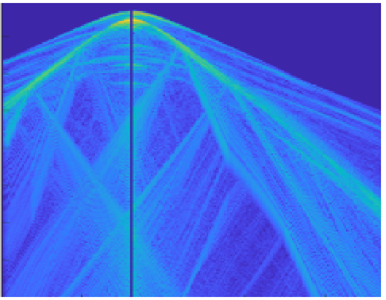
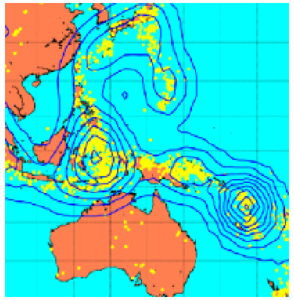
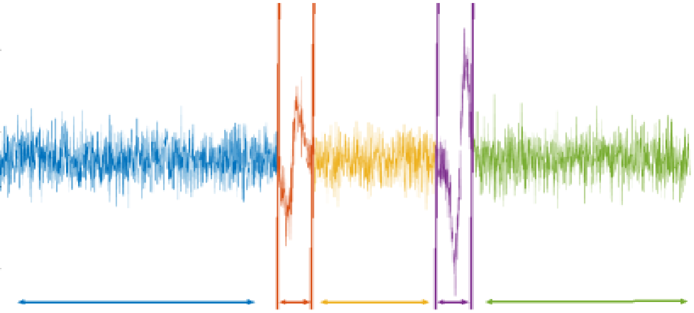
This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

SAND2022-4640C







Sandia National Laboratories

Exceptional service in the national interest

Quantifying the Impact of Simulation Frequency Fidelity on Waveform-Based Bayesian Inference for Seismic Monitoring Using Bayesian Experimental Design

T. A. Catanach

SSA 2022

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

This research was funded by the National Nuclear Security Administration, Defense Nuclear Nonproliferation Research and Development (NNSA DNN R&D). The authors acknowledge important interdisciplinary collaboration with scientists and engineers from LANL, LLNL, MSTS, PNNL, and SNL. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology &Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.__________

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Overview

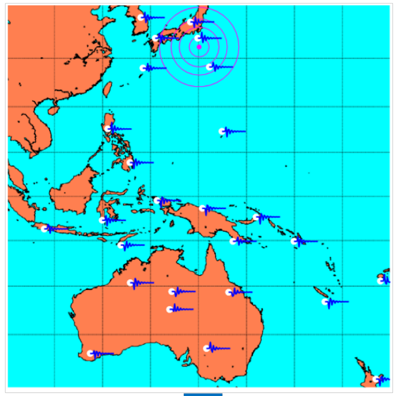
- Computational tools allow high frequency (>1 Hz) seismic simulation; however, modeling uncertainties limits the accuracy of these results.
- How would refining simulations to higher frequencies increase Bayesian seismic monitoring capabilities e.g. improving inference of event parameters with uncertainty?

  Latitude, Longitude, Depth, Origin Time, Source Time Function, and Moment Tensor

- **Target Contribution:**
  1) Outline how to use the Bayesian experimental design to quantify the effect on simulation frequency on waveform-based seismic monitoring.
  2) Apply this method to simple models constructed to explore the effect of frequency content on a simulated inference problem for local monitoring.
  3) Identify future research directions.

