ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-WEPAB244

OPTIMIZATION AND MACHINE LEARNING APPLIED TO THE RF MANIPULATIONS OF PROTON BEAMS IN THE CERN PS

A. Lasheen*, H. Damerau, S. Johnston, CERN, Geneva, Switzerland

Abstract

The 25 ns bunch spacing in the LHC is defined by a sequence of RF manipulations in the Proton Synchrotron (PS). Multiple RF systems covering a large range of revolution harmonics (7 to 21, 42, 84, 168) allow to perform RF manipulations such as beam splitting, and non-adiabatic bunch shortening. For the nominal beam sent to LHC, each bunch is split in 12 in the PS. The relative amplitude and phase settings of the RF systems need to be precisely adjusted to minimize the bunch-by-bunch variations in intensity, longitudinal emittance and bunch shape. However, due to transient beam-loading, the ideal settings, as well as the best achievable beam quality, vary with beam intensity. Slow drifts of the hardware may also affect beam quality. In this paper, automatized optimization routines based on particle simulations with intensity effects are presented, together with first considerations of machine learning. The optimization routines are used to assess the best achievable longitudinal beam quality expected with the PS RF systems upgrades, in the framework of the LHC Injector Upgrade project.

INTRODUCTION

The longitudinal beam structure for the LHC beams is defined by a sequence of RF manipulations in the PS to bring the bunch distance down to 25 ns and a target longitudinal emittance of $\varepsilon_l = 0.35$ eVs. The RF settings (voltage, phase, timings) for each manipulation must be carefully adjusted in order for all bunches to have identical intensity and emittance. Multiple schemes are available in the PS following the same skeleton [1]. A selected example presented on Fig. 1 is the Batch Compression Merging and Splitting (BCMS) beam, where each injected bunch is effectively split in six. For high intensity beams, the settings must be adapted to account for beam loading effects. Since the PS is only partially filled, transient beam loading modulates the RF amplitude and phase along the batch. A spread in the bunches parameters can therefore not be avoided. In the framework of the LHC Injector Upgrade (LIU) project [2] the beam intensity in the PS is doubled, with constant longitudinal emittance. Moreover, the bunch-by-bunch variation of intensity and bunch length at extraction should remain below $\pm 10\%$. The first motivation being to avoid too large fluctuations of the luminosity in the LHC, the second is to ensure a good transmission of the beam to the Super Proton Synchrotron (SPS), the accelerator downstream of the PS.

The longitudinal beam quality needs to be regularly checked before filling the LHC. Adjustments are presently done manually in operation allowing to reach acceptable beam parameters. Nonetheless, it does not ensure the best achievable parameters for each LHC fill. To approach this level of performance, optimization algorithms presented in this paper were tested and in simulations. Additionally, in order to further improve the usability in operation, machine learning techniques were evaluated for the first time in the PS to better adjust RF manipulations.

OPTIMIZATION OF BUNCH SPLITTINGS WITH BEAM LOADING EFFECTS

The first step of the optimization consists of defining the observables. The quality of the splittings is presently evaluated at extraction by a Fourier analysis technique of the

