

## Statistical analysis of RHIC beam position monitors performance

R. Calaga and R. Tomás

*BNL, Upton, New York 11973, USA*

(Received 19 February 2004; published 9 April 2004)

A detailed statistical analysis of beam position monitors (BPM) performance at RHIC is a critical factor in improving regular operations and future runs. Robust identification of malfunctioning BPMs plays an important role in any orbit or turn-by-turn analysis. Singular value decomposition and Fourier transform methods, which have evolved as powerful numerical techniques in signal processing, will aid in such identification from BPM data. This is the first attempt at RHIC to use a large set of data to statistically enhance the capability of these two techniques and determine BPM performance. A comparison from run 2003 data shows striking agreement between the two methods and hence can be used to improve BPM functioning at RHIC and possibly other accelerators.

DOI: 10.1103/PhysRevSTAB.7.042801

PACS numbers: 29.90.+r

### I. INTRODUCTION

Beam position monitors (BPM) are widely used in accelerators to record the average orbit and transverse turn-by-turn displacements of the bunch centroid motion. RHIC consists of two threefold symmetric rings with six interaction regions. There are 160 BPMs per plane per ring (yellow and blue): 72 dual-plane BPMs distributed through the interaction regions, and 176 single-plane BPMs distributed in the arcs [1]. Each BPM channel acquires 1024 consecutive turn-by-turn positions.

It is imperative to understand the functioning of each BPM to obtain reliable and consistent data crucial for any beam dynamics analysis. In this paper we perform a first detailed statistical analysis to evaluate the performance of the BPM system with the aid of numerical tools. A large set of data files recorded after applying a horizontal kick during run 2002–2003 were analyzed using singular value decomposition (SVD) and fast-Fourier transform (FFT) techniques [2,3]. Both techniques were independently employed to identify malfunctioning BPMs from available data sets. Statistical behavior of BPM performance was extracted to characterize each pickup and make a comparison between the two methods.

### II. FFT TECHNIQUE

The Fourier spectrum of turn-by-turn data has already been used to determine faulty BPM signals from the Super Proton Synchrotron of CERN [3]. This technique relies on the fact that the Fourier spectrum of an ideal signal has well-localized peaks while noisy or faulty signals show a randomly populated Fourier spectrum. *A priori* two observables seem to provide information in this regard: the average and the rms of the background of the Fourier spectrum. The average background depends on other parameters apart from the noise of the signal, therefore it is discarded for our purposes. The rms of the background of the Fourier spectrum is larger for noisier BPMs and its dependence with other parameters is neg-

ligible. This observable is therefore used to identify the noisy pickups. It is estimated by computing the rms of the amplitudes of the spectral lines within one or several spectral windows. These windows are chosen in such a way that an ideal pickup would not show any peak within them. It is important to avoid including the zero frequency and the betatron tunes in any window. An illustration of a typical configuration for RHIC is shown in Fig. 1, using a good signal.

A signal is considered faulty if its rms noise observable is larger than a certain threshold. The value of the threshold is extracted from statistics over a large number of signals. A histogram of rms observables from all signals is constructed. Typically a large peak containing the largest percentage of the data is observed in the low rms values of the histogram. This peak contains the set of physical signals, while its long tail with larger rms values contains the faulty signals. A suitable threshold is chosen towards the end of the tail. It will be shown in Sec. IV that particular choices of the cut do not give qualitatively different results.

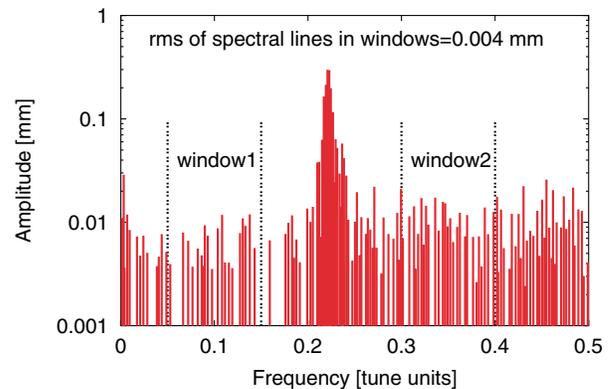


FIG. 1. (Color) Fourier spectrum of a good RHIC BPM signal. The spectral lines within the windows are used to determine the rms noise observable.

