Statistical methods for transverse beam position diagnostics with higher order modes in third harmonic 3.9 GHz superconducting accelerating cavities at FLASH

P. Zhang[1,2,3], N. Baboi[2], R.M. Jones[1,3]
[1]School of Physics and Astronomy, The University of Manchester, Manchester, U.K.
[2]Deutsches Elektronen-Synchrotron (DESY), Hamburg, Germany
[3]The Cockcroft Institute of Accelerator Science and Technology, Daresbury, U.K.

Abstract

Beam-excited higher order modes (HOM) can be used to provide beam diagnostics. Here we focus on 3.9 GHz superconducting accelerating cavities. In particular we study dipole mode excitation and its application to beam position determinations. In order to extract beam position information, linear regression can be used. Due to a large number of sampling points in the waveforms, statistical methods are used to effectively reduce the dimension of the system, such as singular value decomposition (SVD) and k-means clustering. These are compared with the direct linear regression (DLR) on the entire waveforms. A cross-validation technique is used to study the sample independent precisions of the position predictions given by these three methods. A RMS prediction error in the beam position of approximately 50 micron can be achieved by DLR and SVD, while k-means clustering suggests 70 micron.

keywords Higher order mode, Beam position diagnostics, Superconducting accelerating cavity, Singular value decomposition, k-means clustering, Cross-validation

Declaration: This paper has been accepted by Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment (ISSN: 0168-9002). Currently in press.

Nuclear Instruments & Methods In Physics Research A (2012),
DOI: http://dx.doi.org/10.1016/j.nima.2012.11.057

1 Introduction

FLASH is a free-electron laser (FEL) facility at DESY (Hamburg, Germany) providing ultrashort radiation pulses with an unprecedented brilliance [1]. This is a user facility for fundamental light source applications and test facility for accelerator physics. An electron source generates a high quality electron beam, which is then driven by seven 1.3 GHz superconducting TESLA type [2] accelerating modules to maximum energy of 1.2 GeV. In the undulator section the beam produces a laser-like coherent FEL radiation.

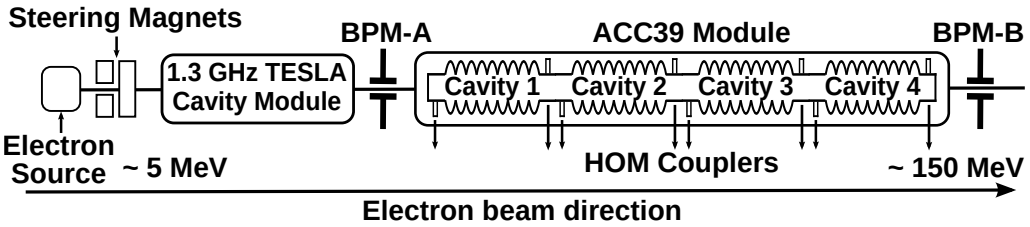

Figure 1: Schematic of the FLASH injector section (not to scale, the 1.3 GHz TESLA cavity is approximately three times longer than the 3.9 GHz cavity in ACC39 module).

The injector section of FLASH is shown schematically in Fig. 1. A third harmonic module (ACC39) containing four 3.9 GHz superconducting cavities is installed in FLASH in order to linearize the longitudinal phase space of the electron bunch [3]. Wakefields are much stronger in these cavities than in the TESLA

1.3 GHz cavities [4]. If left unchecked, these will add adverse effects to the beam resulting in a dilution of the beam quality. Dipole higher order modes (HOM), with linear dependence on the offset of the excitation bunch [5], have the main contribution to the transverse wakefield. By centering the beam in the cavities, one can reduce the excited transverse wakefield. Moreover, monitoring dipole modes provides an indirect measurement of the transverse beam position in the cavity. Such beam diagnostics have been previously studied and electronics have been built for the TESLA 1.3 GHz cavities at FLASH, using a dipole mode at approximately 1.7 GHz with large R/Q value [6]. In the 3.9 GHz cavities, however, no obvious candidate mode exists [7]. The spectrum is dense particularly due to most HOMs being above the cutoff frequency of the inter-connecting beam pipes [8]. Also these inter-coupled modes preclude a local determination of beam position within each cavity in contradistinction to the TESLA 1.3 GHz cavities. There are however a limited number of modes in the fifth dipole band which are not coupled [9]. These have smaller R/Q values compared to those in the inter-coupled region and hence have a reduced resolution for beam position diagnostics [10].

In order to evaluate the various frequency regions which could be used for diagnostics, test electronics have been built by FNAL [10, 11]. This down-converts the HOM signal measured from one of the two HOM couplers of each cavity containing dipole modes between 4 and 9.1 GHz to 70 MHz with a 20 MHz bandwidth. The signal is then digitized with a rate of 216 MS/s and a 14 bit resolution. The HOM data are recorded synchronously to various data from the FLASH control system such as beam position monitors, magnetic steerers and charge monitors. The results obtained with the test electronics are used for the following studies.

In order to diagnose transverse beam position with the HOM signal, a straightforward method is to perform a direct linear regression (DLR). However due to a large number of sampling points present in the HOM signal, a considerable amount of unknown variables need to be determined. This is computationally expensive and has the risk of overfitting. It is therefore advisable to reduce the signal dimension. Arranging HOM signals in a matrix form, a row-based method named singular value decomposition (SVD) and a column-based method named k-means clustering are used. The signal size has been effectively cut down by two orders of magnitude to several tens whilst the prediction accuracies are retained compared to DLR. All three methods are elucidated comprehensively in Section 2. The performances of these methods are analyzed and compared firstly on a fixed sample split. Since the samples are scarce, a cross-validation technique is applied to remove the sample dependence in the performance evaluations. These comparisons are presented in Section 3. In Section 4 we conclude that in the HOM-based beam position diagnostics, DLR is an accurate but time-inefficient method, while k-means clustering is computation-efficient but degenerates the position resolution by approximately one third. In contrast, SVD is a robust method in terms of both computation time and position resolution, and has been therefore chosen to be used in the final electronics.

2 Algorithms for beam position diagnostics

The test electronics have been connected to the upstream HOM coupler of Cavity 3 as shown in Fig. 1. The digitized data has been filtered such that only the trapped cavity modes in the fifth dipole band from 9040 MHz to 9078 MHz [9] are preserved and then normalized to the beam charge. These modes are of particular interest since they can provide localized beam position information. We have steered the beam in a two-dimensional grid manner by using a pair of steerer magnets located upstream of ACC39 module as shown in Fig. 1. The RF of ACC39 and the nearby quadrupole magnets were both switched off in order to produce a straight-line beam trajectory between the two beam position monitors: BPM-A and BPM-B. Thus the transverse beam positions inside the cavity can be determined from the readouts of the two BPMs by linear interpolation. The detailed setup of the FLASH machine for this study can be found in [10]. The goal is to extract transverse beam position from the HOM waveforms.

2.1 Data preparation

The HOM waveforms for 85 different beam positions are shown in Fig. 2(a). The corresponding beam positions are shown in Fig. 2(b) (blue and red dots).