* alexandre.lasheen@cern.ch

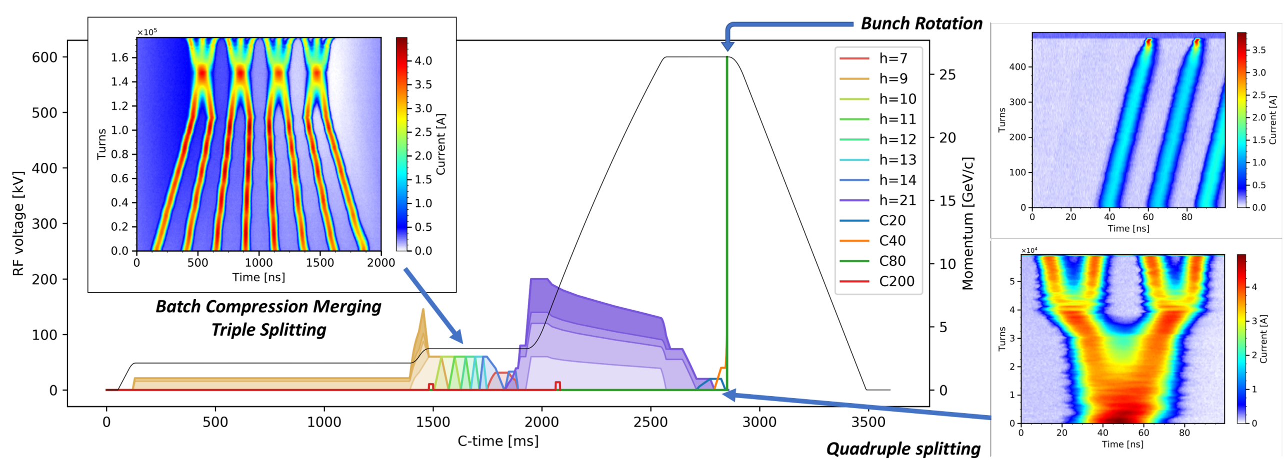

Figure 1: The momentum (black) and RF program (colored) for the BCMS cycle. The RF harmonics $h = 7$ to 21 (10 MHz) are handled by tunable ferrite loaded cavities while the other harmonics $h = 42$ (20 MHz), $h = 84$ (40 MHz) and $h = 168$ (80 MHz) are generated by fixed frequency RF systems. For each manipulation the evolution of the bunch profiles is shown.

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2021). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-WEPAB244
Content from this work may be used under the terms of the CC BY 3.0 licence (© 2021). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI

beam profile of the last turn [3]. The algorithm consists of applying a Discrete Fourier Transform (DFT) to an array containing the bunches amplitude, length, or intensity. The spectrum of the bunch-by-bunch variation is noted $\tilde{\lambda}$. Examples are shown on Fig. 2.

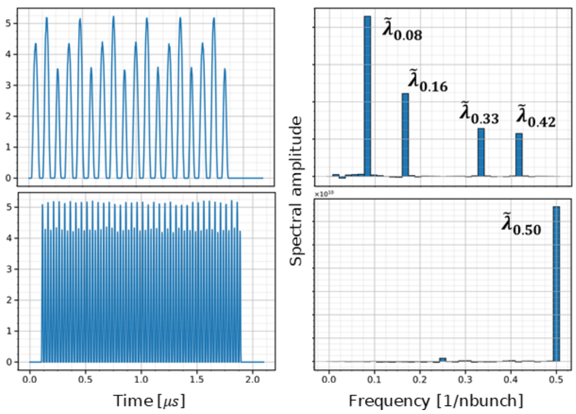

Figure 2: Effect of RF phase errors during the triple (top, $h = 7 \rightarrow 21$) and last (bottom, $h = 42 \rightarrow 84$) double splittings. The beam current (left) is analyzed with DFT (right) to identify the phase errors.

The measured peaks in $\tilde{\lambda}$ can then be used to define the objective function to minimize. Various algorithms such as Nelder-Mead simplex algorithm [4] or Powell's methods [5] were tested. These methods are applicable for single objective optimization and are thus well suited to adjust the splittings. Indeed, each peak of $\tilde{\lambda}$ can be attributed to one specific splitting which can therefore be treated in parallel. The objective functions are defined as

$$\sum_{m\in[2,4]} \tilde{\lambda}^2_{1/m} = f\left(\phi_{h42,\mathrm{err}}, \phi_{h84,\mathrm{err}}\right), \qquad (1)$$

$$\sum_{n\in[1,2,4,5]} \tilde{\lambda}^2_{n/12} = g\left(\phi_{h14,\mathrm{err}}, \phi_{h21,\mathrm{err}}, V_{h14,\mathrm{err}}\right), \qquad (2)$$

where the left hand side is the objective function to minimize obtained from the observable $\tilde{\lambda}$ (m and n are indexes to select the relevant $\tilde{\lambda}$ peaks), and the right hand side the parameters to adjust (errors in phase ϕ_{err} and voltage V_{err}, from settings without beam loading effect). The first objective is used to adjust both double splittings at the same time (faster than two separate optimizations), the second objective is employed to independently adjust the triple splitting.

To evaluate the efficiency of the optimization, a simulated environment was built using the BLonD code [6], with the latest PS impedance model [7]. The convergence of the optimization was satisfactory. The topology of the objective functions f and g was evaluated by doing a parameter scan. In Fig. 3, simultaneous adjustment of both double splittings is shown to be straightforward as the amplitude of the spectral line is linear with the phase errors at 20 MHz and 40 MHz. The algorithms converge, despite the large initial steps leading to search far from the optimum. Note that the objective is not centered at zero phase error, as beam loading effects result in an RF phase shift along the batch.

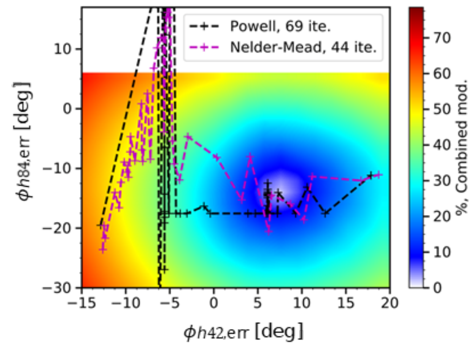

Figure 3: Scan of parameters to obtain the f objective function to minimize bunch intensity spread during the double splittings (colors) with the path of the optimization algorithms (black, purple), including beam loading.

