

Progress Towards Improved Analysis of TES X-ray Data Using Principal Component Analysis

S. E. Busch¹ · J. S. Adams^{1,2} · S. R. Bandler^{1,3} · J. A. Chervenak¹ · M. E. Eckart^{1,2} · F. M. Finkbeiner^{1,4} · D. J. Fixsen¹ · R. L. Kelley¹ · C. A. Kilbourne¹ · S.-J. Lee¹ · S. H. Moseley¹ · J.-P. Porst^{1,5} · F. S. Porter¹ · J. E. Sadleir¹ · S. J. Smith^{1,2}

Received: 14 September 2015 / Accepted: 22 October 2015 / Published online: 6 November 2015
© Springer Science+Business Media New York (outside the USA) 2015

Abstract The traditional method of applying a digital optimal filter to measure X-ray pulses from transition-edge sensor (TES) devices does not achieve the best energy resolution when the signals have a highly non-linear response to energy, or the noise is non-stationary during the pulse. We present an implementation of a method to analyze X-ray data from TESs, which is based upon principal component analysis (PCA). Our method separates the X-ray signal pulse into orthogonal components that have the largest variance. We typically recover pulse height, arrival time, differences in pulse shape, and the variation of pulse height with detector temperature. These components can then be combined to form a representation of pulse energy. An added value of this method is that by reporting information on more descriptive parameters (as opposed to a single number representing energy), we generate a much more complete picture of the pulse received. Here we report on progress in developing this technique for future implementation on X-ray telescopes. We used an ⁵⁵Fe source to characterize Mo/Au TESs. On the same dataset, the PCA method recovers a spectral resolution that is better by a factor of two than achievable with digital optimal filters.

Keywords Low-temperature detectors · Transition-edge sensors (TES) · Principal component analysis

✉ S. R. Bandler
Simon.R.Bandler@nasa.gov

¹ NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

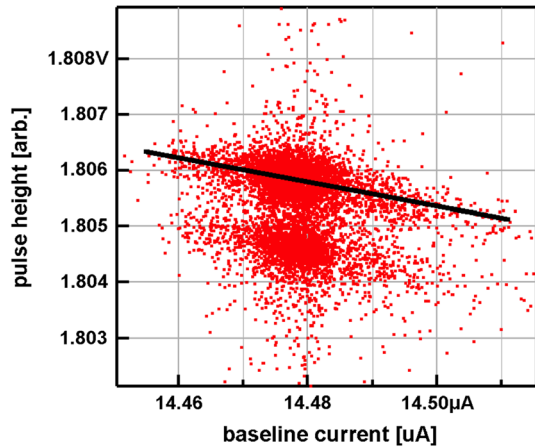
² CRESST and University of Maryland, Baltimore County, MD 21250, USA

³ CRESST and University of Maryland, College Park, MD 20742, USA

⁴ Wyle Information Systems, McLean, VA 22102, USA

⁵ Physics Department, Brown University, Providence, RI 02912, USA

Fig. 1 This plot shows a correlation between optimally filtered pulse height and baseline level. The results are from a TES with X-rays originating from an ^{55}Fe source. The line is a linear fit to the data. Such correlations motivate us to consider that other properties of pulses are changing (Color figure online)



1 Introduction

A standard approach for measuring the energy of an X-ray photon is to apply a digital optimal filter [1]. Optimal filtering provides the most accurate estimate of photon energy only if the noise does not change during the pulse (stationary noise), and if the signals from different energies have the same pulse shape (linear response). In certain cases for our TESs, these conditions may not be met [2] and the energy resolution can thus be significantly degraded. We present an implementation of an approach discussed in previous work [3–5], based on principal component analysis (PCA). The goal of this is to improve energy resolution in the non-linear, non-stationary regime.

