Principal Component Analysis: Reveal Camouflaged Information in X-Ray Absorption Spectroscopy Photoemission Electron Microscopy of Complex Thin Oxide Films

Margret Giesen[1*], Matteo Jugovac[1], Giovanni Zamborlini[1], Vitaliy Feyer[1], Felix Gunkel[2] and David N. Mueller[1]

[1]*Peter Grünberg Institute (PGI-6), Jülich Forschungszentrum GmbH, 52425 Jülich, Germany*

[2]*Institute of Electronic Materials (IWE2), RWTH Aachen University, 52074 Aachen, Germany*

Keywords:

PEEM, XAS, spectromicroscopy, principal component analysis, PCA, multivariate techniques

Abstract:

Principal component analysis (PCA) has become a standard tool in spectromicroscopy and hyperspectral imaging to handle large spectral data sets and to decompose raw data into relevant and residual information. In particular in studies of complex compounds, PCA can be used to disentangle chemical information and thereby deepen the understanding of chemical and physical material properties. Surprisingly, in photoemission electron spectromicroscopy (PEEM), PCA is rarely used. This paper serves to demonstrate how powerful PCA can be to detect hidden chemical information in PEEM data. We demonstrate the capability of PCA in PEEM spectromicroscopy for the case of a thin film of a complex quaternary oxide, $Pr_{0.5}Ba_{0.5}CoO_{3-\delta}$ (PBCO) which is a main contender catalyst material for electrocatalytic water splitting. Upon annealing in air, PBCO decomposes into different phases at the surface. Two of them become obvious from the raw PEEM images, but one is revealed only after PCA.

*Corresponding author: email m.giesen@fz-juelich.de

1. Introduction

Multivariate analysis (MVA) techniques such as Principal Component Analysis (PCA) [1, 2] are very powerful statistical tools which have been used for a long time to reduce the dimensionality of data sets and to analyze spectromicroscopic measurements [3]. PCA decomposes spectral data into new components which represent the largest variance in the data and into other components which form the residual [4-6]. In typical spectromicroscopic data, relevant spectral components that make for about 95% of the variance may account for merely $\sim 10^{-3}$ % of the total data! Therefore, relevant spectral information might be camouflaged under the sheer amount of data with low variance and might not easily be detected.

Given large spectral data sets, interactive analysis becomes tedious and time-consuming; even more critical: in order to handle such large data sets, artificial pre-assumptions on the possible outcome of the analysis may become inevitable. Such pre-assumptions, however, might prevent the observation of unexpected results since the latter may be filtered out in order to handle the large data set. PCA on the other hand, being a proper ab initio method, is for the most part free of artificial pre-assumptions and safely detects all components that explain the largest variances in raw spectral data.

Due to these benefits, statistical methods such as PCA have influenced a wide range of scientific fields and applications [7-9]. The availability of open source as well as commercial software packages support the broad application of PCA and cluster analysis to spectromicroscopic and hyperspectral data [10].

Among the spectroscopy data analyzed using PCA, x-ray photoelectron emission (XPEEM [11]) work is scarce though: An early extended x-ray absorption fine structure (EXAFS)

analysis of Mo/TiO_2 catalysts was described by Fay et al. in 1992 [12]. Tutorial description of PCA applied to XAS (x-ray absorption spectroscopy) and EXAFS studies were published a couple of years later by Wasserman et al. [13, 14]. Multivariate techniques with compositional and lateral resolution were applied to PEEM studies of Ge oxide surfaces [15-17]. Ag, Ni, Ti metallization layers [18] and Au-Sn surfaces [19] were studied using PCA by Walton et al. and Renault et al.. The latter authors used PCA exclusively to detect probably hidden information in the XPS (x-ray photoelectron spectroscopy) data; however, they did not perform a full spectral analysis of their data using PCA [19]. In 2015, Strelcov et al. published studies on complex oxides where they analyzed the spatial variability of electronic transport in $BiFeO_3$-$CoFe_2O_4$ [20]. However, these authors focused on the local electronic transport properties rather than on the chemical aspects or phase domain studies of complex oxides. Very recently, Yang et al. showed using PCA that praseodymium Pr^{3+} accumulates in $SrTiO_3$ grain boundaries [21]. A year later, Hannula et al. employed PCA exclusively to reduce noise in XPS PEEM data of ultra-thin TiO_2 films [22]. Finally, Guo et al. used PCA to determine the reasonable number of components that should be used in peak fitting procedure rather than detecting relevant components directly from PCA [23].

This paper demonstrates that PCA is not restricted to serve as an accessory tool in PEEM spectroscopy on thin solid films. On the contrary, PEEM spectromicroscopy data may be analyzed with high precision based on PCA exclusively. As an example, we report on the analysis of XAS PEEM images of thin $Pr_{0.5}Ba_{0.5}CoO_{3-\delta}$ (PBCO) ordered double-perovskite compound films. PBCO is a mixed conducting oxide which, due to its electronic properties and facile reducibility is proposed as a next generation catalyst for water splitting [24]. It belongs to the family of late 3d transition metal (TM) perovskites known for their highly covalent TM-O bond [25-28]. This feat, also responsible for their fascinating electrochemical

properties (which are not focus of this work though), leads to an information rich X-ray absorption spectroscopy (XAS), especially around the O-K-Edge. The O-K-XAS is a powerful tool to assess the electronic structure in the vicinity of the Fermi level of perovskites. Electronic states probed at the O-K-edge are participating in electrochemical reactions. Their evolution during electrochemical reactions gives valuable insights into atomistic processes [25, 29]. Unfortunately, most studies so far have been carried out in their spatially averaged form, i.e. sample surfaces were considered to be perfectly homogenous at all times notwithstanding thermodynamic conditions. As a consequence, local changes in chemical surface properties based for instance on cation segregation and precipitation are either completely overlooked or cannot be resolved properly. Cation segregation and precipitation, however, are ubiquitous in complex oxides exposed to technologically relevant thermodynamic conditions [30, 31] which in turn will alter the electronic structure locally and with that the reaction mechanisms.

In the case of thin films deposited by pulsed laser deposition, as in this case, it should be noted that due to the typically low process pressure and the involved non-equilibrium growth dynamics [32] the highly reducible oxides are usually grown in a more reduced state than their application counterpart. As the formal oxidation states of Co as well as the concomitantly altered oxygen vacancy concentration are governing factors of the aforementioned applications, it is worthwhile to investigate the materials reaction to exposure to thermodynamic conditions typical for solid state ionics and industrial processing of the catalysts. Though alterations beyond pure oxidation and reduction (the oxygen exchange) may be identified with spatially averaging techniques, especially the more information rich XAS does not give clear indication of e.g. phases as XPS does [31]. Other spatially resolving techniques such as atomic force or electron microscopy do not yield

chemical contrast. The X-PEEM approach in conjunction with sophisticated analysis techniques as presented here yields the necessary spatial and chemical contrast required for a thorough understanding of surface processes of complex oxides. In summary, PBCO as a rather complex oxide film offers itself to be a quite intuitive example to demonstrate the strength of PCA applied to PEEM data.

Here, we report on the observation of a surface decomposition of PBCO films into a Pr lean and Ba rich, and a Ba lean and Pr rich phase after elevated temperature anneal in oxidizing atmosphere, indicating a miscibility gap of the parent material as has been observed before in similar complex perovskites [33]. Details of the physical and chemical implications of this complex domain pattern in terms of proper defect chemistry and thermodynamics are beyond the scope of this article. Here, we focus on how powerful the method can be even in studies of complex oxides. As a further reference for the work at hand, in terms of bulk solid state chemistry, the process can be grossly simplified to:

$$2\ \mathrm{Pr_{0.5}Ba_{0.5}CoO_3} \rightleftharpoons \mathrm{PrCoO_3} + \mathrm{BaCoO_3} \qquad (1)$$

which showcases that the demixing of the A-site being constituted of aliovalent cations (Ba^{2+} and Pr^{3+}) will result in a change of the overall oxidation state of Co, or better the CoO_6 building block due to the rather covalent nature of the late 3d transition metal perovskites. This, of course, neglects the occurrence of any oxygen vacancies, which however are inconsequential for the work presented here. It should be noted, though, that in the late 3d transition metals (TM) due to the negative charge transfer, any charge fluctuation in the TMO_6 building block is observable in the O-K-Edge XAS [25], as will be been seen here.

Here, we focus on the technical aspects of the analysis of the PEEM images and the insights gathered towards the decomposition process. We demonstrate that PCA is a fast analysis method with high spectroscopic resolution and - in contrast to manual methods - not susceptible to false assumptions during analysis.

2. Materials and Methods

2.1. Experimental

Highly oriented thin films of $Pr_{0.5}Ba_{0.5}CoO_{3-\delta}$ (in the following denoted as PBCO) were deposited on Nb doped $SrTiO_3$ (100) substrates as described elsewhere [34]. Crystallinity and formation of the cubic perovskite phase were confirmed by X-ray diffraction.

The pristine thin film sample did not show any obvious surface structure. Preliminary studies of our group indicate that the domain structure evolves with sample annealing. Results are, however, beyond the scope of this paper and will be subject of an upcoming future publication.

In the study presented here, the pristine sample was subjected to a 12h anneal at 1073 K in a stagnant atmosphere of 2×10^4 Pa O_2. After annealing, it was quenched to room temperature in the pertaining atmosphere and transferred to the microscope chamber under UHV conditions.

PEEM data considered in this paper was collected at the NanoESCA beamline of Elettra [35], in a UHV system with a base pressure of $5x10^{-9}$ Pa, using an electrostatic PEEM. PEEM images have a circular field of view (FoV) with an approximate diameter of 10 µm. The total size of the square raw image is 600 × 600 pixels, corresponding to a length of 18 nm per image pixel.

In X-ray absorption spectroscopy (XAS), PEEM images display the secondary electron emission for a distinct kinetic electron energy and the incident photon energy is scanned in order to obtain spatially resolved X-ray absorption spectra. The data set for a certain energy range is therefore a stack of single images, each recorded at a distinct photon energy.

Spectromicroscopic images were recorded and analyzed in a spectral range between (A) 526.0 and 550.0 eV, (B) 772.0 and 809.0 eV and (C) 926.0 and 952.0 eV with an energy step of 0.2 eV for all three regimes. The full data set therefore contained stacks including (A) 121, (B) 185 and (C) 130 images. Regime (A) covers the oxygen K-edge, (B) the Co $L_{3,2}$- and the Ba $M_{5,4}$-edges and (C) the Pr $M_{5,4}$-edges. The energy resolution was (A) 150 meV, (B) 230 meV and (C) 340 meV.

