

ION ACCELERATOR APPLICATIONS, COMPACT RF LINEAR ARCHITECTURES  
AND MACHINE LEARNING FOR REAL-TIME TUNING AND OPTIMIZATION

By

Nicholas Anthony Valverde

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

Physics — Doctor of Philosophy

2024

## Abstract

Linear ion accelerators are widely applicable tools for scientific research and industry advances. The broad scope of applicability demands a wide range of beam and accelerator operating regimes that entail careful design for desired performance. Optimally meeting desired beam properties depend on thorough analysis of the interplay between a charged particle beam and material structures that are responsible for producing the electric and magnetic fields used to guide, focus, and accelerate the beam. This interplay is captured in the study of beam dynamics where the transverse and longitudinal dynamics are studied, to the extent possible, separately under reduced models to gain intuition about the system. The dynamics are then later examined with first principle particle codes that simulate the full 3D dynamics and combined non-linear effects.

This thesis mainly discusses the theoretical and computational work done to support the development of compact, cost effective, multi-beam linear ion accelerator systems. Part I provides a brief description of the theoretical concepts needed to better understand particle dynamics in radio frequency linear accelerators. Part II discusses efforts made to optimize conductor geometries used to focus and accelerate beam arrays, and reduced models developed to analyze and better optimize the longitudinal and transverse dynamics in isolation. These reduced models are used to formulate improved systems that are then self-consistently simulated using a particle-in-cell accelerator code called Warp to model full 3D and non-linear effects to better evaluate system performance. Ideas for improved future systems building off the developed computational tool kit are discussed. Lastly, Part III details a limited scope project that applied a robust auto tuning algorithm to an existing ion induction accelerator NDCX-II. The NDCX-II system has complicated pulse compressions that are difficult to optimize for pulse durations, spot size, and flux on target. Once found, optimal settings will

drift in time as due to various non-static changes (lab conditions, aging components, etc.)

The algorithm was used both as a case study in finding optimum settings for the accelerator, and as a way to sample the accelerator parameter space in order to collect data and build a neural network to act as a surrogate model. With a tuning algorithm, the system can rapidly be brought back to the most-current design optimum and the surrogate model can be used to efficiently probe the parameter space for new optimum, compare performance and identify faults, and more. In this par, the tuning algorithm and data cycle is detailed. A neural network is created to predict accumulated charge on target with limited experimental data—approximately 400 experimental runs.

Dedication: to be written

## ACKNOWLEDGMENTS

I could have not completed this work without the assistance of so many people. Unfortunately, I cannot name everyone as the names and reasons would fill volumes but I do my best to capture the immense support I have received from various persons. Firstly, my mother Connie Nikolaidis, father Robert Valverde, and brother Nathan Valverde for their endless support and for not questioning my path. Though, from an outside perspective my “path” probably looked like the unintelligible meandering of a dolt at best. I was strongly influenced, encouraged, and motivated to pursue higher education by the folks at Complex Adaptive Systems at Arizona State University (ASU), namely: Stephanie Calderone, April Johnson, Dr. Anna Barker, and Dr. George Post who acted as a mentor and an excellent academic to someday emulate. Dr. Sangeeta Malhotra at ASU was incredibly influential and helpful in my early undergraduate years. Before I was a graduate student at Michigan State University (MSU), I was lucky enough to land a Research Experience for Undergraduates (REU) at MSU and was mentored by Dr. Subhendra D. Mahanti who was an excellent mentor and gave me my first exposure to theoretical calculations and computer simulation work. This experience was instrumental to my pursuing of graduate work at MSU. While attending MSU I had nothing but good experiences despite the rigorous classwork and the ever imposing imposter syndrome. I want to acknowledge the graduate department secretary Kim Crosslan; Kim looks out for all of us graduate students and has always looked out for me. It is not currently known if the department can function without her. I’m not sure how she does it but she is one of a kind and I cannot thank her enough. Initially I joined the IMSRG group under Dr. Scott Bogner, Dr. Morten Hjorth-Jensen, and Dr. Heiko Herbert. While in this group I had to evaluate my skillset and design a self-study curriculum

to improve my programming language. The members of this group were incredibly helpful. They were eager to assist and Morten provided me with the best resource for programming in Python, to which I am thankful to him and the late Dr. Hans Petter Langtangen for his inimitable writings and textbooks. I am grateful to Heiko for providing me with additional computational resources when asked. I am also very grateful to Scott for having the difficult conversation with me about my interests and setting me on the path to accelerator physics. When searching for new a research group, Dr. Artemis Spyrou curated a list of interviews for me which I'm especially grateful. This landed me under the advisement of Dr. Steven Lund who I will save for last (best?). I also joined the Accelerator Science Engineering and Technology (ASET) group which has been a tremendous group that never ceases to offer support and opportunities. When I first joined, Dr. Robert Hipple along with Dr. Lund gave me an initial project. Robert was generous with his time and wisdom, and helped me avoid many pitfalls early on in my graduate career. My initial training, and much of my subsequent research, was done using the Warp software and Dr. Dave Grote was incredibly helpful in troubleshooting at all stages of my graduate work. I would like to specifically acknowledge and thank Dr. Peter Ostroumov who is the group lead at MSU and a great one at that. I would also like to thank my comrades in the ASET group who are as humble as they are talented. When forming my committee I was worried about collecting such remarkable scientists to judge my work. Years later I am grateful that this worry was misplaced as my committee consisting of Dr. Steven Lund (Chair), Dr. Yue Hao, Dr. Dean Lee, Dr. Steven Lidia, and Dr. Chong Yu Ruan have been supportive, encouraging, and helpful through my years. Although late to the accelerator world, I was rapidly brought up to speed by taking courses offered by the United States Particle Accelerator School (USPAS) and my sincerest thanks to this wonderful school and the many scientist who volunteer their time to

keep it running. I'd especially like to thank Susan Winchester and Dr. Lund for handling the logistics of the school and navigating the various blocks that arise each session. After completing this coursework, I was lucky to be placed into a research position at Lawrence Berkeley National Lab (LBNL) under Dr. Qing Ji and Dr. Arun Persaud who are both great scientists. I have learned much working with these two over the past few years and their patience and support know no bounds. Graduate school is a difficult journey (an understatement) and my journey would have stopped long before the finish line had it not been for the external support of various friends and communities. My volleyball group consisting of Dr. Zoe Hansen, Dr. Macy Pell, Dr. Adam Kawash, Kurt Walcheske, Steven Hemmerle, Meeshon Rogers, and Dr. Mike Pajkos were an endless source of support and jubilation. I'd like to separately acknowledge Mike for his friendship and indulging me in many discussions. The yoga community at Yoga Connect was an especially important place to me and one that I will not be able to easily replace, if at all. I want to thank the Tucson and East Lansing volleyball community where I can vent frustrations with a great group of people in a fantastic sport. I'd Like to acknowledge Josh Canlas who has indulged me in many discussions and has been a kindred spirit that has given me inspiration. Captain Tyler Dolezal of the United States Air Force who has been my friend and fellow physicist since undergraduate. His help, counsel, and our various discussions have been invaluable. Hang Khuu deserves special thanks. Her support is unrivaled and I'm especially grateful to her ensuring that I ate when research interests held my focus from hunger.

Finally, I'd like to circle back to my adviser and, dare I say, friend Dr. Steven Lund. He has been with my through almost all the tribulations offering endless support, guidance, and wisdom. His mentorship is only matched by his incredible scientific mind. I would consider myself a lucky individual for having the opportunity to be guided by so brilliant an

individual and to have shared so many discussions, thoughts, and time. I doubt I will ever repay the debt I have accrued to this individual but I hope that I can succeed as a notable scientist and that would pay off some of it. Thank you Steve for these last few years and your help through all of it. I hope we can continue to work together and find time to chat.



# CONTENTS

<b>LIST OF TABLES</b> . . . . .	<b>xi</b>
<b>LIST OF FIGURES</b> . . . . .	<b>xii</b>
<b>I Linear Ion Accelerators</b>	<b>1</b>
Chapter 1 Introduction . . . . .	2
Chapter 2 Longitudinal Dynamics . . . . .	4
2.1 The Transit Time Factor . . . . .	5
2.2 RF Synchronism for Particle Acceleration . . . . .	8
2.3 Summary . . . . .	18
Chapter 3 Transverse Dynamics . . . . .	19
3.1 Single Particle Dynamics . . . . .	20
3.1.1 Calculating the Self-Fields for an Elliptical Beam with Uniformly Distributed Charge Density . . . . .	24
3.1.2 Calculating the Forces from Applied ESQ Fields . . . . .	26
3.1.3 The Kapchinskij-Vladimirskij (KV) Envelope Equations . . . . .	29
3.2 Summary . . . . .	37
Chapter 4 Self Consistent Simulations . . . . .	38
<b>II Compact Multi-Beam RF Linear Ion Accelerators</b>	<b>43</b>
Chapter 5 MEMS-RF System . . . . .	44
5.1 Introduction . . . . .	44
5.2 Overview of Compact Multi-Beam MEMS Linac Structures . . . . .	46
5.3 Fields from RF Acceleration Arrays . . . . .	49
5.3.1 Electric Field Profile of the RF-Gap . . . . .	50
5.4 Fields from Optimized Electrostatic Quadrupole (ESQ) Focusing Arrays . . . . .	53
5.4.1 Numerical Methods . . . . .	54
5.4.2 Optimizing ESQ Arrays to Minimize Leading Order Error in Field Expansion . . . . .	58
5.4.2.1 Benchmarking for a Single Axially Long ESQ . . . . .	59
5.4.2.2 Finite Axial Length ESQ Optimization for Full and Chopped Rod Geometries . . . . .	60
5.4.2.3 Voltage Breakdown Limits for ESQ Array Geometry . . . . .	62
5.4.2.4 Optimization Considerations From Realistic Geometry . . . . .	64
5.5 Summary and Concluding Remarks . . . . .	69
Chapter 6 Longitudinal Dynamics in MEMS-RF . . . . .	70
6.1 Numerical Methods . . . . .	71
6.2 Benchmarking the 1D Model with Warp in a 4-Gap System . . . . .	73
6.3 Lattice Design for Optimal Acceleration, Bunching, or Both . . . . .	74
6.3.1 Initial Conditions and Parameter Settings . . . . .	75

6.3.2	Comparing Lattice Designs . . . . .	76
6.4	Summary and Concluding Remarks . . . . .	82
Chapter 7 Transverse Dynamics in MEMS-RF . . . . .		84
7.1	Geometric and Beam Parameters . . . . .	84
7.2	Numerical Integration of KV-Envelope Equations . . . . .	85
7.3	Modeling the RMS-Edge Envelope Evolution Without Acceleration . . . . .	87
7.3.1	Benchmarking with Warp . . . . .	89
7.3.2	Finding Matched Conditions Using the Nelder-Mead Search Algorithm . . . . .	93
7.4	Modeling the RMS-Edge Envelope Evolution with Acceleration . . . . .	97
7.4.1	Creating a Benchmark with Warp . . . . .	97
7.4.2	Using a Thin-gap Model to Model the Transverse Effects from Acceleration . . . . .	98
7.4.3	Using a Thick-gap Model to Model the Transverse Effects from Acceleration . . . . .	101
7.5	Lattice Design For Compact RMS Edge Envelope . . . . .	110
7.5.1	The Pre-matching Section . . . . .	110
7.5.2	The Post-matching Section . . . . .	113
7.5.3	The Matching Section . . . . .	115
7.5.4	The Rest of the Lattice After Matching . . . . .	116
7.6	Summary and Concluding Remarks . . . . .	122
Chapter 8 3D Self-Consistent Particle-In-Cell Simulations of MEMS-RF . . . . .		125
8.1	Aside: Panel Plots and Simulation Snapshots . . . . .	125
8.2	Beam Creation and Loading . . . . .	127
8.3	Simulation Mesh Configuration . . . . .	129
8.4	Simulation Control . . . . .	130
8.5	Diagnostic and Data Management . . . . .	131
8.6	Pulsed Beam Injection . . . . .	132
8.7	Continuous Beam Injection . . . . .	138
8.8	Summary and Concluding Remarks . . . . .	141

### **III Machine Learning and Optimization of an Ion Induction Linac** **143**

Chapter 9	Introduction and System Overview . . . . .	144
Chapter 10	An Auto Tuning Algorithm for Safe Extremum Seeking in NDCX-II . . . . .	148
Chapter 11	Data Collection, Retrieval, and Preprocessing . . . . .	150
Chapter 12	Training a Neural Network to Predict Peak Current and Time of Peak . . . . .	152

<b>Bibliography</b> . . . . .	<b>159</b>	
Appendix A	Beam-Movie Image Panels . . . . .	163
A.1	Pulsed Injection . . . . .	163
A.2	Continuous Injection . . . . .	169

## LIST OF TABLES

Table 5.1: Standard parameter values for the MEMS-RF system. The initial or source values are per beamlet. Values shown as $\sim \#$ have some degree of variability due to operation conditions or changes in beamlet parameters— $L_p$ will change with increasing $\mathcal{E}_b$ . The $R$ listed is the optimal value found in Sect. 5.4. . . . .	49
Table 7.1: Geometric and initial-beam parameter settings used in the matching section simulations. . . . .	85
Table 7.2: Initial conditions (envelope edge radii and envelope edge angles) used for initial condition is shown in the first row. The matched coordinates are shown in the second row. . . . .	93
Table 7.3: Final parameter values for the pre and post-matching sections. The matching section will have to match $\vec{r}_f$ from the pre-matching section to the post-matching section. $Q_f$ and $\epsilon_f$ from the pre-match section are used to correctly model the beam in the post-matching section and find match coordinates listed using the search algorithm. . . . .	115
Table 7.4: Parameter values for a tuned lattice starting at the post-matching section. The tuning was done for each lattice period (each row) by extracting the necessary parameters $\mathcal{E}_f$ , $Q_f$ and $\epsilon_f$ from the previous lattice period to model the beam in the current lattice period and tune. The voltage holding is symmetric with $V_2 = -V_1$ and was hand chosen so that $\bar{r}_x \approx \bar{r}_y \approx 0.28$ mm. This methodology introduces mismatch between lattice periods but due to low phase-advance, we expect the lattice to be partially insensitive and the beam to maintain well-confined. . . . .	118
Table 8.1: Notable parameter values used. This design was for four gaps at max acceleration as noted by $\phi_s$ and forms the baseline for comparing other design choices for pulsed injection schemes. . . . .	133
Table 8.2: Values for the peak current and transmission for different design cases using a $20^\circ$ pulsed beam and four acceleration gaps. . . . .	134
Table 8.3: Values for the peak current and transmission for different design cases using a $20^\circ$ pulsed beam and 16 acceleration gaps. . . . .	136
Table 8.4: Values for the peak current and transmission for different design cases using a continuous beam and 16 acceleration gaps. . . . .	140

## LIST OF FIGURES

Figure 2.1:	Schematic for accelerating cells in linac structure. The particle arrives at each gap center at a phase $\phi_{n-1}$ and is accelerated which changes the particle velocity from $\beta_{n-1}$ to $\beta_n$ . The particle will then progress to the next acceleration center and arrive at a phase of $\phi_n$ . . . . .	10
Figure 2.2:	Operating phase convention typically used for the cosine oscillating electric field in the gap. The design phase is chosen to be between $-\pi \leq \phi_s \leq \pi$ . Examination of the Hamiltonian $H_\phi$ shows that the region between $-\pi/2 < \phi < 0$ is a region of stability. Points 1 and 2 represent operation points for maximum bunching and maximum acceleration respectively. When operating at Point 1 with $\phi_s = -\pi/2$ the maximum amount of particles will be captured in the RF-bucket and there will be no overall acceleration of the beam. At Point 2 with $\phi_s = 0$ the design particle is maximally accelerated but the RF-bucket will be theoretically zero and the minimum amount of particles will be captured. Between these two points there will be both acceleration and bunching. . . . .	15
Figure 2.3:	Image borrowed from Wangler [2]. The left panel shows three plots as a function of $\phi$ . The top plot is the electric field in the gap. The middle plot is the $\Delta\mathcal{E}$ - $\phi$ phase-space contour plot. Lastly, the bottom plot is the potential term $V_\phi$ in the Hamiltonian $H_\phi$ . The phase-space plot labels the region of stable orbits, or RF-bucket, that reside within the outer most stable contour called the separatrix. To find the separatrix one must locate two fixed stable points $-\phi_s$ and $\phi_2$ . If operating at maximum bunching $\phi_s = -\pi/2$ then the phase-space contours form an eye shape as depicted by the right plot. This operating phase has the largest RF-bucket. . . .	17
Figure 2.4:	Image borrowed from Wangler [2]. Acceleration in the system will introduce damping to the phase-space orbits and the fish structure in Fig. 2.4 morphs into a golf club structure. . . . .	18
Figure 3.1:	An elliptical beam of uniformly distributed particles within the ellipse. There is a constant current density $\vec{J} = \rho\beta_b c \hat{z}$ . The beam radii are given by $r_x$ and $r_y$ as shown. Image borrowed from [12]. . . . .	25
Figure 3.2:	Phase space ellipse in $x - x'$ with points of interest labeled in terms of the Twiss Parameters $\alpha$ , $\beta$ , and $\gamma$ . Figure adopted from [9] . . . . .	34

Figure 4.1:	Ideal flow of simulation and experiment. Reduced models are designed and used to rapidly probe the parameters and design space. This narrows the scope so that the more costly simulations can be ran with Warp for more sophisticated treatment and analysis. This first stage is repeated numerous times as signified by the tight double arrow. Once convergence has been met at this stage, the design can be experimented on. Because real experiment is more challenging, this second stage is repeated less frequently as signified by the loose double arrow. . . . .	42
Figure 5.1:	Basic Linac Design. Note that a single RF-unit is comprised of two RF acceleration gaps. Two single RF-units are connected to the same voltage source but multiple voltage sources will allow for frequency changes down the accelerator line. After the matching section, the Linac can be decomposed into a periodic lattice consisting of two RF-units. . . . .	46
Figure 5.2:	An ESQ wafer for a 120 beamlet array design. The 8 acceleration gaps are mounted on a four rods that hold the wafers in place. This design offers an easy way of adjusting the spacing between wafers to manipulate the RF phasing by changing the distance from one gap center to another.	47
Figure 5.3:	Front view of the array of gaps (left) and a close-up of a single gap (right). The conducting material, in our case copper, is shown in grey. In blue is the FR-4 material that the PCB is comprised of. The PCB is copper clad on both sides and the pattern and aperture hole (shown in black) created using laser cutting. Image adapted from [28]. . . . .	50
Figure 5.4:	Extracted On-axis electric field $E_z(r = 0, z)$ using Warp for transverse-periodic boundary conditions (top row) and Dirichlet-boundary conditions (bottom row). Two RF-gaps are needed to correctly simulate the transverse periodic-array geometry. If a single gap is used, then the fields will look like the bottom row. The gap width $g$ is labeled as blue vertical dash lines. The field data can be isolated by extracting the data within the red box. . . . .	51
Figure 5.5:	Transverse view of a single electrostatic quadrupole in a simulation box with grid sizes $\Delta x = \Delta y$ . The conducting rods are shaded in grey with the applied voltage $V_q$ . The rods have a radius of $R$ and the dashed line shows the clear bore aperture with radius $r_p$ . Conducting objects are shaded gray with all white space representing vacuum. . . . .	55
Figure 5.6:	Electric field gradient one grid-cell off-axis for a single isolated ESQ made of solid cylindrical rods. The black vertical dashed line demarcates the physical length of the ESQ $\ell_q$ . The green line shows the hard-edge equivalent with a total field extent equal to the calculated effective length $\ell_{\text{eff}}$ .	56

Figure 5.7: Ratio of physical ESQ length $\ell_q$ to calculated effective length $\ell_{\text{eff}}$ as $\ell_q$ in increased. . . . .	57
Figure 5.8: (Left) The clear bore aperture is displayed as the dotted red circle with radius $r_p$ . The integration path is shown as a dotted block line at radius $r$ and was chosen to be approximately one grid cell within the aperture. The radial distance between $r_p$ and $r$ is represented by $\Delta$ which is chosen to be grid cell away from the ESQ edges shown in solid black lines. (Right) Schematic of the the simulation mesh with the interpolation points passing through the discretized space with simulation cell box sizes of $\Delta x$ and $\Delta y$ . . . . .	59
Figure 5.9: Leading error term $A_6$ as a function of the rod radius $R/r_p$ for an infinitely long quadrupole. . . . .	60
Figure 5.10: Leading error term normalized by the quadrupole term $A_6/A_2$ in the multipole field expansion for finite length ESQ using full (blue circles), half-chopped (green triangles), and quarter-chopped (magenta "x"s) rod geometries. To the right are illustrative examples of what the geometric configuration is and where the chop is being done. . . . .	61
Figure 5.11: (Top) Schematic of a single cell from the realistic transverse-periodic array ESQ design used in simulations. The clear bore aperture is shown as the dotted black line with a radius of $r_p$ . The half-chopped quadrupole rods have radius $R$ . Conducting objects are shaded in grey and white-space represents vacuum. (Bottom) Transversely periodic ESQ array modeled in simulation. Positively biased structures are colored blue and negatively biased structures in red. . . . .	63
Figure 5.12: Leading error term as a function of the rod radius $R/r_p$ for an ESQ with the manufacturedx geometry. . . . .	65
Figure 5.13: Gradient of ESQ Doublet structure. The ESQ elements are drawn below the plot with axial length $\ell_q$ and axial separation distance $d$ . . . . .	67
Figure 5.14: Difference in the calculated effective length $\ell_q$ as the doublet separation distance $d$ is varied. . . . .	68
Figure 6.1: Visual description of the array-stitching process for the position data. An initial $z$ -array is created with $n$ -grid points. The extracted gap data has both field data and position data. Here the position data $\mathbf{z}_{\text{ext}}$ is stitched into the initial array between $z^k$ and $z^{k+1}$ . The resultant mesh will have $n + p$ grid points. An identical algorithm is performed in tandem for the field array and extracted field data. . . . .	72

- Figure 6.2: Comparison of a four-gap simulation with  $V_g = 7$  kV and  $\mathcal{E}_{b,i} = 7$  keV between the 1D code (blue) and Warp (orange). The binning is done over the final energies of the particles after arriving at target. Particle energies are filtered based on arrival time relative to the synchronous particle with  $|\Delta t| = |t - t_s| \leq 0.5\tau_{\text{rf}}$  . . . . . 74
- Figure 6.3: Normalized on-axis electric field for a 4-gap system at  $t = 0$ . The electric fields are normalized by the calculated DC-electric field  $E_{\text{DC}} = V_g/g$  (V/m). The on-axis voltage will be less than the applied voltage in the real conductor. Since the field is extracted from the conductor, the field needs to be scaled to be greater than  $E_{\text{DC}}$  to ensure the on-axis voltage matches the desired voltage hence why the amplitudes in the plot are greater than unity. Each gap center is marked by a grey vertical dashed line and the distances between gap centers increases with increasing energy by  $\beta\lambda_{\text{rf}}/2$ —the RF-resonance condition for gaps operating in the  $\pi$ -mode. The particle advance is terminated at the Fcup marked by a red vertical dashed line and placed 10 mm from the final gap edge. . . . . 76
- Figure 6.4: Phase space plots for different design choices. The main plot is a kernel density estimator (KDE) plot and the estimates at the extreme values have been chopped to prevent misleading data such as energies greater than possible in 1. Each plot also has the  $\mathcal{E}$  and  $\Delta t/\tau_{\text{rf}}$  binned and plotted on top and bottom respectively. The vertical dashed lines represents the design particle and the vertical line will always be at 0 since the  $\Delta t = t - t_s$ . Where the lines intersect is where the design particle is in phase space. . . . . 77
- Figure 6.5: Similar to Fig. 6.4, these plots are for a 16 gap system and D) has been included for a new design where the gaps are tapered for a small number of gaps before the gaps are design for maximum acceleration. . . . . 81
- Figure 7.1: The extracted  $\kappa(z)$  in blue and the hard-edge equivalent in green are both normalized by  $\hat{\kappa}$ , the focusing strength calculated in the 2D case using the initial beam parameters and the half breakdown limit for  $V_q$ . At the top of the figure is a schematic showing the ESQ positioning along the mesh. The case shown has voltage settings (from left to right) of  $(V_1, V_2, V_3, V_4) = (-59.97, 79.94, -39.98, 79.97)$  V. . . . . 88
- Figure 7.2: An example of a continuously injected beam and the discretized grid for which the moments are calculated. The potential is normalized by the half-breakdown limit  $V_{1/2\text{-breakdown}} = 500$  V and is generated by the ESQ conductors in gray. A random selection of 10,000 particles are plotted as black dots. The outer panel is a close-up view of the discretized grid. Particles within  $\Delta z$  of the two grid points are used to calculate the statistical moments which are then used to evaluate the statistical envelope evolution. . . . . 90

Figure 7.3: The rms-edge of the beam envelope boundary A) and C) and the envelope angles  $r'_x$  and  $r'_y$  of the beam edge B) and D). The  $x$ -coordinate is in black and  $y$ -coordinate in blue. Each model is shown on the plot: a full 3D simulation with space-charge done with Warp (solid line), the KV model using the extracted  $\kappa(z)$  (dashed), and the KV model using a hard-edge model (dotted). The full simulation and the extracted  $\kappa(z)$  are overlaid showing good agreement and suggesting that the extracted  $\kappa(z)$  using the KV model can serve in-place of an expensive full 3D simulation. The bottom is a result of adjusting the effective length in the hard-edge model for more focusing. With this adjustment the hard-edge model agrees with the other two models. . . . . 92

Figure 7.4: Matched solutions given by a previously developed Mathematica algorithm [33]. The software assumes a periodic lattice and does not account for acceleration. Panel A) shows the focusing strength from the ESQ conductor  $\kappa_q$  in the transverse direction ( $x$  in black,  $y$  in red) for single lattice period. Panel B) shows the evolution of the matched rms-edge radii  $r_x$  and  $r_y$  in black and red respectively over a single lattice period. Panel C) shows the evolution of the matched rms-edge angle  $r'_x$  and  $r'_y$  in black and red respectively over a single lattice period. . . . . 94

Figure 7.5: Flow diagram of the optimization algorithm for finding a matched condition. There are two types of matching that can be done. In voltage fixed matching, we seek a matched solution given the ESQ voltages and thus  $\kappa(z)$ . The algorithm will then iteratively perform the optimization process and vary the input coordinates  $\vec{r}_i$  until a set tolerance is met. In coordinate fixed matching, we seek a matched solution given  $\vec{r}_i$  and  $\vec{r}_f$ . The algorithm will iteratively perform the optimization process, vary the voltages, and recompute  $\kappa(z)$  until a set tolerance is met. Both schemes use the Nelder-Mead search algorithm and rely on a created cost function to guide the searching. . . . . 96

Figure 7.6: Comparing the envelope rms-edge evolution through a single DC acceleration gap between Warp and a thin-gap model. On the left is the rms-edge envelope radii and on the right the rms-edge envelope angle. The thin-gap model is given by a dashed line and Warp as a solid line. Dashed vertical lines mark the gap edges. . . . . 100

Figure 7.7: Comparing the envelope rms-edge evolution through a single DC acceleration gap with  $V_g = 7$  kV between Warp and a thick-gap model. (Top row): On the left is the rms-edge envelope radii and on the right the rms-edge envelope angle. (Bottom row): An overlay of  $\epsilon(z)$ . The thin-gap model is given by a dashed line and Warp as a solid line. Dashed vertical lines mark the gap edges. . . . . 108



- Figure 7.8: Fractional normalized emittance calculated at each grid-point normalized by the initial value. The gap is at 7 kV with a 7 keV beam representing the initial conditions from the source and where the acceleration is the strongest. Here, the normalized emittance fluctuates between values less than 1% from the initial value and can be considered conserve. . . . . 109
- Figure 7.9: Transverse envelope edge radius and angle in  $x$  (black) and  $y$  (blue) through a six gap lattice at maximum acceleration. The gap centers are marked by vertical dashed lines and due to the injection symmetry the blue curve is directly overlaid the black curve. Transverse focusing from the gap is strong but after about four gaps begins to overfocus and the beam begins to expand. Additionally, the gap acts to decrease the angle of the beam gap-to-gap until the beam expansion after gap 4. . . . . 111
- Figure 7.10: Using the envelope code, the evolution of  $r_x$  ( $r_y$  will be identical) was averaged to give  $\bar{r}_x$  for a design with a given number of lattice periods. For each lattice period,  $\bar{r}_x$  was calculated and plotted. At about 7 lattice periods,  $\bar{r}_x > r_p$  and significant scraping is expected. At about 16 lattice periods,  $\bar{r}_x \approx 2r_p$  and much of the beam can be expected to be lost. . . . 112
- Figure 7.11: Transverse envelope edge radii (left) and angle (right) for  $x$  (black) and  $y$  (blue). Since the initial conditions are symmetric the two curves are overlaid. The gaps are chosen to operate with design phase  $\phi_s = (-\pi/2 \quad -\pi/3 \quad -\pi/6 \quad 30)$  to give both longitudinal focusing and radial confinement. . . . . 113
- Figure 7.12: A matched solution for the first lattice period after the matching section. Acceleration gaps operating a  $\phi_s = 0$  are shown as vertical dashed lines. ESQs are shown as vertical dotted lines with a black rectangle at the top and the voltage used labeled. Parameters from the exit of the pre-matching section are used as the initial parameters in this post-matching section. The voltage setting was then iterated in until  $\bar{r}_x \approx \bar{r}_y \approx 0.28$  mm. Stars mark the initial coordinates values and it is clear that the match condition is reached since the curves start and end at the same points. . . 115
- Figure 7.13: Transverse envelope edge radii (left) and angle (right) for  $x$  (black) and  $y$  (blue) for the matching section where there is no acceleration (bottom row). The four ESQ centers are shown as vertical dotted lines. Star markers are placed at the desired initial coordinates for the post matching section. The top figure shows the focusing strength  $\kappa(z)$  along the matching section and each ESQ voltage is labeled . . . . . 117

- Figure 7.14: Transverse envelope edge radii (left) and angle (right) for  $x$  (black) and  $y$  (blue) for a full lattice design starting after the post-matching section. The acceleration gap centers are marked by vertical dashed lines with a  $\phi_s = 0$  for all. After every two gaps, there is an ESQ doublet to provide focusing—not shown in the graphic. Each period (two gaps) was tuned using the ESQ voltages to keep  $\bar{r}_x$  and  $\bar{r}_y$  approximately 0.28 mm. Star markers are placed at the initial conditions and due to low  $\sigma$ , the beam is well confined despite slight mismatch between lattice periods. . . . . 118
- Figure 7.15: A snapshot of the various parameter calculations that can be saved in a csv file when using the envelope code. The tuning method discussed in this section was used to hand-tune lattice periods until an energy of 1 MeV was reached using approximately 6 kV gap voltage. . . . . 120
- Figure 7.16: Voltage scaling for increasing ESQ stacks. As the beam energy increases, the required voltage will need to increase to focus the beam. The provided ESQ focusing can be increased by making the ESQ length longer. Instead of machining an ESQ with a different length, the system design allows ESQ wafers to be stacked to increase the focusing provided. The red vertically dashed line marks the half breakdown. This gives a guide for the ESQs when needing to focus beams at higher energies up to 1 MeV. 121
- Figure 8.1: Labeled snapshots of the beam in  $x$  and  $z$  from a continuous injection simulation at different times labeled from 1 to 6 starting at the top left. Panel 1 is at  $t = 0$ . The design particle is shown as a red dot that starts at the injection site but is “disabled” until half the beam has been injected (panel 2). At an absorption boundary is placed at  $r = 0.55$  mm and is labeled in panel 1 with horizontal dashed lines. The gaps are labeled  $G_n$  and panel one shows  $G_1$  which starts at a maximum positive  $z$ -field indicated by the color bar to the right of the panels. Panel 3 shows the full length of the beam and the design particle has been switched “on”. Also note the change in the  $z$  values as the simulation window is now moving with the design particle. Between  $G_1$  and  $G_2$  is a BPM monitor that will record particle parameters. Lastly, panel 5 shows the location of the Fcup and the particles have been absorbed at the Fcup as seen in panel 6. . . 127
- Figure 8.2: Fraction of particles at a radial distance normalized by the aperture  $r/r_p$ (left) and calculated current normalized by the source current  $I_s$  (right) for three cases using four gaps ( $N_g = 4$ ). The bin width in both plots are the same for the all three cases. . . . . 134
- Figure 8.3: Fraction of particles at a radial distance normalized by the aperture  $r/r_p$ (left) and calculated current normalized by the source current  $I_s$  (right) for three cases using 16 gaps ( $N_g = 16$ ). The bin width in both plots are the same for the all three cases. . . . . 136

Figure 8.4:	Fraction of particles at a radial distance normalized by the aperture $r/r_p$ (left) and calculated current normalized by the source current $I_s$ (right) for three cases using 16 gaps ( $N_g = 16$ ). The bin width in both plots are the same for the all three cases. The inset plot shows Case 2 and Case 3 together for better comparison of the current. . . . .	139
Figure 9.1:	(A) CAD drawing of NDCX-II with various components labeled. (B) An actual photo of the accelerator housed at LBNL. Image borrowed from [46]. . . . .	146
Figure 9.2:	The nominal beamline configuration for NDCX-II [46]. Each of the 27 cells have solenoidal focusing. The dark red cells are active induction cells which accelerate and compress the beam, for which there are 12. The first seven have custom waveforms designed for optimal performance. Following are two flat waveform cells and then the last through apply a ramp waveform to longitudinal compress the beam before entering a neutral drift plasma region. In light red are inactive cells that only contribute solenoidal focusing. In dark blue are diagnostic cells for beam measurements. . . . .	146
Figure 9.3:	Voltage waveforms for the 12 active induction cells and the injector [47]. The first seven cells are custom waveforms labeled Cell # on the plot. The remaining five active induction cells are also called Blumleins and are labeled with BL on the plot. . . . .	147
Figure 11.1:	Graphical representation of the data collection and retrieval process for NDCX-II. Conductor settings are initialized and the accelerator is commanded through a Python interface at the computer. After each shot, the experimental data is stored into a shotnumber object in the Cassandra database. The database can be queried through Python interface and the data collected through the data classes shotnumber and shot. . . . .	151
Figure 12.1:	Schematic of a dense fully connected neural network architecture. We adopt the notation given by Andrew Ng's Coursera course [41]. The inputs are given as a single column vector $\vec{x}$ of which there are $n_x$ entries or data samples. A single output layer is given by the vector $\hat{y}$ containing $n_y$ targets. The hidden layers are given by $\ell$ and each layer is comprised of $n_\ell$ neurons. Additionally, each layer has an activation function such as rectified linear unit (RELU). . . . .	152
Figure A.1:	A pulsed injection for a $20^\circ$ beam and a 4-gap lattice operating at maximum acceleration $\phi_s = 40$ . . . . .	163

- Figure A.2: A pulsed injection for a  $20^\circ$  beam and a 4-gap lattice using a tapering design for acceleration and longitudinal bunching  $\phi_s = (-\pi/2 \quad -\pi/3 \quad -\pi/6 \quad 0)$ . . . . . 164
- Figure A.3: A pulsed injection for a  $20^\circ$  beam and a 4-gap lattice using a tapering design for acceleration and longitudinal bunching  $\phi_s = (0 \quad -\pi/6 \quad -\pi/6 \quad -\pi/2)$ . The tapering is reversed to avoid defocusing when maximally bunching. Instead, the maximum bunching is done at the last gap. . . . . 165
- Figure A.4: A pulsed injection for a  $20^\circ$  beam and a 18-gap lattice designed for max acceleration  $\phi_s = {}^{18}0$ . To minimize clutter, only the first gap to appear in the frame is labeled and the vertical axis of each plot is the same as those in plots above with maximum extents  $\pm 1.5$  mm. . . . . 166
- Figure A.5: A pulsed injection for a  $20^\circ$  beam and a 16-gap lattice designed for bunching and acceleration  $\phi_s = (0 \quad {}^3-\pi/3 \quad {}^{12}0)$  with a 4-ESQ matching section and periodic doublet focusing. To minimize clutter, only the first gap to appear in the frame is labeled and the vertical axis of each plot is the same as previous plots in this section with the max extents  $\pm 1.5$  mm. Additionally, the first appearing ESQ is labeled either  $M_n$  for a matching ESQ or  $Q_n$  for a doublet ESQ. Note that each there are two ESQs for every  $Q_n$ -label. Also, due to complications with the cgm data writing, some conductors have off-coloring (red, teal). This has no physical significance. 167
- Figure A.6: A pulsed injection for a  $20^\circ$  beam and a 16-gap lattice designed for maximum acceleration  $\phi_s = {}^{16}0$  with a 4-ESQ matching section and periodic doublet focusing. To minimize clutter, only the first gap to appear in the frame is labeled and the vertical axis of each plot is the same as previous plots in this section with the max extents  $\pm 1.5$  mm. Additionally, the first appearing ESQ is labeled either  $M_n$  for a matching ESQ or  $Q_n$  for a doublet ESQ. Note that each there are two ESQs for every  $Q_n$ -label. Also, due to complications with the cgm data writing, some conductors have off-coloring (red, teal). This has no physical significance. . . . . 168
- Figure A.7: A continuous injection and a 16-gap lattice designed for max acceleration  $\phi_s = {}^{16}0$ . To minimize clutter, only the first gap to appear in the frame is labeled and the vertical axis of each plot is the same as those in plots above with maximum extents  $\pm 1.5$  mm. Panel 1 is a snapshot of the beam after approximately one RF cycle (full beam injected). . . . . 169

Figure A.8: A continuous injection and a 16-gap lattice designed for bunching and acceleration  $\phi_s = (0 \quad 3-\pi/3 \quad 120)$  with a 4-ESQ matching section and periodic doublet focusing. To minimize clutter, only the first gap to appear in the frame is labeled and the vertical axis of each plot is the same as previous plots in this section with the max extents  $\pm 1.5$  mm. Additionally, the first appearing ESQ is labeled either  $M_n$  for a matching ESQ or  $Q_n$  for a doublet ESQ. Note that each there are two ESQs for every  $Q_n$ -label. Also, due to complications with the cgm data writing, some conductors have off-coloring (red, teal). This has no physical significance. Panel 1 is a snapshot of the beam after approximately one RF cycle (full beam injected). . . . . 170

Figure A.9: A continuous injection and a 16-gap lattice designed for maximum acceleration  $\phi_s = 160$  with a 4-ESQ matching section and periodic doublet focusing. To minimize clutter, only the first gap to appear in the frame is labeled and the vertical axis of each plot is the same as previous plots in this section with the max extents  $\pm 1.5$  mm. Additionally, the first appearing ESQ is labeled either  $M_n$  for a matching ESQ or  $Q_n$  for a doublet ESQ. Note that each there are two ESQs for every  $Q_n$ -label. Also, due to complications with the cgm data writing, some conductors have off-coloring (red, teal). This has no physical significance. Panel 1 is a snapshot of the beam after approximately one RF cycle (full beam injected). . . . . 171

# Part I

## Linear Ion Accelerators

# Chapter 1. Introduction

The application of accelerator technologies is manifold and individual accelerator devices are diverse [1]. Accelerators are powerful tools that play a pivotal role in advancing scientific research, medical diagnostics and treatments, industrial processes, and our understanding of the fundamental building blocks of matter. At the heart of accelerator technologies is the principle of imparting kinetic energy to charged particles using applied electric and magnetic fields. There are two primary types of accelerators: linear accelerators (linacs) and circular accelerators (synchrotrons and cyclotrons). Linear accelerators propel particles in a straight line, while circular accelerators use magnetic fields to guide particles along a circular path. Particle accelerators find applications in various fields. In high-energy physics, enormous accelerators like the Large Hadron Collider (LHC) at CERN are used to explore the fundamental particles that make up the universe. In medicine, smaller accelerators are employed for cancer treatment through techniques like radiotherapy. Industrial processes benefit from accelerators in areas such as materials science, where they are used for material analysis and modification. Additionally, accelerators have become integral in generating intense beams of light, known as synchrotron radiation, for advanced imaging and spectroscopy applications. The development of accelerator technologies has been marked by continuous innovation and collaboration among scientists and engineers worldwide. From the early days of particle accelerators in the 20th century to the cutting-edge technologies of today, the field has witnessed remarkable advancements. These include the integration of superconducting magnets, precise beam control systems, and the exploration of novel acceleration techniques, such as plasma acceleration. As we delve into the intricacies of accelerators and accelerator

technologies, we discover not only the immense impact they have on scientific discovery but also their vital role in shaping the technological landscape of the modern world. The ongoing pursuit of more efficient and powerful accelerators promises to unlock new frontiers in scientific exploration and technological applications, fueling progress across diverse fields.



## Chapter 2. Longitudinal Dynamics

Longitudinal beam dynamics plays a critical role in understanding and controlling the behavior of charged particle beams in accelerators as the beam gains energy in the acceleration lattice. The longitudinal motion of particles within an accelerator is essential for achieving precise energy and maintaining stability for optimal performance. The longitudinal dynamics of charged particle beams primarily centers on analysis of the gain in energy and its spread as the beam evolves in the lattice. As particles traverse an accelerator, they experience energy changes due to the applied accelerating fields as well as interactions between particles comprising the beam. Understanding and controlling the longitudinal dynamics of the beam is crucial for achieving desired beam characteristics such as beam quality, stability, and beam lifetime.

A single particle with charge  $q$  will gain energy when advanced through a potential difference. The gain in energy  $\mathcal{E}$  will depend on the accelerating gap voltage  $V_g$  by:

$$\mathcal{E} = qV_g.$$

A simple acceleration gap can be a pair of annuli with one of the pair held at ground and the other held at an applied bias. This will establish a static (DC) electric field between the two and as particles enter through the center aperture they will experience an accelerative force. If these acceleration gaps are then distributed in a straight line then this accelerator system is called a linear accelerator or linac which was first proposed by Ralph Wideröe in 1928 [2]. The DC linac is a simple system but is limited by voltage breakdown which occurs at a few

tens of megavolts. To circumvent the problem of voltage breakdown, time-varying fields can be applied which gives rise to the radio-frequency (RF) linac. The history and development of the RF linac is beyond the scope of this dissertation but can be found elsewhere [3]. This chapter is focused on the key equations in analyzing the longitudinal dynamics and the considerations involved in using an RF linac.

## 2.1 The Transit Time Factor

The energy gain in a time-dependent field at time  $t$  is given by

$$\mathcal{E} = qV_g \cos(\omega t + \phi) \quad (2.1)$$

where the angular frequency  $\omega$  is related to the frequency  $f$  through the well-known relation  $\omega = 2\pi f$ , and  $\phi$  is a phase offset. The time-variation of the applied electric field adds complexity to particle transport. If inappropriately timed, the particle can experience an acceleration opposite to the desired direction (deceleration) or arrive at a phase that corresponds to a fractional gain in the maximally achievable energy gain from the electric field. In order to analyze these considerations we establish a framework where the system is designed for a specific particle called the design or synchronous particle. This fictitious particle moves downstream of the linac which we take to be the positive direction  $z$ . It will always remain on-axis with no transverse position  $x = y = 0$ . Here and in later sections, parameters referring to the design particle are subscripted  $s$ . Let the design particle have initial kinetic energy  $\mathcal{E}_{s,i}$ . The change in energy from an RF acceleration gap will then be

$$\mathcal{E}_s = \mathcal{E}_{s,i} + qV_g \cos(\omega t_s). \quad (2.2)$$

with  $\varphi = 0$  for the design particle. The field experienced by the particle will vary depending on the particle's speed. If the particle is slow relative to the RF period ( $\tau_{rf} = 1/f$ ), then the particle will experience a highly oscillatory field and the energy gain is equally likely to be negative or positive. On the flip-side, if the particle is relatively fast, then the field experienced will be similar to a DC field and the the energy gain can be approximated as

$$\mathcal{E} = \mathcal{E}_{s,i} + qV_g \cos(\omega t_s),$$

where  $t_s$  is the time at which the design particle arrives at the gap center. To quantify the time it takes for the particle to traverse an RF gap, a parameter called the transit time factor  $T$  can be used. A complete derivation is given in [4] and only the main steps are shown. The gain in kinetic energy  $\Delta\mathcal{E}$  for the design particle traversing an RF acceleration gap is given by integrating over the gap:

$$\Delta\mathcal{E} = \int_{\text{gap}} \vec{E}(\vec{r}, t) \cdot d\vec{\ell},$$

where  $\vec{E}$  represents the electric field and  $d\vec{\ell}$  is the path element taken by the particle. By definition, the design particle always remains on-axis and moves downstream of the accelerator.

Define the gap width to be  $g$  and the electric field to be of the form

$$\vec{E}(\vec{r}, t) = \vec{E}(r, z) \cos(\omega t + \varphi).$$

Regarding the design particle, the energy gain is then

$$\Delta\mathcal{E} = \int_{-g/2}^{g/2} E_z(r=0, z) \cos(\omega t + \varphi) dz.$$

To evaluate the integral one needs to define  $t$  in terms of  $z$ . The design particle only has velocity  $\vec{v}$  in the  $z$ -direction so that  $\vec{v} = v_z \hat{e}_z$  and the differentials are related by  $dt = dz/v_z(z)$ . Without loss of generality, the lower bound on the time integration can be set to 0 give  $t(z)$  as:

$$t(z) = \int_{z_1}^{z_2} \frac{dz}{v_z(z)}$$

where  $z_1$  and  $z_2$  are two arbitrary points. Additionally, if one defines the average field potential  $V_0$  as

$$V_0 \equiv \int_{-g/2}^{g/2} E(r=0, z) dz,$$

then the energy gain can be written in a deceptively simple equation known as Panofsky's equation:

$$\Delta \mathcal{E} = qV_0 T \cos \varphi. \quad (2.3)$$

The transit time factor  $T$  in Eq. 2.3 can be found by first rewriting the cosine term in terms of cosine and sine using the trigonometric identities(see [2, 4]). Then, multiplying and dividing by  $V_0$  gives

$$T \equiv \frac{\int_{-g/2}^{g/2} E_z(r=0, z) \cos(\omega t(z)) dz}{V_0} - \tan \varphi \frac{\int_{-g/2}^{g/2} E_z(r=0, z) \sin(\omega t(z)) dz}{V_0}. \quad (2.4)$$

The time-variation in the electric field will always reduce the maximum energy gain that can be obtained from a static field in the gap. This can be seen directly by comparing the numerators and denominators in Eq. 2.4. If  $T \approx 1$  then the particle will experience an approximately static field and a lower value of  $T$  implies more variation the particle experiences in the gap. In most cases the electric field profile is symmetric about  $z$  and the second term in Eq. 2.4 evaluates to zero. Furthermore, if the energy gain in the gap is

small compared to the particle's kinetic energy, then the velocity is roughly constant and the integral for  $t(z) \approx (z_2 - z_1)/v_z$ . With these two assumptions Eq. 2.4 reduces to

$$T = \frac{\int_{-g/2}^{g/2} E_z(r=0, z) \cos\left(\frac{2\pi z}{\beta\lambda_{rf}}\right) dz}{V_0}, \quad (2.5)$$

where  $\beta = v_z/c$  and  $\lambda_{rf} = c/f$ . Note that  $T$  is dependent on the width of the gap and the speed of the particle. This suggests that gaps be designed thin or use an applied voltage that is small compared to the particle's kinetic energy. A useful approximation is to assume the electric field is hard-edge so that it is of a constant value  $E_g$  within the gap and zero outside

$$E(r=0, z) = \begin{cases} E_g & -g/2 \leq z \leq g/2, \\ 0 & \text{otherwise.} \end{cases}$$

Computing the integral gives

$$T = \frac{\sin\left(\frac{\pi g}{\beta\lambda_{rf}}\right)}{\left(\frac{\pi g}{\beta\lambda_{rf}}\right)} = \text{sinc}\left(\frac{\pi g}{\beta\lambda_{rf}}\right). \quad (2.6)$$

With Eq. 2.6 one can quickly calculate and understand the interplay between the particle velocity and the acceleration gap geometry. This equation will be used in later sections to justify using a static approximation to the electric field within the gap.

