

# Ground motion prediction using ambient seismic noise on a large-N array in the LA basin

Julian Schmitt<sup>1</sup>, Nan Wang<sup>2</sup>, Tim Clements<sup>1</sup>, Kurama Okubo<sup>1</sup>, Kim Olsen<sup>2</sup>, Marine A. Denolle<sup>1</sup>

<sup>1</sup>Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA., <sup>2</sup>Department of Geological Sciences, San Diego State University, San Diego, CA, USA

## Abstract

Accurate ground motion is critical for seismic hazard mitigation in southern California. The presence of multiple sedimentary basins may, or may not, be responsible for basin-scale seismic amplification during the next San Andreas Earthquake (Olsen et al, 2006, 2014; Denolle et al, 2014).

This project aims to validate the presence of a seismic waveguide and their connectivity in the northern corridor San Bernardino-Pomona-San Gabriel. We compare two methodologies for ground motion prediction: wavefield numerical simulations (from imperfect velocity models) and ambient noise virtual earthquakes (from imperfect approximation to the Green's functions of the noise correlation functions). We perform a systematic comparison of the Green's functions in the northern corridor using broadbands and BASIN node data.

## Methodology

### Ambient noise correlation functions



We develop an awesome high-performance code to process TBs of ambient noise data to produce a library of 100,000s of noise cross correlations based on minimal Julia-native seismology tools (SeisIO.jl J20; SeisNoise.jl C20).

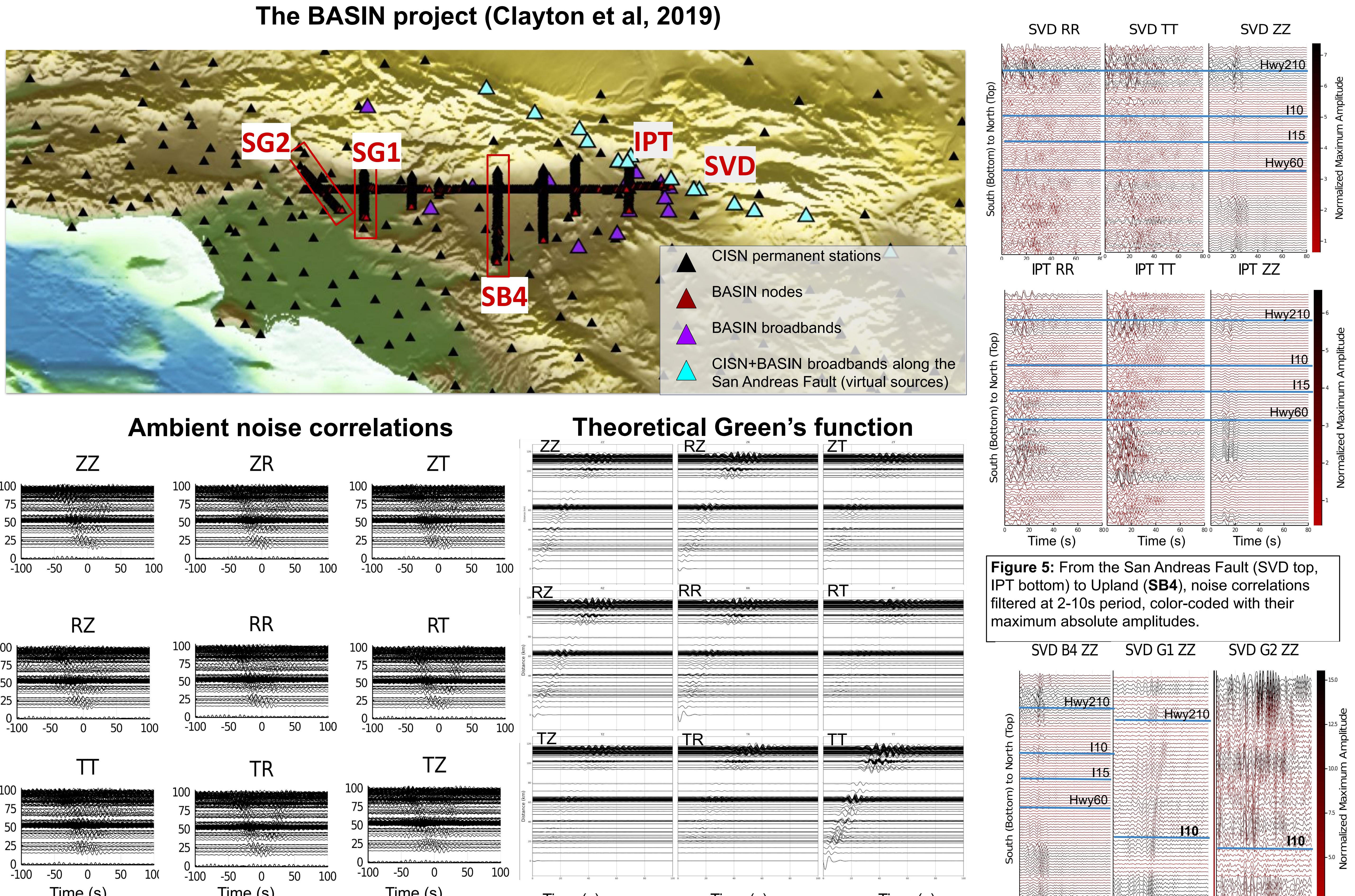
We cut the continuous recordings into 1 non-overlapping hour and perform a weak spectral whitening by computing the coherency (Prieto et al, 2009)

