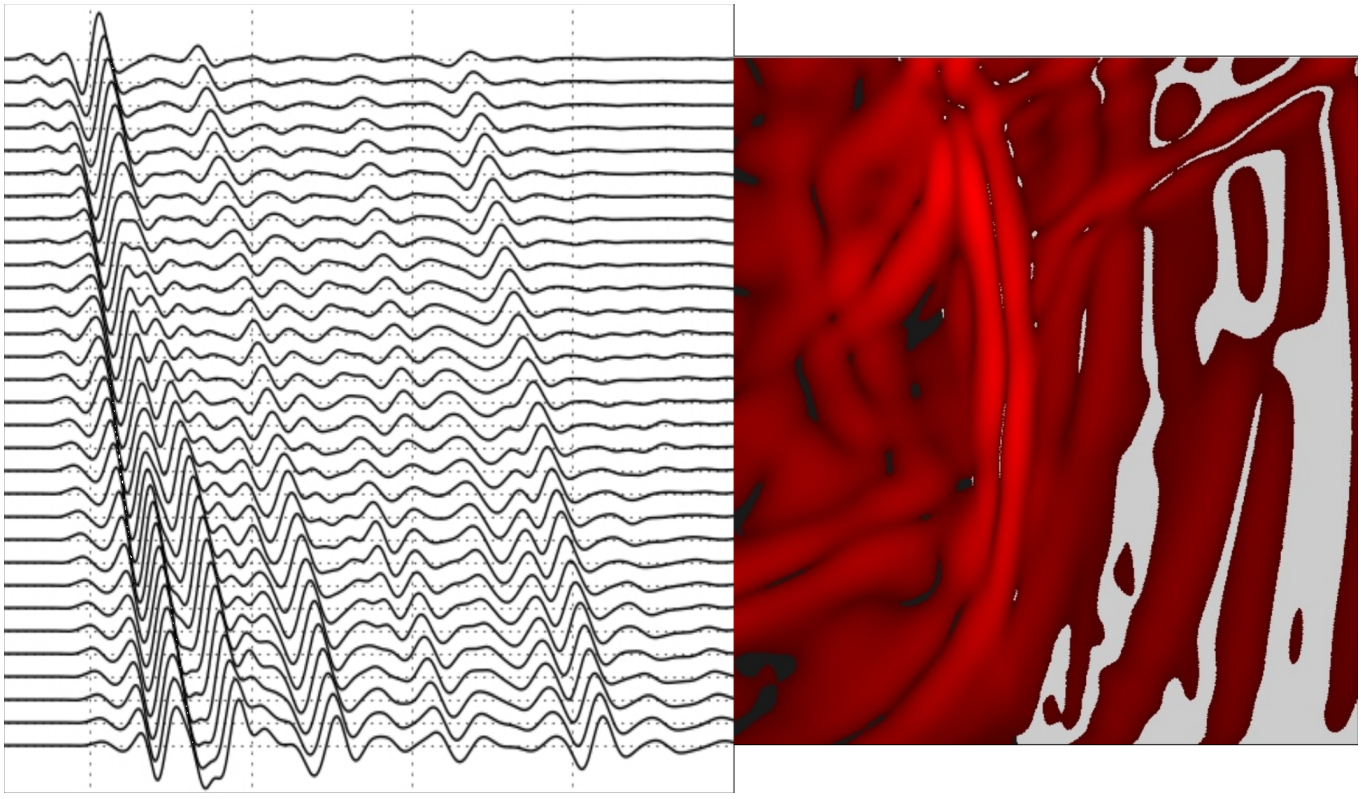
RUHR-UNIVERSITÄT BOCHUM

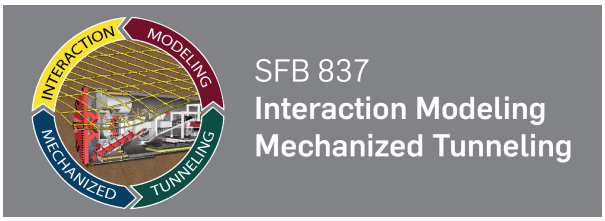
RUB

SFB 837
Interaktionsmodelle für den maschinellen Tunnelbau

Inference of ground condition in mechanized tunneling via inverse analysis using sequential Bayesian filtering

Luan T. Nguyen





RUHR
UNIVERSITÄT
BOCHUM

RUB

Inference of ground condition in mechanized tunneling via inverse analysis using sequential Bayesian filtering

Dissertation

as a requirement for the degree of
Doctor of Enginering (Dr.-Ing.)

at the Faculty of
Civil and Environmental Engineering
Ruhr-University Bochum

submitted by
Luan T. Nguyen

Bochum, July 2017

Date of submission: March 8, 2017
Date of defense: May 18, 2017

Referees:

Prof. Dr.-Ing. Tamara Nestorović
Research group Mechanics of adaptive Systems
Ruhr-University Bochum

Prof. Dr.-Ing. habil. Tom Schanz
Chair of Foundation Engineering, Soil and Rock Mechanics
Ruhr-University Bochum

Prof. Dr. Wolfgang Friederich
Institute of Geology, Mineralogy and Geophysics
Ruhr-University Bochum

Prof. Dr.-Ing. Martin Radenberg
Chair for Highway and Road Engineering
Ruhr-University Bochum

Abstract

The facts that the subsurface condition is very uncertain and that tunneling with tunnel boring machines is a ground-dependent complex controlled process make direct monitoring of the soil and structure responses induced by tunneling necessary. Model-based inference and mapping of the ground conditions driven by monitored data add values to the observed quantities and therefore can assist in choosing appropriate tunneling strategies and applying preventive measures for safe and cost-effective tunneling especially in urban areas. In this thesis, I proposed to use a set of efficient methods based on the sequential Bayesian filtering to solving inverse problems in mechanized tunneling. In particular, the sequential Bayesian inference based on the Kalman filtering is utilized either as a complete inversion algorithm or in combination with other numerical ingredients to achieve fast and reliable inversion results. It is shown that the nonlinear Kalman filters are suitable for estimating the mean values of the elastoplastic soil parameters from tunneling induced settlements. As far as the probability distributions of the soil parameters are of concern, data assimilation scheme based on the particle filter can be applied to quantify uncertainties sequentially along the excavation steps. By utilizing its efficiency in local navigation, coupling the Kalman filter with simulated annealing algorithm results in an accelerated hybrid global optimization method, which is of benefit for solving highly nonlinear inversion of the tunnel seismic waves. Also for tunnel seismic prediction where the geological disturbance is of arbitrary geometries, the level-set method is effectively used in combination with the Kalman filter for identifying both geometric and material properties of the anomalies. The proposed methods are verified through solving back analysis based on the synthetic measurements of the tunneling induced settlements and waveform inversion based on model-generated recordings of the tunnel seismic waves. Concerning the used data types, the results suggest that tunneling induced settlements are well suited for the interpretation of the geomechanical properties in homogeneous subsurface while the seismic waves based inversion is more suitable for identification of anomalies ahead of the tunnel face in heterogeneous subsurface conditions. In addition to the inversion studies, I set up a small-scale wave propagation laboratory experiment in the course of

this work as a basis for validation of the forward modeling and inversion results against imperfect experimental data.

Cuốn luận văn này dành tặng cho ba, mẹ, vợ, và con tôi

Contents

1 Introduction

1.1 Motivation

Tunneling began very early in human histories. Nowadays, tunneling is necessary for transportation and utility supplies especially in highly populated areas. Modern tunneling requires high standards for safety of the construction personnel and low degrees of surface settlement. The later is necessary to protect the critical infrastructure existing on the ground surface. However, due to the increased population in urban areas, tunnel diameters are expected to be larger to accommodate the traffic demand. At the same time, advance rates of the excavation are leveled up as much as possible to save costs. These expectations are in contradiction to the safety requirements mentioned above. Tunneling can be fast and safe in known ground conditions; however, the ground situation in the subsurface is often very uncertain.

Despite much of technological enhancements in tunnel boring machines in recent years to cope with difficult ground conditions, ground and face collapses and damages to surrounding civil structures are not rarely seen. When anomalies ahead of the tunnel faces are concerned, urban subsurface contains, besides natural formations, historic remnants and dense networks of utilities supporting structures, water supply, and electricity for the cities. Ground surface settlements and machine behavior in the case the TBM encounters unknown geological/ artificial objects are much harder to predict than those in the case of homogeneous ground because the incidents often happen without warnings. For example, in 2013 a 30-centimeter diameter steel pipe had blocked a large-size TBM in Seattle which resulted in extended downtime and massive cost overrun. Consequently, tunnel builders are willing to invest in monitoring the tunneling-induced mechanical changes on the surface and mapping ground conditions in the subsurface to assist them in choosing appropriate tunneling strategies and applying preventive measures. Monitoring is becoming easier with advanced sensing technologies and data management platforms. Although monitoring is a topic of itself to study what quantities to measure, which instruments are

needed, where and how frequent it is necessary to measure, etc., I assume in this work that the settings for monitoring are good enough.

Interpretations of the monitored/ measured data remain a demanding task especially if one wishes to look for unobserved quantities indirectly from the observables. For example, we want to know the elastoplastic soil properties from the measured settlements induced by tunnel excavation. Or we would like to reconstruct a map of the actual geological situation just ahead of the tunnel face by measurements of reflections and transmissions of the elastic waves. For computational models aided tunneling, well-determined soil parameters and the geological situation in return increase reliability of the widely used finite element and finite difference methods for further analyses and predictions of the excavation process.

Although the literature contains rich bodies of theories and computational methods for solving inverse problems, no particular method can beat the others for any applications. While gradient-based optimization methods are fast, they can easily be trapped at a local minimum if the starting point does not lie close to the global minimum. Statistical global optimization methods promise to be able to find the global minimum, but their computation burdens are often too much to be applicable because of an overly large number of function evaluations.

Generic inverse problems are naturally interpreted in the form of Bayesian statistical inference, i.e. the posterior probability density $\rho(\boldsymbol{m}|\boldsymbol{d}^{\text{obs}})$ of the concerned uncertain parameters $\boldsymbol{m}$ given the observation data $\boldsymbol{d}^{\text{obs}}$ is inferred from the prior probability density $\rho_{\text{prior}}(\boldsymbol{m})$ and the likelihood function $\rho(\boldsymbol{d}^{\text{obs}}|\boldsymbol{m})$ according to Bayes's theorem:

$$\rho(\boldsymbol{m}|\boldsymbol{d}^{\text{obs}}) \propto \rho(\boldsymbol{d}^{\text{obs}}|\boldsymbol{m})\rho_{\text{prior}}(\boldsymbol{m}). \tag{1.1}$$

In Eq. 1.1, the likelihood function $\rho(\boldsymbol{d}^{\text{obs}}|\boldsymbol{m})$ is a probability function weighting how likely the model parameters set $\boldsymbol{m}$ is to explain the observation data $\boldsymbol{d}^{\text{obs}}$. Given a numerical model $h(\boldsymbol{m})$, one aims to judge how well the model parameters $\boldsymbol{m}$ results in a set of model outputs that match the corresponding on-site measurements. Assume that uncertainty in observation data $\boldsymbol{d}^{\text{obs}}$ and uncertainty in prior estimated parameters $\hat{\boldsymbol{m}}_0$ are modeled as Gaussian quantities with covariance matrices $\boldsymbol{R}$ and $\boldsymbol{P}_0$ respectively, Eq. 1.1 is rewritten as

$$\rho(\boldsymbol{m}|\boldsymbol{d}^{\text{obs}}) \propto \exp\left[-\frac{1}{2}\left(\|h(\boldsymbol{m}) - \boldsymbol{d}^{\text{obs}}\|^2_{\boldsymbol{R}^{-1}} + \|\boldsymbol{m} - \boldsymbol{m}_0\|^2_{\boldsymbol{P}_0^{-1}}\right)\right]. \tag{1.2}$$

One can employ a probabilistic model parameter updating technique such as the Markov-chain Monte Carlo method to estimate the probability distribution of $\rho(\boldsymbol{m}|\boldsymbol{d}^{\text{obs}})$. Note

that $\rho(\boldsymbol{m}|\boldsymbol{d}^{\text{obs}})$ is generally not Gaussian because the model $h(\boldsymbol{m})$ that maps from the model parameter space to the observation space is in most cases nonlinear. Alternatively, one may follow the deterministic approach to seek for the parameters set $\boldsymbol{m}^*$ that minimizes the negative log-posterior in Eq. 1.2, i.e.

$$\boldsymbol{m}^* = \arg\min_{\boldsymbol{m}} \frac{1}{2} \left(\|h(\boldsymbol{m}) - \boldsymbol{d}^{\text{obs}}\|^2_{\boldsymbol{R}^{-1}} + \|\boldsymbol{m} - \boldsymbol{m}_0\|^2_{\boldsymbol{P}_0^{-1}} \right) \tag{1.3}$$

The function to minimize in Eq. 1.3 is so called the data misfit functional regularized by the prior. On the right hand side of Eq. 1.3, the first term is the data misfit and the second term is the regularization. Both of them are weighted by the corresponding inverse uncertainty covariances.

The Kalman filtering approach for off-line parameter estimation aims to estimate iteratively the posterior probability density function $\rho(\boldsymbol{m}|\boldsymbol{d}^{\text{obs}})$ but restricts itself to estimating only the first two statistical moments, i.e. the multivariate mean and covariance of the hidden model parameter set. Because the nonlinear Kalman filters can provide an unbiased posterior estimate (mean), its mean solution is equivalent to the solution of Eq. 1.3 or the maximum a posteriori estimate of Eq. 1.2. In addition, the Kalman filters can approximate the distribution of the posterior density function Eq. 1.2 in the form of Gaussian distributions. The Gaussian approximation is key to achieving feasible Bayesian inference for computationally demanding applications. As long as the full probability distributions are concerned, the particle filter can be used to estimate the full posterior Eq. 1.2.

This dissertation describes how the sequential Bayesian filtering is suitable for solving inverse problems with particular applications for subsurface characterization in mechanized tunneling. In particular, it is shown how the Kalman filters and particle filter are formulated and become effective in solving back analysis problem driven by tunneling-induced settlements. For mapping the geological structure ahead of the tunnel face, a derivative-free variant of the Kalman filtering is used in combination with other numerical ingredients to facilitate waveform inversion of the tunnel seismic waves.

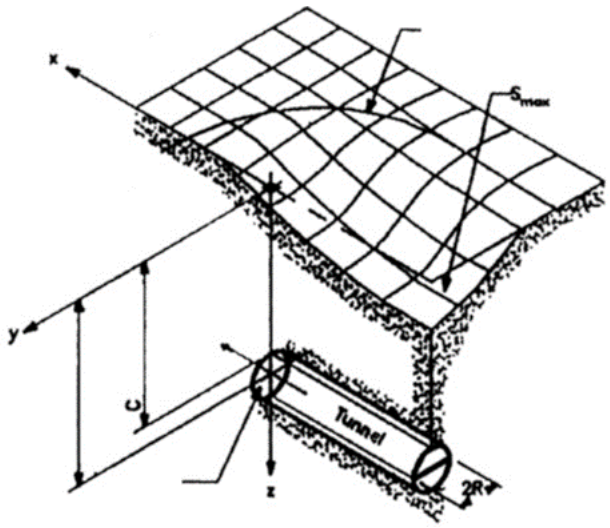


Figure 1.1: Surface settlement trough (Leca & New (2007))

1.2 Overview of inverse analyses for ground condition characterization

1.2.1 Settlements based identification of soil parameters

1.2.1.1 Tunneling-induced settlements

Ground surface settlements induced by tunneling depend on a number of factors, including (1) geological, hydro-geological and geotechnical conditions, (2) tunnel geometry and depth, (3) excavation methods, and (4) the quality of workmanship and management (Leca & New (2007)).

Tunnel excavations modify stress state of the ground, thus inducing surface settlements. Besides the in-situ condition of the subsurface, the degree of settlements is influenced by the tunneling process itself from parameters such as rate of advancement, face-support pressure, grouting pressure, etc. Assuming that the tunneling process is very well controlled with the best engineering practices, then surface settlements are due to geomechanical properties of the soil being excavated. Fig. 1.1 shows the typical settlement trough resulting from tunnel excavation.

Taken into account certain geometrical and geomechanical simplifications, empirical and analytical methods can be very useful for calculating deformations due to tunneling (Loganathan & Poulos (1998); Bobet (2001)). Numerical methods such as finite element and

finite difference methods, however, can take into account complex geomechanical behavior as well as interactions among soil, machine, and surrounding structures (Kasper & Meschke (2004); Lambrughi et al. (2012); Zarev et al. (2011); Do et al. (2014)).

Together with advancement in modeling capability, more fully deployment of state-of-the-art monitoring systems and the model-based inference methods for extracting added information from the measurements have witnessed strong growths as well. An interaction platform to manage and process raw monitoring data, then perform updating of the ground conditions driven by data, and finally assist in the decision-making of tunnel driving process will help make tunneling faster and safer.

1.2.1.2 Settlements based inversion

In numerical modeling of geotechnical problems, the modeling equation is of the form $\mathbf{d} = \mathbf{h}(\mathbf{x}, \mathbf{m}, \mathbf{f})$, where $\mathbf{d}$ stores modeling outputs obtained from the model function $\mathbf{h}(\cdot)$ (usually a finite element (FE) model), $\mathbf{x}$ is the current physical state of the model, $\mathbf{m}$ contains the model parameters (for example material parameters), and $\mathbf{f}$ represents the external loads and boundary conditions. In the numerical model, it is assumed that the applied loads, the initial and boundary conditions are already well determined. It remains the cumbersome task to find a set of actual model parameters $\mathbf{m}$. Beside measuring the model parameters with the help of sophisticated laboratory tests, model calibration by inverse analysis is preferred in many cases to match the model outputs $\mathbf{d}$ to the in-situ measurement data $\mathbf{d}^{obs}$. This inverse problem is mentioned under several technical terms such as parameter identification, model calibration, inverse analysis (or back analysis in geotechnical engineering).

Being viewed from the optimization perspective, inverse analysis can be considered as a general optimization problem, i.e. to find a set of model parameters $\mathbf{m}$ that minimizes a mathematical norm that represents the difference between the model outputs and the measurement data $\mathbf{h}(\mathbf{m}) - \mathbf{d}^{obs}$. However, this optimization problem is very challenging due to following reasons: (i) the fact that geomechanical models are highly nonlinear leads to inverse problems that can have many local minima satisfying the minimization criterion, (ii) model calculation of realistic large-scale elasto-plastic material laws is time-consuming and may not be exact due to model simplification and errors coming from numerical approximation, and (iii) in-situ measurement data may be incomplete and susceptible to noise which exacerbates convergence of inverse analysis to a unique solution. The first difficulty is intrinsic to the inverse analysis of geotechnical problems. Global optimization

methods such as the Monte Carlo simulations and meta-heuristic algorithms can overcome this challenge, but they result in extremely high computation cost. In spite of the advance in mathematical modeling, numerical computation, and parallelization techniques, the second disadvantage is still severe for solving inverse analysis due to the requirement of a large number of forward model calculations. Therefore, it is crucial to employ inverse analysis methods that require only a moderate number of forward model calculations. The third challenge concerns measurement uncertainty which cannot be avoided either. For that reason, an inverse method that is able to incorporate and quantify uncertainty to obtain robust and reliable estimates of the model parameters offers an essential advantage.

In geomechanics, there exist a variety of methods which are suitable to solve inverse problems. Categorization of the commonly employed methods was made early by the work of Gioda & Sakurai (1987) in which inverse analysis methods are surveyed under deterministic and probabilistic viewpoints. The fact that deterministic methods such as conjugate gradient and quasi-Newton methods require very good initial model parameters so as to avoid converging to local minima makes them unappealing to solving inverse problems. Probabilistic approaches, on the contrary, allow for more freedom in the choice of initial model parameters and are more consistent in attaining global minimum. Among the prominent works that employed probabilistic inverse analysis methods, Ledesma et al. (1996) introduced a solid and unified probabilistic framework that utilizes prior information obtained from site investigation stage.

Inverse problems can be viewed as optimization problems whose objective functions are highly nonlinear, non-smooth, and can have multiple local minima. Therefore, derivative-free optimization algorithms based on deterministic and random generated candidate models in the parameter have been used widely. Meier et al. (2008) and Knabe et al. (2013) had performed extensive study of the particle swarm optimization method and used it for calibrating constitutive models and identifying soil parameters for complex geotechnical models. Levasseur et al. (2008) introduced the genetic algorithm for estimating constitutive parameters of the Mohr-Coulomb model from a sheet pile wall retaining an excavation in-situ test. On application in the underground structure, Miranda et al. (2011) and Moreira et al. (2013) have applied the evolutionary algorithm to identify parameters of rock mass surrounding underground space structures.

In general, when an optimization method is applied to minimize the discrepancy between model generated outputs and corresponding in situ data, it is required that the measurements reflect the true in situ changes of the observables and the numerical model used is able to simulate the exact geomechanical phenomena. However, measuring campaigns

in large geotechnical sites are prone to acquiring data with some degree of uncertainties. Besides, it is a fact that constitutive models and numerical methods are limited by theoretical assumptions and numerical approximation that make numerical results not the exact reality. These uncertainties and inaccuracies exacerbate performance of the optimization methods. For such reasons, Bayesian filtering approaches are preferred to optimization methods because they can tolerate uncertainties and even can utilize prior engineering knowledge which is often obtainable by initial site investigations. Early reports on applying Bayesian methods to inverse problems in geomechanics include Cividini et al. (1983) and Gioda & Sakurai (1987). Another Bayesian scheme which is based on sequential filtering of Gaussian density, the Kalman filter, has also been applied successfully for identification of soil parameters. The early extensive works that introduced and adapted the Kalman filter to solving inverse problems performed by Murakami (1991) and Hoshiya & Sutoh (1993) have shown that the Kalman filter is very effective for this kind of problem in geomechnics. Later, Kalman filter and its variants have been reported to be applied in back analyses for Mohr-Coulomb's geomaterial parameters and in-situ stress state by Hommels et al. (2009) and Yang et al. (2011) respectively. A probabilistic Bayesian filtering based on the particle filter has been introduced to identifying elasto-plastic soil properties for soil-water coupled problem in geomechnics (Murakami et al. (2013)).

1.2.2 Elastic waves based mapping ahead of the tunnel face

1.2.2.1 Elastic waves

The motion of wave propagation obeys Newton's second law. The stress-strain relation that describes the small-strain deformation of the continuous media follows Hooke's law of elasticity. Therefore, two elastic constants (e.g. P-wave speed and S-wave speed) and density are required to define material properties of the elastic elements in the wave propagation medium. In addition, internal material damping can be taken into account by the dimensionless quality factor Q. The quality factor is defined as inversely proportional to the energy loss in each cycle caused by anelasticity (Aki & Richards (2002)). It is noted that, compared to the elastic constants and density, anelastic attenuation factor Q is harder to determine and decreases very fast with increasing frequency (physical dispersion).

Elastic waves propagate in the body of a medium in the form of longitudinal waves (P waves) and transversal waves (S waves) with P waves faster than S waves. P waves and S

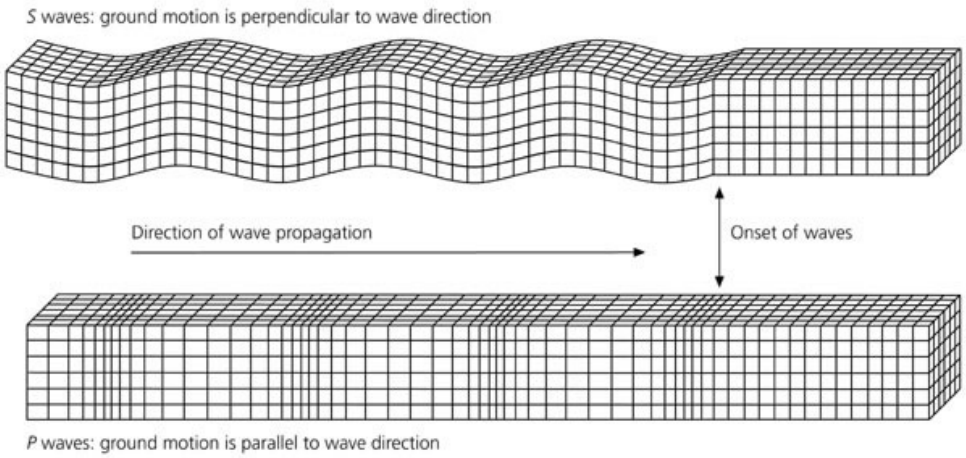


Figure 1.2: Motions of P and S waves (Stein & Wysession (2009))

waves differ in the directions of particle displacement with respect to the direction of wave propagation. As visually illustrated by Stein & Wysession (2009) in Fig. 1.2, in P-wave propagation, material particles oscillate in the same direction as the direction of wave propagation; while in S-wave propagation, material particles oscillate in the direction perpendicular to the direction of wave propagation. S waves themselves are separated into two modes depending on the polarization of the particle displacements in the plane in which the S waves propagate. If the S-wave displacement is in the plane of wave propagation, it is called the shear vertical (SV) waves; if the S-wave displacement is perpendicular to the plane of wave propagation, it is called the shear horizontal (SH) waves. When in interactions with heterogeneities, P and SV waves are coupled together whereas SH waves are separate.

Waves propagating along the surface are Rayleigh waves which travel more slowly than body waves. In addition to Rayleigh waves, Love waves can exist as a result of the interactions of SH waves on the surface in case of a layered half-space. As Rayleigh waves only travel along the surface at the depth no more than several wavelengths, they are less attenuative than P waves and S waves. Concerning the geometrical damping alone, Rayleigh wave magnitude traveled a distance r is decayed to only $1/\sqrt{r}$ compared to $1/r$ decay rule of body waves. The particle motion of Rayleigh waves is a combination of the motions of P waves and SV waves resulting in the motion of a point on the surface a retrograde ellipse. In a layered half-space, different wavelengths of the Rayleigh waves penetrate at different depths and therefore travel with different velocities causing very strong geometrical dispersion. Due to low attenuation and high dispersion characteristics,

Rayleigh waves are useful for geotechnical site investigations (Xia et al. (1999); Park et al. (1999); Socco et al. (2010)); however, for the same reasons, circulating and scattered Rayleigh waves can obscure the low-amplitude reflected body waves and may need to be suppressed artificially (Halliday et al. (2007)). One further advantage of using surface waves is that most energy excited by an impact on the surface is transferred to the medium in the form of surface waves. According to a calculation for soil having Poisson's ratio of 0.25 (Miller & Pursey (1955)), the ratio of wave energy by a vertically osculating source on a homogeneous, isotropic, elastic half-space is transferred to the subsurface with the ratio nearly 67:26:7 with respect to the energy of Rayleigh waves, S waves, and P waves respectively. Wave mode conversion from one wave type to another can happen when the wavefront is incident on an interface at a boundary surface or a heterogeneity with an angle other than normal to the direction of wave propagation.

Depending on scales, geometries, inner structures, and the excited frequencies, elastic wave propagation can result in complex waveforms recorded at certain positions (often on the surface). Interpretation of the information encrypted in the recorded waveforms allows for characterization of the seismic source and the structure of the medium through which the elastic waves traverse. In engineering applications such as in geotechnical site investigation and non-destructive evaluation of civil materials based on seismic and ultrasonic waves, wave source is often induced artificially with a known dynamic function, the task, therefore, pertains to characterizing the internal material structure or the subsurface. Depending on the characterization scales and purposes, various elastic waves based techniques can be employed: multichannel analysis of surface waves (MASW) (Park et al. (1999)) is very successful for vertical shear wave profile mapping based on dispersion characteristics of the surface waves; coda wave analysis (Snieder (2006)) can be applied in characterization of small perturbations in granular materials such as natural geomaterials (Dai et al. (2012)), road pavements (Papadopoulos et al. (2016)), and concrete (Stähler et al. (2011)); acoustic emission (AE) (Grosse & Ohtsu (2008); Saenger et al. (2011)) is used to quantify damage in a brittle structure being loaded quasi-nondestructively or destructively; and for applications in damage detection in plate-like structures the use of Lamb waves (Su & Ye (2009)) is very widespread. Most elastic waves based material characterization methods are attractive in practice because their data acquisition processes require either very minimal invasion into the internal structure of the investigated medium or are completely nondestructive.

The so-called tunnel seismic waves are elastic waves propagating in the subsurface environment influenced by the tunnel under excavation. Bohlen et al. (2007) detailed the

behavior of tunnel seismic wave propagation especially the Rayleigh-to-shear wave mode conversion at the tunnel face. An example of a seismogram resulting from a tunnel seismic survey is seen in Fig. 1.3.

1.2.2.2 Ahead-of-tunnel prediction using tunnel seismic waves

Conventional site investigation by drilling boreholes helps to draw a general image of the geological structures along the borehole axis, but that image is likely to misrepresent the exact geological condition due to limited sampled positions. Due to high costs and restricted boring space which deep tunneling and tunneling under urban areas encounter, many industrial and academic research groups have had a great interest in developing nondestructive ahead of the tunnel face prediction systems. Nondestructive investigation means only excitation sources and receiving sensors, which are usually placed on the ground surface, on the side walls, or on the cutter head of the tunnel boring machine (TBM), are used for the acquisition of measurement data. The recorded signals in the form of seismic or electromagnetic/ electrical waves, which travel through geological structure or back reflected from it, holds the signature of the geological structures of concern such as geological layer changes, fault zones, erratic rocks, boulders, construction remnants, etc. Therefore, these seismic signals can be used to reconstruct the spatial and material properties of those hidden geological structures. Most standard nondestructive ahead-of-tunnel investigation systems to date are based on interpretation of reflected elastic/ acoustic waves (seismic waves) and analysis of electromagnetic/ electrical resistivity signals (Mooney et al. (2012)). However, due to insufficient signal penetration depth and spread of radar and geoelectric sources, the use of seismic waves for ahead of tunnel face prediction (tunnel reconnaissance) is more preferable (Kneib et al. (2000); Bohlen et al. (2007)).

The principle of a generic tunnel seismic prediction system is depicted in Fig. 1.4. In principle, the current tunnel seismic prediction systems require either active source(s) (excited by a sledgehammer or explosive) or passive seismic source (induced by vibrations of the TBM itself) or both kinds of excitation as in Ashida (2001) and an array of receiving geophones placed in the tunnel wall and/ or on the cutter head of the tunnel boring machine. Energy generated from a seismic source propagates the tunnel geomaterials (soils and rocks) in the form of body waves which comprise pressure wave (P-wave) and shear wave (S-wave). When these body waves encounter the geologic heterogeneities (fault zones, erratic rocks, boulders, etc.), the partial energy of the incident waves is reflected/ refracted to the tunnel excavation area and is recorded by the geophones installed there.

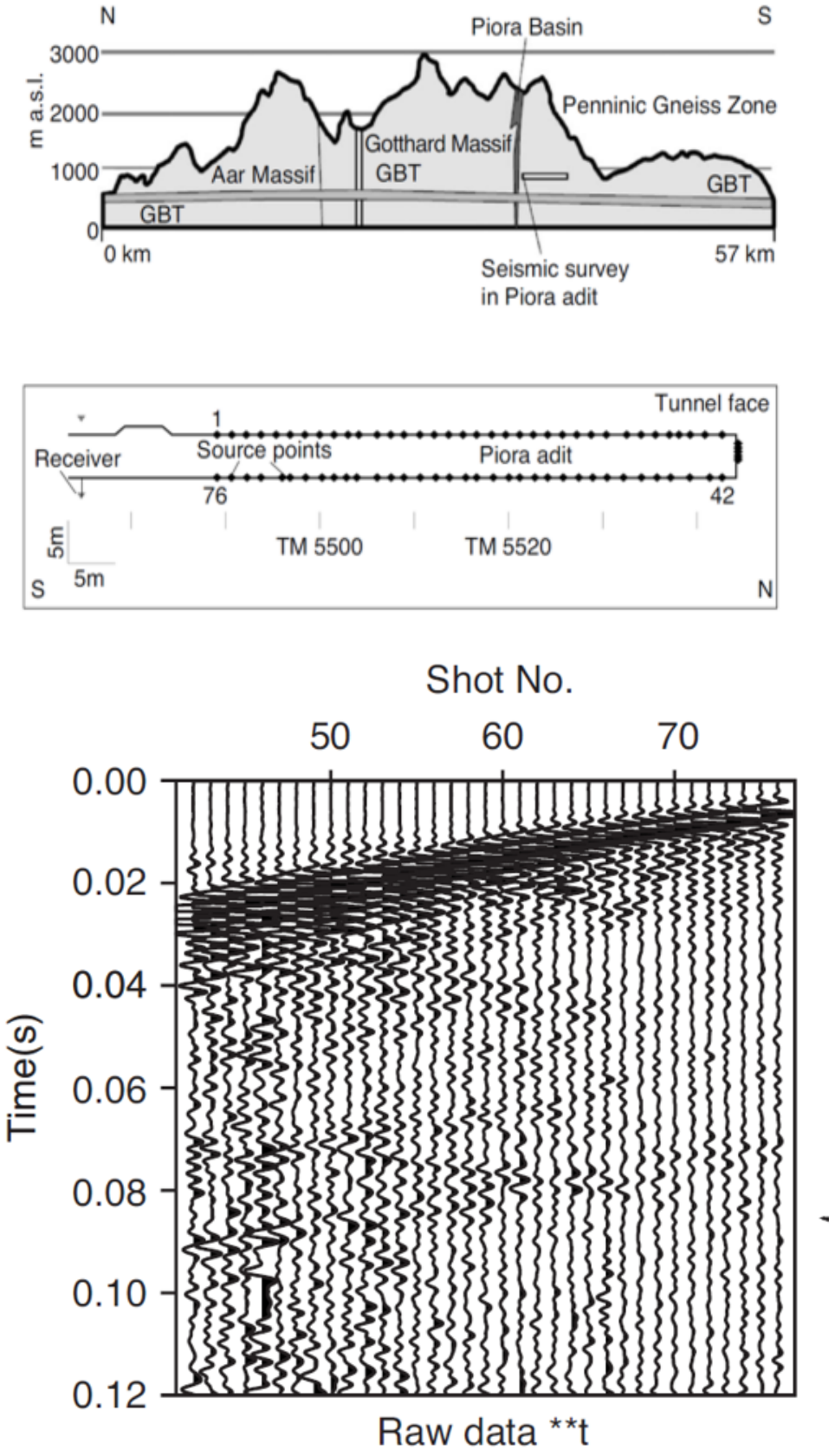


Figure 1.3: Tunnel seismic survey at in the Piora adit of Gotthard Base Tunnel (More details in Bohlen et al. (2007))

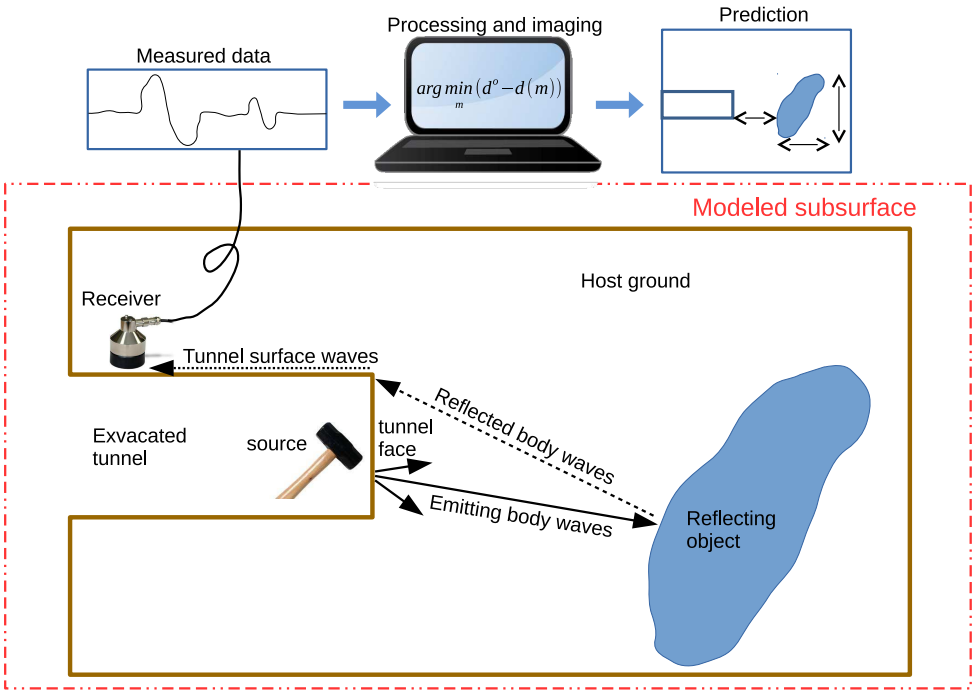


Figure 1.4: Principle of a generic tunnel seismic prediction system

The image of the geological features ahead of the tunnel face is then reconstructed by well-established geophysical interpretation methods such as seismic migration or seismic travel time tomography. An overview of the development of acoustic and seismic waves based ahead of tunnel prediction systems is summarized below in light of the technological development and the development of the imaging strategies.

Early application of seismic methods in tunnel advance exploration made use of the vertical seismic profiling (VSP) technique (Sattel et al. (1992); Brückl et al. (2001)) and adapted it to work in the horizontal direction along the tunnel axis. According to the technical descriptions in the mentioned reports, direct and reflected waves are separated, and the later are used to map hazardous zones ahead of the tunnel face by means of seismic migration. This method was further used for TBM driving and renamed in-tunnel horizontal seismic profiling (HSP) (Inazaki et al. (1999)). One of the most widely used advance exploration systems to date for hard rock environments is the Tunnel Seismic Prediction (TSP) system (Dickmann & Sander (1996)) in which measurement instrumentation and interpretation software are integrated into a compact commercial package by Amberg Technologies. The sources and receivers of the TSP system are placed in small boreholes along the tunnel wall as illustrated in Fig. 1.5. Also for application in hard rock, GeoForschungsZentrum Potsdam cooperated with other research institutes and Herrenknecht AG to further develop an optimized source and receiver system as well as

the seismic interpretation software for the so-called Integrated Seismic Imaging System (ISIS) (Borm et al. (2003)). Other available seismic tomography systems with a similar working principle include the Tunnel Reflector Tracing (TRT) (Neil et al. (1999); Yamamoto et al. (2011)), the True Reflection Tomography (TRT) (Otto et al. (2002)), the True Reflection Underground Seismic Technique (TRUST) (Benecke et al. (2008)), the Tunnel Seismic Tomography (TST) (Zhao et al. (2006)), and the Tunnel Geological Prediction (TGP) (Jiao et al. (2015)).

Targeting at tunneling in soft soils by the earth pressure balance (EPB) boring machines, the Sonic Softground Probing (SSP) was developed (Kneib et al. (2000); Gehrig et al. (2010)). The main difference between the SSP and other tunnel seismic systems is that the sources and the receivers are completely placed on the cutting wheel of the boring machine (Fig. 1.5).

All of the seismic prediction systems mentioned above require the active seismic source(s) to initiate acoustic/ seismic wave propagation. One particular family of the tunnel seismic systems that does not need a separate source is the Tunnel Seismic While Drilling (TSWD) system which uses the seismic waves generated by the TBM cutting wheel as the passive seismic source (Petronio & Poletto (2002); Gehrig et al. (2010); Petronio et al. (2003, 2007); Brückl et al. (2008)).

Besides the conventional seismic migration adopted from geophysics, several other effective imaging strategies based on increasingly advanced computational power have been proposed. With regard to the identification of small obstacles (such as boulders, foundations, and artificial relics) in the vicinity ahead of the tunnel face, Swinnen et al. (2007) applied the refined focusing operators to standard seismic migration used in hydrocarbon exploration to improve the resolution of the imaged structure in front of the TBM.

By performing 3D finite-difference modeling, Bohlen et al. (2007) and Jetschny et al. (2010) discovered that the tunnel surface waves (Rayleigh waves) arriving at the tunnel face are converted to P-wave and S-wave. The converted body waves, of which S-wave is dominant, continue to travel to geological structures ahead of the tunnel face and are reflected/ refracted by the geological heterogeneities. The reflected body waves are scattered back into surface waves at the tunnel face and are guided along the tunnel sidewalls. They proposed to use Rayleigh waves instead of body waves for the imaging purpose to have the advantage of placing explosive source and geophones well behind the cutter head and shield of the TBM. Validation of the developed prediction sequence with tunnel seismic data sets is reported in Jetschny et al. (2011) . Finite difference method

had been used also by Kneib & Leykam (2004) and Essen et al. (2007) to model seismic wave propagation in underground tunnel environment.

Tzavaras et al. (2012) implemented 3D versions of various advanced seismic imaging methods, namely the Kirchhoff Prestack Depth Migration (KPSDM), the Fresnel Volume Migration (FVM), and the Reflection-Image-Spectroscopy (RIS). The authors concluded that the integration of the two approaches (FVM and RIS) exploits their advantages and delivers more pronounced and clearer image of the tunnel environment. Bellino et al. (2013) introduced a simple and fast technique for automatic projection of peaks of the seismic events onto the tunnel space domain. The analysis procedure taking into account multiple sources and sensors is able to detect locations at geological changes and the velocity model is approximated by simple regression analysis of the travel times of direct waves and refracted waves. Cheng et al. (2014) applied the 2D reverse time migration (RTM) by numerically solving the decoupled elastic wave equation to image the primary geological reflectors ahead of the tunnel face.

Lambrecht & Friederich (2013) implemented the discontinuous Galerkin method to model the elastic wave propagation in tunnel environment. The advantage of the discontinuous Galerkin method is that the model can be meshed by triangular (2D)/ tetrahedral (3D) elements. This ability gives flexibility for meshing the highly inhomogeneous subsurface with an embedded tunnel. However, due to long computation time, simulation of elastic wave propagation using this method is preferably executed in parallel mode on a computer cluster. With a similar aim, Musayev et al. (2014) implemented forward modeling and inversion of acoustic waves in the frequency domain for 2D and 3D tunnel models.

Very recently, attempts to perform elastic full waveform tomography have been made to take benefits of utilizing information from the whole seismogram which undoubtedly promises to produce a more detailed image of the geologic heterogeneities. Recent reports regarding applications of full waveform inversion of seismic data recorded in underground galleries Bretaudeau et al. (2013); Lüth et al. (2014) have shown positive results.

To characterize the geological anomalies ahead of the underground tunnel, the authors are interested in inverting the full waveforms because of its potential to produce a high-resolution image of the subsurface. The cost function to minimize in this study is targeted to be the full waveform misfit, which measures differences between the recorded and the modeled information encrypted in the respective waveforms, because of the two main reasons. Firstly, the full waveforms (body waves and surface waves) contain the most information about the hidden geological structure that can be obtained from the reflected seismic waves. Therefore, inversion of the full waveforms promises high-resolution imag-

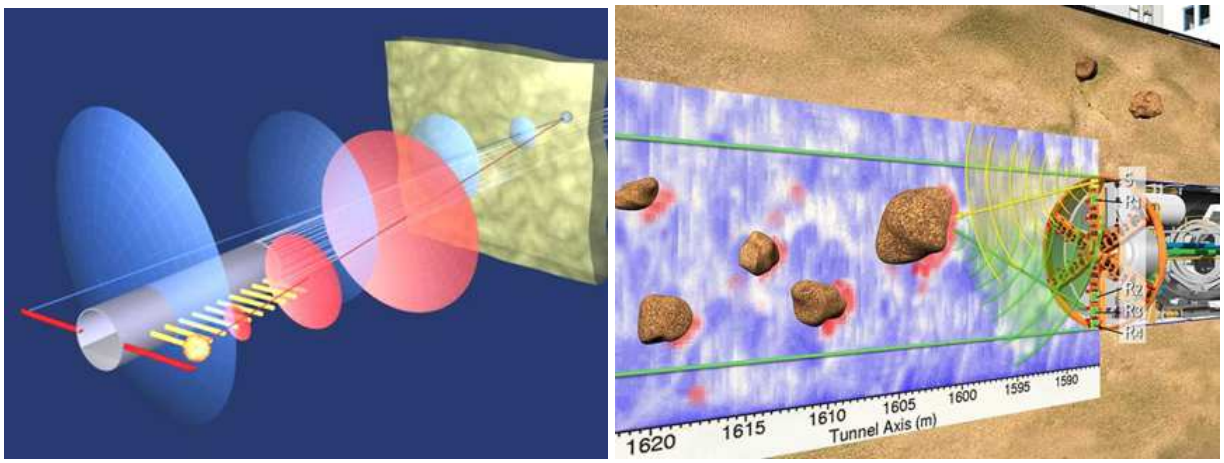


Figure 1.5: Illustrations of the Tunnel Seismic Prediction (TSP) system and the Sonic Softground Probing (SSP) system developed by Amberg Technologies and Herrenknecht AG, respectively

ing of the subsurface. In contrast, travel time tomography and seismic migration rely only on arrival times and/or amplitudes of the reflected waves for imaging the geological structure. The advantage of having rich information is beneficial to the underground tunnel settings because the underground working space is strictly constrained for placing the receivers, usually on the tunnel side wall or the tunnel face, which limits the amount of valuable recorded data. Secondly, waveform inversion is not yet fully developed, for example, not only one answer is accepted to the question what form of misfit functional better represents the mismatch and also facilitates global convergence for the inversion or the question what inversion method guarantees global convergence while keeps the computation time acceptable in engineering applications. Besides, several general questions are yet to answer such as what method of model parameterization can help accelerate convergence, how prior knowledge about the geologic model should be taken into account, and what numerical inversion methods yield high accuracy while remaining computationally feasible despite the fact that measurements are contaminated by noise and that the initial model is far from the true model.

