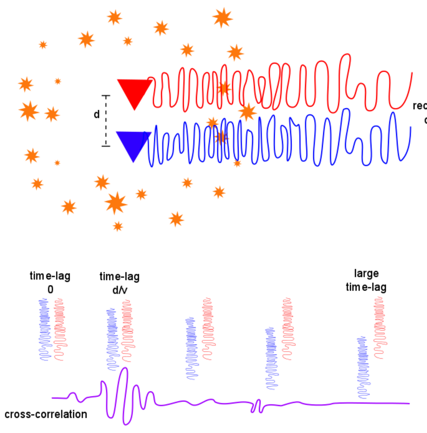
Fast Ambient Noise Imaging of Compressed Data

Eileen R. Martin (Virginia Tech)

Blacksburg, VA: Dept. of Mathematics, Program in Computational Modeling and Data Analytics

Problem: Traditional Ambient Seismic Noise Interferometry Requires too Much Data Movement for Ultra-dense Arrays

What is Ambient Noise Interferometry?



Notation

Number of sensors	N_s
Number of time windows	N_w
Number of time samples per data window	N_t
Number of time-lags per cross-correlation	N_τ
Data at receiver r, time window w, time t	$d_w(r,t)$
Data at receiver r, time window w, frequency ω	$\hat{d}_w(r,\omega)$

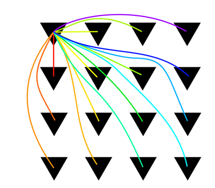
Standard Algorithm

average cross-correlations over time windows

for each time window:

for each virtual source (receiver):

perform cross-correlations w/ all receivers:



time domain cost: $N_s 2(N_t - N_\tau)N_\tau \quad \propto \quad O(N_s N_t N_\tau)$

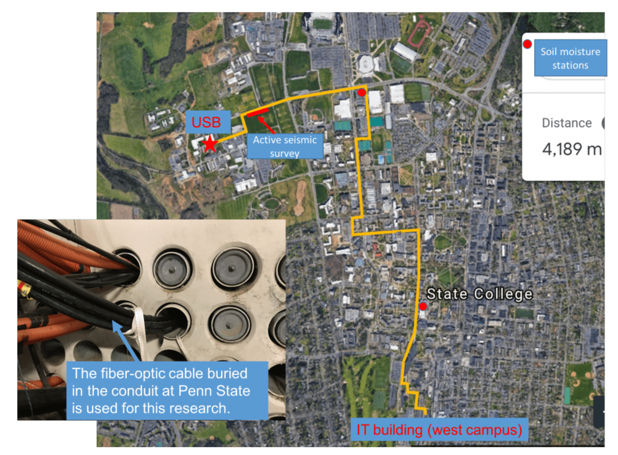
freq. domain cost: $O(N_s N_t) + O(N_t \log(N_t))$

$\times N_s$

$\times N_w$

New Dense Arrays

FORESEE Fiber Array at Penn State



Node Arrays

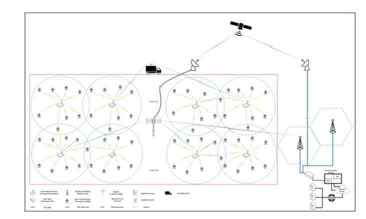


fig. from Jamali-Rad and Campman (2018)

Cell Phone Accelerometers

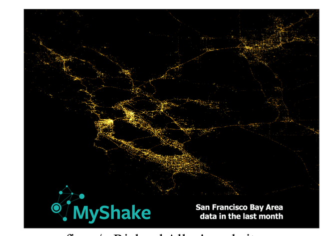


fig. c/o Richard Allen's website

Scales of Interest and Assumptions

100s - 10,000s of sensors

10,000 - 1,000,000 time samples per window

100s - 1,000s of time lags per correlation

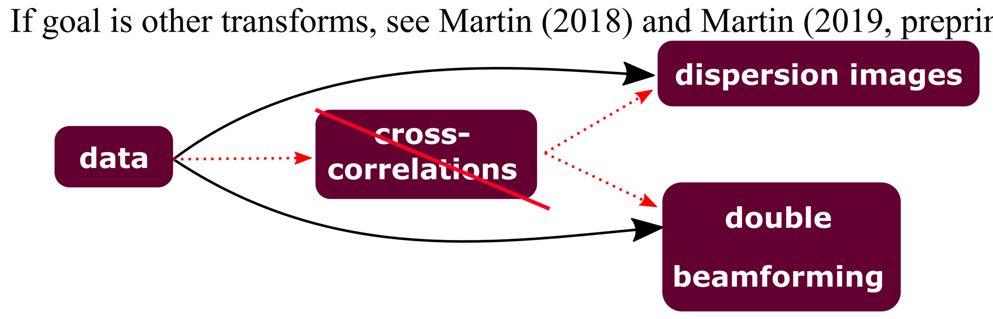
Correlations are much shorter than data windows

