Qualitative exploration of stochastic amplification to XGBoost regression for guitar string tuning

Lyle Kenneth M. Geraldez*

National Institute of Physics, University of the Philippines - Diliman, Philippines
*Corresponding author: lgeraldez@nip.upd.edu.ph

Abstract

Guitar string frequency detection transforms time-domain data to frequency-domain data via Fourier transform methods and identifies the resonant peak frequency. The paper aims to explore the effects of stochastic amplification to a XGBoost regression model on peak frequency identification and resonance detection. Six guitar string frequencies were recorded and frequency-transformed for analysis. By testing an umamplified and amplified training data to test data, the study finds qualitative evidence of reduction of stochastic amplification effects near resonance and qualitative evidence of XGBoost modeling accuracy near resonance for D, G, and B strings.

Keywords: resonance, frequency analysis, xgboost

1 Harmonics and resonance

Nature offers a plethora of physical phenomena which contain periodic behavior. Specifically, mechanical oscillations manifest in the oscillatory dynamics of a medium on which it propagates itself. A ubiquitous phenomenon that can be observed daily is in the form of sound waves which, on an appropriate frequency interval, can be perceived by a human ear. These oscillatory phenomena can be analyzed conveniently using harmonic analysis by representing medium disturbance functions as superposition of basic waves (i.e. sine waves, cosine waves) forming the foundational applications of Fourier analysis.

An important application of harmonics and resonance analysis is present on frequency detection of instrument tuner such as the mobile guitar tuner mobile applicatons. Another important application is the detection of resonant peaks on mechanical systems; identification of resonance has dire consequences. Failure to execute resonance analysis has led to historical disaster such as the Millenium bridge incident which collapsed due to an apparent failure to analyze the consequence of the resonance of small-amplitude oscillations which, tragically, added up to form large-amplitude oscillations eventually destroying the bridge from unexpected driving agents.

Fourier analysis offers a powerful way to time-domain data, which is commonly used to detect and record oscillations, to frequency-domain data containing contribution of each individual frequencies to the overall spectrum. Alternatively, detection of resonance to system response can be conveniently executed by feeding a chirp signal to a system and recording its response.

2 Stochastic amplification

Typically, the presence of noise to a recorded signal weakens the quality of extracted data from the phenomenon hence requiring additional filtration techniques for noise removal. One interesting phenomenon is *stochastic amplification* where the presence of noise in a nonlinear system increases the output signal quality instead [1]. This noise-induced resonance has been applied to different fields such as electronic systems, biology, climate models, and physiological neural populations, among other things. Stochastic amplification works by the nonlinear consequence of noise resonance which causes the original signal to be amplified thereby counterintuitively increasing its signal-to-noise ratio. This paper aims to qualitatively explore the effects of stochastic amplification to the modeling accuracy of XGBoost regression for guitar strings' resonant frequency identification in standard E A D G B E tuning. Due to a possible increase in signal-to-noise ratio, incorporation of stochastic amplification may improve modeling accuracy near resonance. To do so, a one-second sample of each guitar string in standard tuning was recorded. The time-series data of the amplitude levels (in dB) was recorded. Using FFT, the data was converted from time-domain to frequency-domain. A Hann window function was applied to smooth the data for convenience with a sampling size of 1024. Three sets of data was recorded for each 6-string, 1-second sample: one serving as training data, one serving as stochastically amplified training data by incorporating a white-noise background, and one serving as testing data for validation. The peak/resonant frequencies from the results for each string are compared with the corresponding empirical frequencies expected from the standard tuning.

3 Results and Discussion

The time-domain data of different guitar strings were recorded and a 1-second sample for each string was isolated. By applying FFT to the time-series, resulting peak frequencies were compared with empirical frequencies; respectively, low E, A, D, G, B, and high E strings have empirical frequencies of approximately 82 Hz, 110 Hz, 147 Hz, 196 Hz, 247 Hz, and 330 Hz.