1.3 Contributions of this thesis

This thesis reports the findings achieved within this doctoral work. Large pieces of this dissertation are reported in the below articles.

- L. T. Nguyen, T. Nestorović, K. Fujisawa, A. Murakami, *Particle filter-based data assimilation for identification of soil parameters with application in tunneling*, Proceedings of the 14th International Conference of the International Association for Computer Methods and Advances in Geomechanics, Kyoto, Japan, 2014
- L. T. Nguyen, T. Nestorović, *Nonlinear Kalman filters for model calibration of soil parameters for geomechanical modeling in mechanized tunneling*, ASCE Journal of Computing in Civil Engineering, 2015
- L. T. Nguyen, T. Nestorović, *Unscented hybrid simulated annealing for fast inversion of tunnel seismic waves*, Computer Methods in Applied Mechanics and Engineering, 2016

This dissertation is organized as follows.

Chapter 1 has introduced the motivation for the inversion methods studied in this work in which the Bayesian filtering concept is key to the inversion approaches used and developed in the course of this work. Besides, the basis of the forward problems of excavation-induced settlements and elastic wave propagation is summarized, and an overview of inverse problems for solving back analysis of elastoplastic soil parameters and subsurface imaging in the tunneling context has also been detailed.

Chapter 2 discusses the use of *nonlinear Kalman filters for soil parameters identification*. It will be shown that the nonlinear Kalman filters can be applied very effectively to calibration of the geomaterial parameters for geomechanical modeling in mechanized tunneling using tunneling-induced settlements and horizontal displacements. The data curves measured along tunnel excavation steps which exhibit nonlinear relationships with respect to soil parameters and are prone to measurement inaccuracies are utilized in combination with finite element modeling to estimate the underlying soil parameters using sequential inference framework: the nonlinear Kalman filtering. Comparative performance of the two types of nonlinear Kalman filter shows that they are effective for soil parameters identification in terms of convergence speed and accuracy: the extended Kalman filter (EKF) and the sigma-point Kalman filter (SPKF). The effectiveness of the two Kalman filters for inverse analysis is demonstrated through computer simulations for identifying a number of important constitutive parameters of the Hardening Soil model in the context of mechanized tunneling. In details, the state-space representation of the forward model is presented as a basis for deployment of the Kalman filters. After that, sequential Bayesian filtering that leads to algorithms of the two Kalman filters is introduced. Before going to the main application in this work, the capability of the EKF and SPKF is demonstrated

for identification of hidden parameters of a noisy time series and localization of the global optimum of the multimodal Ackley's function. In the application section that follows, a short description of the finite element model for the tunnel excavation and generation of noisy synthetic measurement data are given. Also, the authors discuss inverse analysis outcomes that result from different computer-experiment settings. Finally, several final remarks and suggestions for the optimal use of the two nonlinear Kalman filtering techniques for model calibration of soil parameters in modeling of mechanized tunneling and other geotechnical structures are given.

Chapter 3 presents a data assimilation (DA) method that employs the particle filter for identifying the soil parameters in tunneling application. The idea of DA is to incorporate observation data into numerical modeling to improve prediction capability of the numerical model which is prone to inexactness due to limitations of modeling procedure such as lack of knowledge of physical phenomena and inaccuracy due to mathematical approximation. When applied for parameter identification, the uncertain model parameters are described as stochastic variables whose distributions will come closer to the "true" probability densities through assimilation process as more observation data become available. DA is not an iterative process but rather a sequential process. In other words, identification is done in parallel with acquiring data from fields during loading, construction, or excavation steps. The major advantage of particle filter over the Gaussian Kalman filters is that the resulted estimation quantities are represented in the form of probability distributions rather than their centered means and variances only. In the demonstrative application, data being used for assimilation are settlements and horizontal displacements induced by tunnel excavations. As the tunnel boring machine advances, more information is taken into the assimilation process to reduce estimation bias. By using noisy synthetic data, it is observed that those mean model parameters identified in the end of excavation steps are very close to the true parameter values.

Chapter 4 introduces a global optimization strategy for inversion of tunnel seismic waves. A new hybridized global optimization method that combines simulated annealing global search with unscented Kalman filter minimization is proposed to solve waveform inversion for predicting ahead of the underground tunnel. The authors demonstrate in this work fast and reliable convergence of this new algorithm through validation of an optimization of multi-minima test function and an inversion of synthetic tunnel seismic waveforms to predict geological structure ahead of a tunnel face. Concerning the engineering application, the successful identification of the true model by minimizing a multimodal misfit functional for wide feasible bounds of the model parameters confirms that waveform inver-

sion by the improved global optimization method is promising for practical applications with real measurement data.

Chapter 5 proposes a waveform inversion scheme supported by the level-set method for locating and characterizing the disturbance zones ahead of the underground tunnel face. The inversion process is completely free from gradient calculations as the unscented Kalman filter is used for the efficient inversion. The conceptual methodology is validated through successful reconstructions of single- and multiple-disturbance objects in a simple 2D frequency domain model. In the tunneling application, the special characteristics of the tunnel seismic waves in the time domain are described and the results of SPECFEM2D simulation and a qualitative evaluation of the simulated tunnel seismic waveforms are shown. The computer model and its simulated tunnel seismic waves are eventually used to reconstruct the geological scenarios in which disturbance of a single object and multiple discontinuous objects in a parsimonious and flexible manner. Although further validations with laboratory or in-situ measurements and the use of fully 3D model are needed to prove the practicality of this approach, the current results are encouraging and promising to apply FWI in tunneling practice as an advanced tool for looking ahead of the tunnel face.

Chapter 6 concludes the thesis and points out the potentials of the inference schemes for being integrated into a tunnel information system. Furthermore, a small-scale laboratory experiment set-up is described for validation of the numerical methods against laboratory measurements.

2 Nonlinear Kalman filters for model calibration of soil parameters

2.1 Background

There is one special family of inverse methods that has been successfully adopted from the field of state and parameter estimation of generic dynamic systems to solve parameter identification problem in geotechnical engineering and civil engineering in general – the Kalman filter and its variants. From perspective of state and parameter estimation of nonlinear dynamic and quasi-static models, the state variables and parameters of the nonlinear model are recursively updated as new measurements become available. With this principle, the Kalman filter estimates the Gaussian densities (means and covariances) of the quantities being estimated by repeatedly performing time update and measurement update to reduce the variance of the estimation error and therefore converges the estimated quantities to unbiased true state and model parameters (Kalman (1960); Jazwinski (1970); Gelb (1974)). Since the first closed form solution to the state estimation (filtering) problem for linear models, the Kalman filter (Kalman (1960)), there have been extensive further developments of this filtering scheme. Later, the effort to extend the Kalman filter to be working with even nonlinear system models was successfully achieved by applying Taylor series expansion of the models around the working point (Jazwinski (1970)). The successful extended Kalman filter (EKF) has become until today the standard state/ parameter estimation method for slightly nonlinear dynamic systems and structures. More recently, Julier & Uhlmann (1997) contributed to the development of the Kalman filter with the so-called unscented Kalman filter (UKF) based on the intuition that the approximation of a Gaussian distribution is easier than the approximation of an arbitrary nonlinear function. Because of its deterministic samples (points), it is preferably referred to as sigma-point Kalman filter (SPKF). The SPKF algorithm has been shown to outperform the EKF in different state and parameter estimation problems ranging from machine learning (Van Der Merwe (2004); Haykin et al. (2001); de Oliveira (2012)) to

electric battery management system (Plett (2006)), and structural dynamics (Mariani & Ghisi (2007); Chatzi & Smyth (2009)).

2.1.1 State-space representation of the forward model for parameter identification

State-space representation of the forward model provides a convenient way to describe mathematical models of systems and structures whose outputs depend on the progressive change of the state variables, the model parameters, and external inputs/ loads. State-space models are mostly used to describe dynamical system models for problems of state estimation and control because they allow for intuitive, tractable, and on-line analysis of the systems in the time domain. Examples of using state-space models for this area of engineering can be seen in Nestorović-Trajkov et al. (2005, 2006). This representation leads to an intuition that state-space formulation can be very well used for representing models that require updating of their parameters in a recursive manner. In practice of recursive parameter identification that represents the forward model as state-space model (Hoshiya & Sutoh (1993); Wan & Van Der Merwe (2000); Van Der Merwe & Wan (2001)), state-space model is described by a stationary process transition Eq. 2.1a and the modeling observation equation Eq. 2.1b

$$\mathbf{m}_k = \mathbf{m}_{k-1} + \mathbf{w}_{k-1}, \tag{2.1a}$$

$$\mathbf{d}_k = \mathbf{h}\left(\mathbf{x}_k, \mathbf{m}_k\right) + \mathbf{v}_k, \tag{2.1b}$$

where vector $\mathbf{m}_k$ contains the model parameters and $\mathbf{x}_k$ is the physical state of the model at recursive step t_k. Vector $\mathbf{d}$ extracts modeling outputs at predefined positions in the model from the modeling observation Eq. 2.1b. In Eq. 2.1a, the stationary process transition is added by a modest amount of pseudo-noise $\mathbf{w}_k$ to act like parameter updating from one iterative step to the next. Inexactness of the modeling observation function $\mathbf{h}(\cdot)$ in Eq. 2.1b is characterized by an amount of uncertainty $\mathbf{v}_k$. Additive uncertainties $\mathbf{w}_k$ and $\mathbf{v}_k$ are assumed to be uncorrelated and white Gaussian having zero-mean and time-invariant covariance matrices $\mathbf{Q}$ and $\mathbf{R}^{\mathbf{m}}$ respectively.

The modeling observation function $\mathbf{h}(\cdot)$ for geomechanical models is a highly nonlinear function relating the current physical state $\mathbf{x}$ and the parameters set $\mathbf{m}$ (Murakami (1991); Hoshiya & Sutoh (1993)) to the modeling outputs $\mathbf{d}$. By the state-space presentation described here, the forward model $\mathbf{h}(\cdot)$ can be built separately from the identification

implementation, and it can also be modeled by different numerical methods or programming languages. In this work, forward modeling is done by FE software Plaxis 3D, and model parameters for geomechanical tunnel excavation are those that constitute a complex elastoplastic material model. Solving the modeling function $\mathbf{h}(\cdot)$, which is also termed forward problem, can be done rather fast and reliable nowadays thanks to the use of advanced FE codes and computers. However, because solving the forward problem in a recursive manner, as it is in the context of inverse analysis, needs quite a large number of forward calculations, the decisive factor that makes one inverse analysis method preferable to the others is that convergence can be achieved after an acceptably small number of forward calculations.

2.1.2 Recursive Bayesian filtering and nonlinear Kalman filters for parameter identification

In this section, the authors first introduce how inverse analysis can be solved by parameter identification technique in sequential Bayesian filtering scheme. Later, the adaptation of the EKF and the SPKF algorithms to solving inverse analysis problems will be given.

By in-situ measurement, a set of measurement data contaminated with measurement uncertainty is recorded. The measurement uncertainty may probably be provided by the instrument manufacturer. If not, one can assume measurement uncertainty to have deviated from a recorded value a small percentage of the maximum measured amplitude. In either case, it is reasonable to describe measurement uncertainty as Gaussian distribution

$$\rho_D(\mathbf{d}) \sim \mathcal{N}(\mathbf{d}^{obs}, \mathbf{R}^{obs}). \tag{2.2}$$

All prior knowledge about the model parameters is summoned to provide a good initial model. Prior information is apparently not precise and can, therefore, be defined as Gaussian distribution as well

$$\rho_M(\mathbf{m}) \sim \mathcal{N}(\mathbf{m}^{prior}, \mathbf{P}^{prior}). \tag{2.3}$$

When the given state-space model Eq. 2.1, the measurement data Eq. 2.2, and the prior knowledge about the model parameters Eq. 2.3 are in place, it is ready to perform parameter estimation using the filtering scheme, in particular the Kalman filters, to estimate the model parameters. Bayesian filtering (Ho & Lee (1964); Tarantola (2005)) provides a unifying framework for deriving the Kalman filters by means of conditional probability

$$\rho(\mathbf{m}_k|\mathbf{D}_k) = \frac{\rho(\mathbf{d}_k|\mathbf{m}_k)\rho(\mathbf{m}_k|\mathbf{D}_{k-1})}{\rho(\mathbf{d}_k|\mathbf{D}_{k-1})}, \tag{2.4}$$

where $\mathbf{D}_k$ is a series of modeling data up to and including recursive step k. The term on the left hand side of Eq. 2.4 is the posterior distribution of the estimating parameters given all the modeling data up to recursive step k. On the right hand side of Eq. 2.4, the first term in the numerator is called the *likelihood* which is defined in terms of the modeling Eq. 2.1b. The second term in the numerator is named the *prior* which depends on the prior knowledge Eq. 2.3 and the state transition Eq. 2.1a. The denominator in Eq. 2.4 is labeled the *evidence* which is the integral of the numerator over the parameter space. The evidence is considered a scaling factor and therefore does not need to be calculated explicitly in many cases.

The maximum a posteriori (MAP) is employed to solve Eq. 2.4. With Gaussian assumptions for the statistical quantities, MAP is equivalent to minimization of the misfit function Eq. 2.5

$$S(\mathbf{m}_k) = \frac{1}{2}\left((\mathbf{h}(\mathbf{m}_k) - \mathbf{d}^{obs})^t \mathbf{R}^{-1}(\mathbf{h}(\mathbf{m}_k) - \mathbf{d}^{obs}) + (\mathbf{m}_k - \hat{\mathbf{m}}_{k|k-1})^t \mathbf{P}_{k|k-1}^{-1}(\mathbf{m}_k - \hat{\mathbf{m}}_{k|k-1})\right), \tag{2.5}$$

where $\hat{\mathbf{m}}_{k|k-1}$ is the vector of estimated parameters before the modeling data $\mathbf{d}_k$ are taken into account (also called the a prior estimate) and $\mathbf{P}_{k|k-1}$ denotes the estimation error covariance of the estimation error term $\mathbf{m}_k - \hat{\mathbf{m}}_{k|k-1}$. The covariance matrix $\mathbf{R}$ is the sum of measurement uncertainties and modeling observation uncertainties, i.e. $\mathbf{R} = \mathbf{R}^{obs} + \mathbf{R}^m$ (Tarantola (2005)).

When modeling data are taken into account given the prior estimated parameters $\hat{\mathbf{m}}_{k|k-1}$, the MAP results in a recursive a posterior estimate of the parameters vector $\hat{\mathbf{m}}_k$:

$$\hat{\mathbf{m}}_k = \hat{\mathbf{m}}_{k|k-1} + \mathbf{K}_k\left(\mathbf{d}^{obs} - \mathbf{h}(\hat{\mathbf{m}}_{k|k-1})\right), \tag{2.6}$$

where $\mathbf{K}$ is the Kalman gain matrix resulted from minimization of Eq. 2.5. Calculation of $\mathbf{K}$ is of vital importance and is adapted for the particular approximation scheme. The a posterior error covariance $\mathbf{P}_k^m$ associated with $\hat{\mathbf{m}}_k$ is also dependent on the employed estimation scheme in such a way that the computation is facilitated.

Similar to the scheme applied in state estimation, Kalman filters estimate the model parameters by executing prediction step (time update) and correction step (measurement update) of the estimated parameters $\hat{\mathbf{m}}$ and the corresponding estimation error covariance matrix $\mathbf{P}^m$ recursively. The prediction step is performed based on the prior model and the state transition equation to predict the estimate, whereas the correction step requires measurement data and model calculated data to drive the estimated parameters to a point closer to the misfit's minimum. See Table 2.1 for the notations used to denote the mean estimate and its error covariance in each of the two steps in each iteration.

Table 2.1: Notations of the Gaussian moments in recursive estimation step k

Gaussian moments	Prediction step	Correction step
Mean parameter values : $\hat{\mathbf{m}}$	$\hat{\mathbf{m}}_{k\|k-1}$	$\hat{\mathbf{m}}_k$
Estimation error covariances: $\mathbf{P}^m$	$\mathbf{P}^m_{k\|k-1}$	$\mathbf{P}^m_k$

It is worth mentioning that Monte Carlo simulation based sequential filtering schemes such as the ensemble Kalman filter and the Particle filter can be employed to estimate the posterior distribution Eq. 2.4. Because fully random sampling is used in these schemes, the burden on the computation of forward modeling is enormous. For an exposition of the application of this stochastic filtering approach to geomechanical inverse problems, the reader can refer to Murakami et al. (2013); Shuku et al. (2012); Nguyen et al. (2014). In this work, the authors demonstrate that the use of the two deterministic nonlinear Kalman filters (the EKF and the SPKF) for model calibration of geomaterial parameters in geotechnical FE modeling is effective regarding convergence speed and reliability. In the following, the two nonlinear Kalman filters employed in this work — the extended Kalman filter and the sigma-point Kalman filter — are summarized.

2.1.3 The extended Kalman filter

The EKF is a linearized Kalman filter that can be applied to nonlinear state-space models. The EKF requires linearization of the state transition equation and the modeling equation around the current estimate. In the state-space model Eq. 2.1, because the state transition equation is linear, it needs only to calculate the Jacobian of the modeling function in Eq. 2.1b in order to approximate $\mathbf{h}(\cdot)$ to the linear form. The Jacobian of $\mathbf{h}(\cdot)$ at the current estimate of $\mathbf{m}$ is defined as

$$\mathbf{H}_k = \left.\frac{\partial \mathbf{h}(\mathbf{m}_k)}{\partial \mathbf{m}_k}\right|_{\mathbf{m}_k=\hat{\mathbf{m}}_{k|k-1}}, \tag{2.7}$$

where the current estimate is the a prior estimate $\hat{\mathbf{m}}_{k|k-1}$. In fact, when the modeling function $\mathbf{h}(\cdot)$ is a numerical model, there is no analytic calculation of the partial derivatives. Rather, the finite difference scheme is employed to approximate the Jacobian by taking the ratios of the change of each model output to the perturbation of each parameter by a finite small amount from the current estimate. If the paramter vector $\mathbf{m}$ contains n elements, this approximation of the Jacobian requires n function calls to the modeling function $\mathbf{h}(\cdot)$. This n function calls together with one further function call needed

in the posterior estimation Eq. 2.6 make totally $n+1$ function calls (or forward model calculations) for one EKF iterative step.

When the approximated Jacobian is used as the observation matrix for linear models, the equations for the EKF are the same as those originally derived for the Kalman filter (Kalman (1960); Jazwinski (1970)). The EKF algorithm for identifying parameters of a model represented by the state-space formulation Eqs. 2.1 is summarized in Algorithm 1.

Algorithm 1: The EKF algorithm for parameter identification

Initialization:

$$\begin{aligned} \hat{\mathbf{m}}_0 &= \mathbf{m}^{prior} \\ \mathbf{P}_0^m &= \mathbf{P}^{prior} \end{aligned}$$

$k \leftarrow 1$

while S *(Eq. 2.5)* $< TOLERANCE$ **do**

Prediction step:

$$\begin{aligned} \hat{\mathbf{m}}_{k|k-1} &= \hat{\mathbf{m}}_{k-1} \\ \mathbf{P}_{k|k-1}^m &= \mathbf{P}_{k|k}^m + \mathbf{Q} \end{aligned}$$

Correction step:

$$\begin{aligned} \mathbf{K}_k &= \mathbf{P}_{k|k-1}^m \mathbf{H}_k^t \{\mathbf{H}_k \mathbf{P}_{k|k-1}^m \mathbf{H}_k^t + \mathbf{R}\}^{-1} \\ \hat{\mathbf{m}}_k &= \hat{\mathbf{m}}_{k|k-1} + \mathbf{K}_k \left(\mathbf{d}^{obs} - \mathbf{h}(\hat{\mathbf{m}}_{k|k-1})\right) \\ \mathbf{P}_k^m &= \mathbf{P}_{k|k-1}^m - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k|k-1}^m \end{aligned}$$

$k \leftarrow k+1$

end

As proposed by Hoshiya & Sutoh (1993), to make the EKF's convergence rate faster for inverse analysis problems, the error covariance $\mathbf{P}^m$ is multiplied by a modification weight W in every correction step as expressed in Eq. 2.8. This multiplication is equivalent to making the fictitious difference between the true parameters and the estimated parameters more significant. The inflated covariance matrix $\mathbf{P}^m$ is inherited and magnifies the Kalman gain $\mathbf{K}$ in the successive iterative step. In this way it intensifies the effect of Kalman gain on the residual $\mathbf{d}^{obs} - \mathbf{h}(\hat{\mathbf{m}}_{k|k-1})$ and therefore makes convergence rate of the filter faster. The choice of the weight magnitude is problem-dependent, one can begin with a small

positive number larger than one and increase it as long as the EKF keeps converged.

$$\mathbf{P}_k^m = \left(\mathbf{P}_{k|k-1}^m - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{k|k-1}^m\right) W \tag{2.8}$$

2.1.4 The sigma-point Kalman filter

Although the EKF is one of the most widely used methods for state estimation of nonlinear systems, its inherent linearization makes it difficult to implement and only reliable for slightly nonlinear system models. Julier & Uhlmann (1997) developed an alternative type of the Kalman filter for nonlinear models that obviates the need for linearization of the state-space model: the unscented Kalman filter. Later, Van Der Merwe (2004) has combined and systematized this new filter together with some other independent works that derived the derivativeless deterministic sampling based approximation of the Gaussian statistics to belong to the family of SPKF. This family of the Kalman filter is based on the principle that a set of deterministically sampled points, called sigma-points, can be used to represent the mean and covariance of the estimated quantities. Update mechanisms are inherited from the linear Kalman filter (Kalman (1960)).

The sigma-points consist of $2n+1$ discretely distributed points around the current estimate on the n-dimension space. The distribution of these sigma-points around the current estimate is determined by square-root decomposition of the prior covariance and the spread of them can be assigned by setting a scaling parameter ζ as follows.

$$\boldsymbol{\sigma} = \zeta\sqrt{\mathbf{P}_{k-1}}, \tag{2.9}$$

$$(\boldsymbol{\mathcal{M}}_{k-1})_0 = \hat{\mathbf{m}}_{k-1}, \tag{2.10}$$

$$(\boldsymbol{\mathcal{M}}_{k-1})_i = \hat{\mathbf{m}}_{k-1} + (\boldsymbol{\sigma})_i, \text{ for i = 1:n} \tag{2.11}$$

$$(\boldsymbol{\mathcal{M}}_{k-1})_{i+n} = \hat{\mathbf{m}}_{k-1} - (\boldsymbol{\sigma})_i, \text{ for i = 1:n} \tag{2.12}$$

where the scaling parameter is calculated as $\zeta = \sqrt{n+\lambda}$. The parameter λ remains free to tune. It is suggested to select $n+\lambda = 3$ when $\mathbf{m}$ is assumed Gaussian (Julier & Uhlmann (1997)). The notation $(\cdot)_i$ denotes the i-column of the matrix within parentheses.

Each sigma-point is associated with a weight. The weights are defined as in Eq. 2.13 such that the sum of weights is unity.

$$W_0 = \frac{\lambda}{n+\lambda}, W_i = W_{i+n} = \frac{1}{2(n+\lambda)}, \text{ for i = 1:n} \tag{2.13}$$

Direct nonlinear transfer of the sigma-points through the nonlinear state-space functions helps reserve second order accuracy of the Gaussian approximate estimated quantities as

opposed to first order truncation of the Taylor series employed by the EKF. The recursive prediction step and correction step of the SPKF are summarized in Algorithm 2.

Algorithm 2: The SPKF algorithm for parameter identification

Initialization:

$$\begin{aligned} \hat{\mathbf{m}}_0 &= \mathbf{m}^{prior} \\ \mathbf{P}_0^m &= \mathbf{P}^{prior} \end{aligned}$$

$k \leftarrow 1$

while S *(Eq. 2.5)* $< TOLERANCE$ **do**

Prediction step:

Generate the sigma-points Eq. 2.9-2.12:$(\boldsymbol{\mathcal{M}}_{k-1})_i$

$$(\boldsymbol{\mathcal{M}}_{k|k-1})_i = (\boldsymbol{\mathcal{M}}_{k-1})_i$$

$$\hat{\mathbf{m}}_{k|k-1} = \sum_{i=0}^{2n} W_i (\boldsymbol{\mathcal{M}}_{k|k-1})_i$$

$$\mathbf{P}_{k|k-1}^m = \sum_{i=0}^{2n} W_i \left((\boldsymbol{\mathcal{M}}_{k|k-1})_i - \hat{\mathbf{m}}_{k|k-1} \right) \left((\boldsymbol{\mathcal{M}}_{k|k-1})_i - \hat{\mathbf{m}}_{k|k-1} \right)^t + \mathbf{Q}$$

$$(\boldsymbol{\mathcal{D}}_{k|k-1})_i = \mathbf{h}\left((\boldsymbol{\mathcal{M}}_{k|k-1})_i \right)$$

$$\hat{\mathbf{d}}_{k|k-1} = \sum_{i=0}^{2n} W_i (\boldsymbol{\mathcal{D}}_{k|k-1})_i$$

$$\mathbf{P}_{k|k-1}^d = \sum_{i=0}^{2n} W_i \left((\boldsymbol{\mathcal{D}}_{k|k-1})_i - \hat{\mathbf{d}}_{k|k-1} \right) \left((\boldsymbol{\mathcal{D}}_{k|k-1})_i - \hat{\mathbf{d}}_{k|k-1} \right)^t + \mathbf{R}$$

$$\mathbf{P}_{k|k-1}^{md} = \sum_{i=0}^{2n} W_i \left((\boldsymbol{\mathcal{M}}_{k|k-1})_i - \hat{\mathbf{m}}_{k|k-1} \right) \left((\boldsymbol{\mathcal{D}}_{k|k-1})_i - \hat{\mathbf{d}}_{k|k-1} \right)^t$$

Correction step:

$$\begin{aligned} \mathbf{K}_k &= \mathbf{P}_{k|k-1}^{md} (\mathbf{P}_{k|k-1}^d)^{-1} \\ \hat{\mathbf{m}}_k &= \hat{\mathbf{m}}_{k|k-1} + \mathbf{K}_k \left(\mathbf{d}^{obs} - \hat{\mathbf{d}}_{k|k-1} \right) \\ \mathbf{P}_k^m &= \mathbf{P}_{k|k-1}^m - \mathbf{K}_k \mathbf{P}_{k|k-1}^d \mathbf{K}_k^t \end{aligned}$$

$k \leftarrow k + 1$

end

As it can be seen from the SPKF Algorithm 2, the means of predicted parameters and model outputs are calculated as a weighted summation over the sigma-points and the corresponding predicted model outputs. The predicted model outputs $\boldsymbol{\mathcal{D}}_{k|k-1}$ are the results of direct nonlinear transfers of the representative sigma-points $\boldsymbol{\mathcal{M}}_{k|k-1}$ through the nonlinear model function $\mathbf{h}(\cdot)$. The predicted estimation error covariances of the estimated parameters $\mathbf{P}^m$, of the model outputs $\mathbf{P}^d$, and of the cross-covariance between them $\mathbf{P}^{md}$ are calculated as the weighted summation of the squares of distances between each sigma-point and respective model output to the corresponding mean estimates. The correction equations for the SPKF take into account the means and covariances approximated using the sigma-points and the nonlinear model outputs in the prediction step without calculation of the Jacobians in the state-space model needed.

For application in the calibration of FE model in geomechanics, value for each parameter can be limited to a certain range depending on initial categorization of geomaterial samples. Therefore, it is advantageous to apply a feasible range for each parameter in the identification process. The advantage of giving constraints to the parameter space is threefold: (i) it helps to exclude the unreasonable convergence points, which are local minima, of the objective function, (ii) it allows the FE code to run uninterruptedly during the iterative estimation process as the parameters are clipped in the reasonable regions, and (iii) it may help to accelerate the estimation process by filtering algorithms since the estimation bias is not allowed to explode. Here, the 'clipping' constraints handling for both the EKF and the SPKF described in Kandepu et al. (2008) is applied.

2.2 Application for parameter identification and optimization to test models

2.2.1 Identification of hidden parameter in a nonlinear noisy time series

A time series model previously studied by Van Der Merwe (2004) and Chatzi & Smyth (2013) described in Eq. 2.14 is selected for a study of parameter identification using the EKF and the SPKF.

$$x_k = mx_{k-1} + 1 + \sin\left(\omega\pi(k-1)\right) + u_{k-1}, \tag{2.14}$$

where, $u_k \sim \mathcal{N}(0, C)$ is the additive process noise with negligibly small variance C, $\omega = 0.04$ and m is the hidden parameter that requires identification. A non-stationary

observation model below serves as the measurement equation:

$$d_k = \begin{cases} 0.4(x_k)^2 + n_k & \text{if } k \leq 30, \\ 0.5x_k - 2 + n_k & \text{if } k > 30. \end{cases} \tag{2.15}$$

The measurement noise n_k is drawn from Gaussian distribution $\mathcal{N}(0, \sigma^2)$. The Kalman filters are to estimate the hidden parameter m for the cases of low and high noise levels in the measurements. The state-space equations are reformulated by setting stationary transition of the parameter as described in the last section in the following manner.

$$m_k = m_{k-1} + w_{k-1}, \tag{2.16}$$

where, $w \sim \mathcal{N}(0, Q)$ is the pseudo noise that drives the parameter from one time step to the next time step. According to the authors' practice (Nguyen et al. (2013)), the variance Q can be set so that standard deviation of process noise to be around ten percent that of the initial estimate error, i.e. $Q = 0.01P_0$. The state transition equation is then evaluated as

$$x_k = m_{k-1}x_{k-1} + 1 + \sin(\omega\pi(k-1)). \tag{2.17}$$

The process noise for the state transition is intentionally ignored for the purpose of fitting the model response to noisy measurement data. Estimation results for measurement noise levels are presented in Fig. 2.1. As can be observed, both EKF and SPKF equally well estimate the true parameter in the case of low measurement noise level. For the low measurement level, zero-mean Gaussian noise with standard deviation $\sigma = 10^{-4}$, the estimated values are biased by small distances from the true parameter very soon after estimation starts. At the beginning of the changed observation model at time $t = 30$, it happens that the estimates fluctuate by a large amount but gradually come to unbiased positions when more measurement data are received. For the higher noise level, zero-mean Gaussian noise with standard deviation $\sigma = 10^{-3}$, of the measurements it is observed that more significant bias of the estimates along time steps is resulted both for the EKF and the SPKF.

2.2.2 Optimization of nonlinear multi-minima Ackley's function

In this example, the authors would like to evaluate the capability of the nonlinear Kalman filters in performing optimization of a nonlinear multimodal function. The test function

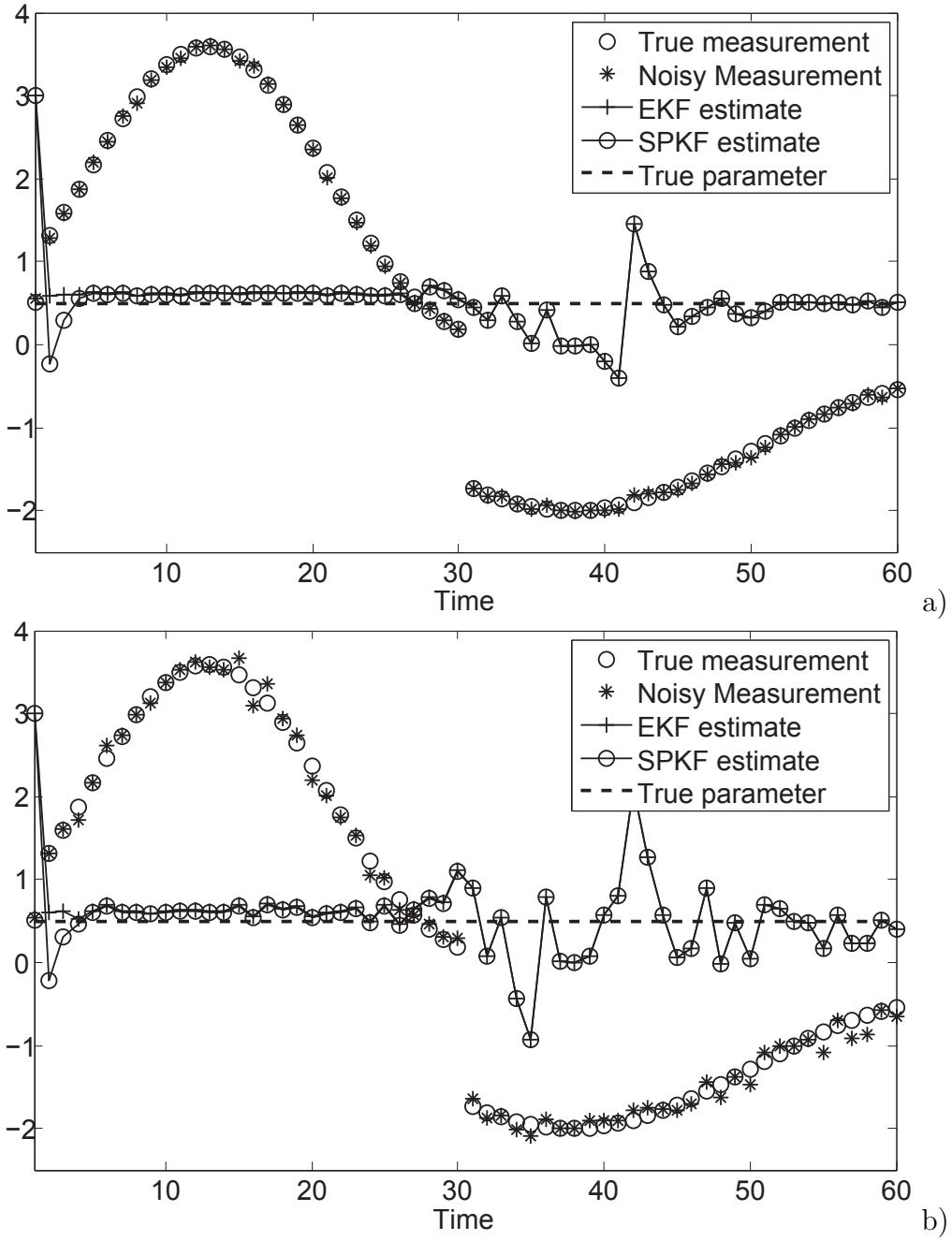


Figure 2.1: Estimate of hidden parameter for a noisy time series - initial parameter value $m = 3$, true parameter value $m = 0.5$, a) Gaussian noise $\sigma = 10^{-4}$, b) Gaussian noise $\sigma = 10^{-3}$

Table 2.2: Position of the estimated minimum point of the Ackley's function after 20 iterations. Points marked with † signify wrong final estimates.

Starting point (x_1, x_2)	EKF estimate	SPKF estimate
(1,1)	$(-7.83 * 10^{-15}, -7.83 * 10^{15})$	$(6.64 * 10^{-05}, 8.03 * 10^{-06})$
(-3,2)	$(-0.89, -1, 18)^{\dagger}$	$(7.47 * 10^{-06}, 1.00 * 10^{-05})$
(3,-2)	$(0.89, 1, 18)^{\dagger}$	$(-7.47 * 10^{-06}, -1.00 * 10^{-05})$
(4,2)	$(-5, -5)^{\dagger}$	$(-1.22 * 10^{-06}, -1.88 * 10^{-07})$
(4,4)	$(3.73, 3.73)^{\dagger}$	$(-0.000107, 0.000191)$

is chosen to be the two-variable Ackley's function which has the following mathematical expression

$$d = -20\exp\left(-0.2\sqrt{0.5(x_1^2 + x_2^2)}\right) - \exp\left(0.5(\cos(2\pi x_1) + \cos(2\pi x_2))\right) + 20 + \exp(1). \tag{2.18}$$

This Ackley's function has a global minimum at $(x_1 = 0, x_2 = 0)$ and many local minima in the variable space which makes finding the global minimum by numerical optimization algorithms a challenging task, specifically for algorithms that are based on gradients.

The EKF and SPKF are identically configured to run in an iterative manner similar to that of the previous parameter identification problem of time series but with the measurement that is evaluated from Eq. 2.18 at the global minimum point. For the purpose of solving the optimization problem, the measurement noise is set to be negligibly small. It has been observed that the EKF can easily be trapped within regional minima, as one example with initial state variables at $x_1 = 4, x_2 = 4$ in Fig. 2.2 shows. The SPKF, however, consistently drives the state variables to the global minimum point despite the arbitrary choice of the starting point as listed in Table 2.2.

2.3 Model calibration of soil parameters in mechanized tunneling

Geomechanical analysis of tunnel excavation that is considered in this work is the quasi-static non-Markov process, i.e. the state of the model is about equilibrium along the analysis, and the current state depends not only on the previous state but also on the

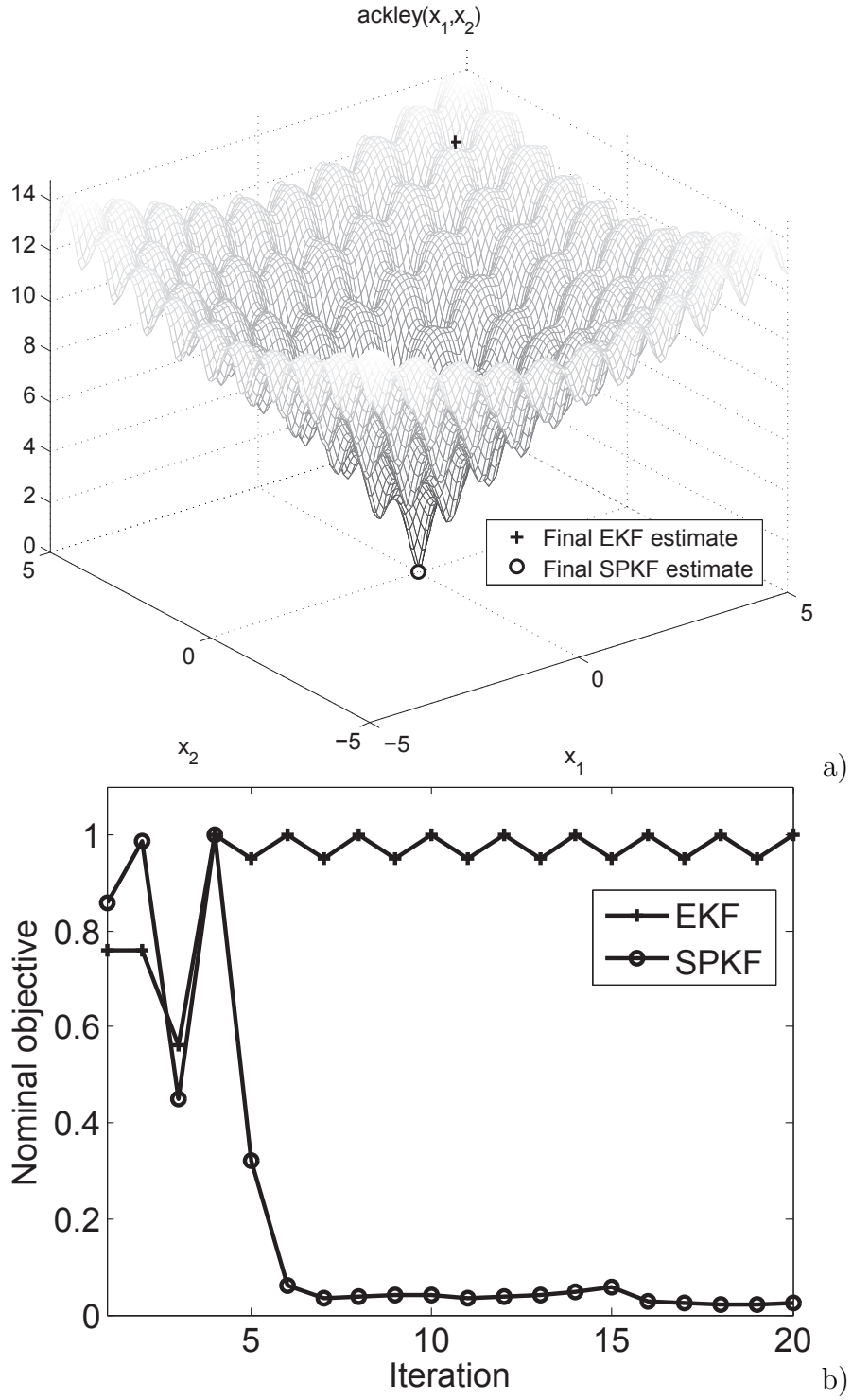


Figure 2.2: a) Ackley's function and final estimates of its global minimum, initial point $x_1 = 4$, $x_2 = 4$. b) History of the normalized objective functions

history of analysis steps before given constant geomechanical parameters. As a consequence, the manner by which the nonlinear Kalman filters are employed to identify the model parameters is adapted from recursive steps in time to entire iterative steps. That means the geomechanical model analysis is performed entirely in one iterative step and the identification process is run until the discrepancy between model calculated outputs and the in-situ measurement data cannot become smaller. Alternatively, the measure of this data misfit can be quantified by Eq. 2.5. The iterative run of the Kalman filter algorithms, thus, can run until this misfit measure is smaller than a predefined tolerance.

2.3.1 Forward model and noisy synthetic measurement data

The FE model was provided by the C2-subproject within the Collaborative Research Center SFB 837 at the Ruhr-University Bochum. In this model, a shield tunnel boring machine (TBM) advancing in homogeneous soil is modeled using the FE analysis software PLAXIS 3D. The model domain is 60 m long, 40 m wide, and 45 m deep. The tunnel of diameter D = 8.5 m and under an overburden depth equal to 1D. The overall model is depicted in Fig. 2.3. Except that the ground surface is stress-free, all other boundaries of the FE model are constrained by zero displacements. Only a half of the model is analyzed because of the symmetric geometrical, loading and material conditions about the vertical plane along the tunnel axis.

The homogeneous soil behaves according to the Hardening Soil constitutive model which has complex stress dependent stiffness moduli, yield surface and flow rule as described in Schanz et al. (1999). The slurry shield TBM is 9 m long modeled by circular elastic plate elements. Each concrete lining ring has the length of 1.5 m also modeled by elastic plate elements. For detailed descriptions regarding support pressure, soil excavation, grouting, erection of tunnel linings, and shield skin-ground and tunnel lining-ground contacts used in the FE model the reader is referred to Zarev et al. (2011).

Assuming that properties of the elastic materials are known, a set of soil's stiffness and failure parameters is chosen for identification: stiffness for un-/reloading G_{ur}, secant stiffness in standard drained triaxial test E_{50}, and friction angle ϕ.

Observation positions and orientations at observation surface at 39 m from beginning position of the tunnel head (before excavation) and perpendicular to the tunnel axis are depicted in Fig. 2.3 as well.