For our type of detector, PCA uses the eigenvectors of a pulse-pulse covariance matrix, where the eigenvectors are the principal components of the information in the signal. It enables the weighting of different parts of the signal within a pulse at different times, according to the ratio of signal-to-noise of the different components during different times. For the case of stationary noise and a linear response with energy, it should be equivalent to the conventional optimal filter. This provides a more robust estimate of the photon energy under non-linear, non-stationary conditions. The successful implementation of such an algorithm will enable use of TES detectors with, for example, low heat capacity, with improved energy resolution over a large range of energies. This approach also extracts information about (and may enable correction for) event-to-event variations in the pulse shape that may arise due to trigger-jitter, position-dependence of the X-ray absorption producing a rise-time variation, or gain drift due to changing environment conditions such as temperature or magnetic field. An example of the pulse height changing with baseline level is shown in Fig. 1.

2 Principal Component Analysis

For our experiment, we measure an ensemble of pulses from an ^{55}Fe source. We select out single-pulse records (not pile-up pulses) around the MnKa line for the following

Fig. 2 *Top* Covariance matrix that is formed from the ensemble of pulses. The units of the color-coded legend in amperes². *Bottom* An example of current pulse from a TES. The abscissa represents time (sampling rate is 1 Ms/s). The vertical axis measures current through the TES (Color figure online)

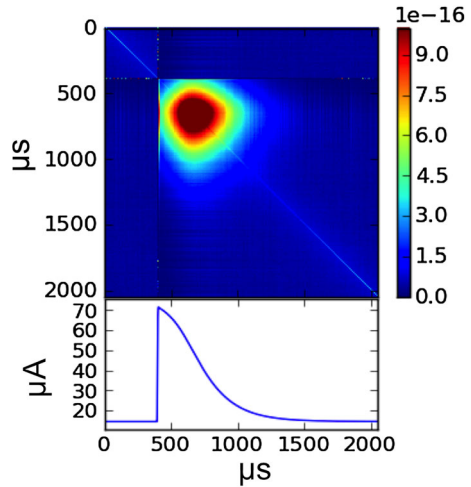
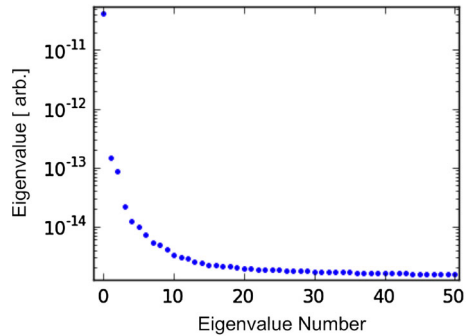


Fig. 3 The spectrum of the largest 50 eigenvalues of our covariance matrix (Color figure online)



analysis, with no separate noise traces included. The fraction of piled-up pulses was small, less than 5 %. We create a variance–covariance matrix of each pulse, essentially creating the matrix from the product of each point in a trace with each other point. For a set of X-ray events, we define a matrix \mathbf{x} ; the matrix elements are x_{ij} where each element represents the j th pulse at the i th time in that pulse. We then create a mean pulse $(1/M)\sum_j x_{ij} = \langle x_i \rangle$, where M is the number of pulses. The mean pulse is used to create \mathbf{D} , the matrix of the residuals of each pulse from the mean: $\mathbf{D}_{ij} = x_{ij} - \langle x_i \rangle$. We then calculate the covariance of the residual matrix $\text{cov}(\mathbf{D}) = \mathbf{D}\mathbf{D}^T$. An example of the covariant matrix is depicted in Fig. 2, and the eigenvalues from this matrix is shown in Fig. 3.

We now want to transform the data to a basis where the covariance matrix is diagonal. We, therefore, determine the eigenvectors of the covariance matrix $\text{cov}(\mathbf{D}) = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}$, where $\mathbf{\Lambda}$ is the diagonal matrix of eigenvalues sorted from highest to lowest eigenvalues, and \mathbf{Q} is the matrix of eigenvectors. We then rotate each pulse into the basis of eigenvectors $\mathbf{R} = \mathbf{Q}^T\mathbf{D}$. The fundamental task at hand is to determine which eigenvectors of the average covariance matrix are responsible for the properties of the pulse with the greatest photon energy information. We identify these as