For the analysis of our data we used the entire FoV (details will be presented in section 2.2). Furthermore, we used PEEM data without any pre-processing, such as normalization or background removal applied. In order to improve visibility of image details for the reader, images shown here, however, are sections of size 180 $\times$ 180 pixels and the images are displayed with increased contrast. The section size corresponds to a sample area of 3.2 $\times$ 3.2 μm^2.

Fig. 1(a)-(c) show average spectra as determined for the 12 h annealed PBCO sample in all three energy regimes (black solid lines): (a) 526.0 to 550.0 eV, corresponding to the oxygen K-edge; (b) 772.0 to 809.0 eV, corresponding to the cobalt $L_{3,2}$- and barium $M_{5,4}$-edges and (c) 920.0 to 960.0 eV, corresponding to the praseodymium $M_{5,4}$-edges. The solid lines in Fig. 1 were obtained by averaging the pixel intensity of all image pixels inside the circular FoV in the image stack and plotting the mean intensity vs. the photon energy (for details see also section 2.2). In the view graphs, spectroscopic features of each dataset prototypical for a

perovskite are indicated: first, the e_g and t_{2g} resonances are the transitions from O1s to the hybridized O2p/Co3d states which are split by the octahedral field (Fig. 1 (a)). Secondly, the spin orbit split 2p to 3d and 3d to 4f resonances for Co and Ba/Pr respectively appear as distinct peaks in the spectra in (b) and (c).

Spectra as obtained at distinct pixels in a single image of the stack may substantially deviate from the mean spectra. As an example, we have plotted two local spectra as found at distinct pixel positions as dotted gray and dark gray curves. Similar deviations are also found for the other two energy ranges but are not shown in Fig. 1 (b), (c).

Fig. 1(d) shows 180 x 180 sections of the raw PEEM data for photon energies corresponding to the typical resonant features of the XAS in the spectra (a) to (c): The O2p/Co3d e_g and t_{2g} states (528.8 eV and 530.2 eV respectively), cobalt (780.2 eV), barium (784.0 eV) and praseodymium (929.4 eV). The raw PEEM images of the 12h-annealed sample show high- and low-intensity areas at 528.8 eV and 530.2 eV. For 528.8 eV, this structure is well-resolved, whereas for 530.2 eV, the PEEM image is noisy. At 780.2 eV, no obvious structure safely distinguishable from noise is observed. At 784.0 eV, a similar structure compared to 528.8 eV is revealed; and finally, at 929.4 eV, the PEEM image appears to have a complementary contrast compared to 528.8 and 784.0 eV. In the first image of Fig. 1(d) we have indicated those pixels as white cross/circle where we have measured the two local spectra displayed in Fig. 1(a) as dark-/light-gray dotted curves.

From the raw PEEM images in Fig. 1 one might conclude that annealing PBCO to 1073 K for 12 h causes a severe restructuring of the surface into two domains: one with high Ba and low Pr content, and the other vice versa, corresponding to a spinodal decomposition observed in the bulk of similar materials [33]. The low contrast in the raw PEEM image at 780.2 eV

indicates that cobalt is almost homogeneously distributed over the surface. However, the Co signal is hardly above the noise level. Hence, a localization of cobalt with respect to the visible domains is difficult if not to say impossible.

In the following we will demonstrate that PCA can not only attribute specific core level states (including those of cobalt) to the visible domains, but rather, it is capable to identify a third domain which is camouflaged in the raw PEEM images.

2.2. Principal Component Analysis of PEEM images

As a benefit to those readers not familiar with PCA, we briefly summarize some basics. Principal component analysis is a multivariate procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components (PC).

Mathematically, a spectromicroscopic image of $(m \times n)$ pixels is described by an intensity matrix I_{mn} and the raw data set of spectromicroscopic images for k energies is a matrix stack of k $(m \times n)$-matrices. Spectral information is obtained from the image stack by plotting the k intensities $I_{\mathrm{mn}}(E_{\mathrm{k}})$ vs. the energy E_k for a fixed distinct value pair (m, n) (see Fig. 2(a)).

Let N be the total number of pixels in a particular image of the stack, i.e. $N = m \cdot n$, the spectral information in the entire spectromicroscopic image stack is described by a $(k \times N)$ matrix $\overline{M}$. In order to keep the spatial information in the spectromicroscopic image accessible, the N matrix columns of $\overline{M}$ must remain uniquely related to the image pixel indices (m, n). In the matrix $\overline{M}$, the spectral frequencies are identified as the experimental variables represented by the matrix rows (Fig. 2(b)). In the matrix columns one finds the spectral intensities which are identified as the experimental observations. In other words,

the observed spectral intensities are the projections of the data on the coordinate axes in the original space.

The PCA projects $\overline{M}$ onto its PCs, i.e. it searches for orthonormal axes of a new principal component space such that the original space-axes decompose into two sets of principal axes: one subset of few principal axes which are oriented along the direction of largest data variances. The second subset contains the large number of remaining principal axes which account for the residual with small data variances. The goal is to describe almost all of the original data variances with a small subset of principal axes in order to reduce data dimensions and separate residual information.

Mathematically, the new principal axes are the eigenvectors of the covariance matrix of $\overline{M}$. PCA calculates new principal axes (principal components) as linear combinations of the original coordinate axes. PCA picks the particular solution where the first PC accounts for the largest variance; the second principal component for the second-largest variance and so forth. The principal components are hence sorted with respect to the variance, largest comes first.

The PCA expresses $\overline{M}$ as the product of two new matrices $\overline{T}$ (scores) and $\overline{P}$ (loadings)

$$\overline{M} = \overline{T}\,\bar{P}^{T} \qquad (2)$$

The scores are the representations of the original data in the new orthogonal space of the new principal axes, i.e. they are the coefficients of linear combinations in the space of the PC. The loadings, on the other hand, reflect the variance along the direction of the PCs, i.e. they show how much the component contributes to the variance of original data for a given variable.

Multiplication of eq. (2) by $\overline{P}$ on both sides yields (with $\bar{P}^T\ \bar{P} = \bar{I}$ for orthonormal matrices)

$$\overline{M}\,\overline{P} = \overline{T}\,\bar{P}^T\overline{P} = \overline{T} \qquad (3)$$

If $N \gg k$, it is sufficient to reduce the number of PCs to k, and hence, the score matrix $\overline{T}$ to a square $(k \times k)$-matrix. The reduction to a square matrix may considerably reduce calculation time, in particular in the case of large pixel numbers such as the $600x600$ pixels PEEM images presented here. Once the loadings matrix $\bar{P}$ is known, the full original data set is regained according to eq.(3) via transposition of $\overline{P}$ and multiplication with the score matrix $\overline{T}$.

As mentioned before, the PCA aims at the description of the experimental data by a small number of linearly independent PCs. That is, the original, and mostly correlated data axes are replaced by the uncorrelated PCs which account for the largest variance in the data. The fact that the PCs are uncorrelated will be of importance when interpreting the results for the PCs since it allows distinguishing correlated and uncorrelated chemical changes in complex materials such as perovskites.

The difficulty in experimental data analysis is to decide how many PCs belong to the small subset of PCs with large variance and which should be taken into account to describe the physical and chemical information in the data. For that purpose it is inevitable to have a look at the total variance explained by the PCs considered. The variances are the eigenvalues λ of the covariance matrix of $\bar{M}$. Hence, the total variance explained (TVE) by the first *j* PCs is given by:

$$\text{TVE} = \frac{\sum_1^j \lambda_j}{\sum_1^N \lambda_N} \qquad (4)$$

Typically, the value for the TVE rises strongly for the first PCs and then gradually flattens to approach a value of 100% when all PCs are taken into account. The question remains, how large the TVE value must be to explain the most relevant part of the data. In 1966, Cattell developed with the Scree-test a simple and straight-forward procedure [36]. Here, the eigenvalues λ_j are plotted vs. the index *j* of the PC. For small *j*, the decrease in the eigenvalue λ_j is almost linear. For a certain value of *j* one observes an obvious kink in the data curve. All values of *j* below the kink are considered to be relevant. The kink criterion leaves some room for interpretation, in particular when there is no obvious kink as will be discussed in more detail for the data presented in Fig. 4.

An interesting side effect of the PCA is that one may generate loading maps using individual PCs only:

$$\overline{M}_{PC} = \overline{T}_{PC}\,\bar{P}_{PC}^{T} \qquad (5)$$

Here, the index "PC" indicates that eq. (2) is evaluated for individual PCs only.

With eq. (5), one identifies those pixels from which the information on the particular principal component loading originated. In case of different chemical components with clearly distinguishable spectra, image generation using individual loading intensity maps yield therefore spatially resolved images of local chemical component concentrations.

For our analysis, we employed the standard algorithms for PCA analysis with centered data as implemented in the Matlab statistics toolbox (version 7.11.0 R2010b, Mathworks Inc.). No additional preprocessing was applied to the raw PEEM image data.

To exclude all points in the 600×600 PEEM image referring to a position outside the FoV of the PEEM, we applied a region of interest (ROI) with a radius of 567 pixels. Hence, the pixel size of a raw PEEM image was reduced from $600 \times 600 = 360000$ to 252448 pixels within the ROI. That is, principal component analysis as reported here was performed on 252448 spectra for each of the energy ranges (A)-(C).

The origin of the various PCs as determined by PCA may be manifold, in particular for complex oxides: first, different chemical elements give rise to different PCs, as employed readily for chemical composition mapping [15]. Second, separate domains with different, nevertheless, homogenous concentrations of one and the same chemical species will cause separate PCs. And third, different oxidation states of a chemical element will yield different Eigen vectors due to the sensitivity of the overall XAS shape to the local electronic structure. Hence, care in the interpretation of the results and a basic knowledge on the particular oxide is essential to interpret PCA results of perovskite oxides.

In the case of PBCO, we have three different energy ranges and various spectral features serving as fingerprints for the chemical states. Furthermore, as it turns out, PCs are sensitive to differences in chemical compounds (Co, Ba), differences in oxidation states (nominally Co^{3+}, Co^{4+}) as well as concentration differences (Co, Ba, Pr).

