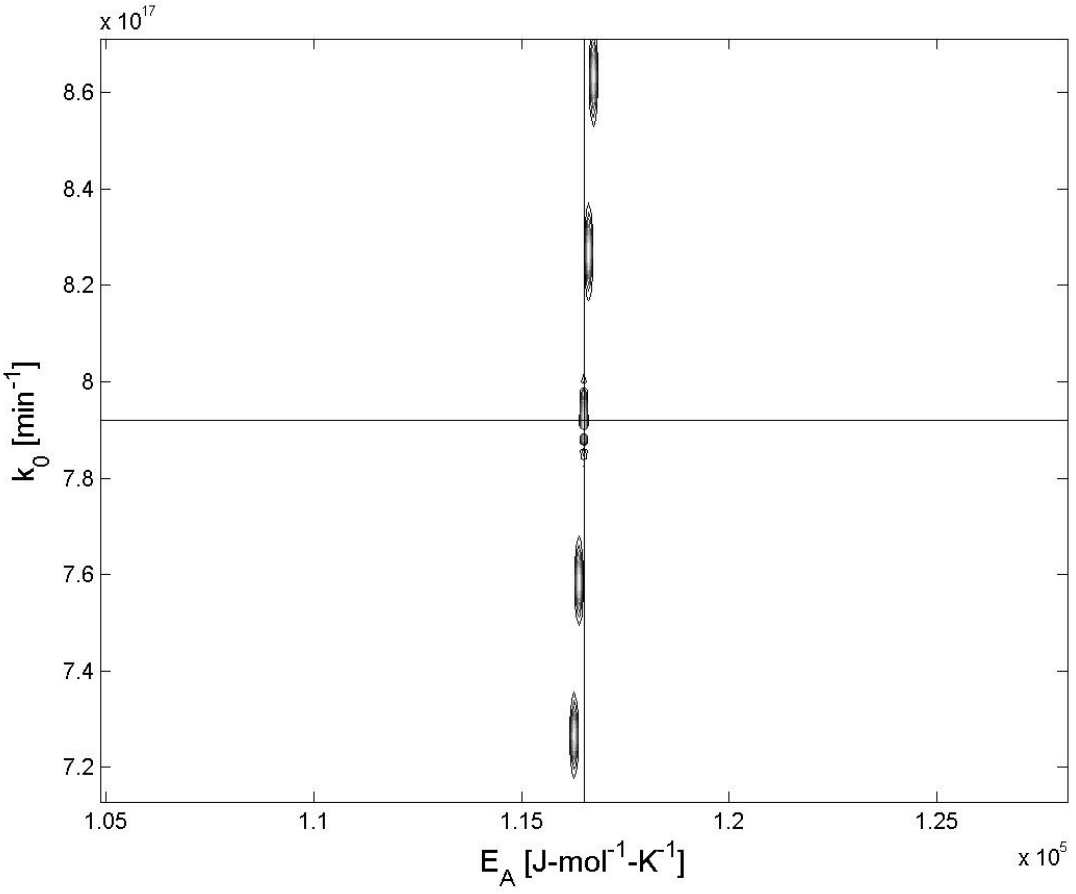

Figure 3.9: Arrhenius Posterior Density (Un-Perturbed Case, 201 Nodes)

Notice that two additional modes have occurred by increasing the discretization of the model space. As the discretization in each Arrhenius parameter coordinate increases to infinity these modes are expected to vanish and a uni-modal density will appear. Figure 3.10 displays the 101

node discretized Arrhenius inverse problem with the IRL inverse problem likelihood standard deviation taken to be 0.001 M. It can be seen from Figure 3.10 that the IRL likelihood variance results in an increase in Arrhenius posterior modal variance.

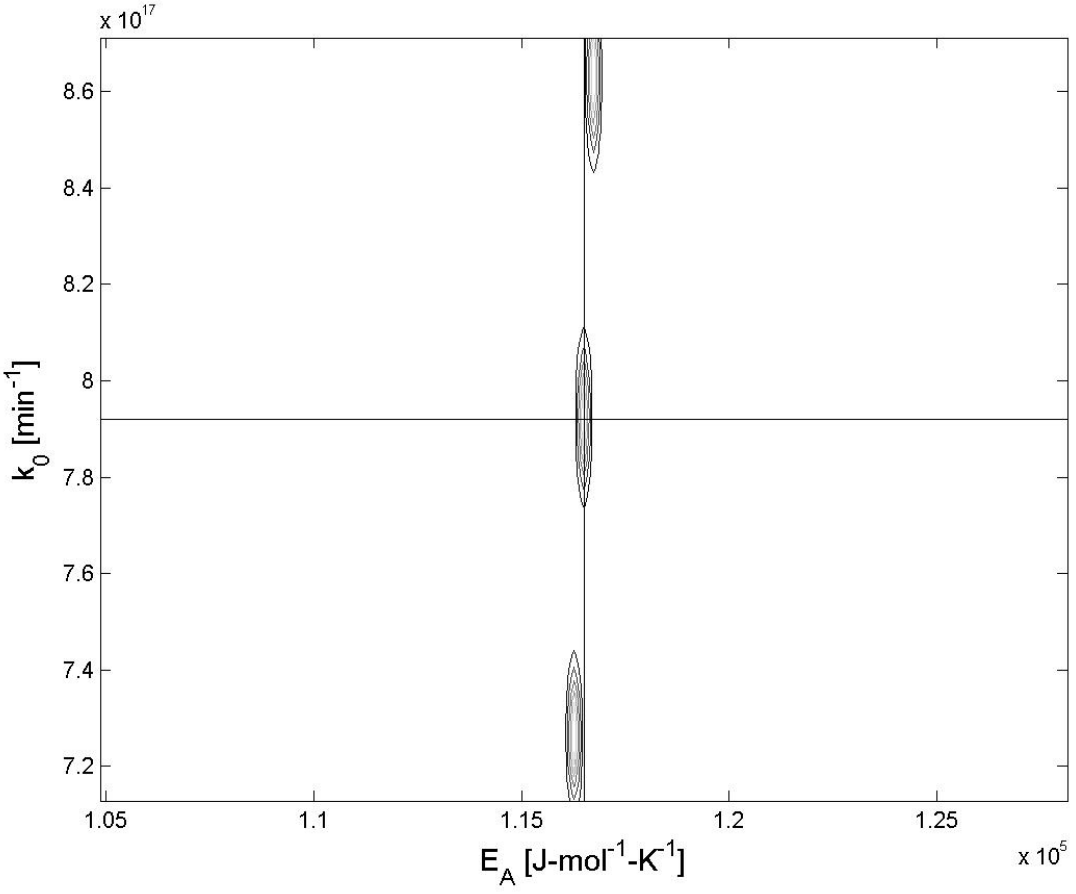

Figure 3.10: Arrhenius Posterior Density (Un-Perturbed Case, STD = 0.001)

These interesting properties of the sequential Bayesian inversion procedure for the estimation of model parameters show that the result of the procedure depends heavily on the discretization of the model space and the confidence which may be placed in the experimental measurement device.