### III. SVD TECHNIQUE

The motion of the beam centroid usually depends on several variables, and it is necessary to identify the dominant variables of the system to understand beam dynamics. The problem reduces to finding dominant patterns in a data set containing displacements at multiple locations around the ring for a large number of turns. Such multivariate statistical data analysis is commonly referred to as principle component analysis and can be solved using many numerical tools.

SVD has proved to be a powerful numerical tool in a wide variety of disciplines and has been recently applied to beam dynamics under the name “model independent analysis” [2]. SVD is used to identify the principle components by maximizing the cross covariance between time-dependent data. It is a well-established technique for efficient and accurate numerical computation for large data sets.

The SVD of an arbitrary matrix  $A$  is given by

$$A = U\Sigma V^T = \sum_i u_i \sigma_i v_i^T, \quad (1)$$

where  $u_i$  and  $v_i$  are normalized eigenvectors of  $AA^T$  and  $A^T A$ , respectively, and are also referred to as left and right singular vectors of  $A$ .  $\Sigma$  is a diagonal matrix of singular values  $\sigma_i$  with  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$ . The number of nonzero singular values gives the effective rank of the matrix. For a matrix containing turn-by-turn data from several BPMs, the singular vectors represent the temporal and spatial variation of physical modes characterizing the motion of the beam. One should note that singular vectors may not uniquely correspond to the real physical modes of the beam.

The potential scope of SVD analysis in accelerators runs deep. This paper attempts to exploit one facet that aids in identifying malfunctioning BPMs. Figure 2 shows typical plots from the application of SVD to simulated BPM data (after subtracting the average orbit), with only linear elements and two noisy BPMs at arbitrary locations. The singular value spectrum shows two dominant modes (“mode 1” and “mode 2”) corresponding to the betatron oscillation in the plane of observation. The Fourier transform of the temporal mode yields the betatron tune. “Mode 3” and “mode 4” show sharp spikes in their spatial vectors. The signal manifested in this mode is localized to a particular BPM location indicating a potentially noisy BPM. The Fourier transform of their temporal modes yields all frequencies, confirming that these are noisy BPMs. In this situation of relatively ideal conditions, we find distinct peaks localized at corresponding noisy BPM locations and a flat signal elsewhere. However, from real data, we observe multiple peaks in the spatial vectors, due to random correlations between the noisy BPMs. One such possibility was simulated by setting the noise amplitudes in the two pickups to be ap-

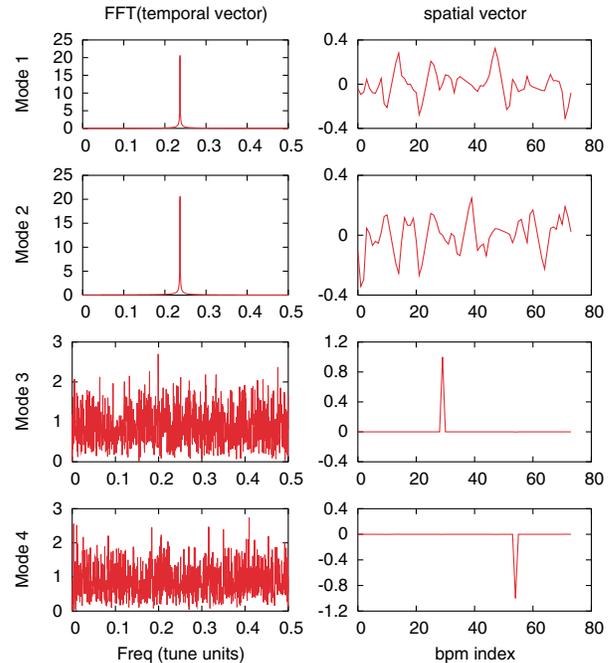


FIG. 2. (Color) Spatial vectors and FFT of temporal vectors of the four dominant modes from simulation data. Note that singular vectors are normalized.

proximately equal. We observe two dominant modes as above, but the spatial vectors now contain two peaks in each mode, as shown in Fig. 3.