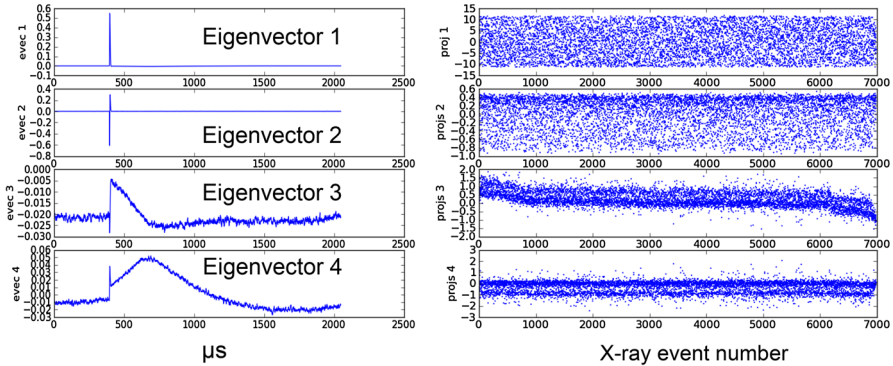


Fig. 4 Left column shows normalized eigenvectors plotted against time. The distribution in this panel helps us to identify what aspect of the pulse each vector is associated with. Right column shows the projection of all pulses onto the corresponding (left) eigenvector (Color figure online)

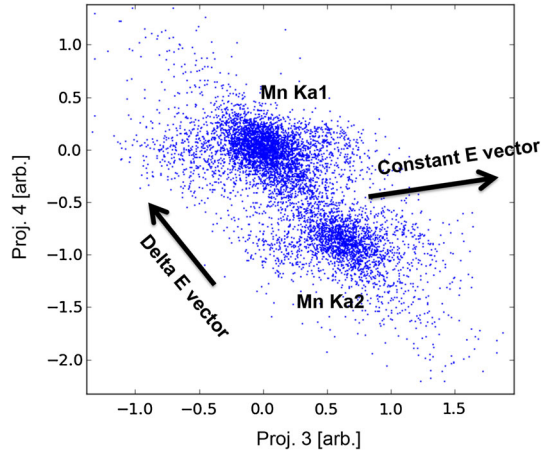
those eigenvectors with the largest eigenvalues in the covariance matrix. The logarithm of the eigenvalue is a measure of the information in the pulse of the corresponding eigenvector.

In Fig. 4, we show an example of this procedure. The third and fourth eigenvectors (left column) have a shape that is reminiscent of the pulse shape. The components of each pulse in our dataset, projected onto these eigenvectors (right column of Fig. 4), show a distribution similar to the pulse height distribution we expect from the Mn $K\alpha$ 1 and 2 lines. The value of eigenvector 3 is completely negative, and conversely, eigenvector 4 has points both positive and negative. These features indicate that eigenvector 4 is likely more sensitive to the pulse height, and eigenvector 3 is more sensitive to the baseline, analogous to that shown in Fig. 1. In the first and second eigenvectors, most of the information is coincident with the pulse arrival. Also we find non-Gaussian distributions of the rotated projections, as is clearly evident in the right panels of vectors 1 and 2 in Fig. 4. Structures like this are indicative of arrival time variations. The fact that the leading eigenvector does not look like a pulse is not problematic. This purely implies that for this dataset, most of the information available is related to arrival time. Our PCA method independently determines corrections that we typically use in our standard optimal filtering processing method, but without a priori knowledge of these variations, or being limited to those assumptions.

In Fig. 3, we show the spectrum of eigenvalues. It is clear that the vast majority of the magnitude in the covariance matrix is coming from the first three or four eigenvectors. We therefore designate these as our principal components. Thus, for this example, it has been determined that each X-ray can now be described by four numbers, which are the first four points in the eigenvector basis. With the knowledge of these principal components of the pulse, the next step is understanding how these components or their combination correlate with energy. This allows us to determine more accurately how the information from the covariance matrix can be used to improve the characterization of photon energy.

