Quantifying uncertainty in tomographic problems with a statistical zipper model

Kris Innanen

ABSTRACT

The statistical "molecular zipper" model used by Kittel to analyze the unraveling of a strand of DNA and its relationship to temperature is ported to the problem of analysis of uncertainty and appraisal in tomographic inversion. Equilibrium methods due to Boltzmann, with an emphasis on the analytically-derived partition function of the zipper, are used to estimate the average contribution of a tomographic cell to the bulk properties of the data. This number is seen to depend on almost all features of the experiment we expect to impact the reliability (or at any rate the importance in explaining the data) of each cell of a tomographic model individually. Specific techniques for analyzing this number, and spatial maps of the number, especially as it varies with an artificial temperature (whose value reflects broadly fast versus slow geological structures), remain to be created.

INTRODUCTION

In this report a method for statistical characterization of uncertainty in tomographic inverse problems is set out. It is inspired by calculations of system averages in equilibrium statistical mechanics. These are transposed to become calculations of average contributions to bulk measurements of traveltimes in an experiment, given acquisition and discretization details. In particular, it brings over from statistical physics a model referred to as the "molecular zipper", devised to predict average behaviours of strands of DNA, and made popular by Kittel (1969). The logic of the zipper, in which links break according to simple discrete energy rules, ports quite smoothly over to that of a discrete slowness model, tomographic cells are assigned discrete slownesses, subject to rules regarding the traveltime cost of adding a discrete element to the slowness of a cell.

The methods of calculation, which focus on the construction and analysis of the partition function of the canonical ensemble, are not reviewed extensively here. See Innanen (2021) for a tutorial style review, which includes discussion of the zipper model.

What is created is a measure of the average contribution of each cell to the total traveltime in the data; this measure is seen to vary depending on the acquisition parameters, discretization features of the model both in space and in slowness dimensions, and on an artificial "temperature", a smoothly tunable parameter which can be set to mimic relatively slow or fast geological structures.

THE TOMOGRAPHY PROBLEM

Traveltime tomography

In the 2D straight-ray cross-well tomography problem, a plane is discretized into J grid cells, and source-receiver pairs are assumed to bracket the grid. A coarse example with a single source-receiver pair is illustrated in Figure 1a. The ray is assumed to comprise a

straight line segment connecting the source and receiver, and the datum associated with the ith ray is the traveltime $\tau(i)$.

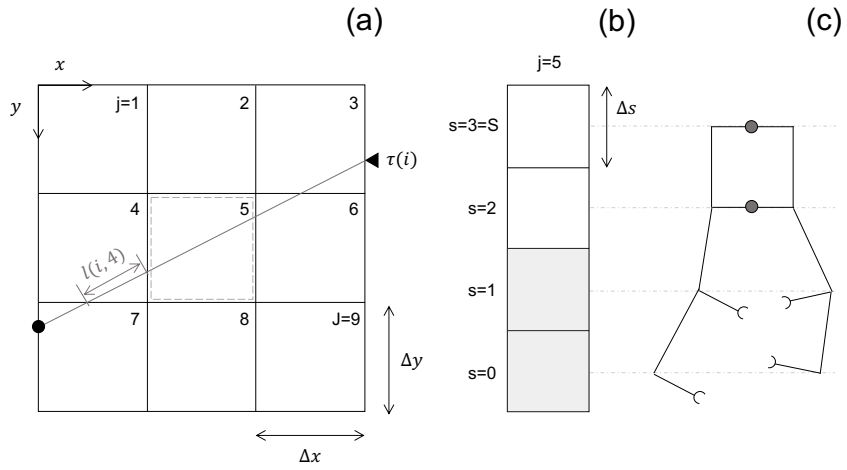

FIG. 1. (a) A coarse slowness grid, with cells numbered $j = 1, ..., J$. An example source and receiver pair, and their ray, are illustrated. Traveltime along the ray gives the ith datum $\tau(i)$. The ith ray has a length $l(i, j)$ in the jth cell; the $l(i, 4)$ segment is illustrated. (b) Each cell (e.g., $j = 5$) is assigned a discrete slowness. In this case $s = 1$, giving the impression of a pile of bricks, two bricks in height. (c) The brick pile plays the role of the molecular zipper, e.g., "2 bricks" $\sim$ "2 broken zipper links".