In order to define the problem, we arrange the waveforms and beam positions in matrix forms. The

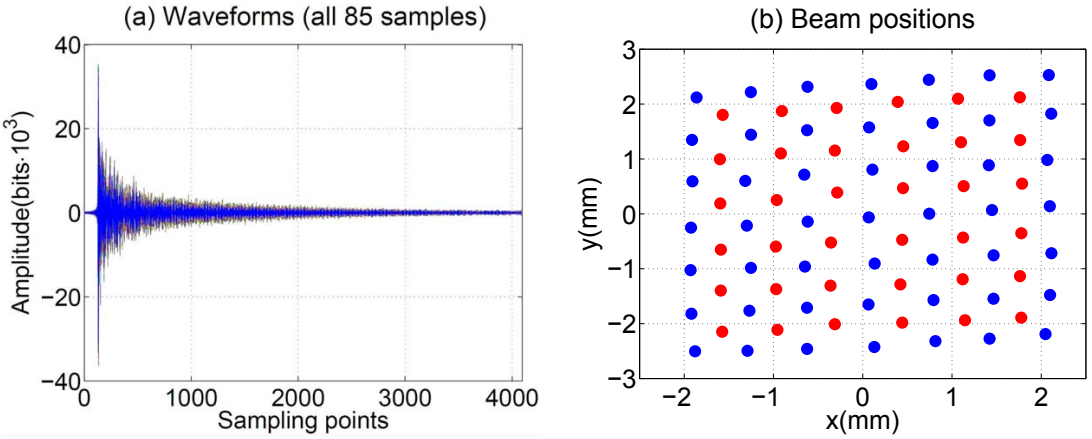

Figure 2: (a) Waveforms of all 85 samples. (b) Fixed sample split. Calibration samples are in blue and validation samples are in red.

matrix A contains the waveforms and is constructed as

$$A = \begin{pmatrix} waveform_1 \\ waveform_2 \\ \vdots \\ waveform_m \end{pmatrix} = \left(a_1, a_2, \cdots, a_n\right) \in \mathbb{R}^{m \times n}. \tag{1}$$

A is a $m \times n$ matrix. $a_i \in \mathbb{R}^m$ is the i^{th} column of A, which contains values of the i^{th} sampling point from all m waveforms. Each waveform has 4096 sampling points, therefore $n = 4096$.

The matrix B contains the interpolated beam positions and is constructed as

$$B = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_m & y_m \end{pmatrix} \in \mathbb{R}^{m \times 2}. \tag{2}$$

B is a $m \times 2$ matrix. x_i and y_i are the i^{th} transverse beam positions at the center of the cavity corresponding to the measurement of $waveform_i$. There are technical difficulties which would be entailed in obtaining angle effects in ACC39 independently. As the contribution of the angle wakefield is much smaller than that of the offset wakefield [6], the angle is not considered in this study.

Now we have built a system consisting of waveforms A and beam positions B. To study this system, we split the samples in a fixed manner as shown in Fig. 2(b). 49 samples are used for calibration (blue) and the remaining 36 samples for validation (red). Therefore, calibration samples have $m = 49$ and validation samples have $m = 36$ in both Eq. 1 and Eq. 2.

2.2 Direct linear regression

To extract beam position information from the waveforms, a straightforward method is *Direct Linear Regression* (DLR),

$$A \cdot M = B, \tag{3}$$

where $A \in \mathbb{R}^{m \times (n+1)}$ with one column of 1 added to the waveform matrix defined in Eq. 1 representing the intercept term, $B \in \mathbb{R}^{m \times 2}$ is the position matrix defined in Eq. 2, $M \in \mathbb{R}^{(n+1) \times 2}$ is the matrix of regression coefficients. M is further represented as (M_1, M_2), where M_k is a $(n+1) \times 1$ vector ($k = 1$ or 2). The matrix M is obtained by solving the linear system composed by matrix A and B.

The linear system is highly under-determined where $m \ll n$, which is vulnerable to overfitting. Therefore, we used a modified least square algorithm to do the linear regression: *Regularized Least Squares* (RLS) [12].

A regularization term is added into the normal error function in order to control the overfitting. The RLS error function is therefore defined as

$$J_k = \frac{1}{2}\sum_{i=1}^{m}(B_i^{pre} - B_i)^2 + \frac{\lambda}{2} \parallel \vec{M}_k \parallel^2, (k = 1, 2) \tag{4}$$

where B_i^{pre} is the predicted beam position calculated by using Eq. 3, B_i^{pre} and B_i represent x when $k = 1$ or y when $k = 2$, $\parallel \vec{M}_k \parallel$ is the 2-norm [13] of M_k, and λ is the regularization coefficient controlling the relative importance of the regularization term. λ provides a shrinkage of the value of the regression coefficients in M. The linear regression is therefore performed by minimizing the RLS error functions while updating the elements of M. A figure of merit has also been defined in order to evaluate the goodness of the regression: the root-mean-square (RMS) error,

$$E_{RMS} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(B_i^{pre} - B_i)^2}. \tag{5}$$

E_{RMS} has the unit of mm, which is the same as B_i^{pre} and B_i.

The impact of the regularization term in Eq. 4 on the prediction error can be seen in Fig. 3 by plotting the RMS errors of both calibration and validation samples against the regularization coefficient λ. A model with a $\lambda < 10^8$ has a stable and promising performance, while in contrast a poor fit emerges for a large λ. The complexity of the model is now controlled with λ overseeing the degree of overfitting. When $\lambda < 10^7$, the RMS errors fall sharply for calibration samples whiles slightly rise for validation samples. This is an indication of overfitting due to the lack of control by a too small λ. Thus we choose $\lambda = 10^7$ for the following analysis.

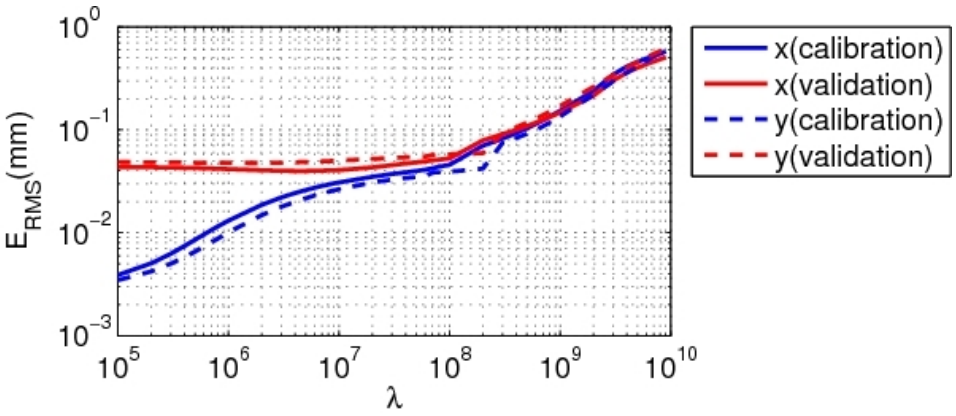

Figure 3: RMS error versus λ.

The value of the RLS error function is reduced in each iteration while performing the linear regression as shown in Fig. 4(a). The difference of the J's between two consecutive iterations is denoted as ΔJ. The values of J's decrease evidently when $\Delta J > 10^{-7}$, and then land on a relatively flat region between 10^{-9} and 10^{-7} of ΔJ. The RMS errors of both calibration and validation samples for each iteration are shown in Fig. 4(b). The prediction power on the validation samples stabilizes after $\Delta J < 10^{-6}$. An overfitting emerges while $\Delta J < 10^{-7}$ as the RMS errors of calibration samples drop sharply, while the RMS errors of validation samples rise slightly. Therefore, we set that the calculation stops when $\Delta J < 10^{-7}$.

Using $\lambda = 10^7$ and $\Delta J < 10^{-7}$, the predicted beam positions are obtained by RLS and then plotted together with the direct measurement for both calibration and validation samples in Fig. 5. To measure the consistency, the coefficient of determination r^2 is calculated ($r^2 = 1$ corresponds to a perfect fit whilst $r^2 = 0$ is a poor fit) [14]. The differences between predicted and measured positions are shown in Fig. 6 as the residuals. The RMS error on the validation samples is 40 μm for x and 50 μm for y.