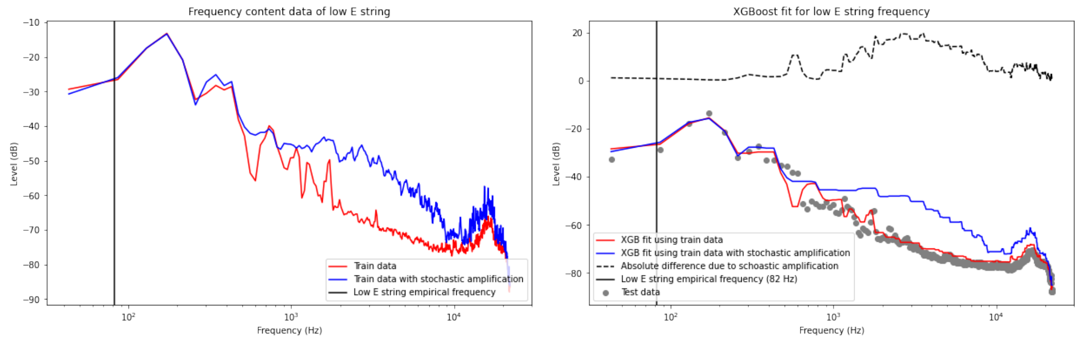

Figure 1: Right plot shows frequency-domain data of stochastically unamplified and amplified with the corresponding empirical frequency for low E string. XGBoost regression fit and the resulting absolute effects of stochastic amplification were shown on the right plot for low E string.

For the low E string, both stochastically amplified and unamplified resonant frequencies are at 172 Hz which is 90 Hz above the empirical frequency corresponding to a 109.76% deviation. The plots show that stochastic amplification has reduced effects on frequencies close to resonance. Observe that amplification effects reduce again at higher frequencies. This behavior also carries over the XGBoost fitting. Of course, in general, stochastic amplification drastically reduces modelling accuracy to unamplified test data.

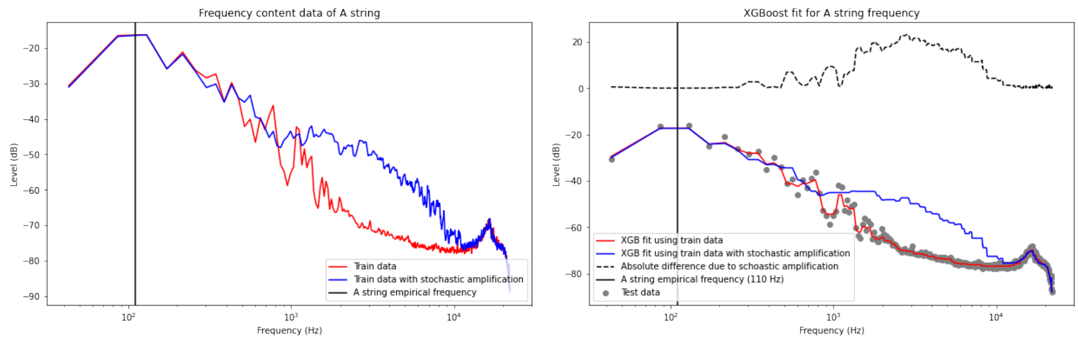

Figure 2: Right plot shows frequency-domain data of stochastically unamplified and amplified with the corresponding empirical frequency for A string. XGBoost regression fit and the resulting absolute effects of stochastic amplification were shown on the right plot for A string.

For the A string, both stochastically amplified and unamplified resonant frequencies are at 86 Hz which is 24 Hz below the empirical frequency corresponding to a 21.82% deviation. Again, stochastic amplification reduces in the frequency intervals close to resonance and for higher values. At the tail end of the plot, the reduction in amplification effects may be caused by a lack of high-frequency signals present in the generated white noise mixture. Without Fourier decomposing the generated white noise, the apparent fitting of amplified and unamplified frequencies at resonance may be considered inconclusive and may only be attribute to lack of low-frequency white noise signals.

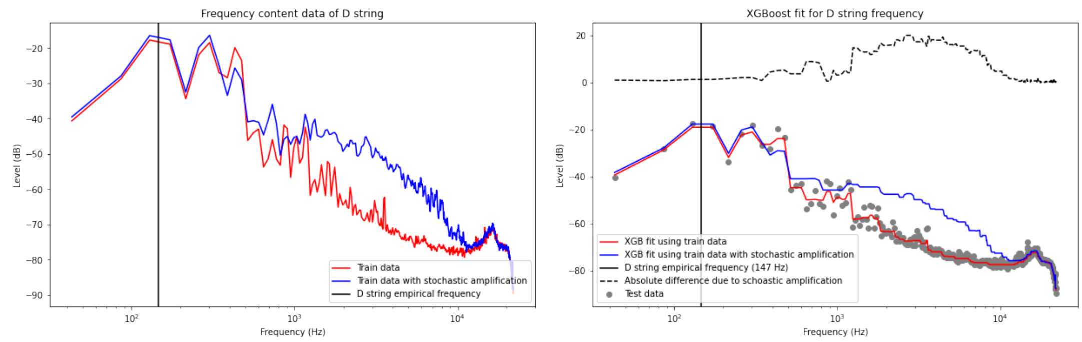

Figure 3: Right plot shows frequency-domain data of stochastically unamplified and amplified with the corresponding empirical frequency for D string. XGBoost regression fit and the resulting absolute effects of stochastic amplification were shown on the right plot for D string.

