Optimization of the Booster Tuner Configuration

Edward Hazelton
Fermi National Accelerator Laboratory

12 August 2022

1 Background

1.1 Booster Ring

The booster ring is a 1500-foot circumference synchrotron ring that exists as one of four main parts of the accelerator complex. It receives beam at 400 MeV from the LINAC, and accelerates it up to an exit energy of 8 GeV. It is the first synchrotron in the sequence of accelerators and directly provides beam to the MicroBooNE, Muon g-2 and Mu2e experiments. It also provides beam to the main accelerator ring.

1.2 Radio-frequency Cavity

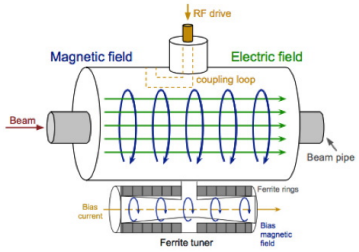

Figure 1: RF cavity with ferrite tuner diagram

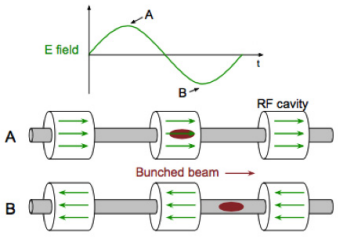

Figure 2: Electric field oscillation synchronized with beam position

Radio-frequency (RF) cavities are the primary accelerators of particles in the booster. They accelerate particles by way of an oscillating electromagnetic field. The particles enter the cavity as the waveform is in its accelerating phase, and are between cavities when the it is in its decelerating phase. To maintain this synchronized relationship, the resonant frequency of the cavities must increase as the particles accelerate. This cycle will happen at 20 Hz in the PIP-II era,

which is too fast for a mechanical system. Therefore, the RF cavities use an electrical tuning system. Each RF cavity has three tuners, each surrounded by pairs of ferrite discs and thin aluminum slugs that act as a dielectric. By themselves, the tuners massively hamper the resonant frequency of the cavity. But, by running a current through the tuners, known as a bias current, we can alter the magnetic properties of the ferrite, reducing its relative permeability, and thus increasing the resonant frequency of the cavity-tuner system.

2 Motivation

The primary motivation for this project is the PIP-II upgrade. Currently, the energy of particles that enter the booster from the LINAC is about 400 MeV. With the PIP-II linear accelerator, that will increase up to 800 MeV. Additionally, this increase in initial injection energy increases the booster ring cycle frequency from 15 Hz to 20 Hz. If we keep the current tuner configuration on the cavities in the booster, this increase in initial injection energy will lead to a direct increase in the root mean squared (RMS) bias current. This is due to the particles starting their cycles through the booster ring at a faster speed, and thus requiring higher average frequencies in the RF cavities. Higher currents in the cavity tuners leads to more power dissipation. The existing tuner configuration was designed and optimized for ranges and timing that includes lower frequencies than would be present in the PIP-II era. Currently, RF cavities must handle a frequency range of about 37.8 MHz to 52.8 MHz. With the PIP-II upgrade, we can reduce the lower bound of that range to 44.7 MHz, a total reduction of 46%. The current tuner configuration is more lossy at these higher frequencies, so finding a better configuration would save on power.

3 Methods

3.1 Tuner Parameters

There are many properties we can modify with the RF cavities, but the primary factors taken into consideration are the tuner length, material properties of the ferrite disks, and the number of disks on each tuner. Each of the other properties can additionally be targeted for optimization, though their usefulness would be limited as that would be a much more involved modification of the cavities and tuners. The properties selected would be much quicker and cheaper to implement than the others.

