A scalable online algorithm for passive seismic tomography in underground mines

Claire Delplancke

University of Bath

NA seminar

Joint work with Joaquín Fontbona and Jorge Prado (U. de Chile)
Collaboration with CODELCO-Chile
Partially supported by the supercomputing infrastructure of NLHPC

Overview

1. Motivation and problem

2. Model

3. Algorithm

4. Results on synthetic data

5. Conclusions and perspectives

1. Motivation and problem
2. Model
3. Algorithm
4. Results on synthetic data
5. Conclusions and perspectives

El Teniente (ET) mine

Figure: Largest underground copper mine in the world. Located 150 kilometers south from Santiago and 2300 a.m.s.l., it began production in 1905 and has today more than 3000 km of underground galleries.

Exploitation method in ET: Block caving

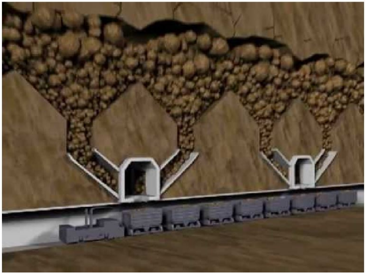


Figure: Block caving

Large scale exploitation as in ET induces:

- potentially dangerous stress redistribution within rock masses
- evolving, localized microseismic activity (magnitude < 2, 10^5 per/day, unknown hypocenters)
- geomechanical instabilities

Hazardous events: rockbursts

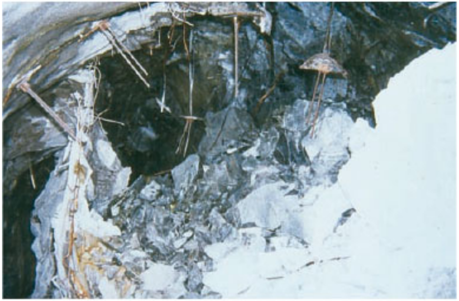


Figure 3—Rockburst damage showing protruding and failed support elements (photograph W.D. Ortlepp)

Figure: Rockbursts are violent explosions of rocks and mine structures owed to a sudden release of mechanical energy, through a large (>3) local seismic event

Figure: Seismic response of rock masses to mining must be taken into account both in design and operation stages.

Microseismic monitoring nowadays used to:
- localize seismic sources
- locally quantify energy released as seismic activity
- detect unstable local mechanical episodes (hardening, softening)

relying on predetermined, homogeneous velocity models.

Figure: Seismic response of rock masses to mining must be taken into account both in design and operation stages.

Microseismic monitoring nowadays used to:

- localize seismic sources
- locally quantify energy released as seismic activity
- detect unstable local mechanical episodes (hardening, softening)

relying on predetermined, homogeneous velocity models.

How can we better estimate the velocity field?

Tomography

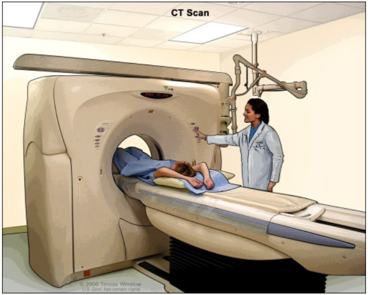


Figure: Medical imaging

Example of **active tomography:**
controlled wave sources (known emission position, time, and wave form...) .

Seismic Tomography: use seismic waves instead !

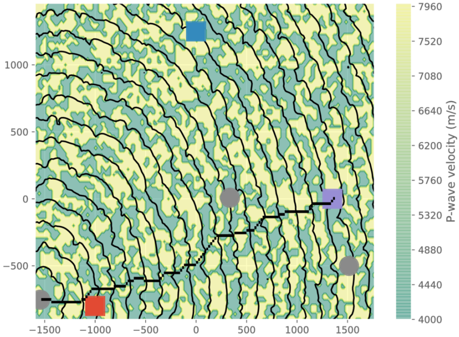


Figure: Minimum raypath and wavefront propagation in non-homogeneous velocity field.