Tunneling-induced displacements are simulated for 30 excavation steps, and each step advances the TBM 1.5 m forward. Calculated outputs obtained from the FE model are

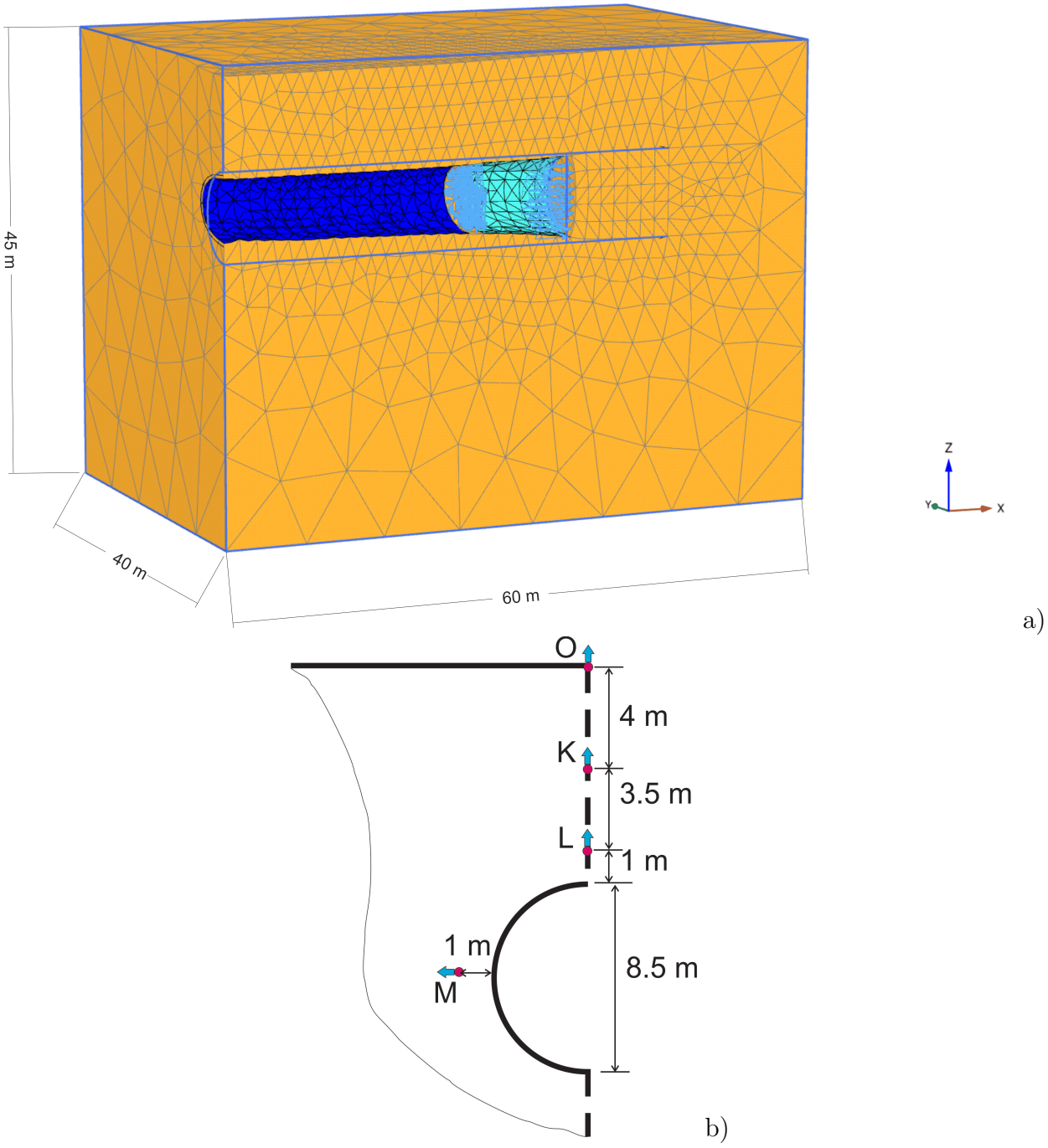


Figure 2.3: a) FE model of the 3D tunnel excavation and b) Measurement points at observation section

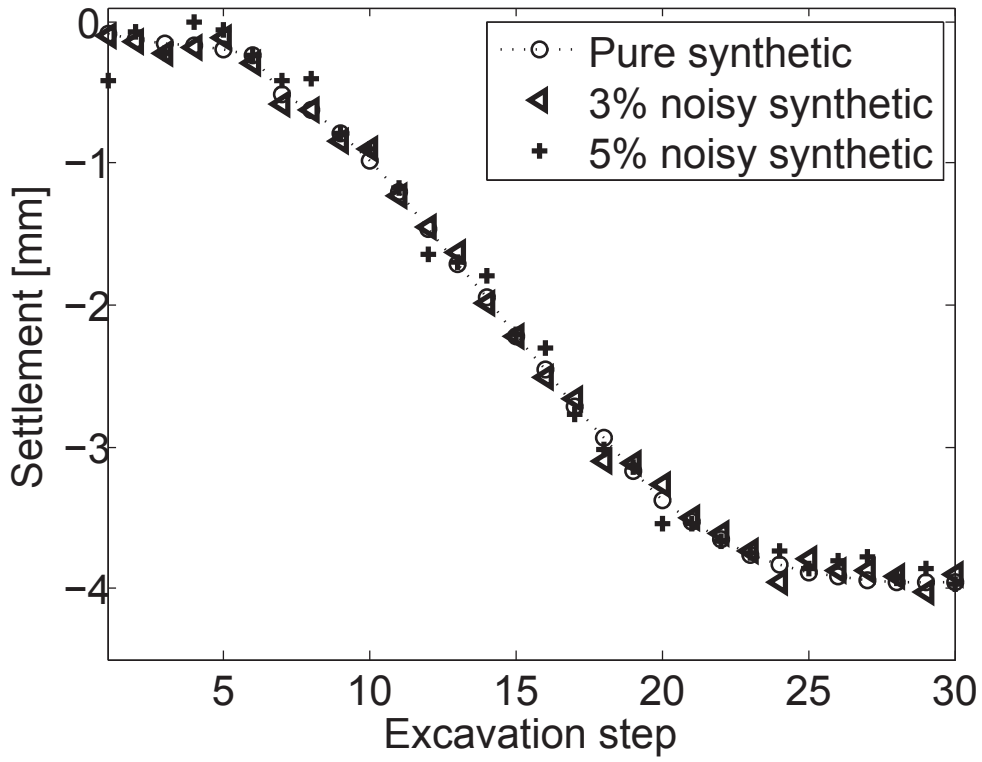


Figure 2.4: Pure calculated and synthetic noisy measurements at Point O

registered at every excavation step. To produce the synthetic observation data, the model calculated displacements, given a set of 'true' model parameters, are added by an amount of zero-mean Gaussian noise. The true model parameters for this model are predefined as

$$G_{ur} = 41600[\mathrm{kN/m^2}], \phi = 35[^\circ], E_{50} = 35000[\mathrm{kN/m^2}]. \tag{2.19}$$

In contrast to the use of field measurement data whose noise level is difficult to estimate, using synthetic data makes it possible to control the amount of noise that contaminates data. Different levels of noise variances are used to examine the robustness of the EKF and the SPKF against noisy measurement data. It is taken into account that measurement noise has standard deviations of three percent and five percent the mean value of the noise-free FE calculated outputs for each measurement point. Synthetic measurement data with three percent additive Gaussian noise at measurement Point O are plotted in comparison with pure FE analysis response in Fig. 2.4.

2.3.2 Results and discussion

Let us assume a very bad situation in which one has quite little knowledge of the soil parameters. The feasible ranges for the considered soil parameters are defined such that the lower bounds and upper bounds are considerably distant from true parameters as follows

$$20833 \leq G_{ur}[\mathrm{kN/m^2}] \leq 62499, \tag{2.20}$$

$$17500 \leq E_{50}[\mathrm{kN/m^2}] \leq 52500, \tag{2.21}$$

$$32 \leq \phi[^\circ] \leq 52.5. \tag{2.22}$$

The same initial estimation error covariance $\mathbf{P}_0$, the covariance matrix of the process noise $\mathbf{Q}$, and covariance matrix of measurement noise $\mathbf{R}$ are set for all the inverse analysis experiments.

The initial estimation error covariance $\mathbf{P}_0$ expresses how much confidence is associated with particular sets of initial parameters in a Gaussian distribution sense. If it is very uncertain about the model parameters, the covariances can be set arbitrarily large to cover the high uncertainty on the initial choice of parameters. Taken into account the assumption that the parameters are uncorrelated, the initial state error covariance matrix for the parameters $\mathbf{m} = [G_{ur}; \phi; E_{50}]$ is set: $\mathbf{P}(0|0) = \mathrm{diag}\left(4000^2, 5^2, 5000^2\right)$.

The authors' practice for setting the covariance matrix of the process noise $\mathbf{Q}$ is that it can be set in such a way that its square roots are ranging from five percent to ten percent the values of initial estimation error standard deviations. In this work, it is chosen to be around seven percent, i.e. the covariance matrix of the process noise can be set as $\mathbf{Q} = 0.005\mathbf{P}_0$.

The covariance matrix of measurement noise $\mathbf{R}$ reflects the accuracies of the measurement data and the modeling observation equation. In fact, the values in $\mathbf{R}$ do not necessarily reflect the true covariances of these sources of uncertainties. The same as for the initial estimation error covariance $\mathbf{P}_0$, if appropriate measurement setup with high-quality measurement devices and a reliable numerical model are given, small values for diagonal elements in $\mathbf{R}$ can be assigned and vice versa. For setting up the nonlinear Kalman filters in this study, every measurement datum is assigned to have the same covariance, i.e. $\mathbf{R} = 0.01^2 * \mathbf{I}(m, m)$, with $\mathbf{I}(m, m)$ the identity matrix having a dimension equal to the size of the measurement data, m.

Two initial sets at the extremes of feasible ranges are chosen for evaluating the performance of the EKF and SPKF. Identification results obtained at the final iterative step

Table 2.3: Identified soil parameters - Initial case 1 with three percent measurement noise

Parameter	Initial		EKF		SPKF	
	Mean	SD	Mean	SD	Mean	SD
G_{ur}[kN/m^2]	20833	4000	41588.02	552.30	41476.61	296.64
E_{50}[kN/m^2]	17500	5000	34856.67	958.45	34643.40	387.50
ϕ[°]	32	5	34.96	0.25	34.91	0.35

Table 2.4: Identified soil parameters - Initial case 2 with three percent measurement noise

Parameter	Initial		EKF		SPKF	
	Mean	SD	Mean	SD	Mean	SD
G_{ur}[kN/m^2]	62499	4000	41633.92	558.93	41735.41	297.10
E_{50}[kN/m^2]	52500	5000	34944.11	974.28	35133.16	389.89
ϕ[°]	52.5	5	34.98	0.25	35.03	0.35

are presented in Table 2.3 and Table 2.4 for two different initial cases. Very good convergence of the means to the true parameter values is observed, whereas the standard deviations (SDs) become significantly small at the end of EKF and SPKF identification processes. Of the estimated results obtained by the EKF or the SPKF, the estimation error covariances (or SDs) does not necessarily represent neighborhood region in which the model parameter may be situated because the initial error covariance matrix $\mathbf{P}_0$ as well as covariances of process noise $\mathbf{Q}$ and measurement noise $\mathbf{R}$ are not necessarily true when setting up the EKF and SPKF. However, besides the misfit function, the estimation error covariance matrix can be helpful for tracking the performance of the EKF and SPKF along iterative steps as the diagonal elements in this matrix should become smaller in converging iterative identification processes.

Graphical plots of the converging soil parameters for the presented two initial cases Figs. 2.5 and 2.6 demonstrate fast and robust convergence. Convergence rates of the EKF and the SPKF are quite the same for the Initial case 1 in which the set of soil parameters is unbiasedly identified at the fifth iteration. Regarding the number of forward calculations needed for five iterative steps, the EKF has used 20 calls to FE analysis and the SPKF has used 35 of them. Few numbers of calls to FE runs are advantageous for back-analysis as each FE analysis of a complex geotechnical model such as in modeling of mechanized tunneling consumes a considerable amount of time.

Table 2.5: Identified soil parameters - Initial case 1 with five percent measurement noise

Parameter	Initial		EKF		SPKF	
	Mean	SD	Mean	SD	Mean	SD
G_{ur}[kN/m^2]	20833	4000	42483.47	6317.98	42288.64	392.62
E_{50}[kN/m^2]	17500	5000	37036.86	1130.50	36646.55	600.89
ϕ[°]	32	5	35.52	2.82	35.42	0.37

For the Initial case 2, Fig. 2.6, the EKF requires several more iterations than the SPKF to obtain convergence. It is observed that the estimated parameter values at a few beginning steps of EKF filtering process exhibit significant oscillations, the value of parameter E_{50} is even out of bounds for the first three iterative steps of the EKF. The SPKF, however, keeps the convergence speed steady for both initial cases.

To experiment how detrimental higher level of measurement noise affects the performance of the EKF and the SPKF identification of Initial case 1 is rerun with five percent measurement noise. It can be observed from Fig. 2.7 that more iterations are required for both the Kalman filters to get to convergence. The converged parameter values are much more biased than those in the three percent measurement noise case. The summary results in Table 2.5 also show significant standard deviations of the identified parameter values which indicate less reliable parameter estimates compared to the previous cases of three percent measurement noise.

Besides the observation that the estimation error covariance gets reduced significantly in either case of measurement noise levels, converging development of the misfit (the first term in Eq. 2.5) during inverse analysis process shown in Figs. 2.8 and 2.9 confirms convergence in a sense that misfit between measured and simulated settlements is minimized.

In general, it does not need so many calls to FE analysis of the investigated geotechnical model for performing model calibration with the EKF or the SPKF. The overall computation time depends mostly on how much time it takes for an analysis of the FE model. For the presented tunneling model in this work, computer experiments for inverse analysis require on average two days to obtain the identified soil parameters on a standard personal computer (Intel(R) Core(TM)2 Duo 1.97 GHz, 4 GB RAM). Speed-up of the inverse analysis process can be achieved by running the FE analyses in parallel by multiple computers for purpose of approximating the Jacobian for the EKF or evaluating the transfer of the sigma-points for the SPKF.

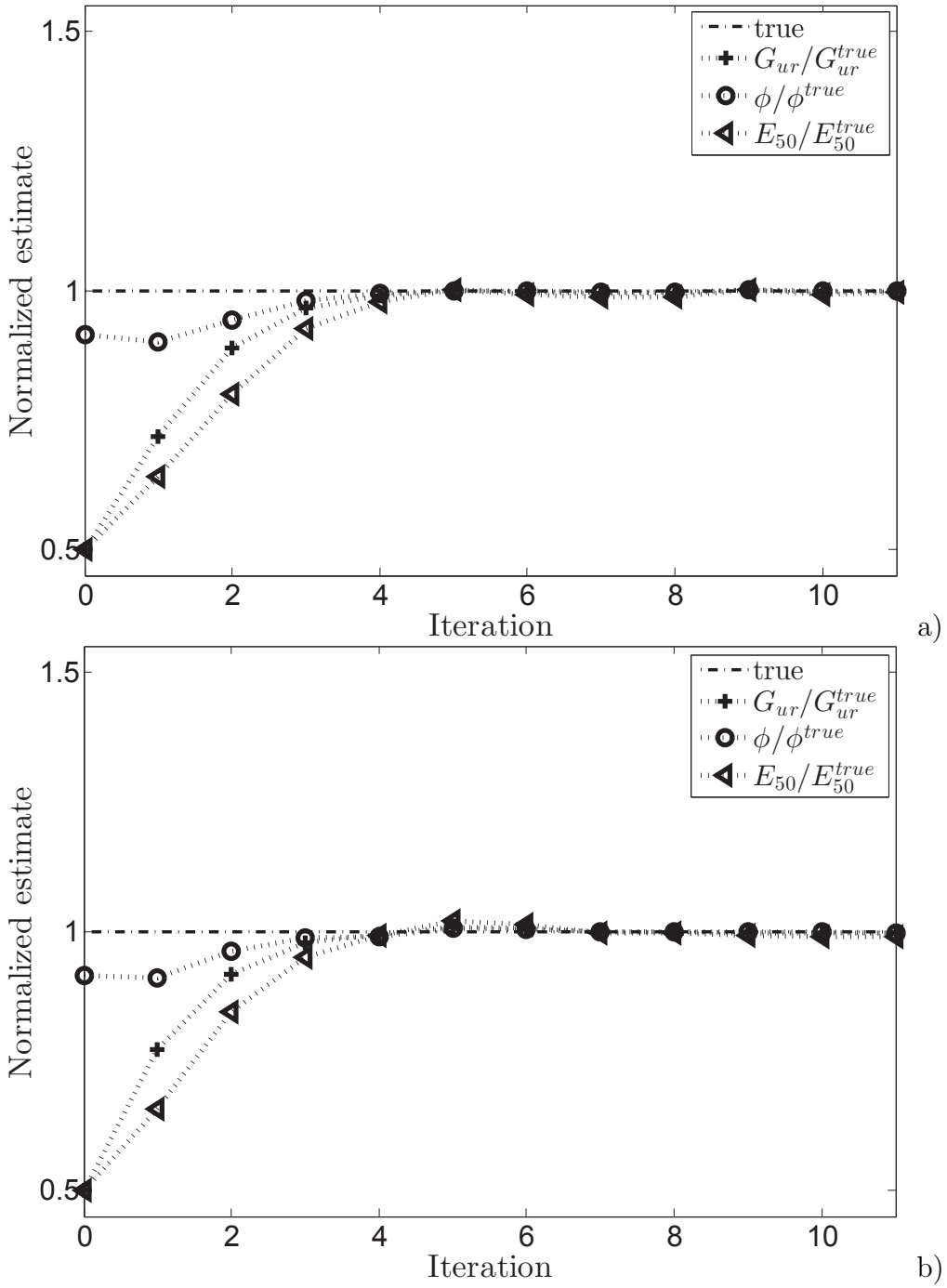


Figure 2.5: Iterative development of soil parameters for Initial case 1 with three percent measurement noise a) EKF and b) SPKF

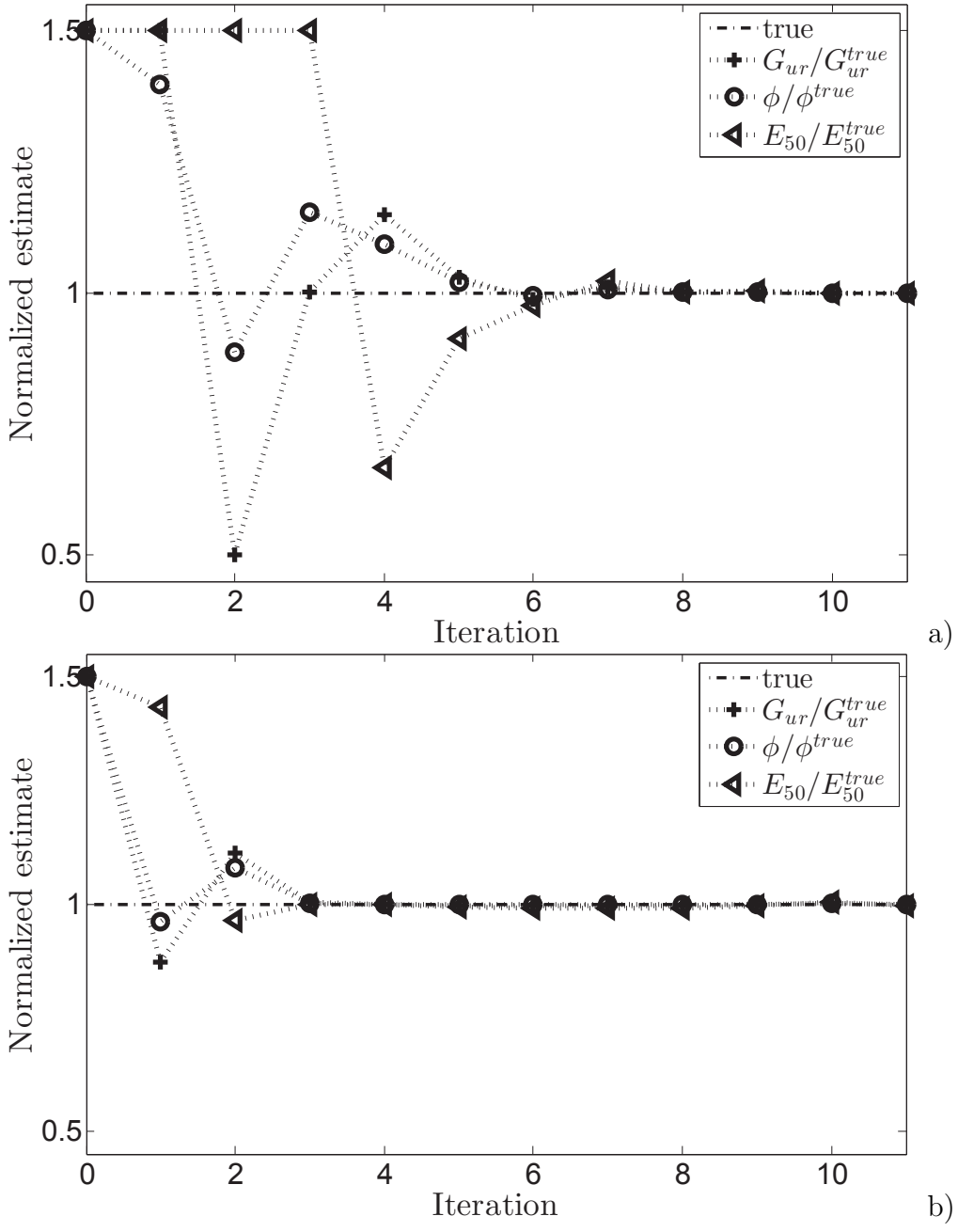


Figure 2.6: Iterative development of soil parameters for Initial case 2 with three percent measurement noise a) EKF and b) SPKF

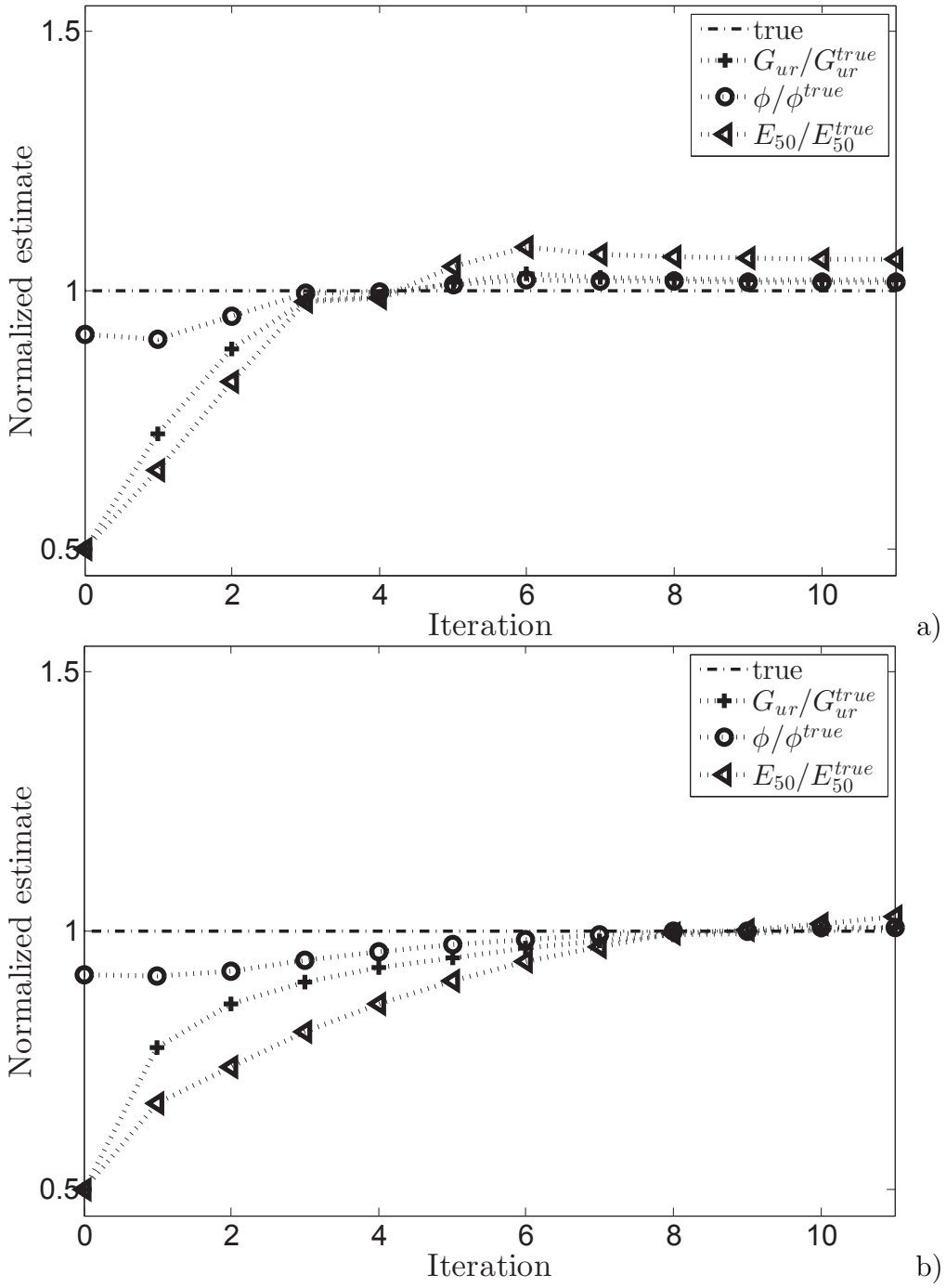


Figure 2.7: Iterative development of soil parameters for Initial case 1 with five percent measurement noise a) EKF and b) SPKF

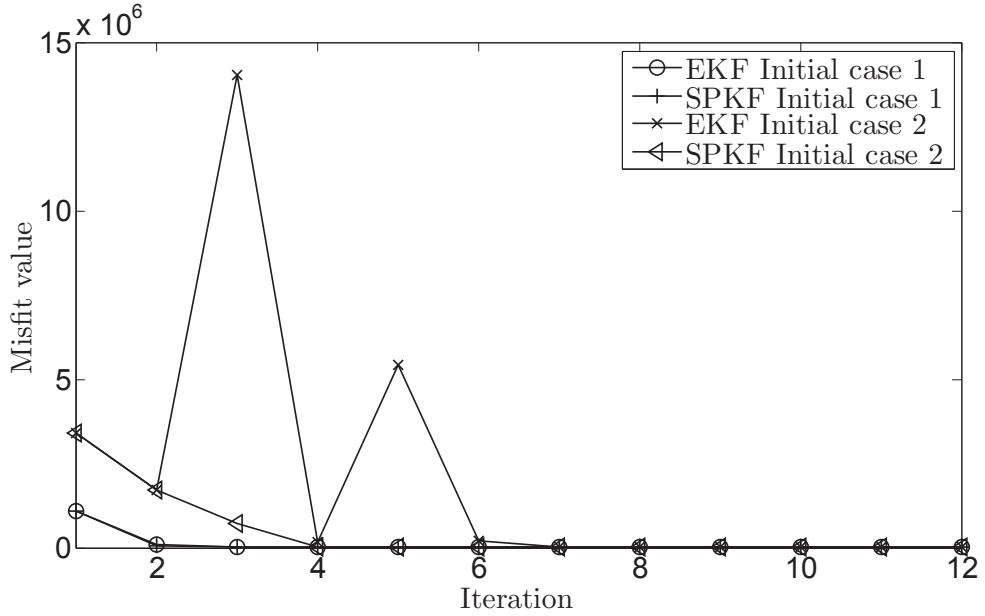


Figure 2.8: Comparison of misfit development for the case three percent measurement noise

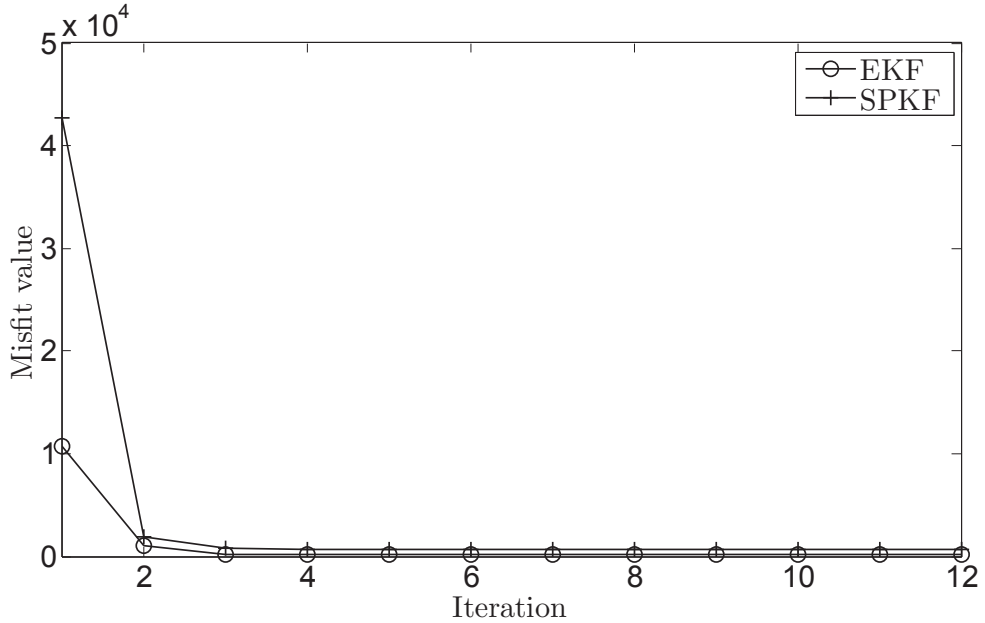


Figure 2.9: Comparison of misfit development for Initial case 1 and with five percent measurement noise

3 Particle filter-based data assimilation for uncertainty quantification

3.1 Data assimilation for inversion of model parameters

3.1.1 State-space formulation

I apply in this work the sequential data assimilation (DA), i.e. DA process is fetched whenever field measurement data become available to update the numerical modeling. This sequential scheme is especially suited for numerical simulation of long time loading or multiple-step geomechanical analyses such as soil consolidation, construction of embankments, and tunnel excavations. To describe the assimilation process, let us represent the forward modeling of geomechanical problems in state-space equations as follows.

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{w}_k; \qquad \mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k), \tag{3.1a}$$

$$\mathbf{d}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k; \qquad \mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k^m), \tag{3.1b}$$

where state vector $\mathbf{x}_k$ stores the state variables (e.g. displacements) at analysis step t_k. This state vector is also augmented by the model parameters (i.e. constitutive parameters). The modeling function $\mathbf{f}(\cdot)$ represents the numerical model (e.g. finite element model) that transfers the state from one calculation step to the next. Output vector $\mathbf{d}_k$ represents the observed degrees of freedom resulted from observation operator $\mathbf{h}(\cdot)$ which is simply a selection matrix in case of linear observation. Additive uncertainties $\mathbf{w}_k$ and $\mathbf{v}_k$ are assumed to be uncorrelated and white Gaussian having zero-mean and covariance matrices $\mathbf{Q}_k$ and $\mathbf{R}_k^m$ respectively.

By in situ observation, a set of observation data contaminated with observation uncertainty $\mathbf{d}_k^{obs}$ is obtained. The amount of measurement noise can be represented also by zero-mean Gaussian with covariance matrix $\mathbf{R}_k^{obs}$. As a consequence, this covariance matrix is added to the one from observation Eq. 3.1b $\mathbf{R}_k^m$ to form a single observation covariance matrix, i.e. $\mathbf{R}_k = \mathbf{R}_k^{obs} + \mathbf{R}_k^m$ (Tarantola (2005)).

Because geomechanical analysis is a non-Markov process, i.e. state vector $\mathbf{x}$ at t_k depends not only on its values at t_{k-1} but also on values from all previous analysis steps, the ensemble Kalman filter (EnKF) and other ensemble based estimation methods that require re-sampling in the parameter space are not suitable for straightforward implementation of the DA process (Evensen (2003); Nakano et al. (2007)). Instead, the particle filter with sequential importance sampling, which can be applied to non-Markov process is proposed for identification in geomechanics (Murakami et al. (2013)). The particle filter requires neither linearization in the state-space equations nor the assumption of Gaussian densities for the state variables and noises therein.

3.1.2 Sequential data assimilation using particle filter

I have employed the particle filter (Gordon et al. (1993)) with sequential importance sampling (Doucet et al. (2000)) for updating the posterior distribution in each assimilation step. As it is a type of sequential Monte Carlo method, the particle filter is very suitable for working with nonlinear models having uncertain parameters represented by arbitrary probability density function. The particle filter operates by executing prediction step and filtering (correction) step in a sequence. As measurements at time t_k are received, the posterior probability density of model parameters given all the data up to time step t_k, $\mathbf{D}_k$, is calculated by the Bayes' rule Eq. 3.2 in the filtering step below.

$$\rho(\mathbf{x}_k|\mathbf{D}_k) = C^{-1}\rho(\mathbf{d}_k|\mathbf{x}_k)\rho(\mathbf{x}_k|\mathbf{D}_{k-1}), \tag{3.2}$$

where $C = \int \rho(\mathbf{d}_k|\mathbf{x}_k)\rho(\mathbf{x}_k|\mathbf{D}_{k-1})d\mathbf{x}_k$ is a normalizing constant. The probability density $\rho(\mathbf{d}_k|\mathbf{x}_k)$ in Eq. 3.2 is called the likelihood that quantifies how likely the model is in explaining the measurement data. Evaluation of this term, which is crucial for the implementation of the particle filter, will be presented below. The third term in Eq. 3.2, $\rho(\mathbf{x}_k|\mathbf{D}_{k-1})$ is the prediction probability density which is resulted from the initial probability density of the model parameters and the forward modeling Eq. 3.1a.

Prediction step to calculate the prior probability density of the state variables at time step t_k given observation data up to the previous time step t_{k-1} is represented by the integral as Eq. 3.3.

$$\rho(\mathbf{x}_k|\mathbf{D}_{k-1}) = \int \rho(\mathbf{x}_{k-1}|\mathbf{D}_{k-1})\rho(\mathbf{x}_k|\mathbf{x}_{k-1})d\mathbf{x}_{k-1}, \tag{3.3}$$

Direct evaluations of integrals in Eqs. 3.2 and 3.3 are not practically possible in high dimensional space and arbitrary probability density functions. Instead, in the particle

filter estimation, the integrals are effectively approximated by direct Monte Carlo simulations. The probability distribution function (PDF) in the model space is represented at time step t_k by an ensemble of Ns particles $\{\mathbf{x}_k^1, \mathbf{x}_k^2, \cdots, \mathbf{x}_k^{Ns}\}$. Each particle in the ensemble is associated with a weighting factor so as to form the corresponding set of weights $\{w_k^1, w_k^2, \cdots, w_k^{Ns}\}$. Sum of weighs at any time step t_k must satisfy unity condition $\sum_{i=1}^{Ns} w_k^i = 1$. Thus, the filtered PDF at at any time step t_{k-1} is approximated by weighted summation Eq. 3.4.

$$\rho(\mathbf{x}_{k-1}|\mathbf{D}_{k-1}) \approx \sum_{i=1}^{Ns} w_{k-1}^i \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1|k-1}^i), \tag{3.4}$$

where δ is Dirac delta function.

Monte Carlo approximation of Eq. 3.3 to the corresponding prior PDF is made using the state transition Eq. 3.1a and the weighted summation Eq. 3.4 as follows.

$$\begin{aligned} \rho(\mathbf{x}_k|\mathbf{D}_{k-1}) &\approx \int \sum_{i=1}^{Ns} w_{k-1}^i \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1|k-1}^i) \\ &\quad \rho(\mathbf{x}_k|\mathbf{x}_{k-1}) d\mathbf{x}_{k-1} \\ &= \sum_{i=1}^{Ns} w_{k-1}^i \delta(\mathbf{x}_k - \mathbf{x}_{k|k-1}^i). \end{aligned} \tag{3.5}$$

In practice, the prediction Eq. 3.5 does not need to be evaluated but most importantly the new state vectors are calculated by the forward modeling Eq. 3.1a for each particle in the ensemble $\mathbf{x}_{k|k-1}^i$. In the sequel, the posterior PDF is approximated from the filtering Eq. 3.2 as observation Eq. 3.1b and measurement data are taken into account, i.e.

$$\begin{aligned} \rho(\mathbf{x}_k|\mathbf{D}_k) &\approx \frac{1}{\sum_{j=1}^{Ns} \rho(\mathbf{d}_k|\mathbf{x}_{k|k-1}^j) w_{k-1}^j} \\ &\quad \sum_{i=1}^{Ns} \rho(\mathbf{d}_k|\mathbf{x}_{k|k-1}^i) w_{k-1}^i \delta(\mathbf{x}_k - \mathbf{x}_{k|k-1}^i) \\ &= \sum_{i=1}^{Ns} \tilde{w}_k^i w_{k-1}^i \delta(\mathbf{x}_k - \mathbf{x}_{k|k-1}^i), \end{aligned} \tag{3.6}$$

where

$$\tilde{w}_k^i = \frac{\rho(\mathbf{d}_k|\mathbf{x}_{k|k-1}^i)}{\sum_{j=1}^{Ns} \rho(\mathbf{d}_k|\mathbf{x}_{k|k-1}^j) w_{k-1}^j}. \tag{3.7}$$

The likelihood $\rho(\mathbf{d}_k|\mathbf{x}_k)$is evaluated for each particle as follows.

$$\begin{aligned} \rho(\mathbf{d}_k|\mathbf{x}_{k|k-1}^i) &= \frac{1}{((2\pi)^r |\mathbf{R}_k|)^{1/2}} \\ &* \exp\{-0.5(\mathbf{d}_k^{obs} - \mathbf{h}(\mathbf{x}_{k|k-1}^i)^t \mathbf{R}_k^{-1} (\mathbf{d}_k^{obs} - \mathbf{h}(\mathbf{x}_{k|k-1}^i))\}, \end{aligned} \tag{3.8}$$

where r is the number of observables. From Eq. 3.6, it can be observed that each particle's weight is updated as a product of $\tilde{w}_k^i$ and the weight in the previous time step w_{k-1}^i, i.e.

$$w_k^i = \tilde{w}_k^i w_{k-1}^i. \tag{3.9}$$

I employ the importance sampling approach without resampling (sequential importance sampling), and the proposal PDF is chosen to be the same as the PDF of the state transition $\rho(\mathbf{x}_k|\mathbf{x}_{k-1})$ for practical implementation of the particle filter. At initialization, the particles are sampled uniformly using the Latin hypercube sampling strategy (Forrester et al. (2008)) in the feasible ranges of parameter values. All initial particles have the same values of weight. The whole DA process employing the particle filter for parameter identification of geomechanical problems is illustrated in Fig. 3.1.

3.2 Identification of elasto-plastic soil parameters for tunnel model

3.2.1 Tunnel model and tunnel-induced settlements

The FE model was provided by the C2-subproject within the Collaborative Research Center SFB 837 at the Ruhr-University Bochum. In this model, a shield tunnel boring machine (TBM) advancing in homogeneous soil is modeled using the FE analysis software PLAXIS 3D 2011. The model dimensions are 60 m long, 40 m wide, and 45 m deep. The tunnel is of diameter D = 8.5 m and under an overburden depth equal to 1D. The overall model is depicted in Fig. 3.2. The homogeneous soil behaves according to the Hardening Soil constitutive model (Schanz et al. (1999)). The slurry shield TBM is 9 m long modeled by circular elastic plate elements. Each concrete lining ring has the length of 1.5 m and is also modeled by elastic plate elements. For detailed descriptions regarding modeling of support pressure, soil excavation, grouting, erection of tunnel linings, to shield skin-ground and tunnel lining-ground contacts used in the FE model the reader is referred to Zarev et al. (2011). Assuming that properties of the elastic materials are known, two important soil parameters are chosen for identification: shear modulus for un-/reloading G_{ur} and friction angle ϕ. Four observation positions and their orientations at observation surface which is located at 39 m from the initial position of the tunnel head (before excavation takes place) and perpendicular to the tunnel axis are depicted in Fig. 3.2.

The tunneling-induced displacements are calculated for 30 excavation steps, and each step advances the TBM 1.5 m forward. Note, data assimilation is carried out sequentially for

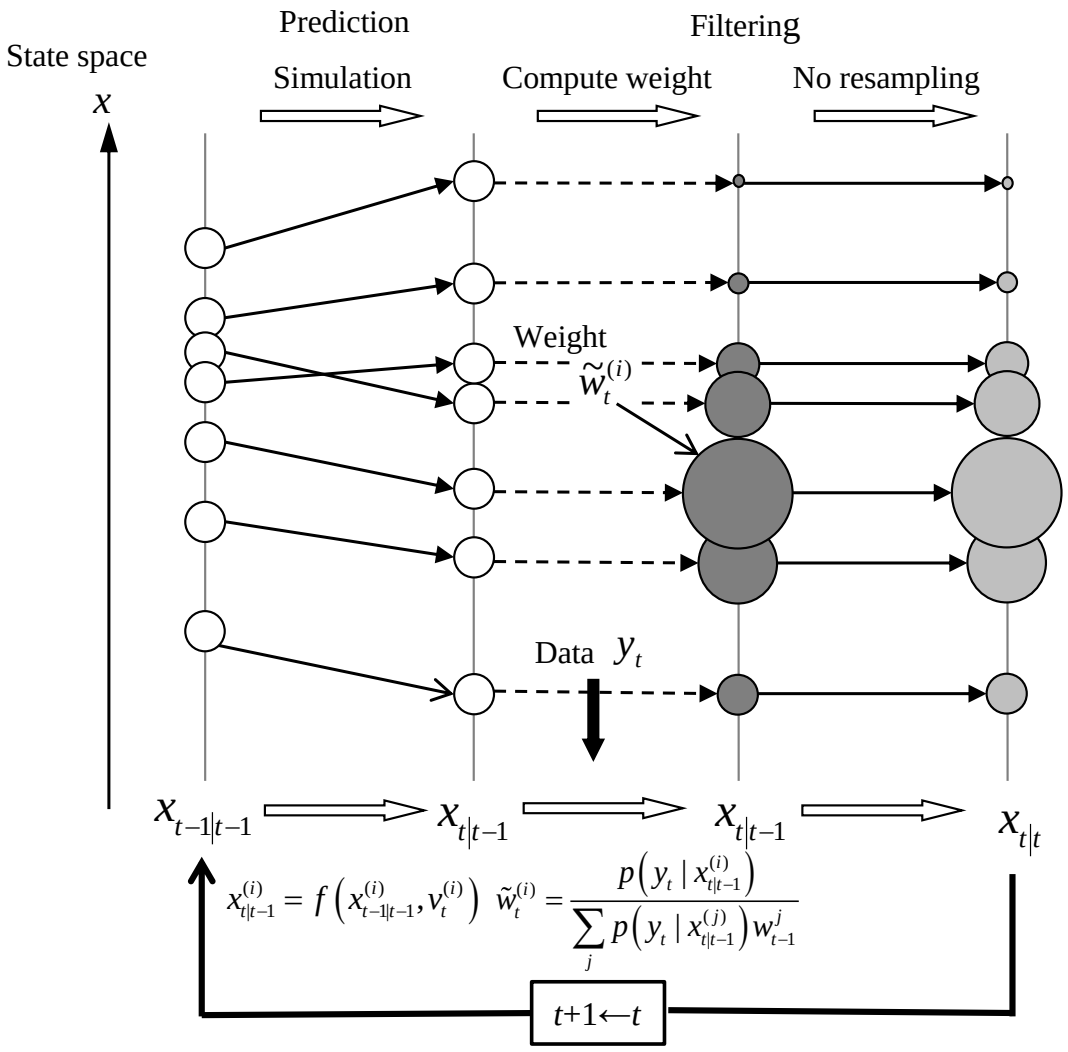


Figure 3.1: Particle filter using the sequential importance sampling (Murakami et al. (2013))

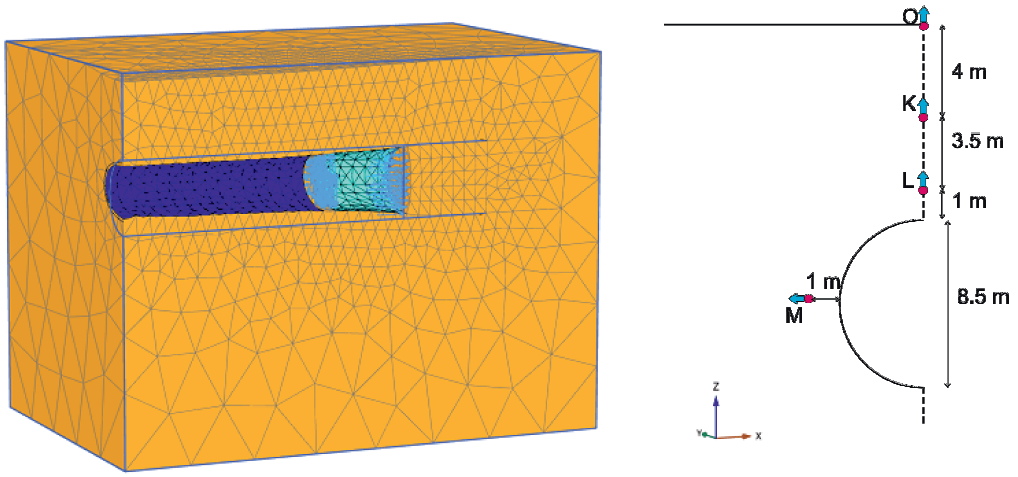


Figure 3.2: FE model of the 3D tunnel excavation and observables at observation cross-section

each tunnel excavation step; so, the total number of filtering steps here equals 30. In addition, the filtering performance is examined by using the noise-contaminated synthetic observation data. To produce the noisy synthetic observation data, the FE model calculated displacements are added by an amount of Gaussian noise having zero mean and standard deviations equal to three percent of root mean square values of them for each observation point.

