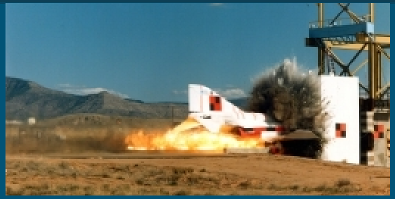
SAND2022-12113C

Sandia National Laboratories

Hampel Filtering of ASDs to Remove Spurious Sine Tones in Random Vibration Data



James Woodall- Dept. 1557
Vit Babuska – Dept. 1557

92ND SHOCK & VIBRATION SYMPOSIUM

SEPTEMBER 18-22, 2022

DENVER, CO

U.S. DEPARTMENT OF ENERGY NNSA National Nuclear Security Administration

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.

Presentation Outline

Introduction
- Problem Description
- Motivation

Background
- Median Filter
- Hampel Filter

Approaches
- Noise Floor Removal
- Filtered Noise Floor
- Difference of Squares
- Least Squares (LS) Model

Conclusions

Introduction

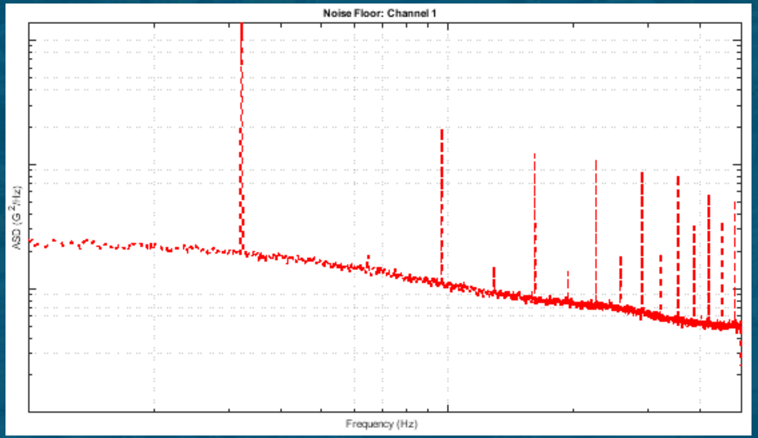
Problem Description

A Transportation test was conducted during which several sensors recorded vibration data at different points in the vehicle as it drove on a variety of roads

The objective of the test was to determine transportation environments for the cargo

The test had 3 phases

- Engine off quiescent phase to determine the noise floor
- Engine on static phase (no vehicle motion)
- Driving phase – Driving on various roads at various speeds



During the engine off quiescent phase, sine-tone spikes were observed in the ASDs

- The tones correspond to no known physical cause, and are assumed to be electrical noise
  - They appear in the ASDs in the other phases

Motivation

The sine-tones should be removed so they do not pollute the ASDs during the Travel phase and skew the measured environments

- Removing the noise tones in the Travel phase ASDs should not affect legitimate data at those frequencies

Since the sine tones are part of the system noise characteristics it is assumed they are additive noise

Isolate the sine tones and subtract their contributions to the Driving Phase ASDs

- A Hampel (median) filter in the frequency domain was used to identify and isolate the tones
- Different subtraction approaches are discussed

Background

Median Filter

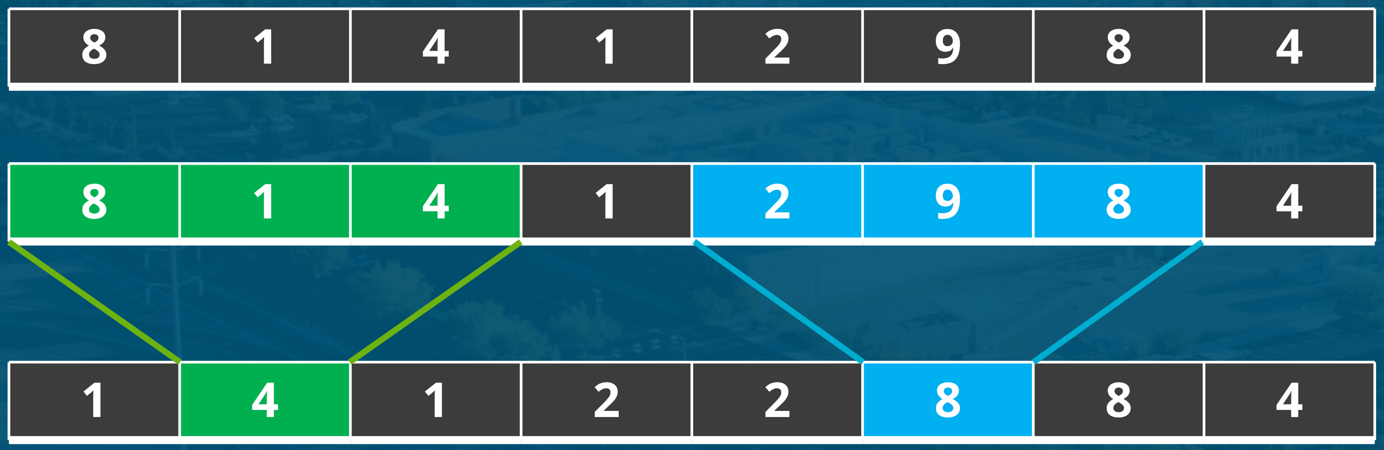
Median filters are used to remove noise from signals

- Used in digital image processing to remove noise artifacts like “popcorn” noise (hot pixels)

A median filters is a local smoothing filter that removes noise and preserves edges

Basic idea:

- Apply a sliding window over the data (usually an odd number of points)
- In each window the center value is replaced with the median value in the window



Source: MATLAB medfilt2 function help file

Hampel Filter

F.R. Hampel
1941-2018

A Hampel Filter is a median filter combined with outlier detection/identification

