SUPPLEMENTAL MATERIAL

Alignment of electron optical beam shaping elements using a convolutional neural network

E. Rotunno[1], A.H. Tavabi[2], P. Rosi[3], S. Frabboni[1,3], P. Tiemeijer[4], R.E. Dunin-Borkowski[2], V. Grillo[1]

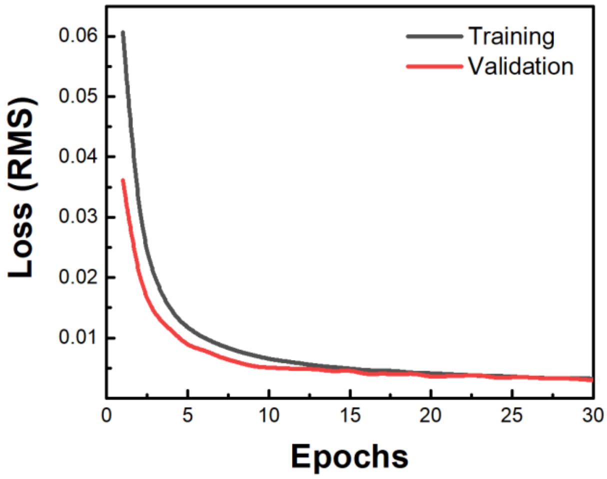
1) *Istituto di Nanoscienze - CNR, 41125 Modena, Italy*
2) *Ernst Ruska-Centre for Microscopy and Spectroscopy with Electrons, Forschungszentrum Jülich, 52425 Jülich, Germany*
3) *University of Modena and Reggio Emilia, via G. Campi 213/A, 41125 Modena, Italy*
4) *Thermo Fisher Scientific, PO Box 80066, 5600 KA Eindhoven, The Netherlands*

S1. Range of the aberrations

Aberration	Lower limit	Higher limit
Defocus (df/f)	-0.13%	+0.13%
Size Mismatch (SM)	-20%	+20%
Shift x (shx)	-500 nm	+500 nm
Shift y (shy)	-500 nm	+500 nm
Rotation (β)	-5°	+5°

Images in the training dataset have been simulated by randomly selecting aberration value in the ranges here reported. Because these ranges greatly differ in units and sensitivities, all the training labels were normalized in the symmetric range (-1,1) before being fed to the learning algorithm.