Figure 5 shows projection 3 versus projection 4, that is to say we take the data from the right panel of eigenvector 3 (in Fig. 4) and plot it against the data from the

Fig. 5 The projection of each pulse onto eigenvector 3 is plotted against the projection of each pulse onto eigenvector 4. The arrows indicate, as labeled, the constant energy and the perpendicular ΔE vector. We also label the two Mn $K\alpha$ lines (Color figure online)



right panel of eigenvector 4. From observation of the correlations in this plot, we determine that the two over-densities of points trace diagonal lines which lie in the vector direction of constant energy. This indicates that they are consistent with the two Mn $K\alpha$ lines. We can therefore optimize the energy resolution of our system by determining what combination of these eigenvectors gives the maximum separation in the Mahalanobis [6] coordinate space, a space that provides a relative measure of a data point's distance from a common point. When the eigenvectors are the basis set, as is the case here, then the Mahalanobis distance is simply the Euclidian distance. We illustrate this maximization of the separation with two arrows. The constant E vector represents the direction of constant energy, and the delta E vector represents the perpendicular direction. Since we know the energy difference between these two lines, we can use this information to calibrate our gain scale. When a broad spectrum of energies needs to be measured, including regions of continuum, it is necessary to calibrate the pulse response as a function of energy in separate measurements, such that the increments are small enough that pulse-shape variations can be accurately approximated by interpolation. A study of the response of a detector as a function of energy, such as from the 3 eV separations that can occur when absorbing pulsed photons from a laser diode [7], could be ideal for this. This type of calibration allows us to know the direction of increasing energy, and the surfaces of constant energy. In general, this analytic approach rigorously optimizes the energy resolution from the study of one projection if the variation in pulse data is purely due to energy deposited in the detector. When there are other contributions to the pulse variation, such as in the dataset here, then it is more complicated to determine the optimal combination of projections, and a rigorous generalized procedure will require further research.

3 PCA Method Improves Measured Resolution

Here we compare the results of PCA analysis with the basic optimal filtering on one particular dataset [2]. This device demonstrated a full-width at half maximum

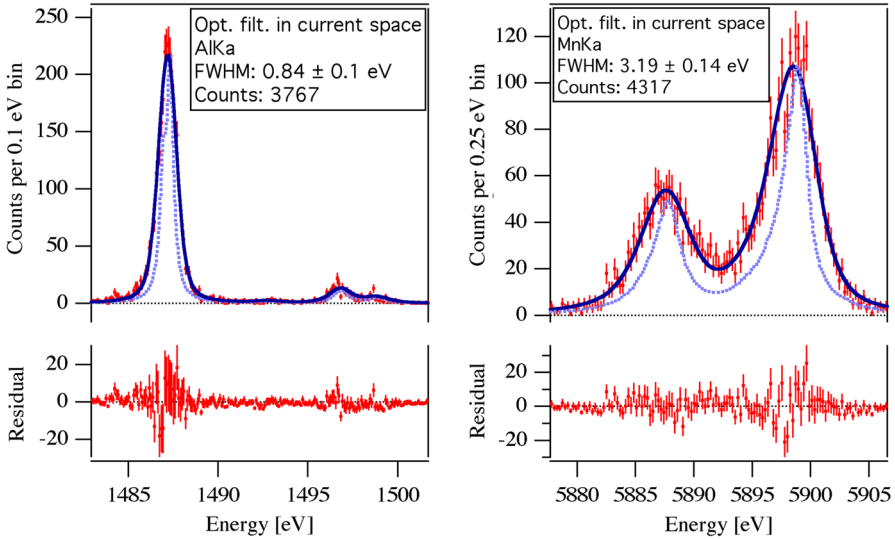


Fig. 6 (left) Al $K\alpha$ spectrum determined using a standard optimal filter. (right) For the same device, Mn $K\alpha$ spectrum processed using a standard optimal filter. The red lines indicate data, the center indicates the number of counts, and the length of the line indicates the statistical error. The light blue line represents the intrinsic line shape, and dark blue line represents a fit to the data and is the convolution of the intrinsic line shape with a Gaussian of the FWHM stated above each panel. Bottom panels show residuals from the fit (Color figure online)