2.3 Domain localization using orthogonal non-negative matrix factorization

If samples under study reveal domain separation of chemical components on the resolution scale of the PEEM, loading intensity maps will form gray scale images with clearly distinguishable high- and low-intensity areas. In order to analyze the average spectral intensity within certain domains, pixels that belong to the specific domain in the loading map must be identified. An easy approach would be to choose threshold gray values to separate

different domains. However, this procedure depends sensitively on the chosen thresholds. Therefore, we favored another well-known multivariate technique, the orthogonal non-negative matrix factorization (ortho-NMF [37, 38]). Here, domains can exclusively be identified using mathematical procedures: in ortho-NMF, the $(m \times n)$ data matrix $\overline{M}$ at the *k*th energy is factorized into non-negative, orthogonal factors, a $(m \times c)$ matrix $\overline{V}$ and a $(c \times n)$ matrix $\overline{H}$. The dimensionality c is determined by the number of data groups, or clusters, in the system. The factorization is quite similar to PCA; however, with the constraint that the basis vectors as well as the coefficients must be all positive and orthogonal. The separation of the data into specific groups (or clusters) can be performed with respect to various criteria. Here, we classify data groups according to similar Euclidian distance from the mean using a Matlab based code according to Andri Mirzal [38, 39].

A major drawback of the NMF compared to PCA is that NMF requires a pre-assumption of the number *c* of data groups to be found. Such pre-assumptions, however, might introduce artificial results; in particular, they may prevent detection of unexpected cluster groups. Therefore, we use the ortho-NMF exclusively for the localization of domains in the loading maps detected by PCA.

3. Experimental Results

The mean spectra as obtained from the average over all spectra in the ROI were already presented in Fig. 1. As demonstrated in Fig. 1(a), spectra as obtained from individual pixels, however, may deviate substantially from the mean spectrum.

These local variations in spectral information were analyzed using PCA: Fig. 3 shows the loading intensity for the lowest five PCs. In all energy ranges (A) – (C) the loading intensity of PC 1 resembles the mean spectrum as displayed in Fig. 1(a) – (c). The loading intensities of PC 2 and higher account for changes in peak intensity and peak positions.

Prior to the discussion of details of these changes we have to decide which of the PCs are sufficient to describe our experimental data. For this purpose we make use of the TVE plot according to eq. (4) and of the Scree-test [36] as described before. Fig. 4(a) shows the total variance explained (TVE) vs. the principal component number for the three energy ranges (A), (B) and (C). For better visibility, the inset is a plot of the same data, however, restricted to the first 9 PCs. For (A) and (B), the first 3 PCs explain about 93% of the total variance, for (C) about 98%. For (A) and (C), the TVE increases steeply between the first 3 PCs, whereas for $j > 3$ the increase is weaker. In contrast, in the energy range (B) the TVE shows no obvious steep increase for $j \leq 3$. The TVE increases gradually over the entire PC number range. Hence, $1 \leq j \leq 3$ seem to be sufficient to describe the data in the energy ranges (A) and (C), whereas in the energy range (B) it is more difficult to decide how many PCs should be taken into account. Further information is gained from the Scree-test [36] which is shown in Fig. 4(b). Here, the Eigenvalues λ_j (eq. (4)) are plotted against the principal component number. The data curves coincide for $j \geq 3$. For the energy ranges (A) and (C) an obvious kink is

observed at *j* = 3, in accordance with the TVE plot. For the energy range (B), the kink is found at *j* = 2.

According to the Scree-test criterion the first three PCs should therefore be enough to fully describe the PBCO PEEM results in the energy ranges (A) and (C). For energy range (B) one may restrict the PCs to the first two. However, we will later see in Fig. 5 that PC 3 indeed contains considerable structural information also in the energy range (B); therefore we will discuss in the following the loading intensities in Fig. 3 up to a maximum number of three for all three energy ranges (A) to (C).

3.1. Loading intensities

3.1.1. (A) O K-edge: 525 eV – 550 eV

The loading intensity of PC 1 (Fig. 3(A)) resembles the mean spectrum (black curve in Fig. 1(a)). Strong peak signals are observed for the loading intensities of PC 2 and 3 at 528.8 and 530.2 eV: The loading intensity of PC 2 accounts for a correlated reduction of both peaks, which would correspond to a reduction of Co-bound oxygen or a decrease in overall covalency [26]. The loading intensity of PC 3 in turn accounts for an uncorrelated sole increase of the peak at 528.8 eV which can be attributed to the electron holes on the CoO_6-manifold [20]. For higher energies between 533 and 550 eV there seems to be a slight contribution of the loading intensity of PC 2 and 3 to the broad peaks around 536 and 542 eV, corresponding to the Ba/Pr4d-O2p hybridized states; however, the contribution is only slightly above the error limit. The local spectral variations seem therefore to be dominated by variations in the oxygen e_g and t_{2g} states.

3.1.2. (B) Co L- and Ba M-edge: 775 eV – 815 eV

The loading intensity of PC 1 reflects the typical spectrum for Co-Ba (Fig. 1(b)), whereas PC 2 and 3 have only small contributions to the raw spectra. Nevertheless, the loading intensities of PC 2 and PC 3 have positive bumps at the $Co\ 2p_{3/2}$ and $Co\ 2p_{1/2}$ levels. Simultaneously, the loading intensities of PC 2 and PC 3 reveal a slight decrease of the $Ba\ 3d_{3/2}$ and the $Ba\ 3d_{5/2}$ signal, respectively. That is, the spectral variations seem to be attributed to a change in the Co-Ba peak ratio and, hence, a change in the relative concentration of both elements.

3.1.3. (C) Pr M-edge: 920 eV – 960 eV

The loading intensity of PC 1 as shown in Fig. 3(c) reflects the typical spectrum of praseodymium (Fig. 1(c)). PC 2 makes a contribution to the peak intensity at 929.4 eV. The loading intensity of PC 3 is hardly above the noise limit, however, it appears to be caused by a peak shift of the $Pr\ 3d_{5/2}$ peak at 929.4 eV.

3.2. Loading intensity maps

In Fig.5, we map the loading intensities as obtained from specific image pixels using eq. (5). The results for the first three PCs are shown in column 1 to 3. The last column in Fig. 5 shows the residual images as obtained by mapping loading intensities higher than PC 3. The residual images have no apparent remaining structure and are dominated by noise; thereby corroborating that PC 1 to PC 3 are sufficient to fully describe the PEEM data.

100% positive loading intensity in the images corresponds to 1, and 100% negative loading intensity to -1. Negative values originate from negative entries in the score matrix $\bar{T}$. For improved visibility, the loading images in Fig. 5 are contrast enhanced such that the

maximum/ minimum intensity values are displayed as white/black. The real numbers of the maximum/ minimum values are denoted together with the scale bars at each loading image. The photon energies are indicated in the left-hand side of the graphics.

For the energy range (B) it becomes obvious that PC 3 contains a non-negligible amount of structural information. Simultaneously, the residual image containing PCs with $j \geq 4$ is almost without any structure. Although the interpretation of the TVE and the Scree-plot in Fig. 4 was not unambiguous, we can finally conclude from Fig. 5 that the first three PCs are required and sufficient to describe the relevant variances in the raw PEEM data in the energy ranges (A), (B) and (C). In the following paragraphs, we therefore discuss the mapping images for the different energy ranges (A) to (C) based on PC 1 – PC 3:

3.2.1. (A): O K-edge: 525 eV – 550 eV

At 528.8 eV, the loading intensity maps of PC 2 and PC 3 are similar to the raw PEEM image at this energy: high and low intensity in the original PEEM image corresponds to positive, respectively negative loading intensity. The loading intensity map of PC 1, the principal component with the largest eigenvalue, corresponds to a third domain whose existence is not obvious in the raw PEEM image: This domain is a thin region separating the two intensity domains as observed for PC 2 and PC 3. At 530.2 eV, the loading intensity maps of PC 1 and PC 2 show a domain structure similar to that at 528.8 eV. The loading intensity map of PC 3, however, is reversed: low-intensity domains at 530.2 eV correspond to high-intensity areas at 528.8 eV and vice versa.

3.2.2. (B): Co L- and Ba M-edge: 775 eV – 815 eV

For PC 1, the loading intensity map for $Co\ 2p_{3/2}$ at 780.2 eV is – safe for differences in noise - similar to that of $Ba\ 3d_{5/2}$ at 784.0 eV. High- and low-intensity areas coincide. For PC 2 and

3 on the other hand, loading intensities for $Co\ 2p_{3/2}$ at 780.2 eV are high where the loading intensities for $Ba\ 3d_{5/2}$ at 784.0 eV are low, and vice versa. The inversed contributions of PC 1 and PC 2,3 to the intensity maps explains the almost structure-free PEEM image at 780.2 eV (Fig. 1). Furthermore, they seem to indicate that high cobalt contribution is found where the barium contribution is reduced and vice versa.

In contrast to energy range (A), we observe only two domains in the Co-Ba energy range (B). No third domain as found for (A) with PC 1 exists for (B). The domain structure in (B) for PC 1, 2 and 3, however, is similar to what is found in (A) for PC 2 and PC 3. While the domain borders seem to be less sharp in (B) compared to (A), the domains in (B) obviously coincide with the high- and low-intensity loading areas of PC 2 and PC 3 in the energy range (A): at 528.8 eV the loading intensity of PC 2 and 3 is high where the loading intensity of PC 2 and 3 at 784.0 eV is high and the loading intensity of PC 2 and 3 at 780.2 eV is low (and vice versa).

3.2.3. (C): Pr M-edge: 920 eV – 960 eV

For the Pr energy range, high-intensity loading areas of PC 1 correspond to the high-intensity domains in the raw PEEM image (Fig. 1). The loading intensities of PC 2 and PC 3 are close to the noise level. Still, the domain structure is visible, in particular in the loading intensity map of PC 3. Here, large negative loading intensities of PC 3 coincide with low-intensity domains in the original PEEM image. Similar to the Co-Ba energy range (B), no third domain is found for the energy range (C). The contrast in the loading map PC 1 in (C) is inversed compared to (B) indicating that domains with enhanced Ba content have reduced Pr concentration and vice versa.

3.3. Spectral decomposition for the different loading intensity domains

In order to determine the mean spectra in the various loading intensity domains in Fig. 3, we use the orthogonal non-negative matrix factorization (ortho-NMF) as described in Section 2.3. Since energy range (A) reveals all three domains (Fig. 5) we use the data from (A) as an input. We chose *c*=3 for the ortho-NMF analysis. The result of the ortho-NMF is presented in Fig. 6: Fig. 6(a) shows the raw PEEM image of PBCO at photon energy 528.8 eV. In Fig. 6(b) white (I), black (II) and gray (III) colored areas indicate the cluster groups as obtained by the ortho-NMF analysis. Pixels of the same color have spectral intensities with similar Euclidian distance. Obviously, the areas as found by ortho-NMF assuming 3 cluster groups coincide perfectly with the domains observed in the loading intensity maps of PC 1, 2 and 3 as determined by PCA in energy range (A) (Fig. 5): the black cluster group corresponds to high-intensity areas in the mapping image of PC 1 at 528.8 eV; the white and gray cluster to high-, respectively, low-intensity areas in the mapping images of PC 2 and PC 3 at this energy.