Figure 3.11 shows a contour of the Arrhenius posterior for the perturbed case. It can be seen that the perturbed case displays the same multi-modal behavior as the un-perturbed case. Table 3.8 shows a comparison of the point estimation results to the true parameter values.

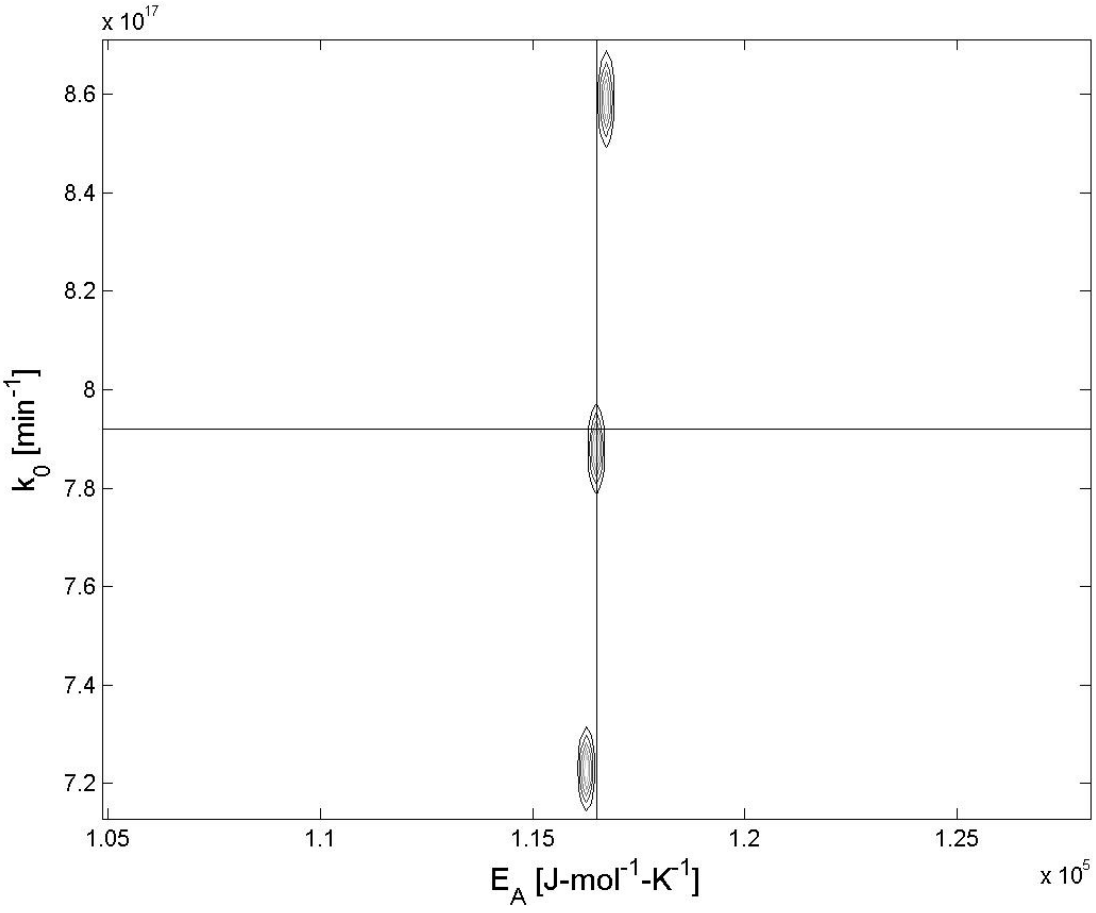

Figure 3.11: Arrhenius Posterior Density (Perturbed Case)

Table 3.8: Arrhenius Posterior Density Point Estimate Comparison (Perturbed Case)

	E_a [kJ-mol^{-1}]	k_0 [min^{-1}]
True	1.165000e+05	7.920000e+17
MAP	1.162670e+05	7.223040e+17
EV	1.165037e+05	7.910266e+17

MAP: maximum a posteriori, EV: expected value

Figure 3.12 shows a side view of a surface plot of the Arrhenius posterior. Table 3.9 shows the values of the Arrhenius parameters associated with each Arrhenius posterior mode peak.