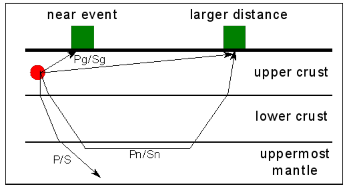
The Bayesian Approach

Data: $\mathcal{D}$

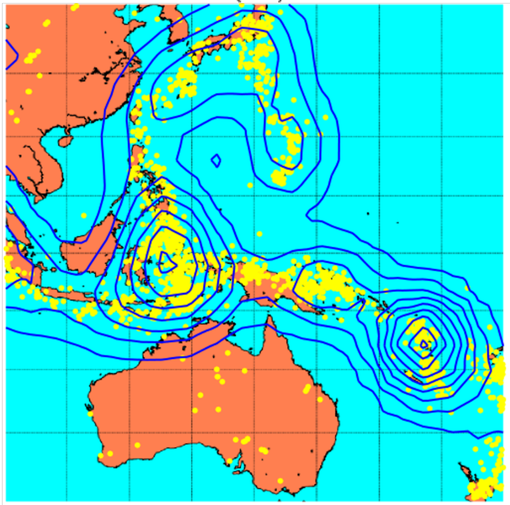


Likelihood: $p(\mathcal{D} \mid \theta)$

Physics Model
Sensor Model
Uncertainty Model



Prior: $p(\theta)$

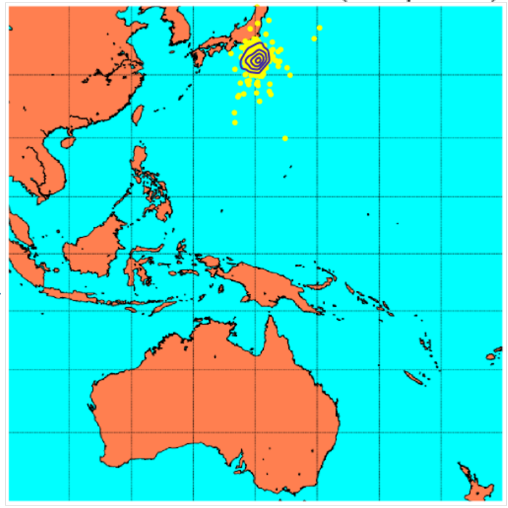


Knowledge about where events are likely to occur

Bayes' Theorem:

$$p(\theta \mid \mathcal{D}) = \frac{p(\mathcal{D} \mid \theta)\, p(\theta)}{p(\mathcal{D})}$$

Posterior: $p(\theta \mid \mathcal{D})$



Updated knowledge about where a specific event occurred

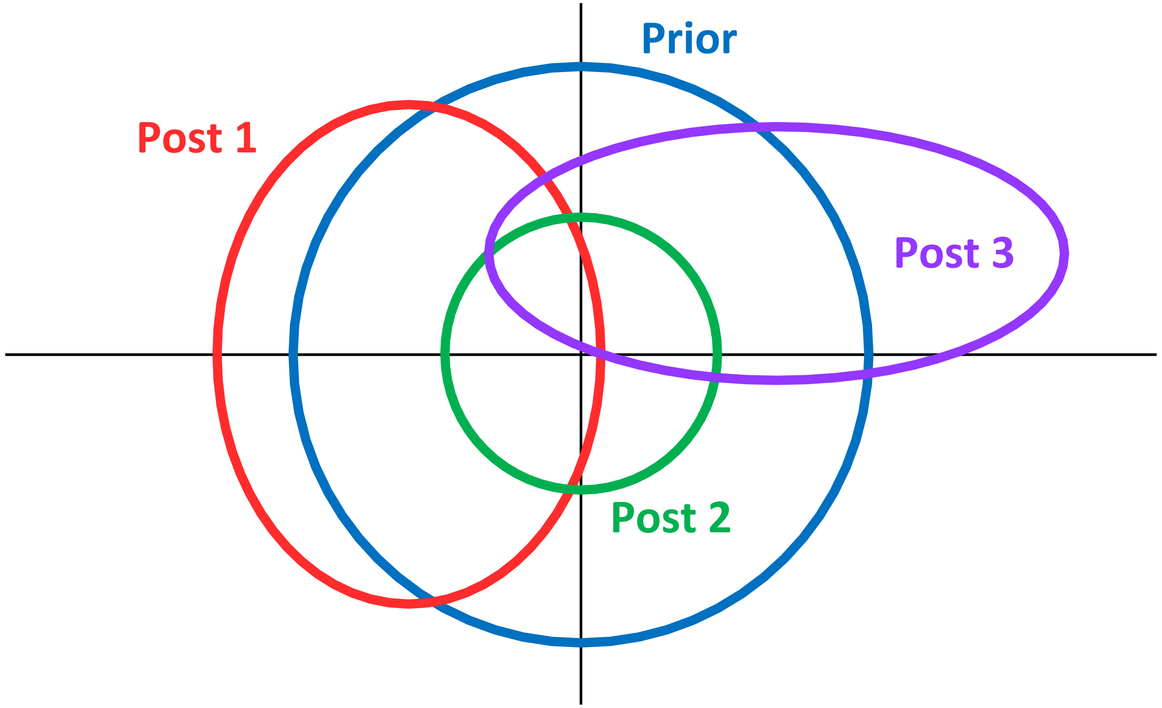
Quantifying Information Gain

Bayesian Inference:

$$p(\theta \mid \mathcal{D}) = \frac{p(\mathcal{D} \mid \theta)\, p(\theta)}{p(\mathcal{D})}$$