A simple approach to identify faulty BPMs would entail finding modes with spatial vectors consisting only of localized peaks. A localized peak threshold value of 0.7 or greater might be sufficient to tag them as noisy BPMs [4]. However, we explore a statistical approach to understand the characteristics of BPM signals from which we choose the appropriate thresholds. A histogram of the largest peaks in each spatial mode for a large set of data is constructed to determine these thresholds. It will

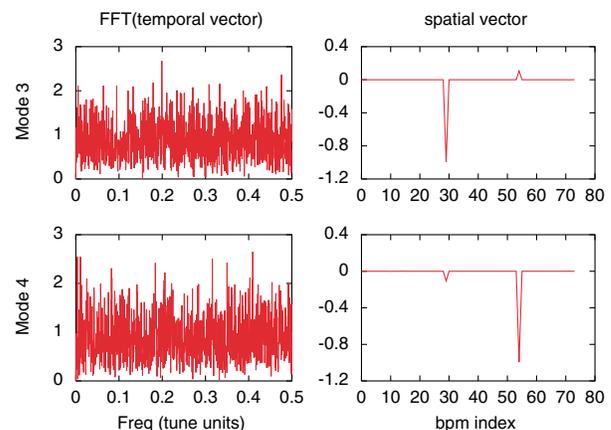


FIG. 3. (Color) Spatial and temporal vectors of modes corresponding to two noisy BPMs with correlation.

also be shown from statistics that an alternate approach using the norm of  $n$  largest peaks in each mode is a more accurate procedure to determine threshold values. One should note that the SVD method is insensitive to flat signals. Some preprocessing of BPM data, using peak-to-peak signal information, will be effective before applying such techniques. This is discussed below.

#### IV. ANALYSIS

About 2000 data sets (1000 for each ring) taken during run 2003 of deuteron-gold collisions were used in this statistical approach. Statistical cuts were applied during data preprocessing before using the two techniques. These data are then used to determine independent thresholds for each technique, to identify noisy BPMs. We discuss each statistical cut in detail to explore the advantages and limitations of such an approach.

##### A. Hardware cut

RHIC BPM hardware internally determines the status for each turn-by-turn measurement in a data set. Status bit information acquired in this way allows us to identify obvious hardware failures. We make a simple histogram of all BPMs that fail (status = 0), as shown in Fig. 4. BPMs failing this cut are removed from the data and are not included in further analysis. This histogram also helps us identify any consistent hardware problems. We also find a number of files in which fewer than 30 BPMs were present, and these files were excluded for this analysis separately for each plane.

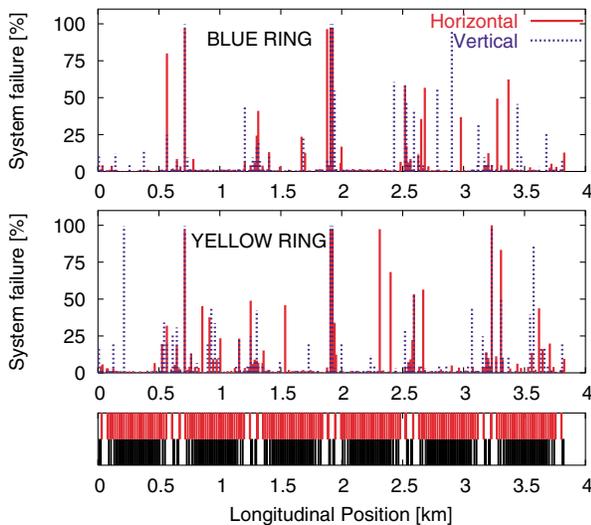


FIG. 4. (Color) Percentage of occurrences of system failure per BPM versus longitudinal location. A representation of the lattice (dipoles in black and quadrupoles in red) is shown in the bottom graph.

##### B. Peak-to-peak thresholds

A peak-to-peak cut is necessary because both techniques become less sensitive as signal oscillations become small. Peak-to-peak signal values for all BPMs in all data sets were calculated and plotted in a histogram in Fig. 5. One has to be careful to choose an appropriate cut. If optics and machine conditions were static, it would be easier to determine the best threshold. However, this is not true in regular operation. A very low cut might help retain good BPMs with a small signal, but it will also retain BPMs that are faulty. A large cut removes faulty BPMs, but will also identify many good BPMs as faulty.

The data sets being analyzed are mostly horizontally kicked data, hence we will discuss horizontal plane features in detail. There is a distinct minimum at 0.3 mm and peaks below that value which we believe are BPMs that do not respond to beam current. If a 0.3 mm cut removes good BPMs for a specific set of unexcited data, they will appear as background in the final identification, and will not be tagged as faulty BPMs. We observe a large peak around 1 mm indicating the typical oscillation amplitude for most pickups. Two smaller peaks are also observed between 1.5 and 2.2 mm, and two more between 0.5 and 1 mm, possibly indicating signals from different sets of optics or different kick amplitudes.