The grid cells are assigned slowness values selected from a discrete list. If the jth cell has slowness $\Delta s \times s(j)$, where $s(j)$ is an integer ranging from 0 to some maximum S, the traveltime $\tau(i)$ is

$$\tau(i) = \Delta s \sum_{j=1}^{J} l(i, j) s(j), \tag{1}$$

where $l(i, j)$ is the length of the ith ray in the jth cell (Figure 1a). A complete data set involves a large number of source receiver pairs, and thereby a large number of data $\tau(i)$, i=$(1, ..., I)$. A less-coarse example, with 100 randomly-chosen source-receiver pairs and their associated raypaths, is illustrated in Figure 2.

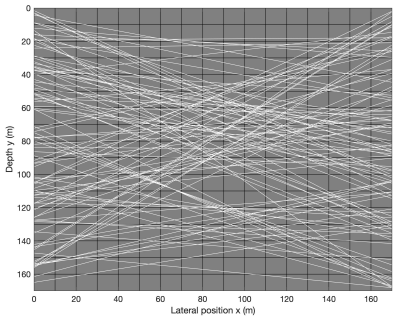

FIG. 2. An illustrative tomography example with a 17$\times$17 grid of slowness cells criss-crossed by the rays associated with 100 randomly selected source-receiver pairs.

Slowness of a grid cell as a "stack of bricks"

It will presently be helpful to have a concrete picture of the discrete slownesses. Let us consider the grid in Figure 1a to be a view from above of a set of 9 stacks of bricks. In Figure 1b we focus on the $j = 5$ cell, and look at it from the side, again choosing a very coarse discretization in which $S = 3$. The slowness ascribed to the $j = 5$ cell is Δs, so the slowness integer is $s(j = 5) = 1$. This can be pictured as a stack of "slowness bricks", at cell $j = 5$, which is two bricks high.

A STATISTICAL MODEL OF THE CONTRIBUTIONS OF SLOWNESS GRID CELLS TO TOMOGRAPHIC DATA

Let us now apply Boltzmann statistical methods to this problem. We will make use of the molecular zipper model set out in the review paper companion to this one (Innanen, 2021).

The molecular zipper

In the molecular zipper, a DNA molecule is modelled as a ladder-shaped superstructure with R rungs, or links. The molecule operates like a zipper, in that links can break, but for link r to break, link $r - 1$ must already be broken. Statistics for a single molecule afloat in a heat bath at temperature T provide the probability for the molecule to have r open links:

$$P(r) = \frac{1}{Z(\beta)} e^{-\beta\varepsilon r}, \tag{2}$$

where $\beta = 1/T$. The partition function Z is the sum over all possible configurations and their energies; because it costs energy ε to break link r, but all links below r must be broken for this amount to be paid, the only possible energy values the zipper can take on are 0, ε, 2ε, ..., $(R-1)\varepsilon$, so:

$$Z(\beta) = \sum_{r=0}^{R-1} e^{\beta\varepsilon r} = \frac{1 - X^R}{1 - X}, \text{ where } X = e^{-\beta\varepsilon}. \tag{3}$$

The upper limit of $R - 1$ is a restriction that the molecule cannot completely separate.

The zipper model applied to the tomographic problem

Let any one selected grid cell in the tomography problem be associated with the single molecule in the DNA model, such that its stack of bricks (Figure 1b) is matched with an upside-down version of the zipper (Figure 1c). This association takes advantage of the logical similarity between a stack and a zipper: bricks can be added to a slot ($\sim$ links can be broken) only if all lower bricks are in place ($\sim$ only if all lower links are broken).

It costs energy to break a link in the zipper, and it costs traveltime to add a discrete element of slowness to a grid cell. This suggests traveltime should play the role of energy in our use of the model. In order to engage the full tomography dataset, let us start with $\mathcal{T}$, the

total traveltime of all of the rays in the experiment, which is, using (1),

$$\mathcal{T} = \sum_{i=1}^{I} \tau(i) = \sum_{i=1}^{I} \left[\Delta s \sum_{j=1}^{J} l(i,j) s(j) \right]. \tag{4}$$