- Identify and replace outliers with the median value of the window in which the outliers occur

$$y_j = \begin{cases} x_j \text{ if } |x_j - m_j| \le k\sigma_j \\ m_j \text{ if } |x_j - m_j| > k\sigma_j \end{cases}$$

- An outlier is a data point that is differs from the window median by more than k standard deviations of the data in the window

There are 2 parameters to be selected – the window length and the exceedance factor, k

- The window length should be long enough so that the outliers do not bias the standard deviation

Our Innovation was to use the Hampel filter in the frequency domain on the PSDs

Approaches

Approaches

Noise floor removal

- Subtracts entire Noise Floor without filtering - reference method

Filtered noise floor removal

- Removes only the noise sine tones as identified by the Hampel Filter

Difference of roots

- Removes only the noise sine tones as identified by the Hampel Filter
- Similar to Filtered noise floor removal but assumes sine tones add to real content as squares rather than linearly

Least Squares (LS) model

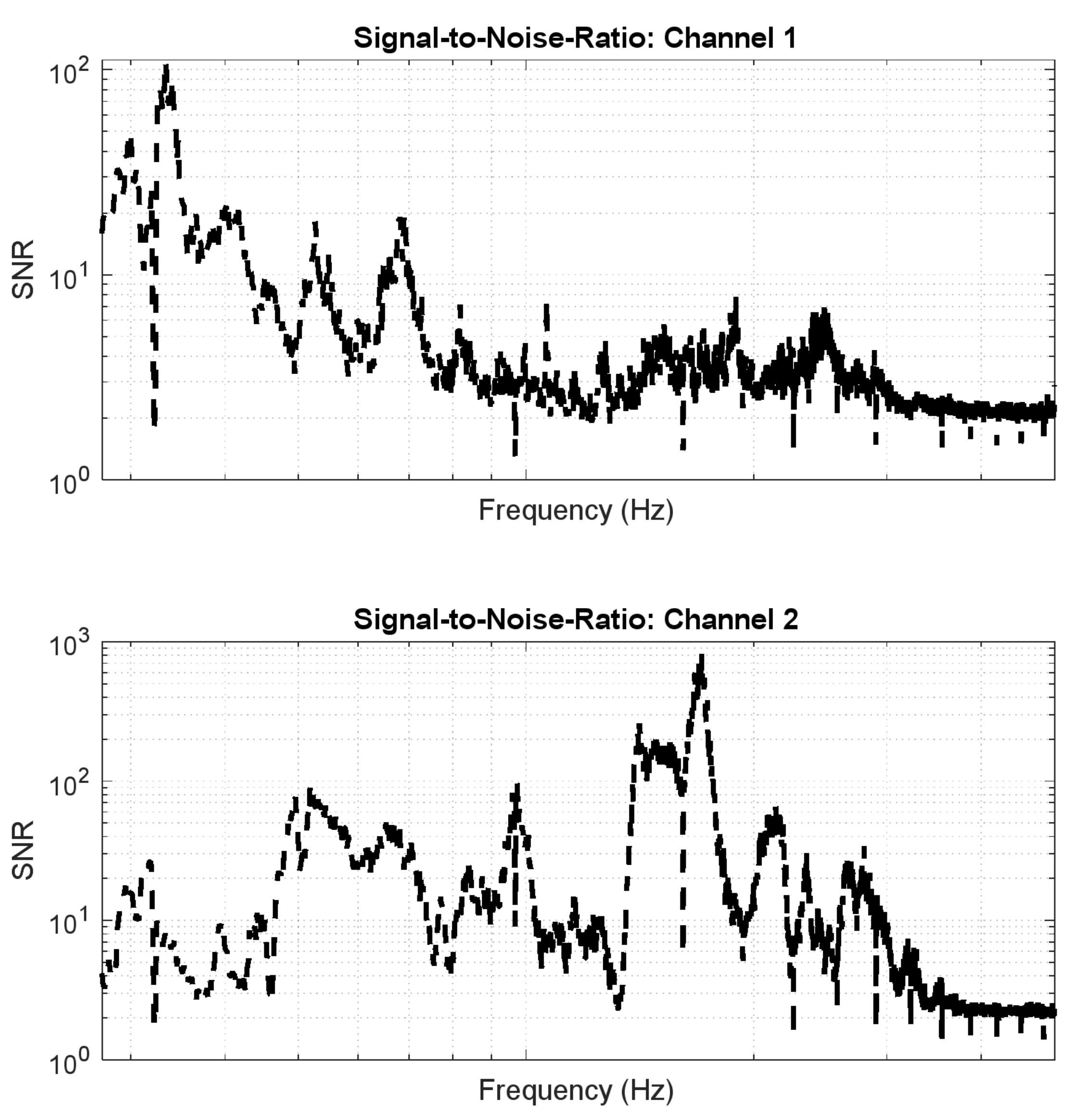
- Removes only the noise sine tones as identified by the Hampel Filter
- Empirically defines relationship between sine tones in the Static Phase and those in the Driving Phase

Noise Floor Removal Method

This method is simply the subtraction of the entire noise floor from the Driving Phase ASDs.

This will ideally have little effect at frequencies with real content, assuming a reasonably high Signal-to-Noise ratio.

The results will look the worst at higher frequencies, where the real signal content dies off and the Signal-to-Noise ratio drops.