3.2 Objective Functions

3.2.1 Q Factor

The Q factor, or quality factor, is a metric for resonators, and it is essentially the ratio between the energy stored and the power dissipation rate. A higher Q

	Parameter	Current Value	
Cavity	Length	2.02	m
	Taper offset	0.3114	m
	Taper length	0.741	m
	Min inner conductor radius	0.0286	m
	Max inner conductor radius	0.1165	m
	Outer conductor radius	0.13	m
	Gap capacitance	2.4	pF
Coupling Stem	Length	0.116	m
	Inner conductor radius	0.0265	m
	Outer conductor radius	0.047	m
Tuner	Length	0.9285	m
	Offset	0	m
	Min inner conductor radius	0.0337	m
	Max inner conductor radius	0.0582	m
	Outer conductor radius	0.1	m
	Taper length	0.45	m
Ferrite Disc	Inner radius	0.0582	m
	Outer radius	0.1	m
	Width	0.0266	m
	Num Stackpole	9	
	Num Toshiba	5	
	Total	14	
	Num Tuners	3	
Aluminum Slug	Width	0.00241	m

Figure 3: Simulation Data with selected variables highlighted in green

factor indicates that a resonator has a lower energy loss rate, which means less wasted power. We can calculate the Q factor of the cavity using its impedance at its resonance frequency, which is all dependent on the supplied bias current. The impedance of a circuit is the complex extension of its resistance. The complex part is known as reactance, which represents the energy storage of the circuit. Due to the complicated nature of the RF cavities' design, calculating the impedance is difficult, and thus so is calculating the associated Q factor.

3.2.2 Root Mean Square Bias Current

The RMS bias current is a metric for the average current over the 20 Hz cycle the booster operates on. By minimizing this value, we lower the total power draw of the RF cavity. This directly lowers the power dissipation of the circuit. To approximate the RMS bias current, we look at the bias current ramp over the 20 Hz cycle. Taking the average of the trough and peak of the ramp will give us a good approximation. We can find the trough and peak easily as they coincide with the lowest and highest resonant frequencies, which are 44.7 and 52.8 MHz respectively. While not accurate to the true RMS bias current, this approximation mostly preserves the relative order of the configurations, allowing the optimization technique to still select the best results.

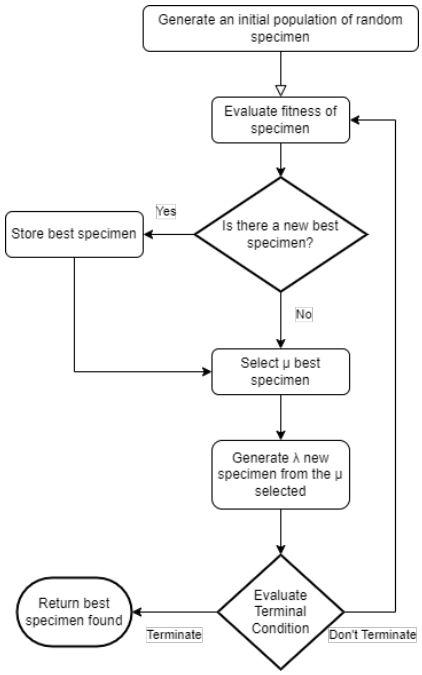

Figure 4: Evolutionary Strategy Flowchart

3.3 Evolutionary Algorithm

3.3.1 Overview

Evolutionary algorithms have a lot of traits we are looking for with an optimization problem such as this. First and foremost, they work well on non-convex and non-concave objective functions due to their stochastic nature.

Convex and concave functions are functions with the property that any line segment between two points does not intersect with the function. For convex functions, the curve lies below the line segment. For concave functions, the curve lies above. This guarantees a global minimum or maximum. It also guarantees that following the gradient will reach the optimum.

Stochastic methods are useful because of how they work with complicated functions. By exploring the problem space with some random variance, we gain the ability to leave local optima by selecting worse-off solutions. These worse solutions are actually ideal because they have a chance to explore areas of the problem space that traditional greedy optimization methods would not be able to reach.

Evolutionary algorithms also do not rely on derivative information, which would be time-consuming to calculate. The calculation for the actual derivative of the Q factor is unavailable to us currently, but we could approximate it using the slope for nearby points. The issue is that we need to calculate the slope of

nearby points in the search space on each axis. As is, we have 6 dimensions to evaluate which is already significant, but it is also easy for more to be added later in future work.