Switching the order of summation, we re-express this as

$$\mathcal{T} = \sum_{j=1}^{J} s(j) L(j) \Delta s, \tag{5}$$

where

$$L(j) = \sum_{i=1}^{I} l(i,j) \tag{6}$$

is the total length of all rays in cell j. The summand in (5) is the contribution to the total traveltime of the jth cell. This will be our choice to replace the energy εr in the zipper model. Thus the probability of the slowness integer s occurring in the jth cell is

$$P(s,j) = \frac{1}{Z(\beta,j)} e^{-\beta s L(j) \Delta s}, \tag{7}$$

where

$$Z(\beta,j) = \sum_{s=0}^{S-1} e^{-\beta s L(j) \Delta s} = \frac{1 - X^S(j)}{1 - X(j)}, \tag{8}$$

and

$$X(j) = e^{-\beta L(j) \Delta s}. \tag{9}$$

Comparing this setup with (2)-(3), the main difference is the presence of j: we evidently should expect different statistical behaviour from grid cell to grid cell, on account of $L(j)$. This places the *total ray length occupying each cell* in a central role in the analysis. It is the quantity that distinguishes one grid cell from another, and that brings together cell dimensions, ray coverage, and slowness values, and relates them to traveltime data. The internal logic of statistical mechanics has been suggestive that we give $L(j)$ this part to play, but we should still consider it to be a sort of hypothesis.

In the zipper model (and in statistical mechanics generally), β is interpreted as the reciprocal temperature T, which operates as a tunable parameter, with units of energy, whose increase or decrease gives or takes away the energy needed by the zipper to break links. Because for us traveltime has replaced energy, $T = 1/\beta$ becomes a tunable parameter, with units of time, which gives the system access to the traveltime reserves needed for slowness values to increase. We will continue to call this the *temperature*, but the change in interpretation should be emphasized.

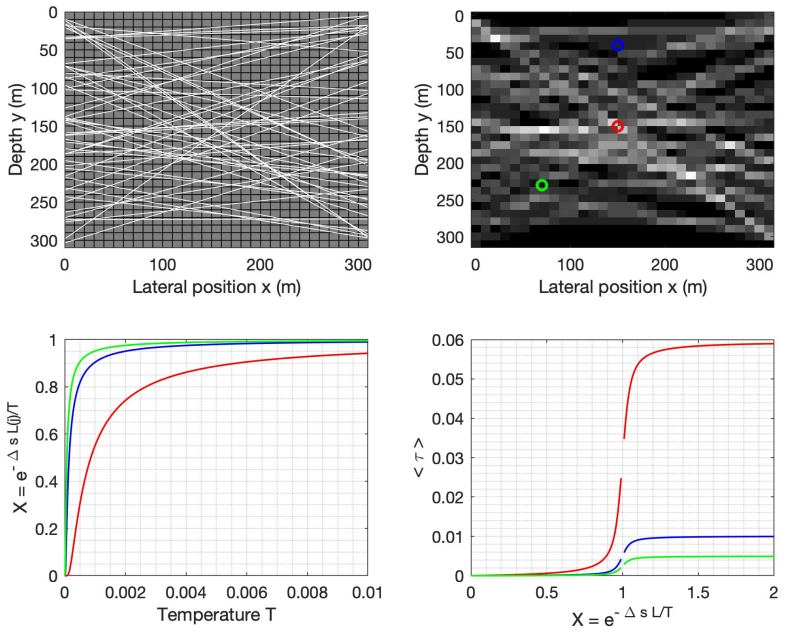

FIG. 3. Panel 1: ray coverage. Panel 2: $L(j)$ with the j cells arranged spatially. Panel 3: X versus T. Panel 4: $\langle\, \tau(j)\, \rangle$ versus X.

PREDICTIONS OF THE MODEL

Average values of the contribution of cell j to the total experimental traveltime $\mathcal{T}$ and the slowness index $s(j)$, emerge quickly from the partition function:

$$\langle\, \tau(j)\, \rangle = -\frac{\partial \log Z(\beta, j)}{\partial \beta} = \Delta s L(j) \left[\frac{X(j)}{1 - X(j)} - S \frac{X(j)^S}{1 - X(j)^S} \right], \tag{10}$$

which from (1) means

$$\langle\, s(j)\, \rangle = \left[\frac{X(j)}{1 - X(j)} - S \frac{X(j)^S}{1 - X(j)^S} \right]. \tag{11}$$

This is the average slowness contribution made by cell j to the total traveltime $\mathcal{T}$. Since this number grows as the relevance to the data of cell j grows — in terms of ray coverage, source/receiver geometry, cell dimensions, etc. — we will allow $\langle\, \tau(j)\, \rangle$ and $\langle\, s(j)\, \rangle$ to assume central roles in the analysis.

