Supporting Information

Unveiling the Chemical Composition of Halide Perovskite Films using Multivariate Statistical Analyses

Stefania Cacovich[1§], Fabio Matteocci[2], Mojtaba Abdi-Jalebi[3], Samuel D. Stranks[3], Aldo Di Carlo[2], Caterina Ducati[1], Giorgio Divitini[1]*.*

[1] Department of Materials Science and Metallurgy, University of Cambridge, 27 Charles Babbage Road, Cambridge CB3 0FS, United Kingdom.

[2] Centre for Hybrid and Organic Solar Energy (CHOSE), University of Rome Tor Vergata, via del Politecnico 1, Rome 00133, Italy.

[3]Cavendish Laboratory, University of Cambridge, JJ Thompson Avenue, Cambridge CB3 0HE, United Kingdom.

Present Address: § Institut Photovoltaïque d'Île-de-France (IPVF) 30, Route départemental 128, 91120 Palaiseau, France.

Corresponding Author

*stefania.cacovich@ipvf.fr (S.C.); *gd322@ cam.ac.uk (G.D.).

Device Fabrication

Fluorine doped Tin Oxide (FTO)-coated glasses (Pilkington, 8 Ωcm -1, 25x25 mm^2) were etched with an infrared Nd:YVO_4 laser beam. The substrate was cleaned in an ultrasonic bath, by using a three-step procedure (10 minutes each step) with de-ionized water, acetone and isopropanol. A 50 nm-thick patterned blocking TiO_2 layer (BL-TiO_2) layer was deposited onto the patterned FTO following a reported procedure by using spray pyrolysis. TiO_2 paste (18 NR-T paste, Dyesol) diluted with ethanol (1:6 w/w) was deposited over the BL-TiO_2 surface by spin coating (2000 rpm for 20 s) and sintered at 450 °C for 30 min to obtain a titania mesoporous structure. The TiO_2-coated substrate was then transferred into a glovebox. A solution 1.1 M of PbI_2 powder (Sigma Aldrich) and 1.1 M of CH_3NH_3I (Dyesol) was dissolved in DMSO and stirred at room temperature overnight. 110 µl of the aforementioned solution was spin coated on the TiO_2 substrate at 1000 rpm for 10 s and 5000 rpm 30 s. 200 µl of toluene were poured on the device 10 s prior the end of the second ramp. In order to obtain a full perovskite conversion the film was annealed at 100 °C for 40 minutes. A doped Spiro-OMeTAD (2,20,7,70-tetrakis(N,N-dip-methoxyphenylamine) 9,9'-spirobifluorene) solution (60 mM, Lumtec) in chlorobenzene was spun at 2000 rpm for 20 s. The molar ratio the cobalt additive (FK102)/Spiro-OMeTAD was 0.03. Finally, the samples were transferred into a high vacuum chamber (10^{-6} mbar) to thermally evaporate Au back contacts having a nominal thickness of 80 nm and leaving an active area of 1.05 cm^2.

Thin Film Fabrication

We purchased the organic cation salts from Dyesol; the Pb compounds from TCI and CsI and KI from Alfa Aesar. Spiro-OMeTAD was purchased from Borun Chemicals and used as

received. Unless otherwise stated, all other materials were purchased from Sigma-Aldrich. We prepared the triple-cation-based perovskite Cs0.06FA0.79MA0.15Pb(I0.85Br0.15)3 films by mixing PbI2 (1.2 mol.L^{-1}), FAI (1.11 mol.L^{-1}), MABr (0.21 mol.L^{-1}) and $PbBr_2$ (0.21 mol.L-1) in a mixture of anhydrous DMF:DMSO (4:1, volume ratio). We then added 5 volume percent CsI stock solution (1.5 mol.L^{-1} in DMSO) to the prepared double cation perovskite solution. The potassium iodide powders was first dissolved in mixed solutions of DMF/DMSO (4:1, volume ratios) to make stock solutions each of 1.5 mol.L^{-1}. The resulting KI solution was added into the triple cation perovskite solution to achieve the desired additive ratios. We then deposited the perovskite films by spin coating the prepared solutions using a two-step program at 2000 and 6000 rpm for 10 and 30 seconds, respectively, and dripping 150 µL of chlorobenzene 30 seconds after the start of the spinning process. We then annealed the films at 100°C for 1 hour. All solutions and films were prepared in a nitrogen-filled glove box.