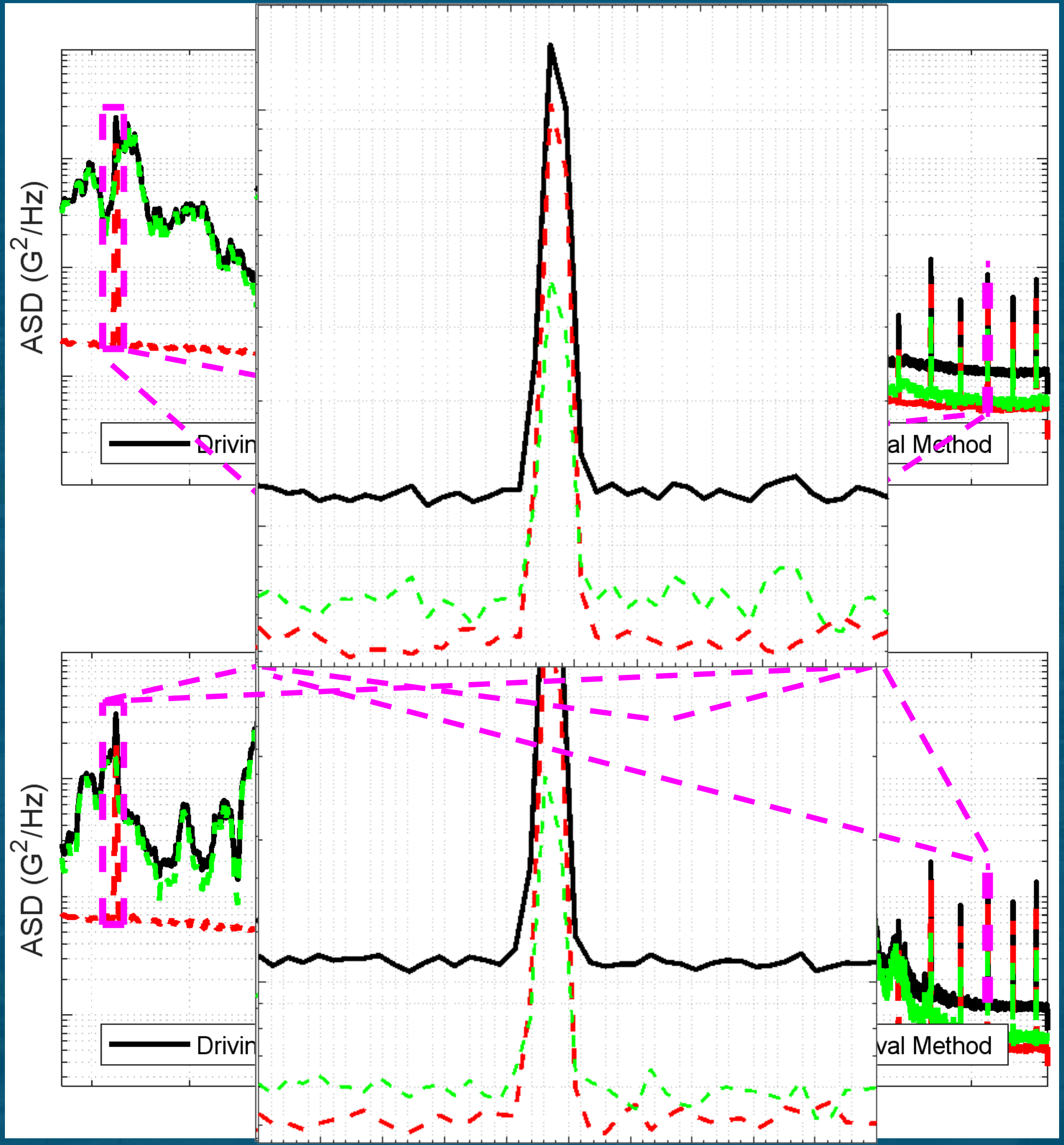
Noise Floor Removal Method

Frequencies with legitimate frequency content were not much affected, either where the spikes are located or otherwise.

Sine tone spikes were not significantly mitigated by this method, especially at low SNR regions.

At lower spectral content magnitudes, the effects are more pronounced- the ASD spikes have a lower magnitude, but the ambient signal around them are likewise noticeably reduced.

This method diminishes the Signal-to-Noise Ratio (SNR) by 1 across all frequencies. Where the SNR is high there is little overall effect, while where it's low the results are more easily seen.



Filtered Noise Floor Method

A Hampel filter was applied to the noise floor, which replaced points higher than 4 standard deviations from a 33 point window with the median value for that window. The difference between the original noise floor (ST) and the filtered noise floor (ST_{DS}) is subtracted from the ASDs of the Driving Phase.

