SANDIA REPORT
SAND201X-XXXX
Unlimited Release
Printed Month and Year

Paracousti-UQ: A Stochastic 3-D Acoustic Wave Propagation Algorithm

Leiph A. Preston

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology and Engineering Solutions of Sandia, LLC.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: https://classic.ntis.gov/help/order-methods/

SAND201X-XXXX
Printed September 2017
Unlimited Release

Paracousti-UQ: A Stochastic 3-D Acoustic Wave Propagation Algorithm

Leiph A. Preston
Geophysics Department
Sandia National Laboratories
P. O. Box 5800
Albuquerque, New Mexico 87185-MS0750

Abstract

Acoustic full waveform algorithms, such as Paracousti, provide deterministic solutions in complex, 3-D variable environments. In reality, environmental and source characteristics are often only known in a statistical sense. Thus, to fully characterize the expected sound levels within an environment, this uncertainty in environmental and source factors should be incorporated into the acoustic simulations. Performing Monte Carlo (MC) simulations is one method of assessing this uncertainty, but it can quickly become computationally intractable for realistic problems. An alternative method, using the technique of stochastic partial differential equations (SPDE), allows computation of the statistical properties of output signals at a fraction of the computational cost of MC. Paracousti-UQ solves the SPDE system of 3-D acoustic wave propagation equations and provides estimates of the uncertainty of the output simulated wave field (e.g., amplitudes, waveforms) based on estimated probability distributions of the input medium and source parameters. This report describes the derivation of the stochastic partial differential equations, their implementation, and comparison of Paracousti-UQ results with MC simulations using simple models.

ACKNOWLEDGMENTS

This work was funded by the U.S. Department of Energy's Wind and Water Power Technologies Office.

TABLE OF CONTENTS

FIGURES

TABLES

NOMENCLATURE

Abbreviation	Definition
UQ	Uncertainty quantification
MC	Monte Carlo
PDE	Partial differential equation
SPDE	Stochastic partial differential equation
FD	Finite difference
3-D	Three dimensional
PDF	Probability density function

1. INTRODUCTION

The standard partial differential equations (PDEs) governing wave propagation in an acoustic medium are deterministic. Once the sources, boundary conditions, and material properties are defined, the solution is well defined and will give a single answer. If you alter any of the above conditions, you will obtain a different solution. Uncertainties in any of the sources, boundary conditions, or material properties cannot be directly addressed by the solution of these equations. One means of assessing the variability in solutions based on changes in any of the starting conditions is via a Monte Carlo (MC) scheme. In this approach hundreds, thousands, or even more models are produced by drawing random samples from these starting conditions and producing a set of solutions. Classical statistical methods can then be used on the solutions set to derive solution mean, variance, or higher order statistical moments. The primary issue with this approach is that for complex numerical models, running hundreds or more models is sometimes computationally intractable.

Another means of attacking this problem is via solution of stochastic partial differential equations (Ghanem, 1999; Xiu and Karniadakis, 2003). In this approach the physical PDEs are altered to incorporate uncertainty into the solutions directly. The uncertainty is introduced by expanding the physical dependent and independent variables in terms of orthonormal chaos polynomials, which are purely functions of random variables. A probability density function or correlation length are defined for the independent variables, which characterize the uncertainty in these variables (medium properties, boundary conditions, etc.). The solutions of the stochastic PDEs are coefficients of these chaos polynomials that can be used to derive a statistical description of the physical solutions, just as in Monte Carlo methods. However, the primary advantage of the stochastic polynomials is that they require only a fraction of the cost of Monte Carlo methods to achieve equivalent statistical convergence (Xiu and Karniadakis, 2003).

In this report, we will first derive the stochastic partial differential equations for a 3-D linear acoustic medium. The following section will describe the numerical implementation of these equations as embodied in the Paracousti-UQ algorithm. Finally, we will provide validation of the method and implementation by comparing output from Paracousti-UQ to MC simulation result.

2. STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS IN AN ACOUSTIC MEDIUM

2.1. Deterministic Acoustic Partial Differential Equations

The derivation of the system of stochastic PDEs governing linear acoustic wave propagation begins with the deterministic system. The deterministic system of equations we choose to use is the coupled first-order system of partial differential equations called the velocity-pressure system

$$\frac{\partial v_x(\mathbf{x},t)}{\partial t} = -\frac{1}{\rho(\mathbf{x})}\frac{\partial p(\mathbf{x},t)}{\partial x} + \frac{1}{\rho(\mathbf{x})}s_x(\mathbf{x},t) \tag{2.1}$$

$$\frac{\partial v_y(\mathbf{x},t)}{\partial t} = -\frac{1}{\rho(\mathbf{x})}\frac{\partial p(\mathbf{x},t)}{\partial y} + \frac{1}{\rho(\mathbf{x})}s_y(\mathbf{x},t) \tag{2.2}$$

$$\frac{\partial v_z(\mathbf{x},t)}{\partial t} = -\frac{1}{\rho(\mathbf{x})}\frac{\partial p(\mathbf{x},t)}{\partial z} + \frac{1}{\rho(\mathbf{x})}s_z(\mathbf{x},t) \tag{2.3}$$

$$\frac{\partial p(\mathbf{x},t)}{\partial t} = -\kappa(\mathbf{x})\left[\frac{\partial v_x(\mathbf{x},t)}{\partial x} + \frac{\partial v_y(\mathbf{x},t)}{\partial y} + \frac{\partial v_z(\mathbf{x},t)}{\partial z}\right] + \frac{\partial m(\mathbf{x},t)}{\partial t} \tag{2.4}$$

where v_x, v_y, and v_z are the particle velocity components; p is pressure; ρ and κ are the medium density and bulk modulus, respectively. All of the dependent variables (v_x, v_y, v_z, and p) are functions of 3-D space ($\mathbf{x}$) and time (t), whereas we are assuming here that the medium parameters (ρ and κ) are functions of 3-D space only. Sources (s_x, s_y, s_z, and m) are also functions of space and time and are defined by

$$s_x(\mathbf{x},t) = f_x(\mathbf{x},t) + \frac{\partial m^a_{xx}(\mathbf{x},t)}{\partial x} + \frac{\partial m^a_{xy}(\mathbf{x},t)}{\partial y} + \frac{\partial m^a_{xz}(\mathbf{x},t)}{\partial z} \tag{2.5}$$

$$s_y(\mathbf{x},t) = f_y(\mathbf{x},t) + \frac{\partial m^a_{xy}(\mathbf{x},t)}{\partial x} + \frac{\partial m^a_{yy}(\mathbf{x},t)}{\partial y} + \frac{\partial m^a_{yz}(\mathbf{x},t)}{\partial z} \tag{2.6}$$

$$s_z(\mathbf{x},t) = f_z(\mathbf{x},t) + \frac{\partial m^a_{xz}(\mathbf{x},t)}{\partial x} + \frac{\partial m^a_{yz}(\mathbf{x},t)}{\partial y} + \frac{\partial m^a_{zz}(\mathbf{x},t)}{\partial z} \tag{2.7}$$

$$m(\mathbf{x},t) = -\frac{1}{3}\left[m^s_{xx}(\mathbf{x},t) + m^s_{yy}(\mathbf{x},t) + m^s_{zz}(\mathbf{x},t)\right] \tag{2.8}$$

where m^a_{ij} are components of the antisymmetric portion of the 3 by 3 moment tensor; m^s_{ii} are the diagonal components of the symmetric portion of the moment tensor; f_x, f_y, and f_z are components of a body force. m then is simply an isotropic pressure source like an explosion or implosion. Typically, all m^a_{ij} are zero, so that only the body force components will act as sources for the velocity equations. Please see Preston (2016) for a more detailed description of these equations.

2.2. Stochastic Acoustic Partial Differential Equations

The first stage in deriving the stochastic acoustic PDEs is to expand all dependent and independent variables as sums over orthogonal basis functions of random variables, which are assumed to be independent of both space and time. For reasons that will be explained below, the orthogonal basis functions used are taken from a class of orthogonal polynomials called chaos polynomials. The dependent and independent variables become

$$v_i(\mathbf{x},t) = \sum_{j=0}^{N} V_{i,j}(\mathbf{x},t)\Phi_j(\theta) \tag{2.9}$$

$$p(\mathbf{x},t) = \sum_{j=0}^{N} P_j(\mathbf{x},t)\Phi_j(\theta) \tag{2.10}$$

$$\frac{1}{\rho(\mathbf{x},t)} = \sum_{j=0}^{N_r} R_j(\mathbf{x},t)\Phi_j(\theta) \tag{2.11}$$

$$\kappa(\mathbf{x},t) = \sum_{j=0}^{N_k} K_j(\mathbf{x},t)\Phi_j(\theta) \tag{2.12}$$

$$s_i(\mathbf{x},t) = \sum_{j=0}^{N_{s_i}} S_{i,j}(\mathbf{x},t)\Phi_j(\theta) \tag{2.13}$$

$$m(\mathbf{x},t) = \sum_{j=0}^{N_m} M_j(\mathbf{x},t)\Phi_j(\theta) \tag{2.14}$$

In these equations, i is an index variable and stands in place of (x, y, or z); $V_{i,j}$, $S_{i,j}$, P_j, R_j, K_j, and M_j are the expansion coefficients for the chaos polynomial basis function of order j, $\Phi_j(\theta)$, of random variable(s) θ. Note that the buoyancy (reciprocal of density) is expanded in Equation 2.11 instead of density itself. This simplifies the derivation of the equations below.

Note that there are five different maximum orders in the sums for Equations 2.9-2.14: N, N_r, N_k, N_{s_i}, and N_m. The latter four are needed for the independent variable expansions. They are independent of each other and are dependent solely on the maximum chaos polynomial order required to adequately represent the underlying probability distribution for that medium parameter or source. As will be discussed in more detail, one chooses which particular class of chaos polynomials to use based on the one that will represent the medium parameter and/or source probability distribution with the fewest number of expansion coefficients. If any of these maximum orders are zero, then it is equivalent to assuming that parameter is known exactly.

The maximum order for the dependent variables, N, is user adjustable. To be formally correct, except in certain very simple situations, N would need to be infinite; however,

for practical reasons, we must choose to truncate the expansion at some N. How large N needs to be depends on the probability density functions, the class of chaos polynomials chosen, the number of independent random variables, and the set of PDEs being solved. This will be subject of future research. For now, however, we will assume some N has been suitably chosen.