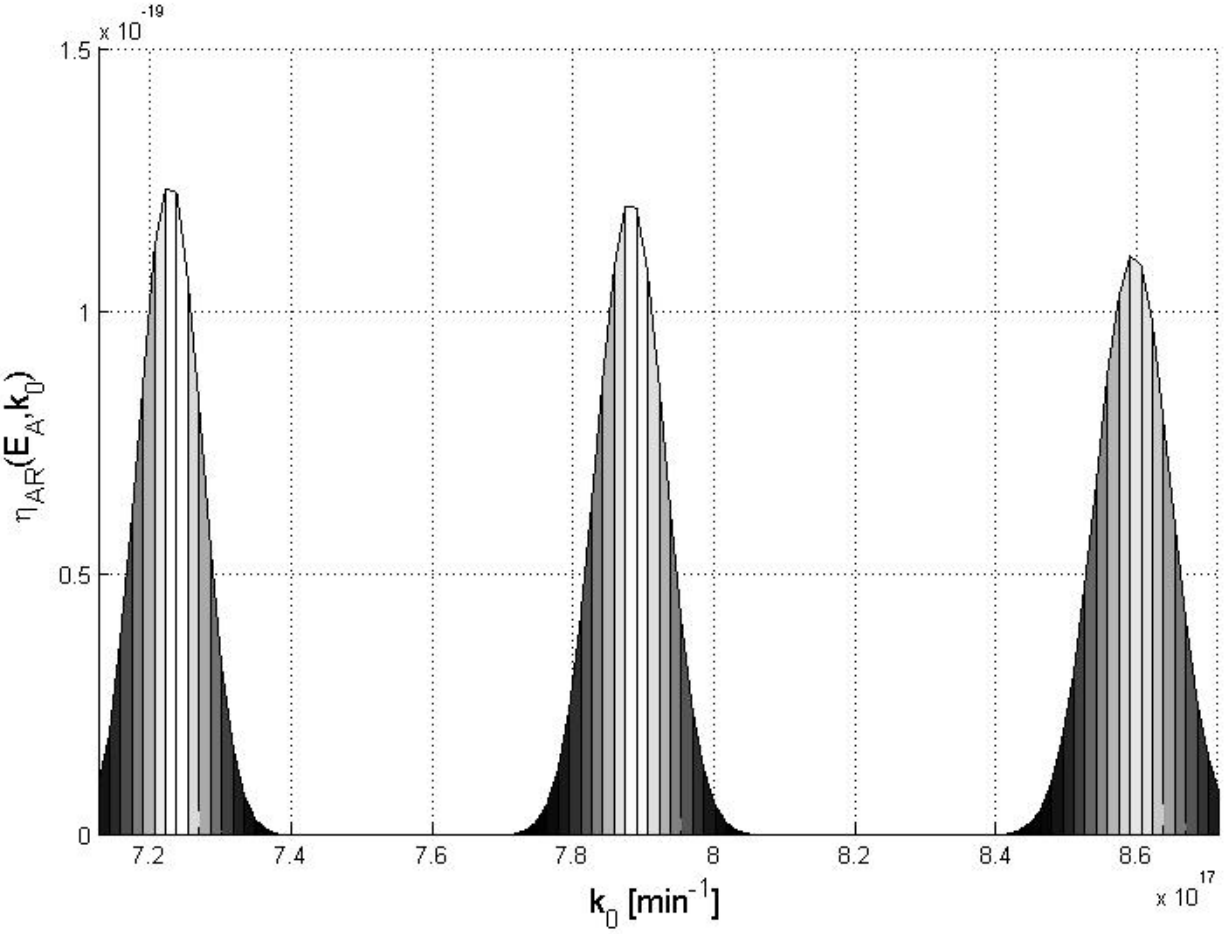

Figure 3.12: Side View of Arrhenius Posterior Surface Plot (Perturbed Case)

Table 3.9: Maximum A Posteriori Point Estimate Peak Comparison (Perturbed Case)

	E_a [kJ-mol^{-1}]	k_0 [min^{-1}]
Peak 1	1.162670e+05	7.223040e+17
Peak 2	1.165000e+05	7.872480e+17
Peak 3	1.167330e+05	8.585280e+17

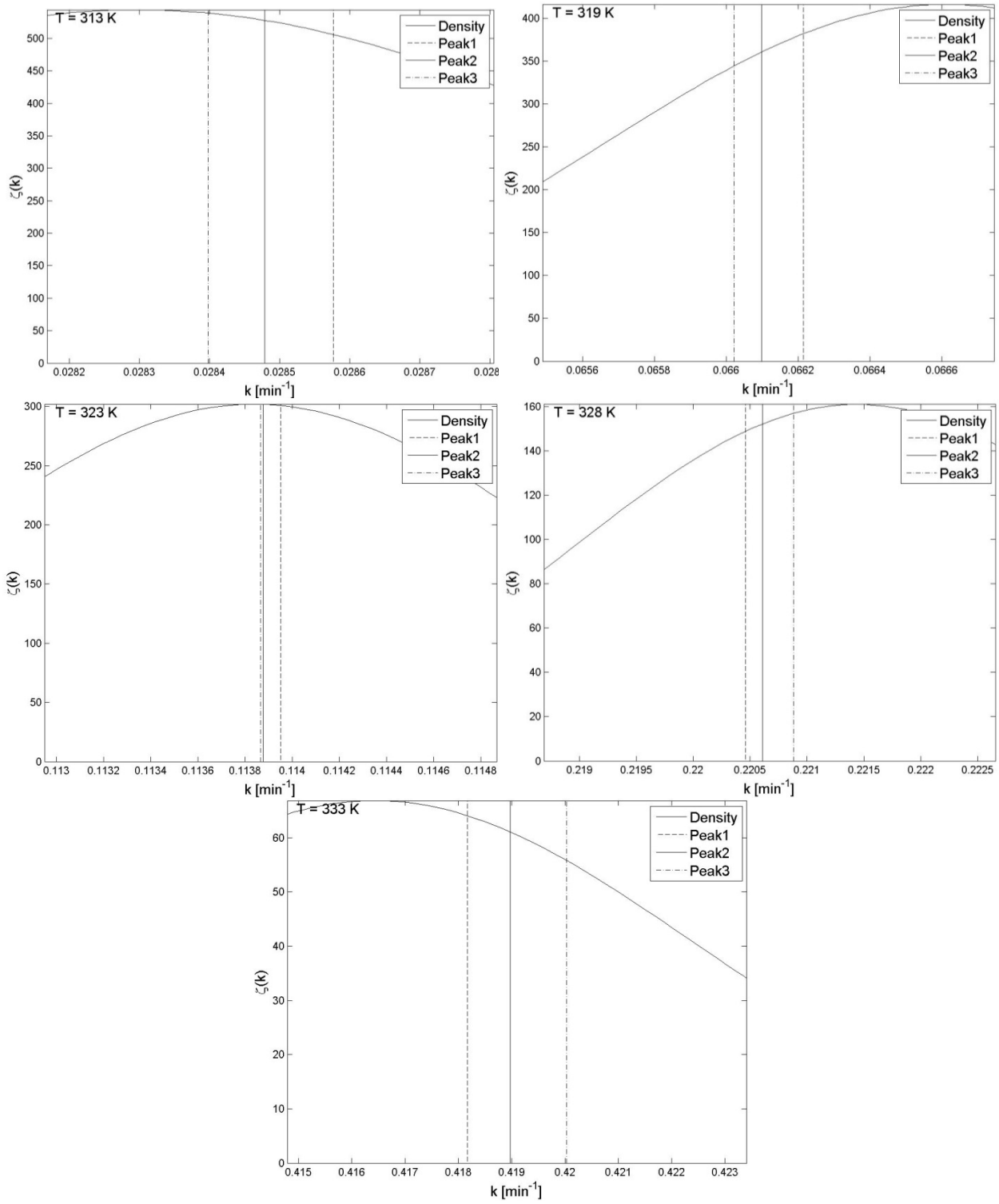

Figure 3.13: Marginalized IRL Posteriors with Peak Probabilities (Perturbed Case)

Inspection of Tables 3.7 and 3.9 shows that the Arrhenius posterior mode peaks reside at approximately the same location for both cases with the difference attributed the perturbation of the data. In the perturbed case the Arrhenius parameters associated with peak 2 are the closest to the values used in the data generation; however, the sequential Bayesian inversion procedure returns with peak 1 presenting the greatest probability, as seen in Table 3.8. This shows that the values used to generate the data were not recovered by the sequential procedure. Figure 3.13 depicts the marginalized IRL posteriors and the value of specific rate constant associated with each Arrhenius posterior mode peak. Inspection of Figure 3.13 shows that the shift in each marginalized IRL posterior caused by the perturbation results in a high probability reported for one of the false peaks in each isothermal case. This results in the method returning the Arrhenius parameter values associated with on the of the false posterior mode peaks as the most probable. This finding questions the legitimacy of the sequential Bayesian procedure; however, it has yet to be seen if the false estimate presented here is still and improvement over other Arrhenius parameter estimation techniques. In actual parameter estimation problems, the true values of the model parameters are not known; therefore, the only way to assess the quality of an estimation technique is to determine how closely the results of the forward model using the estimated parameters match the experimental measurements. Here, three alternative methods of parameter estimation of the Arrhenius parameters are presented and applied to ten different random perturbations of the numerically generated data. This will provide an assessment of the sequential Bayesian method presented here. The methods will be compared based on the residuals between the forward model result using the estimated parameters and the numerically generated data.