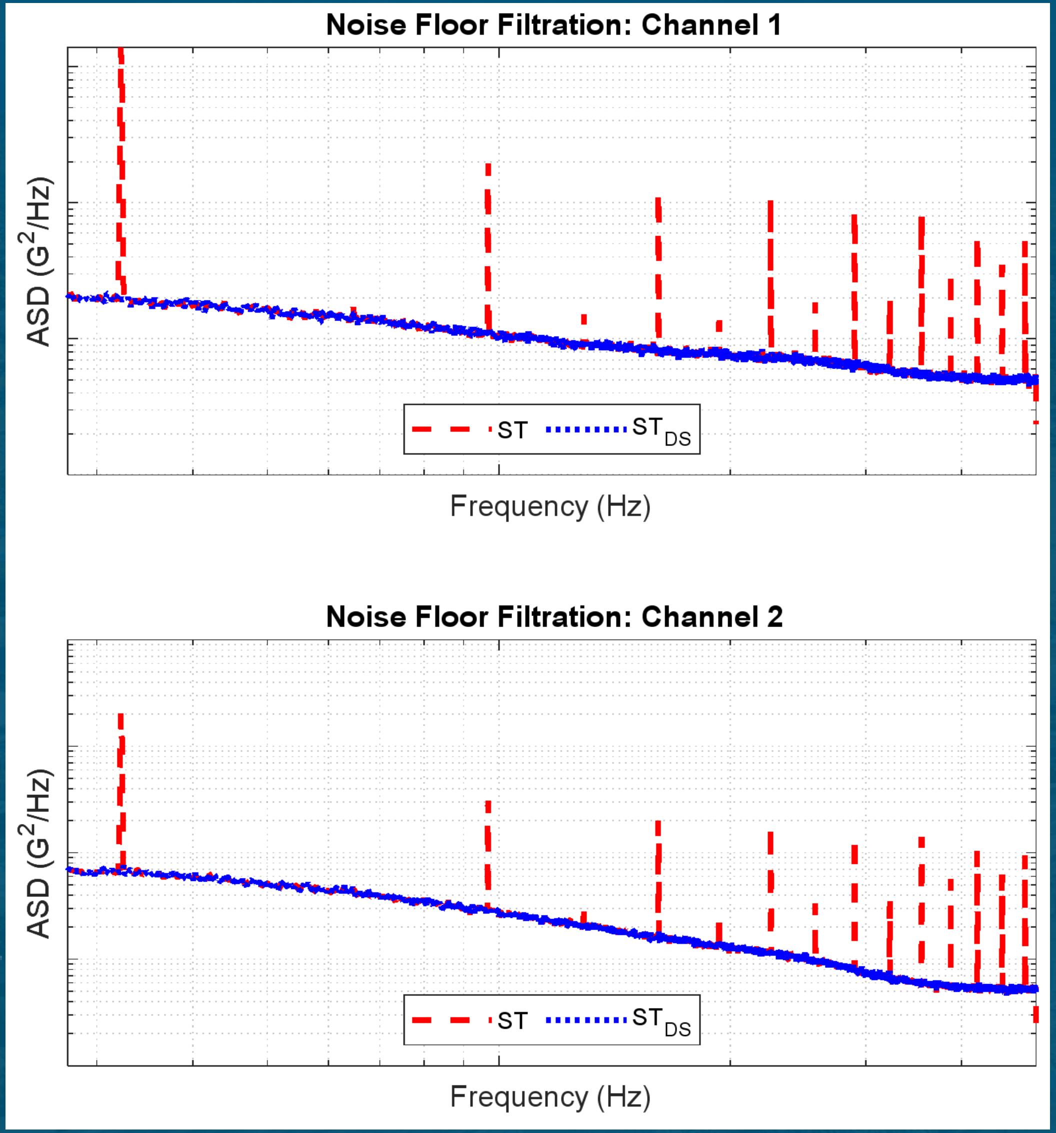
$$ASD_{DS} = ASD - (ST - ST_{DS})$$

Where ASD is the Driving Phase ASD

ASD_{DS} is the desired Driving Phase ASD without the artificial sine tones,

ST is the original noise floor, calculated by averaging the Static phase ASDs,

ST_{DS} is the noise floor after being subjected to a filter to remove the sine tones.