The size of matrix M is 4097×2. A considerable number of unknown variables need to be determined. This is computationally expensive. As all sampling points are used in the regression, each column of matrix

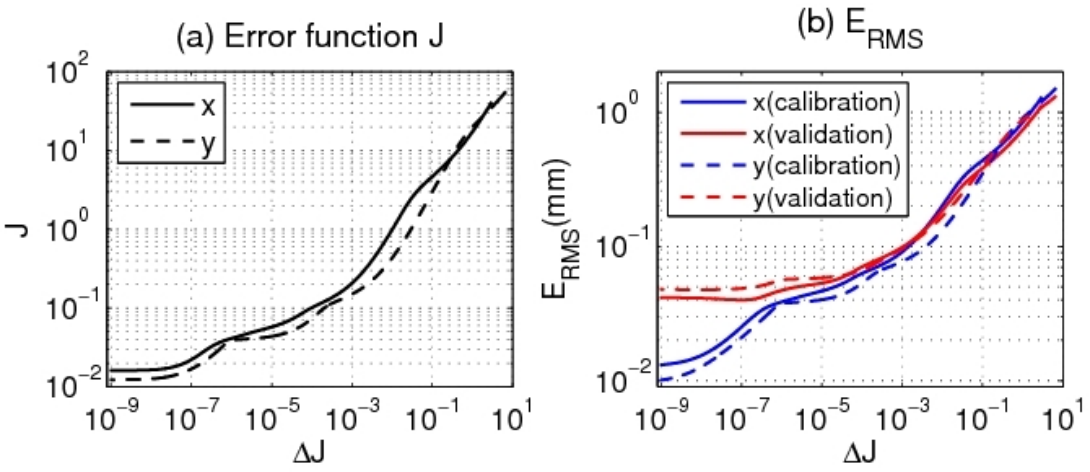

Figure 4: Error function and RMS error in calculating the calibration matrix M.

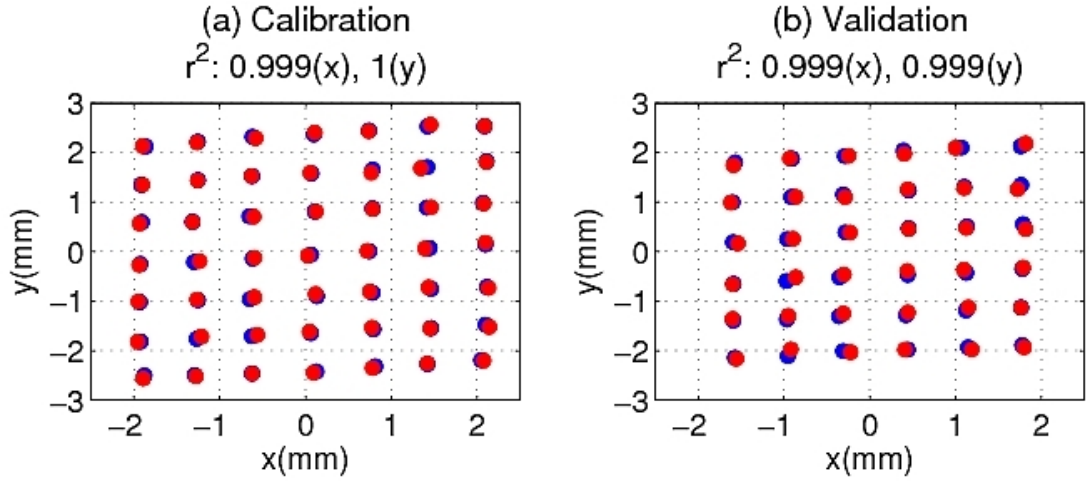

Figure 5: Measurement (blue) and prediction (red) of the transverse beam position from calibration and validation samples. The method applied is DLR.

A, a_i, is a regressor. The correlation coefficients $R(i,j)$ of these regressors are calculated as [14]

$$R(i,j) = \frac{\mathrm{Cov}(a_i, a_j)}{\sqrt{\mathrm{Cov}(a_i, a_i) \cdot \mathrm{Cov}(a_j, a_j)}}, \tag{6}$$

where a_i and a_j are i^{th} and j^{th} column of matrix A as defined in Eq. 1, $\mathrm{Cov}(a_i, a_j)$ is the covariance [15] between two regressors a_i and a_j. In this definition, $R = \pm 1$ corresponds to a strong correlation while $R = 0$ indicates no correlation at all. Fig. 7 shows the correlation coefficients of the calibration samples. Most regressors are strongly correlated and this makes the linear system sensitive to the fluctuations of the calibration samples. In our case, it means that the system is vulnerable to noise. Moreover, even though we have made efforts to control it, overfitting is always a potential problem unless the number of calibration samples is fairly large ($m > n$). To avoid these risks, it is prudent to reduce the system size from multi-thousand to several tens, and then apply linear regression on the reduced system. The methods we used for this purpose are described in the following sections.

2.3 Singular value decomposition

In order to reduce the dimension of the system, a method known as *Singular Value Decomposition* (SVD) [16] is used to find a small number of prominent components from the waveform matrix A. In general, SVD looks for patterns of a matrix in terms of SVD modes without the knowledge of any physical parameters (like dipole mode frequency, quality factor, beam position, etc.). Those SVD modes are natural groupings of the signal matrix, which are not clearly visible or explicitly defined in the matrix itself.

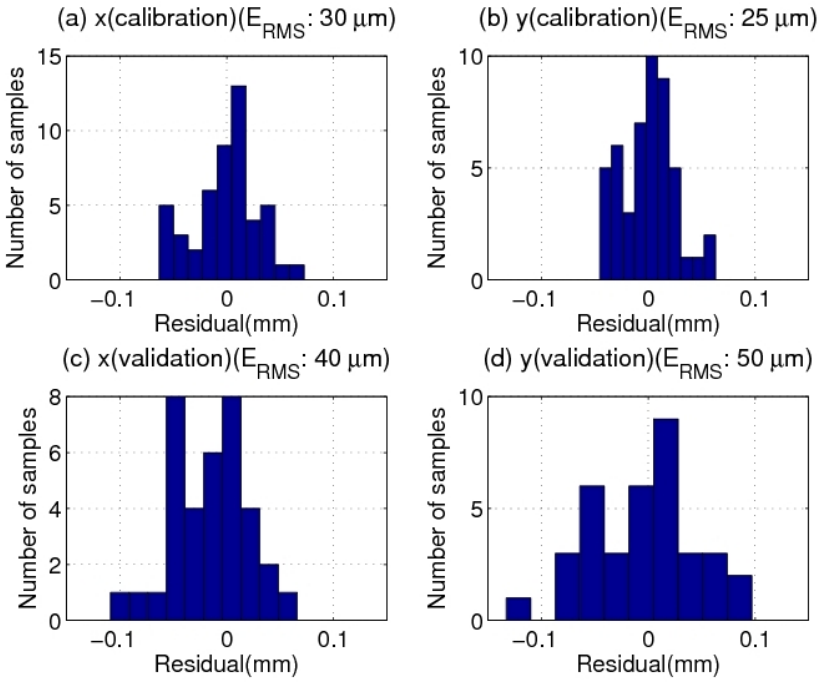

Figure 6: Difference of measured and predicted transverse beam position from calibration and validation samples. The method applied is DLR.

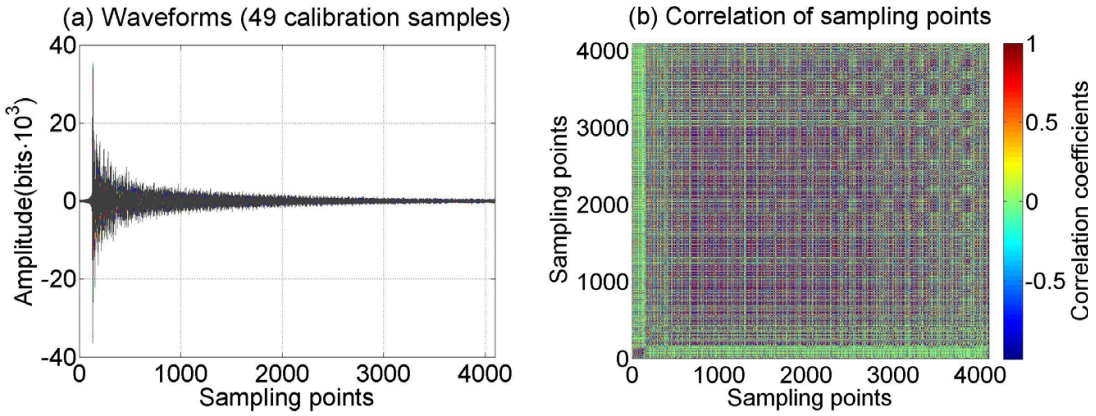

Figure 7: (a) Waveforms of 49 calibration samples; (b) Correlation coefficients between each regressor of DLR.