3.2.2 Results and discussion

I use an ensemble of 150 particles uniformly distributed on the 2-dimensional parameter space. For efficient sampling in the parameter space, a Latin hypercube sampling is used to sample in the feasible ranges: $20000 \leq G_{ur}[kN/m^2] \leq 50000$, and $20 \leq \phi[\circ] \leq 40$.

Covariance matrix of observation uncertainty is set the same value for each observation point, i.e. $\mathbf{R} = diag(1.2, 1.2, 1.2, 1.2)$.

From uniformly distributed weights before DA, the development of distributions of weights, which represent the PDFs of model parameters, are representatively plotted at DA steps 10, 20, and 30 in Fig. 3.3. As can be observed, the PDFs become less scattered as more measurement data, i.e. more tunnel excavation steps, are assimilated. At the end of assimilation analysis, excavation step 30, each PDF is bell-shaped with a prominent peak and a narrow variance which means very certain estimates of model parameters.

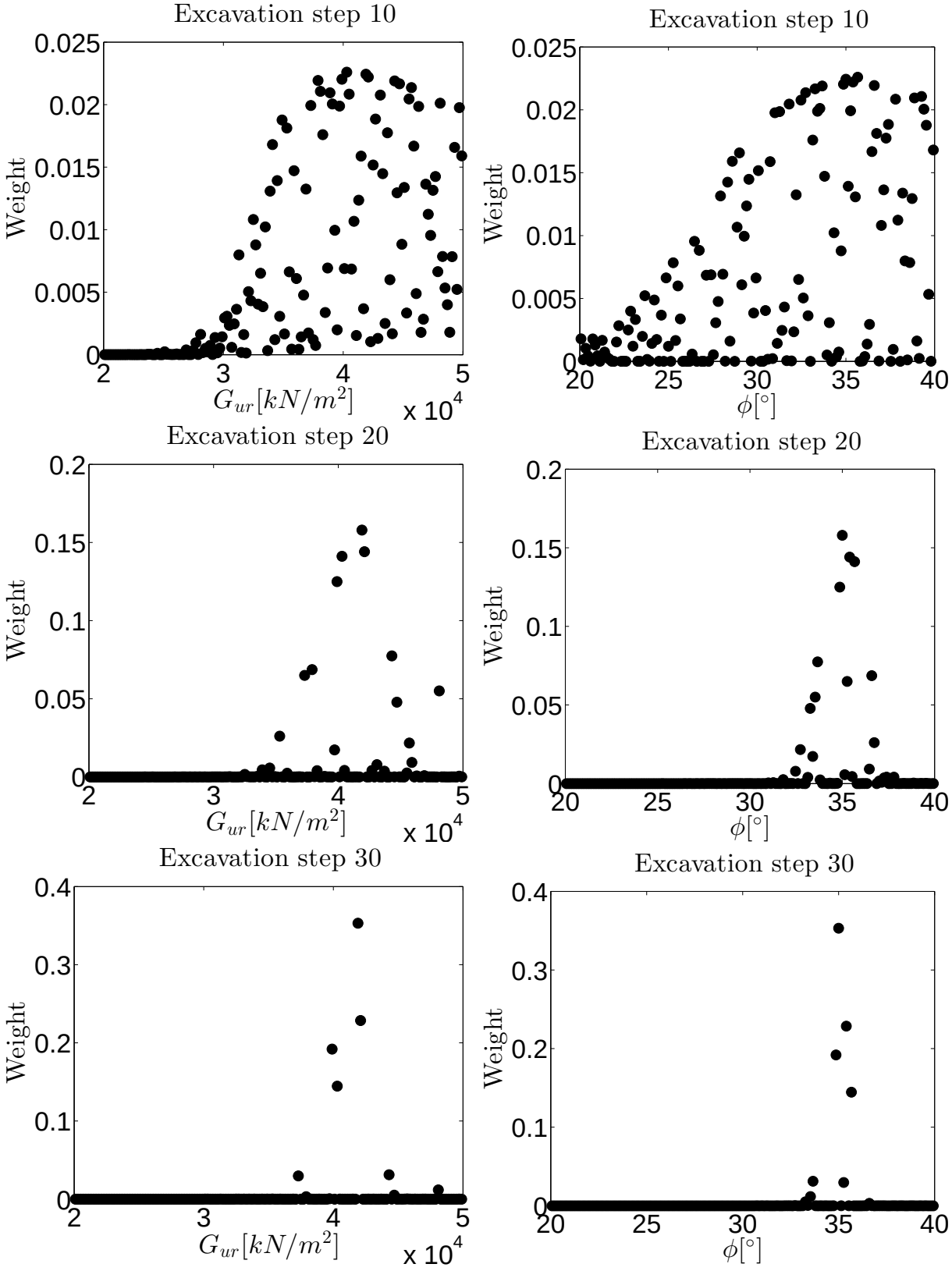


Figure 3.3: Development of model parameters' PDFs along DA process

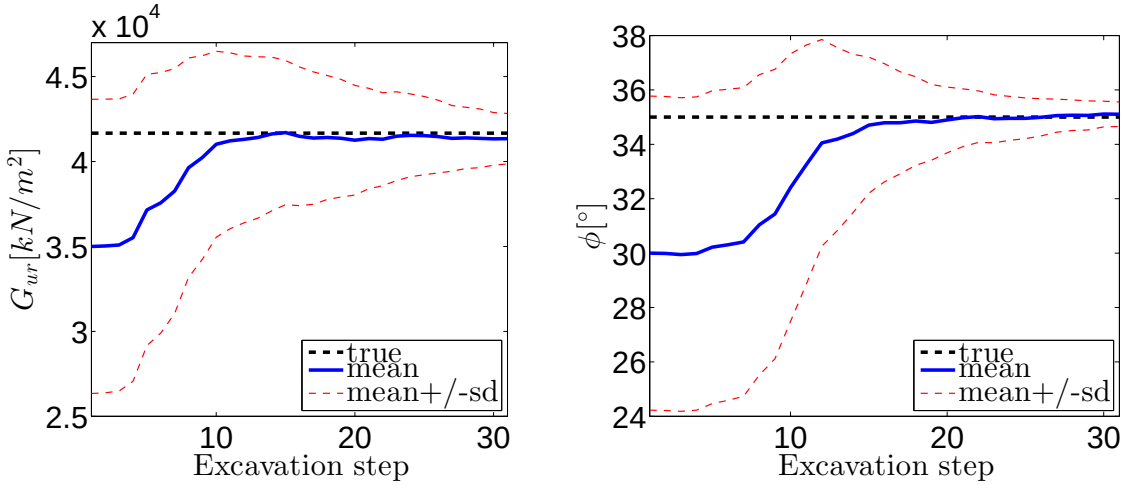


Figure 3.4: Parameter means and SDs along DA process

Identification results in form of PDFs are very advantageous for design engineers to pick up highly probable parameter values for performing further design analyses. Besides, when the resulted PDFs are bell-shaped, means and standard deviations are readily calculated from the ensemble weights. In this study, for example, development of means and standard deviations of the PDFs are shown in Fig. 3.4. It can be observed that as mean values become closer to the true model parameters, the corresponding standard deviations of the distributions also become narrower which reflect converging estimates of the soil parameters.

For the nonlinear inverse problem as in this work, it is not evident to predefine the amount of observed information that is required to ascertain convergence. With DA scheme, the accumulation of information is utilized sequentially, and convergence is recognized when no substantial changes of the mean values occur for several consecutive DA steps. In our application, it can be expected that the more observation points are used, the faster the convergence DA process can result. With the current observation configuration, less than 20 DA steps are needed to attain convergence as can be seen in Fig. 3.4.

It is worth noting that feasible ranges for generating the initial PDFs must include the expected parameter values and have an amount of appropriate margins to allow scattered distribution of the particles. The feasible ranges should be chosen not so wide to avoid using a large number of particles which makes DA computationally expensive.

I have made an effort to include less sensitive parameters such as the stiffness modulus in standard drained triaxial test E_{50} and cohesion c but the results were very scattered distributions of the particles at the end of DA process. One reason could be that with

such limited observable points, the data obtained is not informative enough to accurately identify so many model parameters.

4 A global optimization approach for inversion of the tunnel seismic waves

4.1 Concept

Solving for model parameters of a model described by partial differential equations given a set of noisy measurement data such as the model for wave propagation in the subsurface is a challenging nonlinear inverse problem. A standard approach to solving this problem is to minimize a measure of misfit between model outputs and measurement data by employing an optimization method. To solve inversion problems in geotechnical and geophysical engineering, local optimization methods such as quasi-Newton and conjugate-gradient are most often used to achieve feasible calculation time of the inverse problem because forward modeling of the wave propagation is computationally costly and the number of model parameters in practical applications is formidably large. The recipe for success of local optimization is that only a few number of optimization iterations are required, and the gradients can be approximated by mathematical tools such as the adjoint state method (Tromp et al. (2005)) or the reciprocity principle (Sheen et al. (2006)). The very significant limitation of the local optimization method for waveform inversion is that the initial model must be very close to the sought-for model to avoid the solution being trapped in one of the local minima on the misfit topography.

With regard to the application for predicting geological structure ahead of the underground tunnel face, in which there are a limited number of major fault scenarios that are commonly envisaged such as inclined fault zones and large blocks of erratic rocks, the number of model parameters can be substantially reduced using appropriate parameterization of the model regarding the concerned subsurface situation. The limited number of model parameter gives rise to a feasibility study of the application of global optimization methods for predicting ahead of the tunnel face.

The cost function to minimize in this study is targeted to be the full waveform misfit, which measures differences between the recorded and the modeled information encrypted in the respective waveforms, because of the two main reasons. Firstly, the full waveforms (body waves and surface waves) contain the most information about the hidden geological structure that can be obtained from the reflected seismic waves. Therefore, inversion of the full waveforms promises high-resolution imaging of the subsurface. In contrast, travel time tomography (Krauß et al. (2014)) and seismic migration (Dickmann & Sander (1996); Kneib et al. (2000)) rely only on arrival times and/or amplitudes of the reflected waves for imaging the geological structure. The advantage of having rich information is beneficial to the underground tunnel settings because the underground working space is strictly constrained for placing the receivers, usually on the tunnel side wall or the tunnel face, which limits the amount of valuable recorded data. Secondly, waveform inversion is not yet fully developed, for example, not only one answer is accepted to the question what form of misfit functional better represents the mismatch and also facilitates global convergence for the inversion or the question what inversion method guarantees global convergence while keeps the computation time acceptable in engineering applications. Besides, several general questions are yet to be answered such as what method of model parameterization can help accelerate convergence, how prior knowledge about the geologic model should be taken into account, and what numerical inversion methods yield high accuracy while remaining computationally feasible despite the fact that measurements are contaminated by noise and that the initial model is far from the true model.

It is known that the waveform misfit is often very multimodal. Therefore, minimizing the misfit functional by global optimization methods is a preferable methodology to be applied in near surface engineering including tunnel reconnaissance based on the reflection/ refraction of elastic waves propagation in the subsurface. However, one problem for the deployment of global optimization methods is the excessive number of forward simulations of the computationally expensive elastic/ viscoelastic wave propagation that is required. This difficulty can be overcome by either improving the calculation speed of the forward model through the use of advanced computational methods and high power computing facility or by improving the convergence rate of a concerned global optimization method. In this work, the authors propose a new hybridized global optimization method that is based on heuristic global search and local fine-tuning to invert the tunnel seismic waveforms for characterizing geological structure ahead of the tunnel face. A highly accurate and computationally efficient forward modeling of tunnel seismic waves propagation is achieved by utilizing a robust implementation of an open-source spectral finite element code.

Simulated annealing (SA) is a metaheuristic search algorithm to find near-optimal solutions of an energy function. A probabilistic perturbation and acceptance of the candidate sample with probability give SA the chance to escape from local minima. This perturbation and acceptance/ rejection steps are run for many cycles under control of a gradually decreasing temperature defined by a cooling schedule. The SA algorithm imitates annealing process for metals and glass that consists of heating the material to a high temperature and then cooling the material gradually to allow for rearrangement of the material structures to happen. The best crystalline configuration of the material, which has minimum energy, is expected to be achieved at the end of the annealing process.

Although SA possesses simplicity and physically sound principle from statistical mechanics, this method is not preferred so much as other metaheuristic optimization methods such as particle swarm optimization, genetic algorithm, or other evolutionary strategies to be used for solving inverse problems in engineering and sciences where the forward modeling is computationally expensive. The reason for that is because the SA relies only on rejection/ acceptance of a single candidate at each run cycle which results in running excessively many annealing cycles until the global minimum is found. The global minimum can easily be missed if the annealing temperature decreases too fast or the probability distribution to generate sample candidates is set inappropriately.

Efforts to hybridize SA algorithm with other global and local optimization methods have been made elsewhere in order to improve convergence or accuracy for specific applications. Martin & Otto (1996) proposed combining SA with local search heuristics to solve combinatorial optimization problems in that a search heuristic is employed to improve each random candidate by SA to a lower energy position before acceptance/ rejection decision is made. Jeong & Lee (1996) combined genetic algorithm with SA to create an adaptive simulated annealing genetic algorithm that can fix poor hill-climbing capability of the genetic algorithm for applications in system identification. An effort is made in this work to improve the performance of the SA algorithm based on the unscented Kalman filter for use in tunnel seismic inversion. It will be shown in this work that execution time and accuracy of SA algorithm can be improved by incorporating the unscented Kalman filter (UKF) as an effective local minimization method into it. The UKF, which belongs to the sigma-point Kalman filter family, is a derivative-free estimator for nonlinear state-space models that can be efficiently adapted for solving parameter identification of static and dynamic models (Van Der Merwe (2004)). The new unscented hybrid simulated annealing (UHSA) algorithm is robust because it combines the exploration power of SA with the exploitation efficiency of UKF. Herein, the authors imply exploration as the capability

to randomly sample the full parameter space and exploitation is the power to navigate regionally to a lower point in the local area with the help of information provided by the data misfit landscape.

The working concept of the new hybridized optimization is illustrated in an exemplary multimodal misfit landscape in 1-dimensional space, Fig. 4.1. At any annealing cycle, if a random move (1) is accepted by the SA algorithm, this point is then improved locally by the UKF procedure to attain to the near local minimum point (1*). This improved point is biased from the centered local minimum because the misfit value evaluated in this local minima region is larger than the minimum of misfit and also because the UKF runs only for some iterations before SA proposes the next random move. SA proposes the next random move (2) which can be accepted for sure or with probability according to a chosen Markov chain Monte Carlo algorithm that will be discussed in the following section. If this accepted move is in the global minima region, the UKF will exploit the region and the global minimum (2*) is found with high accuracy. If the global optimal point is found, acceptance rates in later annealing cycles become very small owing to the Metropolis sampling (Metropolis et al. (1953)) that is used in the SA algorithm. The UHSA algorithm bases on early convergence to the minima to achieve fast convergence to the global minimum. The early convergence to the local minima is achieved thanks to the UKF local improvements of the random moves and the Metropolis's acceptance/ rejection mechanism that will be explained in the next section. Thus, applying the UHSA algorithm to invert for the complicated model does not require very long annealing cycles compared with the standard SA.

The advantages of the new algorithm are:

- Simplicity: The UHSA can be implemented quite easily by a modern programming language because the standard procedure for SA algorithm is well known. The "plug-in" of the UKF in the SA is straightforward and requires very minimal implementation effort as the UKF consists of only a few matrix operations.
- Speed and accuracy: The intrinsic slow convergence rate of SA to solve global nonlinear optimization problems is overcome thanks to effective local search supported by the UKF. Furthermore, rather than collecting random candidates in the final estimation results, the UHSA seeks to improve them locally, which results in a much higher accuracy of the final estimated quantities.
- Guarantee: The statistical assurance of SA algorithm for finding the global minimum is maintained for the UHSA.

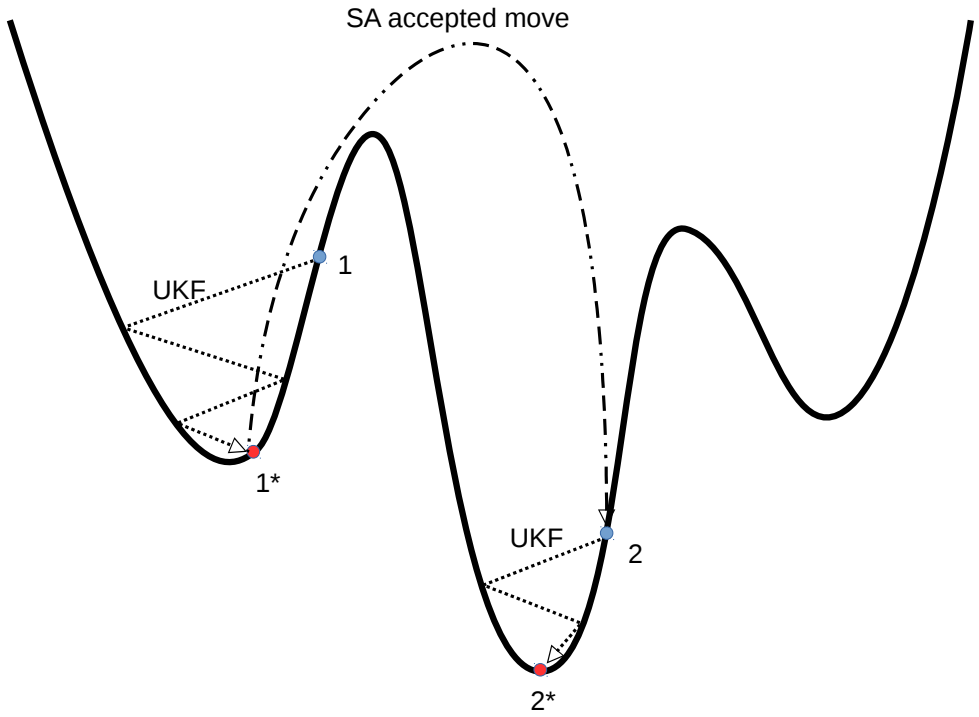


Figure 4.1: Global-local optimization concept of the new hybridized algorithm

- Derivativelessness: UHSA is a nonlinear derivative-free inverse method which is suitable for non-convex, non-smooth misfit topographies.

Compared with the standard SA algorithm, the UHSA has additional algorithmic parameters belonging to the UKF to be tuned. However, most of them can be set utilizing prior knowledge about the model parameters and accuracy of the measured data. Attention should be paid to the designation of the sigma-points for the unscented transformation.

4.2 Unscented hybrid simulated annealing

The forward model is described by a nonlinear black-box function

$$\boldsymbol{d} = h\left(\boldsymbol{m}\right) + \boldsymbol{v}, \tag{4.1}$$

where $\boldsymbol{m} \in \mathbb{R}^n$ denotes n parameters of the parametrized model, $\boldsymbol{d} \in \mathbb{R}^r$ stores the model outputs at r observation points resulting from the nonlinear model $h : \mathbb{R}^n \mapsto \mathbb{R}^r$. For time-dependent problem such as wave propagation, the model $h(\cdot)$ is solved in time steps for the simulation period; and the resulting batch data are stored in $\boldsymbol{d}$. $\boldsymbol{v} \in \mathbb{R}^r$ represents modeling error caused by assumptions made in building the mathematical model and

by the numerical approximations. This error can be assumed to be zero-mean Guassian distribution with covariance $\boldsymbol{R}^m$.

The misfit $s : \mathbb{R}^n \mapsto \mathbb{R}$ to be minimized is constructed for each receiver j by squares of differences between the measured data $d^{obs}(t)$ and the modeled outputs $d(t)$ in the time window of interest $[0, \tau]$:

$$s_j = \frac{1}{2} \int_0^\tau \|d_j^{obs}(t) - d_j(\boldsymbol{m}, t)\|^2 dt, j \in \{1, 2, .., r\}. \tag{4.2}$$

Before describing the new derivative-free global optimization algorithm, an adaptation of the UKF procedure for local exploitation in the parameter space is developed next.

4.2.1 Unscented Kalman filter for local minimization

UKF is a very successful state and parameter estimator for nonlinear dynamic models in which the posterior estimates are inferred from noisy observation data. It aims to overcome the inherent limitations of the traditional extended Kalman filter that works only for slightly nonlinear models and requires linearization of the nonlinear models. The linearization of highly nonlinear models is often numerically non-trivial and inaccurate. Owing to its high accuracy and simplicity, the UKF has attracted great attention and has been applied to state and parameter estimation problems in fields ranging from training neural networks in machine learning (Van Der Merwe (2004); Haykin et al. (2001); de Oliveira (2012)) to monitoring electrical parameters of a battery (Plett (2006)), identifying model parameters and hidden state variables for nonlinear structural dynamics (Mariani & Ghisi (2007); Chatzi & Smyth (2009)), state and parameter estimation of nonlinear biological dynamics (Attarian et al. (2013)), and calibration of soil parameters for geotechnical modeling in mechanized tunneling (Nguyen & Nestorović (2015*a*)).

As for the Kalman filter family, the quantities being estimated in the form of a joint Gaussian distribution are inferred from noisy measurement data by updating the numerical model recursively. For linear models, the linear Kalman filter (Kalman (1960)) is the optimal estimator. For nonlinear models, a kind of linearization is made to solve the suboptimal estimation problem. The linearized Kalman filter results from applying the Taylor series expansion to linearize the model (Jazwinski (1970)). Mapping of the Gaussian mean and covariance through the linearized model suffers badly from linearization inaccuracy. A more accurate way to keep high accuracy of the linearized mapping is to

employ the unscented transformation (Julier & Uhlmann (2004)). The unscented transformation is a statistical linearization method to approximate a multivariate Gaussian distribution by a set of sampled points. Unlike Monte Carlo sampling, the number of samples required by the unscented transformation is minimal, and the sample positions are set deterministically about the current estimate and its covariance. The samples used in the unscented transformation are the so-called sigma-points.

The predicted mean and covariance before measurement update are denoted as $\hat{\boldsymbol{m}}$ and $\boldsymbol{P}^m$ respectively. If n model parameters are to be estimated, $2n+1$ sigma-points are defined as follows for the unscented transformation to approximate a Gaussian distribution of the predicted estimate centered at $\hat{\boldsymbol{m}}$ with covariance $\boldsymbol{P}^m$. The $2n+1$ sigma-point vectors $\boldsymbol{\mathcal{M}}_i$ are assigned as:

$$\boldsymbol{\mathcal{M}}_0 = \hat{\boldsymbol{m}}, \tag{4.3}$$

$$\boldsymbol{\mathcal{M}}_i = \hat{\boldsymbol{m}} + \left(\sqrt{(n+\lambda)\boldsymbol{P}^m}\right)_i, \text{ for } i = 1:n \tag{4.4}$$

$$\boldsymbol{\mathcal{M}}_{n+i} = \hat{\boldsymbol{m}} - \left(\sqrt{(n+\lambda)\boldsymbol{P}^m}\right)_i, \text{ for } i = 1:n. \tag{4.5}$$

The parameter λ remains free to tune so as to adjust the spread of the sigma-points about the predicted mean estimate in directions of principle variances. It is suggested to select $n+\lambda = 3$ if $\boldsymbol{m}$ is Gaussian (Julier & Uhlmann (1997)). The notation $(\cdot)_i$ denotes the i-column of the matrix within parentheses.

Each sigma-point is associated with a weight. The weights are defined as in Eq. 4.6 such that the summation of weights is unity:

$$W_0 = \frac{\lambda}{n+\lambda}, W_i = W_{i+n} = \frac{1}{2(n+\lambda)}, \text{ for } i = 1:n. \tag{4.6}$$

Misfit measures for all r receivers are obtained by mapping the sigma-points through the misfit functional Eq. 4.2. The mapping for r multiple receivers $\boldsymbol{s} = (s_1, s_2, \ldots, s_r)$, $\boldsymbol{s} : \mathbb{R}^n \mapsto \mathbb{R}^r$, is done for each sigma-point to create the corresponding misfit vectors $\boldsymbol{\mathcal{S}}_i$ as written in Eq. 4.7.

$$\boldsymbol{\mathcal{S}}_i = \boldsymbol{s}\left(\boldsymbol{\mathcal{M}}_i\right), \text{ for } i = 0:2n. \tag{4.7}$$

This direct nonlinear mapping of the sigma-points helps to preserve second order accuracy of the Gaussian distribution as opposed to the first-order truncation of the Taylor series employed by the linearized Kalman filter.

A fictitious stationary state transition equation perturbed by a minimal amount of zero-mean Gaussian noise with covariance $\boldsymbol{Q}$ emulating iterative time updating of the parameter vector being estimated completes the state-space model formulation that allows

straightforward application of the Kalman filtering scheme. After the nonlinear mapping of the sigma-points through the nonlinear model, the mean of the predicted misfit $\hat{\boldsymbol{s}}$, the covariance of the predicted parameters $\boldsymbol{P}^m$, the covariance of the misfit at the predicted model $\boldsymbol{P}^s$, and cross-covariance between them $\boldsymbol{P}^{ms}$ are approximated by weighted summation rule as follows:

$$\hat{\boldsymbol{s}} = \sum_{i=0}^{2n} W_i \boldsymbol{\mathcal{S}}_i, \tag{4.8}$$

$$\boldsymbol{P}^m = \sum_{i=0}^{2n} W_i \left(\boldsymbol{\mathcal{M}}_i - \hat{\boldsymbol{m}}\right)\left(\boldsymbol{\mathcal{M}}_i - \hat{\boldsymbol{m}}\right)^t + \boldsymbol{Q}, \tag{4.9}$$

$$\boldsymbol{P}^s = \sum_{i=0}^{2n} W_i \left(\boldsymbol{\mathcal{S}}_i - \hat{\boldsymbol{s}}\right)\left(\boldsymbol{\mathcal{S}}_i - \hat{\boldsymbol{s}}\right)^t + \boldsymbol{R}, \tag{4.10}$$

$$\boldsymbol{P}^{ms} = \sum_{i=0}^{2n} W_i \left(\boldsymbol{\mathcal{M}}_i - \hat{\boldsymbol{m}}\right)\left(\boldsymbol{\mathcal{S}}_i - \hat{\boldsymbol{s}}\right)^t. \tag{4.11}$$

A set of measurement data contaminated by measurement noise can practically be described by Gaussian distribution $\rho_D(\boldsymbol{d}) \sim \mathcal{N}(\boldsymbol{d}^{obs}, \boldsymbol{R}^{obs})$. Then, covariance matrix $\boldsymbol{R}$ in Eq. 4.10 is obtained by an addition of the covariance for measurement uncertainty and the covariance for modeling uncertainty, i.e. $\boldsymbol{R} = \boldsymbol{R}^{obs} + \boldsymbol{R}^m$ (Tarantola (2005)). Posterior mean $\hat{\boldsymbol{m}}_+$ and covariance $\boldsymbol{P}^m_+$ of the estimated model parameters are updated following the Kalman filter's measurement update step:

$$\hat{\boldsymbol{m}}_+ = \hat{\boldsymbol{m}} + \boldsymbol{K}\left(\boldsymbol{0} - \hat{\boldsymbol{s}}\right) \tag{4.12}$$

$$\boldsymbol{P}^m_+ = \boldsymbol{P}^m - \boldsymbol{K}\boldsymbol{P}^s\boldsymbol{K}^t \tag{4.13}$$

with the Kalman gain $\boldsymbol{K}$ calculated as

$$\boldsymbol{K} = \boldsymbol{P}^{ms}\left(\boldsymbol{P}^s\right)^{-1}. \tag{4.14}$$

The zero vector term in the innovation term $\boldsymbol{0} - \hat{\boldsymbol{s}}$ in the measurement update Eq. 4.12 results from the fact that the misfit for each receiver (described in Eq. 4.2) is expected to be zero. Each element in the innovation term represents the least square difference between the expected and the current estimated model.

The authors suggest iterative runs instead of a single run of the UKF procedure Eqs. 4.3-4.13 to benefit from the superior convergence of the Kalman filter that is similar to that of the Gauss-Newton method for least squares optimization (Bell & Cathey (1993); Bertsekas (1996)).

To initialize the UKF procedure, some prior knowledge about the model is provided in the form of Gaussian distribution $\rho_0(\boldsymbol{m}) \sim \mathcal{N}(\boldsymbol{m}_0, \boldsymbol{P}_0)$. In the first iteration of the UKF, this prior information is assigned to the predicted mean and covariance: $\hat{\boldsymbol{m}} = \boldsymbol{m}_0$, and $\boldsymbol{P}^m = \boldsymbol{P}_0^m$. In the successive iterations, the predicted estimated mean and covariance are updated by taking the respective values from the posterior estimate, i.e. $\hat{\boldsymbol{m}} = \boldsymbol{m}_+$ and $\boldsymbol{P}^m = \boldsymbol{P}_+^m$.

This completes the UKF procedure for minimizing the misfit functional locally. The combined algorithm for fast global optimization is described next by plugging the UKF procedure in the SA algorithm.

4.2.2 Unscented hybrid simulated annealing

The introduction of simulation of annealing process in statistical mechanics for solving computer optimization problems by the work of Kirkpatrick et al. (1983) has been very fruitful. Applications of the standard SA in seismic waveform inversion can be found in Sen & Stoffa (1991); Sambridge & Mosegaard (2002); Tran & Hiltunen (2011).

On the core of the SA algorithm is the Metropolis algorithm which is designed to sample a probability distribution in high dimensional space whose mathematical description is unknown and only its function values can be evaluated. The Metropolis procedure is based on the energy change ΔE of the system resulting from a small random perturbation to decide if the candidate is accepted or not. Probability of acceptance follows the Boltzmann distribution $P(\Delta E) = \exp\left(-\frac{\Delta E}{K_B T}\right)$, where K_B is the Boltzmann constant and T is the current temperature of the system. If $\Delta E \leq 0$, the move is always accepted, otherwise the move is accepted with the probability $P(\Delta E)$.

For solving inverse problems, the misfit functional is used in place of the energy function. The misfit functional is defined for r receivers as the summation of misfits calculated from each receiver (Eq. 4.2), i.e. $S : \mathbb{R}^n \mapsto \mathbb{R}$,

$$S = \sum_{j=1}^{r} s_j. \tag{4.15}$$

The misfit functional can take values of arbitrary magnitude depending on the type of measurement data and the concerned misfit definition. In order to make the acceptance probability independent of the misfit magnitude, I adapt normalized misfit differences as proposed in Balling (1991),

$$P = \exp\left(-\frac{\Delta S}{\Delta S_{avrg} T_c})\right), \tag{4.16}$$

where ΔS is a misfit difference between the proposed move compared to that from the previous accepted value, and T_c is the current temperature. The Boltzmann constant ΔS_{avrg} is set to be the averaged misfit differences of all accepted moves up to the current temperature.

It is guaranteed that the global convergence can be achieved if annealing temperature at cycle c decreases according to $T_{c+1} = \frac{T_0}{ln(c+1)}$. But this slow annealing schedule leads to too slow convergence for practical applications. The faster cooling schedule following the exponential relation

$$T_{c+1} = \alpha T_c, \tag{4.17}$$

where $\alpha \in (0, 1)$, is used instead (Blum & Roli (2003)).

Random moves simulating perturbation are generated according to a distribution function that is chosen based on proposal by Ingber (1993) under consideration of the upper bound U^i and the lower bound L^i of each dimension m^i, for $i = 1, .., n$.

$$m^i_{c+1} = m^i_c + y^i \left(U^i - L^i\right), \tag{4.18}$$

with $y^i \in [-1, 1]$ which is generated by the uniform distribution variable $u^i \in [0, 1]$ as follows:

$$y^i = \operatorname{sgn}\left(u^i - \frac{1}{2}\right) T_c \left[\left(1 + \frac{1}{T_c}\right)^{|2u^i - 1|} - 1\right]. \tag{4.19}$$

Eq. 4.18 is repeated until $m^i_{c+1} \in [L^i, U^i]$, for $i = 1, .., n$.

A limitation that hinders the use of the SA algorithm for solving inverse problems of time-consuming forward models is that global convergence of SA is not secured if the number of annealing cycles is not sufficient. On the contrary, fast annealing by decreasing the temperature rapidly (quenching) has a high risk of convergence at a local minimum.

A new scheme for accelerating convergence speed of SA is proposed in this work. The authors coupled the UKF described above with the standard SA simulation if the random move is accepted to improve locally the intermediate candidate to a better position in the parameter space. Improved accepted samples give rise for the annealing simulation to more quickly reject high energy random moves which are less important in the search for the global minimum. Moreover, The UKF local search increases the accuracy of the found model parameters because it utilizes information of the sample points (sigma-points) in the neighborhood of the sampled candidate to find the best possible low energy point in the vicinity. Therefore, the sensitivity of SA to annealing scheduling and proposal distribution is relaxed when SA runs in combination with the UKF local search. To

maximize the utilization of the local search, the SA should be able to sample in the entire search space, and its acceptance rate should be kept relatively large to increase the probability of escaping from local minima. However, the number of sampling points needs not to be as many as that of the standard SA because the UKF can improve the bad sample points. The new UHSA algorithm is summarized in Algorithm 3 for running N_c annealing cycles enhanced by N_k iterations of the UKF.

Algorithm 3: Unscented hybrid simulated annealing

```
// Initialization
Settings for the UKF: P_0, R, Q, N_k
Settings for the SA: T_0, α, N_c
Choose a starting point for annealing process m_c
c ← 1
// run N_c cycles of SA
while c < N_c do
    Random move to propose m_p following Eq. 4.18
    Calculate acceptance probability P: Eq. 4.16
    // Check if the proposed move is accepted
    if rand.uniform(0,1) < P then
        Set initial estimate for the UKF: m̂_0 = m_p
        // Improve the accepted move by N_k iterations of UKF
        for k = 1..N_k do
            // Run the UKF with function s(·) defined in Eq. 4.2
            m̂_k, P_k = UKF(s(·), m̂_{k-1}, P_{k-1}, R, Q)
        end
        Store the improved point m_c = m̂_k
    end
    Lowering temperature T_c: Eq. 4.17
    c ← c + 1
end
```

The balance between exploration and exploitation in the parameter space can be adjusted by choosing the number of UKF iterations. A fewer number of UKF iterations allows for more exploration capability of the SA process. In a limit case, setting zero UKF iteration converts the UHSA to the standard SA algorithm. Over-exploitation for each accepted random move allows the algorithm to attain the minimum of the misfit very early but leads to low acceptance probability for later annealed moves. When applied in solving inverse

problems involving complex numerical model, the computation time depends mostly on time to run the forward model. As described in Algorithm 3, in each annealing step, the UHSA requires one forward model run for calculating the acceptance probability and a fixed number of forward model runs for local exploitation by the UKF if the random move is accepted. The number of forward model runs called by the UKF is determined by (the number of sigma-points) $\times$ (the number of iterations), which is $(2n+1) \times N_k$.

When initializing the UKF, it is worth noting that setting of the initial covariance $\boldsymbol{P}_0$ needs not to be close to the true estimation error covariance because the UKF only aims for attaining the local minima in the vicinity of the randomly accepted candidate model. Practically, $\boldsymbol{P}_0$ is a diagonal matrix whose square roots of the elements are set inversely proportional to the degree of multimodality of the misfit landscape, but typically not exceeding one tenth of the range for each dimension. Elements in the covariance $\boldsymbol{Q}$ can be set so that their standard deviations should be around ten percent of the corresponding standard deviations in the $\boldsymbol{P}_0$ as suggested in Nguyen & Nestorović (2015*a*). Data uncertainty matrix $\boldsymbol{R}$ is assigned according to the noise level of the data recorded by the receivers.

4.3 Finding global minimum of a multimodal optimization test function

In order to test the efficiency of the UHSA algorithm, the new algorithm is employed to find the global minimum of the multimodal Schwefel's function. The n-dimensional Schwefel's function has the form

$$f(x_1, x_2, .., x_n) = 481.982887n + \sum_{j=1}^{n} \left(-x_i \sin\left(\sqrt{|x_i|} \right) \right). \tag{4.20}$$

Finding the global minimum of the Schwefel's function is challenging because there exist several local minima on the parameter space and the second minimal point is deceptively very distant from the global minimum. The global minimum is located at $\boldsymbol{x}_{min} = (420.968746, 420.968746, .., 420.968746)$, where $f(\boldsymbol{x}_{min}) = 0$. Parameter space of the Schwefel's function takes feasible range in $[-512, 512]$ for each parameter. Contour plot of the 2-dimensional Schwefel's function is plotted in the background of Fig. 4.4.

Five comparative SA and UHSA runs, each with the same random seed and algorithmic settings, are set to evaluate the performance of the UHSA algorithm compared with the

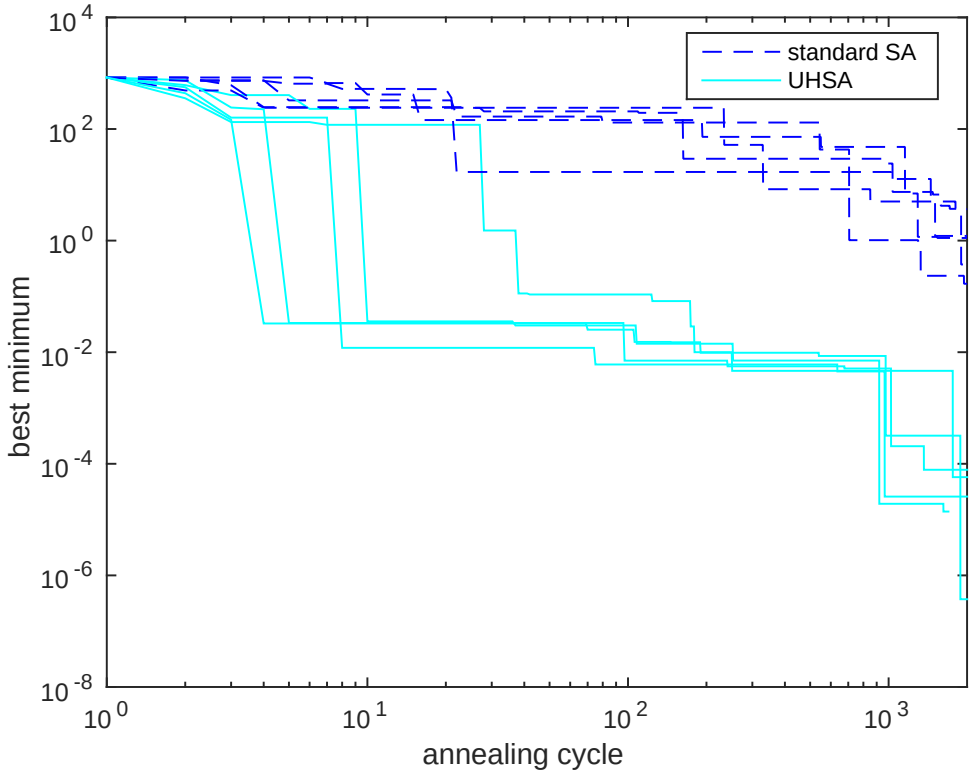


Figure 4.2: Comparison of best minimal values of Schwefel's function with siminar runs of standard SA and UHSA

standard SA. The standard SA is run using the same computer code but with the number of UKF iterations set to zero. The UHSA is run with 10 UKF iterations. The annealing processes are run for 2000 cycles. The fast cooling schedule Eq. 4.17 is set to begin with 6 degrees and to finish at 0.01 degree. The covariance matrices are set as follows for the UKF: $\boldsymbol{P}_0 = \text{diag}(4^2, 4^2)$, $\boldsymbol{Q} = \text{diag}(0.4^2, 0.4^2)$, $R = 0.1^2$.

Starting the algorithm with initial point $\boldsymbol{m}_0 = (x_1 = 0, x_2 = 0)$, convergence of various annealing runs plotted in Fig. 4.2 demonstrates that UHSA converges much earlier and more precise than the standard SA thanks to the ability to exploit the local neighborhood of the UKF procedure. Early acceptance of the candidates at the local minima results in lower acceptance rate for the whole annealing process. The 'no-free-lunch' rule holds in that a large number of function evaluations are needed in the early annealing cycles to have this early convergence advantage. However, the use of a large number of function evaluations in the early stage pays off for the later annealing cycles as the acceptance probability according to the Boltzmann distribution defined in Eq. 4.16 has been substantially reduced.

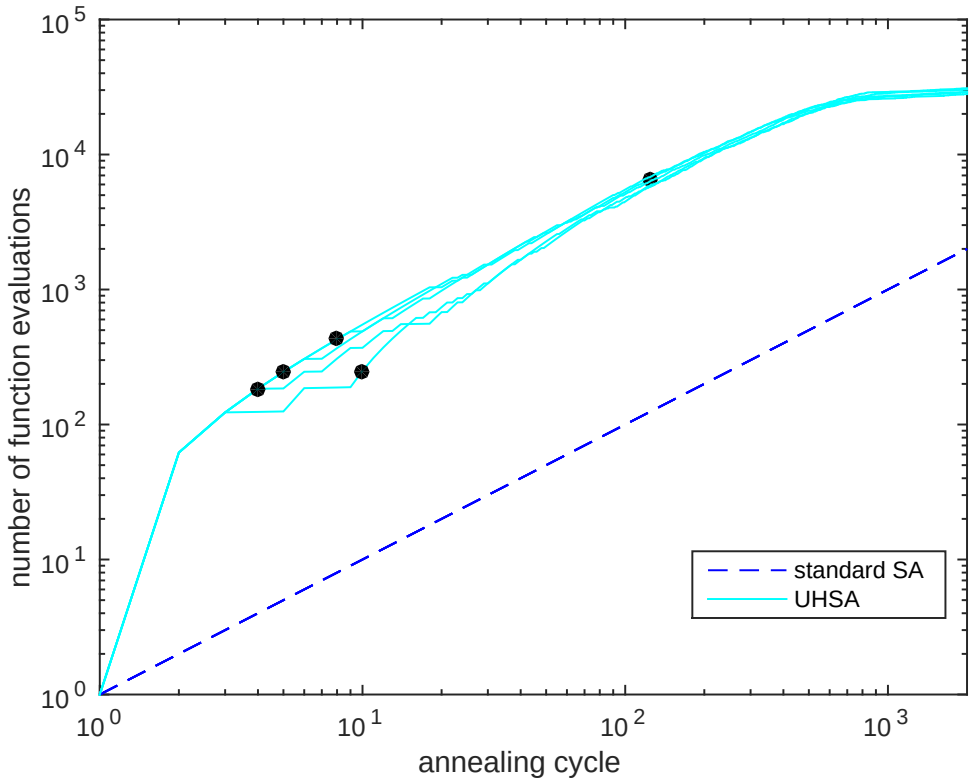


Figure 4.3: Comparison of number of function calls of standard SA and UHSA. Dark points indicate cycles at which convergence criterion is met. Note that the convergence criterion is not met by the standard SA

The number of evaluations of Schwefel's function during the annealing process is showed in Fig. 4.3. Overall, the UHSA requires a larger number of function calls than the standard SA along the annealing cycles. However, high accuracy convergence of the UHSA is achieved very early while the standard SA can not reach convergence criterion even until the end of the annealing process. The number of function calls until convergence for each annealing run is marked with a black dot in Fig. 4.3. Here, the convergence criterion is met when the value of Schwefel's function is less than 0.1. As can be seen in Fig. 4.3, the convergence criterion is not achieved by the standard SA within the set number of annealing cycles although the function is evaluated for 2000 times (one in each annealing cycle). Many more annealing cycles are needed for the standard SA to converge, but the number of function evaluations may explode greatly. Conversely, the UHSA needs in the worst case 6604 functions evaluations, and only 184 function evaluations in the best case.

