Optimizing Simulation Times of SPS Slow Extraction using MAD-X

Philipp M. Schicho

ETH Zurich
Raemi Street 101, 8092 Zurich, Switzerland

Advisor: Matthew A. Fraser
Co-Advisor: Francesco M. Velotti
CERN, Geneva, Switzerland

July 31, 2015

Abstract

The simulation time of a MAD-X script employing particle tracking to model the slow-extraction process of 400 GeV protons from the SPS was studied by batching the code on the CERN computing batch service. The behaviour of the code for large numbers of particles and turns was examined and batching significantly reduced simulation times to $\mathcal{O}(1000\; s)$. Convergence studies showed that simulated quantities that are time-averaged across the spill, e.g. the rms extracted beam emittance, can be reliably simulated with more than 400 particles and in a simulation time of less than an hour.

1 Introduction

Resonant extraction yields a steady flux of particles from synchrotrons over timescales of a few milliseconds to several seconds. The third-integer resonance draws out particles over a large number of turns ($N_{\text{tu}} \sim 10^6$) equivalent to extraction times (spill times) of about one second in the SPS [4]. The circulating particles are driven unstable along the arms of the non-linear phase space separatrix when the betatron tune in the extraction plane respects the condition $3Q_h = n$ with $n \in \mathbb{Z}$ for the resonance excited by sextupole magnets. The tune spread induced by chromaticity in the machine is exploited in the SPS to extract the beam smoothly during the spill by changing the strength of the quadrupoles and thus the tune. The strength of the extraction sextupoles are held constant during the spill. As the tune is swept, particles of different momenta are extracted as a function of time as those with the resonant tune are excited. As the amplitude of the beam grows, particles reach and cross the thin electrostatic septum (ZS) into a region of high electric field, which deflects the beam into the extraction channel and to the experiments. A detailed description of the slow-extraction process that is employed at the SPS can be found in [3].

2 Analysis and Results

This chapter focuses on the practical details of its simulation regarding CPU time and convergence of the extracted beam emittance.

2.1 Simulation Time Optimization

An existing MAD-X [2] script describing the SPS was wrapped in python to batch the simulations from lxplus.cern.ch to the CERN batch computing service. The particle number was swept in
the range \( N_p = [10, 1000] \) while keeping the number of turns \( N_{tu} = [1000, 200000] \) fixed; 200000 turns in the SPS corresponds to a spill length of over 4 seconds. In order to control the simulation speed the total number of processes \( N_{proc} = 7 \) was held fixed during the batch submission (\texttt{bsub}). Statistical fluctuations in the simulation time is thought to stem from the different computing resources allocated to the job at the time of submission onto the batch service. The fluctuations were accounted for by sufficiently large statistics.

As a first inspection we included extraction in the simulation with the extraction sextupoles set to their nominal normalised value of \( \text{extr}_{\text{sext}} = -0.11992 \text{ m}^{-3} \) with the aperture model included. As a result, the number of extracted particles is the defining parameter determining the total simulation time. In other words, the more particles that get extracted, the more the simulation gathers speed. To this end, we only studied the simulation time in the case that the extraction sextupoles were turned off. We observe a distinct linearity in simulation time as a function of the number of input particles with a spread of \( \text{std} = \mathcal{O}(0.1) \) for \( N_p > 20 \), as presented in Figure 1. The variation in the simulation time for each data series appears random and dependent on the resources available on the computing cluster when the job is submitted. The simulation time as a function of the number of turns is also closely linear, with a steady time measured per turn as shown in Figure 2. The heatmap in Figure 3 helps to illustrate the linearity in the simulation time for both \( N_p \) and \( N_{tu} \).

As the simulation time is linear in the number of turns, it makes sense to simulate the entire extraction process by splitting the input particle ensemble into small batches processed independently and in parallel on different CPUs. This is correct under the assumption that particles do not interact with one another. In order to minimise the simulation time one would split up the particle ensemble into as many jobs as possible. In practice, 100 jobs was feasible such that an average of one turn per job could be handled on the cluster when the job is submitted. The 

\[ \langle \cdots \rangle \]

heating in Figure 3 helps to illustrate the linearity in the simulation time for both \( N_p \) and \( N_{tu} \).

