

# Hybrid CMAES

Brian Hoenes  
Colorado School of Mines  
Department of Mathematics and Computer Science  
1700 Illinois Street  
Golden, Colorado, USA  
bhoenes@mines.edu

May 13, 2011

## **Abstract**

TODO - write abstract

# 1 Introduction

The objective of this project is to improve the performance of parallel version of the Covariance Matrix Adaption Evolution Strategy (CMA-ES) algorithm on the Alamode cluster through an application of a hybrid OpenMP and MPI approach. The Colorado School of Mines SmartGeo program is attempting to use complex resistivity and seismic geophysical methods to identify and localize biological activity in its Intelligent Remediation Project. Localization of phenomena using these geophysical methods requires the solution of the inverse problem based on the collected field data and a forward model of the physics of the system. This paper focuses on improving the time performance of a numerical method to solve a subset inverse problem of the overall feedback control loop.

## 1.1 Intelligent Bioremediation

Currently, bioremediation is performed at the project test using an electron donor liquid amendment injection system with open loop controls. The current bioremediation system injects a concentrated electron donor solution to facilitate reduction of Uranium from a soluble state to an insoluble state. In order to provide data to adjust the system operating parameters, water samples are taken and analyzed at an off-site laboratory. The results of the off-site chemical analysis are generally available in several months. Upon receipt of the data, adjustments to the bioremediation system parameters, such as injection rate or concentration, are made by manually turning valves based. The Intelligent Remediation Project is attempting to identify and localize the electrical signal from the microbial reduction of Uranium using near real-time geophysical methods in order to actuate the bioremediation system controls in a near real-time feedback manner.

In order to localize the microbial reduction signal, it is necessary to identify and subtract other phenomena that produce an overlapping electrical signal. In order to solve the desired problem of microbial reduction activity localization, it is necessary to localize and subtract the signal contribution of ionic solution injection. Both of these problems are inverse prob-

lems.

## 1.2 Inverse Problems

Ideally, an inverse problem could be solved using strictly analytical methods, however, ([?]) has identified following three primary issues make the solution of the inverse problem in a analytical manner impractical in practice:

1. Exact inversion techniques are usually only applicable for idealistic situations that may not hold in practice due to inherently noisy data sets because of the small magnitude of the physical phenomena measured and the ambient noise in the relatively uncontrolled environment.
2. Exact inversion techniques are often very unstable.
3. Most importantly, the model that one aims to determine is a continuous function of the space variables. This means that the model has infinitely many degrees of freedom. However, in a realistic experiment the amount of data that can be used for the determination of the model is finite. In addition to finite data sets, the Smart-Geo physical systems exhibit significant spatial and temporal heterogeneity and have complex interactions of the system components (e.g. geology, geochemistry, hydrology, biology, etc.)

Because an analytical solution is impractical, inverse problems are often solved using iterative numerical methods that estimate and appraise trial solutions. In this case, the Intelligent Bioremediation Project has chosen to utilize the iterative numerical method. The following more formal description of the injection dipole localization inverse problem formulation taken from ([?]). A specific requirement for analysis of self potential data is the isolation of component sources that generate the parts of the total self potential response. Unfortunately, there is no way to directly determine the location of phenomena from a set of electrode observations. Many of the signals can be estimated as dipoles (a source that generates an electrical field based on having two poles). In order to decompose the signal, any dipole signals must

be identified and localized. Specifically, localization for a dipole must identify the spatial localization and a vector describing the direction and strength of the dipole. Thus, each dipole has six defining features: its location in x,y,z space; and the definition of the direction and strength of the dipole, defined either as a vector (strength in x,y,z directions) or as angles (direction in plane and azimuth) with magnitude. The available sensor readings for self potential provide the electrical potential field at a time instant as measured by the electrodes. Because of these properties, the problem dimensional space increases by a factor of 6 for each added dipole.

By considering the geology, resistivity, hydrogeology, geochemistry, and other parameters of a site, a model can be developed that provides an approximation of what observations would be seen if a given dipole (or dipoles) was present. In order to use measurements from self potential electrodes for dipole localization, the inverse problem of finding dipoles from samples of an electrical field must be solved. The Intelligent Bioremediation Project has chosen the

## 2 Related Work

### 2.1 CMA-ES

### 2.2 pCMAES

identified the generation and evaluation of the next generation as the processing intensive task. created separate MPI nodes that generated and evaluated

## 3 Hybrid OpenMP/MPI Approach

reduce the number of MPI nodes to match the number of independent machines in target architecture perform the

---

### Algorithm 1 CMAES

---

```

initialize Random Seed on  $p$ 
initialize mean  $m \in \mathbb{R}^n$ , covariance  $C$ , step size  $\sigma$ ,
path(s)  $path$ 
 $\mu \leftarrow \frac{\lambda}{2}$ 
while NotDone do
  for  $i \leftarrow 1$  to  $\lambda$  do
     $q_i \leftarrow N(m, C)$ 
     $v_i \leftarrow Objective(q_i, observations)$ 
  end for
   $Sort(v, q)$ 
   $m \leftarrow mean(q_1 : q_\mu)$ 
  Update  $path$  using  $m$ 
  Update  $\sigma$  using  $path$ 
  Update  $C$  using  $path, \sigma$ 
  Determine NotDone
end while
Return  $(m, Objective(mean, Observations))$ 

```

---



---

### Algorithm 2 Basic pCMAES

---

```

Require:  $obs \in \mathbb{R}^O, Evaluate(\mathbb{R}^D, \mathbb{R}^O) \Rightarrow \mathbb{R}, \lambda \in \mathbb{Z}_+$ 
for all Processors  $p \in P$  do
  initialize Random Seed on  $p$ 
  initialize mean  $m$ , covariance  $C$ , step size  $\sigma$ ,
  path(s)  $path$ 
   $\mu \leftarrow \frac{\lambda}{2}$ 
  while NotDone do
    for  $i \leftarrow 1$  to  $\frac{\lambda}{\|P\|}$  do
       $q_i \leftarrow N(m, C)$ 
       $v_i \leftarrow Evaluate(q_i, observations)$ 
    end for
     $Sort(v, q)$ 
     $m \leftarrow mean(q_1 : q_\mu)$ 
    Update  $path$  using  $m$ 
    Update  $\sigma$  using  $path$ 
    Update  $C$  using  $path, \sigma$ 
    Determine NotDone
  end while
  Reduce Best Solution
end for

```

---

## 4 Test Configuration

### 4.1 Issues

Shared lab. Other processes running including time intensive

## **5 Results**

## **6 Conclusions**

## **7 Future Work**

LAPACK implementation for generation of next generation. Streamline the sorts. Evaluate overprescribing threads on OpenMP portion.