For analysis of the optimization results, not only is it important that the global minimum is found, but the local minima also help provide a clear insight into the landscape of

the function in the exploratory parameter space. This information is very well presented by the distribution of all accepted moves. Representative distributions of the accepted candidates displayed in Fig. 4.4 show clearly that the SA assisted by UKF local search is more powerful than the standard SA in locating global as well as local minimal regions. The standard SA can find the near-global solution, but many of the accepted moves are scattered over unimportant regions. The UHSA, however, is able to locate very well the function's minimal regions. Interestingly, the UHSA results in a very high concentration of the candidates for the dominant minimal regions. And at the global minimum, many moves are reached precisely at the center of the minimal valley.

4.4 Tunnel seismic waveform inversion

4.4.1 Forward modeling of tunnel seismic waves

In principle, the current tunnel seismic prediction systems require either active source(s) (excited by sledgehammers or explosives) or passive seismic source (induced by vibrations of the tunnel boring machine) or both kinds of excitation as in Ashida (2001) and an array of receiving geophones placed in the tunnel wall and/or on the cutter head of the tunnel boring machine. Energy generated from a seismic source is propagated to the tunnel geomaterials (soils and rocks) in the form of body waves which consist of pressure waves (P-wave) and shear waves (S-wave). When these body waves encounter geologic heterogeneities (fault zones, erratic rocks, boulders, etc.), the partial energy of the incident waves is reflected/ refracted to the tunnel area and is recorded by the receivers installed there. It is noted that a portion of the body waves is converted to surface waves (Rayleigh waves) when the waves come into contact with the free surfaces such as the surface in the tunnel chamber area and the ground surface. The rest of this Subsection 4.4.1 describes the forward modeling of tunnel seismic waves using the spectral element method implemented in the open-source SPECFEM2D software (Komatitsch & Vilotte (1998)).

4.4.1.1 The spectral element method

SPECFEM2D by Komatitsch & Tromp (1999), a computer code that implements the spectral element method (SEM), is employed to simulate viscoelastic wave propagation. The peculiarity of SEM for solving waves propagation problem is that its solver is very

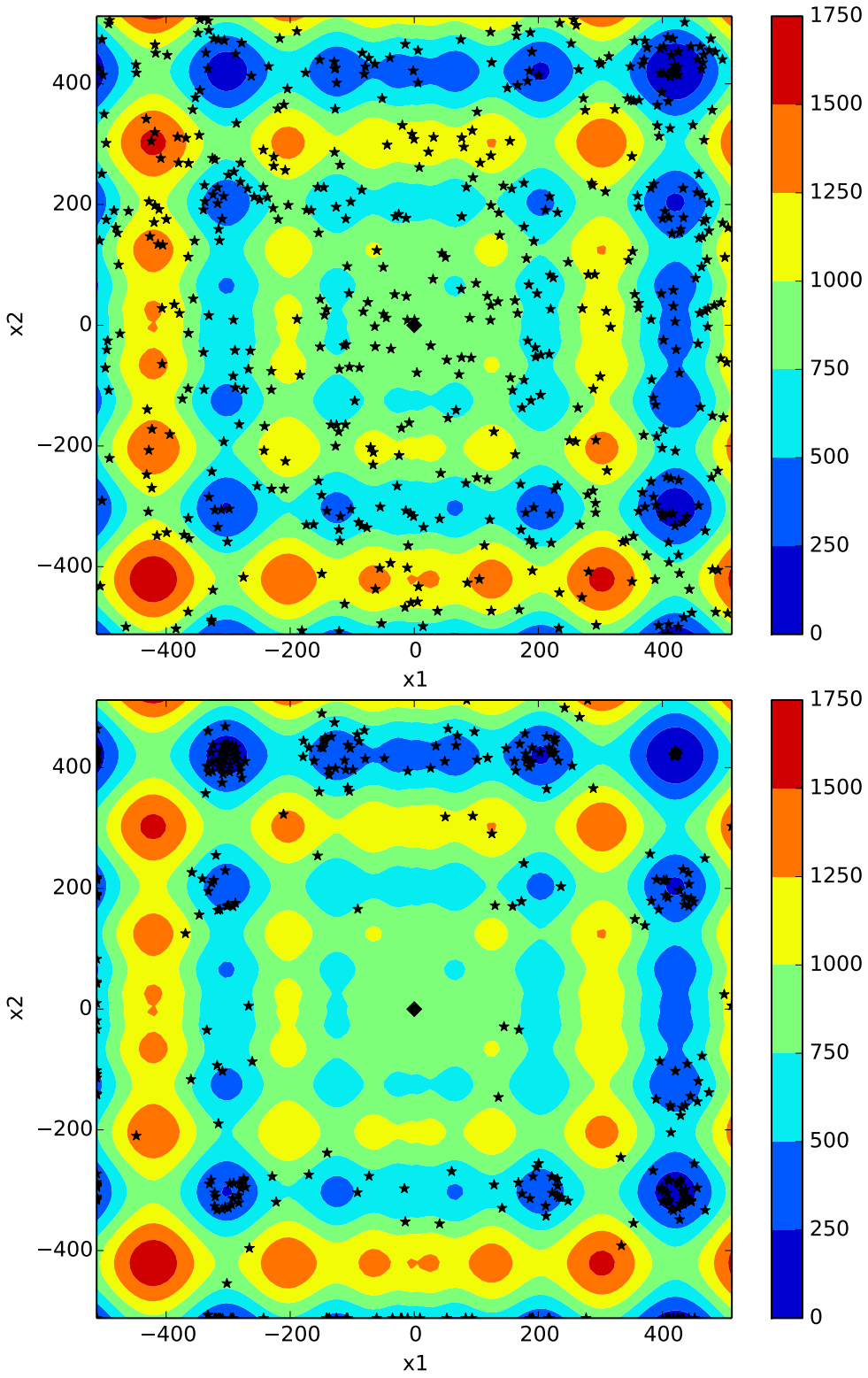


Figure 4.4: Representative final distributions of the accepted moves (black points) after 700 annealing cycles of standard SA (upper) and UHSA (lower) of the 2-dimensional Schwefel's function

time-saving thanks to effective high-order finite element interpolation technique that is used. In geotechnical engineering applications, SEM has been applied to engineering scale models such as elastic wave propagation as in Zheng et al. (2014) and elasto-plastic behaviour of geotechnical structures as in Gharti et al. (2012).

SEM solves the linear elastic wave equation, Eq. 4.21, for displacement vector **u**:

$$\rho \partial_t^2 \boldsymbol{u} = \nabla \cdot \boldsymbol{\sigma} + \boldsymbol{f}, \tag{4.21}$$

where ρ is density distribution in the body Ω, $\boldsymbol{f}$ denotes external excitation. The stress tensor $\boldsymbol{\sigma}$ is related to the small strain tensor $\boldsymbol{\varepsilon}$ by Hook's law for isotropic materials as written in Eq. 4.22:

$$\boldsymbol{\sigma} = 2\mu\boldsymbol{\varepsilon} + \lambda \mathrm{tr}(\boldsymbol{\varepsilon})\boldsymbol{\delta}. \tag{4.22}$$

The two constants λ and μ are the first and second Lamé constants. The other equivalent pairs of parameters such as Young's modulus E and Poisson's ratio ν, or pressure wave velocity v_p and shear wave velocity v_s can also be used in place of the Lamé constants depending on parameterization of the model. In case the material is not perfectly elastic, attenuation is taken into account such that the stress is determined by the entire strain history and the anelastic attenuation is characterized by the quality factor Q (Tromp et al. (2008)).

Similar to the standard finite element method, the weak form formulation derived from Eq. 4.21 on the modeled domain Ω is discretized into ne non-overlapping elements Ω_e, such that $\Omega = \cup_e^{ne} \Omega_e$. Since the surface integrals need to be calculated at the absorbing boundary Γ, a number of nb surface elements Γ_b, such that $\Gamma = \cup_b^{nb} \Gamma_b$, resulting from element discretization of the domain are taken into account. The SEM implemented in SPECFEM2D package is restricted to quadrilateral finite elements of which the reference element is defined on a unit square. Geometrical mapping between points in physical coordinates and points in the reference coordinates is adequately approximated by Lagrangian polynomials of degree 1 or 2.

Displacement field and its derivatives defined on the elements are, however, effectively approximated by higher-order Lagrangian interpolants defined over the Gauss-Lobatto-Legendre points. The Lagrangian polynomials of degree 4 to 10 are optimal to use in a SEM implementation for wave propagation problem because this choice provides a very good trade-off between accuracy and computation cost (Komatitsch & Tromp (1999); Tromp et al. (2008)). Combination of the higher-degree Lagrangian interpolants and the Gauss-Lobatto-Legendre integration rule leads to diagonal structure of the mass matrix

in the assembly system

$$\boldsymbol{M}\ddot{\boldsymbol{U}} + \boldsymbol{C}\dot{\boldsymbol{U}} + \boldsymbol{K}\boldsymbol{U} = \boldsymbol{F}, \tag{4.23}$$

where $\boldsymbol{U}$ denotes displacement field, $\boldsymbol{M}$ the global mass matrix, $\boldsymbol{C}$ the global absorbing matrix, $\boldsymbol{K}$ the global stiffness matrix, and $\boldsymbol{F}$ the source term. The diagonal mass matrix $\boldsymbol{M}$ allows for simple explicit time-marching scheme to solve Eq. (4.23).

To ensure spatial accuracy and time stepping stability, attention should be paid in designing the mesh and choosing of minimal time step. The smallest element size $\triangle h$ should be designed so as to resolve the smallest wave period regulated by the relation expressed in Eq. (4.24)

$$\triangle h < \frac{\min(v_s)}{\kappa f_c}, \tag{4.24}$$

where $\min(v_s)$ is the smallest shear wave velocity, f_c denotes central frequency of the source. The value of κf_c defines the upper power frequency. For Ricker wavelet, κ is 2.5. It is found that approximately 5 mesh points per minimum wavelength is very accurate when the Lagrangian polynomial degree of order 8 is used (Komatitsch & Vilotte (1998)).

Given the minimum mesh size $\triangle h$ defined in Eq. (4.24), the time stepping size $\triangle t$ must satisfy the condition Eq. (4.25) in order for the explicit time-marching scheme to be stable.

$$\triangle t < C\frac{\triangle h}{\max(v_p)}, \tag{4.25}$$

where $\max(v_p)$ is the maximum pressure wave velocity. C denotes the Courant constant. By heuristic rule $C \simeq 0.5$ for regular meshes, C reduces to approximately 0.3 to 0.4 for irregular meshes (Komatitsch et al. (2005)).

4.4.1.2 2D tunnel seismic model

A simple geometrically parameterized model of the underground influenced by a tunnel is set up for the purpose of studying the tunnel seismic wavefields. This model also serves as a basis to perform a feasibility study of the characterization of the position and dip angle of the geological layer change ahead of the tunnel face by means of waveform inversion. The 2-dimensional modeled subsurface domain consists of a host ground and a fault zone which are supposed to be in contact along a straight interface (Figure 4.5). The tunnel diameter D is 8.5 m and the tunnel length is 30 m. Distance from the ground surface to the tunnel crown is $4D$. The model extends 125 m along the tunnel axis and the overall height is 76.5 m. The ground surface, the tunnel side walls, and tunnel face are modeled as free surfaces, all other faces in the model are absorbing boundaries. A point source

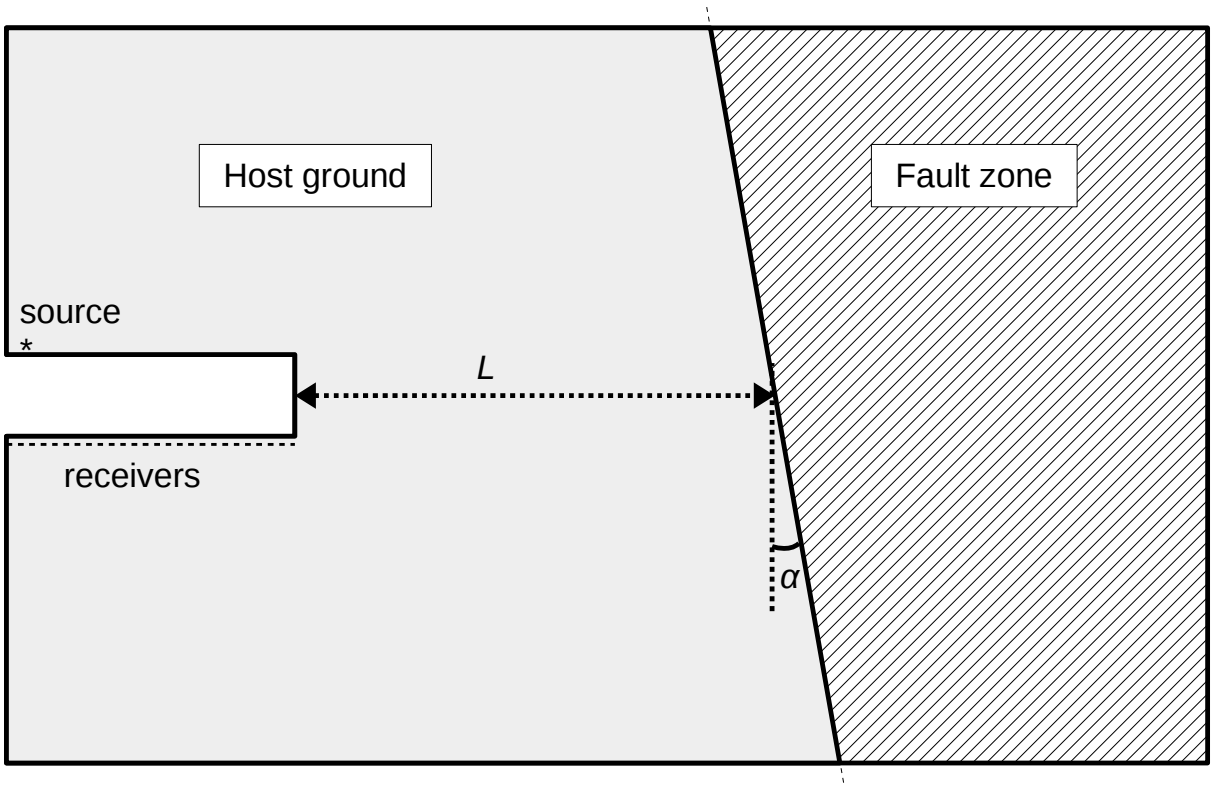


Figure 4.5: 2-dimensional tunnel model in ground with a dip fault zone for SEM simulation of wave propagation

with central frequency $f_c = 200$ Hz is located at the tunnel crown at 2 m from the left boundary of the model. Thirty one receivers are placed in a line along the tunnel bottom at interval spacing of 1 m. A geological layer change occurs 50 m ahead of tunnel face and with dip angle of -10 degree. Material properties are defined as in Table 4.1. To account for attenuation, materials are assumed to have viscoelastic quality factors $Q_P = Q_S = 50$.

Table 4.1: Material properties for simulation of tunnel seismic waves

Properties	Host ground	Fault zone
ρ [kg/m^3]	2700	1900
v_p [m/s]	3500	2200
v_s [m/s]	2000	1200

Simulated wavefields at selected time frames having important tunnel seismic events are illustrated in Figure 4.6. Although this model does not take into account elastic wave propagation in full 3-dimension and the dispersion effect caused by the excavation damage zone is neglected, the important seismic events agree well with those from the detailed

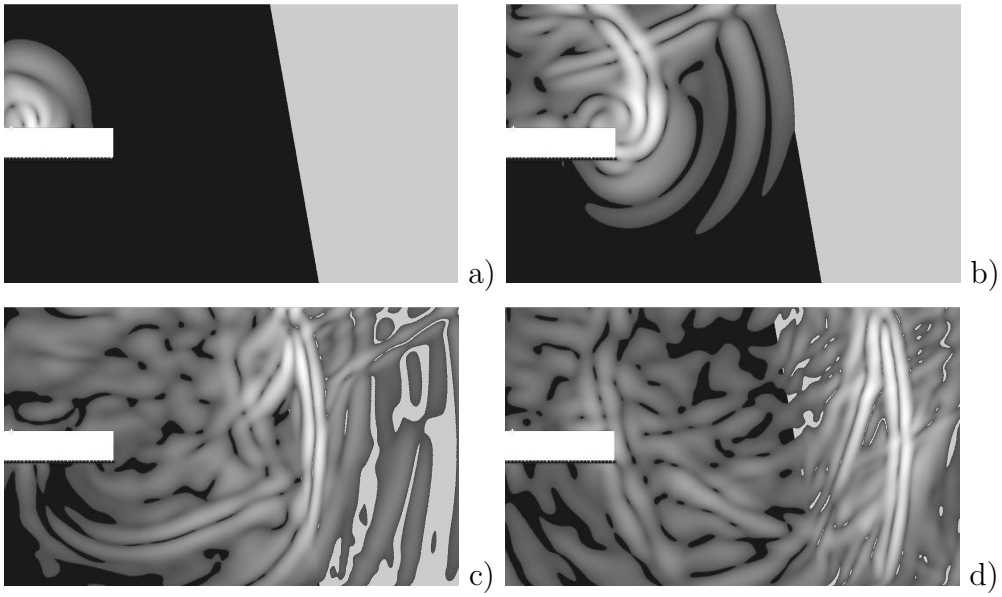


Figure 4.6: Simulated total displacement wavefields: a) waves just after starting to propagate from the source (t=10 ms), b) waves conversion at the tunnel face (t=26 ms), c) waves being reflected at the layer change (t=52 ms), and d) reflected waves reaching the tunnel face (t=72 ms)

3-dimensional finite difference modeling performed by Bohlen et al. (2007). Direct body waves (P-wave and S-wave) travel in all directions into the subsurface and become attenuated gradually. The Rayleigh waves travel strongly along the tunnel wall to the tunnel face where most of its energy is scattered into the secondary body waves and the rest scattered Rayleigh waves keep circulating around the tunnel wall. The converted body waves (mainly S-wave) propagate further ahead of the tunnel. Part of the propagating wave energy is reflected at the heterogeneity (dip layer interface) and travels back to the tunnel face in the form of body waves. As these reflected body waves reach the tunnel face, much of their energy is converted back to Rayleigh waves. These waves travel along the tunnel wall and are recorded by the receivers.

Synthetic reference data are generated with distance from the tunnel face to the geological layer change $L = 50$ m (measured at center line of the tunnel axis) and dip angle $\alpha = -10°$(counterclockwise from the perpendicular line to the tunnel axis).

4.4.2 Inverse analysis

4.4.2.1 Misfit functional

The waveform misfit between measured waveforms $\boldsymbol{d}^{obs}$ and simulated seismic data $\boldsymbol{u}$ is defined as

$$S(\boldsymbol{m}) = \frac{1}{2}\sum_{j=1}^{r}\int_{0}^{\tau} \|\boldsymbol{d}^{obs}(\boldsymbol{x}_j, t) - \boldsymbol{u}(\boldsymbol{x}_j, t, \boldsymbol{m})\|^2 dt. \tag{4.26}$$

In Eq. 4.26, model parameters are elements of vector $\boldsymbol{m} = (L, \alpha)$. The misfit functional is calculated in the concerned duration τ and for r receivers. Other efficient measures of the waveform misfit based on separation of phase and amplitude information are proposed in Bozdağ et al. (2011). It has been shown in Chi et al. (2014) that the envelope misfit is more convex than the traditional waveform misfit Eq. 4.26. The envelope misfit is defined as

$$S(\boldsymbol{m}) = \frac{1}{2}\sum_{j=1}^{r}\int_{0}^{\tau} \|\boldsymbol{D}^{obs}(\boldsymbol{x}_j, t) - \boldsymbol{U}(\boldsymbol{x}_j, t, \boldsymbol{m})\|^2 dt. \tag{4.27}$$

In Eq. 4.27, $\boldsymbol{D}^{obs}(\boldsymbol{x}_j, t)$ and $\boldsymbol{U}(\boldsymbol{x}_j, t, \boldsymbol{m})$ are the instantaneous amplitudes of the measured and modeled waveforms respectively. The instantaneous amplitude of a waveform is a complex norm of its analytic signal whose real part is the signal itself, and the imaginary part is the signal's Hilbert transform. The plots of one waveform and its envelope are compared in Fig. 4.7. Owing to its milder convexity, the inversion in this study is based on the envelope misfit Eq. 4.26.

By simulations over a gridded parameter space, a contour of the envelope misfit landscape can be viewed in Fig. 4.8. It can be seen that the misfit topography exposes a dominant region containing the global minimum point at the reference model and several areas of local minima elsewhere in the parametrized model space. The multimodal property of the waveform misfit can be explained by the so-called cycle-skipping artifact, i.e. the wrong cycle of the modeled and measured waveforms are fitted together when the phase difference between the two waveforms is greater than half of a cycle (Virieux & Operto (2009)). This problem happens because smaller local mismatch is likely achieved by fitting waveforms to the closest wave cycle. Also, it can be seen in Fig. 4.8 that the valley that contains the global minimum is very hilly along the parameter L while it elongates with gentle slopes along the parameter α. The known characteristics of the misfit landscapes prompt difficult convergence behavior in the inversion process (Nguyen & Nestorović (2015*b*)).

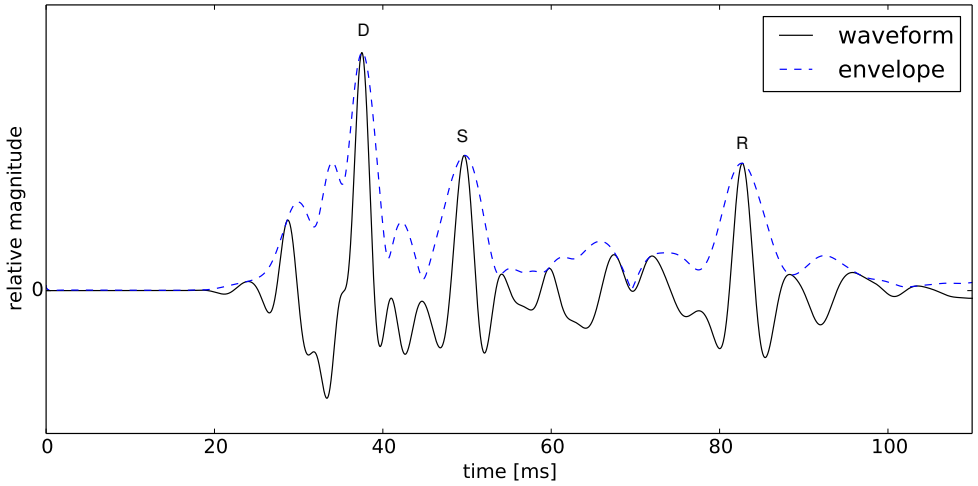


Figure 4.7: Vertical displacement waveform and its envelope at one receiver placed on the tunnel wall, D: direct surface wave, S: scattered surface wave, R: reflected surface wave

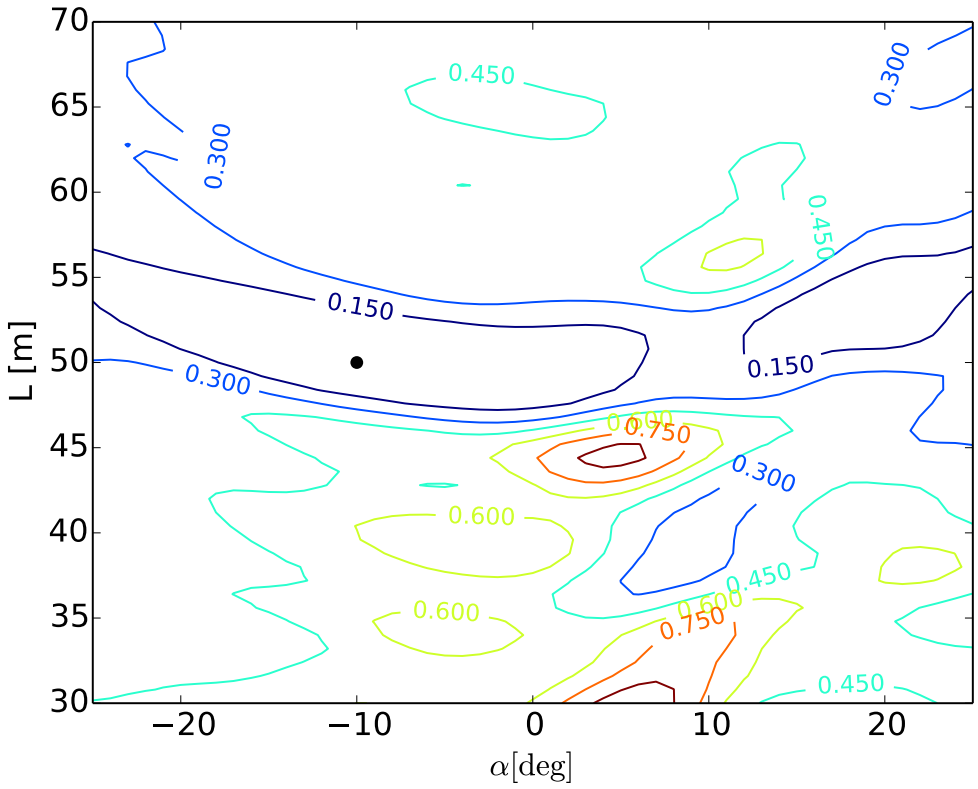


Figure 4.8: Contour of envelope misfit topography with respect to parameters L and α. A dot denotes the reference model

4.4.2.2 Test 1: Inversion for an inclined layer change

The UHSA algorithm is set forth to invert the synthetic tunnel seismic waveforms of the above described underground scenario. With respect to time-consuming forward simulation of wave propagation problem and the multimodal property of the misfit landscape, the authors aim to run a limited number of annealing cycles and focus more heavily on the local minimization of the UKF procedure to fast achieve the misfit minimum. Thus, 50 annealing cycles and 8 UKF iterations are set for the inverse simulation. Annealing simulation is performed with high initial temperature, $T_0 = 10$ degrees, and ends with $T_{50} = 0.1$ degree to allow for high acceptance ratio. Parameters of the UKF are configured as follows: $\boldsymbol{P}_0 = \mathrm{diag}(0.1^2, 0.1^2)$, $\boldsymbol{Q} = \mathrm{diag}(0.01^2, 0.01^2)$, $\boldsymbol{R} = 0.1^2\boldsymbol{I}_r$, where $\boldsymbol{I}_r \in \mathbb{R}^{r\times r}$ is an identity matrix with size equal to the number of receivers $r = 31$. The initial model ($L = 60$ m, $\alpha = 15°$) is intentionally placed distantly from the true model parameters in order to challenge the search capability of the UHSA algorithm on a multimodal misfit landscape. The parameters L and α take values in feasible ranges in $[30\text{ m}, 50\text{ m}]$ and $[-25°, 25°]$ respectively.

Results for one run of the UHSA for inversion of tunnel seismic waves plotted in Fig. 4.9 show typical convergence behavior of the UHSA algorithm. The misfit is minimized quickly in a few cycles, but the global minimum cannot be found until the 38-th cycle. Until this global convergent state, 900 forward model runs of the computationally demanding wave propagation have been used, mostly by the UKF procedure for local exploitation of the misfit topography. Since global minimum is achieved negligible increments of the number of forward model runs are observed because of very low acceptance rates.

Waveforms at every fifth receiver are plotted in Fig. 4.10 for a visual comparison between the reference model and the initial model, and the reference model and the identified model. The starting model generates waveforms which are noticeably different from those generated by the reference model in that the reflected waves arrive much later. The direct waves and waves converted at the tunnel face, which arrive much earlier to the receiver array, generated by the two models show similar waveforms because they are not affected by the position and angle of the layer change far ahead of the tunnel face. It can be seen that waveforms resulting from the identified model are different from those of the initial model in that the reflected waves are shifted to the true model. In addition to small misfit value at the end of the inversion process, insignificant residual between waveforms of the identified model and the true model confirms global convergence of the inverse analysis.

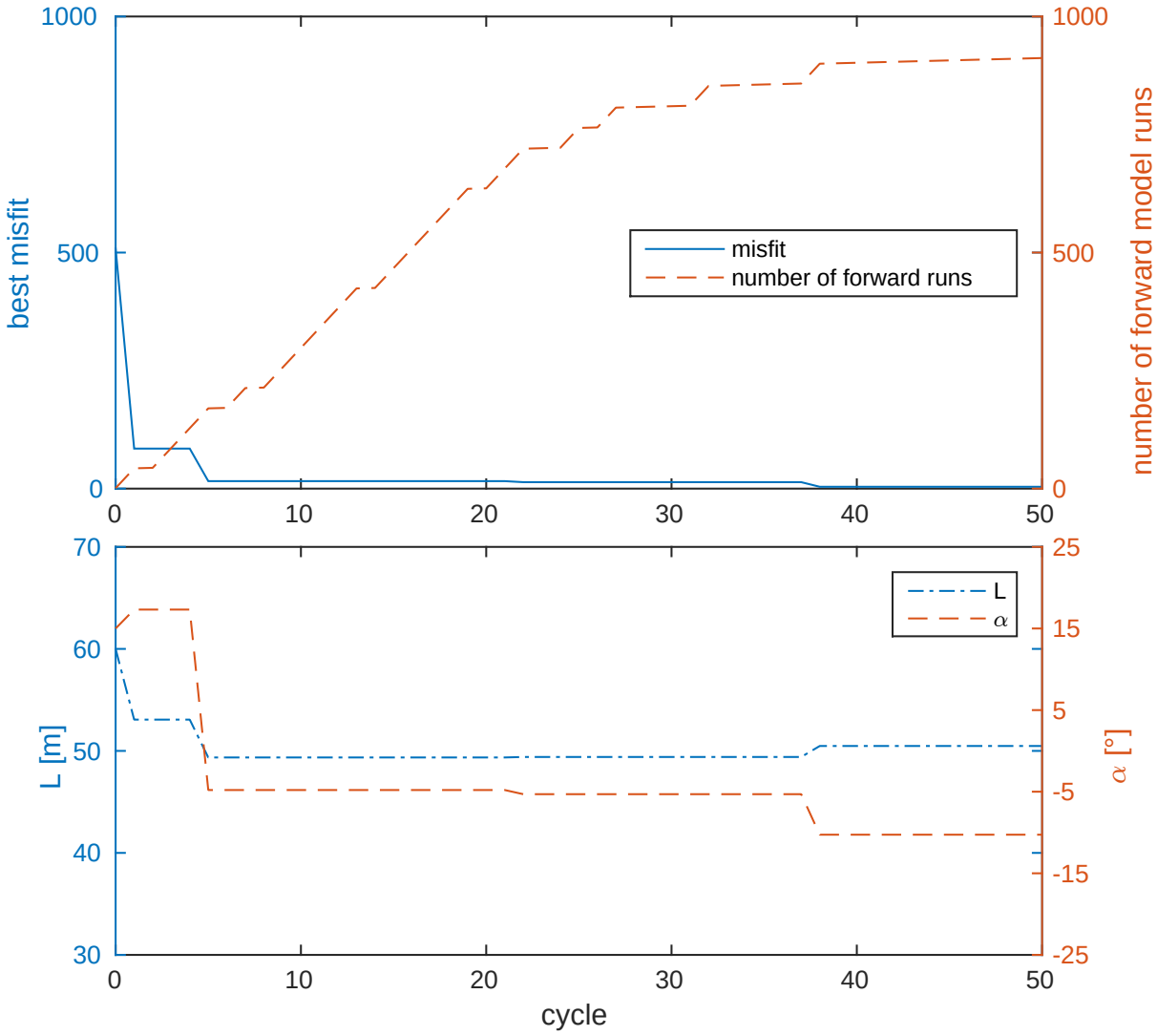


Figure 4.9: Results for one run of UHSA's 50 cycles: best minimum misfit and number of forward model runs (upper) and values of model parameters (lower) along annealing cycles

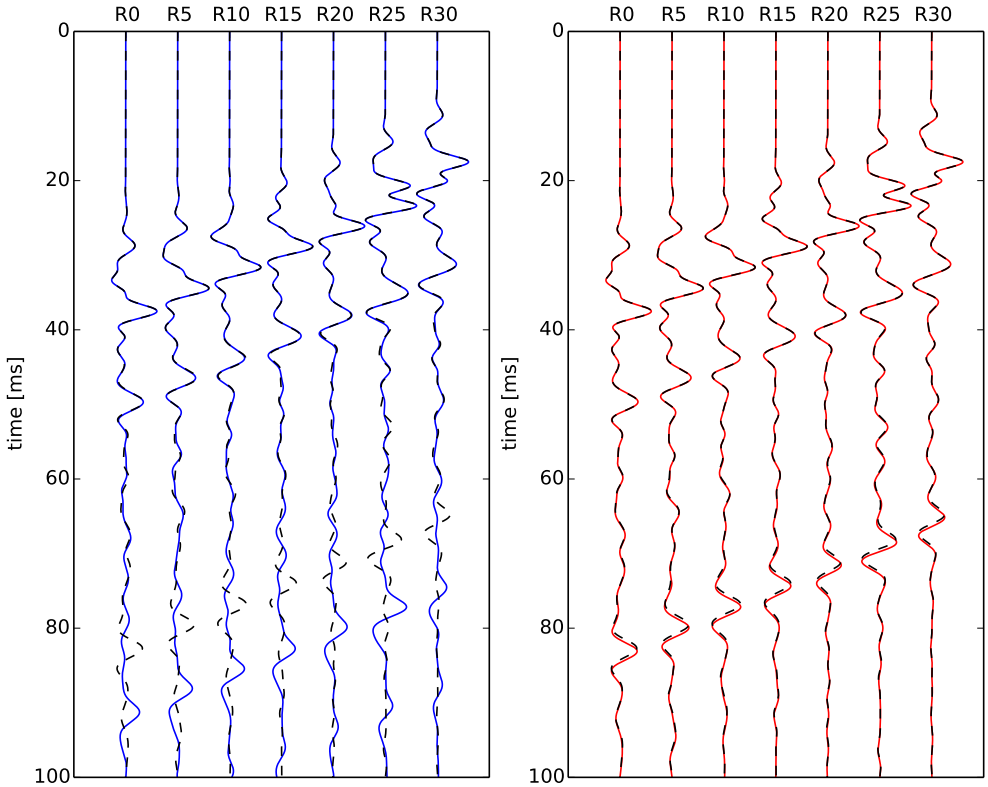


Figure 4.10: Selected waveforms resulting from the reference model (dashed-black line), the starting model (blue line), and the identified model (red line)

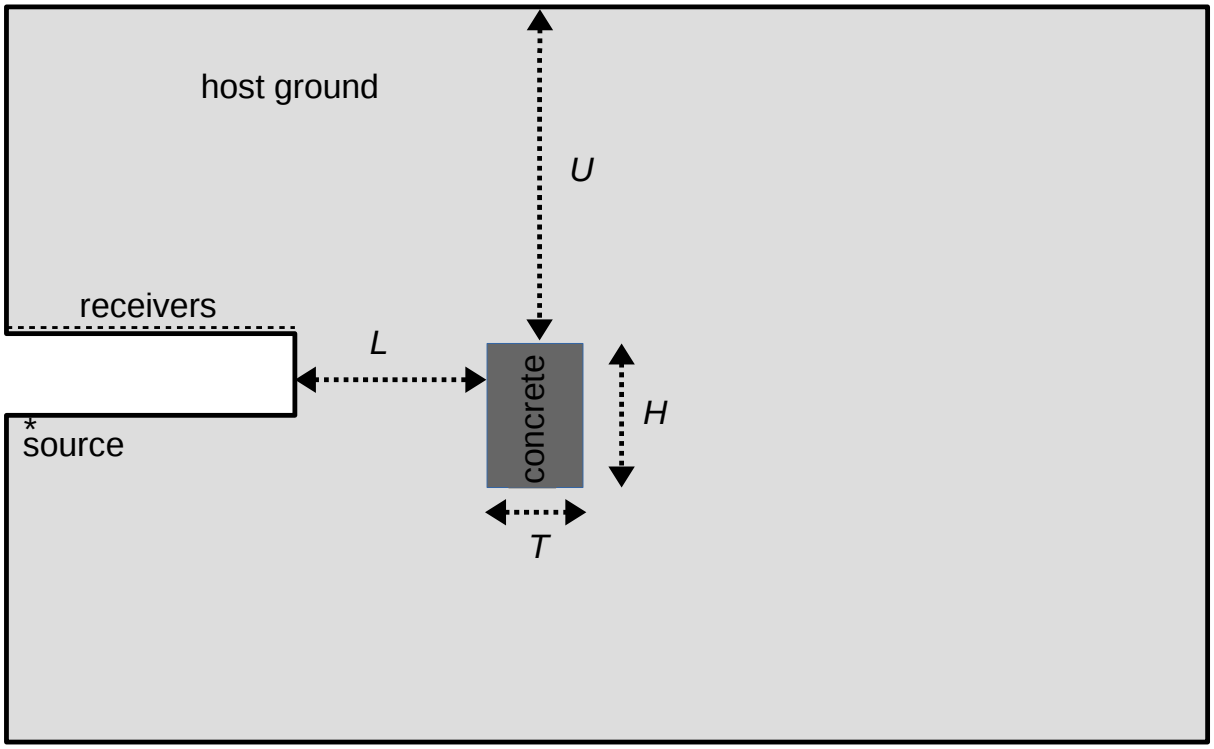


Figure 4.11: 2-dimensional tunnel model in ground an artificial concrete remnant. The overall modeled area is 125 m × 76.5 m. The concrete size and position are represented by the geometrical parameters $L = 20$ m, $U = 35$ m, $T = 10$ m, $H = 15$ m

4.4.2.3 Test 2: Inversion for a concrete remnant

It is usually encountered when tunneling in urban areas also remnants of civil engineering structures remain after they had been demolished. Here, it is assumed that a large concrete block placing in proximity to the tunnel face as shown in Fig. 4.11. In this case, the parameters of interest to the tunnel engineers are the anomaly's position (L and U) and size (T and H) as well as its material properties. With this more complicated hindrance situation, the authors aim to test the algorithm capability further to deal with a problem of a relative higher number of parameters.

An arbitrarily shaped anomaly is more realistic for a geological disturbance. However, the number of parameters required to parametrize the anomalous structure increases quite an extent and therefore the runtime the inversion takes will be tremendously significant for the computer hardware available to the authors (Linux PC, 3.3 GHz).

The amount of energy reflected back by a small block of remnant concrete is expected to be much lower than the geological situation in Test 1 above. In order to mitigate reverberating waves traveling back and forth between the ground surface and the free

surface at the tunnel crown, the source is placed at the tunnel invert. Thus, most of the excitation energy propagates in direction ahead of the tunnel face and the bottom absorbing boundary. The receivers are placed along the tunnel crown wall to reduce displacement magnitudes of the direct waves from the source.

Material properties of the reference model are shown in Table 4.2. No attenuation is taken into account in this Test 2 to reduce computational time of the forward model because of a hurdle in a high number of forward model evaluations required for this rather high dimensional inverse analysis.

Table 4.2: Material properties of the reference model in Test 2

Properties	Host ground	Concrete remnant
ρ [kg/m^3]	1900	2300
v_p [m/s]	1000	4100
v_s [m/s]	600	2300

The overall modeled subsurface and the tunnel area have the same dimensions as in the Test 1. The parameter values of the reference model are given as $\boldsymbol{m}^{true} = (L = 20\text{ m}, U = 35\text{ m}, T = 10\text{ m}, H = 15\text{ m}, v_p = 4100\text{ m/s}, v_s = 2300\text{ m/s})$. The feasible bounds for the model parameters are taken to be in ranges of $[5\text{ m}, 50\text{ m}]$, $[5\text{ m}, 40\text{ m}]$, $[5\text{ m}, 40\text{ m}]$, $[5\text{ m}, 30\text{ m}]$, $[3500\text{ m/s}, 5000\text{ m/s}]$, and $[1000\text{ m/s}, 2500\text{ m/s}]$ for L, U, T, H, v_p, and v_s respectively.

The UHSA is set to run 200 annealing cycles and with only 4 UKF iterations to achieve tractable computation time in approximately one week run on a standard computer. Starting and end temperatures of the annealing process are set the same values as for the Test 1. The UKF is configured to have rather scattered sigma-points with $\boldsymbol{P}_0 = \text{diag}(1.0^2, 1.0^2, 1.0^2, 1.0^2, 10.0^2, 10.0^2)$. The other covariance matrices are set accordingly $\boldsymbol{Q} = \text{diag}(0.2^2, 0.2^2, 0.2^2, 0.2^2, 2.0^2, 2.0^2)$, $\boldsymbol{R} = 0.1^2\boldsymbol{I}_r$, where $\boldsymbol{I}_r \in \mathbb{R}^{r\times r}$ is an identity matrix with size equal to the number of receivers $r = 31$. The inverse analysis is started with the initial estimate $\boldsymbol{m}_0 = (L = 40.0\text{ m}, U = 20.0\text{ m}, T = 30.0\text{ m}, H = 5.0\text{ m}, v_p = 3300.0\text{ m/s}, v_s = 1200.0\text{ m/s})$.

With this parsimonious configuration of the UHSA, the globally convergent solution cannot be found at the end of the annealing cycle as the misfit value shown in Fig. 4.12 is not minimized. This misfit value decreases significantly in the first until the 83rd cycle, but no further improvement can be made until the end of the 200 annealing cycles. For the complete inverse analysis, only 4547 calls to the forward model are used.