To proceed with the derivation, we place Equations 2.9-2.14 into Equations 2.1-2.4, recalling that the chaos polynomial basis functions are independent of time and space

$$\sum_{j=0}^{N} \frac{\partial V_{x,j}}{\partial t}\Phi_j(\theta) = -\sum_{i=0}^{N}\sum_{j=0}^{N_r} R_j \frac{\partial P_i}{\partial x}\Phi_i(\theta)\Phi_j(\theta) + \sum_{i=0}^{N_{s_x}}\sum_{j=0}^{N_r} R_j S_{x,i}\Phi_i(\theta)\Phi_j(\theta) \tag{2.15}$$

$$\sum_{j=0}^{N} \frac{\partial V_{y,j}}{\partial t}\Phi_j(\theta) = -\sum_{i=0}^{N}\sum_{j=0}^{N_r} R_j \frac{\partial P_i}{\partial y}\Phi_i(\theta)\Phi_j(\theta) + \sum_{i=0}^{N_{s_y}}\sum_{j=0}^{N_r} R_j S_{y,i}\Phi_i(\theta)\Phi_j(\theta) \tag{2.16}$$

$$\sum_{j=0}^{N} \frac{\partial V_{z,j}}{\partial t}\Phi_j(\theta) = -\sum_{i=0}^{N}\sum_{j=0}^{N_r} R_j \frac{\partial P_i}{\partial z}\Phi_i(\theta)\Phi_j(\theta) + \sum_{i=0}^{N_{s_z}}\sum_{j=0}^{N_r} R_j S_{z,i}\Phi_i(\theta)\Phi_j(\theta) \tag{2.17}$$

$$\begin{aligned}\sum_{j=0}^{N} \frac{\partial P_j}{\partial t}\Phi_j(\theta) = &-\sum_{i=0}^{N}\sum_{j=0}^{N_k} K_j \left[\frac{\partial V_{x,i}}{\partial x} + \frac{\partial V_{y,i}}{\partial y} + \frac{\partial V_{z,i}}{\partial z}\right]\Phi_i(\theta)\Phi_j(\theta) \\ &+ \sum_{i=0}^{N_m} \frac{\partial M_i}{\partial t}\Phi_i(\theta)\end{aligned} \tag{2.18}$$

Now we determine the inner product of each equation with $\Phi_k(\theta)$ using the definition of inner product

$$\langle f(x)g(x)\rangle = \int_a^b f(x)g(x)w(x)\mathrm{d}x \tag{2.19}$$

where $w(x)$ is the weight function specific to the chosen chaos polynomial and a and b confine the integration to the support space of the polynomial.

Applying Equation 2.19 with $\Phi_k(\theta)$ and assuming that we have normalized the basis functions relative to this inner product definition so that we have orthonormal chaos polynomials, we can use the orthonormality of the basis functions to obtain

$$\frac{\partial V_{x,k}}{\partial t} = -\sum_{i=0}^{N}\sum_{j=0}^{N_r} e_{ijk} R_j \frac{\partial P_i}{\partial x} + \sum_{i=0}^{N_{s_x}}\sum_{j=0}^{N_r} e_{ijk} R_j S_{x,i} \tag{2.20}$$

$$\frac{\partial V_{y,k}}{\partial t} = -\sum_{i=0}^{N}\sum_{j=0}^{N_r} e_{ijk} R_j \frac{\partial P_i}{\partial y} + \sum_{i=0}^{N_{s_y}}\sum_{j=0}^{N_r} e_{ijk} R_j S_{y,i} \tag{2.21}$$

$$\frac{\partial V_{z,k}}{\partial t} = -\sum_{i=0}^{N}\sum_{j=0}^{N_r} e_{ijk} R_j \frac{\partial P_i}{\partial z} + \sum_{i=0}^{N_{s_z}}\sum_{j=0}^{N_r} e_{ijk} R_j S_{z,i} \tag{2.22}$$

$$\frac{\partial P_k}{\partial t} = -\sum_{i=0}^{N}\sum_{j=0}^{N_k} e_{ijk} K_j \left[\frac{\partial V_{x,i}}{\partial x} + \frac{\partial V_{y,i}}{\partial y} + \frac{\partial V_{z,i}}{\partial z}\right] + \frac{\partial M_k}{\partial t} \tag{2.23}$$

where

$$e_{ijk} = \left\langle \Phi_i \Phi_j \Phi_k \right\rangle \tag{2.24}$$

Equation 2.24 is known and can be tabulated for each i, j, k triple beforehand.

The system of equations defined by Equations 2.20-2.23 form the stochastic PDEs for an acoustic medium. These are similar in form to Equations 2.1-2.4; indeed, if one assumes that all independent variables are exactly known (equating $N_r = N_k = N_{s_i} = N_m = 0$) then one recovers Equations 2.1-2.4 as one should. Despite their similarity, it is obvious that there are many additional terms in each equation and that there are N sets of the four equations that must be solved, one set for each index k. However, the number of additional terms is not as great as one might first think. e_{ijk} tends to be relatively sparse so many of the additional terms in the sums are trivially zero.

2.2.1. Choice of Chaos Polynomial

The choice of the class of chaos polynomials to use as basis functions should depend upon the class of probability distributions that describe the medium parameters and/or source. An improper choice will lead to larger numbers of expansion coefficients being needed to adequately represent the probability distributions of the independent variables and slow convergence to the true dependent variable distributions (i.e., higher order will be needed for dependent variable expansion)(Xiu and Karniadakis, 2003).

Xiu and Karniadakis (2003) provide a table that relates the optimal chaos polynomial for specific probability distribution functions. We repeat the most relevant ones in Table 1.

Table 1. Optimal Chaos Polynomials

PDF	chaos polynomial	Support	weight function
Gaussian	Hermite	$(-\infty, +\infty)$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
Gamma	Laguerre	$[0, +\infty)$	$\frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)}$
Uniform	Legendre	$[a, b]$	$\frac{1}{b-a}$

The reason these polynomials are optimally suited for their particular probability distribution function (PDF) can be seen in their respective weight functions. Each weight function is similar, if not identical, to the generic PDF. For example, the weight function for Hermite polynomials is the PDF for a gaussian with zero mean and unit standard deviation. Likewise, the weight function for Laguerre polynomials is a scaled version of the gamma distribution.

Generally for acoustic models, the gaussian or gamma distributions would seem to be the PDFs that would represent the independent variables best. The medium parameters density and bulk modulus must be positive numbers; thus, the gamma distribution is likely the best representation for these parameters given the support space of Laguerre chaos polynomials. Sources, on the other hand, could have positive or negative values; thus, a gaussian distribution would be a proper choice.

When both sources and medium parameters must be represented with uncertainty, one must choose a single chaos polynomial class for the entire problem. In this case, the choice will be based on which PDF best represents the problem as a whole. For example, when the standard deviation for a gamma distribution is small relative to its mean, a gaussian distribution can adequately approximate that gamma distribution.

2.2.2. Expansion Coefficients for Independent Variables

According to Equations 2.20-2.23, the expansion coefficients for the medium parameters and sources are needed to solve the system of equations. Provided that one has a functional form of the PDFs for the medium parameters and source, one can find the coefficients in the manner typical for polynomial expansion with weight functions over a support space:

$$f_j = \left\langle f(z)\Phi_j(z) \right\rangle = \int_a^b f(z)\Phi_j(z)w(z)\mathrm{d}z \tag{2.25}$$

where $f(z)$ is the probability distribution function for the independent variable, $w(z)$ is the weight function associated with the particular chaos polynomial, and integration limits a and b are the limits of the support space.

If one chooses the correct chaos polynomial basis functions for the given PDF, one will minimize the required number of coefficients necessary to accurately represent that PDF. For example, if a medium parameter conforms to a Gaussian distribution with a given mean and variance, Hermite chaos polynomials can *exactly* represent it with two coefficients. The coefficients for this particular case are

$$f_0 = \mu \tag{2.26}$$

$$f_1 = \sigma \tag{2.27}$$

where μ is the mean of the gaussian distribution and σ is its standard deviation. All other coefficients are zero.

Similarly, for a medium parameter that has a gamma PDF, Laguerre chaos polynomials can exactly represent that distribution with only two coefficients

$$f_0 = \mu \tag{2.28}$$

$$f_1 = -\sigma \tag{2.29}$$

where again μ is the mean of the gamma distribution and σ is its standard deviation (note the minus sign). However, gamma distributions are not typically functionally defined based on their mean and standard deviation. One typical definition uses parameters α and β and defines the gamma distribution as

$$f(z) = \frac{\beta^{\alpha} e^{-\beta z} z^{\alpha-1}}{\Gamma(\alpha)} \tag{2.30}$$

where $\Gamma(\alpha)$ is the gamma function and $\alpha > 0$ and $\beta > 0$ (e.g., Kurtz, 1991). With this definition, the mean and variance are related to α and β via

$$\mu = \frac{\alpha}{\beta} \tag{2.31}$$

$$\sigma^2 = \frac{\alpha}{\beta^2} \tag{2.32}$$

3. STOCHASTIC FINITE DIFFERENCE IMPLEMENTATION

3.1. Finite Difference Scheme

The first stage in implementing the system of partial differential equations listed in Equations 2.20-2.23 is discretization. Like Paracousti (Preston, 2016a), we utilize a standard staggered grid as a basis for discretizing these equations. In this scheme medium parameters (density and bulk modulus) and the pressure variables reside at the corners of grid cells. Velocity variables reside on the centers of the edges of the cell: v_x is midway in the x-direction of the cell edges; v_y is midway in the y-direction; and v_z is midway along the z-direction edges (Figure 1). Time discretization is also staggered with pressures being updated at the integer time steps and velocities at the half-integer time steps. Staggering of the space and time variables allows compact, centered finite-difference (FD) operators to be used. The equations are discretized with fourth order accurate spatial operators and second order temporal accuracy using standard Taylor series coefficients.

3.2. Finite Difference Equations

The following equations are the finite difference equivalents to Equations 2.20-2.23 without sources. Sources will be addressed separately. In these equations i, j, and k refer to spatial indices; l is the temporal index; m, n, and p refer to the polynomial coefficient orders of the terms. The fourth order finite difference coefficients are

$$c_0 = \frac{9}{8}\frac{1}{h} \tag{3.1}$$

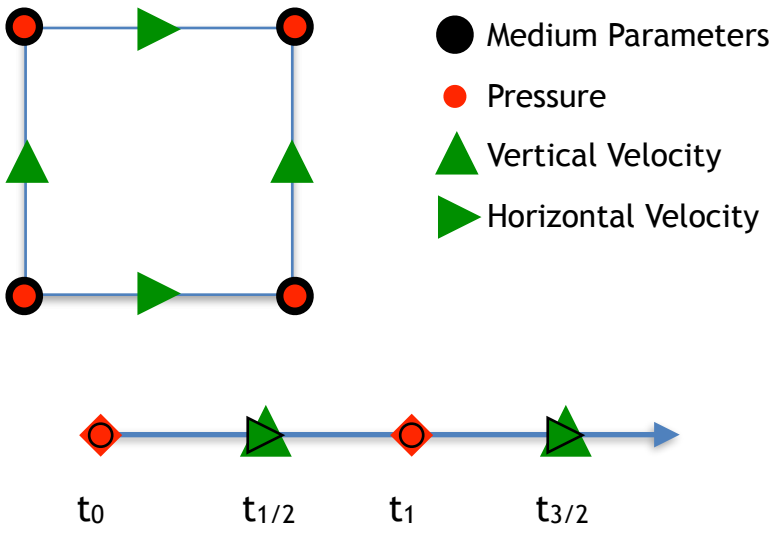

Figure 1: Unit cell (top) and time axis (bottom) for the staggered finite-difference scheme.

$$c_1 = -\frac{1}{24}\frac{1}{h} \tag{3.2}$$

where h is the grid spacing.

Note that since density is provided on the pressure nodes, it must be interpolated onto the velocity node points using second order interpolation. The pressure updating equations do not require any interpolation of medium parameters.