The fluctuation of $\tau(j)$ is also immediately computable from the model:

$$\Delta\tau(j)^2 = L(j)^2 \Delta s^2 \big[S K_S(j) - K(j) \big], \tag{12}$$

where

$$K(j) = \left(\frac{X}{1 - X} \right) \left(1 + \frac{X}{1 - X} \right), \tag{13}$$

and

$$K_S(j) = \left(\frac{X^S}{1-X^S}\right)\left(1+\frac{X^S}{1-X^S}\right). \tag{14}$$

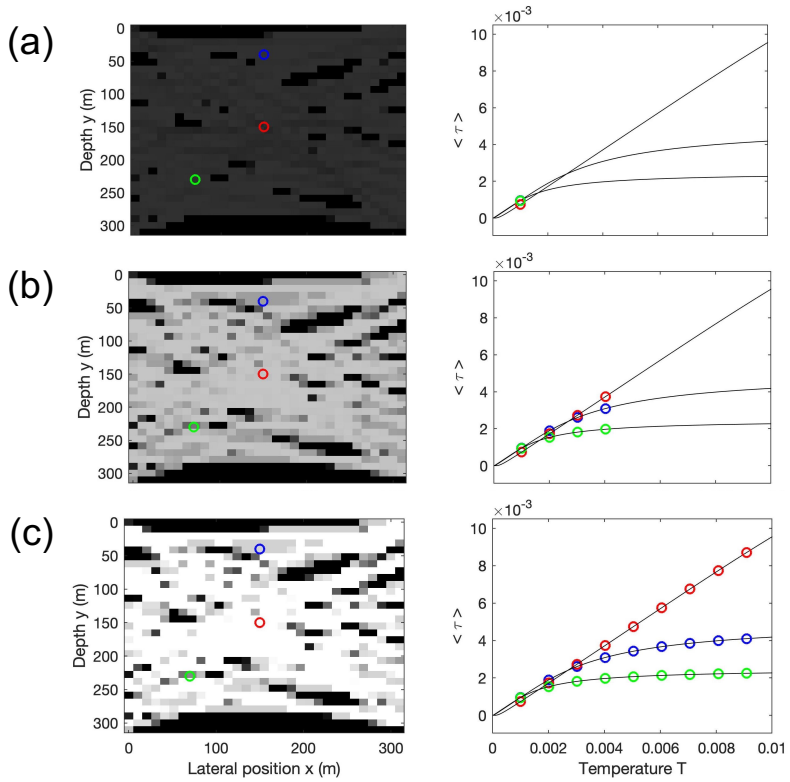

FIG. 4. (a)-(c) Change in relative average contribution of cells as temperature grows.

Numerical examples

In Figure 3 we set out a simple example tomography problem with a 17×17grid of slowness cells of dimension $\Delta x = \Delta y = 10$m bounded by two vertical boreholes, in which 100 source-receiver pairs with depths selected randomly produce straight rays.

The slowness values assigned to each cell are drawn from a discrete array of $S = 101$ values between 0 and $S \times \Delta s = 1/1000$. With the raypaths drawn in Figure 3 in hand the values $l(i,j)$ and then $L(j)$ can be computed, producing all of the input needed to create $X(j)$ in (11). The raypaths and the $L(j)$ for each cell are plotted in Figure 3, with $L(j)$ arranged spatially and plotted as a gray-scale image.

Each of the cells in Figure 3 produce their own statistical behaviour (i.e., $\langle\, s(j)\,\rangle$). We select three representative cells (red, blue, and green) for specific study. The temperature $T = 1/\beta$ is chosen to range between 0s and 0.01s, and we observe in the bottom left panel that over that range X grows monotonically but only approaches 1, doing so more slowly

for the red cell, which has a relatively large $L(j)$ value, than green or blue. As it does so, the contributions to the total traveltime of the cells also grows, with the red cell's growing at a significantly higher rate than those of the green and blue cells. The contribution at $X = 1$ itself is not defined.

In addition to acquisition details, discretization is a strong feature of this difference in growth rates — as the number of possible slowness values S grows, the jump at $X = 1$ becomes more and more rapid; coarse slowness discretizations have the opposite effect. For example, in Figure 5 we repeat the experiment with $S = 20$ but all else held fixed. Differences in rate of change of the relative contributions of the three example cells appear at much lower temperatures.

