A Bayesian Optimization-Assisted, High-Performance Simulator for Modeling RF Accelerator Cavities

Aman Rani
Department of Mathematics and Statistics
Texas Tech University
Lubbock, USA
aman.rani@ttu.edu

Yang Liu
Applied Mathematics and Computational Research Division
Lawrence Berkeley National Labs
Berkeley, USA
liuyangzhuan@lbl.gov

Tianhuan Luo
Accelerator Technology and Applied Physics Division
Lawrence Berkeley National Labs
Berkeley, USA
TLuo@lbl.gov

Hengrui Luo
Applied Mathematics and Computational Research Division
Lawrence Berkeley National Labs
Berkeley, USA
hrluo@lbl.gov

Xiaoye S. Li
Applied Mathematics and Computational Research Division
Lawrence Berkeley National Labs
Berkeley, USA
xsli@lbl.gov

***Abstract*—Accelerator cavity modeling can yield objective functions with multiple sharp minima, which can make their efficient optimization very challenging. We present an efficient ML algorithm for the modeling of of particle accelerator cavities by combining Gaussian Process with Downhill-SIMPLEX methods. The algorithm is tested for a single-cell cavity design to locate all resonant modes quickly and accurately.**

***Index Terms*—Cavity modeling, Integral equation method, Resonance modes, Gaussian process, Down-hill SIMPLEX**

I. Introduction

High-quality radio-frequency (RF) cavities are key components for high energy particle accelerators, synchrotron light sources, dark matter searches, quantum computing, etc. Designing these cavities comes along with the computational challenges such as multi-objective optimization, the requirement for high performance computing (HPC) for handling large-sized cavities etc. To be more precise, its multi-objective optimization requires an efficient 3D full-wave electromagnetic simulator. Traditional simulators for cavity modeling are finite element method (FEM)-based which requires a volumetric discretization and can leads to challenging nonlinear eigen problems [5]. In contrast, we rely on the integral equation (IE) method [4], [7] which leads to linear eigen problem with many fewer number of degrees of freedoms as it requires only surface discretization. IE requires a fast solver with high performance computing and ML algorithms to search for resonance modes.

In this work, we propose an HPC-based fast direct matrix solver for IE, combined with hybrid optimization algorithms to attain an efficient simulator for accelerator cavity modeling. First we solve the linear eigen problem for each trial frequency by a distributed-memory parallel, fast direct solver package called ButterflyPACK [8], which implements several rank-structured hierarchical matrix solvers to achieve reduced matrix factorization complexities. Second, we propose the combination of the global optimizer Gaussian process with the local optimizer Downhill-SIMPLEX methods to generate the trial frequency samples which successfully optimize the corresponding 1D objective function with multiple sharp minima. We test the proposed simulator for single-cell cavity design to quickly and accurately locate all resonant modes.

II. Proposed IE-based formulation for accelerator modeling

The IE method for cavity modeling boils down the problem to find the resonance frequency f such that the smallest eigenvalue of the corresponding system matrix is close to 0:

$$Z[f]I \approx 0. \quad (1)$$

Here Z is the discretized IE operator, whose dimension is approximately the number of degrees of freedom on the surface of the cavity. Z can possibly model the effects of beam ports and damping ports as well. The eigenvector I represents the discretized surface electric current density at resonance.

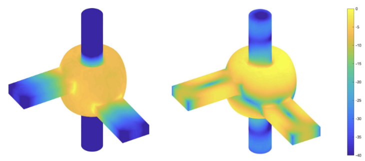

Fig. 1. Mode vector I at $f = 1.5149$ GHz (left) and $f = 2.3329$ GHz (right)

Each trial sample f requires solving a linear eigen problem, and we rely on the proposed optimization strategy in Section III to propose the samples f. For each fixed f, we leverage a distributed-memory eigen solver package P_ARPACK [9], which calls the distributed-memory fast direct matrix solver ButterflyPACK [8]. One typically requires only the lowest 20 - 30 eigenvalues to identify whether one or several resonance modes have been found at a given frequency f.

III. Proposed Optimization Strategy for Searching Resonance Frequencies

To locate the resonance frequencies, we note that the objective function is a 1D function with multiple sharp minima (see Fig. 2 for an example), making their efficient optimization very challenging. Global optimizer like Gaussian process (e.g., GPTune, an optimizer designed for objective functions in exascale computing project (ECP) application codes [1]) has smoothing effects and cannot capture the sharp minima. Such a difficulty has been partially remedied by [10]. On the other hand, local optimizer like Downhill-SIMPLEX can obtain very accurate function minimum, but requires a good initial guess.

We propose combining the Gaussian process (GPTune) with Downhill-SIMPLEX algorithms to leverage the exploration nature of Gaussian process and the exploitation nature of Downhill-SIMPLEX. The proposed algorithm first generates a few frequency samples using GPTune to identify the number of resonance modes within a prescribed frequency range. For each mode, we then leverage Downhill-SIMPLEX with a refined search region and initial guess defined by the GPTune samples, to locate a more accurate resonance frequency. In practice, the number of GPTune samples should dominate over that of the Downhill-SIMPLEX ones to ensure that the correct number of modes can be discovered.