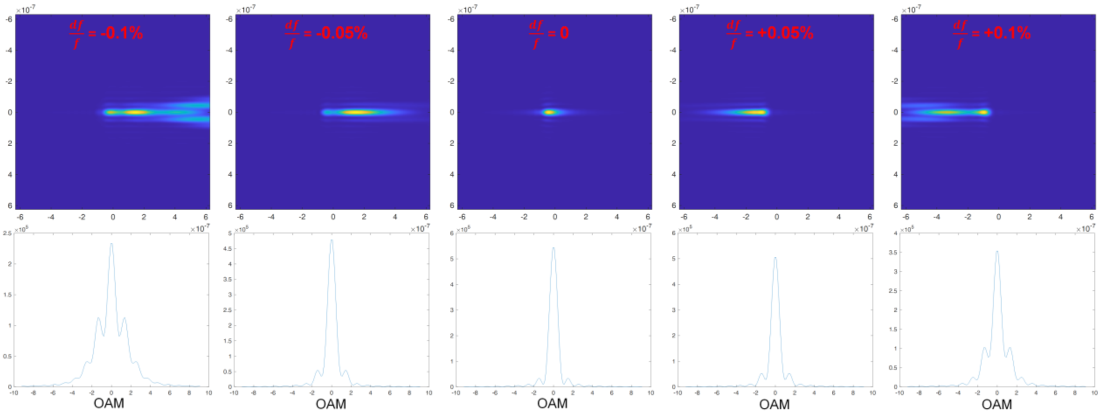
S2. Training and validation loss curves



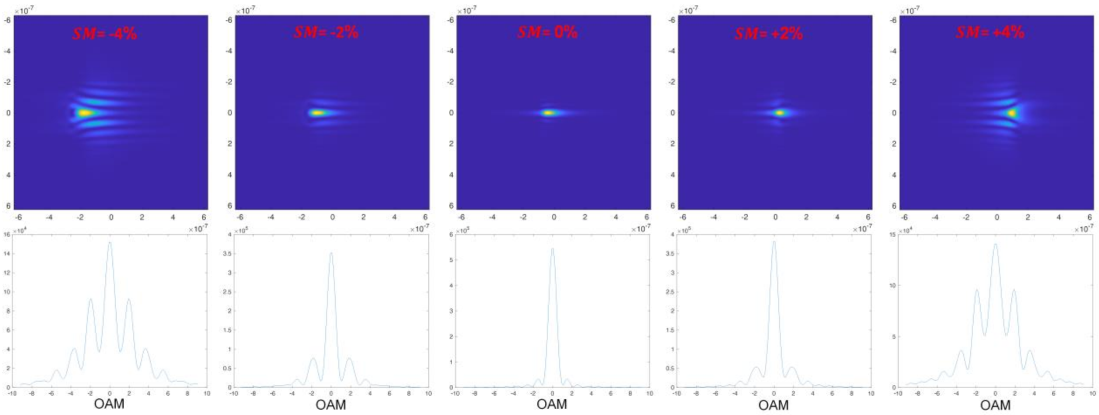
Here we report the evolution of the loss function during the training. It has been plotted both for the training dataset (black curve) and the validation (red curve) dataset.
The training loss is intended as the average over the entire epoch, while the validation loss is evaluated at the end of each epoch. Therefore the training losses are expected to be higher than the validation ones. Typically the two curves are translated by half an epoch.
It can be noted that after 20 epochs the two curves converge to the same values. This is one of the evidences that the neural network is overfitting. The network begins to memorize features in the training dataset losing its generalization ability. Training further is not only useless (the learning curve is almost flat) but even harmful.

S3. Sorter misalignments

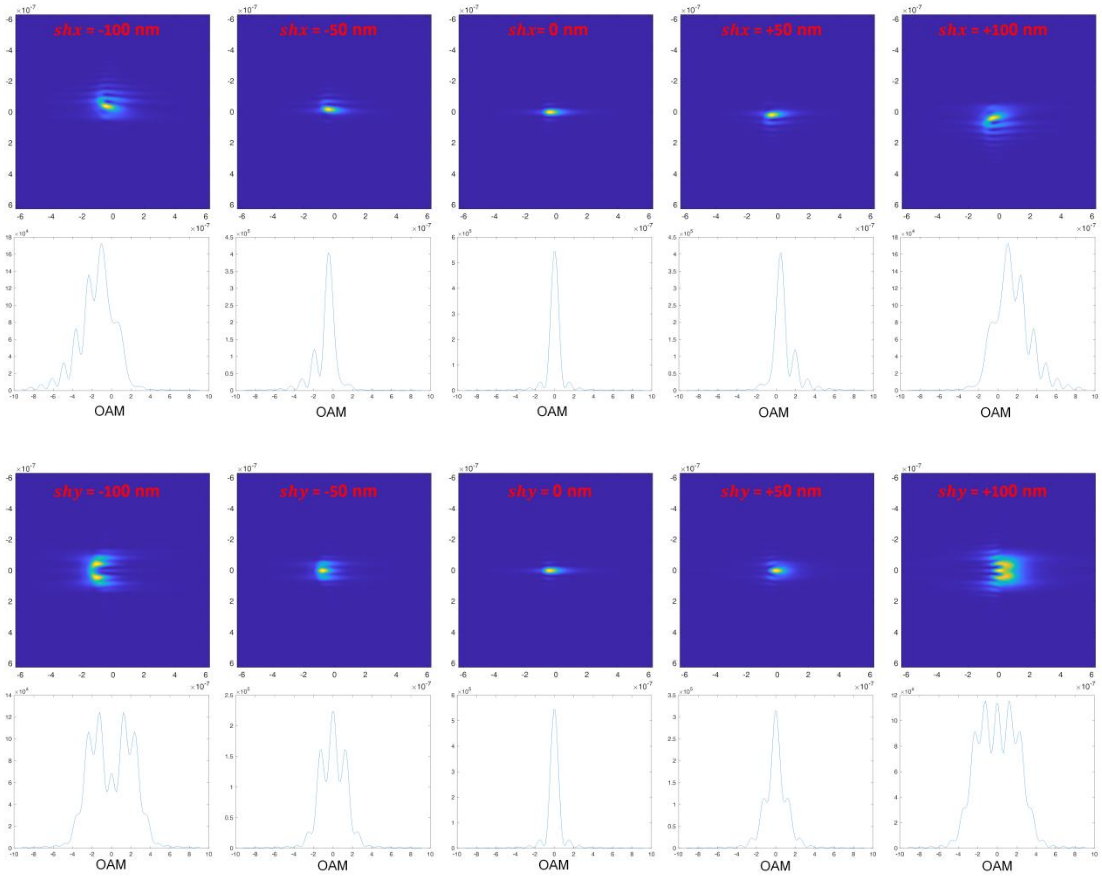
We systematically studied the effect of the different aberration values on the OAM spectrum.