3.3.2 Evolutionary Strategy

There are many different evolutionary algorithms available, but an evolutionary strategy was an ideal approach for this problem. By employing an evolutionary strategy, we can quickly converge on a local optimum of the problem space. Evolutionary strategies are the standard evolutionary algorithm for continuous search spaces, which the tuner configuration is logically modeled as. Evolutionary strategies work very simply on an abstract level. You start with a population of possible solutions. Then, we evaluate their fitness using the objective functions. Next, select the μ best specimen and generate λ children from them. We evenly generate λ/μ children per parent. Each child is a mutation of one of the μ specimen. We then take these λ children and make them the next parent population. We continue the process until some terminal condition is reached. Usually, it is a number of generations, but it can also be based on the rate of improvement or if the best value found has passed a threshold. Provided with good enough hyperparameters, we can almost guarantee the probability of terminating with the global optimum.

3.3.3 Mutation

An uncorrelated mutation with n step sizes, where n denotes the number of dimensions in the data, was used. This is to allow differing movement across the problem space in different dimensions as the space can have different slopes along any given dimension. To do this, the specimen $< x_1, x_2, \ldots, x_n >$ is extended to include the n step sizes as additional dimensions, resulting in $< x_1, x_2, \ldots, x_n, \sigma_1, \sigma_2, \ldots, \sigma_n >$. Then, the mutation is calculated as

$$\sigma_i' = \sigma_i e^{\tau' N(0,1) + \tau N_i(0,1)}, \tag{1}$$

$$x_i' = x_i + \sigma_i N_i(0,1). \tag{2}$$

Where $\tau' \propto 1/\sqrt{2n}$ and $\tau \propto 1/\sqrt{2\sqrt{n}}$, and $N(0,1)$ represents one standard normal distribution draw shared across all coordinates, and $N_i(0,1)$ representing a standard normal distribution draw per-coordinate. To prevent standard deviations too close to 0, we apply a boundary rule

$$\sigma_i' < \epsilon_0 \implies \sigma_i' = \epsilon_0.$$

3.4 Calculations

The calculation of the impedance of a specific tuner configuration is an iterative process. We split the geometry into discrete slices, and calculate the total

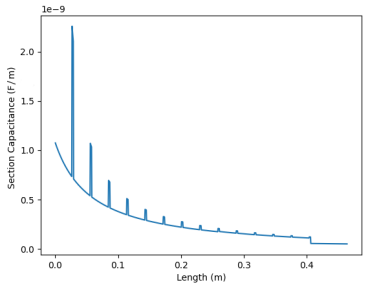

Figure 5: Section Capacitance vs Length

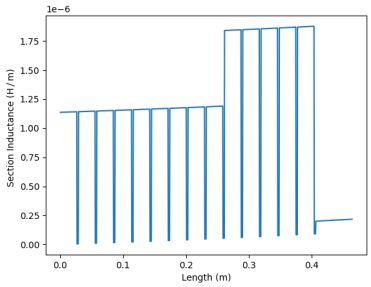

Figure 6: Section Inductance vs Length

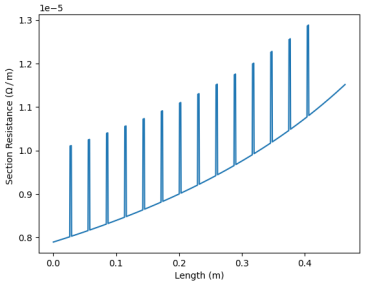

Figure 7: Section Resistance vs Length

impedance as a function of the cumulative impedance of the previous section and the current section's properties. The formula is as follows:

$$Z_{in}(l-\Delta l) = Z_i(l)\frac{Z_{in}(l) + Z_i(l)tanh(\gamma_i(l)\Delta l)}{Z_i(l) + Z_{in}(l)tanh(\gamma_i(l)\Delta l)} \tag{3}$$