First arrival times of seismic waves to sensors depend on non homogeneous medium and velocity field and contain information about it

Passive seismic tomography

- Use mining-induced microseisms to estimate velocity field.

Passive seismic tomography

- Use mining-induced microseisms to estimate velocity field.
- Temporal-spatial seismic sources are **unknown**.
  → Requires simultaneous inversion of sources and velocity field

Passive seismic tomography

- Use mining-induced microseisms to estimate velocity field.
- Temporal-spatial seismic sources are **unknown**.
  $\rightarrow$ Requires simultaneous inversion of sources and velocity field
- El Teniente: **Lots of data: 5 seisms per minute,** $\sim 10^6$ **a year,** ~ 100 **sensors.**

Non-linear inverse problem, with latent variables and large dataset

1 Motivation and problem

2 Model

3 Algorithm

4 Results on synthetic data

5 Conclusions and perspectives

Data

Data: First arrival time of *P*-wave.

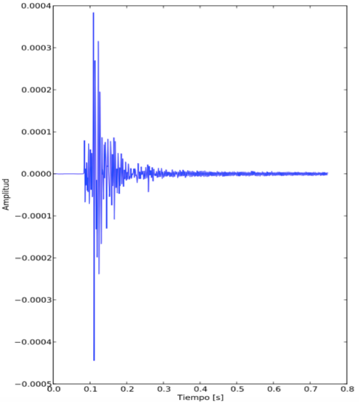


Figure: Sismogram

Forward problem

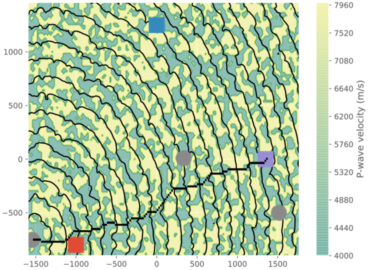


Figure: Propagation of a seismic wave in an isotropic, non-homogeneous medium

Let s be the (scalar) slowness field, x^0 the hypocenter of the microseism, and x any point in the spatial domain. Travel time:

$$(x^0, x, s) \to F(x^0, x, s)$$

Forward problem

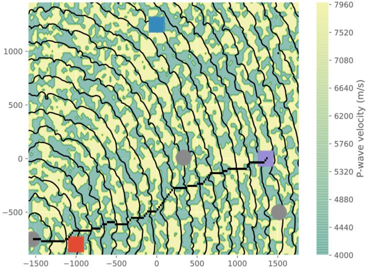


Figure: Propagation of a seismic wave in an isotropic, non-homogeneous medium

- Fermat's principle

$$F(x^0, x, s) = \min_{c \in \mathcal{C}(x^0, x)} \int_c s(y) dy$$

Forward problem

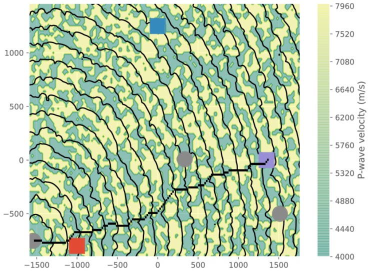


Figure: Propagation of a seismic wave in an isotropic, non-homogeneous medium

- Fermat's principle

$$F(x^0, x, {\color{red}s}) = \min_{c \in \mathcal{C}(x^0, x)} \int_c {\color{red}s}(y)dy$$

- Eikonal equation

$$\left|\nabla_x F(x^0, x, {\color{red}s})\right| = {\color{red}s}(x), \qquad F(x^0, x^0, {\color{red}s}) = 0.$$

Forward model

Let's discretize the spatial domain D on a finite grid ($\approx 10^6$ cells).

- **Efficient solvers** based on Eikonal equation. For a given x^0 and s, provides the value of $F(x^0, x, s)$ for all x.
  M. Noble et al, Accurate 3-d finite difference computation of traveltimes in strongly heterogeneous media. *Geophysical Journal International*, 2014.

Forward model

Let's discretize the spatial domain D on a finite grid ($\approx 10^6$ cells).

