Synthesizing Ground Motion Using Bayesian Optimization in Variational Auto-encoder

Kaiyuan HSU[1], Di LIN[2], Riki HONDA[3]

[1] Dept. of International Studies, GSFS, The University of Tokyo.
[2] Dept. of Civil Eng., The University of Tokyo.
[3] Member of JSCE, Professor, Dept. of International Studies, GSFS, The University of Tokyo.

Background

- When designing structures, it is necessary to use ground motions (GMs) that accurately reflect the potential impact of seismic events on structures.
- Conventional methods fitting ground motions to a specific response spectrum to fit their expectation of simulation analysis.
- However, even if GMs match the design spectrum, they could still be different to trigger their nonlinear response, which can influence the reliability of structural design.

Synthesis of Design GM Using Variational Auto-encoder

- The method has been proposed (Lin, 2024) to synthesize GMs that not only matched the design spectrum but also exhibited strong nonlinear intensity.
- To achieve this, they utilized a Variational Auto-encoder as a method:

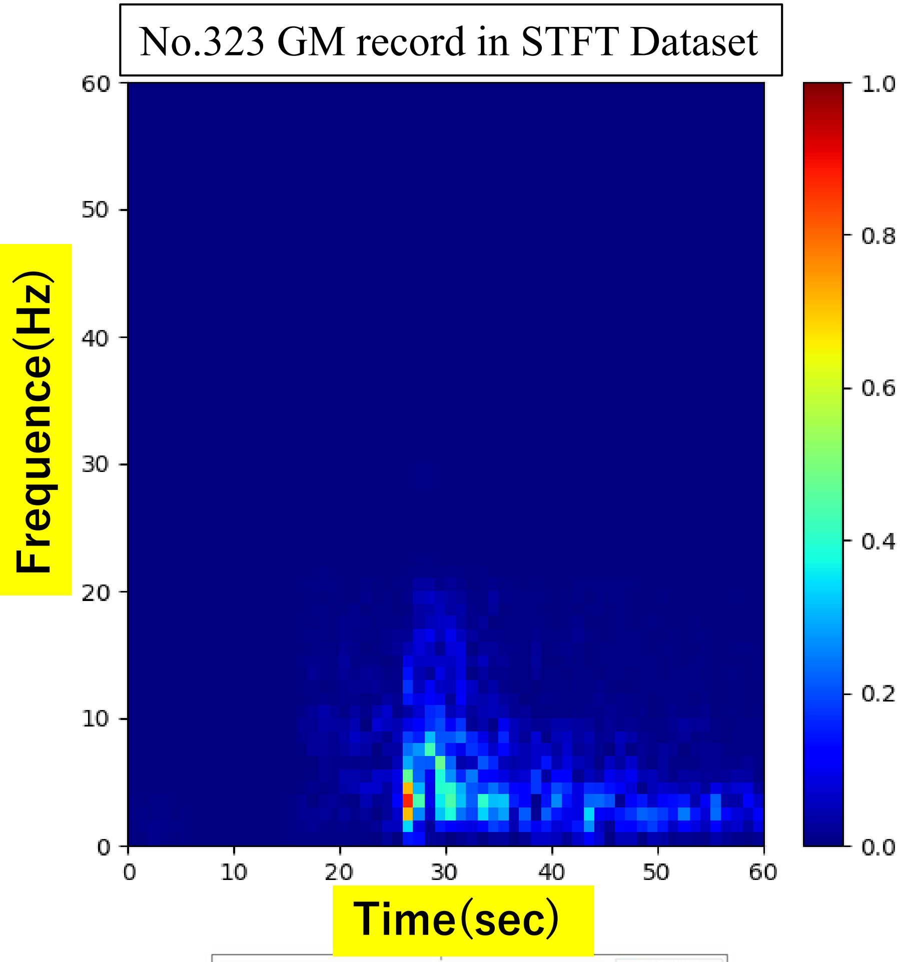
A) Data: STFT (Short Time Fourier Transform), not time history, of GMs are used as the training data, since it gives stable results.

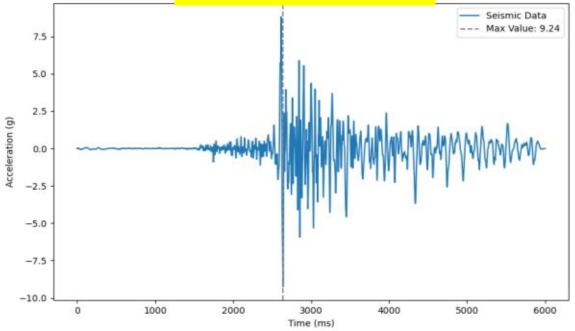
B) Train VAE: Learn the underlying patterns within the data.

C) Feature Extraction: Represent the key features for synthesis GMs. Features do not have physical meanings.

STFT Dataset

- 20,000 GM records from KNET.
- GMs were spectrum-fitted.
- Short-Time Fourier Transform (STFT) was estimated to show Energy Distribution and Frequency Variations in time.