Applying SVD, the matrix A is decomposed into the product of three matrices,

$$A = U \cdot S \cdot V^T, \tag{7}$$

where V^T indicates the transpose of V. The columns of U and V are singular vectors of A, which form the bases of the decomposition. S is a diagonal matrix whose non-zero elements are known as singular values. Fig. 8(a) shows the singular value of each SVD mode decomposed from the calibration samples. It can be seen that the first few SVD modes have relatively large singular values, in other words, they are the dominant patterns of matrix A.

Each of the SVD modes can be used to produce a waveform matrix by

$$A_i = U_i \cdot S_{ii} \cdot V_i^T, \tag{8}$$

where U_i is the i^{th} column of U, and V_i^T is the i^{th} row of V^T. S_{ii} denotes the i^{th} diagonal element of S. A_i has the same size as the original waveform matrix A. Fig. 9 shows one waveform for each of the first nine

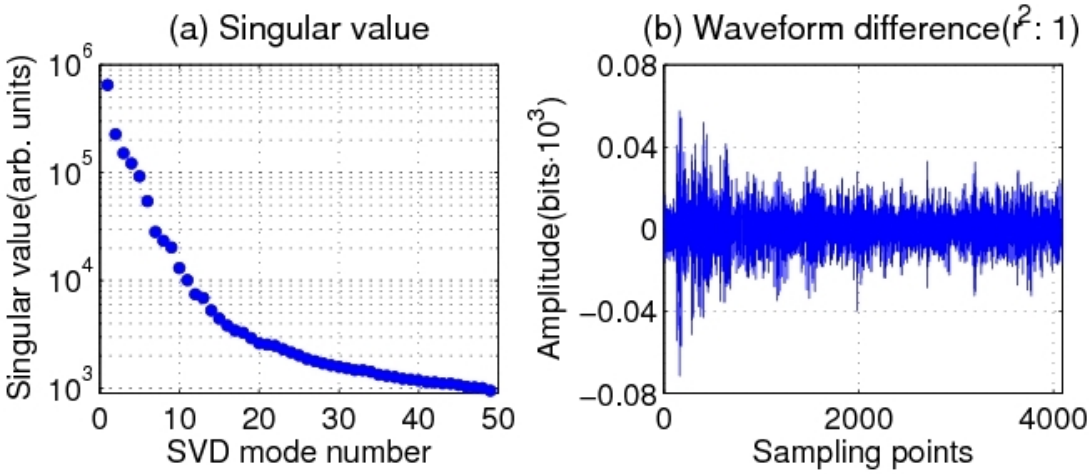

Figure 8: (a) Singular value of each SVD mode; (b) Difference of the original and the reconstructed waveform using the first 26 SVD modes. Note the scale of the waveform amplitude when comparing with Fig. 2(a).

A_i. The approximation A_{reco} can be calculated by simply summing over the related A_i's by

$$A_{reco} = \sum_{i=1}^{p} A_i. \tag{9}$$

Fig. 8(b) shows the difference between the original waveform A and A_{reco} for $p = 26$. One can see that the waveform can be well represented by using only the first 26 SVD modes, which is also illustrated by $r^2 \approx 1$. The choice of p will be discussed later.

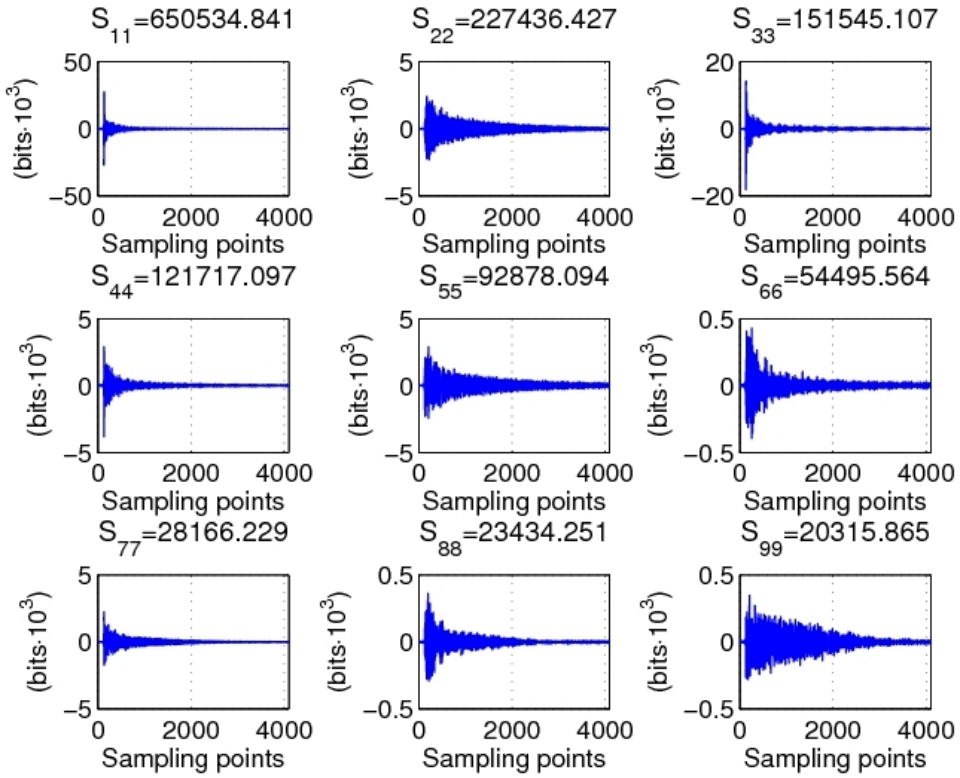

Figure 9: Waveform reconstructed from each of the first nine SVD modes.

With the SVD base vectors from the decompositions of waveform matrix A, the amplitude of all SVD modes, $A_{svd,full}$, can be obtained for each beam position by

$$A_{svd,full} = A \cdot V = (A_1^{svd}, A_2^{svd}, \cdots). \tag{10}$$

$A_i^{svd} \in \mathbb{R}^{m\times 1}$ contains the mode amplitude for each beam position of the i^{th} SVD mode. Fig. 10(a) shows the mode amplitude of the first 49 SVD modes with non-zero singular values obtained from the calibration samples. The correlation coefficient between each pair of SVD mode is also calculated using Eq. 6 by replacing a_i and a_j by A_i^{svd} and A_j^{svd} and it is shown in Fig. 10(b). The first SVD mode has relatively strong correlations with several SVD modes. This is because the mean of the 49 calibration waveforms is not zero.

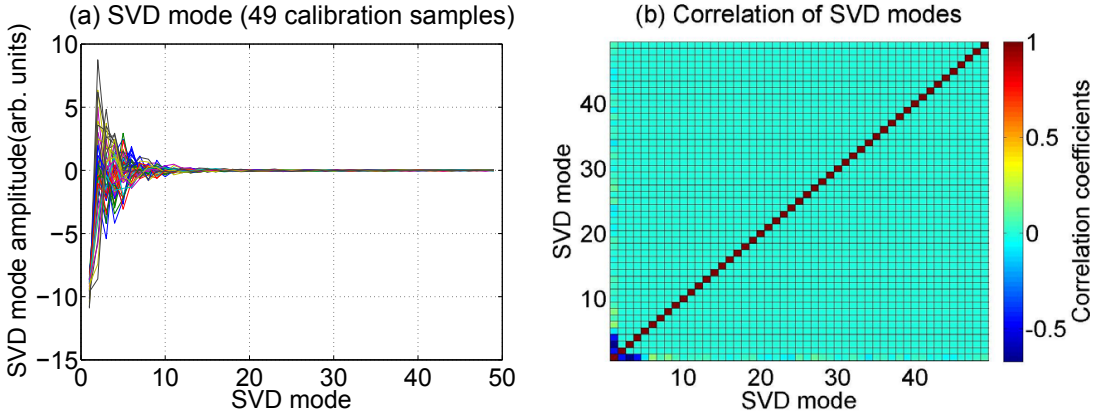

Figure 10: SVD mode amplitudes and their correlations.

