Automated Survey Design for Nonlinear Problems

(1) Université Pierre et Marie Curie
Laboratoire Jacques-Louis Lions
75252 Paris Cedex 05, FRANCE
winterfors@ann.jussieu.fr

Emanuel Winterfors(1) and Andrew Curtis(2)

(2) Edinburgh University
Department of Geology and Geophysics
Grant Institute, West Mains Road
Edinburgh EH9 3JW, UK

Why?

Enormous sums of money are invested annually on academic and industrial geophysical remote sensing surveys. Yet, methods used to design such surveys are rudimentary, usually based on heuristics (rules of thumb) and linearised physics due to the complexity of nonlinear design problems. Introducing methods to improve the basic information content of recorded geophysical data should lead to greater efficiency and value for money (Curtis and Maurer, 2000).

The field of Statistical Experimental Design deals with the problem of finding the experimental (equivalently, survey) design (e.g. detector types and locations) that gives the best estimate of the model parameters in terms of accuracy, robustness, or any other desirable design criterion. It is a well-established branch of statistics (Atkinson and Donev, 1992; Pukelsheim, 1993) and has been applied to a number of different scientific investigation techniques, including geophysics (Rabinowitz and Steinberg, 1990; Steinberg et al., 1995, Curtis, 1999).

Design theory based on a linearised approximation of the relation between model parameters and measured data does not take into account changes in the design criterion due to nonlinear effects which are common in geophysics, making resulting designs non-robust. This work provides a systematic approach for detecting and quantifying such effects, as well as a design optimisation algorithm for their reduction and -- if possible -- elimination.

Survey goals are usually to estimate material properties, or to locate sources (or receivers) of wave-based energy using detectors at a few locations. We focus here on geophysical location problems, examples of which include monitoring seismic or nuclear test activity, surface geodetic or subsurface active source or receiver positioning, and localising secondary sources of wave scattering.

The Problem

When there is a nonlinear relation between the parameters of intrest and the observed data, it is possible to have observations that correspond to several different parameter combinations. These ambiguous observations will corresond to self-intersectinos of the function predicting observations.

Two schematic examples are shown to the right (figures 1 and 2).

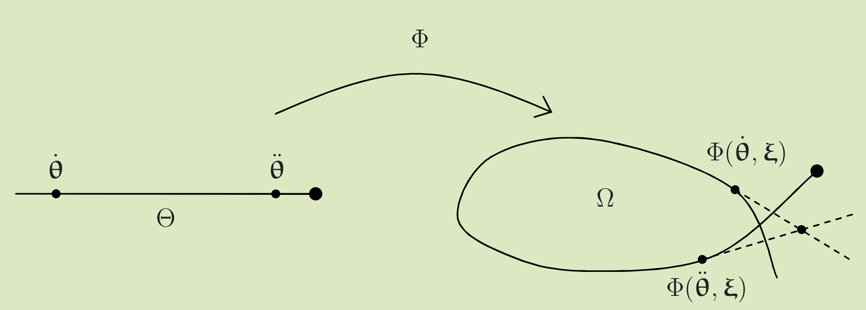


Figure 1: A simple case with one single parameter of interst and with a nonlinear relation to two obsereved variables is depicted to the right, with one probematic observation at the self-intersection of the curve corresponding to two different points in the range of the parameter of interest.

Finding Ambiguous Observations

It is possible to search for intersection point pairs using a modified version of Newton's method for finding roots of an equation. Starting with an initial guess of an intersection point pair, a linear approximation around each of the points can be used to find an approximate intersection (see figure 1). Updating the initial guess with the approximation and repeating the procedure, an iterative algorithm has been constucted that will converge to a point pair on the real intersection, provided the initial guess is close enough.

By repeating the intersection search algorithm many times using different random starting points, a sample of the intersection maifolds can be created (in figure 2 these is a 3-point pair sample of an intersection line manifold).