3.2.1. X-Component of Velocity

$$\begin{aligned} v_x^{(p)}\left(x_{i+1/2}, y_j, z_k, t_{l+1/2}\right) &= v_x^{(p)}\left(x_{i+1/2}, y_j, z_k, t_{l-1/2}\right) \\ &-d_t \sum_{m=0}^{N} \sum_{n=0}^{N_r} e_{mnp} R^{(n)}\left(x_{i+1/2}, y_j, z_k\right) \times \\ &\left[c_0\left[p^{(m)}\left(x_{i+1}, y_j, z_k, t_l\right) - p^{(m)}\left(x_i, y_j, z_k, t_l\right)\right]\right. \\ &\left. +c_1\left[p^{(m)}\left(x_{i+2}, y_j, z_k, t_l\right) - p^{(m)}\left(x_{i-1}, y_j, z_k, t_l\right)\right]\right] \end{aligned} \tag{3.3}$$

3.2.2. Y-Component of Velocity

$$\begin{aligned} v_y^{(p)}\left(x_i, y_{j+1/2}, z_k, t_{l+1/2}\right) &= v_y^{(p)}\left(x_i, y_{j+1/2}, z_k, t_{l-1/2}\right) \\ &-d_t \sum_{m=0}^{N} \sum_{n=0}^{N_r} e_{mnp} R^{(n)}\left(x_i, y_{j+1/2}, z_k\right) \times \\ &\left[c_0\left[p^{(m)}\left(x_i, y_{j+1}, z_k, t_l\right) - p^{(m)}\left(x_i, y_j, z_k, t_l\right)\right]\right. \\ &\left. +c_1\left[p^{(m)}\left(x_i, y_{j+2}, z_k, t_l\right) - p^{(m)}\left(x_i, y_{j-1}, z_k, t_l\right)\right]\right] \end{aligned} \tag{3.4}$$

3.2.3. Z-Component of Velocity

$$\begin{aligned} v_z^{(p)}\left(x_i, y_j, z_{k+1/2}, t_{l+1/2}\right) &= v_z^{(p)}\left(x_i, y_j, z_{k+1/2}, t_{l-1/2}\right) \\ &-d_t \sum_{m=0}^{N} \sum_{n=0}^{N_r} e_{mnp} R^{(n)}\left(x_i, y_j, z_{k+1/2}\right) \times \\ &\left[c_0\left[p^{(m)}\left(x_i, y_j, z_{k+1}, t_l\right) - p^{(m)}\left(x_i, y_j, z_k, t_l\right)\right]\right. \\ &\left. +c_1\left[p^{(m)}\left(x_i, y_j, z_{k+2}, t_l\right) - p^{(m)}\left(x_i, y_j, z_{k-1}, t_l\right)\right]\right] \end{aligned} \tag{3.5}$$

3.2.4. Pressure

$$
\begin{aligned}
p^{(p)}\left(x_i, y_j, z_k, t_{l+1}\right) &= p^{(p)}\left(x_i, y_j, z_k, t_l\right) \\
&\quad -d_t \sum_{m=0}^{N} \sum_{n=0}^{N_k} e_{mnp} K^{(n)}\left(x_i, y_j, z_k\right) \times \\
&\quad \left[c_0 \left[v_x^{(m)}\left(x_{i+1/2}, y_j, z_k, t_{l+1/2}\right) - v_x^{(m)}\left(x_{i-1/2}, y_j, z_k, t_{l+1/2}\right) \right] \right. \\
&\quad + c_1 \left[v_x^{(m)}\left(x_{i+3/2}, y_j, z_k, t_{l+1/2}\right) - v_x^{(m)}\left(x_{i-3/2}, y_j, z_k, t_{l+1/2}\right) \right] \\
&\quad + c_0 \left[v_y^{(m)}\left(x_i, y_{j+1/2}, z_k, t_{l+1/2}\right) - v_y^{(m)}\left(x_i, y_{j-1/2}, z_k, t_{l+1/2}\right) \right] \\
&\quad + c_1 \left[v_x^{(m)}\left(y_i, y_{j+3/2}, z_k, t_{l+1/2}\right) - v_y^{(m)}\left(x_i, y_{j-3/2}, z_k, t_{l+1/2}\right) \right] \\
&\quad + c_0 \left[v_z^{(m)}\left(x_i, y_j, z_{k+1/2}, t_{l+1/2}\right) - v_z^{(m)}\left(x_i, y_j, z_{k-1/2}, t_{l+1/2}\right) \right] \\
&\quad \left. + c_1 \left[v_x^{(m)}\left(y_i, y_j, z_{k+3/2}, t_{l+1/2}\right) - v_y^{(m)}\left(x_i, y_j, z_{k-3/2}, t_{l+1/2}\right) \right] \right]
\end{aligned}
\tag{3.6}
$$

3.3. Sources

Sources may be placed at any location within the domain; however, it is ill advised to place sources within the absorbing boundary zones. When a source does not lie directly on a node, it is trilinearly extrapolated to the eight surrounding grid nodes. The finite difference updating formulae given in Equations 3.3-3.6 can be augmented with the source terms from Equations 2.20-2.23. These assume that Equations 3.3-3.6 have already been applied so the sources are added to the current time step index.

3.3.1. X-Component of Force

$$
\begin{aligned}
v_x^{(p)}\left(x_{i+1/2}, y_j, z_k, t_{l+1/2}\right) &= v_x^{(p)}\left(x_{i+1/2}, y_j, z_k, t_{l+1/2}\right) \\
&\quad + d_t \sum_{m=0}^{N_{s_x}} \sum_{n=0}^{N_r} e_{mnp} R^{(n)}\left(x_{i+1/2}, y_j, z_k\right) S_x^{(m)}\left(x_{i+1/2}, y_j, z_k, t_l\right)
\end{aligned}
\tag{3.7}
$$

3.3.2. Y-Component of Force

$$
\begin{aligned}
v_y^{(p)}\left(x_i, y_{j+1/2}, z_k, t_{l+1/2}\right) &= v_y^{(p)}\left(x_i, y_{j+1/2}, z_k, t_{l+1/2}\right) \\
&\quad + d_t \sum_{m=0}^{N_{s_y}} \sum_{n=0}^{N_r} e_{mnp} R^{(n)}\left(x_i, y_{j+1/2}, z_k\right) S_y^{(m)}\left(x_i, y_{j+1/2}, z_k, t_l\right)
\end{aligned}
\tag{3.8}
$$

3.3.3. Z-Component of Force

$$v_z^{(p)}\left(x_i, y_j, z_{k+1/2}, t_{l+1/2}\right) = v_z^{(p)}\left(x_i, y_j, z_{k+1/2}, t_{l+1/2}\right) + d_t \sum_{m=0}^{N_{s_z}} \sum_{n=0}^{N_r} e_{mnp} R^{(n)}\left(x_i, y_j, z_{k+1/2}\right) S_z^{(m)}\left(x_i, y_j, z_{k+1/2}, t_l\right) \tag{3.9}$$

3.3.4. Isotropic Source

$$p^{(p)}\left(x_i, y_j, z_k, t_{l+1}\right) = p^{(p)}\left(x_i, y_j, z_k, t_{l+1}\right) + M^{(p)}\left(x_i, y_j, z_k, t_{l+1}\right) - M^{(p)}\left(x_i, y_j, z_k, t_l\right) \tag{3.10}$$

3.3.5. Source Time Functions

For deterministic sources, one can provide any arbitrary source time function for the source or specify a delta function source. In the latter case, the output traces are Greens Functions and can be convolved with any source time function to obtain results just as if that source time function had been used originally. This provides a computationally efficient mechanism to try a variety of source time functions with a single run of the algorithm. Besides being able to provide any arbitrary source time function via an input text file, there are some ready-made source time functions that Paracousti-UQ can generate for the user.

Source uncertainty is actually uncertainty in the source time function. The source location is always assumed known. With uncertainty of the source time function, a separate "source coefficient time function" must be supplied for each expansion coefficient. For example, if the source time function conforms to a gaussian distribution, the first coefficient time function would be the mean of the source amplitude as a function of time. The second coefficient time function specifies the standard deviation of the source amplitude as a function of time.

3.4. Receivers

Receivers are implemented similarly to those in Paracousti. The primary difference is that instead of recording one single component (v_x, v_y, v_z, or p), each receiver records all N orders for each component. Receivers are the primary means of ascertaining the uncertainty in or distribution of the output velocity or pressure at a particular point in space. Interpreting the results will be discussed in the next chapter.

Receivers can be placed anywhere in the domain, but they are ill advised in the absorbing boundary zone or at the source point. These restrictions are due to numerical issues that will cause the traces to be erroneous or hard to interpret. For receivers not directly located on a node, the values are trilinearly interpolated from the surrounding eight nodes.

3.5. Absorbing Boundary Conditions

Due to the finite size of a computational domain, absorbing boundary conditions are required to mitigate unrealistic, numerical reflections from the domain boundaries. Just as in Paracousti, we utilize convolutional perfectly matched layers (CPML; Komatitsch and Martin, 2007) as absorbing boundary zones. These can be used exactly as they are in Paracousti, except that there are now N times the number of memory variables: a separate one for each dependent variable coefficient. Additionally, care must be exercised to ensure that each memory variable is updated only once per time step. The sums in Equations 3.3-3.6 have to be handled carefully in the absorbing boundary zones to ensure this.

3.6. Grid Spacing and Time Steps

The optimal grid spacing and time step is an area that needs further research for the stochastic PDE solver. However, in the tests we have done thus far, guidelines borrowed from deterministic solvers such as Paracousti work well. The appropriate node spacing to use in deterministic models is based upon the minimum wave speed in the domain (V_{min})and the maximum frequency (f_{max})that one desires to be simulated,

$$h = G\frac{V_{min}}{f_{max}} \tag{3.11}$$

where G is a constant that is based upon the desired accuracy. Based on numerical phase and group speed curves for elastic media, which in the limit of shear modulus going to zero approaches an acoustic medium, the optimal G is between 0.1 and 0.16 (Preston, 2016b; Aldridge and Haney, 2008). The best value depends on the time step and the model to a certain extent. For models with topography, bathymetry, and other complications, experience shows that a value for closer to 0.1 is optimal, whereas larger values can be used in simpler models.

Since this is an explicit leap-frog time-stepping algorithm, there exists a maximum time step for stability ($\max d_t$), called the Courant-Friedrichs-Lewy (CFL) condition. It is determined by the grid spacing and maximum seismic velocity (V_{max})via

$$\max d_t = \frac{h}{\sqrt{2}V_{max}\sum\left|c_i\right|} \tag{3.12}$$

where the c_i are the finite-difference coefficients given in Equations 3.1-3.2.

Although this is the maximum time step allowed for stability, it is not the optimal choice in deterministic solvers since smaller d_t will provide superior accuracy. Based on von Neumann analysis of the numerical wave speeds (Preston, 2016b) versus computational runtime as a function of d_t, the optimal d_t is approximately $0.6\max d_t$, again for deterministic solvers.

Some obvious questions arise when looking at Equations 3.11 and 3.12: what does V_{min}, V_{max}, or for that matter, f_{max}, mean in a stochastic solver since wave speed and source terms can be from probability distributions? Are the mean values sufficient to

define the extrema or do the width of the probability distributions affect stability and accuracy. Again, these are areas where further research is needed.