Electron Microscopy Characterisation

Cross-sectional specimens were extracted as lamellae for TEM analysis using focused ion beam (FIB) milling. EDX data were acquired using a FEI Tecnai Osiris TEM equipped with a high brightness Schottky X-FEG gun and a Super-EDX system composed of four silicon drift detectors, each approximately 30 mm^2 in area and arranged symmetrically around the optical axis to achieve a collection solid angle of 0.9 sr. Spectrum images were acquired with an acceleration voltage of 200 kV, a probe current of 0.6 nA and a spatial sampling of 10 nm per pixel.

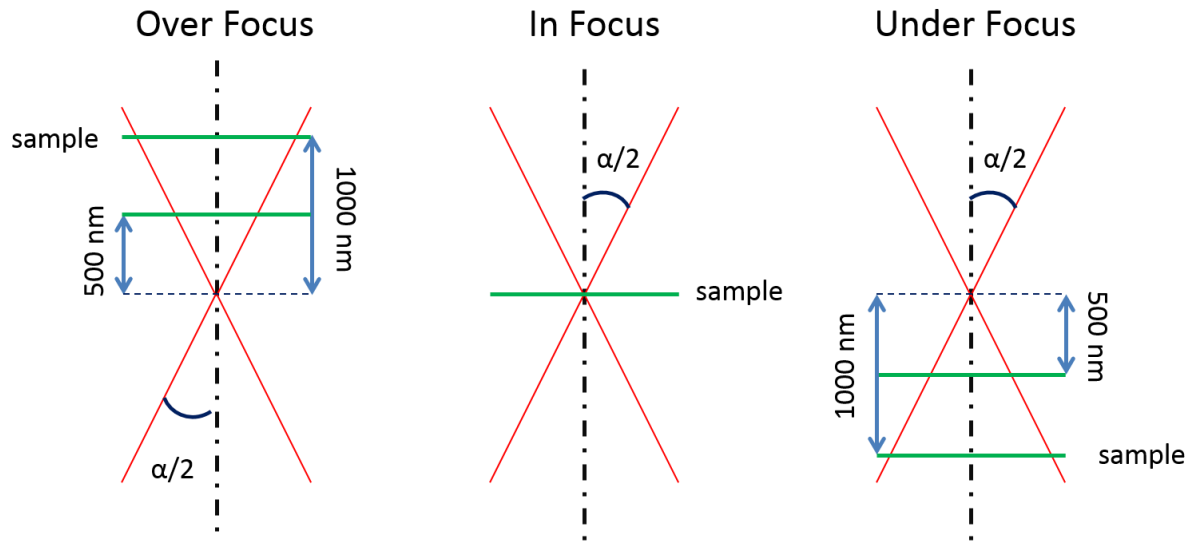

Figure S1. Variation in the probe size for different focus conditions. The convergence angle α/2 is equal to 13 mrad.

Principal Component Analysis (PCA) is one of the most popular multivariate analysis methods and it is used to identify correlations in a given dataset. In particular, PCA aims to find the covariance structure of a set of components, i.e the directions that explain the highest data variation. These directions are called principal components, from which the name of the technique. The simple case of a dataset of two variables (X,Y) is shown in Figure S2.