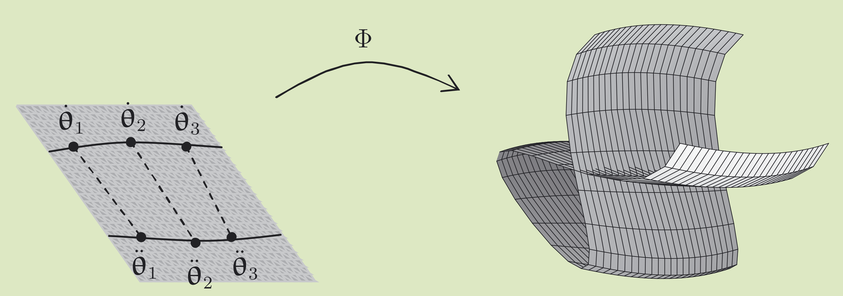


Figure 2: A 2D-space spanned by two model parameters is observed by three variables, where a self-intersection gives a line of ambiguous observations corresponding to a manifold of distant point pairs (of which tree are depicted), yielding the same observation.

Improving the Design

By studying how the intersection point pairs move when changing the design of the investigation technique, it is possible to constuct an algorithm for reducing the distance (on average) between the two points of each pair. The goal is to get them as close as possible, or to completely coincide - equivalent to the ambiguity of the corresponding observation to disappear.

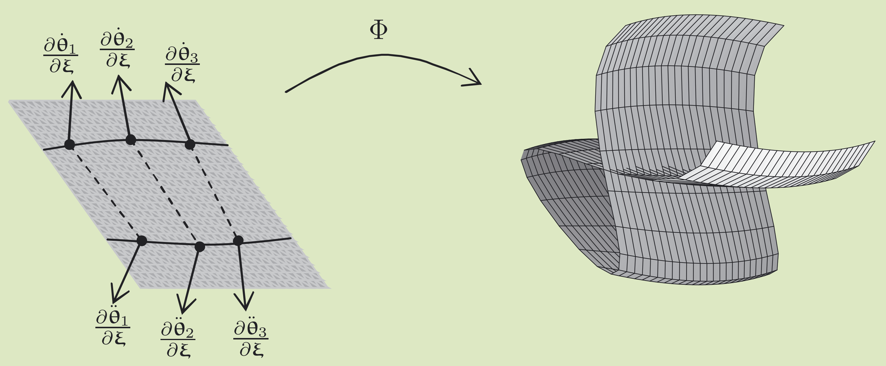
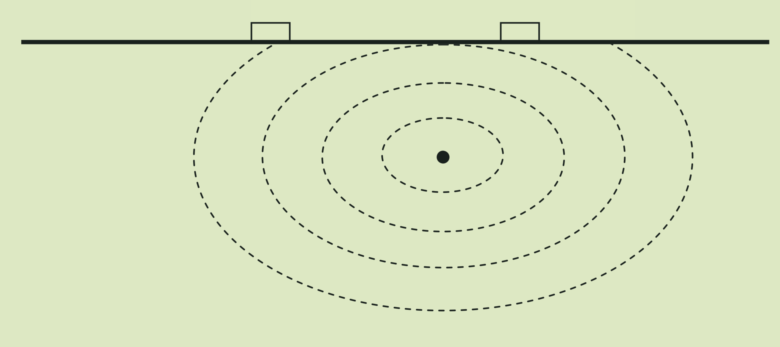


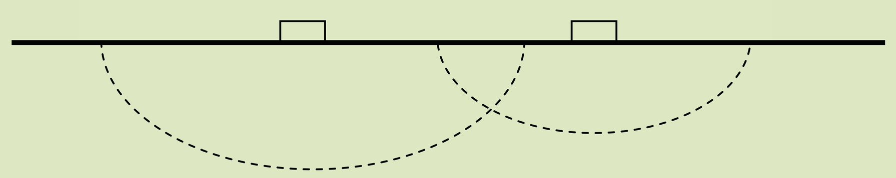
Figure 3: Evolution of intersection point pairs under perturbations of a design parameter, affecting the shape of the relation between model parameter (af interest) and observed variables.

We propose a novel measure of the quality of any investigation technique that accounts for effects on parameter estimation due to nonlinear model-data relationships. Computational algorithms are derived that enable its efficient numerical estimation, as well as the optimisation of an experimental design by maximising the quality measure. Thus, this work makes designing robust geophysical remote sensing surveys and experiments feasible. The efficacy of the approach is illustrated by designing a seismic location experiment.