- **Efficient solvers** based on Eikonal equation. For a given x^0 and s, provides the value of $F(x^0, x, s)$ for all x.
  M. Noble et al, Accurate 3-d finite difference computation of traveltimes in strongly heterogeneous media. *Geophysical Journal International*, 2014.
- Fermat's principle: **gradient** of the forward function with respect to slowness field s.

$$F(x^0, x, s) = \min_{c \in \mathcal{C}(x^0, x)} \int_c s(y)dy = \min_{c \in \mathcal{C}(x^0, x)} \int_D \mathbf{1}_c(y) s(y) dy.$$

By the envelope theorem,

$$\nabla_s F(x^0, x, s) = \mathbf{1}_{\text{raypath}_{\text{min}}(x^0, x, s)}.$$

Model

M microseismic events, N sensors. Dataset:

$$\mathcal{D} = \left(\overline{t}^i\right)_{1 \leq i \leq M},$$

where $\overline{t}^i$ stands for the vector of first arrival times of i-th event:

$$\overline{t}^i = \begin{pmatrix} t_1^i \\ t_2^i \\ \dots \\ t_N^i \end{pmatrix} = t^{0(i)} + \begin{pmatrix} F\left(x^{0(i)}, x_{r(1)}, s\right) \\ F\left(x^{0(i)}, x_{r(2)}, s\right) \\ \dots \\ F\left(x^{0(i)}, x_{r(N)}, s\right) \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \dots \\ \epsilon_N \end{pmatrix},$$

with independent Gaussian recording errors $\epsilon_j \sim \mathcal{N}(0, \sigma^2)$.

Conditional likelihood of one microseismic event:

$$p\left(\overline{t} \mid t^0, x^0, s\right) \sim \exp\left(-\frac{1}{2\sigma^2} \sum_{j=1}^{N} \left(t_j - t^0 - F\left(x^0, x_{r(j)}, s\right)\right)^2\right).$$

Model

By integrating out the latent variables $\left(t^0, x^0\right)$ corresponding to hypocenter time and location,

$$
\begin{aligned}
& p\left(\bar{t} \mid s\right) \\
= & \int p\left(\bar{t} \mid \left(t^0, x^0\right), s\right) p_{\text{prior}}\left(t^0, x^0\right) dt^0\, dx^0 \\
\sim & \int \exp\left(-\frac{1}{2\sigma^2} \sum_{j=1}^{N} \left(t_j - t^0 - F\left(x^0, x_{r(j)}, s\right)\right)^2\right) p_{\text{prior}}\left(t^0, x^0\right) dt^0\, dx^0.
\end{aligned}
$$

Under assumption of independence of microseismic events:

$$
p\left(\mathcal{S} \mid s\right) = \prod_{\bar{t} \in \mathcal{S}} p\left(\bar{t} \mid s\right).
$$

Model

Then, by Bayes' theorem,

$$p\,(s\,|\,\mathcal{S}) = \prod_{\bar{t}\in\mathcal{S}} p\,(\bar{t}\,|\,s)\,p_{\text{prior}}(s).$$

Mathematical formulation of the problem

$$\arg\max_{s} \quad \mathcal{O}(s|\mathcal{S}),$$

where the objective function is the (renormalized) log-posterior distribution on s:

$$\mathcal{O}(s,\mathcal{S}) := \frac{1}{M}\sum_{\bar{t}\in\mathcal{S}}\left(\log p\,(\bar{t}\,|\,s) + \frac{1}{M}p_{\text{prior}}(s)\right).$$

Gradient of the objective

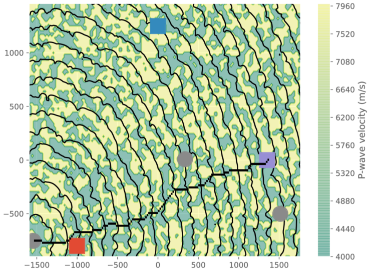


Figure: Minimum raypath in non-homogeneous velocity field. Vector $\text{raypath}_{\text{min}}$ indicates which cells of the discretized mine are visited by the minimum raypath.