3.4.5 Sequential Linear Least-Squares

3.4.5.1 Linear Least-Squares of the IRL Model

Inspection of the IRL expression shows that the mathematical model may be linearized by taking the natural logarithm of both sides the equation. This results in:

$$\ln(C_{BDC}) = -kt + \ln(C_{BDC,0})$$

This formulation of the IRL expression allows for the application of the linear least squares technique by treating the natural logarithm of the isothermal concentration data as the response data, the times as the control, the natural logarithm of initial concentration as the intercept and the specific rate of reaction as the slope. Applying this technique to each of the isothermal data sets results in a set of specific reaction rates corresponding to the set of experimental temperature levels.

3.4.5.2 Linear Least-Squares of the Arrhenius Model

The Arrhenius equation may be linearized in a similar manner resulting in:

$$\ln(k) = \frac{-E_A}{\tilde{R}}\frac{1}{T} + \ln(k_0)$$

Here, treating the natural logarithm of the isothermal specific reaction rates as the response data, the inverse temperature as the control data, the natural logarithm of the pre-exponential factor as the intercept and the ratio of the activation energy and universal gas constant as the slope allows for the application of the liner least squares method of parameter estimation. This is the traditional method of Arrhenius parameter estimation [15]. Tables 3.10, 3.11and Figures 3.14, 3.15 show the result of this method applied to the same perturbed data used in the sequential Bayesian inverse problem presented in subsections 3.4.3 and 3.4.4.

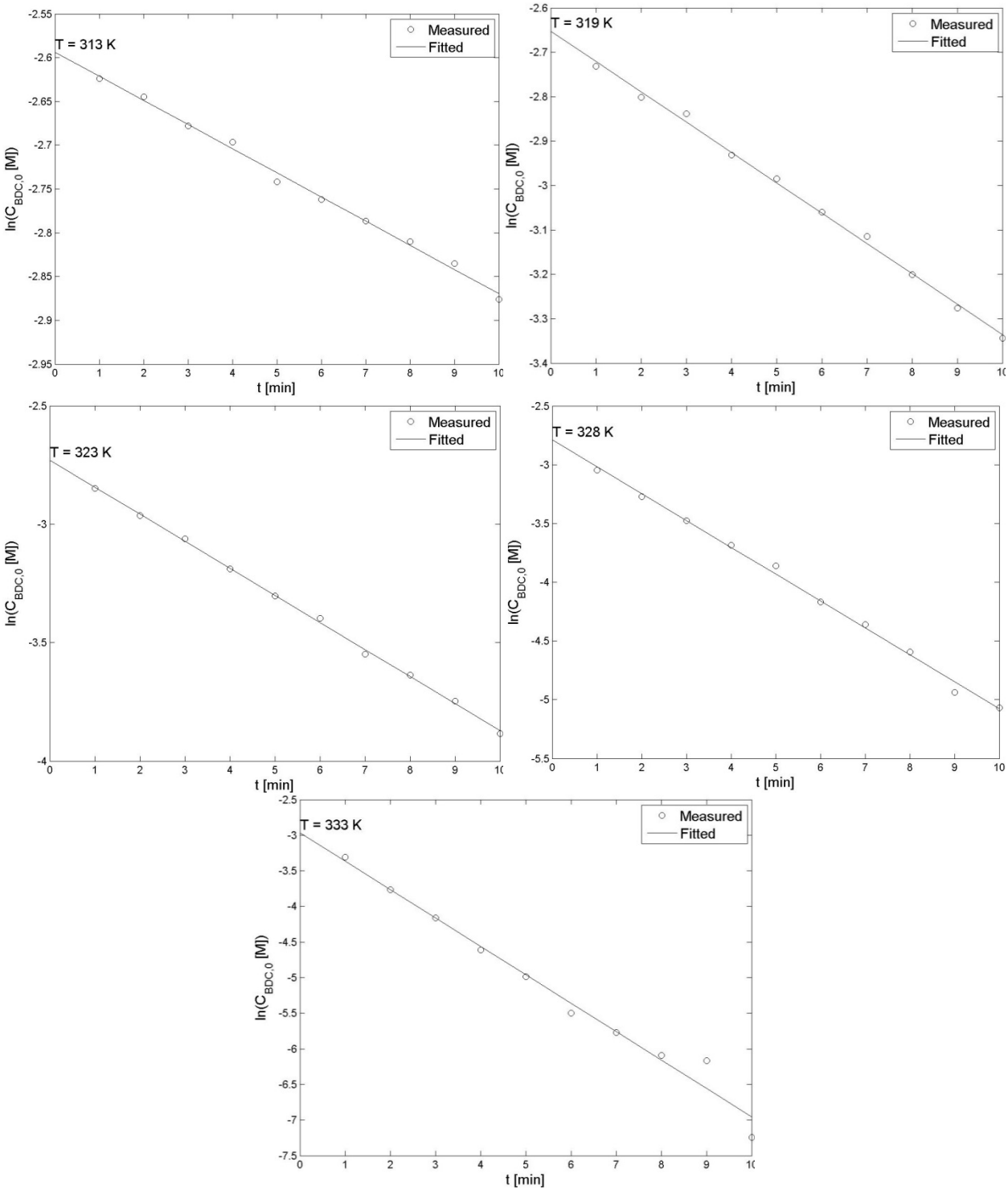

Figure 3.14: IRL Linear Least-Squares Regression Plots (Perturbed Data)

