ETH zürich

EIDGENÖSSISCHE TECHNISCHE HOCHSCHULE ZÜRICH

Master Degree course in Physics

Master Degree Thesis

On understanding the differences in measured and calculated energy spread in the SwissFEL injector.

Supervisor: Dr. Andreas Adelmann
Candidate: Garðar Árni Skarphéðinsson

Academic Year 2021-2022

Acknowledgements

I would like the extend my deepest gratitude to my supervisor, Dr. Andreas Adelmann, for his constant support and guidance, as well as for his endless patience. I would also like to sincerely thank Dr. Sriramkrishnan Muralikrishnan for the dedicated assistance and advice he provided me with during the course of this thesis project.

I also wish to thank Sven Reiche, Thomas Geoffrey Lucas, and Paolo Craievich for their time and insights during all our meetings over the months.

Lastly, special thanks go out to my partner Maria, whose support was invaluable during my whole time of studies at ETH, and especially during the thesis project.

Abstract

In this thesis we endeavored to simulate the SwissFEL injector using the OPAL library. We investigated the effects of grid size and particle number on the calculated longitudinal energy spread in an attempt to find an explanation for the discrepancy between previous numerical calculations and measurements. Additionally, we developed algorithms to calculate values such as slice energy spread and slice emittance, which had previously not been done with OPAL data.

We examined and dealt with problematic effects which arose due to the heightened resolution of the simulation, but determined that simply increasing the resolution of the simulated mesh was not enough to explain the aforementioned discrepancy. We performed simulations using different field-solvers and concluded that only calculating particle-mesh interactions leads to a significant underestimate of the contribution of intrabeam scattering to the beam's energy spread.

Contents

Chapter 1

Introduction

The SwissFEL facility at the Paul Scherrer Institut in Villigen, Switzerland, began construction in 2013 and has been operational since the end of 2016. The function of a free-electron laser (FEL) such as this is to produce high brilliance X-ray pulses for the purposes of research and applied science in a multitude of fields, including biology and material science.

In the first component of the machine, the RF gun, a laser pulse is used to induce the emission of an electron bunch from a metal plate via the photo-electric effect. This bunch passes through linear accelerator sections to gain longitudinal momentum and is focused transversely, as well as longitudinally, by use of a multitude of magnets and chicanes. It then passes through a section of alternating dipole magnets, or undulators, which gives the electrons alternating transverse acceleration, leading to the release of X-rays.

Recent measurements of the beams longitudinal energy spread give values that far exceed those predicted by simulations. Specifically, measurements give values close to 10 keV. This is inconsistent with the numerical predictions made with codes such as Astra and Elegant, which predict energy spread values smaller than 1 keV.

In chapter 2 of this thesis we will discuss the parts of the machine we focused on simulating, as well as OPAL, the code library we used for our simulations. Additionally we will give an overview of the theory of intrabeam scattering, how it affects the energy spread, and how we attempted to account for these effects in our simulations.

In chapter 3 we will explain the simulations we performed. We will discuss our initial goal of matching the output of OPAL with that of Astra, as well as our work in gradually increasing the resolution of our simulations, and other methods with which we hoped would explain the previously mentioned discrepancy.

In chapter 4 we will then go over the numerical evaluations of our simulation outputs. We will discuss how the increased resolution affected our output, and we will go over some issues that arose during that process and how we overcame them. This chapter will then be followed by a summary of our results and conclusions.

Chapter 2

Theory & Model

2.1 The SwissFEL Injector

While the total length of the SwissFEL facility is about 719 m, the primary focus of this thesis was specifically on the first 13 m of the FEL injector. In this part of the machine the beam still has a low enough energy for the beam quality to be easily improved or destroyed by small changes to input parameters. These first 13 m consist of the following components [**SwissFELDesign**]: an RF electron gun, a focusing solenoid, two linear accelerator cavities, and four solenoids along each of those cavities.

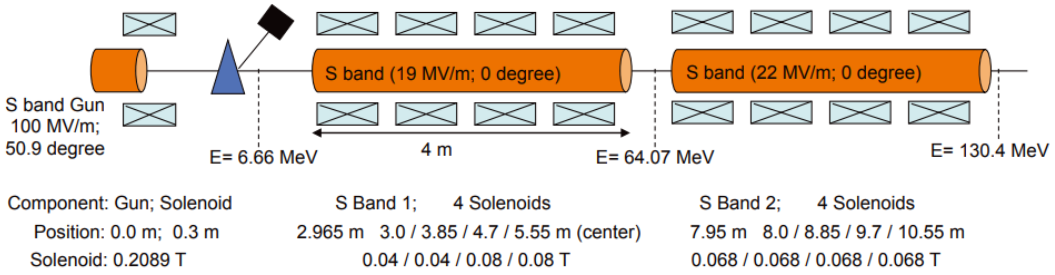

Figure 2.1: A diagram of the first 13 m of the SwissFEL injector [**SwissFELDesign**].

As shown in the diagram in figure 2.1, this early part of the structure will accelerate the beam to about 130 MeV energy by the time it reaches the 13 m mark.

The measurement mentioned in the introduction is performed at over 100 m downstream from the RF gun. While this means that we would not expect the energy spread to reach values near 10 keV when we simulate only the first 13 m, the measurement still implies that the energy spread should have reached a value well above 1 keV at the end of this section [**Meetings**].

2.2 OPAL

The object oriented parallel accelerator library, or OPAL for short, is an open source library written in the MAD programming language. Its main purpose is to simulate

charged particle beams in linear accelerators and cyclotrons [**2019arXiv190506654A**].

OPAL is a fully three dimensional electrostatic code that is specifically built for parallel processing, making it relatively simple to simulate a high number of particles and to use very fine grids for the particle-in-cell (PIC) field-solver. The field-solver used by Astra, on the other hand, has reduced dimensionality, specifically, it is said to be 2.5 dimensional. In addition to this, the capacity of Astra to simulate huge numbers of particles is not as great as that of OPAL, making the latter better suited for the kind of study we aimed to perform.

The PIC field-solver we used calculates particle-mesh (PM) interaction, where the density of charged particles is projected onto a mesh, or grid. The fields from said particles are then calculated on those grid-points and interpolated onto the position of each particle. For systems containing a large quantity of particles this is a very efficient way to compute each step, however, there is a certain loss of information. Firstly, the simulated particles are so called macro-particles, which stand in for a large number of physical particles. Secondly, within each grid-point, or cell, the forces between such neighboring particles are underestimated, since directly computing all such pairwise forces would be very computationally expensive.

The OPAL simulations done for the purposes of this thesis were mostly run on the MERLIN cluster at PSI, while a few, more computationally expensive ones, were run at NERSC (the National Energy Research Scientific Computing Center), at the Lawrence Berkeley National Laboratory in California.

2.3 Energy Spread

The energy spread of a particle beam is, simply put, the standard deviation of the beam's energy, i.e.

$$dE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(E_i - \mu)^2} \tag{2.1}$$

Where N is the number of particles in the beam, E_i is the energy of the particle labeled i, and μ is the mean energy of the beam.

For a given dimension, the energy spread can be visualized as the width in energy of said dimension's phase-space curve. In this thesis we imagine the electron beam travelling along the z-axis, perpendicular to the xy plane. The energy spread we concern ourselves with is the z-directional, or longitudinal, energy spread. Note that from here on z will refer to the longitudinal coordinate within the electron beam, while the longitudinal coordinate along the SwissFEL injector shall be called s.

2.3.1 Intrabeam scattering

Energy spread in an accelerator beam mainly arises from a few separate effects, those being: intrabeam scattering, micro-bunch instability and RF curvature. Intrabeam scattering is a process of multiple coulomb scatterings within a particle beam that causes the

gradual diffusion of the beam in both transverse and longitudinal dimensions [**Piwinski**]. To counteract this diffusion focusing magnets are used to keep the particles together.

Micro-bunch instability is an effect caused by shot noise, and characterised by the formations of self-fields within the electron beam and, consequently, the appearance of micro-structures in the longitudinal phase-space. The subjects of micro-bunch instabilities and RF curvature are beyond the scope of this thesis. In the first stages of a beam's acceleration the main contributor to energy spread is thought to be intrabeam scattering, so the other effects will not be considered further [**Meetings**].

The theory of intrabeam scattering and how it affects longitudinal energy spread in synchrotrons has been laid out by Piwinski [**Piwinski**], and applied to LINACs by Huang [**Huang**]. In the latter's formulation, the instantaneous growth rate of the energy spread in a transversely symmetrical system, such as that of SwissFEL, is given as

$$\frac{1}{\tau_\delta} = \frac{r_e^2 c N}{8\pi^2 \gamma^2 \epsilon_x^n \sigma_x \sigma_z \sigma_\delta^2} \ln\left(\frac{\Delta\gamma_{max}}{\Delta\gamma_{min}}\right)$$

where r_e is the classical electron radius, N is the number of particles in the beam, γ is the Lorentz factor, ϵ_x^n is the normalized transverse beam emittance, σ_x and σ_z are the RMS (Root Mean Square) transverse and longitudinal beam sizes respectively and $\sigma_\delta = dE$ is the longitudinal energy spread. The factors in the logarithm are the maximum and minimum energy exchanges

$$\Delta\gamma_{max} \sim \gamma^2 \sigma_{x'} \qquad \Delta\gamma_{min} \sim r_e/(\sigma_x \sigma_{x'}).$$

In his article, Huang argues that further approximations can be made on the basis that the intrabeam scattering energy spread has a near Gaussian core. Since we only care about this core, he argues, we can somewhat arbitrarily limit the maximum energy transfer. Making a few more approximations and assumptions on the behaviour of the beam, Huang finally arrives at an integrated equation for the energy spread

$$\sigma_\delta = \sqrt{\sigma_{\delta,0}^2 + \frac{2 r_e^2 N}{\gamma^2 \epsilon_x^n \sigma_x \sigma_z} \Delta s} \equiv \sqrt{\sigma_{\delta,0}^2 + \sigma_{\delta,IBS}^2}$$

where $\sigma_{\delta,0}$ is the initial longitudinal energy spread, and Δs is the distance over which the beam travels and accumulates energy spread via the intrabeam scattering.

Di Mitri [**DiMitri**] showed that, for the parameters of a state-of-the-art photoinjector, this model predicts an accumulation of longitudinal energy spread of about 3 keV over a distance $\Delta s = 30$ m. It is important to note that this formulation assumes no energy dispersion nor particle acceleration to take place, making it not directly applicable to our simulations of SwissFEL. Nonetheless, it is interesting that Di Mitri's prediction is of around the same order of magnitude as the measurement from SwissFEL mentioned in section 2.1 above.

2.3.2 Slice energy spread

The thickness of a curve is a phrase that can be interpreted in a few different ways. If the energy spread of the beam is taken to simply be the whole range of values in which the beam's energy lies, i.e. equation 2.1 is used, one gets what is called the 'projected' energy spread. While this quantity can be of interest, it is not representative of the average energy spread along the beam.

The energy spread of more relevance to us is the so called slice energy spread. It is the width in energy of a slice of the longitudinal phase-space curve at a given point along the beam. At a certain location in the beamline, the beam's energy spread can be said to be the mean of these slice energy spreads

$$dE = \frac{1}{M}\sum_{j=1}^{M}\sqrt{\frac{1}{N_j}\sum_{i=1}^{N_j}(E_i - \mu_j)^2}$$

where M is the number of slices, N_j is the number of particles in the slice labeled j and μ_j is the mean energy of the same slice.

For a heavily sloped phase-space curve however, the energy width calculated in any given slice will be an overestimate due to the slope of the curve contained in said slice. In averaging over the whole beam, this additional slope factor can be decreased by increasing the number of slices the beam is split into, but this can quickly become computationally expensive and inefficient.