Fig. 6(c) is a plot of the mean spectra determined as an average over all raw spectra in the white (I), black (II) and gray (III) colored areas indicated in (b). Obviously, the white (I), black (II) and gray (III) colored domains differ substantially in the spectra: white areas have a prominent e_g peak, negligible $Co\ 2p_{3/2}$ and $2p_{1/2}$ signals and a low $Pr\ 3d_{5/2}$ peak. Gray domains have a strongly reduced e_g peak, considerable $Co\ 2p_{3/2}$ and $2p_{1/2}$ contributions and an intermediate $Pr\ 3d_{5/2}$ intensity. The black areas show intermediate e_g, $Co\ 2p_{3/2}$ and $2p_{1/2}$ peaks and the highest $Pr\ 3d_{5/2}$ intensity.

The final average spectra in the three domains shown in Fig. 6(c) indicate that PBCO is best described by assuming a phase separation between different chemical oxidation states, element concentrations and stoichiometry. Gray areas (II) have a lower Ba-to-Co ratio than

white areas (I). Black areas (II) have an intermediate Ba-to-Co ratio, an intermediate O e_g peak height, but maximum Pr concentration. A large Ba-to-Co ratio is correlated with a low Pr concentration and maximum O e_g peak. The maximum cobalt concentration with respect to Ba is associated with the lowest O e_g peak. These observations are well compatible with the rudimentary decomposition reaction as shown in eq. (1), albeit it appears that the decomposition is not a complete demixing of Ba and Pr, i.e. there is still some solubility of either A-site cation in the other lattice.

Apparently, neither a simple A-site decomposition of the perovskite as observed in bulk counterparts nor a precipitation of a secondary phase can describe the PCA results well: it appears that a residual parent phase with a small tolerance of Pr/Ba ratio is still present, whereas the decomposition products do exhibit such a phase breadth. Granted, this might be to a non-complete conversion of the reaction, however these thermodynamic details are, as stated above, the scope of a subsequent work.

In Fig. 7, we present reconstructed images according to eq. (5) (denoted as "PCA") where we restricted the scores $\bar{T}$ and loadings $\bar{P}$ to the relevant first three PC indices $1 \leq j \leq 3$. We show the PCA images at the energies of distinct core level peaks. PCA images reproduce the raw PEEM data (denoted as "PEEM"); however, noise is reduced, resolution is improved and local intensity can be unambiguously attributed to specific spectral features, thereby providing spatially resolved chemical information. From Fig. 7 one can separate the various spectral contributions to the image contrast in the raw PEEM image for distinct photon energies: Areas where e_g is large have reduced $Co\ 2p_{3/2}$ and $Co\ 2p_{1/2}$ and enhanced $Ba\ 3d_{5/2}$ and $Ba\ 3d_{3/2}$ contribution. The blurred contrast of the PEEM image at 530.2 eV is unmasked by PCA as high O t_{2g} signals at the rims of the high intensity e_g domains. The

uncorrelated accumulation of $O\ t_{2g}$ at the rims of high-intensity e_g domains can be detected only in the PCA images. Finally, areas with the highest Pr concentration correspond to areas with lower $Ba\ 3d_{5/2}$ / $Co\ 2p_{3/2}$and $Ba\ 3d_{3/2}$ / $Co\ 2p_{1/2}$ ratio.

4. Discussion

As mentioned in the beginning of this article, the interpretation of the spectroscopic information with regard to the actual surface defect chemistry of PBCO is beyond the scope of the paper and will be the focus of a future publication. Here, we want to emphasize the strength of the PCA method since we believe that multivariate tools have not yet received the deserved attention for the analysis of spectromicroscopic PEEM data.

The major strength of PCA when applied to PEEM images is that data is analyzed while being fully detached from researchers' pre-assumptions about the possible outcome of the experiment. As a main consequence, the scientific outcome does not suffer from unintentional artificial input and data manipulation. As demonstrated nicely in this work, the identification of a third domain revealed in the first principal component PC 1 for the oxygen K-edge (A) (Fig. 5) would hardly have been possible without PCA. The raw PEEM data merely shows two obvious domains and the third domain camouflages as an intermediate contrast in the border region between those domains. An intermediate contrast in raw spectromicroscopic images; however, will be most probably interpreted as a gradient in gray levels between low and high intensity areas rather than as a real third chemical domain.

In addition, PCA is capable of identifying correlated and uncorrelated variances in the data. This information can provide real physical and chemical information since one may be able to identify correlated and uncorrelated changes in oxidation states. This is demonstrated in Fig. 3 in the data for the oxygen K-edge (A). The loading intensity of PC 2 corresponds to a correlated change of e_g and t_{2g}. However, the loading intensity of PC 3 shows that there is

an additional contribution to e_g which is not correlated with changes in t_{2g}. This becomes obvious also from the mapping images in Fig. 7. Here, the uncorrelated t_{2g} loading intensity is localized in the rim regions of high intensity e_g areas. The uncorrelated change of e_g can be easily interpreted in the framework of the defect chemistry here: a demixing of divalent Ba and trivalent Pr impacts the overall hole concentration, which is measured by the e_g electron count [25]. The correlated change of e_g and t_{2g} signals – a measure of the number of empty t_{2g}/e_g hybrid orbitals – shows the concomitant change of covalency with oxidation state [25]. What is most interesting, though, is the fact that these two effects are linearly independent, as the PCA shows, leading to the interpretation that the decomposition products do not simply act as the end members of the $Pr_{1-x}Ba_xCoO_3$ solid solution series. This is crucial information not provided easily without multivariate techniques and may be missed if one embarks on analysis with a simple decomposition in mind.

As an extra which may not be underestimated though, chemical information is achieved with an improved signal-to-noise ratio and with spatial resolution. For instance in Fig. 7 in the energy range (B) at 780.2 and 794.6 eV (corresponding to $Co\ 2p_{3/2}$ and $Co\ 2p_{1/2}$), raw PEEM data is hardly above the noise level. After PCA, the domains are clearly identified and obviously correlated with lower barium intensity. This is chemical information which might have been easily missed without PCA.

Finally, the PCA evaluation is systematic and fast and can handle large spectroscopic image stacks which can be quite tedious to analyze manually. This is an advantage which one should not belittle. Time-consuming, manual evaluation procedures may lead to researchers' subliminal self-restrictions in data acquisition and evaluation and, therefore, in unintentional reduction of statistical data base and precision and relevance of scientific outcome.

One should emphasize though that care has to be taken when interpreting loading intensities of principal components. The most critical researcher's decision is how many principal components have to be taken into account. Although methods such as the Scree-test may help, there is still some room for interpretation. Therefore, evaluation of spectromicroscopic images using PCA may be fast, interpretation of the results may still be the bottle neck on the way to publication of reasonable results.

Finally we address the question why we did not use the ortho-NMF method to determine the different domains and their spectral differences in the first place. Application of ortho-NMF requires assumptions on the algorithms used, the starting values and the convergence criteria of the iterations which may be manifold [40]. Hence, great care has to be taken to avoid artificial results. The advantage of the ortho-NMF vs. PCA is that the former yields positive basis vectors which appear more appropriate for physical and chemical problems. Furthermore, the inherent clustering property [41] is an interesting tool to analyze phase separation behavior. As mentioned before, however, the clustering tool requires quite a number of pre-assumption, in particular, the number of cluster groups to be found. In the case of our data on PBCO, the raw PEEM data apparently display merely 2 different domains. Assuming 2 cluster groups for an ortho-NMF may therefore be the obvious decision and the ortho-NMF would yield cluster images similar to the loading maps of PC 2 and PC 3 in Fig. 5 at 528 eV. As a consequence, one would be tempted to accept the finding of two different domains and one would not detect the third phase hidden in the raw PEEM images which is described by PC 1. In contrast, PCA revealed the camouflaged third phase without any further input or pre-assumptions.

5. Summary

In this paper we demonstrated the strength of multivariate techniques such as principal component analysis in evaluating spectromicroscopic images. We showed that even in the case of raw data with very low signal-to-noise ratio, PCA may safely yield low-noise spectroscopic data without any data pre-processing or any influence by pre-assumptions on the scientific outcome. PCA is fast and makes large statistical data sets easily accessible.

PCA may detect unexpected features that are not obvious from the raw spectromicroscopic images and might therefore lead to a far deeper understanding of the systems under investigation. Here, we demonstrated that PBCO separates into three phases upon prolonged annealing. Two of the phases are visible in the raw PEEM images. The third phase, however, was revealed by PCA without neither additional pre-assumptions nor pre-processing of the raw data.

Interpretation of PCA results must be performed with due care though, since the result for the different core levels might be as complex as the sample under investigation. We hope that we nevertheless convinced more scientists from the spectromicroscopic community that the application of multivariate methods to the data is worth the effort.

Figure Captions:

Fig. 1: Mean XAS spectra as obtained as an average over all pixels in the raw PEEM images for three energy regimes (solid black lines): (a) 526.0 to 550.0 eV, corresponding to the oxygen K-edge; (b) 772.0 to 809.0 eV, corresponding to the cobalt L3,2- and barium M5,4-edges and (c) 920.0 to 960.0 eV, corresponding to the praseodymium M5,4-edges of PBCO. As an example for deviations of local spectra from the mean, in (a) strongly deviating spectra as measured at two distinct pixel positions are plotted as dotted gray and dark gray curves. Similar deviations are found also for the other two energy ranges but are not shown in (b), (c). In (d) we show the PBCO sample after 12 h annealing. The images are $180x180$ pixels $3.2 \times 3.2\ \mu m^2$ sections of $600x600$ pixels raw PEEM images measured at the indicated energies. The white circle/cross in the left-hand image indicates the positions where the light-/dark-gray curves in (a) have been measured.

Fig. 2: Graphical illustration of how spectral information is deduced from raw spectromicroscopic images with intensity values I_{mn} in the image matrix with $m \times n$ entries. (a) Raw spectromicroscopic data is stored in an image stack with k levels where each image level is recorded at a distinct photon energy E_k. The spectral information is obtained by plotting the intensity values at individual image pixels $(m, n)_k$ vs. energy E_k. (b) For PCA, the spectral information is written in a matrix with *k* rows and *N* columns; $N = m \cdot n$ is the total number of pixels in a spectromicroscopic image of the stack of size *k*.

Fig. 3: Loading intensities for the first 5 principal components (PC) for the energy ranges (A) to (C). The loading intensity of PC 1 resembles the measured mean spectrum as displayed in Fig. 1. Loading intensities of higher PCs indicate deviations from the mean spectrum.