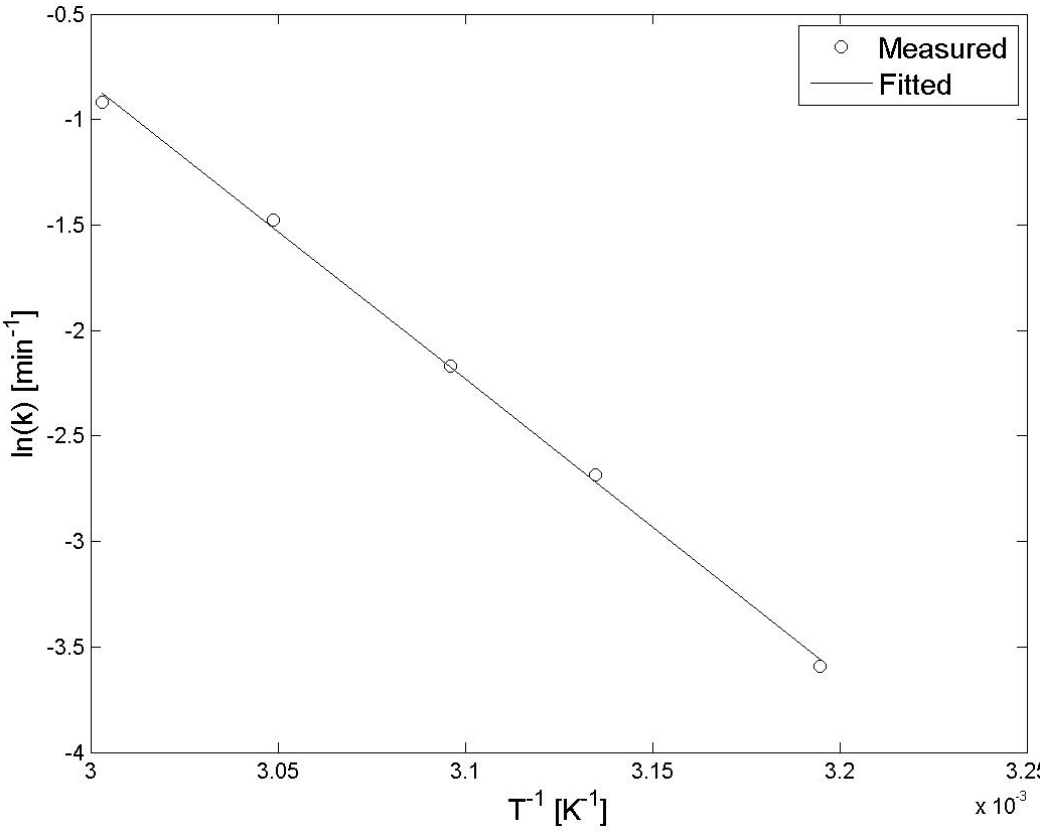

Figure 3.15: Arrhenius Linear Least-Squares Plot (Perturbed Data)

Table 3.10: IRL Linear Least-Squares Results (Perturbed Data)

Temperature	313 K	319 K	323 K	328 K	333 K
k	0.0276	0.0684	0.1142	0.2287	0.3992
R^2	0.9952	0.9970	0.9988	0.9961	0.9804

Table 3.11: Arrhenius Linear Least-Squares Results (Perturbed Data)

Parameter/Fit Quality	Value
E_A	1.161853e+05
k_0	6.990067e+17
R^2	0.9987

3.4.6 Direct Least Squares Estimation

Substitution of the IRL expression into the Arrhenius equation results in:

$$C_{BDC}(t,T) = C_{BDC,0}(T)exp\left[-tk_0\left(\frac{-E_A}{\tilde{R}T}\right)\right]$$

Taking the activation energy, pre-exponential factor, and each of the isothermal initial concentrations to be model parameters, the least squares parameter estimation problem may be formulated as:

$$\underset{\boldsymbol{m}}{minimize}\ f(\boldsymbol{m}) = \sum_{j=1}^{m}(C_A(t,T) - \bar{C}_A(t,T))^2$$

As stated in chapter 1, this is an unconstrained optimization problem typically evaluated through search methods. Here, the problem is solved using the built-in MATLAB function *fminsearch* which uses the Nelder-Mead simplex direct search algorithm. Because the true values are known, as they were used to generate the perturbed data, these are taken to be the initial guess for the parameter values. Table 3.12 shows the results of this method applied to the perturbed data used in the sequential Bayesian case.

Table 3.12: Result of Direct Optimization Least-Squares Problem (Perturbed Data)

Parameter/Fit Quality	Value
$\boldsymbol{E_A}$	1.1682e+05
$\boldsymbol{k_0}$	8.8674e+17
Sum of Residual Squares	1.1191e-05

3.4.7 Direct Bayesian Inversion

The last method of Arrhenius parameter estimation presented in this work is the application of the Bayesian approach to the combined Arrhenius-IRL model:

$$C_{BDC}(t,T) = C_{BDC,0}(T)exp\left[-tk_0\left(\frac{-E_A}{\tilde{R}T}\right)\right]$$

This direct formulation involves the evaluation of the posterior density over a seven dimensional model space; two dimensions associated with the Arrhenius parameters and five associated with each isothermal initial concentration. The model space may be constructed as:

$$\mathbb{M}_{AR} = \left[E_A^{min}, E_A^{max}\right] \times \left[k_0^{min}, k_0^{max}\right] \times \prod_{j=1}^{5}\left[\bar{C}_{BDC,0}^{j} - 4\sigma, \bar{C}_{BDC,0}^{j} + 4\sigma\right]$$

The model space may be discretized in a manner similar to that of the sequential Bayesian approach. Here, the Arrhenius relative prior densities are taken to be uniform while the initial concentration relative priors are taken to be Gaussian with known variance. The likelihood in this case is the same as the likelihood for the IRL inverse problem as the form of the likelihood is determined by the measurement technique and the combined Arrhenius-IRL model predicts the value of concentration. In the numerical evaluation of the posterior density the Arrhenius space is bounded by ∓10% of the true values, similar to the sequential Bayesian case. Each Arrhenius coordinate is discretized using 101 nodes and each initial concentration coordinate is discretized using 5 nodes. Figure 3.16 and Table 3.13 show the results of this direct inverse problem formulation for the same perturbed data used in the sequential Bayesian case. It can be seen from Figure 3.15 that the direct Bayesian posterior suffers from the same multimodal phenomenon as the sequential Bayesian. This multimodal occurrence may be attributed to the highly non-linear nature of the combined Arrhenius- IRL model [8].