Synthesis of Design GM Using Variational Auto-encoder

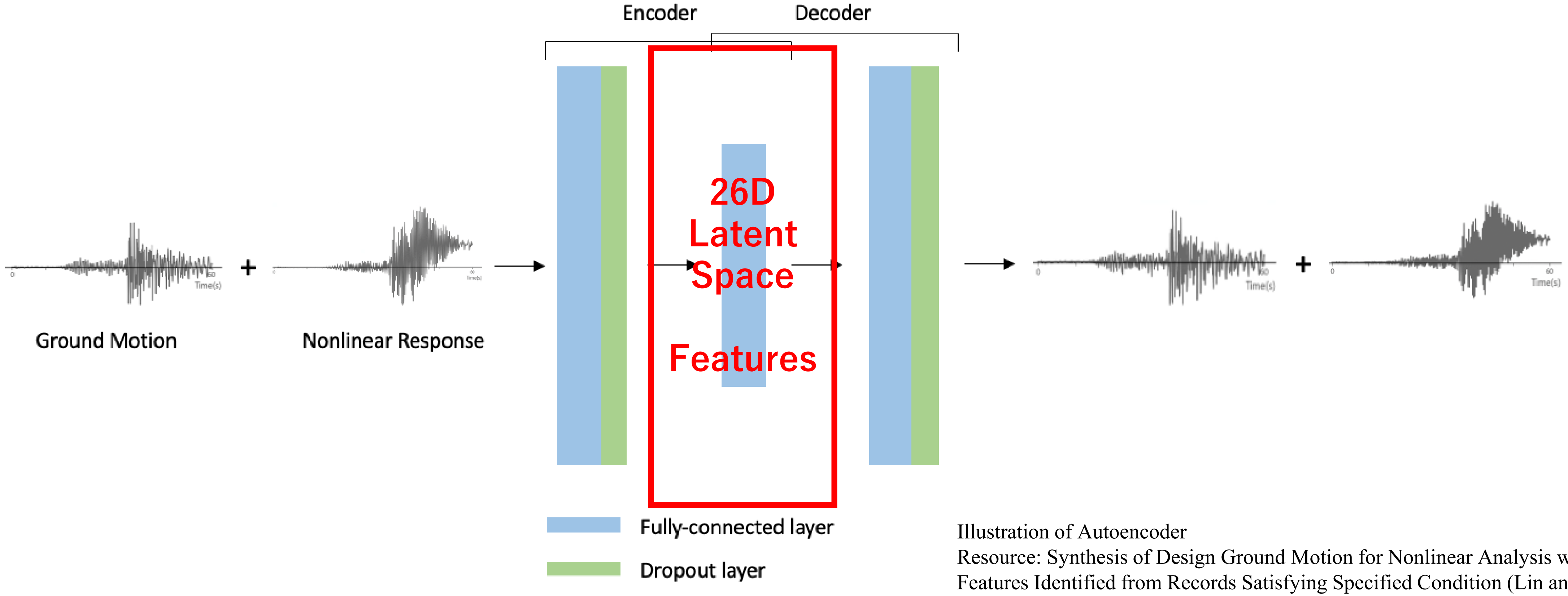


Illustration of Autoencoder
Resource: Synthesis of Design Ground Motion for Nonlinear Analysis with Features Identified from Records Satisfying Specified Condition (Lin and Honda, 2022).

Objective of Simulation

- Search the optimal GM in the set of GMs compatible with the design spectrum.
- In the latent space spanned by the features identified by Encoder, gradient method is not applicable, because gradient is not defined.
- Lin update the GM by finding the optimal increment among the candidate assumed by ad-hoc approach.
- →The optimal solution is not assured.

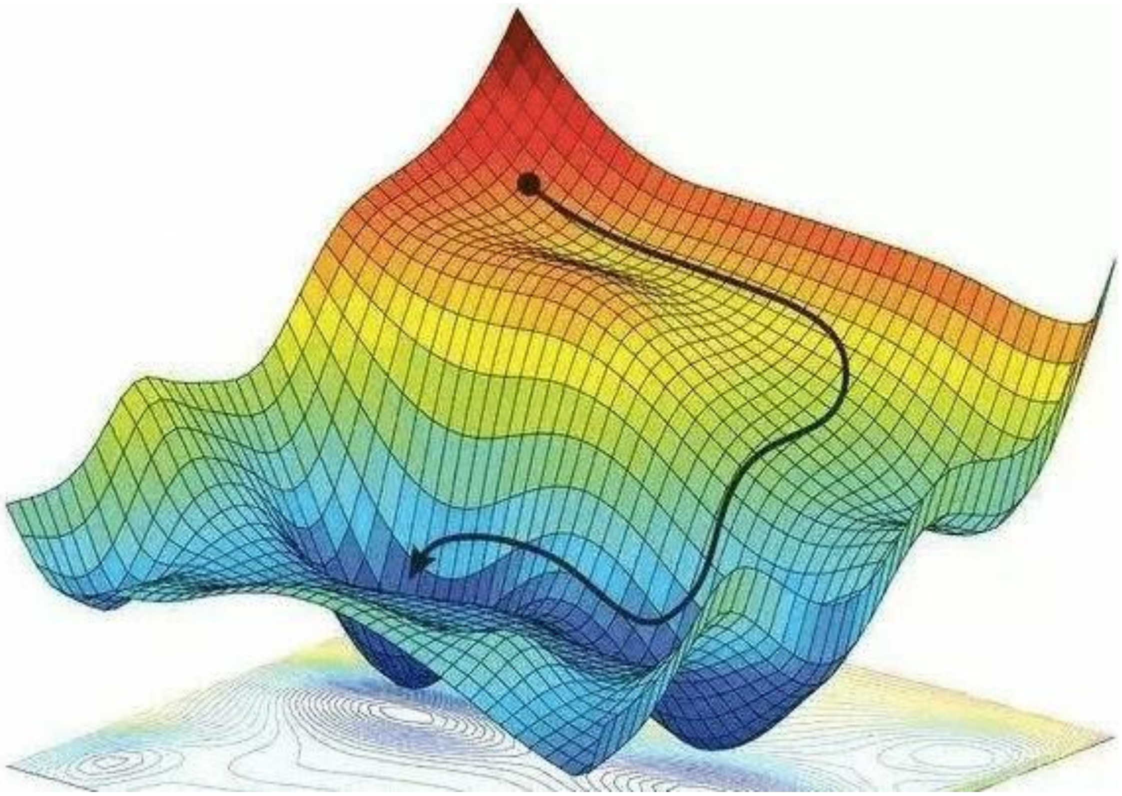


Illustration of Gradient Method.
(Resource: https://easyai.tech/ai-definition/gradient-descent/)

Research Objective

- To develop a method to synthesize ground motions efficiency by integrating Variational Auto-encoder with Bayesian optimization.

  - To improve the performance of a method based on Variational Autoencoder, which had been proposed by Lin et al., gradient to update the GM is obtained by Bayesian Optimization.

Bayesian Optimization (BO)

- BO can be used effectively on the VAE in GM synthesis to maximize the specific features by the Exploitation-Exploration Trade-off.
- Exploitation: Focus on assessing already identified good results to optimize.
- Exploration: Focus on assessing new, unstudied areas with potential for good results.