Application: Seismic location

In a seismic location experiment, the waves from a seismic event propagate through some inhomogeneous medium for which the structure is assumed to be known. The waves are recorded by detectors at the surface or in boreholes. The seismic event will generate both pressure and shear waves that will travel at different speeds through the medium. It is therefore possible to deduce the distance from the detector to the source by recording the difference in arrival times for the two wave types.





In isolation, the data from each detector will thus yield a manifold of possible source locations, with size dependent on the difference between arrival times for pressure and shear waves. Combining the data from several detectors rules out all locations but those on the intersection between the manifolds.

With an inhomogeneous velocity distribution in the medium, the manifolds are less regular and might intersect at several different locations. In that case there will be several candidate locations for the source. It is of great interest to be able to analyse the geometry of the problem to detect such combinations of arrival times that give rise to ambiguous locations, as well as to quantify how distant the multiple possible locations might be from each other. The detector locations might then be chosen to minimise the expectation of this distance, post-experiment.

The intersection search algorithm was applied repeatedly with different starting points to find pairs of locations that would yield the same observed time differences between pressure and shear wave arrival time at the detectors.

The velocity profile that was used was a background linear increase in velocity difference between pressure and shear wave velocity with respect to depth with a significant perturbing random variation of on the order of order 1500 m/s. The profile only varies with depth and has no variation in the lateral directions.

For a distance of 2048 m between the two simulated detectors, a sample of the resulting ambiguous locations are shown in figure 4a.

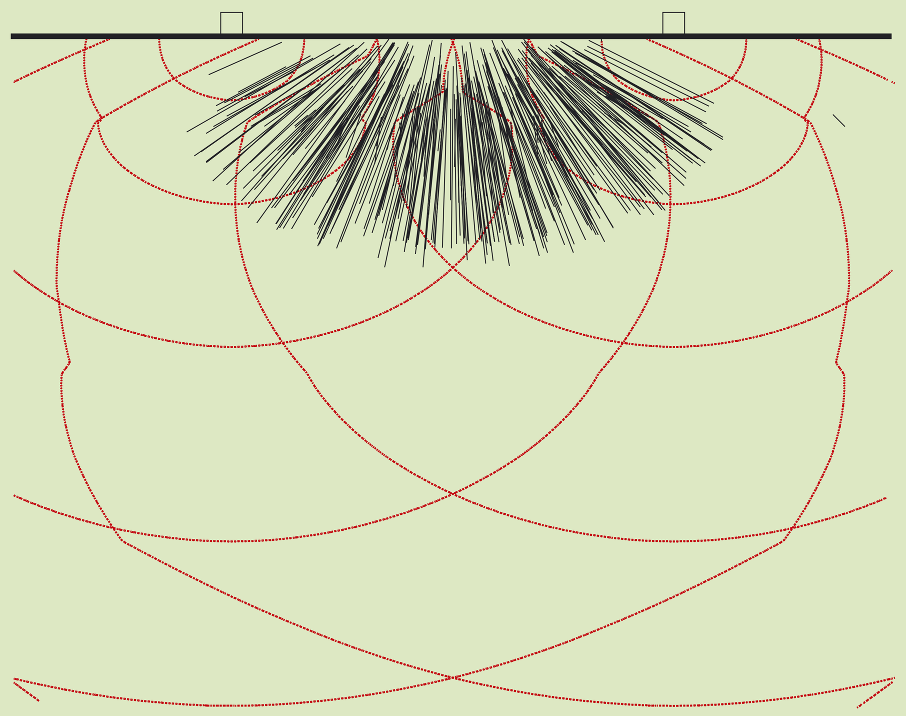


Figure 4a: Initial two-detector configuration at distance of 2048m

Figures 4a-c: Surface detectors for source location, each with travel time contours in red. Solid lines connect pairs of ambiguously resolved seismic source locations, sampled randomly using the intersection search algorithm.

