Deep nets for making sense of ambient noise?

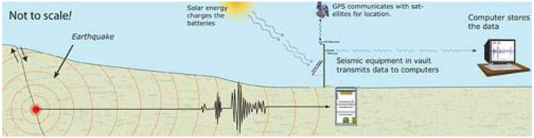
Julien Clancy

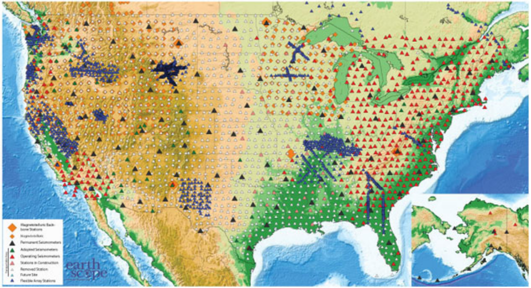
Massachusetts Institute of Technology
Dept of Mathematics

Joint work with:
Laurent Demanet (MIT), Jonathan Helland (CSM), Zongbo Xu (Boise)

May 2018

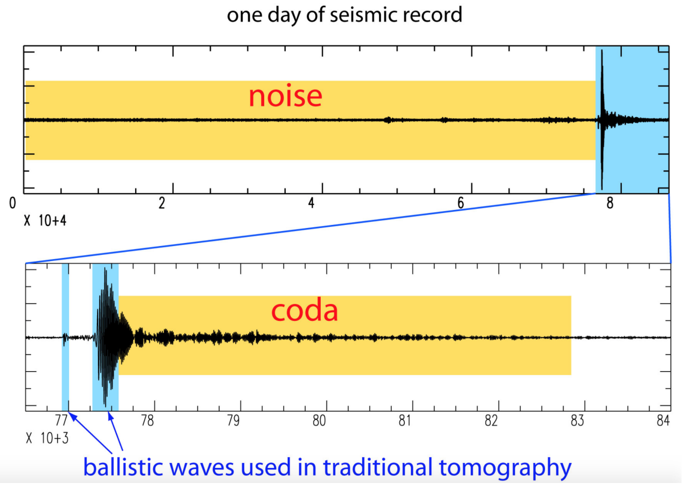
Example: ambient seismic noise





Credit: EarthScope, usarray.org

Noise, ballistic waves, and coda



Credit: Nikolai Shapiro, Michel Campillo

What information is present in ambient seismic noise?

Wave equation with random forcing

$$\left(\frac{1}{c^2(x)}(a+\partial_t)^2-\Delta\right)u(x,t)=n(x,t)$$

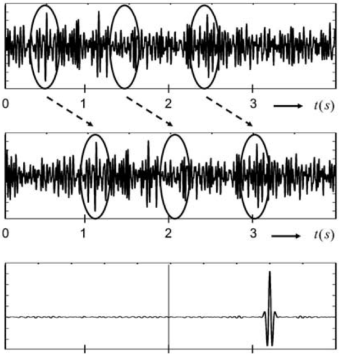
Space-time white noise: Gaussian with
$\langle n(x,t)n(y,s)\rangle=\delta(t-s)\delta(x-y)$.

- Pointwise: $u(x,t)$ has Gaussian statistics in t
- Pairwise: $u(x,t)$ and $u(y,s)$ may be strongly correlated
  $\Rightarrow$ may contain info about $c(x)$

Cross-correlations

Cross-correlations are robust to source randomness.
Their peaks indicate traveltimes.

- Time reversal (phys): Fink et al. 1993.
- Time reversal (math) Bal, Papanicolaou, Ryzhik, 2002; Bal, Ryzhik 2003
- CINT imaging: Borcea, Papanicolaou, Tsogka, 2003, 2005
- Seismic interf. (phys): Weaver et al. 2001; Campillo, Paul 2003; Snieder, 2004; Wapenaar et al. 2006
- Seismic interf. (math): Bardos et al. 2008; Colin de Verdiere 2009; Garnier, Papanicolaou, Solna 2009+



Wapenaar et al. 2010

Sample mathematical result (Bardos et al. 2008)

Assume a space-time white noise source. Let

$$C_T(\tau, x, y) = \frac{1}{T} \int_0^T u(x,t) u(y, t+\tau)\, dt.$$

Then

$$\partial_\tau \langle C_T(\tau, x, y) \rangle = -\frac{e^{-a|\tau|}}{4a} \mathrm{sgn}(\tau) G(|\tau|, x, y)$$

where

$$\left(\frac{1}{c^2(x)} \partial_t^2 - \Delta \right) G(t, x, y) = \delta(t)\delta(x-y).$$