3.7. High Contrast Media

Strong contrasts, especially in density, from one grid point to the next is known to cause instabilities and/or inaccuracies in deterministic simulation results. Since the stochastic PDEs reduce to the deterministic ones as our certainty increases, it is reasonable to suppose that the stochastic solvers would suffer from the same instability issues. One solution that produces accurate and stable results uses the order-switching methodology outlined in Preston et al. (2008). In this method, the earth model is scanned prior to time stepping and high contrast points are diagnosed. Finite-difference updating formulae only in the vicinity of these points are altered from fourth to second order accuracy; all other points remain at fourth order accuracy. By limiting the reach of the operators to second order, stability and accuracy can be maintained, while keeping the majority of the dependent variable updates at fourth order accuracy. Currently, only the zeroth order coefficients are scanned for high

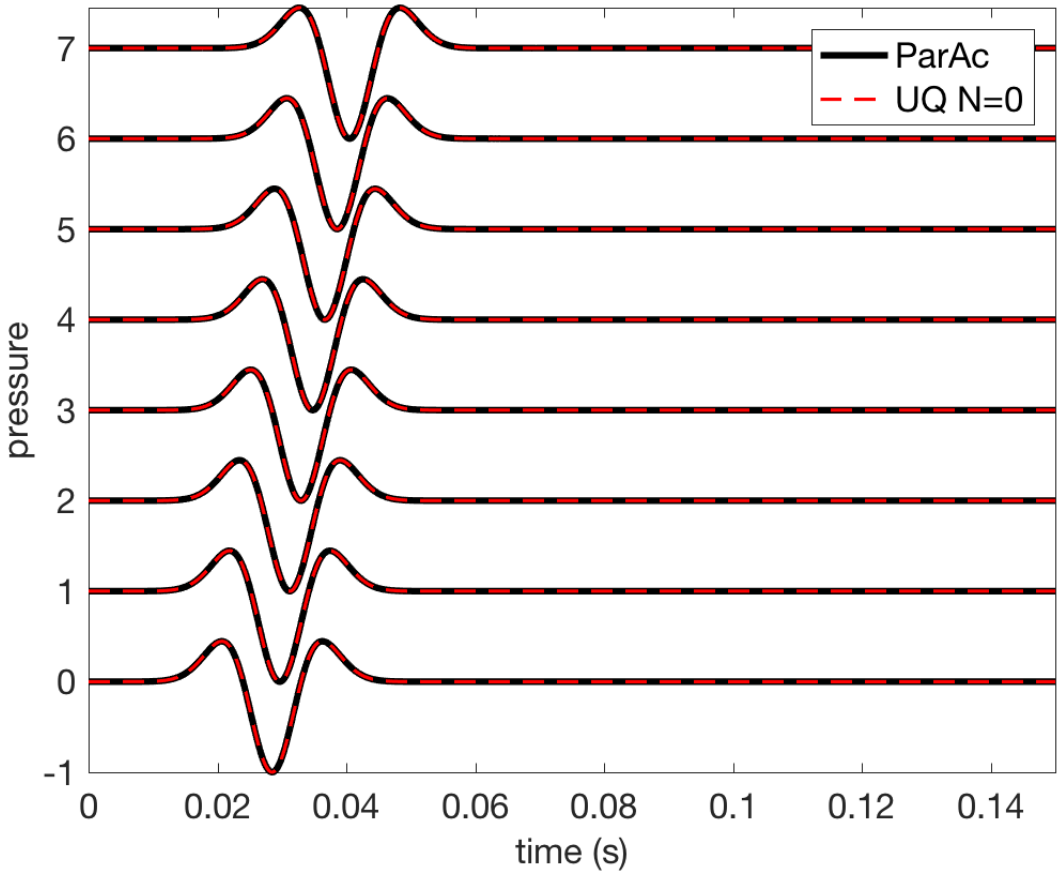

Figure 2: Comparison of pressure traces from deterministic Paracousti (black) and Paracousti-UQ (red dashed) with all maximum chaos polynomial orders set to zero.

contrasts between adjacent grid points. Further research is needed to ascertain whether this is sufficient for stability and accuracy, or whether higher order coefficients must be incorporated into the scanning procedure.

3.8. Massively Parallel Design

Paracousti-UQ is designed to use the Message Passing Framework (MPI). This allows one to utilize multiple cores and physical machines to distribute the workload and reduce runtime. Paracousti-UQ can run on anything from a laptop to 1,000's or more cores on an institutional cluster. This allows one to run small simple models as well as very large complex models with the same code.

This is especially important for Paracousti-UQ due to the nature of the stochastic PDEs that it solves. As noted in Chapter 2, there are many more equations and several more terms in each of those equations compared to a deterministic solver such as Paracousti. Distributing the load across many machines provides the necessary computational power as well as the RAM necessary to hold the extra data required for Paracousti-UQ.

3.9. Verification With Paracousti

When all maximum chaos polynomial orders are set to zero (i.e., $N = 0$ and all medium parameters and source terms are known exactly ($N_r = N_k = N_{s_i} = N_m = 0$)), Equations 3.3-3.6 reduce to those used in Paracousti. To verify that Paracousti and Paracousti-UQ give the same results in this case, we compare pressure traces between the two algorithms from an explosion source in a homogenous medium (Figure 2). The agreement between the two algorithms is excellent, demonstrating that Paracousti-UQ does indeed reduce to Paracousti in this special limiting case.

4. COMPARISON WITH MONTE CARLO SIMULATIONS

4.1. Modeling Scenario

In all of the test cases presented below homogeneous acoustic models were utilized. The baseline problem consisted of a homogeneous medium with sound speed of 2500 m/s and density of 2000 kg/m^3. This translates to a bulk modulus of 12.5e9 Pa and buoyancy of 0.0005 m^3/kg. The model domain extended from -15 m to 100 m on the x-axis, -50 m to 50 m on the y-axis, and -200 m to 200 m on the z-axis, with uniform 1 m grid node spacing. CPML absorbing boundary conditions 10 nodes thick were utilized on all eight sides. The source was placed at position (0,0,0) m with pressure receivers arrayed in a line from 5 m to 40 m, every 5 m, on the x-axis offset 10 m in z from the source. In each case, one of the independent variables (bulk modulus, buoyancy, or source) was assigned a probability distribution while all other independent variables were assumed to be known exactly.

4.2. Monte Carlo Simulations

Although this is a relatively small domain size, due to the estimated number of Monte Carlo runs that would be required to obtain a stable distribution, we did not use Paracousti as the forward solver. Instead, we utilized a 2-D axisymmetric acoustic full waveform solver, an acoustic version of the algorithm axiElasti (Preston, 2017) called axiAcousti, for the Monte Carlo simulations. Given the simple homogeneous models and geometries we are using for these scenarios, a 2-D axisymmetric solver is appropriate. Multiple tests demonstrate that in these simple modeling cases axiAcousti gives identical results to Paracousti and is much faster computationally.

In each case, a well-defined probably distribution was assigned to one of the medium parameter independent variables (bulk modulus or buoyancy), while the other medium parameter is fixed (assumed known). The mean of the distribution is equal to the corresponding value stated in Section 4.1 for that medium parameter in the baseline model. The standard deviation for the medium parameter was approximated to provide a standard deviation in sound speed of about 200 m/s. For bulk modulus this gave a standard deviation of 2e9 Pa; for buoyancy the standard deviation was set to 8e-5 m^3/kg. For these cases, a gamma distribution was assumed.

For tests involving source time function distributions, bulk modulus and buoyancy were fixed at the mean values stated in Section 4.1. The only random variable in this case was the peak amplitude of the source time function. A gaussian random variable with mean 1.0 (J or N) and standard deviation of 0.1 (J or N) was utilized for explosion and force sources.

Since axiAcousti is a deterministic solver, fixed values for all the input values are required. Thus, random samples were drawn from the probability distribution and each sample constituted a single run of axiAcousti. By drawing many samples, we can aggregate all the pressure trace results to obtain an estimate of the mean and standard deviation of the pressure waveforms at the receiver points. To ensure convergence we ran approximately 800,000 axiAcousti runs. Figure 3 shows the

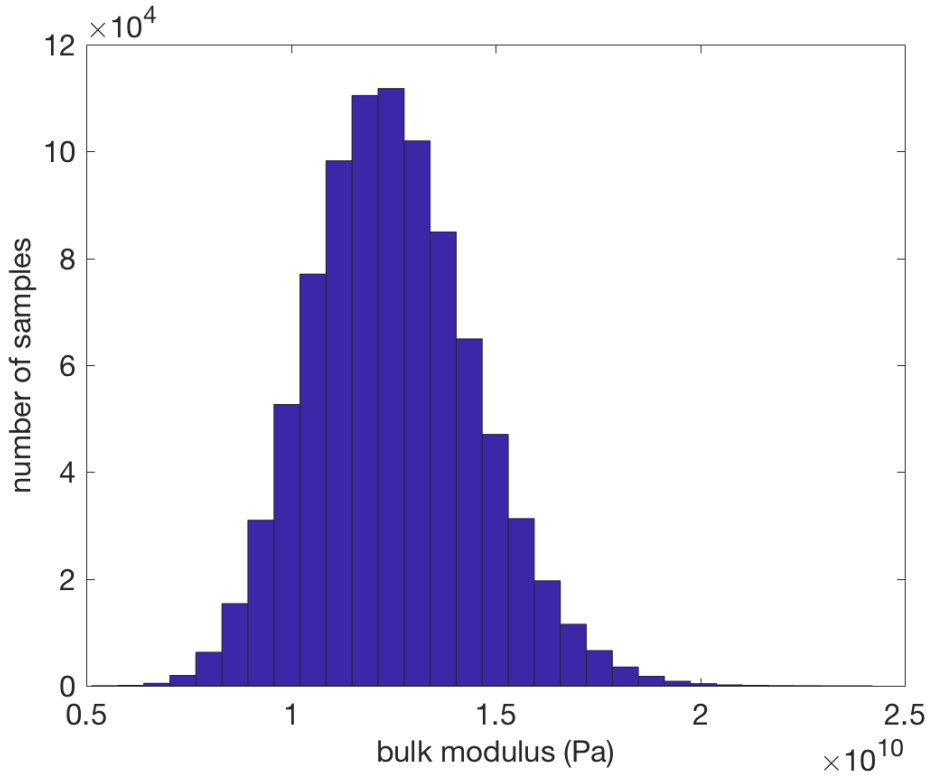

Figure 3: Distribution of bulk modulus used in Monte Carlo simulations.

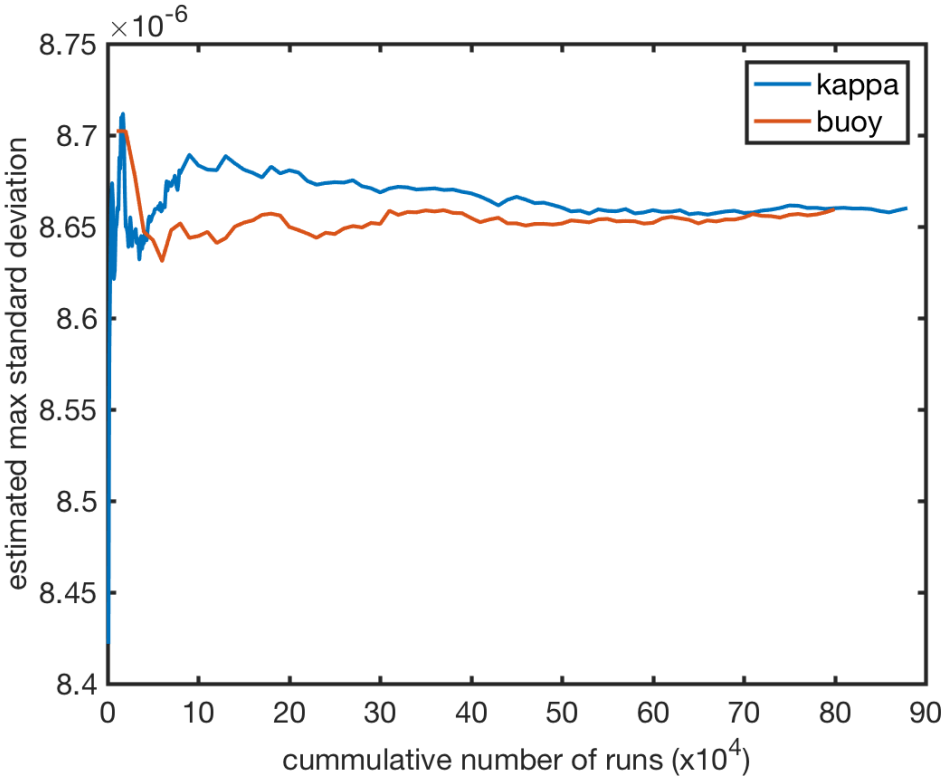

Figure 4: Estimated maximum of standard deviation in pressure at the nearest receiver as a function of cumulative number of Monte Carlo simulations for kappa and buoyancy distributions.

distribution of the actual samples used in the bulk modulus distribution test. The mean of the actual distribution is 12.498e9 Pa with a standard deviation of 1.9994e9 Pa, very close to the designed values.