Along with a column of 1, $I \in \mathbb{R}^{m\times 1}$, representing the intercept term, we use the first p ($p \leq m$) SVD modes to build the SVD mode amplitude matrix A_{svd}:

$$A_{svd} = (I, A_1^{svd}, A_2^{svd}, \cdots, A_p^{svd}), \tag{11}$$

The size of the matrix A_{svd} is significantly smaller ($p+1$ columns, representing p SVD modes and 1 intercept term) compared to the original waveform matrix A (4097 columns, representing 4096 sampling points and 1 intercept term). Replacing A by A_{svd} in the regression (Eq. 3), the linear system composed by A_{svd} and B is now over-determined, and has a best solution in a least square sense. The prediction for the validation samples has two steps: first, project A onto the base vectors obtained from calibration samples to get the amplitude of the SVD modes using Eq. 10; second, predict positions using the SVD amplitude matrix and the coefficient matrix M.

The contribution of combining the first p SVD modes to determine the transverse beam position x and y is measured by the RMS error as shown in Fig. 11. Zooming into the region where $12 \leq p \leq 32$, the E_{RMS} of x and y for validation samples are shown in Fig. 12. Using the first 26 SVD modes is seen to give an optimal performance for both x and y plane. This confirms, with Fig. 8, that the first few SVD modes contain the majority of the beam position information.

Fig. 13 shows the HOM response from calibration and validation samples by using the first 26 SVD modes to determine the transverse beam position x and y. The RMS errors remain small and comparable for calibration and validation samples as shown in Fig. 14. The RMS error is 44 μm for x and 45 μm for y when predicting positions from validation samples.

2.4 *k*-means clustering

An information redundancy is present in the linear system composed of A (Eq. 1) and B (Eq. 2), as the regressors a_i's from the waveform matrix A are highly correlated (Fig. 7). A technique named *k-means clustering* [12] is used in order to group the data into a few cohesive *clusters*. Each a_i is considered as one point in the m-dimensional space, and the goal is to find k groups of points based on Euclidean distance. The cluster centroid μ_j is defined as the average value of a_i's

$$\mu_j = \frac{1}{M}\sum_{i=j_1}^{j_M} a_i, \tag{12}$$

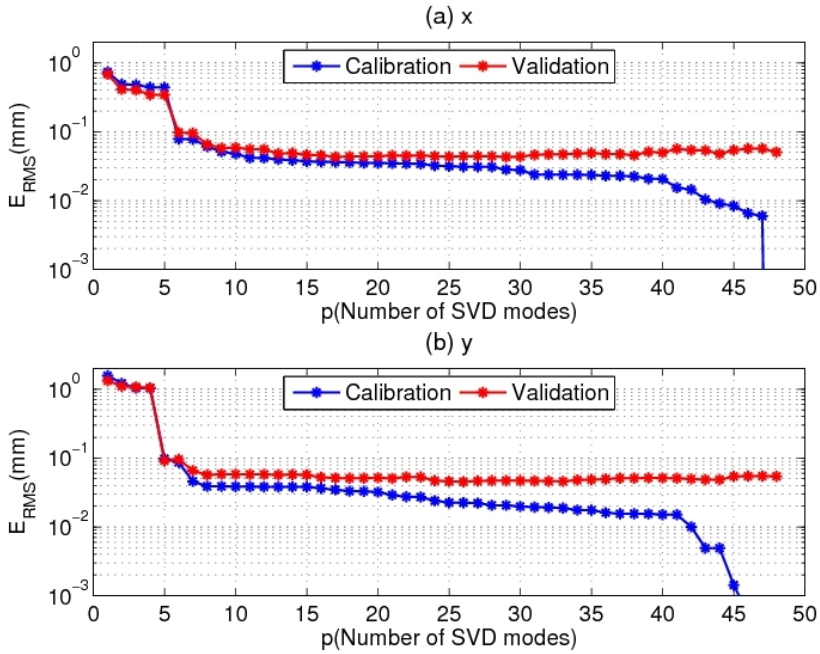

Figure 11: Contribution of the first p SVD modes to determine the transverse beam position x and y measured by E_{RMS}.

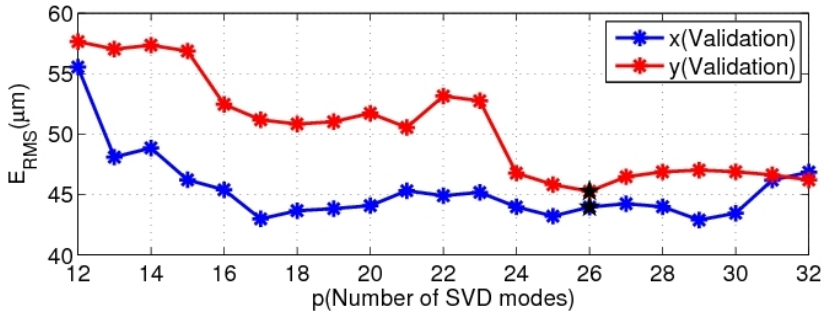

Figure 12: Contribution of the first p SVD modes ($12 \leq p \leq 32$) to determine the transverse beam position x and y measured by E_{RMS}.

where $(j_1, j_2, \cdots, j_M)$ are M indices of the m-dimensional points a_i's which belong to the j^{th} cluster whose centroid is μ_j. Fig. 15 shows 22 clusters in red partitioned from the calibration samples in blue for one beam position. The values of the corresponding 22 centroids are shown in Fig. 16(a) for all 49 beam positions. The correlation coefficients of these clusters can be calculated by replacing a_i with μ_j in Eq. 6 and are shown in Fig. 16(b). As expected, correlations can be observed among clusters.

The contribution of different number of clusters to determine the transverse beam position x and y is measured by the RMS error as shown in Fig. 17. Using 22 clusters is seen to give an optimal performance for both x and y.

Fig. 18 shows the HOM response from calibration and validation samples by using 22 clusters to determine the transverse beam position x and y respectively. The RMS errors remain small and comparable for calibration and validation samples as shown in Fig. 19. The RMS error is 68 μm for x and 56 μm for y when predicting positions from validation samples.

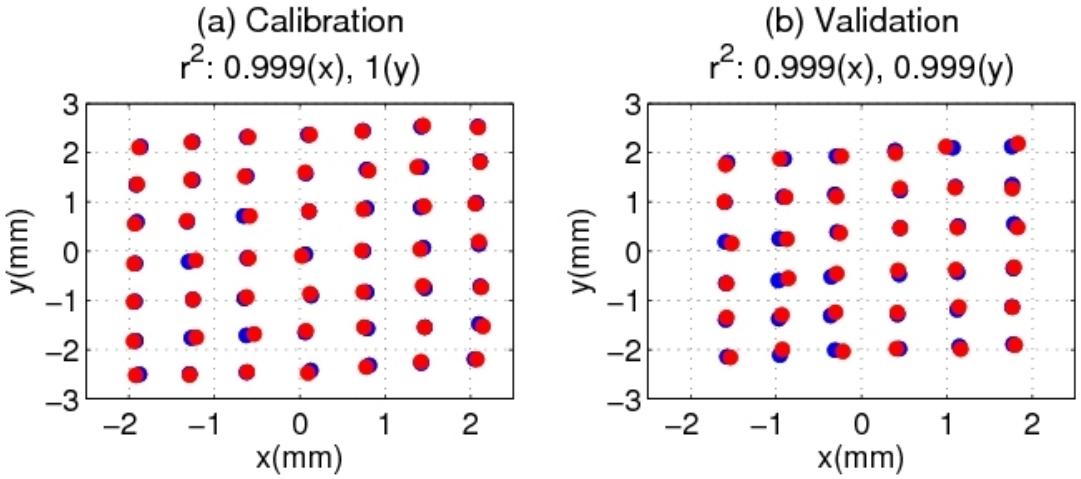

Figure 13: Measurement (blue) and prediction (red) of the transverse beam position from calibration and validation samples. The method applied is SVD with the first 26 SVD modes.