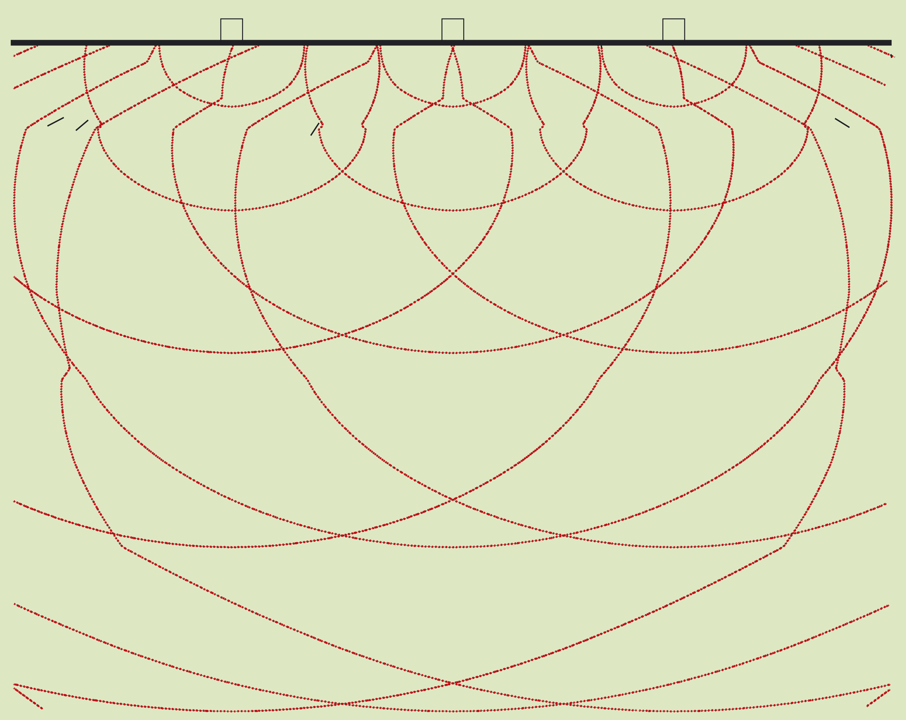


Figure 4b: 3-detector version

If a third detector is added in between the two existing, almost all ambiguity in possible detected locations is eliminated (figure 3b).

Finally an iterative optimisation algorithm was applied to minimise the sum of the squared distances between the two points of each intersection point pair. It converged to an optimal distance of 1280 m between the detectors, which gave the distribution of ambiguous location pairs shown in figure 4c. The lines are shorter than in figure 4a, meaning that although ambiguity still exists, the alternative possible location will always be relatively proximal no matter where the true event occurred. The short lines at that have appeared at the deeper range correspond to lateral uncertainty due to the angles to the two detectors becoming more and more similar for events at increasing depth.