$$N_\tau << N_t$$

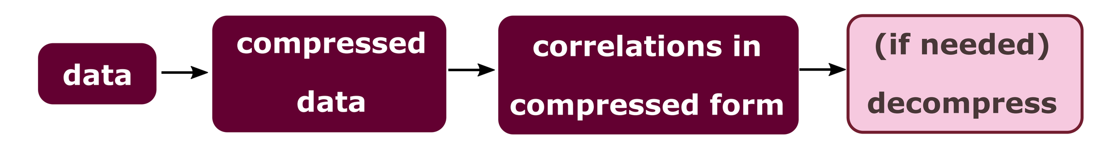
Only one window of data is available at a time (streaming paradigm)

Roadmap of Recent Advances

If goal is other transforms, see Martin (2018) and Martin (2019, preprint).



If cross-correlations are required (**this work**):

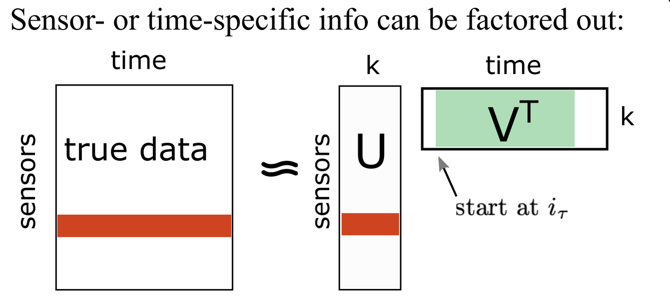


One Solution: Interferometry of Low-rank Compressed Data

Reference: Martin (2019, SEG Ann. Mtg.)

Intuition: Assuming data in one time window, D, compressed as UVT we want to calculate cross-correlation tensor C

Sensor- or time-specific info can be factored out:

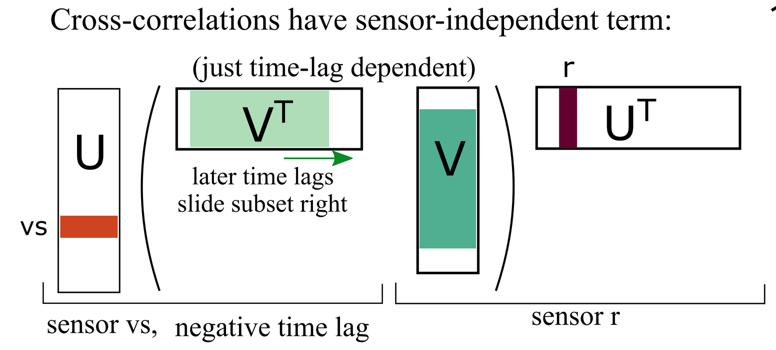


$$D(vs, 1:N_t) \approx U(vs, 1:k)V^T(1:k, 1:N_t)$$

$$\tilde{D}^{vs,i_\tau} = D\left(vs, i_\tau + \frac{N_\tau}{2} : i_\tau + \frac{N_\tau}{2}\right)$$

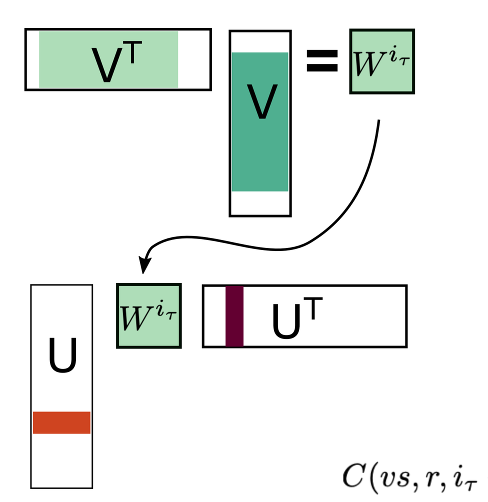
$$\tilde{U}^{vs}\tilde{V}^{T(i_\tau)} = U(vs,1:k)V^T\left(1:k, i_\tau + \frac{N_\tau}{2} : i_\tau + N_t - \frac{N_\tau}{2}\right)$$