This project examines the subset of this large correlation library, selecting 87 SB4, 59 G1, and 45 G2 nodal stations from January and February 2017. We correlated station pairs using a bandpass from 0.05 to 9.9 Hz, a coherency filter with a water level of 0.01 and half window of 30, removing any correlation window with more than 10 times the median window before stacking daily and storing in the cloud. In our post-processing after discarding a few gapped nodes, we apply a simple linear stack to 34 days of data, shorten correlations to a maximum time lag of 300 seconds, and then rotate correlations to the transverse and radial directions.

The plots we display here for the BASIN project exemplify the enormous computational and cost advantage the framework we have developed will provide. We preconfigure a 32-core AWS EC2 c5 instance, pull 19 GB/day of data from two S3 buckets (scedc-seisbasin), preprocess and correlate (as described above), and push back over 9,000 correlations to S3 in under 120 seconds of compute time. The instance is priced at \$1.53 per hour meaning **we can process over 0.5TB of data into a 0.25M correlations for less than the cost of a cup of coffee**. We consider this to be the future of ambient noise seismology.

### Numerical wavefield simulations: AWP-ODC solver

Domain	
Velocity model	UCVM-S4.26
Length	167400 m
Width	79200 m
Depth	21600 m
Southwest corner	-118.2, 33.8
Spatial resolution	
Maximum frequency	1 Hz
Minimum Vs	500 m/s
Points per minimum wavelength	6
Grid discretization	75 m
Temporal resolution	
Time discretization	0.002 s
Simulation time	150.0 s
Number of timesteps	75000