$$\nabla_s \log\, p\left(\bar{t}\,|\,s\right) = \frac{1}{\sigma^2}\sum_{j=1}^{N}\int \text{raypath}_{\text{min}}\left(x^0, x_{r(j)}, s\right)\left(t_j - t^0 - F\left(x^0, x_{r(j)}, s\right)\right)$$

$$\cdot\; p_{\text{post}}\left(t^0, x^0 \middle| \bar{t},\, s\right) dt^0\, dx^0.$$

1. Motivation and problem
2. Model
3. Algorithm
4. Results on synthetic data
5. Conclusions and perspectives

Stochastic gradient descent

To perform a classical gradient descent, we need the gradient of the objective function

$$\nabla_s \mathcal{O}(s, \mathcal{S}) = \frac{1}{M} \sum_{\bar{t} \in \mathcal{S}} \nabla_s \log\, p\,(\bar{t}\,|\,s)\,.$$

Too heavy to compute at each iteration!

→ noised version of the gradient, computed on a mini-batch of data $\mathcal{B}$ with $|\mathcal{B}| \ll M$:

$$\nabla_s \mathcal{O}(s, \mathcal{S}) \approx \nabla_s \mathcal{O}(s, \mathcal{B}) = \frac{1}{|\mathcal{B}|} \sum_{\bar{t} \in \mathcal{B}} \nabla_v \log\, p\,(\bar{t}\,|\,v)\,.$$

Algorithm

Input: sensors' locations $\left(x_{r(j)}\right)_{1\leq j\leq M}$, $P-$wave triggers data $\left(\bar{t}^i\right)_{1\leq i\leq M}$.

Parameters: step size γ, mini-batch size $m = |\mathcal{B}|$.

```
Initialize slowness field s .


Iterate: update s .
```

Algorithm: update rule

- Under current slowness field s , compute for all x, x_r travel-times $F(x, x_r, s)$ (back-propagation, FORTRAN).
- Pick a new mini-batch of data $\mathcal{B}$.
- For each event in $\mathcal{B}$, compute the a posteriori distribution of the hypocenter, $p_{\text{post}}\left(t^0, x^0 | \bar{t}, s\right)$.
- Compute the minimum raypath between each likely hypocenter and sensor, $\text{raypath}_{\text{min}}\left(x^0, x_{r(j)}, s\right)$ (FORTRAN).

Algorithm: update rule

- Compute the noised gradient:

$$\nabla_s \mathcal{O}(s,\mathcal{B}) = \frac{1}{|\mathcal{B}|\,\sigma^2} \sum_{\bar{t}\in\mathcal{B},\,1\le j\le N} \int dt^0\,dx^0\,\left(t_j - t^0 - F\left(x^0, x_{r(j)}, s\right)\right) \cdot \mathrm{raypath}_{\mathrm{min}}\left(x^0, x_{r(j)}, s\right) p_{\mathrm{post}}\left(t^0, x^0 | \bar{t},\, s\right).$$

- Update s :

$$s \leftarrow s + \gamma\,\nabla_s \mathcal{O}(\mathcal{S}, s).$$

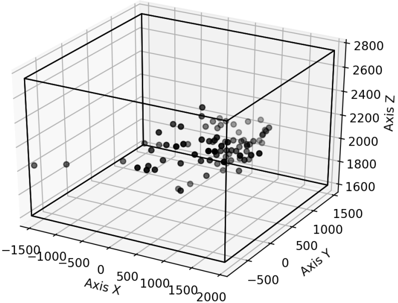
1. Motivation and problem
2. Model
3. Algorithm
4. Results on synthetic data
5. Conclusions and perspectives

Generation of synthetic data

The synthetic dataset $\mathcal{D}$ is generated with:

- 100^5 hypocenters sampled uniformly in spatial domain. We use batches of size $m = 500$.

- Sensor array: **real** one of El Teniente mine.