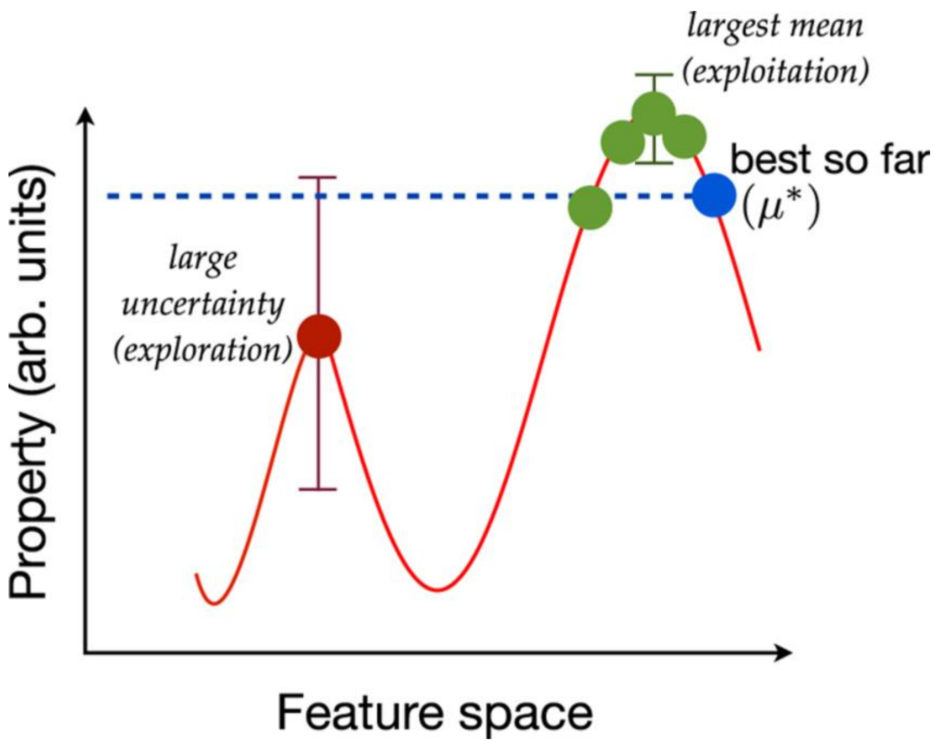


Illustration of Exploration and Exploitation.
Adaptive Strategies for Materials Design using Uncertainties(Balachandran et al, 2016).

Bayesian Optimization Procedure

The function that we want to optimize

Define the Objective Function

The probability distribution over the possible values of the objective function.

Define Surrogate Model

Determine where to evaluate next.

Choose the Acquisition Function

Evaluate the Objective Function

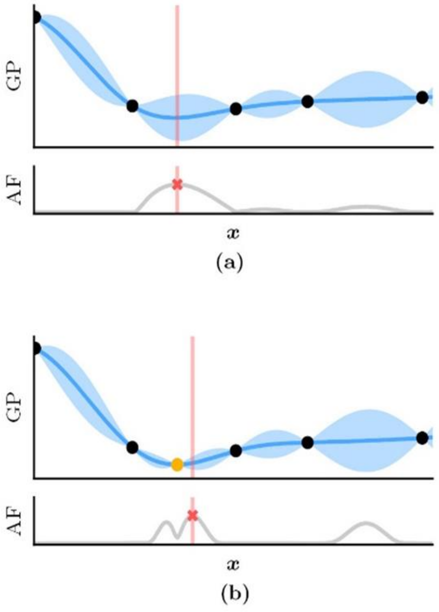
Provide new data to reduce the uncertainty range of the objective function.

Update the Posterior

Update the Posterior distribution of objective function with new data.

Iteration until meets the requirement

Ex. without renewing optimized value



This is the plot of Gaussian Process and Acquisition in an iteration (Öcal et al., 2019).

Tree-structured Parzen Estimator(TPE)

- A method is used in the BO.
- More computational efficiency and adaptability to high-dimensional data.
- Use of two different probabilistic models to estimate the probability distribution.
- For a given hyperparameter x, the model chooses the probability density function $p(x|y)$ to be categorized into **high-performance** regional model $l(x)$ and **low-performance** regional model $g(x)$ according to the sampling value of the objective function $y(x)$.

Numerical Simulations

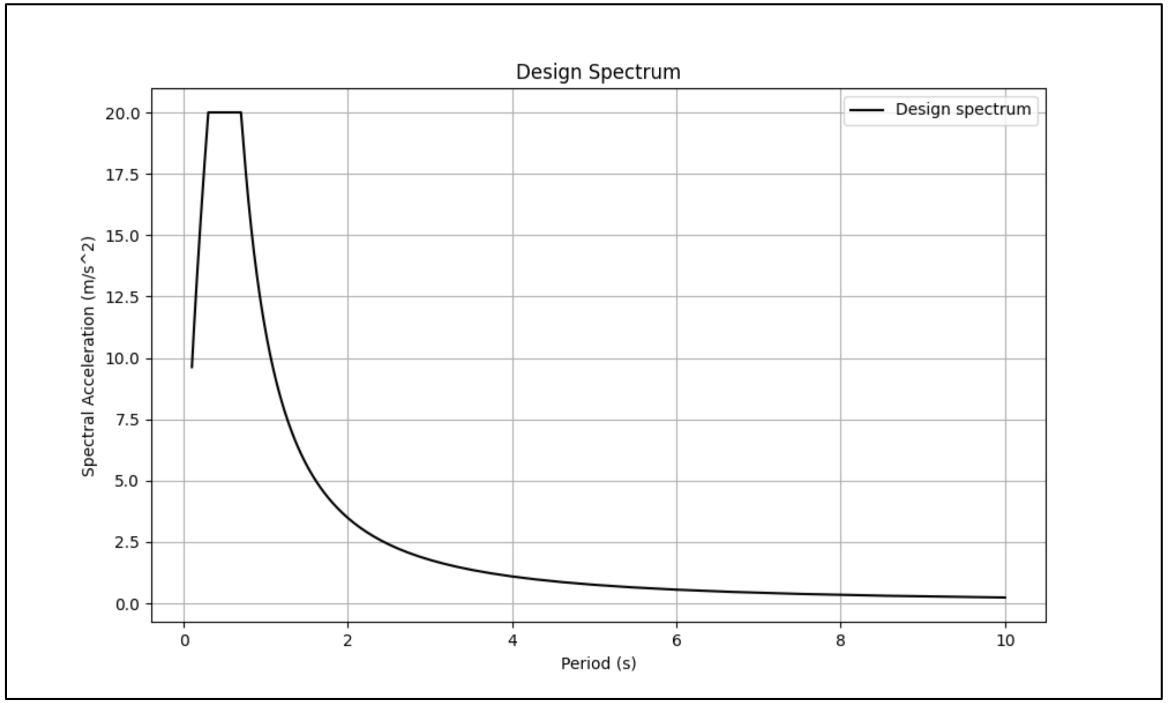
Optimize the maximum displacement of an SDOF system within a 95% confidence interval of Peak Ground Acceleration.

Under the same condition, the only difference is:

- Case 1: Search in One Shot in the Upper and Lower Bound of the 95% confidence interval.
- Case 2: Gradually Search in the Upper and Lower Bound of the 95% confidence interval.