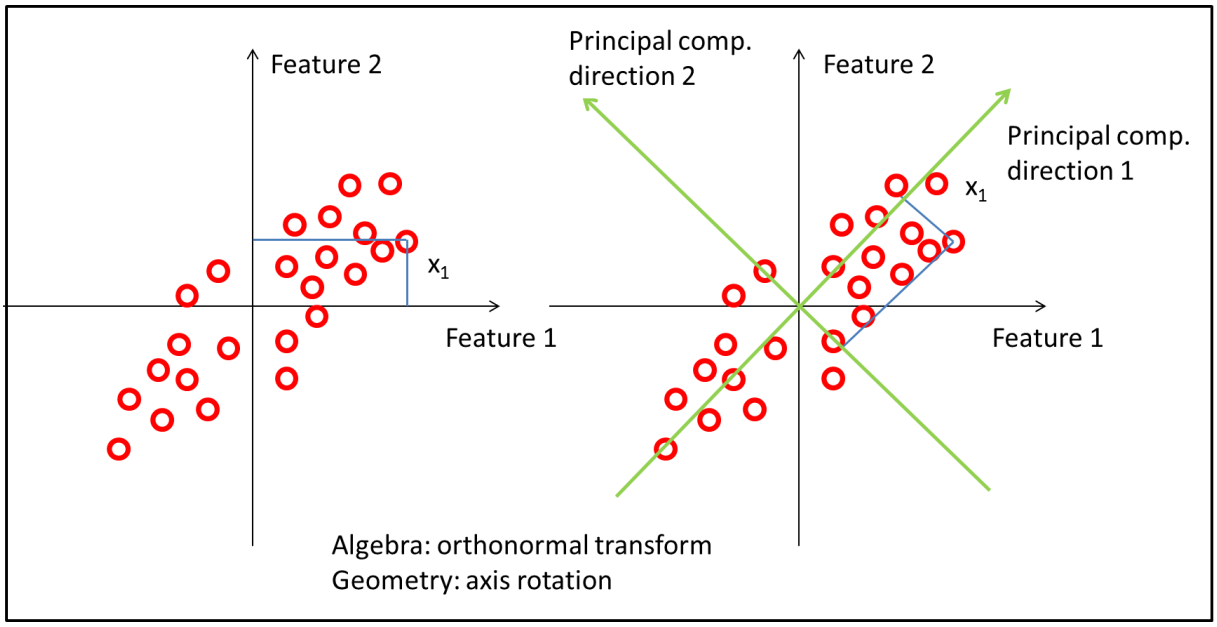

Figure S2. Data set of two variables. Direction 1 is responsible for most of the variance. Direction 2 is perpendicular to direction 1.

Applying PCA two components are obtained (green arrows).

To rigorously define PCA, we can consider a random vector

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_P \end{pmatrix}$$

and its population variance-covariance matrix E:

$$E(X) = var(X) = \begin{pmatrix} \sigma_1^2 & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_p^2 \end{pmatrix}$$

where

$\sigma_{jj} = E([X_j - \mu_j]^2)$ is the population variance of the j-th variable

$\sigma jk = E([X_j - \mu_j][X_k - \mu_k])$ is the population covariance between the j-th and k-th variables

$\mu_j = E(X_j)$ is the population mean of the j-th variable

The following linear combinations are linear regressions, predicting Y_i from X_1, X_2, ... , X_p. Y_i is function of random data, therefore also Yi is random.

$Y_1 = e_{11}X_1 + e_{12}X_2 + \cdots + e_{1p}X_p$

$Y_2 = e_{21}X1 + e_{22}X_2 + \cdots + e_{2p}X_p$

$Y_p = e_{p1}X1 + e_{p2}X_2 + \cdots + e_{pp}X_p$

Y_i has a population variance of

$$\mathrm{var}(Y_i) = \sum_{k=1}^{p}\sum_{l=1}^{p} e_{ik}e_{il}\sigma_{il}$$

Y_1 is the first principal component and is the linear combination of x variables which has the maximum variance. All the coefficients e_{11}, e_{12}, ... , e_{1p} will be defined in order to maximise the variance, with the additional constraint that the sum of the squared coefficients is unitary. This constraint is required in order to obtain a unique answer. The second principal component (Y_2) is the linear combination of x-variables that accounts for as much of the remaining variation as possible. There is no correlation between first and second component. All subsequent principal components are characterised by the same properties: being linear combinations that account for as much of the remaining variation as possible and having no correlations with the other principal components. Considering the problem from a geometrical point of view, PCA projects the data along the directions where the data varies the most as shown in Figure S2. The first direction is defined by v_1 corresponding to the largest eigenvalue d_{12}, whereas the second direction is defined by v_2 corresponding to the second largest eigenvalue d_{22}. The variance of the data along the principal component directions is related to the magnitude of the eigenvalues.