The vertical plane signals appear at smaller amplitudes because the data analyzed are mostly kicked in the horizontal plane. In such cases smaller peak-to-peak thresholds must be chosen. A summary of thresholds and files analyzed are given in Table I.

##### C. FFT analysis

The histograms of the FFT rms observable for the two rings and two planes are shown in Fig. 6. It is remarkable that the four peaks show almost identical features. This confirms the fact that the hardware system of the pickups is very similar for the two rings and planes. From these

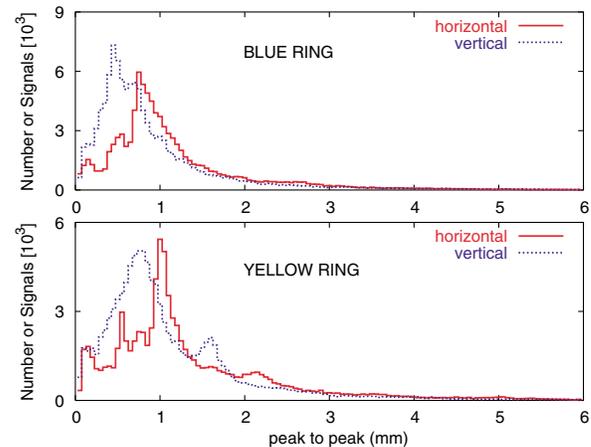


FIG. 5. (Color) Peak-to-peak values for all BPMs in all data files.

TABLE I. Thresholds for peak-to-peak values

Plane	Ring	Peak-to-peak	No. of files analyzed
H	Yellow	0.3	678
V	Yellow	0.15	815
H	Blue	0.3	708
V	Blue	0.15	833

plots it is concluded that suitable cuts lie in the range between 1.5 and 3  $\mu\text{m}$ . Signals with rms noise above the cut are labeled faulty.

To study the performance of particular BPMs, we record the number of faulty signals provided by each pickup. Figure 7 shows these occurrences plotted versus longitudinal locations for two different cuts (1.5 and 2.7  $\mu\text{m}$ ). Pickups that perform worse than the others are clearly identified from either cut.

To obtain more information on the performance of particular BPMs, the average of the rms noise observables coming from each pickup is computed and plotted in Fig. 8 versus the longitudinal location of the pickup. The picture shows few spikes that correspond to those systematically bad pickups. These spikes happen to be dense in the interaction regions. Figure 8 also shows that the BPMs within certain sextants of the machine have consistently larger rms noise than in the rest of the ring. This will be discussed in Sec. V.

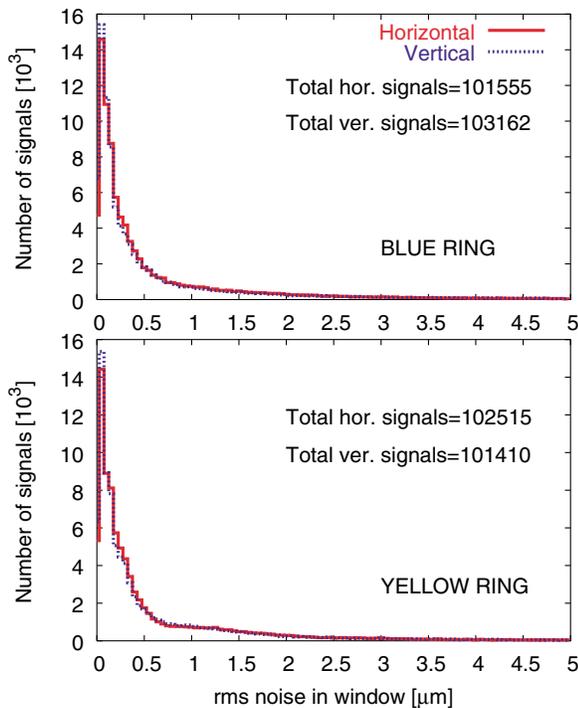


FIG. 6. (Color) Histograms of the rms observable for the two rings and the two planes. The total number of signals used in the histograms are shown.