Kullback-Leibler (KL) Divergence measures the information due to inference:

$$\mathrm{KL}\left[p(\theta \mid \mathcal{D}) \,||\, p(\theta)\right] = \int p(\theta \mid \mathcal{D}) \log \frac{p(\theta \mid \mathcal{D})}{p(\theta)} d\theta$$



Prior → Post 1	0.5 Bits
Prior → Post 2	1 Bit
Prior → Post 3	1 Bit

Quantifying Information Gain

Expected Information Gain (EIG) from an experiment ($\mathcal{S}$):

$$\mathcal{I}(\mathcal{S}) = \mathrm{E}\left[\mathbf{KL}\left[p(\theta \mid \mathcal{D}) \,||\, p(\theta)\right] \mid \mathcal{D} \sim p(\mathcal{D} \mid \mathcal{S})\right]$$

$$= \int p(\mathcal{D} \mid \mathcal{S}) \int p(\theta \mid \mathcal{D}, \mathcal{S}) \log \frac{p(\theta \mid \mathcal{D}, \mathcal{S})}{p(\theta)} d\theta d\mathcal{D}$$

Distribution of hypothetical data

$$p(\mathcal{D} \mid \mathcal{S}) = \int p(\mathcal{D} \mid \theta', \mathcal{S})\, p(\theta')\, d\theta'$$

KL Divergence to measure information gain

Quantifying Information Gain

Distribution of hypothetical data

$$p\,(\mathcal{D} \mid \mathcal{S}) = \int p\,(\mathcal{D} \mid \theta', \mathcal{S})\, p\,(\theta')\, d\theta'$$

θ: Source parameters in the full-waveform simulations e.g. origin time, latitude, longitude, depth, magnitude, source time function

$\mathcal{S}$: Simulation frequency fidelity modeled as a low pass filter

$\mathcal{D}$: Simulated waveform plus additive background noise after filtering

Evaluation Algorithm Outline

1) Create a representative set of seismic sources with different locations and source properties
2) Simulate high-frequency waveforms for each of these sources and add background noise from a known noise model
3) For each waveform apply a set of low pass filters with different cutoff frequencies
4) For each representative event and filtered waveform solve the Bayesian inference problem to find the posterior distribution on the event parameters and compute the information gain
5) Average the information gain over all events for each of the filters to capture the effect of frequency content on seismic monitoring.