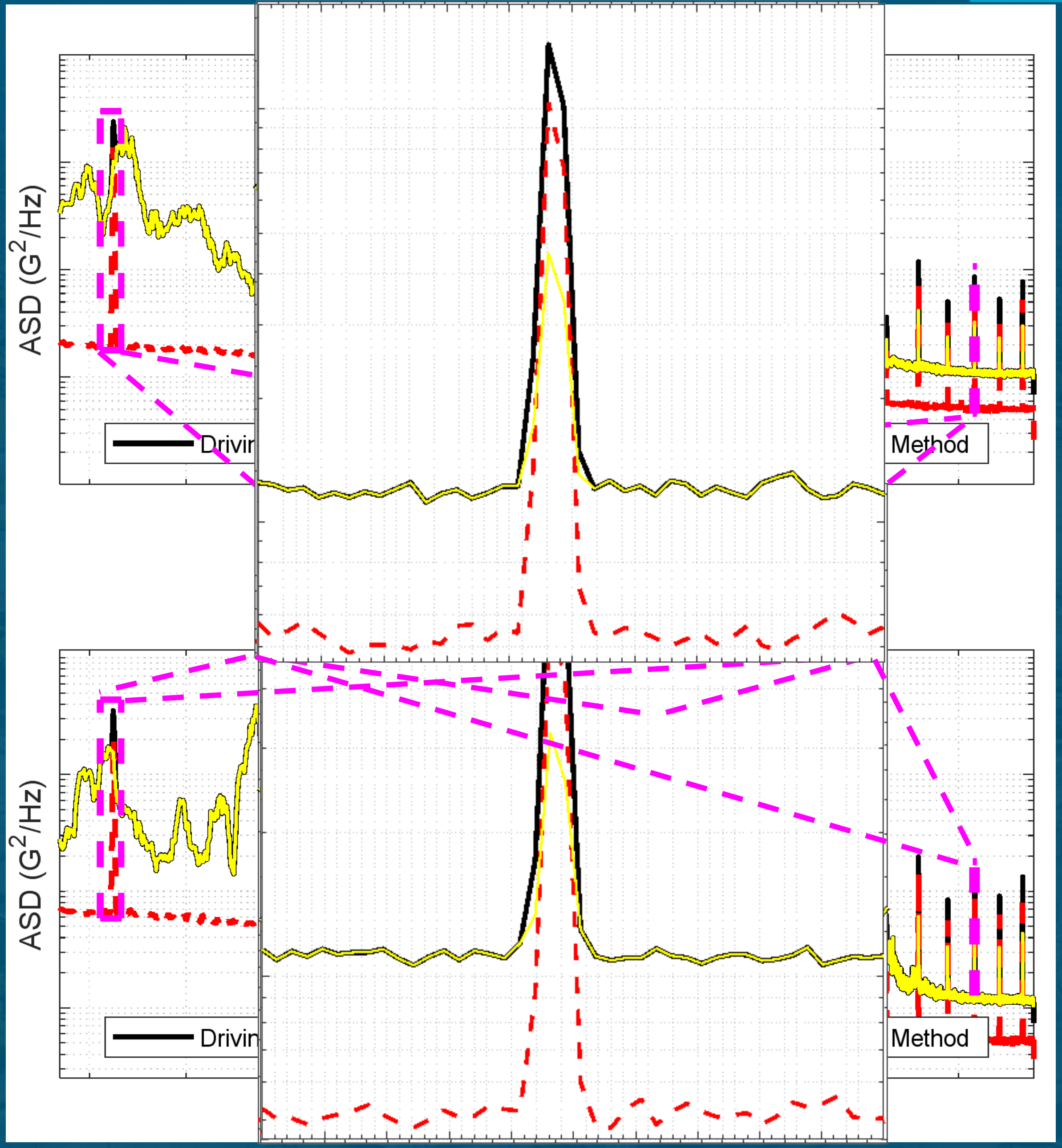


Filtered Noise Floor Method

As a result of using the Hampel Filter to identify and isolate the frequencies with sine tones, the ASDs experience less alteration.

In most cases, there still tend to be residual spikes in the ASDs after applying this method.

This is easiest to see at frequencies with low SNR.



Difference of Roots Method

This approach assumes that the relationship between the peaks in the ASDs and those in the Static phase is not linear:

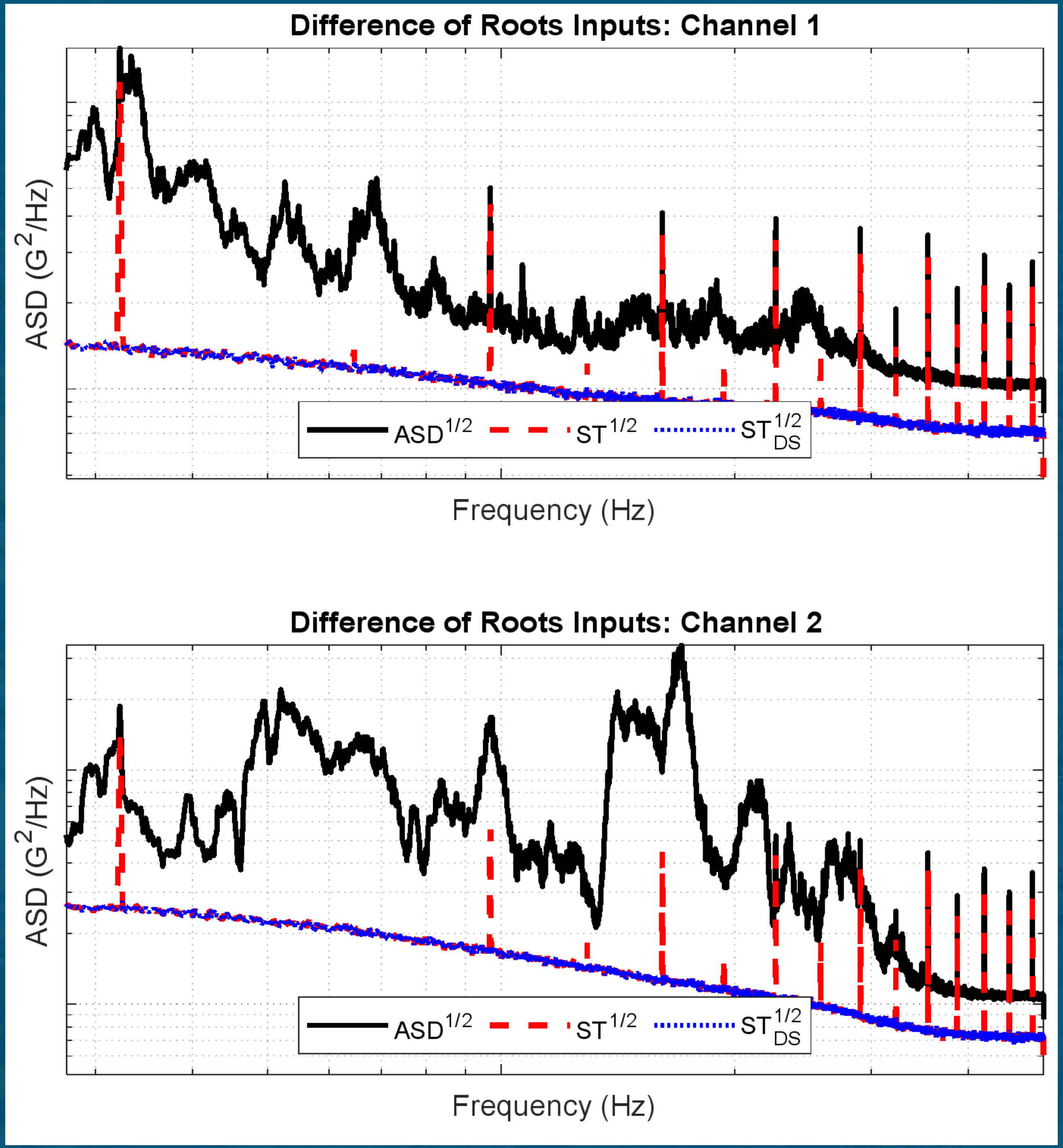
$$ASD_{DS} = \left(ASD^{1/2} - \left(ST^{1/2} - ST_{DS}^{1/2}\right)\right)^2$$

where ASD is the original ASD for a given phase,

ASD_{DS} is the desired ASD without the artificial sine tones,

ST is the original noise floor, calculated by averaging the Static phase ASDs,

ST_{DS} is the noise floor after removing the sine tones.