In a multidimensional dataset with p variables the data are arranged in a matrix C. According to the matrix notation the system can be represented by:

$C = AT^X$

where C is the principal components matrix, X is a m×p matrix having orthogonal columns, called score matrix (the image) and A is a n×p matrix having orthonormal columns, called loading matrix (the spectra). The rows of matrix A are the eigenvectors of the covariance matrix. The elements of this matrix, called loadings, are the weights a_{ij}. The elements in the diagonal of the covariance matrix are called eigenvalues. The total variance in the data is defined as the sum of the eigenvalues. Although the basic aim in PCA is to decorrelate the data by performing an orthogonal projection, this method is also often employed to reduce the dimension of the data removing unwanted components in the

signal. PCA is usually applied in the post processing analysis of EDX spectral datasets with two different goals. Firstly, it can be employed as a data mining technique to interpret the principal components in a qualitative analysis. Secondly, PCA can act as a noise-filter to improve the accuracy of quantitative analysis. However, the main drawback of this technique is that loadings and scores can have negative peaks, therefore they have no direct physical meaning.

Non-negative matrix factorisation (NMF) is a computational method aimed to approximate high dimensional data, where the data are comprised of non-negative components. Similarly to PCA the main goal of non-negative matrix factorization (NMF) is to explain the observed data by using only a limited number of components. However, NMF performs the decomposition imposing different constrains. First of all, both the matrix representing the basis components and the matrix of mixture coefficients are imposed to have non-negative entries. Additionally, there are no constraints of orthogonality or independence on the basis components, but the components are allowed to overlap. The fact that the matrix factors are non-negative results in a direct and intuitive interpretation of decomposition results. The constraint of non-negativity induces sparsity. A sparse matrix is a matrix in which most of the elements are zero. Therefore, the mixture coefficients have only a few non zero-entries, leading to more compact and local representation. To mathematically define the problem, we can consider r >0 be an integer, and X a matrix with n rows - the measured features - and p columns - the samples - with only non-negative entries. Non-negative Matrix Factorization consists in finding an approximation:

$X \approx WH$

where W, H are $n \times r$ and $r \times x_p$ non-negative matrices, respectively.