SDOF System Properties

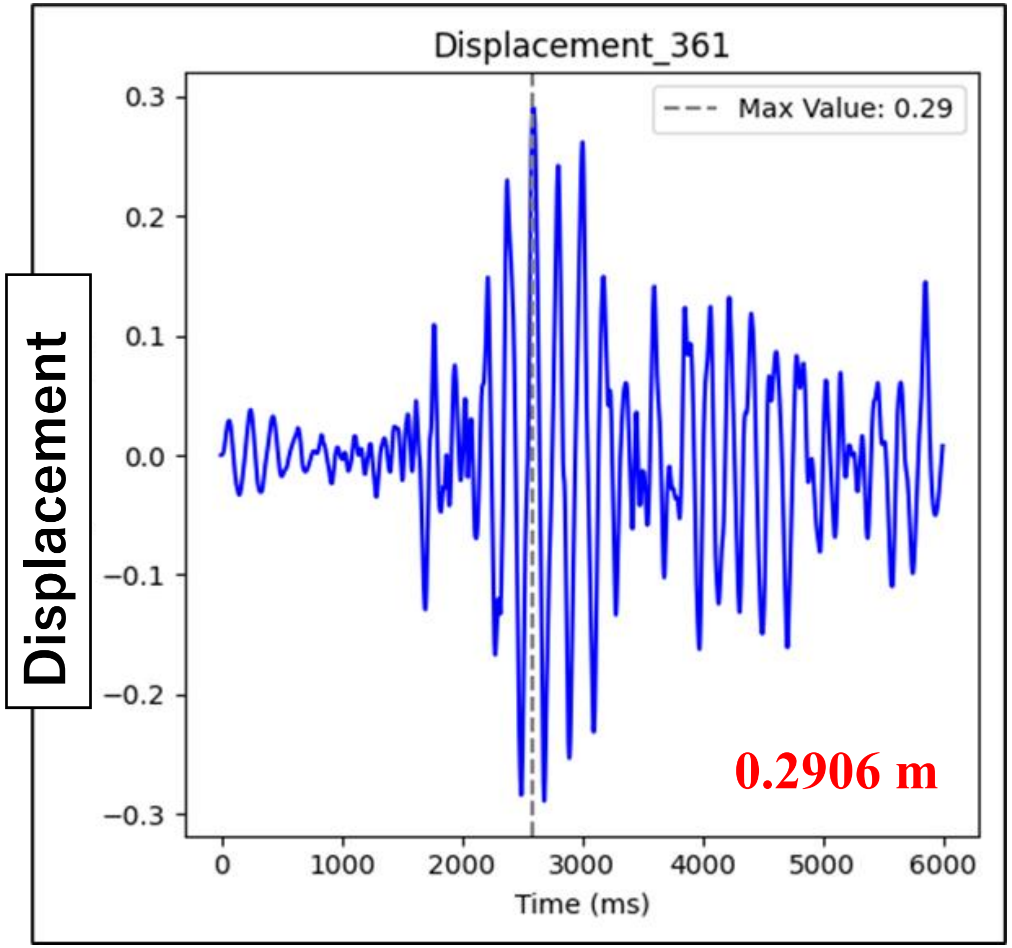
- ➢Target SDOF Structure
  - ➢Natural Frequency: 0.5Hz
  - ➢Damping coefficient: h = 0.05
- ➢Target Design Spectrum

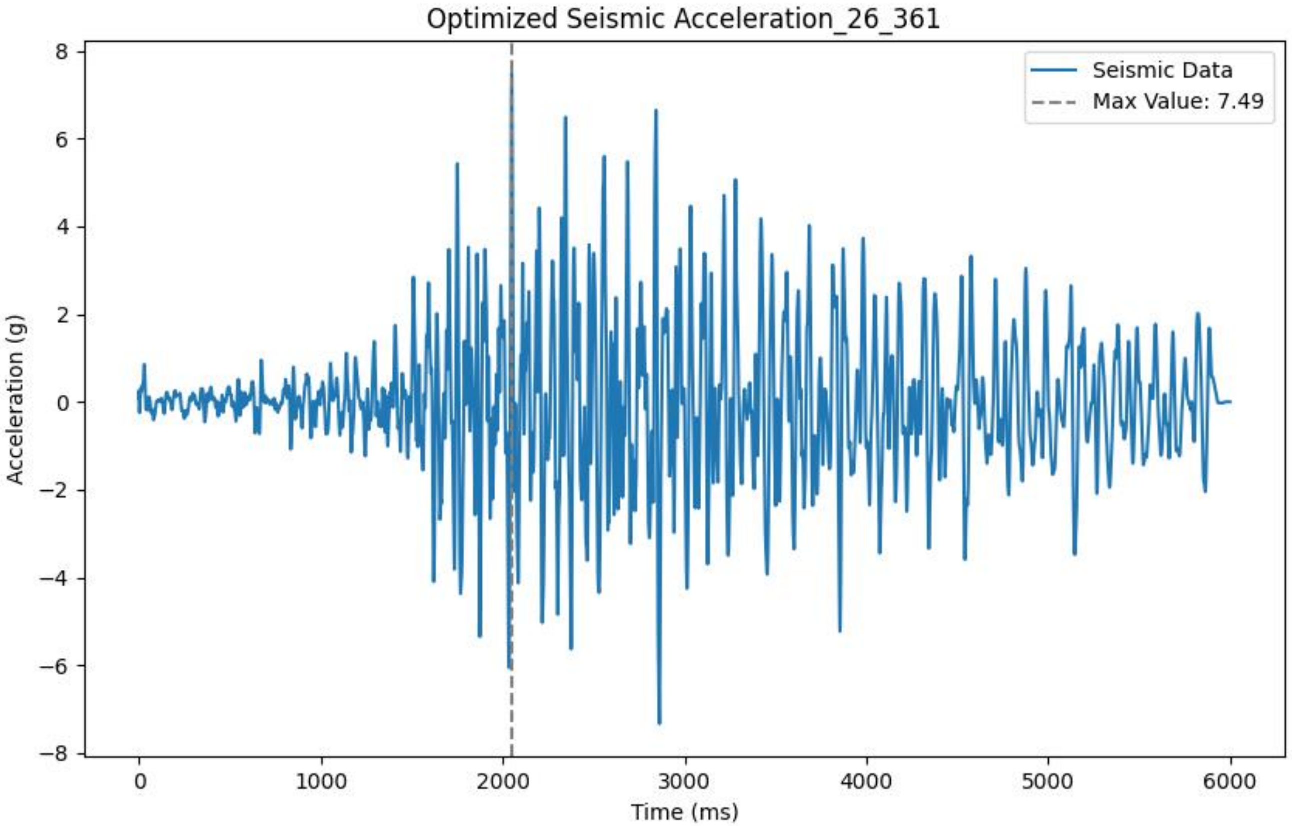


Case 1: Setting

A) Objective Function: Maximum **Displacement** of SDOF.
B) Searching Space: the 95% confidence interval of **26-dimension** latent space.
C) Assumed the Upper Bound is bigger than the Lower Bound for easily apply on the BO without any influence on the result.
D) Early Stopper: if the optimized result does not renew after 200 iterations, the current result is considered the optimized result.