A more efficient way to eliminate the overestimate due to the slope is to fit a polynomial to the phase-space curve, and then subtract said fit from the data. So instead of plotting P_z against z, we would plot $P_z - p(z)$ against z, where $p(z) = p_0 z^n + \ldots + p_{n-1}z + p_n$. This will eliminate the slope while preserving the energy width along the beam. This is called the uncorrelated energy spread, and is presented in figures 2.2 and 2.3.

The $n+1$ coefficients of the fitting polynomial p were obtained by minimizing the error

$$E = \sum_{j=0}^{N-1} |p(z_j) - P_{z,j}|^2$$

in the Vandermonde matrix equation

$$\begin{bmatrix} z_0^n & z_0^{n-1} & \cdots & z_0 & 1 \\ z_1^n & z_1^{n-1} & \cdots & z_1 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ z_{N-1}^n & z_{N-1}^{n-1} & \cdots & z_{N-1} & 1 \end{bmatrix} \cdot \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_n \end{bmatrix} = \begin{bmatrix} P_{z,0} \\ P_{z,1} \\ \vdots \\ P_{z,N-1} \end{bmatrix}$$

where z and P_z are the $N \times 1$ position and momentum arrays of the beam respectively, N being the number of simulated particles. The order of the polynomial used was empirically selected to be $n = 14$. These calculations were performed using [**harris2020array**].

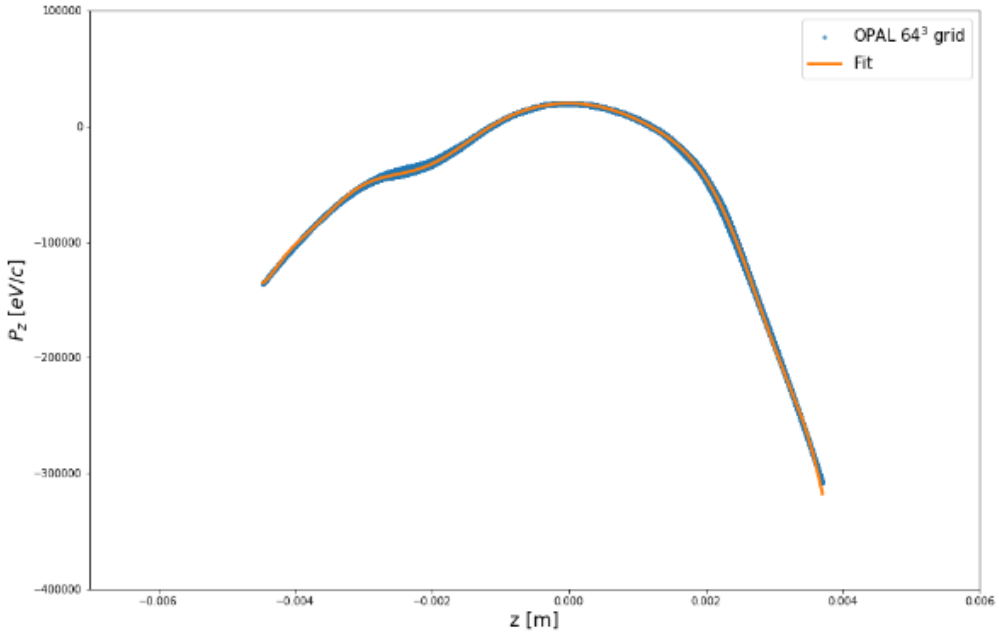

Figure 2.2: The longitudinal phase-space curve of an electron beam simulated in OPAL with a 64^3 grid and 6 particles per grid-point. The data is a snapshot of the phase-space at 2.95 m along the injector and is fitted with a polynomial.

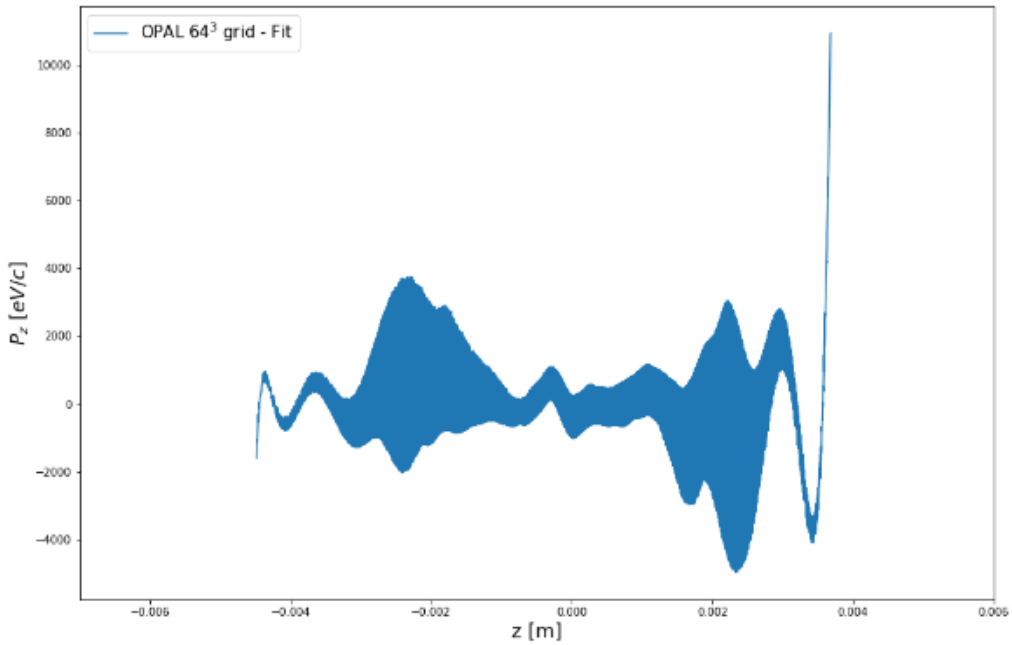

Figure 2.3: The phase-space data shown in figure 2.2, but now the polynomial fit has been subtracted.

The remaining structure can now be split into a number of slices, the energy width of each slice calculated, and finally the energy spread of the beam is said to be the average energy spread of the slices. The average taken is weighted by the number of simulated particles contained in each slice to minimize the effects of sparsely populated slices. The uncorrelated slice energy spread is then

$$dE = \frac{1}{N}\sum_{j=1}^{M}\sqrt{N_j\sum_{i=1}^{N_j}(E_i-\mu_j)^2}$$

where $N = \sum_j^M N_j$ is, once again, the total number of particles in the beam.

From figure 2.3 it can be seen that at the front and back ends of the phase-space data the fit starts to diverge from the curve. This creates regions that would contribute unreasonably high values to the energy spread average due to their slope. To avoid this issue we note that, as was previously mentioned, the particle density is far from uniform across the beam, and is in fact very low at the front and back ends. This is demonstrated in figure 2.6.

We therefore allow ourselves to simply drop the first and last few slices from our calculations, leaving us with only the core of the data, which is centered around $P_z = 0$. If we split a typical OPAL beam into 25 slices and drop the first and last 9 slices, leaving only 7 left, we will only have excluded $20\% - 25\%$ of the beam's total particle number from our calculation. This process is depicted in figures 2.4 and 2.5.

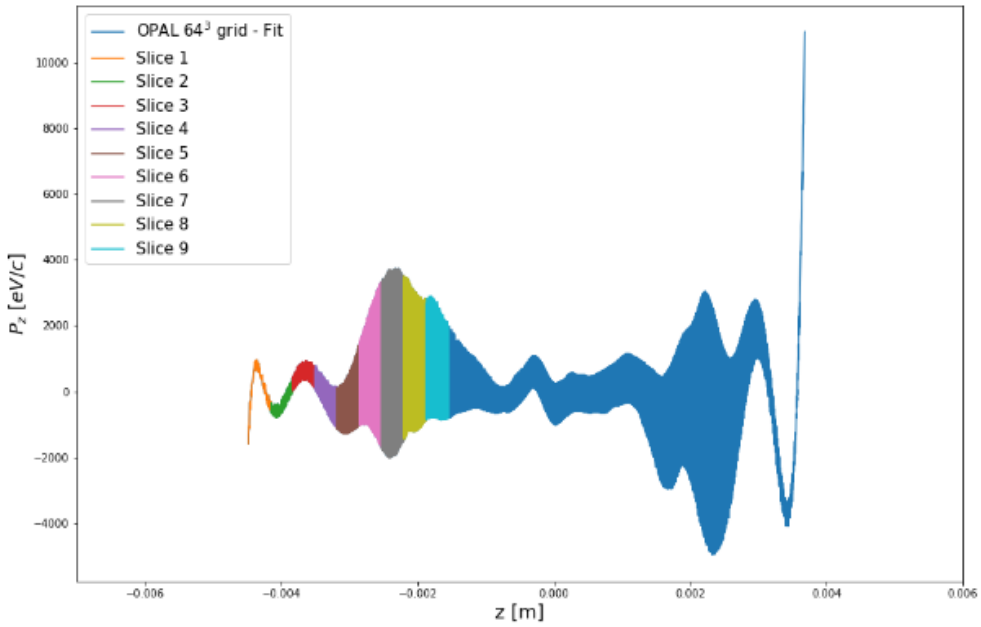

Figure 2.4: The data from Figure 2.3 with the first 9 slices colored in. Note how the data forms bulbs and then diverges from $P_z = 0$ at the front and back ends.

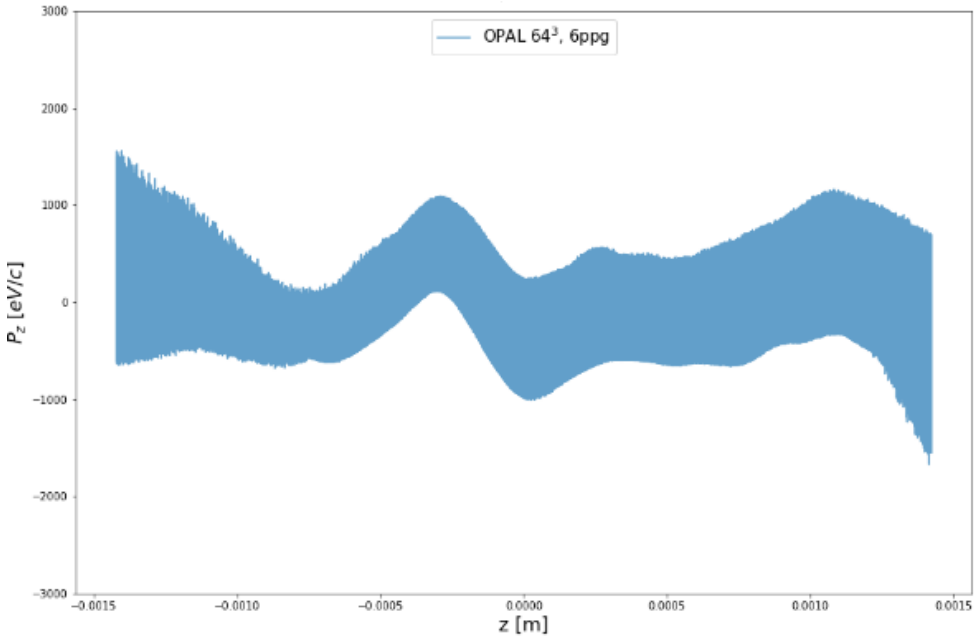

Figure 2.5: The core of the data from figure 2.4. The first 9 and last 9 slices removed.