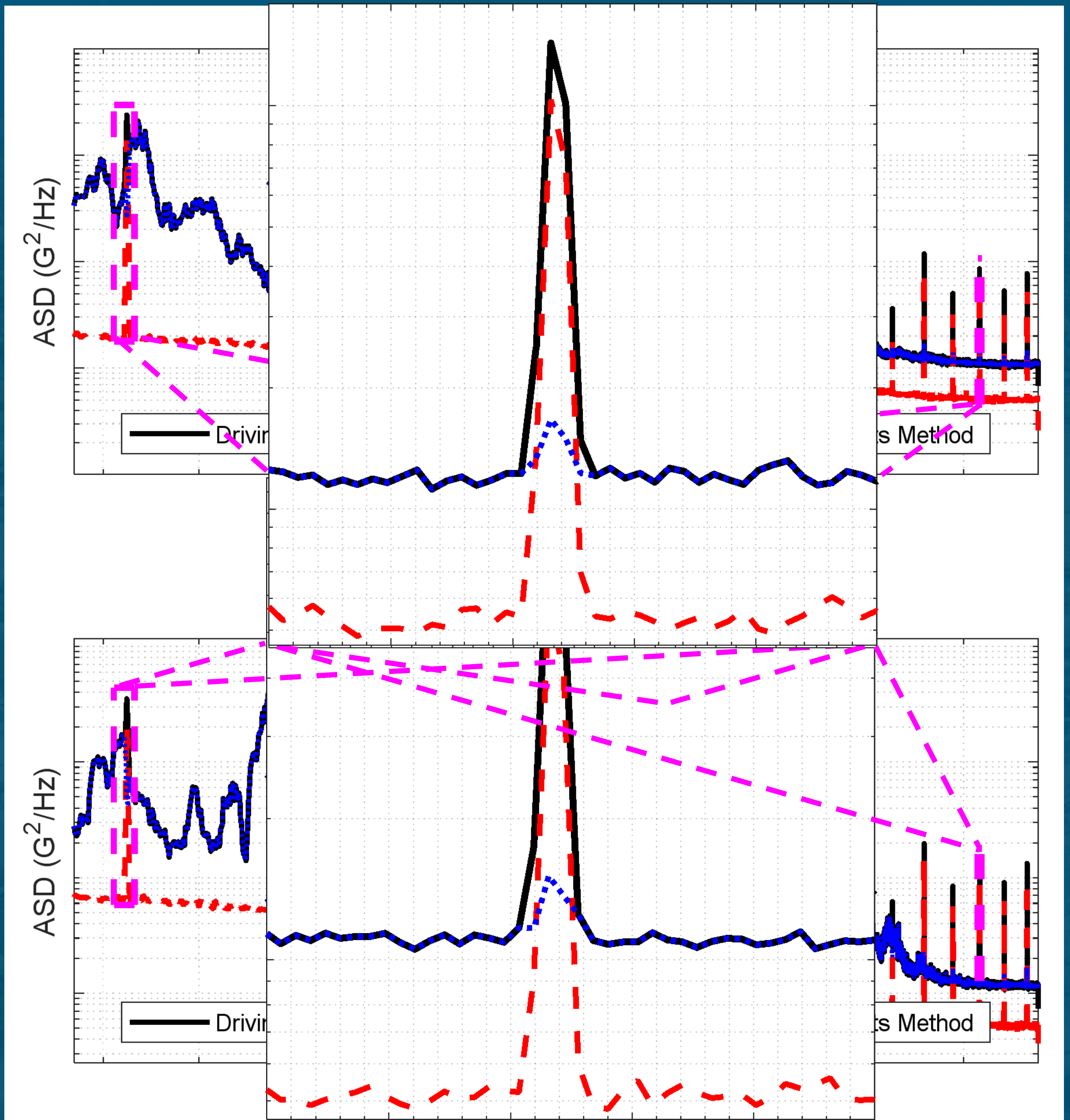


Difference of Roots Method

This method performs reasonably well where there is low signal content, although it still sometimes leaves small residual peaks.

At frequencies with high SNR, this method tends to overcompensate, bringing the ASDs lower than the surrounding frequencies would indicate is appropriate.

On the log scale where ASDs are traditionally viewed this method would appear to out-perform the previous methods, but at the frequencies with real spectral content where accuracy is most important, it can be worse than not subtracting the tones.