Cross-correlations have sensor-independent term:



$$C\left(vs, r, i_\tau + N_\tau/2\right) = \tilde{D}^{(vs,i_\tau)}\tilde{D}^{T(r,0)}$$

$$\approx \tilde{U}^{vs}\tilde{V}^{T(i_\tau)}\tilde{V}^{(0)}\tilde{U}^{T(r)}$$



Two approaches:

Phase 1:
For a given time lag index calculate W^{iT}

$$W^{i_\tau} = \tilde{V}^{T(i_\tau)}\tilde{V}^{(0)}$$

Phase 2:
Multiply the sensor-dependent matrices by W^{iT}

$$C(vs, r, i_\tau + N_\tau/2) \approx \tilde{U}^{vs}W^{i_\tau}\tilde{U}^{T(r)}$$

for each time lag

for each time lag

for each time lag

Cost Analysis per Window:

per time lag

To calculate each W^{i_τ}:
$O(k^2(N_t - N_\tau)) \approx O(k^2 N_t)$

To calculate each $UW^{i_\tau}U^T$:
$O(N_s k^2) + O(N_s^2 k)$

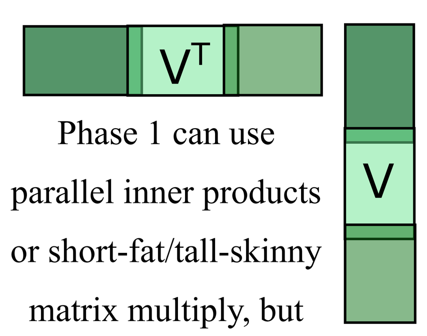
Total: $O(N_\tau k^2 N_t) + O(N_\tau N_s^2 k)$

time dependent + sensor dependent

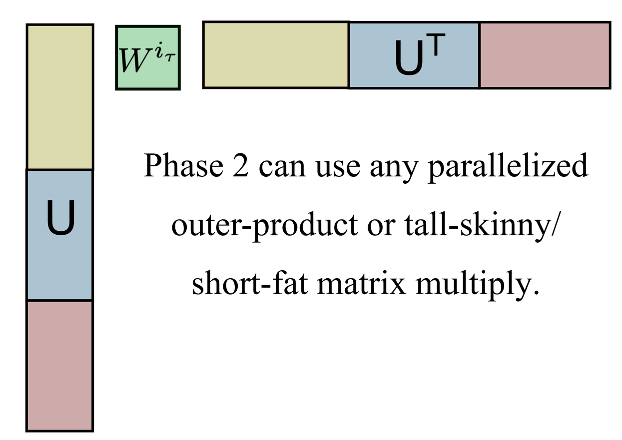
Benefits:

- Fewer operations (memops and flops)
- Can use any UV^T (randomized SVD or QR for example)
- Can be adapted to any cheap-to-multiply factorization
- Can utilize existing parallel or GPU matrix-multiply
- Extension to tensor compression

Sketch for Parallelism:

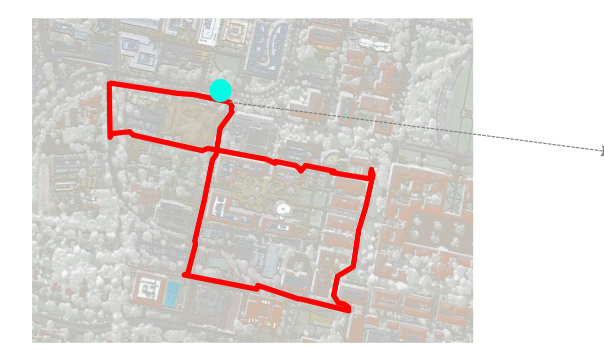


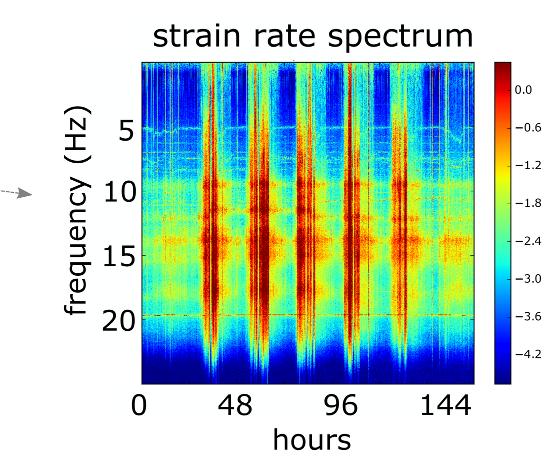
Phase 1 can use parallel inner products or fast-fat/tall-skinny matrix multiply, but needs some overlap.