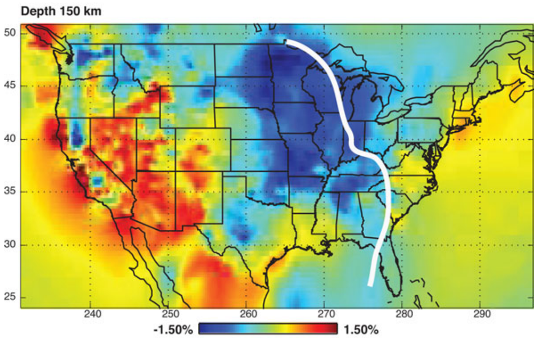
(Extensions: ergodicity, or large T limit, in the high freq regime.)

Seismic interferometry

In turn,

- $G(t, x, y)$ is singular near $t = \text{traveltime}(x, y)$
- $\text{traveltime}(x, y)$ is related to $c(x)$

Hence: determine $c(x)$ by traveltime tomography



Credit: EarthScope, usarray.org

Open questions

Are c.c always the right idea?

What if the source distribution is not isotropic?

Can we avoid traveltime tomography to get $c(x)$?

Is there physics that the lag of a cross-correlation peak doesn't see?

What to do with combinations of 3 or more sensors?

Open questions

Are c.c always the right idea?

What if the source distribution is not isotropic?

Can we avoid traveltime tomography to get $c(x)$?

Is there physics that the lag of a cross-correlation peak doesn't see?

What to do with combinations of 3 or more sensors?

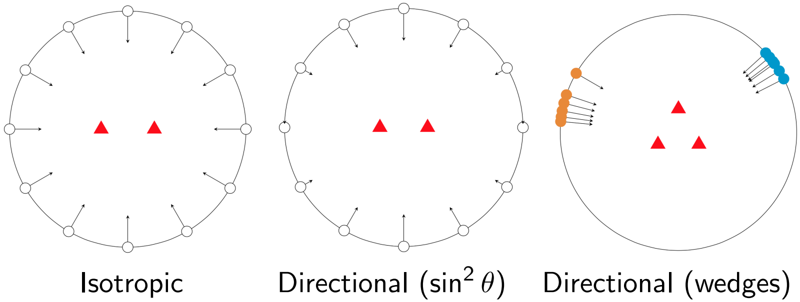
Deep nets provide "cheap" answers to most such questions
(no guarantees, no physics)

Experimental setups

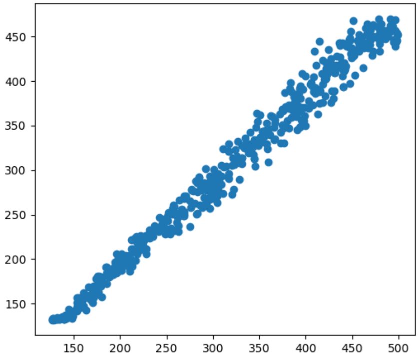
1. Traveltime inversion with anisotropic sources, homogeneous media
2. Identification of source distributions
3. Inhomogeneous wavespeed inversion

Training: For homogeneous media, generate new data on the fly, seen once and never again. For inhomogeneous media, direct solve with conventional SGD + data augmentation

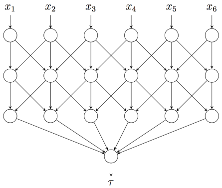
Traveltime inversion



Traveltime inversion: 2 sensors suffice

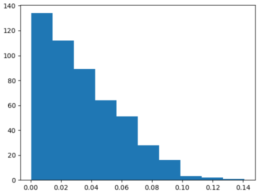


Estimated vs true traveltime (test samples)



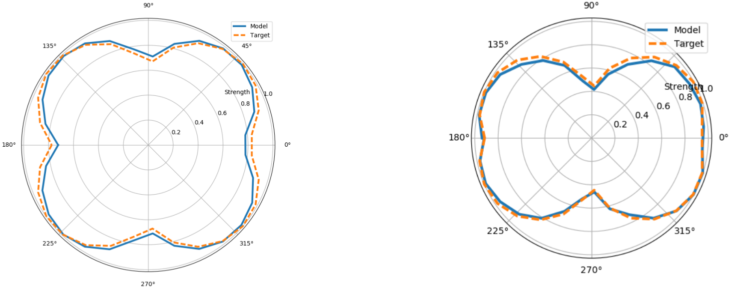
Locally connected

(also works with CLSTM)