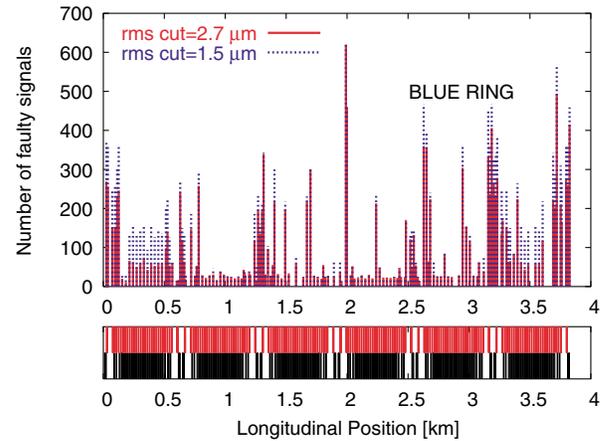


FIG. 7. (Color) Comparison between two different rms cuts showing qualitatively similar results.

### D. SVD analysis

The SVD of BPM data with oscillations above the peak-to-peak threshold can be processed to determine spatial modes with localized peaks as shown in Fig. 3. The largest peak values of all modes were determined for a large set of data and plotted in the histogram shown in Fig. 9. This figure shows a Gaussian-like distribution representing a coherent signal in good BPMs and a sharp peak close to 1 indicating obviously faulty BPMs. In the absence of any correlations between noisy BPMs, the second peak unambiguously identifies the complete set of faulty BPMs. However, for machine data we observe multiple peaks in each mode due to correlations between noisy BPMs. There exist many faulty BPMs between 0.6 and 1. Figure 9 does not provide a clear way to determine a good threshold value. A large threshold might be too

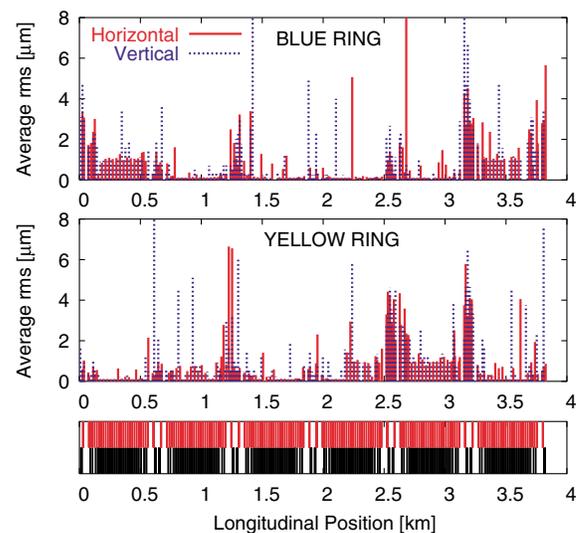


FIG. 8. (Color) Average rms observable versus longitudinal position of the BPM.

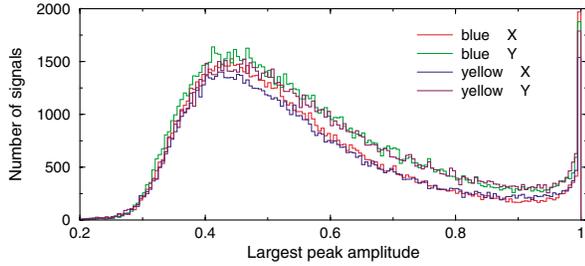


FIG. 9. (Color) Histogram of the largest peak amplitudes of all spatial modes for both planes in both rings. The Gaussian-like distributions represent a coherent signal from good BPMs and peaks at 1 represent faulty BPMs.

pessimistic, leaving unidentified faulty BPMs, while a smaller threshold might risk losing good data.

Since the singular vectors are normalized, we explore an alternate approach taking advantage of the unit norm of spatial vectors. In the presence of multiple peaks, the norm of the  $n$  largest peaks for noisy BPMs,

$$\sqrt{\sum_{j=1}^n v_{ij}^2}, \quad (2)$$

is close to 1, where  $2 \geq n \geq 5$ . However, a mode consisting of coherent signals usually has a norm  $< 0.85$  for the same  $n$ . Thus, the range of the choice of threshold values is considerably smaller than in the largest peak approach. The norms of each mode using the  $n$  largest peaks were calculated for all data sets and plotted in the histograms shown in Fig. 10. As we increase  $n$ , a clear minimum becomes apparent helping to determine the threshold