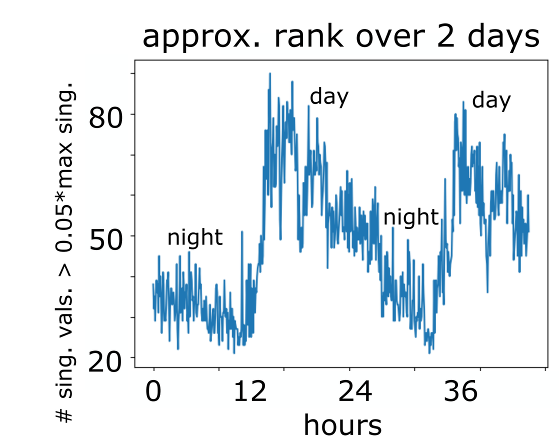


Phase 2 can use any parallelized outer-product or tall-skinny/ short-fat matrix multiply.

Application to Real Data: Stanford Fiber Optic Seismic Observatory







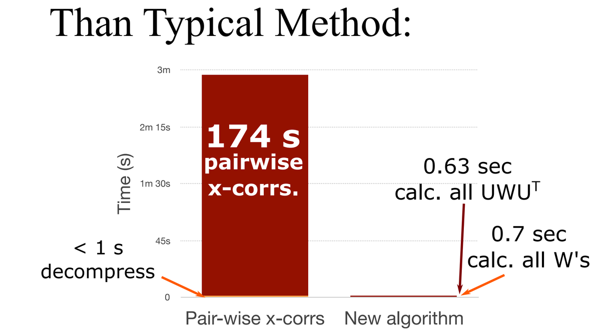
The Stanford Fiber Optic seismic observatory has continuously acquired seismic noise throughout campus since Sep. 2016 led by Biondo Biondi. 2.5 km of fiber optic cable were run through existing telecommunications conduits under campus.

Although noises are caused by random sources, many repeat, and some noise sources are more dominant than others (left top). Given the singular value decomposition of every 620x15000 sample data window matrix, only roughly 10% of singular values are > 1/20th the largest singular value (top right).

Relevant Scales:

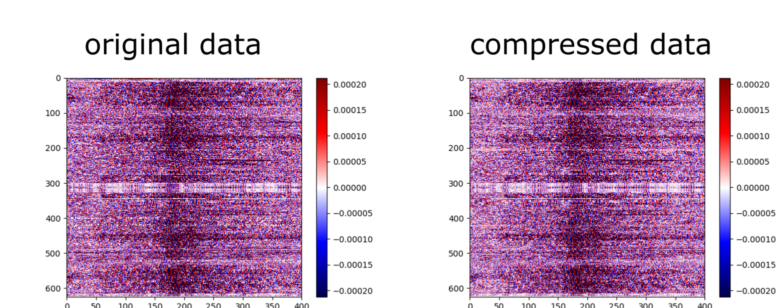
$N_s = 620$ sensors
$N_t = 15{,}000$ samples (50 samp./sec. x 300 sec.)
$N_\tau = 100$ samples
$N_w = 288$ windows/day (recording for 3+ years now)
$k = 38$ (preserves sing. vals. > 5 % largest)

Low-rank Algorithm is Faster Than Typical Method:



Moveouts Same:

Data during a small earthquake show close moveouts for the original normalized and compressed data



Another Solution: Interferometry of Wavelet Compressed Data

Motivation:
One downside of low-rank compression is that it tends to preserve lower frequencies, while we are motivated to use higher density sensor arrays in order to capture both low and high frequencies. An alternative compression scheme that tends to better preserve multiscale trends in seismic data is the wavelet transform. Here, we consider a 1D transform in time for simple notation, but we have extended to 2D for additional compressibility.