Case 1: Generated GM and time history of the system





Case 2: Setting

A) Objective Function: Maximum **Displacement** of SDOF.
B) Searching Space: the 95% confidence interval of **26-dimension** latent space.
C) Assumed the Upper Bound is bigger than the Lower Bound for easily apply on the BO without any influence on the result.
D) Early Stopper: if the optimized result does not renew in the next 10 times BO, the current result is considered the optimized result.

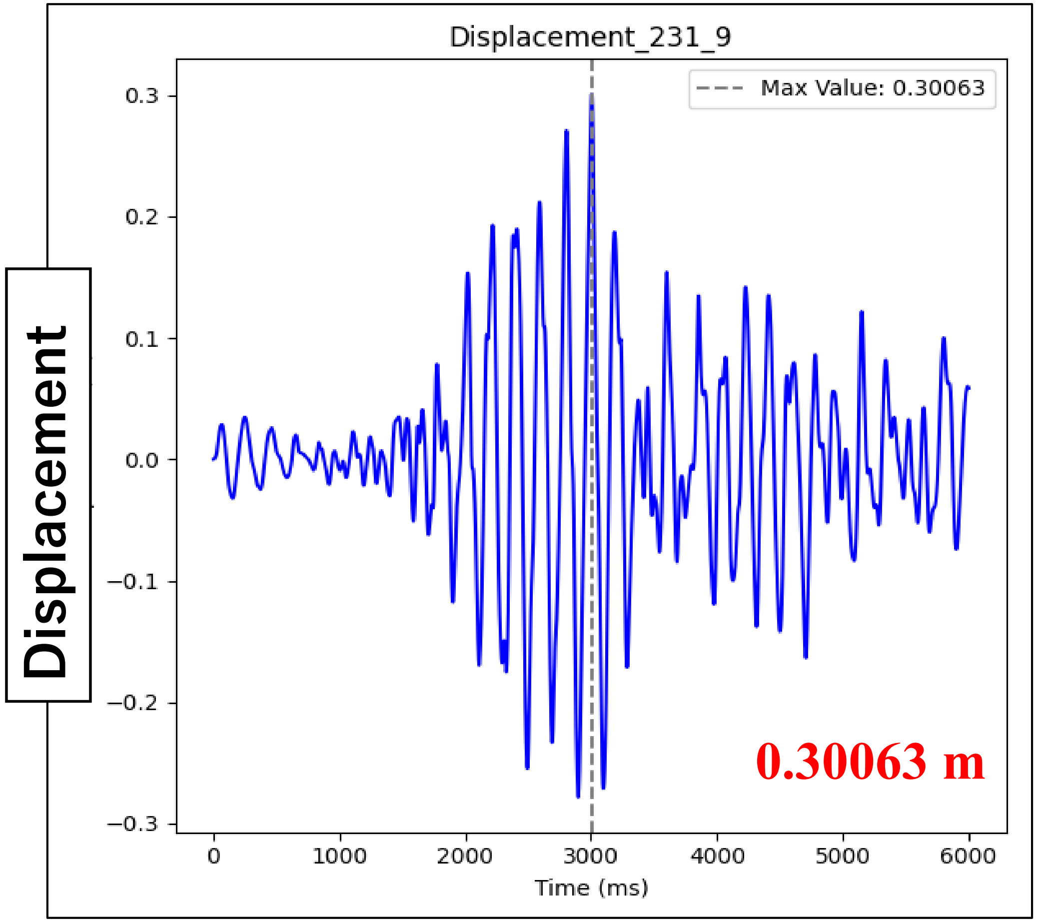
Case 2: Bound Setting

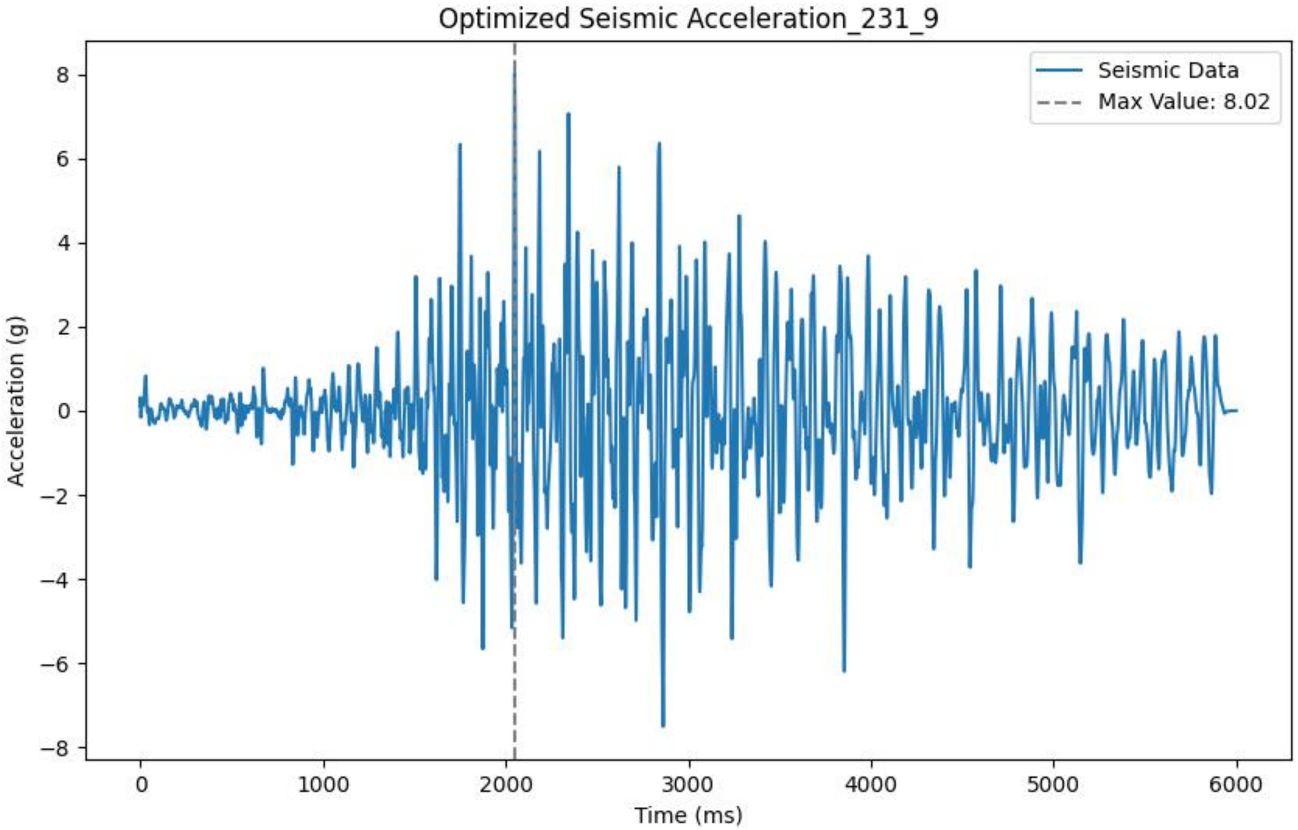
A) Set a parameter $\boldsymbol{\Delta} = \dfrac{(\boldsymbol{Upper\ Bound\ of\ 95\%} - \boldsymbol{Lower\ Bound\ of\ 95\%})}{\mathbf{20}}$

B) Renew: After each time of BO renew the upper bound and lower bound of new iterations with the optimized result +DELTA and –DELTA.