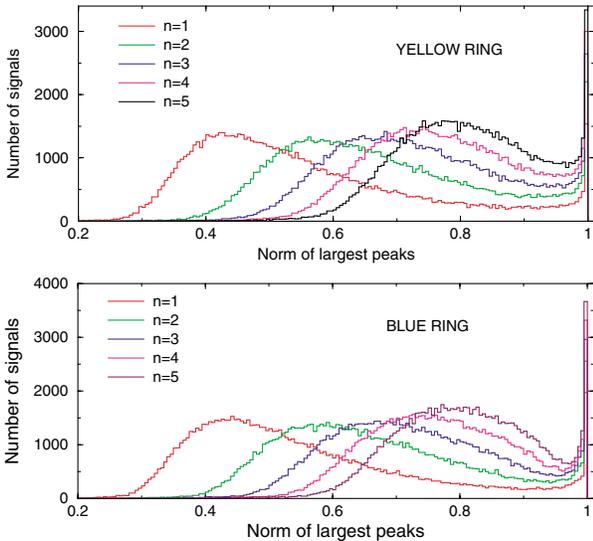


FIG. 10. (Color) Histograms of the norm of  $n$  largest peak values in each mode in the horizontal plane. Gaussian-like distributions represent coherent signal from good BPMs and peaks at 1 represent faulty BPMs.

values more accurately. Thus, this approach is more reliable than using the largest peak approach, where the optimum threshold value is ambiguous. One can determine a suitable threshold value from inspection. For the horizontal plane in the yellow ring, we find the threshold value lies between 0.85 and 1. It is important to note that signals from all four peaks contribute to fail a chosen cut, but we tag only the BPM with the largest amplitude in the norm as faulty to exclude degeneracies. The other correlated BPMs will appear dominant in subsequent modes and will be identified. A similar analysis for both planes and both rings was performed.

These threshold values and the peak-to-peak cut presented earlier are applied to data to construct a final histogram showing the number of occurrences of BPMs failing these cuts. Figure 11 shows such a histogram with two different SVD thresholds (0.85 and 0.95) for the blue ring horizontal data. This comparison allows us to understand the sensitivity of the SVD cut and hence determine an optimum range. It is clear from Fig. 11 that the regions corresponding to more noisy BPMs (mostly interactions regions) are significantly enhanced with a lower threshold compared to that of the arc regions. However, certain arc regions show different behavior from the others. This disagreement will be discussed in Sec. V.

### E. SVD and FFT comparison

To compare the SVD and FFT techniques, the number of identified faulty signals provided by each BPM is plotted versus its location, in Fig. 12. This plot contains the horizontal BPMs of the yellow ring. The reader can also compare the results for horizontal BPMs of the blue ring from Figs. 7 and 11. For a set of optimized cuts, the agreement is excellent between the two different techniques. This confirms the feasibility of identifying non-physical signals provided by beam position monitors.

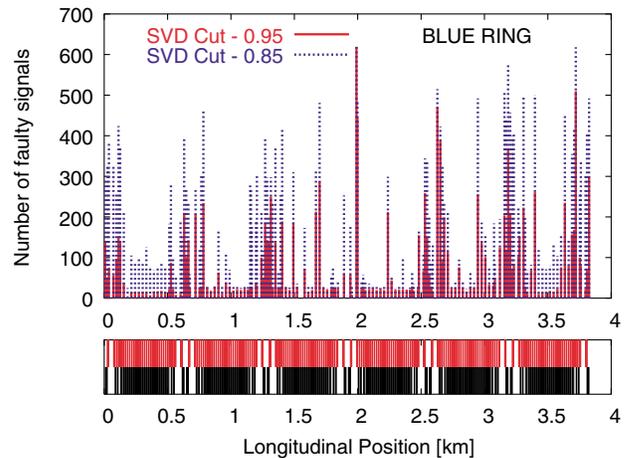


FIG. 11. (Color) Comparison between two SVD thresholds (0.85 and 0.95) for the blue ring-horizontal plane.

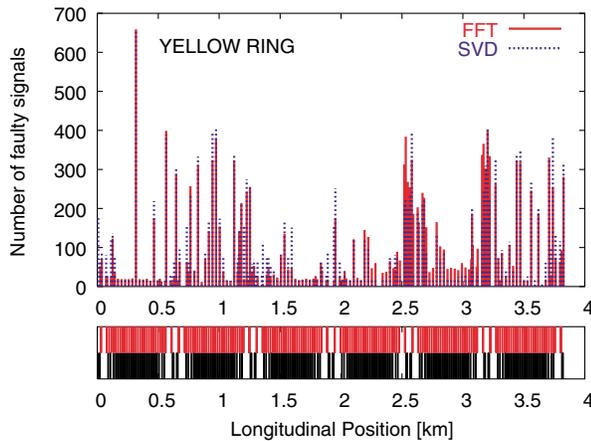


FIG. 12. (Color) Comparison of the FFT and SVD techniques for 678 files from the yellow ring showing a good agreement. The number of times a BPM provides a faulty signal is plotted versus its location.