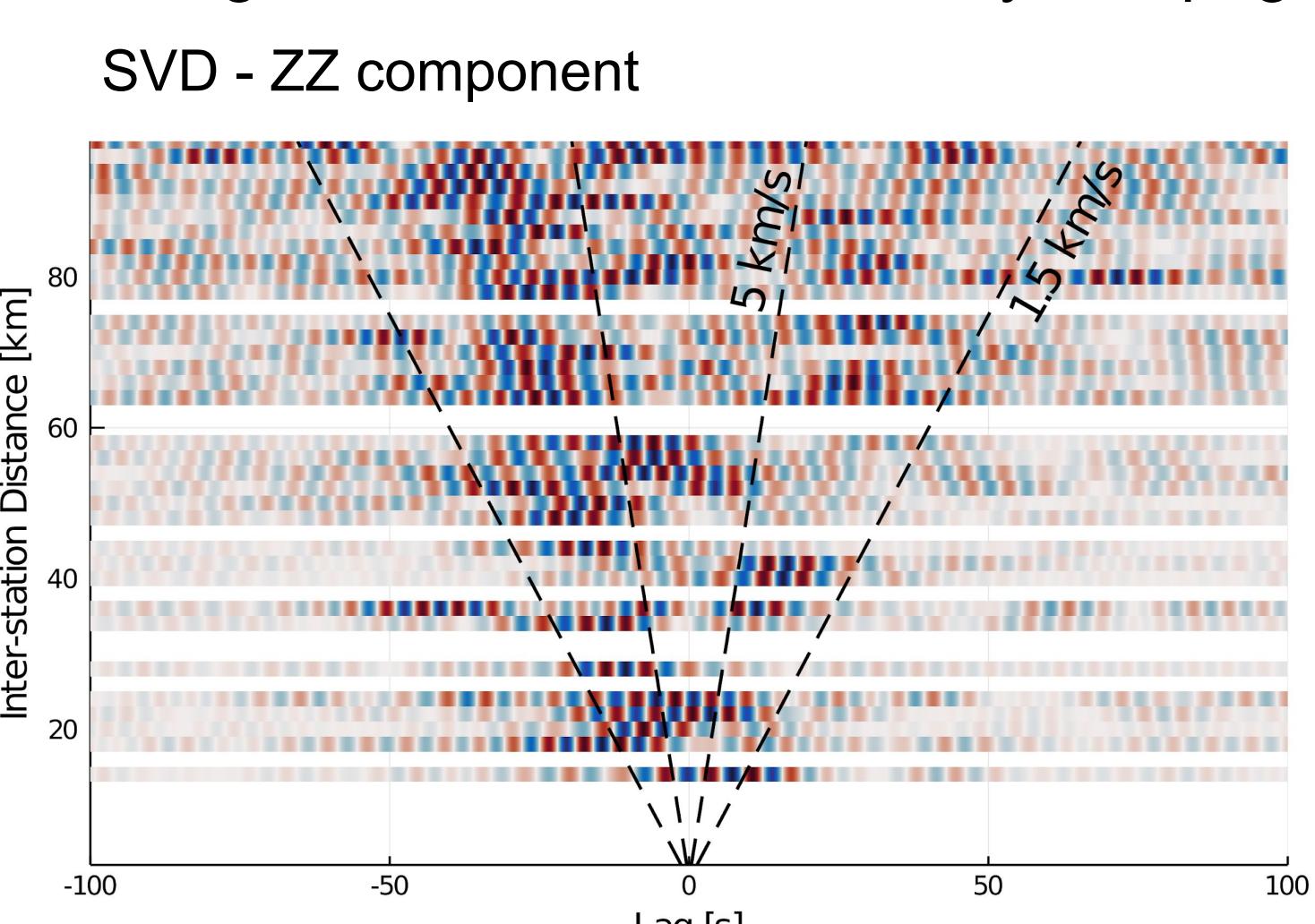


**Figure 1:** (top panel) raw noise signals simultaneously recorded at 2 stations. (bottom panel) the cross correlation of the two yield a function that approximate the surface-to-surface impulse response.

