Hybrid Genetic Deflated Newton Method for Distributed-Source Optimization

Marcus M. Noack[a,b,c,d,1], Steven M. Day[e,2]

a: Kalkulo AS, P.O.Box 134, 1325 Lysaker, Norway
b: Simula Research Laboratory, P.O.Box 134, 1325 Lysaker, Norway
c: Department of Informatics, University of Oslo, Gaustadalleen 23 B, 0373 Oslo, Norway
d: High Performance Geocomputing Lab, San Diego Supercomputer Center, UCSD, 9500 Gilman Drive, La Jolla CA 92093-0505, USA
e: Department of Geological Sciences, San Diego State University, 5500 Campanile Dr.,San Diego, CA 92182-1020, USA

1: mcn@simula.no, 2: sday@mail.sdsu.edu

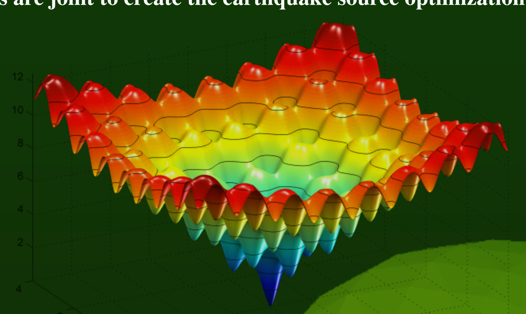
Abstract

Earthquakes are a hazard which compromises many lives every decade around the planet. Finding out more about the source of earthquakes can help to make decisions to avoid fatalities. To investigate the source of an earthquake, wave motion data can be inverted for the fault parameters. The inversion process comprises an optimization of a highly non-linear misfit function. There are in general two issues when optimizing highly non-linear functions: the function shows in general many local optima and the global solution might not be unique. Standard optimization procedures, like the common Newton optimization or the genetic algorithm, can not be used since they either find only one stationary point or are inefficient. Recently, we introduced a new hybrid optimization method which is able find many optima and can effectively remove them from the function. The resulting algorithm finds, lists, and removes optima until the global optimum is found or until no improvement of the misfit can be made. This new optimization procedure can be used to invert for many fault parameters and could therefore deliver a clear picture of the fault dynamics. As a first step, in this work, we are outlining the theory for the acoustic approximation of the wave equation and a simplified source term. The simplifications are made to direct the focus on the optimization procedure and to keep the mathematical procedure as simple as possible. Future work will, step by step, consider more complex models.

1. Introduction

Earthquake fault parameter inversion is a promising field of research. An accurately defined source leads in turn to accurate ground motion predictions. The promising perspective comes at a cost; local and global optimization procedures, like the Newton method and the genetic algorithm, offer a trade-off between accuracy and computational cost. Using the renowned Newton method cannot lead to success finding the global optimum because the misfit function to be optimized exhibits many local optima due to the natural sine shape of a wave. Ackley's function (see below) represents a famous example for this problem. Furthermore, the global optimum might not be unique. On the plus side, the Newton method is very efficient in finding a stationary point. As a global method, the genetic algorithm suffers from high computational costs in high dimensional search spaces. Also, since no local information of the function is used, the algorithm cannot determine whether a stationary point is found. Therefore, only one solution can be found. An advantage of the genetic algorithm is, however, that the global optimum will be found eventually [1]. Genetic algorithm and the Newton method are only examples of many local and global optimization methods but they are representative in terms of advantages and disadvantages. To join the advantages of both methods, hybrid methods have been developed [2, 3, 4, 1]. Hybrid methods are using a global optimization scheme like the genetic algorithm to explore the parameter space and a local scheme to find stationary points. One major drawback persists; one optimum can be found several times which leads to biased behavior and a premature convergence. Another algorithm worth mentioning in the context of earthquake source inversion is the neighborhood algorithm [5]. This scheme places a multiplicity of individuals randomly in the search space. The function values of the individuals determine regions in which the global optimum could be located. The neighborhood algorithm does not use local information. For many dimensions of the search space, the algorithm becomes inefficient because many individuals are necessary to preserve the unbiased behavior. Recently, we proposed a new method with the name Hybrid Genetic Deflated Newton method (HGDN) which uses the advantages of a hybrid method. The new method places several individuals randomly in the search space. All individuals perform a Newton method to a stationary point. The found points are transferred to a list and the points are removed from the function by deflation. Afterwards, the entire population of individuals is set back to its starting position and the search starts again. Already found optima cannot be found again. Therefore, the individuals converge in new stationary points. The procedure continues until no more stationary points can be found in the vicinity. A genetic algorithm replaces the fittest individuals and the procedure starts over. In the context of earthquake source optimization, the function to be optimized is a functional, mapping a real number to the difference of a computed and a measured wave motion field. The mentioned individuals are different source models. Since it is unfeasible to obtain the gradients and the Hessian by finite differencing, the adjoint equations of the problem have to be derived. The location of the optimum gives the parameters of the source of the investigated earthquake.