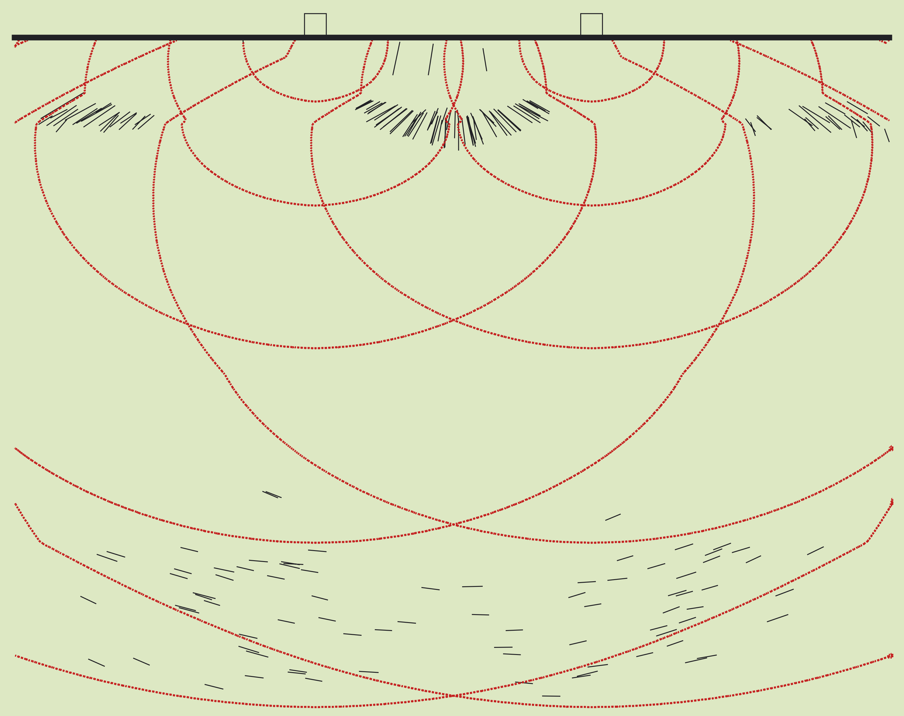


Figure 4c: After optimisation, surface detector distance = 1280m

Application: Travel Time Tomography

In travel time (cross well) tomography, seismic waves will travel trough inhomogeneous medium just as in the location problem, with the main difference being that the source position is known but velocity profile unknown. We used a 3x3 point velocity grid (in two dimensions) with linear interpolation of the velocity in between the defined points. Travel times were calculated for at six detector locations for waves originating at two different sources, making up a total of 12 arrival times. One could believe that 12 observed variables should be enough to determinate the 9 velocity parameters of interest (which is also what is assumed in standard seismic inversion methods), but our algorithm for finding ambiguous observations shows this is not necessarily the case.

The intersection search algorithm was applied on pairs of random velocity profiles, in search of pairs of profiles that give rise to the same observed combination of arrival times. Three examples of such pairs that were found is shown below. Even though the velocity profiles and travel paths are radically different, the travel times are identical.

If standardised tomographic inversion methods based on local linearization was applied to calculate the velocity profile for one of the ambiguous observed travel times below, it would either one or other of the two possible velocity profiles, depending on what profile was used as starting point.

In high dimensional nonlinear problems — such as travel time tomography — there will in fact almost always be an infinite number of intersection point pairs. This is a consequence of the multitude of possible ways for a function to "curl up on itself" in high dimension.

Our method provides an efficient way of detecting the extent of such problems.

In many cases, the ambiguity of certain observations can be resolved just by increasing the amount of observed data (as in figure 4b), but it may sometimes not be the best option due to cost or other constraints. Is such cases, an optimisation approach as described may provide a design with satisfactory precision.

