# Towards Combining Interval and Probabilistic Uncertainty in Finite Element Methods

Vladik Kreinovich and Roberto Araiza
Department of Computer Science

Matthew G. Averill
Department of Geological Sciences

University of Texas at El Paso
El Paso, TX 79968, USA
emails vladik@utep.edu
raraiza@utep.edu, averill@geo.utep.edu

# 1. Uncertainty in FEM: Probabilistic Approach

- *Traditional approach to FEM:* we know the *exact* equations and the *exact* values of the parameters $x_i$ of these equations.
- *In practice:* we only know the *approximate* values of the corresponding parameters.
- *Question:* estimate how the uncertainty in the parameters of the system can affect the result $y$ of applying the FEM techniques.
- *Probabilistic approach – assumption:* we know the exact *probability* distributions of all $x_i$.
- *Probabilistic approach – algorithm:* e.g., use Monte-Carlo simulations.

## 2. Uncertainty in FEM: Interval Approach and Remaining Problem

- *Interval approach – assumption:* sometimes, we only know the lower bounds $\underline{x}_i$ and the upper bounds $\overline{x}_i$ on $x_i$.
- *Example:* the probabilities of different values of Young modulus may depend on the manufacturing process.
- *Interval approach – objective:* find the interval

$$[\widetilde{y} - \Delta, \widetilde{y} + \Delta]$$

  that contains $y$.
- *Interval approach – algorithms:* presented in talks by Muhanna, Nakao, Neumaier, Pownuk.

## 3. Uncertainty in FEM: Remaining Problem

- *In practice:* we sometimes have *both* interval and probabilistic uncertainty.
- *Example:*
  - for manufacturing-related parameters, we may only know intervals of possible values;
  - for weather-related parameters, we also know the probabilities of different values (e.g., from the weather records).
- *Objective:* for different $p \in [0, 1]$, find the value $\Delta(p)$ that bound $\Delta y$ with probability $p$.

## 4. General Algorithm

- *Problem – reminder:*
  - for some $x_i$, we know intervals $[\underline{x}_i, \overline{x}_i]$;
  - for some $x_j$, we know the probability distribution;
  - we want to find the value $\Delta(p)$ that bound $\Delta y$ with probability $p$.
- *Algorithm:*
  - use Monte-Carlo techniques to simulate parameters $x_j$ with known probability distributions;
  - for each such simulation, use interval FEM techniques to get an upper bound $\Delta$ for $|y - \widetilde{y}|$;
  - after several simulations, we get the resulting bounds distribution;
  - from this distribution, we can find the desired bound $\Delta(p)$.

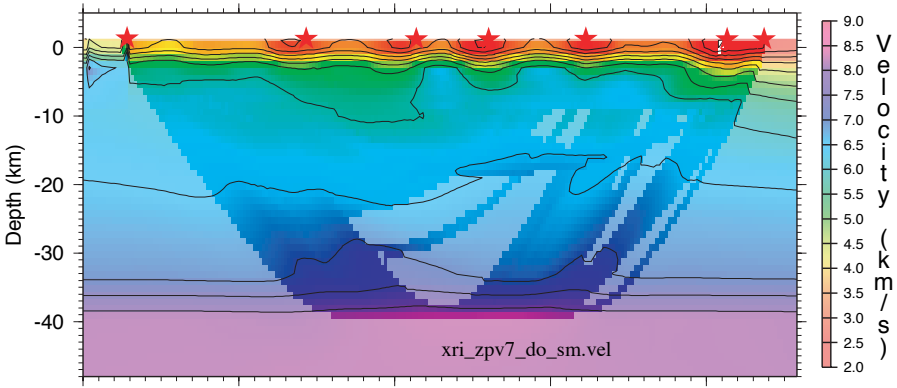
## 5. Case Study: Determining Earth Structure

- *Importance:* civilization greatly depends on the things we extract from the Earth: oil, gas, water.
- *Need* is growing, so we must find new resources.
- *Problem:* most easy-to-access mineral resources have been discovered.
- *Example:* new oil fields are at large depths, under water, in remote areas – so drilling is very expensive.
- *Objective:* predict resources before we invest in drilling.
- *How:* we know what structures are promising.
- *Example:* oil and gas concentrate near the top of (natural) underground domal structures.
- *Conclusion:* to find mineral resources, we must determine the structure at different depths $z$ at different locations $(x, y)$.

# 6. Data that We Can Use to Determine the Earth Structure

- *Available measurement results:* those obtained without drilling boreholes.
- *Examples:*
  - gravity and magnetic measurements;
  - travel-times $t_i$ of seismic ways through the earth.
- *Need for active seismic data:*
  - passive data from earthquakes are rare;
  - to get more information, we make explosions, and measure how the resulting seismic waves propagate.
- *Resulting seismic inverse problem:*
  - we know the travel times $t_i$;
  - we want to reconstruct velocities at different depths.