In this work, we outline the work flow of using HGDN for fault parameter optimization, illustrating the concept for the case of a distributed acoustic source. Future work will consider the corresponding elasto-dynamic source inverse problem. The poster is organized as follows. Each box gives an introduction to the basics of the new algorithm. The center box outlines how the methods are joint to create the earthquake source optimization method.



2. Hybrid Optimization Schemes

To combine the strengths of global and local optimization methods, hybrid methods have been developed in the past. Hybrid methods use a global search algorithm to explore the search space on a global level. At this stage, we are focusing on hybrid schemes that use the genetic algorithm as global search scheme and the Newton method as local scheme. After each iteration of the genetic algorithm, all individuals perform a Newton search to find a stationary point of the function. When all individuals have converged, the genetic algorithm choses the fittest individuals and creates offspring.

The genetic algorithm works as follows. A random population is created and placed in the search space. We refer to a population as a multiplicity of chosen points (individuals) in the search space. The fittest individuals have the best chance to procreate and produce offspring. The fitness, in this context, is the function value at the point that is associated with a certain individual. The offspring are built by crossover of the genome (the location in the search space) of the parent individuals. There are many different types of crossover ranging from one-point crossover to completely random methods. After the crossover, mutation can happen randomly, which gives the genetic algorithm the ability to find the global optimum eventually. Mutation means a random change of the genome (the location) of an affected individual. Mutation and the chance for individuals that are not among the fittest to procreate, give the genetic algorithm an unbiased behavior.

The offspring generation is, in general, fitter than the last one, which leads to the convergence of the algorithm. After a new generation is created, all individuals start again the search using the Newton scheme. A Newton scheme uses first and second derivatives to optimize a function locally as explained in section 5. An obvious drawback of the traditional hybrid method is that the local search method might compute the same optima for different individuals; also, these optima might be re-identified repeatedly in each genetic iteration. Therefore, a significant amount of computational effort is potentially utilized identifying already known optima. Only if an individual is positioned sufficiently close to a new optimum, then the local search will converge to this new optimum.

3. The Adjoint Problem

As outlined earlier, the new hybrid optimization scheme uses local information about the function to reliably find stationary points. Therefore, it is inevitable to compute the first and second partial derivatives of the function at the respective locations. Since finite differencing is unfeasible, the adjoint problem has to be derived. The derivation starts with the formulation of a misfit or cost function

$$\Phi[u(\mathbf{x},t;f(\xi,\tau))] = \frac{1}{2}\int_T\int_\Omega [u(\mathbf{x},t;f(\xi,\tau)) - u_0(\mathbf{x},t)]^2 M(\mathbf{x}) d^3\mathbf{x}\, dt$$

The interesting physical entity is the functional derivative of this function with respect to the displacement (or in this case pressure) field u. The functional derivative of the objective function with respect to u is given by

$$\delta\Phi[\delta u(\mathbf{x},t;f(\xi,\tau))] = \int_T\int_\Omega [u(\mathbf{x},t;f(\xi,\tau)) - u_0(\mathbf{x},t)]\delta u(\mathbf{x},t,\delta f(\xi,\tau))M(\mathbf{x}) d^3\mathbf{x}\, dt$$

The optimization is subject to the wave equation

$$\ddot{u}(\mathbf{x},t) - \kappa\nabla\cdot(\rho^{-1}\nabla u(\mathbf{x},t)) = f(\xi,\tau)$$

The solution can be written by using Green's function

$$u(\mathbf{x},t;f(\xi,\tau)) = \int_\Sigma\int_\Lambda G(\mathbf{x},t,\xi_\mathbf{0},\tau_0)f(\xi,\tau)d^3\xi\, d\tau$$

Inserting this equation in the functional derivative of the misfit function and using the symmetry properties of Green's function leads to

$$\delta\Phi[u(\mathbf{x},t;\delta f(\xi,\tau))] =$$
$$\delta f(\xi,\tau)\int_T\int_\Omega [u(\mathbf{x},T-t;f(\xi,\tau)) - u_0(\mathbf{x},T-t)]M(\mathbf{x})G(\mathbf{x},t,\xi,\tau)d^3\mathbf{x}\, d\tilde{\tau}$$