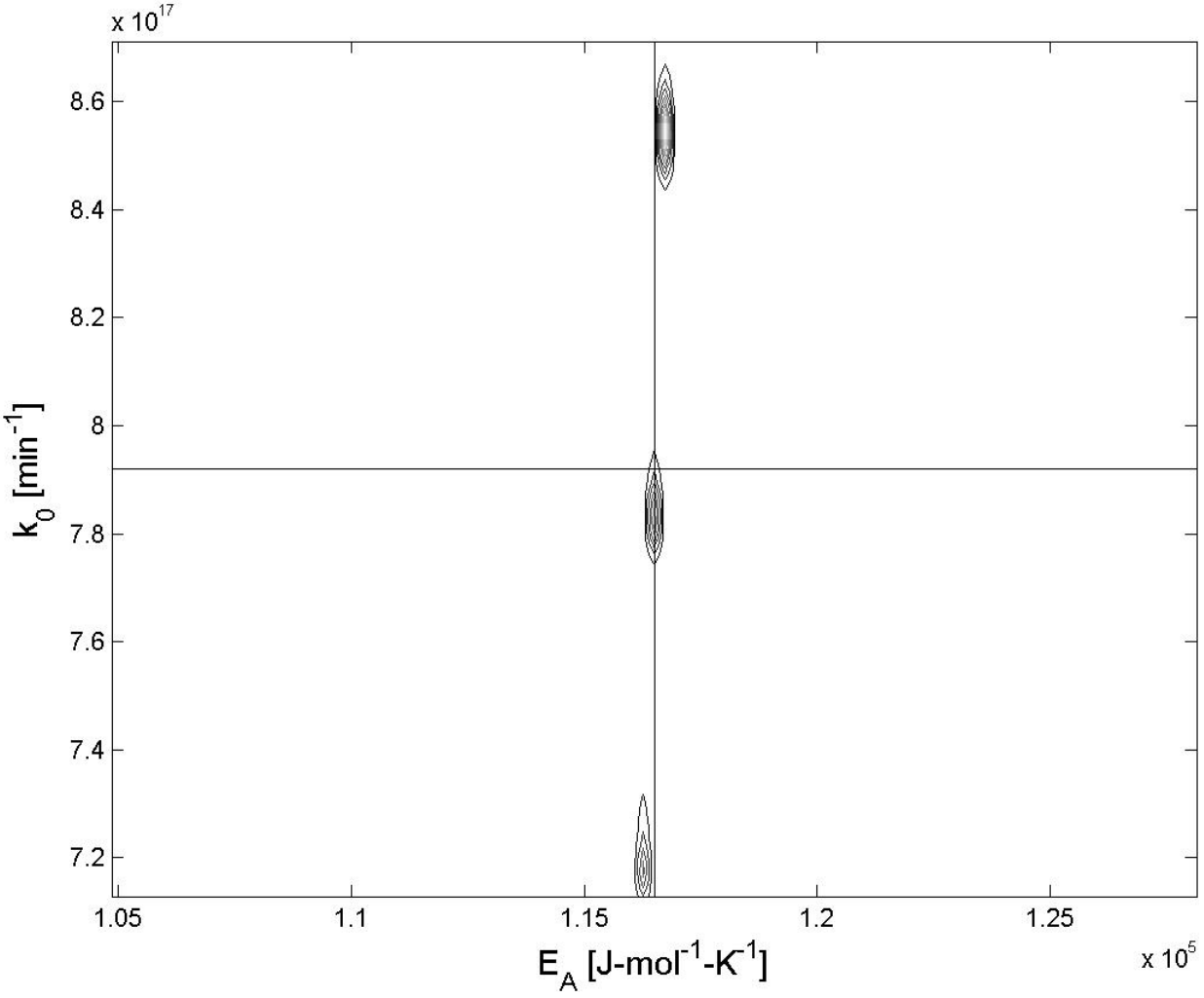

Figure 3.16: Posterior Contour for Direct Bayesian Formulation (Perturbed Data)

Table 3.13: Results of Direct Bayesian Inversion (Perturbed Data)

Parameter	Value
E_A	1.1673e+05
k_0	8.5536e+17

3.4.8 Comparison of Techniques

Each of the four methods presented here was applied to ten different random perturbations of the data. The parameter estimates, taken to be the MAP estimator in the Bayesian cases, were used in the evaluation of the forward problem to generate isothermal concentration time data sets. The residuals between these forward model data sets and the randomly perturbed data were computed as the Euclidean norm of the difference in each concentration value. Table 3.14 shows the means and variances of the residuals associated with each method.

Table 3.14: Estimation Technique Comparison

Technique	Mean	Variance
Bayesian Sequential	3.1564e-03	4.7437e-07
Sequential Least-Squares	4.1837e-03	2.5852e-06
Direct Least-Squares	3.3157e-03	4.9609e-07
Direct Bayesian	3.2054e-03	6.2029e-07

The sequential Bayesian formulation is observed to result in the lowest mean residual and the narrowest variance. Both the Bayesian and least-squares direct formulations performed marginally worse than in the sequential Bayesian in terms of mean and variance. The sequential least squares formulation performs the worst with the highest overall residual mean and variance. The results presented in Table 3.14 show that the sequential Bayesian approach developed here yields results of a similar quality to that of direct problem while significantly reducing the computational cost. Furthermore, it can be seen that the current method typically employed in Arrhenius parameter estimation performs the worst of any of the methods investigated.

3.4.9 Combination and Utilization of Arrhenius Parameter Estimation Methods

The application of Bayesian statistics to inverse problems is not driven by the desire for more accurate point estimates. It is driven by the pursuit of information concerning uncertainty quantification. Furthermore, the Bayesian formulation of a given inverse problem is useless in the absence of adequate prior information concerning the parameters of the model. In the numerical example presented here the true values, i.e. the values used in the generation of the data, were known. This allowed for several of the relative prior probabilities, which are used to constrain the model space, to be constructed using these known true values. In actual parameter estimation problems the true values are not known, requiring an alternative means of prior construction and model space truncation. In this study, the relative prior probabilities were primarily used to constrain the model space to a region of expectable model vectors. If the true values of the parameters are not known then how may the model space be constrained to a region of expectable values? This task may be accomplished through the successive refinement of information obtained from both the deterministic and probabilistic approaches. The sequential least-squares, the direct least-squares, and the sequential Bayesian formulations may be used in tandem to obtain a quality state of information concerning the model parameters while still being significantly more computationally tractable than the direct Bayesian approach. The direct least-squares problem requires an initial guess of the optimal vales of the model parameters. In this study the true values of the parameters were taken to be the initial guess in the interest of simplicity; however, in practice the true values will not be known. The sequential least squares is a simplistic approach which provides a rough estimate of the model parameters. This approach involves the application of linear least squares which has a unique, analytic solution for the model parameters and requires no initial guess. The Arrhenius parameters and isothermal initial

concentrations predicted through the sequential approach may be used as the initial guess for the optimal values of the direct least squares approach. The sequential Bayesian formulation of the problem requires some prior information in each isothermal specific rate constant and both Arrhenius parameters to allow for the truncation of the model space. The specific rate constant estimates from the sequential least squares and the Arrhenius parameter estimates from the direct least squares provide likely values for all these quantities. Taking some uniform region centered at these values allows for the definition of a finite model space. Figure 3.17 depicts a flow chart of this coupled method approach.

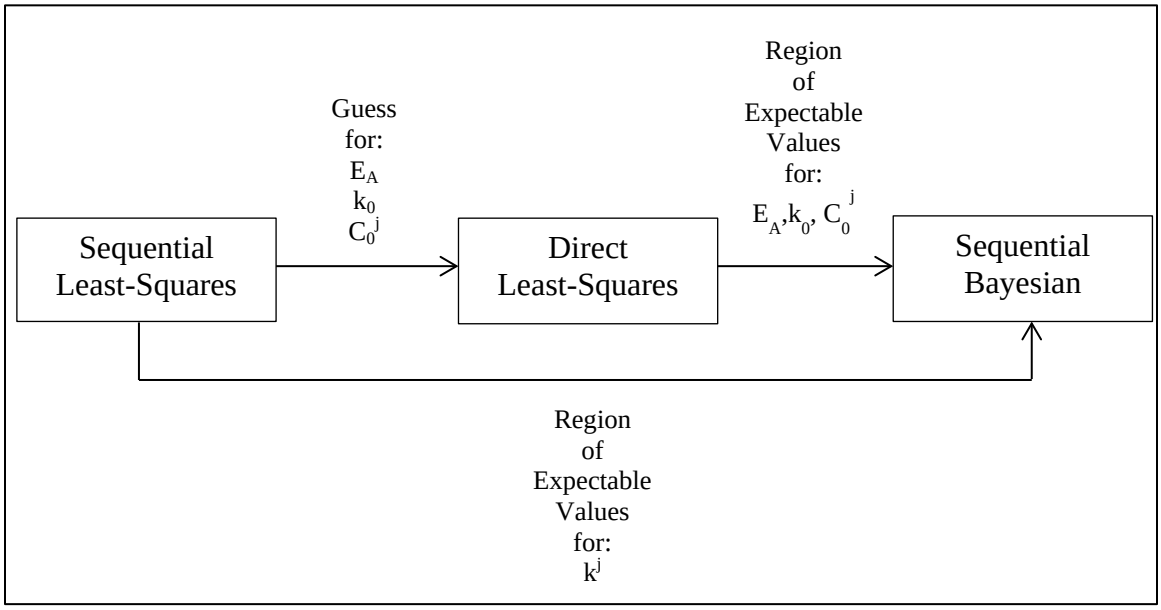

Figure 3.17: Method Combination Flow Chart

Using these three estimation techniques allows for the estimation of probable values of the Arrhenius parameters as well as quantifiable confidence information at a reduced computational cost compared to the direct Bayesian formulation.