IV. Numerical Results

We tested our algorithm for modeling of an ALS Higher Harmonic Cavity (HHC) with a single cell [6]. The cells can be cascaded to provide continuous beam acceleration. Our IE formulation leads to a 25000×25000 system Z which is much smaller than system sizes of the FEM-based simulator, called Omege3P/ACE3P [5]. In both the proposed IE solver and Omege3P, we set the frequency range to $[1.5, 3]$ GHz to search for the resonance modes. Fig. 1 shows the surface current, i.e., the eigen vectors I from (1) discovered at frequencies $f = 1.5149$ GHz and $f = 2.3329$ GHz.

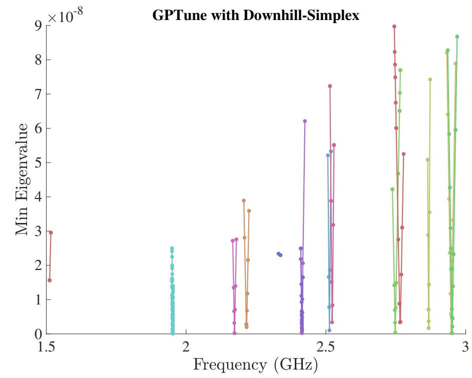

Fig. 2. 15 resonance modes found by GPTune+Downhill-SIMPLEX

Detected Frequencies		
GPTune	GPTune + Simplex	Omege3P
1.5164	1.5141	1.51
1.9516	1.9516	1.953
1.9526	1.9526	1.954
2.1733	2.1733	2.172
2.2166	2.2166	2.216
2.3306	2.3376	2.329
2.4176	2.4149	2.413
2.4176	2.4141	2.414
2.5108	2.5127	2.512
2.529	2.5223	2.522
2.7446	2.7481	2.747
2.7664	2.7657	2.764
2.867	2.8674	2.868
2.9482	2.9482	2.945
2.9482	2.9519	2.948

Fig. 3. Table of first 15 resonance frequencies found by GPTune, GPTune + Down-hill Simplex, and the reference data from a FEM code developed by SLAC called Omege3P

In the proposed simulator (GPTune + Downhill-SIMPLEX), we set budget of 30 GPTune samples followed by 8 Downhill-SIMPLEX samples for each located resonance region. Each eigen solver (i.e., function evaluation during optimization) in (1) uses 16 NERSC Cori Haswell nodes (512 CPU cores) requiring about 1-2 minutes of CPU time. Fig. 2 shows the function evaluation samples. Each dot represents one sample (frequency and corresponding smallest eigenvalue), and each color represents one resonance mode. There are a total of 15 resonance modes discovered by the proposed solver.

As a comparison, we tested the HHC cavity with the Omege3P solver, the discovered resonance frequencies by the IE solver (GPTune samples only), IE solver (GPTune+SIMPLEX), and the FEM solver (Omege3P) are listed in Fig. 3. Note the GPTune+SIMPLEX match well with Omege3P results.

References

[1] Y. Liu, W. Sid-Lakhdar, O. Marques, X. Zhu, C. Meng, J. Demmel, X. Li, "GPTune: Multitask Learning for Autotuning Exascale Applications," PPoPP21, pp. 234–246, Feb 2021.

[2] Y. Cho, J. W. Demmel, X. S. Li, Y. Liu, and H. Luo, "Enhancing Autotuning Capability with a History Database," IEEE MCSoC, Feb 2022.

[3] H. Luo, J.W. Demmel, Y. Cho, X. S. Li, and Y. Liu, "Non-smooth Bayesian optimization in tuning problems," arxiv.

[4] Q. I. Dai, J. Wu, H. Gan, Q. S. Liu, W. C. Chew and W. E. I. Sha, "Large-Scale Characteristic Mode Analysis with Fast Multipole Algorithms," IEEE TAP, pp. 2608–2616, Feb 2016.

[5] R. Van Beeumen, O. Marques, E.G. Ng, C. Yang, Z. Bai, L. Ge, O. Kononenko, Z. Li, C.-K. Ng, and L. Xiao, "Computing resonant modes of accelerator cavities by solving nonlinear eigenvalue problems via rational approximation," J. of Comp. Physics, pp. 1031–1043, Dec 2018.

[6] Feng H, Luo T, De Santis S, Li D, Baptiste K, Huang W, Tang C, Liu Z. Design study on higher harmonic cavity for ALS-U. InProceedings of the 10th International Particle Accelerator Conference (IPAC2019) 2019 May (pp. 3011-3013).

[7] Y. Liu, T. Luo, and S. Tan, A Fast Direct IE Solver for Characterization of Accelerator Cavities with Waveguide Ports, in ICCEM, 2020.

[8] Liu, Yang, and USDOE. ButterflyPACK. Computer software. https://www.osti.gov//servlets/purl/1564244. USDOE. 28 Nov. 2018. Web. doi:10.11578/dc.20181213.1.

[9] Maschhoff, Kristyn J., and Danny C. Sorensen. "P_ARPACK: An efficient portable large scale eigenvalue package for distributed memory parallel architectures." International workshop on applied parallel computing. Springer, Berlin, Heidelberg, 1996.

[10] Eriksson, David, Michael Pearce, Jacob Gardner, Ryan D. Turner, and Matthias Poloczek. "Scalable global optimization via local bayesian optimization." Advances in neural information processing systems 32 (2019).