The triple splitting case is more complex as illustrated in Fig. 4. There, another important parameter to adjust is the amplitude of the RF voltage at $h = 14$ to have enough beam intensity on the central bunch. Adjusting all three parameters lead to a non-monotonic behavior of the objective function. Additionally, the topology of the objective function is different whether the bunch length or intensity spread is chosen for minimization. This explains the experienced difficulty to optimize settings manually in operation. The intensity spread is usually minimized, but can lead to a rather large spread in bunch lengths.

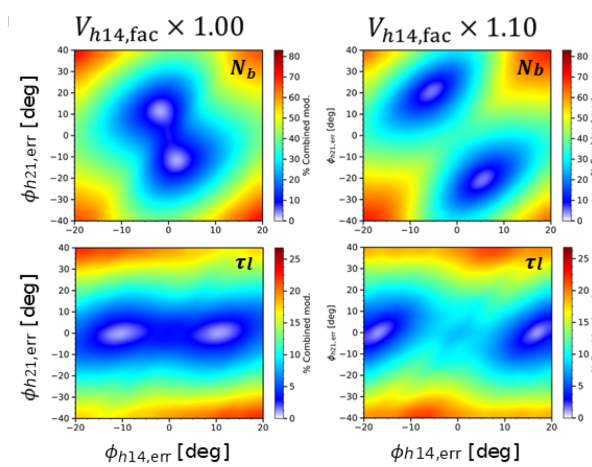

Figure 4: Map of the objective function minimizing $\sum \tilde{\lambda}^2$ for the triple splitting, without beam loading. The target spread to reduce is intensity (top) or length (bottom), for the nominal voltage at $h = 14$ (left) or increased by 10% (right).

With the LIU RF upgrades [8,9], the obtained spread after minimization of the bunch intensity and lengths spread in simulation remains below the $\pm 10\%$ budget, except for a few bunches at the head of the batch due to strong transient beam loading effects. Further work is required to obtain an efficient optimization of both bunch intensity and length spread at once. A second issue is that the number of iterations needed (tens) is unacceptably large to optimize each LHC fill.

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-WEPAB244

COMPUTER VISION APPLIED TO BUNCH ROTATION

In order to improve the efficiency of the optimization, machine learning techniques were also explored aiming at extracting further observables for the bunch shortening prior to extraction. The bunch is rotated in the longitudinal phase space by applying a non-adiabatic step in voltage and is extracted when it is the shortest. In the PS, the operation is done in two steps by pulsing two RF cavities tuned at 40 MHz, then by two RF cavities at 80 MHz (300 kV per cavity, see end of cycle in Fig. 1).

An important parameter to adjust is the relative phase between 40 MHz and 80 MHz voltages to have a symmetric bunch rotation and avoid losses downstream due to uncaptured particles in the SPS buckets [10]. This is presently achieved by evaluating the tilt of a single bunch profile, which limits the accuracy. Other methods like longitudinal beam tomography were also demonstrated to be sub-optimal [11].

A new approach was hence to train a Convolutional Neural Network (CNN) to quantify the phase error from several profiles, following the same methods applied in computer vision. The CNN is displayed in Fig. 5. It is based on the LeNet architecture [12], using rectified linear units for activation functions. The first stage consists of extracting features with two convolution layers, while the second connects all the resulting nodes into an output quantifying the phase error.

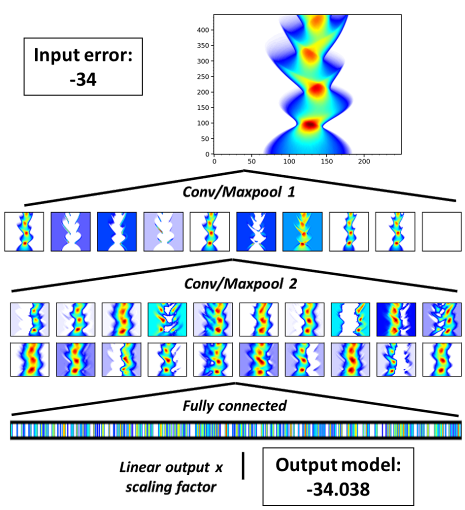

Figure 5: Convolutional Neural Network trained to identify relative phase errors between 40 MHz and 80 MHz voltages during bunch rotation from bunch profiles.

To train the model, a simulated dataset was produced with BLonD by applying a phase error in the range $[-80, +80]$ degrees and recording the bunch profiles. Instead of stopping the simulation when the bunch is the shortest, like in normal operation, the bunch is left to oscillate for a few synchrotron periods to have enough samples for the CNN to characterize the phase error. Then, 80% of the dataset is used for training and the remaining 20% for validation of the model. In this configuration, the model is successfully trained as seen for the example in Fig. 5, with an accuracy of the error detection below one degree.