## 2.2 RF Synchronism for Particle Acceleration

If  $T \approx 1$  then the energy gain of a particle progressing through the gap will only depend on the the phase of arrival  $\phi = \omega t$  and for simplicity we assume  $\varphi = 0$ . The distance from the

$(n - 1)^{\text{th}}$  gap center to the  $n^{\text{th}}$  gap center will be  $d_n$ . Before arriving, the design particle will have a velocity  $\beta_{s,n-1}c$  from the previous gap at a phase of  $\phi_{s,n-1}$ . At the next gap center, the particle will arrive at  $\phi_{s,n}$  based on the equation:

$$\phi_{s,n} = \phi_{s,n-1} + \frac{\omega d_n}{\beta_{s,n-1}c}. \quad (\text{design particle})$$

Similarly, for an arbitrary particle the equation is:

$$\phi_n = \phi_{n-1} + \frac{\omega d_n}{\beta_{n-1}c}. \quad (\text{arbitrary particle})$$

An RF gap can operate in two modes. If all RF gaps are in phase with each other then this is referred to as the O-mode (“oh”-mode). The RF gaps can also operate such that each successive gap is  $180^\circ$  out of phase relative to the previous gap and is called the  $\pi$ -mode (“pi”-mode) which will produce a more compact structure. The mode type can be incorporated into the previous equations by:

$$\begin{aligned} \phi_{s,n} &= \phi_{s,n-1} + \frac{\omega d_{n-1}}{\beta_{s,n-1}c} + \begin{cases} 0, & \text{O-Mode} \\ \pi, & \pi - \text{mode} \end{cases}, \\ \phi_n &= \phi_{n-1} + \frac{\omega d_{n-1}}{\beta_{n-1}c} + \begin{cases} 0, & \text{O-Mode} \\ \pi, & \pi - \text{mode} \end{cases}. \end{aligned} \quad (2.7)$$

The system this dissertation is concerned with uses the  $\pi$ -mode operation for a more compact structure. Therefore, from here on the  $\pi$ -mode convention will be used and case notation dropped. The synchronous particle, by definition, arrives at each gap at the synchronous

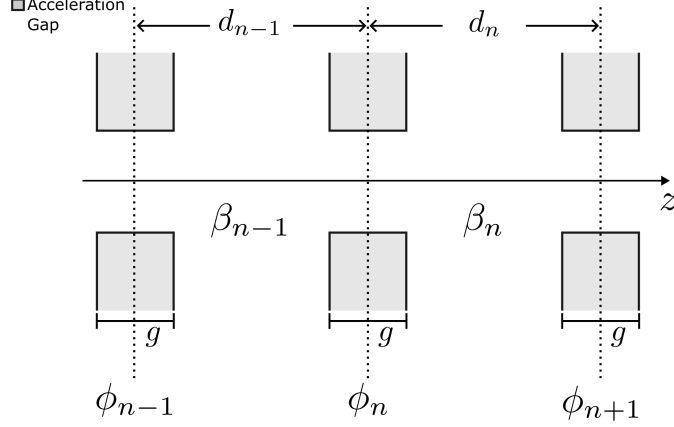


Figure 2.1: Schematic for accelerating cells in linac structure. The particle arrives at each gap center at a phase  $\phi_{n-1}$  and is accelerated which changes the particle velocity from  $\beta_{n-1}$  to  $\beta_n$ . The particle will then progress to the next acceleration center and arrive at a phase of  $\phi_n$ .

phase and thus the phase difference between each pair of gaps should remain  $\pi$ . This condition can be used to rewrite  $d_{n-1}$  in Eq. 2.7 for the arbitrary particle in terms of the design particle parameters. First, using the definition of the synchronous particle:

$$d_{n-1} = \frac{\beta_{s,n-1} c \pi}{\omega}$$

Substituting for  $d_{n-1}$  in Eq. 2.7 gives:

$$\phi_n = \phi_{n-1} + \pi \frac{\beta_{s,n-1}}{\beta_{n-1}} + \pi.$$

It is useful to find the relative phase difference as particles travel from gap to gap. This can be done by subtracting the synchronous phase equation from the arbitrary phase condition yielding

$$(\phi_n - \phi_{s,n}) - (\phi_{n-1} - \phi_{s,n-1}) = \pi \beta_{s,n-1} \left( \frac{1}{\beta_{n-1}} - \frac{1}{\beta_{s,n-1}} \right).$$

Next, we are more concerned with the particle energy than its velocity and a conversion from velocity to energy is usually done. First, the  $\beta$ -terms on the right can be manipulated to as follows:

$$\begin{aligned}\beta_{s,n-1} \left( \frac{1}{\beta_{n-1}} - \frac{1}{\beta_{s,n-1}} \right) &= - \left( 1 - \frac{\beta_{s,n-1}}{\beta_{n-1}} \right) \\ &= - \left( \frac{\beta_{n-1} - \beta_{s,n-1}}{\beta_{n-1}} \right) \\ &= \frac{\Delta\beta_{n-1}}{\beta_{s,n-1} + \Delta\beta_{n-1}}.\end{aligned}$$

Here,  $\Delta\xi = \xi - \xi_s$  denotes the relative difference between some measure  $\xi$  of an arbitrary particle and the design particle. Generally, the difference  $\Delta\beta_{n-1}$  will be small compared to  $\beta_{s,n-1}$  and the denominator in the last line of the equation above can be approximated as  $\beta_{s,n-1} + \Delta\beta_{n-1} \approx \beta_{s,n-1}$ . Using this approximation and substituting into Eq. 2.7 gives:

$$\Delta\phi_n - \Delta\phi_{n-1} \approx \pi \frac{\Delta\beta_{n-1}}{\beta_{s,n-1}}.$$

The  $\beta$ -terms can now be written in terms of the kinetic energy  $\mathcal{E} = (\gamma - 1)mc^2$  with  $\gamma = (1 - \beta^2)^{-1/2}$ . Using these identities and some algebraic manipulations one can show that  $\Delta\mathcal{E} = \gamma_s^3 \beta_s mc^2 \Delta\beta$ . The relative change in the gap-to-gap phase can be written in terms of the relative change in kinetic energy:

$$\Delta\phi_n - \Delta\phi_{n-1} = \frac{-\pi}{\gamma_{s,n-1}^3 \beta_{s,n-1}^2} \left( \frac{\Delta\mathcal{E}_{n-1}}{mc^2} \right) \quad (2.8)$$

A similar equation can be derived for the gap-to-gap kinetic energy difference  $\Delta\mathcal{E}_n - \Delta\mathcal{E}_{n-1}$ . To do this, use the Panofsky equation given by Eq. 2.3 and perform the same differencing



as was done here. Doing so results in

$$\Delta\mathcal{E}_n - \Delta\mathcal{E}_{n-1} = qV_{0,n}T_n(\beta_{s,n})[\cos\phi_n - \cos\phi_{s,n}]. \quad (2.9)$$

Together, Eq. 2.8 and Eq. 2.9 form a pair of coupled finite difference equations that can be used to model the relative change in phase, which can be converted to time with correct bookkeeping of the design particle parameters and energy. Additionally, the equations can be converted into coupled continuous differential equations under the assumption that the gap-to-gap changes are small. The steps to do so can be found in multiple places (in particular see file 09 in [5]). Making the conversion to differential form results in

$$(\gamma_s\beta_s)^3 \frac{d\Delta\phi}{dz} = \frac{-2\pi}{\lambda_{\text{rf}}} \left( \frac{\Delta\mathcal{E}}{mc^2} \right), \quad (2.10a)$$

$$\frac{d\Delta\mathcal{E}}{dz} = qE_0T(\cos\phi - \cos\phi_s), \quad (2.10b)$$

with the average electric field over the gap defined as  $E_0 = V_0/g$  being used in the second line. Taking another derivative of Eq. 2.10a and substituting  $\Delta\mathcal{E}/dz$ , results in a second order nonlinear differential equation for  $\Delta\phi$ :

$$\frac{d}{dz} \left[ (\gamma_s\beta_s)^3 \frac{d\Delta\phi}{dz} \right] = \frac{-2\pi qE_0T}{\lambda_{\text{rf}} mc^2} [\cos(\phi_s + \Delta\phi) - \cos\phi_s]. \quad (2.11)$$

Further analysis can be done by making two approximations. Assume that  $\Delta\phi$  is small ( $\Delta\phi \ll 1$ ) and the energy change is small  $d(\gamma_s\beta_s)/dz \approx 0$ . Making these approximations allows one to derive a conserved Hamiltonian  $H_\phi$ :

$$H_\phi = \frac{1}{2}(\Delta\phi')^2 + \frac{k_s^2}{2}(\Delta\phi)^2, \quad (2.12)$$

where the synchrotron wave number  $k_s$  is given by:

$$k_s = \sqrt{\frac{2\pi}{\lambda_{rf}} \frac{qE_0 T \sin(-\phi_s)}{mc^2(\gamma_s\beta_s)^3}}. \quad (2.13)$$

Note the sine argument under the square-root is  $-\phi_s$ . Here, the convention that the RF-accelerating field in the gap ranges from  $-\pi \leq \phi \leq \pi$  and that from  $-\pi \leq \phi < 0$  the electric field is rising. At  $\phi = 0$  the field is a maximum and then decreases to the minimum value at  $\phi = \pi$ . Eq. 2.12 immediately lends intuition about the dynamical behavior of a particle subject to the assumptions made (small  $\Delta\phi$  and small energy gain). Additionally, regions of stable oscillations can be identified by applying those same approximations to Eq. 2.11 in which the equation is reduced to:

$$\frac{d^2\Delta\phi}{dz^2} + k_s^2\Delta\phi = 0.$$

This is the standard harmonic oscillator equation and one can immediately see that amplitude oscillations around the design particle occur for  $k_s^2 > 0$  or  $-\pi < \phi_s < 0$  while the oscillations are unstable for  $0 < \phi_s < \pi$ . Armed with these insights, additional analysis can be done to include acceleration and the subsequent effects on the beam dynamics. Reverting back to equation Eq. 2.11 and only assuming what is needed for the continuous limit, the derivatives can be taken to give:

$$\frac{d^2\Delta\phi}{dz^2} + 3\frac{(\gamma_s\beta_s)'}{(\gamma_s\beta_s)}\frac{d\Delta\phi}{dz} = \frac{-2\pi}{\lambda_{rf}}\frac{qE_0T}{mc^2\gamma_s^3\beta_s^3}\left[\cos(\phi_s + \Delta\phi) - \cos(\phi_s)\right] \quad (2.14)$$

The synchrotron Hamiltonian can again be redefined with some clever algebra and by assuming the acceleration rate is small  $(\gamma_s\beta_s)' \approx 0$  (again see file 09 in [5]). With the following

variable definitions:

$$A = \frac{2\pi}{\lambda_{rf}(\gamma_s\beta_s)^3}, \quad B = \frac{qE_0T}{mc^2},$$

the non linear form of  $H_\phi$  with acceleration is then

$$H_\phi = \frac{A}{2} \left( \frac{\Delta\mathcal{E}}{mc^2} \right)^2 + B(\sin\phi - \phi \cos\phi_s). \quad (2.15)$$

Examining  $H_\phi$  we see that the first term represents the kinetic energy while the second term is the potential. This identification allows one to find stable regions for the potential energy by examining the first and second derivatives of  $V_\phi = B(\sin\phi - \phi \cos\phi_s)$ :

$$\begin{aligned} V_\phi &= B(\sin\phi - \phi \cos\phi_s), \\ \frac{\partial V_\phi}{\partial \phi} &= B(\cos\phi - \cos\phi_s), \\ \frac{\partial^2 V_\phi}{\partial \phi^2} &= -B \sin\phi. \end{aligned}$$

Upon examining the potential and its derivatives above, the following operating regions for  $\phi_s$  can be identified:

- $-\pi < \phi_s < 0$ : stable region
- $-\pi/2 < \phi_s < \pi/2$ : acceleration
- $-\pi/2 < \phi_s < 0$ : acceleration and longitudinal focusing
- $0 \leq \phi_s \leq \pi$ : unstable region

Two operating points for  $\phi_s$  are useful landmarks for navigating the operating region. The

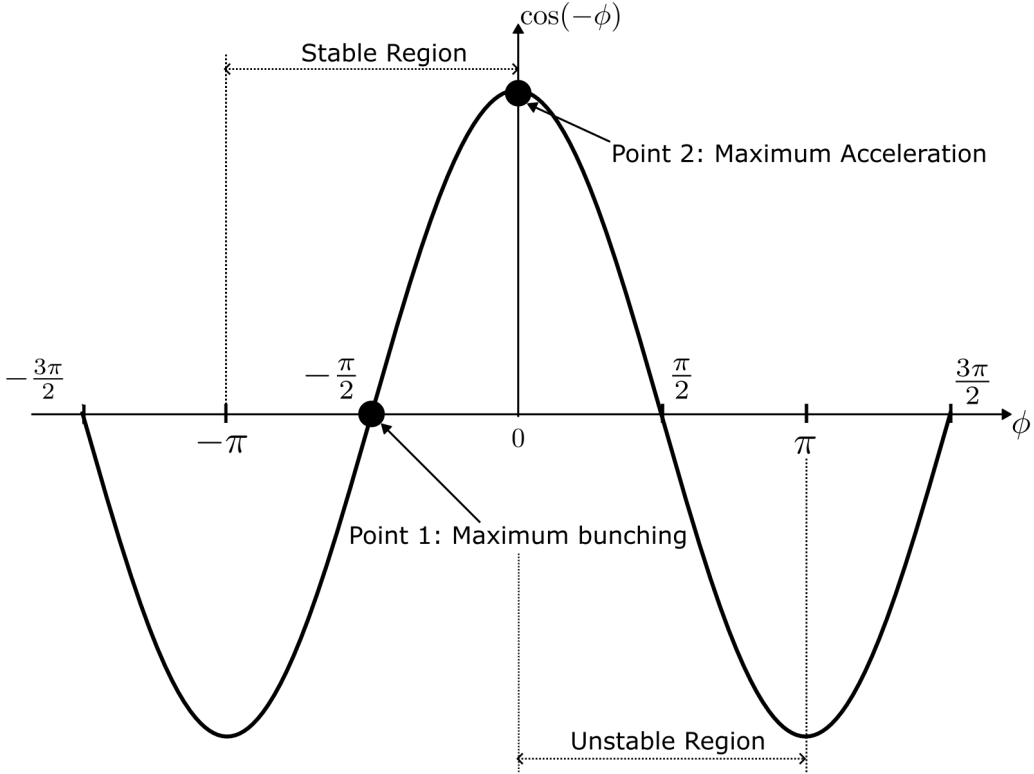


Figure 2.2: Operating phase convention typically used for the cosine oscillating electric field in the gap. The design phase is chosen to be between  $-\pi \leq \phi_s \leq \pi$ . Examination of the Hamiltonian  $H_\phi$  shows that the region between  $-\pi/2 < \phi < 0$  is a region of stability. Points 1 and 2 represent operation points for maximum bunching and maximum acceleration respectively. When operating at Point 1 with  $\phi_s = -\pi/2$  the maximum amount of particles will be captured in the RF-bucket and there will be no overall acceleration of the beam. At Point 2 with  $\phi_s = 0$  the design particle is maximally accelerated but the RF-bucket will be theoretically zero and the minimum amount of particles will be captured. Between these two points there will be both acceleration and bunching.

first point (Point-1 in Fig. 2.2) is when  $\phi_s = -\pi/2$  and is used for maximum focusing or bunching of the beam. To see this, consider three particles. The first particle will have slightly less energy than the design particle ( $\Delta\mathcal{E} < 0$ ) and will be trailing behind the design particle as the particles progress to the next gap ( $\Delta\phi > 0$ ). Call this slower particle particle-S. The second particle will have the opposite conditions and will be leading the design particle, this will be particle-F. Lastly, there is the design particle. At  $\phi_s = -\pi/2$ , the design particle will not be accelerated. However, particle-F will arrive before the design

particle and see a decelerating field. This will cause the design particle to pass particle-F. On the flip-side, particle-S will arrive late and see an accelerating field. Particle-S will gain energy and surpass the design particle. Particle-F and particle-S have now swapped positions relative to the design particle. Extrapolating this logic to a multitude of particles, this operating phase will prevent particle losses by bunching the slow and fast particles around the design particle. The other landmark is when the design particle is maximally accelerated at  $\phi_s = 0$  (Point-2 in Fig. 2.2). Applying the same logic, both particle-F and particle-S will see a lower magnitude electric field than the design particle. The design particle will surpass all particles and  $\Delta\phi_s$  will tend to negative infinity for all particles.

The conservation of  $H_\phi$  implies that the phase-dynamics of particles will evolve on nested closed-contours within a bounding contour that partitions stable phase-space evolution from unstable. This bounding contour is called the separatrix in accelerator jargon and can be found by solving for the stable fixed points in phase-space  $\phi_1$  and  $\phi_2$  where  $\Delta\mathcal{E} = 0$ . Evaluating the first  $\partial V_\phi/\partial\phi = 0$  will give the maximum potential and thus  $\Delta\mathcal{E} = 0$ . This first point can then be used to find the constant value of  $H_\phi$  and then, upon substitution into Eq. 2.15 one has the following equation for  $\phi_2$ :

$$\frac{A}{2} \left( \frac{\Delta\mathcal{E}}{mc^2} \right)^2 + B[\sin \phi_2 - \phi_2 \cos \phi_s] = -B[\sin \phi_s] - \phi_s \cos \phi_s.$$

With  $\phi_2$  and  $\phi_s$  the separatrix is defined and the stable region within the separatrix is referred to as the RF-bucket. The stable region defined by the separatrix will form a fish-shape structure as depicted in the central plot on the left panel in Fig. 2.3. The two stable points defining the separatrix are shown as dashed-lines intersecting the  $\phi - axis$  at  $\phi_2$  and  $-\phi_s$ . To the right is the shape taken when  $\phi_s = -\pi/2$  or maximum bunching as discussed

before. Here, the contours make an eye-shape and this is the maximum width bucket that can be achieved with a full range phase-acceptance of  $2\pi$ . Lastly, if the acceleration term

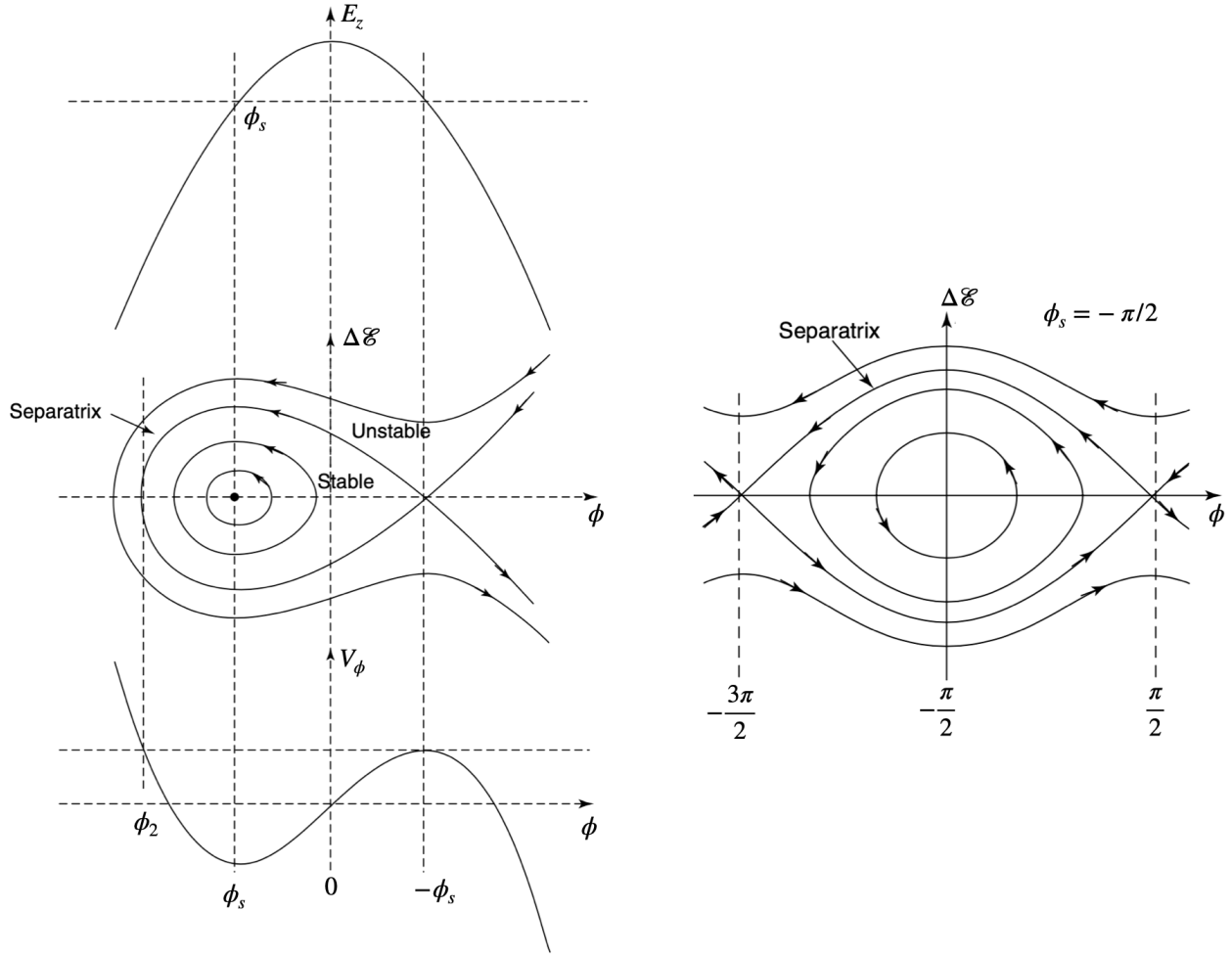


Figure 2.3: Image borrowed from Wangler [2]. The left panel shows three plots as a function of  $\phi$ . The top plot is the electric field in the gap. The middle plot is the  $\Delta\mathcal{E}$ - $\phi$  phase-space contour plot. Lastly, the bottom plot is the potential term  $V_\phi$  in the Hamiltonian  $H_\phi$ . The phase-space plot labels the region of stable orbits, or RF-bucket, that reside within the outer most stable contour called the separatrix. To find the separatrix one must locate two fixed stable points  $-\phi_s$  and  $\phi_2$ . If operating at maximum bunching  $\phi_s = -\pi/2$  then the phase-space contours form an eye shape as depicted by the right plot. This operating phase has the largest RF-bucket.

in Eq. 2.14 is not neglected then there is damping in the equation and  $H_\phi$  is no longer conserved. When this term is present, the phase-space plot looks more like a golf club structure as shown in Fig. 2.4. Additionally, when  $H_\phi$  is no longer conserved the separatrix

can no longer be identified. Though the RF-bucket is no longer defined, the analysis done for the non-acceleration case can be employed to give a prescription for optimizing transport.

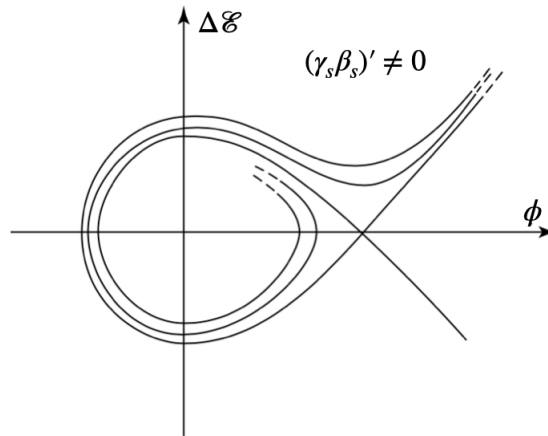


Figure 2.4: Image borrowed from Wangler [2]. Acceleration in the system will introduce damping to the phase-space orbits and the fish structure in Fig. 2.4 morphs into a golf club structure.

## 2.3 Summary

This chapter laid out the conceptual ideas for longitudinal dynamics in a linac system. In doing so, important parameters, as well as the technical jargon, were defined. Additionally, the conventions used here and later are to be emphasized. The concepts described here will be a useful guide for navigating the challenging operational landscape we are to face (see Ch. 6). Lastly, this material is by no means complete. More thorough discussion can be found in many texts with [2, 5, 6] being particularly good.

# Chapter 3. Transverse Dynamics

Transverse particle dynamics focuses on the behavior of particles in the transverse plane. It encompasses the study of particle trajectories, beam dynamics in accelerators, and the principles governing the focusing and defocusing of particle beams. In the realm of particle accelerators, transverse particle dynamics plays a crucial role in the design, operation, and optimization of these complex machines. Understanding the transverse dynamics of particles is essential for controlling and manipulating their trajectories within accelerators. Defocusing effects arise from interactions with various electromagnetic fields in the system such as the inter-particle (space-charge) forces. These forces drive particles off-axis resulting in a degraded beam and possible system damage from charged particles impinging on the system walls. It is essential to comprehend these effects to achieve control over the beam's characteristics such as its intensity, size, and shape. Transverse focusing is critical to ensuring that particles remain confined within the system. Focusing elements, such as quadrupole magnets and electromagnetic lenses, are employed to control the beam divergence thus allowing for the concentration of particles to maintain a tight and collimated beam. The primary focusing elements the system in this thesis discusses are the electro static quadrupoles (ESQs) and RF focusing provided by the acceleration gaps.

This chapter will review the theory of transverse beam dynamics that is necessary for understanding the models and results used in later sections. It is by no means a comprehensive review. For a more thorough treatment of the subject matter there are many fine textbooks [2, 6] as well as open access course notes provided by the United States Particle Accelerator School (USPAS) [7] and the CERN document server [8]. The chapter is organized as follows.



We start by examining the equations of motion in the transverse plane for a single particle in applied and self electric and magnetic fields. Assumptions are made to calculate both the applied fields and the self-fields in ideal cases. The single-particle case will yield a set of second order non-linear differential equations that can be used to model the particle motion in the presence of applied fields and a charged particle beam. To describe a system of particles, a statistical approach is built from the single particle approach.

### 3.1 Single Particle Dynamics

To understand the transverse dynamics of a beam it is convenient to first consider a single charged particle of mass  $m$  and charge  $q$ . This particle is moving in the presence of electric and magnetic fields for which there are two sources. The first source comes from the applied fields that arise from the various conducting objects in the system used to accelerate, focus, and steer the beam. These fields can be denoted  $\vec{E}^a$  and  $\vec{B}^a$  for the applied electric and magnetic fields respectively. The second source comes from the surrounding beam that the single particle under consideration is moving with. This beam is comprised of other charged particles that each produce an electric and magnetic field. Solving each individual particle's field is intractable. In order to progress with the analysis, various assumptions are made and the culmination of charges are assigned a distribution. This distribution of charge can be analyzed and the self-field can be calculated from the designated distribution. These self-fields can be denoted  $\vec{E}^s$  and  $\vec{B}^s$  for the self-electric and self-magnetic field respectively. The two sources can be added together so that the individual particle moves in a total electric field  $\vec{E} = \vec{E}^a + \vec{E}^s$  and a total magnetic field  $\vec{B} = \vec{B}^a + \vec{B}^s$ . Denoting the particle's momentum as  $\vec{p}$  and the speed of light as  $c$ , the particle's dynamics are given by the Lorentz force

equation:

$$\begin{aligned}
\frac{d\vec{p}}{dt} &= q[\vec{E}(\vec{r}, t) + c\vec{\beta} \times \vec{B}(\vec{r}, t)], \\
&= q[\vec{E}^a(\vec{r}, t) + c\vec{\beta} \times \vec{B}^a(\vec{r}, t)] + q[\vec{E}^s(\vec{r}, t) + c\vec{\beta} \times \vec{B}^s(\vec{r}, t)], \\
&= \vec{F}^a(\vec{r}, t) + \vec{F}^s(\vec{r}, t),
\end{aligned}$$

where the applied and self-field forces are given by  $\vec{F}^a(\vec{r}, t)$  and  $\vec{F}^s(\vec{r}, t)$  respectively. The particles velocity is given in terms of  $\beta$  where  $\vec{v} = \beta c$ . The particle momentum can be written as

$$\vec{p} = \gamma mc\vec{\beta},$$

where  $\gamma = (1 - 1/\beta^2)^{-1/2}$ . Thus far, no approximations have been made. The first will be the so called paraxial approximation. In this approximation the beam, and therefore the particle being considered, is assumed to have small transverse velocities relative to the axial velocity—axial meaning along the accelerator system. Symbolically this means  $v_{\perp} \ll v_z$  where the  $\perp$  symbol stands for the perpendicular components which in this case is  $x$  and  $y$  components. Additionally, it is more informative to transform the time derivative into a derivative along the beam line also known as Frenet-Serret coordinates. Note that, in the case of a Linac system, the beam line coordinate is the same as  $z$ . Denoting  $' \equiv d/dz$  and analyzing the transverse components only, the momentum can be rewritten as:

$$p_{\perp} = \gamma m v_z x'_{\perp}.$$

Note that the time derivative becomes  $d/dt = v_z d/dz$  due to the paraxial approximation.

The time derivative of the transverse momentum in the new coordinates is:

$$mv_z \frac{d(\gamma v_z x'_\perp)}{dz} = \gamma m v_z^2 x''_\perp + m v_z (\gamma v_z)' x'_\perp.$$

Rewriting in terms of  $\beta = v_z/c$  and including the force terms gives:

$$x''_\perp + \frac{(\gamma\beta)'}{(\gamma\beta)} x'_\perp = \frac{1}{\gamma m \beta^2 c^2} (F_\perp^a + F_\perp^s). \quad (3.1)$$

To progress, information is needed about both the applied and self fields. A reasonable assumption is to assume that the beam is comprised of a uniform distribution of particles that is cylindrically symmetric. To calculate the self-field, Gauss' law can be applied. For a cylinder of uniform charge with radius  $r_b$ , beam current  $I_b$ , and beam velocity  $v_b$ , Gauss' law gives<sup>1</sup>:

$$E_r^2 = \begin{cases} \frac{I_b r}{2\pi\epsilon_0 r_b^2 v_b}, & r \leq r_b \\ \frac{I_b}{2\pi\epsilon_0 v_b r}, & r > r_b, \end{cases} \quad (3.2)$$

where  $\epsilon_0$  is the vacuum permittivity. For a cold beam the particles' longitudinal velocity spread is small and thus  $v_b \approx v_z$ . Under the assumption that the charge is uniformly distributed, it follows that the current density is of the form  $\vec{J}^s \approx \beta_b c \rho(x_\perp, t) \hat{z}$  with  $\beta_b = v_b/c$ . The self-magnetic field is obtained through applying Ampère's law:

$$\begin{cases} B_\theta^s = \frac{\mu_0 I_b r}{2\pi r_b^2}, & r \leq r_b \\ B_\theta^s = \frac{\mu_0 I_b}{2\pi r}, & r > r_b, \end{cases} \quad (3.3)$$

---

<sup>1</sup>The derivation here closely follows steps taken in other sources [6, 9, 10, 11] are used in this section

where  $\mu_0$  is the vacuum permeability. At this point, it is convenient to convert the equations of motion into cylindrical coordinates where  $r$  is the radius and  $\theta$  is the angle in  $x$ - $y$  plane. Denoting the time derivative  $\dot{\chi} = d\chi/dt$ , the radial component of the Lorentz force equation gives:

$$\frac{d(\gamma_b m \dot{r})}{dt} - \gamma_b m r \dot{\theta}^2 = q(E_r^s + r \dot{\theta} B_z^s - \dot{z} B_\theta^s).$$

A few assumptions can be made. First, since the applied fields are not being considered then there is no acceleration and  $\gamma \approx \text{const}$ . Second, we take the beam to be slowly rotating such that  $r \dot{\theta}$  can be assumed negligible. These assumptions reduce the above equation to

$$\gamma m \ddot{r} = q E_r^s - q \dot{z} B_\theta^s.$$

Using  $c^2 = 1/(\epsilon_0 \mu_0)$  the self-fields can be shown equal

$$E_r^s - \beta_b c B_\theta^s = \frac{I_b r}{2\pi \epsilon_0 \gamma_b^2 v_b r_b^2} \frac{r}{r_b^2}.$$

Additionally, transforming the time derivative to a spatial one gives  $\ddot{r} = \beta_b^2 c^2 r''$ . Combining and simplifying gives

$$r'' = \frac{q I_b}{2\pi \epsilon_0 m \gamma_b^3 \beta_b^3 c^3} \frac{r}{r_b^2} \quad (3.4)$$

The right-hand-side of Eq. 3.4 can be wrapped into a parameter named the generalized perveance  $Q$ :

$$Q = \frac{q I_b}{2\pi \epsilon_0 m \gamma_b^3 \beta_b^3 c^3}. \quad (3.5)$$

Rewriting Eq. 3.4 in terms of  $Q$  gives:

$$r'' = \frac{Q}{r_b^2} r.$$

Since  $Q > 0$  for ions it will act to expand an ion beam. Because it is only dependent on a handful of constants, the beam current, and beam energy,  $Q$  is a useful parameter in describing the space-charge forces present. The  $r_b^2$  term is a geometry factor and thus left out of  $Q$  as it will change depending on the beam shape taken.

In addition to the paraxial approximation it was assumed that the particles were uniformly distributed and cylindrically symmetric. It was further assumed that the beam is slowly-rotating so that the term  $r\dot{\theta}$  was negligible. These assumptions and subsequent derivations are common when analyzing different beam distributions. As an example of a beam that fulfills these approximations we will review a uniform density elliptical beam next but will have to reanalyze since we no longer have cylindrical symmetry.

### 3.1.1 Calculating the Self-Fields for an Elliptical Beam with Uniformly Distributed Charge Density

The previous section provides a recipe for calculating the self-field of the beam. Fig. 3.1 gives a schematic of the an elliptical beam. The beam has edge lengths of  $r_x$  and  $r_y$  with a uniform number density of particles given by  $n(z)$ . Given the area of the ellipse as  $\pi r_x(z)r_y(z)$ , the line-charge density is a constant given by  $\lambda = qn(z)\pi r_x(z)r_y(z)$ . A derivation of the self-field

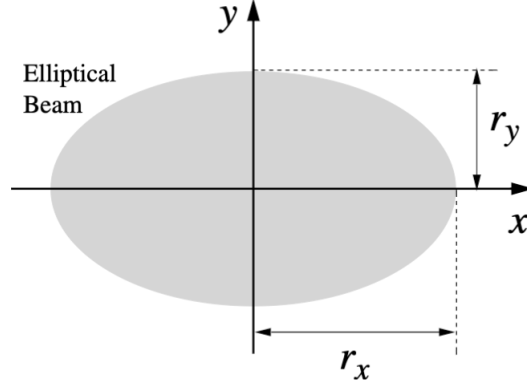


Figure 3.1: An elliptical beam of uniformly distributed particles within the ellipse. There is a constant current density  $\vec{J} = \rho\beta_b c \hat{z}$ . The beam radii are given by  $r_x$  and  $r_y$  as shown. Image borrowed from [12].

potential is given in [13] and borrowing the equation, the self-potential is<sup>2</sup>:

$$\phi^s = -\frac{\lambda}{2\pi\epsilon_0} \left[ \frac{x^2}{(r_x + r_y)r_x} + \frac{y^2}{(r_x + r_y)r_y} \right]. \quad (3.6)$$

Taking the negative gradient gives the transverse components of the electric field for the self-field:

$$E_x^s = \frac{\lambda}{\pi\epsilon_0} \frac{x}{(r_x + r_y)r_x}, \quad (3.7)$$

$$E_y^s = \frac{\lambda}{\pi\epsilon_0} \frac{y}{(r_x + r_y)r_y}. \quad (3.8)$$

---

<sup>2</sup>At the time of writing, the URL provided in the references also has a video of the USPAS lecture given by Dr. Lund. For the especially curious this video is helpful in explaining where the non-trivial potential comes from.

Repeating the steps used for the cylindrical beam, the self-fields can be plugged into the Eq. 3.1 resulting in:

$$x''(z) = \frac{q\lambda}{\pi\epsilon_0 m \gamma_b^3 \beta_b^2 c^2} \frac{x}{(r_x + r_y)r_x},$$

$$y''(z) = \frac{q\lambda}{\pi\epsilon_0 m \gamma_b^3 \beta_b^2 c^2} \frac{y}{(r_x + r_y)r_y}.$$

In the cylindrical case,  $Q$  was defined and its usefulness comes from only depending on the beam characteristics  $I_b$  and  $\mathcal{E}_b$ . Thus,  $Q$  should be present in the elliptical case and indeed it is. Noting that  $\lambda = I_b/(\beta_b c)$  the dynamical equations can be rewritten in terms of  $Q$  as:

$$x''(z) = \frac{2Q}{(r_x + r_y)r_x} x,$$

$$y''(z) = \frac{2Q}{(r_x + r_y)r_y} y. \tag{3.9}$$

The right-hand-side of Eq. 3.9 is the contribution from the self-fields of an elliptical beam and will be used in following sections.

### 3.1.2 Calculating the Forces from Applied ESQ Fields

We begin by considering a purely transverse electric field in vacuum. Following the steps in [11], the purely transverse potential can be expanded as a complex series since it satisfies the Cauchy-Riemann conditions. Using  $\underline{z} = (x + iy)$  and denoting the complex potential  $\underline{P}(\underline{z})$  the series is written as:

$$\underline{P} = \sum_{n=0}^{\infty} C_n \underline{z}^n,$$

where the coefficients are, in general, complex numbers such that  $\underline{C}_n = (a_n + ib_n)$ . The complex conjugate of the electric field  $\underline{E}^* \equiv E_x - iE_y$  is compliant with Maxwell's equations. Taking the derivative of the potential gives the electric field:

$$\underline{E}^* = E_x - iE_y = i \frac{\partial P}{\partial \underline{z}}.$$

Taking derivative of the series gives:

$$i \frac{\partial P}{\partial \underline{z}} = i \sum_{n=1}^{\infty} n \underline{C}_n \underline{z}^{n-1} = i \underline{C}_1 + 2i \underline{C}_2 \underline{z} + 3i \underline{C}_3 \underline{z}^2 + \dots$$

The first coefficient can be found by looking at the field at  $\underline{z} = 0$ ,

$$\underline{C}_1 = -i \underline{E}^*(\underline{z} = 0).$$

Equating the real and complex coefficient parts to the field components gives:

$$a_1 = -E_y(x=0, y=0); \quad b_1 = -E_x(x=0, y=0).$$

To find the rest of the components, successive derivatives can be taken and evaluated at  $\underline{z} = 0$ . The linear component of the field comes from expanding  $\underline{z}$  and then taking the derivative with respect to  $x$  and  $y$ . Expanding  $\underline{z}$  first gives:

$$E_x - iE_y = (ia_1 - b_1) + 2(ia_2 - b_2)(x + iy) + 3(ia_3 - b_3)(x + iy)^2.$$



From the expansion,  $a_2$  and  $b_2$  can be seen to be:

$$a_2 = \frac{1}{2} \frac{\partial E_y}{\partial y} \Big|_{x=0, y=0}; \quad b_2 = -\frac{1}{2} \frac{\partial E_x}{\partial x} \Big|_{x=0, y=0}.$$

Lastly, it is common to normalize the fields so that they are independent of energy. The factor  $[B\rho]$  (“Brho” or rigidity) is common and defined as:

$$\begin{aligned} [B\rho]_{\text{magnetic}} &\equiv \frac{\gamma_b \beta_b m c}{q}, \quad \text{for magnetic fields} \\ [B\rho]_{\text{electric}} &\equiv \frac{\gamma_b \beta_b^2 m c^2}{q} = \beta_b c [B\rho]_{\text{magnetic}}, \quad \text{for electric fields.} \end{aligned}$$

Since we are concerned with electric fields, we will drop the subscripts and use  $[B\rho]$  for electric fields. For a pure quadrupole field, the dipole term coefficients  $a_1$  and  $b_1$  will be zero and the leading term will be the  $a_2$  and  $b_2$  coefficients. This will give the quadrupole term and the corresponding normalized transverse fields will be:

$$\begin{aligned} \frac{E_x}{\beta_b c [B\rho]} &= -\frac{q}{\gamma_b \beta_b^2 m c^2} \frac{\partial E}{\partial x} \Big|_{x=0, y=0} x, \\ \frac{E_y}{\beta_b c [B\rho]} &= \frac{q}{\gamma_b \beta_b^2 m c^2} \frac{\partial E}{\partial y} \Big|_{x=0, y=0} y. \end{aligned}$$

For a perfect quadrupole, the gradients are equal and opposite. Denoting the gradient  $G$  as  $G = -\partial E_x / \partial x = \partial E_y / \partial y$ , the linear applied fields ( $E_x^a$  and  $E_y^a$ ) from a perfect pure

transverse quadrupole are

$$\begin{aligned}\frac{E_x^a}{\beta_b c [B\rho]} &= -\frac{qG(z)|_{x=0,y=0}}{\gamma_b \beta_b^2 mc^2} x = -\kappa(z)x, \\ \frac{E_y^a}{\beta_b c [B\rho]} &= \frac{qG|_{x=0,y=0}}{\gamma_b \beta_b^2 mc^2} x = \kappa(z)y,\end{aligned}\tag{3.10}$$

where the focusing function  $\kappa(z)$  has been introduced and is defined as:

$$\kappa(z) = \frac{qG|_{x=0,y=0}}{\gamma_b \beta_b^2 mc^2}\tag{3.11}$$

for a purely transverse ESQ.

### 3.1.3 The Kapchinskij-Vladimirskij (KV) Envelope Equations

Eq. 3.1 is the general form for the equations of motion with the applied forces  $F^a$  and self-field forces  $F^s$ . Using Eq. 3.10 and Eq. 3.9 the full single particle equation of motion can be written for the transverse directions:

$$\begin{aligned}x'' + \frac{(\gamma_b \beta_b)'}{(\gamma_b \beta_b)} x' + \kappa(z)x - \frac{2Q}{(r_x + r_y)r_x} x &= 0, \\ y'' + \frac{(\gamma_b \beta_b)'}{(\gamma_b \beta_b)} y' - \kappa(z)y - \frac{2Q}{(r_x + r_y)r_y} y &= 0.\end{aligned}\tag{3.12}$$

Recall that the particle is assumed to be moving within an elliptical beam of uniform particle density with major axes  $r_x$  and  $r_y$ . Additionally, the ESQ focusing is from a pure quadrupole with purely transverse fields. Note that the focusing is equal and opposite in the two planes so that while one plane is being focused by the ESQ the other plane is defocused. In general, the focusing can be from another device with higher-order terms and the equations of motion

can be rewritten as

$$\begin{aligned} x'' + \frac{(\gamma_b \beta_b)'}{(\gamma_b \beta_b)} x' + \kappa_x(z)x - \frac{2Q}{(r_x + r_y)r_x} x &= 0, \\ y'' + \frac{(\gamma_b \beta_b)'}{(\gamma_b \beta_b)} y' + \kappa_y(z)y - \frac{2Q}{(r_x + r_y)r_y} y &= 0. \end{aligned} \quad (3.13)$$

Additional insight can be gleaned by analyzing Eq. 3.13 for different cases.

**Case A:** Assume a perfect quadrupole such that  $\kappa_y = -\kappa_x$  and  $\kappa_x = \kappa(z)$ . Assume no acceleration  $\gamma_b \beta_b \approx \text{const.}$  and that space-charge effects are negligible  $Q \approx 0$ . Lastly, take  $\kappa(z)$  to be periodic such that, for a lattice period  $L_p$ ,  $\kappa(z + L_p) = \kappa(z)$ .

With these assumptions, the particle equations of motion take the form of Hill's equations [14]:

$$\begin{aligned} x(z)'' + \kappa(z)x(z) &= 0, \\ y(z)'' - \kappa(z)y(z) &= 0. \end{aligned}$$

Hill's equation can be solved using a phase-amplitude formulation [6, 9, 11]. We assume a solution of the form  $x(z) = A(z) \cos \psi(z)$  which we then plug into Hill's equation to give:

$$x'' + \kappa(z)x = [A'' + \kappa A - A\psi'^2] \cos \psi - [2A'\psi' + A\psi''] \sin \psi = 0.$$

Since we have used two functions,  $A(z)$  and  $\psi(z)$ , to describe the single function  $x(z)$ , we are free to introduce a constraint. We will choose the coefficient for  $\sin \psi$  to be 0 giving the

first constraint:

$$2A'\psi' + A\psi'' = 0. \quad \text{Constraint 1}$$

This, in turn, leads to a second constraint for the  $\cos \psi$  coefficient in order to solve Hill's equation:

$$A'' + \kappa A - A\psi^2 = 0. \quad \text{Constraint 2}$$

Using Constraint 1, one can show that  $A^2\psi' = \text{const}$ . We then rescale  $A(z)$  such that  $A(z) = A_i w(z)$  where  $A_i$  is the initial amplitude and  $w(z)$  is a real-valued function. Constraint 1 then gives  $w^2\psi' \equiv 1$ . Rewriting  $x(z)$  in terms of these functions gives:

$$\begin{aligned} x(z) &= A_i w(z) \cos \psi(z), \\ x'(z) &= A_i w'(z) \cos \psi(z) - \frac{A_i}{w(z)} \sin \psi(z), \end{aligned} \quad (3.14)$$

with a similar solution for  $y(z)$ . Here  $A_i$  is the initial amplitude,  $w(z)$  is a real-valued function, and  $\psi(z)$  is a real-valued phase function. A useful dimensionless parameter called the undepressed phase-advance  $\sigma_0$  can be calculated by integrating Constraint 1 over a single lattice period  $L_p$ :

$$\sigma_0 \equiv \Delta\psi(z) = \int_{z_i}^{z_i+L_p} \frac{d\tilde{z}}{w^2(\tilde{z})} \quad (3.15)$$

where  $\Delta\psi(z) = \psi(z) - \psi_i(z_i)$  and  $z_i$  is the starting  $z$ -location. One can intuitively understand  $\sigma_0$  by taking what is called a smooth approximation to  $\kappa(z)$ . In this approximation,  $\kappa(z)$  is “smeared” out such that it is a constant value over an arbitrarily defined  $L_p$ . In this case, the solution to Hill's equations are the common harmonic solutions and  $\kappa(z)$  is redefined as the undepressed betatron wavenumber  $\kappa_{\beta 0}$  which is constant ( $\kappa(z) = \kappa_{\beta 0}^2 = \text{const}$ ). Then  $\sigma_0 = \kappa_{\beta 0} L_p$ . It should be stressed that  $\sigma_0$  only depends on the lattice properties and

not the initial conditions of the particle. It can be shown that the single-particle orbits are stable when  $\sigma_0 < 180^\circ/\text{period}$ . The usefulness of  $\sigma_0$  is put succinctly in Dr. Lund's lecture notes [9]:

The phase advance  $\sigma_0$  is an extremely useful dimensionless measure to characterize the focusing strength of a periodic lattice. Much of conventional accelerator physics centers on focusing strength and the suppression of resonance effects. The phase advance is a natural parameter to employ in many situations to allow ready interpretation of results in a generalizable manner.

**Case B:** We keep the same assumptions as Case A except we include space-charge so that  $Q \neq 0$ .

Incorporating the space-charge shifts  $\kappa_{\beta 0}$  resulting in the depressed betatron wavenumber  $\kappa_\beta^2$  and depressed phase-advance  $\sigma$ :

$$\kappa_\beta^2 = \left( \kappa_{\beta 0} - \frac{2Q}{(r_x + r_y)r_x} \right)^2.$$

Space-charge will act to dampen the transverse oscillations. One can quantify the degree of space charge in system by calculated  $\sigma/\sigma_0$ . A value close to zero represents a space-charge dominated beam (cold beam) while a value close to 1 represents a thermally dominated beam.

The phase-amplitude formulation can be used to form the Courant-Snyder (CS) invariant [13, 15]. Recall that  $A_i$  was the initial amplitude and some manipulation of Eq. 3.14 will yield the invariant:

$$\left( \frac{x}{w} \right)^2 + (wx' - w'x)^2 = A_i^2. \tag{3.16}$$

From the CS invariant we can see that the particle's trajectory in  $x - x'$  phase space is

described by a set of nested ellipses. The CS invariant is commonly rescaled by defining the so called Twiss Parameters:

$$\begin{aligned}
 \alpha(z) &\equiv -w(z)w'(z), \\
 \beta(z) &\equiv w^2(z), \\
 \gamma(z) &\equiv \frac{1}{w^2(z)} = \frac{1 + \alpha(z)^2}{\beta(z)^2}.
 \end{aligned}
 \tag{3.17}$$

Rewriting Eq. 3.16 in terms of the Twiss Parameters gives:

$$\gamma x^2 + 2\alpha x x' + \beta x'^2 = A_i^2.
 \tag{3.18}$$

The Twiss Parameters are functions of the lattice only and are thus incredibly useful in characterizing a lattice. One can integrate over the ellipse to get the phase-space area obtaining:

$$\begin{aligned}
 \int_{\text{ellipse}} dx dx' &= \pi A^2, \\
 A_i^2 &\equiv \epsilon,
 \end{aligned}$$

where the last line defines the single particle emittance  $\epsilon$ . Thus, the single particle phase-space area is a conserved quantity and is widely used in accelerator physics. Fig. 3.2 is an drawing of the  $x - x'$  ellipse with various points of interest labeled in terms of the Twiss Parameters. In terms of  $\epsilon$ , the solutions to Hill's equation is

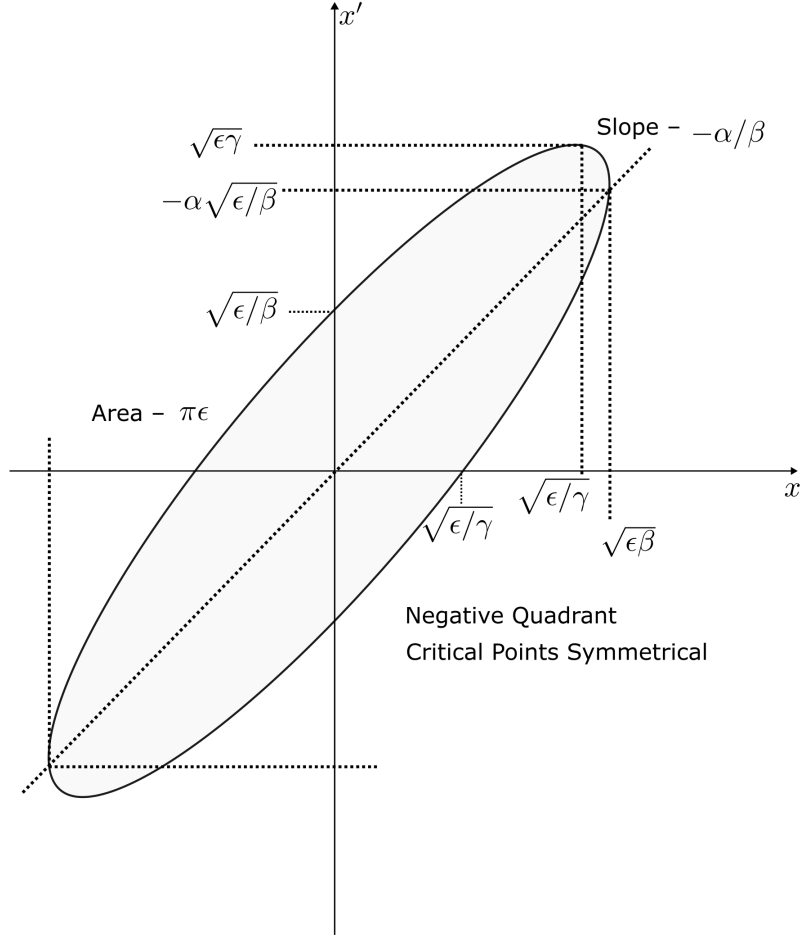


Figure 3.2: Phase space ellipse in  $x - x'$  with points of interest labeled in terms of the Twiss Parameters  $\alpha$ ,  $\beta$ , and  $\gamma$ . Figure adopted from [9]

$$x(z) = \sqrt{\epsilon_x} w_x(z) \cos \psi(z),$$

$$y(z) = \sqrt{\epsilon_y} w_y(z) \cos \psi(z).$$

The boundary in  $x$  and  $y$  can be found by setting  $\cos \psi(z) = \pm 1$  in the solution which gives the beam edge as:

$$r_x(z) = \sqrt{\epsilon_x} w_x(z),$$

$$r_y(z) = \sqrt{\epsilon_y} w_y(z).$$

Rewriting Hill's equations in terms of  $r_x$  and  $r_y$  results in a set of coupled, second-order, nonlinear differential equations that describe the evolution of the envelope edge known as the Kapchinskij-Vladimirskij (KV) envelope equations [16]. If the acceleration is included (see [9] for transformation needed to include acceleration in KV equations), then the full equations for linear applied fields are

$$\begin{aligned} r_x''(z) + \frac{(\gamma_b\beta_b)'}{(\gamma_b\beta_b)}r_x + \kappa_x(z)r_x(z) - \frac{2Q}{r_x(z) + r_y(z)} - \frac{\epsilon_x^2}{r_x^3(z)} &= 0, \\ r_y''(z) + \frac{(\gamma_b\beta_b)'}{(\gamma_b\beta_b)}r_y + \kappa_y(z)r_y(z) - \frac{2Q}{r_x(z) + r_y(z)} - \frac{\epsilon_y^2}{r_y^3(z)} &= 0. \end{aligned} \quad (3.19)$$

A statistical approach can be accomplished and statistical moments for the beam calculated if one were to find a distribution that satisfies the underlying assumptions: uniform density of an elliptical beam. Such a distribution exists and is called the KV distribution [16] after the same Kapchinskij and Vladimirskij. The KV distribution is the only solution to the time-dependent Vlasov equation (see [6, 9]) when the applied and self-fields are linear, and the transverse and longitudinal equation of motions are decoupled. Since it is a solution to Vlasov's equation, the 4D transverse phase-area is conserved, i.e., the transverse emittance.

The details of the KV distribution (its derivations and physical interpretations) are fascinating but fairly complex. We therefore neglect this discussion as it would take up some significant page length with no immediate benefit to this dissertation. However, complete lecture notes and a video discussion can be found in the space-charge course taught at USPAS [9] or in a review paper authored by one of the course lecturers [17]. Using the KV distribution, the statistical root-mean-square (rms) envelope edges and angles can be



calculated from the statistical moments giving:

$$\begin{aligned}
 r_x &= 2\sqrt{\langle x^2 \rangle_{\perp}}, & r_y &= 2\sqrt{\langle y^2 \rangle_{\perp}}, \\
 r'_x &= 2\frac{\sqrt{\langle xx' \rangle_{\perp}}}{\sqrt{\langle x^2 \rangle_{\perp}}}, & r'_y &= 2\frac{\sqrt{\langle yy' \rangle_{\perp}}}{\sqrt{\langle y^2 \rangle_{\perp}}},
 \end{aligned}
 \tag{3.20}$$

where  $\langle \dots \rangle_{\perp}$  denotes a transverse statistical average over the beam phase-space area. Lastly, the statistical rms-edge emittance is given by:

$$\epsilon_{\text{rms}} = 4[\langle x^2 \rangle \langle x'^2 \rangle - \langle xx' \rangle^2]^{1/2}.
 \tag{3.21}$$

Generally the emittance will not be conserved in an acceleration lattice and the acceleration gaps will dampen the emittance. Furthermore, we cannot hope to achieve a KV distribution as it is a 4D hyper-ellipsoid with particles existing only at the beam's transverse edges. The issue is stated succinctly by Reiser [6]:

The K-V theory is a good and very useful approximation for beams where the current remains well below the space-charge limit. This is true for practically all accelerators and other devices [...] It must be kept in mind, however, that the KV model does not include nonlinear effects that increase the emittance.

This prompts the question: how can we possibly model a real accelerator system with equations that assume such a distribution and that will change the emittance due to intrinsic non-linear effects? Lapostolle [18] and Sachere [19] developed the concept of equivalent beams (discussed in [6]) which allows one to study the KV equivalent of a laboratory beam of the same species, current, and kinetic energy if second moments of the distribution are the same. Furthermore, this approximate equivalence can be further justified if the emittance stays approximately the same over the lattice. To counteract the dampening affects

of acceleration,  $\epsilon_{\text{rms}}$  is rescaled to give the normalized rms-edge emittance  $\epsilon_{n,\text{rms}}$ :

$$\epsilon_{n,\text{rms}} \equiv \gamma_b \beta_b \epsilon_{\text{rms}}. \quad (3.22)$$

## 3.2 Summary

This chapter presented a narrow overview of the key concepts and equations that are of importance to the work that will be presented in Part II and especially Ch. 7. Derivations for the transverse equations of motion in the presence of linear applied and self field were sketched. Key parameters such as the generalized perveance  $Q$  and rms-edge emittance  $\epsilon_{\text{rms}}$  were defined. Of particular importance are the KV envelope equations which will be the primary equations used in Ch. 7 to model the desired accelerator system in the presence of accelerating gaps and ESQ focusing. We will also show that although the beam does not follow a KV distribution, we can form an equivalent beam and further justify this approach by computing  $\epsilon_{n,\text{rms}}$  and checking that it remains roughly constant.