For the D string, both stochastically amplified and unamplified resonant frequencies are at 129 Hz which is 18 Hz below the empirical frequency corresponding to a 12.24% deviation. Similar stochastic amplification behavior can be observed. However, at this frequency, data shows some mismatch of amplified and unamplified data at resonance although the frequency peak itself remained unchanged. One can qualitatively eyeball that XGBoost model for amplified data "hits" more test data points closer than the unamplified counterpart.

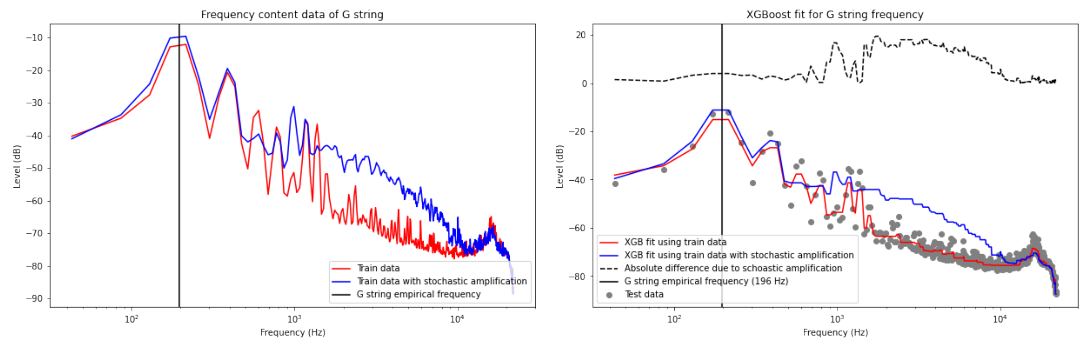

Figure 4: Right plot shows frequency-domain data of stochastically unamplified and amplified with the corresponding empirical frequency for G string. XGBoost regression fit and the resulting absolute effects of stochastic amplification were shown on the right plot for G string.

For the G string, both stochastically amplified and unamplified resonant frequencies are at 172 Hz which is 24 Hz below the empirical frequency corresponding to a 12.24% deviation. At the G frequency, the qualitative increase in accuracy near resonance can be seen more prominently. Also, one can observe that stochastically amplified data seems to adjust near resonance to hit more data points as it dips in amplitude level even when compared to the unamplified counterpart. This may be qualitative evidence of a stochastic amplification-induced increase in signal-to-noise ratio. Further analysis may help in providing an analytical or numerical supporting rationalization for the qualitative increase in the regression accuracy.

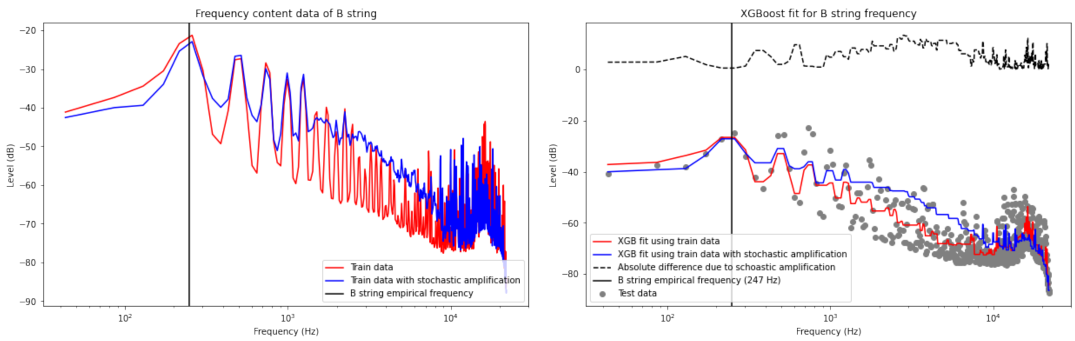

Figure 5: Right plot shows frequency-domain data of stochastically unamplified and amplified with the corresponding empirical frequency for B string. XGBoost regression fit and the resulting absolute effects of stochastic amplification were shown on the right plot for B string.