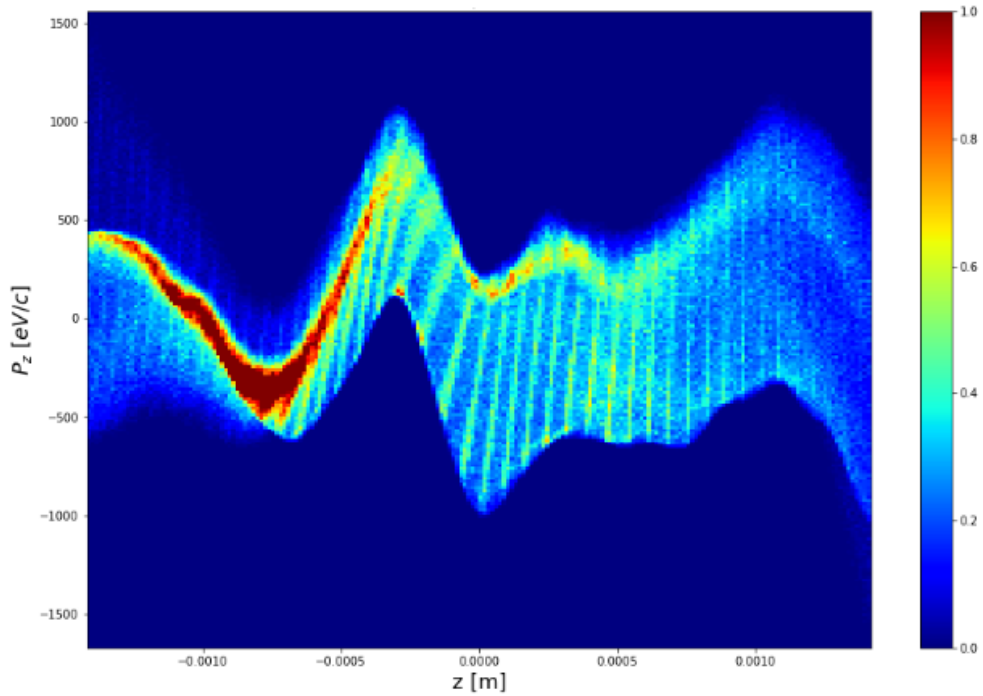

Figure 2.6: The normalized phase-space density for the data in Figure 2.5. Most of the beam's particles are concentrated in a relatively small section of the core.

2.4 Slice Emittance

Though emittance is not the main subject of this thesis, it is nonetheless an important quantity in accelerator physics in general. It refers to the area occupied by the beam in phase space and has the units of [length · angle]. One can calculate the RMS emittance of a particle beam at a given point, for each of the spatial dimensions, with the equation

$$\epsilon_i = \sqrt{\langle x_i^2 \rangle \left\langle p_{x_i}^2 \right\rangle - \langle x_i p_{x_i} \rangle^2}, \qquad \{x_1, x_2, x_3\} = \{x, y, z\} \tag{2.2}$$

where x_i and p_{x_i} are, respectively, the position and momentum vectors of the beam at said given point for the relevant dimension. Of interest to us is the transverse emittance, that is the emittance $\epsilon_x = \epsilon_y$.

The calculations of the transverse emittance for the simulated beam is, in many ways, analogous to that of the longitudinal energy spread. We could simply use equation 2.2 to calculate at the projected RMS values of the emittance for the whole beam at a certain point along the injector, but this would result in an overestimate. We can instead take the beam data at said point, split it into slices, and use the same equation to calculate the RMS emittance of each slice. These slices are depicted in figures 2.7 and 2.8.

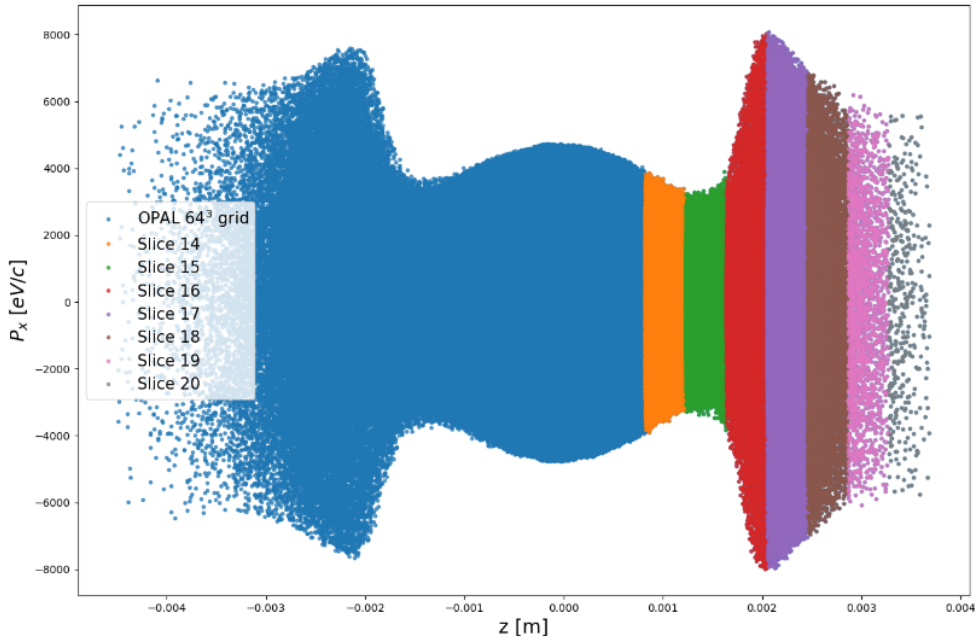

Figure 2.7: The transverse momentum of the beam described in figure 2.2, plotted along the longitudinal dimension z. The last 7 slices are colored in.

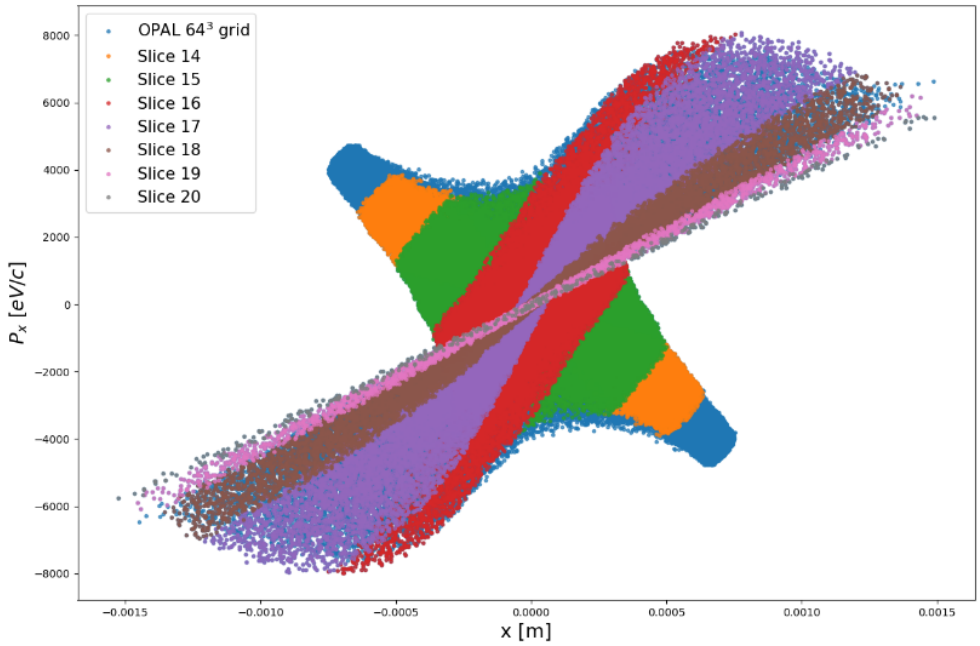

Figure 2.8: A 'frontal view' of the beam data in Figure 2.7. Note how each slice covers only a small part of the total transverse phase-space area of the beam.

For the transverse momentum there is no need for any clever tricks like in the section above; we simply slice the data along the longitudinal dimension and calculate the emittance of each slice using equation 2.2. As with the longitudinal energy spread, a weighted average is then taken over the slices. The transverse slice emittance is then calculated as

$$\epsilon_x = \frac{1}{N} \sum_{j}^{K} N_j \epsilon_{x,j} \tag{2.3}$$

where $N = \sum_i^K N_i$ is the total number of particles in the beam, K is the number of slices, and $\epsilon_{x,i}$ is the RMS transverse emittance in the slice labeled with i, as calculated with equation 2.2.

Chapter 3

Simulations

So far we have described the length of the simulations in terms of the distance the beam travels, i.e. the first 13 m of the injector section. However, it would be more accurate to describe it in terms of time, as OPAL simulates the beam one timestep after another.

The length of each timestep is defined in the initial input, and for the simulations discussed in this thesis we have stuck to a constant pair of values throughout. For the first 0.2 m of the injector, the timestep is set to be $dt = 100\,\text{fs}$, while after that threshold, it is set to be $dt = 2\,\text{ps}$.

A simulation that we say covers 13 m could instead be said to cover 43.4 ns consisting of about 28,000 timesteps. About a quarter of those steps cover the first 0.2 m of the distance travelled, so we could say the first couple of centimeters of the injector are in fact simulated in more detail than the rest. This is done because the initial emission and propagation of the beam is crucial. If a coarseness in the timesteps caused any kind of anomaly or error at the start of the simulation, those effects could then propagate with the beam through the rest of the simulation, spoiling the output data.

3.1 Proof of Concept

Before diving right in, we first needed to show that we could reproduce the output of Astra simulations using OPAL. It is important to show that we can indeed simulate the same system with the two codes so the comparison between the two can be seriously considered. To do this we define a parameter-space in which we vary the configuration of OPAL inputs and then compare the outputs of the following simulations to that of Astra.

We chose three input parameters to define our parameter-space: the longitudinal and transverse beam dimensions, that is to say the RMS beam radius and length, as well as the number of energy bins the beam is split into. The splitting of the beam by energy into bins affects how costly the calculations of the field-solver are. Each bin is approximated as an electrostatic object, so while an increased number of bins is more accurate, it can quickly become computationally expensive.

Using the OPAL sampler tool, we defined a range for each of these parameters and assigned them ten values each within said ranges, creating a total of 1000 distinct configurations. A simulation was then done for each of these configurations, the output of

which was compared to that of Astra.

To determine what we consider to be the best configurations, we look at specific output quantities of both Astra and the 1000 sampler runs and see which OPAL configuration minimizes the L_2 discrepancy for said quantities. We do this for four different output quantities, those being the projected RMS values of: the beam radius, the beam length, the energy spread and the transverse beam emittance.

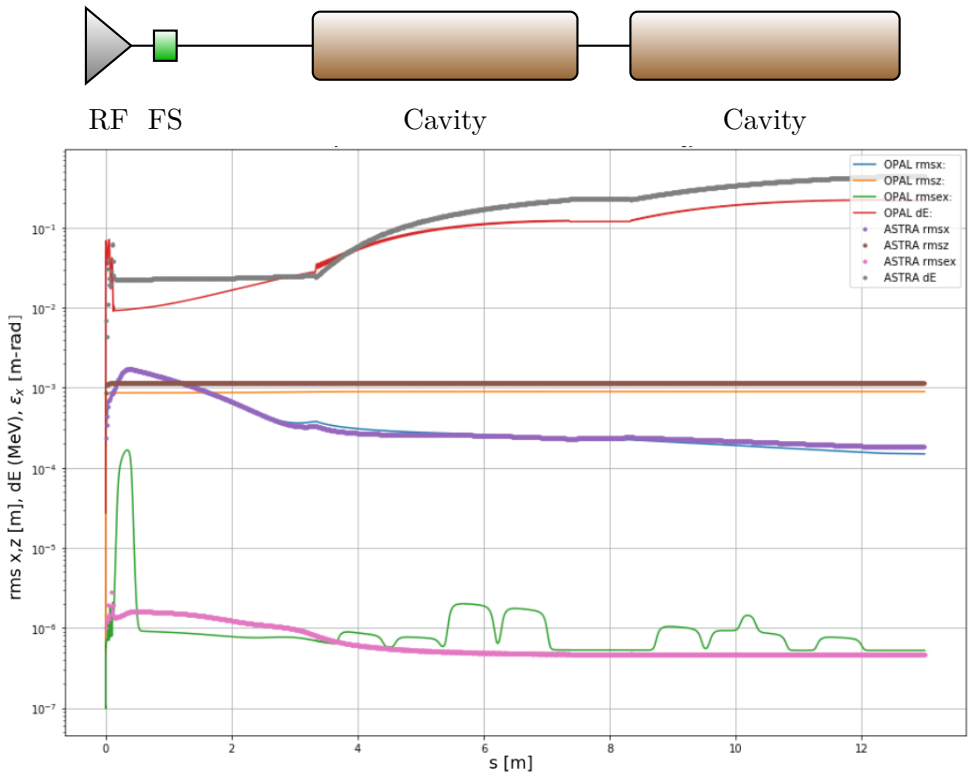

Figure 3.1: The comparison of four output quantities from both OPAL and Astra over 13 m. The quantities are the projected RMS values of: the beam radius (rmsx), the beam length (rmsz), the transverse emittance (rmsex) and the longitudinal energy spread (dE). Above the plot is a simple diagram of the main SwissFEL components covered, roughly positioned along the plot's s-axis. The RF gun followed by a focusing solenoid, and two S-band cavities.