Fig. 4: (a) Total variance explained (TVE) vs. the principal component number for the energy ranges (A), (B) and (C) as dashed, solid and dot-dashed curves, respectively. The inset shows the same data restricted to PCs up to 9 (A: dashed line with light-gray circles; B: solid line with dark-gray triangles; C: white squares with dot-dashed line). (b) Eigenvalues of the covariance matrix of $\bar{M}$ (eq. (2)) versus the principal component number j for $j \leq 6$ (Scree-test [36]). Data is plotted using the same notation as in the inset of (a). Observations of kinks in the data curves indicate an upper limit of the number of relevant PCs to describe the data. See Section 3 for further discussion.

Fig. 5: Loading intensity mapping images for PC 1, 2 and 3 calculated according to eq. (5). Images are calculated using the FoV of the $600x600$ pixels raw PEEM images. Here, merely $180x180$ pixels ($3.2 \times 3.2\ \mu m^2$) sections are displayed. The mapping images are contrast enhanced to maximum contrast between black/white. Minimum/maximum loading intensities as found in the mapping images are denoted together with the scale bars at the image side. 100% positive/ negative loading corresponds to values +1, respectively -1. The last column shows residual images, i.e. mapping images using principal component numbers higher than 3. The residual images are dominated by noise.

Fig. 6: (a) Raw PEEM image section (3.2 µm × 3.2 µm) of PBCO at photon energy 528.8 eV. (b) Cluster group image as obtained by an ortho-NMF analysis (see 2.3) showing three different domains I (white), II (black) and III (gray). (c) Average of raw spectral PEEM data. Spectra related to white, black and gray colored areas are denoted I, II and III, respectively.

Fig. 7: Reconstruction of PEEM image data at distinct energies for (A) the oxygen K-edge, (B) the cobalt L- and barium M-edge, and (C) the praseodymium M-edge according to eq. (4) using PC 1, 2 and 3 (indicated as "PCA", surface area 3.2 µm × 3.2 µm). Note that the PCA images are similar to the raw PEEM images (indicated as "PEEM"); however, noise is reduced, resolution is improved and local intensity can be unambiguously attributed to specific core levels, thereby providing locally resolved chemical information.

6. References:

[1] I.T. Jolliffe, Principal component analysis, Springer, New-York, NY, 1986.

[2] J.E. Jackson, A User's Guide to Principal Components, John Wiley & Sons, Inc.2004.

[3] K. Artyushkova, Structure determination of nanocomposites through 3D imaging using laboratory XPS and multivariate analysis, Journal of Electron Spectroscopy and Related Phenomena, 178–179 (2010) 292-302.

[4] J. Walton, N. Fairley, Noise reduction in X-ray photoelectron spectromicroscopy by a singular value decomposition sorting procedure, Journal of Electron Spectroscopy and Related Phenomena, 148 (2005) 29-40.

[5] E. Folven, Y. Takamura, J.K. Grepstad, X-PEEM study of antiferromagnetic domain patterns in LaFeO3 thin films and embedded nanostructures, Journal of Electron Spectroscopy and Related Phenomena, 185 (2012) 381-388.

[6] T. Furnival, R.K. Leary, P.A. Midgley, Denoising time-resolved microscopy image sequences with singular value thresholding, Ultramicroscopy, 178 (2017) 112-124.

[7] B. Lavine, J. Workman, Chemometrics, Analytical Chemistry, 82 (2010) 4699-4711.

[8] B.K. Lavine, J. Workman, Chemometrics, Analytical Chemistry, 85 (2013) 705-714.

[9] H. Shinzawa, K. Awa, W. Kanematsu, Y. Ozaki, Multivariate data analysis for Raman spectroscopic imaging, Journal of Raman Spectroscopy, 40 (2009) 1720-1725.

[10] H. Grahn, P. Geladi, Techniques and Applications of Hyperspectral Image Analysis, John Wiley & Sons, Incorporated, Chichester, UK, 2007.

[11] E. Bauer, A brief history of PEEM, Journal of Electron Spectroscopy and Related Phenomena, 185 (2012) 314-322.

[12] M.J. Fay, A. Proctor, D.P. Hoffmann, M. Houalla, D.M. Hercules, Determination of the Mo surface environment of Mo/TiO2 catalysts by EXAFS, XANES and PCA, Microchimica Acta, 109 (1992) 281-293.

[13] S. Wasserman, R., The Analysis of Mixtures: Application of Principal Component Analysis to XAS Spectra, J. Phys. IV France, 7 (1997) 203-205.

[14] S.R. Wasserman, P.G. Allen, D.K. Shuh, J.J. Bucher, N.M. Edelstein, EXAFS and principal component analysis: a new shell game, Journal of Synchrotron Radiation, 6 (1999) 284-286.

[15] E.F. Smith, D. Briggs, N. Fairley, Further developments in quantitative X-ray photoelectron spectromicroscopy: preliminary results from the study of germanium corrosion, Surface and Interface Analysis, 38 (2006) 69-75.

[16] O. Renault, High-resolution XPS spectromicroscopy, Surface and Interface Analysis, 42 (2010) 816-825.

[17] A.J. Roberts, S.C. Page, K. Takahashi, Getting More from XPS Imaging: Multivariate Analysis for Spectromicroscopy, Journal of Surface Analysis, 14 (2007) 160.

[18] J. Walton, N. Fairley, XPS spectromicroscopy: exploiting the relationship between images and spectra, Surface and Interface Analysis, 40 (2008) 478-481.

[19] O. Renault, A. Garnier, J. Morin, N. Gambacorti, F. Bertin, High-resolution XPS spectromicroscopy study of micro-patterned gold–tin surfaces, Appl. Surf. Sci., 258 (2012) 10077-10083.

[20] E. Strelcov, A. Belianinov, Y.-H. Hsieh, S. Jesse, A.P. Baddorf, Y.-H. Chu, S.V. Kalinin, Deep Data Analysis of Conductive Phenomena on Complex Oxide Interfaces: Physics from Data Mining, ACS Nano, 8 (2014) 6449-6457.

[21] H. Yang, H.S. Lee, P.G. Kotula, Y. Sato, Y. Ikuhara, N.D. Browning, Amphoteric doping of praseodymium Pr3+ in SrTiO3 grain boundaries, Appl. Phys. Lett., 106 (2015) 121904.

[22] M. Hannula, K. Lahtonen, H. Ali-Löytty, A.A. Zakharov, T. Isotalo, J. Saari, M. Valden, Fabrication of topographically microstructured titanium silicide interface for advanced photonic applications, Scripta Materialia, 119 (2016) 76-81.

[23] H. Guo, E. Strelcov, A. Yulaev, J. Wang, N. Appathurai, S. Urquhart, J. Vinson, S. Sahu, M. Zwolak, A. Kolmakov, Enabling Photoemission Electron Microscopy in Liquids via Graphene-Capped Microchannel Arrays, Nano Letters, 17 (2017) 1034-1041.

[24] A. Grimaud, K.J. May, C.E. Carlton, Y.-L. Lee, M. Risch, W.T. Hong, J. Zhou, Y. Shao-Horn, Double perovskites as a family of highly active catalysts for oxygen evolution in alkaline solution, Nature Communications, 4 (2013) 2439.

[25] D.N. Mueller, M.L. Machala, H. Bluhm, W.C. Chueh, Redox activity of surface oxygen anions in oxygen-deficient perovskite oxides during electrochemical reactions, Nature Communications, 6 (2015) 6097.

[26] J. Suntivich, W.T. Hong, Y.-L. Lee, J.M. Rondinelli, W. Yang, J.B. Goodenough, B. Dabrowski, J.W. Freeland, Y. Shao-Horn, Estimating Hybridization of Transition Metal and Oxygen States in Perovskites from O K-edge X-ray Absorption Spectroscopy, The Journal of Physical Chemistry C, 118 (2014) 1856-1863.

[27] Y. Orikasa, T. Ina, T. Nakao, A. Mineshige, K. Amezawa, M. Oishi, H. Arai, Z. Ogumi, Y. Uchimoto, X-ray Absorption Spectroscopic Study on La0.6Sr0.4CoO3−δ Cathode Materials Related with Oxygen Vacancy Formation, The Journal of Physical Chemistry C, 115 (2011) 16433-16438.

[28] A.E. Bocquet, T. Mizokawa, K. Morikawa, A. Fujimori, S.R. Barman, K. Maiti, D.D. Sarma, Y. Tokura, M. Onoda, Electronic structure of early 3d-transition-metal oxides by analysis of the 2p core-level photoemission spectra, Phys. Rev. B, 53 (1996) 1161-1170.

[29] J. Suntivich, H.A. Gasteiger, N. Yabuuchi, H. Nakanishi, J.B. Goodenough, Y. Shao-Horn, Design principles for oxygen-reduction activity on perovskite oxide catalysts for fuel cells and metal–air batteries, Nature Chemistry, 3 (2011) 546.

[30] D.N. Mueller, R.A. De Souza, H.-I. Yoo, M. Martin, Phase Stability and Oxygen Nonstoichiometry of Highly Oxygen-Deficient Perovskite-Type Oxides: A Case Study of (Ba,Sr)(Co,Fe)O3−δ, Chemistry of Materials, 24 (2012) 269-274.

[31] W. Lee, J.W. Han, Y. Chen, Z. Cai, B. Yildiz, Cation Size Mismatch and Charge Interactions Drive Dopant Segregation at the Surfaces of Manganite Perovskites, J. Am. Chem. Soc., 135 (2013) 7909-7925.

[32] C. Xu, C. Bäumer, R.A. Heinen, S. Hoffmann-Eifert, F. Gunkel, R. Dittmann, Disentanglement of growth dynamic and thermodynamic effects in LaAlO3/SrTiO3 heterostructures, Scientific Reports, 6 (2016) 22410.

[33] D.N. Mueller, R.A. De Souza, T.E. Weirich, D. Roehrens, J. Mayer, M. Martin, A kinetic study of the decomposition of the cubic perovskite-type oxide BaxSr1-xCo0.8Fe0.2O3-[small delta] (BSCF) (x = 0.1 and 0.5), Phys. Chem. Chem. Phys., 12 (2010) 10320-10328.

[34] F. Gunkel, L. Jin, D.N. Mueller, C. Hausner, D.S. Bick, C.-L. Jia, T. Schneller, I. Valov, R. Waser, R. Dittmann, Ordering and Phase Control in Epitaxial Double-Perovskite Catalysts for the Oxygen Evolution Reaction, ACS Catalysis, 7 (2017) 7029-7037.