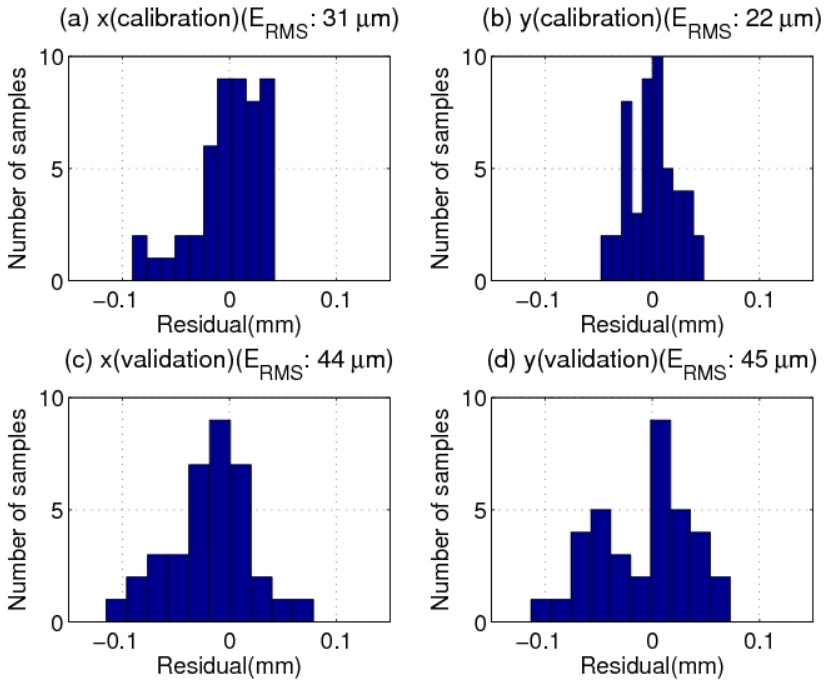

Figure 14: Difference of measured and predicted transverse beam position from calibration and validation samples. The method applied is SVD with the first 26 SVD modes.

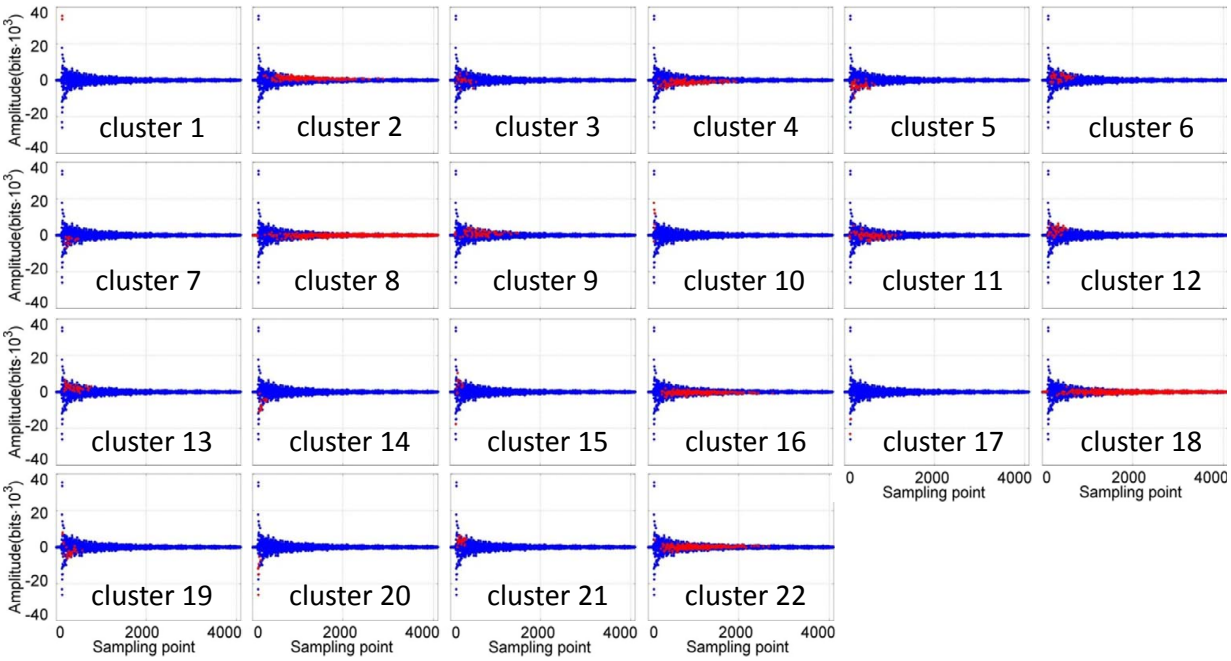

Figure 15: 22 clusters (red) partitioned from the calibration samples (blue) for one beam position.

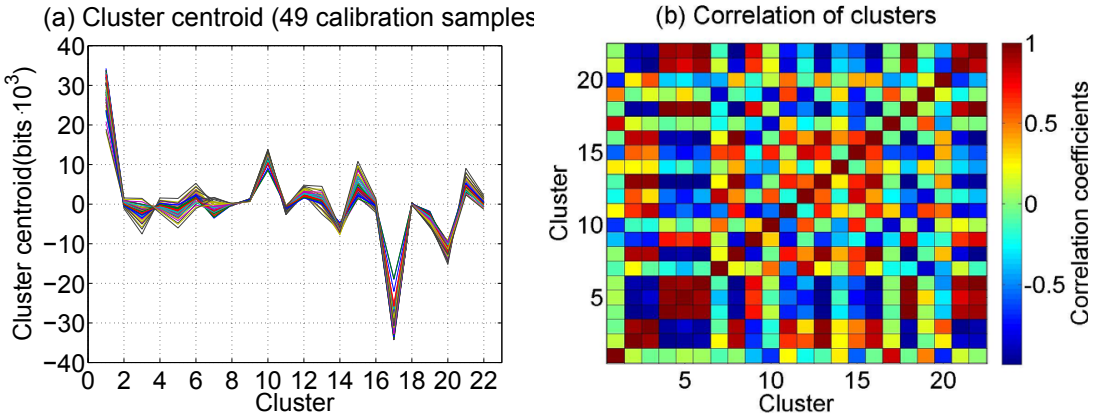

Figure 16: Cluster centroids and their correlations.

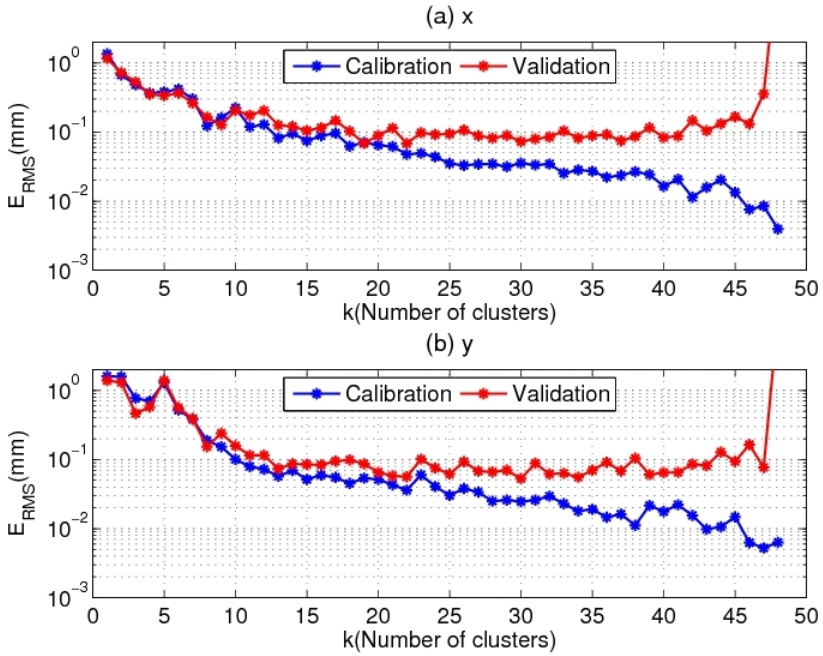

Figure 17: Different number of clusters to determine the transverse beam position x and y measured by E_{RMS}.