## V. OBSERVATION OF SYSTEM IMPROVEMENTS

As seen in Figs. 7, 8, and 11 BPMs in the arc regions between 0–0.5 km and 3.2–3.8 km have a strikingly larger background than the rest of the arcs. It was found that BPMs in these arcs were exhibiting sporadic noise (“hairs”) of ten to thousands of  $\mu\text{m}$  on BPM position data. This effect was caused due to unforeseen and untested conditions in the low-level digital signal processing (DSP) code leading to noise jumps on some raw sampled data [5].

Low-level software upgrades were implemented as a part of BPM system improvement during the commissioning period of run 2004. A change of DSP timing parameters was particularly important in resolving the phenomena of hairs. Kicked data during run 2004 were regularly recorded for the analysis of BPM performance. A large set of data similar to the one from 2003 was analyzed using the above numerical techniques to understand the behavior of the BPMs after system improvements.

Figure 13 shows the percentage of occurrences of system failure per BPM versus longitudinal location for run 2004. These percentages of system failure are in general larger than those observed in the previous year as seen in Fig. 4. This is partly due to the fact that the system improvements lead to a better recognition of system failures. The abnormal abundance of system failures at the location of 3 km is presently under investigation.

Figure 14 shows the average of the rms noise observables coming from each BPM plotted versus the longitudinal location of the BPM for run 2004. This figure is to be compared to Fig. 8 from the previous year. We observe a clear improvement in BPM resolutions. The larger background at particular sextants has disappeared yielding a consistent BPM performance in all the arcs. This confirms

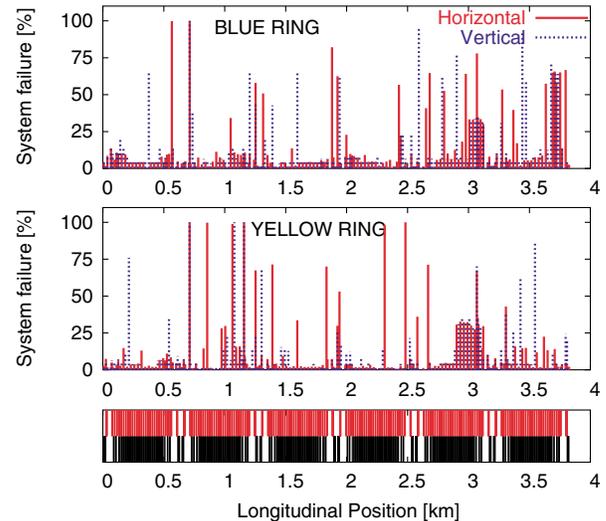


FIG. 13. (Color) Percentage of occurrences of system failure per BPM versus longitudinal location for run 2004. A representation of the lattice (dipoles in black and quadrupoles in red) is shown in the bottom graph.

that the problem of hairs has been resolved. Nevertheless BPMs located within the interaction regions still show a poorer performance than the rest.

Figure 15 shows faulty BPMs identified by FFT and SVD techniques with two different thresholds. The thresholds for the FFT technique are smaller than the previous year since the rms observable is smaller. These histograms help us determine the sensitivity of the cuts in each method. It is again clear from both techniques that the problem of hairs in the two sextants has been resolved.

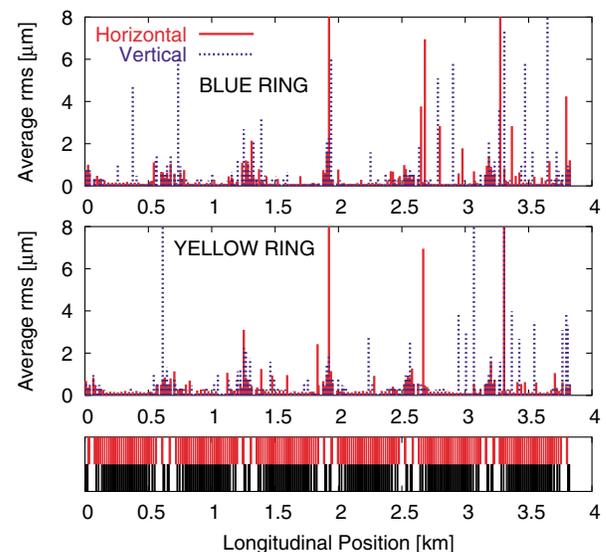


FIG. 14. (Color) Average rms observable versus longitudinal position of the BPM for run 2004.

