ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-WEPOMS035



# HARPY: A FAST, SIMPLE AND ACCURATE HARMONIC ANALYSIS WITH ERROR PROPAGATION

L. Malina, DESY, Hamburg, Germany

*Abstract*

Traditionally, in the accelerator physics field, accurate harmonic analysis has been performed by iteratively interpolating the result of Fast Fourier Transform (FFT) in the frequency domain. Such an approach becomes computationally demanding when relatively small effects are being studied, which is especially evident in the typical example of harmonic analysis of turn-by-turn beam position monitor data, i.e. many correlated but noisy signals. A new harmonic analysis algorithm, called Harpy, is about an order of magnitude faster than other methods, while often being also more accurate. Harpy combines standard techniques such as zero-padded FFT and noise-cleaning based on singular value decomposition. This combination also allows estimating errors of phases and amplitudes of beam-related harmonics calculated from cleaned data.

## INTRODUCTION

In accelerator physics, accurate harmonic analysis is one of the critical numerical methods. Nowadays, the tighter tolerances shift the attention from the determination of frequencies, such as betatron or synchrotron frequency, to the measurement of phases of even smaller spectral lines. A good example and this paper's motivation is the beam optics measurement in storage rings. One of the ways to measure the beam optics in a storage ring is to analyse beam position monitor (BPM) orbit readings of coherently-excited beams recorded turn-by-turn (TBT) [1]. The calculations of the actual optical functions (for example, phase advances, $\beta$-functions and resonance driving terms) are the last steps in the analyses, which require frequencies, amplitudes and phases of the different spectral lines, commonly referred to as frequency spectra.

Traditionally, in the analysis process, TBT BPM data is first cleaned of noise using methods [2–4] based on Singular Value Decomposition (SVD). Then, frequency spectra of cleaned TBT data are computed for every BPM independently employing methods [5,6] based on frequency interpolation in the output of the Fast Fourier Transform (FFT). The computation is an iterative process, and in each iteration, the strongest signal in the FFT output is interpolated (i.e. found) and subtracted after Gram-Schmidt orthonormalisation. With typically hundreds of iterations the frequency analysis is computationally expensive. The phase accuracy of frequency-interpolation based methods [5,6] was found to be worse than FFT [7].

We presented an intermediate algorithm [8], which calculated the frequency spectra of the data decomposed by SVD and recomposed the BPM spectra in the frequency domain. As a result, it was more than an order of magnitude faster, which is essential for efficient beam operation, for example, in automatic coupling correction [9]. However, the accuracy issue remained unresolved for weak spectral lines.

The work presented here leverages modern open-source scientific libraries such as NumPy [10], SciPy [11] and pandas [12] used across various domains. Together with the current computational power, they allow for efficient high-level analyses, such as refined frequency analysis performing FFT only once, i.e. not in a number of iterations.

This paper describes Harpy, which combines SVD with zero-padded real FFT (RFFT) to clean the noise and compute frequency spectra efficiently, as shown in Figure 1. Such an approach addresses the performance issues (both speed and accuracy), makes the error propagation transparent, and allows for simpler use of windowing functions to trade frequency, phase, and amplitude accuracy.