A question that arises with Monte Carlo methods is how many runs do you need to adequately represent the output distribution. Figure 4 demonstrates the convergence of the Monte Carlo runs for the bulk modulus and buoyancy distribution tests. The image shows the variation in the estimated peak value of the standard deviation waveform at the nearest receiver as a function of the cumulative number of runs. As expected, there is more variation in the estimate for fewer runs, but it stabilizes as runs accumulate. To achieve 0.1% accuracy (assuming the 880,000th run of bulk modulus is closest to the actual distribution) it would require 460,000 runs in the buoyancy test and 390,000 for bulk modulus. Less accuracy requires fewer MC runs. For example, 0.2% accuracy requires 120,000 runs for buoyancy and 210,000 runs for bulk modulus, while 17,000 runs is needed to achieve 0.5% accuracy for bulk modulus (buoyancy was aggregated every 10,000 runs whereas bulk modulus was aggregated every 1,000 runs for the first 80,000 runs). The difference in convergence for bulk modulus and buoyancy is probably due to sampling, but they do obviously converge to the same value. The estimate of the peak of the mean is about an order of magnitude more accurate for the same number of runs relative to the peak of the standard deviation.

4.3. Paracousti-UQ Mean and Standard Deviation

For Monte Carlo simulations, the means and standard deviations can be computed using well-known methods. However, for Paracousti-UQ output the computation of mean and standard deviation cannot be performed using these well-known methods since the solution is a weighted sum of chaos polynomial basis functions.

The general definition of statistical moment is

$$m_n = \int_a^b (x - c)^n \, f(x) \mathrm{d}x \tag{4.1}$$

where $f(x)$ is a probability distribution, a and b are the limits of the support space for that probability distribution, and c and n are constants that define the type and kind of moment we are seeking (e.g., Spanos, 1999). The mean is a moment where $n = 1$ and $c = 0$, so

$$\mu = \int_a^b x f(x) \mathrm{d}x \tag{4.2}$$

For stochastic PDEs, $f(x)$ is simply the weight function (assuming it has been scaled such that $\int_a^b f(x) \mathrm{d}x = 1$, which is true of our definitions in Table 1), and

$$x = \sum_{i=0}^{N} c_i \Phi_i(\theta) \tag{4.3}$$

Placing Equation 4.3 into Equation 4.2, and using the fact that $\Phi_0 = 1$ and the orthonormality properties of the chaos polynomials, one obtains

$$\mu = c_0 \tag{4.4}$$

The variance, or squared standard deviation, is defined by setting $c = \mu$ and $n = 2$ in Equation 4.1:

$$\sigma^2 = \int_a^b \left(x - \mu\right)^2 f(x) \mathrm{d}x \tag{4.5}$$

After some manipulation, one finds for stochastic PDEs

$$\sigma^2 = \sum_{i=1}^{N} c_i^2 \tag{4.6}$$

Note that the sum does not include c_0.

The output from Paracousti-UQ are the c_i up to order N for pressure and/or velocities at each receiver as a function of time. Thus, one obtains the mean and standard deviation at each receiver as a function of time. Higher order moments can also be computed using the same methodology outlined in this section. For this report, we will only be investigating the mean and standard deviation of the output signals.

Table 2: Approximate Number of Deterministic Runs for Maximum Polynomial Orders

Max Order	Equivalent Number of Runs (Gamma)	Equivalent Number of Runs (Gaussian)
0	1	1
1	5	4
2	9	7
3	13	10
4	17	13
5	21	16
6	25	19

4.4. Comparison Cases

We compare the accuracy of the mean and standard deviations obtained via MC methods with those computed from Paracousti-UQ output. In all cases, Paracousti-UQ used up to 5th order chaos polynomials for the dependent variables. This is approximately equivalent to running the deterministic Paracousti 21 times for a gamma distribution. Table 2 gives the approximate number of deterministic runs that is equivalent to running Paracousti-UQ for polynomial orders up to 6 with one random variable for gamma and gaussian distributions. This tabulation is based upon the number of non-zero elements in the e_{ijk} matrix.

4.4.1. Bulk Modulus Distribution with Known Explosion Source

The bulk modulus distribution case uses a gamma distribution with mean of 12.5e9 Pa and standard deviation of 2e9 Pa. Buoyancy is fixed at 5e-4 m^3/kg and an explosion source with a known gaussian pulse source time function with 50 Hz central frequency. The means and standard deviations for the eight pressure receivers are compared between MC and Paracousti-UQ in Figures 5 and 6. The agreement is excellent. Figure 7 shows the effect of maximum polynomial order on the standard deviation at the nearest receiver. On the visual plot scale only maximum order 1 (clearly) and order 2 (in a few places) are distinguishable from the MC standard deviation. As an aside, it should be noted that the mean signal is not exactly the same

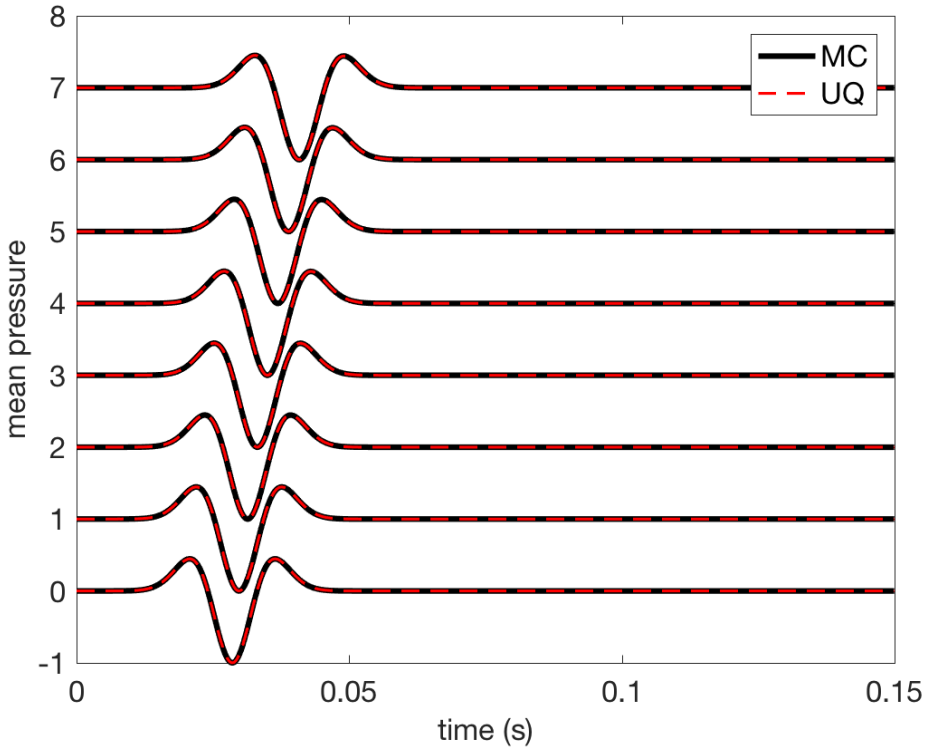

Figure 5: Comparison of mean pressure traces between MC and Paracousti-UQ for bulk modulus distribution.

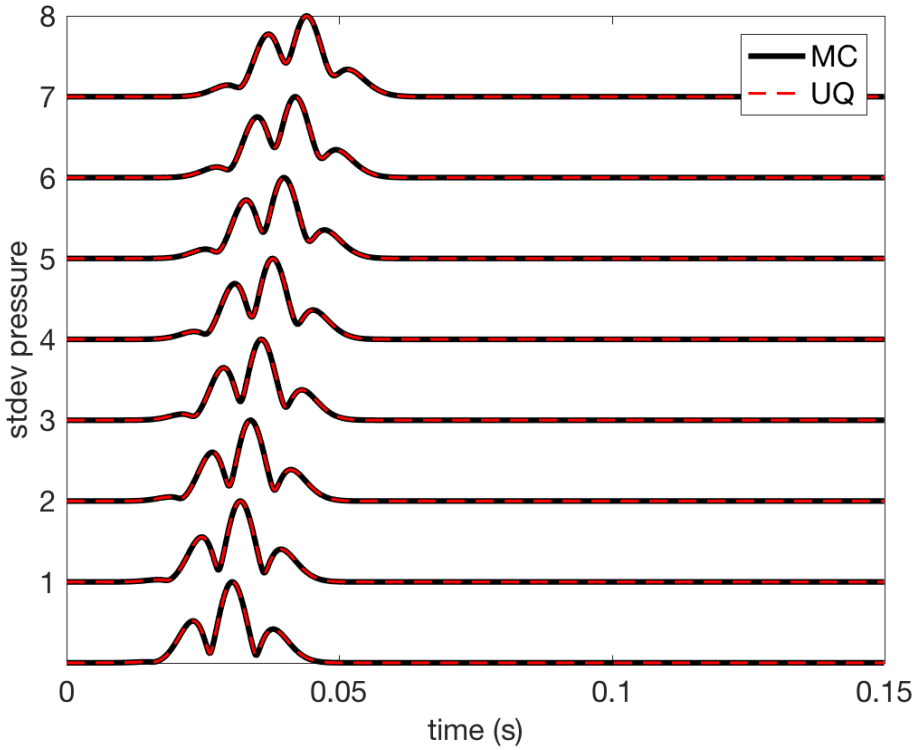

Figure 6: Comparison of standard deviation pressure traces between MC and Paracousti-UQ for bulk modulus distribution.