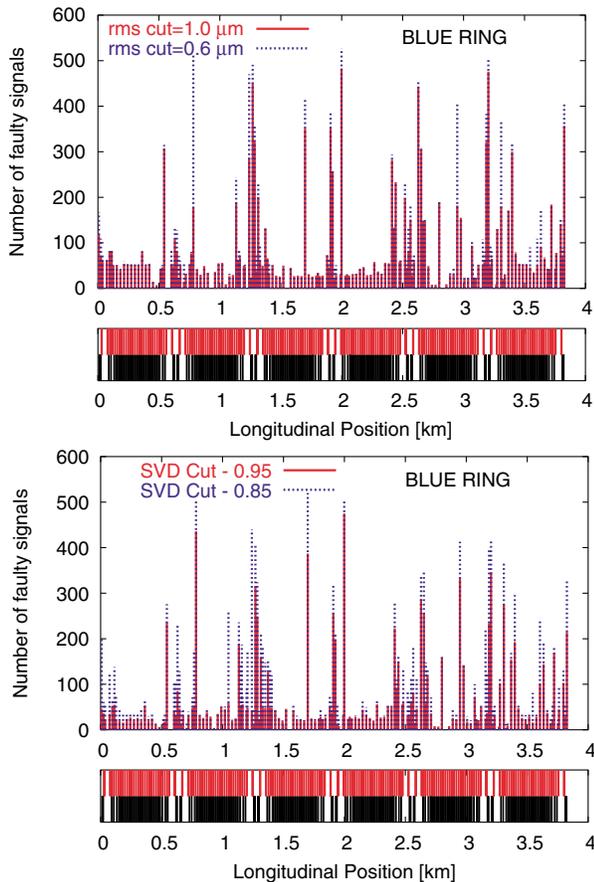


FIG. 15. (Color) Comparison between two different rms and SVD cuts for run 2004.

## VI. CONCLUSION

Two numerical algorithms, the Fourier transform and the singular value decomposition, have been used to identify faulty BPM signals. Appropriate observables characterizing the noise content of a BPM signal have been defined for both techniques. The observables'

thresholds above which a signal is identified as faulty are obtained from statistics over a large set of RHIC BPM data. These cuts are used to assess the global performance of every BPM, thus identifying those BPMs that systematically provide faulty data. A comparison between the results from both independent techniques has been presented showing an excellent agreement. The analysis of run 2004 data shows clear improvements in the BPM system and the successful elimination of hairs.

## ACKNOWLEDGMENTS

We are greatly indebted to T. Satogata, W. Fischer, and the operations crew for acquiring the data and making this analysis possible. We would also like to thank P. Cameron, P. Cerniglia, C. Dawson, C. Degan, J. Mead, and T. Satogata for valuable discussions on the BPM system. The authors also thank W. Fischer, W. McKay, S. Peggs, and T. Satogata for reading the manuscript and making valuable comments. This work was performed under the auspices of the U.S. Department of Energy.

- 
- [1] M. Bai, P. Cameron, P. Cerniglia, R. Connolly, J. Cupolo, C. Degen, A. Drees, R. Fliller, D. Gassner, J. Mead, V. Ptitsyn, T. Satogata, T. Shea, R. Sikora, P. Thompson, and R. Witkover, *Nucl. Instrum. Methods Phys. Res., Sect. A* **499**, 372 (2003).
  - [2] J. Irwin, C. X. Wang, Y. Yan, K. Bane, Y. Cai, F. Decker, M. Minty, G. Stupakov, and F. Zimmerman, *Phys. Rev. Lett.* **82**, 1684 (1999).
  - [3] R. Tomás, Ph.D. thesis, University of Valencia, Spain, 2003.
  - [4] C. X. Wang, M. Borland, V. Sajaev, and K. Kim, in *Proceedings of the PAC2001, Chicago, IL* (IEEE, Piscataway, NJ, 2001), p. 1354.
  - [5] T. Satogata (private communication).