Bayesian Problem Setup

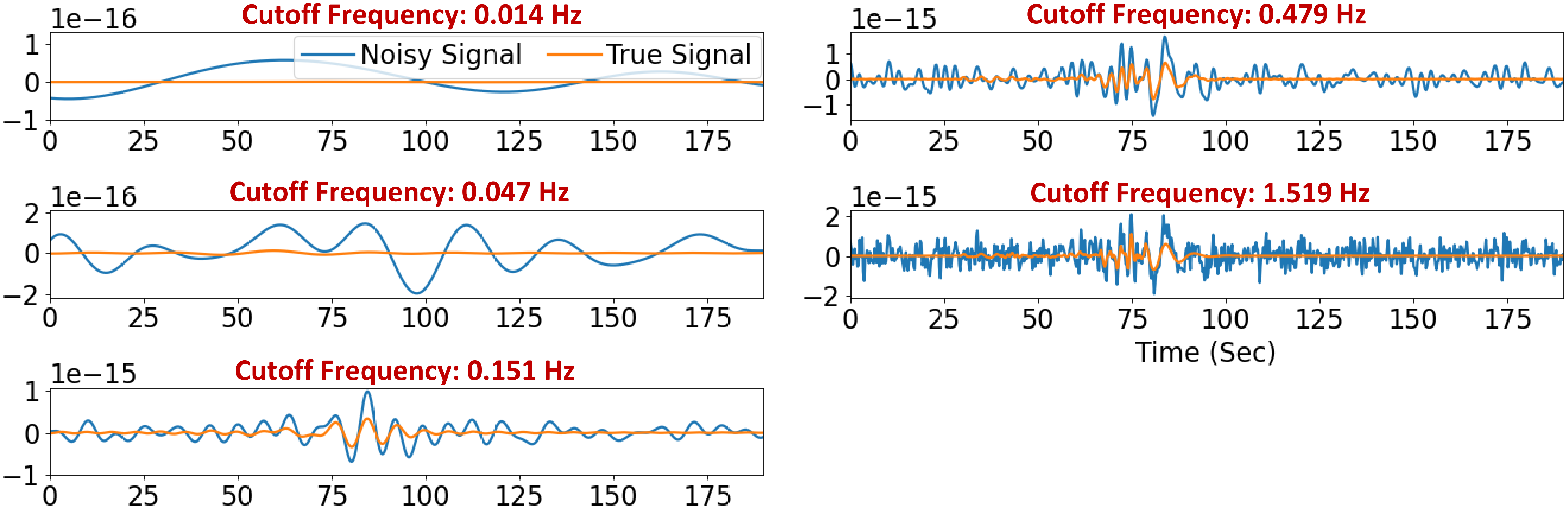
Model Parameters

- Regional domain ±2°in latitude and longitude, 40km in depth
- Exponentially distributed log10 moment magnitude factor between -2 to 2 and isotropic moment tensor
- Source time function width 0 to 5 seconds
- Origin time 0 to 152 seconds

Simulated Waveforms

- AxiSEM/Instaseis simulations with an AK135 earth model
- Sample rate 3.95 Hz
- Additive white background noise
- Low pass filtered using sinc filter