3.5 CLOSING REMARKS

This numerical example conveys the complexities and ambiguities associated with the application of Bayesian inversion to non-linear inverse problems. The multimodal nature of the resulting posterior density makes this a difficult problem to analyze; however, such is the nature of subjective probability. The Bayesian approach describes a *belief* of information concerning the values of the model parameters as well as the confidence which may be placed in their estimation. The primary advantage of the Bayes' formulated inverse problem is information concerning the uncertainty in a given parameter estimate. In the multimodal case presented here the typical uncertainty quantifiers such as variance and covariance may not be strictly applied, as doing so would result in a grossly conservative confidence estimate. Such uncertainty quantifiers would only hold meaning by treating the mode of interest as a single, unimodal probability density and computing the uncertainty quantification indicators for the mode. This interpretation of single mode uncertainty quantification is necessary to give physical meaning and utility to the resulting posterior density. While this single mode selection technique many lack mathematical rigor, subjective probability is as its name implies; subjective.

4.0 CONCLUSIONS AND FURTHER DEVELOPMENTS

In the previous chapter, the direct Arrhenius inverse problem was introduced and the entirety of the posterior probability density was resolved through direct computation of the likelihood over the whole of the discrete model space. This direct formulation of the inverse problem suffers from the *curse of dimensionality* in that the number of model space vectors increases exponentially with number of parameter coordinates. In the case of high dimensionality inverse problems, even modestly resolved discretizations of individual parameter coordinates result in an extremely high number of model space vectors. Computing the posterior probability at each model space vector places high dimensionally inverse problems out of the range of computational tractability; however, the majority of the computations associated with this direct procedure provide little information about characteristics of the posterior probability density. This is because high dimensionality model spaces tend to be very empty, *i.e.* there exist large regions of extremely low probability throughout the model space. It is desirable to develop a procedure to find locations of high probability within the model space without sampling the entirety of the space. In this chapter the topics of Monte Carlo sampling and sparse grid construction as they apply to inverse problem solutions, are discussed in modest detail. This chapter serves as mild introduction to the handling of high dimensionality inverse problem using these two methods of probability sampling.

4.1 THE METROPOLIS-HASTINGS ALGORITHM

Monte Carlo methods involve the random sampling of the posterior probability over the model space in an effort to locate regions of high probability. One such method is the Metropolis-Hastings algorithm, a Markov chain, Monte Carlo Method which moves through the model space by accepting move directions most likely to result in a higher value of the posterior probability and rejecting moves which will likely result in a lower value of the posterior probability. The algorithm is initiated by selecting a point in the model space believed to reside near a region of high probability. The selection of this initial point is left to interpretation as it involves an understanding of the problem, leading to an expectation of the location of the highly probable parameter regions. This starting point will be called $\boldsymbol{m}_i$. From this initial point a move to model vector $\boldsymbol{m}_j$ is randomly selected. If $\lambda(\boldsymbol{m}_j) \geq \lambda(\boldsymbol{m}_i)$ then the move is accepted. If $\lambda(\boldsymbol{m}_j) < \lambda(\boldsymbol{m}_i)$ then decide randomly to move to $\boldsymbol{m}_j$ or stay at $\boldsymbol{m}_i$, with the probability of moving to $\boldsymbol{m}_j$ given by [8]:

$$P_{i\to j} = \frac{\lambda(\boldsymbol{m}_j)}{\lambda(\boldsymbol{m}_i)}$$

This procedure is followed until the region of high probability is located and sufficiently sampled such that meaningful information about the region may be inferred. This method of posterior probability resolution is well suited for inverse problems where the posterior density is expected to be unimodal; however, in the case of non-linear inverse problems the method may fail to locate other regions of high probability as non-linear inverse problems tend to be multimodal [8]. Application of Monte Carlo methods to non-linear inverse problems requires knowledge of the physics of the given problem to appropriately sample the model space to determine sufficient information concerning the behavior of the posterior probability density.

4.2 SPARSE GRIDS

The method of sparse grids handles the problem of high dimensionality in a more deterministic manner by performing a hierarchical subspace-splitting procedure and interpolating the value of the desired function, which in the case of Bayesian inversion is the posterior probability density, between the sparse grid points [16]. Let h be the grid mesh size, defined by: $h = 2^{-n}$ where n is the discretization level. For a *k*-dimensional space the number of grid points utilized in the sparse grid procedure to obtain 2nd order accuracy is described by:

$$O(h^{-1} \cdot \log(h^{-1})^{k-1})$$

This may be compared to the number of grid points employed by a standard tensor product grid to achieve 2nd order accuracy, which is given by:

$$O(h^{-k})$$

The method of sparse grids is analogous to the method if finite elements in that the function is approximated using linear-piecewise shape functions to approximate the value of the function within the hierarchical subspaces between the grid points.

4.3 CONCLUSIONS

Here, the application of Bayesian statistics to the general discrete inverse problem has been presented. The application of the Bayesian inversion procedure was applied to two scientifically interesting problems: the reversible reaction-diffusion inverse problem and the Arrhenius inverse problem. The reversible-reaction diffusion inverse problem served as a well behaved example problem to introduce the procedure of Bayesian inversion. The initial artificial experiment

produced adequate data to resolve the true values of the model parameters with high confidence. It was observed that initial condition and measurement frequency affected the quality of knowledge concerning the model parameters, thus showing that Bayesian inversion allows for the tailoring of experimental methods for a desired parameter estimate confidence. The Arrhenius inverse problem was not a simple problem to formulate due to the inability to observe the specific rate of reaction. A novel procedure was developed here to sequentially solve isothermal IRL inverse problems and take the marginalized IRL posteriors to be the relative likelihoods in the Arrhenius inverse problem. The estimates produced from the novel approach were capable of replicating the data with quality comparable to that of the least-squares optimization and Bayesian inversion of the direct model, while providing uncertainty information and maintaining small scale computing tractability. This sequential Bayesian approach significantly reduces the total computational cost of Arrhenius parameter estimation by reducing the dimensionally of the problem and replacing dimensions with separate inverse problems, making the number of operations additive as opposed to exponential. On the whole, Bayesian inversion provides a means of quantifying the confidence which may be placed in a parameter estimate; however, an understating of the physics of the inverted model is required to interpret the resulting posterior density to a useful state of knowledge.