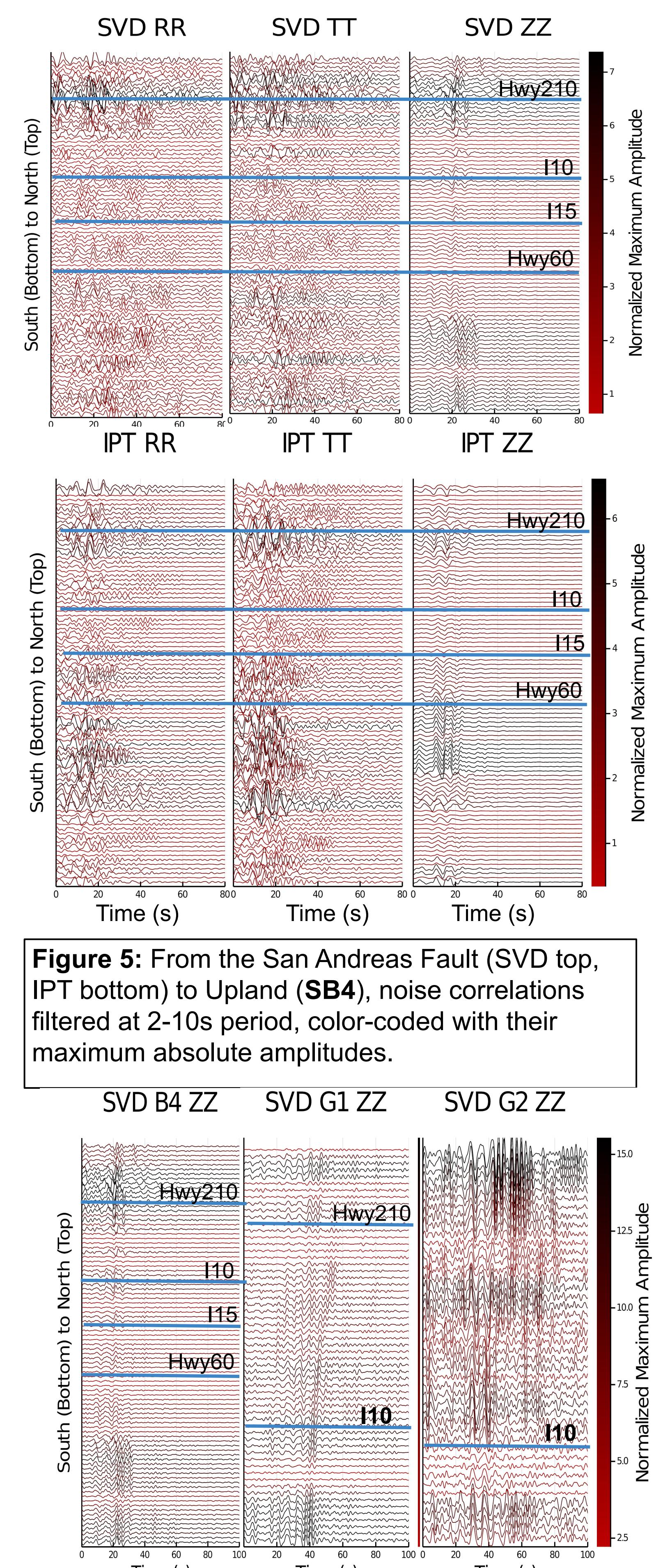
**Figure 2:** SVD 9-component Green tensor as a function of distance (in km) filtered 0.1-0.2Hz (5-10s). Each waveform here is normalized to its peak amplitude.

**Figure 3:** SVD 9-component Green tensor as a function of distance (in km) filtered 0.1-0.2Hz (5-10s). Each waveform is normalized to the ZZ peak amplitude.

Here we explore the 3D seismic wavefield by comparing the noise correlation tensor and the theoretical Green's functions. In a laterally homogeneous "flat" Earth, Love waves should be only on the TT component, and Rayleigh waves should be on the ZZ-ZR-RZ-RR components. Here, because of wavefront distortion due to laterally heterogeneous structure and maybe topography, Love waves appear on the radial components.



**Figure 4:** Heatmap with all 3 nodal deployments (SB4-SG1-SG2) filtered 0.1-0.2 Hz (5-10s). Dashed lines represent wave velocities of 5 and 15 km/s. The most energetic wave packets mostly represent contributions by surface waves.



**Figure 5:** From the San Andreas Fault (SVD top, IPT bottom) to Upland (SB4), noise correlations filtered at 2-10s period, color-coded with their maximum absolute amplitudes.

## Our next steps

Cross correlate the rest of the BASIN nodes with the broadband data on the San Andreas Fault.

Validate simulations and virtual earthquakes using local moderate recorded earthquakes.

Perform large M6-7 simulations along the San Andreas Fault to compare both methods

## Acknowledgements

We would like to acknowledge all of the BASIN deployment team, and its leaders: Robert Clayton, Patricia Persaud, Jascha Polet, Fan-Chi Lin, Nori Nakata, and many other awesome volunteers.