# Chapter 4. Self Consistent Simulations

Warp [20, 21] is a particle-in-cell (PIC) code that was developed to model high current, high brightness beams for heavy-ion driven inertial confinement fusion (HIF). However, the software is not limited to HIF modeling. Warp has extensive capabilities and can self-consistently model various accelerator systems. Systems can be time-varying or static with various options for the transverse and/or longitudinal boundary conditions. In a nutshell, the software will first perform a field-solve on the user-defined mesh for all the conductor objects at subgrid resolution. Particles, either single species or mixed, will then be injected and deposited on the mesh based on user settings. Once deposited, the self-field is calculated by finding a global solution to Poisson's equation (including conductors) and the particles are advanced using a 2<sup>nd</sup> order leap-frog algorithm. A variety of symmetry options are available to reduce run-times and options exist to implement adaptive mesh refinement (AMR) for more efficient simulations by increasing grid resolution in regions where finer detail is needed while keeping less interesting areas at a grainer resolution to save computational time. Warp also has pre-defined conductor objects as well as the capability to create conductors through its conductor building tool-kit. The back-end of the code is written in Fortran but is wrapped in a Python interface for easier use. A typical script outline is given in pseudo-code in Lst. 4.1.

```
1  import warp as wp
2  # Mesh Creation and time step
3  wp.w3d.zmin = minimum z
4  wp.w3d.zmax = maximum z
5  wp.w3d.nz = number of grid points to use
6  wp.top.dt = time step
```

```

7   # ... same setup for x and y
8
9   # Set boundary conditions
10  wp.bound0 = boundary condition at minimum z (dirchlet, neumann,
    reflect)
11  wp.boundnz = boundary condition at maximum z (dirchlet, neumann,
    reflect)
12  wp.boundxy = transverse boundary conditions (dirichlet, neumann,
    periodic,)
13
14  # Can also set particle boundary conditions
15  wp.top.pbound0 = ...
16  wp.top.pboundnz = ...
17
18  # Beam Specifications
19  beam = wp.Species(type=wp.Argon, charge_state=1, name="Ar+")
20  beam.ekin = beam kinetic energy
21  beam.ibeam = current of beam
22  ... # More specifications available like beam radius and
23  ... # angle, emittance, etc.
24
25  # Specify injection. The number of particles injected, particle weight
    ,
26  # injection type, beam distribution of injection and
27  # much more can be specified.
28
29
30  # Create and install conductors. Once

```

```

31 conductor = wp.Annulus(...)
32 wp.installconductors(conductor)
33
34 # Specify solver type. RZ, four-fold, two-fold, or
35 # full 3D options available.
36 solver = wp.MRBlock3D() # full 3D symmetry
37 solver.ldosolve = True # Solve for self-fields, i.e. include space-
charge
38 wp.register(solver)
39
40 # Generate the PIC code, allocate storage for arrays,
41 # and do initial field solve
42 wp.package("w3d")
43 wp.generate()
44
45 # Advance particles. Depending on the settings, diagnostics will be
collected
46 # each time step and plots can be generated in a cgm file.
47 for i in range(desired_time_steps):
48     wp.step() # One time step advance
49     ...
50
51 # Do analysis. Various outputs are specified in the top.v file for a
multitude of different calculations Warp performs based on the user-
setting for the information to be gathered.
52 do analysis of simulation

```

Listing 4.1: A prototypical example of a Warp script in Python syntax.

Additionally, Warp has multiple diagnostic tools and will store statistical moment histories if the options are set to do so. In particular, the `ZCrossingParticles` tool will be used in the data collection in later sections (see Ch. 8). This diagnostic acts as a beam position monitor (BPM) and will record particle data (transverse position, time of crossing, velocities, momentum, etc.) at the given  $z$ -location for particles crossing it.

As invaluable as Warp is, it can be computationally expensive and, as will be discussed later, using Warp exclusively to optimize the system this dissertation is concerned with is intractable. Thus, Warp acts as the last step in the design and optimization before testing in real-experiment. Fig. 4.1 gives an ideal flow of simulation to experiment. Reduced models focusing on distinct aspects of the system are first developed and used to probe system parameters and design. Optimization cases can be developed quickly and multiple analyses performed. This then acts as a guide to use Warp for more sophisticated treatment. This initial stage is done repeatedly as signified by the small double arrow icon. Once there is established confidence in design, the experiment can be done and compared. Since experiments are even more difficult and time consuming, this cycle is repeated less as signified by the more open double linking arrows. Ideally, this whole process is repeated until convergence is reached. This dissertation will be focused on the first stage in the cycle, using reduced models to probe the parameter/design space which then guides simulations in Warp.

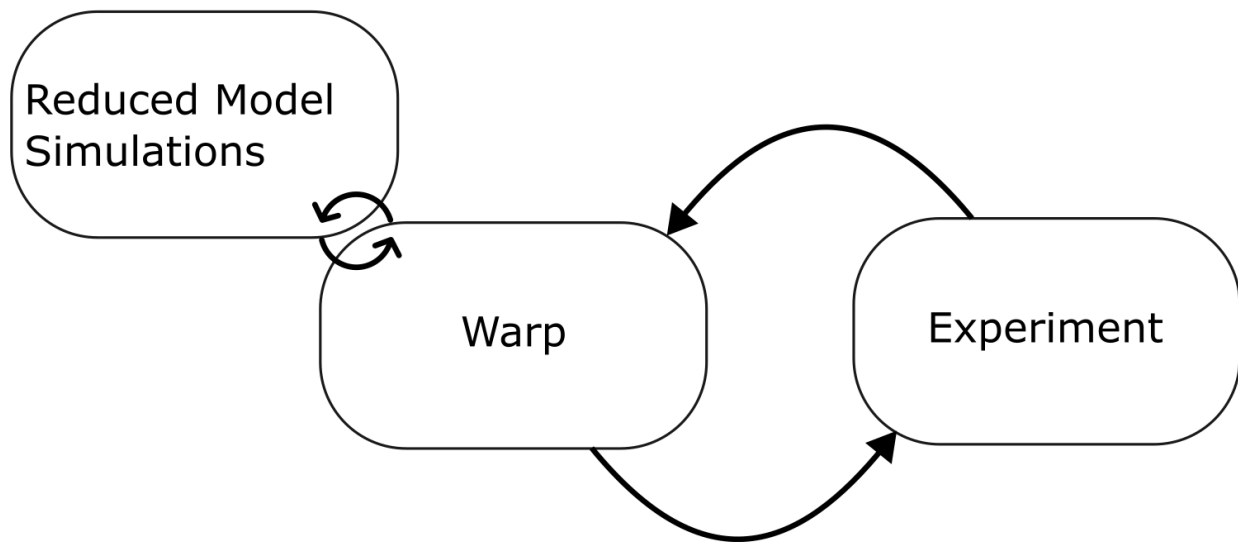


Figure 4.1: Ideal flow of simulation and experiment. Reduced models are designed and used to rapidly probe the parameters and design space. This narrows the scope so that the more costly simulations can be ran with Warp for more sophisticated treatment and analysis. This first stage is repeated numerous times as signified by the tight double arrow. Once convergence has been met at this stage, the design can be experimented on. Because real experiment is more challenging, this second stage is repeated less frequently as signified by the loose double arrow.

## Part II

# Compact Multi-Beam RF Linear Ion Accelerators



# Chapter 5. MEMS-RF System

## 5.1 Introduction

The accelerator system this part of the dissertation is concerned is the RF-linac. RF-linacs are pervasive in industrial and medical applications. Within the realm of application there are multitudes of required performance needs. For example, medical applications for cancer therapy using linac accelerators require ions to be of the order of 10 MeV and focused on a spot size that is on the order of  $1 \text{ cm}^2$  [22]. Ion implantation is another common use of RF-linac for which the range of energy and currents vary. Technologies require highly stable and collimated beams with currents ranging from a few  $\mu\text{A}$  to 100 mA and energies from 100 eV to  $\sim 10 \text{ MeV}$  [23]. Additionally, accelerator technology is used in food sterilization, nondestructive testing and imaging, and propulsion systems for low mass  $\sim 1 \text{ kg}$  satellite systems. With this in mind it is immediately obvious the benefit of a compact and cost-effective RF-linac device with some variability in the delivered current and energy. Reducing cost and size of a linac machine (or any complex system) has gained momentum due to advances in the field of micro electromechanical systems (MEMS). MEMS technology has revolutionized many fields and accelerator technologies are an excellent arena for application. The development of compact and cost effective RF-linac using MEMS technology (MEMS-RF) to accelerate an array of ion beams is the focus of Part II of this dissertation.

Part II is organized as follows. This chapter will introduce the MEMS-RF system. An overview of the system will be given with various references for material about the system that do not fit the scope here. The acceleration gaps and the ESQs will be discussed along with

their associated fields. We will discuss the work done to optimize the ESQ geometries such that the leading order error in the quadrupole field expansion is minimized. Once we have an optimum geometry we investigate the ESQs further to quantify other useful properties for design such as the voltage breakdown limit and spacing properties. With a firm grasp of the overall system architecture and individual conductors, we proceed by discussing reduced models used to efficiently simulate the system. This discussion begins with the longitudinal motion in Ch. 6. In this chapter, the computational tools developed and used to model the longitudinal phase-space is the primary focus along with how this model was benchmarked using Warp. We then move on to discussing the transverse motion in Ch. 7 and the reduced models used to simulate the physics therein. We will show how the developed model is in great agreement with Warp and will subsequently describe a prescription for how to design a lattice with improved transverse focusing to radially confine the beam. Additionally, the model can be used to quantify descriptive metrics such as the phase advance  $\sigma$  and various measures of the beam envelope. Lastly, we combine the previous chapters to perform 3D self consistent simulations with Warp in Ch. 8. Due to the complex nature of the accelerator system, the resulting self consistent simulations are themselves complex and we begin this chapter by discussing the simulation itself. We then proceed gently by performing simulations in a short lattice structure comprised of four acceleration gaps and a pulsed beam injection. This allows us to capture and observe much of the physics involved that is otherwise obscured by longer lattice structures and continuous injection. Finally, we perform a simulation of a 16 gap design with continuous injection and discuss results. Each chapter ends with section that summarized the discussion and also briefly describes areas of improvement and/or further research avenues.

## 5.2 Overview of Compact Multi-Beam MEMS Linac Structures

Although a linac is a technologically complex machine, its basic construction is relatively simple. A schematic of the MEMS-RF device is shown in Fig. 5.1 [24, 25, 26]. At the left

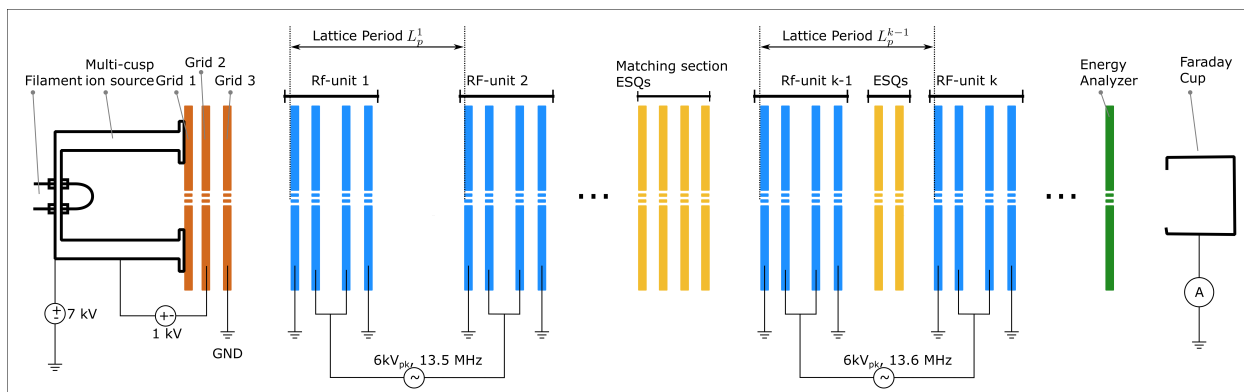


Figure 5.1: Basic Linac Design. Note that a single RF-unit is comprised of two RF acceleration gaps. Two single RF-units are connected to the same voltage source but multiple voltage sources will allow for frequency changes down the accelerator line. After the matching section, the Linac can be decomposed into a periodic lattice consisting of two RF-units.

of the schematic is the ion source. The source is housed in an insulated “cage” and floated to range of 7–10 kV. A collection of ion gas is then heated by a filament source to create a plasma which is extracted to produce an array of beams (beamlets) of energetic ions at about 7 keV and a current of  $\sim 10 \mu\text{A}$  per beamlet. A distinguishing feature of the MEMS-RF system is that an array of beamlets are accelerated compared to the acceleration of a single beam in conventional linac devices. The injection is continuous (CW) unless stated otherwise. Fig. 5.2 shows a close up image of an accelerating wafer with a 120 beamlet array. Following the source are RF-units which are comprised of two RF-acceleration gaps operating  $180^\circ$  out of phase relative to one another. Recall that this design choice is for longitudinal compactness of the system. Note that this first section of RF-gaps does not have periodic

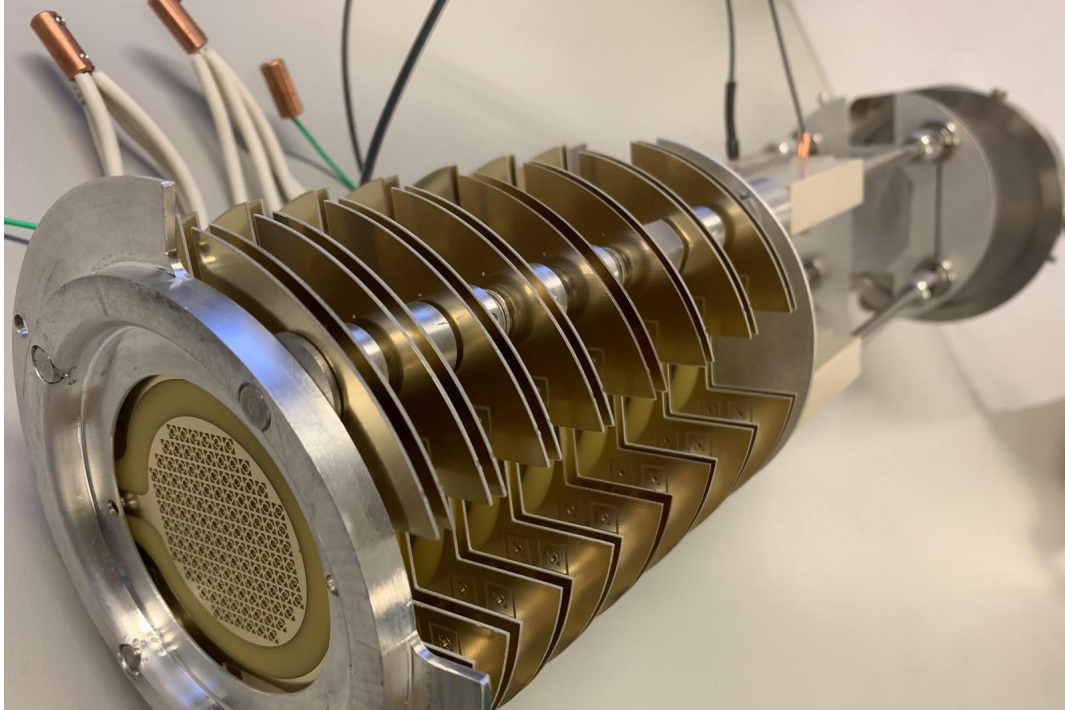


Figure 5.2: An ESQ wafer for a 120 beamlet array design. The 8 acceleration gaps are mounted on a four rods that hold the wafers in place. This design offers an easy way of adjusting the spacing between wafers to manipulate the RF phasing by changing the distance from one gap center to another.

doublet focusing but acceleration gaps after the matching section do. This unorthodox design is chosen because the acceleration gaps provide strong transverse focusing and can be used immediately following extract. This design will be further explained in the Ch. 7. The acceleration gap width is about 2 mm in length with the individual gap wafers being about 700  $\mu\text{m}$  thick. A typical accelerating voltage is 7 kV but is operation dependent and can be within the range of 5–7 kV. Accelerating voltage is applied with a sinusoidal variation at a frequency of 13.6 MHz.

After the matching section are RF-units with periodic doublet focusing that lies in a field-free region where ESQs can be inserted. The schematic shows a doublet configuration but if there is space and the need, more ESQs can be inserted or “stacked”. The Focusing strength depends on the beam energy and the transverse dynamics. Therefore  $V_q$  will likely be

different in different lattice periods. Breakdown voltage is estimated to be at approximately 1 kV and limits the operating voltage to be within the range of a few hundred volts. With a thickness of about  $0.695\ \mu\text{m}$  the ESQs are thin compared to conventional ESQs. Lastly, in between the two acceleration portions of the system is a matching section that will match the beam from the initial phase where there is acceleration and no ESQ focusing to the second phase where both are present. The schematic shows four ESQs being used but more can be inserted. The whole system is in vacuum with a pressure of  $\sim 3\ \text{mTorr}$ . At the end of the acceleration lattice is a collimating slit. After passing through the slit, the beam enters a deflector plate that acts as an energy analyzer deflecting ions of a certain energy into a faraday cup.

Table 7.1 collects some of the key parameters per beamlet for the system which are organized into three groups: Source, Geometric, and Operational. Source parameters are initial values of the individual beamlets after extraction. The geometric parameters are values of various components of the system such as the length of the acceleration gap and the radius of the clear bore aperture of each beamlet. Note that the lattice period is for the first period of the acceleration lattice and is generally not a fixed quantity. The listed value is for the initial kinetic energy of 7 keV and maximum acceleration in the first two gaps. Operational quantities relate to the system settings for acceleration and focusing. Also note that the gap voltage  $V_g$ , ESQ focusing voltage  $V_q$ , and RF-frequency  $f_{\text{rf}}$  are not fixed quantities in experiment since they are operation dependent as mentioned earlier. The MEMS-RF system can be separated into two project areas. One area focuses on the application of the RF power amplifier developed by Airity technologies [27]. Employing these RF power amplifiers to deliver consistent and stable accelerating voltages without damaging the circuitry is challenging. Another area focuses on the analysis and optimization of the

Initial/Source Values per Beamlet	
Ion Species	Ar <sup>+</sup>
Beam Energy $\mathcal{E}_{b,i}$	7.0 keV
Beam Current $I_b$	10 $\mu$ A
Temperature $T_b$	0.1 eV
RMS Edge Emittance $\epsilon_{\text{rms}}$	1.34 mm-mrad
Geometric Design Values	
Source Aperture Radius $r_s$	0.25 mm
Gap Aperture Radius $r_p$	0.55 mm
Gap Width $g$	2.0 mm
ESQ Rod Radius $R$	0.717 mm
ESQ Length $\ell_q$	0.695 mm
Lattice Period $L_p$	$\sim 20$ mm
Operational	
Gap Voltage $V_g$	$\sim 6$ kV
ESQ Voltage $V_q$	$\sim 100$ V
RF Frequency $f_{\text{rf}}$	$\sim 13.6$ MHz
RF Period $\tau_{\text{rf}}$	$\sim 73.5$ ns
RF Wavelength $\lambda_{\text{rf}}$	$\sim 22.04$ m

Table 5.1: Standard parameter values for the MEMS-RF system. The initial or source values are per beamlet. Values shown as  $\sim \#$  have some degree of variability due to operation conditions or changes in beamlet parameters— $L_p$  will change with increasing  $\mathcal{E}_b$ . The  $R$  listed is the optimal value found in Sect. 5.4.

accelerated beam along with the accelerating and focusing components (RF acceleration gaps and focusing ESQs). It is this latter area that this dissertation is directed toward.

### 5.3 Fields from RF Acceleration Arrays

The acceleration gaps are created by first starting with an FR-4 printed circuit board(PCB) with copper cladding on both sides. Laser micro-machining is then used to pattern the PCB and drill the aperture holes (see [28] for more details). Fig. 5.3 is an example schematic show the array of acceleration apertures and a close up of an individual gap. The whole PCB is excluded from simplicity. Conducting material (in this case copper) is shown in grey and

the FR-4 material shown in blue. In black is the vacuum aperture where the beam will be. The FR-4 material is not simulated and instead the blue area is treated as vacuum. The

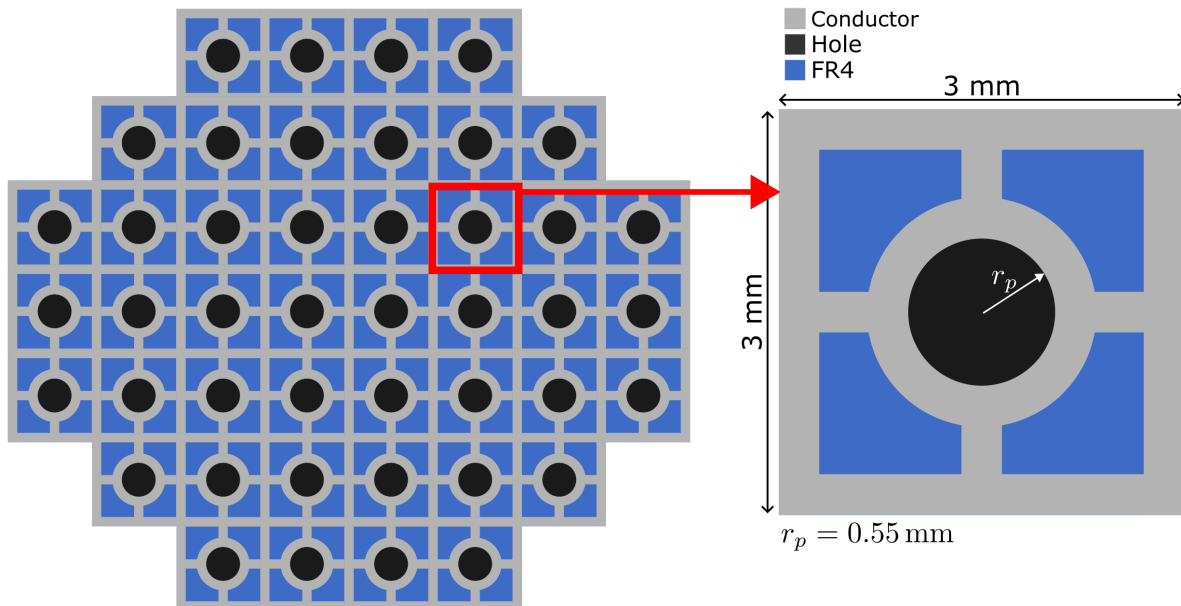


Figure 5.3: Front view of the array of gaps (left) and a close-up of a single gap (right). The conducting material, in our case copper, is shown in grey. In blue is the FR-4 material that the PCB is comprised of. The PCB is copper cladded on both sides and the pattern and aperture hole (shown in black) created using laser cutting. Image adapted from [28].

conducting ring has thickness of 0.3 mm. The connecting prongs have a length of 0.5 mm and width of 0.3 mm. Lastly, the surrounding square has a thickness of 0.5 mm. With the dimensions quantified the conductor can be created in Warp and the field analyzed. The gap geometry is relatively simple and there was not a priority to investigate the gap properties for optimization. Thus, these values can be considered fix quantities unless stated otherwise.

### 5.3.1 Electric Field Profile of the RF-Gap

In later sections the extracted on-axis electric field will be used to advance particles in a dimensionally reduced simulation that can be quickly ran and used to gain insight into the longitudinal and transverse dynamics. It is crucial that periodic boundary condi-

tions are used to simulate the array geometry of the gaps. Fig. 5.4 shows the Warp-

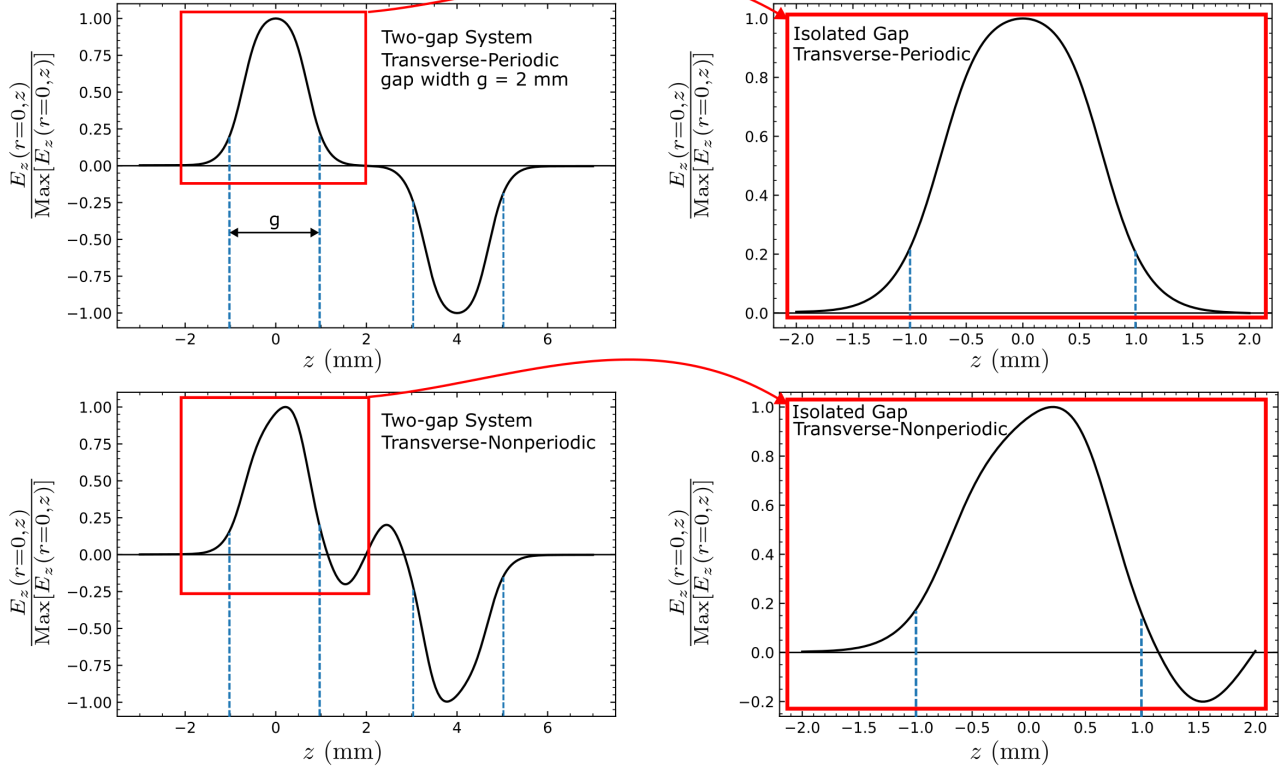


Figure 5.4: Extracted On-axis electric field  $E_z(r = 0, z)$  using Warp for transverse-periodic boundary conditions (top row) and Dirichlet-boundary conditions (bottom row). Two RF-gaps are needed to correctly simulate the transverse periodic-array geometry. If a single gap is used, then the fields will look like the bottom row. The gap width  $g$  is labeled as blue vertical dash lines. The field data can be isolated by extracting the data within the red box.

extracted  $z$ -component of the on-axis electric field normalized by the maximum amplitude ( $E_z(r = 0, z)/\text{Max}[E_z(r = 0, z)]$ ) for two gaps that are  $180^\circ$  out of phase. The top row shows the profile when the transverse conditions are set to periodic. The row below is when the transverse conditions are set to Dirichlet instead (setting Dirichlet or Neumann does not make a difference). It is apparent that these two profiles are drastically different and the choice will significantly impact the longitudinal dynamics, as well as the acceleration, of the particle. Periodic is the correct choice when modeling the beam *within* the array, such as that shown in Fig. 5.3. However, the non-periodic selection shows the field profile of the cells



that surround the array. This suggests that it would be wise to restrict beam acceleration to cells within the array and use the outer arrays as dummy cells that still have an applied voltage but no beam current. Doing so will ensure a desirable profile. An analysis was done to see how many layers are needed to surround a cell so that the field is periodic. Fortunately, only one surrounding layer is needed. The top right image of Fig. 5.4 shows a sample extraction of  $E_z(r = 0, z)$  in the periodic array. In later sections (see Ch. 6) the extracted field is used to advance particles in a 1D simulation—that is, all particles are assumed on-axis and particles can only move in  $\pm z$  along the accelerator axis. The normalized field can be scaled linearly to represent any applied voltage. Particles closer to the conductors (further off-axis) will generally experience a stronger magnitude field. This difference was less than a percent for approximately 80% of the aperture from the axis center. Further from the center than that the difference was only approximately 5%. Therefore, we can safely ignore the added complexity of off-axis particles experience a difference in field strength and assume the field-strength is uniform different  $r$ .

In addition to choosing the correct boundary conditions, to properly model the periodicity and extract the correct electric profile, a two-gap system must be simulated. Without the second gap, the periodicity is not captured does and the field profile is incorrect. Therefore, in order to extract a single gap profile, two gaps, are simulated. The extraction can then be done by properly isolating either field. This is done by extracting the upstream field from  $[z_1, z_2]$  where  $z_1$  is the start of the mesh boundary and  $z_2$  is the geometric center of the two gaps. If the first gap is centered on  $z = 0$  then  $z_1 = -z_2$ . With an isolated field a relationship can be established between the applied voltage and peak voltage and electric field that results. Assuming a linear relationship, the applied voltage  $V_a$  will be related to the on-axis peak voltage  $V_p(r = 0)$  by  $V_a = \xi V_p(r = 0)$  where  $\alpha$  is some scaling parameter.

Similarly, the applied voltage and peak electric field  $\text{Max}[E_z(r = 0, z)] = E_{z,p}(r = 0)$  will be linearly related by  $V_a = \eta E_{z,p}(r = 0)$ . Both will contain the points  $(0, 0)$ . Choosing any voltage, executing Warp, and then extracting the peak values yields the relationships:

$$V_a = 0.9962V_p(r = 0), \quad (5.1)$$

$$V_a = 656.879E_{z,p}(r = 0).$$

These relations can be used on the extracted normalized field to determine the scale factor for a desired voltage. Note that the on-axis voltage is slightly lower than the applied voltage and therefore, to get a desired energy gain a slightly higher voltage will have to be used. This does not account for non-linear effects that may be present in the system, such as conductor defects, sparking, etc.

## 5.4 Fields from Optimized Electrostatic Quadrupole (ESQ) Focusing Arrays

As the beam is transported through the system thermal expansion, space-charge forces, and transverse forces from the acceleration gap will cause the beam to expand. Beam expansion will reduce the beam intensity and lead to particle losses as the beam expands and begins to scrape the system walls. To prevent these effects, ESQ focusing is added in the field free region of the gaps. The ESQs here are comprised of cylindrical rods with rod radius  $R$  and longitudinal extent  $\ell_q$ . Arbitrary choice of  $R$  may result in significant contributions from the leading order error term in the multipole expansion giving undesirable nonlinear focusing. In this section we discuss the development and application of a computational tool for modeling

and optimizing the design for an array of ESQs to minimize the leading order error term in the multipole field expansion. This section is organized as follows. In Sec. 5.4.1, the theoretical and numerical background for calculating the multipole coefficients is discussed. Sec. 5.4.2 uses the computational tool to reproduce a known result to first benchmark the algorithm. We then discuss various geometrical settings and the impact on the field quality. We also discuss the voltage breakdown and spacing considerations to prevent field overlap between the doublet ESQs.

### 5.4.1 Numerical Methods

We seek to minimize the leading order error term  $A_6$  in the multipole expansion of the electric field by choosing the optimized geometrical configurations for the ESQ arrays. Fig. 5.5 shows a face-on view of a typical ESQ design comprised of cylindrical rods. These rods have radius  $R$ , axial length  $\ell_q$ , and applied voltage  $V_q$ . The beam passes through a circular clear-bore aperture of radius  $r_p$  shown as a dotted circle in the figure.

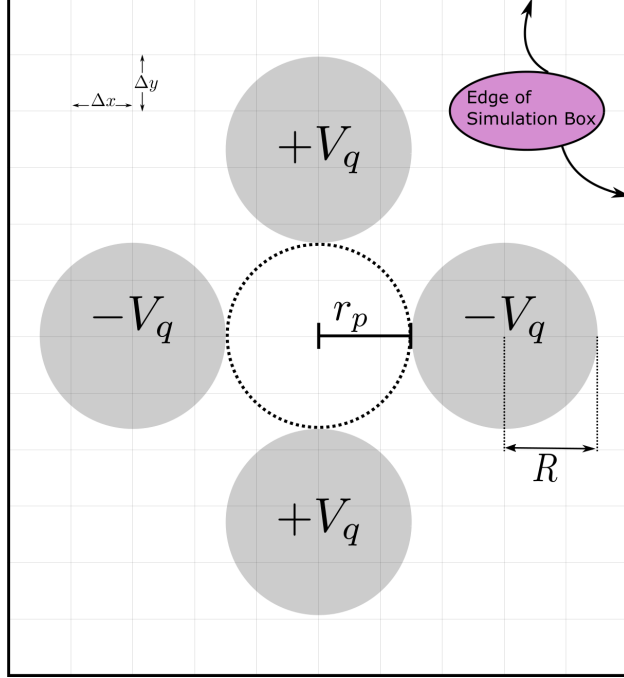


Figure 5.5: Transverse view of a single electrostatic quadrupole in a simulation box with grid sizes  $\Delta x = \Delta y$ . The conducting rods are shaded in grey with the applied voltage  $V_q$ . The rods have a radius of  $R$  and the dashed line shows the clear bore aperture with radius  $r_p$ . Conducting objects are shaded grey with all white space representing vacuum.

The short axial length  $\ell_q$  of the ESQ arrays, relative to  $r_p$ , gives rise to a significant axial fringe field. We used Warp to model an ESQ made of solid cylinders with  $r_p = 0.55$  mm,  $\ell_q = 1.264r_p$ , and rod radius  $R = 1.14r_p$  to extract the 3D electric field  $\vec{E}(x, y, z)$ . The electric field gradient  $G = \partial E_x / \partial x$  is calculated one grid cell off-axis to minimize nonlinear field content. Fig. 5.6 shows the electric field gradient. Denoting the max gradient at the geometric center of the quadrupole as  $G^* = G(z = 0)$ , we define the effective length  $\ell_{\text{eff}}$  as:

$$\int_{z_1}^{z_2} G dz = G^* \ell_{\text{eff}}. \quad (5.2)$$

Here,  $z_1$  and  $z_2$  are sufficiently large to capture the axial extent of the field upstream and downstream of the ESQ. As illustrated in Fig. 5.6, this system has  $\ell_{\text{eff}} = 1.306$  mm, an 88% increase from  $\ell_q$ . The result  $\ell_{\text{eff}} > \ell_q$  shows that this design will provide more focusing

than expected based on  $\ell_q$ . We parametrically evaluated how the physical length of the quadrupole relative to the aperture radius impacts the effective length (see Fig. 5.7). For  $\ell_q/r_p < 3$ ,  $\ell_{\text{eff}}$  can be  $\sim 40\%$  or more greater than  $\ell_q$ . For  $\ell_q/r_p > 4$ ,  $\ell_{\text{eff}}$  approaches  $\ell_q$ . This gives some guidance on the trade offs between creating a more compact system by shortening  $\ell_q$  and the amount of focusing that can be achieved.

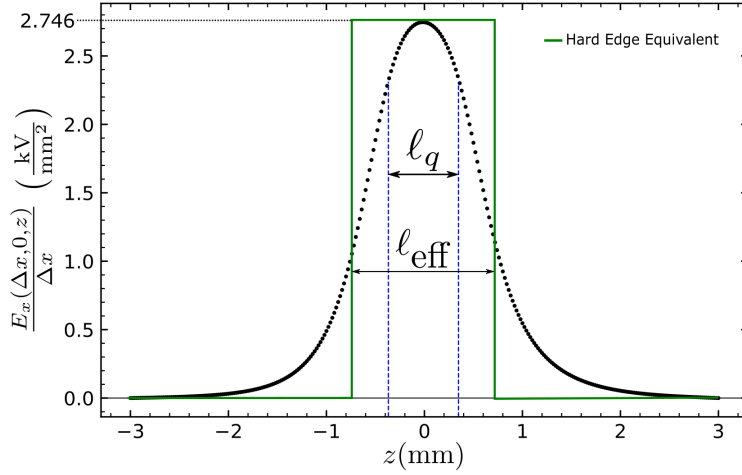


Figure 5.6: Electric field gradient one grid-cell off-axis for a single isolated ESQ made of solid cylindrical rods. The black vertical dashed line demarcates the physical length of the ESQ  $\ell_q$ . The green line shows the hard-edge equivalent with a total field extent equal to the calculated effective length  $\ell_{\text{eff}}$ .

ESQ arrays are optimized to minimize the multipole field errors. We integrated the 3D electric field  $\vec{E}(\vec{x})$  over the axial coordinate  $z$  and normalized by  $\ell_{\text{eff}}$  [29] to obtain the averaged 2D transverse field components  $\bar{E}_x$  and  $\bar{E}_y$

$$\begin{aligned}\bar{E}_x &= \frac{1}{\ell_{\text{eff}}} \int_{z_1}^{z_2} E_x(x, y, z) dz, \\ \bar{E}_y &= \frac{1}{\ell_{\text{eff}}} \int_{z_1}^{z_2} E_y(x, y, z) dz.\end{aligned}\tag{5.3}$$

Again,  $z_1$  and  $z_2$  are again sufficiently large to capture axial extent of the field upstream and

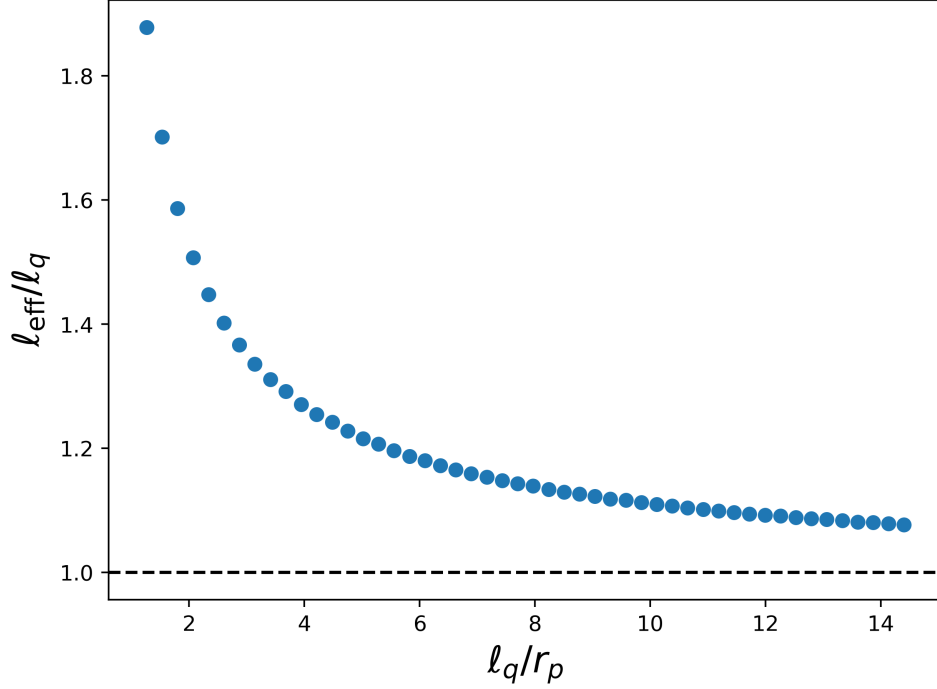


Figure 5.7: Ratio of physical ESQ length  $\ell_q$  to calculated effective length  $\ell_{\text{eff}}$  as  $\ell_q$  is increased.

downstream of the ESQ field. The Maxwell Equations of the 2D-field satisfy the Cauchy-Riemann conditions for a complex field  $\bar{E}^* = \bar{E}_x(x, y) - i\bar{E}_y(x, y)$  with  $i = \sqrt{-1}$  and complex coordinate  $\underline{z} = x + iy$ . Variables that are underlined are considered complex except for  $\bar{E}^*(x, y)$  to avoid awkward notation. Thus, expanding the field in a Laurent series we have:

$$\bar{E}(x, y) = \bar{E}_x(x, y) - i\bar{E}_y(x, y) = \sum_{n=1}^{\infty} \underline{b}_n \left( \frac{\underline{z}}{r_p} \right)^{n-1}. \quad (5.4)$$

Here,  $\underline{b}_n$  is a complex number representing the multipole moments taken so that the  $\underline{b}_n$  coefficients are measured in field units. In a polar representation with  $\underline{z} = r e^{i\theta}$ , Eq. (5.4) is expressed as

$$\bar{E}(x, y) = \bar{E}_x(x, y) - i\bar{E}_y(x, y) = \sum_{n=1}^{\infty} \underline{b}_n \left( \frac{r}{r_p} \right)^{n-1} e^{[i(n-1)\theta]}. \quad (5.5)$$

Expressing  $b_n = A_n - iB_n$ , orthogonality can be exploited to calculate:

$$\begin{aligned} A_n &= \frac{1}{2\pi} \left(\frac{r_p}{r}\right)^{n-1} \int_0^{2\pi} \left[ \bar{E}_x \cos[(n-1)\theta] - \bar{E}_y \sin[(n-1)\theta] \right] d\theta \\ B_n &= \frac{1}{2\pi} \left(\frac{r_p}{r}\right)^{n-1} \int_0^{2\pi} \left[ \bar{E}_x \sin[(n-1)\theta] + \bar{E}_y \cos[(n-1)\theta] \right] d\theta. \end{aligned} \quad (5.6)$$

To evaluate the coefficients in Eq. (5.6) the Warp fields on a 3D Cartesian rectangular parallel-piped mesh were axially integrated using Eq. 5.3 to obtain 2D fields on a transverse square mesh. Field values at mesh points are linearly interpolated to desired coordinates using volume weighting. The multipole coefficients are then calculated using Eq. 5.5. Enough azimuthal mesh points ( $N_\theta$ ) are taken along a circular interpolation path to resolve the maximum order  $n$  of the extracted multipole by taking

$$N_\theta = \frac{\sqrt{2\pi nr}}{\Delta x}. \quad (5.7)$$

Here  $\Delta x$  is the grid spacing in  $x$ ,  $\Delta y = \Delta x$ , and  $r$  is the integration radius. We chose  $r$  to be approximately one grid cell within  $r_p$  since the multipole error terms will be strongest in regions close to the ESQ surfaces (see Fig. 5.8). Symmetry considerations show that only  $n = 2$  (Quadrupole), and  $n = 2 + 4p$  for  $p = 1, 2, 3, \dots$  are allowed [30]. Therefore, the leading order error term is  $n = 6$  or 12-pole terms. Simulations are consistent with these symmetry considerations to within expected numerical errors.

## 5.4.2 Optimizing ESQ Arrays to Minimize Leading Order Error in Field Expansion

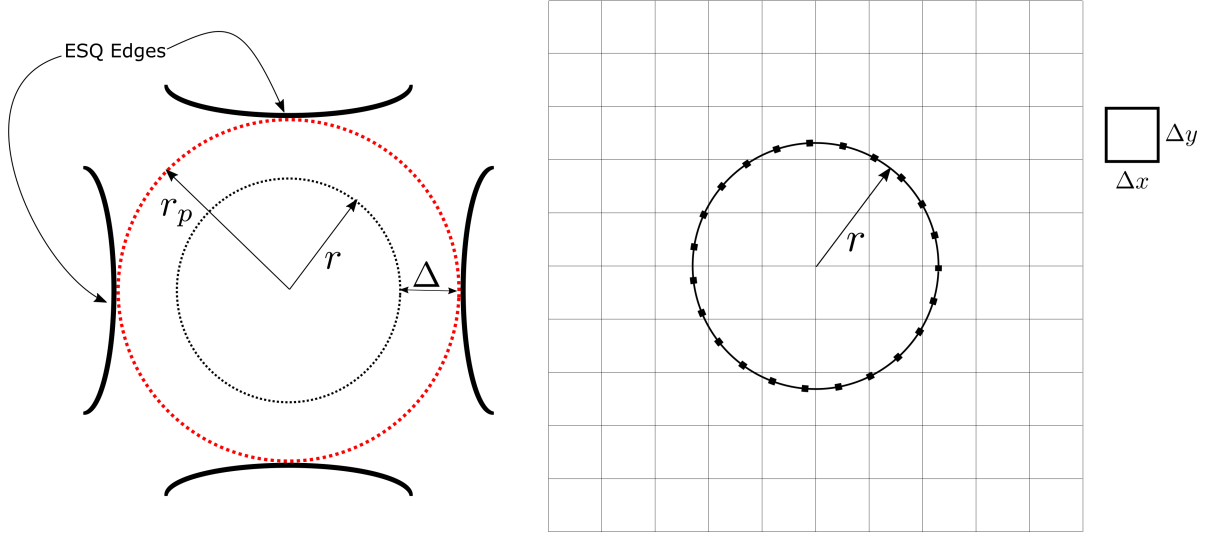


Figure 5.8: (Left) The clear bore aperture is displayed as the dotted red circle with radius  $r_p$ . The integration path is shown as a dotted black line at radius  $r$  and was chosen to be approximately one grid cell within the aperture. The radial distance between  $r_p$  and  $r$  is represented by  $\Delta$  which is chosen to be grid cell away from the ESQ edges shown in solid black lines. (Right) Schematic of the the simulation mesh with the interpolation points passing through the discretized space with simulation cell box sizes of  $\Delta x$  and  $\Delta y$ .

#### 5.4.2.1 Benchmarking for a Single Axially Long ESQ

To simulate a long quadrupole ( $\ell_q \gg r_p$ ) the longitudinal mesh boundaries were set to periodic and placed on the axial rod extents to model an infinitely extended quadrupole with pure 2D transverse fields. Meanwhile, the transverse boundary conditions were set to Dirichlet with the simulation box large enough to contain the rods. To resolve the mesh elements and provide enough sample points to accurately interpolate points along the integration path at radius  $r$ , the resolution on the mesh was fixed at  $5\mu\text{m}$ . To analyze the correlation between the rod radius to aperture radius ratio ( $R/r_p$ ) and the leading order error term ( $A_6$ ), we calculated the normalized error,  $A_6/|A_2|$ , for several rod radii while keeping the aperture radius fixed at  $r_p = 0.55\text{ mm}$ . Referencing Fig. 5.9, it was determined that  $A_6$  is zeroed when  $R/r_p = 1.145$ . The result for a single isolated long quadrupole is within 0.1% of the result previously found by Faltens and verifies the algorithm [1].



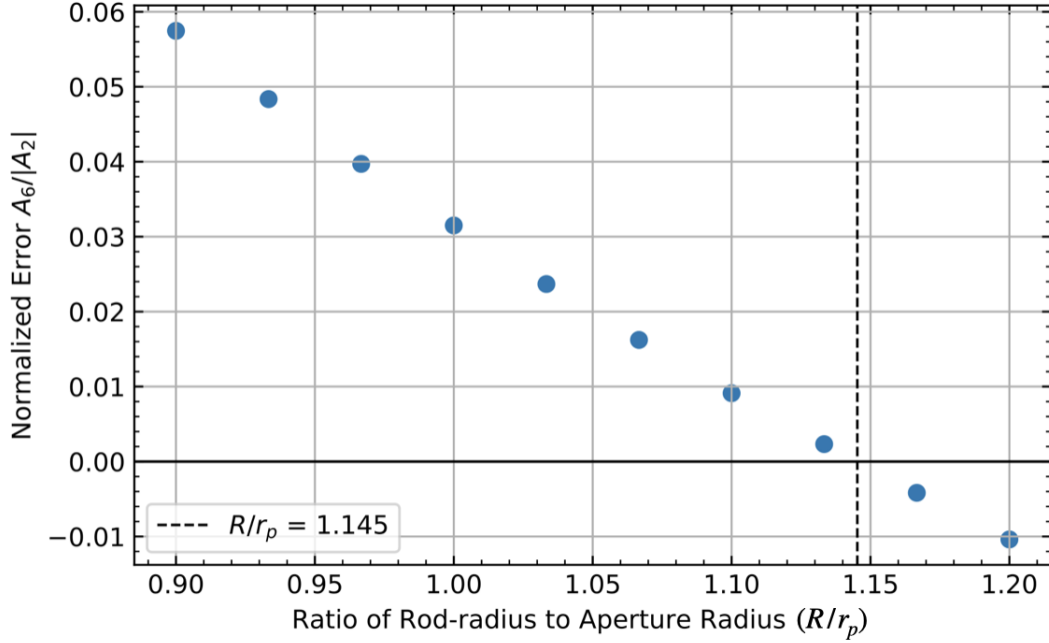


Figure 5.9: Leading error term  $A_6$  as a function of the rod radius  $R/r_p$  for an infinitely long quadrupole.

#### 5.4.2.2 Finite Axial Length ESQ Optimization for Full and Chopped Rod Geometries

Fig. 5.10 shows three sets of curves for a finite length quadrupole with  $\ell_q = 0.695$  mm,  $r_p = 0.55$  mm, and three different rod geometries. The blue-dotted curve corresponds to a geometry using full rods, the green curve with triangle markers is for half-chopped rods, and the third magenta curve with “x” markers is a quarter-chopped rod. All three curves were created with Dirichlet boundary conditions for both transverse and longitudinal directions with boundaries placed sufficiently far away not to significantly impact the results shown. The curves for the full and half-chopped rods share the same optimum at  $R/r_p = 1.318$ . However, the quarter-chopped rod has a new optimum at  $R/r_p = 1.330$ , a one percent difference. Although a relatively small percent difference, this is notably different from the other two curves and we deduce that the curvature of the rod along with the finite axial length

are primary factors that shift from the infinitely long ESQ optimum  $R/r_p = 1.145$  discussed more thoroughly in Sect. 5.4.2.1. This suggests that one could create a more transversely compact accelerator by using chopped rods while maintaining the same curvature and varying the rod radius to minimize field errors. However, transverse chop locations may give strong field enhancement at edges and contribute to voltage breakdown at a lower voltage (Sect. 5.4.2.3).

Lastly, we used periodic boundary conditions instead of Dirichlet boundary conditions and used the simulation box to chop the rods in half. The optimum value for this setup was found to be  $R/r_p = 1.336$ , a less than 0.5% difference from the same setup with Dirichlet transverse conditions. This suggests that the periodic array geometry does not significantly shift the optimum point to minimize field errors. Note that half-chopped rods in a periodic array eliminate transverse edges that may impact voltage holding.

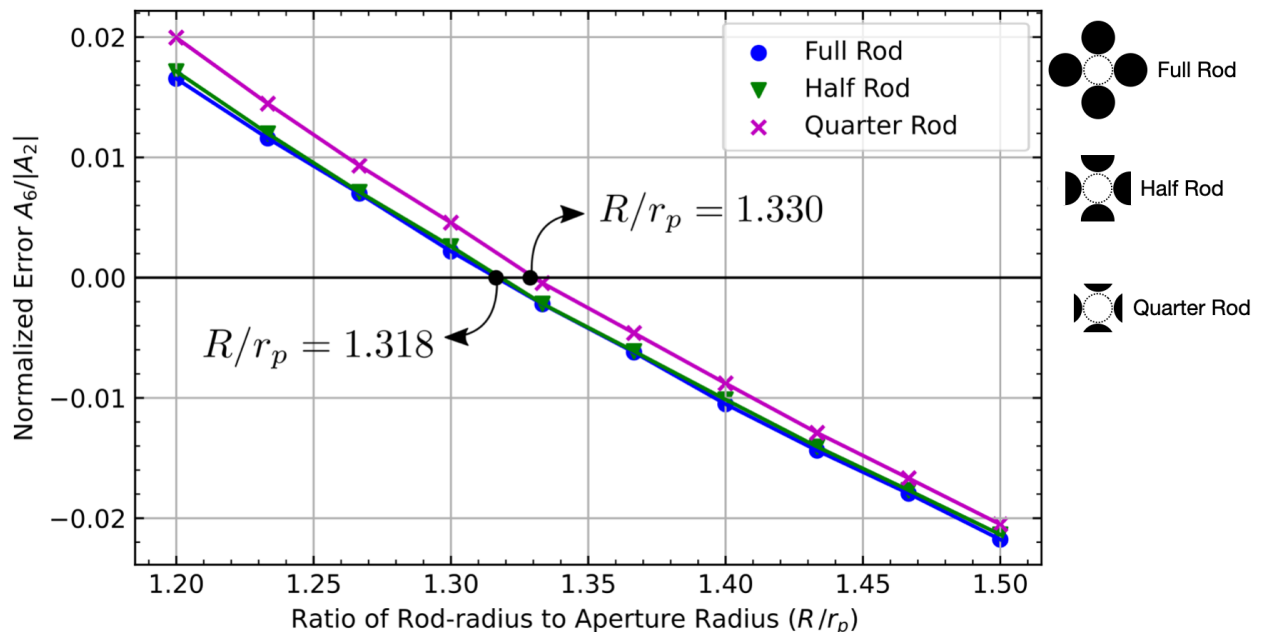


Figure 5.10: Leading error term normalized by the quadrupole term  $A_6/A_2$  in the multipole field expansion for finite length ESQ using full (blue circles), half-chopped (green triangles), and quarter-chopped (magenta “x”s) rod geometries. To the right are illustrative examples of what the geometric configuration is and where the chop is being done.