For the B string, both stochastically amplified and unamplified resonant frequencies are at 215 Hz containing which is 32 Hz below the empirical frequency corresponding to a 12.96% deviation. Here, one can observe a more apparent fitting of amplified and unamplified frequencies. Moreover, stochastic amplification fitting of more test data points near resonance became more apparent as resonant frequency increases.

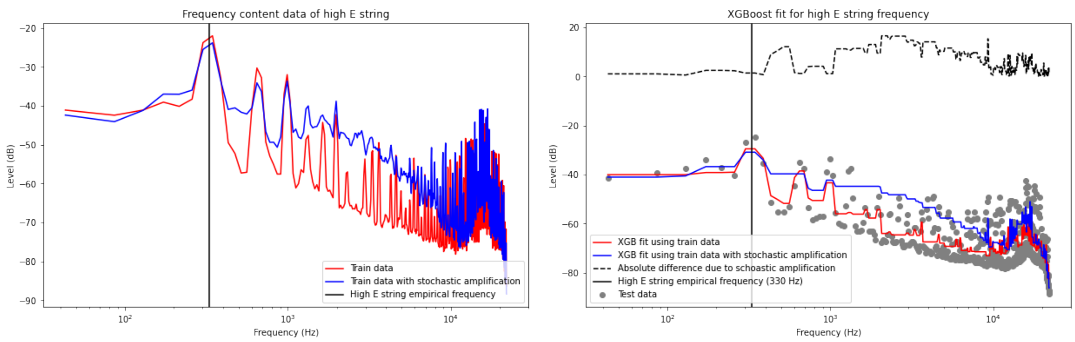

Figure 6: Right plot shows frequency-domain data of stochastically unamplified and amplified with the corresponding empirical frequency for high E string. XGBoost regression fit and the resulting absolute effects of stochastic amplification were shown on the right plot for high E string.

For the high E string, both stochastically amplified and unamplified resonant frequencies are at 301 Hz which is 29 Hz below the empirical frequency corresponding to a 8.79% deviation. Here, the close-fitting of stochastic amplification appears to be qualitatively inconclusive.

4 Conclusion

The paper aims to explore the effects of stochastic amplification to the regression accuracy of XGBoost regression model without hyperparameter tuning as a possible consequence of an increase in signal-to-noise ratio. By analyzing frequency-domain data of all six guitar string-generated audio, the results show evidence of a reduction in the effects of stochastic amplification near resonant frequencies. This may be used in peak-detection applications by finding points of minimum deviations in amplified and unamplified setup. Moreover, results for higher frequency strings such as D, G, and B strings show qualitative evidence of an increase in XGBoost model regression accuracy on the generated test data near resonance.

References

[1] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, Stochastic resonance, *Reviews of modern physics* **70**, 223 (1998).

5 Codes and Raw Data

All files regarding the experiment can be found here: `https://github.com/schwarzschlyle/electronics-and-instrumentation/tree/master/frequency-response`

6 Reflections

6.1 Technical correctness

Although the frequency decomposition of a time-series data was accomplished, the paper has not managed to tackle frequency response detection of a specific material's natural frequency by recording the output response from an input chirp signal. However, the application of DFT to a time-series data of an audio signal was properly carried out resulting in resonance peaks with reasonable proximity to the actual empirical resonance frequencies of the individual strings. All results were written and displayed on the paper as line and scatter plots.
Self-score: 25/35

6.2 Presentation quality

The plots were made to be as visually clear as possible using suggestive color combinations such as blue and red for amplified and unamplified data, respectively, gray for the background test data scatter plot, and black theme for shared metrics such as absolute difference effects of amplification. Plots were labeled properly with captions that can stand alone.
Self-score: 35/35

6.3 Self reflection

The validity of results was properly discussed in each guitar string results and discussion. Attribution of the mitigation in the effects of stochastic amplification near resonance and in the high-frequency domain may have been attributed to the limitations of the frequency range of the generated white noise background. Some of the results, even qualitative, have inconclusive implications. Resources on stochastic resonance were cited on the introduction part of the paper.
Self-score: 30/30

6.4 Initiative

Even though the frequency response recording objective from the chirp signal was not tackled, the study aims to apply the lessons learned from frequency-domain analysis of audio signals to a possible exploration of stochastic amplification effects outside the scope of the course. Moreover, this was done on a machine learning modelling via XGBoost that is also outside the scope of the course.
Self-score: 20/20

Total self-score: 110