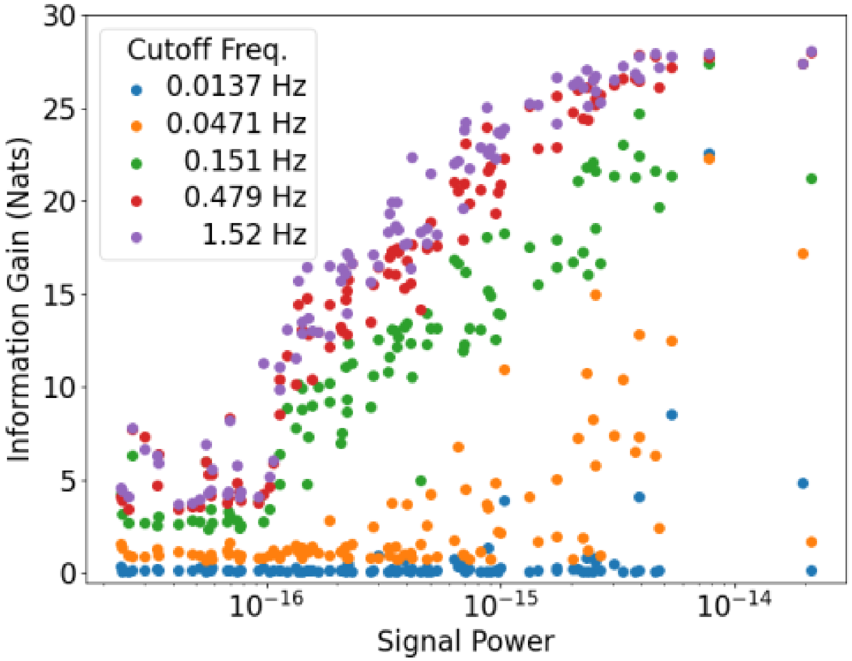
Illustration of filtered waveforms

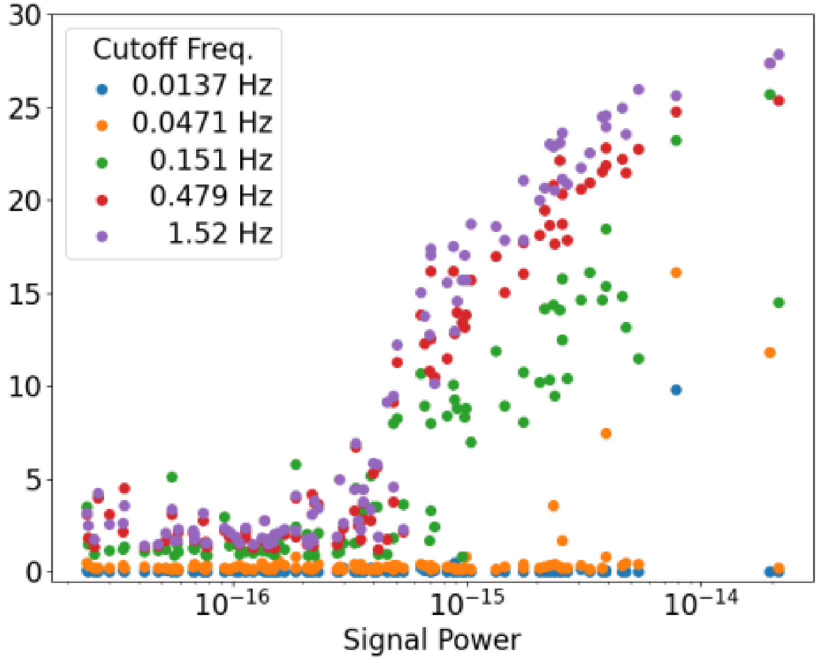


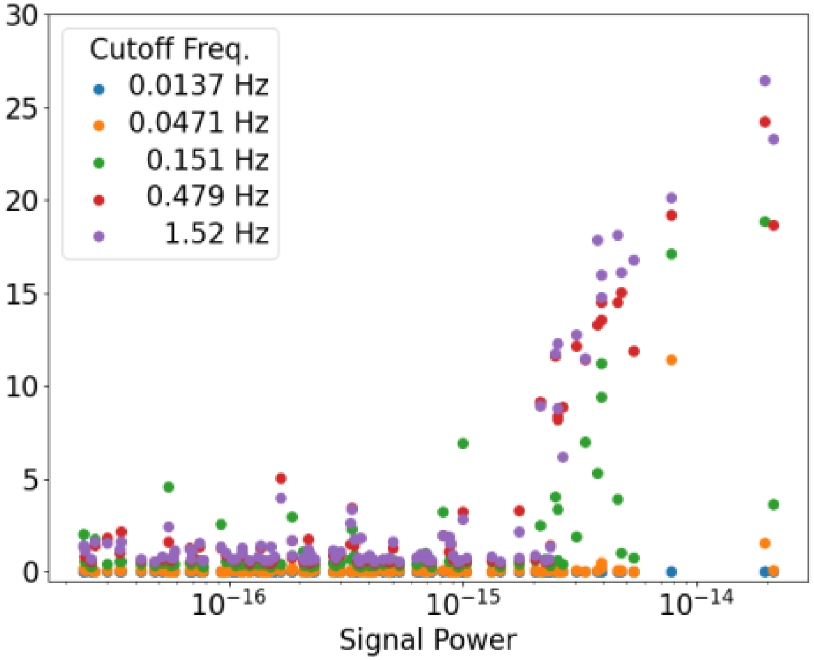
Results: Information gain for each representative event for the five filters vs the true signal power.