The phase space area extracted at the ZS is given by

\[ A = \pi \varepsilon \]

where for the correlation between \( x \) and \( x' \) we proceed similarly,

\[ \overline{xx'} = \langle xx' \rangle = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} (x - \overline{x})(x' - \overline{x}) \]

The phase space area extracted at the ZS is given by \( A = \pi \varepsilon \). A typical scatter plot of the extracted particles at the septum is shown in phase space \((x, x')\) in Figure 4. The distance between the circulating beam on its extraction bump and the ZS constrains the beam size in the extraction channel (\textit{spiral-step}) and was tuned to 90% of the nominal.

The following simulations were batched on the computer cluster. We used particle numbers of \( N_p = [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000] \) and split the processes by a maximum particle number of \( N_{batch} = 10 \). Furthermore, we swept through different numbers of turns, i.e. \( N_{tu} = [100000, 200000] \), to observe how the emittance evolves through the spill at \( t = 2 \) and 4 s. The initial transverse distribution at \( t = \text{top} \) was fixed as a Gaussian with an energy of 400 GeV, normalised rms emittances of \( \varepsilon_h = 12 \text{ mm mrad} \) and \( \varepsilon_v = 8 \text{ mm mrad} \). The momentum spread was distributed uniformly between \( \frac{\Delta p}{p} = \pm 4 \varepsilon \).

The results of the convergence study are illustrated in Figure 5 for both the horizontal and vertical emittances with the aperture model suppressed. The rms emittance of the extracted beam
Figure 1: Simulation time as a function of $N_p$ for SPS slow extraction with extraction sextupoles turned off. From top to bottom as functions of $N_p$ respectively (a) CPU time, (b) CPU time per turn (c) normalised variation (rms) of CPU time.
Figure 2: Simulation time as a function of $N_{tu}$ for SPS slow extraction with extraction sextupoles turned off.

Figure 3: Heatmap of the simulation time as a function of $N_{tu}$ and $N_p$ for sextupoles turned off.
Figure 4: Phase-space \((x, x')\) of particles lost at the ZS septum with \(N_p = 1000\) for an extraction bump amplitude scaled by 90% of the nominal.

Table 1: Converged values of rms emittance \(\varepsilon_{\text{rms}}\) for \(N_{tu} = 200000\) in both \(x\)- and \(y\)-directions.

<table>
<thead>
<tr>
<th>(\varepsilon_{\text{rms}})</th>
<th>input [(\mu\text{m rad})]</th>
<th>extracted [(\mu\text{m rad})]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x)</td>
<td>12</td>
<td>15.8</td>
</tr>
<tr>
<td>(y)</td>
<td>8</td>
<td>9.8</td>
</tr>
</tbody>
</table>

converges above \(N_p \approx 400\) for \(N_{tu} = 200000\). The vertical emittance is seen to grow during the extraction due to filamentation caused by the non-linearity of the sextupole fields during extraction. The horizontal emittance is determined by the machine parameters, e.g. the amplitude of the extraction bump. We observe a horizontal beam size of approximately 15 mm at the septum. Looking closer at the spill in the \(x\) direction one measures a rather sharp cut at approximately \(x_s = 82.5\) mm. The distribution itself exhibits a plateau region around \(x_s = [76.0 - 80.0]\) mm, which becomes more distinct with higher number of particles extracted. The shape resembles the tail of a sharply peaked Cauchy distribution as shown in Figure 6 and compared to measured data found in [3]. The rms emittance values at input and extraction are compared in Table 1.
Figure 5: Convergence studies of the slow extracted rms emittance in both $x$- and $y$ planes for an extraction bump amplitude tuned to 90% of the nominal.

### 3 Conclusion

The study illustrated the feasibility of simulating slow-extraction from the SPS within simulation times of $\mathcal{O}(1000 \text{ s})$ by batching MAD-X on to the CERN computer cluster with approximately 2% of processes crashing. MAD-X was shown to behave linearly with both the number of particle and number of turns simulated. Convergence studies showed that simulated quantities that are time-averaged across the spill, e.g. the rms extracted beam emittance, can be reliably simulated with more than 400 particles and in a simulation time of less than an hour. This lends the code for use in optimisation studies of the slow-extraction process. The main drawback of using the cluster was found to be the unpredictable simulation time.
Figure 6: Horizontal beam profile of particles extracted at the septum (ZS) with bump amplitude 90\% for (a) simulation and (b) measurement presented in [3].
References