The second step was to consider using the same model, but under the normal operational conditions where the bunch is extracted when it is the shortest, at the vertical trace index ≈ 100 (Fig. 5). Unfortunately, the accuracy of the model is greatly reduced with the missing data. Moreover, training the model only with data where the bunch is extracted does not allow to reach good convergence. The strategy applied to circumvent the issue was to combine both datasets, with and without extraction, equivalent to data augmentation. The result is shown on Fig. 6, bringing the accuracy of the model output back to the order of a degree.

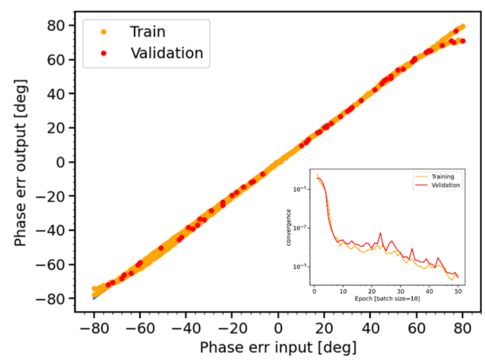

Figure 6: The phase of the 80 MHz voltage as obtained from the CNN compared to the expected phase error. The convergence (inlay plot) is displayed for both training and validation data with no sign of overfitting.

CONCLUSIONS

Optimizers and machine learning techniques were applied in simulations in view of better adjusting the PS RF manipulations in operation. The optimization with established algorithms was shown to converge to the optimal solution, even in presence of transient beam loading. Nonetheless, the number of iterations remains large and techniques like reinforcement learning to reduce the number of iterations are considered in the future. Computer vision was tested for the bunch rotation to extract features from the bunch profiles, and could be generalized to improve the observables as input for optimization. However, these methods were studied with smooth simulated bunch profiles and will be applied to measured ones. Eventually, CNN and reinforcement learning could be combined to maximize the efficiency of the optimization to systematically reach the best achievable beam parameters within a few cycles.

ACKNOWLEDGMENTS

The authors would like to thank S. Hancock and R. Maillet for introduction on the Fourier harmonics analysis, as well as S. Hirlaender, G. Trad and T. Argyropoulos for fruitful discussions on reinforcement learning and computer vision.

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2021). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-WEPAB244

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2021). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI

REFERENCES

[1] H. Damerau *et al.*, "LIU: exploring alternative ideas", in *Proc. RLIUP: Review of LHC and Injector Upgrade Plans*, Achamps, France, Oct. 2013, pp. 127–137. doi:10.5170/CERN-2014-006.127

[2] "LHC Injectors Upgrade, Technical Design Report", CERN, Geneva, Switzerland, Rep. CERN-ACC-2014-0337, 2017.

[3] S. Hancock, "LHC Beams in the PS: Reliability and Reproducibility Issues", in *Workshop on LHC performance*, Chamonix, France, Mar. 2003, paper AB-2003-008, pp. 38-40.

[4] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization", *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965. doi:10.1093/comjnl/7.4.308

[5] M. J. D. Powell, "Direct search algorithms for optimization calculations", *Acta Numerica*, vol. 7, pp. 287–336, Jan. 1998. doi:10.1017/s0962492900002841

[6] CERN Beam Longitudinal Dynamics simulator BLonD, https://blond.web.cern.ch

[7] A. Lasheen, "PS Longitudinal Impedance Model", Zenodo, Mar. 2020. doi:10.5281/zenodo.4722835

[8] G. Favia *et al.*, "The PS 10 MHz High Level RF System Upgrade", in *Proc. 7th Int. Particle Accelerator Conf. (IPAC'16)*, Busan, Korea, May 2016, pp. 622-625. doi:10.18429/JACoW-IPAC2016-MOPOR013

[9] F. Bertin, Y. Brischetto, H. Damerau, A. Jibar, and D. Perrelet, "Impedance reduction of the High-frequency Cavities in the CERN PS by Multi-harmonic Feedback", presented at the Low-Level Radio Frequency Workshop, Chicago, IL, USA, Set.-Oct. 2019. arXiv:1910.06874

[10] H. Timko *et al.*, "Longitudinal transfer of rotated bunches in the cern injectors", *Physical Review Special Topics - Accelerators and Beams*, vol. 16, p. 051004, May 2013. doi:10.1103/physrevstab.16.051004

[11] S. C. Johnston, "Tomography of Rotated Bunches at PS Extraction", CERN, Geneva, Switzerland, Rep. CERN-STUDENTS-Note-2020-032, Oct. 2020.

[12] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition", *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998. doi:10.1109/5.726791