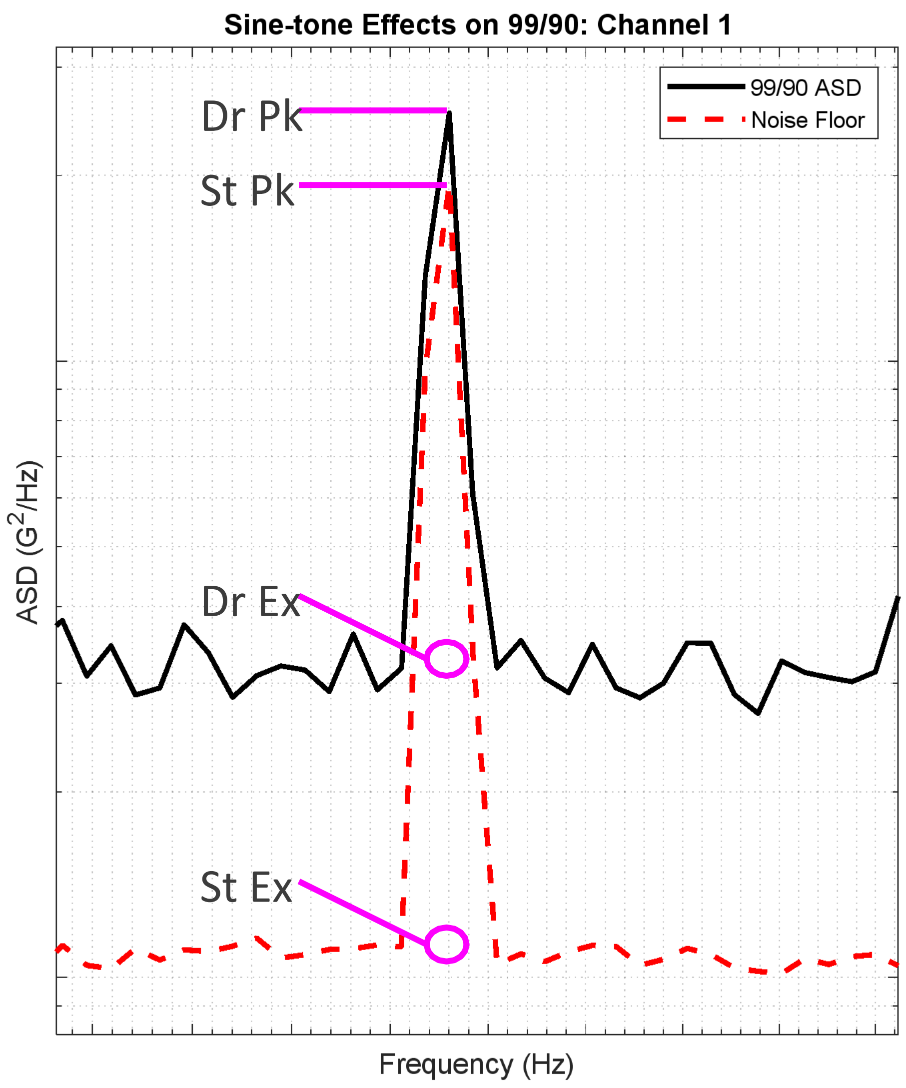
LS Model Method: Background

This model quantifies the relationship between the ASD spikes and those in the Static Phase with a Least Squares Model.

A collection of sine spike properties was recorded:

- The magnitude of the spike in the Static Phase (St Pk)
- The magnitude of the spike in the Driving Phase ASD (Dr Pk)
- The approximate real expected value for the Static Phase (St Ex)
- The approximate real expected value for the Driving Phase ASD (Dr Ex)

From these parameters the difference of the spike peak and the expected value (St Dif, Dr Dif) were also computed.



LS Model Method: Application

Dr Dif is linearly related to St Dif in the loglog scale, so a linear regression model was created for the logs of St Dif and Dr Dif.

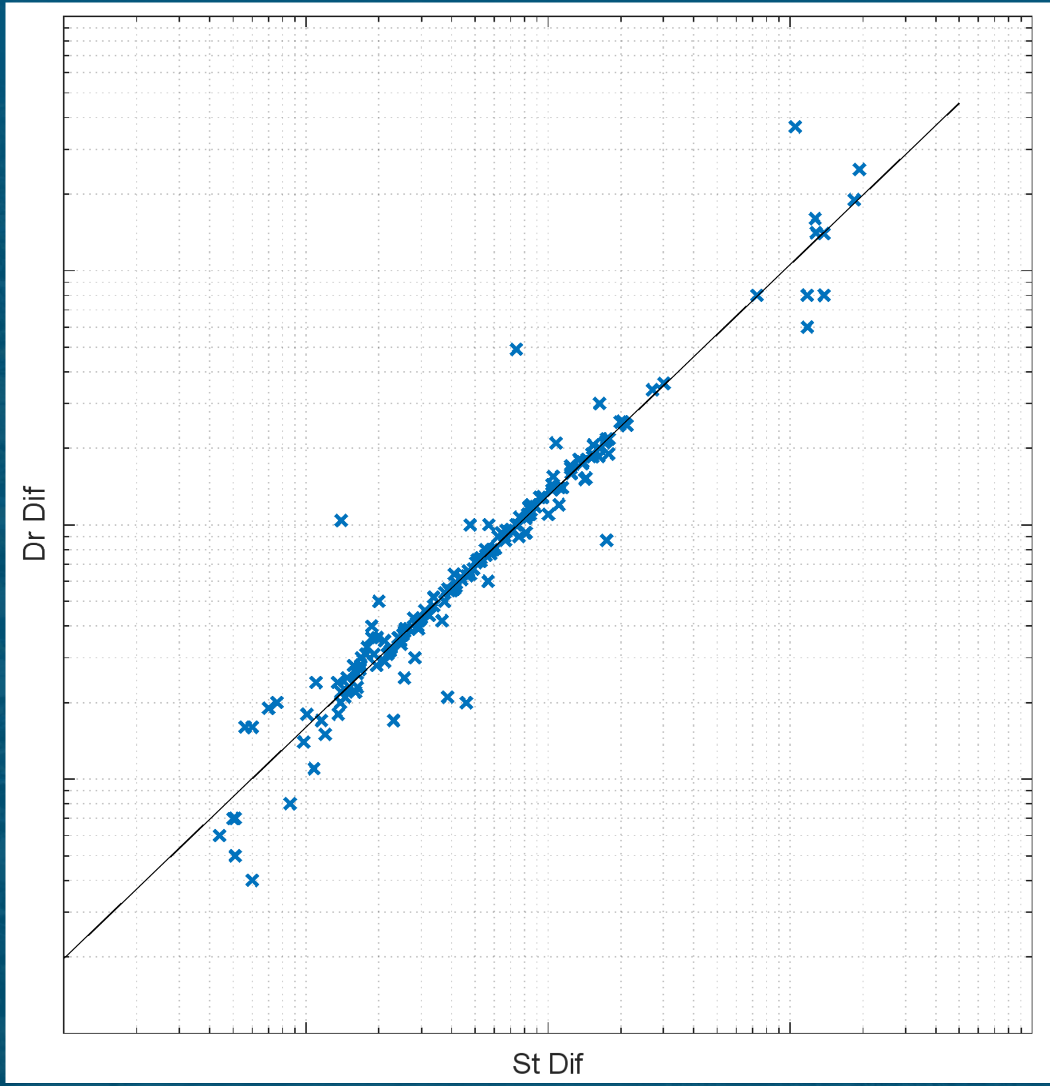
- This equation was then transformed to be implemented in the linear space.