Once we determined that we had an OPAL simulation whose output values for these quantities matched that of Astra's sufficiently well, we could move on. Choosing a configuration where the L_2 discrepancy between the OPAL and Astra RMS beam radius was minimized yielded favourable result, which are shown in figure 3.1. Note that in the figure, the OPAL data exhibits localized bumps, or hills, in the transverse emittance. These structures correspond to the solenoids along the cavities, and only appear in the OPAL data due to the fact that OPAL considers momentum canonically, whereas Astra looks only at kinetic momentum.

3.2 Beam Distribution

While matching the projected RMS outputs of the simulations is important, one must not neglect comparing the phase-spaces as well. When doing this we found that, while the projected values of certain quantities seem to match, corresponding phase-space structures can still look quite different.

After selecting an OPAL input configuration as described in the section above, we noted that the longitudinal phase-space curves of Astra and OPAL were significantly mismatched as seen in figure 3.2. While the projected RMS energy spread of the simulations seemed relatively similar, the structure of the longitudinal phase-space of the OPAL run showed a relatively flat structure with a sharp dip and peak at the left and right side respectively. This is in contrast to the parabolic shape we observe from the Astra output.

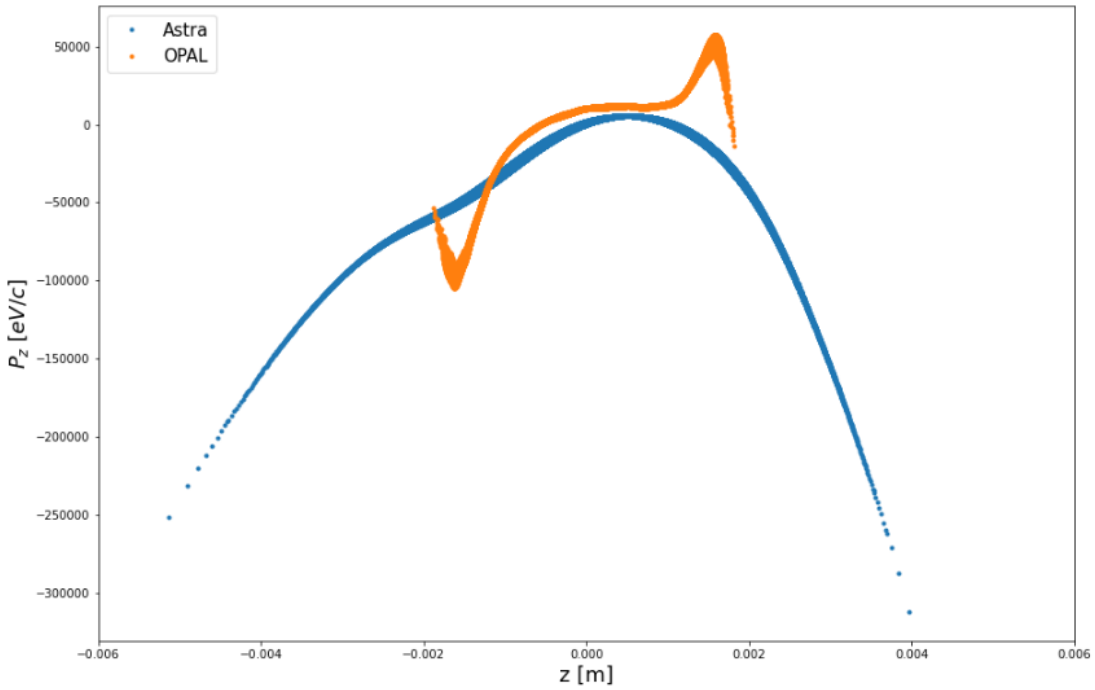

Figure 3.2: The longitudinal phase-space of two beams, one from an Astra simulation and another from an OPAL simulation, at $s = 2.95\,\mathrm{m}$.

The issue here had to do with the initial spatial distribution of the emitted beam. In Astra the beam is approximated to have a cylindrical shape, with a truncated Gaussian transverse distribution and a Gaussian longitudinal distribution. Up until this point we had defined the distribution of the OPAL beam as being Gaussian transversely and a flat-top Gaussian longitudinally. The idea had been for the OPAL beam to be more true to a physical beam.

Adjusting the distribution of the OPAL beam, changing it to be both longitudinally and transversely Gaussian, and adding a screen with a circular opening, or a collimator, to the start of beamline, fixed this issue. After then adjusting the initial distribution

values to match that of Astra, both projected RMS values and phase-space figures of the OPAL simulation were consistent with the Astra data. The longitudinal phase-space of OPAL with this new distribution is compared with that of Astra in figure 3.3.

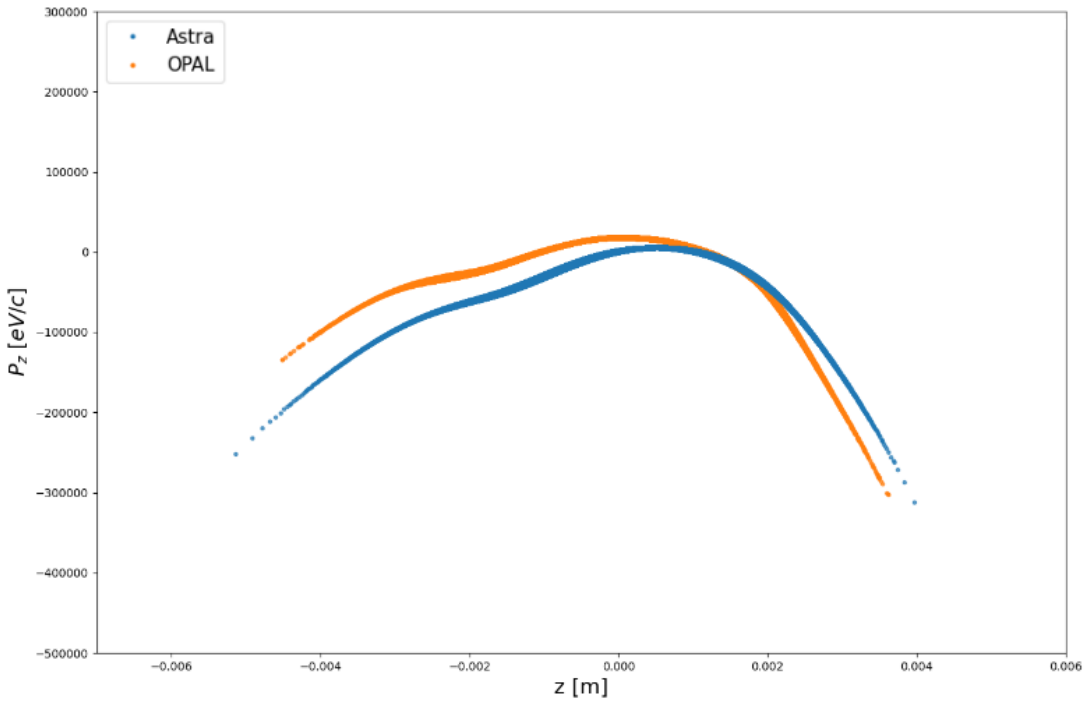

Figure 3.3: The longitudinal phase-space of two beams, one from an Astra simulation and another from an OPAL simulation with the same initial distribution as Astra, at $s = 2.95\,\mathrm{m}$.

3.3 Higher resolutions

With the beam input parameters selected, we moved on to doing simulations of higher resolution. Up until this point most of our simulations had been done on a cubic grid consisting of 16^3 grid-points, with 200,000 simulated particles. For a beam with a total charge of 200 pC, that comes down to each simulated macro-particle representing about $6 \cdot 10^3$ physical electrons, with about 49 such particles being projected onto each grid-point.

Steps were taken to increase the resolution, both by increasing the number of simulated particles per grid-point, as well as making the grid finer, that is to say, increasing the total number of grid-points. In doing this we could move ever closer to even doing $1:1$ simulations. The hope was that with these higher resolutions particle-particle interaction would be better simulated and we would see more contribution from this intrabeam scattering to the energy spread.

The steps in which these changes were made, and the total number of particles in

Table 3.1: Total number of simulated particles for a given grid size and number of particles per grid-point.

	6 p.p.g.	12 p.p.g.	24 p.p.g.	48 p.p.g.
32^3	196,608	393,216	786,432	1,572,864
64^3	1,572,864	3,145,728	6,291,456	12,582,912
128^3	12,582,912	25,165,824	50,331,648	100,663,296

the corresponding simulations, can be seen in table 3.1. This exponential increase in simulated particles did of course increase the computational cost of our simulations, and so we had to increase the number of CPU cores used accordingly.

As discussed in section 2.2, OPAL is designed for parallel processing. A simulation with a 64^3 grid and 48 particles per grid-point takes just under one hour to finish on 44 CPU cores, that is, about 0.125 s per simulated timestep. But as we continued to refine the grid, the issue of computational constraints became ever more apparent.

Refining the grid even further, we now started increasing only the number of grid-points along the longitudinal dimension of the beam independently of the number of transverse grid-points. This was done with the intention of continuing the raising of resolution without incurring too high of a cost when it came to computation time. As we increased our grid resolution past $128^2 \cdot 512$ we started running the simulations on the NERSC cluster, as the computational capacity of MERLIN became insufficient.

As shown in table 3.2, we raised the resolution as far as a $256^2 \cdot 1024$ grid with 6 particles per grid-point, which gives us a ratio between simulated and physical particles of about $1 : 3$. This is just over 2000 times higher than the ratio we started out with.

Table 3.2: Total number of simulated particles for given grid sizes with 6 particles per grid-point. The third column shows the ratio between simulated macro-particles and physical particles in a 200 pC beam.

Grid size	6 p.p.g.	Macro-particle ratio
32^3	196,608	1 : 6144
64^3	1,572,864	1 : 768
128^3	12,582,912	1 : 96
$128^2 \cdot 256$	25,165,824	1 : 48
$128^2 \cdot 512$	50,331,648	1 : 24
$128^2 \cdot 1024$	100,663,296	1 : 12
256^3	100,663,296	1 : 12
$128^2 \cdot 2048$	201,326,593	1 : 6
$256^2 \cdot 512$	201,326,593	1 : 6
$256^2 \cdot 1024$	402,653,184	1 : 3

3.4 Booster 2

Although this thesis focused heavily on properly simulating the first 13 m of the injector, we also wanted to experiment with simulating larger parts of the machine. We therefore decided to perform simulations that reached up to about 66 m. This means we covered the laser heater, the whole of the section referred to as Booster 2, and the first bunch compressor section up to the compressor chicane.