Synthetic velocity fields: checkerboards

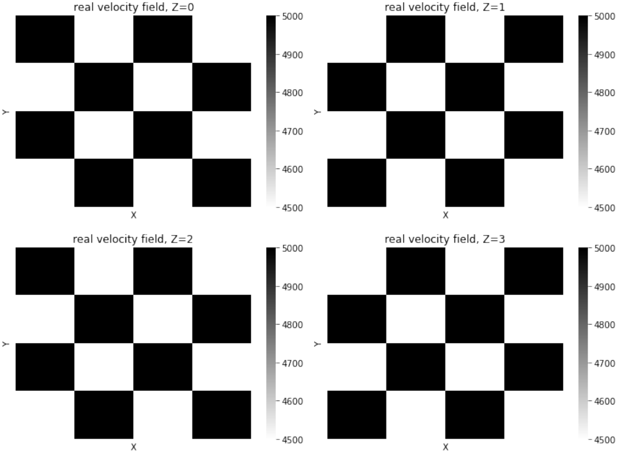
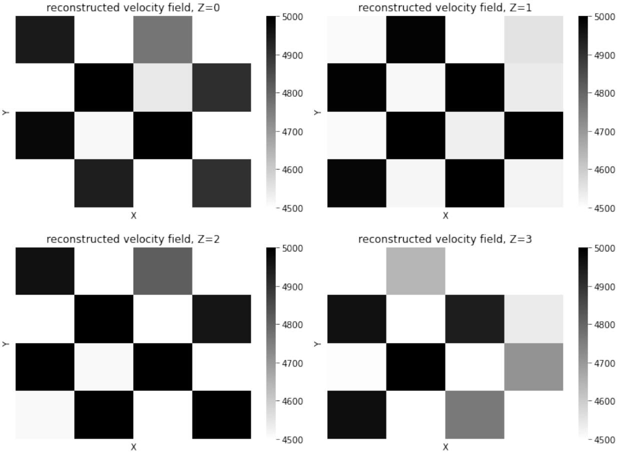
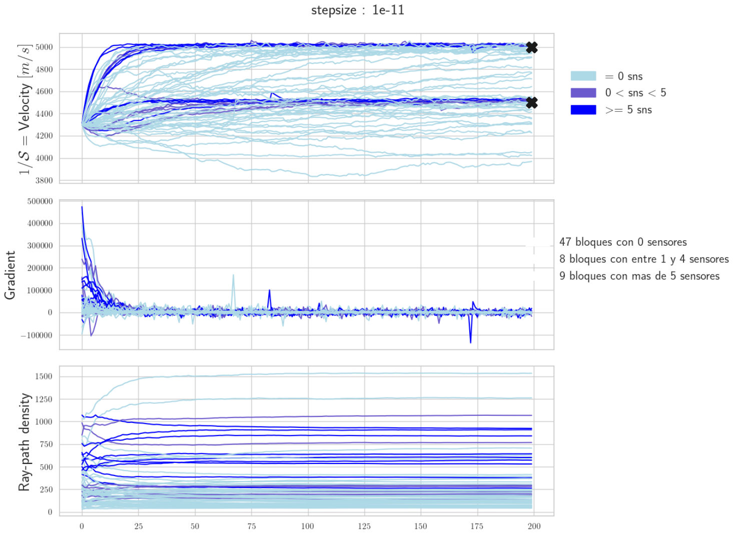


Figure: Checkerboard: two possible values

Checkerboard reconstruction, 64 blocks, El Teniente sensors array



Checkerboard reconstruction, 64 blocks, El Teniente sensors array



47 bloques con 0 sensores
8 bloques con entre 1 y 4 sensores
9 bloques con mas de 5 sensores

Checkerboard accuracy reconstruction, El Teniente sensors array

Reconstruction of the checkerboard velocity model		
number of blocks	constant learning rate	adaptive learning rate
8	63 %	100 %
27	37 %	81 %
64	34 %	78 %