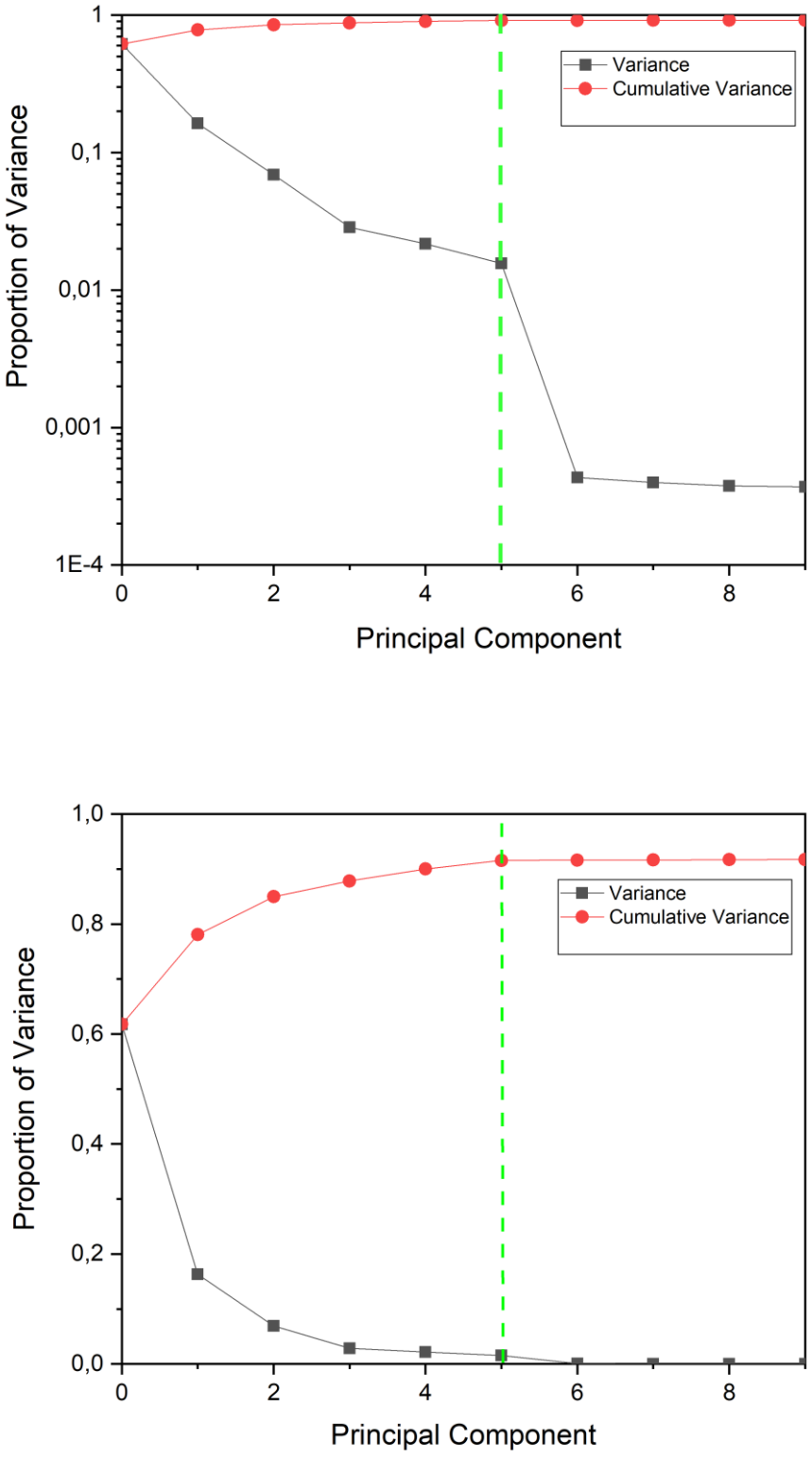

Figure S3. Variance and cumulative variance on the dataset analysed on Fig 2 represented in logarithmic and linear scale.

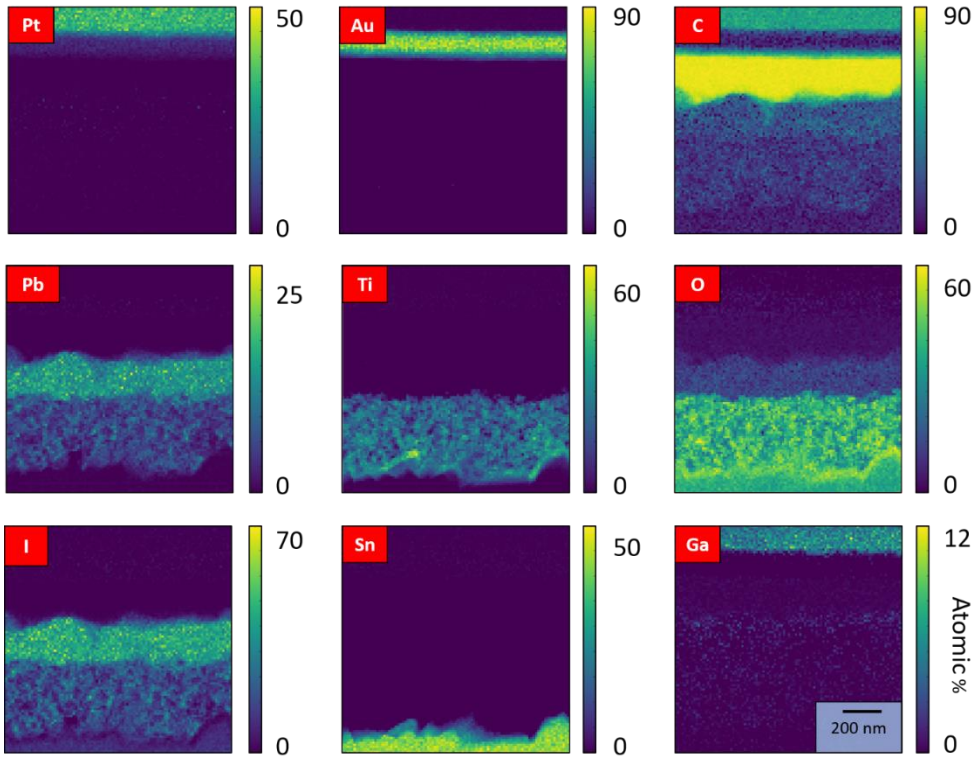

Figure S4. EDX elemental maps calculated on the raw data.