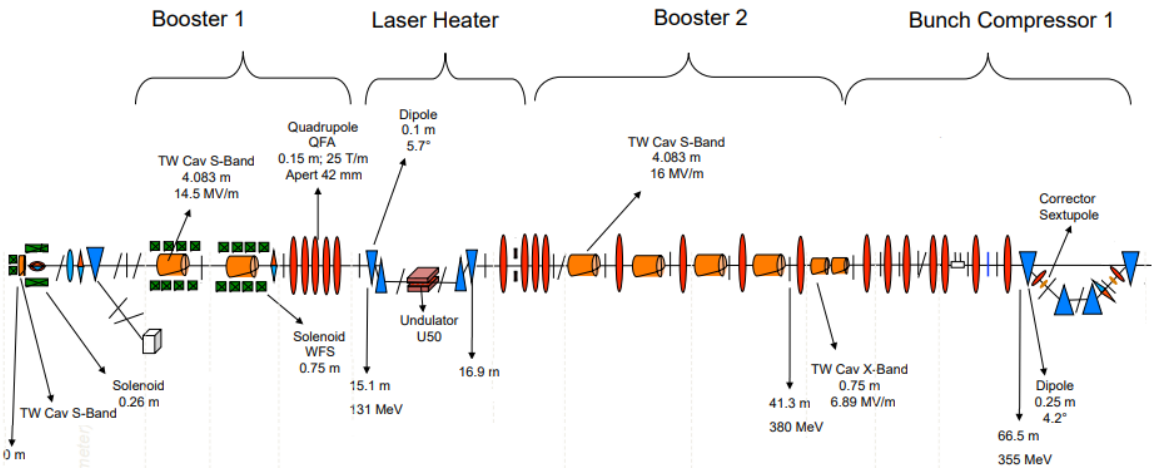

Figure 3.4: A diagram of the SwissFEL injector, up to and including the first bunch compressor [**SwissFELDesign**].

A variety of components needed to be added to the simulated beamline for it to accurately describe 66 m of the machine. Four new S-band cavities were added, followed by two X-band cavities operated at the fourth harmonic frequency of the S-bands. These cavities were calibrated so the beam's energy was raised to around 350 MeV when it reached the bunch compressor. Amidst the cavities a large number of quadrupole magnets were also simulated.

The whole of the laser heater section, which can seen along with the other components just listed in figure 3.4, was simply simulated as a drift space. We justify this simplification by saying that we consider only the special case in which the external laser is not applied and the undulator is inactive. While an inactive undulator does provide some beam focusing [**Meetings**], we accept its loss as a justifiable expedient.

As was already mentioned in the introduction and in section 2.1, the measurement which motivates this thesis is performed at around $s = 113\,\text{m}$ and is close to 10 keV. Specifically, a measurement performed when the laser heater was turned off gave an energy spread value of about 6.7 keV. While this measurement is performed far downstream from the sections we focus on in this thesis, these values would still suggest that the energy spread at $s = 13\,\text{m}$ should be nearly 2 keV [**Meetings**]. In increasing the length of our simulations to reach further downstream, we will see how the energy spread develops as the simulated beam propagates, and whether or not it rises to approach the measured values.

3.5 P^3M

If increasing the resolution of the simulations doesn't produce the predicted effects of intrabeam scattering on the energy spread, another approach would have to be considered. If the reason for the lack of properly represented particle scattering events is the PM method itself, a way to achieve this representation could be adjusting the field-solver to compute a mixture of particle-particle (PP) and particle-mesh (PM) interaction. This is the so called P^3M method.

Using this method, an interaction radius must be selected. For each particle, pairwise interactions will then be computed for every other particle within said radius, whereas interaction with particles outside the radius will be treated as interaction with the grid like before. The total force on each particle will then be a sum of the interaction with the mesh and the direct interactions with all other close-by particles. In doing this we could bypass the problem mentioned in section 2.2, where the forces between particles within the same cell are underestimated.

The obvious issue here is that of computational complexity, both in terms of computation time and disc storage. While the time cycle of a PM interaction timestep is $\mathcal{O}(N)$, where N is the number of simulated particles, the additional PP interactions adds a factor $\mathcal{O}(N^2)$ to the total time cycle. A more comprehensive review of the working of the P^3M algorithm can be found in [**book:9047**].

Another issue is that of the macro-particles. If each simulated particle carries the charge of multiple electrons, the P3M solver would overestimate the pairwise particle-particle forces. For reliable results one would therefore want to use this solver only with a $1:1$ simulation setup.

For a 200 pC beam, a $1:1$ setup would entail simulating 1,248,439,451 particles, which would be enormously time consuming. A compromise could therefore be to simulate a beam of lower charge. This yield data which differs significantly from the 200 pC simulations, but it would nonetheless enable us to examine the differing effects of the PM and P^3M field-solvers on the longitudinal energy spread.

Another way in which we can cut down on our simulations run time is to examine where precisely in our beamline intrabeam scattering should be of significant importance, and only apply the P^3M solver there. A way to determine this would be to use the distinction between strongly and weakly coupled plasma.

An important quantity used to characterize the screening of electric potential in a plasma is the Debye length

$$\lambda_D = \sqrt{\frac{\epsilon_0 k_B T}{q^2 n}} \tag{3.1}$$

where ϵ_0 is the vacuum permittivity, k_B is the Boltzmann constant, T is the plasma temperature, and n and q are the density and charge of the plasma particles respectively. We can use this fundamental length to define what we call the plasma parameter

$$\Lambda = \frac{4}{3}\pi n \lambda_D^3 \tag{3.2}$$

This parameter gives us the number of particles contained within the so called Debye sphere. Based on this parameter we say that a plasma is strongly coupled if there are few particles within each Debye sphere, i.e. if $\Lambda \ll 1$, and weakly coupled if $\Lambda \gg 1$ [**Plasma**].

We can use this distinction to determine how to most effectively utilize the P^3M field-solver in our simulations. If we examine how the plasma parameter of our beam develops along the beamline, we can determine where precisely the plasma is strongly coupled, i.e. where intrabeam scattering should contribute most to the rise in energy spread.

3.6 Gitlab Repository

The files used and created during this thesis were organized in the following Gitlab repository: https://gitlab.psi.ch/adelmann/swissfel-de.git. This repository contains the input files used for the simulations laid out in chapter 3 and the analysis notebooks applied in chapter 4.

The Jupyter notebooks used to perform the numerical evaluations are stored in the Analysis directory while the input files and setups for the simulations have been organized in the Overview directory.

Chapter 4

Numerical evaluation

4.1 Higher resolutions

As we increased the grid size from 32^3 to 64^3 and 128^3, we examined how the projected RMS quantities of interest mentioned in section 3.1 were affected. At first we compared simulations with increasingly fine grids but a constant number of particles per grid-point, and noticed that the only substantial change that could be seen was a decrease in transverse emittance. As can be seen in figure 4.1, this change was most pronounced in the jump from a 32^3 grid to a 64^3 grid.

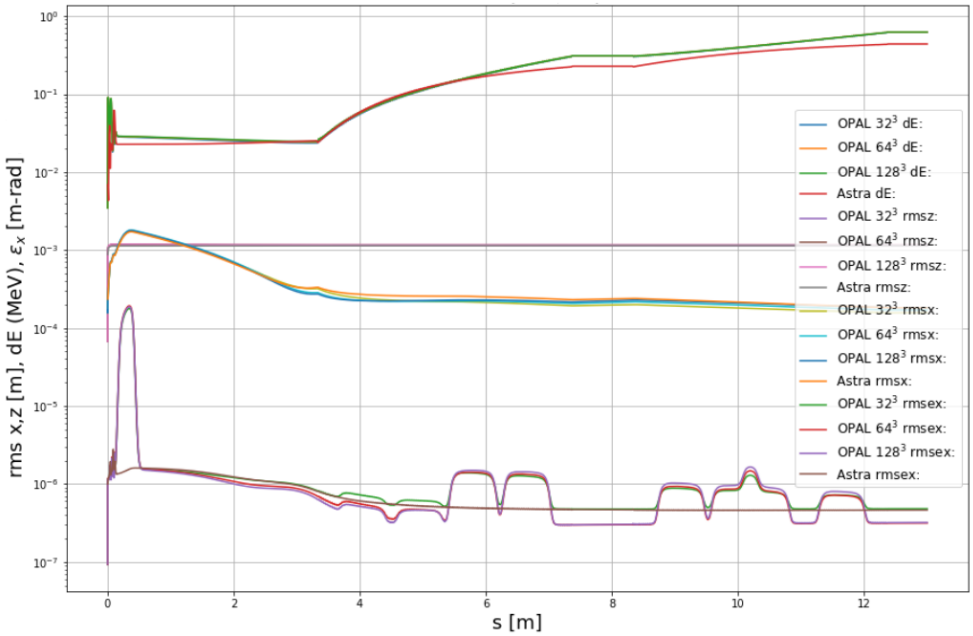

Figure 4.1: The quantities of interest described in figure 3.1, from an Astra simulation and three OPAL simulations, each with different grid-sizes. The OPAL simulations all had 12 particles per grid-point.

We then examined how a change in the number of simulated particles for a constant grid size affected the RMS quantities. We saw that, regardless of the given grid size, changing the number of particles seemed to have no effect whatsoever. This can be seen in figure 4.2, where the plotted output quantities of the different OPAL simulations cannot be distinguished because they lie on top of one another.

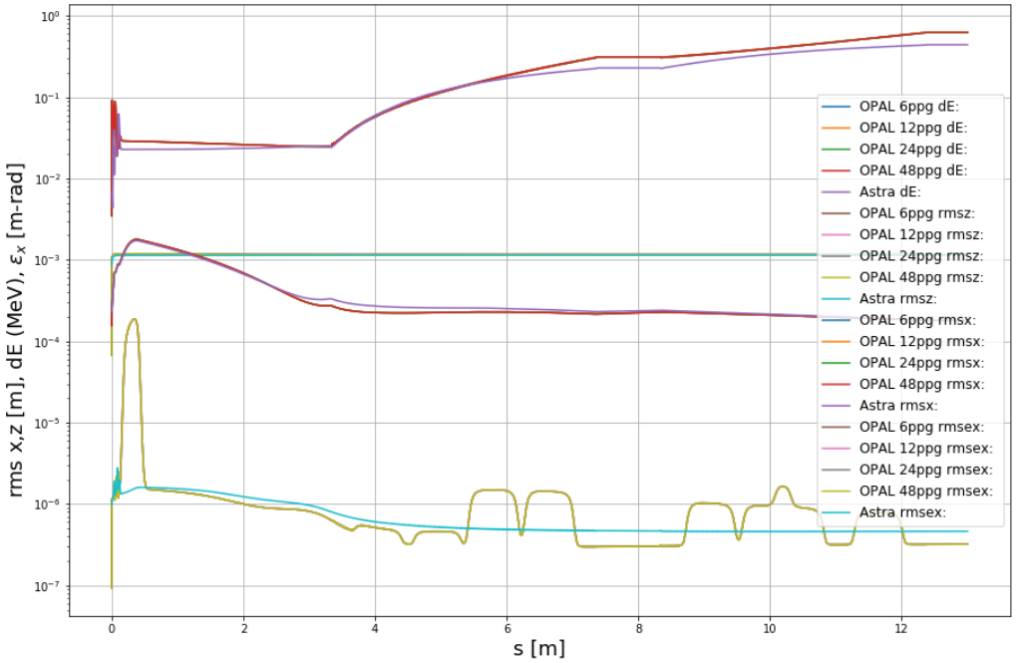

Figure 4.2: The quantities of interest described in Figure 3.1, from an Astra simulation and four OPAL simulations, each with different number of particles per grid-point. The OPAL simulations all had a 64^3 grid.

Wanting to see if these effects were similar for slice values described in sections 2.3.2 and 2.4, we selected a number of points along the injector at which the beam's phase-space data would be dumped, and calculated said slice values there. The points along the injector we selected were $s = \{0.1\,\text{m}, 2.95\,\text{m}, 8.0\,\text{m}, 13.0\,\text{m}\}$.