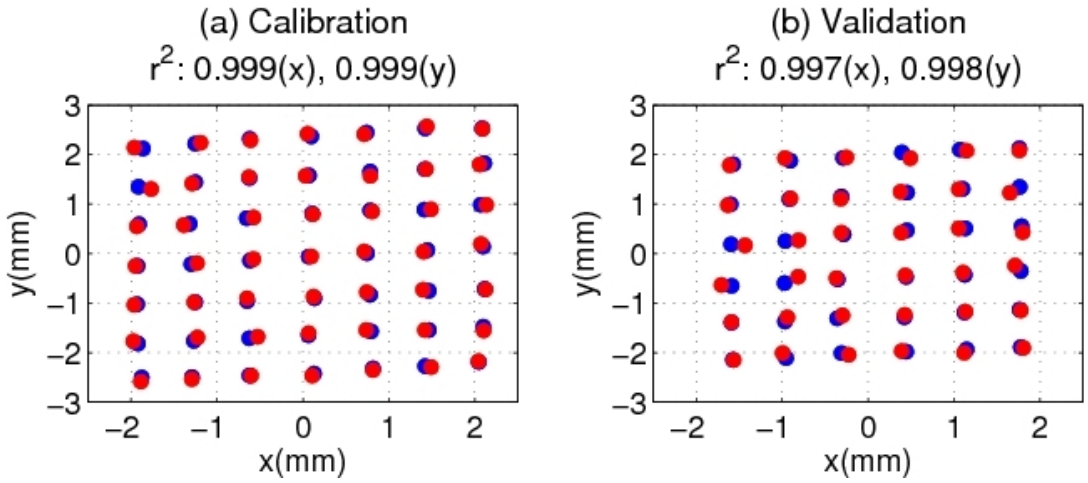

Figure 18: Measurement (blue) and prediction (red) of the transverse beam position from calibration and validation samples. The method applied is k-means clustering with 22 clusters.

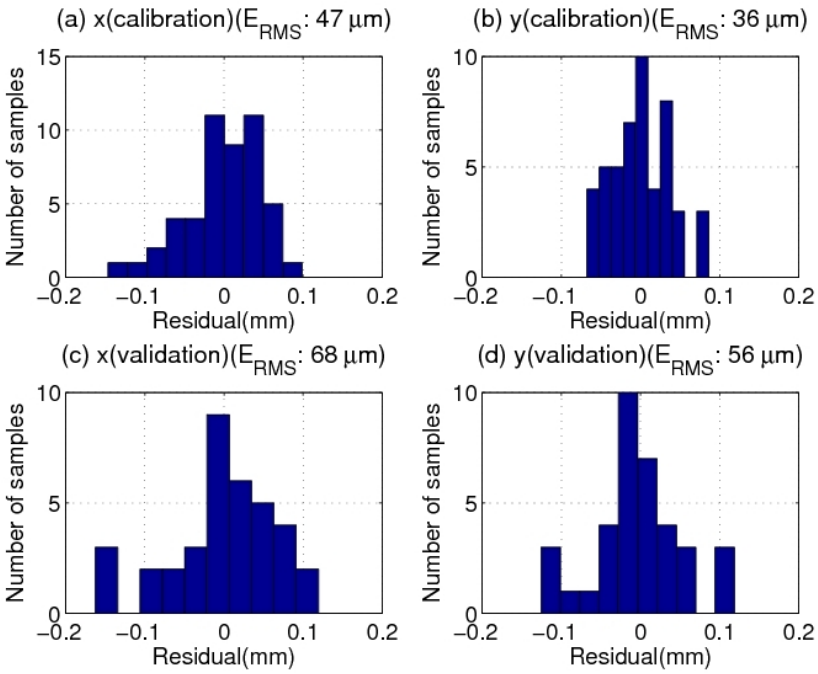

Figure 19: Difference of measured and predicted transverse beam position from calibration and validation samples. The method applied is k-means clustering with 22 clusters.

3 Performance Evaluation

The performance of the three methods used to extract the beam position described in the previous section is evaluated in this section by two different techniques: the fixed sample split and cross-validation. A figure of merit, position resolution, is defined as the RMS error when predicting the validation samples (Eq. 5).

3.1 Fixed sample split

For the fixed sample split shown in Fig. 2(b), the position resolutions obtained from the three different methods are listed in Table 1 along with the calibration RMS errors. The RMS errors are comparable for DLR and SVD of both calibration and validation samples. Both methods suggest an approximately 50 μm position resolution. The k-means clustering suggests a worse resolution of 60–70 μm.

Table 1: Direct comparison of DLR, SVD and k-means clustering for the fixed sample split.

	Calibration		Validation	
	E_{RMS} (x)	E_{RMS} (y)	E_{RMS} (x)	E_{RMS} (y)
DLR	30 μm	25 μm	40 μm	50 μm
SVD	31 μm	22 μm	44 μm	45 μm
Clustering	47 μm	36 μm	68 μm	56 μm

3.2 Cross-validation

Until now, our analysis is based on a specific sample split (Fig. 2(b)). To remove the sample dependence, a technique named *cross-validation* is used [12]. As our total sample size is quite small (85 samples), we use the technique called *leave-one-out cross-validation* (LOOCV) [12]. For each sample split, LOOCV uses only one sample from the total samples as validation, and the remaining 84 samples are used for calibration. This is repeated 85 times such that each sample out of all the samples is used once for validation. The RMS error is then calculated on the residuals for the validation sample from all 85 different sample splits.

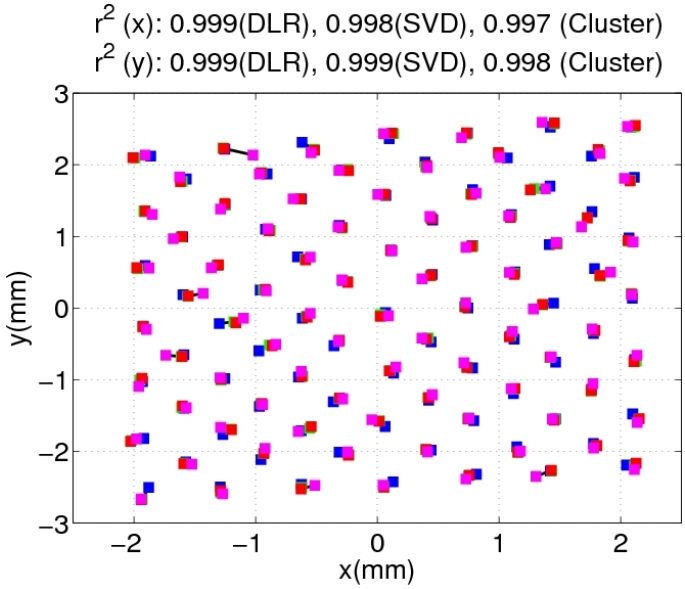

Figure 20: Measurement (blue) and prediction of the transverse beam position from each of 85 validation samples using LOOCV. The DLR is in red. The SVD is in green with the first 26 SVD modes. The k-means clustering is in magenta with 22 clusters. Points connected with black lines belong to the same beam position.