$$Z_i(l) = \sqrt{\frac{j\omega L_i(l) + R_i(l)}{j\omega C_i(l)}} \tag{4}$$

$$\gamma_i = \sqrt{(j\omega L_i(l) + R_i(l))(j\omega C_i(l))} \tag{5}$$

$$R_i(l) = \frac{\mu_0}{2}\sqrt{\frac{f}{\pi\sigma}}(\frac{1}{b(l)} + \frac{1}{a(l)}) \tag{6}$$

$$L_i(l) = \frac{\mu_0 ln(\frac{r_a}{a(l)}) + \mu(l) ln(\frac{b}{r_a})}{2\pi} \tag{7}$$

$$C_i(l) = \frac{2\pi}{\frac{1}{\epsilon_0} ln(\frac{r_a}{a(l)}) + \frac{1}{\epsilon(l)} ln(\frac{b}{r_a})} \tag{8}$$

where $a(l)$ and $b(l)$ represent the inner and outer conductor radii respectively,

and r_a is the inner radius of the ferrite discs. With left-right symmetry about the coupling stem, we can calculate the impedance for half the tuner and then use the formula for parallel resistance, $\frac{1}{Z_{total}} = \frac{1}{Z_1} + \frac{1}{Z_2}$. This calculation is similar for the impedance of the cavity, with the caveat that many parts are simplified due to a lack of ferrite discs. Then, we can calculate the total impedance for the cavity-tuner system using the fact that the cavity and its three tuners are also in parallel, so $\frac{1}{Z_{total}} = \frac{1}{Z_{cavity}} + \frac{3}{Z_tuner}$.

The Q factor can be obtained from the impedance using the formula

$$Q = \frac{\omega \mathrm{Re}(Z_{cavity}(\omega))}{2} \frac{\mathrm{d}(\mathrm{Im}(\frac{1}{Z_{cavity}(\omega)}))}{\mathrm{d}\omega}\Big|_{\omega=\omega_0} \tag{9}$$

where ω_0 is the angular resonant frequency.

3.4.1 RMS Bias Current

Due to highly non-trivial values that have only been found empirically, we cannot calculate the RMS bias current using traditional analysis. We approximate the RMS bias current using the average between the currents when the resonant frequency of the cavity is at 44.7 MHz and 52.8 MHz.

To calculate the true RMS bias current, we first start with the frequency with respect to the bias current.

$$f(I) = \Delta f(1 - e^{\frac{-(I-I_{min})}{I_d}}) + f_0 \tag{10}$$

where f_0 is the frequency at the minimum bias and Δf is the difference between the frequency at the minimum bias, I_{min}, and saturation, I_d.

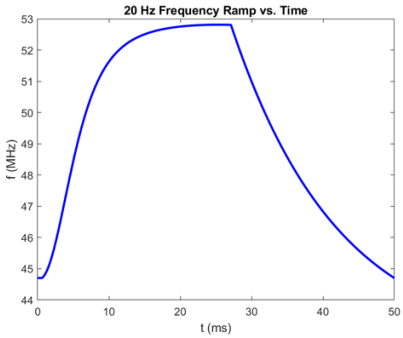

Figure 8: Frequency Ramp vs Time

Using (10) and a predicted frequency ramp between 44.7 and 52.8 MHz with a 50% duty cycle at 20 Hz and an exponential ramp back down to 44.7 MHz, we can then find the bias current with respect to time.

$$I(t) \approx -I_d \ln(1 - \frac{f(t) - f_0}{\Delta f}) + I_{min} \tag{11}$$

We can then use (11) to calculate the RMS bias current via

$$RMS_I = \sqrt{\frac{1}{T}\int_0^T I(t)^2 \,\mathrm{d}t}. \tag{12}$$

3.5 Lookup Table

For performance, A lookup table was generated as an approximation for the traditional Q factor calculation. Lookup tables are useful as they transform a computationally-intense problem into a spatially-intense problem by running the calculations over a dense enough subset of the problem space. The table would pre-compute values for a large amount of permutations evenly dispersed throughout the problem space, then, an approximate value for any given specimen can be obtained by interpolating between nearby values.