[35] C. Wiemann, M. Patt, I.P. Krug, N.B. Weber, M. Escher, M. Merkel, C.M. Schneider, A New Nanospectroscopy Tool with Synchrotron Radiation: NanoESCA@Elettra, e-Journal of Surface Science and Nanotechnology, 9 (2011) 395-399.

[36] R.B. Cattell, The Scree Test For The Number Of Factors, Multivariate Behavioral Research, 1 (1966) 245-276.

[37] C. Ding, T. Li, W. Peng, H. Park, Orthogonal nonnegative matrix t-factorizations for clustering, Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, Philadelphia, PA, USA, 2006, pp. 126-135.

[38] A. Mirzal, Orthogonal Nonnegative Matrix Factorization for Blind Image Separation, in: H.B. Zaman, P. Robinson, P. Olivier, T.K. Shih, S. Velastin (Eds.) Advances in Visual Informatics: Third International Visual Informatics Conference, IVIC 2013, Selangor, Malaysia, November 13-15, 2013. Proceedings, Springer International Publishing, Cham, 2013, pp. 25-35.

[39] A. Mirzal, A convergent algorithm for orthogonal nonnegative matrix factorization, Journal of Computational and Applied Mathematics, 260 (2014) 149-166.

[40] A. Cichocki, Nonnegative matrix and tensor factorizations : applications to exploratory multi-way data analysis and blind source separation, Wiley, Chichester, 2009.

[41] C. Ding, X. He, H.D. Simon, On the Equivalence of Nonnegative Matrix Factorization and Spectral Clustering, Proceedings of the 2005 SIAM International Conference on Data Mining, pp. 606-610.

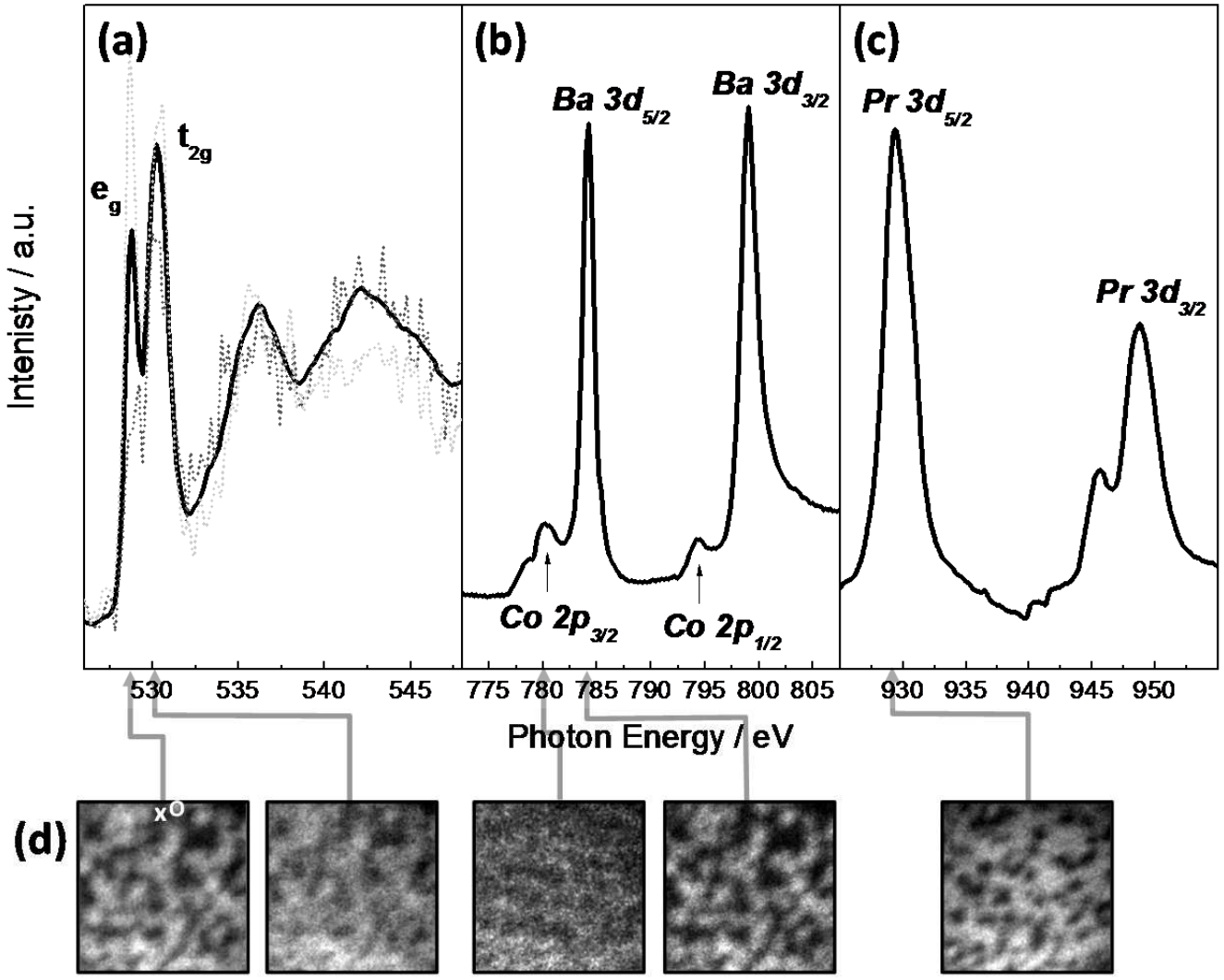
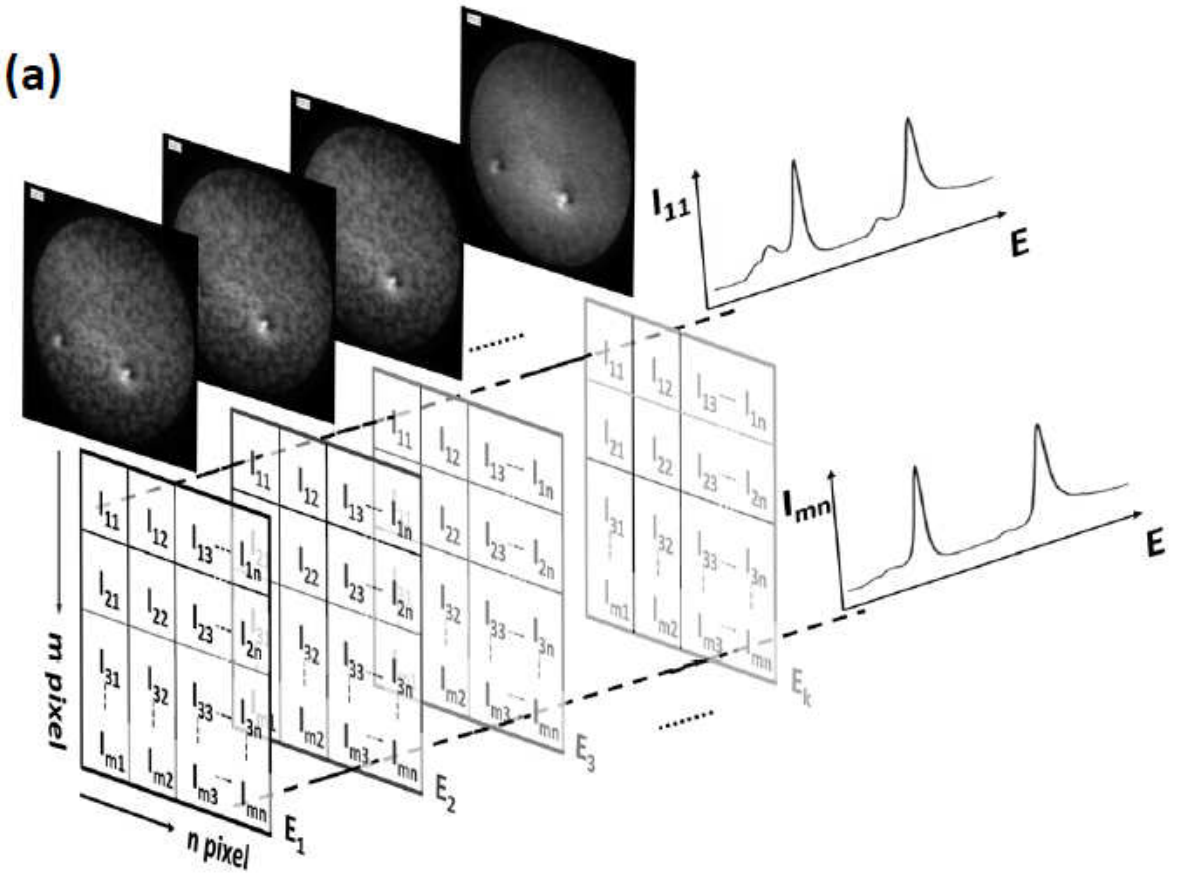
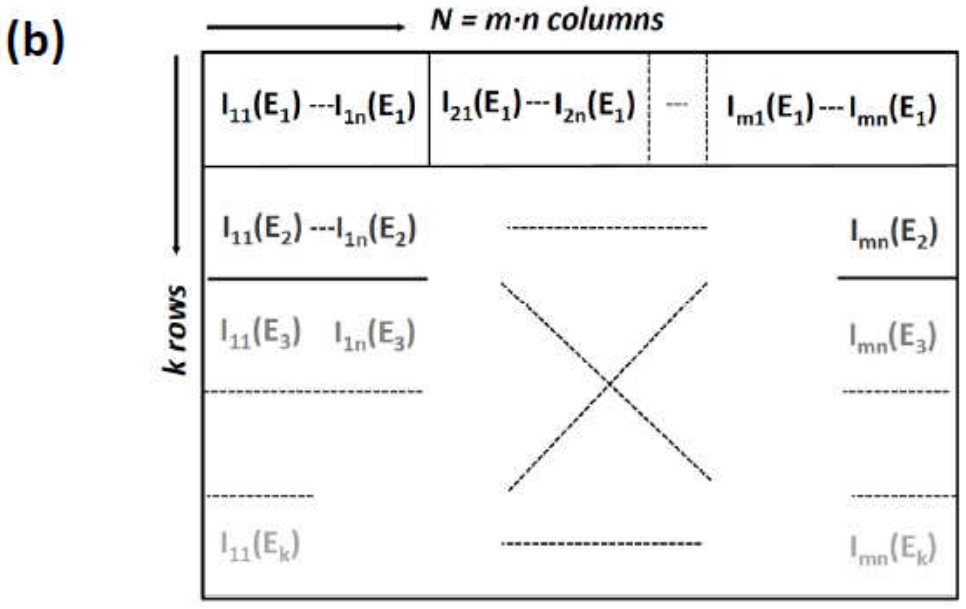