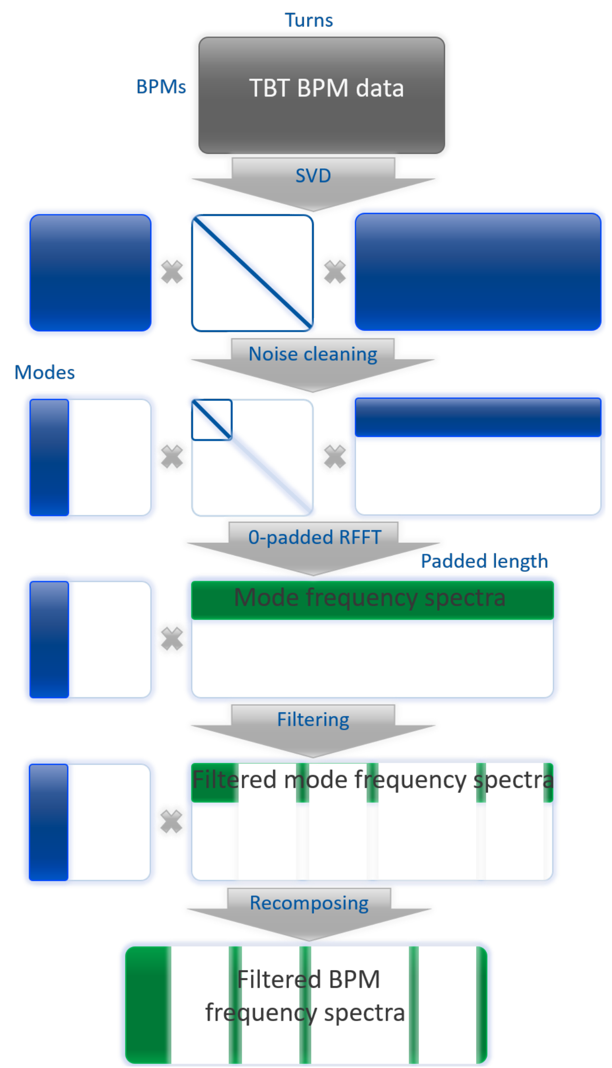


Figure 1: Schematic illustration of main Harpy principle.

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-WEPOMS035

## NOISE CLEANING USING SVD

All the BPMs observe the same beam oscillations, which grants large correlations between the different BPM signals, at the same time, BPM signals are not fully correlated due to, for example, BPM electronic noise. In order to improve analysis precision and accuracy, the BPM noise is reduced using SVD.

For each of the two planes, the TBT BPM matrix $\mathbf{A}$ is decomposed as $\mathbf{A} = \mathbf{USV^T}$, where columns of the $\mathbf{U}$ and $\mathbf{V}$ matrices are orthonormal vectors. The columns of $\mathbf{V}$ are being called “modes” in the following. $\mathbf{S}$ is a diagonal (positively definite) matrix with non-negative elements, the singular values, sorted in decreasing order. The following holds for elements $a_{jn}$ of $\mathbf{A}$ with $j$ and $n$ indexing BPMs (up to $N_{BPMs}$) and turns (up to $N_{turns}$), respectively:

$$a_{jn} = \sum_{k,l=1}^{min(N_{BPMs},N_{turns})} u_{jk} s_{kl} v_{nl} \,, \tag{1}$$

The least correlated signals represented by large $|u_{jk}|$ are removed, the columns of $\mathbf{U}$ matrix renormalised and singular values rescaled accordingly, typically in 3 iterations. The BPM noise is reduced by recomposing cleaned TBT data $\mathbf{C}$ using only the first $N_{modes}$ modes with the largest singular values (after the rescaling):

$$c_{jn} = \sum_{k,l=1}^{N_{modes}} u_{jk} s_{kl} v_{nl} \,. \tag{2}$$

The amplitude of the removed signals depends on the BPM hardware, beam parameters, and the actual choice of $N_{modes}$. Rms of the difference between raw and cleaned data estimates the BPM resolution. The estimate of noise remaining in the cleaned data will be discussed later.

## ZERO-PADDED RFFT

Once the data are cleaned their frequency spectra are computed using FFT. For a signal $x$ of length $N_{turns}$, we get a complex coefficient $X_m$:

$$X_m = \sum_{n=0}^{N_{turns}-1} x_n e^{-i2\pi mn/N_{turns}} \,, \tag{3}$$

for $m \in \{0, 1, \ldots, N-1\}$, where $m/N_{turns}$ denotes corresponding frequency. The equation has the form of inner product of $x_n$ and $e^{-i2\pi mn/N_{turns}}$. We utilise two well known facts that follow from Eq. (3):

- for real signal $x$, $X_{N-m} = X_m^*$ (complex conjugate), i.e. only half of the spectra needs to be computed by FFT.
- signal $x$ can be formally extended (padded) with zeros covering the same frequency range by larger number $N_{padded}$ of distinct frequencies.

Signal $x$ may be multiplied by a normalised windowing function $w$ to manipulate spectral leakage. Harpy utilises the output of RFFT of zero-padded (powers of 2) signal $x$:

$$X_m = \sum_{n=0}^{N_{turns}-1} x_n w_n e^{-i2\pi mn/N_{padded}} \,, \tag{4}$$

for integer $m < N_{padded}/2$ covering the frequency range from 0 to 0.5.