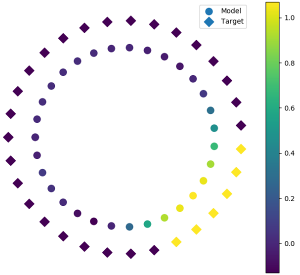
Distribution of relative errors

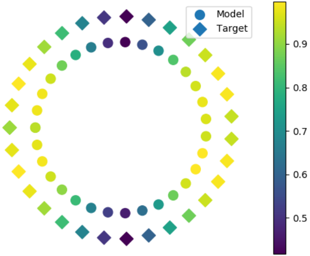
Inversion of source directionality



Estimated vs true source strength of the form $a \sin\theta + b \cos\theta$ (test samples). Need CLSTM architecture.

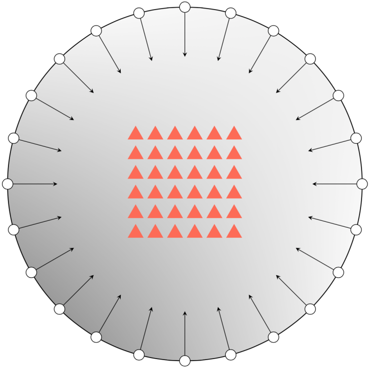
More Source Directionality



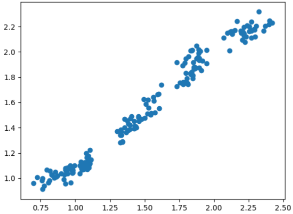


Left: network trained in a wedge distribution (with three sensors).
Right: same as previous slide.

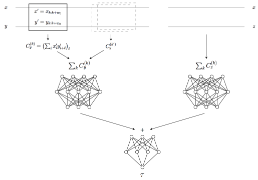
Inversion of heterogeneous wavespeeds



$$c(x) = ax_1 + bx_2 + c$$

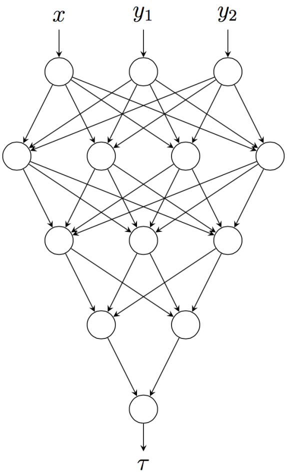


Estimated vs true traveltime

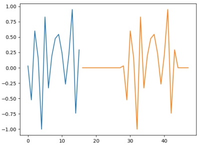


Relational network

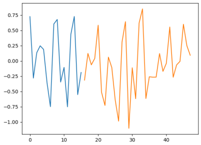
What are the nodes actually computing? (1)



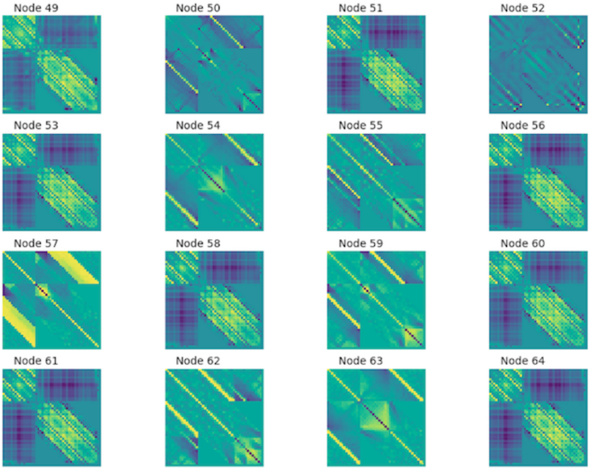
Noiseless:



Noisy:

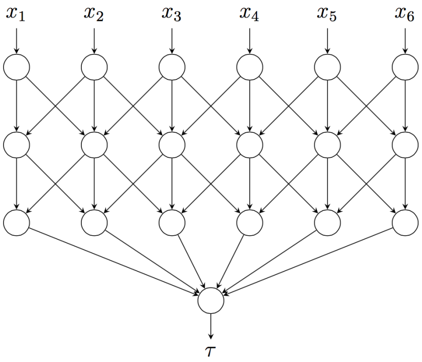


What are the nodes actually computing? (2)