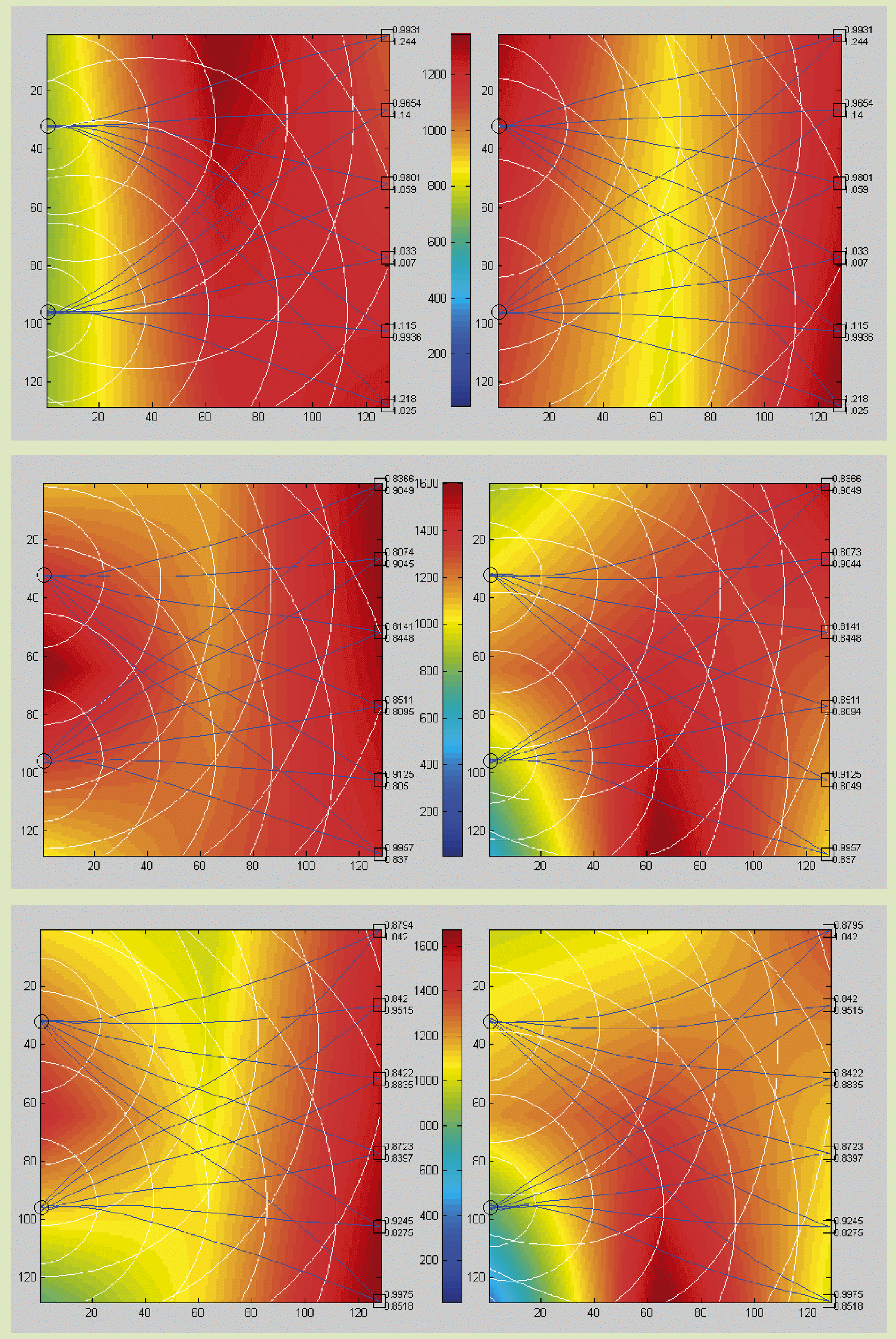


Figure 5: Pairs of velocity profiles that give rise to the same obsered travel times: Two simulated sources to the left and six simulated recievers to the right, along with observed travel times (two for each reciever) in seconds. Colour-velocity scale (in m/s) is in between each pair. Shortest travel paths are shown in blue and travel time contours in white.

References

A. C. Atkinson and A. N. Donev. *Optimum Experimental Designs*. Oxford University Press, 1992. ISBN 0198522541.

A. Curtis and H. Maurer. Optimizing the design of geophysical experiments: Is it worthwhile? *EOS, Trans., Amer. Geophys. Union*, 81(20):224–225, 2000. Reprinted in The Leading Edge, October 2000.

A. Curtis, A. Michelini, D. Leslie, and A. Lomax. A deterministic algorithm for experimental design applied to tomographic and microseismic monitoring surveys. *Geophysical Journal International*, 157(2):595–606, 2004.

A. Curtis and C. Spencer. Survey design strategies for linearized nonlinear inversion. In *69th Annual International Meeting, Expanded Abstracts*, pages 1775–1778. Society of Exploration Geophysics, 1999.

Andrew Curtis. Optimal design of focused experiments and surveys. *Geophysical Journal International*, 139(1):205–215, 1999. 11

A. Michelini and A. Lomax. The effect of velocity structure errors on double-difference earthquake location. *Geophys. Res. Lett.*, 31(L09602): doi:10.1029/2004GL019682, 2004.

Friedrich Pukelsheim. *Optimal Design of Experiments*. Wiley and Sons, New York, 1993. ISBN 047161971X.

N. Rabinowitz and D.M. Steinberg. Optimal configuration of a seismographic network: a statistical approach. *Bull. Seis. Soc. Am.*, 80:187–196, 1990.

D.M. Steinberg, N. Rabinowitz, Y. Shimshoni, and D. Mizrachi. Configuring a seismographic network for optimal monitoring of fault lines and multiple sources. *Bull. Seis. Soc. Am.*, 85:1847–1857, 1995.

A. Tarantola and B. Valette. Inverse problems = quest for information. *Journal of Geophysics-Zeitschrift Fur Geophysik*, 50(3):159–170, 1982.

Acknowledgements

Parts of this work was carried out under the HPC-Europa project (RII3-CT-2003-506079), with the support of the European Community - Research Infrastructure Action under the FP6 "Structuring the European Research Area" Programme. The Eikonal solver for calculating travel times was provided by Enru Liu of the British Geological Survey.