When looking at the uncorrelated longitudinal energy spread as the beam moves along the injector, the first thing to note is how far off the values are from the projected RMS energy spread. The reason for this is the large slope of the phase-space curve, as was explained in section 2.3. The second thing to note, is that while the projected RMS energy spread rises as the beam moves along the injector, that does not seem to be the case for the uncorrelated slice values, at least not consistently. The energy spread of a beam decreasing as the beam is accelerated is not only strange, it is simply un-physical behaviour [**Meetings**].

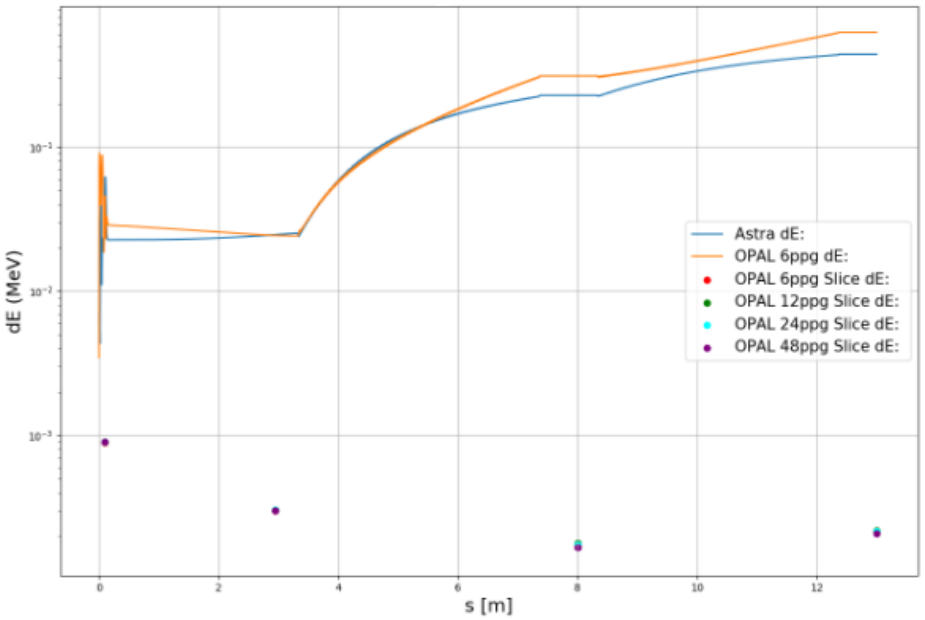

Figure 4.3: Uncorrelated slice energy spread values for four different OPAL simulations, each with a 64^3 grid but different number of particles per grid-point. The solid lines show the projected RMS energy spread of both Astra and OPAL.

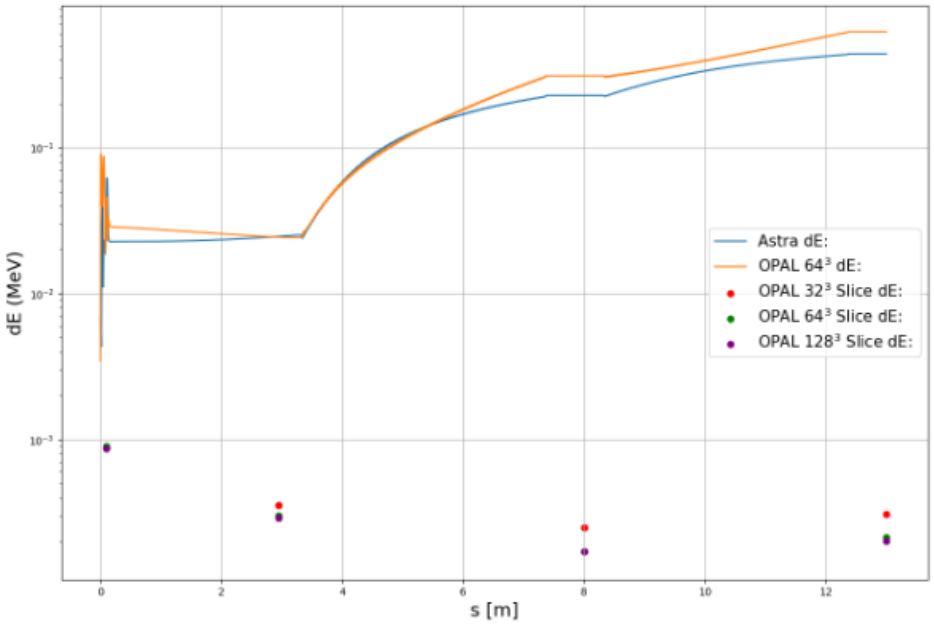

Figure 4.4: Uncorrelated slice energy spread values for three different OPAL simulations, each with 6 particles per grid-point, but different grid sizes. The solid lines show the projected RMS energy spread of both Astra and OPAL.

At $s = 0.1\,\mathrm{m}$ the beam is still inside of the RF gun, and so that data point is not to be taken too seriously. Additionally, while equation 2.1 describes a coasting beam which undergoes neither acceleration nor energy dispersion, we are simulating a more complicated system, so we would not expect the energy spread to behave exactly as described in said equation. However, we would still expect the energy spread to rise as the beam travels downstream and accelerates.

In figures 4.3 and 4.4 we see the projected energy spread of Astra and a single OPAL simulation along the 13 m of the injector, as well as the uncorrelated energy spread of OPAL simulations with different grid sizes and particle numbers per grid-point.

For a given grid size and a rising number of particles per grid-point we see that, like for the RMS value, there is no substantial change in the uncorrelated energy spread. Unlike the projected RMS energy spread though, the uncorrelated values seem to decrease substantially when the grid is refined, as shown in figure 4.4.

A similar behaviour can also be observed when looking at the slice emittance of the beam as it moves along the injector. This is shown in figures 4.5 and 4.6, with the projected emittance of both OPAL and Astra also shown for reference.

We continued to examine the RMS and slice quantities of interest and noticed that where before the slice energy spread had decreased with a rising number of grid-points, the values started significantly increasing as we move beyond the 128^3 grid. This can be seen in figure 4.7 and table 4.1.

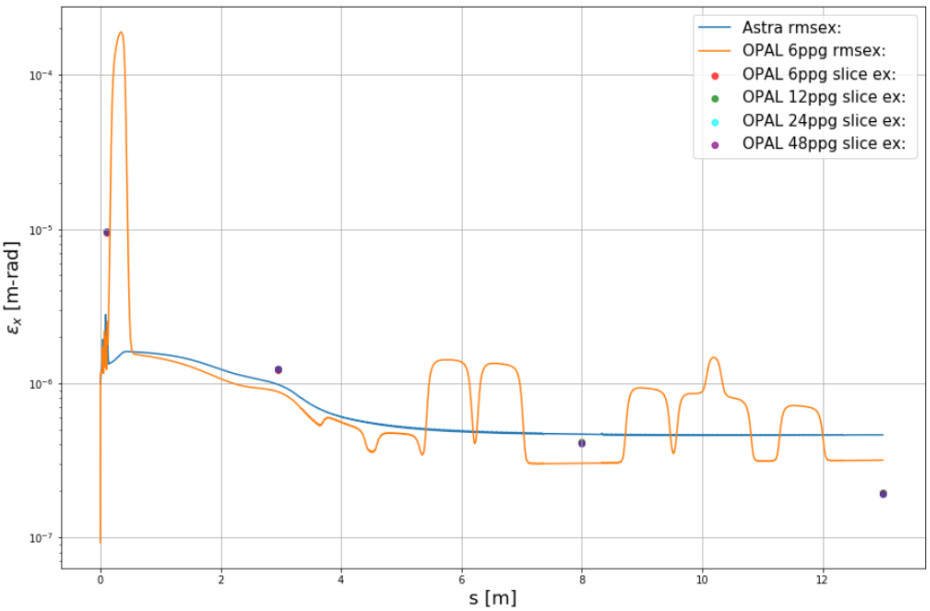

Figure 4.5: Slice emittance values for four different OPAL simulations, each with a 64^3 grid but different number of particles per grid-point. The solid lines show the projected RMS emittance of both Astra and OPAL.

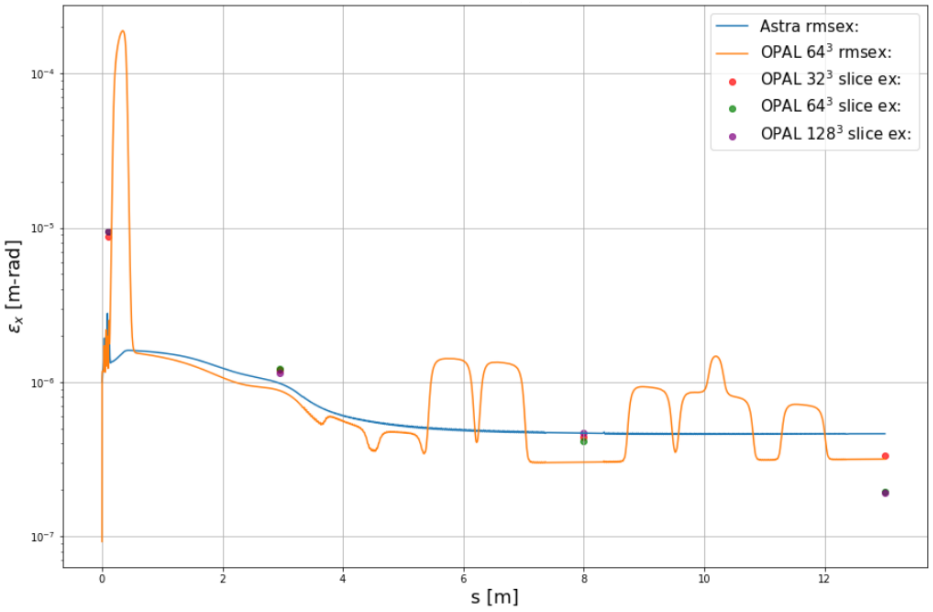

Figure 4.6: Slice emittance values for three different OPAL simulations, each with 6 particles per grid-point, but different grid sizes. The solid lines show the projected RMS emittance of both Astra and OPAL.

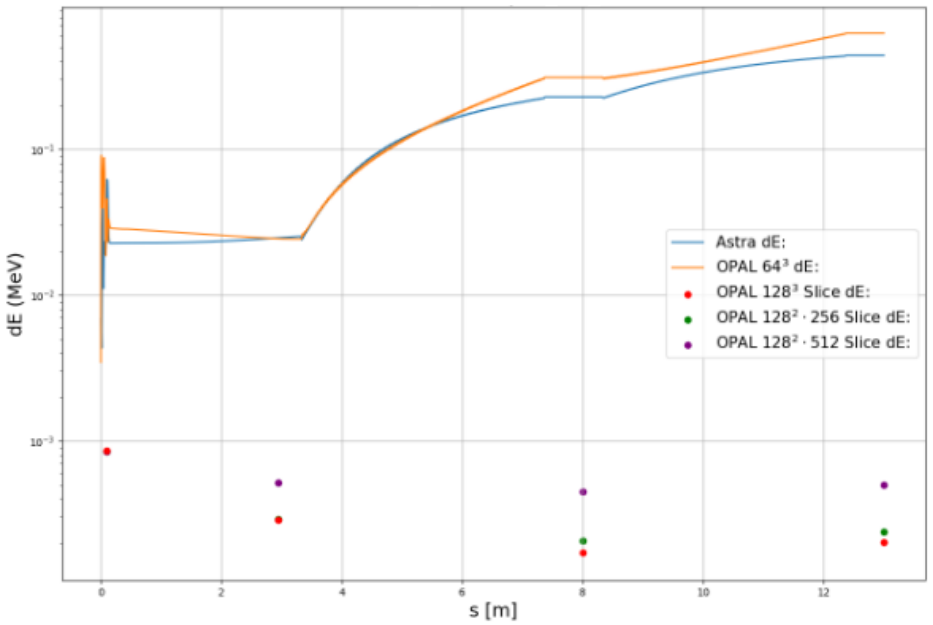

Figure 4.7: Uncorrelated slice energy spread values for three different OPAL simulations, each with 6 particles per grid-point, but different grid sizes. The solid lines show the projected RMS energy spread of both Astra and OPAL.