### 5.4.2.3 Voltage Breakdown Limits for ESQ Array Geometry

Focusing provided by the ESQs will require higher voltages to compensate for increased kinetic energy of the ions as they advance through the acceleration lattice. As kinetic energy increases, the lattice period will also increase allowing for axially longer ESQs (length is increased by stacking wafers and by machining a new ESQ design) to help with voltage holding. For distances less than 1 cm, breakdown in typical low energy transport conditions is expected at electric field strengths of  $10^7$  V/m [1]. The max 2D electric field for an ideal quadrupole occurs at  $R = r_p$  and is given by

$$E_{2D} = \frac{2V_q}{r_p}. \quad (5.8)$$

Substituting the breakdown field limit of  $10^7$  V/m and solving for the quadrupole voltage  $V_q$  gives a maximum limit of  $V_{q,br} = 2.75$  kV; it is recommended that fields be kept to half the breakdown, giving a practical limit of  $V_{q, \frac{1}{2}br} = 1.37$  kV. To evaluate the impact of transverse and longitudinal rod chopping, we used Warp to perform a voltage scan of an isolated ESQ with  $\ell_q = 0.695$  mm and  $r_p = 0.55$  mm using periodic transverse boundary conditions on a 3D grid. The simulations calculated the electric field for a mesh with  $10 \mu\text{m}$  in the transverse direction and  $20 \mu\text{m}$  in the longitudinal direction. The voltage scan was performed over a range of voltages from 0.1 kV to 1 kV and the magnitude of the 3D electric field  $|E(x, y, z)|$  was computed at each point on the mesh. The half breakdown limit occurred at an applied voltage of 531 kV, a 61% decrease from estimated 2D field estimated above. This breakdown value occurred at the longitudinal edge of the ESQ and transversely at the point closest to the oppositely biased rod.

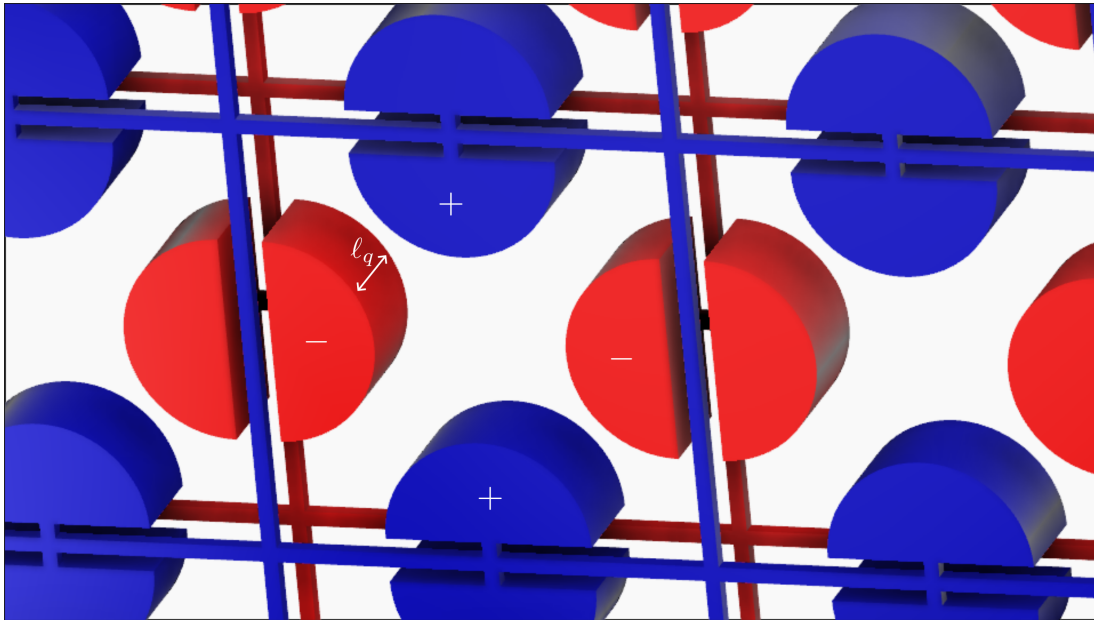
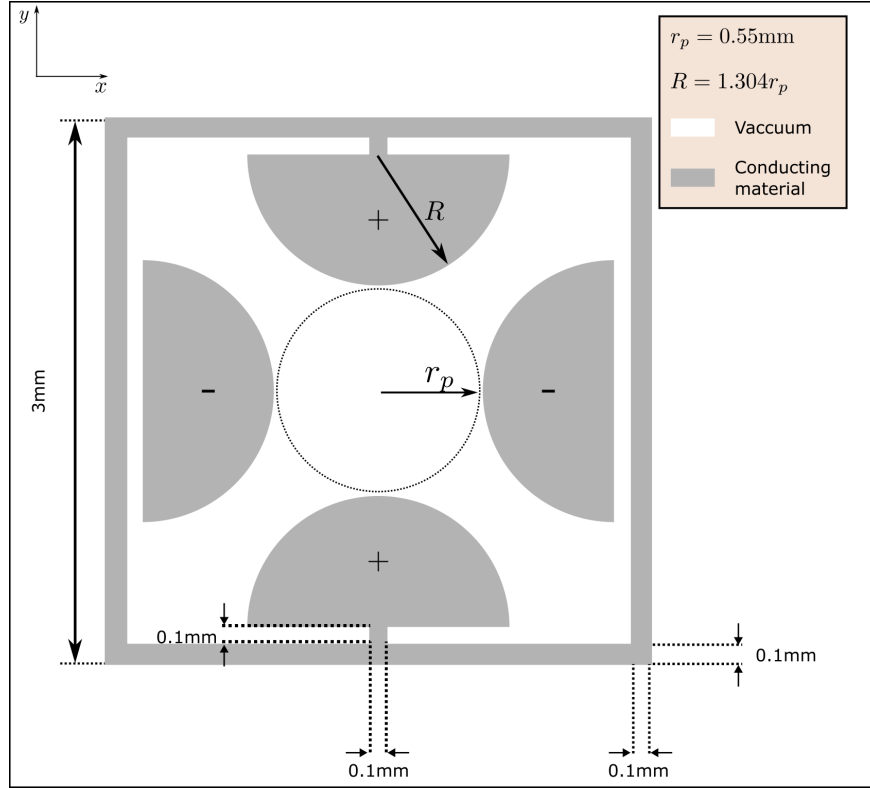


Figure 5.11: (Top) Schematic of a single cell from the realistic transverse-periodic array ESQ design used in simulations. The clear bore aperture is shown as the dotted black line with a radius of  $r_p$ . The half-chopped quadrupole rods have radius  $R$ . Conducting objects are shaded in grey and white-space represents vacuum. (Bottom) Transversely periodic ESQ array modeled in simulation. Positively biased structures are colored blue and negatively biased structures in red.

#### 5.4.2.4 Optimization Considerations From Realistic Geometry

The manufactured geometry varies from the idealized geometry in both the transverse and longitudinal directions as can be seen in Fig. 5.11. Transversely, there is a surrounding conducting square with prong attachments to the rods. This structure is printed on each longitudinal side of a single wafer. On one side, the conducting structures are positively biased while on the other end the prong attachments are switched to the other rods with all conductors being negatively biased. Note that the surrounding material is plastic. Because WARP does not have the capability to simulate dielectric material in the full 3D solver, we take the surrounding material to be vacuum. The conducting square that surrounds the rods and the attaching prongs place a physical limit on the size of the poles. This limit ensures that the biased rods do not touch the surrounding conducting square and that the rods do not cross into the aperture space. We take the half-chopped rod geometry as our design to create a more compact structure. With these physical constraints the physical rod radius must satisfy  $R_{\max} \leq 0.75 \text{ mm} = 1.36r_p$ . Using the same analysis as for the ideal geometry and a half-chopped structure with  $r_p = 0.55 \text{ mm}$  and  $\ell_q = 0.695 \text{ mm}$  an optimum ratio of  $R/r_p = 1.304$  was found (see Fig. 5.12).

Following the same procedure as before for the ideal case, the half breakdown-limit was found to be at  $V_{q,\text{br}} = 0.531 \text{ kV}$ . This places the safe operating limit at a much lower voltage of  $V_{q,\frac{1}{2}\text{br}} = 0.267 \text{ kV}$ . Before, the breakdown point occurred at the longitudinal chopped ends of the ESQs. In the realistic geometry, the point shifts to the space between the oppositely charge rod and surrounding square at the prong tips. Due to the aforementioned inability to simulate the dielectric material in WARP's full 3D solver, the operating limit is likely higher since this area is surrounded by dielectric and breakdown will likely occur at

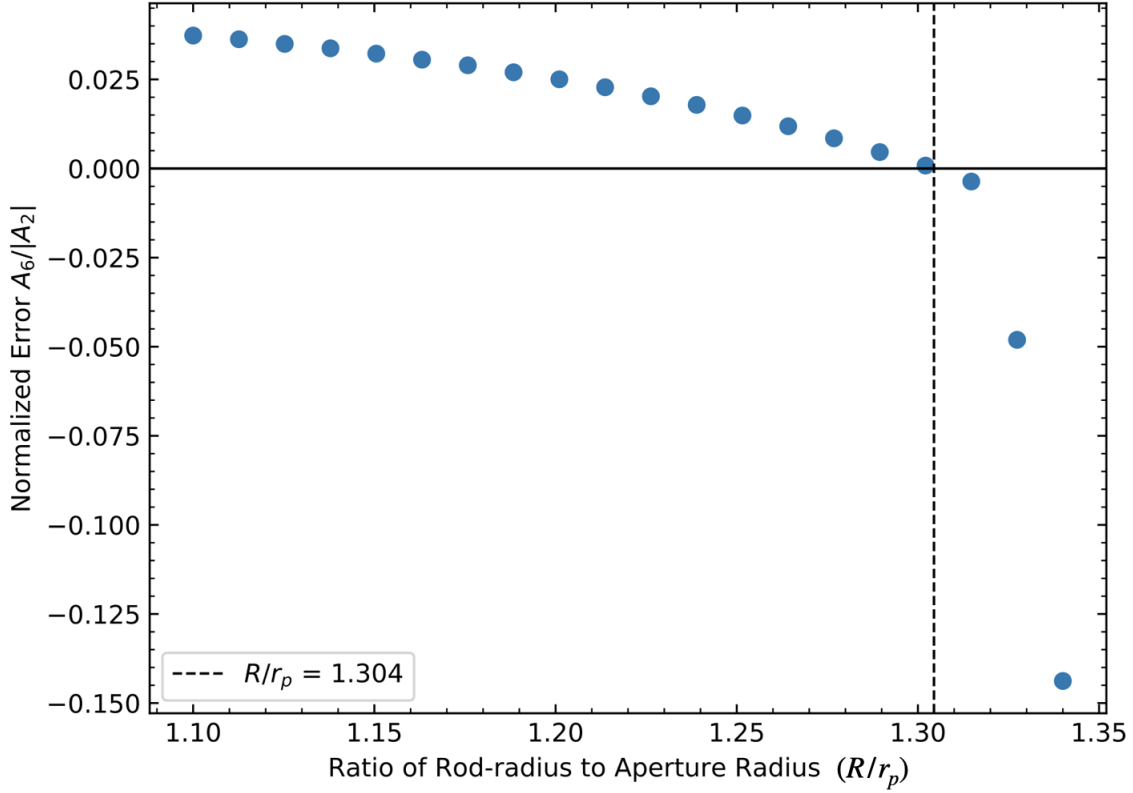


Figure 5.12: Leading error term as a function of the rod radius  $R/r_p$  for an ESQ with the manufacturedx geometry.

the axial rod ends. For example, the dielectric material used in this system is FR-4. This material has a dielectric constant ranging from 3.8 to 4.8. Thus, assuming the electric field scales linearly with the dielectric constant, we can assume the safety limit to occur within the range of  $0.7 \text{ keV} \leq V_q \leq 0.9 \text{ keV}$ . This safety limit defines the operating point at 50% the breakdown; therefore, this limit can be pushed higher with caution. Since it is unclear which breakdown limit we should use but we know the dielectric material will increase the limit, we will take the breakdown limit found in vacuum to be the half-breakdown limit from here on: that is,  $V_{q, \frac{1}{2}\text{br}} = 0.531 \text{ kV}$ .

In addition to the geometrical differences, the acceleration lattice will be comprised of a doublet with oppositely biased neighboring ESQs. The gradient of the two ESQs will be zero

at the symmetry point located axially between the geometric center of the two ESQs. Due to the close separation and opposite bias, the close axial separation  $d$  will have an impact on the effective length and therefore the overall focusing strength obtained. This is illustrated in Fig. 5.13 with  $r_p = 0.55$  mm and  $\ell_q = 0.695$  mm for two ESQs separated by  $d = 0.695$  mm. To analyze the effects of  $d$  in the doublet structure we performed simulations by placing an isolated doublet in a simulation box with periodic boundary conditions in the transverse direction and Dirichlet boundary conditions in the longitudinal direction. The upstream quadrupole was held fixed at a sufficient distance from the upstream grounded plane to ensure that there was no significant impact on the field. The downstream quadrupole was then placed at varying distances and the effective length for the upstream quadrupole was calculated by integrating the gradient from the upstream grid extent to the downstream zero point and applying Eq. 5.2 with  $G^*$  corresponding to the max field in the upstream ESQ. The results are shown in Fig. 5.14. In the figure, we plot  $\Delta\ell_{\text{eff}} = \ell_{\text{eff}}|_{\text{doublet}} - \ell_{\text{eff}}|_{\text{isolate}}$  normalized by the  $\ell_{\text{eff}}|_{\text{isolate}}$  using the effective length of the isolated ESQ and varying  $d$  between the two ESQ extents. Results show that close ESQ spacing with  $d \lesssim 2.5$  mm can significantly degrade effective focusing strength. This result provides guidance on spacing requirements for the ESQ doublet system both between the ESQ pairs and the ESQ and neighboring grounded acceleration gap wafer.

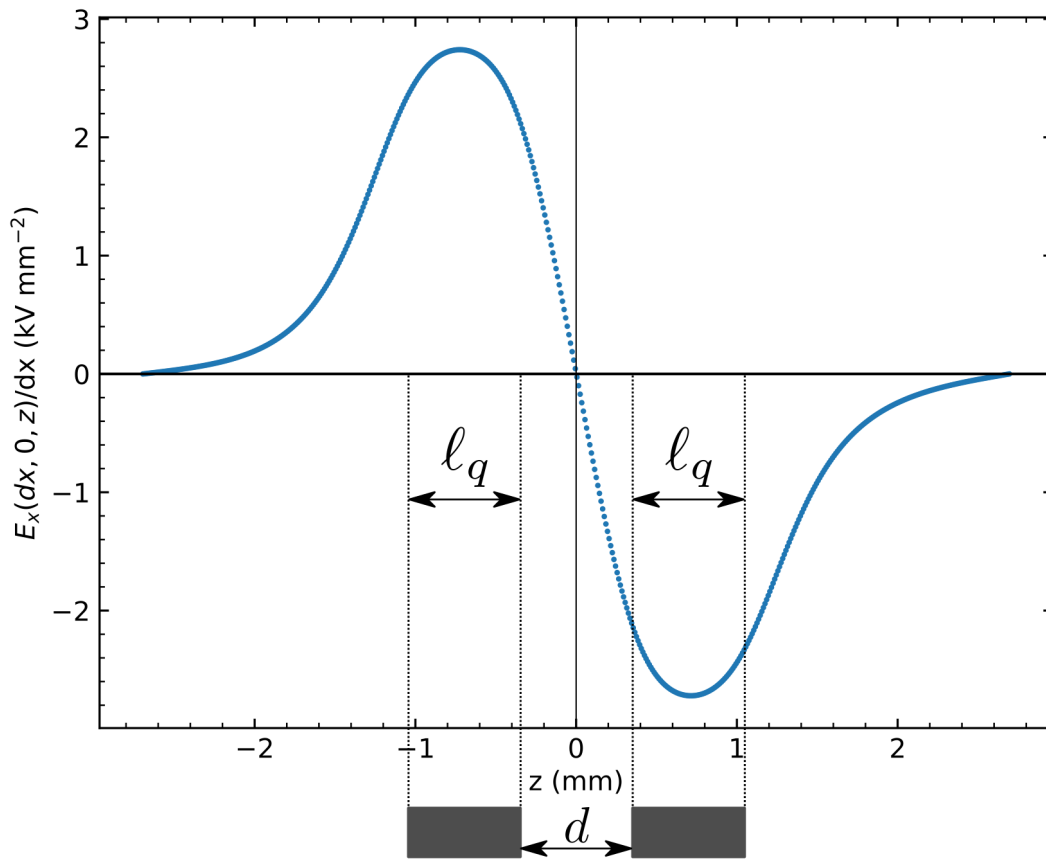


Figure 5.13: Gradient of ESQ Doublet structure. The ESQ elements are drawn below the plot with axial length  $\ell_q$  and axial separation distance  $d$ .



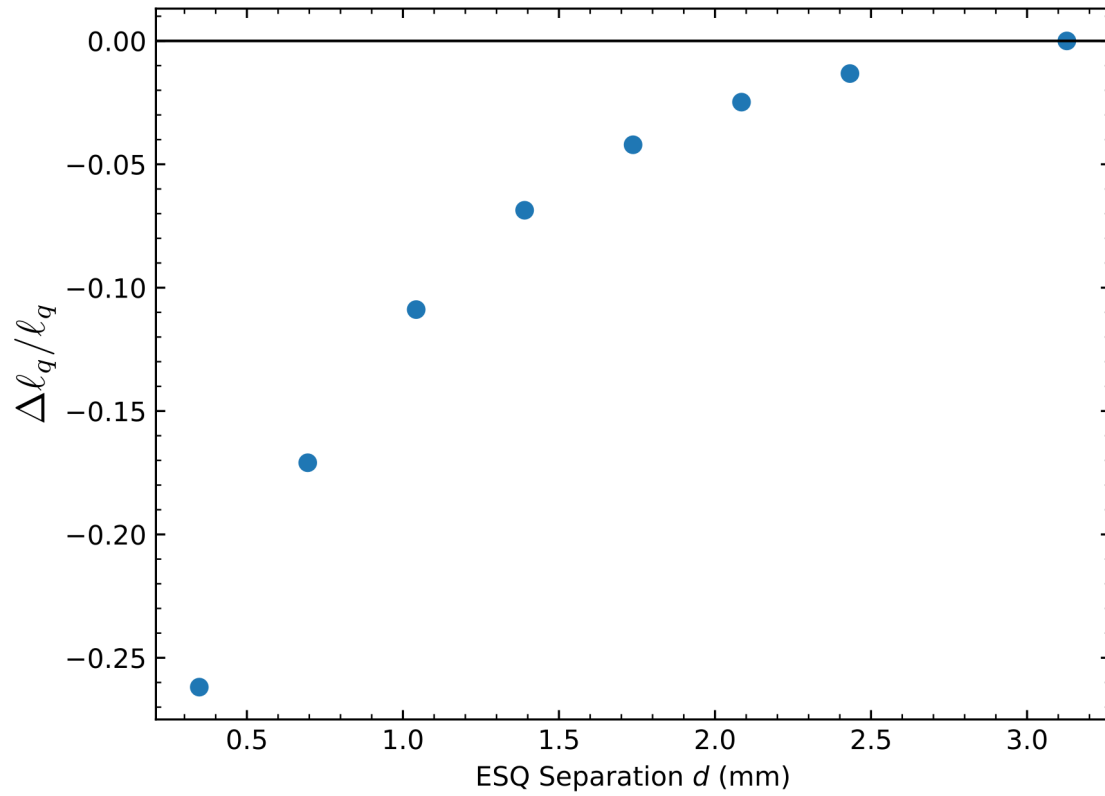


Figure 5.14: Difference in the calculated effective length  $\ell_q$  as the doublet separation distance  $d$  is varied.

## 5.5 Summary and Concluding Remarks

The ESQs in our system will provide transverse focusing to increase the number of particles delivered on target. Creating a pure quadrupole field is not possible and there will be contributions to the field from high order error terms. We developed a computational tool to simulate ESQ geometries and optimize geometric choices to minimize the leading order error term  $A_6$  in the quadrupole field expansion. We found that for  $r_p = 0.55$  mm and  $\ell_q = 0.695$  mm that using a half-chopped rod of radius  $R = 1.304r_p$  is the optimal choice. Our tool also enabled us to predict where voltage breakdown is likely to occur and at what voltage. Due to the inability of Warp to simulate dielectrics in 3D, the breakdown limit is probably significantly lower than when the FR-4 dielectric is present. We therefore will take the breakdown limit found in vacuum to be the half-breakdown limit so  $V_{q, \frac{1}{2}\text{br}} = 0.531$  kV. Lastly, we were able to characterize the effects of axial separation in the doublet structure for our ESQ design has on the field. We showed that for an axial separation of  $d \geq 2.5$  mm the fields from the opposing doublets do not interfere with each other. The developed tool is robust and can be applied to new geometries with the limitation being the ability to accurately create the conductor in Warp.

# Chapter 6. Longitudinal Dynamics in MEMS-RF

In Ch. 2 the longitudinal dynamics in the MEMS-RF system are discussed. The main difficulty arises due to the strong acceleration because  $qV_g = \mathcal{E}_b$  out the source which prevents a perturbative treatment. The low beam energy at the source means that we will want to accelerate as much as possible and keep the system compact. With non perturbative acceleration, defining an RF bucket is challenging and the  $\mathcal{E}_b$ - $t$  phase-space evolution will have to be analyzed numerically. Using a self-consistent 3D simulation like Warp is possible but impractical. Varying the synchronous phase condition to balance longitudinal focusing and acceleration can be done by varying the gap-to-gap distances in the lattice. Using a fixed set of injection conditions this limits the parameter tuning to one variable. However, the synchronous phase is a continuous variable between  $-\pi/2 \leq \phi_s \leq 0$  and would need to be tuned for each gap. This brute force approach approaches infeasibility for higher energies. Furthermore, the complexities captured in the Warp simulation (space-charge dynamics, transverse dynamics, particle scraping) make analysis of the longitudinal dynamics difficult to extract. For these reasons, a rapid and reduced model is beneficial. The model should be able to rapidly execute for any number of gaps and designs. It should also be comparable to Warp when making a comparison for a test case. In this chapter, the development and application of such a model is discussed.

## 6.1 Numerical Methods

In Sect. 5.3 the extraction of the gap array fields was discussed. To simulate particles moving RF-varying fields, the normalized RF-fringe field is extracted. Since the gap geometry is constant throughout the lattice, this extraction only needs to be done once. However, as noted in the earlier section, care must be taken to include two gaps in the Warp simulation or else the periodic array geometry will not be incorporated into the field-solve. This will give a vastly different field and will alter the results in the simulation. With the extracted fringe data, the position and field array can be stitched into a complete lattice. Fig. 6.1 is an illustration to help explain the process for stitching in the extracted  $z$ -data. Initially, the complete  $z$ -array is created with some specified resolution  $dz$  and  $n$ -points. The gap center is located in this initial array between  $z^k$  and  $z^{k+1}$ . Here, the extracted  $p$ -points of  $z$ -data ( $\mathbf{z}_{\text{ext}}$ ) is incorporated into the lattice. In general,  $p \neq n$  and  $dz \neq dz_{\text{ext}}$ . A completed  $z$ -mesh will now have  $n + m$  grid points. An identical algorithm is performed in tandem to incorporate the field data  $E_{z,\text{ext}}$  between  $E_z^k$  and  $E_z^{k+1}$ . This process is repeated for the specified number of gaps and the normalized  $E_{z,\text{ext}}$  is scaled to represent the acceleration voltage with the scaling factors discussed in Sect. 5.3.

With the completed position and field data, a simulation for transporting particles down the lattice with RF-varying fields can be performed. To explain the algorithm consider a single particle at the  $(j - 1)$ -th point in the advancement. The particle of mass  $m$  at current position  $z^{j-1}$  and time  $t^{j-1}$  with energy  $\mathcal{E}^{j-1}$ . To advance the particle to  $z^j$  the particle velocity  $v^{j-1}$  must be calculated. Then, the velocity is used to find the time it takes to travel  $\Delta z^j = z^j - z^{j-1}$ . Note that this distance should be calculated each step since it is not guaranteed that the mesh is uniformly discretized. This time is then used to compute

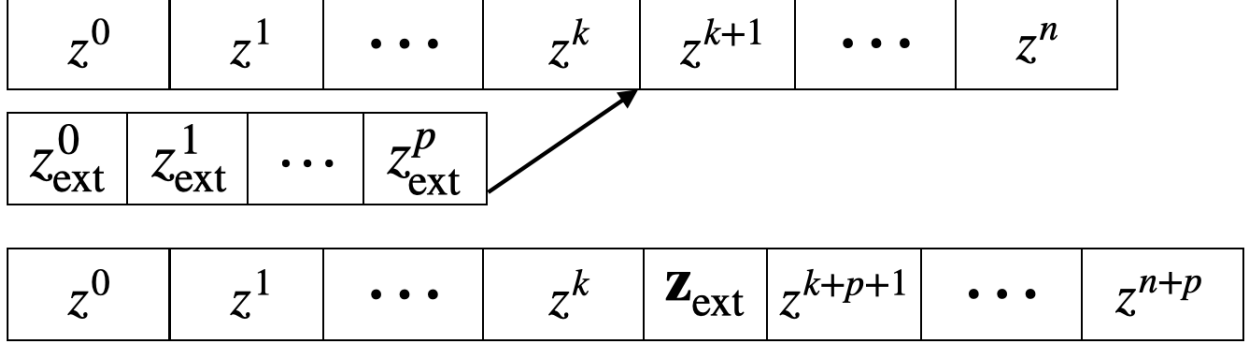


Figure 6.1: Visual description of the array-stitching process for the position data. An initial  $z$ -array is created with  $n$ -grid points. The extracted gap data has both field data and position data. Here the position data  $\mathbf{z}_{\text{ext}}$  is stitched into the initial array between  $z^k$  and  $z^{k+1}$ . The resultant mesh will have  $n + p$  grid points. An identical algorithm is performed in tandem for the field array and extracted field data.

the future electric field  $E_z(z^j, t^j)$  and calculate the work done on the particle by the field.

Symbolically, the algorithm is:

$$\begin{aligned}
 v_z^{j-1} &= c \sqrt{\frac{2\mathcal{E}^{j-1}}{mc^2}}, \\
 t^j &= t^{j-1} + \frac{\Delta z^j}{v_z^{j-1}}, \\
 \mathcal{E}^j &= \mathcal{E}^{j-1} + q\Delta z^j E_z(z^j, t^j)
 \end{aligned} \tag{6.1}$$

where  $c$  is the speed of light. The electric field at the  $j$ -th grid point at  $t^j$  is given by  $E_z(z^j, t^j) = E_z^j \cos(\omega t^j)$  with  $\omega - \text{rf} = 2\pi f_{\text{rf}}$  and  $E_z^j$  is the  $j$ -th element of the field data.

It is possible for particles to get decelerated and acquire negative velocity in the simulation.

When this is done, these particles are “scraped,” meaning that a filter is applied so that

particles with  $\mathcal{E} \leq 0$  are ignored. Particles are loaded by initially distributing about the

design particle which is always placed at  $z = 0$ . Thus, particles will be between  $|z_{\text{load}}/2|$

where  $z_{\text{load}}$  is some distance set by the user. Then, the time for each particle to reach  $z = 0$

is calculated and saved. This initializes all particles to  $z = 0$  with an initial time.

## 6.2 Benchmarking the 1D Model with Warp in a 4-Gap System

To verify that the 1D code was properly describing the longitudinal dynamics of the MEMS-RF system, a simulation of the 1D code was compared against Warp. A four gap system with  $V_g = 7$  kV was modeled for maximum acceleration. A continuous and uniformly distributed beam was injected with initial energy  $\mathcal{E}_{b,i} = 7$  keV. Fig. 6.2 shows a comparison in the final energy distribution between the Warp simulation in orange and the 1D code in blue for a selection of particles satisfying  $|\Delta t| = |t - t_s| \leq 0.5\tau_{\text{rf}}$ . If this filtering is not done then the comparison is misleading. This is because Warp advances particle in time and particles will be lost for various reasons: absorbed at transverse boundary, decelerated to negative velocity, to slow to have made to target before the simulation is terminated. In the 1D code this is not an issue since particles are advanced in space and thus, all particles that maintained a positive kinetic energy will be collected. Despite Warp simulating space-charge effects and advancing particles in a 3D space, the 1D model does well to capture the energy distribution with no space-charge effects. Additionally, this code is rapid in comparison with Warp: the 1D code takes less than a second whereas the Warp simulation can take up to ten minutes for a low-resolution (large spatial discretization, large time step) run.

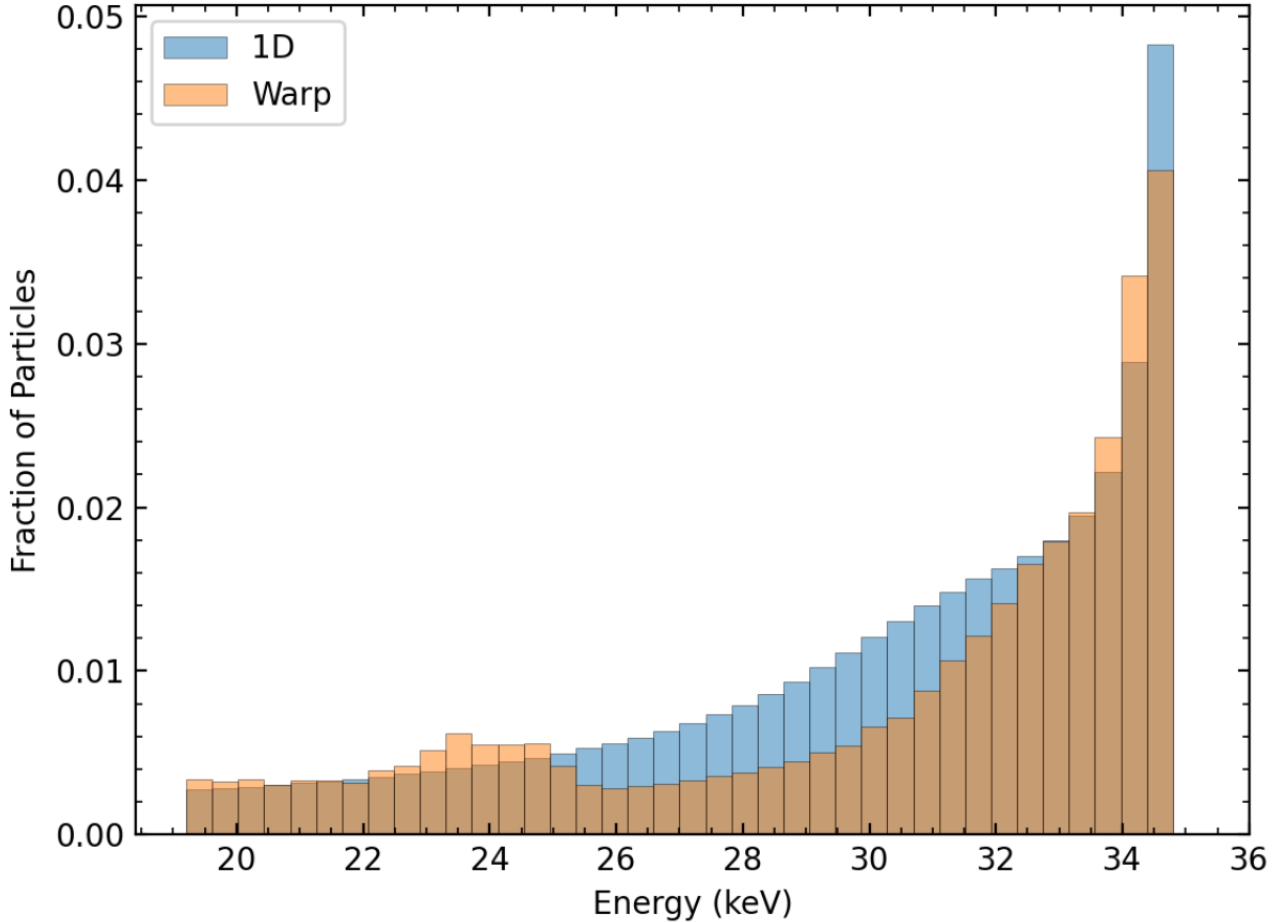


Figure 6.2: Comparison of a four-gap simulation with  $V_g = 7$  kV and  $\mathcal{E}_{b,i} = 7$  keV between the 1D code (blue) and Warp (orange). The binning is done over the final energies of the particles after arriving at target. Particle energies are filtered based on arrival time relative to the synchronous particle with  $|\Delta t| = |t - t_s| \leq 0.5\tau_{rf}$

### 6.3 Lattice Design for Optimal Acceleration, Bunching, or Both

With the algorithm validated the parameter space can be explored to find optimal lattice design. The definition of optimal will change based on the application. In most cases users want the most particles delivered with the highest energy. Especially in cases of linacs, users want both this criterion met with the shortest lattice possible. These two optimums

theoretically conflict with one another—maximum acceleration gives least particle capture and vice versa. As a demonstration of the algorithm, we will compare three different lattice designs: one for maximum acceleration, one for maximum bunching, and one that tapers the acceleration so that there is particle capture and bunching. This will be done for a four gap system to keep the analysis easier.

### 6.3.1 Initial Conditions and Parameter Settings

A cold (iso energetic) continuous beam of uniformly distributed  $\text{Ar}^+$  ions will be simulated with  $\mathcal{E}_b = 7 \text{ keV}$  from the source. The injection will be for one RF-period giving a beam length of  $\tau_{\text{rf}}$  where  $L_b = v_b \tau_{\text{rf}}$ . At the target, only particles with  $|\Delta t| \leq 0.4 \tau_{\text{rf}}$  will be involved in the analysis. Note, by tightening the range on  $\Delta t$  we have made it more difficult to capture particles in the end diagnostic. This drives design to increased particle capture when tuning design parameters. We use an applied bias of  $V_g = 6 \text{ kV}$  with the first gap starting at a maximally negative field. The voltage drop is done as a design choice that will prevent possible damage to the source. At extraction, such a strong acceleration will stagnate and decelerate particles which will impinge on the source. Fig. 6.3 shows the extracted  $E_z(r = 0, z)$  at  $t = 0$  for a  $N_g = 4$  where the fields are normalized by the calculated DC-field  $E_{\text{DC}} = V_g/g$ . Note that the field amplitude is greater than unity. When applying  $V_g$  to the conductors, the on-axis potential for the real conductor is less than this value. In order to ensure the desired on-axis voltage, the applied bias must be slightly greater. Because the fields shown are extracted from the real conductors, these fields are scaled to achieve the correct  $E_z(r = 0, z)$ . If this scaling is not done, then the RF-matching will be slightly mismatched for the design particle and desynchronization will grow gap-to-gap. Gap centers are marked by vertical gray dashed lines and note the increase in spacing to maintain



synchronism as the beam energy increases. Particles are advanced from  $z = 0$  to a Faraday cup (Fcup) marked by a vertical red dashed line that is placed 10 mm from the edge plate of the final gap.

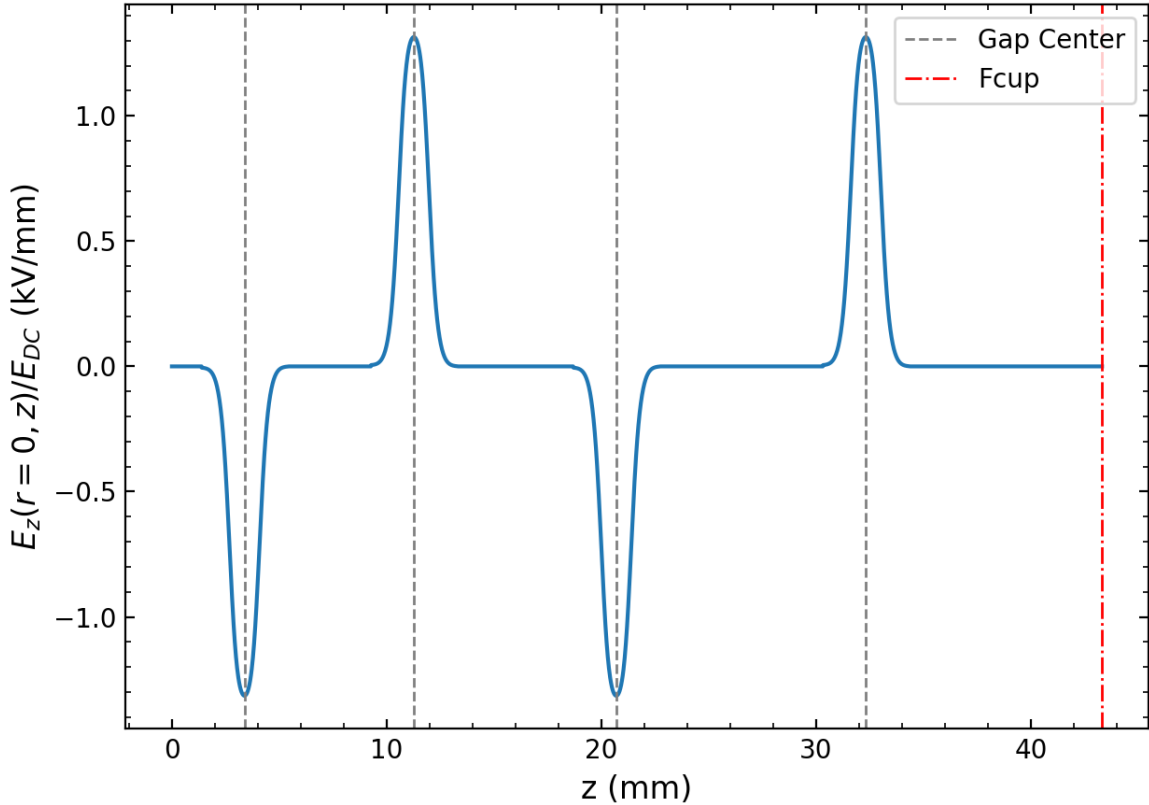


Figure 6.3: Normalized on-axis electric field for a 4-gap system at  $t = 0$ . The electric fields are normalized by the calculated DC-electric field  $E_{\text{DC}} = V_g/g$  (V/m). The on-axis voltage will be less than the applied voltage in the real conductor. Since the field is extracted from the conductor, the field needs to be scaled to be greater than  $E_{\text{DC}}$  to ensure the on-axis voltage matches the desired voltage hence why the amplitudes in the plot are greater than unity. Each gap center is marked by a grey vertical dashed line and the distances between gap centers increases with increasing energy by  $\beta\lambda_{\text{rf}}/2$ —the RF-resonance condition for gaps operating in the  $\pi$ -mode. The particle advance is terminated at the Fcup marked by a red vertical dashed line and placed 10 mm from the final gap edge.

### 6.3.2 Comparing Lattice Designs

Fig. 6.4 shows a side-by-side comparison of the  $\mathcal{E}$ - $t$  phase space for the three design choices at the Fcup. Each plot is created using a kernel density estimator (KDE) to plot the 2D data

of  $\mathcal{E}$  and scaled relative time difference  $\Delta t/\tau_{rf}$ . The KDE plot is chopped at the extreme points where the data does not exist—a KDE plot will estimate data in these regions and can be misleading like in 1. where there should not be particles with higher energy than possible. The  $\mathcal{E}$  and  $\Delta t/\tau_{rf}$  histograms are plotted and added to the main plot at the top and bottom respectively. Marked by vertical dashed lines is the design particles energy a relative. Since  $\Delta t = t - t_s$ , the vertical line will always be at 0. The intersection of the two lines is where the design particle is found in phase-space. Final design energy and particle transmission is noted in the top right corner of each plot.

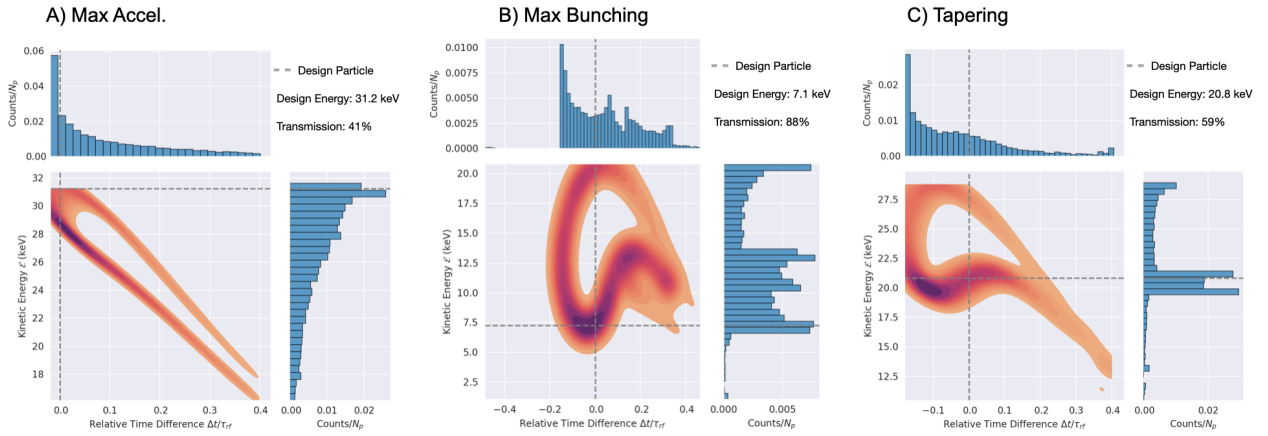


Figure 6.4: Phase space plots for different design choices. The main plot is a kernel density estimator (KDE) plot and the estimates at the extreme values have been chopped to prevent misleading data such as energies greater than possible in 1. Each plot also has the  $\mathcal{E}$  and  $\Delta t/\tau_{rf}$  binned and plotted on top and bottom respectively. The vertical dashed lines represents the design particle and the vertical line will always be at 0 since the  $\Delta t = t - t_s$ . Where the lines intersect is where the design particle is in phase space.

Design for maximum acceleration and maximum bunching are shown in A) and B) respectively with, unsurprisingly, drastically different phase spaces. In A), particles are mostly clumped near the design particle with a monotonic decrease in the particle density towards lower energy and higher  $\Delta t$ . This is expected. Since the design particle is maximally accelerated it will be at the very tip of phase space at the highest energy possible. Particles close to the design particle will follow the design particle but, at each gap, receive less ac-

celeration. With each pass through a gap, the density of particles near the design particle is reduced and the design particle continues to get further away in energy from the rest. If the simulation were to be run for more and more gaps, one would see the bin near the design particle shrink more and more while bins at higher  $\Delta t$  and lower energy would increase from these particles drifting away from the design particle. Eventually, only the design particle will remain near  $\Delta t = 0$  at maximum energy. The transmission is also the lowest we should expect assuming we are not actively trying to lose particles. Maximum bunching is shown in B) with 88% transmission compared to A)'s 41% transmission, although B) has the lowest energy gain with most particles being at 15 keV or lower. Although the phase space in B) is malformed, there are some telling features. The phase space wrapping shown is indicative of the focusing effect. Recall that in a perturbative acceleration case, this phase space would look like an eye. Such strong acceleration distorts the shape but we still see this wrapping around the design particle. We verified this by running cases where the acceleration was perturbative for many gaps and were able to see the iris shape. By iteratively simulating the lattice with increasing  $V_g$  we saw the iris distort into the present shape. Also note the large portion of particles missing between  $-0.4\tau_{\text{rf}} \leq \Delta t \leq 0.2$ . This is due to the strong acceleration since these earlier particles are going to see a strong decelerating field and will get pushed to negative energy values thus being filtered from the diagnostics. The small number of particles that made it through the lattice and were not filtered by time can only be due to the fortuitous combination of phasing. These particles were able to keep up with the design particle in time till, most likely, the last gap where they were mostly out of phase and experienced a strong deceleration once more. Although maximum bunching occurs when choosing  $\phi_s = -\pi/2$  one must be cautious when in the non-perturbative regime. We will show momentarily that one cannot rely on this design to give the best particle capture.

To the far right is C) where we attempted to strike the middle ground between acceleration and bunching by tapering the acceleration. In this case, the RF phasing was chosen to be  $\phi_s = (-\pi/2, -\pi/3, -\pi/6, 0)$ . The familiar golf club structure is present as discussed in Ch. 2. With this design, we can accomplish bunching and acceleration as noted by the 50.9% transmission and 20.1 keV. This has not been rigorously optimized to find the global optimum however and is merely a test case to show the methodology.

All three cases have a void in the phase space where there seems to be no data in contradiction to the histograms. This is due to the cold beam injection and was verified by simulating a hot beam with a thermal spread over the individual particle velocity that was a few percent of the beam velocity. When simulating the hot beam, there were no voids in the phase space distribution. This void arises because certain parts in phase space are not accessible due to the cold beam, strong acceleration, and high transit time factor. These conditions create a thin band in phase space rather than a contiguous distribution and these voids are part of the physics involved.

As a final example we simulated a 16-gap lattice and the results are shown in Fig. 6.5 where a new design D) Fast Tapering is also shown. This design is inspired from C) where, although there is excellent particle capture, the tapering is slow and the final energy is about half the maximum acceleration case. Therefore, we end up with a longer lattice without much acceleration which leads to a design where tapering is done for a few gaps and then we immediately switch to maximum acceleration. This is shown in D) where the first four gaps are tapered and the remaining gaps are at maximum acceleration. The phasing is  $\phi_s = (-\pi/2, -\pi/3, -\pi/6, -\pi/8, \dots, 0)$  where we introduce a useful notation of  $^k\phi$  which means that  $k$ -gaps are operating at the same phase  $\phi$ . A longer lattice exacerbates the phase space morphing as can be seen in A), B), and C). In A), the golf club structure is clearly

present and we see significant particle losses with clumping near the design particle as before and a longer band where these clumped particles eventually fall into after so many gaps. B) now shows a spiral shape which is indicative of non-linear effects. Still, the majority of particles are concentrated near the design particle as expected. C) is almost the same as before with some additional structure that is not too informative. The results of D) are surprising. Note the particle capture is better than B) and is why the phase choice of  $\phi_s = -\pi/2$  does not guarantee maximum particle capture when non-linearities are present. Additionally, the majority of the lattice is operating at maximum acceleration and we should expect a significant loss in particles. This effect is not an artifact of the reduced model and is present in the self consistent Warp simulations later on in (Ch. 8). This suggests the system is sensitive to longitudinal overfocusing and this fast tapering will inspire design choices later on. However, we will see that the gaps have strong transverse focusing effects (see Ch. ??) and this tapering design will have to be tweaked (see Ch. ??). Nonetheless, we have an efficient code that allows us to probe this reduced parameter space and gain insight into the non-linear and strong acceleration effects present. These results also show the complex physics involved that show up in a reduced model and are only a small portion of the complexities that arise in the transverse motion and the self consistent treatment. Thus, one must be cautious when tuning these parameters since intuition, like those formed in Ch. 2 will only get one so far.

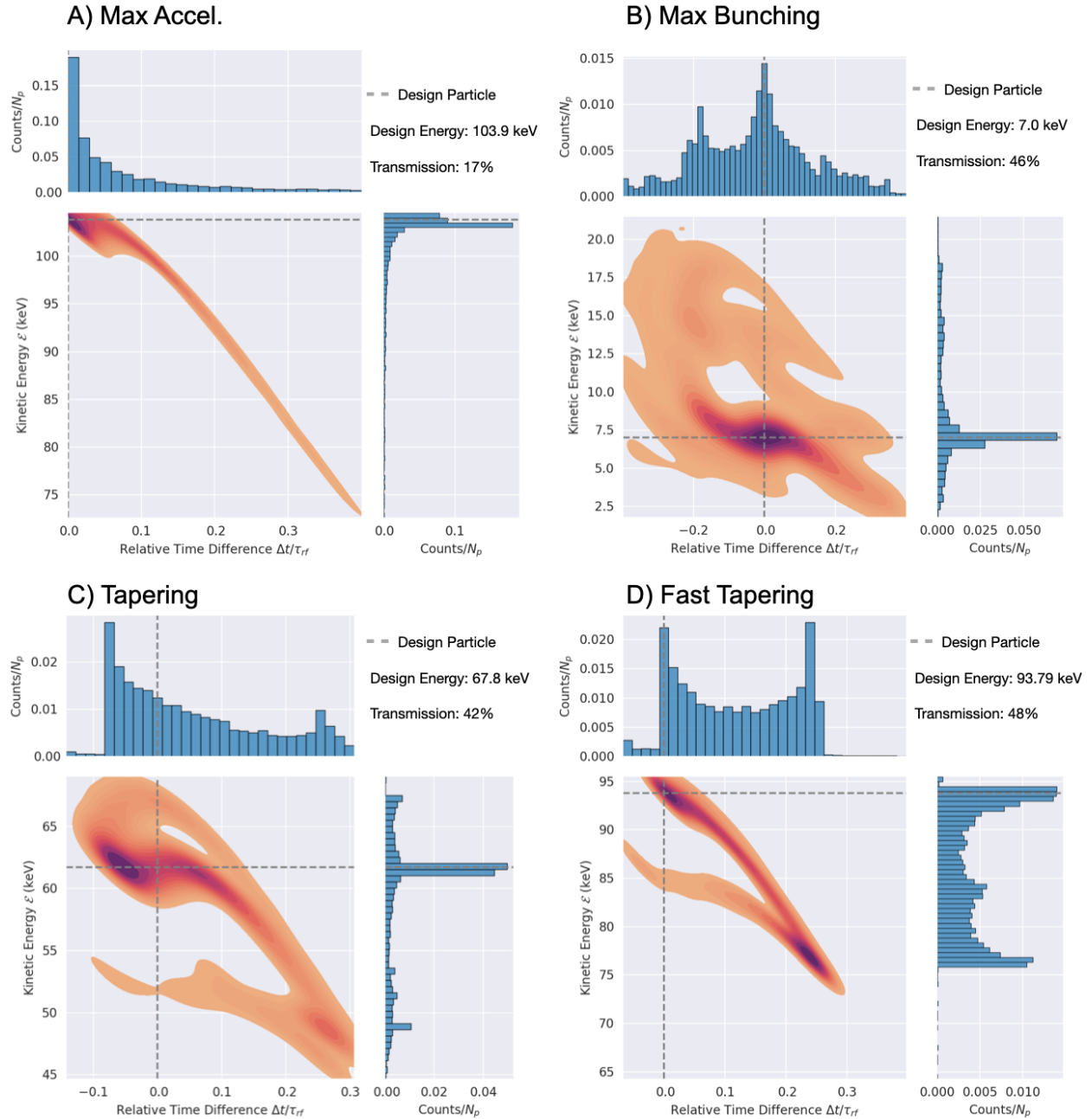


Figure 6.5: Similar to Fig. 6.4, these plots are for a 16 gap system and D) has been included for a new design where the gaps are tapered for a small number of gaps before the gaps are design for maximum acceleration.