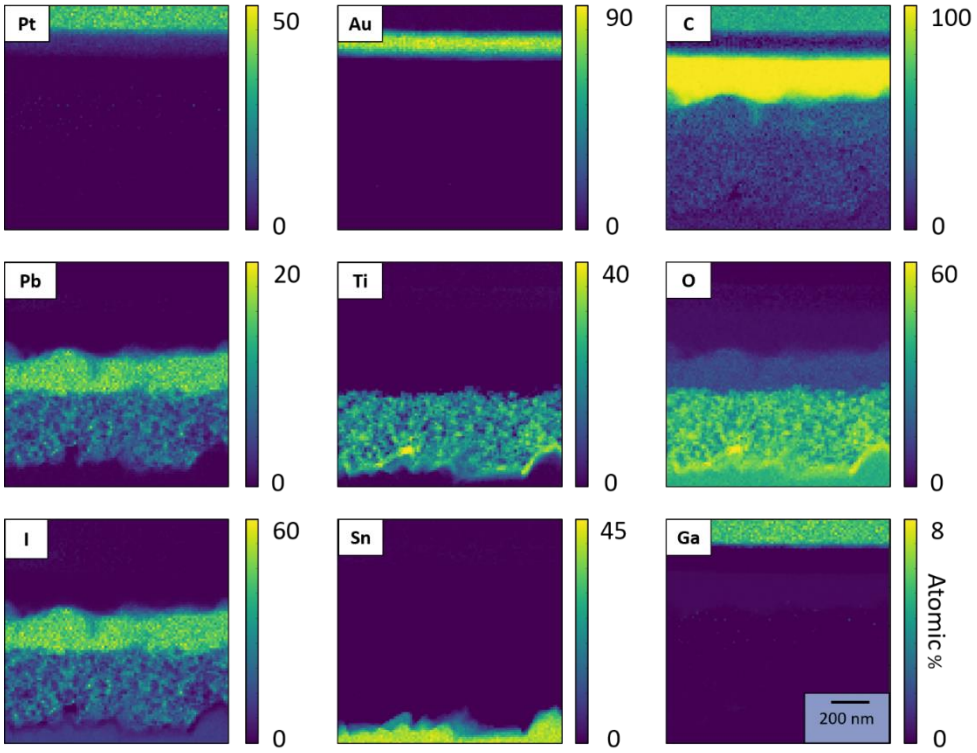

Figure S5. EDX elemental maps calculated after PCA denoising.

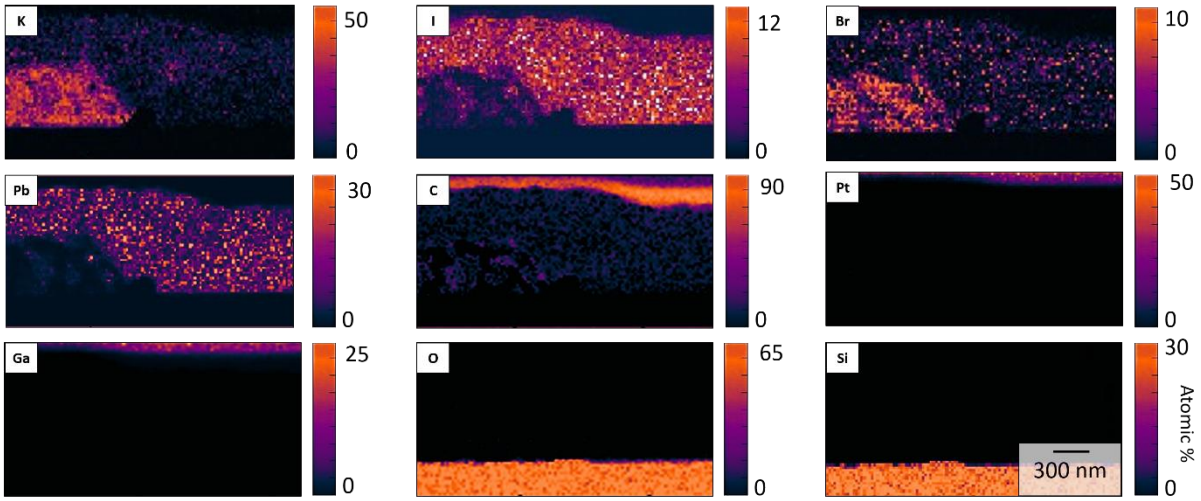

Figure S6. EDX elemental maps calculated on the raw data.