Table 4.1: The number of particles simulated in OPAL with 6 particles per grid-point and an increasingly refined longitudinal grid. The uncorrelated slice energy spread at $s = 13\,\mathrm{m}$ is seen to rise by over 50% when moving from the 128^3 grid to $128^2 \cdot 1024$.

Grid size	Simulated particles (6 p.p.g.)	dE [keV] ($s = 13\,\mathrm{m}$)
128^3	12,582,912	0.564
$128^2 \cdot 256$	25,165,824	0.568
$128^2 \cdot 512$	50,331,648	0.705
$128^2 \cdot 1024$	100,663,296	0.865
$128^2 \cdot 2048$	201,326,592	0.811

4.1.1 Emission Steps

At first the rising energy spread values obtained by increasing the number of grid-points along the z-axis seemed promising. However, when we examined the actual uncorrelated longitudinal phase-space for these simulations we could see ripples forming in the structures. These ripples were periodic modulations that seemed to rise in amplitude and frequency as we increased the number of longitudinal grid points from 128 to 256 and 512.

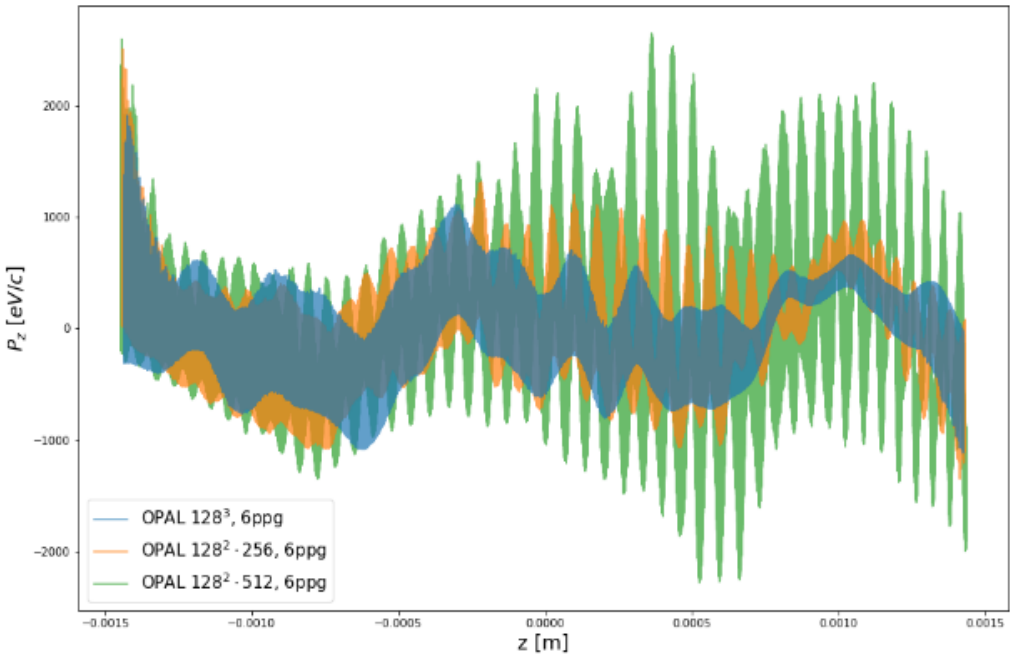

Figure 4.8: The core of the uncorrelated longitudinal phase-space for OPAL beams with grid sizes 128^3, $128^2 \cdot 256$ and $128^2 \cdot 512$, with 6 particles per grid-point each, taken at $s = 13\,\mathrm{m}$.

The question now was whether these ripples were an actual physical effect or simply computational artifacts resulting from the grid sizes. In order to examine this a number of tests were conducted.

At first we aimed to find the frequencies/wavelengths of the ripples and compare them to the longitudinal mesh sizes of the simulations. The idea was that if the two were commensurate, the ripples would obviously be artificial. Here longitudinal mesh size refers simply to the total length of the beam divided by the number of longitudinal grid-points.

Comparing the wavelengths of the ripples that appeared in different simulations to the longitudinal mesh sizes of said simulations showed that they were not commensurate. In fact, as we kept increasing the simulation resolution, the wavelength and amplitude of the ripples seemed to stop increasing after a certain point. This can be seen in table 4.1, where the energy spread significantly increases as the grid size is increased to $128 \cdot 1024$, but then seems to stabilize.

We did, however, find that the wavelength was commensurate with the size of the simulation timestep during the initial emission of the beam. This timestep, which we shall refer to as dt_E, is adjusted in OPAL during emission so that the whole of the beam is emitted in a given number of emission steps. Up until this point we had been running simulations with 100 such steps, but we now decided to adjust that number and see if it had any effect on our data.

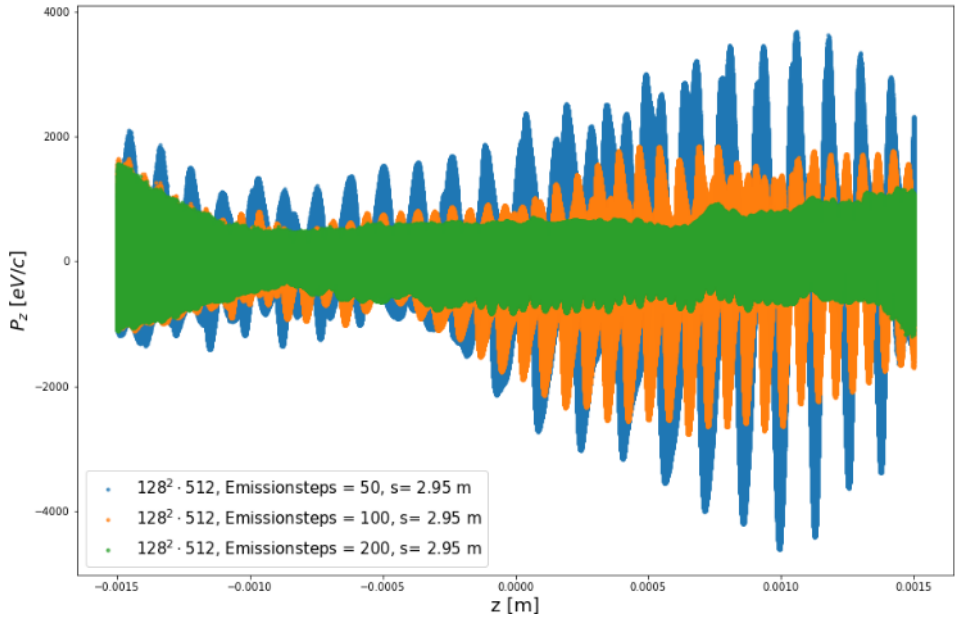

Figure 4.9: The core of the uncorrelated longitudinal phase-space of three OPAL simulations with a varying number of emission steps, taken at $s = 2.95\,\mathrm{m}$.

Figure 4.9 shows that when the number of emission steps are increased, the amplitude and wavelength of the ripples drops. In fact, as we can read from table 4.2, when we double the number of emission steps the frequency seems to just about double as well.

Table 4.2: The ripple wavelength λ and emission timestep dt_E of OPAL simulations with a $128 \cdot 512$ grids, but a varying number of emission steps. We see how λ is commensurate with $dt_E \cdot c$, where c is the relativistic speed of light.

Emission steps	dt_E [ps]	$dt_E \cdot c$ [μm]	λ [μm]
50	0.451603	135.48	130.01
100	0.225802	67.74	71.94
200	0.112901	33.87	39.34
300	0.0752672	16.94	25.56

We see then that these ripples are artifacts caused by a coarseness in the initial emission timestep. As we increase the number of longitudinal grid-points, we get rid of this ripple effect by increasing the number of emission steps as well, until the ripple amplitudes die down and the longitudinal energy spread converges.

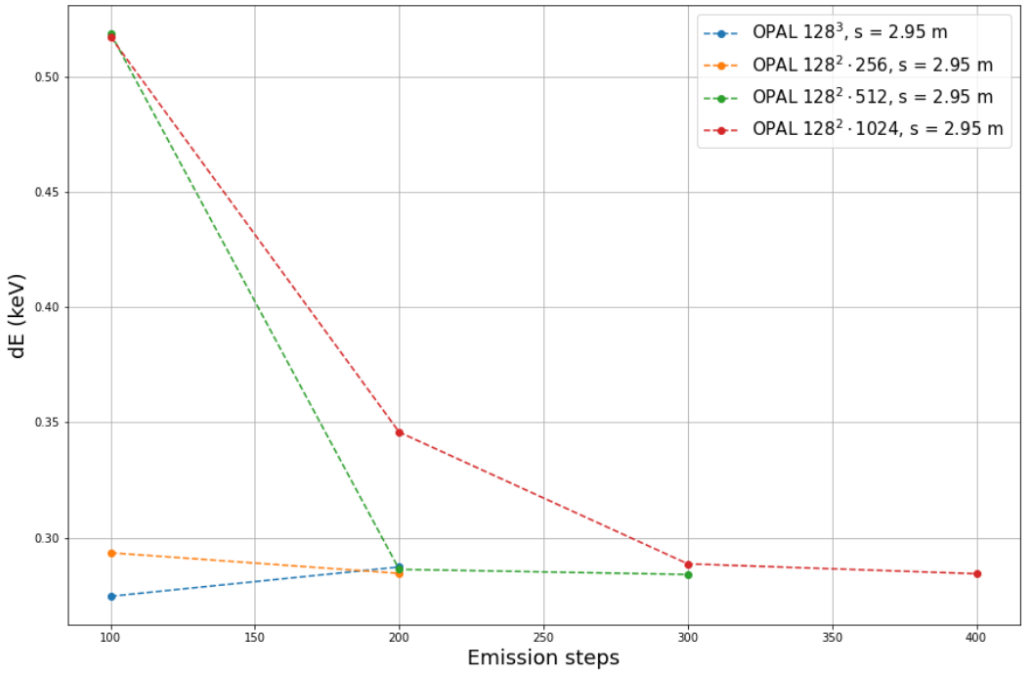

Figure 4.10: The energy spread of OPAL beams at $s = 2.95\,\mathrm{m}$, simulated for different grids converging to about 0.28 keV as the number of emission steps is increased.

Having resolved this issue we are faced with a predicament. The resolution of the simulation has been refined many times over, and we have gotten tantalizingly close to a $1:1$ simulation, yet the slice energy spread remains the same as before. We must conclude that the mesh resolution is not, or at the very least not the only, issue with the previous simulations.

4.2 Booster 2

As we saw in the previous sections 4.1 the uncorrelated energy spread did not rise substantially as the beam travelled along the first 13 m of the machine. This is not in line with what we would expect. Equation 2.1, simplified though it might be, predicts that energy spread will grow in line with the square root of the distance travelled by the beam.