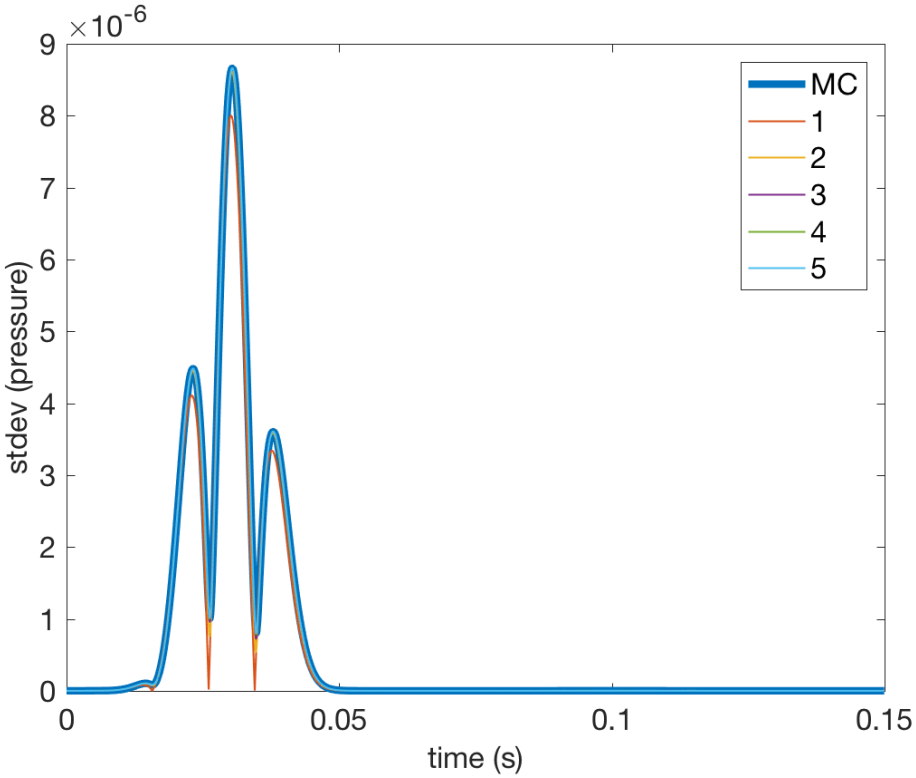

Figure 7: Convergence of Paracousti-UQ based on maximum polynomial order for standard deviation of pressure at the nearest receiver.

as the signal that would be generated for a deterministic model run with the mean bulk modulus and buoyancy, although they are close.

As a comparison of accuracy with the MC standard deviation, the differences in the maximum of the standard deviation at the nearest station is 7.6% for N=1, 0.5% for N=2, and less than 0.05% for N>2. As a speed comparison, one observes that it takes 17,000 MC runs to achieve the same accuracy as the equivalent of 9 runs using Paracousti-UQ. At greater distances from the source, somewhat higher orders are needed to obtain the same accuracy. At these wavelengths at the farthest receiver, N=3 is required instead of N=2 to achieve 0.5% error. These comparisons indicate the efficiency of the stochastic PDE method is at least on the order of a factor of 1,000 greater than MC methods.

4.4.2. Buoyancy Distribution with Known Explosion Source

The buoyancy distribution uses a gamma distribution with a mean of 5e-4 m^3/kg and standard deviation of 8e-5 m^3/kg. The bulk modulus was fixed at 12.5e9 Pa. An explosion source with a known gaussian source time function with 50 Hz central frequency was utilized. Figures 8 and 9 show the mean and standard deviation of the pressure signals at all eight receivers. Note that these are indistinguishable from the mean and standard deviation images for a bulk modulus distribution and explosion source. Again the match between MC and Paacousti-UQ is excellent.

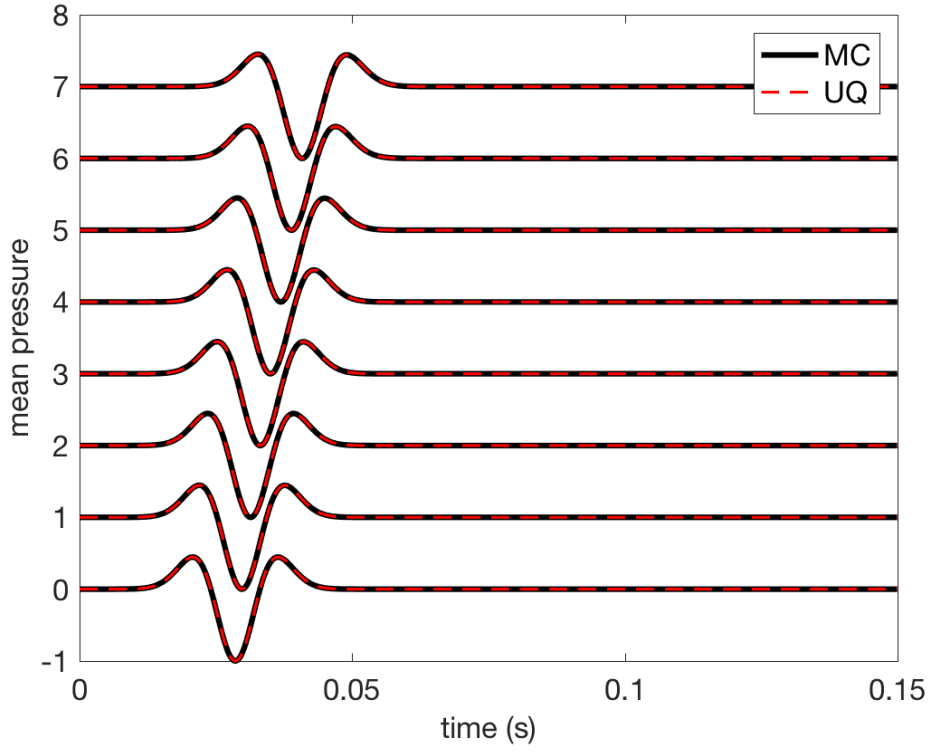

Figure 8: Comparison of mean pressure traces between MC and Paracousti-UQ for buoyancy distribution.

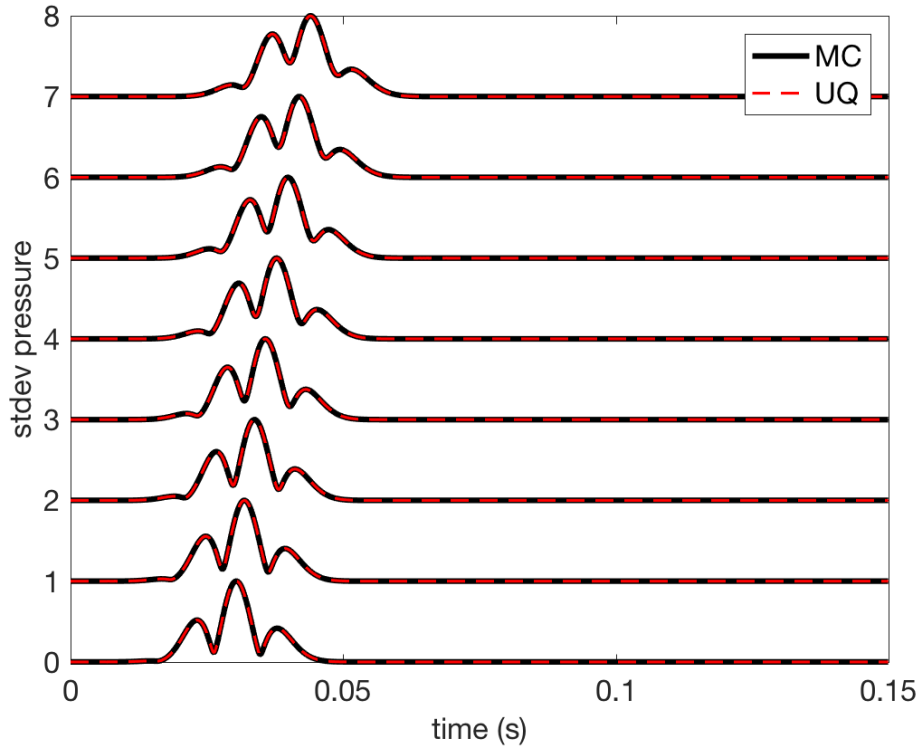

Figure 9: Comparison of standard deviation pressure traces between MC and Paracousti-UQ for buoyancy distribution.

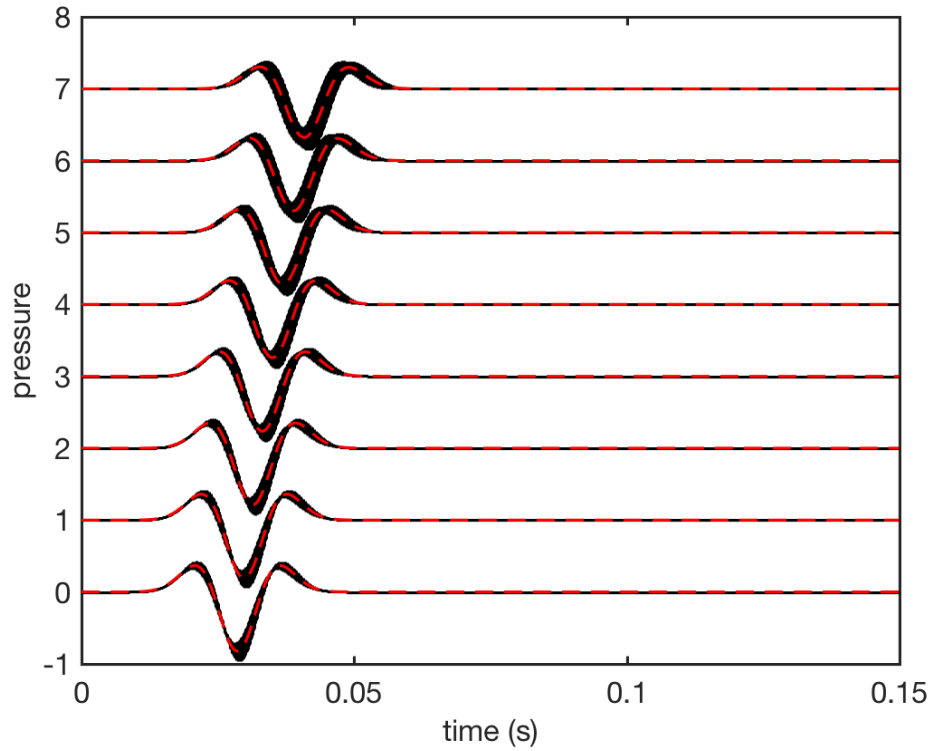

Figure 10: Mean (red dash) and one standard deviation (black) as a function of time for pressure traces.

Figure 10 shows the same mean signal as Figure 8, but also shows the one standard deviation region as a function of time around the mean. Standard deviations are smaller relative to the signal at closer distances from the source, but grow at greater distances. The largest standard deviations lag slightly in time relative to the peaks and troughs of the mean signals

4.4.3. Explosion Source Time Function Distribution

For the explosive source time function distribution, only the peak amplitude is considered to be a random variable. The peak amplitude is drawn from a gaussian distribution with a mean of 1 J and standard deviation of 0.1 J. The source time function is a 50 Hz central frequency gaussian pulse. The bulk modulus and buoyancy were fixed in this test at 12.5e9 Pa and 5e-4 m^3/kg, respectively. In this particular case where only the source is taken from a probability distribution and the medium parameters are considered known, Equations 2.20-2.23 completely decouple in term of polynomial order. In other words, each polynomial order is like a separate solution that does not depend on the behavior of the other orders. For this source time function distribution, there are only two orders required: the mean and standard deviation. Thus, all dependent variable orders greater than 1 should be zero for all times. This is indeed the case; only N=0 and N=1 show any signal at the pressure receivers. These are shown in Figures 11 and 12 for the mean and standard deviation. The true