## 6.4 Summary and Concluding Remarks

This chapter discussed the development of a reduced longitudinal code that uses the extracted on-axis electric field to simulate the acceleration of particles along  $z$  in 1D (no transverse motion). The good agreement with Warp gives confidence to the simulation and we can expect that the model is a good approximation to the longitudinal dynamics in a self-consistent simulation for particles near the axis. Since the simulation advances particles spatially, we used filtering in time and energy to remove particles that drifted too far from the design particle and would be lost in a Warp code or in lab. We showed design cases for maximum acceleration, maximum bunching, and tapering. These results showed the nuances involved when designing in the non perturbative regime of this system and is a caution against using the intuitions formed in Ch. 2.

A great improvement to the reduced model would be to create an envelope model using the Neuffer distribution [31]. Recall that the presented model does not include space charge while creating an envelope model from the Neuffer distribution would be self consistent and analogous to the work done in Ch. 7. Thus, both the longitudinal and transverse motion would be modeled with an almost identical approach creating a nice symmetric simulation suite to model the dynamics of the system. In addition to an envelope model, better testing could be done to fine tune choices to experiment. That is, what time filtering should we be applying that best fits experiment and/or desired outputs? Is there a better criteria for considering particles lost? Lastly, a more rigorous optimization scheme would be very beneficial. The optimization is currently done by hand using intuition developed by looking at numerous plots and testing various designs and this is impractical and perhaps wrong. However, this depends on defining what is best. Although transmission and energy are two

intuitive metrics, we have seen how intuition can lead astray and the optimum found here will not generally translate to optimums in the transverse motion or the full 3D simulations in Warp.



# Chapter 7. Transverse Dynamics in MEMS-RF

In order to efficiently transport the extracted ion beam array emerging from the injector we wish to pre-shape the beam using a matching quadrupole section in order to rms-match the transverse beam envelope array to the acceleration lattice. The lattice is aperiodic due to the strong acceleration experienced by the beam in the early stages of acceleration. This will prevent us from remaining perfectly matched as the beam accelerates. However, pre-shaping the beam to *almost* match will help radially confine the beam and limit beam loss from radial excursions produced by the defocusing space-charge and thermal forces arising from the distribution of beam particles.

## 7.1 Geometric and Beam Parameters

The ESQs have geometric parameters: aperture radius  $r_p$ , rod radius  $R$ , physical length  $\ell_q$ , and effective length  $\ell_{\text{eff}}$ . Sect.5.4 found the optimal geometry had  $R = 1.304r_p = 0.717$  mm when choosing  $r_p = 0.55$  mm and  $\ell_q = 0.695$  mm. These values were used in the matching section unless otherwise specified. Also shown previously was that the fringe fields did not overlap when ESQ pairs were held at a distance (ESQ edge to ESQ edge) of  $d > 2.5$  mm. Choosing  $d = 3$  mm ensures that successive quadrupole fringe fields will not overlap. The injected beam parameters were calculated in an earlier section and are used here. These values are conveniently listed in Table 7.1.

Geometric Parameters	
$r_p$ (mm)	0.55
$R$ (mm)	0.717
$\ell_q$ (mm)	0.695
Beam Parameters	
$\mathcal{E}_b$ (keV)	7
$I$ $\mu\text{A}$	10
$T_b$ (eV)	0.1
$Q$	$6.99 \times 10^{-5}$
$\epsilon$ (mm-mrad)	1.34

Table 7.1: Geometric and initial-beam parameter settings used in the matching section simulations.

## 7.2 Numerical Integration of KV-Envelope Equations

We apply the Kapchinskij-Vladimirskij (KV) envelope equations (discussed in Sect. 3 to simply model the DC beam evolution in the matching section. Ignoring acceleration and denoting the transverse rms-edge radii of the envelope as  $r_x$  and  $r_y$ , the KV-envelope equations are:

$$\begin{aligned}
r_x'' - \kappa_x(z)r_x - \frac{2Q}{r_x + r_y} - \frac{\epsilon^2}{r_x^3} &= 0, \\
r_y'' - \kappa_y(z)r_y - \frac{2Q}{r_x + r_y} - \frac{\epsilon^2}{r_y^3} &= 0
\end{aligned} \tag{7.1}$$

with generalized perveance  $Q$ , rms-edge emittance  $\epsilon$ , and focusing strength in the transverse directions  $\kappa_x(z)$  and  $\kappa_y(z)$ . For a uniform round beam,  $Q$  can be defined in term of the beam current  $I$ , species mass  $m$ , and free-space permittivity  $\epsilon_0$  as

$$Q = \frac{qIm^{1/2}}{2^{5/2}\pi\epsilon_0\mathcal{E}_b^{3/2}}. \tag{7.2}$$

Similarly, the rms-edge emittance can be written as

$$\epsilon = 2r_b \left( \frac{k_B T_b}{\mathcal{E}_b} \right)^{1/2}, \quad (7.3)$$

where  $r_b$  is the beam radius, and  $T_b$  the beam temperature (see Sect. 3). Unless otherwise specified we assume that the focusing strength is equal and opposite in the two planes such that  $\kappa_y(z) = -\kappa_x(z)$  and  $\kappa_x(z) = \kappa(z)$ .

These coupled equations can be numerically integrated in the usual way by transforming them into a set of coupled first-order differential equations. We first define the following variables:

$$\begin{aligned} u_x &= r_x, & v_x &= u'_x = r'_x, \\ u_y &= r_y, & v_y &= u'_y = r'_y. \end{aligned}$$

Using the newly defined variables Eq.7.1 transforms to:

$$\begin{aligned} v'_x &= \kappa(z)u_x - \frac{2Q}{u_x + u_y} - \frac{\epsilon^2}{u_x^3}, \\ v'_y &= -\kappa(z)u_y - \frac{2Q}{u_x + u_y} - \frac{\epsilon^2}{u_y^3}. \end{aligned}$$

Discretizing the above formulas and rearranging leads to the update equations for a forward

differencing scheme:

$$\begin{aligned} v_x^{n+1} &= v_x^n + \Delta z \left( \kappa(z^n) u_x^n - \frac{2Q}{u_x^n + u_y^n} - \frac{\epsilon^2}{(u_x^n)^3} \right), \\ v_y^{n+1} &= v_y^n - \Delta z \left( \kappa(z^n) u_y^n + \frac{2Q}{u_x^n + u_y^n} + \frac{\epsilon^2}{(u_y^n)^3} \right). \end{aligned}$$

The above equations can be written more compactly if the terms in the parantheses are wrapped into a function  $f(z; \theta)$  where  $\theta$  represents a set of parameters  $\theta = \{\kappa(z), Q, \epsilon\}$  that are fed into the function. The coupled first order differential equations can then be written

$$\begin{aligned} v_x^{n+1} &= v_x^n + f^n, \\ u_x^{n+1} &= u_x^n + \Delta z v_x^{n+1}. \end{aligned}$$

Solving for  $v_y, u_y$  results in a similar expression. Note that that updated  $v_x$  is used in updating  $u_x$  which gives increased accuracy compared to a forward/backward Euler algorithm at little computational complexity. This scheme is known as the Euler-Cromer or Semi-implicit Euler scheme [32] and is used to solve the KV-envelope equations.

### 7.3 Modeling the RMS-Edge Envelope Evolution Without Acceleration

We construct a matching section from four ESQs. To prevent the fringe fields from overlapping, each ESQ is given 3 mm of spacing on both sides, i.e.  $d = 3$  mm. This spacing is done relative to  $\ell_q$  and not  $\ell_{\text{eff}}$ . An example of a simulation mesh is shown in Fig. 7.1. The

extracted  $\kappa(z)$  is shown in blue with the hard-edge equivalent model overlaid in green. At the top of the plot the physical ESQ extents are represented by black rectangles and the geometric variables shown. The normalization factor  $\hat{\kappa}$  is calculated from the 2D equation using the half-breakdown voltage limit  $V_q = 500$  V and the values listed in Table 7.1:

$$\hat{\kappa} = \frac{V_q}{\mathcal{E}_b r_p^2} \Big|_{V_q=500\text{V}} = 4.72 \times 10^5 \text{ m}^{-2}.$$

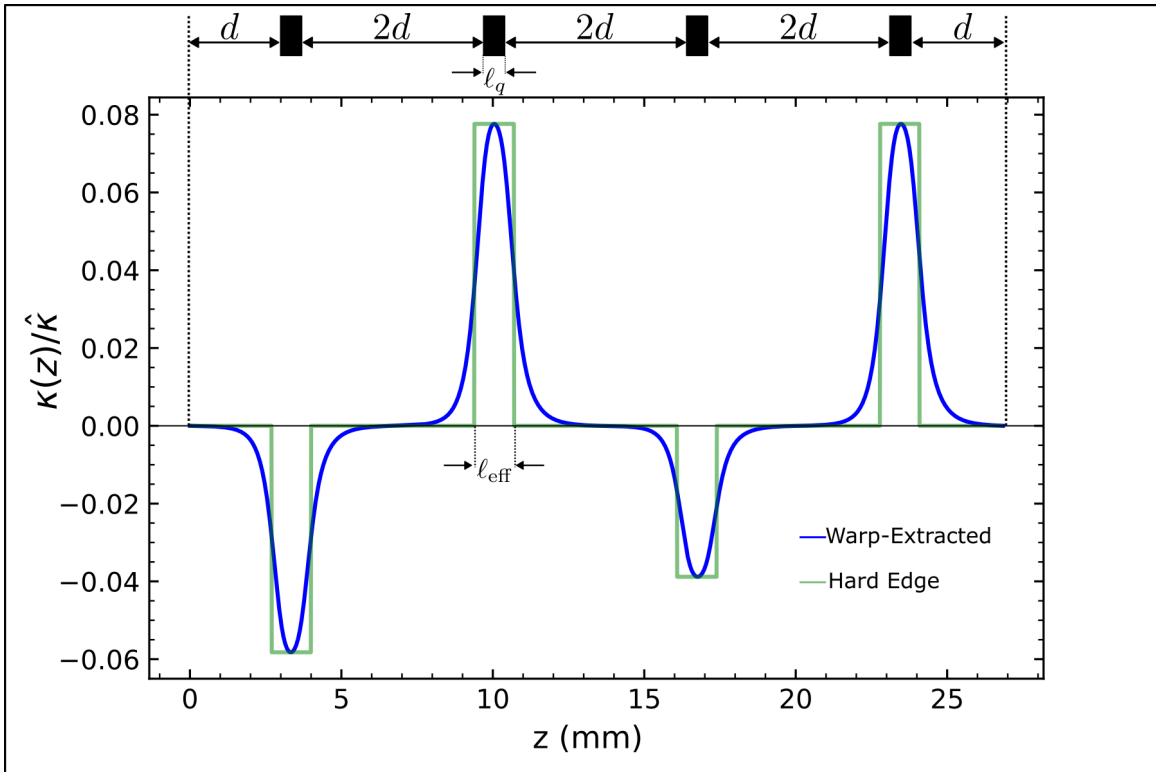


Figure 7.1: The extracted  $\kappa(z)$  in blue and the hard-edge equivalent in green are both normalized by  $\hat{\kappa}$ , the focusing strength calculated in the 2D case using the initial beam parameters and the half breakdown limit for  $V_q$ . At the top of the figure is a schematic showing the ESQ positioning along the mesh. The case shown has voltage settings (from left to right) of  $(V_1, V_2, V_3, V_4) = (-59.97, 79.94, -39.98, 79.97)$  V.

### 7.3.1 Benchmarking with Warp

Using the placement scheme shown in Fig. 7.1 the conductors are loaded onto a 3D mesh in Warp with periodic transverse boundary conditions for an array of beams and Dirichlet boundary conditions at longitudinal ends of the simulation box—simulations include beam space charge. A continuous beam of  $\text{Ar}^+$  ions are injected with the parameters listed in Table 7.1 along with zero initial angle to model the beam exiting the source aperture  $r_s = 0.25$  mm. Particles are injected uniformly and continuously at each time step to create a uniformly filled cylindrical beam that extends the length of the simulation box. The grid is discretized with  $dx = dy = 30 \mu\text{m}$ ,  $dz = 63 \mu\text{m}$ , and  $dt = 0.24$  ns—the grid resolution was so that higher resolution runs resulted in differences within numerical noise. To calculate the statistical evolution of the beam envelope in the full 3D beam simulation, the rms-edge statistics are calculated at each grid cell using the particle data for particles within the selected grid cell. Fig. 7.2 shows an example of the discretized mesh and an injected beam. Within each grid-cell, the moments are calculated and stored for that  $z$ -location. At the end of the simulation there will be of order  $10^6$  particles injected. Particles whose transverse position exceed  $r_p$  are absorbed and not included the statistical calculation. This could influence the statistical measure calculations. However, voltages were kept relatively low so that the envelope excursion was well with  $r_p$  and particle losses were not significant enough to impact the calculated measures.

We wish to compare the statistical evolution of the beam edge in a full 3D simulation with the evolution calculated by integrating the KV-envelope equations given in Eq. 7.1. The statistical moments of interest are the rms-edge radii and the change in the rms-edge radii along  $z$  given by Eq. 3.20. Each statistical measure can be extracted from Warp using

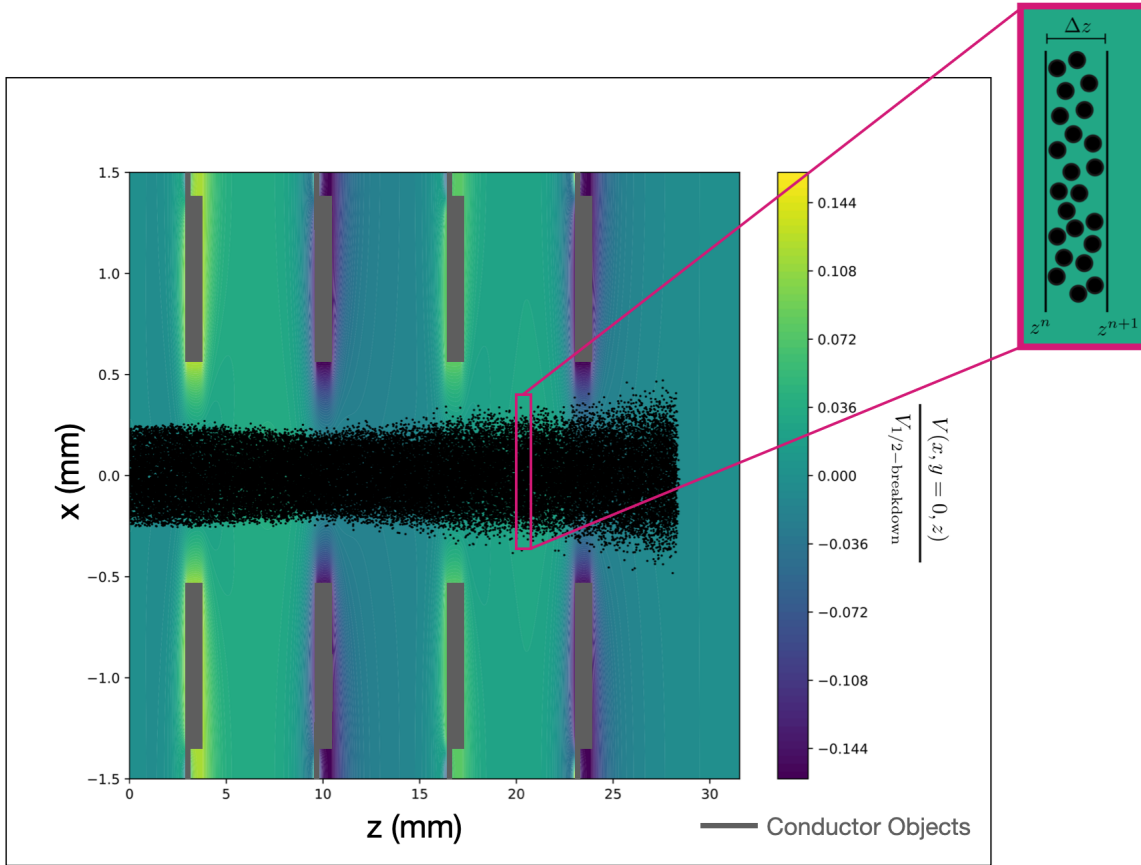


Figure 7.2: An example of a continuously injected beam and the discretized grid for which the moments are calculated. The potential is normalized by the half-breakdown limit  $V_{1/2\text{-breakdown}} = 500$  V and is generated by the ESQ conductors in gray. A random selection of 10,000 particles are plotted as black dots. The outer panel is a close-up view of the discretized grid. Particles within  $\Delta z$  of the two grid points are used to calculate the statistical moments which are then used to evaluate the statistical envelope evolution.

the correct call found in the “Z\_moments dump” section in the top.v file. For example,  $2\sqrt{\langle x^2 \rangle_{\perp}}$  can be extracted by calling `warp.top.xrmsz` after importing the Warp module.

Using the aforementioned tools and schemes we analyze the rms-edge evolution using the KV envelope model for an extracted Warp  $\kappa(z)$  and a hard-edge model; and a full 3D simulation using Warp that will advance particles in a potential field generated by ESQ conductors while including space-charge forces. The results for each model and the statistical measures in Eq. 3.20 are shown in Fig. 7.3. There is excellent agreement between the extracted  $\kappa(z)$  model (dashed line) and the full 3D Warp simulation (solid line). This suggests that rapid testing can be done with the extracted  $\kappa(z)$  and the results used without having to adjust for model differences and fine-tune with a full 3D simulation. The rapidity with the solver model will be especially advantageous when optimizing the voltages to pre-shape the beam for a matched condition. The hard-edge model shows some agreement but is slightly offset as is shown by the dotted line. Note that because the acceleration lattice will be asymmetric it will not be possible to perfectly match the beam to the acceleration lattice. In this regard, the slight offset in the hard-edge model may not be so detrimental or noticeable after a few lattice periods. The hard-edge model is the most computationally efficient since it does not require an initial solve of the ESQs to extract the gradient and therefore be the superior model in some cases. However, if the beam can be perfectly matched, then the small error could result in a mismatched beam and an inefficient transport early on in the lattice. Lastly, the disagreement is more apparent early in the simulation and with the  $x$ -coordinate. It is not fully known why but a potential reason is that the hard-edge model is not sufficiently capturing the focusing encountered in the extracted  $\kappa(z)$ . To test this hypothesis,  $\ell_{\text{eff}}$  was increased by 10% to provide more focusing and the bottom row of Fig. 7.3 shows the result. With the increased focusing the hard-edge model was in good agreement with the other



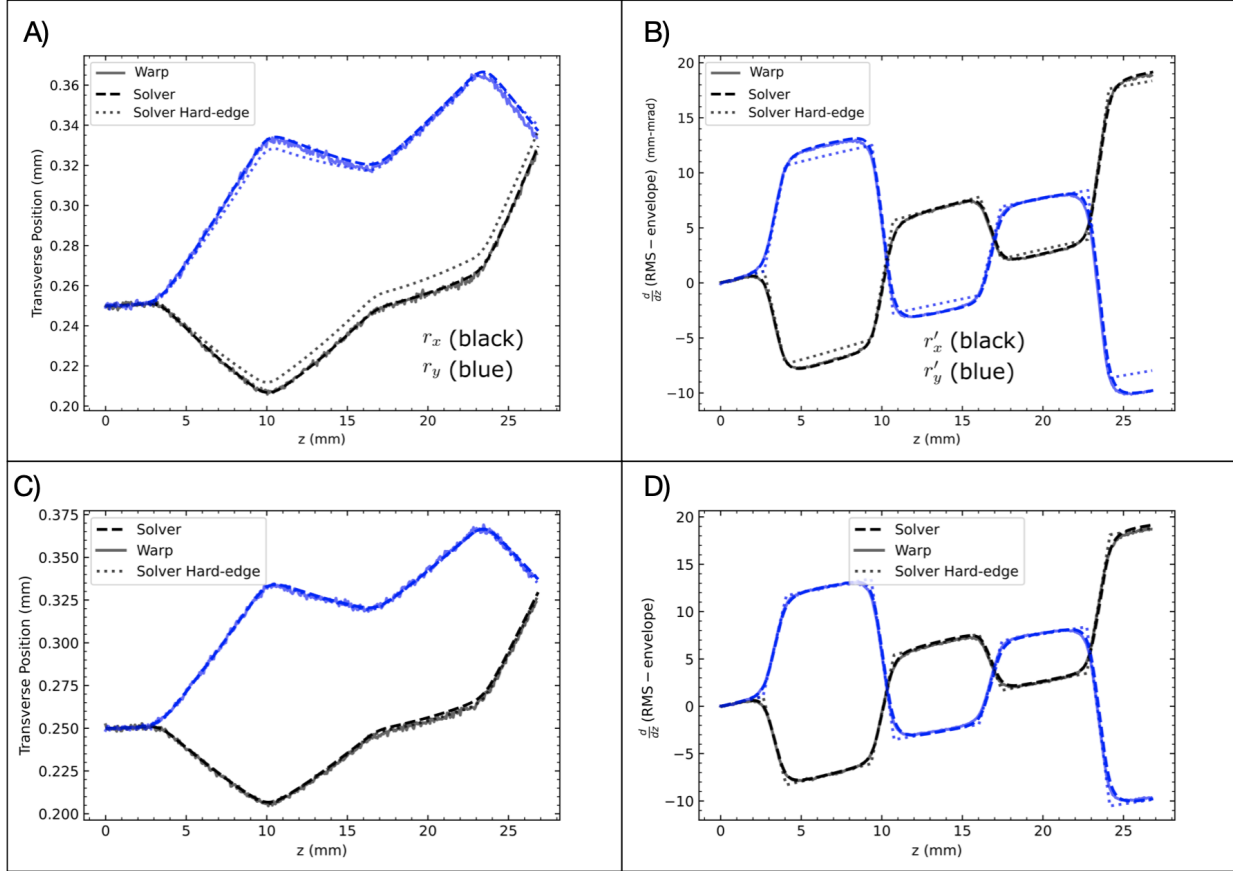


Figure 7.3: The rms-edge of the beam envelope boundary A) and C) and the envelope angles  $r'_x$  and  $r'_y$  of the beam edge B) and D). The  $x$ -coordinate is in black and  $y$ -coordinate in blue. Each model is shown on the plot: a full 3D simulation with space-charge done with Warp (solid line), the KV model using the extracted  $\kappa(z)$  (dashed), and the KV model using a hard-edge model (dotted). The full simulation and the extracted  $\kappa(z)$  are overlaid showing good agreement and suggesting that the extracted  $\kappa(z)$  using the KV model can serve in place of an expensive full 3D simulation. The bottom is a result of adjusting the effective length in the hard-edge model for more focusing. With this adjustment the hard-edge model agrees with the other two models.

two and suggests that the hypothesis was correct. If computation time was an optimization criterion these results suggest that one could first compare the extracted  $\kappa(z)$  model with a full simulation to show good agreement. From there, the hard-edge model could be fine-tuned with the extracted model and then used from then on to save computational time. In this case the computational time is not an issue and the extracted model is preferred since it captures the fringe field for analysis.

### 7.3.2 Finding Matched Conditions Using the Nelder-Mead Search Algorithm

With a tool capable of rapidly integrating the KV-envelope equations we now wish to extend the functionality to providing a matched condition for a given lattice. A previously developed Mathematica algorithm was used to benchmark [33] the developed algorithm. The Mathematica algorithm assumes that the lattice is periodic and, given various input parameters, the user can extract the matched solution among other quantities. Additionally,  $\kappa_q(z)$  can be given directly as an array of values along with the  $z$ -mesh and a match condition for  $r_x$ ,  $r_y$ ,  $r'_x$ , and  $r'_y$  will be found if one exists. The Mathematica tool does not allow for providing acceleration in the lattice and thus we are limited to comparing lattices with only ESQ focusing and constant  $Q$  and  $\epsilon$ . Fig. 7.4 shows an example output for a doublet configuration operating at  $V_q = 400$  V for the Mathematica algorithm. The top panel A) shows  $\kappa_q(s)$  where  $s \equiv z$ . Note that as previously stated  $\kappa_y = -\kappa_x$  and  $\kappa_x = \kappa_q$ . The bottom panels shows  $r_x$  (black) and  $r_y$  (red) in panel B). In panel C) are  $r'_x$  (black) and  $r'_y$  (red).

The matching was done for the first lattice period. The gap voltages were set to  $V_g = 5$  kV with an initial beam energy of  $\mathcal{E}_{b,i} = 7$  keV. Using maximum acceleration design the resulting

	$r_x$ (mm)	$r_y$ (mm)	$r'_x$ (mrad)	$r'_y$ (mrad)
Initial Condition	0.250	0.250	0	0
Match Solution	0.155	0.280	0.847	-11.146

Table 7.2: Initial conditions (envelope edge radii and envelope edge angles) used for initial condition is shown in the first row. The matched coordinates are shown in the second row.

lattice period was of length  $L_p = 25.148$  mm. A voltage of  $V_q = \pm 200$  V was used for the focusing doublet. Using the values for  $Q = 6.99 \times 10^{-5}$  and  $\epsilon = 1.34$  mm-mrad the

algorithm was able to find the matched solutions listed in Table 7.2. Fig. 7.4 shows  $\kappa(z)$  along the lattice in panel A), and the matched envelope evolution for the transverse position and angle in panel B) and panel C) respectively. With the matched condition found we now have a simulation experiment to compare to.

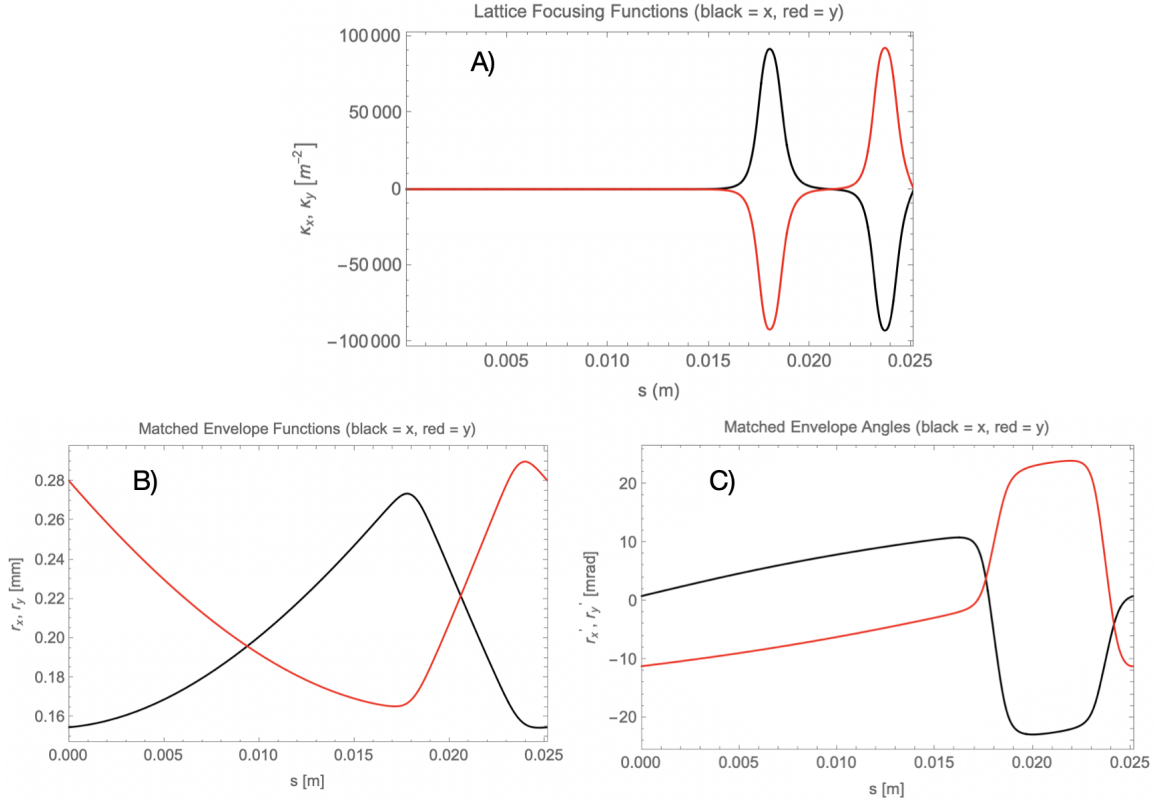


Figure 7.4: Matched solutions given by a previously developed Mathematica algorithm [33]. The software assumes a periodic lattice and does not account for acceleration. Panel A) shows the focusing strength from the ESQ conductor  $\kappa_q$  in the transverse direction ( $x$  in black,  $y$  in red) for single lattice period. Panel B) shows the evolution of the matched rms-edge radii  $r_x$  and  $r_y$  in black and red respectively over a single lattice period. Panel C) shows the evolution of the matched rms-edge angle  $r'_x$  and  $r'_y$  in black and red respectively over a single lattice period.

The algorithmic procedure is shown schematically in Fig. 7.5. We wish to find the matched coordinates given in Table 7.2 given the symmetric doublet voltages  $V_q = \pm 200$  V. Thus, the algorithm will be voltage fixed and the initial coordinates varied until matched condition is achieved. Writing the 4D coordinates in a compact vector form as  $\vec{r} = (r_x, r_y, r'_x, r'_y)$ ,

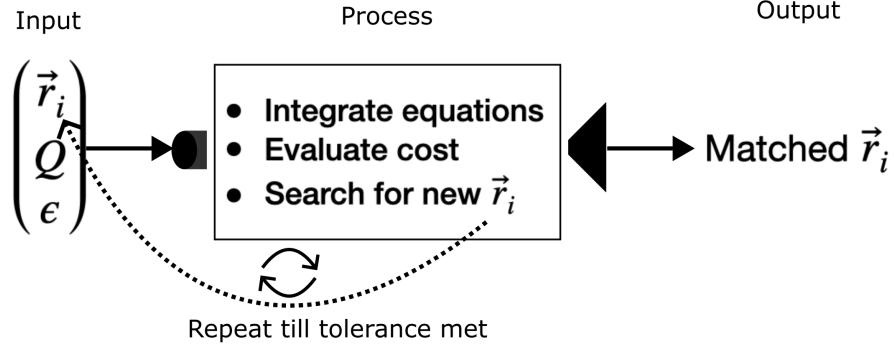
we can define the final coordinates extracted from integrating Eq. 7.1 as  $\vec{r}_f = (r_{x,f}, r_{y,f}, r'_{x,f}, r'_{y,f})$  and a set of target coordinates we wish to match as  $\hat{\vec{r}} = (\hat{r}_x, \hat{r}_y, \hat{r}'_x, \hat{r}'_y)$ . Note, in this case  $\hat{\vec{r}} \equiv \vec{r}_i$ . A cost function can then be defined as:

$$J(\vec{r}_f, \hat{\vec{r}}, \vec{S}; \vec{V}) = \frac{1}{2} \sum_{j=1}^4 \frac{(r_{j,f} - \hat{r}_j)^2}{S_j^2} \quad (7.4)$$

where the index  $j$  runs over the components of the vectors  $\vec{r}_f$ ,  $\hat{\vec{r}}$ , and  $\vec{S}$ . The vector  $\vec{S}$  is a set of normalization coefficients chosen so the algorithm can better converge. We made the natural choice of  $r_p$  to be the normalization factor for the position and through simulation we found a maximum envelope excursion value of  $r'_{\max} = 21$  mrad. With this we can define the vector of normalization values to be  $\hat{\vec{S}} = (r_p, r_p, r'_{\max}, r'_{\max})$ . We used the Nelder-Mead search algorithm to minimize the cost function [34, 35]. This is a robust and derivative-free searching algorithm that is advantageous here because of the low computational burden [36]. With an initial guess at  $\vec{r}_i$ , the algorithm will integrate the KV-equations with the set  $\kappa(z)$ ,  $Q$ , and  $\epsilon$ . Once integrated, the cost function will be computed using the extracted  $\vec{r}_f$ . This cost function informs the search algorithm on how to pick the next set of coordinates for minimizing the cost. This procedure will be repeated until the change in the cost function is below some set tolerance. Executing the algorithm produced the matched coordinates to within numerical noise verifying the algorithm. Additionally, the problem was switched to a coordinate fixed optimization where instead of finding the coordinates that matched to a given set of voltages, the algorithm instead searched for the voltages that produced the desired  $\vec{r}_f$  given  $\vec{r}_i$ . We obtained the necessary voltage of  $V_q = \pm 200$  V verifying that this algorithm can be used to also find the voltage required for a given condition. Different parameter settings were also used and then compared to the Mathematica algorithm to

## Voltage Fixed

$$\vec{r}_i = (r_{x,i}, r_{y,i}, r'_{x,i}, r'_{y,i})$$



## Coordinate Fixed

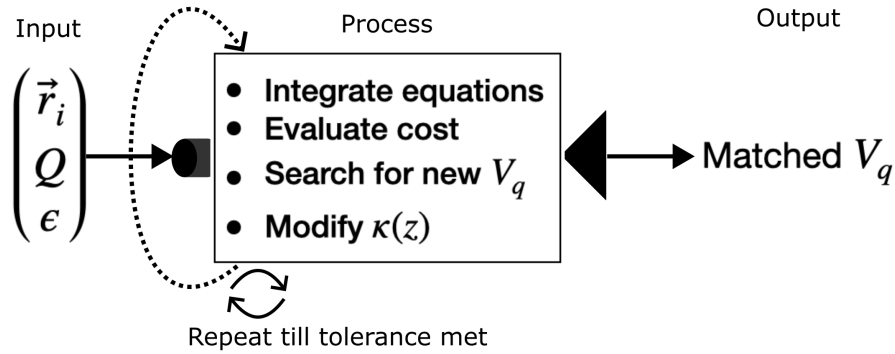


Figure 7.5: Flow diagram of the optimization algorithm for finding a matched condition. There are two types of matching that can be done. In voltage fixed matching, we seek a matched solution given the ESQ voltages and thus  $\kappa(z)$ . The algorithm will then iteratively perform the optimization process and vary the input coordinates  $\vec{r}_i$  until a set tolerance is met. In coordinate fixed matching, we seek a matched solution given  $\vec{r}_i$  and  $\vec{r}_f$ . The algorithm will iteratively perform the optimization process, vary the voltages, and recompute  $\kappa(z)$  until a set tolerance is met. Both schemes use the Nelder-Mead search algorithm and rely on a created cost function to guide the searching.

increase confidence.

## 7.4 Modeling the RMS-Edge Envelope Evolution with Acceleration

### 7.4.1 Creating a Benchmark with Warp

In the previous section Warp was used as a benchmark to establish credibility in the modeling of ESQ effects in the transverse evolution. We wish to do the same in modeling the acceleration. Two acceleration gaps are simulated in order to get the correct fields from the periodic array geometry (see Sect. 5.3) and the gaps are placed symmetrical about the midpoint of  $z = 0$  with enough spacing between the simulation box-ends and the midpoint to ensure there is no interference or fringe overlap. In stead of an RF-varying electric field, the time variation is turned off giving a static DC-field which is valid if the particle's gap traversal time is small compared to  $\tau_{rf}$ . Comparing the traversal time at the injection energy gives:

$$\frac{\tau_{rf}}{\Delta t_{\text{gap}}} \approx 0.1,$$

a weak assumption initially. However, the transit-time factor (see Sect. 2) can also be used as a gauge. Using the hard-edge approximation, the transit time factor at injection energy is:

$$T = \frac{\sin\left(\frac{\pi g}{\beta\lambda}\right)}{\left(\frac{\pi g}{\beta\lambda}\right)} = 0.964,$$

which suggests that the DC-assumption is valid. As before, a stream of particles is continuously injected into the system until the beam-head reaches  $z = 0$ . Like the previous section, the particle statistics can be extracted. The simulation extends from  $[-z_{\text{max}}, z_{\text{max}}]$

and so the extraction of Eq. 3.20 is done from  $[-z_{\max}, 0]$ . However, when plotting, the  $z$ -axis bounds are shifted to  $[0, z_{\max}]$  in order to avoid any confusion that may arise from a negative distance.

## 7.4.2 Using a Thin-gap Model to Model the Transverse Effects from Acceleration

The KV-envelope in Eq. 7.1 can also model the acceleration of the beam which changes the equations to:

$$\begin{aligned} r_x'' + \frac{(\gamma_b \beta_b)'}{(\gamma_b \beta_b)} r_x' - \kappa(z) r_x - \frac{2Q}{r_x + r_y} - \frac{\epsilon^2}{r_x^3} &= 0, \\ r_y'' + \frac{(\gamma_b \beta_b)'}{(\gamma_b \beta_b)} r_y' + \kappa(z) r_y - \frac{2Q}{r_x + r_y} - \frac{\epsilon^2}{r_y^3} &= 0. \end{aligned} \quad (7.5)$$

Here,  $\beta_b = v_b/c$  and the usual relativistic factor  $\gamma_b = (1 - \beta_b^2)^{-1/2}$ . In earlier sections  $Q$  and  $\epsilon$  were assumed constant as the beam progresses through the lattice. However, acceleration will change the beam energy and will therefore change the value of these parameters according to Eqs. 7.2 and 7.3. Scaling equations for these parameters can be derived if a thin gap model resulting in a thin lens kick to the energy. In this approximation, the beam radius does not change but the angle and energy are abruptly changed. Here, the notation is used so that a parameter value  $\chi^-$  is the value of  $\chi$  before the gap and  $\chi^+$  the value after the gap.

Starting with  $Q$ , the value before the gap is

$$Q^- = \frac{qIm^{1/2}}{2^{5/2}\pi\epsilon_0(\mathcal{E}_b^-)^{3/2}}.$$

and after

$$Q^+ = \frac{qIm^{1/2}}{2^{5/2}\pi\epsilon_0(\mathcal{E}_b^+)^{3/2}}.$$

The energy gain can be written as  $\mathcal{E}_b^+ = \mathcal{E}_b^- + \Delta\mathcal{E}_g$ , where  $\Delta\mathcal{E}_g$  is the energy change from the acceleration gap. Dividing  $Q^+$  by  $Q^-$  and substituting for  $\mathcal{E}_b^+$  gives:

$$Q^+ = \frac{Q^-}{(1 + \Delta\mathcal{E}_g/\mathcal{E}_b^-)^{3/2}}. \quad (7.6)$$

From the derivation of  $\epsilon$  in Sect. 3, one sees that for unchanged  $T_b$  and  $r_b$ ,  $\epsilon \propto \mathcal{E}_b^{-1/2}$ . Using the same steps for  $\epsilon^+$  and  $\epsilon^-$  gives the scaled rms-edge emittance as:

$$\epsilon^+ = \frac{\epsilon^-}{(1 + \Delta\mathcal{E}_g/\mathcal{E}_b^-)^{1/2}}. \quad (7.7)$$

Lastly,  $x'^+$  can be calculated:

$$\begin{aligned} x'^+ &\equiv \frac{v_x^+}{v_z^+}, \\ &= \frac{v_x^-}{v_z^- + \sqrt{2\Delta\mathcal{E}_g/m}}, \\ &= \frac{v_x^-/v_z^-}{1 + \sqrt{2\Delta\mathcal{E}_g/m(v_z^-)^2}}, \\ \therefore x'^+ &= \frac{x'^-}{1 + \sqrt{\Delta\mathcal{E}_g/\mathcal{E}_b^-}}. \end{aligned}$$

Note that  $y'^+$  will be of the same form. Eq. 3.20 gives  $r'_x = 2\langle xx' \rangle / \sqrt{\langle x^2 \rangle}$ . Since  $x^+ = x^-$  and  $y^+ = y^-$  is assumed, the update equation for  $x$  and  $y$  can be directly applied to  $r_x$  and



$r_y$ . Doing so and repeating the steps taken for  $Q$  and  $\epsilon$  gives:

$$r_x'^+ = \frac{r_x'^-}{1 + \sqrt{\Delta\mathcal{E}_g/\mathcal{E}_b^-}}, \quad (7.8)$$

with a similar expression for  $r_y$ . Eqs. 7.6, 7.7, 7.8 can be used to update  $Q$ ,  $\epsilon$ ,  $r_x'$  and  $r_y'$  respectively at the center of each gap. The initial envelope edge radii and edge angle were set to  $r_x = r_y = 0.25$  mm and  $r_x' = r_y' = 2.7$  mrad. The comparison is shown in Fig.7.6. With equal initial conditions the envelope evolution of  $r_x$  and  $r_y$  will be identical and therefore only  $r_x$  and  $r_x'$  plots are shown. The thin-gap model is shown as a dashed line while Warp is shown as a solid line. The gap edges are marked by thick vertical dashed lines. Clearly, the thin gap model does not capture the physics as can be seen by the monstrous disagreement between the two curves. A new model will have to be developed.

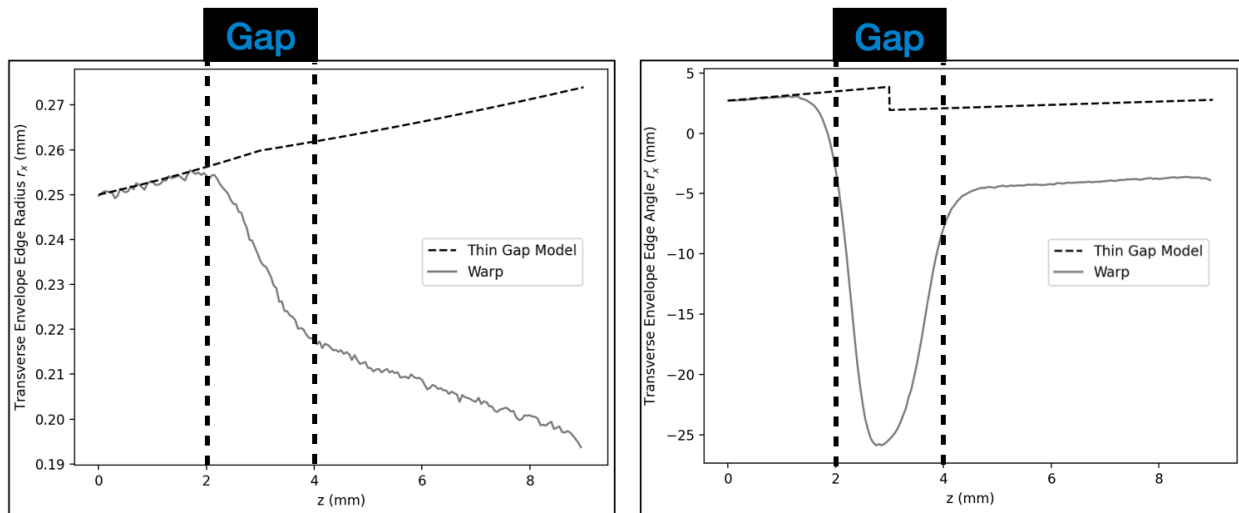


Figure 7.6: Comparing the envelope rms-edge evolution through a single DC acceleration gap between Warp and a thin-gap model. On the left is the rms-edge envelope radii and on the right the rms-edge envelope angle. The thin-gap model is given by a dashed line and Warp as a solid line. Dashed vertical lines mark the gap edges.

### 7.4.3 Using a Thick-gap Model to Model the Transverse Effects from Acceleration

In the thin-gap model Eqs. 7.6 and 7.7 are related to  $\mathcal{E}_b$ . In the electrostatic approximation, the energy changes as a function of  $z$  by

$$\mathcal{E}_b(z) = \mathcal{E}_{b,i} - qV_0(z) \cos(\phi_s)$$

where  $V_0(z)$  is the on-axis electric potential. As the beam traverses the acceleration gap the charge density  $\lambda$  will remain constant. Rewriting  $Q$  in terms of the charge density gives:

$$Q = \frac{q\lambda}{2\pi\epsilon_0 m \gamma_b^3 \beta_b^2 c^2}$$

where  $\epsilon_0$  is the vacuum permittivity. In the nonrelativistic limit,  $\gamma_b = 1$  and using the constancy of  $\lambda$  a scaling relation can be found. To do so, define the initial perveance before the gap in the non-relativistic limit as

$$Q_i = \frac{q\lambda}{2\pi\epsilon_0 m \beta_{b,i}^2 c^2}$$

Then take similiar steps as before by finding  $Q/Q_i$  to give:

$$\begin{aligned} Q &= \frac{Q_i}{\beta_b^2 / \beta_{b,i}^2}, \\ &= \frac{Q_i}{1 + \frac{qV_0(z) \cos(\phi_s)}{\mathcal{E}_{b,i}}}. \end{aligned}$$

where, the last line comes from using  $\beta_b^2 = 2\mathcal{E}_b/(mc^2)$  and the equation for  $\mathcal{E}_b$  above. It is useful to define a normalized potential as  $\tilde{V}_0(z) = V_0(z)/V_g$ . Rewriting with  $\tilde{V}_0(z)$  gives:

$$Q(z) = \frac{Q_i}{1 + \frac{V_g \cos \phi_s}{\mathcal{E}_{b,i}} \tilde{V}_0(z)}. \quad (7.9)$$

The rms-edge emittance can be found in a similar fashion. The constant is the normalized rms-edge emittance  $\epsilon_n = (\gamma_b \beta_b) \epsilon$ . Once again, the non-relativistic limit is used and using  $\epsilon_i = \epsilon_{n,i}/\beta_{b,i}$  to take a ratio gives:

$$\begin{aligned} \epsilon(z) &= \frac{\epsilon_i}{\beta_b/\beta_{b,i}}, \\ &= \frac{\epsilon_i}{\left(1 + \frac{qV_0(z) \cos \phi_s}{\mathcal{E}_{b,i}}\right)^{1/2}}. \end{aligned}$$

Rewriting with  $\tilde{V}_0(z)$  gives:

$$\epsilon^2(z) = \frac{\epsilon_i^2}{1 + \frac{V_g \cos \phi_s}{\mathcal{E}_{b,i}} \tilde{V}_0(z)}, \quad (7.10)$$

where the square was taken since it is  $\epsilon^2$  in the KV-equation.

Finally, the acceleration term must be treated:

$$\frac{(\gamma_b \beta_b)'}{\gamma_b \beta_b} r_x'.$$

We will consider  $r_x$  since the process will yield similar results for  $r_y$ . Again, use the non-relativistic limit to write:

$$\frac{(\gamma_b \beta_b)'}{\gamma_b \beta_b} r_x' \approx \frac{\beta_b'}{\beta_b} r_x'.$$

Writing  $\beta_b$  in terms of  $\mathcal{E}_b$  gives

$$\begin{aligned}\beta_b(z) &= \sqrt{2\mathcal{E}_b(z)/(mc^2)}, \\ &= \sqrt{\frac{2}{mc^2} \left( \frac{\mathcal{E}_{b,i} + qV_0(z) \cos(\phi_s)}{mc^2} \right)^{1/2}}.\end{aligned}$$

Taking the derivative with respect to  $z$  gives:

$$\beta'_b(z) = \sqrt{\frac{2}{mc^2} \frac{(q/2)V'_0(z) \cos(\phi_s)}{\left( \mathcal{E}_{b,i} + qV_0(z) \cos(\phi_s) \right)^{1/2}}}.$$

The acceleration term can now be written as:

$$\frac{\beta'_b r'_x}{\beta_b} = \frac{(q/2)V'_0(z) \cos(\phi_s)}{\mathcal{E}_{b,i} + qV_0(z) \cos(\phi_s)} r'_x.$$

Lastly, rewriting in terms of  $\tilde{V}_0$  gives:

$$\frac{\beta'_b r'_x}{\beta_b} = \frac{\frac{V_g \cos \phi_s}{2\mathcal{E}_{b,i}} \tilde{V}'_0(z)}{1 + \frac{V_g \cos \phi_s}{\mathcal{E}_{b,i}} \tilde{V}_0(z)} r'_x. \quad (7.11)$$

So far, all that has been done was take the thin-gap model scaling equations and mold them into a thick-gap formulation. An additional term comes from the radial component of the electric field. Here, the steps in [6] will be paralleled to derive the transverse components of the applied gap fields. Assume the applied static electric field is rotationally symmetric such that  $\partial/\partial\theta = 0$ ,  $\partial/\partial t = 0$ . Additionally, Using Maxwell's equations in vacuum, the applied electric field can be written as a scalar potential  $f^a(r, z)$  which obeys Laplace's equation. In

spherical coordinates, Laplace's equation is

$$\nabla^2 f^a(r, z) = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial f^a}{\partial r} \right) + \frac{\partial^2 f^a}{\partial z^2} = 0.$$

To solve, take a power series solution of the form

$$f^a(r, z) = \sum_{n=0}^{\infty} f_{2n}^a(z) r^{2n}, \quad \text{with } n = 0, 1, 2, \dots$$

An even series is assumed to ensure that the radial component of the electric field is zero on-axis  $E_r^a(r = 0, z) = 0$ . To see this, note that taking the gradient of the series solution will give the electric field. If there is a linear  $r$ -term, then taking the gradient with respect to  $r$  will give a constant field meaning that there will be a non-zero field at  $r = 0$ . Denoting  $' = \partial/\partial z$  and plugging the solution into Laplace's equation gives

$$\sum_{n=0}^{\infty} (2n)^2 f_{2n} r^{2(n-1)} + \sum_{n=0}^{\infty} f_{2n}'' r^{2n} = 0,$$

where the superscript  $a$  has been dropped reduce clutter. The first term is 0 for  $n = 0$  and the index can be advanced by 1 giving

$$\sum_{n=1}^{\infty} (2n)^2 f_{2n} r^{2(n-1)} + \sum_{n=0}^{\infty} f_{2n}'' r^{2n} = 0.$$

Reindexing the first term by letting  $m = n - 1$  gives

$$\sum_{m=0}^{\infty} 4(m+1)^2 f_{2m+2} r^{2m} + \sum_{n=0}^{\infty} f_{2n}'' r^{2n} = 0.$$

Since  $n$  and  $m$  are dummy indices, both sums can be written in terms of the same index  $m$ .