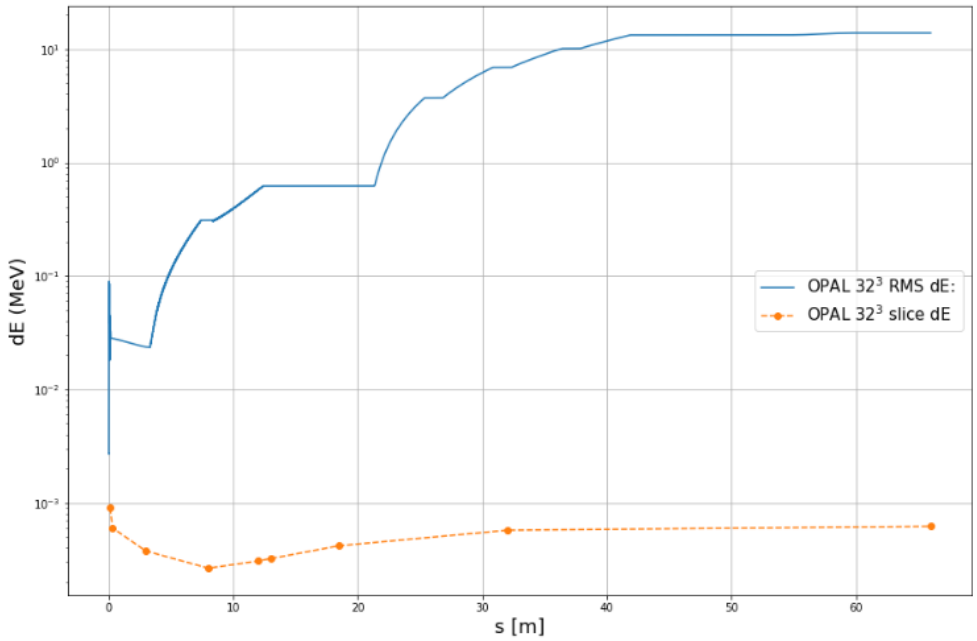

Figure 4.11: Uncorrelated slice energy spread values of an OPAL beam with a 32^3 simulated up to $s = 66$ m. The solid line shows the projected RMS energy spread of the same simulation.

We made preliminary attempts to extend our OPAL simulations to reach further downstream. In figure 4.11 we see that even as we extend our simulation to $s = 66$ m the uncorrelated slice energy spread does not even rise about 0.7 keV. The problem with our current approach seems to be a more fundamental one than either a lack of resolution or the length of our simulations.

4.3 P^{3}M

We simulated a beam of total charge 2 pC, with a 1 : 1 particle ratio, i.e. 12,484,395 simulated particles, using a PM field-solver with 64^3 grid-points. Note that this comes down to about 48 particles per grid-point. We examined how the plasma parameter discussed in section 3.5 behaved as the beam moved downstream.

As depicted in figure 4.12, we saw that the parameter dipped below 1 between $s = 11$ m and $s = 12$ m. We therefore decided to stop the simulation at $s = 11$ m and, using the beam's phase-space data at that point, restart it using the P^{3}M field-solver.

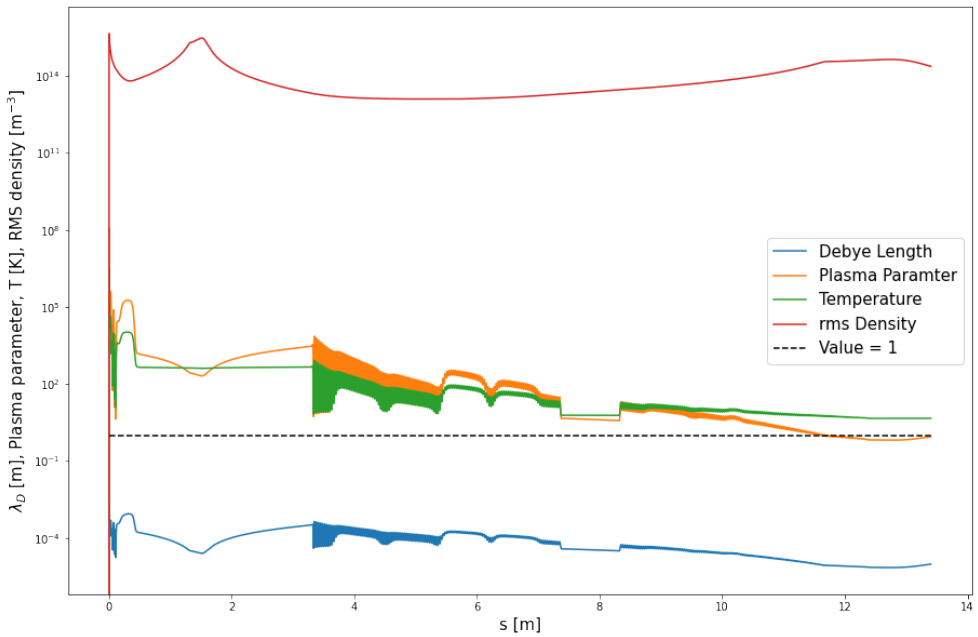

Figure 4.12: The Debye length, plasma parameter, temperature and RMS density of a 2 pC OPAL beam as it moves along the injector.

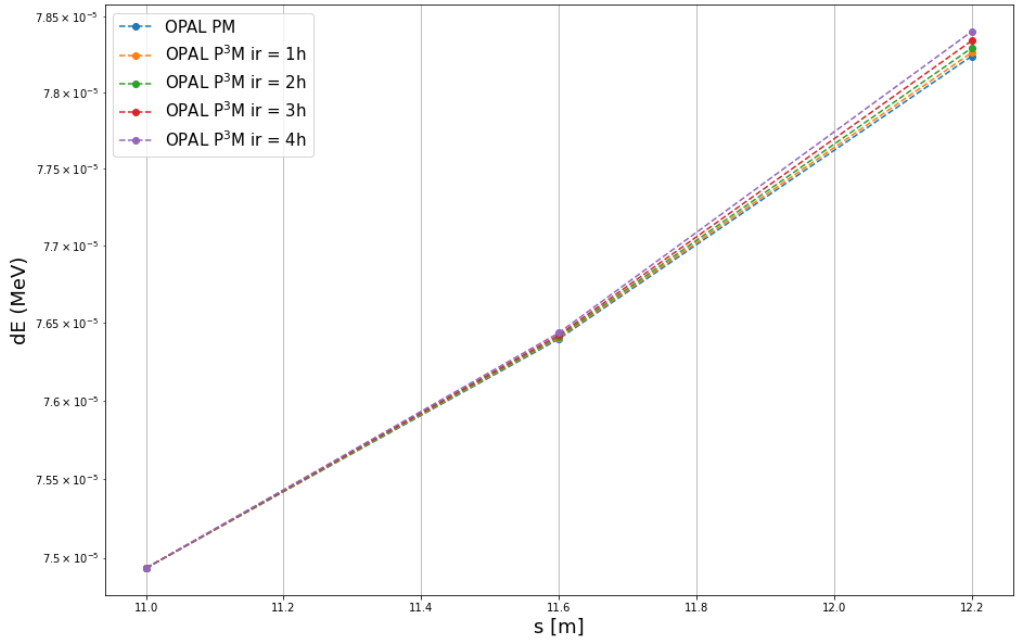

Figure 4.13: The uncorrelated slice energy spread of a 2 pC OPAL beam. The simulation using a PM solver was run up to 12.2 m, but from 11 m on simulations using a P^{3}M solver with varying interaction radii were conducted.

We performed four instances of this restart, varying the size of the P^3M interaction radius. We set the radius the be an integer multiple of the system's transverse grid size h, i.e. the width of the beam at $s = 11\,\mathrm{m}$ divided by the number of transverse grid-points. This was calculated to be $h = 334\,\mu\mathrm{m}/64 = 5.219\,\mu\mathrm{m}$.

Figure 4.13 shows how, from $s = 11\,\mathrm{m}$ to $s = 12.2\,\mathrm{m}$, the rise in energy spread is greater, though only slight, when the P^3M solver is employed. As the interaction radius is increased, so too does this effect seem to increase. The magnitude of this increase is less than 1%, however, this can easily be attributed to the fact that the distance travelled in which we examine it is very small as well.

Despite its low magnitude, the increase in energy spread is clearly present, and its connection with the size of the interaction radius seems clear. The number of pairwise particle interactions simulated is correlated with the gain in energy spread of the beam.

Chapter 5

Results

As described in section 4.1, we increased the resolution of our OPAL simulations in two ways. Firstly, we increased the number of grid-points for our field solver mesh, and secondly, we increased the number of simulated particles per grid point.

As we did this, we noted that the projected RMS energy spread was highly insensitive to both approaches, while the uncorrelated slice energy spread was only unaffected by the change in particle number. As we initially increased the number of grid-points up to 128^3, we saw a decrease in the slice energy spread values at $s = 13\,\mathrm{m}$ of over 30%, though we note that this decrease mostly took place as we moved from the 32^3 grid to the 64^3 one.

We noted a similar behaviour in the transverse emittance output. In the case of the emittance, changing the number of simulated particles alone had no apparent effect on the output, whilst when we changed the grid size, both projected RMS values as well as slice values decreased by about 40%. Again, this change was mostly confined to the change from a 32^3 grid to a 64^3 one.

When the grid sizes are increased further, specifically, when the number of grid-points along the longitudinal dimension of the beam are increased, we see that both the slice energy spread and the slice emittance increase. While we only saw an increase of about 7% in the slice emittance, the uncorrelated slice energy spread at $s = 13\,\mathrm{m}$ increased by over 55% as the simulation grid was changed from 128^3 to $128^2 \cdot 512$.

We discovered that this increase was, however, a numerical effect caused by a coarseness in the simulation's timestep during the initial emission of the beam, dt_E. As the longitudinal mesh size approached the length $dt_E \cdot c$, periodic micro-structures started forming in the beam's longitudinal phase-space. The growing amplitude of these ripples cause our calculations to return increasing values for the slice energy spread. For these grids the issue was resolved by increasing the number of timesteps taken during the emission of the beam. This increase in emission steps caused the wavelengths and amplitudes of the periodic ripples to decrease sharply, correcting our energy spread predictions.

As our output still showed uncorrelated energy spread below 1 keV, we extended the simulations to encompass the first 66 m of the SwissFEL injector. When we looked at the energy spread of the beam as it travelled downstream, we noted that even at 66 m the slice value was still less than 1 keV. We concluded that the issue was not that of low

resolution in the PIC mesh, but the particle-mesh solver itself.

When we simulated a 2 pC beam and applied the P^3M field-solver in the section where we determined the plasma to be strongly coupled, we saw that in doing so the longitudinal energy spread increased. This increase was, admittedly, very small, but that can be explained by the fact that the simulated distance over which the solver was applied was also very small. On a fundamental level, this shows that properly calculating the pairwise interactions of neighboring electrons, as opposed to only looking at their interaction with the mesh, does contribute to a rise in energy spread.

Chapter 6

Conclusions

We have shown that OPAL can be reliably used to simulate the SwissFEL injector, and that achieving very high resolutions in said simulations is feasible. We have implemented methods to calculate uncorrelated slice energy spread values as well as slice emittance values from the OPAL outputs, which had not been done before.

In addition, we have shown the effects that increasing the simulation resolution has on the primary output quantities of interest, and specifically, we have shown that this alone is not enough to account for the discrepancy between the previous numerical predictions and measurements.

We demonstrated that neither applying a high resolution to the PM mesh nor simulating the beam over long distances returned energy spread values close to that of the measurements. After making very computationally costly simulations with the P^3M field-solver though, we noted a clear, if small, rise in the energy spread. We therefore conclude that the discrepancy between the measurements and previous simulations is due to the lack of properly simulated intrabeam scattering of electrons provided by a particle-mesh solver.

Declaration of authorship

I, Garðar Árni Skarphéðinsson, hereby declare that this thesis titled "On understanding the differences in measured and calculated energy spread in the SwissFEL injector", and the work presented in it, are my own. I confirm that this work was done solely while in candidature for a masters degree at ETH Zuerich.

Except where reference is made in the text of the thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis accepted for the award of any other degree or diploma. No other person's work has been used without due acknowledgment in the main text of the thesis. This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution.

Garðar Árni Skarphéðinsson 28. September, 2022