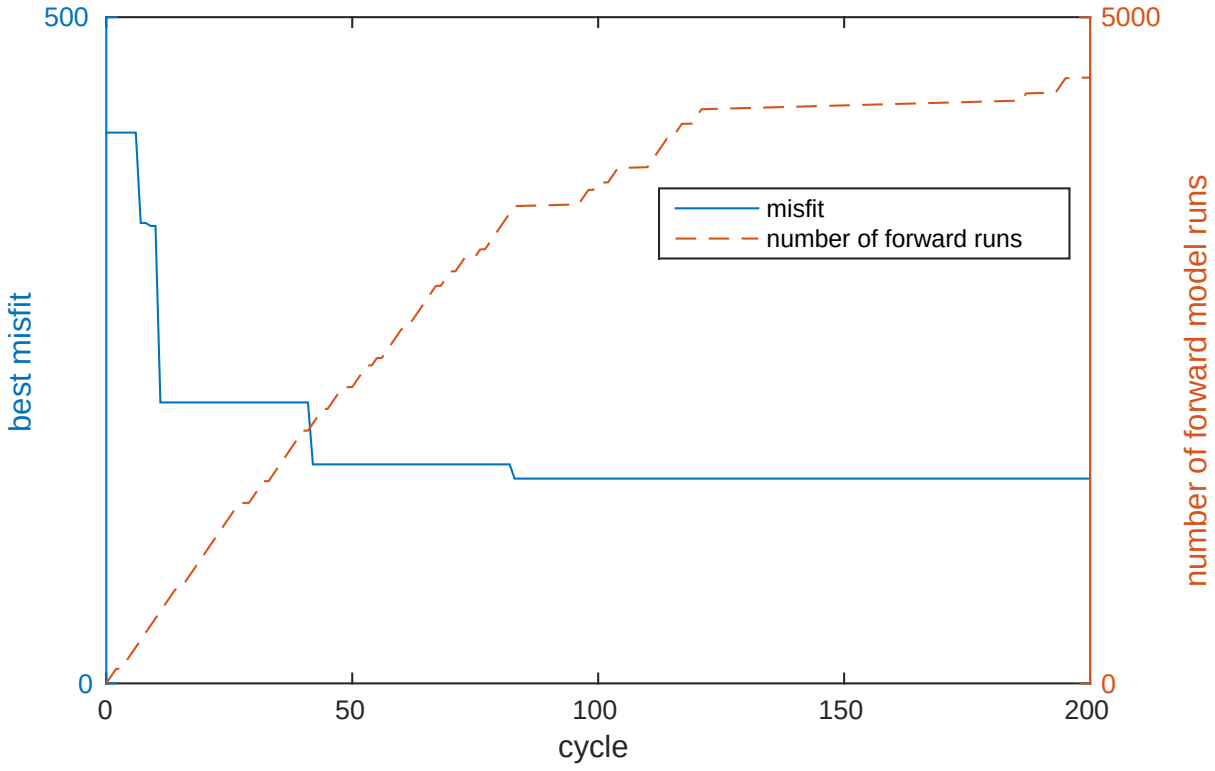


Figure 4.12: Results for one run of UHSA's 200 cycles: best minimum misfit and number of forward model runs along annealing cycles

Fig. 4.13 shows positions, sizes and wave velocities of the concrete remnant indicating the starting model and the identified model in comparison with the reference model. It can be seen that the distance and depth to the disturbance are found. Considering the height, the thickness and the velocities of the concrete remnant, the identified state is not exactly those of the reference model. In other words, a local minimum solution is found at the end of the annealing process. Wrong global convergence of the last four parameters can be explained by their less sensitivity to affect the reflected waves. More SA cycles and more UKF iterations may be needed to achieve higher accuracy of the identified model. However, the use of higher SA cycles and UKF can be realized when a high power computing facility is available. One other way to improve the overall inversion speed of the UHSA is to parallelize the mappings of the sigma-points; mappings of all the sigma-points are done simultaneously using multiple computers.

Placement positions of the receivers have a strong influence on the convergence due to the information coverage of the received signals. However, placement of receivers in this particular application is constrained to be on the tunnel sidewall. In the case drilling boreholes is allowed, receivers can be placed well a distance from the tunnel face to record the direct propagating waves. It is then expected that higher accuracy imaging of the anomaly is achieved.

Nevertheless, the waveforms resulting from the identified model displayed in Fig. 4.14 show significant mismatch reduction through the inversion process. In a similar manner to the Test 1, the reflected waves are improved the most by the inversion. The low-amplitude wiggly waves resulting from the reverberations between the interface and the free surfaces can be hardly fitted. Because the envelope misfit definition is used, the effect of these low-amplitude waves to the inversion is lessened.

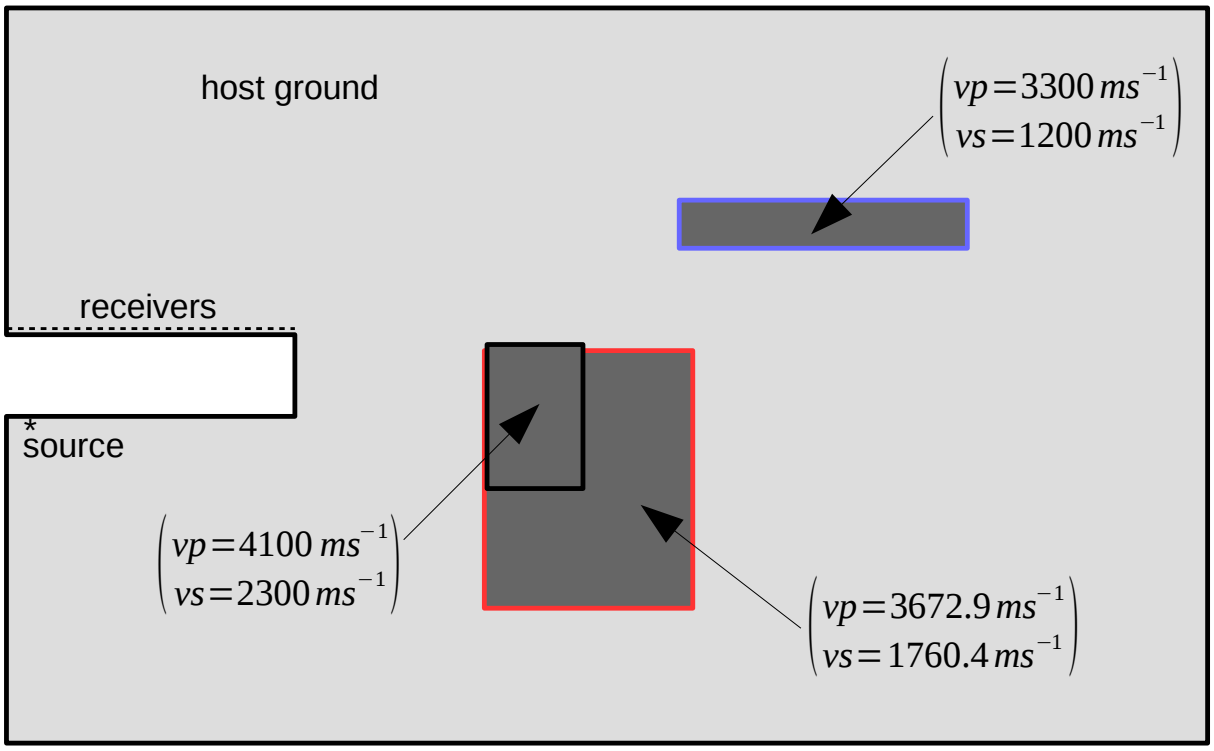


Figure 4.13: Positions, sizes and wave velocities of the concrete remnant for the starting model (blue) and the identified model (red) in comparison with the reference model (black)

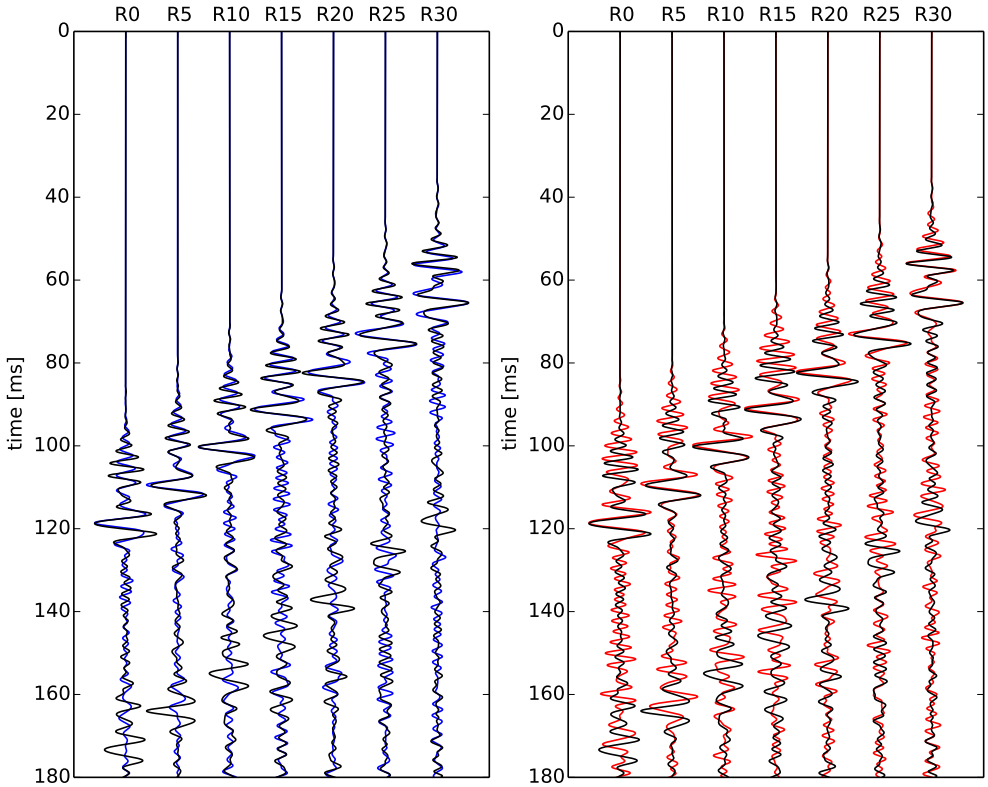


Figure 4.14: Selected waveforms resulting from the reference model (black line), the starting model (blue line), and the identified model (red line)

5 Full-waveform inversion supported by parametric level-set representation

5.1 An effective approach to reconstructing multiple geological disturbed zones

As full-waveform inversion belongs to the model-based reconstruction methodology, which is usually run in an iterative manner to update material properties of the subsurface, a considerable amount of time and computation power is needed to obtain the reconstruction results. Engineering applicability, on the other hand, demands reliable and immediate reconstruction of the disturbance scenarios. Speedy inversion can be achieved by starting the inversion process with a very good initial model and using a derivative-based method such as conjugate gradient, Levenberg-Marquardt, Broyden-Fletcher-Goldfarb-Shanno (BFGS) methods, or other local optimization methods. This family of gradient-based optimization can produce an entirely wrong reconstruction result (local minimum) if little is known about the structure of the disturbance, which is usually the case, or if the numerical calculations of the high dimensional derivatives are not exact enough.

At the other end of the inversion methodology are the global optimization methods that can avoid the solution being trapped at one of the local minima and the statistical Bayesian inference that can construct the posterior probability density of the concerned hidden quantities. Stochastic algorithms can feasibly realize this family of methods via a meta-heuristic optimization process or a Markov-chain Monte Carlo (MCMC) sampling strategy. Although being actively researched, these approaches are still very computationally demanding and therefore are not preferred for large models and models of complex physics that require high computational power.

Most often, the continuous material properties of the inversion domain are discretized depending on the spatial discretization and sometimes on the numerical integration scheme used to solve the forward problem. This method for assigning the material properties leads

to overly too many spatial unknowns in the corresponding inverse problem. For such a pixel based representation of the material distribution, gradient-based optimization is the method of choice to minimize the data misfit given the condition that the initial model must be close enough to the true scenario to avoid the solution being trapped in one of the local minima. The gradients of the misfit functional in high dimensions can be economically computed by the adjoint-wavefield approach (Virieux & Operto (2009)). However, the pixel-based FWI by this approach often leads to blurred reconstructions, i.e. the borders surrounding regions of sharp material contrasts cannot be clearly delineated.

At the other extreme, the hidden structure can be directly parameterized with a few parameters using a set of lines or the basic geometric functions. However, this way of parameterization is limited to very simple objects only, and it requires careful constraining the parameters to avoid self-intersecting of the geometric elements. Moreover, re-arrangement of the geometric components in the computational domain during the inversion process often requires re-meshing the computational grid to assure convergence of the forward solves.

Therefore, it is very beneficial for the inversion process to have a parameterization method that can flexibly define arbitrarily irregular objects using only a limited number of parameters. Such an efficient method helps to free the researchers and users from using only the gradient-based optimization option and encourages the use and development of other possible holistic inversion approaches based on global optimization and MCMC procedures.

5.1.1 Level-set method for flexible representation of the disturbance

The original idea of the level set method is to define a smooth function $\phi(\boldsymbol{x}, t)$ that represents the interface ∂D of the multiply connected region D in the computational domain Ω as a set where $\phi(\boldsymbol{x}, t) = 0$. The subsequent motion of the interface ∂D is analyzed by convecting the level-set function $\phi(\boldsymbol{x}, t)$ influenced by a velocity field $\boldsymbol{v}$ and under control of the Hamilton-Jacobi Eq. 5.1. The velocity field $\boldsymbol{v}$ is dependent on space, time, and the coupled external physics (Osher & Fedkiw (2001)).

$$\frac{\partial \phi}{\partial t} + \boldsymbol{v} \cdot \nabla \phi = 0. \tag{5.1}$$

One of the advantageous features of the level set method is that topological breaking and merging of the region D are taken into account naturally by the evolution of the level

set function. In inverse problems, this means no additional parameter(s) is assigned to control the number of multiple objects, which is unknown, during the iterative updating process. One further advantage of the level set method over the pixel-based representation is that it results in crisp delineation between the background and the objects.

For the purpose of identifying multiple disturbances, which are defined as interfaces of inclusions in a host domain, I only use the level sets to represent the arbitrarily multiple disturbances. It is noted that one can approach to solve for the evolution Eq. 5.1 provided that shape derivatives are calculated based on minimizing a data misfit functional.

Material distribution is assumed to have piece-wise constant values that define the background domain and the possibly multiple disturbance regions embedded in the background. Then, the subsurface region can be characterized by the spatial distribution function $\chi_D(\boldsymbol{x})$ as

$$\chi_D(\boldsymbol{x}) = \begin{cases} 1, & \text{if } \boldsymbol{x} \in D, \\ 0, & \text{if } \boldsymbol{x} \in \Omega \backslash D. \end{cases} \tag{5.2}$$

Given the characteristic function Eq. 5.2, the material properties are assigned to the whole investigated domain by the following simple multiplication rule

$$m(\boldsymbol{x}) = m_i \chi_D(\boldsymbol{x}) + m_o \left(1 - \chi_D(\boldsymbol{x})\right), \tag{5.3}$$

where m_o is the material properties belonging to the background, and m_i are those of the disturbance. It is noted that m_o and m_i can have multiple components, for example material's dynamic elasticity can be defined by pressure wave speed v_p, shear wave speed v_s and density ρ.

The level-set function comes in play to flexibly define the interface ∂D between the background and the disturbance. The level-set function $\phi : \Omega \mapsto \mathbb{R}$ is a smooth function defined on $\boldsymbol{x} \in \Omega$. Often, the zero-level of ϕ is used to define the interface ∂D. Then, the regions inside and outside of the disturbance are delineated such that

$$\begin{cases} \phi(\boldsymbol{x}) > 0, & \forall \boldsymbol{x} \in D, \\ \phi(\boldsymbol{x}) = 0, & \forall \boldsymbol{x} \in \partial D, \\ \phi(\boldsymbol{x}) < 0, & \forall \boldsymbol{x} \in \Omega \backslash D. \end{cases} \tag{5.4}$$

Similar to Eq. 5.3, material distribution in the whole space Ω is assigned by transferring the level-set function through the Heaviside function $H(\cdot) = 0.5\,(1 + \text{sign}(\cdot))$:

$$m(\boldsymbol{x}) = m_i H\left(\phi(\boldsymbol{x})\right) + m_o\{1 - H\left(\phi(\boldsymbol{x})\right)\}. \tag{5.5}$$

Because the level set function $\phi(\boldsymbol{x})$ is used in this application only for representing the geometries, the Hamilton-Jacobi Eq. 5.1 is not evolved in time. Instead, the level set function $\phi(\boldsymbol{x})$ is updated based on minimizing the discrepancy between the measured data and the modelled output. To facilitate the model updating process supported by the level-set representation, I adopt here the parametric level set (PaLS) representation suggested by Aghasi et al. (2011) which is based on aggregating a number of weighted instants of a predefined radial basis function (RBF) $\psi(\cdot)$. The RBF in this work is chosen to be the C^2-smooth Wendland's function $\psi(r) = (1-r)^4(4r+1)$ defined on $r \in [0,1]$ and equal zero if $r > 1$. The selected PaLS function is then written as

$$\phi(\boldsymbol{x}) = \sum_{i=1}^{Nb} \alpha_i \psi\left(\frac{1}{\beta_i^2}\|\boldsymbol{x} - \boldsymbol{x}_i\|\right), \tag{5.6}$$

where $\boldsymbol{x}_i$ are centers of the RBFs - conveniently called as bumps, α_i controls the magnitudes and directions of the bumps, and β_i is dilation factor defining the radial size of each of Nb bumps. The variables $\boldsymbol{x}_i$ can be set as free to tune if precise reconstruction of highly curved interfaces is required. In this work, in order to save computation power while achieving a fairly good representation of the geological situation, I populate these centers at fixed grid positions in region that are important to investigate, e.g. in tunneling — some of ten meters in the planned course ahead of the tunnel face. Therefore, in addition to the material properties of the disturbance, $\boldsymbol{\alpha} \in \mathbb{R}^{Nb}$ and $\boldsymbol{\beta} \in \mathbb{R}^{Nb}$ are the free parameters to be tuned during the inversion process in the applications in this work.

5.1.2 The proposed inversion technique

Generic inverse problems are naturally interpreted in the form of Bayesian statistical inference, i.e. the posterior probability density $\rho(\boldsymbol{m}|\boldsymbol{d}^{\text{obs}})$ of the concerned uncertain parameters $\boldsymbol{m}$ given the observation data $\boldsymbol{d}^{\text{obs}}$ is inferred from the prior probability density $\rho_{\text{prior}}(\boldsymbol{m})$ and the likelihood function $\rho(\boldsymbol{d}^{\text{obs}}|\boldsymbol{m})$ according to Bayes's theorem:

$$\rho(\boldsymbol{m}|\boldsymbol{d}^{\text{obs}}) \propto \rho(\boldsymbol{d}^{\text{obs}}|\boldsymbol{m})\rho_{\text{prior}}(\boldsymbol{m}). \tag{5.7}$$

The likelihood function $\rho(\boldsymbol{d}^{\text{obs}}|\boldsymbol{m})$ is a probability function weighting how likely the model parameters set $\boldsymbol{m}$ is to explain the observation data $\boldsymbol{d}^{\text{obs}}$. Assume that uncertainty in observation data $\boldsymbol{d}^{\text{obs}}$ and uncertainty in prior estimated parameters $\hat{\boldsymbol{m}}_0$ are modeled as Gaussian quantities with covariance matrices $\boldsymbol{R}$ and $\boldsymbol{P}_0$ respectively, Eq. 5.7 is rewritten as

$$\rho(\boldsymbol{m}|\boldsymbol{d}^{\text{obs}}) \propto \exp\left[-\frac{1}{2}\left(\|h(\boldsymbol{m}) - \boldsymbol{d}^{\text{obs}}\|^2_{\boldsymbol{R}^{-1}} + \|\boldsymbol{m} - \boldsymbol{m}_0\|^2_{\boldsymbol{P}_0^{-1}}\right)\right]. \tag{5.8}$$

One can employ a probabilistic model parameter updating technique such as the MCMC method to estimate the probability distribution of $\rho(\boldsymbol{m}|\boldsymbol{d}^{\text{obs}})$. Note that $\rho(\boldsymbol{m}|\boldsymbol{d}^{\text{obs}})$ is generally not Gaussian because the model $h(\boldsymbol{m})$ that maps from the model parameter space to the observation space is in most cases non-linear. Or one may follow the deterministic approach to seek for the parameters set $\boldsymbol{m}^*$ that minimizes the negative exponential term in Eq. 5.8, i.e.

$$\boldsymbol{m}^* = \arg\min_{\boldsymbol{m}} \frac{1}{2}\left(\|h(\boldsymbol{m}) - \boldsymbol{d}^{\text{obs}}\|^2_{\boldsymbol{R}^{-1}} + \|\boldsymbol{m} - \boldsymbol{m}_0\|^2_{\boldsymbol{P}_0^{-1}}\right) \tag{5.9}$$

The function to minimize in Eq. 5.9 is so called the data misfit functional regularized by the prior. The data misfit term and the regularization term are weighted by the corresponding inverse uncertainty covariances. For more exposition to the connection between the statistical approach and deterministic approach, the reader is advised to refer to Martin et al. (2012).

The Kalman filtering approach for off-line parameter estimation aims to estimate iteratively the posterior probability density function $\rho(\boldsymbol{m}|\boldsymbol{d}^{\text{obs}})$ but restricts itself to estimating only the first two statistical moments, i.e. the multivariate mean and covariance of the hidden model parameter set. Because the nonlinear Kalman filter can provide an unbiased posterior estimate (mean), its mean solution is equivalent to the solution of Eq. 5.9 or the maximum a posteriori estimate of Eq. 5.8. In addition, the Kalman filter can approximate the distribution of the posterior density function Eq. 5.8 in the form of Gaussian covariance. Therefore, with respect to the fullness of the estimated quantities, we can state Kalman filtering is in between the deterministic minimization for a single solution and the statistical inference for the complete posterior density function.

To employ the Kalman filtering in FWI, I assign the data misfit for each sensing node $\boldsymbol{x}_j$ as a nonlinear mapping which integrates the squared magnitude discrepancy between the measured trace and the modeled trace along the entire recorded time $t \in [0, T]$ (in time domain inversion):

$$s_j = G(\boldsymbol{m}, \boldsymbol{x}_j) = \frac{1}{2}\int_0^T \|h(\boldsymbol{m}, \boldsymbol{x}_j, t) - d^{\text{obs}}(\boldsymbol{x}_j, t)\|^2 dt. \tag{5.10}$$

For a number of r receivers, the collective data misfits form the observation vector $\boldsymbol{s} \in \mathbb{R}^r$ given the model parameters $\boldsymbol{m} \in \mathbb{R}^n$ is formed by the nonlinear mapping $\boldsymbol{G} : \mathbb{R}^n \mapsto \mathbb{R}^r$:

$$\boldsymbol{s} = \boldsymbol{G}(\boldsymbol{m}) = \{s_1, s_2, \cdots, s_r\}. \tag{5.11}$$

Then, evaluation of the data misfit function at the updating step k is cast as a pseudo-dynamic state-space system in which the state transition is stationary (Wan & van der Merwe (2001)):

$$\boldsymbol{m}_k = \boldsymbol{m}_{k-1} + \boldsymbol{w}_{k-1}, \tag{5.12a}$$

$$\boldsymbol{s}_k = \boldsymbol{G}(\boldsymbol{m}_k) + \boldsymbol{v}_k. \tag{5.12b}$$

In Eqs. 5.12a and 5.12b, $\boldsymbol{v}_k \in \mathbb{R}^r$ is the summed modeling and measurement uncertainty and $\boldsymbol{w}_{k-1} \in \mathbb{R}^n$ is the pseudo-uncertainty of the state transition. The uncertainties $\boldsymbol{v}_k$ and $\boldsymbol{w}_{k-1}$ are uncorrelated zero-mean Gaussian quantities. The covariance of $\boldsymbol{v}_k$, $\boldsymbol{R}$, and ovariance of $\boldsymbol{w}_{k-1}$, $\boldsymbol{Q}$, are assumed fixed during the filtering process. If the components in uncertain quantities $\boldsymbol{v}$ and $\boldsymbol{w}$ are statistically independent, leading to diagonal covariance matrices $\boldsymbol{R}$ and $\boldsymbol{Q}$ respectively. A set of measurement data contaminated by measurement noise can practically be described by Gaussian distribution $\rho_D(\boldsymbol{d}) \sim \mathcal{N}(\boldsymbol{d}^{obs}, \boldsymbol{R}^{obs})$. Then, covariance matrix $\boldsymbol{R}$ is obtained by adding the covariance for measurement uncertainty and the covariance for modeling uncertainty, i.e. $\boldsymbol{R} = \boldsymbol{R}^{obs} + \boldsymbol{R}^{m}$ (Tarantola (2005)).

I propose to use a special variant of nonlinear Kalman filtering — *the unscented Kalman filter* (UKF) (Julier & Uhlmann (1997)) — to estimate the hidden parameters so as to obtain minimized value of the data misfit function Eq. 5.11. The proposed methodology is written shortly as the UKF-PaLS. Hence, the FWI task is accomplished by the use the UKF for updating the model $m(\boldsymbol{x})$ described in Eq. 5.5 which is equivalent to tuning a set of uncertain parameters consisting of PaLS parameters and material properties $\boldsymbol{m} \in \mathbb{R}^n$:

$$\boldsymbol{m} = \{\alpha_1, \alpha_2, \cdots, \alpha_{Nb}, \beta_1, \beta_2, \cdots, \beta_{Nb}, m_i, m_o\}. \tag{5.13}$$

The state-space model/ filtering process is initialized with an Gaussian estimate centered at $\hat{\boldsymbol{m}} = \boldsymbol{m}_0$ with covariance $\boldsymbol{P}^m = \boldsymbol{P}_0$. Then the value for the covariance $\boldsymbol{Q}$ of the pseudo-uncertainty quantity $\boldsymbol{w}_{k-1}$ is set to be as a small fraction of $\boldsymbol{P}_0$ (Nguyen & Nestorović (2015*a*)).

The UKF improves the linearization error of the conventional extended Kalman filter by directing mapping of the estimated quantity through the nonlinear model using the *unscented transformation*. The unscented transformation is a statistical linearization method to approximate a multivariate Gaussian distribution by a set of sampled points. Unlike Monte Carlo sampling, the number of samples required by the unscented transformation is minimal, and the sample positions are set deterministically about the current estimate and its covariance. The samples used in the unscented transformation are so called the *sigma-points*.

If n model parameters are to be estimated, $2n+1$ sigma-points are defined as follows for the unscented transformation to approximate a Gaussian distribution of the current estimate centered at $\hat{\boldsymbol{m}}$ with covariance $\boldsymbol{P}^m$:

$$\boldsymbol{\mathcal{M}}_0 = \hat{\boldsymbol{m}}, \tag{5.14}$$

$$\boldsymbol{\mathcal{M}}_i = \hat{\boldsymbol{m}} + \left(\sqrt{(n+\eta)\boldsymbol{P}^m}\right)_i, \text{ for } i = 1:n \tag{5.15}$$

$$\boldsymbol{\mathcal{M}}_{n+i} = \hat{\boldsymbol{m}} - \left(\sqrt{(n+\eta)\boldsymbol{P}^m}\right)_i, \text{ for } i = 1:n. \tag{5.16}$$

The parameter η remains free to tune so as to adjust the spread of the sigma-points about the predicted mean estimate in directions of principle variances. The notation $(\cdot)_i$ denotes the i-column of the matrix within parentheses.

Each sigma-point is associated with a weight. The weights are defined as in Eq. 5.17 such that the summation of weights is unity:

$$W_0 = \frac{\eta}{n+\eta}, W_i = W_{i+n} = \frac{1}{2(n+\eta)}, \text{ for } i = 1:n. \tag{5.17}$$

Misfit measures for all r receivers are obtained by mapping the sigma-points through the data misfit function Eq. 5.11, i.e.

$$\boldsymbol{\mathcal{S}}_i = \boldsymbol{G}\left(\mathcal{M}_i\right), \text{ for } i = 0:2n. \tag{5.18}$$

This direct nonlinear mapping of the sigma-points helps to preserve second order accuracy of the Gaussian distribution as opposed to first order truncation of the Taylor series employed by the linearized Kalman filter.

After the nonlinear mapping of the sigma-points through the nonlinear model, the mean of the predicted misfit $\hat{\boldsymbol{s}}$, covariance of the predicted parameters $\boldsymbol{P}^m$, covariance of the misfit at the predicted model $\boldsymbol{P}^s$, and cross-covariance between them $\boldsymbol{P}^{ms}$ are approximated by weighted summation rule as follows:

$$\hat{\boldsymbol{s}} = \sum_{i=0}^{2n} W_i \boldsymbol{\mathcal{S}}_i, \tag{5.19}$$

$$\boldsymbol{P}^m = \sum_{i=0}^{2n} W_i \left(\boldsymbol{\mathcal{M}}_i - \hat{\boldsymbol{m}}\right)\left(\boldsymbol{\mathcal{M}}_i - \hat{\boldsymbol{m}}\right)^t + \boldsymbol{Q}, \tag{5.20}$$

$$\boldsymbol{P}^s = \sum_{i=0}^{2n} W_i \left(\boldsymbol{\mathcal{S}}_i - \hat{\boldsymbol{s}}\right)\left(\boldsymbol{\mathcal{S}}_i - \hat{\boldsymbol{s}}\right)^t + \boldsymbol{R}, \tag{5.21}$$

$$\boldsymbol{P}^{ms} = \sum_{i=0}^{2n} W_i \left(\boldsymbol{\mathcal{M}}_i - \hat{\boldsymbol{m}}\right)\left(\boldsymbol{\mathcal{S}}_i - \hat{\boldsymbol{s}}\right)^t. \tag{5.22}$$

The posterior mean $\hat{\boldsymbol{m}}_+$ and covariance $\boldsymbol{P}^m_+$ of the estimated model parameters are updated following the Kalman filter's measurement update step:

$$\hat{\boldsymbol{m}}_+ = \hat{\boldsymbol{m}} + \boldsymbol{K}\left(\boldsymbol{0} - \hat{\boldsymbol{s}}\right) \tag{5.23}$$

$$\boldsymbol{P}^m_+ = \boldsymbol{P}^m - \boldsymbol{K}\boldsymbol{P}^s\boldsymbol{K}^t \tag{5.24}$$

with the Kalman gain $\boldsymbol{K}$ calculated as

$$\boldsymbol{K} = \boldsymbol{P}^{ms}\left(\boldsymbol{P}^s\right)^{-1}. \tag{5.25}$$

The zero vector term in the innovation term $\boldsymbol{0} - \hat{\boldsymbol{s}}$ in the measurement update Eq. 5.23 results from the fact that the data misfit function Eq. 5.11 is expected to be zero.

To initialize the UKF procedure, some prior knowledge about the model is provided in form of Gaussian distribution $\rho_0(\boldsymbol{m}) \sim \mathcal{N}(\boldsymbol{m}_0, \boldsymbol{P}_0)$. In the first iteration of the UKF, this prior information is assigned to the current estimated mean and covariance: $\hat{\boldsymbol{m}} = \boldsymbol{m}_0$, and $\boldsymbol{P}^m = \boldsymbol{P}_0$. In the successive iterations, the estimated mean and covariance are updated by taking the respective values from the posterior estimate, i.e. $\hat{\boldsymbol{m}} = \boldsymbol{m}_+$ and $\boldsymbol{P}^m = \boldsymbol{P}^m_+$.

5.2 Inversion tests with cross-hole acoustic wave measurements

5.2.1 Acoustic waves in frequency domain

I first test the proposed approach to reconstruct the internal structure based on the acoustic wave in the frequency domain, the Helmholtz equation, because its forward model takes considerably less time to compute. The 2D model with an embedded disturbance and the resulting wavefield at 25 Hz is illustrated in Fig. 5.1. The cross-hole experiment is set up with 51 sources on the left side and 51 receivers on the right side of the model. The forward modeling code accompanying the paper van Leeuwen & Herrmann (2013) is used to compute the simulated wavefields utilized in the inversion. The 2D model is discretized into 101×101 grid points. The synthetic measurements along the depth are added by 3% of the RMS to simulate the inherent measurement noise.

The first reference model with a square disturbance at the center and the resulting scattered wavefield are plotted in Fig. 5.1

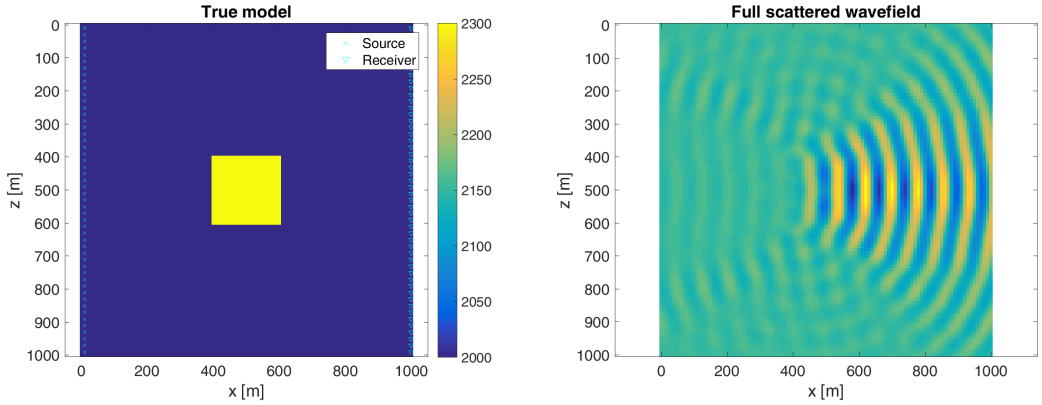


Figure 5.1: The true model and the scattered wavefield resulting from one source at the center of the height. The filled colors in the model indicate acoustic wave velocities; the filled colors in the scattered wavefield show the real part of the wavefield magnitudes.

5.2.2 Inversion results

The initial model is scattered in the square *region of investigation* (ROI), which covers an area much larger than the 'suspected' disturbance, with a grid of 8×8 equidistant bumps ($Nb = 64$) as shown in the upper-left sub-figure in Fig. 5.2. The total number of hidden parameters for the inverse problem is 129 (64 for $\boldsymbol{\alpha}$, 64 for $\boldsymbol{\beta}$, and 1 for unknown acoustic wave speed of the disturbance v_i). The wave velocity of the disturbance is set very close to the background wave velocity. The UKF-PaLS is set with the following initial and uncertainty settings in particular:

- Initial estimate: $\boldsymbol{m}_0 = \{\boldsymbol{\alpha}_0, \boldsymbol{\beta}_0, (v_i)_0\}$. In that, $\boldsymbol{\alpha}_0 = 10 \times \boldsymbol{I}(Nb, 1)$, $\boldsymbol{\beta}_0 = 20 \times \boldsymbol{I}(Nb, 1)$, $(v_i)_0 = 2020$. $\boldsymbol{I}(Nb, 1)$ is an Nb-dimensional vector of all ones.
- Initial estimation error covariance: $\boldsymbol{P}_0^m = \text{diag}(\boldsymbol{P}_0^\alpha, \boldsymbol{P}_0^\beta, P_0^{v_i})$. In that, $\boldsymbol{P}_0^\alpha = 40^2 \times \boldsymbol{I}(Nb, 1)$, $\boldsymbol{P}_0^\beta = 20^2 \times \boldsymbol{I}(Nb, 1)$, $P_0^{v_i} = 30^2$. diag foms a diagonal matrix composed the given elements.
- Data and model uncertainty covariances: $\boldsymbol{R} = (0.4)^2 \times \boldsymbol{I}(r, r)$, $\boldsymbol{Q} = (0.1)^2 \times \boldsymbol{P}_0^m$. $\boldsymbol{I}(r, r)$ is an $r \times r$ identity matrix.

The reconstructions after 1, 8, and 16 iterations are shown in the other sub-figures in Fig. 5.2 as well. It is observed that the reconstructions of the shape and material properties of the disturbances improve over iterations. The satisfactory reconstruction after the 16th UKF iteration has consumed only $16 \times (2 \times 129 + 1) = 4144$ forward model solves, which is very positive for solving an inverse problem of 129 dimensions.

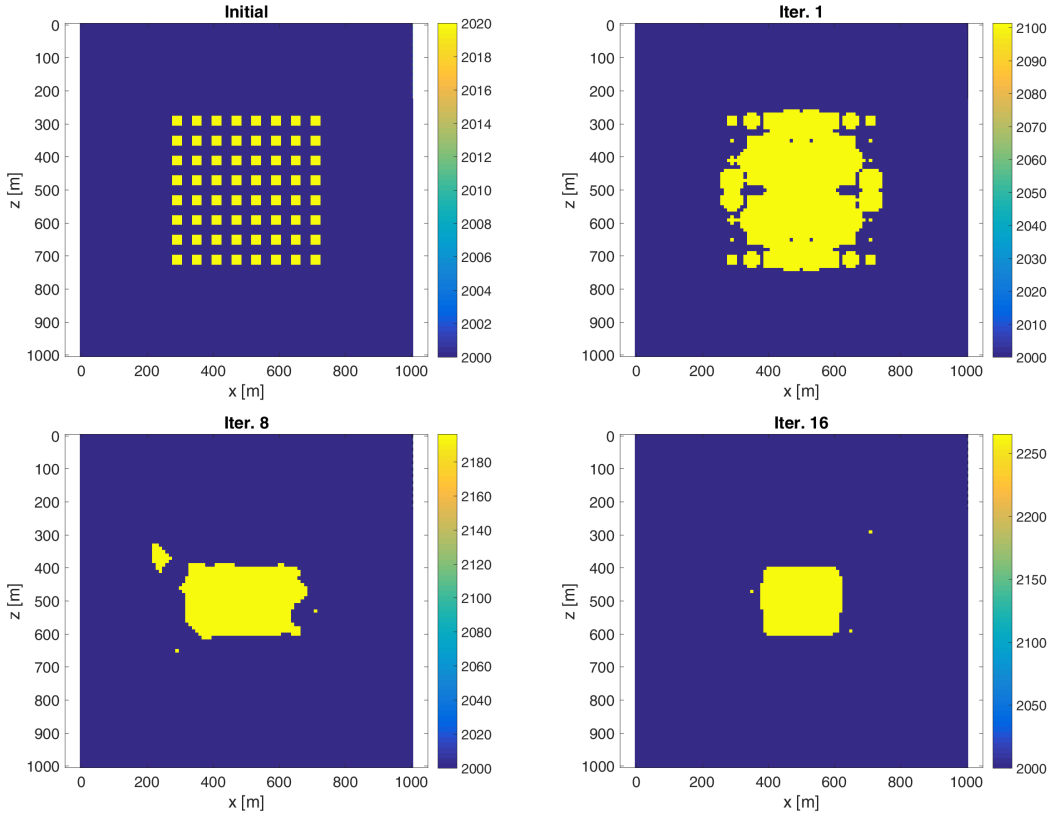


Figure 5.2: Initial model and the reconstructions after 1, 8, and 16 iterations

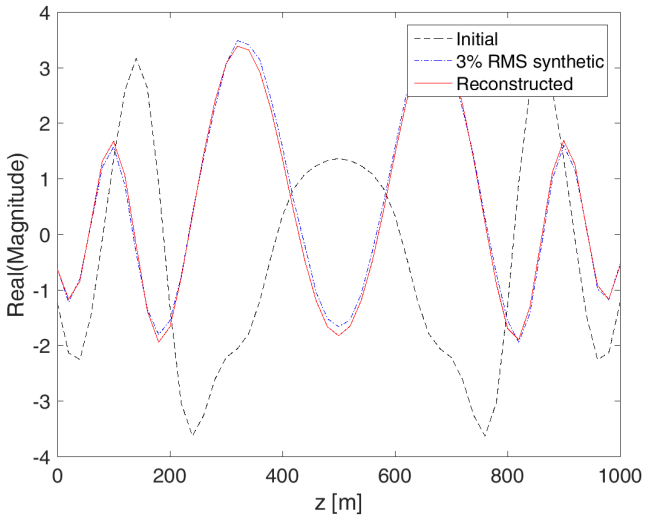


Figure 5.3: Noisy synthetic measurements and model outputs of the initial and the reconstructed models

Fig. 5.3 shows the noisy synthetic measurements, the observation data resulting from the initial model and the reconstructed model. It can be seen that the response of the reconstructed model can fit to a great extent the corresponding noisy measured data.

Next, to test the reconstruction capacity of the method with reconstructing multiple objects, in the second reference model, the square disturbance is broken into two triangles which are placed at a short distance to create two objects of the same material as shown in Fig. 5.4. The algorithmic settings of the inversion process are kept the same as in the case of reconstruction of the square disturbance; thus, the computational cost is not changed compared to the previous inversion test.

The reconstructed wave velocity field is observed to improve in the iterative inversion steps as can be seen in Fig. 5.5. The final reconstruction resembles the true model shown in Fig. 5.4 to a great extent although there exist several small-size objects which are not sensitive to the output wavefield due to the very small object-size to wavelength ratios. Because of the same reason, the areas at the small-angle corners of the disturbance objects cannot be resolved very well. The amount of data fit in the measurements can be observed in Fig. 5.6. Overall, the convergence for the second reference model is slower, and the resulting final image contains more artifacts compared to that for the first reference model.

It is noted that for waveform inversion in the frequency domain, the reconstructed image of the disturbance can be improved by starting the inversion at a lower frequency. The inversion result is then used as an initial model for the next inversion at a higher frequency.

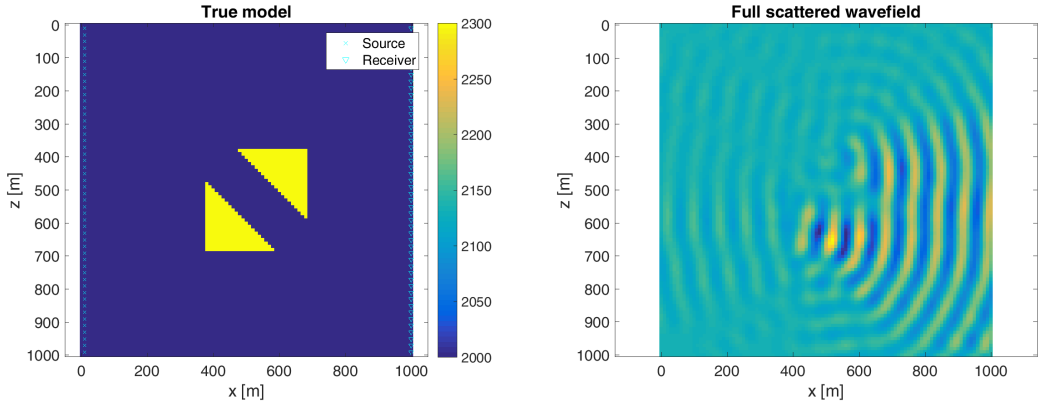


Figure 5.4: The true model and the scattered wavefield resulting from one source at the center of the height

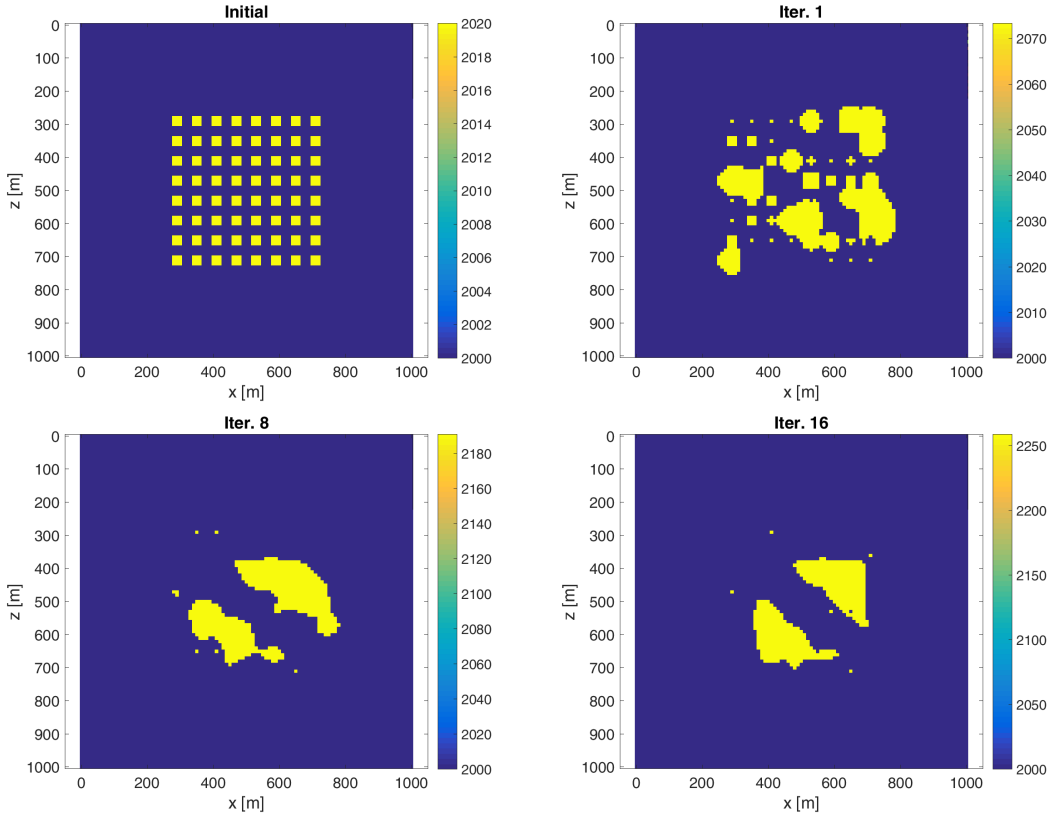


Figure 5.5: Initial model and the reconstructions after 1, 8, and 16 iterations

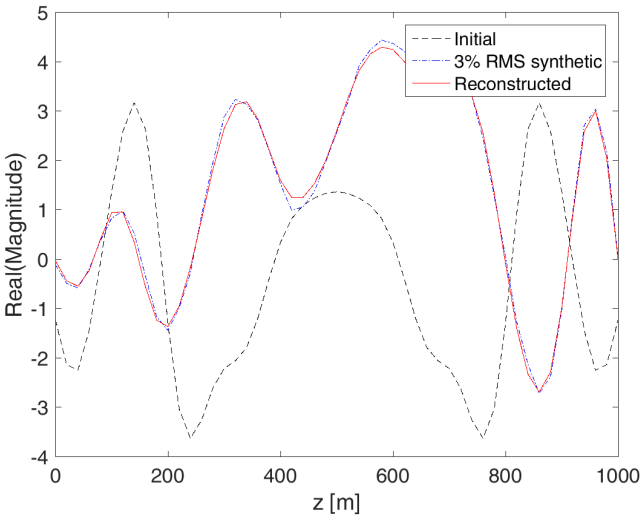


Figure 5.6: Noisy synthetic measurements and model outputs of the initial and the reconstructed models

The inversion process is then repeated with increased frequencies and ends at a frequency that can resolve the small and highly curved objects. But at the same time, this multiple frequencies inversion requires finer discretization of the model and longer computational time as the frequency increases.