Fig. 20 shows the measurement and the prediction of each sample when it is used for validation in that specific sample split. The methods applied are DLR (red), SVD (green) with the first 26 SVD modes, and

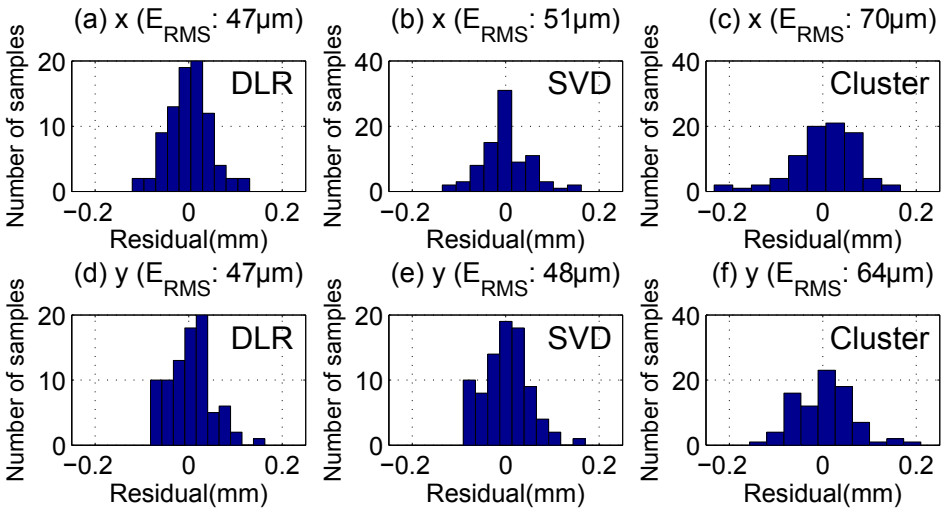

Figure 21: Difference of measured and predicted transverse beam position from each of 85 validation samples using LOOCV.

k-means clustering (magenta) with 22 clusters. Fig. 21 shows the corresponding residuals for these three methods. The sample-independent RMS errors are listed in Table 2. Since the calibration samples for each of the 85 sample splits are similar, the RMS error for the validation sample from all 85 sample splits is a good estimation of the position resolution. The results are similar to the fixed sample split case shown in Table 1. Although the results are comparable for all three methods, SVD and k-means clustering are far more efficient than DLR in computing calibration coefficients, since the number of unknown variables is much smaller, as shown in Table 2. However, once the calibration matrix M is obtained, the time required for position prediction is generally equally fast.

Table 2: Direct comparison of DLR, SVD and k-means clustering using cross-validation.

	E_{RMS} (x)	E_{RMS} (y)	Number of unknowns
DLR	47 μm	47 μm	4097
SVD	51 μm	48 μm	27
Clustering	70 μm	64 μm	23

4 Summary and Conclusions

We have studied various methods to extract the transverse beam position from the HOMs excited by an electron beam in a third harmonic superconducting cavity. A linear system has been composed from HOM waveforms in the fifth dipole band with frequencies in the range 9040–9078 MHz and measured beam positions. DLR on the entire waveforms suggests a position resolution of 40 μm for x and 50 μm for y. Due to the large number of sampling points occurred from the waveforms, the system has a considerable number of unknown variables to be determined. This leads to a long computation time in the calibration phase. Therefore, two methods are used to reduce the dimension of the system effectively: SVD and k-means clustering. SVD gives a comparable position resolution to DLR of 44 μm for x and 45 μm for y, while k-means clustering provides a reasonable resolution of 68 μm for x and 56 μm for y. In order to eliminate the sample dependence, a cross-validation technique has been applied with these three methods. The RMS prediction error of approximately 50 μm can be achieved by both DLR and SVD, while a 30%–40% worse precision is suggested by the k-means clustering.

In this study, SVD is seen to be an effective and computation-efficient method to reduce the system dimension while preserving good position resolutions in both x and y. Therefore it has been decided to implement SVD into the final HOM electronics for beam position diagnostics in third harmonic cavities for FLASH and the European XFEL.

Acknowledgements

We thank Nathan Eddy, Brian Fellenz and Manfred Wendt who developed the test electronics and for useful discussions on the data analysis. We are also grateful to members of MEW group from the University of Manchester for enlightening ideas and to the FLASH crew for supporting the measurements. This work was partially funded by the European Commission under the FP7 Research Infrastructures grant agreement No.227579.

References

[1] W. Ackermann *et al.*, "Operation of a free-electron laser from the extreme ultraviolet to the water window," *Nature Photonics*, vol. 1, pp. 336–342, 2007.

[2] J. Sekutowicz, *Multi-cell Superconducting Structures for High Energy e+e- Colliders and Free Electron Laser Linacs*. Warsaw, Poland: Warsaw University of Technology Publishing House, first ed., 2008.

[3] K. Flottmann, T. Limberg and Ph. Piot, "Generation of Ultrashort Electron Bunches by Cancellation of Nonlinear Distortions in the Longitudinal Phase Space," TESLA Note: TESLA-FEL 2001-06, 2001.

[4] J. Sekutowicz, R. Wanzenberg, W.F.O. Müller and T. Weiland, "A Design of a 3rd Harmonic Cavity for the TTF 2 Photoinjector," TESLA Note: TESLA-FEL 2002-05, 2002.

[5] K.L.F. Bane, "Wake Field Effects in a Linear Collider," SLAC Note: SLAC-PUB-4169, 1986.

[6] S. Molloy *et al.*, "High precision superconducting cavity diagnostics with higher order mode measurements," *Phys. Rev. ST Accel. Beams*, vol. 9, p. 112802, 2006.

[7] P. Zhang, N. Baboi, R.M. Jones, I.R.R. Shinton, T. Flisgen and H.W. Glock, "A study of beam position diagnostics using beam-excited dipole modes in third harmonic superconducting accelerating cavities at a free-electron laser," *Rev. Sci. Instrum.*, vol. 83, p. 085117, 2012.

[8] P. Zhang, N. Baboi and R.M. Jones, "Eigenmode simulations of third harmonic superconducting accelerating cavities for FLASH and the European XFEL," DESY Report: DESY 12-101, 2012.

[9] P. Zhang, N. Baboi, R.M. Jones and I.R.R. Shinton, "Study of Beam Diagnostics with Trapped Modes in Third Harmonic Superconducting Cavities at FLASH," in *Proceedings of IPAC2011*, no. THPPA00, (San Sebastian, Spain), pp. 2891–2893, 2011.

[10] P. Zhang, N. Baboi N. Eddy, B. Fellenz, R.M. Jones, B. Lorbeer, T. Wamsat and M. Wendt, "A Study of Beam Position Diagnostics with Beam-excited Dipole Higher Order Modes using a Downconverter Test Electronics in Third Harmonic 3.9 GHz Superconducting Accelerating Cavities at FLASH," DESY Report: DESY 12-143, 2012.

[11] N. Baboi *et al.*, "Higher Order Modes for Beam Diagnostics in Third Harmonic 3.9 GHz Accelerating Modules," in *Proceedings of SRF2011*, no. MOPO060, (Chicago, USA), pp. 239–243, 2011.

[12] C.M. Bishop, *Pattern Recognition and Machine Learning*, ch. 3.1.4. Springer Publishing, first ed., 2006.

[13] G.H. Golub and C.F. Van Loan, *Matrix Computations*, ch. 2.1, p. 12. The John Hopkins University Press, second ed., 1984.

[14] R.L. Ott and M. Longnecker, *An Introduction to Statistical Methods and Data Analysis*, ch. 11.7. Duxbury Press, sixth ed., 2008.

[15] M.R. Spiegel, J.J. Schiller and R.A. Srinivasan, *Schaum's Outline of Theory and Problems of Probability and Statistics*, ch. 8, p. 281. McGraw-Hill, second ed., 2000.

[16] G.H. Golub and C.F. Van Loan, *Matrix Computations*, ch. 2.3, pp. 16–20. The John Hopkins University Press, second ed., 1984.