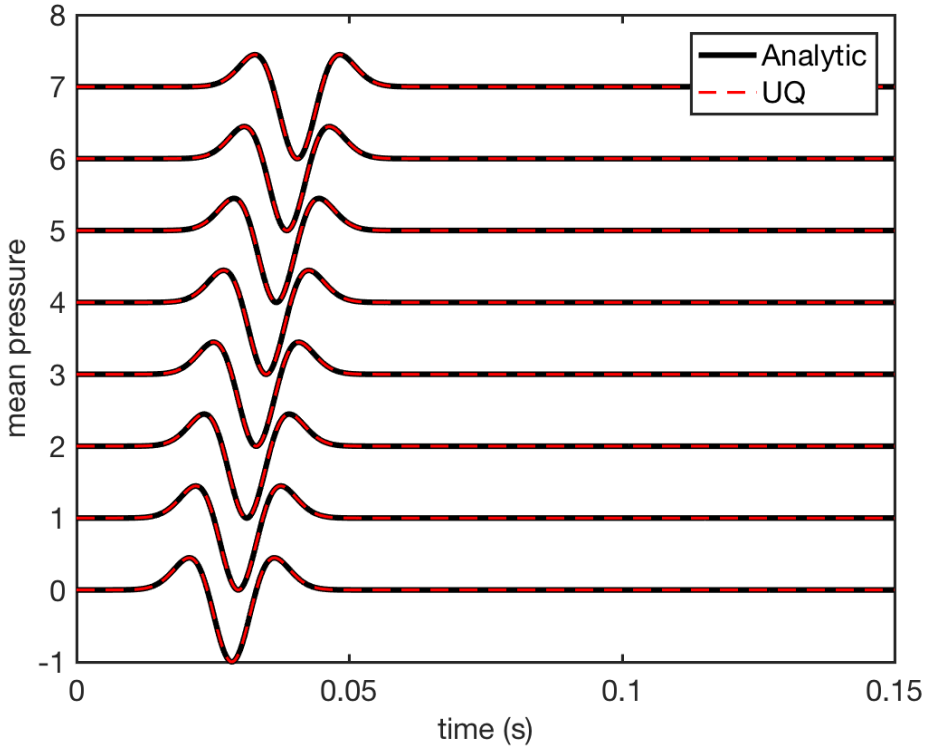

Figure 11: Comparison of the mean pressure traces between the analytic solution and Paracousti-UQ for the explosive source time function distribution.

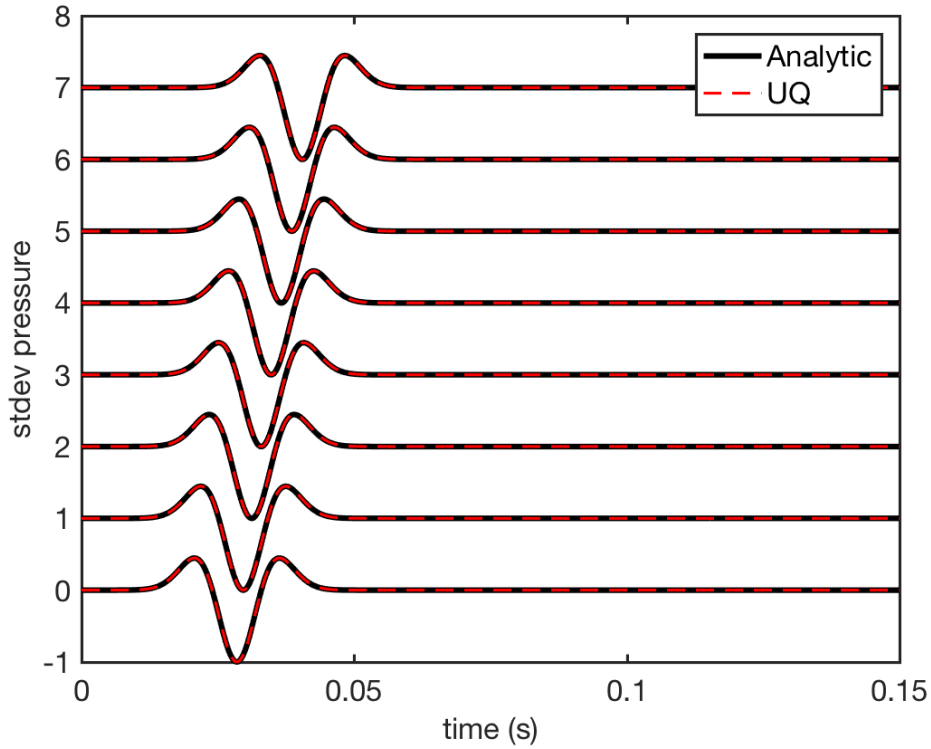

Figure 12: Comparison of the standard deviation pressure traces between the analytic solution and Paracousti-UQ for the explosive source time function distribution.

solutions are known analytically in this case. The mean will be the deterministic solution using the mean source amplitude and the standard deviation will be exactly a scaled version of this with the scale being the ratio of the input standard deviation to its mean, in this case 0.1. The waveforms for the mean and standard deviation are thus the same, just scaled relative to each other. The agreement is again excellent.

4.4.4. Bulk Modulus/Buoyancy with Known Force Source

We also evaluated the accuracy of Paracousti-UQ using a force source. Similar to the explosion source, we performed two tests with a known force source with a 50 Hz central frequency gaussian wavelet. The force was directed parallel to the line of receivers. For the 2-D axisymmetric code, axiAcousti, this required a vertical force source to be coincident with the symmetry axis, and the line of receivers was rotated so that the line is directed in the positive z-direction. Otherwise, the relative distances and offset from the source location is the same as previous tests.

In one test we assume that the bulk modulus is uncertain but buoyancy is known, and in the second we assume the buoyancy is uncertain but bulk modulus is known. Both bulk modulus and buoyancy were again assumed to conform to a gamma distribution. The mean and standard deviation of the bulk modulus was identical to the previous test with a mean of 12.5e9 Pa and standard deviation of 2e9 Pa, with the buoyancy

fixed at 5e-4 m^3/kg. Likewise, for the buoyancy case the mean buoyancy was set to 5e-4 m^3/kg and standard deviation to 8e-5 m^3/kg, with the bulk modulus fixed at 12.5e9 Pa.

The results from these two tests are identical to each other, as they were for the explosion source. A comparison of the mean and standard deviation of the output pressure traces as a function of time between the MC and Paracousti-UQ are displayed in Figures 13 and 14. Additionally, we also show the mean with one standard deviation pressure traces in Figure 15. The difference in waveforms between the source types is expected based on theoretical differences between force and explosive sources (Aldridge, 2000). Explosion sources will produce far-field pressure waveforms proportional to the second time derivative of the source time function, whereas force sources produce waveforms proportional to the first time derivative of the source time function. The results presented here agree with theoretical expectations, and Paracousti-UQ results again match the MC results extremely well.

The mean and standard deviation traces (Figure 15) demonstrate similar properties to those from the explosion source (Figure 10). The standard deviation is relatively small compared to the mean traces at the closest receiver, but the standard deviation grows in relation to the mean as distance increases. Also, the maximum standard deviation lags slightly behind the peak of the mean trace.

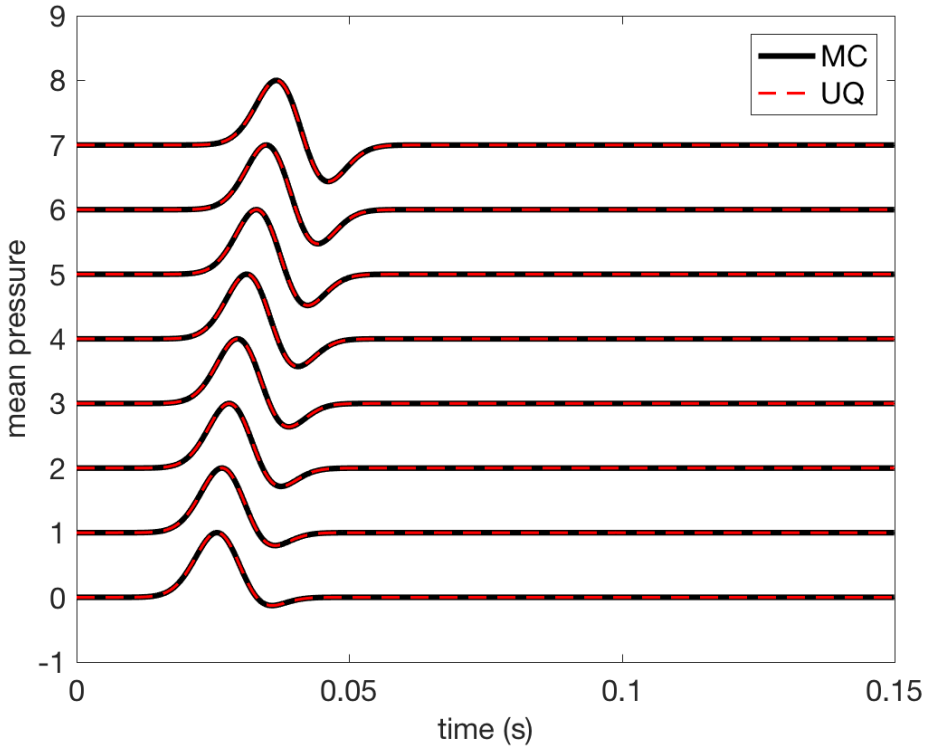

Figure 13: Comparison of the mean pressure traces between Monte Carlo and Paracousti-UQ simulations for a buoyancy distribution with a force source.

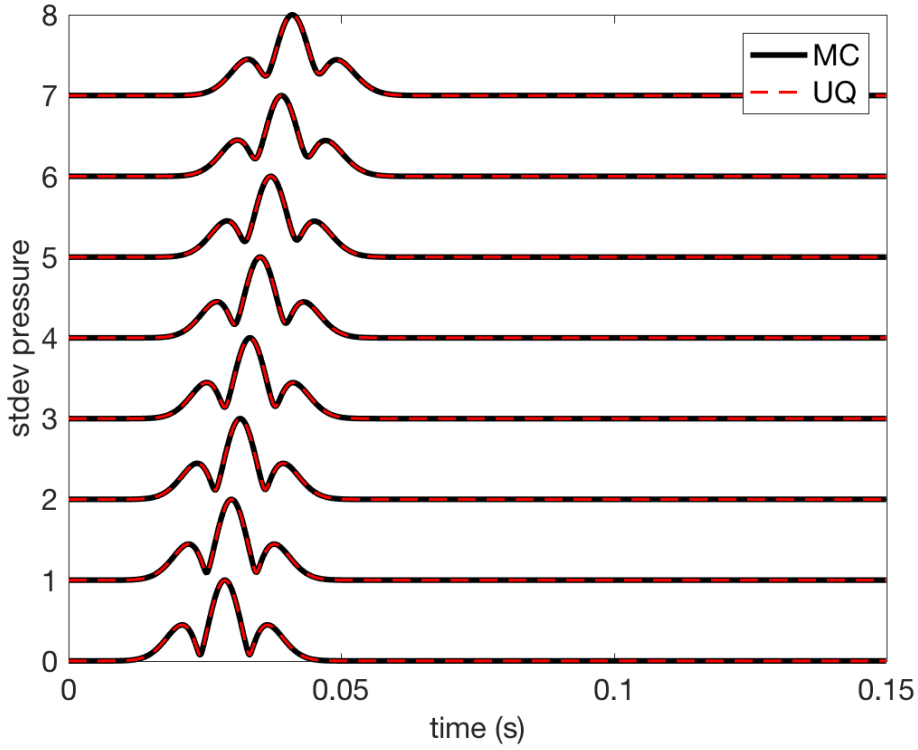

Figure 14: Comparison of the standard deviation of pressure traces between Monte Carlo and Paracousti-UQ simulations for a buoyancy distribution with a force source.

