PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Approximate nonnegative matrix factorization algorithm for the analysis of angular differential imaging data

Arcidiacono, C., Simoncini, V.

C. Arcidiacono, V. Simoncini, "Approximate nonnegative matrix factorization algorithm for the analysis of angular differential imaging data," Proc. SPIE 10703, Adaptive Optics Systems VI, 1070331 (11 July 2018); doi: 10.1117/12.2311681

SPIE.

Event: SPIE Astronomical Telescopes + Instrumentation, 2018, Austin, Texas, United States

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 23 Nov 2020 Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Approximate nonnegative matrix factorization algorithm for the analysis of angular differential imaging data

Arcidiacono C.[a,c] and Simoncini V.[b]

[a]INAF – Osservatorio Astrofisico e scienza dello Spazio di Bologna, Via P. Gobetti 93/3, 40129 Bologna, Italy
[b]Università degli Studi di Bologna, Dipartimento di Matematica, Piazza di Porta S. Donato, 5, 40126 Bologna BO, Italy
[c]ADONI – Laboratorio Nazionale di Ottica Adattiva, Italy

ABSTRACT

The angular differential imaging (ADI) is used to improve contrast in high resolution astronomical imaging. An example is the direct imaging of exoplanet in camera fed by Extreme Adaptive Optics. The subtraction of the main dazzling object to observe the faint companion was improved using Principal Component Analysis (PCA). It factorizes the positive astronomical frames into positive and negative components. On the contrary, the Nonnegative Matrix Factorization (NMF) uses only positive components, mimicking the actual composition of the long exposure images.

Keywords: Extreme Adaptive Optics, Exoplanets, Angular Differential Imaging

1. INTRODUCTION

The front-line of the technological advancement for astronomical imaging can be reduced to the achievement of a better resolution and a better contrast. Larger and larger telescopes are being deployed in space or built on the ground. In both cases the major aim is to reduce the fundamental limit given by the diffraction, enhancing both the theoretical resolution and contrast. High contrast imaging in astronomy is a key tool for most of the science cases of future Extremely Large Telescopes.[1–3] In the following, we study the case of high contrast technique such the Angular Differential Imaging (see section 2) used to exploit the power of extreme adaptive optics system. In this framework we propose the use of Nonnegative Matrix Factorization algorithms[4,5] to improve the data reduction, see section 3.

2. ANGULAR DIFFERENTIAL IMAGING

The angular differential imaging[6] (ADI) is a data reduction and data collection strategy that improves contrast in astronomical imaging. An example is direct imaging of exoplanet in cameras fed by Extreme Adaptive Optics modules. Considering the case of Adaptive Optics (AO) observations from a ground based telescope, the ADI uses a large set of short exposures taken in pupil tracking mode, see Figure 1. In its original version ADI the images are median subtracted and recombined according to the actual field rotation experienced by the particular frame. Following ADI concept, the images (star+planets) centered on the main star rotate with the sky (pupil tracking mode). The idea is subtracting the median of the images in order to wipe out the footprint of the star (the PSF), counter-rotating the residuals containing the planets.

Using mathematical notation, we can write each image as a vector and pile up them on a matrix, see Figure 2. The residual images $\mathbf{R}$ to be counter-rotated and summed in the ADI process can be written as:

$$\mathbf{R} = \sum_i \mathbf{A}_i - \hat{A} \tag{1}$$

Further author information: (Send correspondence to Carmelo Arcidiacono)
C.A.: E-mail: carmelo.arcidiacono@inaf.it, