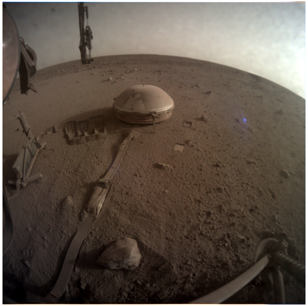
Unearthing InSights into Mars: Unsupervised source separation with limited data

Ali Siahkoohi [1] Rudy Morel [2] Maarten V. de Hoop [1] Erwan Allys [2] Grégory Sainton [3] Taichi Kawamura [3]

[1]Rice University

[2]École Normale Supérieure de Paris

[3]Institut de Physique du Globe de Paris



NASA/JPL-Caltech. *One of InSight's Last Images*. Accessed on June 29, 2023. 2022. URL: https://mars.nasa.gov/resources/27209.

Objective. Retrieve unknown source $\mathbf{s}_1(t)$ such that

$$\mathbf{x}(t) = \sum_{i=1}^{N} \mathbf{s}_i(t) + \boldsymbol{\nu}(t) = \mathbf{s}_1(t) + \sum_{i=2}^{N} \mathbf{s}_i(t) + \boldsymbol{\nu}(t)$$

$$= \mathbf{s}_1(t) + \mathbf{n}(t)$$

underlying unknown sources $\{\mathbf{s}_i(t)\}_{i=1}^{N}$

measurement noise $\boldsymbol{\nu}(t)$

only access to limited realizations of $\mathbf{n}(t)$

thermal relaxations

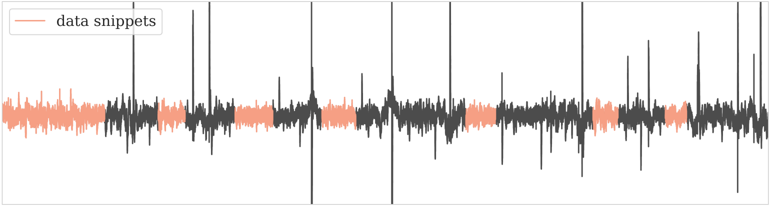
atmospheric disturbances

instrument and mechanical noise

Proposed method

unsupervised source separation in a low-dimensional representation space

based on the second-order moments of scattering coefficients



Wavelet scattering covariance

$$\Phi(\mathbf{x}) = \mathbb{E}\left[S(\mathbf{x})S(\mathbf{x})^\top\right] \quad \text{with} \quad S(\mathbf{x}) = \begin{bmatrix} \mathbf{W}\mathbf{x} \\ \mathbf{W}|\mathbf{W}\mathbf{x}| \end{bmatrix}$$

wavelet transform $\mathbf{W}$

low-dimensional representation

rich inductive bias

captures non-Gaussian properties

sparsity, intermittency, and time-asymmetry

data snippet size selected a priori

wavelet selected a priori

Rudy Morel, Gaspar Rochette, Roberto Leonarduzzi, Jean-Philippe Bouchaud, and Stéphane Mallat. "Scale dependencies and self-similarity through wavelet scattering covariance". In: *arXiv preprint arXiv:2204.10177* (2022).

Proposed method. Find $\widehat{\mathbf{s}}_1$ such that $\Phi(\mathbf{x} - \widehat{\mathbf{s}}_1)$ matches the statistics of $\Phi(\mathbf{n})$

$$\widehat{\mathbf{s}}_1 := \arg\min_{\mathbf{s}_1} \left[\mathcal{L}_{\mathsf{data}}(\mathbf{s}_1) + \mathcal{L}_{\mathsf{prior}}(\mathbf{s}_1) + \mathcal{L}_{\mathsf{cross}}(\mathbf{s}_1)\right]$$

formulated using limited realizations of $\mathbf{n}(t)$

carefully normalized, eliminating loss weighting hyperparameters

with probabilistic recovery guarantees

data misfit loss: $\mathcal{L}_{\text{data}}(\mathbf{s}) = \frac{1}{K}\sum_{k=1}^{K} \frac{\left\|\Phi(\mathbf{s}+\mathbf{n}_k) - \Phi(\mathbf{x})\right\|_2^2}{\sigma^2(\Phi(\mathbf{x}+\mathbf{n}))}$

prior loss: $\mathcal{L}_{\text{prior}}(\mathbf{s}) = \frac{1}{K}\sum_{k=1}^{K} \frac{\left\|\Phi(\mathbf{x}-\mathbf{s}) - \Phi(\mathbf{n}_k)\right\|_2^2}{\sigma^2(\Phi(\mathbf{n}))}$

cross-covariance loss: $\mathcal{L}_{\text{cross}}(\mathbf{s}) = \frac{1}{K}\sum_{k=1}^{K} \frac{\left\|\Phi(\mathbf{s},\mathbf{n}_k)\right\|_2^2}{\sigma^2(\Phi(\mathbf{x},\mathbf{n}))}$

number of data snippets K

data misfit loss: $$\mathcal{L}_{\text{data}}(\mathbf{s}) = \frac{1}{K}\sum_{k=1}^{K} \frac{\left\| \Phi\left(\mathbf{s}+\mathbf{n}_k\right) - \Phi\left(\mathbf{x}\right) \right\|_2^2}{\sigma^2\left(\Phi\left(\mathbf{x}+\mathbf{n}\right)\right)}$$

prior loss: $$\mathcal{L}_{\text{prior}}(\mathbf{s}) = \frac{1}{K}\sum_{k=1}^{K} \frac{\left\| \Phi\left(\mathbf{x}-\mathbf{s}\right) - \Phi\left(\mathbf{n}_k\right) \right\|_2^2}{\sigma^2\left(\Phi\left(\mathbf{n}\right)\right)}$$

cross-covariance loss: $$\mathcal{L}_{\text{cross}}(\mathbf{s}) = \frac{1}{K}\sum_{k=1}^{K} \frac{\left\| \Phi\left(\mathbf{s},\mathbf{n}_k\right) \right\|_2^2}{\sigma^2\left(\Phi\left(\mathbf{x},\mathbf{n}\right)\right)}$$

number of data snippets K

data misfit loss: $\mathcal{L}_{\text{data}}(\mathbf{s}) = \frac{1}{K} \sum_{k=1}^{K} \frac{\left\| \Phi(\mathbf{s} + \mathbf{n}_k) - \Phi(\mathbf{x}) \right\|_2^2}{\sigma^2\big(\Phi(\mathbf{x} + \mathbf{n})\big)}$

prior loss: $\mathcal{L}_{\text{prior}}(\mathbf{s}) = \frac{1}{K} \sum_{k=1}^{K} \frac{\left\| \Phi(\mathbf{x} - \mathbf{s}) - \Phi(\mathbf{n}_k) \right\|_2^2}{\sigma^2\big(\Phi(\mathbf{n})\big)}$

cross-covariance loss: $\mathcal{L}_{\text{cross}}(\mathbf{s}) = \frac{1}{K} \sum_{k=1}^{K} \frac{\left\| \Phi(\mathbf{s}, \mathbf{n}_k) \right\|_2^2}{\sigma^2\big(\Phi(\mathbf{x}, \mathbf{n})\big)}$

number of data snippets K

Theorem
(informal)

Let $\mathbf{x} = \mathbf{s} + \mathbf{n}$ *with* $\mathbf{s}$ *and* $\mathbf{n}$ *two independent processes. Given two independent processes* $\widehat{\mathbf{s}}$ *and* $\widehat{\mathbf{n}}$ *with* $\mathbf{x} = \widehat{\mathbf{s}} + \widehat{\mathbf{n}}$*, if:*

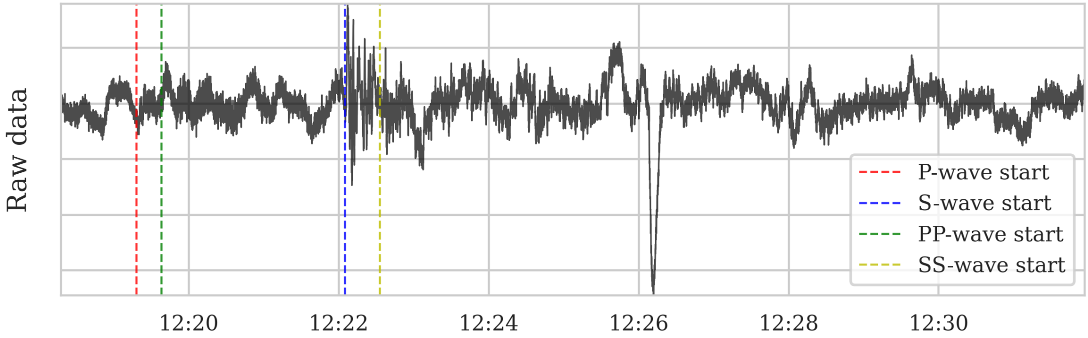
- $\mathbf{n}$ *and* $\widehat{\mathbf{n}}$ *have a maximum entropy under moment constraints* $\mathbb{E}\{\Phi(\mathbf{n})\} = \mathbb{E}\{\Phi(\widehat{\mathbf{n}})\}$
- *The Fourier transform of the distribution of* $\mathbf{n}$ *is non-zero everywhere*

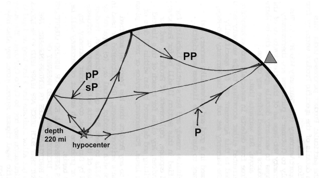
$\Rightarrow \mathbf{n} \stackrel{d}{=} \widehat{\mathbf{n}}$ *and* $\mathbf{s} \stackrel{d}{=} \widehat{\mathbf{s}}$ *where the equality is on the distribution of the processes.*

Example: Separating non-seismic noise from marsquakes

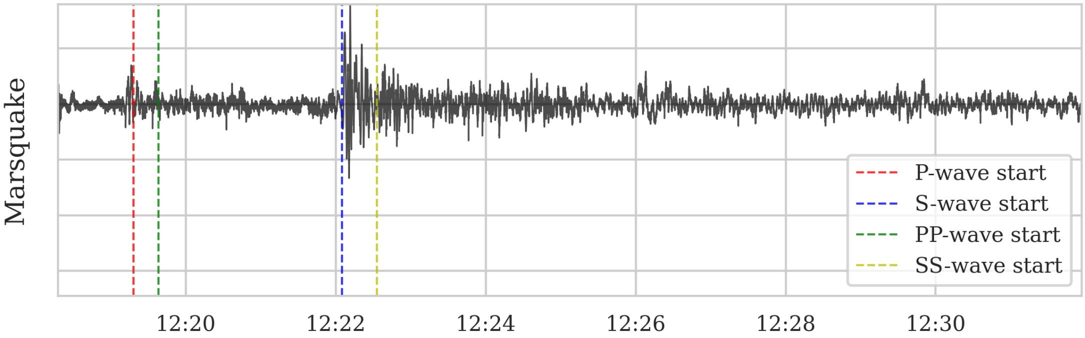
recorded by the InSight lander's seismometer on July 26, 2019

InSight Marsquake Service. *Mars Seismic Catalogue, InSight Mission; V13 2023-01-01*. en. 2023. DOI: 10.12686/a19. URL: https://www.insight.ethz.ch/seismicity/catalog/v13.





Berkeley Seismology Lab. *Idealized path of various seismic waves*. Accessed on June 30, 2023. URL: http://seismo.berkeley.edu/gifs/20120101wavepaths.jpg.



Marsquake
P-wave start
S-wave start
PP-wave start
SS-wave start
12:20
12:22
12:24
12:26
12:28
12:30

Conclusions

unsupervised source separation made possible with limited access to data

enabled by the inductive bias of wavelet scattering covariances

minimal assumptions on data and sources

otherwise impossible with existing unsupervised methods

https://github.com/alisiahkoohi/insight_src_sep