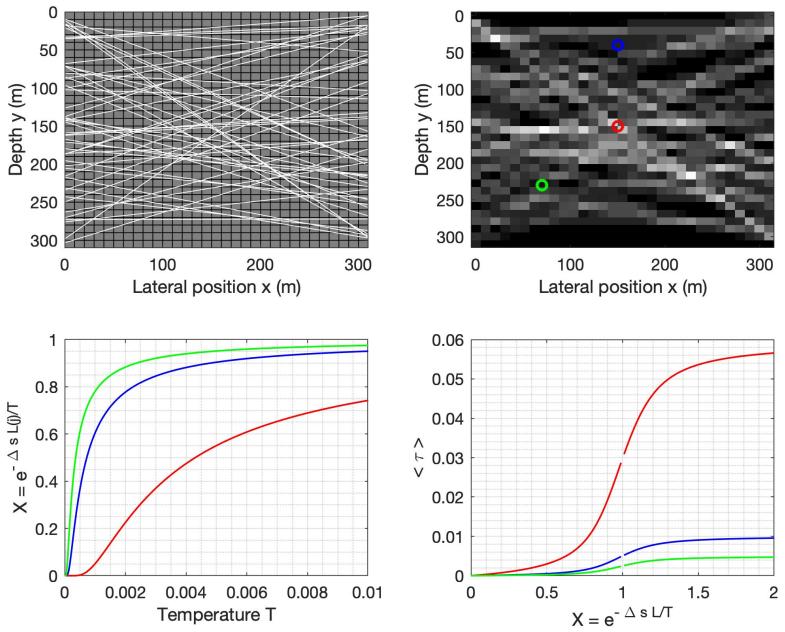

FIG. 5. Similar to Figure 3 but with coarse ($S = 20$) slowness discretization. Panel 1: ray coverage. Panel 2: $L(j)$ with the j cells arranged spatially. Panel 3: X versus T. Panel 4: $\langle\, \tau(j)\, \rangle$ versus X.

DISCUSSION AND CONCLUSIONS

Generalization

By the time we get to the calculation of the key features of the zipper model for uncertainty quantification, it is apparent that most of the details of the experiment we started the derivation with have been diminished, leaving only a few central aspects left. Specifically, although we set the problem up (and exemplified it) with a crosswell tomography problem, in the end all the acquisition details have done is defined how to compute L, i.e., the total length of raypaths transiting a cell. It follows that this approach is not tied to any particular mode of tomography— VSP, reverse VSP, cross-well, early iterations of FWI, etc., all of which are driven by crossing raypaths can all be subjected to this analysis without any particular change being necessary. All that is required is an ability to estimate ray lengths in model cells.

It also follows immediately that finite elements can be included in this analysis without changes in the formulation.

Finally, although we have designed the method to appraise traveltime tomography, the approach is at a level of abstraction where the differences between it and attenuation tomography (e.g., X-ray tomography) are inessential.

Temperature and "fast" versus "slow" geologies

The temperature $T = 1/\beta$ is an artifice of the model, but one which plays an important role and which represents a simple way to fit bulk features of the data into the appraisal scheme. Explicitly, the data do not enter into the averages in the model – the scheme is set up to act absent data (or rather, prior to data being acquired), and be driven by physics, experimental/acquisition parameters, and discretization.

In standard statistical mechanics, the temperature emerging from the Boltzmann theory is treated as a kind of parameter, or dial, which can be smoothly varied to observe the behaviour of the system in warmer versus cooler environments, i.e., in environments where elements of the system have access to more versus less energy. In the case of the molecular zipper as applied to the DNA model, a higher temperature gives the strand of DNA access to a greater capacity to un-ravel. In analogy, the temperature here should be considered a parameter or dial that can be smoothly varied to change the access slowness cells have to "available traveltime" to distribute amongst its realizations in the ensemble. A larger T in our scheme sets as the goal the explanation of larger bulk traveltimes. This means increasing T for cell j produces $\langle\, \tau(j) \,\rangle$ numbers appropriate for relatively slow geologies; decreasing T produces $\langle\, \tau(j) \,\rangle$ interpretable for faster geologies.

ACKNOWLEDGEMENTS

The sponsors of CREWES are gratefully thanked for continued support. This work was funded by CREWES industrial sponsors, NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19, and in part by an NSERC-DG.

REFERENCES

Innanen, K. A., 2021, A selective review of methods of statistical mechanics: CREWES Research Reports, **33**.

Kittel, C., 1969, Phase transition of a molecular zipper: American Journal of Physics, **37**, 917.