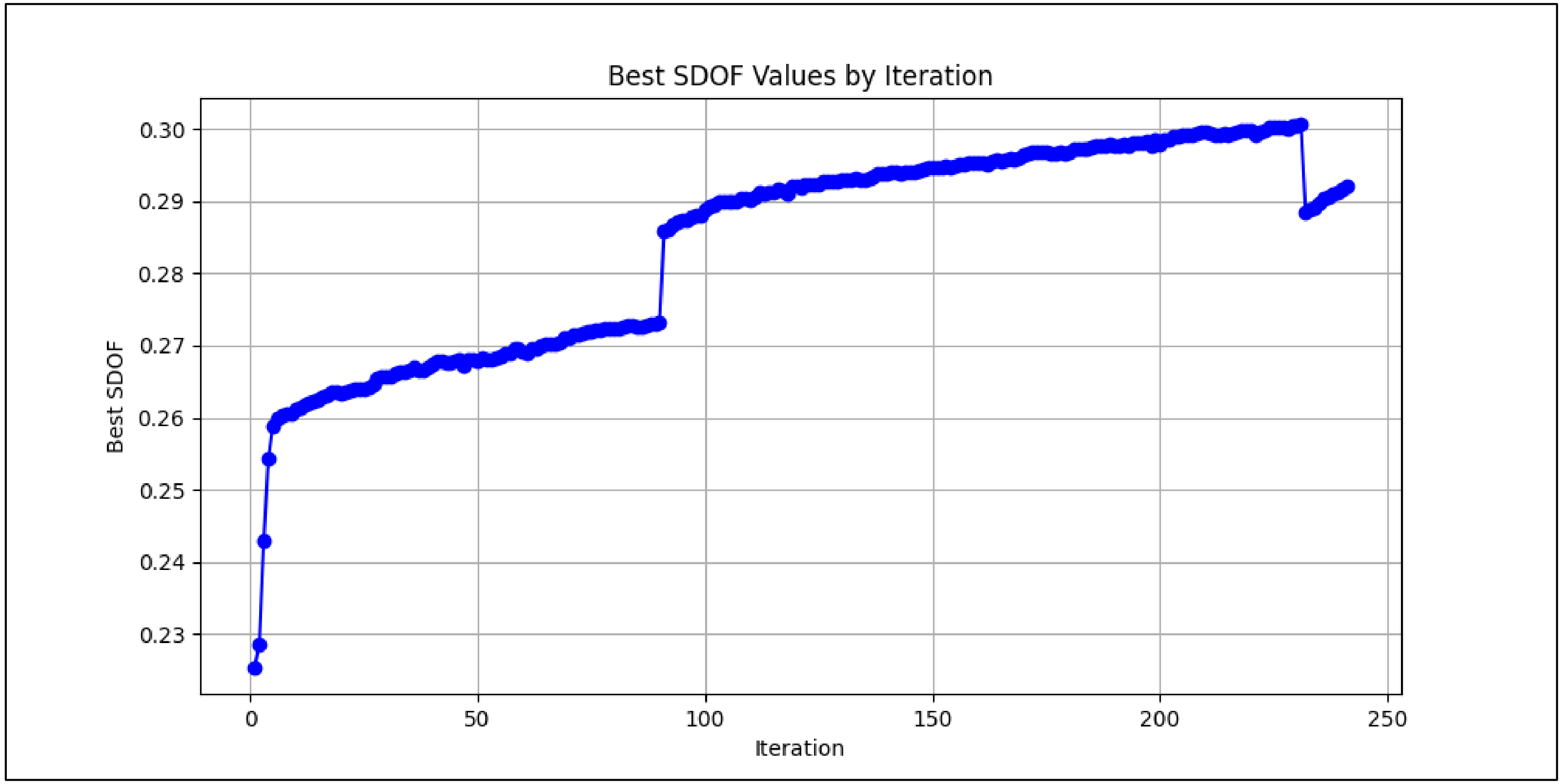
C) Optimized iterations in each BO: 100 iterations.

Case 2: Generated GM and time history of the system





Case 2: Update path of the peak displacement



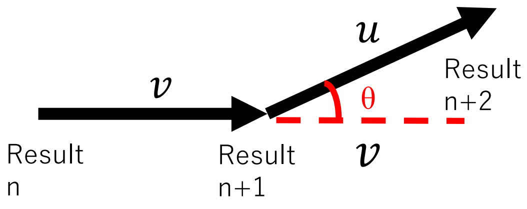
- Result: a total of 241 times of BO.
- The best result happened: the 231 times of BO.
- The optimized result is 0.30063(m).