While generating the lookup table takes a very long time, it can now be cached and re-used no matter what changes happen to the evolutionary strategy. The goal is to have the results of this evolutionary strategy be piped into another meta-heuristic, likely another evolutionary strategy. That secondary meta-heuristic would use the direct calculation with the now much smaller problem space to find the best solution for the objective functions.

There is a major issue with using a lookup table. Adding any new dimension to the data has a multiplicative effect on the lookup table's generation time and storage cost. It is an $O(n^k)$ time and space complexity, where n is the highest resolution axis and k is the dimensions of the input data.

3.6 Discretized Gradient Descent

All data that the frequency column of the lookup table provides that is directly usable is a small slice at or around the resonant frequency, so we can calculate the Q Factor at that point. That means there is a substantial amount of unnecessary data being calculated and stored in the table. Therefore, a modified gradient descent algorithm was devised to find the resonant frequency for a specific tuner configuration at a fixed bias current. Gradient descent simply follows the gradient of a function to always move towards a local optimum. The algorithm is as follows[1]

$$x_n = x_{n-1} + \alpha \frac{\mathrm{d}}{\mathrm{d}x} f(x) \tag{13}$$

where α is the step size.

The primary issue is that gradient descent requires a convex curve to obtain a globally optimum solution, and the impedance follows a resonance pattern. However, if we limit our view to only the peaks of the resistance, we can see that they follow a concave curve. So, using the period of the impedance, we

[1] Note that this formula is technically a gradient *ascent*, however the terminology of gradient descent is much more colloquially accepted.

can use a gradient descent on the peaks to find the true resonant frequency. The discretized gradient descent uses the slope between neighboring peaks as the gradient, and locks the step size to the period of the impedance. To obtain the period of the impedance, we can run two disparate gradient descents two find two peaks, then take a series of integer divisions to test generated periods until every point lies on a peak. To prevent erroneously finding a multiple of the actual period, we can use an assumed bound on the maximum distance between resonances, $|f_n - f_{n-1}| \leq \epsilon_0$.

4 Results

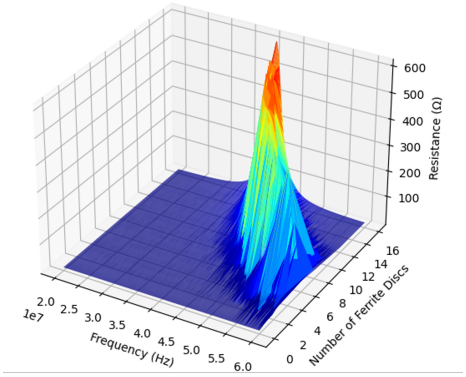

Figure 9: Resistance vs Frequency vs Number of Ferrite Discs

This project is hasn't produced any exact tuner specifications. However, it has verified some general trends in the data.

- Decreasing tuner length (or the representation as such via shorting) drops Q Factor but maintains frequency range.
- Offsetting discs (by filling the excess space with a dielectric) boosts Q Factor but lowers the tuning range. Notably though, with the PIP-II upgrade we only need 54% of the normal tuning range.
- Q Factor has a mostly inverse relation with the number of ferrite discs on the tuners.
- The tuning range drops as we remove discs.
- The resonant frequency range increases as the number of discs decreases.