Doing so gives

$$\sum_{m=0}^{\infty} 4(m+1)^2 f_{2m+2} r^{2m} + \sum_{m=0}^{\infty} f_{2m}'' r^{2m} = 0.$$

For  $r \neq 0$  the only solution on where the coefficients equate to zero which establishes the recursive relationship:

$$f_{2m+2} = -\frac{f_{2m}''}{4(m+1)^2}.$$

The first terms in the series are:

$$f(r, z) = f_0(z) - \frac{f_0''(z)}{4} r^2 + \frac{f_0^{(4)}(z)}{64} r^4,$$

where  $f_0(z) = f(r=0, z)$ . The transverse electric field can now be extracted by taking the gradient. Doing so and truncating at the linear term in  $r$  gives:

$$E_r = -\nabla_r f(r, z) = \frac{f_0''}{2} r = -\frac{r}{2} \frac{\partial E_z}{\partial z}.$$

Converting to cartesian coordinates the transverse electric field components are:

$$E_x^a = \frac{1}{2} f''|_{r=0} x, \quad E_y^a = \frac{1}{2} f''|_{r=0} y. \quad (7.12)$$

This result is for any rotationally symmetric potential that meets the assumptions. Thus, substituting  $V_0$  for  $f_0$  gives:

$$E_x = \frac{V_g \cos(\phi_s) \tilde{V}_0''(z)}{2} x, \quad E_y = \frac{V_g \cos(\phi_s) \tilde{V}_0''(z)}{2} y. \quad (7.13)$$

From Sect.3 the ESQ focusing term was derived. Comparing the ESQ optic with Eq. 7.13 shows that another optic will be introduced from the applied gap field. Following the same steps a  $\kappa_g(z)$  can be found. The gradient is given by

$$G_{g,x}(z) = -\frac{\partial E_x}{\partial x} = -\frac{V_g \cos(\phi_s)}{2} \frac{\partial^2 \tilde{V}_0(z)}{\partial z^2}.$$

Plugging the gradient in to the optic equation gives:

$$\kappa_{x,g}(z) = \frac{qG_{g,x}(z)}{2\mathcal{E}_b(z)} = -\frac{qV_g \cos(\phi_s)}{4\mathcal{E}_b} \frac{\partial^2 \tilde{V}_0(z)}{\partial z^2}.$$

Lastly, expanding  $\mathcal{E}_b$  gives the focusing optic for the acceleration gap:

$$\kappa_g(z) = \frac{\frac{-qV_g \cos(\phi_s)}{4\mathcal{E}_{b,i}} \frac{\partial^2 \tilde{V}_0(z)}{\partial z^2}}{1 + \frac{qV_g \cos(\phi_s)}{\mathcal{E}_{b,i}} \tilde{V}_0(z)}, \quad (7.14)$$

with  $\kappa_{y,g} = \kappa_{x,g} = \kappa_g$ . It is useful to define coefficients to collect all the new terms for the thick-gap model. Using the following definitions:

$$\begin{aligned} C_{x,2} &= \frac{\frac{V_g \cos \phi_s}{2\mathcal{E}_{b,i}} \tilde{V}'_0(z)}{1 + \frac{V_g \cos \phi_s}{\mathcal{E}_{b,i}} \tilde{V}_0(z)} r'_x, \\ C_{x,2} = \kappa_q(z) + \kappa_g(z) &= \frac{\frac{q}{2\mathcal{E}_{b,i}} G(z) \Big|_{\text{quad}}}{1 + \frac{qV_g \cos(\phi_s)}{\mathcal{E}_{b,i}} \tilde{V}_0(z)} - \frac{\frac{qV_g \cos(\phi_s)}{4\mathcal{E}_{b,i}} \frac{\partial^2 \tilde{V}_0(z)}{\partial z^2} \Big|_{\text{gap}}}{1 + \frac{qV_g \cos(\phi_s)}{\mathcal{E}_{b,i}} \tilde{V}_0(z)}, \\ C_{x,3} = 2Q(z) &= \frac{2Q_i}{1 + \frac{V_g \cos \phi_s}{\mathcal{E}_{b,i}} \tilde{V}_0(z)}, \\ C_{x,4} = \epsilon^2(z) &= \frac{\epsilon_i^2}{1 + \frac{V_g \cos \phi_s}{\mathcal{E}_{b,i}} \tilde{V}_0(z)}. \end{aligned} \quad (7.15)$$

The coefficients in  $y$  are almost identical except that  $C_{y,2}$  will have a  $-\kappa_q(z)$ . Eq. 7.5 can then be written in terms of the coefficients as:

$$\begin{aligned} r_x'' + C_{x,1}(z)r_x' + C_{x,2}(z)r_x - \frac{C_{x,3}(z)}{r_x + r_y} - \frac{C_{x,4}(z)}{r_x^3} &= 0, \\ r_y'' + C_{y,1}(z)r_y' + C_{y,2}(z)r_y - \frac{C_{y,3}(z)}{r_x + r_y} - \frac{C_{y,4}(z)}{r_y^3} &= 0. \end{aligned} \quad (7.16)$$

Fig. 7.7 shows the  $r_x$  and  $r_x'$  from integrating Eq. 7.16 (dashed-lined) and extracted from Warp (solid line) in the top row as was done in Fig. 7.6. The lower panel shows  $\epsilon(z)$  from the model (dashed-line) and from Warp (solid line). Note that space-charge is also included in the Warp simulation. The agreement is excellent and a tool that can rapidly model the transverse dynamics of beam's rms-edge evolution through both acceleration gaps and ESQs is readily available. Recall the discussion on the KV distribution from Ch. 3. Here, we discussed the validity in applying the KV envelope equations to a beam that does not have a KV distribution. Although the beam is not KV, we can form an equivalent beam and use the KV envelope equations as a low order approximation to the real beam so long that the assumptions are satisfied. A strong indicator that this approximation is valid is if the normalized rms edge emittance  $\epsilon_{n,x}$  is a conserved quantity. Fig. 7.8 shows  $\epsilon_{n,x}$  through a single gap with source conditions and  $V_g = 7$  kV. These are where the conditions for a conserved  $\epsilon_{n,x}$  will be least likely. According to the figure, we have a constant with fluctuations below 1% and we can therefore feel confident that using the KV envelope equations is valid as an approximation to the lab beam.



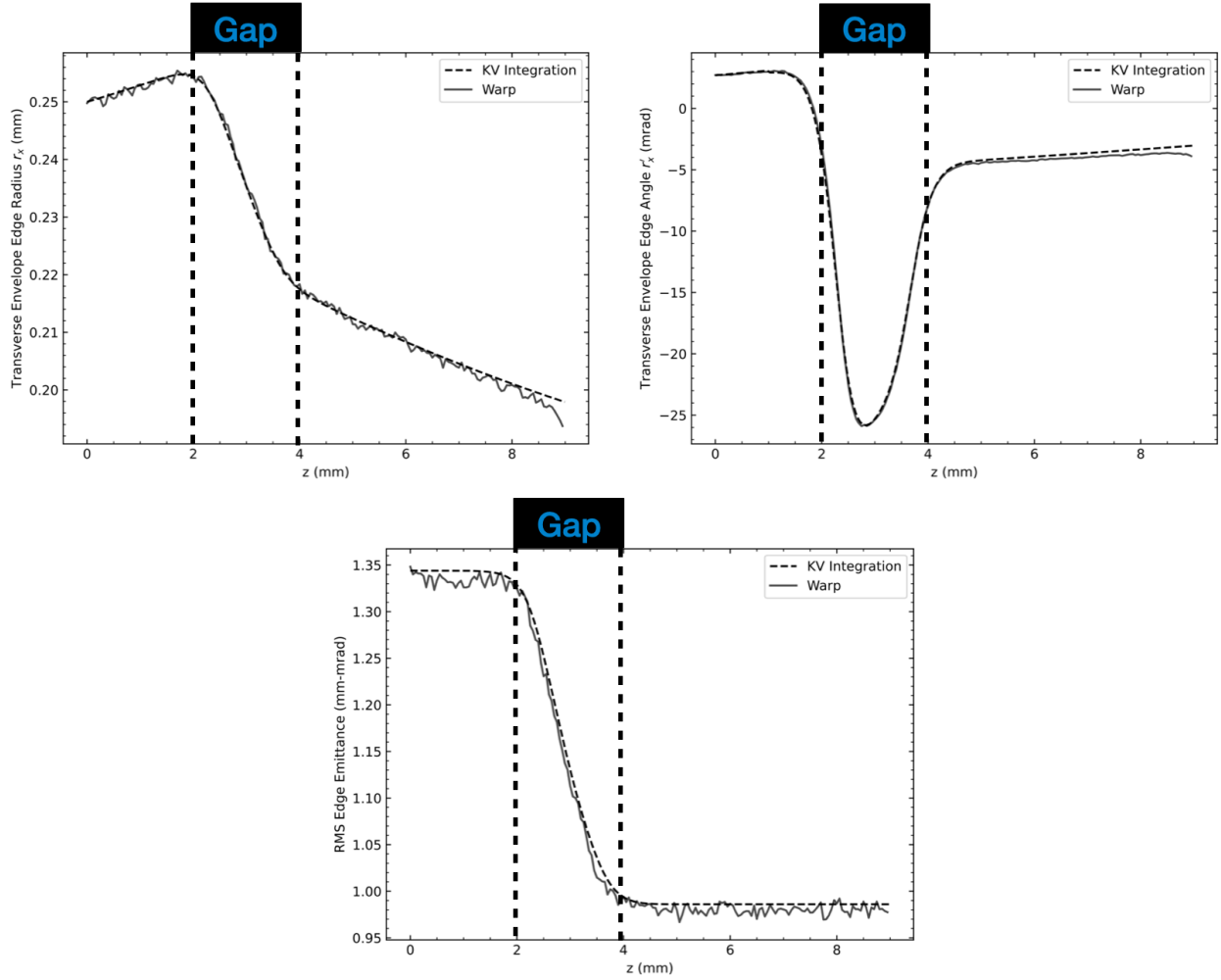


Figure 7.7: Comparing the envelope rms-edge evolution through a single DC acceleration gap with  $V_g = 7$  kV between Warp and a thick-gap model. (Top row): On the left is the rms-edge envelope radii and on the right the rms-edge envelope angle. (Bottom row): An overlay of  $\epsilon(z)$ . The thin-gap model is given by a dashed line and Warp as a solid line. Dashed vertical lines mark the gap edges.

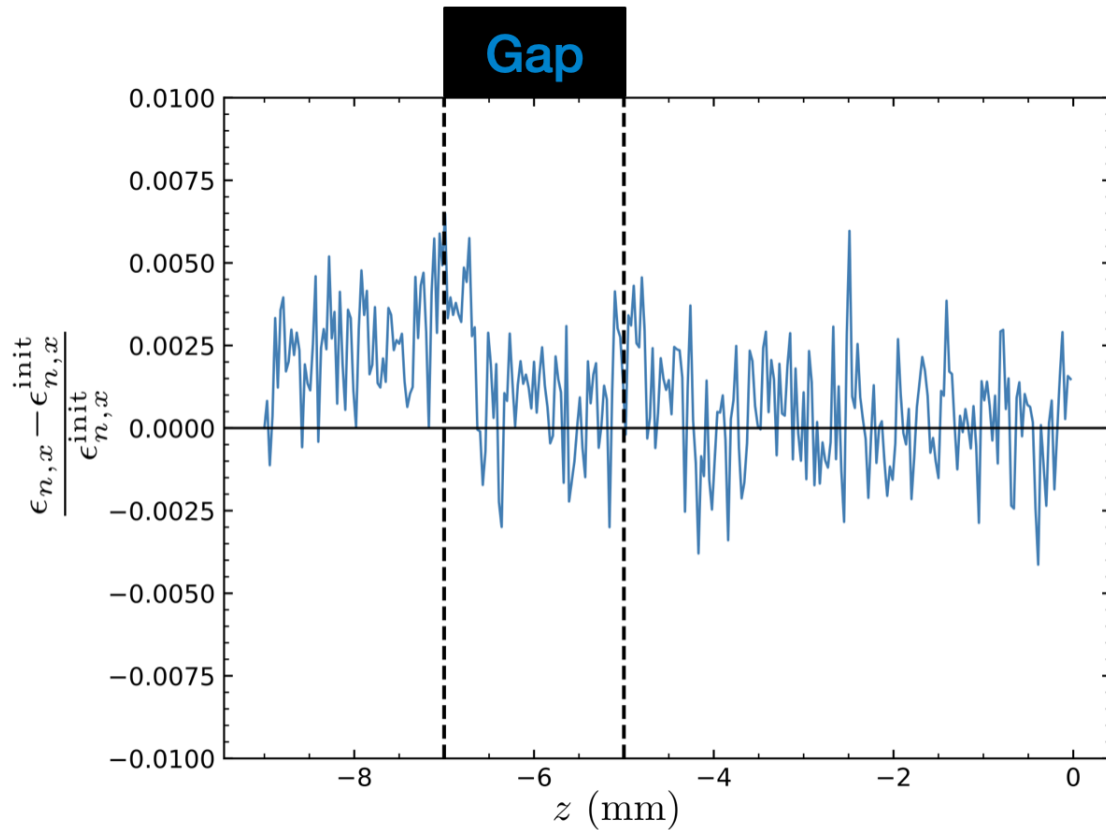


Figure 7.8: Fractional normalized emittance calculated at each grid-point normalized by the initial value. The gap is at 7 kV with a 7 keV beam representing the initial conditions from the source and where the acceleration is the strongest. Here, the normalized emittance fluctuates between values less than 1% from the initial value and can be considered conserve.

## 7.5 Lattice Design For Compact RMS Edge Envelope

The previous two sections demonstrated a model for capturing the physics of the beam's rms-edge envelope evolution and a search algorithm to find a matched condition. With these two tools available, a thorough investigation into the lattice design can be done. The following sections will create a point design for a 14-gap system. The system can be separated into three parts: 1) the pre-matching section comprised of only acceleration gaps, 2) the matching section comprised of four ESQs, and 3) the post-matching section comprised of acceleration gaps and ESQ focusing. Due to the high variability in the system, the design can be changed in multiple ways and the design presented is not proven to be the global optimum. What constitutes the best case is up to the user and these sections are meant to show how one may tune the lattice and quantify improvement.

### 7.5.1 The Pre-matching Section

From Fig. 7.7 we see that the acceleration gaps will provide a great deal of focusing in  $r_x$  and  $r_y$ . This suggests a lattice design that uses only acceleration gaps (no ESQ focusing) in the pre-matching section. The focusing effect from the gap does not last forever as depicted in Fig. 7.9. Here, the  $r_x$  and  $r_y$  are shown on the left panel in black and blue respectively, while  $r'_x$  and  $r'_y$  are shown on the right panel. With symmetric initial conditions the evolution in  $r_x$  is identical to that in  $r_y$  and similarly for  $r'_x$  and  $r'_y$ . Vertically dashed lines are the locations of the gap centers which are operating at maximum acceleration. We observe the strong focusing effect up to the fourth gap (approximately 31 keV beam energy by then) and here the beam begins to expand. This prompts the question: how long can we maximally accelerate until we can expect significant scraping of the beam? To answer this question,

the envelope code was used to iteratively calculate the  $r_x$  evolution ( $r_y$  will be the same) for a number of 16 lattice periods. Recall that each lattice period is comprised of two gaps and so this results in 32 gaps total. For each lattice period, the average  $r_x$  over the curve was calculated giving  $\bar{r}_x$ . Plotting  $\bar{r}_x$  over 16 lattice periods gives Fig. 7.10. We see that, after about 7 lattice periods (14 gaps and about 100 keV beam energy)  $\bar{r}_x > r_p$  and we can expect significant scraping. At 16 lattice periods (32 gaps and about 200 keV beam energy)  $\bar{r}_x \approx 2r_p$  and there little chance of getting a sufficient beam delivered to target.

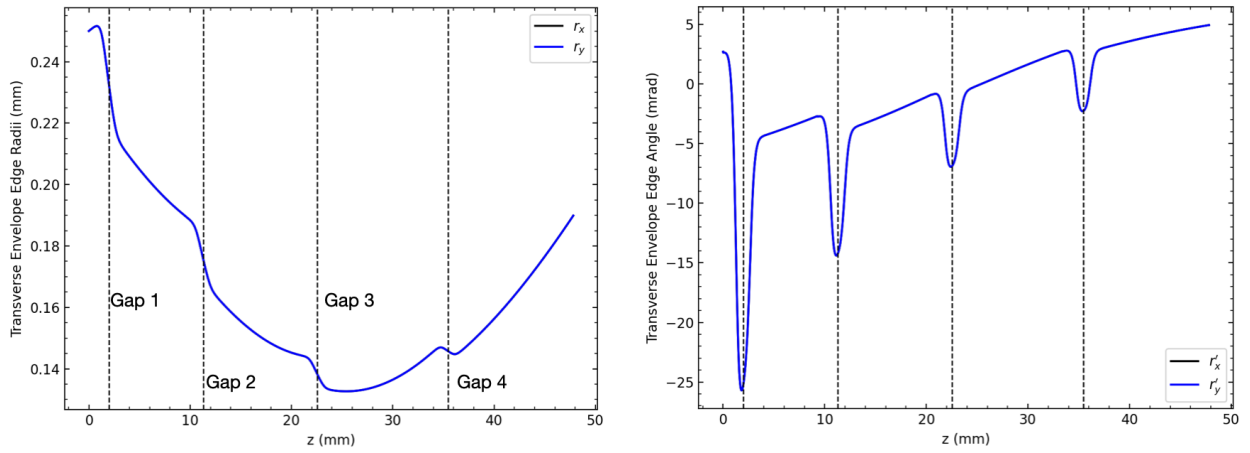


Figure 7.9: Transverse envelope edge radius and angle in  $x$  (black) and  $y$  (blue) through a six gap lattice at maximum acceleration. The gap centers are marked by vertical dashed lines and due to the injection symmetry the blue curve is directly overlaid the black curve. Transverse focusing from the gap is strong but after about four gaps begins to overfocus and the beam begins to expand. Additionally, the gap acts to decrease the angle of the beam gap-to-gap until the beam expansion after gap 4.

We sought to find a design that would longitudinally focus the beam and well-confine the beam. The design was found by hand-tuning the parameters to find a working case. Using the 1D longitudinal script, the design phase was chosen to be  $\phi_s = \begin{pmatrix} -\pi/2 & -\pi/3 & -\pi/6 & 30 \end{pmatrix}$ , where the notation  $^k\phi$  denotes an operating phase of  $\phi$  for  $k$  gaps. The resulting transverse dynamics are shown in Fig. 7.11. With this design choice, the necessary parameters can be extracted and are shown in Table 7.3 under the Pre-matching column. Parameter values

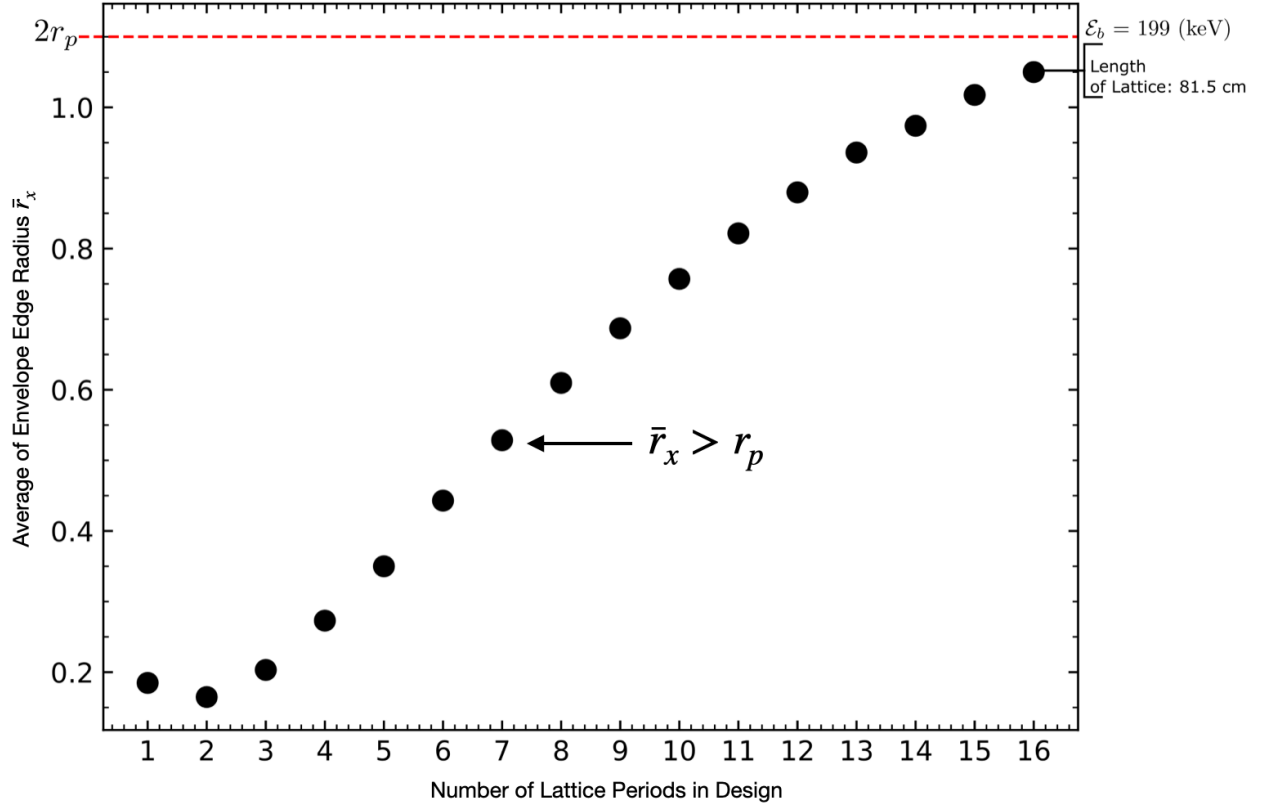


Figure 7.10: Using the envelope code, the evolution of  $r_x$  ( $r_y$  will be identical) was averaged to give  $\bar{r}_x$  for a design with a given number of lattice periods. For each lattice period,  $\bar{r}_x$  was calculated and plotted. At about 7 lattice periods,  $\bar{r}_x > r_p$  and significant scraping is expected. At about 16 lattice periods,  $\bar{r}_x \approx 2r_p$  and much of the beam can be expected to be lost.

$Q_f$  and  $\epsilon_f$  are necessary to correctly model the pre-matching beam in the post-matching section. Once modeled correctly, the matched-values  $\vec{r}_i$  in the post-matching section and the final coordinates in the pre-matching section can be connected via the matching section.

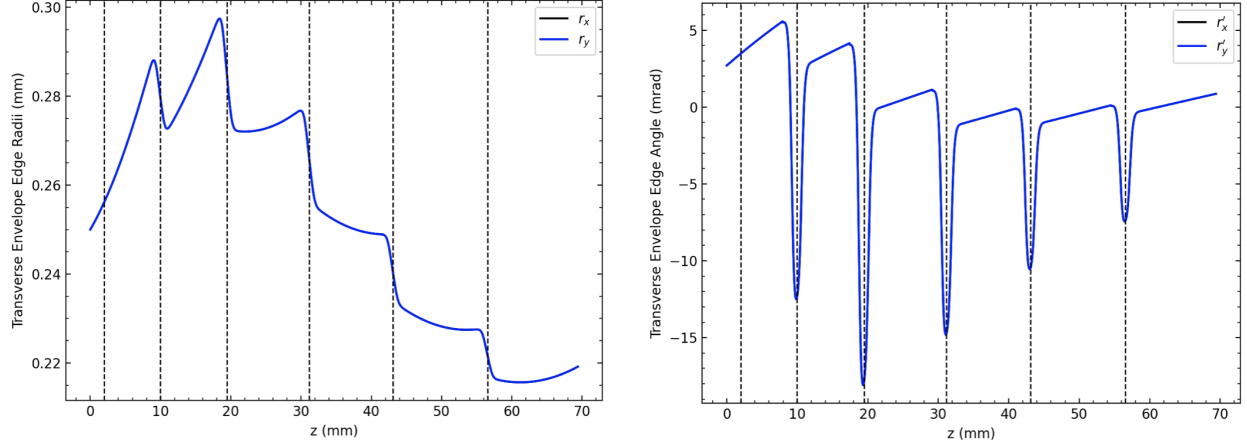


Figure 7.11: Transverse envelope edge radii (left) and angle (right) for  $x$  (black) and  $y$  (blue). Since the initial conditions are symmetric the two curves are overlaid. The gaps are chosen to operate with design phase  $\phi_s = (-\pi/2 \ -\pi/3 \ -\pi/6 \ 30)$  to give both longitudinal focusing and radial confinement.

## 7.5.2 The Post-matching Section

With the pre-matching section designed, we only need the values of  $Q$  and  $\epsilon$  of the beam that will be entering the post-matching section. These values are only changed by changes in energy (see Eq. 7.15). Using the values given in Table. 7.3 the search algorithm can be used to find a matched condition for a set of provided voltages. Note that the reverse conditions can be found, i.e. the voltages found to give the desired match coordinates. This is a more difficult approach for two reasons. Firstly, finding the voltages for the provided match coordinates will generally lead to asymmetric voltage solutions as was found in a cursory exploration into this approach. This can be hard to accomplish in experiment. Second, without knowing a priori what match coordinates are desired, we are forced into exploring a 4D parameter space (edge radii and angle) at worst. At best, a 2D parameter space if the edge radii and angle are kept symmetric. In the best case, one still faces the asymmetric voltage setting. Thus, the best approach in this instance is to use symmetric voltage parameters and search for the matched edge radii and angles. This reduces the hand-

tuning to a 1D parameter space and this space is constrained by the breakdown voltage limit which we are knowledgeable of. In tuning the design, the user must have a desired metric in order to establish what “good” is. It is not immediately obvious what this metric to use.

Once could use:

- Flutter:  $\frac{\text{Max}[r_x] - \text{Min}[r_x]}{\langle r_x \rangle}$
- Maximum rms-edge angle:  $\text{Max}[r_x]$
- Phase advance per lattice period:  $\sigma$
- Average rms-edge angle:  $\bar{r}_x$

We decided to use the average rms-edge angle  $\bar{r}_x$ . To do so, a voltage would be set and the KV equations integrated. We then calculated  $\bar{r}_x$  and  $\bar{r}_y$  from the envelope evolution. We sought voltages that would give  $\bar{r}_x \approx 0.28$  mm and  $\bar{r}_y \approx \bar{r}_x$ . Based on the calculated average, the voltage would be adjusted until the desired result. This method could be automatized but was easy enough to do by hand.

Fig. 7.12 shows match solution which satisfies  $\bar{r}_x \approx \bar{r}_y \approx 0.28$  mm. In the post-matching section, the gaps are operating at maximum acceleration ( $\phi_s = 0$ ). The ESQ centers are marked by black squares at the top and a vertical dotted line on the plot. On the left is the first focusing ESQ  $V_1 = 193$  V and the symmetric setting  $V_2 = -V_1$  following. Initial conditions are marked by stars and it is clear that the final coordinates match the initial. We can extract the match coordinates (Table 7.3 under Post-matching) and use the matching section to mold the beam so that the final coordinates of the pre-match section match the initial coordinates here.

Parameter	Pre-matching	Post-matching
$r_{x,f}$	0.219 mm	0.28 mm
$r_{y,f}$	0.219 mm	0.30 mm
$r'_{x,f}$	0.859 mrad	1.91 mrad
$r'_{y,f}$	0.859 mrad	-0.36 mrad
$Q_f$	$1.472 \times 10^{-5}$	$1.081 \times 10^{-5}$
$\epsilon_f$	0.617 mm-mrad	0.529 mm-mrad

Table 7.3: Final parameter values for the pre and post-matching sections. The matching section will have to match  $\vec{r}'_f$  from the pre-matching section to the post-matching section.  $Q_f$  and  $\epsilon_f$  from the pre-match section are used to correctly model the beam in the post-matching section and find match coordinates listed using the search algorithm.

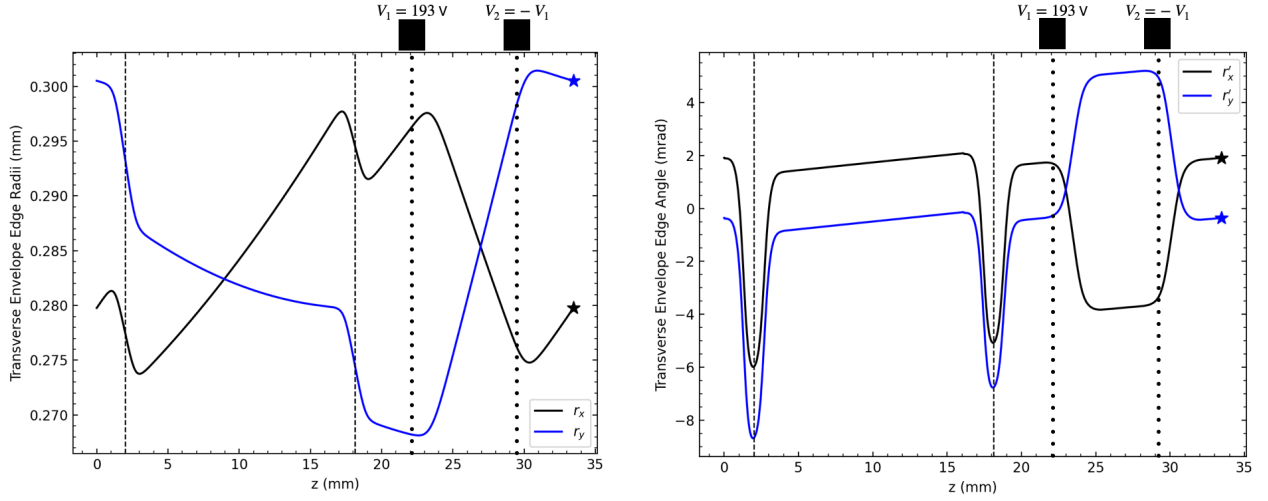


Figure 7.12: A matched solution for the first lattice period after the matching section. Acceleration gaps operating a  $\phi_s = 0$  are shown as vertical dashed lines. ESQs are shown as vertical dotted lines with a black rectangle at the top and the voltage used labeled. Parameters from the exit of the pre-matching section are used as the initial parameters in this post-matching section. The voltage setting was then iterated in until  $\bar{r}_x \approx \bar{r}_y \approx 0.28$  mm. Stars mark the initial coordinates values and it is clear that the match condition is reached since the curves start and end at the same points.

### 7.5.3 The Matching Section

Using the values in Table 7.3, the search algorithm can be used to match the pre-matching section to the post-matching section by finding the necessary voltages. In favor of compactness, we chose a matching section of four ESQs. Sect.5.4.2 gave an approximate 3 mm



separation distance between ESQ ends so that fields did not overlap. This distance was used to set a minimum bound on the placement. Additional space was added so that the effects of focusing in one plane had some space to take hold before being defocused in the opposite plane when entering the subsequent ESQ. A space of about 12 mm was chosen. This spacing was not optimized and there is room for improvement to possibly make the matching section shorter. Optimizing this design is not the priority however, and we merely wanted to establish the possibility of matching. The matched curves are shown in Fig. 7.13 on the bottom row. Since there are only ESQs present in this section, the dotted lines mark ESQ centers. The top image shows  $\kappa(z)$  along with the voltages found.

#### 7.5.4 The Rest of the Lattice After Matching

With the pre and post-matching section successfully interfaced via the matching section, the design analysis can proceed to the rest of the lattice occurring after the post-match section. Once again, there are two approaches:

1. Match neighboring lattice periods one-by-one for the entire lattice by providing the match coordinates and searching for the required ESQ voltages
2. Looking at a single lattice period, provide the ESQ for the beam parameters arriving at that lattice period and find a match condition using some metric as a guide

Due to the same reasons before, we proceed with the second approach and continue using  $\bar{r}_x$  and  $\bar{r}_y$  as a metric. Also, because  $\sigma \approx 13^\circ/\text{period}$  we expect the lattice to be insensitive to slight mismatch that will occur with this approach.

Table 7.4 gives parameter values for each tuned lattice. Each row represents a lattice period starting from the post-matching section. Using values from the previous lattice period

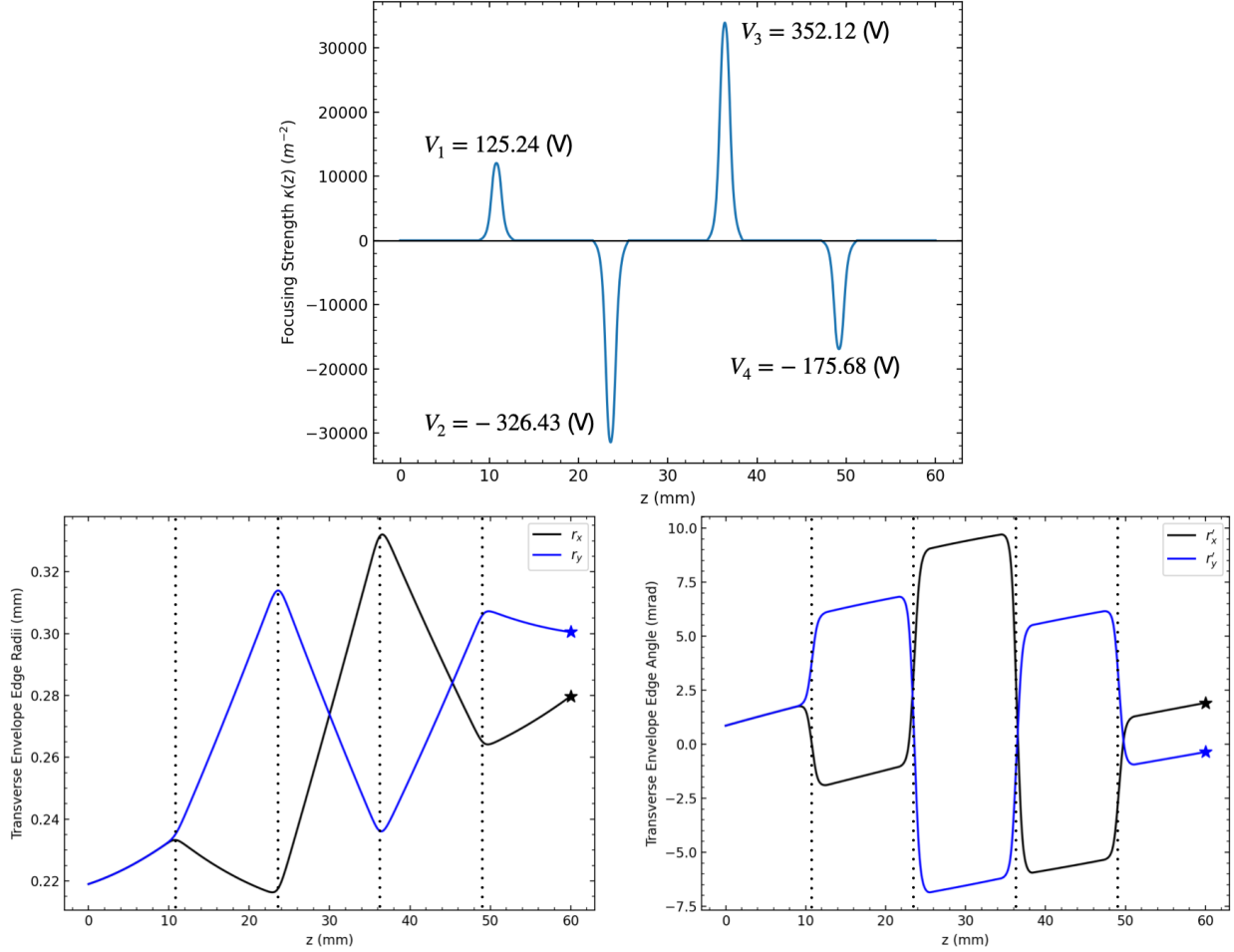


Figure 7.13: Transverse envelope edge radii (left) and angle (right) for  $x$  (black) and  $y$  (blue) for the matching section where there is no acceleration (bottom row). The four ESQ centers are shown as vertical dotted lines. Star markers are placed at the desired initial coordinates for the post matching section. The top figure shows the focusing strength  $\kappa(z)$  along the matching section and each ESQ voltage is labeled

$\mathcal{E}_f$ ,  $Q_f$ , and  $\epsilon_f$  the beam can be modeled in the current lattice period and the tuning done to find the voltage requirements. Each lattice period consists of two gaps operating at max acceleration. The transverse evolution of the beam for this tuned lattice is shown in Fig. 7.14. Due to mismatch, the beam envelope expands slightly which can be seen by comparing the terminal points of the curves with the star markers placed at the initial conditions. The stronger expansion in  $y$  is due to the mismatch being exacerbated as the beam is firstly defocused in  $y$ . This expansion can be countered by reversing the doublet polarity so that

$r_x$ (mm)	$r_y$ (mm)	$r'_x$ (mrad)	$r'_y$ (mrad)	$\mathcal{E}_i$ (keV)	$\mathcal{E}_f$ (keV)	$Q_i$	$\epsilon_i$ (mm-mrad)	$V_1$ (V)
0.2798	0.3005	1.9114	-0.3615	33.22	45.232	$1.4729 \times 10^{-5}$	0.6169	193
0.2745	0.3064	1.780	-1.1752	45.232	57.244	$1.0811 \times 10^{-5}$	0.5287	310
0.270	0.3056	1.581	-1.398	57.244	69.256	$0.8542 \times 10^{-5}$	0.5484	375
0.271	0.308	1.431	-1.437	69.256	81.268	$0.706 \times 10^{-5}$	0.4986	415

Table 7.4: Parameter values for a tuned lattice starting at the post-matching section. The tuning was done for each lattice period (each row) by extracting the necessary parameters  $\mathcal{E}_f$ ,  $Q_f$  and  $\epsilon_f$  from the previous lattice period to model the beam in the current lattice period and tune. The voltage holding is symmetric with  $V_2 = -V_1$  and was hand chosen so that  $\bar{r}_x \approx \bar{r}_y \approx 0.28$  mm. This methodology introduces mismatch between lattice periods but due to low phase-advance, we expect the lattice to be partially insensitive and the beam to maintain well-confined.

the first ESQ in the doublet is focusing in  $y$ . Although there is mismatch, the assumption made about the lattice sensitivity due to low  $\sigma$  is reasonable. This procedure can be repeated

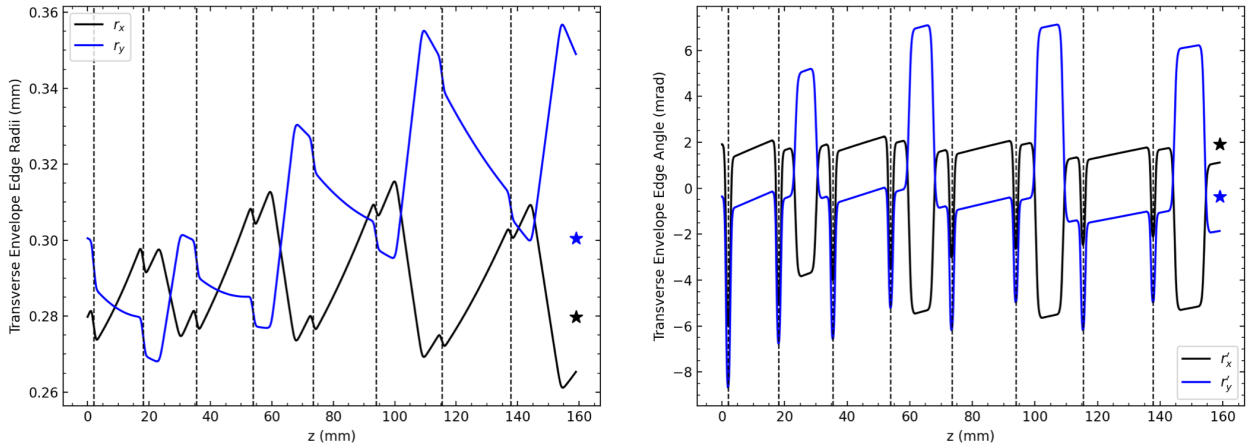


Figure 7.14: Transverse envelope edge radii (left) and angle (right) for  $x$  (black) and  $y$  (blue) for a full lattice design starting after the post-matching section. The acceleration gap centers are marked by vertical dashed lines with a  $\phi_s = 0$  for all. After every two gaps, there is an ESQ doublet to provide focusing—not shown in the graphic. Each period (two gaps) was tuned using the ESQ voltages to keep  $\bar{r}_x$  and  $\bar{r}_y$  approximately 0.28 mm. Star markers are placed at the initial conditions and due to low  $\sigma$ , the beam is well confined despite slight mismatch between lattice periods.

to get up to higher energies. We included additional calculations in the envelope code that are exported to a csv file if the user chooses. Fig. 7.15 is a snap shot of the data table where, due to the numerous columns, multiple snapshots had to be taken. This data was collected by performing the same  $\bar{r}$  tuning procedure and keeping  $\bar{r}_x \approx \bar{r}_y \approx 0.28$  mm. When

the focusing voltage approaching the half-breakdown limit, another “stack” was added to increasing focusing strength and decrease voltage requirements. This tuning was done up to 1 MeV and only a subset is shown here. Fig. 7.16 shows how the voltage scales at higher energies. The vertically dashed line shows the half breakdown limit and we can see that for beam energies above 200 keV, we will need to start incorporating ESQ stacks to increase the focusing provided. The plot goes up to 1 MeV which shows that focusing the beam at this energy will need to five stacks. It is immediately apparent the difficulty in deciding a good metric based on the amount of parameters that can be calculated which are all informative on design. This suggests an area of future study in deciding, and why, what the best metric is for guiding our lattice design.

max rx (mm)	min rx (mm)	max ry (mm)	min ry (mm)	max rxp (mrad)	min rxp (mrad)	max ryp (mrad)	min ryp (mrad)					
0.2965	0.2768	0.2988	0.2727	1.8727	-7.0230	4.6880	-8.9999					
0.3001	0.2707	0.3032	0.2660	2.0215	-5.2032	6.2610	-7.6829					
0.3028	0.2674	0.3056	0.2641	1.9322	-5.6454	6.4259	-6.6383					
0.3039	rx (mm)	ry (mm)	rxp (mrad)	ryp (mrad)	avg. rx (mm)	avg. ry (mm)	avg. rxp (mrad)	avg. ryp (mrad)				
0.3046	0.2845	0.2988	1.8202	0.1566	0.2865	0.2848	-0.0813	0.1110				
0.3061	0.2762	0.3007	1.6889	-0.7237	0.2845	0.2832	-0.1269	0.1343				
0.3082	0.2730	0.3014	1.4927	-1.0208	0.2838	0.2831	-0.1264	0.1246				
0.3104	0.2706	0.3010	1.3390	-1.1489	0.2829	0.2824	-0.1202	0.1140				
0.3107	0.2691	0.3006	1.2180	-1.1978	0.2823	0.2820	-0.1127	0.1053				
0.3114	0.2693	0.3014	1.1211	-1.1997	0.2830	0.2828	-0.1005	0.0930				
0.3102	0.2704	0.3028	1.0445	-1.1855	0.2845	0.2844	-0.0937	0.0856				
0.3058	0.2720	0.3046	0.9803	-1.1595	0.2864	0.2864	-0.0863	0.0785				
0.3133	0.2720	0.3046	0.9226	-1.1308	0.2865	0.2866	-0.0803	0.0730				
0.3108	0.2719	init E (keV)	final E (keV)	Lattice Period (mm)	V1 (V)	V2 (V)	Nstack	Vg (kV)	init Q	init emit (mm-mrad)	sigmax (deg/per)	sigmay (deg/per)
0.3111	0.2701	<b>30.0</b>	42.0120	32.1530	150.0	-150.0	1	6.042	1.629E-05	0.649	12.9461	13.1051
0.3132	0.2648	<b>40.0</b>	52.0121	36.0442	275.0	-275.0	1	6.042	1.222E-05	0.562	13.1166	13.2421
0.3100	0.2726	<b>50.0</b>	62.0120	39.5530	345.0	-345.0	1	6.042	9.76E-06	0.503	13.1757	13.2500
0.3139	0.2697	<b>60.0</b>	72.0121	42.7743	400.0	-400.0	1	6.042	8.13E-06	0.459	13.2539	13.3072
0.3133	0.2697	<b>70.0</b>	82.0121	45.7890	445.0	-445.0	1	6.042	6.97E-06	0.425	13.3049	13.3431
0.3142	0.2715	<b>80.0</b>	92.0121	48.5792	480.0	-480.0	1	6.042	6.1E-06	0.398	13.2519	13.2712
	0.2676	<b>90.0</b>	102.0121	51.2353	510.0	-510.0	1	6.042	5.42E-06	0.375	13.0978	13.1136
	0.2717	<b>100.0</b>	112.0120	53.7603	535.0	-535.0	1	6.042	4.88E-06	0.356	12.9308	12.9358
	0.2708	<b>110.0</b>	122.0121	56.1717	560.0	-560.0	1	6.042	4.4E-06	0.339	12.9016	12.8963
	0.2717	<b>120.0</b>	132.0121	58.4838	590.0	-590.0	1	6.042	4.07E-06	0.325	12.8995	12.9019
		<b>130.0</b>	142.0121	60.7078	620.0	-620.0	1	6.042	3.76E-06	0.312	13.0302	13.0312
		<b>130.0</b>	142.0121	60.7078	420.0	-420.0	2	6.042	3.76E-06	0.312	13.4930	13.4938
		<b>140.0</b>	152.0121	62.8531	420.0	-420.0	2	6.042	3.49E-06	0.3	12.7429	12.7425
		<b>150.0</b>	162.0121	64.9276	440.0	-440.0	2	6.042	3.25E-06	0.29	12.9970	12.9917
		<b>160.0</b>	172.0121	66.9378	455.0	-455.0	2	6.042	3.05E-06	0.281	12.9982	12.9954
		<b>180.0</b>	192.0120	70.7870	480.0	-480.0	2	6.042	2.74E-06	0.265	12.8296	12.8204
		<b>200.0</b>	212.0121	74.4375	515.0	-515.0	2	6.042	2.47E-06	0.251	13.1445	13.1397
		<b>220.0</b>	232.0121	77.9170	530.0	-530.0	2	6.042	2.25E-06	0.239	12.7586	12.7509
		<b>240.0</b>	252.0121	81.2477	555.0	-555.0	2	6.042	2.06E-06	0.229	12.8393	12.8317
		<b>260.0</b>	272.0121	84.4471	575.0	-575.0	2	6.042	1.9E-06	0.22	12.7578	12.7525

Figure 7.15: A snapshot of the various parameter calculations that can be saved in a csv file when using the envelope code. The tuning method discussed in this section was used to hand-tune lattice periods until an energy of 1 MeV was reached using approximately 6 kV gap voltage.

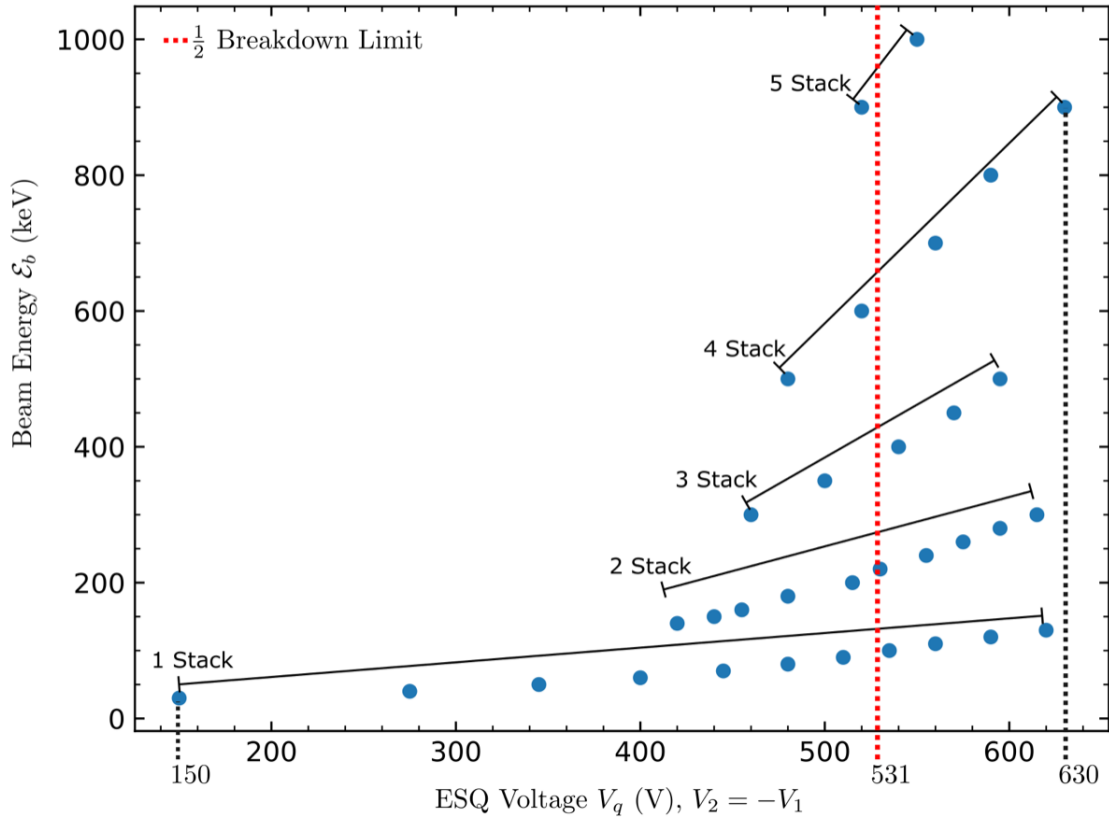


Figure 7.16: Voltage scaling for increasing ESQ stacks. As the beam energy increases, the required voltage will need to increase to focus the beam. The provided ESQ focusing can be increased by making the ESQ length longer. Instead of machining an ESQ with a different length, the system design allows ESQ wafers to be stacked to increase the focusing provided. The red vertically dashed line marks the half breakdown. This gives a guide for the ESQs when needing to focus beams at higher energies up to 1 MeV.

## 7.6 Summary and Concluding Remarks

This chapter detailed the development of a transverse envelope model to efficiently simulate the evolution of the beam envelope through a lattice design comprised of accelerating gaps and ESQs. Inspired by the 1D longitudinal model, the ESQ near-axis gradient was extracted and used to build  $\kappa(z)$  used in Eq.7.1. We benchmarked the developed tool against Warp in a gap-free design where only the ESQ effects were compared. The agreement was excellent as depicted in Fig. 7.3. We then used a previously developed Mathematica code [12] to assist in incorporating a search algorithm that would find matched conditions for the design using a defined cost function. The algorithm uses the Nelder-Mead algorithm [34] to search and allows the user to find a match condition in one of two ways. One way is to hold the voltages fixed and the algorithm will find the 4D coordinates (transverse position and angle) that will match to the provided  $\kappa(z)$ . The second way is to provide the 4D coordinate and the algorithm will find the necessary voltages. We then attempted to incorporate the focusing effects from the acceleration gaps by using a thin-gap model where the coordinates  $r_x$  and  $r_y$  were unaffected after the gap, but the parameters  $r'_x$ ,  $r'_y$ ,  $Q$ , and  $\epsilon$  would all receive a “kick” and change values after the gap. This model showed no agreement forcing a reevaluation and reformulation using a thick gap model leading to Eq. 7.16. Benchmarking with Warp once more in a gap-only design, the model showed excellent agreement as can be seen in Fig. 7.7.

With established confidence in our abilities to model the transverse effects from ESQ focusing and the acceleration gaps along with matching algorithm, we pressed on to find a lattice design for a compact envelope. In our search, we found that there is strong transverse focusing provided by the gaps up to a certain point before the beam begins to expand due to

over-focusing. Further investigation showed that using only acceleration gaps, the beam will expand so that  $\bar{r}_x > r_p$  and  $\bar{r}_y > r_p$  at around seven lattice periods ( $N_g = 14$ ) as shown in Fig. 7.10. Strong focusing from the gaps and the eventual re-expansion of the beam inspired a system design where the first six gaps are acceleration gaps only, followed by a matching section, and then subsequent lattice periods with periodic doublet focusing up to target. The 1D longitudinal script was used to find a candidate design that would provide bunching and acceleration in the first part of the lattice (Fig. 7.11). The matching algorithm was then used to match the first gap-only part of the lattice to the rest of the lattice that had periodic doublet focusing (Fig 7.13). Finally, the rest of the lattice was optimized to keep  $\bar{r}_x \approx \bar{r}_y \approx 0.28$  mm when designed for maximum acceleration and resulted in the voltage values listed in Table 7.4. This approach will lead to slight mismatch period-to-period but we expect the lattice to be insensitive to mismatch due to the small phase advance.

The envelope model can be improved by involving time-dependence. Recall that to benchmark the model, a static acceleration gap was used. This was justified since the RF-wavelength is much greater than the portion of beam we wish to model, i.e. we sought to model the physics of a portion of the beam rather than the entire CW beam. It is not obvious how to include time-dependence. One approach may be to first use the 1D script to advance a particle through a single gap with some design choice of  $\phi_s$ . For this design, the advancement can be done and the potential profile  $\tilde{V}_0(z)$  can be calculated. Note the coefficients in Eq.7.16 are functions of  $z$  through  $\tilde{V}_0(z)$  and its derivatives only (see Eq. 7.15). Once  $\tilde{V}_0(z)$  and its corresponding derivatives are computed, they can be used to integrate Eq. 7.16. This will only represent a small fraction of the overall beam since this was done for a single  $\phi_s$ . The process can be computed for a range of  $\phi_s$  to represent a pulse of a continuously injected beam and then combined to give a time-varying profile of the rms-edge



envelope. Another difficulty comes from how one would go about designing and testing this approach in Warp, for which the author does not currently have an idea.

# Chapter 8. 3D Self-Consistent Particle-In-Cell Simulations of MEMS-RF

## 8.1 Aside: Panel Plots and Simulation Snapshots

Before detailing various parts of the Warp simulations, it would be beneficial to thoroughly explain Fig. 8.1 since these plots will be used frequently without extensive labeling. Warp can output particle-plots to a computer graphics metafile (cgm). These cgms are collected and can be paged through to see the evolution of the beam. A separate computational tool was created to extract these frames and create an mp4 file. This file can be saved for later viewing and screenshots taken of the video used like the one here. The panels are numbered starting at the top left corner with 1 and progressing to the right before starting at the left column again on the second row at panel 4. Each panels is a contour plot for the electric field  $E_z(x, y = 0, z)$  and each plot is of the same scale. Although the  $x$ -axis is the same for each plot (and will be that way for all subsequent panel plots of this kind), the  $z$ -axis will change since the simulation window will move with the design particle (red dot in the plots). Additionally, we used this small two-gap simulations to highlight various objects that are in the simulation but not visible. In panel 1, the injection boundary is shown at  $z = 0$  ranging from  $\pm r_p$ . The horizontal lines mark an absorption boundary that acts as a “scraper” in the simulation—particles with  $r \geq r_p$  are considered lost. Acceleration gaps are labeled  $G_n$  with  $G_1$  being shown in panel 1. All panel plots will have conductors labeled  $G_n$  for gaps,  $MQ_n$  for a matching ESQ, and  $Q_n$  for a focusing ESQs in periodic lattice periods. Panel 2 and panel 3 show the injected beam. The injection is continuous for a whole RF period and

the full  $L_b$  is shown in panel 3 where  $t \approx \tau_{rf}$ . Panel 3 shows where the BPM for particle diagnostics is placed in the simulations at the midway point between  $G_2$  and  $G_1$ . Additional BPMs, which are not labeled to avoid confusion, are placed right at the injection point and right before the Fcup. Lastly, the Fcup is shown in panel 5 where particles are absorbed (panel 6) and the simulation continues to run for a specified time after the design particle reaches this point.