## HARMONIC ANALYSIS OF DECOMPOSED DATA

Multiplication by elements of the matrix $\mathbf{U}$ in Eq. (2) is a linear combination of $\mathbf{S} \cdot \mathbf{V^T}$ rows and zero-padded RFFT is a linear operation. The two operations can be swapped to reduce the amount of computation, yet leading to the identical frequency spectra. Combining the Eqs. (2) and (4), we obtain complex coefficients $C_{jm}$ of zero-padded RFFT of matrix $\mathbf{C}$ corresponding to BPM No. $j$ and frequency $m/N_{padded}$:

$$C_{jm} = \sum_{k=1}^{N_{modes}} u_{jk} \sum_{n=1}^{N_{turns}} s_{kk} v_{nk} w_n e^{-i2\pi m(n-1)/N_{padded}} \,. \tag{5}$$

For precision measurements (high $N_{padded}$), the amount of computation in left summation of Eq. (5) is reduced by keeping only the frequency ranges of interest (around multiples of betatron and synchrotron tunes). The tunes with tolerances are either provided by a user or more often calculated automatically as the frequency of the strongest line in the spectra of an average row of cleaned $\mathbf{S} \cdot \mathbf{V^T}$. Betatron tunes are found in a higher frequency range in a given plane, whilst synchrotron tune is found in a low-frequency range in the horizontal plane. For example, the calculation of RDTs up to octupolar terms requires about 6% of the frequency spectra, i.e. the left summation of Eq. (5) is performed only for 6% of $m$ values. At last, the data size is reduced again by keeping only the complex coefficients (and corresponding frequencies) with largest amplitude among a number (again powers of 2) of coefficients in same-sized bins.

Harpy identifies beam-related harmonics as the strongest lines in given frequency intervals around multiples of the (driven [13] or natural) tunes in the BPM frequency spectra. The same value of betatron tune is expected to be measured by all BPMs. Statistics-based cleaning [14] is utilised to remove BPMs, which measured betatron tunes too different from the average value.

The calculation time naturally scales with the requested precision ($N_{padded}$), the number of singular modes $N_{modes}$, and the portion of the frequency spectrum covered. There are also other effects playing a role: vectorisation, which increases the granularity of some of the parameters, and the speed of memory allocation. The 6% coverage of frequency spectra of LHC’s double-plane TBT BPM data ($N_{BPMs} \approx 500$, $N_{turns} = 6600$, $N_{modes} = 12$, $N_{padded} = 2^{21}$) is calculated in about 2 seconds, compared to about 30 seconds

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-WEPOMS035

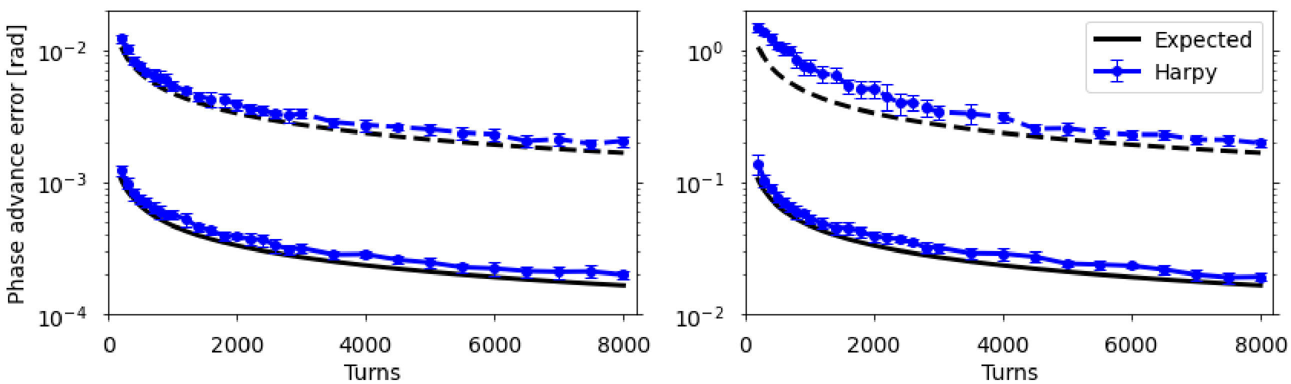


Figure 2: Accuracy of simulated phase advance measured between pairs of BPMs as a function of a number of turns. It uses supposedly the same spectral lines: the main spectral line (left plot) and two orders of magnitude weaker spectral line in its vicinity (right plot). Two sets correspond to the uncleaned noise of 0.5-1% (full lines) and 5-10% (dashed lines) compared to the amplitudes of main line. Accuracy expected from Eq. (6), which does not take the spectral leakage into account, is shown in black and measured Harpy's accuracy using Hann window is shown in blue.