LS Model: $log(Dr\ Dif) = A * log(St\ Dif) + B$

Expression Applied: $Y = X^A * e^B$

- where X is the St Dif axis, Y is the Dr Dif axis, and A and B are the slope and intercept for the least squares linear regression model using the logs of St Dif and Dr Dif respectively

The LS Model method subtracts Y from the Driving Phase ASDs where X is the sine-tone spike magnitude (which will be 0 at most frequencies).



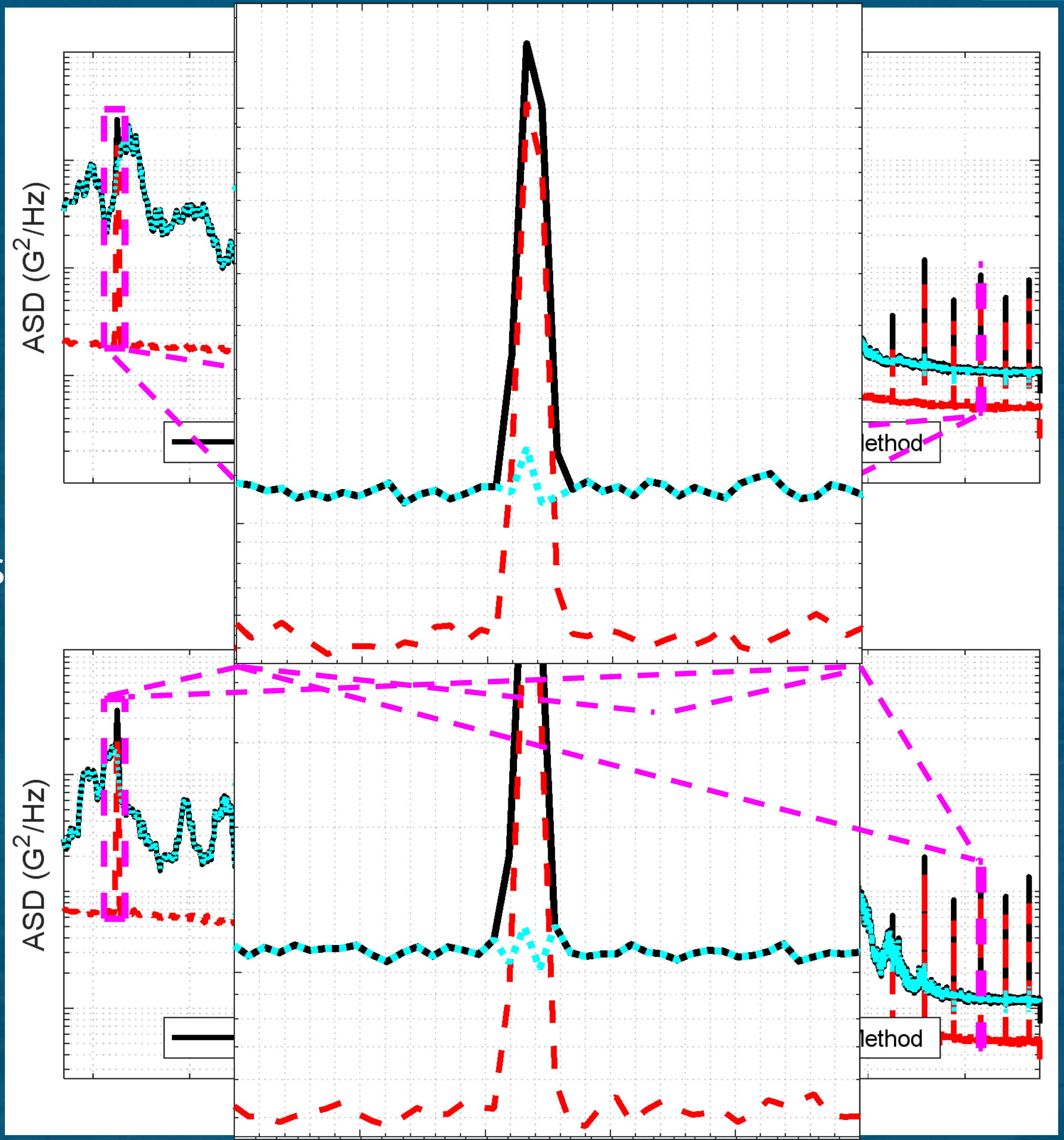
LS Model Method

This method seems to perform reasonably well at frequencies with higher real spectral content.

It brings down the artificial peaks to magnitudes that are believably within the range of expected values given behavior at surrounding frequencies.

This method also performs well at frequencies with low spectral content.

It tends to leave a distinct “jag” pattern where the artificial peak was, but the range of values are generally much closer to the ambient Driving Phase ASD levels than the other trialed methods can reliably achieve.



Conclusions

Method Comparison

- Noise Floor Removal affects every frequency and does little to diminish the relative sine tone magnitudes.
- Filtered Noise Floor affects only frequencies with artificial sine tones, but still leaves residual peaks.
- Difference of Roots tends to reduce the ASDs excessively where there is a high SNR, but is generally reasonable where there is a low SNR.
- LS Model tends to give the best overall estimates, but leaves a telltale “jag,” and depends on more upfront work to collect a reasonable sample-size to generate the model.