In this formulation, we find a solution of the wave equation that uses the time reversed misfit as source. The solution to this equation is the adjoint field. The end result can therefore be formulated as

$$\delta\Phi[u(\mathbf{x},t;\delta f(\xi,\tau))] = \int_\Sigma\int_\Gamma \delta f(\xi,\tau)\Psi(\xi,T-\tau)d^3\xi\, d\tau.$$

If we define the source spatially by linear basis functions and temporarily by delta distributions the gradient with respect to the model parameters (m) is given by

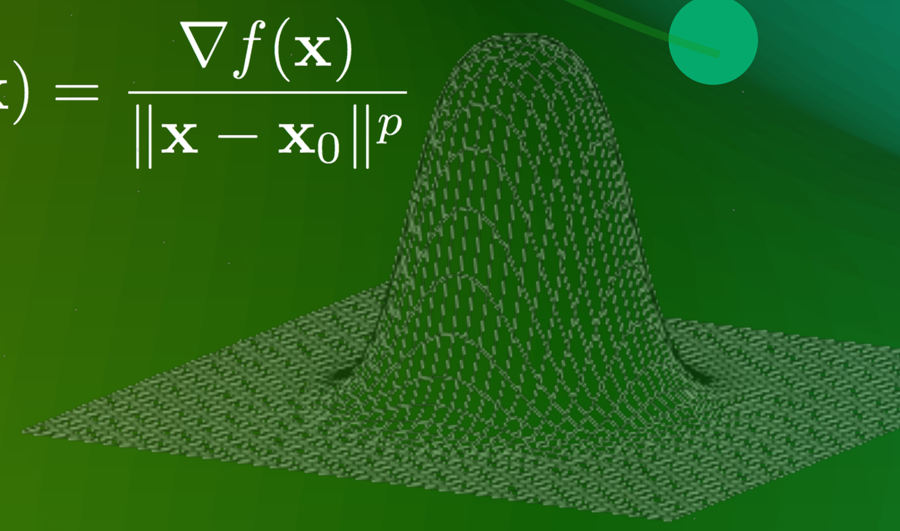
$$\frac{\partial\Phi}{\partial m_{ij}} = \int_\Sigma \Psi(\xi,T-\tau_j)\beta_i(\xi)d^3\xi$$

4. Deflation Operators

The core of the novel optimization method is the deflation. After a new optimum is found, its location is transferred to a list. In the next iteration, the function will be deflated at the stored locations. The function being deflated means that the gradient of the function is altered as shown in the equation below. Deflating the function effectively removes the stationary point since a Newton scheme does not converge anymore in deflated optima. Therefore, already found optima cannot be found again, which represents the main advantage compared with traditional hybrid methods.

The deflation operator presented below cannot directly be used when multiple deflations occur since the gradient of the function will be altered globally in this case. Therefore, a new deflation technique had to be developed, namely the localized deflation. The localized deflation operator uses bump functions to alter the gradient of the objective function only in a well defined area. Bump functions have compact support and are differentiable for all degrees of differentiation. They allow for a highly shapable deflation in a well defined radius.

$$\nabla f_{\mathbf{x_0}}(\mathbf{x}) = \frac{\nabla f(\mathbf{x})}{\|\mathbf{x}-\mathbf{x}_0\|^p}$$