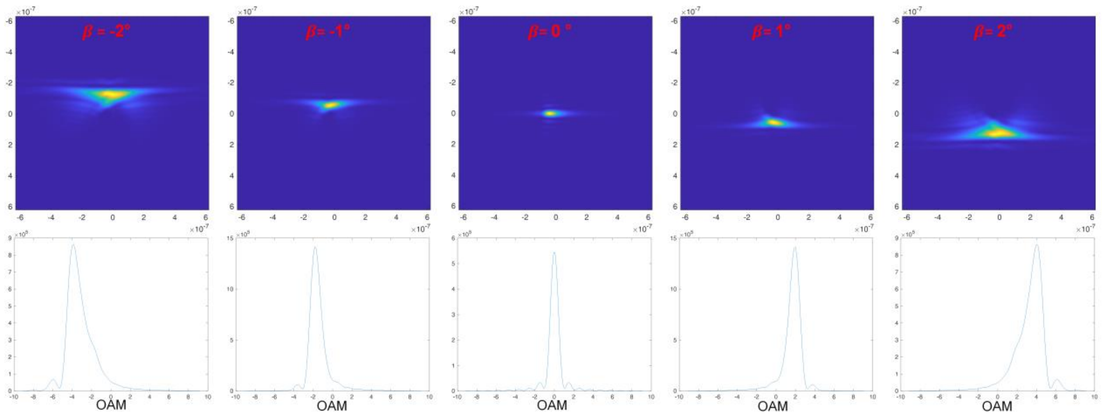
Defocus: among the different aberrations studied in this article, defocus is the less problematic in term of OAM resolution because its effect is mainly in the radial direction. Values below 0.05% are deemed acceptable.



Size mismatch: size mismatch aberration is directly related to the lensing effect of the sorter 1. It manifests itself in an intense background fringe. We estimate a tolerance of about 2%.



Beam shift: the effect of beam shift is different in the two directions. We define the x axis as the direction parallel to the array of needle composing the second element of the sorter. Y axis is orthogonal. The effect of a small rigid shift along the x direction of 50 nm does not drastically reduce the OAM resolution. A similar shift in the y direction is instead much more detrimental. In general we set the tolerance of the shift to 30 nm.



Rotation: the effect of a different orientation between the two elements of the sorter is here reported. The alignment should be better than 1° for this aberration to have a tolerable effect.