# Hole Tomography Smashed Masked Velocity Models



## 7. Seismic Inverse Problem: Towards Mathematical Formulation

- *General description:* wave equation with unknown $v(x)$.
- *Difficulty:* due to noise, we only know $t_i$.
- *Ray approximation:* a seismic wave follows the shortest path $t = \int \frac{d\ell}{v} \to \min$; Eikonal equation $|\nabla t| = \frac{1}{v}$.
- *Discontinuity:* $v$ is only piece-wise continuous; Snell's law describes the transition $\frac{\sin(\varphi)}{v} = \frac{\sin(\varphi')}{v'}$.
- *Ill-posed problem:* a change in $v$ outside paths does not affect observed travel-times $t_i$; hence, many drastically different $v(x)$ are consistent with observations.
- *Current solution:*
  - start with a meaningful first approximation;
  - use physically motivated iterations.

## 8. Seismic Inverse Problem: FEM Approach

- $v(x) = v_j$ is constant within each element $j$.
- *We know:* travel-times $t_i$ between known points $A_i$, $B_i$.
- *We want to find:* velocities $v_j$ for which $t_i = t_i(v) \stackrel{\text{def}}{=} \min \sum_j \frac{\ell_{ij}}{v_j}$, where:
  - min is taken over all paths between $A_i$ and $B_i$, and
  - $\ell_{ij}$ is the length of the part of $i$-th path within element $j$.
- *Shortest path:* straight inside each element, Snell's law on each border.
- *Simplification:* use slownesses $s_j \stackrel{\text{def}}{=} \frac{1}{v_j}$; $t_i = \sum_j \ell_{ij} \cdot s_j$.
- *Problem:* system is under-determined.

# 9. Existing Algorithm for the Seismic Inverse Problem: General Description

- *The most widely used:* John Hole's iterative algorithm.
- *Starting point:* reasonable initial slownesses $s_j^{(0)}$.
- *Starting an iteration:* we use current (approximate) slownesses $s_j^{(k)}$ to:
  - find the shortest paths and
  - compute the corr. travel-times $t_i = \sum_j \ell_{ij} \cdot s_j^{(k)}$.
- *Fact:* measured travel-times $\widetilde{t}_i$ are somewhat different: $\Delta t_i \stackrel{\text{def}}{=} \widetilde{t}_i - t_i \neq 0$.
- *On each iteration:*
  - we find $\Delta s_j$ for which $\sum_j \ell_{ij} \cdot (s_j + \Delta s_j) = \widetilde{t}_i$;
  - we take $s_j^{(k+1)} = s_j^{(k)} + \Delta s_j$.

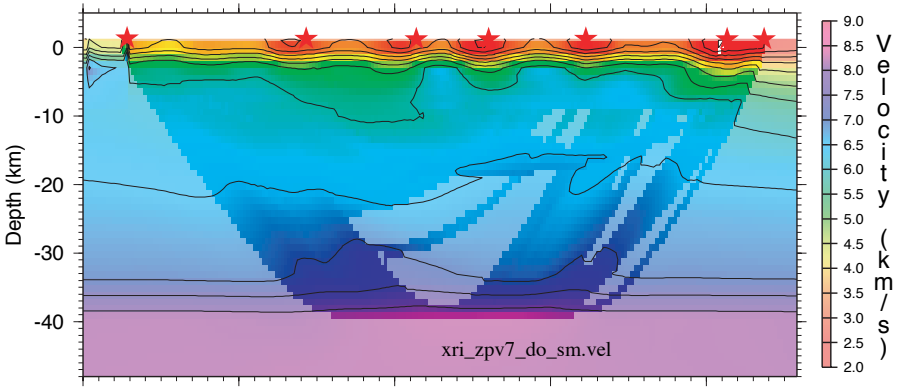
## 10. Algorithm for the Inverse Problem: Details

- *Objective (reminder):* find $\Delta s_j$ s.t. $\sum \ell_{ij} \cdot \Delta s_j = \Delta t_i$.
- *Simplest case:* one path.
- *Specifics:* under-determined system: 1 equation, many unknowns $\Delta s_j$.
- *Idea:* no reason for $\Delta s_j$ to be different: $\Delta s_j \approx \Delta s_{j'}$.
- *Formalization:* minimize $\sum\limits_{j,j'} (\Delta s_j - \Delta s_{j'})^2$ under the constraint $\sum \ell_{ij} \cdot \Delta s_j = \Delta t_i$.
- *Solution:* $\Delta s_j = \dfrac{\Delta t_i}{L_i}$ for all $j$, where $L_i = \sum\limits_j \ell_{ij}$.
- *Realistic case:* several paths; we have $\Delta s_{ij}$ for different paths $i$.
- *Idea:* least squares $\sum\limits_i (\Delta s_j - \Delta s_{ij})^2 \to \min$.
- *Solution:* $\Delta s_j$ is the average of $\Delta s_{ij}$.