Increasing the number of RBFs in the level-set Eq. 5.6 and freeing the RBFs' centers help resolve small-size, highly curved, or small-angle objects but at the same time increases the number of hidden parameters to a great extent. In the application for predicting ahead of the tunnel face, mapping these very detailed geometric features are not as important as the prediction of the existence of the disturbance and its quantitative estimate regarding the overall shape, distribution, and material property. Therefore, I follow this parsimonious parameterization approach in the main application that follows.

5.3 Application to tunnel reconnaissance

5.3.1 Tunnel seismic waves

The term *tunnel seismic waves* in this work is understood in a broad sense to refer to all kinds of elastic waves recorded on the tunnel walls and acquisition borehole short ahead of the tunnel face. I use the term *tunnel surface waves* to exclusively mention the surface waves recorded on the tunnel wall only which result from wave mode conversion

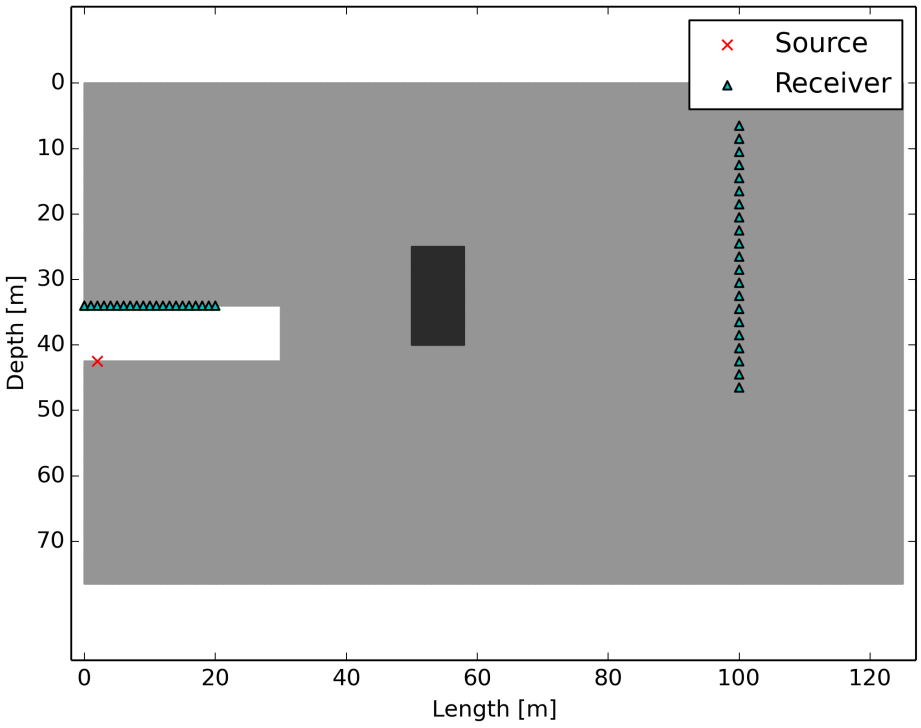
happening at the tunnel face as detailed in the works of Bohlen et al. (2007) and Jetschny et al. (2010, 2011).

The tunnel surface waves are preferably excited on the tunnel wall. The tunnel surface waves that propagate ahead to the tunnel face are converted into body waves (P-wave and S-wave). If there exists impedance change caused by a geological heterogeneity, a partial energy of the body waves are reflected/ refracted and travel back to the tunnel and the rest continues to travel through the heterogeneity and beyond. The reflected body waves undergo mode conversion when they encounter the tunnel face again. This mode conversion generates the tunnel surface waves that travel along the tunnel walls and are recorded by the geophones installed there. Analysis of the tunnel surface waves can result in a good reconstruction of the first geological reflector but likely to miss the further impedance interface(s) of the object because multiple reflections/ refractions of elastic waves in an attenuating medium result in very small amplitudes of the recorded events.

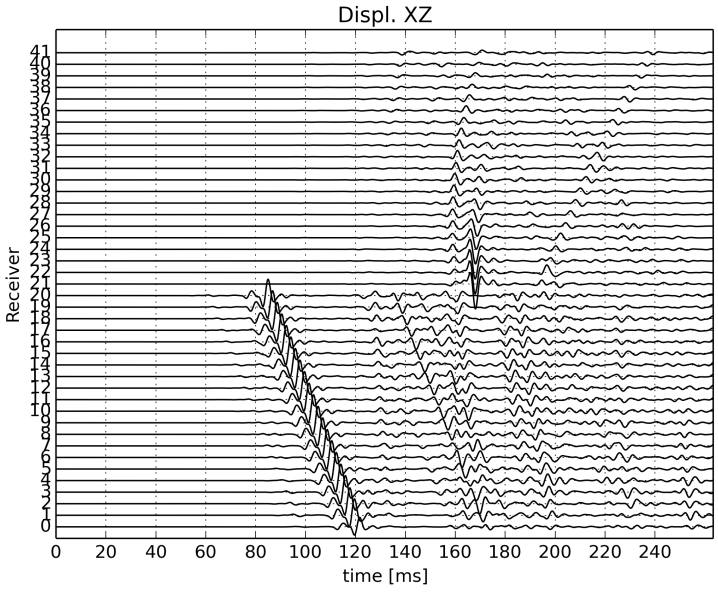
To have a better setting for FWI, I suggest also measuring direct transmission waves at a line of receivers buried in a vertical borehole at an appropriate distance ahead of the tunnel face. Besides, several sources placed at possible positions are excited to illuminate the ROI from different angles.

The fact that the elastic waves in a heterogeneous medium undergo reflections, refractions, and mode conversions and possibly multiples of them complicate the direct interpretation of the recorded waveforms. This section features the special characteristics of the tunnel seismic waves by using numerical simulations using the *spectral element method* (SEM) implemented in the open-source SPECFEM2D software (Komatitsch & Vilotte (1998)). Because SEM is more efficient in computational time and accuracy than finite element or finite difference methods in solving the elastic wave equation, the SPECFEM2D is a preferred tool to study elastic wave propagation problems. Wang & Cai (2015) used SPECFEM2D to study the influence of wavelength of the seismic waves generated from fault slip rock bursts to excavation diameter on ground motion around mine tunnels. Nguyen & Nestorović (2016) used SPECFEM2D as a forward solver to study the effectiveness of a proposed hybrid global optimization method in solving low dimensional tunnel seismic inverse problems.

Fig. 5.7a illustrates a representative disturbance situation in which a hard anomaly is located just a short distance from the tunnel face. As only a small area around and ahead of the tunnel face is taken into account in the model, all boundaries on the left, right, and bottom of the model are modeled with absorbing boundary conditions. Surface



a)



b)

Figure 5.7: A representative disturbance situation: a) the model with a single anomaly of rectangular shape and b) the resulting seismogram. Receivers 0 - 20 are placed on the tunnel wall, and receivers 21 - 41 are placed in the borehole at 70 m ahead of the tunnel face.

boundaries and the tunnel wall surrounding the excavated tunnel are modeled as free reflection surfaces. The Ricker seismic source of 180 Hz and receivers in the excavated tunnel are placed at some distance behind the tunnel face for convenience as the working space in the excavation chamber just behind the tunnel face is very occupied. In addition to the line of receivers along the tunnel wall, a series of receivers placed in a borehole well distant ahead record the direct waves traveling through possible disturbances.

The material surrounding the tunnel is assumed to be homogeneous soft ground, and the disturbance is a hard material. Wave speeds used in the simulations of elastic tunnel seismic wave propagation are tabulated in Table 5.1.

Table 5.1: Material properties for the simulations of tunnel seismic wave propagation

Properties	Host ground	Disturbance
ρ [kg/m^3]	1900	1900
v_p [m/s]	1000	3300
v_s [m/s]	600	1800

The resulting seismogram in Fig. 5.7b shows that direct interpretation of the subsurface structure is hardly possible even in this simplified geological structure. The waveforms in a seismogram contain the signature of the hidden structure we want to map. However, the seismograms become harder to interpret visually as the degree of heterogeneity of the subsurface increases. That is why automatic mapping of the subsurface region of interest by FWI is attractive. FWI benefits from the high fidelity sensing instrumentation, the increasing power of computing facility, and on-going development in advanced mathematical and computational methods.

5.3.2 Qualitative analysis of the recorded waves

To signify the content of the recorded waveforms in Fig. 5.7b, we can analyze the *scattered waveforms* resulting from subtracting the *recorded waveforms* to the *background waveforms*. The background waveforms, which are generated from the model without disturbance, is plotted in Fig. 5.8a. The recorded waveforms are plotted again in Fig. 5.8b for comparison. In the obtained scattered waveforms in Fig. 5.8c, the events that are not caused by the disturbance are muted and therefore reveal very clearly the events caused by the disturbance. It can be seen from the scattered waveforms that, for receivers in the tunnel walls, the reflection from the front boundary of the disturbance is much very

strong while that of the back boundary can be hardly seen. For the waveforms received in the borehole, it can be observed that transmission waves arrive sooner and much more weakly because the fast material disturbance accelerates the waves. Therefore, it can be seen in the scattered waveforms that very strong direct waves are 'masked' in the recorded waveforms. By visual comparison of the peak amplitudes, it can be stated that, at least for this simulation case, waveforms received in the borehole are more sensitive to the geological disturbance than waveforms received on the tunnel wall.

The scattered waveforms in Fig. 5.8c shows the strong first reflected wave (Rayleigh wave) registered by receivers on the tunnel wall (receivers 0-20) and the masked direct wave (mainly S-wave) registered by receivers in borehole (receivers 21-41). Also, the registered waveforms on the tunnel wall also show no significant reflected wave from the back interface of the disturbance object but more multiple reflections traveling back and forth between the object and the tunnel face. Therefore, a line of receivers placed in the borehole is necessary to reconstruct the anomalous object.

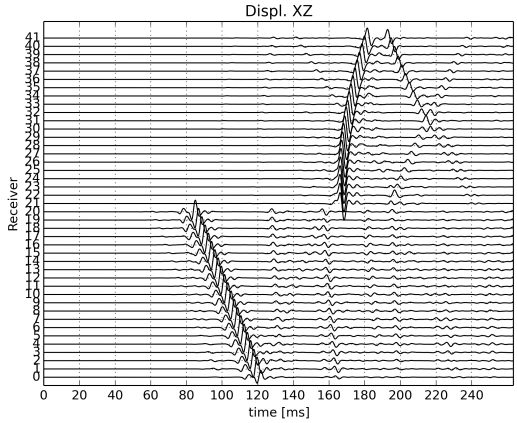
5.3.3 Inversion results and discussions

Two geological scenarios shown in Fig. 5.9 are set up for FWI tests. The first scenario (Fig. 5.9a) is the same as the situation shown previously in Fig. 5.7. The second scenario (Fig. 5.9b) is different in that the disturbance consists of two separate areas. The second scenario is expected to be harder for the reconstruction process because the area in between the disturbed structure does not have direct reflection contact with the incident waves. The wave paths, in this case, are reflected multiple times in this area, making amplitudes of the transmission waves weaker.

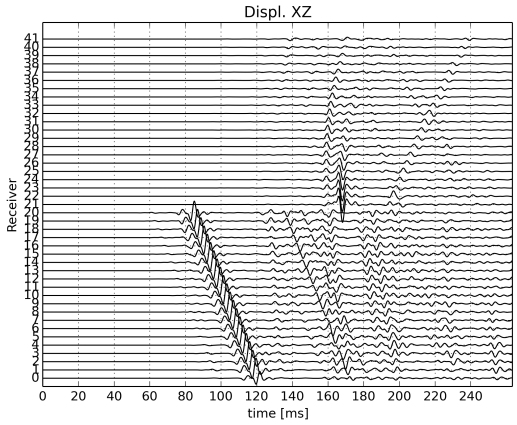
Wave velocities of the background soil can be readily derived from the first arrivals of the direct waves within known source and receiver positions; therefore, I assume in this study that the material properties of the background is known and seek for the velocities as well as the location and the geometry of the geological disturbance.

To avoid 'inverse crime,' the synthetic waveforms are generated not using the PaLS but by exact definition of the disturbance using geometric parameters of Gmsh software (Geuzaine & Remacle (2009)). This means the best model that is achieved by PaLS assignment can hardly resemble the true model.

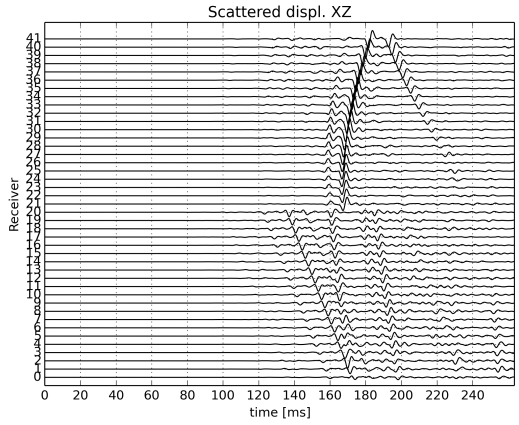
The source frequency is selected such that wavelength of the fastest S-wave is smaller than the size of the disturbance objects. For clear reconstruction of high-curvature sections and



a)

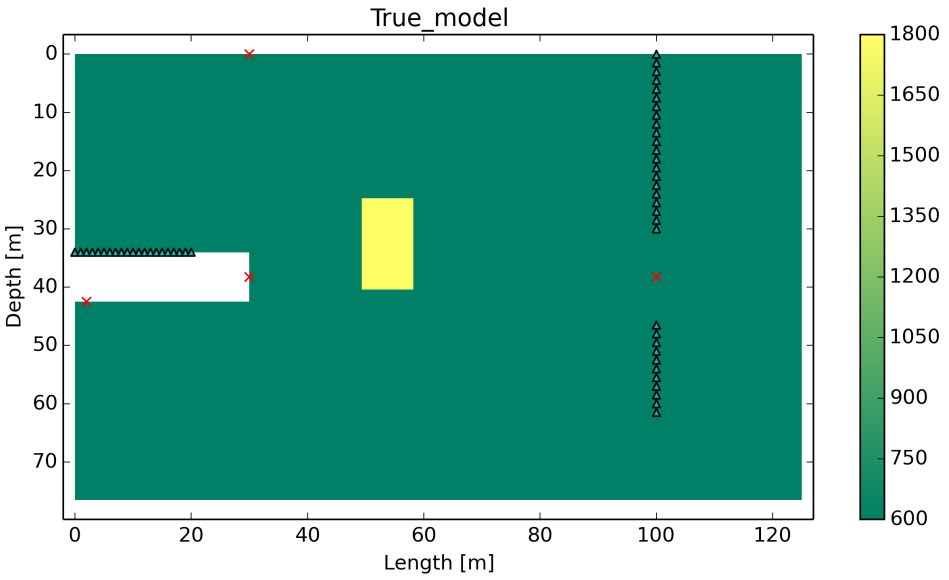


b)

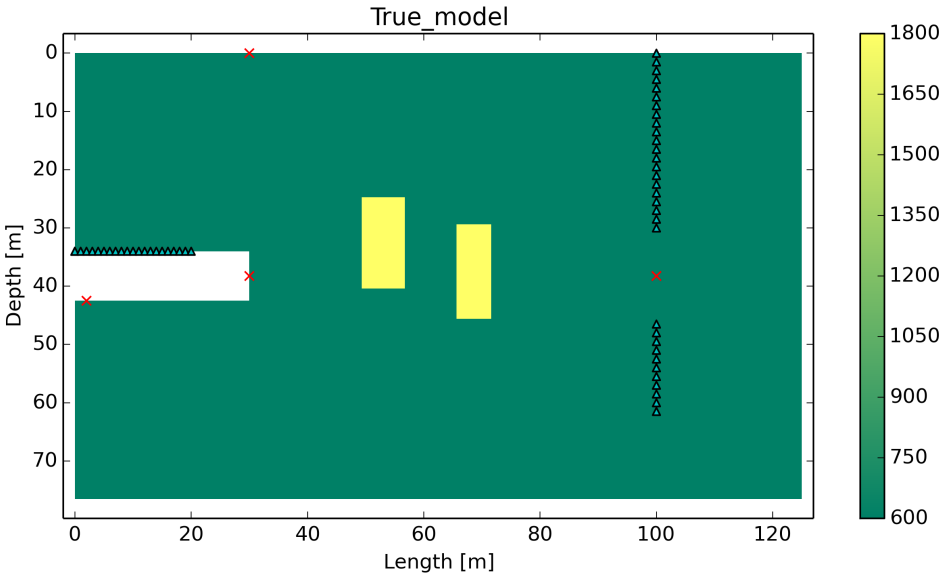


c)

Figure 5.8: Relative-amplitude seimograms of a) the background waveforms (without disturbance), b) the recorded waveforms (with disturbance), and c) the scattered waveforms generated from the disturbance scenario in Fig. 5.7



a)



b)

Figure 5.9: True models for FWI tests: a) Test 1: one-object disturbance, and b) Test 2: two-object disturbance. The filled colors represent S-wave velocities, crosses denote positions of the seismic sources and triangles the receivers

sharp-angled corners of the disturbance, the frequency of the Ricker source function is centered at 360 Hz. First inversion tests with a single source placed on the tunnel wall failed to reconstruct the expected disturbance (I did not show the results due to limited space). To improve the reconstruction quality, additional sources are placed at the tunnel face, in the borehole at a depth of the tunnel axis, and on the ground surface as shown in Fig. 5.9 to illuminate the disturbance object from different angles.

For FWI in time domain, direct use of the measured waveforms $d^{\text{obs}}(\boldsymbol{x}_j, t)$ and the modelled waveforms $h(\boldsymbol{m}, \boldsymbol{x}_j, t)$ to define data misfit in Eq. 5.10 can lead to highly multi-modal challenge in the topography of the misfit function because of the cycle-skipping effect. The cycle-skipping artifact is caused by wrong cycle of the modeled and measured waveforms being fitted together when the phase difference between the two waveforms is greater than one half of a cycle (Virieux & Operto (2009)). To reduce the possible cycle-skipping in the FWI process, the envelope misfit (Bozdağ et al. (2011)) is used instead of the waveform misfit Eq. 5.10. The envelope $D(\boldsymbol{x}_j, t)$ of the original waveform $d(\boldsymbol{x}_j, t)$ at a receiver point $\boldsymbol{x}_j$ is defined as

$$D(\boldsymbol{x}_j, t) = \sqrt{\Re(\tilde{d}(\boldsymbol{x}_j, t)) + \Im(\tilde{d}(\boldsymbol{x}_j, t))}, \tag{5.26}$$

where $\tilde{d}(\boldsymbol{x}_j, t)$ is the analytic signal constructed from the signal itself and its Hilbert transform $\mathcal{H}(\cdot)$:

$$\tilde{d}(\boldsymbol{x}_j, t) = d(\boldsymbol{x}_j, t) + \mathrm{i}\mathcal{H}(d(\boldsymbol{x}_j, t)), \tag{5.27}$$

with the Hilbert transform is defined using the Cauchy principle value P_v:

$$\mathcal{H}(d(\boldsymbol{x}_j, t)) = \frac{P_v}{\pi} \int_{-\infty}^{+\infty} \frac{d(\boldsymbol{x}_j, \tau)}{t - \tau} d\tau. \tag{5.28}$$

5.3.3.1 FWI test 1 — single disturbance

The model has a structured mesh of 3864 rectangular 4-node elements. By level-set parameterization of 5×6 bumps distributed at the ROI just ahead of the tunnel face, the dimension of the inverse problem is reduced to 61 only (30 for $\boldsymbol{\alpha}$, 30 for $\boldsymbol{\beta}$, and 1 for S-wave velocity of the disturbance v_s). P-wave velocity of the disturbance is assumed to be $v_p = \sqrt{3.36} \times v_s$. Other material properties are constant as shown in Table 5.1.

The UKF for time domain FWI in this application is configured with the following initial and uncertainty settings:

- Initial estimate: $\boldsymbol{m}_0 = \{\boldsymbol{\alpha}_0, \boldsymbol{\beta}_0, (v_s)_0\}$. In that, $\boldsymbol{\alpha}_0 = 3 \times \boldsymbol{I}(Nb, 1)$, $\boldsymbol{\beta}_0 = 2 \times \boldsymbol{I}(Nb, 1)$, $(v_s)_0 = 1000$.

- Initial estimation error covariance: $\boldsymbol{P}_0^m = \text{diag}(\boldsymbol{P}_0^\alpha, \boldsymbol{P}_0^\beta, P_0^{(v_s)_0})$. In that, $\boldsymbol{P}_0^\alpha = 4^2 \times \boldsymbol{I}(Nb, 1)$, $\boldsymbol{P}_0^\beta = 3^2 \times \boldsymbol{I}(Nb, 1)$, $P_0^{(v_s)_0} = 120^2$.
- Data and model uncertainty covariances: $\boldsymbol{R} = (0.1)^2 \times \boldsymbol{I}(r, r)$, $\boldsymbol{Q} = (0.1)^2 \times \boldsymbol{P}_0^m$.

The iterative updates of the model resulting from the inversion process are plotted in Fig. 5.10 for several selected iterative steps. It is evident that the reconstructed disturbance progresses to expected reference model as the iteration increases. After 25 iterations, the object has been adequately reconstructed. In details, the object is well reconstructed at its vertical interfaces because most interactions of the incident wavefronts happen there. The particulars of the horizontal interfaces are not in good agreement, and the wave velocity of the disturbance is slightly over-estimated compared to the reference model.

Measures of the total data misfit, which is the misfit summed over all receivers, indicates a sharp reduction in the misfit value. The envelope misfit value is reduced by nearly 70% after 25 iterations as shown in Fig. 5.11.

Comparison of the selected waveforms resulting from the initial model, the estimated model, and the synthetic reference model is plotted in Fig. 5.12. It can be seen that a considerable amount of the waveform mismatch in the initial model is reduced in the identified model.

Computational cost for the UKF-PaLS depends on the number of forward solves it takes to complete the iterative inversion steps. In this FWI test, in which the inverse problem's dimension is $n = 61$, it takes $(2 \times n + 1) \times 30 = 3690$ forward solves to complete the 30 iterative steps.

5.3.3.2 FWI test 2 — multiple disturbances

With the same initial and algorithmic settings, I continue to test the UKF-PaLS approach with solving the more challenging FWI test 2 (Fig. 5.9b). The iterative reconstructions displayed in Fig. 5.13 show that the final reconstruction can locate the overall position of the disturbance but fails to detail the object's geometry, i.e. to delineate the two discontinuous geometries. The difficulty of using FWI in this scenario is attributed to elastic wave transmission through multiple heterogeneities and elastic wave multiple refractions in the space between the two distinct objects. The total data misfit showed in Fig. 5.14 can not be reduced by more than 67% of the initial misfit value.

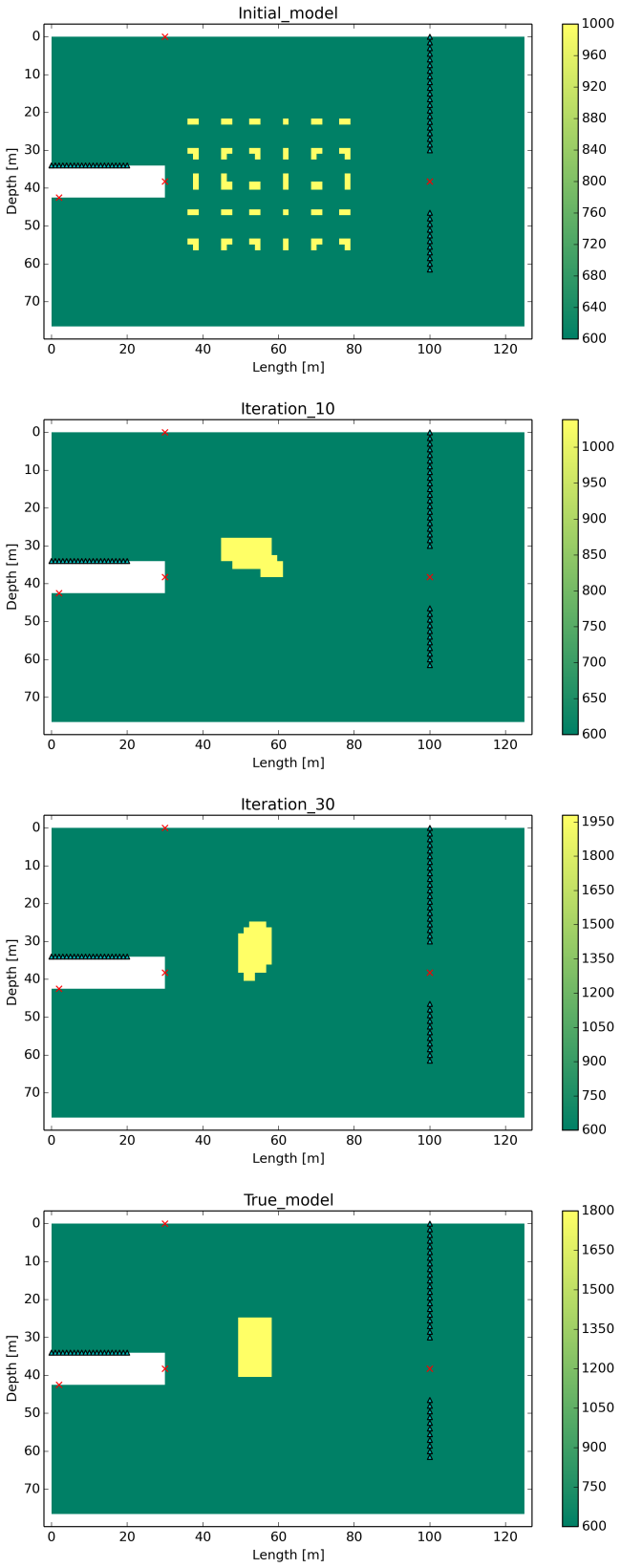


Figure 5.10: The initial model, the reconstructions after 10 and 30 iterations, and the true model resulting from the FWI test 1

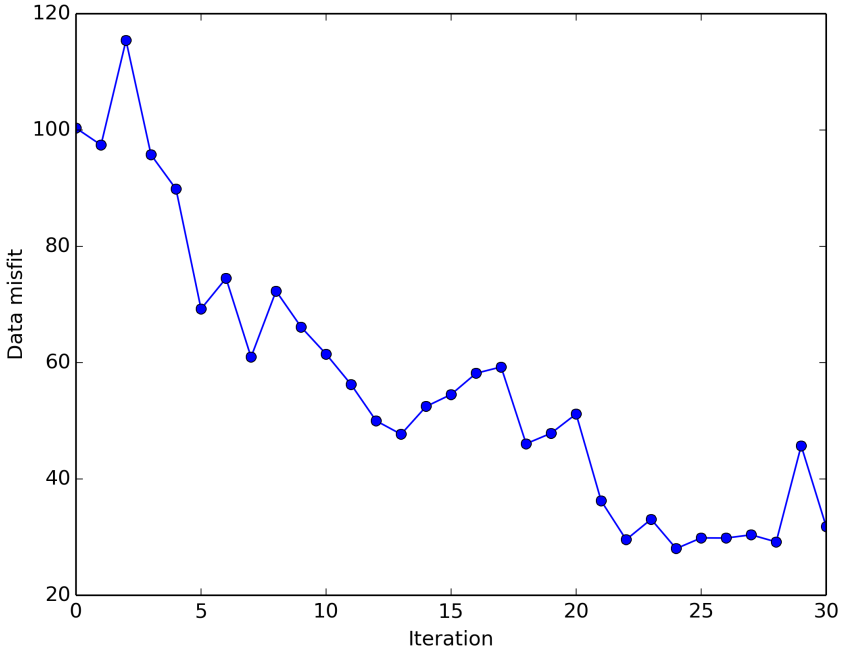


Figure 5.11: Data misfit along UKF iterations resulting from the FWI test 1

The waveforms resulting from the reconstructed model show reduced mismatch compared the waveforms resulting from the initial model; however, 23% mismatch still remains compared to the initial model. Fig. 5.15 visualizes the amount of fit to the target model achieved the waveform inversion compared to that of the initial model at selective observed positions. The FWI in this test consumes $(2 \times 61 + 1) \times 40 = 4920$ forward solves as the UKF-PaLS inversion process is run in 40 iterations.

5.3.3.3 Discussion

For the successful application of FWI in tunneling practice, one needs above all a very efficient inversion method besides a high-fidelity forward solver of the wave propagation problem and a complete set of high signal-to-noise ratio measurement data. The choice of an efficient and flexible parameterization method like the level-set representation helps to achieve accelerated inversion speed and quality reconstruction.

After all, the proposed approach is associated with some limitations. If there are some objects located outside of the ROI (including the case of geological layer change), the FWI may not be able to reconstruct the object inside the ROI. This limitation can be remedied by either expanding the fixed spatial allocations of the RBFs or allowing the

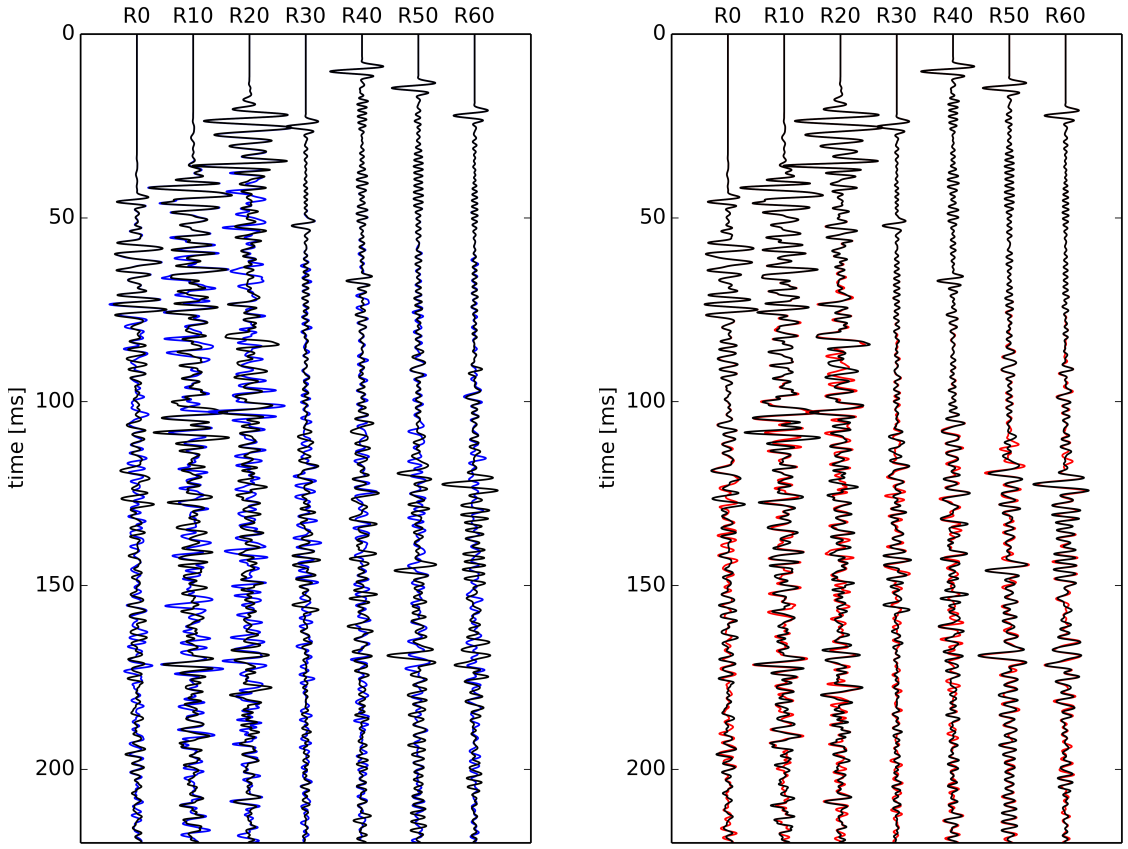


Figure 5.12: Selected waveforms resulting from the reference model (black line), the starting model (blue line), and the identified model (red line) resulting from the FWI test 1. Receivers 0, 10, and 20 are placed on the tunnel wall; receives 30, 40, 50, and 60 are placed in the borehole

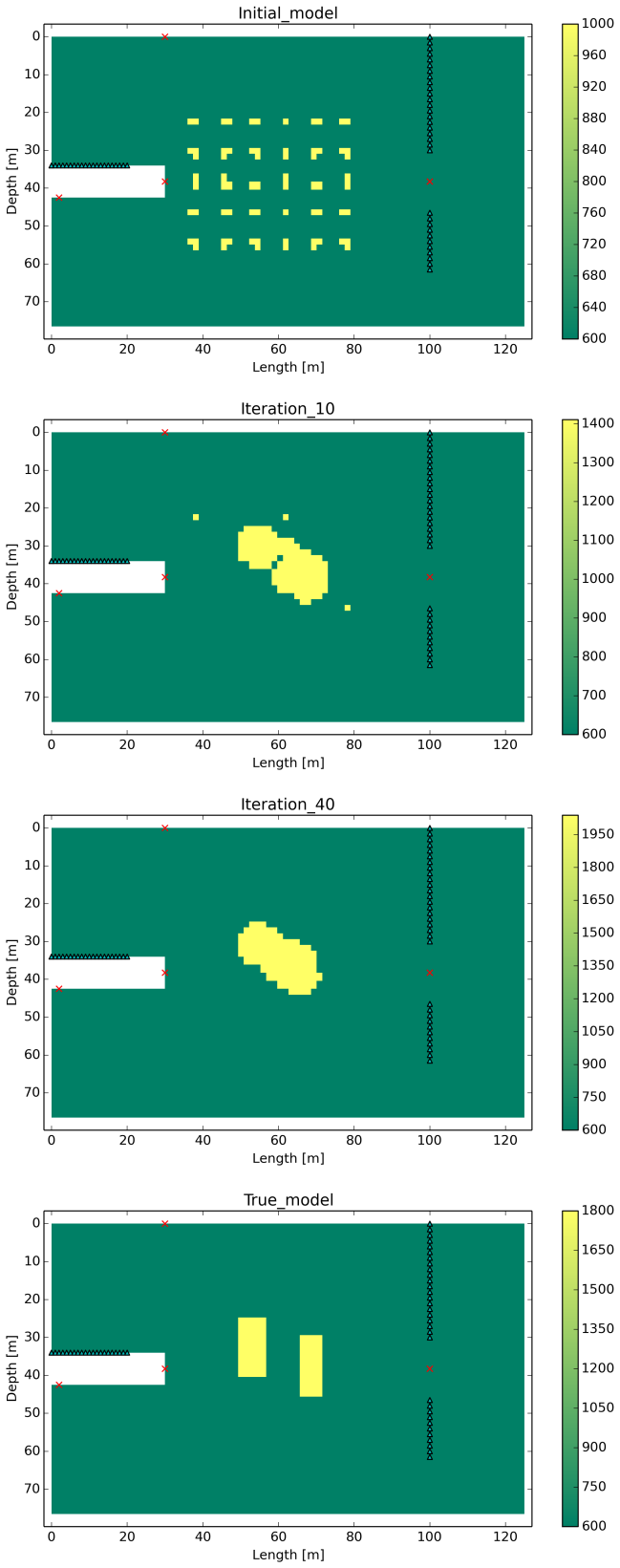


Figure 5.13: The initial model, the reconstructions after 10 and 40 iterations, and the true model resulting from the FWI test 2

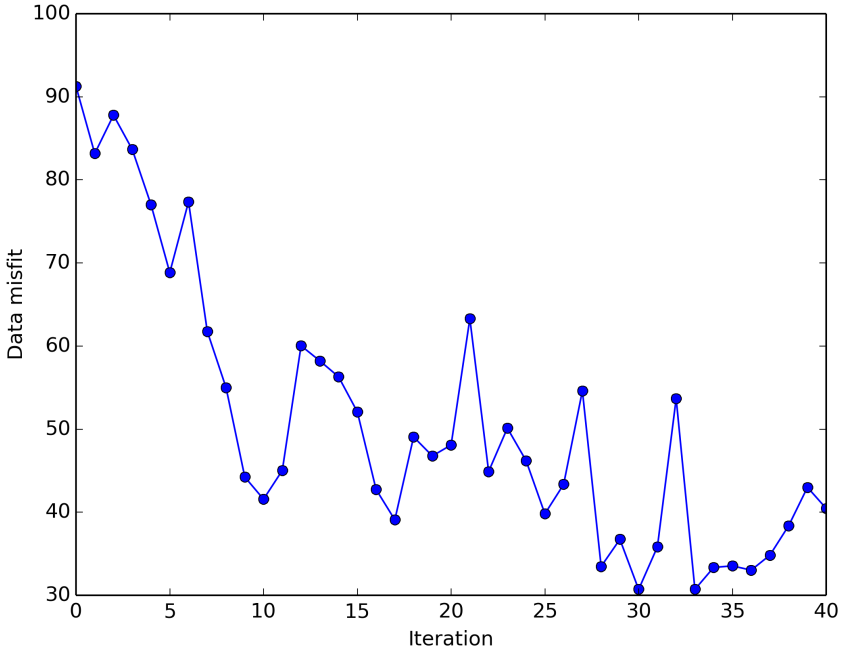


Figure 5.14: Data misfit along UKF iterations resulting from the FWI test 2

centers of all RBFs to be updated in the inversion process. However, this expansion comes with much increasing computational demand. Also, this FWI approach does not invert for heterogeneity of the wave velocities, e.g. gradient material change in a particular direction, within the disturbance as material properties are assigned to the disturbance uniformly.

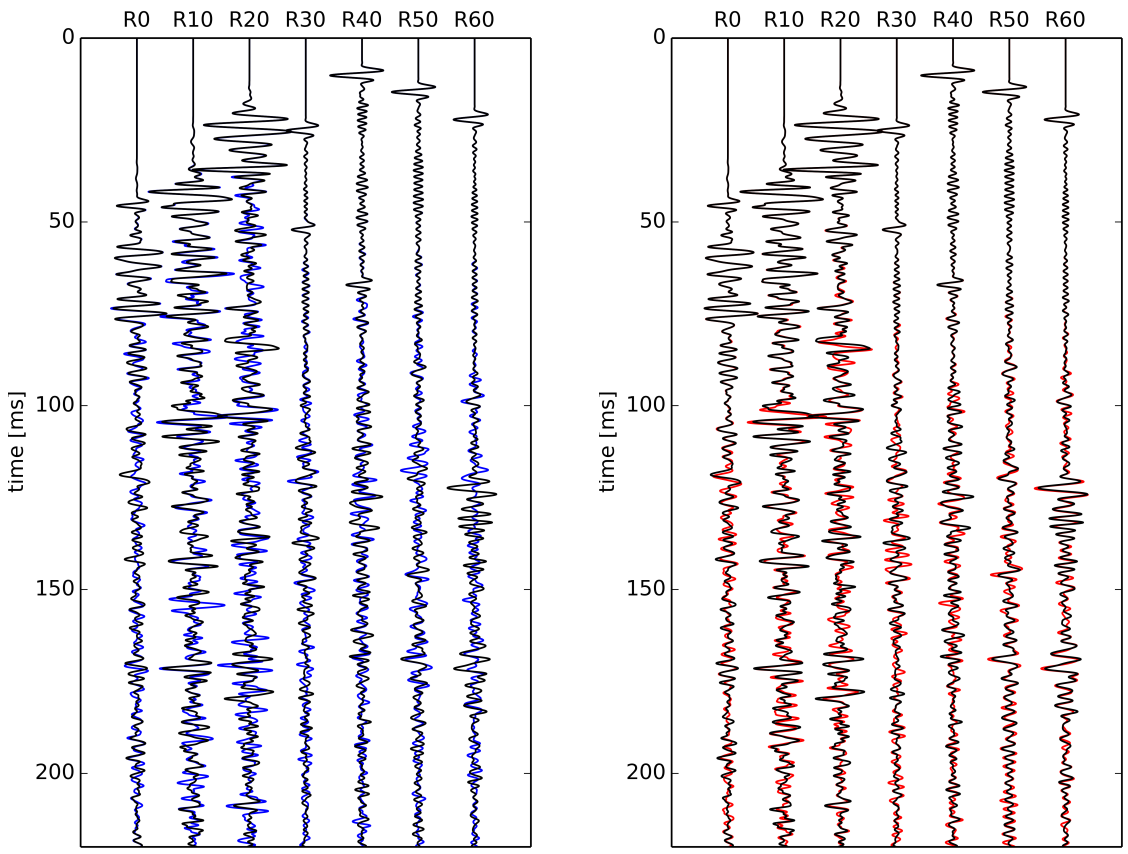


Figure 5.15: Selected waveforms resulting from the reference model (black line), the starting model (blue line), and the identified model (red line) resulting from the FWI test 2

6 Conclusions and outlook

6.1 Conclusions

Nonlinear Kalman filters and particle filter for soil parameters identification

Overall, the EKF and SPKF presented in this work are very capable of doing model calibration for geomechanical modeling in mechanized tunneling. Fast convergence of the parameters being identified and ability to tolerate measurement uncertainty are the main advantages of the two Kalman filter algorithms. One other prominent advantage is that the choice of initial model parameters must not be very close to the true parameter set as for other gradient-based optimization methods. These nonlinear Kalman filters are potential to apply for calibrating model parameters to modeling geotechnical structures other than tunneling as well.

In particular, the use of the EKF requires approximations of the derivatives of measurement data with respect to the model parameters due to identification. These approximations may be inaccurate for highly nonlinear and non-smooth FE responses. On the other hand, the use of SPKF for parameter identification frees us from doing this brute-force numerical calculation of the derivatives. But, in order to have this great advantage, more cost is paid for forward calculation of the model: $2n + 1$ runs for the SPKF compared with $n + 1$ runs for the EKF in every iteration, with n number of parameters. However, the number of forward simulations required is still much less than those required by other global optimization algorithms such as metaheuristic and generic Monte Carlo optimization methods.

Also, particle filter based data assimilation is shown to be suitable for identification of unknown constitutive parameters in modeling of mechanized tunnel excavation. The assimilated results in the form of probability distributions are very convenient for judgment and utilization of the identified parameters for further prediction and analysis of the forthcoming tunnel excavation steps. The advantage of using the particle filter as a mechanism

to update the posterior distributions is that it can intrinsically work with nonlinear models and non-Gaussian densities representing model parameters and uncertainties. Besides, this method is very noise tolerant as it shows high accuracy of the identified parameters even with a level of observation noise.

Although noisy information can be incorporated into the filters in the form of covariance of zero-mean Gaussian distribution, the authors have observed that a large amount of measurement uncertainty can cause great difficulty in achieving the true parameters set especially for the parameters that contribute low sensitivity to the model outputs.

Careful design of measurement campaigns in tunneling and other geotechnical projects will help provide Kalman filters inverse analysis with quality data. As the tunnel boring machine advances, the nonlinear Kalman filtering schemes introduced in this work can be continuously employed for updating the relevant hidden parameters about the soils and rocks surrounded the TBM.