for [6], both using 32 CPUs. A single parameter change at a time extends the calculation time: $N_{modes}$ = 120 to 4 seconds, covering 27% of spectra as well as $N_{padded} = 2^{23}$ to 7 seconds.

## ACCURACY AND ERROR ESTIMATES

The expected accuracy levels (without effect of spectral leakage) for phase and relative amplitude $\sigma_{phase,rel\_amp}$ of any given line in frequency spectrum calculated from Eq. (5) and with amplitude $\mathcal{A}$ is given by e.g. [15]:

$$\sigma_{phase,rel\_amp} \approx \sqrt{\frac{2}{N_{turns}}}\frac{\sigma_{orbit}}{\mathcal{A}}, \qquad (6)$$

where $\sigma_{orbit}$ is the accuracy of the position measurements, i.e. after noise reduction by SVD Eq.(2).

When subtracting the phases of the same spectral line measured in two BPMs (e.g. phase advance calculation), systematic error may arise from slightly different measured frequency of spectral line, i.e. coefficients indexed by unequal $m$s are compared. The phase of a given spectral line is the phase at "reference" turn No. 1. Following the time-wise translation symmetry of underlying physics processes, the systematic error $(\Delta m N_{turns}/2N_{padded})$ is corrected by shifting a "reference" turn to the middle of the analysed sample.

Figure 2 shows the phase advance accuracy of a stronger line and two orders of magnitude weaker secondary line in its vicinity (random frequency difference from 0.03 to 0.05) as a function of a number of turns. RMS of the difference to the value defined (in a set of 100 BPMs, and ten frequency differences) estimates the accuracy. The two sets of samples (full resp. dashed) correspond to uncleaned Gaussian noise with RMS of 0.01 and 0.1, compared to the amplitude of the main spectral line ranging from 1 to 2. The phase accuracy of Harpy was found to be better than standard method, for example, in experimental data from and CERN's PS [16].

The uncleaned noise, position error $\sigma_{phase,rel\_amp}$, is usually estimated empirically, i.e. as a certain fraction of the noise in original data. However, the following algorithm leads to a more robust error estimate after the noise reduction. Similarly to the reduction of frequency error propagating into phase error. The underlying effects have time-wise translation symmetry, i.e. the result is fundamentally not different if we remove starting or ending columns of the TBT BPM matrix. Thus we can construct, for example, three different and mostly overlapping sub-matrices of $\mathbf{A}$: $\mathbf{A^0}$, $\mathbf{A^1}$ and $\mathbf{A^2}$ shifted by a single turn. They are two turns shorter than $\mathbf{A}$. $\mathbf{A_{intersect}}$ matrix being matrix $A$ without the first two and last two columns is a sub-matrix of each of them. Stacking them up, we obtain a matrix with dimensions $3 \cdot N_{BPMs}$ times $N_{turns} - 2$. On such a matrix, we perform noise reduction by SVD. By comparison the elements of cleaned sub-matrices corresponding to elements of the matrix $\mathbf{A_{intersect}}$, we get a more error estimate after noise reduction.

## CONCLUSIONS AND OUTLOOK

Harpy, a simple method based on standard and well-established software components [10–12] has been developed. Harpy reduces the noise, computes the frequency spectra of decomposed data and filters the frequency regions of interest before recomposing BPM frequency spectra. Then, it searches the beam-related harmonics in the filtered spectra and outputs frequencies, amplitudes and phases of the harmonics, together with estimates of their errors. Optionally, Harpy removes mal-functioning BPMs using SVD and statistical cleaning [14] of tune frequency.