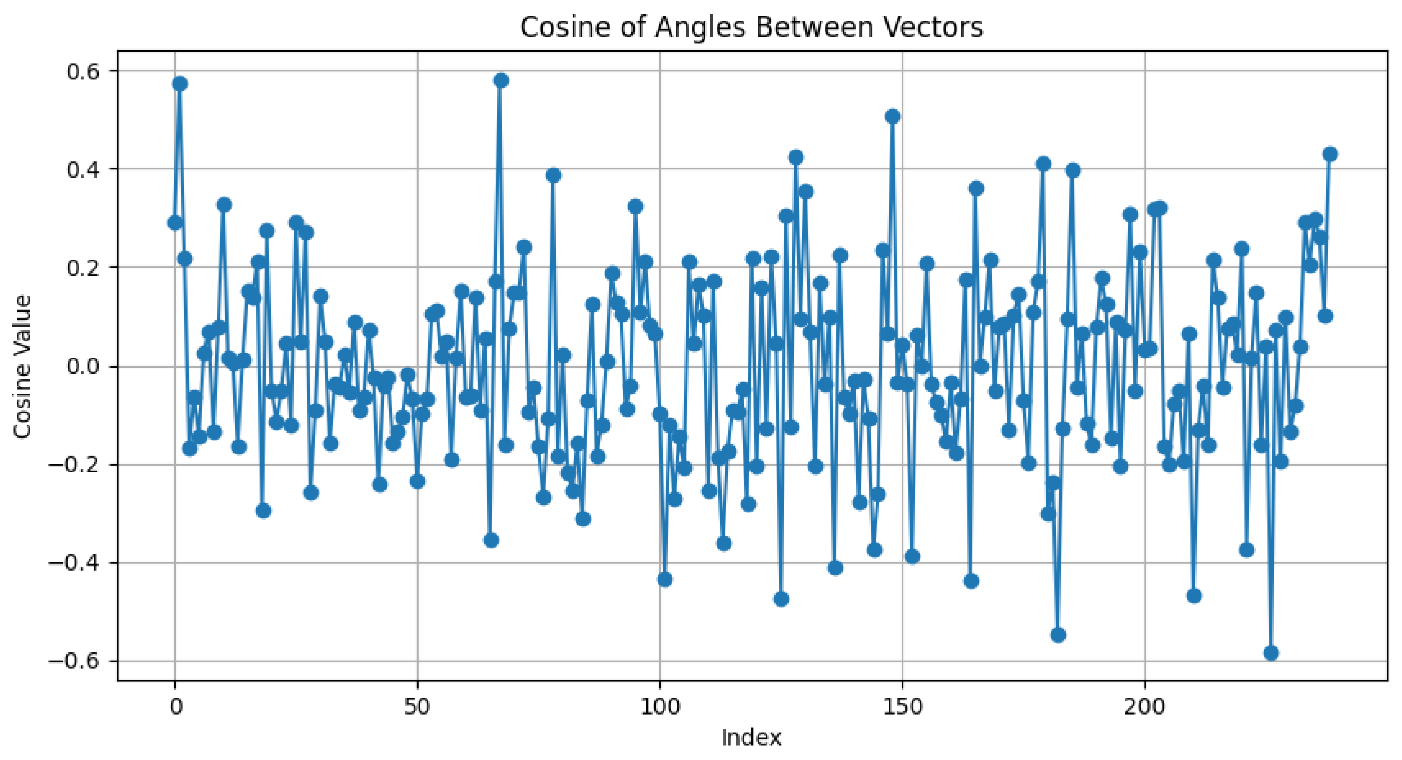
Case 2:
Path of update

To check if the updating path is straight. Since 26-dimension space cannot be plotted, the temporal change of the direction (angle difference).

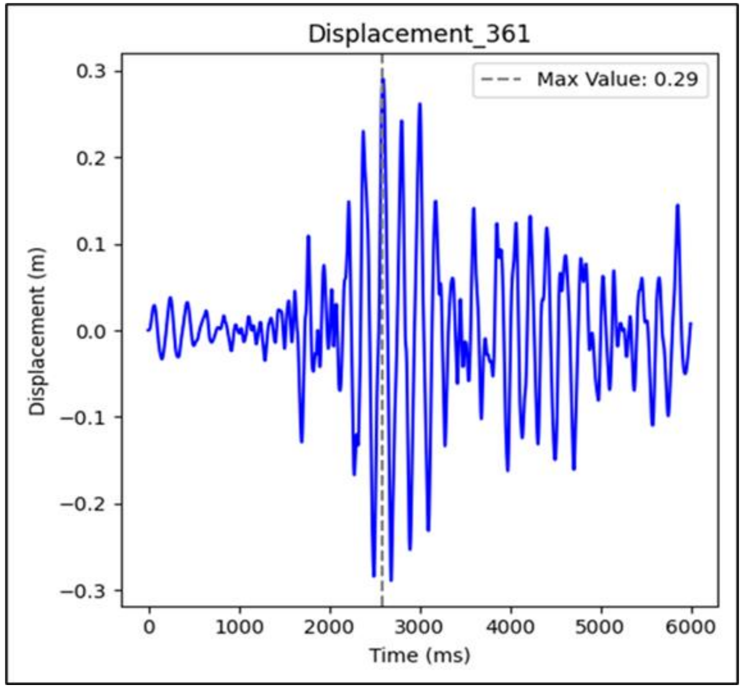
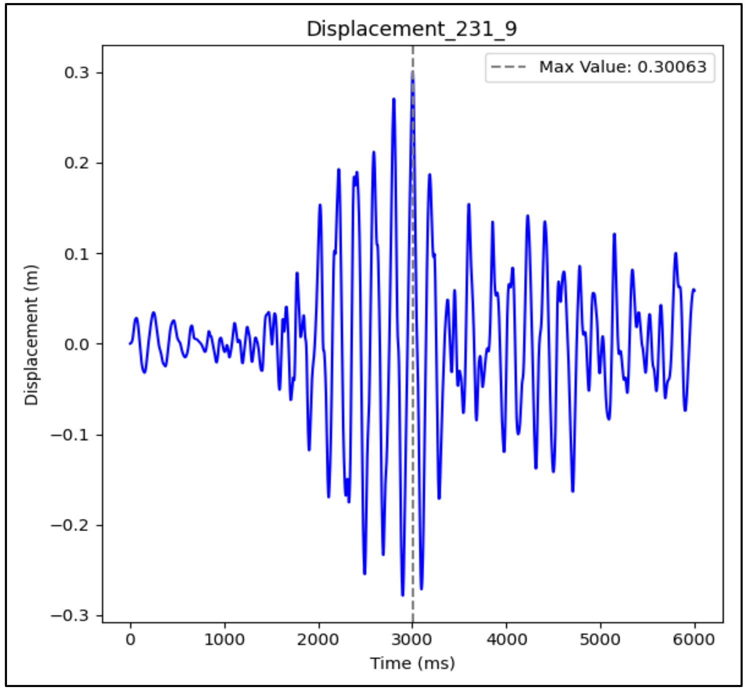
$$cos(\theta) = \frac{u \cdot v}{\| u \| \| v \|}$$

- cos(θ)=1 ➡ Parallel.
- cos(θ)=0 ➡ Orthogonal.
- cos(θ)=−1 ➡ Antiparallel.