(a)
(b)
(c)
Intenisty / a.u.
e_g
t_2g
Ba 3d_5/2
Ba 3d_3/2
Pr 3d_5/2
Pr 3d_3/2
Co 2p_3/2
Co 2p_1/2
530 535 540 545
775 780 785 790 795 800 805
930 935 940 945 950
Photon Energy / eV
(d)
xO

Fig. 1


(a)
I11
E
Imn
E
m pixel
n pixel
E1
E2
E3
Ek


(b)
N = m·n columns
k rows
I11(E1) ---I1n(E1)
I21(E1)---I2n(E1)

Im1(E1) ---Imn(E1)
I11(E2) ---I1n(E2)
Imn(E2)
I11(E3) I1n(E3)
Imn(E3)
I11(Ek)
Imn(Ek)

Fig. 2

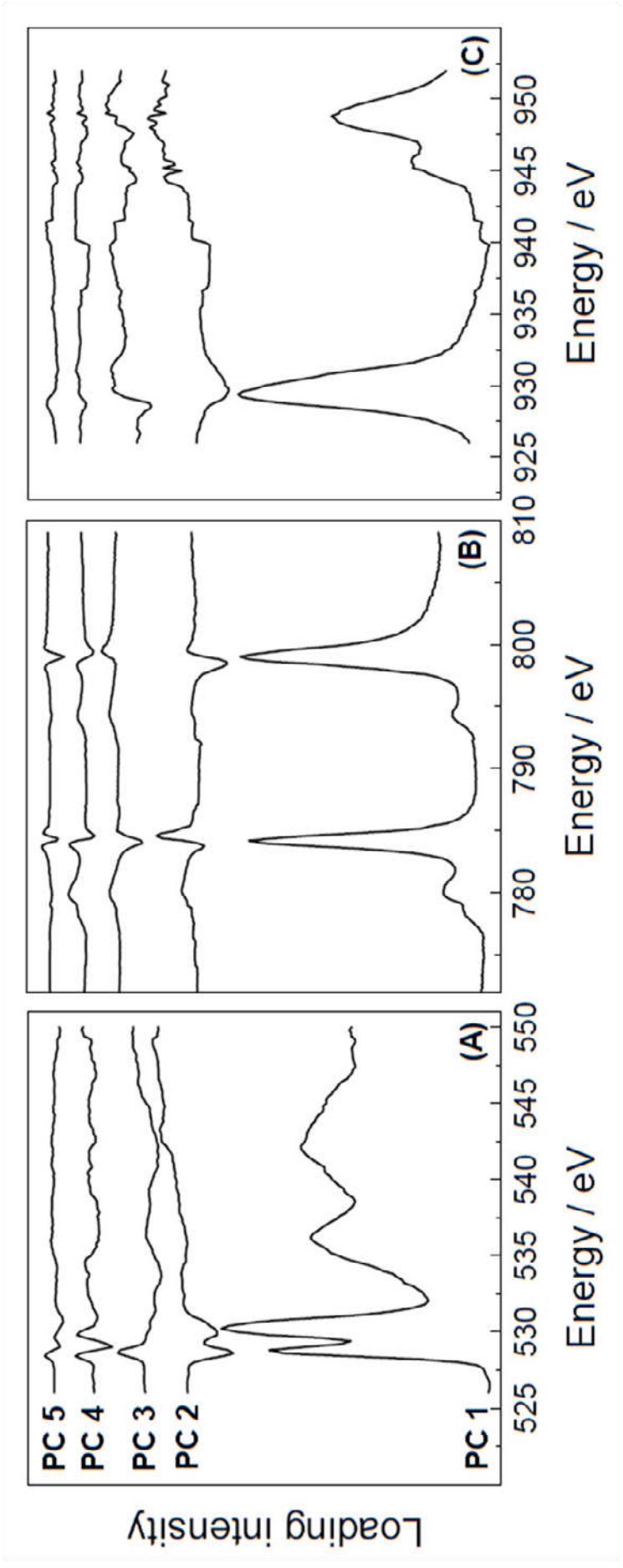

Loading intensity
PC 5
PC 4
PC 3
PC 2
PC 1
(A)
525 530 535 540 545 550
Energy / eV
(B)
780 790 800
Energy / eV
(C)
810 925 930 935 940 945 950
Energy / eV

Fig. 3

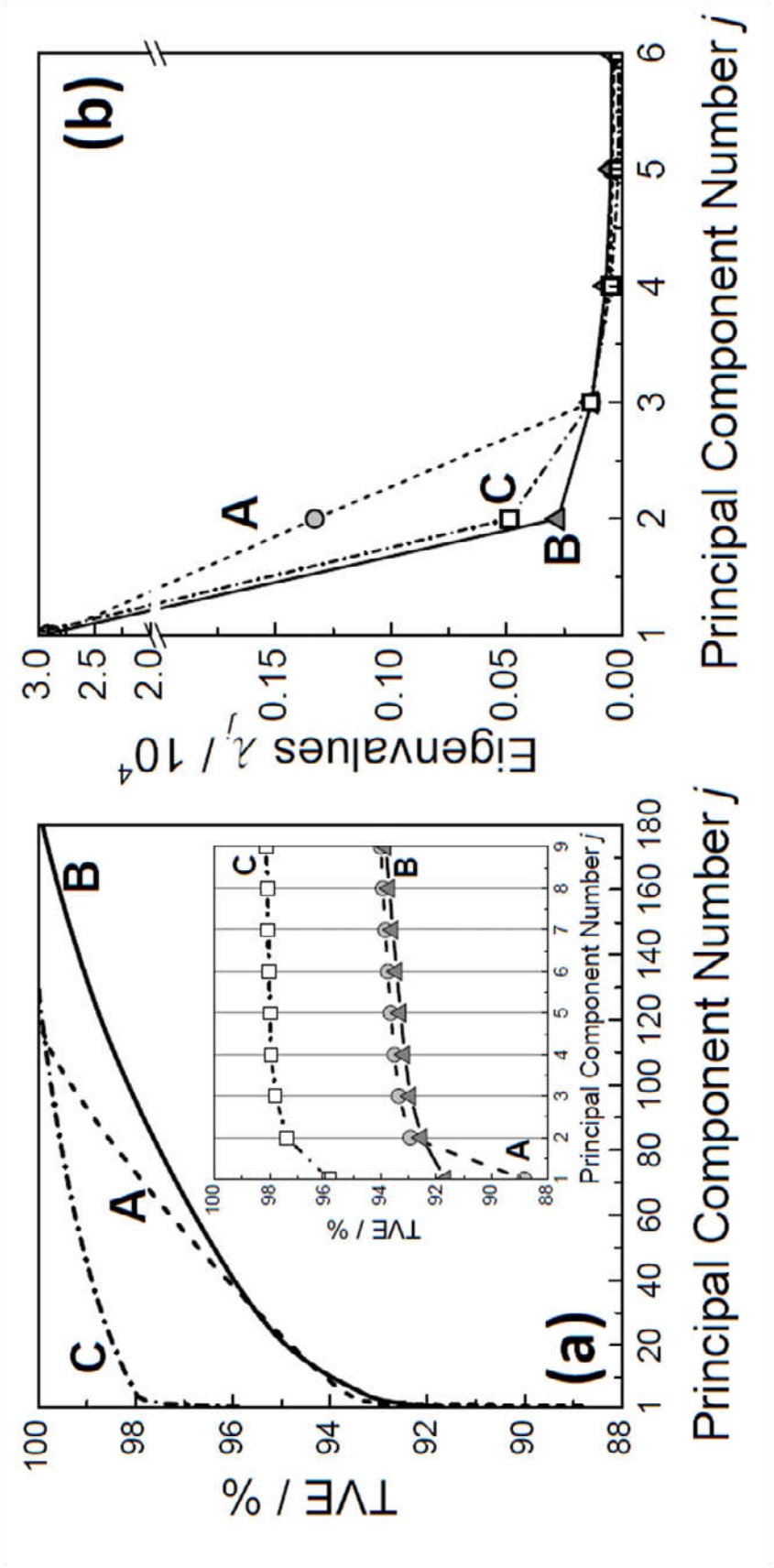

(a)
TVE / %
Principal Component Number j
A
B
C
(b)
Eigenvalues λ_j / 10^4
Principal Component Number j
A
B
C