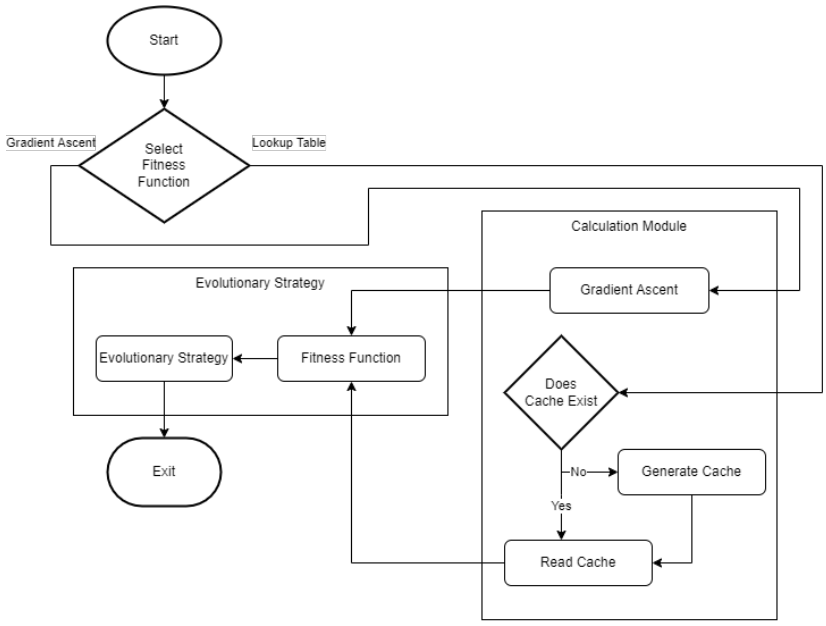

Figure 10: Program Flowchart

4.1 Program

Arguably more important than a specific tuner configuration, this code has become a tool with a solid base for further iterations to develop. We constantly obtain more and more accurate models and simulation data on these RF cavities, and thus the program is still developing. Each time new information is found, we want to be able to combine it with prior art such that we can quickly re-run the pipeline and receive new results with minimal additional work. That is what this program aims to be.

5 Future Work

The next goal is to implement a form of multi-objective optimization. The goal is to have each fitness function; Q factor, tuning range, RMS bias current, and tuning sensitivity contribute to the selection process and help find the Pareto front, or, the set of the most optimal solutions. This front is a set of solutions that are considered Pareto efficient, that is, any change would either cause a loss in one or more of the fitness values or would become one of the other optimal solutions.

Additionally, we would like to develop models for other metrics. Finding a metric for the tuning sensitivity of a given tuner configuration provides a

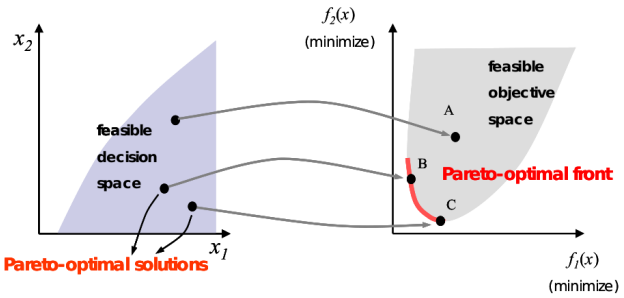

Figure 11: Finding the Pareto-optimal front for a multi-objective minimization problem

much different dimension to the problem's constraints. Fitting a model for the unknown variables in the RMS bias current calculation would allow the dropping of the current weak approximation. And finally, finding an analytic approach to the currently iterative calculation of cavity and tuner impedances would allow a much faster calculation time.

Another consideration for future work is to replace the uncorrelated mutation with n step sizes with a correlated mutation with a covariance matrix. This is known as a CMA-ES (Covariance Matrix Adaptation Evolution Strategy). It is similar to an evolutionary strategy with uncorrelated mutations with n step sizes, except the covariance matrix allows the step sizes $< \sigma_1, \sigma_2, \ldots, \sigma_n >$ to interplay with each other and rotate the space of possible moves about the specimen in a way that creates a higher probability of movement along the axis towards an optimum.

References

[1] DIXON, R. (2015, September 3). *Resonant cavities for the acceleration of charged particles*. Fermilab at work. Retrieved August 12, 2022, from https://news.fnal.gov/2015/09/resonant-cavities-acceleration-charged-particles/

[2] SHVAB, I. (2021, June 9). *Optimization modelling in python: Multiple objectives*. Medium. Retrieved August 12, 2022, from https://medium.com/analytics-vidhya/optimization-modelling-in-python-multiple-objectives-760b9f1f26ee

[3] VAUGHN, B. J. (2021). *Booster Cavity MATLAB Analysis* [PowerPoint slides]. Accelerator Division, Fermi National Accelerator Laboratory.