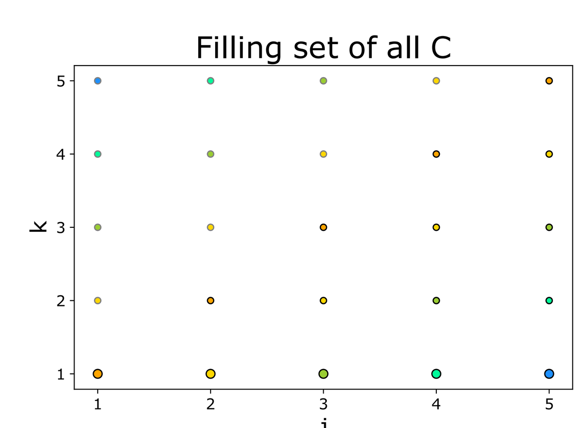
Idea: Express a time series f in terms of a multiresolution analysis:

$$f_N(t) = P_{V_L}f(t) = \sum_{j=L+1}^{J}\sum_{n=0}^{2^{-j}-1}\langle f, \psi_{j,n}\rangle\psi_{j,n}(t) + \sum_{n=0}^{2^{-J}-1}\langle f, \phi_{J,n}\rangle\phi_{J,n}(t)$$

Cross-correlation between pairs of sensors can be expressed as a sum of their wavelet coefficient pairs multiplied with the cross-correlations of their wavelet basic functions (ψ or ϕ, which are represented generally by θ below). Threshold small coefficients so this problem becomes sparse, and wavelet coefficient products can be accumulated over time windows, avoiding reconstruction in time domain at each step.

Properties of Wavelet Cross-correlations:

1. Time-shift transform: $(\theta^{(g)}_{j,n'} \star \theta^{(h)}_{k,m'})(\tau) = (\theta^{(g)}_{j,n}(t) \star \theta^{(h)}_{k,m})(\tau + 2^k(m-m') - 2^j(n-n'))dt$
2. Rescaling transform: $\left(\theta^{(g)}_{j+a,n} \star \theta^{(h)}_{k+a,m}\right)(\tau) = \left(\theta^{(g)}_{j,n} \star \theta^{(h)}_{k,m}\right)(2^{-a}\tau)$
3. Time reversal (true for all cross-correlations): $\left(\theta^{(g)}_{j,n} \star \theta^{(h)}_{k,m}\right)(\tau) = \left(\theta^{(h)}_{k,m} \star \theta^{(g)}_{j,n}\right)(-\tau)$
4. (1+2+3) generate all cross-correlations needed from 1D subset



Using the properties of cross-correlations of wavelet transforms, one can decide how many wavelet basis function cross-correlations to pre-compute, and how many to compute on the fly via simple transforms. The diagram above shows how one might start with a 1D set of (j,k,n,m) = (*,1,0,0), then fill in all (j+a,k+a,0,0) values based on the rescaling transform, then fill in all (k,j,0,0) values based on time reversal, and finally extend for all n and m values based on the time-shift transform.

Acknowledgements

Extensions of this research have been supported in part by contract 4000175567 from UT-Battelle through Oak Ridge National Laboratory. Thank you to Will Ray at Oak Ridge National Laboratory, Joseph Kump and Sarah Morgan at Virginia Tech for helpful conversations related to extensions of this research. Startup funds from the Virginia Tech College of Science partially supported this research. Thank you to Biondo Biondi and the Stanford Exploration Project for sharing data from the SFSO. The FORESEE project is funded by a seed grant, team led by Tieyuan Zhu, from the Penn State Instutes of Energy and the Environment.

References

Jamali-Rad and Campman, "Internet of Things-based wireless networking for seismic applications," Geophysical Prospecting, 2018.

Martin, "Passive Imaging and Characterization with Distributed Acoustic Sensing," Stanford PhD Dissertation, 2018.

Martin, "Scalable Algorithm for Cross-correlations of Compressed Ambient Seismic Noise," 2019 Expanded Abstracts of the SEG 89th Annual International Meeting, doi: 10.1190/segam2019-3216637.1

Martin, "A Linear Algorithm for Ambient Seismic Noise Double Beamforming Without Explicit Cross-correlations," submitted 2019 to Geophysics, preprint at https://vtechworks.lib.vt.edu/handle/10919/96246.