Background noise level

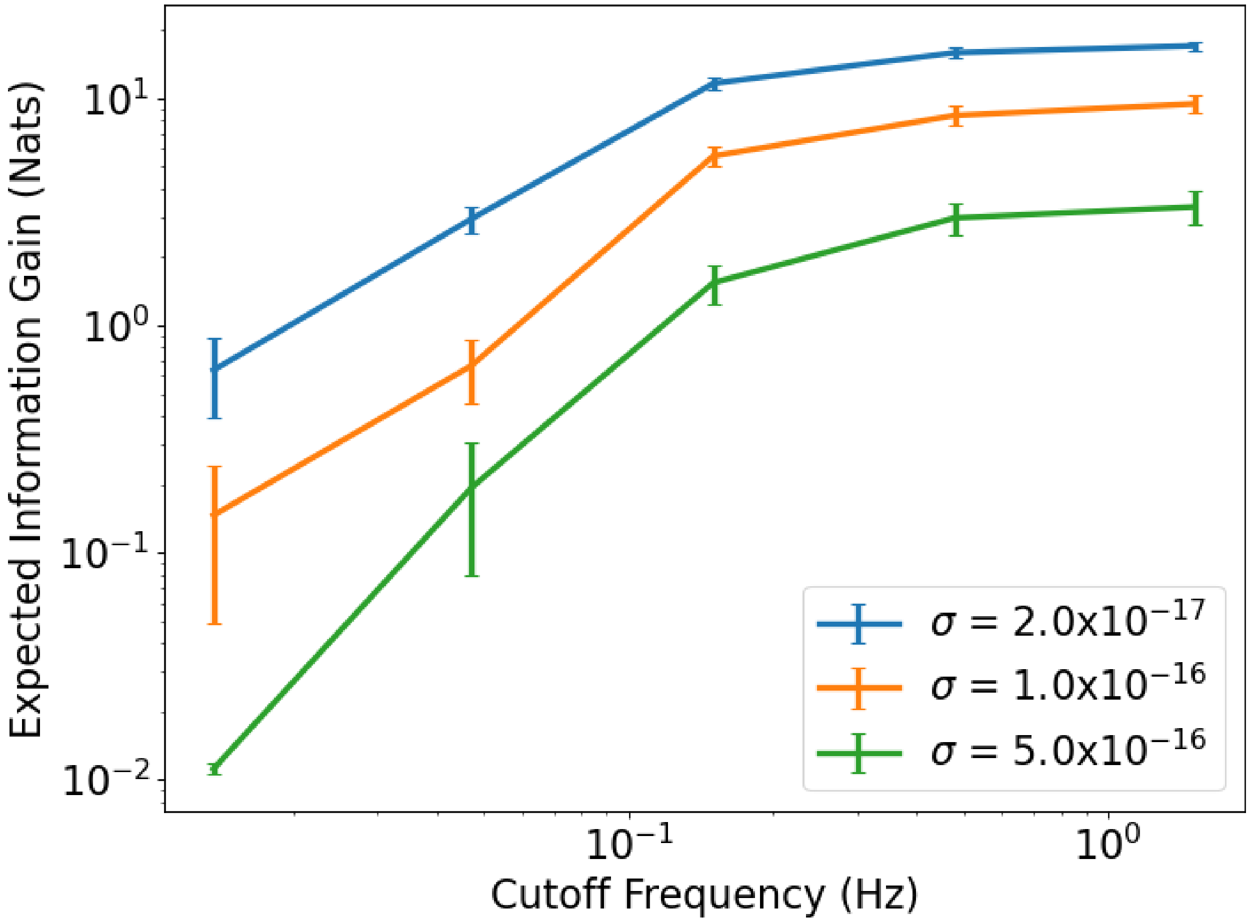
σ=2x10^{-17} σ=1x10^{-16} σ=5x10^{-16}







Results: Expected information gain for different cutoff frequencies at different background noise levels



Conclusion

Discussion

- Using the framework of Bayesian experimental design we can quantify the utility of frequency content in seismic simulation for seismic monitoring.
- Under the assumptions of an AK135 earth model and white background noise process, we observe limited contribution of frequency information above 0.5 Hz.

Future Directions

- Extend this analysis to frequencies up to 10Hz.
- Leverage higher fidelity simulation codes and choose a representative earth model with more complexity than AK135.
- We made simple assumptions about the background and source mechanism to facilitate computation, relaxing the assumptions and using more realistic models may influence these results.

Backup: Mathematica Model of Likelihood

Sandia National Laboratories

Likelihood model in the frequency domain:

- Let ω_j be the Discrete Fourier Transform (F) of the predicted waveform at frequency j for an event characterized by (Lat-Lon L, Depth z, Magnitude m, Origin Time t_o, and STF width λ)

$$\omega_j\left(\mathcal{L}, z, m, t_o, \lambda\right) = F_j \bar{w}\left(\mathcal{L}, z, m, t_o, \lambda\right) \mathbb{1}\left[s_j < f\right]$$

- Then the likelihood of the observed Discrete Fourier Transform (ξ) up to frequency f given the predicted waveform is

$$p\left(\xi_1 \ldots \xi_{n_f} \mid \mathcal{L}, z, m, t_o, \lambda\right) = p\left(\gamma_{01} = \text{Real}\left[\xi_1 - \omega_1\left(\mathcal{L}, z, m, t_o, \lambda\right)\right]\right) \times \prod_{j=2}^{n_f} p\left(\gamma_{0j} = \text{Real}\left[\xi_j - \omega_j\left(\mathcal{L}, z, m, t_o, \lambda\right)\right]\right) p\left(\gamma_{1j} = \text{Imag}\left[\xi_j - \omega_j\left(\mathcal{L}, z, m, t_o, \lambda\right)\right]\right)$$

- Where

$$\gamma_{01} \sim \mathcal{N}\left(0, n\sigma^2\right), j = 1 \qquad \gamma_{0j} \sim \mathcal{N}\left(0, \frac{n}{2}\sigma^2\right), j = 2 \ldots n$$

$$\gamma_{11} = 0, j = 1 \qquad \gamma_{1j} \sim \mathcal{N}\left(0, \frac{n}{2}\sigma^2\right), j = 2 \ldots n$$