APPENDIX A

SEQUENTIAL ARRHEMIUS INVERSE PROBLEM PROGRAM

Integrated Rate Law Posterior Pseudo-Code and Function Information

IRL Posterior Pseudo-Code:

Load Concentration Time Data
for j = 1:J (Loop over Temperature Levels
Define Model Space Grid
for p = 1:P, q = 1:Q (Loop over Model Space Grid)
Run IRL_prior_solver
Run IRL_model_solver
Run IRL_likelihood_solver
Compute Nodal Value of Non-Normalized Posterior
end (Loop over Model Space Grid)
Normalize Posterior
Export Data
end (Loop over Temperature Levels)

IRL Posterior Function Descriptions:

```
function psi = IRL_prior_solver(m,C0_D,k_true,sigma)
%Computes the Nodal Prior Probability for Integrated Rate Law
Inverse Problem
%psi = IRL_prior_solver(m,C0_D,k_true,sigma)
%
%psi is the nodal prior probability
%m is a column vector whose elements are the model parameters
given by:
%   m = [k;C0]
%   k is the specific rate of reaction
%   C0 is the initial concentration
%C0_D is the measured value of initial concentration
%k_true is the value of k used in data generation
%sigma is the standard deviation of the measurement device

function [C] = IRL_model_solver(m,t)
%First Order Integrated Rate Law Forward Model
%[C] = IRL_model_solver(m,t)
%
%C is a column vector containing the concentration data over
time
%m is a column vector whose elements are the model parameters
given by:
%   m = [k;C0]
%   k is the specific rate of reaction
%   C0 is the initial concentration
%t is a column vector containing the corresponding time values

function lambda = IRL_likelihood_solver(C_D,C_M,sigma)
%Computes the Nodal Likelihood for Integrated Rate Law Inverse
Problem
%lambda = IRL_likelihood_solver(C_D,C_M,sigma)
%
%lambda is the nodal likelihood
%C_D is a column vector containing the measured concentration-
time values
%C_M is a column vector containing the model concentration-time
values
%sigma   is   the   standard   deviation   of   the   concentration
measurement device
```

IRL Marginalization Pseudo Code:

for j = 1:J (Loop over Temperature Levels)
Load Isothermal IRL Posterior Density
Run IRL_marginalizer
Normalize Marginal Posterior Density
Export Marginal Posterior Density
end (Loop over Temperature Levels)

IRL Marginalization Function Descriptions:

```
function ETA_k = IRL_margnializer(ETA,mStep,mRange)
%Computes the Marginal Probability in k of the IRL posterior
%ETA_k = IRL_margnializer(ETA,mStep,mRange)
%
%ETA_k is the marginal probability in k
%ETA is an array containing the posterior density
%mStep is a column vector containing the stepsizes used in the
posterior
%   computation of the form mStep = [dk;dC0]
%mRange is an array whose rows are the upper and lower bounds of
the
%   individual parameter spaces of the form
%   mRange = [k_min,k_max;C0_min,C0_max]
```

Arrhenius Posterior Pseudo-Code:

Define Model Space Grid
Run AR_prior_solver
Load Marginalized IRL Posterior
for r = 1:R, s = 1:S
Run AR_model_solver
Run AR_likelihood_solver
Compute Nodal Value of Non-Normalized Posterior
end (Loop over Model Space Grid)
Normalize Posterior Density
Export Arrhenius Posterior Density

Arrhenius Posterior Function Descriptions

```
function psi = AR_prior_solver(Ea_min,Ea_max,k0_min,k0_max)
%Computes Uniform Prior Probability for Arrhenius Inverse
Problem
%psi = AR_prior_solver(Ea_min,Ea_max,k0_min,k0_max)
%
%psi is the prior probability
%Ea_min/Ea_max are the lower and upper bounds of the Ea density
%k0_min/k0_max are the lower and upper bounds of the k0 density

function [k] = AR_model_solver(m,T)
%Arrhenius Forward Model
%[k] = AR_model_solver(m,T)
%
%k is a column vector containing the specific rate constants
%m is a column vector containing the Arrhenius model parameters
of the
%   form: m = [Ea;k0];
%T is a column vector containing the temperatures
```

```
function lambda = AR_likelihood_solver(k,L_k,DK,KRANGE)
%Computes the Nodal Likelihood for the Arrhenius Inverse Problem
%lambda = AR_likelihood_solver(k,T)
%
%lambda is the nodal likelihood
%k is a column vector containing the model values of specific
rate constant
%L_k is an array whose columns are the marginalized isothermal
posterior
%   probabilities
%DK is a column vector whose elements are the dk for each
isothermal
%   posterior
%KRANGE is an array whose columns are the lower and upper bound
for each
%   isothermal posterior
```

BIBLIOGRAPHY

1. *Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification*. 2012: The National Academies Press.

2. Nocedal, J. and S. Wright, *Numerical optimization, series in operations research and financial engineering.* Springer, New York, 2006.

3. Felder, R.M. and R.W. Rousseau, *ELEMENTRY PRINCIPLES OF CHEMICAL PROCESSES, (With CD)*. 2008: Wiley. com.

4. Bolstad, W.M., *Introduction to Bayesian statistics*. 2007: Wiley. com.

5. Grimmett, G.R., *Probability: an introduction*. 1986: Oxford University Press.

6. Allmaras, M., et al., *Estimating Parameters in Physical Models through Bayesian Inversion: A Complete Example.* SIAM Review, 2013. **55**(1): p. 149-167.

7. Calvetti, D. and E. Somersalo, *Introduction to bayesian scientific computing*. 2007: Springer Science+ Business Media.

8. Tarantola, A., *Inverse problem theory and methods for model parameter estimation*. 2005: siam.

9. Kaipio, J.P. and E. Somersalo, *Statistical and computational inverse problems*. Vol. 160. 2005: Springer.

10. Chorin, A.J. and O.H. Hald, *Stochastic tools in mathematics and science*. Vol. 1. 2006: Springer.

11. Pletcher, R.H., D.A. Anderson, and J.C. Tannehill, *Computational fluid mechanics and heat transfer*. 2012: CRC Press.

12. Quarteroni, A., R. Sacco, and F. Saleri, *Numerical mathematics*. Vol. 37. 2007: Springer.

13. Fogler, H.S., *Essentials of chemical reaction engineering*. 2010: Pearson Education.

14. Hill, J.W., R.H. Petrucci, and M.D. Mosher, *General chemistry*. 2005: Pearson Prentice Hall Upper Saddle River, NJ.

15. Espenson, J.H., *Chemical kinetics and reaction mechanisms*. 1995: McGraw-Hill New York.

16. Garcke, J. and M. Griebel, *Sparse grids and applications*. Vol. 88. 2012: Springer.