The next several sections will detail different parts of the simulation and make frequent reference to this plot. Subsequent sections will use plots of identical layout but with reduced labeling. Additionally, to enhance readability, these plots have been moved to Appendix A. Hopefully, this brief aside will enhance the understanding of the following sections.

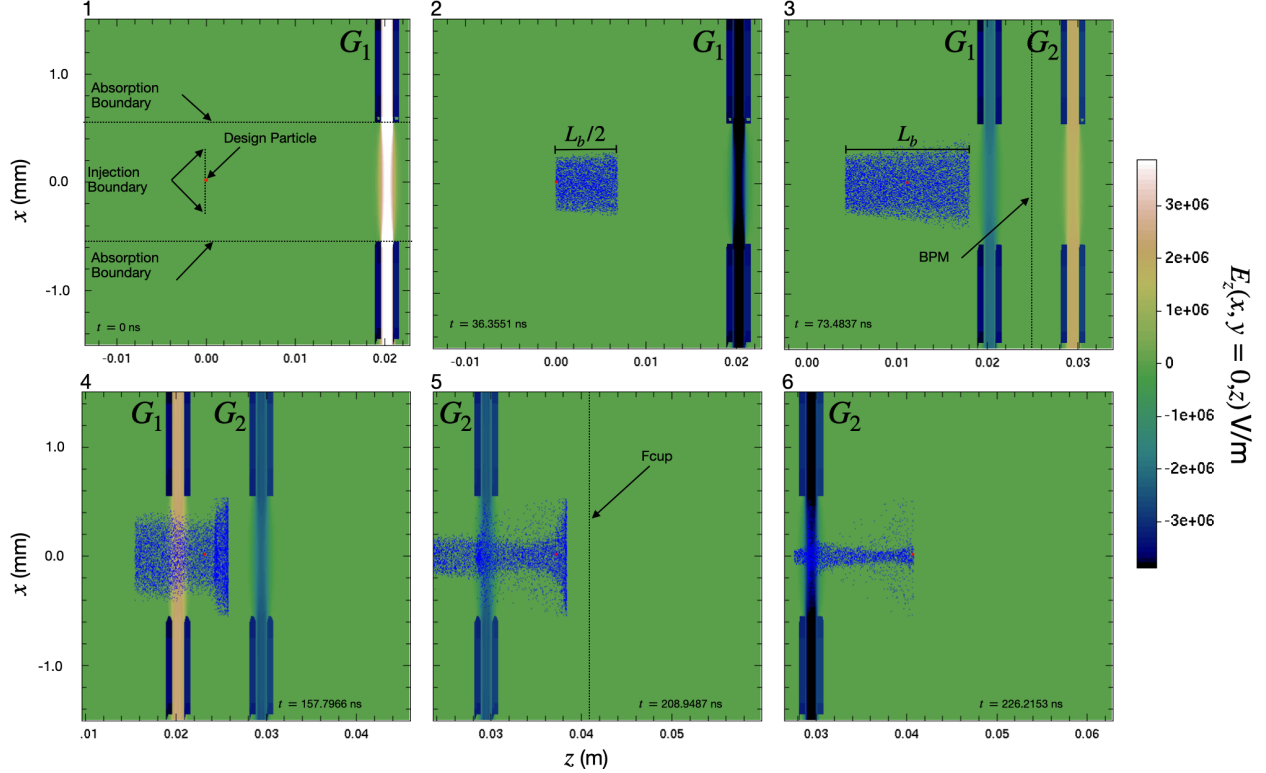


Figure 8.1: Labeled snapshots of the beam in  $x$  and  $z$  from a continuous injection simulation at different times labeled from 1 to 6 starting at the top left. Panel 1 is at  $t = 0$ . The design particle is shown as a red dot that starts at the injection site but is “disabled” until half the beam has been injected (panel 2). At an absorption boundary is placed at  $r = 0.55$  mm and is labeled in panel 1 with horizontal dashed lines. The gaps are labeled  $G_n$  and panel one shows  $G_1$  which starts at a maximum positive  $z$ -field indicated by the color bar to the right of the panels. Panel 3 shows the full length of the beam and the design particle has been switched “on”. Also note the change in the  $z$  values as the simulation window is now moving with the design particle. Between  $G_1$  and  $G_2$  is a BPM monitor that will record particle parameters. Lastly, panel 5 shows the location of the Fcup and the particles have been absorbed at the Fcup as seen in panel 6.

## 8.2 Beam Creation and Loading

A uniform cylindrical beam with flat longitudinal edges is injected (see panel 2 of Fig 8.1). Thus, for a continuous injection, the length of the beam will be  $\tau_{\text{rf}}$  which corresponds to a physical length of  $L_b = \beta_b c \tau_{\text{rf}}$ . The injection is done symmetrically around the design particle such that there is beam of particles of length  $L_b/2$  ahead and behind the design

particle. Using the Warp method `add_uniform_cylinder`, the beam loading can be done by providing the desired arguments as shown in Listing 8.1. By first determining how many macro-particles to put in the simulation ( $N_{p,\max}$ ), the first argument `np=Np_inject` is calculated by solving  $N_{p,\text{inject}} = N_{p,\max}dt/L_b$ . This will then inject the integer of  $N_{p,\text{inject}}$  every time step during injection. The radius of the beam is given by `rmax`. Each time step injects an axial slice of the beam with length `[zmin, zmax]`. `zmin` is given by the the specification of where the injection is done which is always set to  $z = 0$  in the simulation (i.e. `wp.top.zinject[0] = 0`). Therefore, the physical length of the injected slice is given by `zmax` which is calculated using the initial design velocity and time step. This setting ensures that there is no overlap of injected particles. Arguments `vthx`, `vthy`, `vthz` specify the thermal velocity spread to apply which were calculated earlier and showed that we have a cold beam. The design velocity is given by the argument `vzmean`. Out of the source there is no transverse velocity and these arguments are set to zero. Lastly, the angle of the beam is given by `vrmax` and has a radial scaling applied via Warp.

```

1 beam.add_uniform_cylinder(
2     np=Np_inject ,
3     rmax = 0.25mm ,
4     zmin= wp.top.zinject[0] ,
5     zmax = wp.top.zinject[0] + beam.vbeam + wp.top.dt ,
6     vthx = vthx ,
7     vthy = vthy ,
8     vthz = vthz ,
9     vzmean = beam.vbeam ,
10    vxmean = 0.0 ,
11    vymean = 0.0 ,
12    vrmax = beam.vbeam * div_angle ,

```

Listing 8.1: Sample of the input parameters for specifying a uniform cylindrical beam using Warp.

To correctly represent the physical particles in the system given  $N_{p,\max}$  macro particles, the particle charge weight  $w_p$  must be calculated correctly and specified. We use the desired source current  $I_s$  and calculate the  $w_p$  using:

$$w_p = \frac{I_s dt}{ZqN_{p,\text{inj}}},$$

where  $Z$  is the charge state and equal to one for singly charged ions. This will correctly specify the charge of each macro particle such that at each time step the slice of beam injected gives a current equal to  $I_s$ .

### 8.3 Simulation Mesh Configuration

To scale up in energy the length of the linac will increase to hold more gaps. In order to preserve the simulation quality and keep conductors resolved, the number of mesh points will need to increase as length is added to the system. This quickly becomes computational intractable and a moving window simulation is used instead to circumvent the issue. Warp allows one to specify a simulation box that can move at a specified speed. Conductors created downstream will be created once within the simulation window. This significantly reduces computational cost since the field solve for conductors does not need to be done in its entirety every time step. The window size is specified at initialization and is maintained throughout. Since the particles are set to be absorbed at the mesh boundaries (as well

as on conductors), it is important to set the upstream window distance enough so as not to immediately begin absorbing particles. The downstream window must be far enough to allow particle acceleration and evolution without being absorbed. Additionally, recall that in Sect. 5.3 it was shown that in order to get the right field structure two gaps must be in the simulation window. With these considerations, the upstream distance is set to be  $-\beta_b c \tau_{\text{rf}}$ . For the downstream distance, avoiding particle absorption can be done by specifying the same distance for the upstream window,  $\beta_b c \tau_{\text{rf}}$ . To ensure that there will always be two gaps within the window, the largest gap-to-gap distance is used which is the distance between the final two gaps within the lattice. Denoting this distance  $D$ , the upstream window is set to be at  $D + \beta_b c \tau_{\text{rf}}$ . This gives a total window size of  $D + 2\beta_b c \tau_{\text{rf}}$ , with the range being  $z = [-\beta_b c \tau_{\text{rf}}, D + \beta_b c \tau_{\text{rf}}]$ . Lastly, mesh refinement is used to increase resolution around the conductors. At each conductor (RF gaps and focusing elements if used) the transverse mesh is refined to give three times more grid points in  $x$  and  $y$  within the aperture. Longitudinally, the mesh is refined to give four times more grid points and extends over the physical longitudinal length of the conductor object.

## 8.4 Simulation Control

The simulation is controlled through a series of five while-loops. The first loop injects particles each time step with the injection scheme previously discussed. Once the time step is equal to or greater than  $L_b/2$  (hence the usefulness in describing the beam length in time) the design particle is activated and simulation window is set to move (see panels 2 and 3 in Fig. 8.1). While moving with the design particle, another while loop will continue injecting particles until the simulation time is equal to or greater than  $L_b$  which then turns off the

injection. A third while-loop is then entered which is the longest execution. This while-loop execution is wrapped in a for-loop. The while-loop executes until the design particle reaches a gap center. After reaching a gap center, the time-step is recalculated based on the maximally attainable velocity in order to satisfy the Courant-Friedrichs-Lewy (CFL) condition condition [37]:

$$dt = 0.7 \frac{dz}{\beta_{\max} c}. \quad (8.1)$$

Once the time-step is recalculated, the for-loop increments to the next gap-center and the while-loop repeated. After the design particle passes each gap, the next while-loop executes until the design particle reaches the Fcup and is absorbed (see panel 5 and 6 in Fig. 8.1). Finally, the last while-loop executes for an RF period to ensure slower particles are absorbed. The time to ensure absorption can be decreased to speed up performance. However, most particles are absorbed at a small fraction of the RF period and the remaining time-stepping is done relatively instantly compared to the main advancement sequence through the gaps.

## 8.5 Diagnostic and Data Management

Data collection is done through Warp's `ZCrossingParticles` class. This tool is analogous to a BPM monitor. The user specifies the  $z$ -location of the device and the longitudinal width. As particles cross the grid cells where the BPM is located the particle data is stored. Data for the BPM can be accessed through the getter attributes for which there are: `getx`, `y`, `ux`, `uy`, `uz`, `vx`, `vy`, `vz`, `xp`, `yp`, `r`, `theta`, `rp`, `t`, `n`. We chose to extract transverse coordinates (`x`, `y`), transverse angle (`xp`, `yp`), transverse and longitudinal velocities (`vx`, `vy`, `vz`), and the time of crossing (`t`). These diagnostics are placed at the midpoint between each two neighboring gaps (see panel 3 in Fig. 8.1). If a matching section is included, diagnostics are placed



at the midpoint between two neighboring ESQs. An additional diagnostic is placed a few grid cells from the injection point to capture the injection beam data and a final diagnostic is placed just before the Fcup to capture the delivered beam data. In general, there are  $(N_g - 1) + 2$  BPMs not including additional diagnostics placed for the matching section which depends on the number of ESQs used ( $N_g$ ). Once the particle advancement is completed (see Sect. 8.4) the diagnostics are looped through and particle data arrays stored in a dictionary since the data will generally be inhomogeneous—different size data arrays due to particle losses throughout simulation. The dictionary is then written to a pickle file for storage. Various data is calculated from the BPM monitors, plotted, and then saved into a single PDF file named “diagnostic\_plots.pdf.” These data routines can be easily modified by the user for desired metrics.

## 8.6 Pulsed Beam Injection

An initial study of the self-consistent simulations was done with a pulsed beam in order to better observe the dynamics. A  $20^\circ$  injection around the design particle was used with initial energy  $\mathcal{E}_i = 7$  keV and an initial angle of 3.78 mm-mrad in both transverse directions (diverging beam). The macro particle charge was calculated in order to ensure a  $10 \mu\text{A}$  when injecting  $10^5$  particles. To ensure that the on axis potential gain is 6 kV,  $V_g$  is slightly scaled up to 6.19 kV. In short lattice simulations this scaling is not as important but will muddy results in longer simulations due to “slippage” in the design particle synchronism.

We used a four gap lattice operating at maximum acceleration  $\phi_s = {}^40$ , where the notation  ${}^n\phi$  indicates  $n$  gaps at operating phase  $\phi$ , to establish a baseline. This design will be designated Case 1. Case 2 incorporates longitudinal focusing with  $\phi_s = (-\pi/2 \quad -\pi/3 \quad -\pi/6 \quad 0)$ .

Design Matrix

Beam	
$\mathcal{E}_b$	7 keV
Injection Type	Pulsed
$L_b$	$\tau_{rf}/18$
Lattice	
$N_g$	4
$V_g$	6.19 kV
$\phi_s$	$\phi_s = {}^40$

Table 8.1: Notable parameter values used. This design was for four gaps at max acceleration as noted by  $\phi_s$  and forms the baseline for comparing other design choices for pulsed injection schemes.

Case 3 is a modification of Case 2 with  $\phi_s = (0 \quad -\pi/6 \quad -\pi/6 \quad -\pi/2)$  for reasons that will be discussed momentarily. When comparing designs, we decided to compare three attributes of the delivered beam:

1. Transmission: Fractional number of particles, relative to injection, delivered on target
2. Current:  $I = N_{\text{bin}} w_p q / \Delta t$  calculated by binning the transmitted particles into equal time bins  $\Delta t$  (not to confused with the relative time difference used in longitudinal phase-space analysis) across simulations and then scaling by the macro-particle charge  $w_p q$
3. Radial distribution:  $r_i = \sqrt{x_i^2 + y_i^2}$  calculated for every  $i$ -particle at target and then binned.

Fig. 8.2 gives two plots for the radial distribution (left) and the calculated current (right). Each plot gives a histogram for the three cases: Case 1 in blue, Case 2 in orange, and Case 3 in green. The bins in each plot are of equal width.

An immediate question is what happened in Case 2? The answer is twofold. When the beam is compressed longitudinally it will expand transversely due to the increased space-

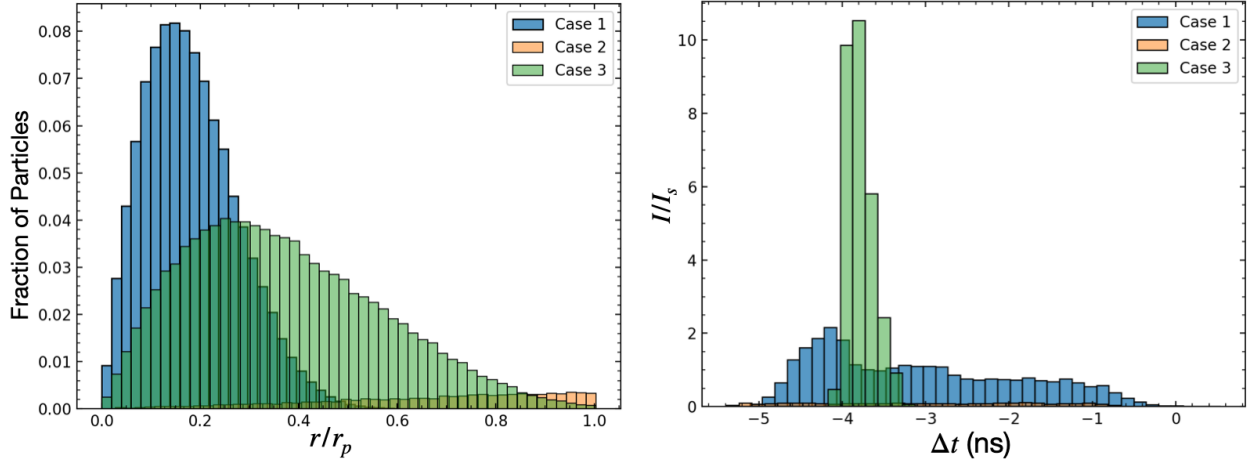


Figure 8.2: Fraction of particles at a radial distance normalized by the aperture  $r/r_p$ (left) and calculated current normalized by the source current  $I_s$  (right) for three cases using four gaps ( $N_g = 4$ ). The bin width in both plots are the same for the all three cases.

	Peak Current ( $\mu\text{A}$ )	Transmission
Case 1	2.1	1.0
Case 2	0.1	0.09
Case 3	10.4	0.99

Table 8.2: Values for the peak current and transmission for different design cases using a  $20^\circ$  pulsed beam and four acceleration gaps.

charge. Additionally, in Ch. 7 we showed that the gap provides strong transverse focusing for an accelerative field and thus will also give strong transverse defocusing when the field is negative. At  $\phi_s = -\pi/2$  particles that arrive in the gap just before will experience defocusing. Furthermore, the RF synchronism is designed for the when the design particle reaches the center of the gap and so particles lagging the design particle will also experience this defocusing. This was unaccounted for in Ch. 7 and to mitigate this effect, the RF synchronism should be moved away from  $-\pi/2$  the head of the beam does not get defocused. With this in mind, Case 3 was chosen with  $\phi_s = -\pi/6$  being sufficiently far so that the head of the beam is not defocused and there is still longitudinal bunching. In Case 3, the last gap was dedicated to maximally focusing longitudinally to produce a compressed beam just before the  $F_{\text{cup}}$ . The intuition in Case 3 worked out: although there is more radial spread

than Case 1, the beam is well-confined and the calculated current on target more than double Case 1. Additionally, there was only a 1% particle loss (see Table 8.2). Appendix A gives snapshots of the beam for the three cases (Case 1 in Fig. A.1, Case 2 in Fig. A.2, and Case 3 in Fig. A.3). Panel 2 and 3 in Fig. A.2 shows clearly the effects of the first gap set to maximally bunch. Initially, the particles are not loss but significant losses occur after gap 3. We can see from Fig. A.3 how reversing Case 2 and employing a weaker bunching design keeps the beam well confined until the final gap maximally bunches the beam (panel 5). However, due to the previous gap radially focusing the beam, the expansion is not deleterious to the beam over the drift distance to the Fcup and thus we find a good medium between radially and longitudinal focusing.

To observe effects of the matching section, and beam dynamics for many gaps, we repeated the analysis for  $N_g = 16$  (two additional gaps were added in one case which will be explained momentarily) with a four ESQ matching section after gap 6 using the voltage and design shown in Fig. 7.13. We compared three cases:

1. Case 1: No matching section or periodic doublet focusing and max acceleration  $\phi_s = {}^{18}0$
2. Case 2: Matching with subsequent periodic doublet focusing and initial tapering acceleration  $\phi_s = (0 \quad {}^3-\pi/6 \quad {}^{12}0)$
3. Case 3: Matching with subsequent periodic doublet focusing and max acceleration  $\phi_s = {}^{16}0$ .

Case 1 again represents a baseline for comparison. Two additional gaps were added to Case 1 to see if there was a persistent trend. Case 2 represents the point design from Ch. 7. Case 3 is a “see what happens” scenario by trying to marry the previous two cases. Recall that the system has low phase advance ( $\sigma \approx 13$  deg/per) and so one would expect Case 3 to yield

preferable results especially considering the radial confinement advantage discussed in the four gap analysis. Fig. 8.3 shows the radial distribution (left) and the normalized current

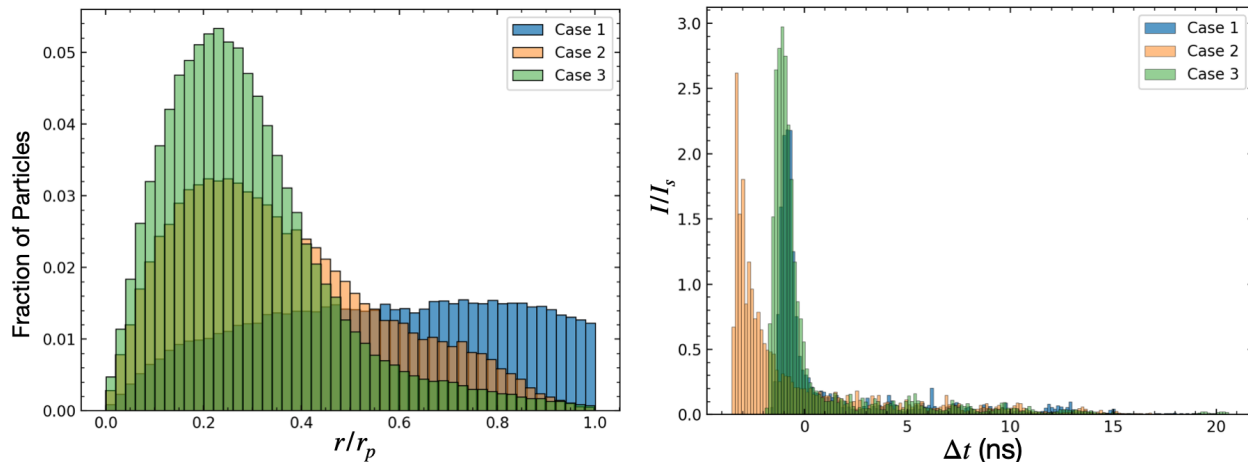


Figure 8.3: Fraction of particles at a radial distance normalized by the aperture  $r/r_p$ (left) and calculated current normalized by the source current  $I_s$  (right) for three cases using 16 gaps ( $N_g = 16$ ). The bin width in both plots are the same for the all three cases.

	Peak Current ( $\mu\text{A}$ )	Transmission
Case 1	2.1	0.62
Case 2	2.6	0.79
Case 3	2.9	0.94

Table 8.3: Values for the peak current and transmission for different design cases using a  $20^\circ$  pulsed beam and 16 acceleration gaps.

(right) for the three cases (Case 1 in blue, Case 2 in orange, and Case 3 in green). The radial distribution in Case 1 is drastically different but not surprising. Recall in Ch. 7 that, after some gaps, the beam begins to expand from overfocusing. This is precisely the long-term effects shown here as the beam has expanded and almost uniformly fills the aperture. Almost half the particles are lost after transport as seen in Table 8.3 and this trend in decreasing particles was persistent gap-to-gap after gap 8. It appeared that the particle losses were beginning to plateau at around gap 14. However, the additional two gaps showed this was not the case. Case 2 and Case 3 are clearly superior to Case 1, but choosing a winner between

the two is non-obvious. At face-value, Case 3 appears superior. This design has better radial confinement, a higher peak current, and higher transmission. Closer inspection of the current profile hints at a possible problem with Case 3. The peak, although highest, is also narrow compared to Case 2. Also, notice that the peak for Case 2 is well in the negative meaning most particles are arriving earlier. This hints at the possibility that a maximum bunching gap at the end of the lattice would significantly increase the peak current just as was shown in the 4-gap design. Since Case 3 has thin peak that is only slightly in the negative, it is suspected that a maximum bunching gap would not be as substantial and possibly even over-compress the beam leading to transverse particle losses. Also, the particle plots shown for Case 2 (Fig. A.5) and Case 3 (Fig. A.6) give further insight when viewing the final panels for each (panel 21). In Case 2, the beam is more longitudinally confined and the overall beam looks intact. Case 3 has a significant tail which will no doubt be absorbed in a couple gaps adding to particle loss. Also, the over beam looks to be breaking apart evident by the jagged outline. This suggests that over the next few gaps, the beam will begin to lose cohesion in Case 3. This is speculation still, and to test one would have to find the matching conditions for the next couple lattice periods as was done in Ch. 7. We leave this as an area for further investigation and highlight the intricacies and subtle observations needed in judging design choices.

The pulsed beam analysis has proved insightful for design choices to be made when continuously injecting the beam. When picking a design phase, bunching phases should be far enough so that the beam is not defocused. It will not be possible to prevent this over a continuous injection. Therefore we have a design tradeoff: we can prevent a larger section of the beam from being defocused by the RF gap by choosing synchronous phases close to 0 but in doing so receive smaller longitudinal focusing. This tradeoff has not been thoroughly

investigated and an optimal design is an open question. Furthermore, we should avoid strong bunching early in the lattice as this will tend to beam expansion and potential significant losses in particles. Using the final gap in the system to maximally compress the beam before the target worked wonderfully and suggests a strategy for design. With the pulsed injection analysis in mind, we can probe the case of continuous injection.

## 8.7 Continuous Beam Injection

When simulating a continuously injected (CW) beam, we injected for a whole RF period and then terminate injection. This differs from a CW beam in the lab where there will be particles ahead of the beam (past the head) and particles behind the beam (before the tail). Particles from these different injection cycles can influence beam dynamics and “jump” into the beam of interest. Therefore, there is the potential that the CW injection scheme of the simulation is missing important physics. To test this, we injected particles for a full period before and after the interest beam and compared final outputs. There was no significant deviations and we can assume that the physics is accurately captured in the simulation scheme. As in the previous section, we test three design cases using a 16 gap lattice:

1. Case 1: No matching section or periodic doublet focusing and max acceleration  $\phi_s = {}^{16}0$
2. Case 2: Matching with subsequent periodic doublet focusing and initial tapering acceleration  $\phi_s = (0 \quad {}^3_{-\pi/6} \quad {}^{12}0)$
3. Case 3: Matching with subsequent periodic doublet focusing and max acceleration  $\phi_s = {}^{16}0$ .

The radial distribution and calculated current at the Fcup are shown in Fig. 8.4 for each

case. Table. 8.4 shows lists the peak current and transmission values for each case. One again, the maximum acceleration case (Case 1) is inferior to Case 2 and Case 3—no radial confinement, lowest peak current, lowest transmission. And once again, the decision between Case 2 and Case 3 is non-obvious. Some differences should be noted compared to the pulse injection cases. One can argue that the radial distribution for Case 2 is the same, if not better, than Case 3. The tail ends are approximately the same and the beam is not so heavily concentrated as it is in Case 3. The wider peak in Case 3 could be advantageous in the case of space-charge where a large concentration of particles would eventually act as a strong transverse defocusing force. Similarly, the inset image on the current plot shows Case 2 and Case 3 alone and it is easier to see that Case 2 has a slimmer peak. Also, Case 2 has less particles in the tail ends ( $\Delta t > 0$ ) which suggests more of a bunched beam compared to Case 3. It is difficult to discern which design performs better between Case 2 and Case 3. It is clear, maybe unsurprisingly, that the matching section and periodic doublet focusing yields better performance.

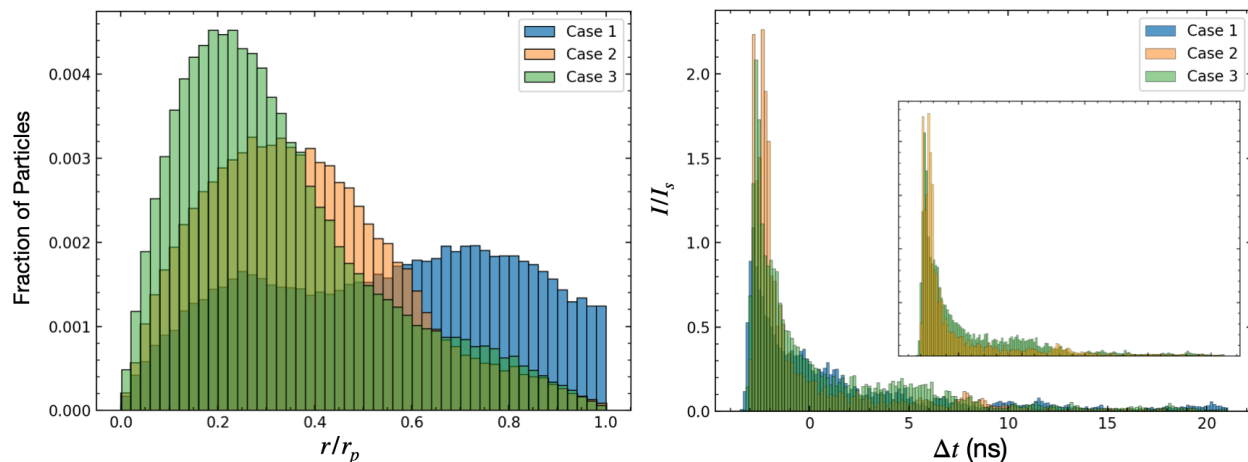


Figure 8.4: Fraction of particles at a radial distance normalized by the aperture  $r/r_p$ (left) and calculated current normalized by the source current  $I_s$  (right) for three cases using 16 gaps ( $N_g = 16$ ). The bin width in both plots are the same for the all three cases. The inset plot shows Case 2 and Case 3 together for better comparison of the current.



	Peak Current ( $\mu\text{A}$ )	Transmission
Case 1	1.1	0.073
Case 2	2.3	0.082
Case 3	2.1	0.096

Table 8.4: Values for the peak current and transmission for different design cases using a continuous beam and 16 acceleration gaps.

The transmission values may be slightly alarming and raises the question: what transmission should we expect? For a uniform CW injection we expect there to be particle losses. In Sec. 2, longitudinal dynamics were discussed and recall that the region of  $0 < \phi_s < -\pi/2$  provided particle capture and acceleration. Since this region is 25% the RF period, we can expect an upper bound for particle transmission of 0.25, i.e. we can at best capture 25% of a uniform CW beam <sup>1</sup>. To establish a lower bound, one could iteratively simulate the beam for different pulse lengths using the desired design lattice. Starting with a pulse length of  $L_b = \tau_{rf}$ , one could decrease the pulse length until a transmission of 1.0 is attained giving a pseudo RF bucket. The end pulse-length  $\ell_b$  will give the fraction of a uniform CW beam one could expect to fully capture and the expected transmission will be  $\ell_b/L_b$ . The transmission values in CW injections should thus fall between  $\ell/L_b$  and 0.25. In the pulsed injection study, a  $20^\circ$  beam was used (5% of uniform CW beam) and the transmission was 0.96 (see Table 8.3). This comes close to the lower bound and assuming such, we can expect that in a uniform CW injection we should be able to attain a 0.05 transmission and design for higher values. Indeed all three cases yield a high transmission with Case 1 being close to the lower bound and Case 3 almost doubling it. These higher transmission results suggest that even when operating at max acceleration (recall this means theoretically no particle capture) we are still capturing some particles in this pseudo RF bucket due to the dynamics of a CW

---

<sup>1</sup>This upper bound assumes no particle transfer from areas outside the operating phase used. If the gap accelerations were perturbative then this would be a reasonable lower bound.

injection.

It is interesting to observe where in the lattice such significant particle losses occur. In all three cases, 50% of the beam is lost after the second gap. Before the matching section (or after gap 6 in these cases), 80-90% of the beam is lost. Recall that Warp does not have the capabilities to simulate dielectric effects in a 3D simulation. The substantial loss of particles is no doubt inducing a charge on the surrounding materials as they are lost to the transverse walls. The effects of this could be significant and warrant further investigation. Additionally, this suggests that the accelerator would benefit from shielding being placed early in the lattice for protection.

## 8.8 Summary and Concluding Remarks

This chapter discussed the developed Warp simulation to self-consistently simulate the ion beam array. Using what was learned from the reduced models we were able to quickly find designs that were vastly improved compared to our base case of a lattice designed for maximum acceleration. Using a combination of particle plots and data analysis, we observed better beam profiles, increased transmission, and higher calculated currents delivered to target. Our current optimal design consisted of a 16 gap lattice. The first six gaps were accelerator gap only and the optimal design phase is debatable—yet, the simplest choice is to keep all six gaps operating at  $\phi_s = 60^\circ$ . A four ESQ matching section with voltages and spacing shown in Fig. 7.13 was used to match the beam to the rest of lattice which consisted of acceleration gaps and periodic doublet focusing. The doublet voltages given in Table. 7.4 were used and the remaining gaps were spaced for maximum acceleration ( $\phi_s = 100^\circ$ ). We also found that 80-90% of the beam was lost after the first six gaps and that roughly 50% of

the beam is lost after the second gap suggesting that the system would benefit from shielding at these locations to prevent damage from lost beam.

This is only a starting point however. We did not find a global optimum and it is hopefully apparent that such a task would warrant a herculean investigation that is outside the scope of the current project. The techniques discussed and simulation itself form the fundamental pieces into taking such an endeavor and is a logical next step. There are areas of improvement that further investigations would benefit from. Simulation run time can be significantly reduced by using gridded electric fields. To do this, one would have to extract the 3D fields for a fixed mesh size and then restructure the current code to use the gridded fields. Using gridded fields means Warp will no longer need to recalculate the fields each time step. It is important to realize that these fields need to be correct for the given system and thus any changes in conductor spacing where there is some field overlap, like the ESQ and neighboring wafer plate in the periodic doublet focusing, would need to be correctly loaded on to the grid. Once this is completed the simulation should be “tuned” to experiment using some controlled experiments to test validity and establish confidence that the simulation is describing the experimental system to desired accuracy. From this point, there are many directions that can be taken for additional work. Perhaps, the simulation can be used to create synthetic experimental data which can then be fed into a neural network.

## Part III

# Machine Learning and Optimization of an Ion Induction Linac

## Chapter 9. Introduction and System Overview

Accelerators are highly complex machines rich with physics and also rich in the production of physical data sets. These devices are ripe for machine learning (ML) applications such as fault detection, anomaly detection, modeling and optimization, and more [38]. The remaining chapters are dedicated to a short-lived ML project applied to an existing induction accelerator called the Neutral Drift Compression Experiment II (NDCX-II). We assume general knowledge of ML and neural network (NN) architectures since there is a plethora of educational resources. Of the many resources, [39] give a practical overview directed at physicists and has accompanying Jupyter notebooks for practice. There is also the free textbook [40] that gives a more rigorous approach to the mathematics behind machine learning. There are also relatively cheap courses, of which, Andrew Ng's [41] and Jose Portilla's [42] are highly recommended. For MSU students, Dr. Wade Fisher's course on Data Analysis is a must take. Lastly, the two textbooks [43, 44] are excellent sources for practical knowledge on machine learning techniques and both have accompanying Jupyter notebooks and downloadable Python scripts to practice with. A final textbook [45] is a balanced approach to machine learning providing practical and theoretical knowledge.

Commissioned in 2012, NDCX-II was designed to probe materials in the warm dense matter regime by delivering an energetic and compressed beam to a target. Fig. 9.1 shows a CAD drawing labeling certain components (A) and an actual picture of the device housed at LBNL (B). The accelerator measures approximately ten meters long and consists of 27 cells that perform acceleration, compression, and radial focusing. Fig. 9.2 shows the nominal beamline configuration. From the source, an ion beam of  $\approx 600$  ns in pulse length with kinetic

energy of  $\approx 100$  keV is extracted. Each of the 27 cells are equipped with a solenoid to provide radial focusing. Active induction cells, labeled A in the figure, provide a carefully tailored compression and acceleration voltage waveforms to the beam. Two active induction cells have a flat-top waveform followed by the final three induction cells that apply a ramp waveform. The waveforms are shown in Fig. 9.3 for the 12 active induction cells—note the last five active induction cells are also called blumleins. Labeled I in Fig. 9.2 are the inactive induction cells that provide radial focusing with no acceleration or compression. After traversing the cells the beam enters a neutral drift plasma for final longitudinal compression. The result is an ion beam that is both compressed to less than 1 ns and accelerated to greater than 1 GeV with a spot size of  $\approx 1$  mm.

Our intent was to use NDCX-II to explore ML applications. Initially, we sought to both test an auto tuning algorithm (see Ch. 10) while also sampling the large parameter space and collecting data to be used for ML applications such as surrogate modeling and optimization. We were operating on a shoestring budget that was enough to cover limited operation of NDCX-II to begin outlining a proposal and garner more funding. However, we were not awarded more funding in the funding cycle and some major components of the machine begin to fail in 2023. In 2023, NDCX-II was officially pronounced dead. Thus, this project was limited to the collected data. We use the remaining chapters to both document the work that was done and to show how an ML project in accelerator physics might look.

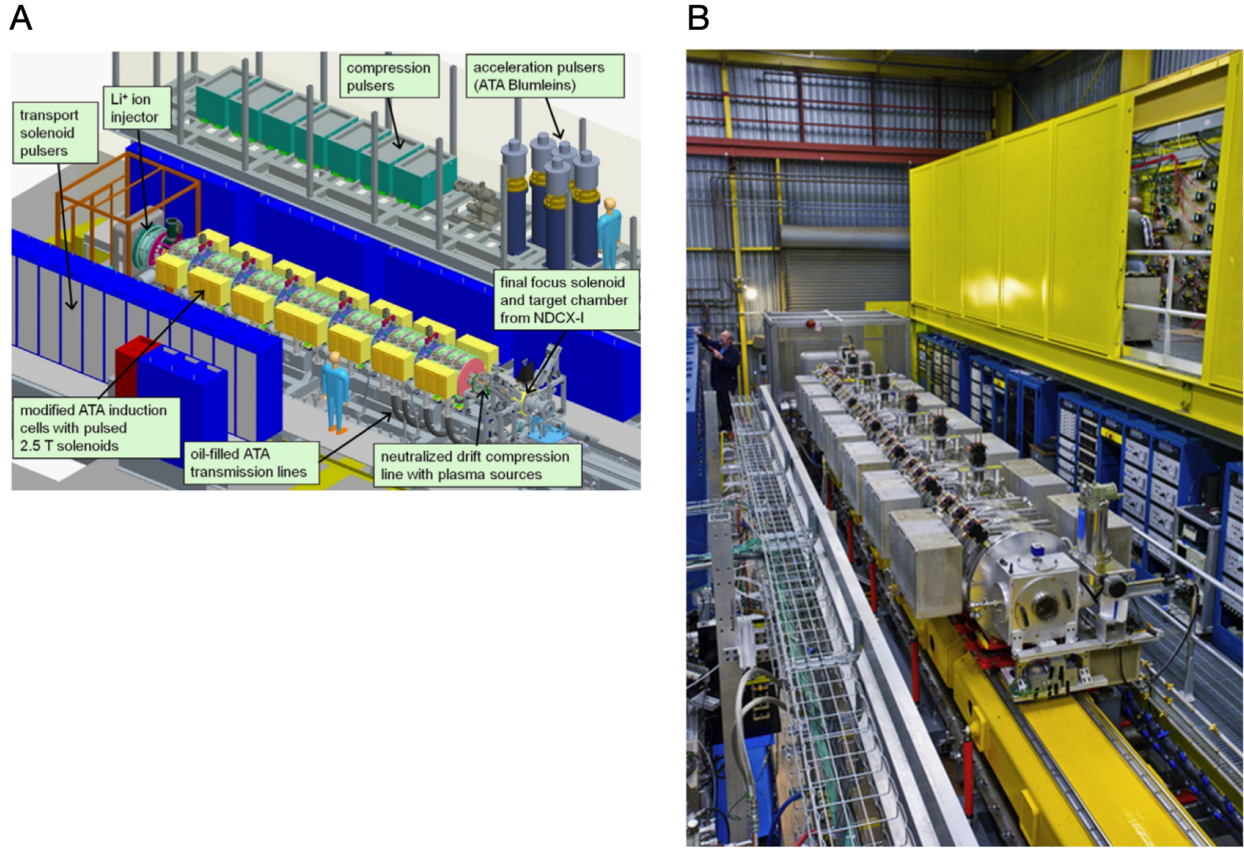


Figure 9.1: (A) CAD drawing of NDCX-II with various components labeled. (B) An actual photo of the accelerator housed at LBNL. Image borrowed from [46].

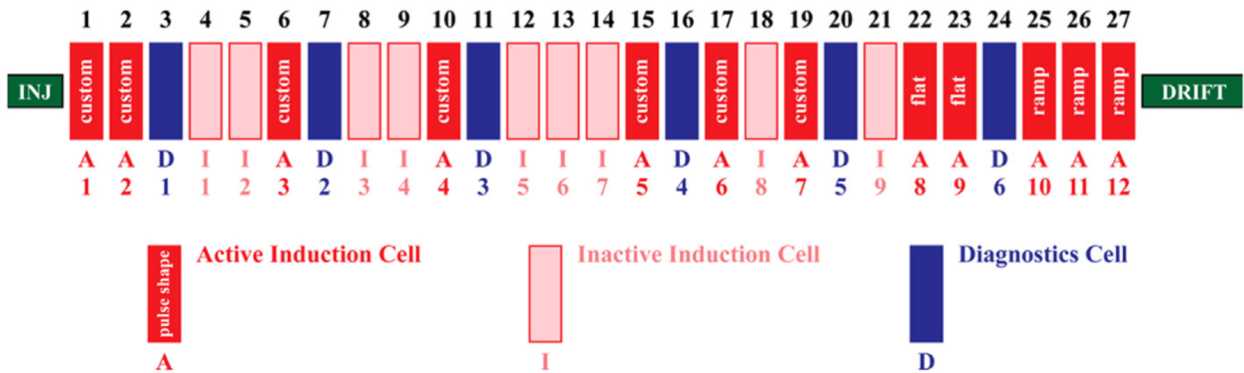


Figure 9.2: The nominal beamline configuration for NDCX-II [46]. Each of the 27 cells have solenoidal focusing. The dark red cells are active induction cells which accelerate and compress the beam, for which there are 12. The first seven have custom waveforms designed for optimal performance. Following are two flat waveform cells and then the last through apply a ramp waveform to longitudinal compress the beam before entering a neutral drift plasma region. In light red are inactive cells that only contribute solenoidal focusing. In dark blue are diagnostic cells for beam measurements.

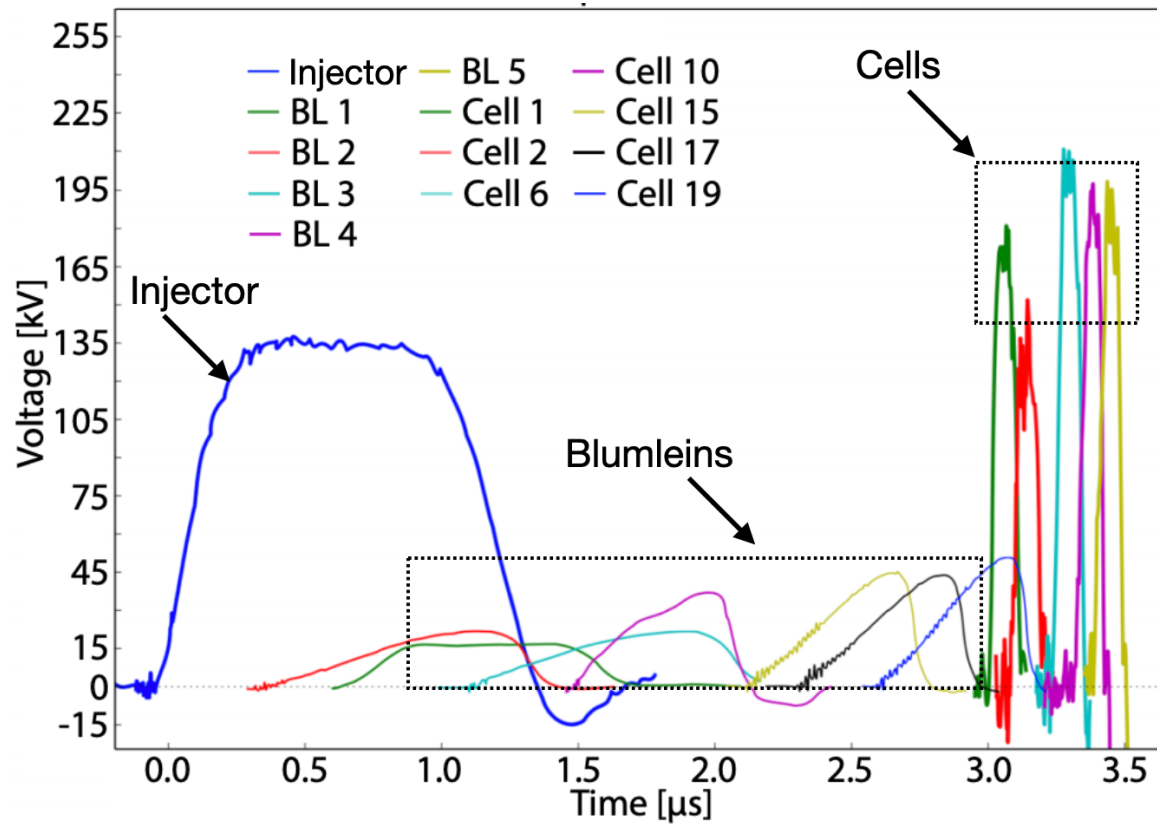


Figure 9.3: Voltage waveforms for the 12 active induction cells and the injector [47]. The first seven cells are custom waveforms labeled Cell # on the plot. The remaining five active induction cells are also called Blumleins and are labeled with BL on the plot.



# Chapter 10. An Auto Tuning Algorithm for Safe Extremum Seeking in NDCX-II

Once a device has been optimized for performance there is no guarantee that the optimum settings are constant in time—i.e. there is no certainty the solenoidal settings for the 27 cells in Fig. 9.2 will be the same. This “drift” is due to changes in the equipment over time. However, unless significant damage has occurred the drift should not be substantial and, in principle, can be found again by sampling the parameter space near the previous optimum. When sampling the parameter space one must be careful since some settings can be deleterious to the machine. To autotune NDCX-II, we used a local, model-independent extremum seeking (ES) algorithm [48, 49, 50]. The algorithm’s simple implementation and quick convergence made it highly desirable since we could monitor the tuning in real-time. Consider a time-varying and analytically unknown cost function that we seek to maximize  $C(\vec{p}, t)$  along with the noise-corrupted measurement  $\hat{C}(\vec{p}, t) = C(\vec{p}, t) + \eta(t)$ . The varied parameters are given by the vector  $\vec{p} = (p_1, p_2, \dots, p_n)$  which, in the case of NDCX-II, will be the application time and the applied bias for each of the 27 cells. The parameters are adjusted by the dynamical equation:

$$\frac{dp_j}{dt} = \sqrt{\alpha\omega_j} \cos[\omega_j t - k\hat{C}(\vec{p}, t)] \quad (10.1)$$

where  $\omega_j$  is the dither,  $\alpha$  is the dithering amplitude, and  $k$  is the feedback gain. Local minima can be escaped by increasing  $\alpha$ , and  $k$  is analogous to the learning rate as seen in gradient based methods. When the dynamics have settled the algorithm will oscillate around

$p_j$  with frequency  $\sqrt{\alpha/\omega_j}$ . Intuitively, the algorithm will converge because by applying a different dither to each parameter, the parameters are orthogonal in Hilbert space since

$$\lim_{\omega_1, \omega_2 \rightarrow \infty} \int_0^t \cos(\omega_1 \tau) \cos(\omega_2 \tau) d\tau = 0, \quad (10.2)$$

# Chapter 11. Data Collection, Retrieval, and Preprocessing

Using Eq. 10.1 the parameters can be varied to find optimum settings while simultaneously sampling the parameter space within a bounded region. The parameters we varied were the solenoid timing, i.e. when the solenoid voltage was applied to give radial focusing, and the applied voltage on the blumleins. In reality, the applied voltage and time of application can be varied for both the solenoids and active induction cells. However, we decided to keep the custom waveforms and only vary the voltage on the blumleins (final five active induction cells). Control of NDCX-II is wrapped in a Python interface where the tuning algorithm can be easily implemented to vary the conductor settings. For each injection or “shot”, the conductor settings as well as voltage readings from the arcsource and faraday cup (among other devices) are stored in a Cassandra database which is a noSQL database. Functionality to query the database has been written in Python making data extraction relatively easy. Individual shot data is collected in a shotnumber data object that contains individual shot classes holding data attributes. After querying the database, a shotnumber or list of shotnumbers can be collected, the shots iterated, and data for individual conductor objects extracted. The process is illustrated graphically in Fig. 11.1.

Once extracted, the data is filtered and stored in a comma separated value (csv) file. This format is easily manipulated by a the Pandas library and a file can be loaded into a dataframe object which is useful for data analysis and preprocessing for machine learning algorithms. Data was mainly filtered by looking at the total charge delivered on target. In our experiments, typical values ranged between 0.1 nc and 10 nc. Thus, values that were

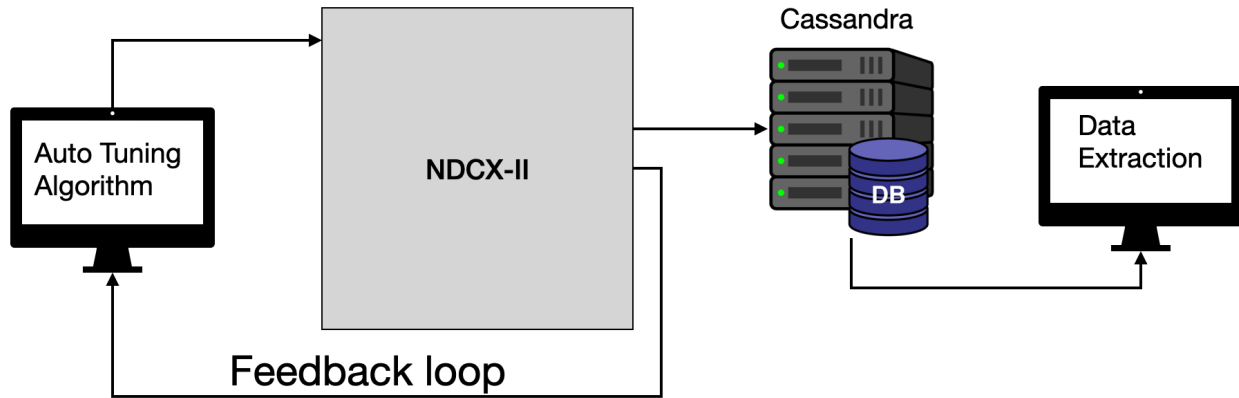


Figure 11.1: Graphical representation of the data collection and retrieval process for NDCX-II. Conductor settings are initialized and the accelerator is commanded through a Python interface at the computer. After each shot, the experimental data is stored into a shotnumber object in the Cassandra database. The database can be queried through Python interface and the data collected through the data classes shotnumber and shot.

significantly different were ignored. Bad shots typically resulted from a misfiring in NDCX-II and resulting in no beam. This was verified by looking at the voltage waveform measured by the arcsource which indicates a successfully extracted beam from the source. Once this initial screening for bad shots was done, a scatter plot was made for each column to visualize the outputs. This allowed us to quickly spot any outliers that we may have missed, track down the shot data, and determine if we should remove or not. After filtering the data we are left with 382 data points—although incredibly low, this amount of data will serve the purposes outlined in the introduction.