Global optimization approach for inversion of the tunnel seismic waves

An improvement of the SA algorithm is proposed to solve tunnel seismic waveform inversion whose purpose is to develop an efficient method for prediction ahead of underground tunnel face. The multimodal misfit functional can be effectively searched for the global minimum by combining the exploration and exploitation capabilities of SA algorithm and UKF procedure respectively. The benefits of using waveform inversion by the proposed derivative-free global optimization include i) information of the full waveforms can be utilized which promises high-resolution of the imaged geological structure ahead of the tunnel face, ii) no good initial model and no gradient calculations are required to run waveform inversion, and iii) not only the global minimum model is found but also other dominant local minima which may help in expert judgment of the likely geological situations.

Overall, the UHSA algorithm does not require so many sample points in the parameter space compared with the standard SA algorithm because 'bad' random sample points can be improved locally by the UKF. This local improvement is essential to accelerate the convergence of the proposed algorithm.

Although the demonstrated example of the tunnel seismic model is very simply parameterized, inversion based on modeling of tunnel seismic wave propagation with a standard PC still takes considerable calculation time. A high power computing facility is likely to allow the application of this method to the characterization of more complicated geological scenarios, or even for more realistic waves propagation in 3-dimensional space.

The proposed method in this work can benefit from future development of the SA and UKF algorithms. Within the current standard implementation of the two constituted algorithmic components, the method can perform better by improving local convergence behavior of the UKF. For example, measurement data can be partitioned into several data blocks, then in each iterative run of the UKF, the data blocks are successively fed to the filter. This data partition scheme may work well because with little data used at the beginning of the filtering ease convergence. Later data increments are expected to guide the solution to the global minimum.

Full-waveform inversion supported by parametric level-set representation

This work presents an inversion scheme enhanced by the PaLS parameterization of the geometry to achieve effective and fast inversion of multiple arbitrarily shaped disturbances ahead of the tunnel face. This parameterization technique is very well suited for reconstructing the object of complex geometries while allows for implicit handling of the object's interfacial boundary on a fixed computational mesh. The use of the UKF as an inversion method results in an efficient derivative-free inversion process. The proposed combination of parameterization and inversion methodology for FWI — the UKF-PaLS — is effective in reconstruction capability and computing requirements. As computational time is one of the largest difficulties preventing the application of FWI in tunneling practice, the use of the proposed inversion scheme can help to achieve reliable reconstruction of arbitrary objects while keeping the computational time acceptably low.

More validations with experimental or in-situ data are important to prove the robustness of the method against measurement errors. One other important work is to take into account the wave propagation simulation in 3D with a detailed study of the influence of the excavation damaged zone close to the tunnel wall and the hydrostatic confining pressure for deep tunneling to build the high fidelity initial model.

Fast and quality reconstruction results achieved using only a few seismic sources and receivers are encouraging to use FWI in tunneling practice and is especially appropriate for tunneling projects in high-risk ground conditions where repeated data acquisition and ground reconnaissance are required as the TBM advances.

6.2 Outlook

6.2.1 Integration into a tunnel information system

It can be foreseen that inverse problems encountered in tunneling and geotechnical site investigation continue to benefit from the booming developments in mathematical inverse theory, the increasing availability of high power computing facilities, and sensing technologies (fiber optical sensor and others) as well as the new real-time monitoring concepts.

From the computational viewpoint, advanced numerical approximations and material models being continuously developed increase the capability to mimic physical phenomena using computer simulations. This ability is very decisive for the model-based identification realized by inverse analysis. The second computational advance that is vital for this identification methodology lends itself to intensive developments of efficient computational inversion methods based either on deterministic optimization or probabilistic sampling approach or a combination of the two.

When dealing with data measured on construction sites, sophisticated data processing tools are needed to increase signal-to-noise ratios and to filter out the unwanted frequency content in data.

An ultimate system to facilitate data exchange and to automate the modeling and updating of the geological conditions are desired for modern tunneling. Such a system has started to exist in the name of tunnel information modeling (TIM). Reconnaissance can be integrated as an advanced module in such a TIM system.

6.2.2 Validation against laboratory measurements

Validation plays a crucial role in developments of numerical methods. The most convincing performance evaluations are to validate the inversion results again in-situ measurements. However, those real measurement data either are often not available, incomplete, or the in-situ situation is not well documented. Therefore, it is very helpful to carry out small-scale laboratory controlled experiments by which it allows to set up the test scenarios of interest. Such an experiment set-up has been developed in the framework of the sub-project A2 *Development of effective concepts for tunnel reconnaissance using acoustic methods* within the Collaborative Research Center SFB 837 'Interaction model-

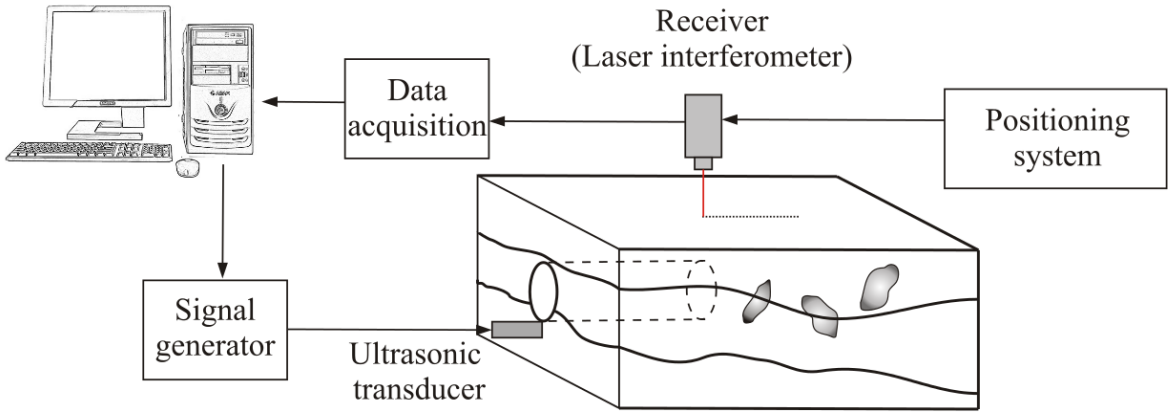


Figure 6.1: Sketch of the small-scale wave propagation experiment

ing in mechanized tunneling' at the Ruhr-University Bochum. The small-scale laboratory set-up for tunnel seismic experiments is sketched in Fig. 6.1.

Because of the small scales of the laboratory experiments, the chosen frequencies are in the range of ultrasound waves. The below instruments are used for ultrasound excitations, receivers, and acquisition of the ultrasound waveforms.

- Laser interferometer: Bossa Nova Quartet 500mW, frequency range 100 kHz – 50 MHz, 24-bit resolution, stand-off distance 100mm (DOF 5.5mm), aperture diameter 50mm, spot size 60 micro (wavelength 532 nm). The device can measure sub-nanometer displacements.
- Ultrasonic transducers: Various low-frequency (in the range of 50 kHz to 600 kHz) piezoelectric transducers manufactured by Karl Deutsch, General Electrics, Olympus are acquired. Arbitrary source function is loaded to the signal generator (Keysight 33500B Series) and then is amplified by the Amplifier (E&I Model 1040L, 53dB fixed gain) before feeding to the ultrasonic transducer.
- A self-built scanning system placed on an isolation table (Newport S-2000 Series).
- Acquisition hardware (NI PCI-5922) with a user interface built with Labview.

With the data acquisition set-up in Fig. 6.1, the laser sensor is used to scan in a line on the specimen's top surface. For visual investigation of the effect of the fault in front of the tunnel face, two corresponding samples of sandstone are prepared: one without fault and one with an inclined cut mimicking the geological dipped fault or layer change as shown in Fig. 6.2. I make surface scans of a length of 120 mm with an interval of 2 mm. Using a 100 kHz ultrasonic transducer, a Ricker wavelet source is transmitted into the

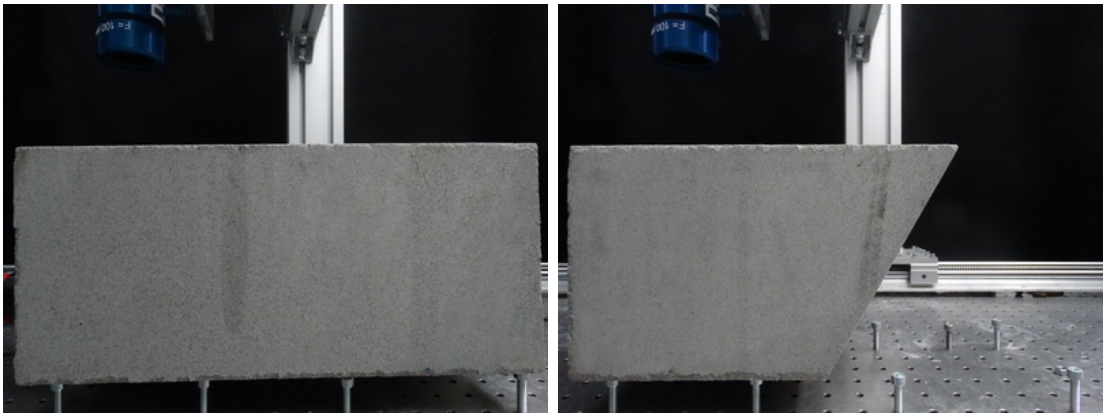


Figure 6.2: Small-scale specimens without fault (left) and with a dipped fault (right)

specimens. The two experiments are carried out in atmospheric condition under the same contact condition between the ultrasonic transducer and the specimen.

The measured vertical displacement waveforms for the finite no-fault sandstone sample is is shown in Fig. 6.3. Fig. 6.4 shows the resulting seismogram measured on the faulted sandstone specimen. It can be seen in both the measured seismograms resemble in early direct arrivals, but the later arrivals differ substantially between the two scenarios as the result of reflections from the fault surface.

I note that before the ultrasonic data can be used in waveform inversion, it is important to match the modeled waveforms to the measured waveforms on a simple test specimen such as a homogeneous cylinder of aluminum, granite, or sandstone. It is needed to characterize the actual source wavelet coming out of the piezoelectric transducers because the transducer-specimen coupling condition and the characteristics of the transducer itself modify the input source wavelet.

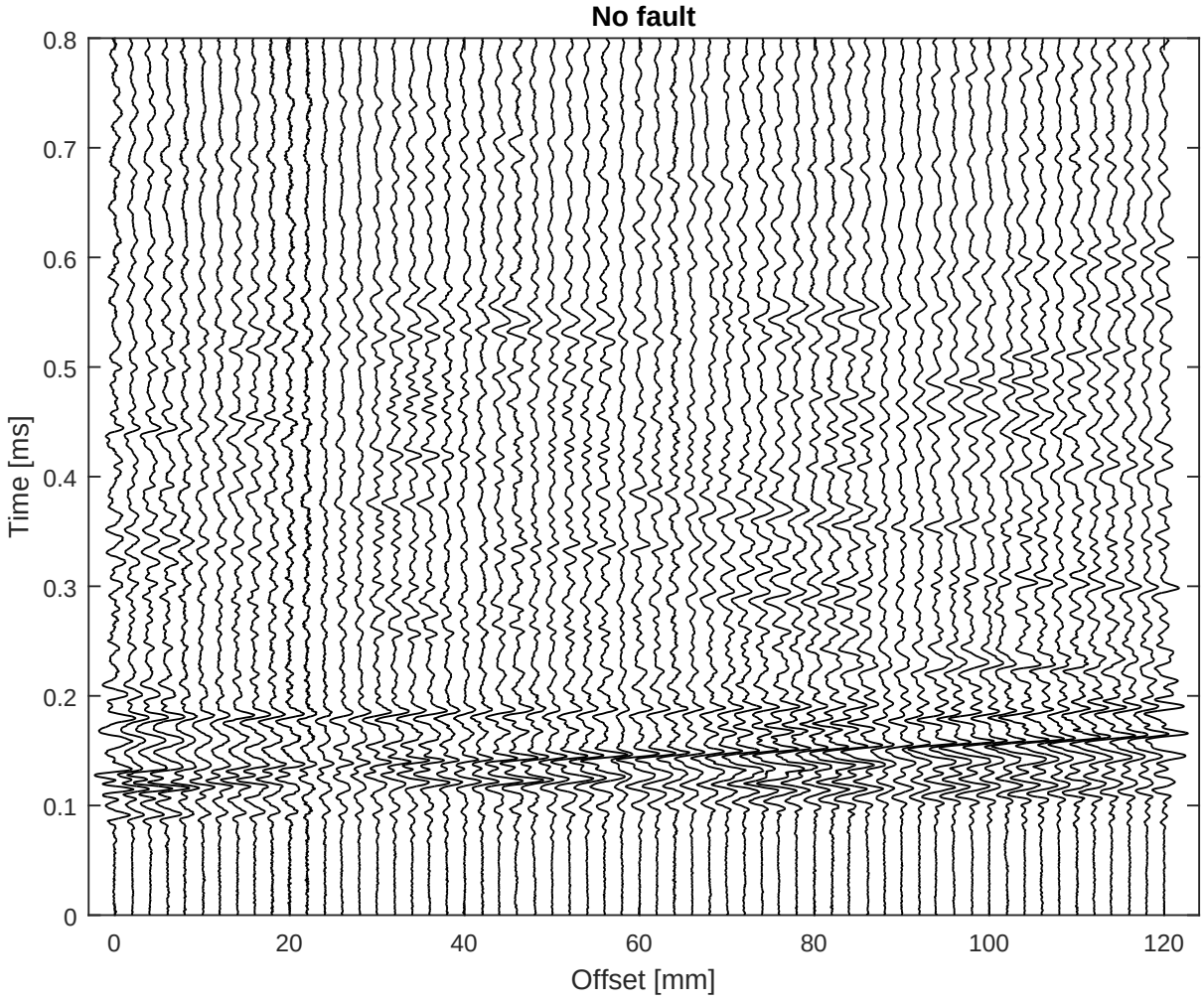


Figure 6.3: Measured seismogram on the no-fault sandstone specimen

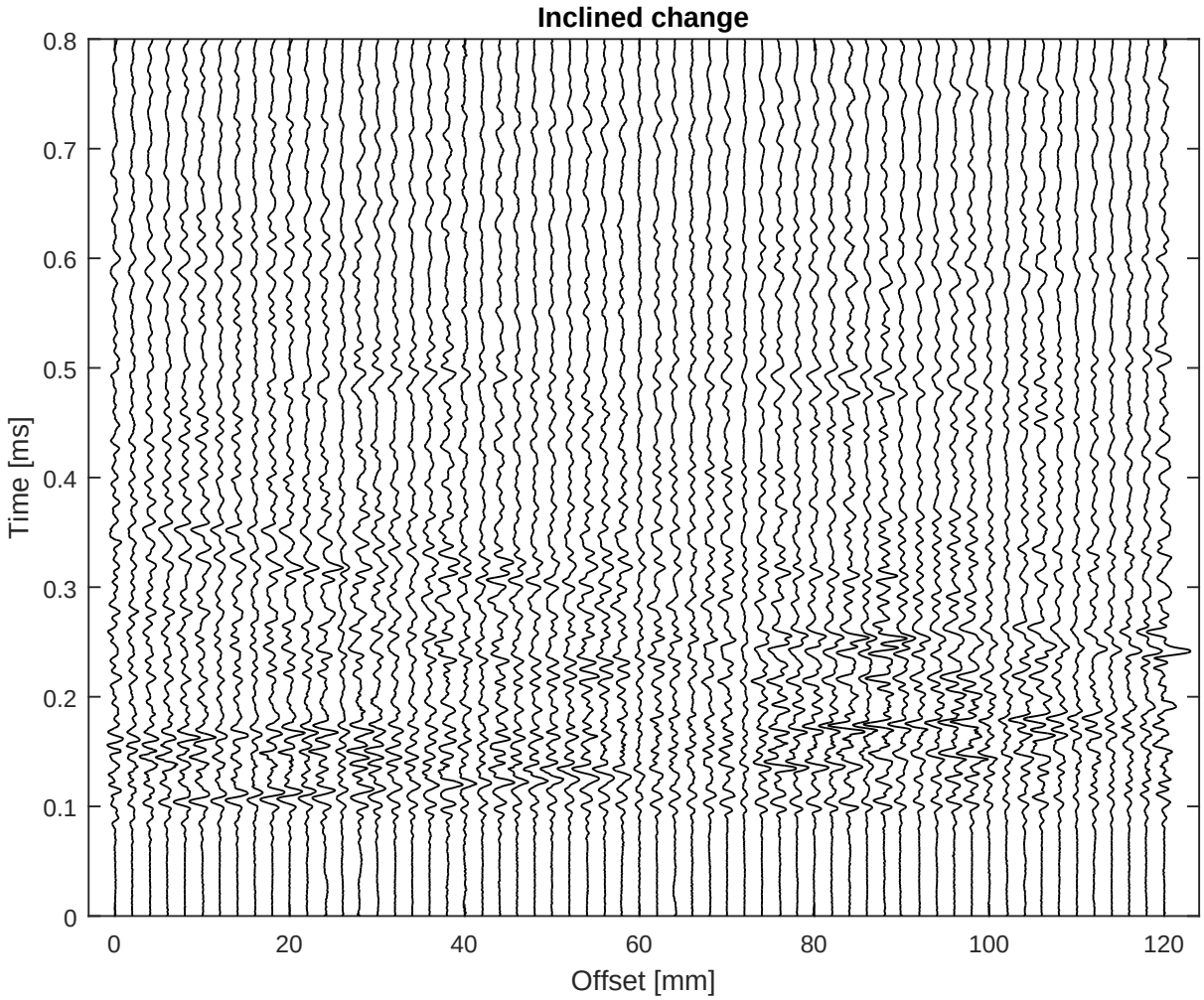


Figure 6.4: Measured seismogram on the sandstone specimen with the inclined layer change

Bibliography

Aghasi, A., Kilmer, M. & Miller, E. L. (2011), 'Parametric level set methods for inverse problems', *SIAM Journal on Imaging Sciences* **4**(2), 618–650.

Aki, K. & Richards, P. G. (2002), *Quantitative seismology*, University Science Books.

Ashida, Y. (2001), 'Seismic imaging ahead of a tunnel face with three-component geophones', *International Journal of Rock Mechanics and Mining Sciences* **38**(6), 823–831.

Attarian, A., Batzel, J. J., Matzuka, B. & Tran, H. (2013), Application of the unscented Kalman filtering to parameter estimation, *in* 'Mathematical Modeling and Validation in Physiology', Springer, pp. 75–88.

Balling, R. J. (1991), 'Optimal steel frame design by simulated annealing', *Journal of Structural Engineering* **117**(6), 1780–1795.

Bell, B. M. & Cathey, F. W. (1993), 'The iterated Kalman filter update as a Gauss-Newton method', *IEEE Transactions on Automatic Control* **38**(2), 294–297.

Bellino, A., Garibaldi, L. & Godio, A. (2013), 'An automatic method for data processing of seismic data in tunneling', *J. of Applied Geophysics* **98**, 243–253.

Benecke, N., Dombrowski, B. A. & Lehmann, B. (2008), TRUST — exploration ahead of the tunnel face for reducing tunneling risks and supporting decision-making, *in* 'World Tunnel Congress', pp. 1124–1129.

Bertsekas, D. P. (1996), 'Incremental least squares methods and the extended Kalman filter', *SIAM Journal on Optimization* **6**(3), 807–822.

Blum, C. & Roli, A. (2003), 'Metaheuristics in combinatorial optimization: Overview and conceptual comparison', *ACM Computing Surveys (CSUR)* **35**(3), 268–308.

Bobet, A. (2001), 'Analytical solutions for shallow tunnels in saturated ground', *Journal of Engineering Mechanics* **127**(12), 1258–1266.

Bohlen, T., Lorang, U., Rabbel, W., Müller, C., Giese, R., Lüth, S. & Jetschny, S. (2007), 'Rayleigh-to-shear wave conversion at the tunnel face-from 3D-FD modeling to ahead-of-drill exploration', *Geophysics* **72**(6), T67–T79.

Borm, G., Giese, R., Klose, C., Mielitz, S., Otto, P. & Bohlen, T. (2003), ISIS integrated seismic imaging system for the geological prediction ahead of underground construction, *in* 'EAGE the 65th Conference and Exhibition'.

Bozdağ, E., Trampert, J. & Tromp, J. (2011), 'Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements', *Geophysical Journal International* **185**(2), 845–870.

Bretaudeau, F., Brossier, R., Leparoux, D., Abraham, O. & Virieux, J. (2013), '2D elastic full-waveform imaging of the near-surface: application to synthetic and physical modelling data sets', *Near Surface Geophysics* **11**(3), 307–316.

Brückl, E., Chwatal, W., Dölzlmüller, J. & Jöbstl, W. (2001), 'A study of the application of VSP to exploration ahead of a tunnel', *International Journal of Rock Mechanics and Mining Sciences* **38**(6), 833–841.

Brückl, E., Chwatal, W., Mertl, S. & Radinger, A. (2008), 'Exploration ahead of a tunnel face by TSWD–tunnel seismic while drilling', *Geomechanics and Tunnelling* **1**(5), 460–465.

Chatzi, E. N. & Smyth, A. W. (2009), 'The unscented Kalman filter and particle filter methods for nonlinear structural system identification with non-collocated heterogeneous sensing', *Structural Control and Health Monitoring* **16**(1), 99–123.

Chatzi, E. N. & Smyth, A. W. (2013), 'Particle filter scheme with mutation for the estimation of time-invariant parameters in structural health monitoring applications', *Structural Control and Health Monitoring* **20**(7), 1081–1095.

Cheng, F., Liu, J., Qu, N., Mao, M. & Zhou, L. (2014), 'Two-dimensional pre-stack reverse time imaging based on tunnel space', *Journal of Applied Geophysics* **104**, 106–113.

Chi, B., Dong, L. & Liu, Y. (2014), 'Full waveform inversion method using envelope objective function without low frequency data', *Journal of Applied Geophysics* **109**, 36–46.

Cividini, A., Maier, G. & Nappi, A. (1983), 'Parameter estimation of a static geotechnical model using a Bayes' approach', *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* **20(5)**, 215–226.

Dai, S., Wuttke, F. & Santamarina, J. C. (2012), 'Coda wave analysis to monitor processes in soils', *Journal of Geotechnical and Geoenvironmental engineering* **139**(9), 1504–1511.

de Oliveira, M. A. (2012), 'An application of neural networks trained with Kalman filter variants (EKF and UKF) to heteroscedastic time series forecasting', *Applied Mathematical Sciences* **6**(74), 3675–3686.

Dickmann, T. & Sander, B. (1996), 'Drivage concurrent tunnel seismic prediction', *Felsbau-rock and Soil Engineering* **14**(1), 406–411.

Do, N.-A., Dias, D., Oreste, P. & Djeran-Maigre, I. (2014), 'Three-dimensional numerical simulation of a mechanized twin tunnels in soft ground', *Tunnelling and Underground Space Technology* **42**, 40–51.

Doucet, A., Godsill, S. & Andrieu, C. (2000), 'On sequential Monte Carlo sampling methods for Bayesian filtering', *Statistics and Computing* **10**(3), 197–208.

Essen, K., Bohlen, T., Friederich, W. & Meier, T. (2007), 'Modelling of Rayleigh-type seam waves in disturbed coal seams and around a coal mine roadway', *Geophysical Journal International* **170**(2), 511–526.

Evensen, G. (2003), 'The Ensemble Kalman Filter: theoretical formulation and practical implementation', *Ocean Dynamics* **53**, 343–367.

Forrester, A., Sobester, A. & Keane, A. (2008), *Engineering design via surrogate modelling: a practical guide*, John Wiley & Sons.

Gehrig, M., Kassel, A. & Lorenz, K. (2010), 'Advance exploration accompanying the drive: mixshield with vision', *Tunnel* **3**, 14–29.

Gelb, A. (1974), *Applied optimal estimation*, MIT Press, Cambridge, Massachusetts.

Geuzaine, C. & Remacle, J.-F. (2009), 'Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities', *International Journal for Numerical Methods in Engineering* **79**(11), 1309–1331.

Gharti, H. N., Oye, V., Komatitsch, D. & Tromp, J. (2012), 'Simulation of multistage excavation based on a 3d spectral-element method', *Computers & Structures* **100**, 54–69.

Gioda, G. & Sakurai, S. (1987), 'Back analysis procedures for the interpretation of field measurements in geomechanics', *International Journal for Numerical and Analytical Methods in Geomechanics* **11**, 555–583.

Gordon, N. J., Salmond, D. J. & Smith, A. F. (1993), Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *in* 'IEE Proceedings F (Radar and Signal Processing)', IET, pp. 107–113.

Grosse, C. U. & Ohtsu, M. (2008), *Acoustic emission testing*, Springer Science & Business Media.

Halliday, D. F., Curtis, A., Robertsson, J. O. & van Manen, D.-J. (2007), 'Interferometric surface-wave isolation and removal', *Geophysics* **72**(5), A69–A73.

Haykin, S. S., Haykin, S. S. & Haykin, S. S. (2001), *Kalman filtering and neural networks*, Wiley Online Library.

Ho, Y. C. & Lee, R. (1964), 'A Bayesian approach to problems in stochastic estimation and control', *IEEE Transactions on Automatic Control* **9**, 333 – 339.

Hommels, A., Murakami, A. & Nishimura, S.-I. (2009), 'A comparison of the ensemble Kalman filter with the unscented Kalman filter: application to the construction of a road embankment', *Geotechniek* **13**(1), 52.

Hoshiya, M. & Sutoh, A. (1993), 'Kalman filter - finite element method in identification', *Journal of Engineering Mechanics* **119(2)**, 197–210.

Inazaki, T., Isahai, H., Kawamura, S., Kurahashi, T. & Hayashi, H. (1999), 'Stepwise application of horizontal seismic profiling for tunnel prediction ahead of the face', *The Leading Edge* **18**(12), 1429–1431.

Ingber, L. (1993), 'Simulated annealing: Practice versus theory', *Mathematical and Computer Modelling* **18**(11), 29–57.

Jazwinski, A. H. (1970), *Stochastic processes and filtering theory*, Academic Press.

Jeong, I.-K. & Lee, J.-J. (1996), 'Adaptive simulated annealing genetic algorithm for system identification', *Engineering Applications of Artificial Intelligence* **9**(5), 523–532.

Jetschny, S., Bohlen, T. & De Nil, D. (2010), 'On the propagation characteristics of tunnel surface-waves for seismic prediction', *Geophysical Prospecting* **58**(2), 245–256.

Jetschny, S., Bohlen, T. & Kurzmann, A. (2011), 'Seismic prediction of geological structures ahead of the tunnel using tunnel surface waves', *Geophysical Prospecting* **59**(5), 934–946.

Jiao, Y.-Y., Tian, H.-N., Liu, Y.-Z., Mei, R.-W. & Li, H.-B. (2015), 'Prediction of tunneling hazardous geological zones using the active seismic approach', *Near Surface Geophysics* **13**(4), 333–342.

Julier, S. J. & Uhlmann, J. K. (1997), A new extension of the Kalman filter to nonlinear systems, *in* 'Proc. AeroSense: 11th Int. Symp. Aerospace/Defense Sensing, Simulation and Controls', pp. 182–193.

Julier, S. J. & Uhlmann, J. K. (2004), 'Unscented filtering and nonlinear estimation', *Proceedings of the IEEE* **92**(3), 401–422.

Kalman, R. E. (1960), 'A new approach to linear filtering and prediction problems', *Trans. ASME Ser. D: J. Basic Eng.* **82(1)**, 35–45.

Kandepu, R., Imsland, L. & Foss, B. A. (2008), Constrained state estimation using the unscented Kalman filter, *in* 'Proceedings of the 16th Mediterranean Conference on Control and Automation', Citeseer, IEEE, pp. 1453–1458.

Kasper, T. & Meschke, G. (2004), 'A 3D finite element simulation model for TBM tunnelling in soft ground', *International journal for numerical and analytical methods in geomechanics* **28**(14), 1441–1460.

Kirkpatrick, S., Vecchi, M. et al. (1983), 'Optimization by simmulated annealing', *science* **220**(4598), 671–680.

Knabe, T., Datcheva, M., Lahmer, T., Cotecchia, F. & Schanz, T. (2013), 'Identification of constitutive parameters of soil using an optimization strategy and statistical analysis', *Computers and Geotechnics* **49**, 143–157.

Kneib, G., Kassel, A. & Lorenz, K. (2000), 'Automatic seismic prediction ahead of the tunnel boring machine', *First Break* **18**(7), 295–302.

Kneib, G. & Leykam, A. (2004), Finite difference modelling for tunnel seismology, *in* '10th European Meeting of Environmental and Engineering Geophysics'.

Komatitsch, D. & Tromp, J. (1999), 'Introduction to the spectral element method for three-dimensional seismic wave propagation', *Geophysical Journal International* **139**(3), 806–822.

Komatitsch, D., Tsuboi, S. & Tromp, J. (2005), 'The spectral-element method in seismology', *Geophysical Monograph Series* **157**, 205–227.

Komatitsch, D. & Vilotte, J.-P. (1998), 'The spectral element method: An efficient tool to simulate the seismic response of 2D and 3D geological structures', *Bulletin of the seismological society of America* **88**(2), 368–392.

Krauß, F., Giese, R., Alexandrakis, C. & Buske, S. (2014), 'Seismic travel-time and attenuation tomography to characterize the excavation damaged zone and the surrounding rock mass of a newly excavated ramp and chamber', *International Journal of Rock Mechanics and Mining Sciences* **70**, 524–532.

Lambrecht, L. & Friederich, W. (2013), Simulation of seismic wave propagation for reconnaissance in mechanized tunneling using spectral element and nodal discontinuous Galerkin method, *in* G. Meschke, J. Eberhardssteiner, T. Schanz, K. Soga & M. Thewes, eds, 'EURO:TUN 2013 Compuatational Methods in Tunneling and Subsurface Engineering', Aedificatio Publishers, pp. 769–777.

Lambrughi, A., Rodríguez, L. M. & Castellanza, R. (2012), 'Development and validation of a 3D numerical model for TBM–EPB mechanised excavations', *Computers and Geotechnics* **40**, 97–113.

Leca, E. & New, B. (2007), 'Settlements induced by tunneling in soft ground', *Tunnelling and Underground Space Technology* **22**(2), 119–149.

Ledesma, A., Gens, A. & Alonso, E. E. (1996), 'Parameter and variance estimation in geotechnical backanalysis using prior information', *International Journal for Numerical and Analytical Methods in Geomechanics* **20**(2), 119–141.

Levasseur, S., Malecot, Y., Boulon, M. & Flavigny, E. (2008), 'Soil parameter identification using a genetic algorithm', *International Journal for Numerical and Analytical Methods in Geomechanics* **32**(2), 189–213.

Loganathan, N. & Poulos, H. (1998), 'Analytical prediction for tunneling-induced ground movements in clays', *Journal of Geotechnical and Geoenvironmental Engineering* **124**(9), 846–856.

Lüth, S., Bohlen, T., Giese, R., Heider, S., Hock, S., Jetschny, S., Polom, U., Wadas, S. & Rechlin, A. (2014), Seismic tomography and monitoring in underground structures: Developments in the Freiberg Reiche Zeche underground lab (Freiberg, Germany) and

their application in underground construction (SOUND), *in* 'Tomography of the Earthś Crust: From Geophysical Sounding to Real-Time Monitoring', Springer, pp. 115–133.

Mariani, S. & Ghisi, A. (2007), 'Unscented Kalman filtering for nonlinear structural dynamics', *Nonlinear Dynamics* **49**(1-2), 131–150.

Martin, J., Wilcox, L. C., Burstedde, C. & Ghattas, O. (2012), 'A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion', *SIAM Journal on Scientific Computing* **34**(3), A1460–A1487.

Martin, O. C. & Otto, S. W. (1996), 'Combining simulated annealing with local search heuristics', *Annals of Operations Research* **63**(1), 57–75.

Meier, J., Schaedler, W., Borgatti, L., Corsini, A. & Schanz, T. (2008), 'Inverse parameter identification technique using PSO algorithm applied to geotechnical modeling', *Journal of Artificial Evolution and Applications* **2008**, 3.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953), 'Equation of state calculations by fast computing machines', *The journal of chemical physics* **21**(6), 1087–1092.

Miller, G. & Pursey, H. (1955), On the partition of energy between elastic waves in a semi-infinite solid, *in* 'Proceedings of the Royal Society of London a: mathematical, physical and engineering sciences', The Royal Society, pp. 55–69.

Miranda, T., Dias, D., Eclaircy-Caudron, S., Correia, A. G. & Costa, L. (2011), 'Back analysis of geomechanical parameters by optimisation of a 3D model of an underground structure', *Tunnelling and Underground Space Technology* **26**(6), 659 – 673.

Mooney, M. A., Walter, B. & Frenzel, C. (2012), Real-time tunnel boring machine monitoring: A state of the art review, *in* 'North American Tunneling 2012 Proceedings', SME.

Moreira, N., Miranda, T., Pinheiro, M., Fernandes, P., Dias, D., Costa, L. & Sena-Cruz, J. (2013), 'Back analysis of geomechanical parameters in underground works using an evolution strategy algorithm', *Tunnelling and Underground Space Technology* **33**(0), 143 – 158.

Murakami, A. (1991), *Studies on the application of Kalman filtering to some geomechanical problems related to safety assessment*, Kyoto University. Doctoral dissertation.

Murakami, A., Shuku, T., Nishimura, S. I., Fujisawa, K. & Nakamura, K. (2013), 'Data assimilation using the particle filter for identifying the elasto-plastic material properties of geomaterials', *International Journal for Numerical and Analytical Methods in Geomechanics* **37**(11), 1642–1669.

Musayev, K., Hackl, K. & Baitsch, M. (2014), 'Identification of the velocity field of 2D and 3D tunnel models with frequency domain full waveform inversion', *PAMM* **14**(1), 781–782.

Nakano, S., Ueno, G. & Higuchi, T. (2007), 'Merging particle filter for sequential data assimilation', *Nonlinear Processes in Geophysics* **14**(4), 395–408.

Neil, D., Haramy, K., Descour, J. & Hanson, D. (1999), 'Imaging ground conditions ahead of the face', *World Tunneling* **12**(9), 425–429.

Nestorović-Trajkov, T., Köppe, H. & Gabbert, U. (2005), 'Active vibration control using optimal LQ tracking system with additional dynamics', *International Journal of Control* **78(15)**, 1182–1197.

Nestorović-Trajkov, T., Köppe, H. & Gabbert, U. (2006), 'Vibration control of a funnel-shaped shell structure with distributed piezoelectric actuators and sensors', *Smart Materials and Structures* **15**, 1119–1132.

Nguyen, L. T., Datcheva, M. & Nestorović, T. (2013), 'Identification of a fault zone ahead of the tunnel excavation face using the extended Kalman filter', *Mechanics Research Communications* **53**, 47–52.

Nguyen, L. T. & Nestorović, T. (2013), 'The extended Kalman filter and the unscented Kalman filter for material parameter identification with application in tunneling', *PAMM* **13**(1), 393–394.

Nguyen, L. T. & Nestorović, T. (2015*a*), 'Nonlinear Kalman filters for model calibration of soil parameters for geomechanical modeling in mechanized tunneling', *Journal of Computing in Civil Engineering* p. 04015025.

Nguyen, L. T. & Nestorović, T. (2015*b*), A study of misfit topography towards waveform inversion for prediction ahead of underground tunnel, *in* 'Near Surface Geoscience 2015 – 21st European Meeting of Environmental and Engineering Geophysics', EAGE.

Nguyen, L. T. & Nestorović, T. (2016), 'Unscented hybrid simulated annealing for fast inversion of tunnel seismic waves', *Computer Methods in Applied Mechanics and Engineering* **301**, 281–299.

Nguyen, L. T., Nestorović, T., Fujisawa, K. & Murakami, A. (2014), Particle filter-based data assimilation for identification of soil parameters with application in tunneling, *in* 'Computer Methods and Recent Advances in Geomechanics', CRC Press, pp. 1241–1246.

Osher, S. & Fedkiw, R. P. (2001), 'Level set methods: an overview and some recent results', *Journal of Computational physics* **169**(2), 463–502.

Otto, R., Button, E., Bretterebner, H. & Schwab, P. (2002), 'The application of TRT (true reflection tomography) at the Unterwald tunnel', *Felsbau* **20**(2), 51–56.

Papadopoulos, E., Cortes, D. D. & Carlos Santamarina, J. (2016), 'In-situ assessment of the stress-dependent stiffness of unbound aggregate bases: application in inverted base pavements', *International Journal of Pavement Engineering* **17**(10), 870–877.

Park, C. B., Miller, R. D. & Xia, J. (1999), 'Multichannel analysis of surface waves', *Geophysics* **64**(3), 800–808.

Petronio, L. & Poletto, F. (2002), 'Seismic-while-drilling by using tunnel boring machine noise', *Geophysics* **67**(6), 1798–1809.

Petronio, L., Poletto, F. & Schleifer, A. (2007), 'Interface prediction ahead of the excavation front by the tunnel-seismic-while-drilling (TSWD) method', *Geophysics* **72**(4), G39–G44.

Petronio, L., Poletto, F., Schleifer, A., Morino, A. et al. (2003), Geology prediction ahead of the excavation front by tunnel-seismic-while-drilling (TSWD) method, *in* 'Society of Exploration Geophysicists, 73rd Annual International Meeting, Expanded Abstracts'.

Plett, G. L. (2006), 'Sigma-point Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 2: Simultaneous state and parameter estimation', *Journal of Power Sources* **161**(2), 1369–1384.

Saenger, E. H., Kocur, G. K., Jud, R. & Torrilhon, M. (2011), 'Application of time reverse modeling on ultrasonic non-destructive testing of concrete', *Applied Mathematical Modelling* **35**(2), 807–816.

Sambridge, M. & Mosegaard, K. (2002), 'Monte Carlo methods in geophysical inverse problems', *Reviews of Geophysics* **40**(3), 3–1.

Sattel, G., Frey, P. & Amberg, R. (1992), 'Prediction ahead of the tunnel face by seismic methods-pilot project in Centovalli tunnel, Locarno, Switzerland', *First Break* **10**(1).

Schanz, T., Vermeer, P. A. & Bonnier, P. G. (1999), The hardening soil model: Formulation and verification, *in* 'R.B.G Brinkgreve, Beyond 2000 in Computational Geotechnics, Balkema, Rotterdam', pp. 281–290.

Sen, M. K. & Stoffa, P. L. (1991), 'Nonlinear one-dimensional seismic waveform inversion using simulated annealing', *Geophysics* **56**(10), 1624–1638.

Sheen, D.-H., Tuncay, K., Baag, C.-E. & Ortoleva, P. J. (2006), 'Time domain Gaussian–Newton seismic waveform inversion in elastic media', *Geophysical Journal International* **167**(3), 1373–1384.

Shuku, T., Murakami, A., Nishimura, S.-i., Fujisawa, K. & Nakamura, K. (2012), 'Parameter identification for Cam-clay model in partial loading model tests using the particle filter', *Soils and Foundations* **52**(2), 279–298.

Snieder, R. (2006), 'The theory of coda wave interferometry', *Pure and Applied Geophysics* **163**(2-3), 455–473.

Socco, L. V., Foti, S. & Boiero, D. (2010), 'Surface-wave analysis for building near-surface velocity models — established approaches and new perspectives', *Geophysics* **75**(5), 75A83–75A102.

Stähler, S. C., Sens-Schönfelder, C. & Niederleithinger, E. (2011), 'Monitoring stress changes in a concrete bridge with coda wave interferometry', *The Journal of the Acoustical Society of America* **129**(4), 1945–1952.

Stein, S. & Wysession, M. (2009), *An introduction to seismology, earthquakes, and Earth structure*, John Wiley & Sons.

Su, Z. & Ye, L. (2009), *Identification of damage using Lamb waves: from fundamentals to applications*, Vol. 48, Springer Science & Business Media.

Swinnen, G., Thorbecke, J. & Drijkoningen, G. (2007), 'Seismic imaging from a TBM', *Rock Mechanics and Rock Engineering* **40**(6), 577–590.

Tarantola, A. (2005), *Inverse problem theory and methods for model parameter estimation*, SIAM.

Tran, K. T. & Hiltunen, D. R. (2011), 'Two-dimensional inversion of full waveforms using simulated annealing', *Journal of Geotechnical and Geoenvironmental Engineering* **138**(9), 1075–1090.

Tromp, J., Komatitsch, D. & Liu, Q. (2008), 'Spectral-element and adjoint methods in seismology', *Communications in Computational Physics* **3**(1), 1–32.

Tromp, J., Tape, C. & Liu, Q. (2005), 'Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels', *Geophysical Journal International* **160**(1), 195–216.

Tzavaras, J., Buske, S., Groß, K. & Shapiro, S. (2012), 'Three-dimensional seismic imaging of tunnels', *International Journal of Rock Mechanics and Mining Sciences* **49**, 12–20.

Van Der Merwe, R. (2004), Sigma-point Kalman filters for probabilistic inference in dynamic state-space models, PhD thesis, Oregon Health & Science University.

Van Der Merwe, R. & Wan, E. A. (2001), Efficient derivative-free Kalman filters for online learning, *in* 'The 9th European Symposium on Artifical Neural Networks (ESANN'2001), Bruges, Belgium', ISBN 2-930307-01-3, pp. 205–210.

van Leeuwen, T. & Herrmann, F. J. (2013), 'Mitigating local minima in full-waveform inversion by expanding the search space', *Geophysical Journal International* **195**(1), 661–667.

Virieux, J. & Operto, S. (2009), 'An overview of full-waveform inversion in exploration geophysics', *Geophysics* **74**(6), WCC1–WCC26.

Wan, E. A. & Van Der Merwe, R. (2000), The unscented Kalman filter for nonlinear estimation, *in* 'Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000', IEEE, pp. 153–158.

Wan, E. A. & van der Merwe, R. (2001), The unscented Kalman filter, *in* S. Haykin, ed., 'Kalman Filtering and Neural Networks', Wiley, New York, chapter 7.

Wang, X. & Cai, M. (2015), 'Influence of wavelength-to-excavation span ratio on ground motion around deep underground excavations', *Tunnelling and Underground Space Technology* **49**, 438–453.

Xia, J., Miller, R. D. & Park, C. B. (1999), 'Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves', *Geophysics* **64**(3), 691–700.

Yamamoto, T., Shirasagi, S., Yokota, Y. & Koizumi, Y. (2011), 'Imaging geological conditions ahead of a tunnel face using three-dimensional seismic reflector tracing system', *International Journal of the JCRM* **6**(1), 23–31.

Yang, C. X., Wu, Y. H., Hon, T. & Feng, X.-T. (2011), 'Application of extended Kalman filter to back analysis of the natural stress state accounting for measuring uncertainties', *International Journal for Numerical and Analytical Methods in Geomechanics* **35**(6), 694–712.

Zarev, V., Schanz, T. & Datcheva, M. (2011), Three-dimensional finite-element analysis of mechanized shield tunneling in urban areas, *in* 'Proceedings of The Thirdteenth International Conference on Civil, Structural and Environmental Engineering Computing, Greece', pp. 1–10.

Zhao, Y., Jiang, H. & Zhao, X. (2006), 'Tunnel seismic tomography method for geological prediction and its application', *Applied Geophysics* **3**(2), 69–74.

Zheng, L., Zhao, Q., Milkereit, B., Grasselli, G. & Liu, Q. (2014), 'Spectral-element simulations of elastic wave propagation in exploration and geotechnical applications', *Earthquake Science* **27**(2), 179–187.

List of Figures

About the author

Personal information

Name: Luan T. Nguyen (Nguyễn Thành Luân)

Birth: July 13, 1985 in Long Xuyen, Vietnam

Email: thanh.nguyen-w7t@rub.de

Educational history

Nov. 2011 – Apr. 2017: Ph.D. candidate in Civil Engineering, Ruhr-University Bochum, Germany

Oct. 2009 – Sep. 2011: M.Sc. program in Computational Engineering, Ruhr-University Bochum, Germany

Sep. 2003 – Mar. 2008: B.Eng. program in Automation, Can Tho University, Vietnam

Work history

Nov. 2011 – Apr. 2017: Scientific assistant, Ruhr-University Bochum, Germany

Mar. 2008 – Aug. 2009: Hardware engineer, Renesas Design Vietnam Co., Ltd., Vietnam

Mitteilungen aus dem Institut für Mechanik
RUHR-UNIVERSITÄT BOCHUM
Nr. 175