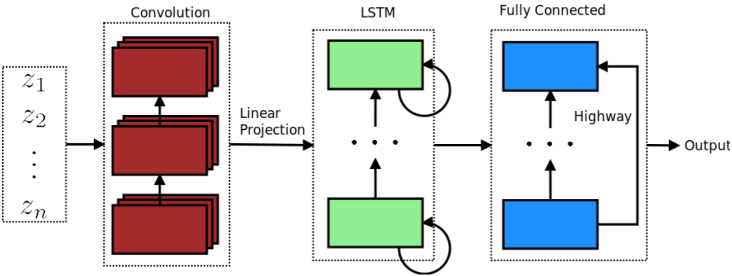


Regress quadratic monomial on hidden nodes.

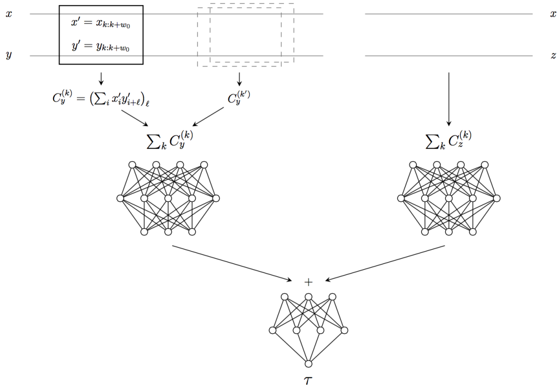
Architectures: Locally connected



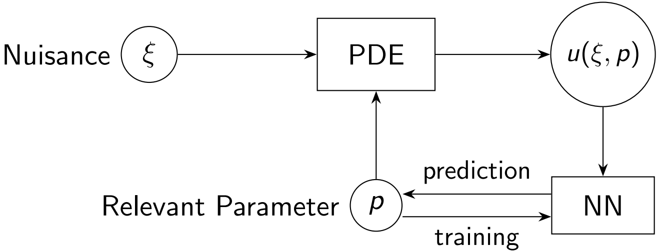
Architectures: CLSTM



Architectures: Relational



Automatic De-parameterization



- ξ: random source
- p: local wave speed, source distribution, etc
- $u(\xi, p)$: passive seismic data

Goal: create representations of the data $u(\xi, p)$ that are mostly *invariant* to ξ and have enough information to recover p. In limited situations, this is what cross-correlations does.

Conclusions

- *Blessing*:
  Data-centric predictions without a forward or inverse physical model (automatic de-parametrization of forward simulations that involve nuisance parameters that cannot be estimated during the inversion)
- *Curse*:
  Physical content is (at best) buried in the network
- *Curse*:
  Lack of performance guarantees
  (possible overfitting, testing out of sample)

Outlook: characterization of scatterers, better performance in the inhomogeneous case

References I

[1] Abduljabbar Asadi et al. "Spatially continuous probabilistic prediction of sparsely measured ground properties constrained by ill-posed tomographic imaging considering data uncertainty and resolution". 2017.

[2] Richard Baraniuk and Ali Mousavi. "Learning To Invert: Signal Recovery Via Deep Convolutional Networks". 2017.

[3] Ashish Bora et al. "Compressed Sensing using Generative Models". June 2017.

[4] Ian Goodfellow et al. "Generative Adversarial Nets". 2014.

[5] Sepp Hochreiter and Jürgen Schmidhuber. "Long Short-Term Memory". Nov. 1997.

[6] Günter Klambauer et al. "Self-Normalizing Neural Networks". 2017.

References II

[7] Alex Krizhevsky et al. "ImageNet Classification with Deep Convolutional Neural Networks". May 2017.

[8] Yann Lecun et al. "Gradient-based learning applied to document recognition". 1998.

[9] Tara N Sainath et al. "Convolutional, long short-term memory, fully connected deep neural networks". 2015.

[10] Adam Santoro et al. "A simple neural network module for relational reasoning". 2017.

[11] David Silver et al. "Mastering the game of Go with deep neural networks and tree search". 2016.

[12] Rupesh Kumar Srivastava et al. "Training Very Deep Networks". 2015.