## 11. Successes, Limitations, Need for Prior Knowledge

- *Successes:* the algorithm usually leads to reasonable geophysical models.
- *Limitations:* often, the resulting velocity model is not geophysically meaningful.
- *Example:* resulting velocities outside of the range of reasonable velocities at this depth.
- *What is currently done:* trying different initial models (hacking).
- *Problem with this approach:* there is no algorithm for selecting a good starting model; often, dozens of tries are needed – and each try requires hours of computations.
- *It is desirable:* to incorporate the expert knowledge into the algorithm for solving the inverse problem.

# Hole Tomography Smashed Masked Velocity Models



## 12. Case of Interval Prior Knowledge

- *Additional information:* an interval $[\underline{s}_j, \overline{s}_j]$ that contains the (unknown) actual value $s_j$.
- *Problem:* for $\Delta s_{ij} = \dfrac{\Delta t_i}{L_i}$, we may have $s_j^{(k)} + \Delta s_{ij} \notin [\underline{s}_j, \overline{s}_j]$.
- *Idea:* as $s_j^{(k+1)}$, take the value from $[\underline{s}_j, \overline{s}_j]$ which is the closest to $s_j^{(k)} + \Delta s_{ij}$, i.e., cut off at $\overline{s}_j$ (or at $\underline{s}_j$).
- *Problem:* since we decreased $\Delta s_{ij}$, we have a remaining discrepancy $\Delta t_i' \stackrel{\text{def}}{=} \Delta t_i - \sum\limits_j \ell_{ij} \cdot \Delta s_{ij}$.
- *Solution:* repeat the same process for $\Delta t_i'$, etc.
- *Problem:* many iterations instead of one – increase in computation time.
- *Our result:* a new linear time algorithm that computes the final result of these additional iterations.

# 13. A New Linear-Time Algorithm

- At each iteration, we have three sets:
  - $J^- = \{j : \text{we know} \Delta s_{ij} = \overline{\Delta}_j \stackrel{\text{def}}{=} \overline{s}_j - s_j^{(k)}\}$;
  - $J^+ = \{j : \text{we know} \Delta s_{ij} < \overline{\Delta}_j\}$,
  - $J = -(J^- \cup J^+)$.

  and quantities $A^- \stackrel{\text{def}}{=} \sum\limits_{j \in J^-} \ell_{ij} \cdot \overline{\Delta}_j$ and $\mathcal{L}^+ \stackrel{\text{def}}{=} \sum\limits_{j \in J^+} \ell_{ij}$.
- We start with $J^- = J^+ = \emptyset$ and $J = \{1, \ldots, c\}$.
- At each iteration:
  - we compute the median $m$ of the set $J$ (median in terms of sorting by $\overline{\Delta}_j$);
  - then, by analyzing the elements of the undecided set $J$ one by one, we divide them into subsets

$$P^- \stackrel{\text{def}}{=} \{j : \overline{\Delta}_j \le \overline{\Delta}_m\}, \quad P^+ \stackrel{\text{def}}{=} \{j : \overline{\Delta}_j > \overline{\Delta}_m\}.$$

## 14. A New Linear-Time Algorithm (cont-d)

- we compute $a^- \stackrel{\text{def}}{=} A^- + \sum\limits_{j \in P^-} \ell_{ij} \cdot \overline{\Delta}_j$ and $\ell^+ \stackrel{\text{def}}{=} \mathcal{L}^+ + \sum\limits_{j \in P^+} \ell_{ij}$;
- then, we compute $\Delta s = \dfrac{\Delta_i - a^-}{\ell^+}$; also, among all the values from $P^+$, we select the smallest value, which we will denote by $\overline{\Delta}_{(p+1)}$;
- if $\Delta s > \overline{\Delta}_{(p+1)}$, then we replace $J^-$ with $J^- \cup P^-$, $A^-$ with $a^-$, and $J$ with $P^+$;
- if $\Delta s \le \overline{\Delta}_m$, then we replace $J^+$ with $J^+ \cup P^+$, $\mathcal{L}^+$ with $\ell^+$, and $J$ with $P^-$;
- finally, if $\overline{\Delta}_m < \Delta s \le \overline{\Delta}_{(p+1)}$, then we replace $J^-$ with $J^- \cup P^-$, $J^+$ with $J^+ \cup P^+$, and $J$ with $\emptyset$.

- Iterations continue until $J = \emptyset$.
- Return $\Delta s_{ij} = \overline{\Delta}_j$ when $\overline{\Delta}_j \le \overline{\Delta}_m$, else $\Delta s_{ij} = \Delta s$.