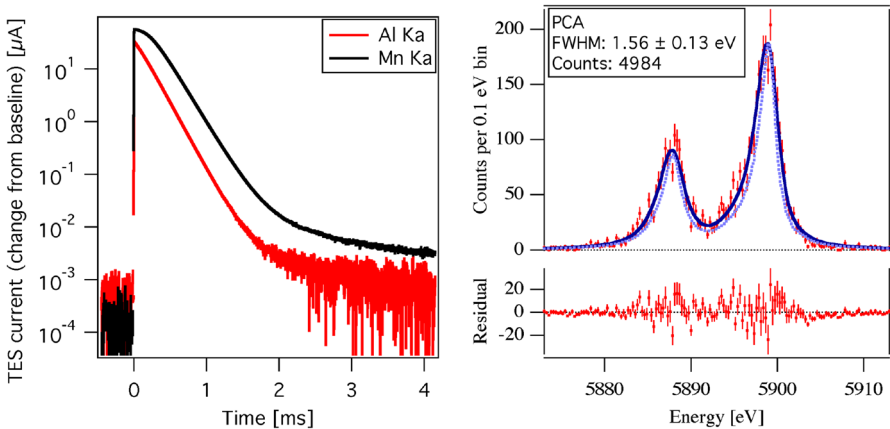


Fig. 7 (left) Pulse shapes for Al $K\alpha$ and Mn $K\alpha$, illustrating a saturated response for the higher energy. (right) Mn $K\alpha$ spectrum processed using the PCA technique, showing dramatic improvement in FWHM. This spectrum uses the same data as is shown in Fig. 6 (right) (Color figure online)

(FWHM) of 0.9 eV at 1.5 keV, using optimal filter for the processing (Fig. 6 left). For the same device and processing technique, but constructing the filter based upon the 6 keV pulse shape and assuming a linear response, we achieve a FWHM of 3.2 eV at 6 keV (Fig. 6 right). Figure 7 left shows the two pulse shapes for the different energies, illustrating a saturated response for the higher energy. Figure 7 right uses the same

6 keV data but processing with PCA. In this example, the PCA method demonstrates a factor of two times improvement over optimal filter techniques, yielding a FWHM of 1.6 eV. We note that this result is the best case that we have measured. The level of improvement that is observed will depend upon how much the assumptions of optimal filter techniques are violated. Because one does not know a priori if the assumptions required for optimal filter methods are violated, methods, such as ours, that provide greater flexibility seem to be necessary to recover the best energy resolution. For example, reprocessing the 1.5 keV data with the PCA method did not improve the measured resolution.

4 Conclusion

We have demonstrated that the use of a PCA method to analyze data from X-ray pulses on TESs has the capacity to significantly improve the energy resolution. We outlined how this procedure can be implemented. To make this procedure broadly useful, it will be necessary to develop prescriptions for selecting an appropriate number of principal components, and for weighting the components to provide the best possible energy estimator. Future efforts will be put into determining how to automate these methods, and thus facilitate implementation in real instruments, and how it would be implemented in a real-time processing situation. This general approach has great potential for yielding improved spectroscopy in the future.

Acknowledgments This research was supported in part by an appointment to the NASA Postdoctoral Program at Goddard Space Flight Center (S.E. Busch, S.-J. Lee), administered by Oak Ridge Associated Universities through a contract with NASA.

References

1. Moseley et al., IEEE Trans. Nucl. Sci. **35**, 59–64 (1988)
2. S.R. Bandler et al., IEEE Trans. Appl. Supercond. **23**(3), 705–2100 (2013)
3. D.J. Fixsen, S.H. Moseley, B. Cabrera, E. Figueroa-Feliciano, AIP Conf. Proc. **605**, 339–342 (2002)
4. D.J. Fixsen, S.H. Moseley, B. Cabrera, E. Figueroa-Feliciano, Nucl. Instrum. Methods Phys. Res. A **520**, 555–558 (2004)
5. C.H. Whitford, Nucl. Instrum. Methods Phys. Res. A **555**, 255–259 (2005)
6. P.C. Mahalanobis, Proc. Natl. Inst. Sci. India **2**(1), 49–55 (1936)
7. D.J. Fixsen et al., J. Low Temp. Phys. **176**(1), 16–26 (2014). doi:[10.1007/s10909-014-1149-x](https://doi.org/10.1007/s10909-014-1149-x)