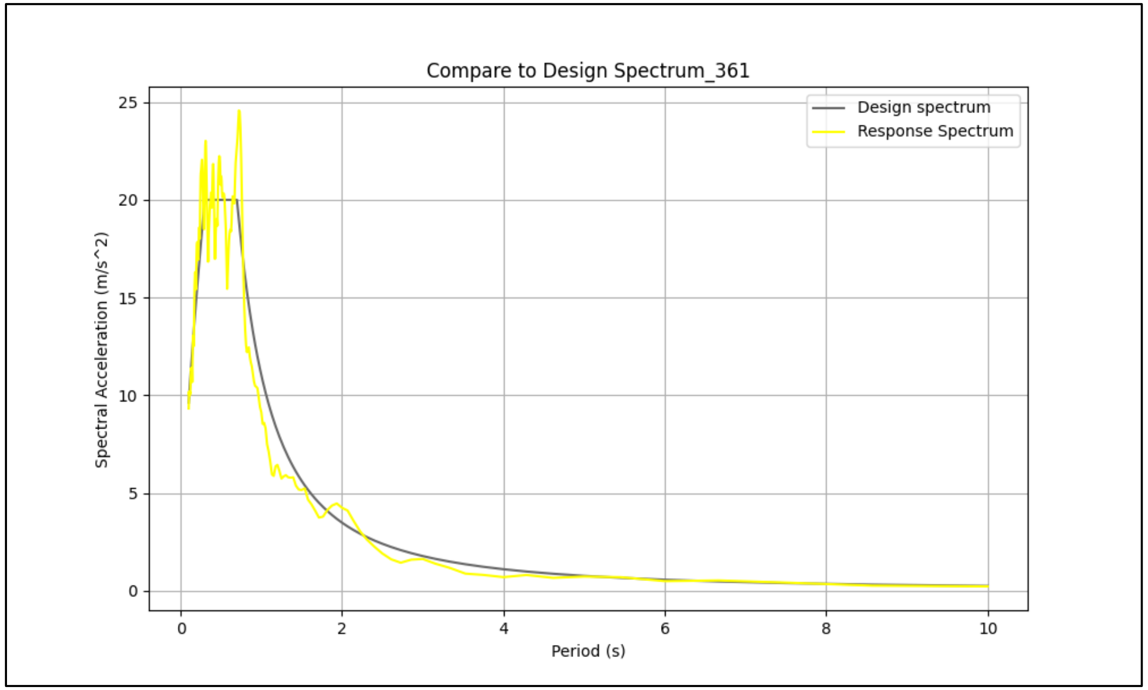


Comparison of Case 1 & Case 2

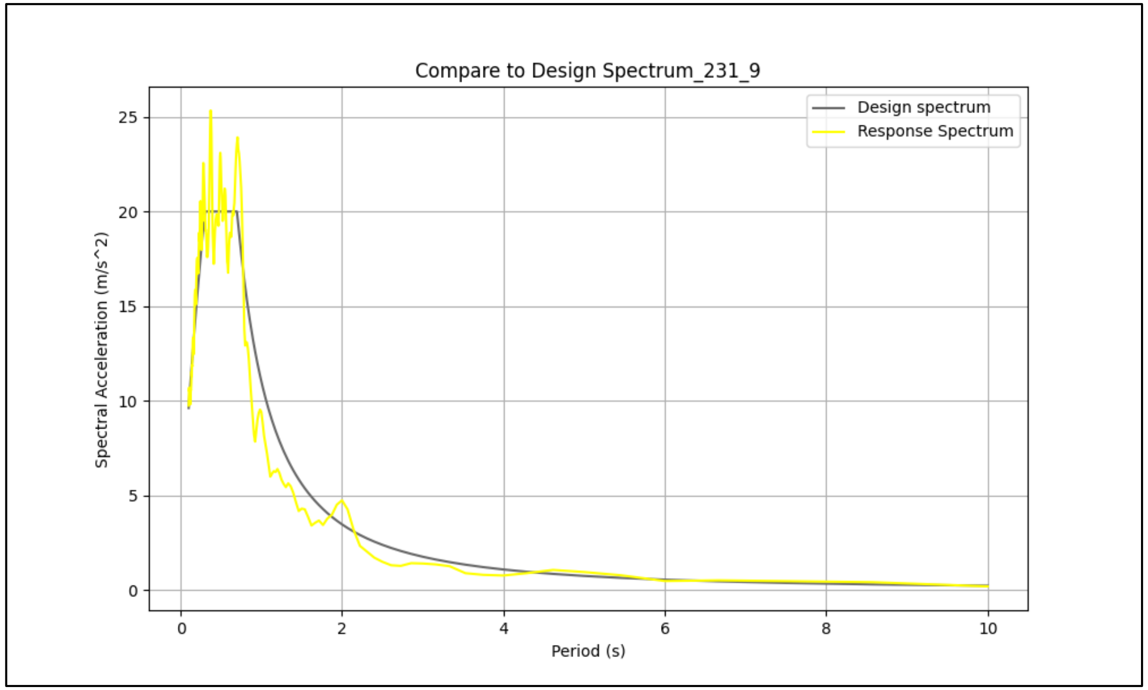
	
Case 1 Total iterations to the best result: 361. The optimized result is 0.2906(m).	Case 2 Total iterations to the best result: 23009. The optimized result is 0.30063(m).

Compare to Design Spectrum

Case 1



Case 2



Summary

- We try to improve the method to synthesize GMs that are not only compatible with the **design spectrum** but also strong in terms of **nonlinear intensity**.
- The original scheme adopts the best among the randomly selected candidates in the updating process. We proposed to use Bayesian Optimization to find the best update of the GM.
- The proposed method generated good GM. The updating path was not straight to the final GM, indicating the importance of **stepwise search**.
- It was also found, however, that one-shot search with Bayesian Optimization can give closely good GM, too.
- Further study is necessary to see if we can improve the performance of the proposed scheme, such as increasing the number of iterations, etc.

Reference

- Balachandran, Prasanna & Xue, Dezhen & Theiler, James & Hogden, John & Lookman, Turab. (2016). Scientific Reports. 6. 19660. 10.1038/srep19660.
- Deierlein, G. G., Reinhorn, A. M., & Willford, M. R. (2010). NEHRP seismic design technical brief, 4, 1-36.
- Di Lin, & Riki Honda. (2022). In Proc. of the 8th International Symposium on Reliability Engineering and Risk Management, 2022, 329-336.
- Frazier, P. I. (2018). A tutorial on Bayesian optimization. arXiv preprint arXiv:1807.02811.
- Guan M. EERI, X., Burton M. EERI, H., & Shokrabadi, M. (2021). A database of seismic designs, nonlinear models, and seismic responses for steel moment-resisting frame buildings. Earthquake Spectra, 37(2), 1199-1222.
- Hu Y, Lam N, Khatiwada P, Menegon SJ, Looi DTW. Site-Specific Response Spectra: Guidelines for Engineering Practice. CivilEng. 2021; 2(3):712-735. https://doi.org/10.3390/civileng2030039
- Kaan Öcal, Ramon Grima, and Guido Sanguinetti. Parameter estimation for biochemical reaction networks using wasserstein distances. Journal of Physics A: Mathematical and Theoretical, 53(3):034002, 2019.
- Satyabrata Choudhury. *Performance-based seismic design of structures*. CRC Press, NewYork, May 2024.
- Watanabe, S. (2023). Tree-structured Parzen estimator: Understanding its algorithm components and their roles for better empirical performance. arXiv preprint arXiv:2304.11127.