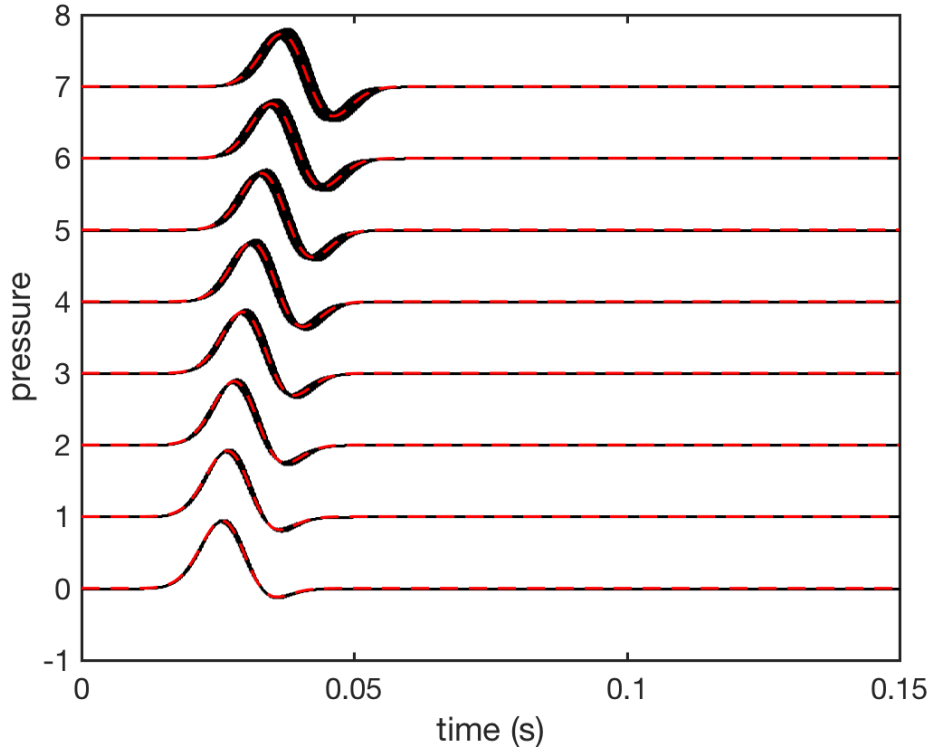

Figure 15: Mean (red dash) and one standard deviation region (black) for pressure traces from a force source.

4.4.5. *Force Source Time Function Distribution*

The final test case uses fixed bulk modulus and buoyancy, but the maximum amplitude of the source time function derives from a gaussian distribution. Similar to the explosion case, the mean of the distribution is set to 1 N with a standard deviation of 0.1 N. In an identical manner to the explosion source, when the only uncertainty is the source and the medium parameters are known exactly, the stochastic PDEs decouple with respect to the polynomial order. Thus, all traces with output polynomial order greater than 1 are identically zero, as expected. The mean of the output pressure traces is just the solution to the deterministic acoustic PDEs using the mean amplitude of 1 N, and the standard deviation has identical waveforms to the mean but scaled by 0.1 due to the ratio between input standard deviation and mean. These predictions are confirmed by the Paracousti-UQ output. Figure 16 shows the mean and one standard deviation in a single figure. In this case, the peak of the standard deviation aligns with the peak of the mean and the relative size of the standard deviation to the mean remains unchanged as source-receiver distance increases.

4.5. Discussion

We have demonstrated with several simple examples that Paracousti-UQ gives results in excellent agreement with Monte Carlo simulations. Additionally, we have indicated that Paracousti-UQ is at least 1,000 times faster than Monte Carlo at achieving

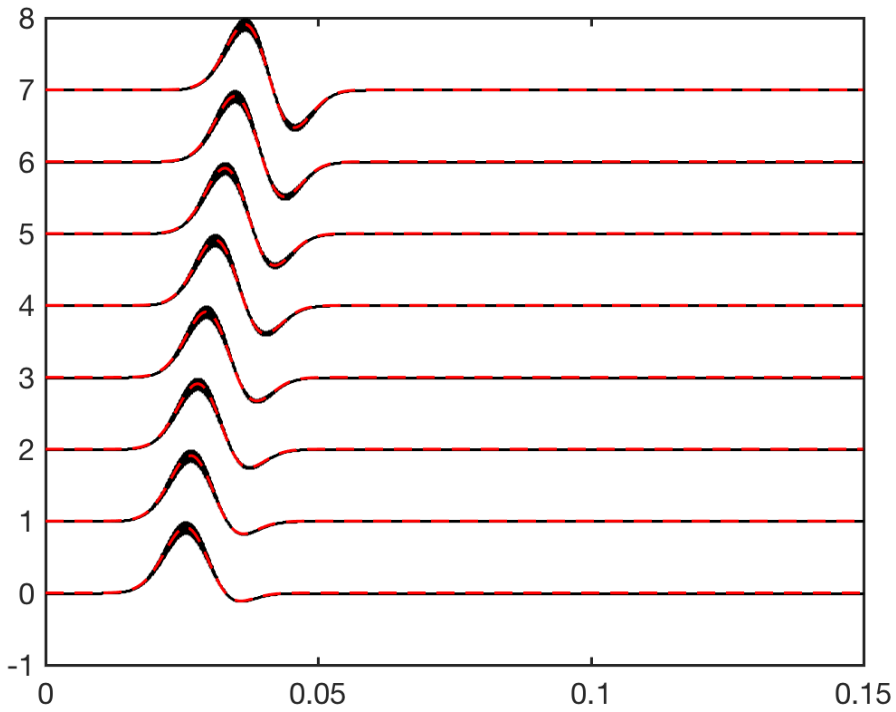

Figure 16: Mean (red dash) and one standard deviation region (black) from pressure traces from a force source distribution.

similarly accurate results. As mentioned in Section 4.2 we utilized a 2-D axisymmetric algorithm in order to obtain the MC results shown with at least 800,000 MC trials. As a comparison with the 3-D case using Paracousti, a single Paracousti run of this model requires ~50 s of compute time. If we did 10,000 MC runs (an absolute minimum number of runs to achieve a semblance of the standard deviation distribution of the output pressure traces) it would have required about 5.8 days of compute time. A 4th order maximum polynomial order run of Paracousti-UQ that we determined was adequately accurate required about 7.5 minutes of runtime. This amounts to Paracousti-UQ being ~1,100 times faster than MC.

An additional advantage of Paracousti-UQ output as opposed to MC simulations is that we immediately have access to estimates of higher order moments as well as the output probability distribution itself. For the MC simulations performed here, we did not retain output of every simulation from the MC suite as this would have required a large amount of disk space. Instead, we only kept that information necessary to compute the mean and standard deviation in an aggregate summary of the traces as the MC suite was running, greatly reducing memory requirements. If we decided we really wanted an estimate of the 3rd order moment of the output, we would have to do another suite of MC simulations. With the Paracousti-UQ output, as long as one has run the algorithm to sufficient maximum polynomial order to obtain such a higher order estimate, one can compute that higher order moment without having to re-run the algorithm.

5. CONCLUSION AND FUTURE WORK

We have derived the coupled set of stochastic partial differential equations for linear wave propagation in an acoustic medium. These equations are more complex than the deterministic equations in that there are more equations and more terms in each equation, but the same numerical principles used to solve the deterministic set of equations can be applied to the stochastic PDEs with minor modification. We developed Paracousti-UQ to numerically solve the SPDEs using 4^{th} order in space and 2^{nd} order in time finite differencing on a standard staggered grid. In the limit that all independent medium parameters and the source are exactly known, the SPDEs reduce to their deterministic counterparts, and we verified that Paracousti-UQ does indeed give identical answers to Paracousti in this special case.

To test the validity of the SPDEs as implemented in Paracousti-UQ, we compared Paracousti-UQ results to those of a large suite of Monte Carlo simulations for simple homogeneous models with a single random variable. The bulk modulus, buoyancy, and source time functions were individually assigned a probability distribution while the other parameters we considered known. In all cases Paracousti-UQ output agrees extremely well with MC simulation estimates. Additionally, Paracousti-UQ was demonstrated to be at least 1,000 times faster at achieving the same results as MC to high accuracy.

There are many research topics one could pursue to advance our understanding of the acoustic stochastic PDEs. One major area for further research is determination of the maximum polynomial order needed to accurately represent the probability distribution of the output wavefield. We touched on this topic in these tests, but these are only simple homogeneous test cases. What effect does realistic 3-D variations have on the maximum required polynomial order? A related topic area involves how to adequately represent the uncertainty in 3-D varying media. These test cases have only considered a single random variable. As the number of random variables increases the dimensionality of the problem grows likewise. The necessary polynomial order is based upon the product of chaos polynomials of each of the random variables, which rapidly increases the number of terms in this product. Monte Carlo methods also suffer from this issue in that greater and greater numbers of simulations are required to adequately represent the true distribution as the number of random variables increases. So the question is, how do we adequately represent the realistic 3-D uncertainty of the medium with as few of random variables as possible? One method that holds promise is by recasting the problem into one based on correlation length instead of point-by-point variations. This matches our intuition of how variability within a continuum would behave physically. These and many other questions await to be addressed with future research into this promising area of uncertainty quantification and propagation.

REFERENCES

1. Aldridge, D.F., *Radiation of elastic waves from point sources in a uniform wholespace*, SAND2000-1767, Sandia National Laboratories, Albuquerque, NM, July, 2000.
2. Aldridge, D.F. and M.M. Haney, *Numerical dispersion for the conventional-staggered-grid finite-difference elastic wave propagation algorithm*, SAND2008-4991, Sandia National Laboratories, Albuquerque, NM, July 2008.
3. Ghanem, R., Ingredients for a general purpose stochastic finite elements implementation, *Comput. Methods Apply. Mech. Engrg.*, 168 (1999), 19-34, 1999.
4. Komatitsch, D., and R. Martin, An Unsplit Convolutional Perfectly Matched Layer Improved at Grazing Incidence for the Seismic Wave Equation, *Geophys.*, 72 (5), SM155-SM167, doi 10.1190/1.2757586, 2007.
5. Kurtz, M., *Handbook of applied mathematics for engineers and scientists*, McGraw-Hill, Inc., New York, NY, 1991.
6. Preston, L.A., *Paracousti user guide*, SAND2016-9291, Sandia National Laboratories, Albuquerque, NM, September, 2016a.
7. Preston, L.A., *Optimized finite-difference coefficients for acoustic modeling*, SAND2016-10333, Sandia National Laboratories, Albuquerque, NM, October, 2016b.
8. Preston, L.A., *Nonlinear to linear elastic code coupling in 2-D axisymmetric media*, SAND2017-8848, Sandia National Laboratories, Albuquerque, NM, August, 2017.
9. Preston, L.A., D.F. Aldridge, N.P. Symons, Finite-Difference Modeling of 3D Seismic Wave Propagation in High-Contrast Media, *Soc. Expl. Geophys. 2008 Annual Meeting Extended Abstracts*, 2008.
10. Spanos, A., *Probability theory and statistical inference*, Cambridge University Press, Cambridge, UK, 1999.
11. Xiu, D. and G.E. Karniadakis, Modeling uncertainty in flow simulations via generalized polynomial chaos, *J. Comp. Phys.*, 187 (2003), 137-167, doi:10.1016/S0021-9991(03)00092-5, 2003.

DISTRIBUTION

1	MS0750	Leiph Preston	8861 (electronic copy)
1	MS1124	Jesse Roberts	8822 (electronic copy)
1	MS0899	Technical Library	9536 (electronic copy)


Sandia National Laboratories