Fig. 4

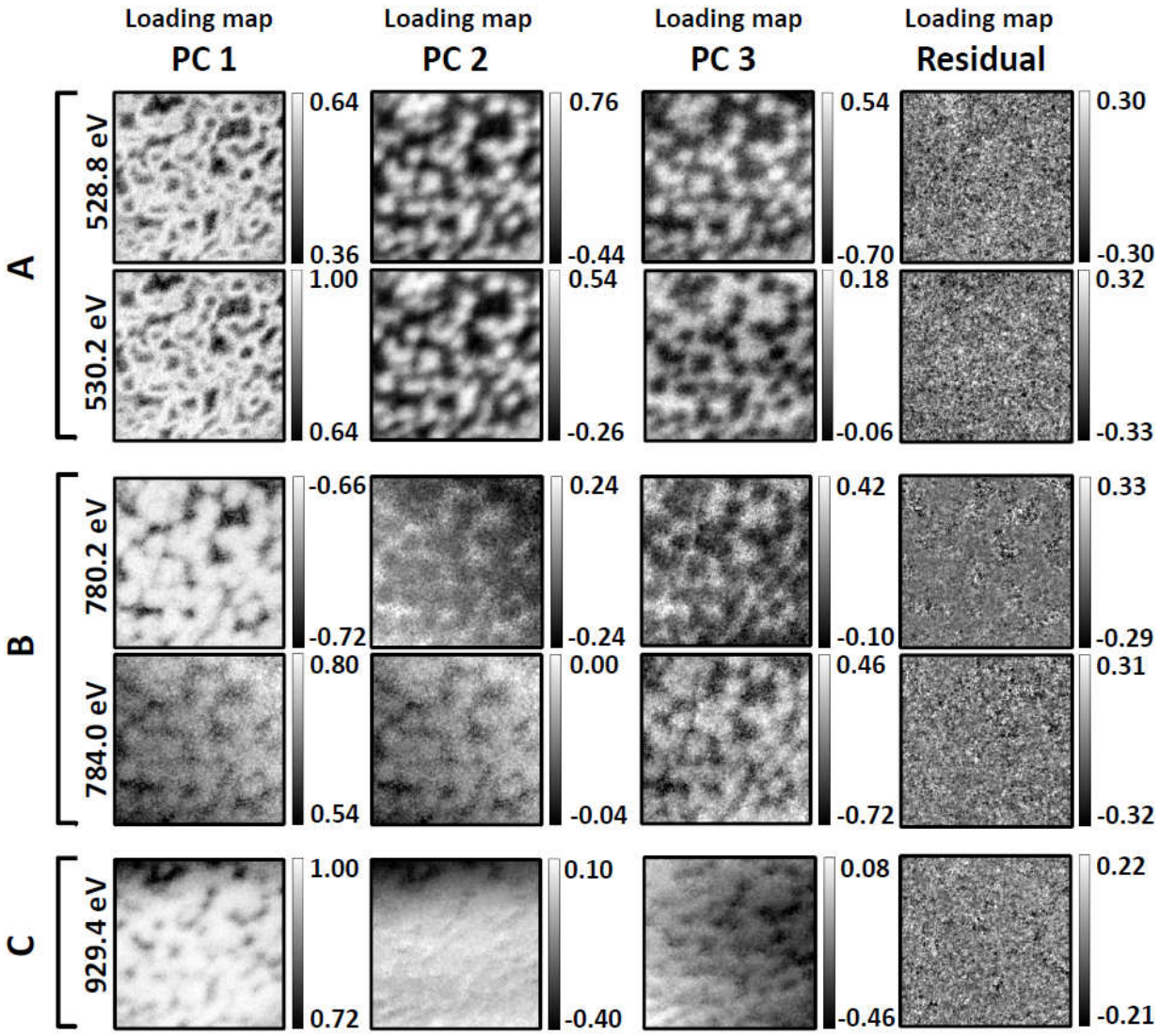

Loading map
PC 1
Loading map
PC 2
Loading map
PC 3
Loading map
Residual
A
528.8 eV
0.64
0.36
0.76
-0.44
0.54
-0.70
0.30
-0.30
530.2 eV
1.00
0.64
0.54
-0.26
0.18
-0.06
0.32
-0.33
B
780.2 eV
-0.66
-0.72
0.24
-0.24
0.42
-0.10
0.33
-0.29
784.0 eV
0.80
0.54
0.00
-0.04
0.46
-0.72
0.31
-0.32
C
929.4 eV
1.00
0.72
0.10
-0.40
0.08
-0.46
0.22
-0.21

Fig. 5

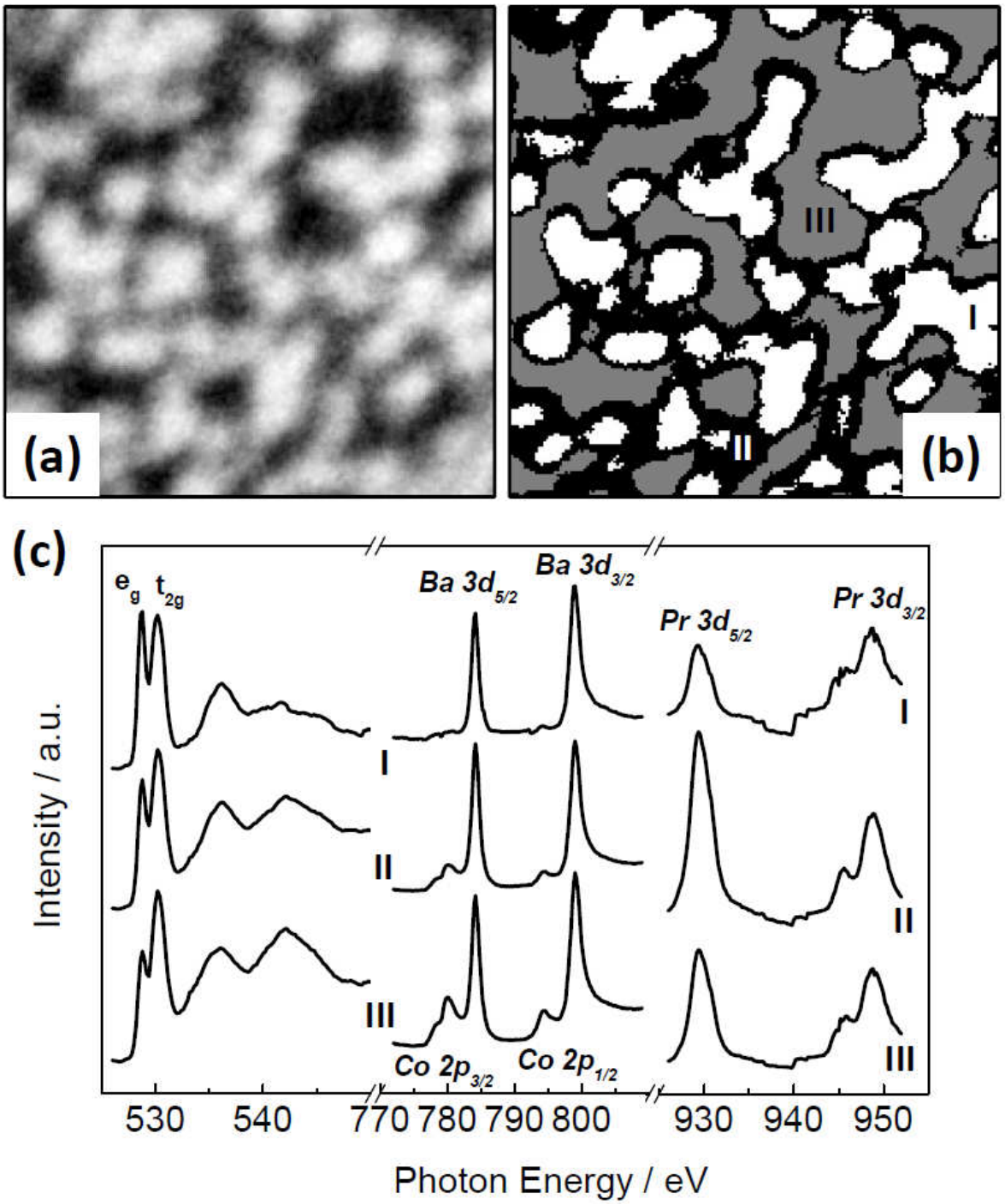

(a)
(b)
III
I
II
(c)
Intensity / a.u.
e_g t_2g
Ba 3d_5/2
Ba 3d_3/2
Pr 3d_5/2
Pr 3d_3/2
I
II
III
Co 2p_3/2
Co 2p_1/2
530
540
770
780
790
800
930
940
950
Photon Energy / eV

Fig. 6

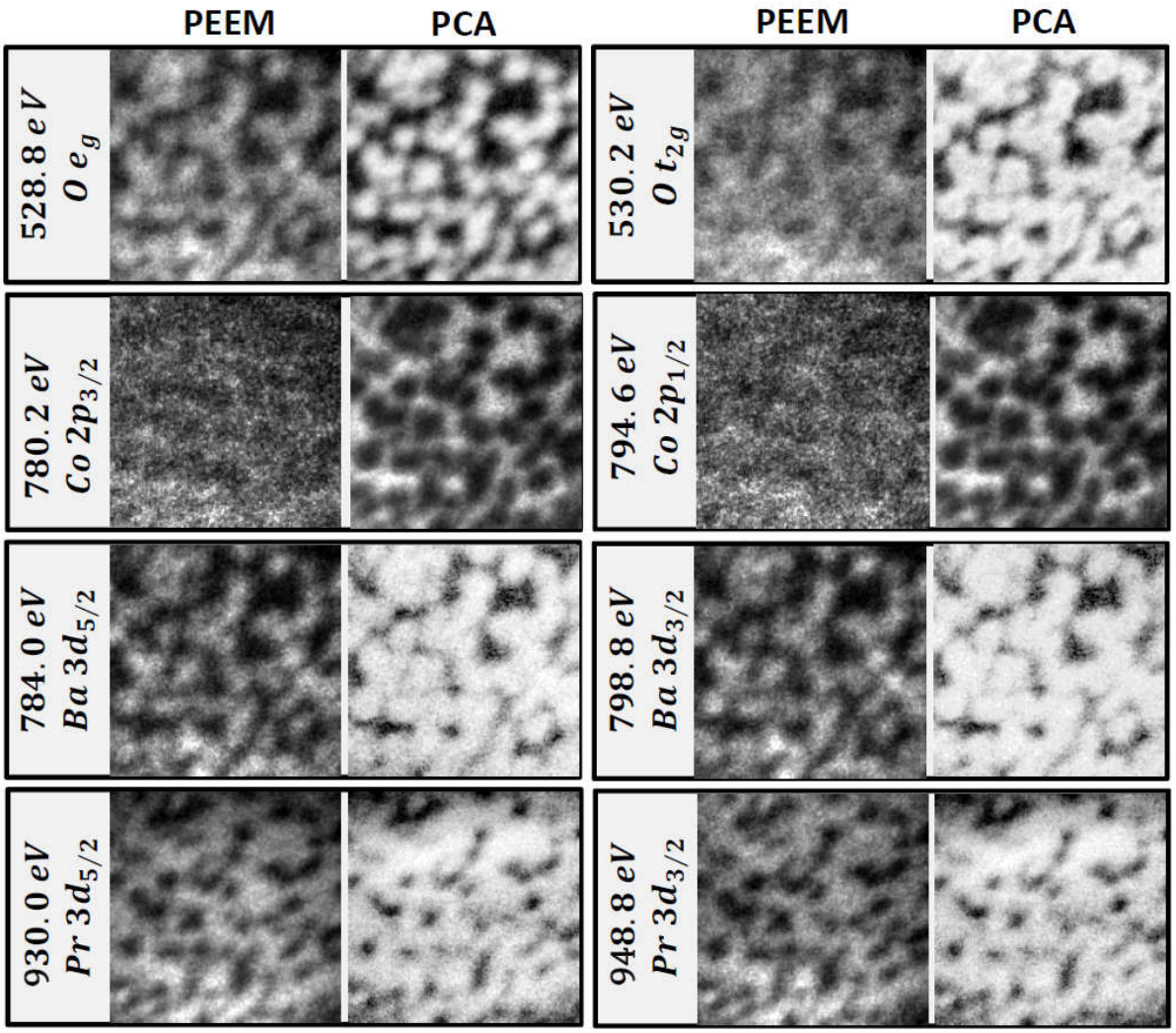

PEEM
PCA
PEEM
PCA
528.8 eV
O e_g
780.2 eV
Co $2p_{3/2}$
784.0 eV
Ba $3d_{5/2}$
930.0 eV
Pr $3d_{5/2}$
530.2 eV
O t_{2g}
794.6 eV
Co $2p_{1/2}$
798.8 eV
Ba $3d_{3/2}$
948.8 eV
Pr $3d_{3/2}$

Fig. 7