Table: Accuracy of the reconstruction up to a 2% discrepancy

Marmousi model

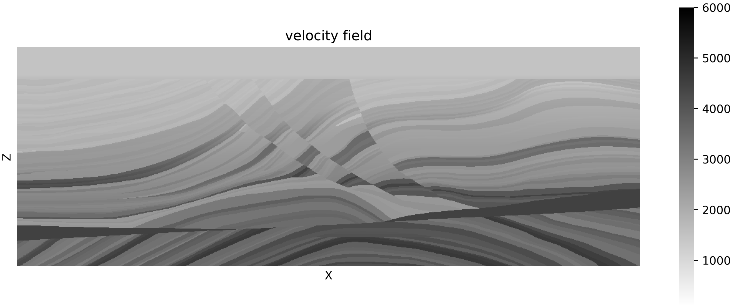


Figure: True Marmousi velocity field

Marmousi model

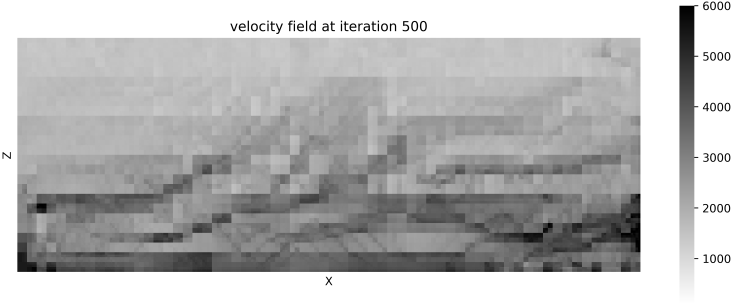


Figure: Reconstruction with our algorithm

Marmousi model: inner workings

True number of cells in Marmousi model: 351000.

Marmousi model: inner workings

True number of cells in Marmousi model: 351000.

- The number of iterations is divided into loops. In the first loop, the velocity field model used in the reconstruction consists of 3 blocks. At the beginning of each successive loop, the number of blocks is multiplied by 4. Total number of blocks at the end is 98304.
- We start the iterations assuming a larger value of the standard deviation than the true one, and decrease it at the beginning of each new loop.

1. Motivation and problem

2. Model

3. Algorithm

4. Results on synthetic data

5. Conclusions and perspectives

Conclusions

- Algorithm is able to quickly reconstruct synthetic velocity field from a streaming of simulated noisy first-arrival time records
- Reconstruction accuracy depends on raypath density.
- Algorithm can extend to other (deterministic) parametrization of the velocity field
- Parallel implementation enables systematic use.

Perspectives

- Mathematical and ML improvements
  - Include regularization in the prior on the slowness field
  - Bayesian sampling on the a posteriori on the slowness field
  - More astute parametrization of the velocity field
- **What about real data?**

Perspectives

- Mathematical and ML improvements
  - Include regularization in the prior on the slowness field
  - Bayesian sampling on the a posteriori on the slowness field
  - More astute parametrization of the velocity field
- **What about real data?** no data quality standard, black-box preprocessing, systematic errors badly understood

Monitoring on test dataset

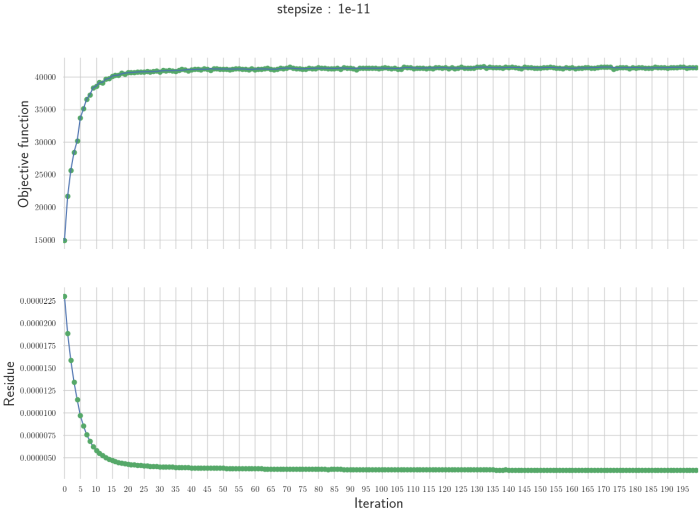


Figure: Monitoring on test dataset