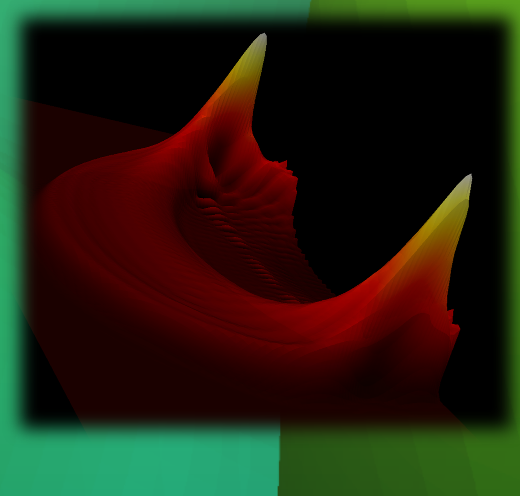


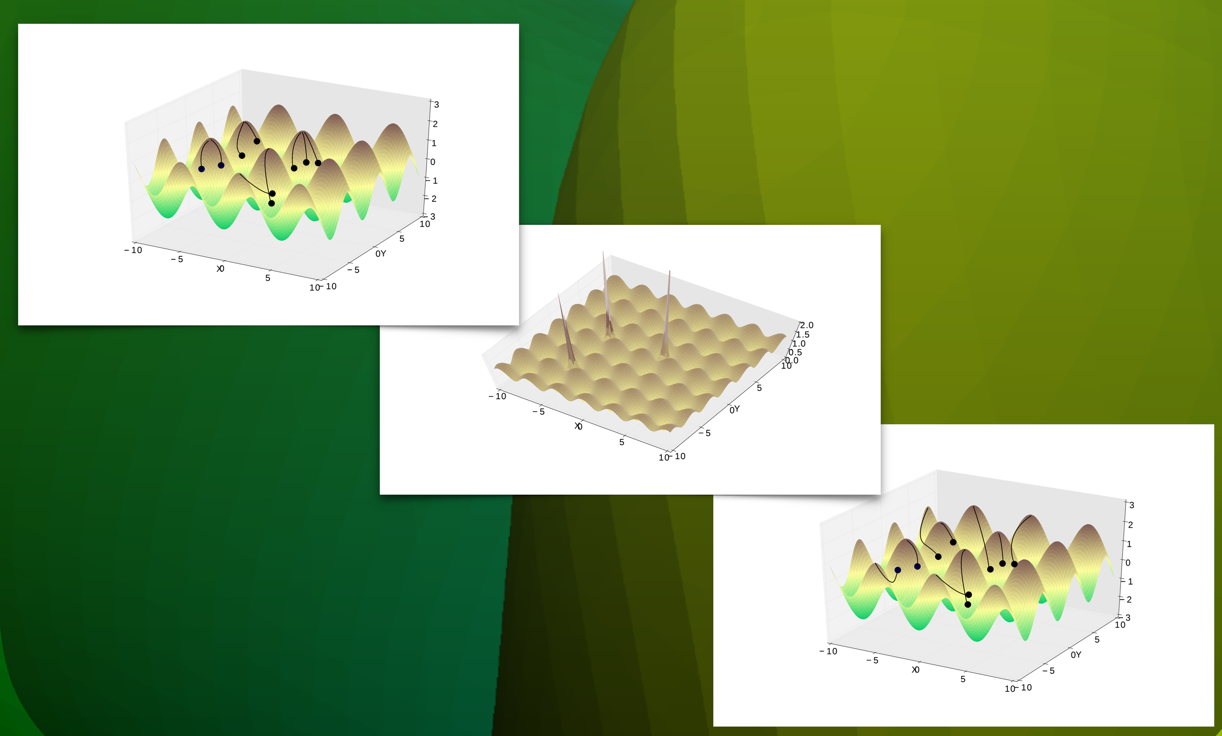
5. Source Parameter Optimization

The entire procedure of the source optimization can be outlined as follows. In the first step, a multiplicity of possible source models (individuals) are selected. For each of the source models, a forward wave simulation is computed. A snapshot of the simulation is shown on the right side. According to the given equations in section 3, the computed residuals between simulated and measured wave fields have to be reverted in time and back propagated to the source. This will yield the gradient for each individual. The Hessian can be computed by using the Gauss-Newton method. The update of the position of each individual can be computed by solving the system

$$H\gamma = -\nabla f$$

for gamma. The described procedure can be repeated until all individuals have converged to a stationary point. The found stationary points are transferred to a list and stored. The gradient of the function is deflated at the stored points. The entire population of individuals can now be set back to their initial position and the search for nearby optima can start over. Since earlier found stationary points were removed by deflating the gradient of the function, exclusively new stationary points are found (see figures below). This procedure continues until most individuals cannot find a stationary point. In this case, a subsequent genetic algorithm replaces the population by procreation of the fittest. The procedure starts over at the new locations. The method can find multiple local and global optima. Furthermore, the algorithm can detect non uniqueness by assessing the null-space at the optima. The end result is a range of possible source models with given null spaces.





References

[1] J.-M. Renders, S. Flasse, Hybrid methods using genetic algorithms for global opti- mization, Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on 26 (2) (1996) 243–258, doi:10.1109/3477.485836

[2] W. Wan, J. B. Birch, An Improved Hybrid Genetic Algorithm with a New Local Search Procedure, Journal of Applied Mathematics 2013, doi: 10.1155/2013/103591.

[3] P. Xu, A hybrid global optimization method: The multi-dimensional case, Jour- nal of Computational and Applied Mathematics 155 (2) (2003) 423–446, doi: 10.1016/S0377-0427(02)00878-6.

[4] Efficient global optimization using hybrid genetic algorithms, 9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, 2002.

[5] M. Sambridge, Geophysical inversion with a neighbourhood algorithm—II. Ap- praising the ensemble, Geophysical Journal International 138 (3) (1999) 727–746.