## 15. Case of Probabilistic Prior Knowledge

- *Description:* from prior observations, we know $\widetilde{s}_j \approx s_j$, and we know the st. dev. $\sigma_j$ of this value.
- *Minimize:* $\sum\limits_{j,j'}(\Delta s_{ij} - \Delta s_{ij'})^2$ s.t. $\sum\limits_{j=1}^{c} \ell_{ij} \cdot \Delta s_{ij} = \Delta t_i$ and

$$\frac{1}{n} \cdot \sum_{j=1}^{c} \frac{((s_j^{(k)} + \Delta s_{ij}) - \widetilde{s}_j)^2}{\sigma_j^2} = 1.$$

- *Solution* (Lagrange multipliers): $\overline{\Delta s} \stackrel{\text{def}}{=} \dfrac{1}{n} \cdot \sum\limits_{j=1}^{c} \Delta s_{ij}$,

$$\frac{2}{n} \cdot \Delta s_{ij} - \frac{2}{n} \cdot \overline{\Delta s} + \lambda \cdot \ell_{ij} + \frac{2\mu}{n \cdot \sigma_j^2} \cdot (s_j^{(k)} + \Delta s_{ij} - \widetilde{s}_j) = 0.$$

- *Fact:* $\Delta s_{ij}$ is an explicit function of $\lambda$, $\mu$, $\overline{\Delta s}$.
- *Algorithm:* solve 3 non-linear equations (above one + 2 constraints) with unknowns $\lambda$, $\mu$, $\overline{\Delta s}$.

# 16. Combination of Different Types of Prior Knowledge

- *Need:* we often have both:
  - prior measurement results – i.e., *probabilistic* knowledge, and
  - expert estimates – i.e., *interval* knowledge.
- *Minimize:* $\sum\limits_{j,j'}(\Delta s_{ij} - \Delta s_{ij'})^2$ s.t. $\sum\limits_{j=1}^{c} \ell_{ij} \cdot \Delta s_{ij} = \Delta t_i$,

$$\frac{1}{n} \cdot \sum_{j=1}^{c} \frac{((s_j^{(k)} + \Delta s_{ij}) - \widetilde{s}_j)^2}{\sigma_j^2} \le 1,$$

  and $\underline{s}_j \le s_j^{(k)} + \Delta s_{ij} \le \overline{s}_j$.
- *Idea:* we minimize a convex function under convex constraints; efficient algorithms are known.

## 17. Combination of Different Types of Prior Knowledge: Algorithm

- *Idea* – method of alternating projections:
  - first, add a correction that satisfy the first constraint,
  - then, the additional correction that satisfies the second constraint,
  - etc.
- *Specifics:*
  - first, add equal values $\Delta s_{ij}$ to minimize $\Delta t_i$;
  - restrict the values to the nearest points from $[\underline{s}_j, \overline{s}_j]$,
  - find the extra corrections that satisfy the probabilistic constraint,
  - repeat until converges.

## 18. Effects of Discretization: Technique

- *Worst-case bound:* $\infty$.
- *Geophysical approach – "checkerboard" method:*
  - add a doubly periodic function $\Delta v(x)$ to the solution $v(x)$;
  - compute corresponding $t'_i$;
  - reconstruct $v'(x)$ from $t'_i$;
  - compare $v'(x)$ with $v(x) + \Delta v(x)$.
- *Conclusion:*
  - For small-step $\Delta v(x)$, we will not see the difference – hence details of this size in $v(x)$ are not reliable.
  - If for some step $h$, we see the difference, this means that details of size $h$ are more reliable.

# 19. Effects of Discretization: Mathematical Results

- *Problem:* how to select functions $\Delta v_i(x)$?
- *Idea:* we want a shift-invariant process.
- *Analysis:* in linear approximation, what matters is the linear hull of the functions $\Delta v_i(x)$.
- *First result:* for bounded $\Delta v_i(x)$, we get

$$\Delta v_i(x_1, x_2) = \sin(a_1 \cdot x_1 + b_1) \cdot \sin(a_2 \cdot x_2 + b_2).$$

- *Alternative formulation:* we require that the family of functions $\{\Delta v_i(x)\}$ is optimal w.r.t. a shift-invariant optimality criterion.
- *Second result:* for each such criterion, we get the same functions

$$\Delta v_i(x_1, x_2) = \sin(a_1 \cdot x_1 + b_1) \cdot \sin(a_2 \cdot x_2 + b_2).$$

## 20. Acknowledgments

This work was supported in part by:

- NASA under cooperative agreement NCC5-209,
- NSF grants EAR-0225670 and DMS-0532645,
- Star Award from the University of Texas System, and
- Texas Department of Transportation grant No. 0-5453.