Harpy is currently being utilised in the optics measurements analyses across circular accelerators at CERN, PETRA III at DESY and SuperKEK-B in KEK. However, it may be helpful for frequency analysis of any set of many correlated noisy signals. In the beam optics measurements, it is typically an order of magnitude faster than standard methods while often providing more accurate results, especially for noisy data.

The code [17] was initially integrated into [18], currently its development continues as a separate package, and future releases aiming toward more general use cases can be found in [19].

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-WEPOMS035

## REFERENCES

[1] R. Tomás, M. Aiba, A. Franchi, and U. Iriso, "Review of linear optics measurement and correction for charged particle accelerators", *Phys. Rev. Accel. Beams*, vol. 20, p. 054801, 2017.

[2] J. Irwin *et al.*, "Model-Independent Beam Dynamics Analysis", *Phys. Rev. Letters*, vol. 82, p. 1684, 1999.

[3] R. Calaga and R. Tomás, "Statistical Analysis of RHIC beam position monitors performance", *Phys. Rev. ST Accel. Beams*, vol. 7, p. 042801, 2004.

[4] X. Huang, S. Y. Lee, E. Prebys and R. Tomlin, "Application of independent component analysis to Fermilab Booster", *Phys. Rev. ST Accel. Beams*, vol. 8, p. 064001, 2005.

[5] J. Laskar, "Frequency analysis for multi-dimensional systems. Global dynamics and diffusion", *Physica D*, vol. 67, p. 257-281, 1993.

[6] R. Bartolini and F. Schmidt, "A Computer Code for Frequency Analysis of Non-Linear Betatron Motion", CERN-SL-NOTE-98-017-AP, 1998.

[7] N. Biancacci and R. Tomás, "Using ac dipoles to localize sources of beam coupling impedance", *Phys. Rev. Accel. Beams*, vol. 19, p. 054001, 2016.

[8] L. Malina, J. M. Coello de Portugal, J. Dilly, P. K. Skowronski, R. Tomas, and M. S. Toplis, "Performance Optimisation of Turn-by-Turn Beam Position Monitor Data Harmonic Analysis", in *Proc. IPAC'18*, Vancouver, Canada, Apr.-May 2018, pp. 3064–3067. doi:10.18429/JACoW-IPAC2018-THPAF045

[9] T. Persson *et al.*, "Transverse Coupling Measurements With High Intensity Beams Using Driven Oscillations", in *Proc. IPAC'18*, Vancouver, Canada, Apr.-May 2018, pp. 208–211. doi:10.18429/JACoW-IPAC2018-MOPMF047

[10] C. R. Harris *et al.*, "Array programming with NumPy", *Nature*, vol. 585, pp. 357–362, 2020.

[11] P. Virtanen *et al.*, "SciPy 1.0: fundamental algorithms for scientific computing in Python", *Nat. Methods*, vol. 17, pp. 261–272, 2020.

[12] W. McKinney, "Data structures for statistical computing in Python", in *Proc. 9th Python in Science Conf.*, Austin, TX, USA, June 2010, pp. 56–61.

[13] L. Malina, J. M. Coello de Portugal, H. Timko, and R. Tomás García, "Driven 3D Beam Oscillations for Optics Measurements in Synchrotrons", in *Proc. IPAC'21*, Campinas, Brazil, May 2021, pp. 3704–3707. doi:10.18429/JACoW-IPAC2021-THXA07

[14] L. Malina, PhD thesis: "Novel beam-based correction and stabilisation methods for particle accelerators", University of Oslo, ISSN 1501-7710/No.2041 and CERN-THESIS-2018-426, 2018.

[15] Y. Alexahin and E. Gianfelice-Wendt, "Determination of linear optics functions from turn-by-turn data", *J. Instrum.* 6, no. 10, p. 10006, 2011.

[16] P. Skowronski, "RDTs in PS vs frequency analysis", Software development - LS2 IV, Mar. 2019, https: //indico.cern.ch/event/805164/

[17] L. Malina, "Harpy", https://github.com/pylhc/omc3/tree/master/omc3/harpy

[18] L. Malina, *et al.*, "OMC3", https://pypi.org/project/omc3, https://doi.org/10.5281/zenodo.5705625, https://github.com/pylhc/omc3

[19] L. Malina, "Harpy", https://pypi.org/project/harpy, https://github.com/lmalina/harpy