# Chapter 12. Training a Neural Network to Predict Peak Current and Time of Peak

We used a fully connected neural network architecture as depicted in Fig. 12.1. The first

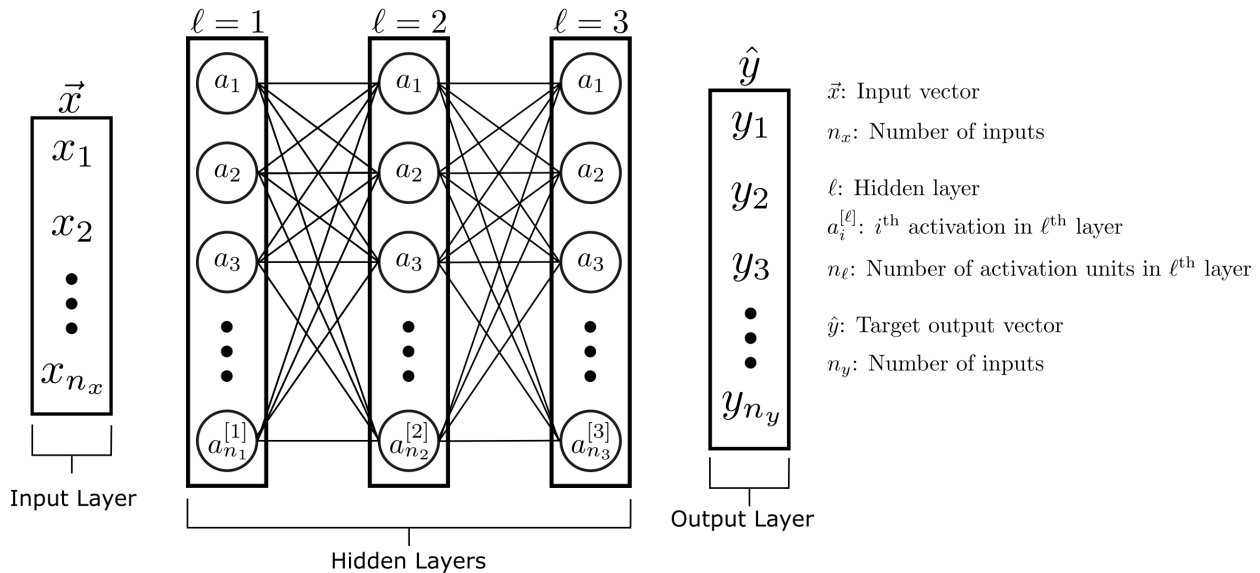


Figure 12.1: Schematic of a dense fully connected neural network architecture. We adopt the notation given by Andrew Ng’s Coursera course [41]. The inputs are given as a single column vector  $\vec{x}$  of which there are  $n_x$  entries or data samples. A single output layer is given by the vector  $\hat{y}$  containing  $n_y$  targets. The hidden layers are given by  $\ell$  and each layer is comprised of  $n_\ell$  neurons. Additionally, each layer has an activation function such as rectified linear unit (RELU).

layer is the input layer which will have all the data we wish to train on. In our case we used the peak voltage and time of peak voltage for the voltage profiles of all 28 solenoids, 7 compression cells, and 5 blumleins. This gives  $2 \times (28 + 7 + 5) = 80$  parameters to train in order to predict peak current and the time of the peak. We can calculate the input and

output shapes ( $n_x$  and  $n_y$ ) given the amount of data we partition for training  $N_{\text{data}}$ :

$$n_x = 80 \times N_{\text{train}},$$

$$n_y = 2 \times N_{\text{train}}.$$

Sandwiched between the input and output layers are the hidden layers denoted  $\ell$ . Each hidden layer is comprised of activations  $a_i$  and has an associated activation function per layer. In a vanilla ML project, the data is from a single source and this makes the NN model construction rather easy. The data is shaped into the input tensor, the layers created, and so on. As an example, consider the input data coming from the seven compression cells alone. Lst. 12.1 is pseudocode to help illustrate the process. After collecting the data from a csv file, we create two dense hidden layers that both use the scaled exponential linear unit (SELU) activation function with a L2 regularizer. Hidden layer 1 has 100 activations while the second has 80. The last layer predicts two outputs using a linear activation function. After creating the model, additional hyper parameters such as the batch size, number of epochs to train, the optimization function to be used, etc. are created and then the `cc_data` is used to fit the model. Now consider training on both the compression cells and the blumleins. When specifying the inputs, we would have to calculate the shapes and then construct the layers accordingly. This does not seem like an issue at first sight. However, what if we wanted to specify a different activation function for the blumleins? We could not do this by combining all the data into single input at the start. To circumvent this, we instead create the layers independently and then use Tensorflow to concatenate the layers into one at the end. Lst. 12.2 abbreviates how this might be done by adding the blumlein data into the procedure from Lst. 12.1. The process is the same as the simple network, although now it

is repeated for both conductors. We therefore, create two NNs for both conductors and the final output network is then fed into a NN that trains on all the conductor inputs. Thus, we go from individual training to coupled training which allows more flexibility in architecture. This is helpful for projects with differing source distributions but that are being combined to predict an output.

```
1  import tensorflow as tf
2  import pandas
3  import numpy as np
4
5  data = pd.readcsv('path_to_data')
6  cc_data = np.zeros(N_data, 7) # Data matrix for 7 cells
7  # Loop through cc_data and assign data to column
8  for i in range(7):
9      # Assume cell data is under column cc_#
10     cc_data[:,i] = data[f"cc_{i+1}"]
11
12 # Specify input shape for model. Note cc_data.shape = (num samples, 7)
13 input_cc = tf.keras.Input(shape=(7,))
14
15 # Create layers
16 dense_l1 = tf.keras.layers.Dense(
17     100,
18     activation = "selu",
19     kernel_regularizer = tf.keras.regularizers.L2(l2=l2_weight),
20     )(input_cc)
21 dense_l2 = tf.keras.layers.Dense(
22     80,
23     activation = "selu",
```

```

24 kernel_regularizer = tf.keras.regularizers.L2(l2=l2_weight),
25         )(dense_l1)
26 dense_l3 = tf.keras.layers.Dense(2, activation="linear")(dense_l2)
27
28 # Create model
29 model = tf.keras.Model(
30     inputs = [input_cc],
31     outputs = [dense_l3],
32     name = "Simple Model"
33 )
34
35 # With created model we can then fit to the reshaped cc_data.

```

Listing 12.1: Pseudocode for constructing a simple NN from a single data source which is taken to be seven compression cells.

```

1  import ...
2  data = ...
3
4  # Specify input shapes for both now
5  input_cc = tf.keras.Input(shape=(7,))
6  input_bl = tf.keras.Input(shape=(5,))
7
8  # Create layers for cells same as before and add cc_ prefix to
variable names
9  ...
10 cc_dense_l3 = tf.keras.layers.Dense(2, activation="linear")(dense_l2)
11
12 # Now create layers for blumleins

```



```

13 bl_dense_l1 = tf.keras.layers.Dense(
14     200,
15     activation = "relu",
16     kernel_regularizer = tf.keras.regularizers.L2(l2=l2_weight),
17     )(input_bl)
18 bl_dense_l2 = tf.keras.layers.Dense(
19     150,
20     activation = "relu",
21     kernel_regularizer = tf.keras.regularizers.L2(l2=l2_weight),
22     )(bl_dense_l1)
23 bl_dense_l3 = tf.keras.layers.Dense(2, activation="linear")(dense_l2)
24
25 # Concatenate layers and then build complete model
26 dense_input = tf.concat([cc_dense_l3, bl_dense_l3], 1)
27
28 # Create layers to train on outputs from both conductor networks
29 dense_l1 = tf.keras.layers.Dense(...)(dense_input)
30 dense_l2 = ...(dense_l1)
31 dense_l3 = tf.keras.layers.Dense(2, activation='linear')(dense_l2)
32
33
34 # Create model
35 model = tf.keras.Model(
36     inputs = [input_cc, input_bl],
37     outputs = [dense_l3],
38     name = "Complex Model"
39 )
40

```

41

```
# With created model we can then fit to the reshaped cc_data.
```

Listing 12.2: Pseudocode for constructing a simple NN from a single data source which is taken to be seven compression cells.



## BIBLIOGRAPHY

- [1] A. W. Chao, K. H. Mess, M. Tigner, and F. Zimmermann, *Handbook of Accelerator Physics and Engineering*. WORLD SCIENTIFIC, 2nd ed., 2013.
- [2] T. P. Wangler, *RF Linear Accelerators Second, Completely Revised and Enlarged Edition*. Wiley-VCH, 2008.
- [3] A. Sessler and E. Wilson, *Engines of Discovery*. WORLD SCIENTIFIC, 2014.
- [4] S. Lund, “Phy 905 lectures 09 longitudinal physics: Beam acceleration.”
- [5] S. Lund, “Longitudinal physics: Beam accelerator.” [https://people.nslc.msu.edu/~lund/msu/phy905\\_2020/lec\\_lund/](https://people.nslc.msu.edu/~lund/msu/phy905_2020/lec_lund/), 2020.
- [6] M. Reiser, *Theory and Design of Charged Particle Beams Second, Updated and Expanded Edition*. Wiley-VCH, 2008.
- [7] “United states particle accelerator school.” <https://uspas.fnal.gov/>.
- [8] “Cern document server.” <https://cds.cern.ch/?ln=en>.
- [9] S. Lund and J. Barnard, “Transverse particle dynamics.” [https://people.nslc.msu.edu/~lund/uspas/bpisc\\_2020](https://people.nslc.msu.edu/~lund/uspas/bpisc_2020), 2020.
- [10] J. Barnard and S. Lund, “Introduction.” [https://people.nslc.msu.edu/~lund/uspas/bpisc\\_2020/lec\\_set\\_01/01.introduction.pdf](https://people.nslc.msu.edu/~lund/uspas/bpisc_2020/lec_set_01/01.introduction.pdf), 2020.
- [11] H. Wiedemann, *Particle Accelerator Physics*. Springer International Publishing, 2015.
- [12] S. M. Lund and B. Bukh, “Stability properties of the transverse envelope equations describing intense ion beam transport,” *Phys. Rev. ST Accel. Beams*, vol. 7, p. 024801, Feb. 2004.
- [13] J. Barnard and S. Lund, “Introduction.” [https://people.nslc.msu.edu/~lund/uspas/bpisc\\_2020/lec\\_set\\_01/01.introduction.pdf](https://people.nslc.msu.edu/~lund/uspas/bpisc_2020/lec_set_01/01.introduction.pdf), 2020.
- [14] W. Magnus, N. Y. U. I. of Mathematical Sciences, and U. S. A. F. O. of Scientific Research, *Hill’s Equation: General theory*. No. pt. 1 in United States. Air Force Office of Scientific Research Technical Note No, New York University, Institute of Mathematical Sciences, Division of Electromagnetic Research, 1957.
- [15] E. Courant and H. Snyder, “Theory of the alternating-gradient synchrotron,” *Annals of Physics*, vol. 3, no. 1, pp. 1–48, 1958.

- [16] I. M. Kapchinskij and V. V. Vladimirskij, “Limitations Of Proton Beam Current In A Strong Focusing Linear Accelerator Associated With The Beam Space Charge,” in *2nd International Conference on High-Energy Accelerators*, pp. 274–287, 1959.
- [17] S. M. Lund, T. Kikuchi, and R. C. Davidson, “Generation of initial kinetic distributions for simulation of long-pulse charged particle beams with high space-charge intensity,” *Phys. Rev. ST Accel. Beams*, vol. 12, p. 114801, Nov. 2009.
- [18] P. M. Lapostolle, “Possible emittance increase through filamentation due to space charge in continuous beams.,” *IEEE Trans. Nucl. Sci.*, vol. 18, pp. 1101–1104, 1971.
- [19] F. J. Sacherer, “Rms envelope equations with space charge.,” *IEEE Trans. Nucl. Sci.*, vol. 18, pp. 1105–1107, 1971.
- [20] D. P. Grote, A. Friedman, J. Vay, and I. Haber, “The WARP Code: Modeling High Intensity Ion Beams,” *AIP Conference Proceedings*, vol. 749, pp. 55–58, 03 2005.
- [21] “Warp wiki home.” <https://warp.lbl.gov/home>.
- [22] S. Hanna, *RF Linear Accelerators for Medical and Industrial Applications*. Artech, 2012.
- [23] R. W. Hamm and M. E. Hamm, *Industrial Accelerators and Their Applications*. WORLD SCIENTIFIC.
- [24] M. Network\*, S. Sinha, Y. Hou, D. Ni, Q. Ji, A. Persaud, P. Seidl, T. Schenkel, A. Lal, and K. K. Afridi, “A 27.12-mhz 10-kv power amplifier for compact particle accelerators utilizing an optimized,” in *2020 IEEE Energy Conversion Congress and Exposition (ECCE)*, pp. 5452–5457, 2020.
- [25] K. B. Vinayakumar, S. Ardanuc, Q. Ji, A. Persaud, P. Seidl, T. Schenkel, and A. Lal, “Demonstration of waferscale voltage amplifier and electrostatic quadrupole focusing array for compact linear accelerators,” *Journal of Applied*, vol. 125, p. 194901, May 2019.
- [26] A. Persaud, Q. Ji, E. Feinberg, P. A. Seidl, W. L. Waldron, T. Schenkel, A. Lal, K. B. Vinayakumar, S. Ardanuc, and D. A. Hammer, “A compact linear accelerator based on a scalable microelectromechanical-system RF-structure,” *Rev. Sci. Instrum.*, vol. 88, p. 063304, June 2017.
- [27] “The future of compact high voltage power is here.” Retrieved May 1, 2021, from <https://www.airitytech.com/>.
- [28] Q. Ji, K. K. Afridi, T. Bauer, G. Giesbrecht, Y. Hou, A. Lal, D. Ni, A. Persaud, Z. Qin, P. Seidl, S. Sinha, and T. Schenkel, “Beam power scale-up in micro-electromechanical

- systems based multi-beam ion accelerators,” *Rev. Sci. Instrum.*, vol. 92, p. 103301, Oct. 2021.
- [29] S. Lund, “Phy 905 lectures 04.sup equations of motion and applied fields,” September 2018.
- [30] A. Wolski, “Maxwell’s equations for magnets,” 2011.
- [31] D. Neuffer, “Longitudinal motion in high current ion beams - a Self-Consistent phase space distribution with an envelope equation,” *IEEE Trans. Nucl. Sci.*, vol. 26, pp. 3031–3033, June 1979.
- [32] S. Linge and H. P. Langtangen, *Programming for Computations-Python: A Gentle Introduction to Numerical Simulations with Python 3.6*. Springer Nature, 2020.
- [33] S. M. Lund, S. H. Chilton, and E. P. Lee, “Efficient computation of matched solutions of the Kapchinskij-Vladimirskij envelope equations for periodic focusing lattices,” *Phys. Rev. ST Accel. Beams*, vol. 9, p. 064201, June 2006.
- [34] F. Gao and L. Han, “Implementing the Nelder-Mead simplex algorithm with adaptive parameters,” vol. 51, no. 1, pp. 259–277.
- [35] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python,” *Nature Methods*, vol. 17, pp. 261–272, 2020.
- [36] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. USA: Cambridge University Press, 3 ed., 2007.
- [37] R. Courant, K. Friedrichs, and H. Lewy, “On the partial difference equations of mathematical physics,” *IBM J. Res. Dev.*, vol. 11, pp. 215–234, Mar. 1967.
- [38] A. Edelen, C. Mayes, D. Bowring, D. Ratner, A. Adelman, R. Ischebeck, J. Snuverink, I. Agapov, R. Kammering, J. Edelen, I. Bazarov, G. Valentino, and J. Wenninger, “Opportunities in machine learning for particle accelerators,” *TBD*.
- [39] P. Mehta, M. Bukov, C.-H. Wang, A. G. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, “A high-bias, low-variance introduction to machine learning for physicists,” *Physics Reports*, vol. 810, 2019.

- [40] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [41] A. Ng, Y. Mourri, and K. Katanforoosh, “Neural networks and deep learning,” in *Deep Learning Specialization*, Coursera, 2023. DeepLearning.AI.
- [42] J. Portilla, “Python for machine learning and data science masterclass,” Udemy, 2023. Pierian Training.
- [43] A. Geron, *Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems*. Sebastopol, CA: O’Reilly Media, 2017.
- [44] F. Chollet, *Deep Learning with Python*. Manning, 2017.
- [45] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006.
- [46] W. L. Waldron, W. J. Abraham, D. Arbelaez, A. Friedman, J. E. Galvin, E. P. Gilson, W. G. Greenway, D. P. Grote, J.-Y. Jung, J. W. Kwan, M. Leitner, S. M. Lidia, T. M. Lipton, L. L. Reginato, M. J. Regis, P. K. Roy, W. M. Sharp, M. W. Stettler, J. H. Takakuwa, J. Volmering, and V. K. Vytla, “The NDCX-II engineering design,” *Nucl. Instrum. Methods Phys. Res. A*, vol. 733, pp. 226–232, Jan. 2014.
- [47] P. A. Seidl, J. J. Barnard, E. Feinberg, A. Friedman, E. P. Gilson, D. P. Grote, Q. Ji, I. D. Kaganovich, B. Ludewigt, A. Persaud, C. Sierra, M. Silverman, A. D. Stepanov, A. Sulyman, F. Treffert, W. L. Waldron, M. Zimmer, and T. Schenkel, “Irradiation of materials with short, intense ion pulses at NDCX-II,” *Laser Part. Beams*, vol. 35, pp. 373–378, June 2017.
- [48] A. Scheinker and M. Krstic, “Extremum seeking with bounded update rates,” *Syst. Control Lett.*, 2014.
- [49] A. Scheinker, A. Edelen, D. Bohler, C. Emma, and A. Lutman, “Demonstration of Model-Independent control of the longitudinal phase space of electron beams in the Linac-Coherent light source with femtosecond resolution,” *Phys. Rev. Lett.*, vol. 121, p. 044801, July 2018.
- [50] A. Scheinker, D. Bohler, S. Tomin, R. Kammering, I. Zagorodnov, H. Schlarb, M. Scholz, B. Beutner, and W. Decking, “Model-independent tuning for maximizing free electron laser pulse energy,” *Phys. Rev. Accel. Beams*, vol. 22, p. 082802, Aug. 2019.

## Appendix A. Beam-Movie Image Panels

### A.1 Pulsed Injection

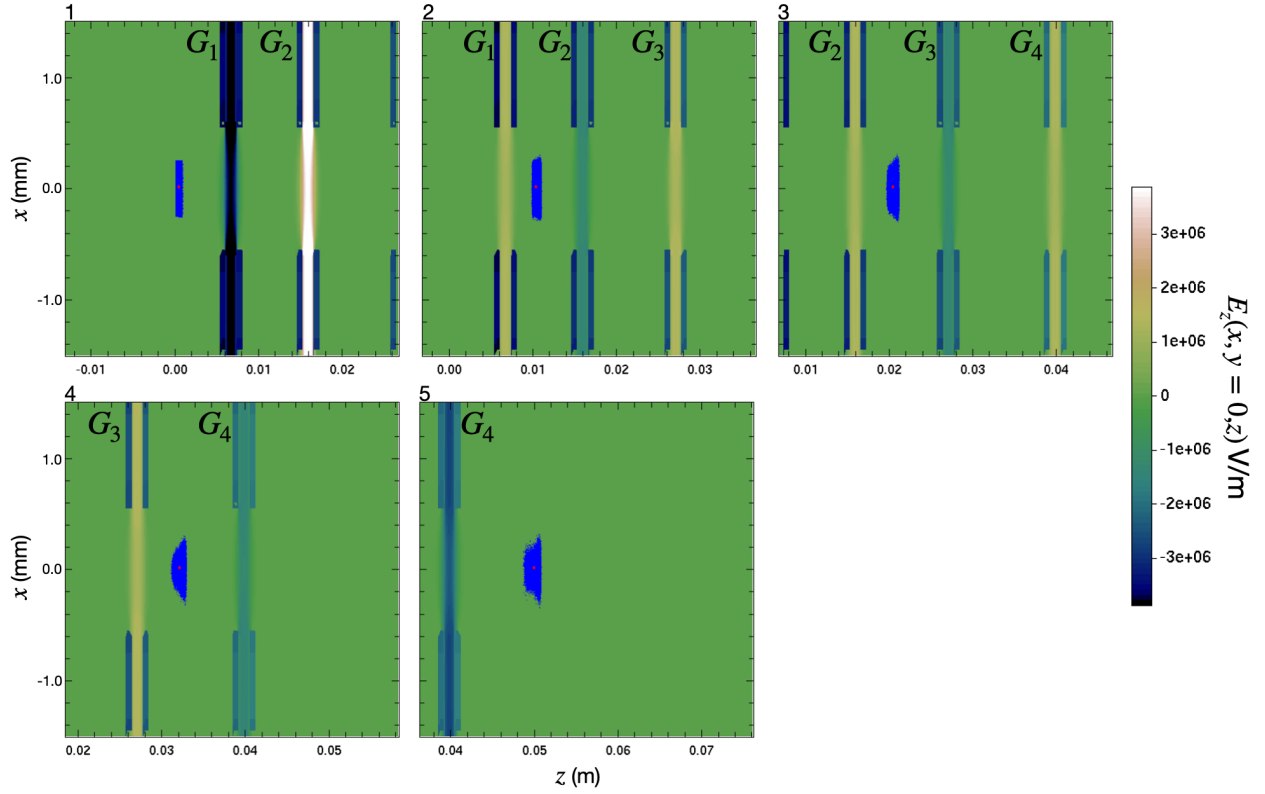


Figure A.1: A pulsed injection for a  $20^\circ$  beam and a 4-gap lattice operating at maximum acceleration  $\phi_s = 40$ .



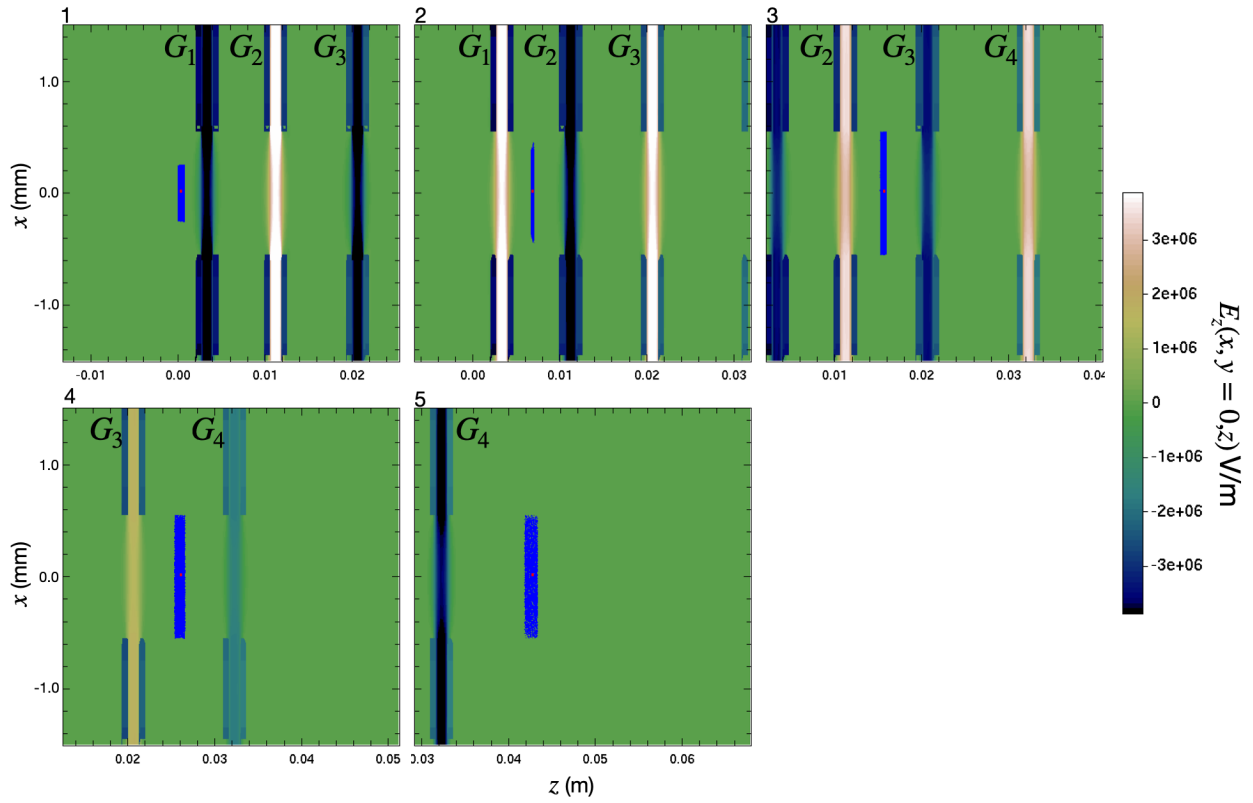


Figure A.2: A pulsed injection for a  $20^\circ$  beam and a 4-gap lattice using a tapering design for acceleration and longitudinal bunching  $\phi_s = (-\pi/2 \ -\pi/3 \ -\pi/6 \ 0)$ .

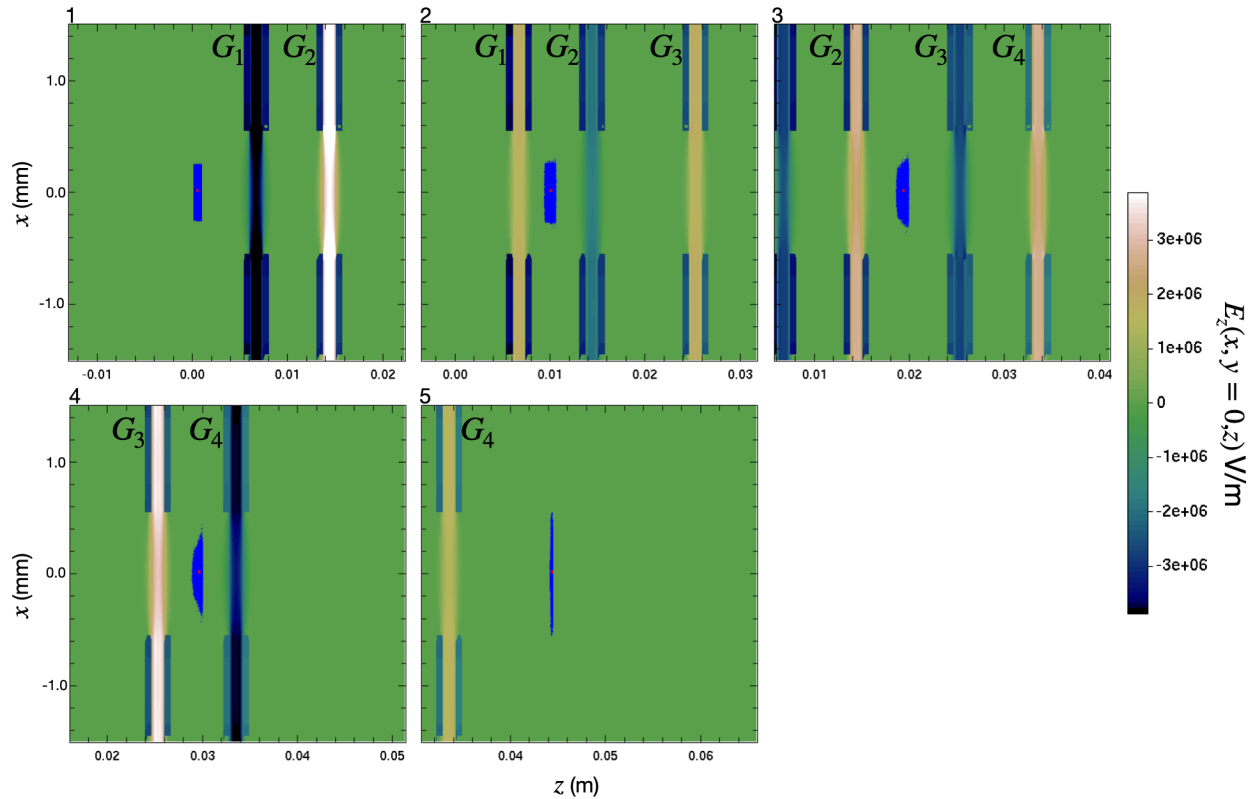


Figure A.3: A pulsed injection for a  $20^\circ$  beam and a 4-gap lattice using a tapering design for acceleration and longitudinal bunching  $\phi_s = (0 \quad -\pi/6 \quad -\pi/6 \quad -\pi/2)$ . The tapering is reversed to avoid defocusing when maximally bunching. Instead, the maximum bunching is done at the last gap.

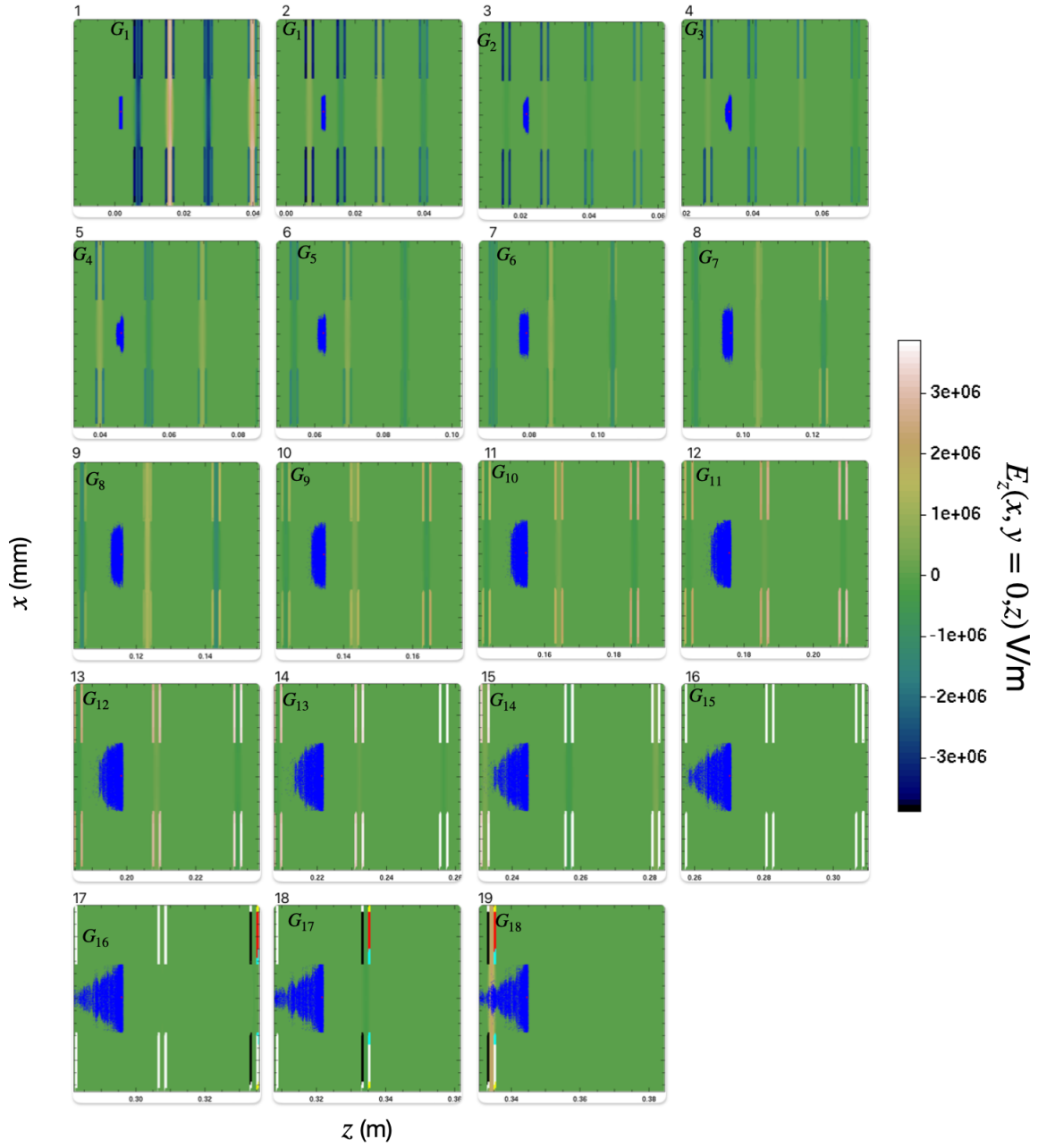


Figure A.4: A pulsed injection for a  $20^\circ$  beam and a 18-gap lattice designed for max acceleration  $\phi_s = 180$ . To minimize clutter, only the first gap to appear in the frame is labeled and the vertical axis of each plot is the same as those in plots above with maximum extents  $\pm 1.5$  mm.

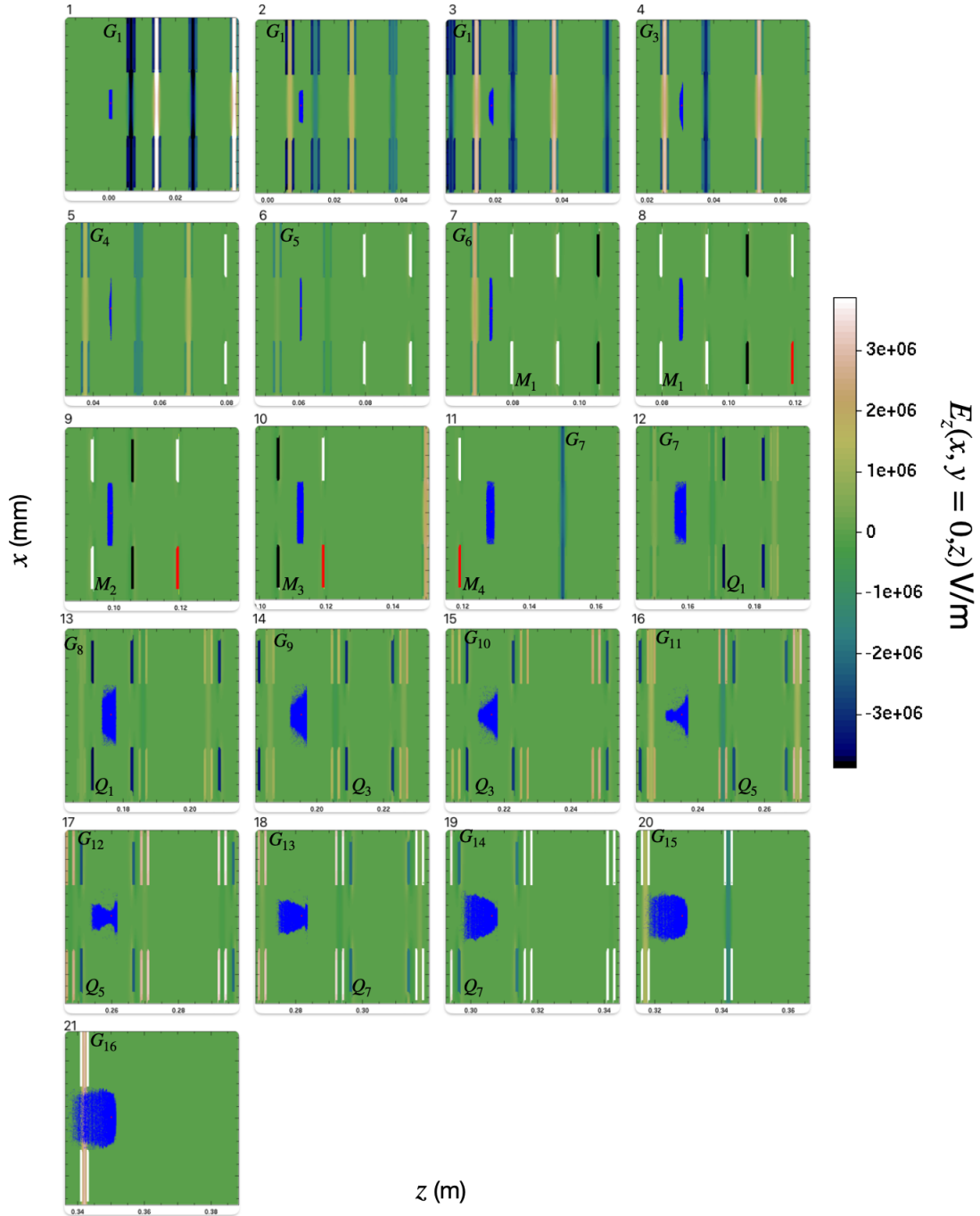


Figure A.5: A pulsed injection for a  $20^\circ$  beam and a 16-gap lattice designed for bunching and acceleration  $\phi_s = (0 \quad 3-\pi/3 \quad 120)$  with a 4-ESQ matching section and periodic doublet focusing. To minimize clutter, only the first gap to appear in the frame is labeled and the vertical axis of each plot is the same as previous plots in this section with the max extents  $\pm 1.5$  mm. Additionally, the first appearing ESQ is labeled either  $M_n$  for a matching ESQ or  $Q_n$  for a doublet ESQ. Note that each there are two ESQs for every  $Q_n$ -label. Also, due to complications with the cgm data writing, some conductors have off-coloring (red, teal). This has no physical significance.

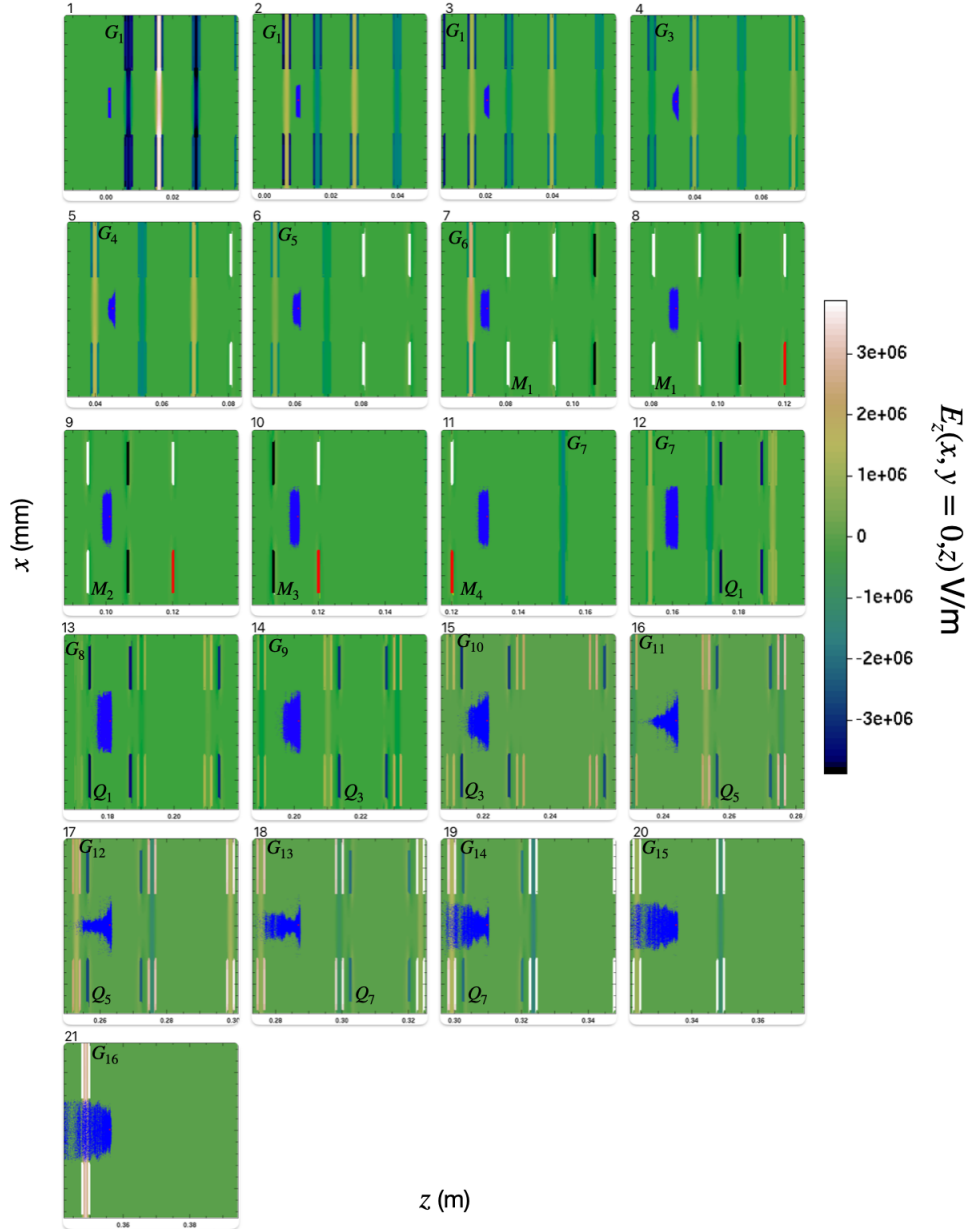


Figure A.6: A pulsed injection for a  $20^\circ$  beam and a 16-gap lattice designed for maximum acceleration  $\phi_s = 160$  with a 4-ESQ matching section and periodic doublet focusing. To minimize clutter, only the first gap to appear in the frame is labeled and the vertical axis of each plot is the same as previous plots in this section with the max extents  $\pm 1.5$  mm. Additionally, the first appearing ESQ is labeled either  $M_n$  for a matching ESQ or  $Q_n$  for a doublet ESQ. Note that each there are two ESQs for every  $Q_n$ -label. Also, due to complications with the cgm data writing, some conductors have off-coloring (red, teal). This has no physical significance.

## A.2 Continuous Injection

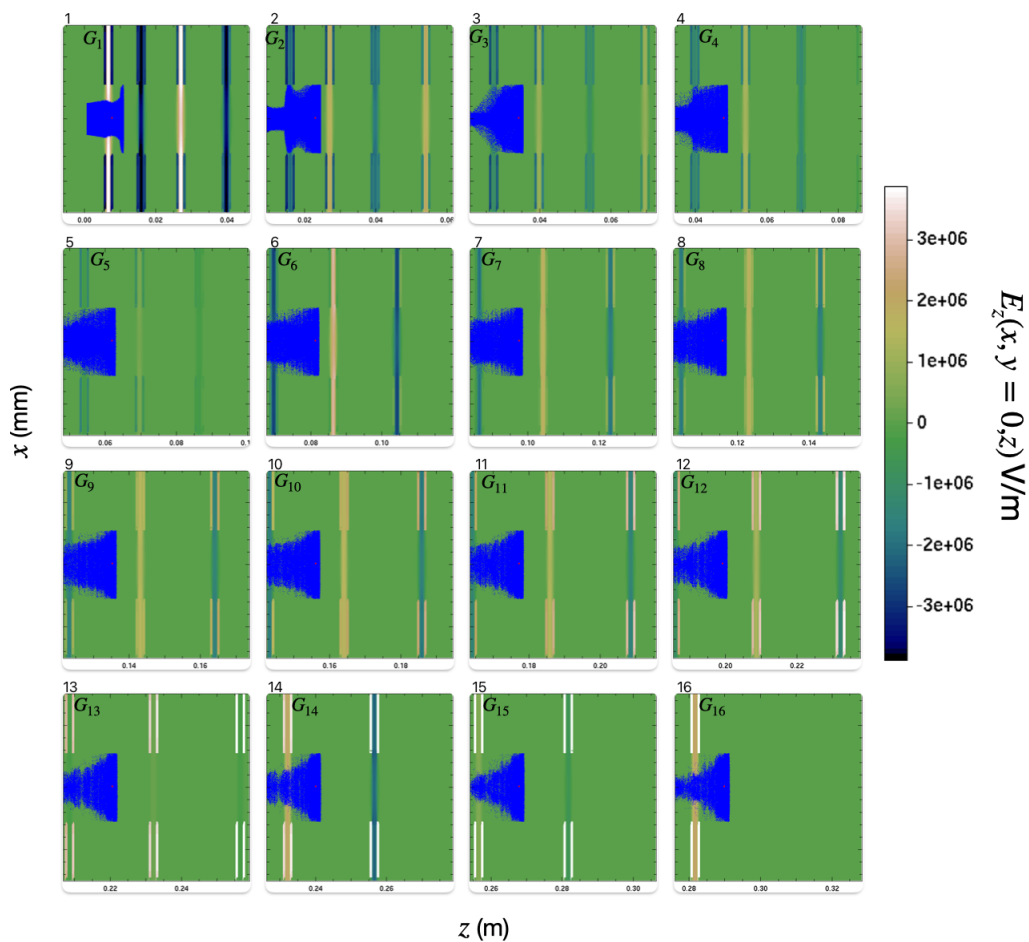


Figure A.7: A continuous injection and a 16-gap lattice designed for max acceleration  $\phi_s = 160$ . To minimize clutter, only the first gap to appear in the frame is labeled and the vertical axis of each plot is the same as those in plots above with maximum extents  $\pm 1.5$  mm. Panel 1 is a snapshot of the beam after approximately one RF cycle (full beam injected).

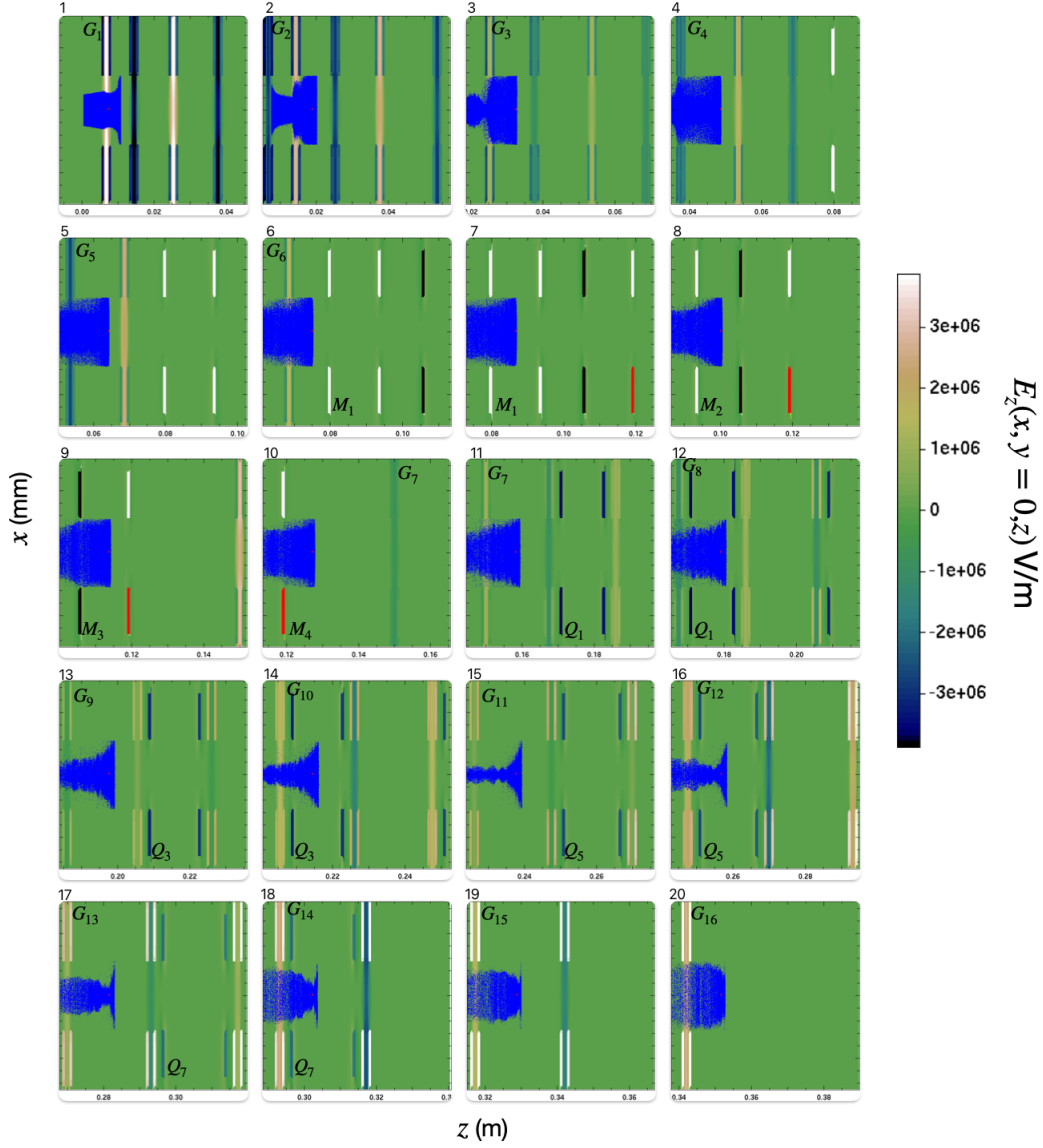


Figure A.8: A continuous injection and a 16-gap lattice designed for bunching and acceleration  $\phi_s = (0 \quad 3-\pi/3 \quad 120)$  with a 4-ESQ matching section and periodic doublet focusing. To minimize clutter, only the first gap to appear in the frame is labeled and the vertical axis of each plot is the same as previous plots in this section with the max extents  $\pm 1.5$  mm. Additionally, the first appearing ESQ is labeled either  $M_n$  for a matching ESQ or  $Q_n$  for a doublet ESQ. Note that each there are two ESQs for every  $Q_n$ -label. Also, due to complications with the cgm data writing, some conductors have off-coloring (red, teal). This has no physical significance. Panel 1 is a snapshot of the beam after approximately one RF cycle (full beam injected).

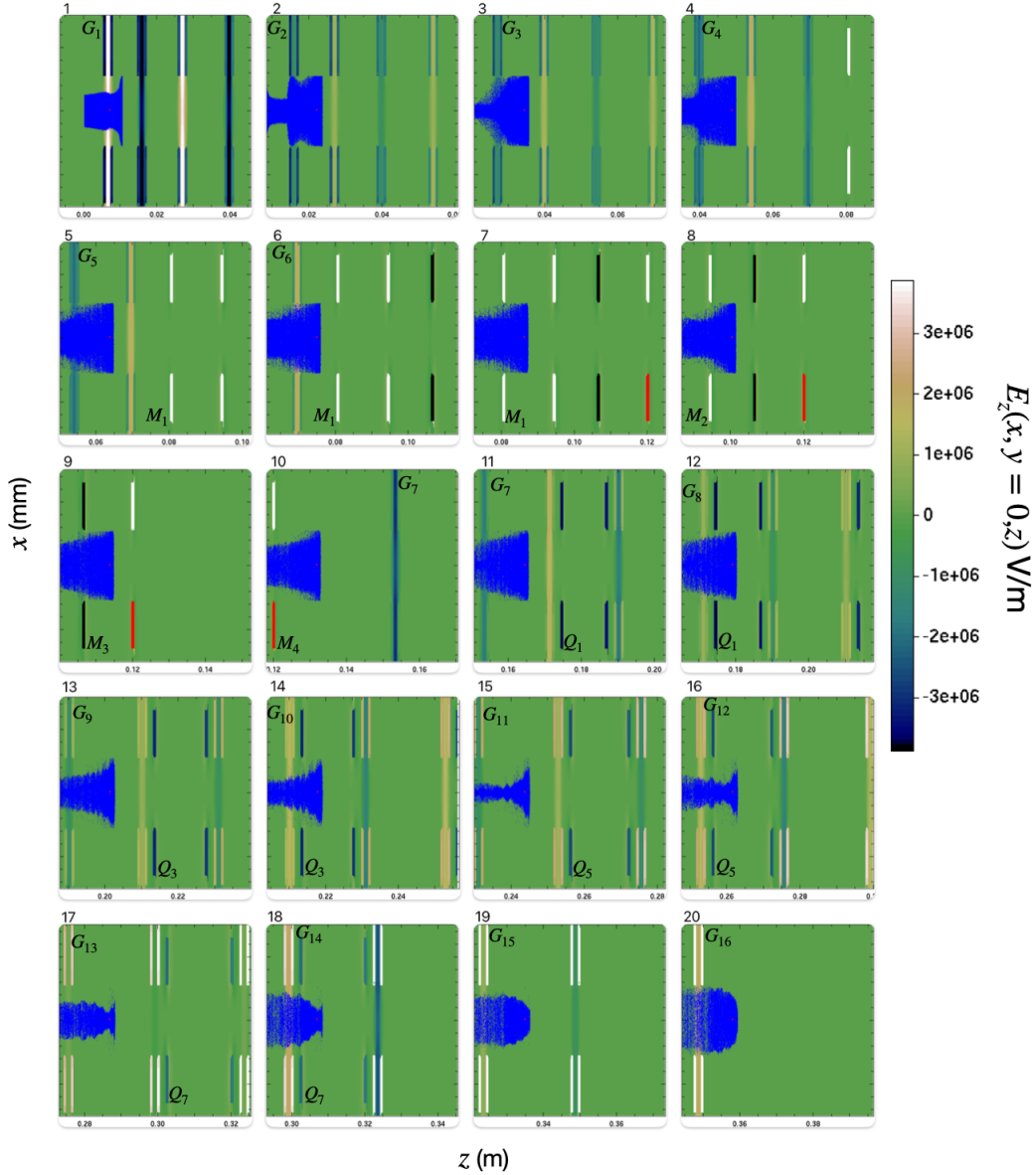


Figure A.9: A continuous injection and a 16-gap lattice designed for maximum acceleration  $\phi_s = 160$  with a 4-ESQ matching section and periodic doublet focusing. To minimize clutter, only the first gap to appear in the frame is labeled and the vertical axis of each plot is the same as previous plots in this section with the max extents  $\pm 1.5$  mm. Additionally, the first appearing ESQ is labeled either  $M_n$  for a matching ESQ or  $Q_n$  for a doublet ESQ. Note that each there are two ESQs for every  $Q_n$ -label. Also, due to complications with the cgm data writing, some conductors have off-coloring (red, teal). This has no physical significance. Panel 1 is a snapshot of the beam after approximately one RF cycle (full beam injected).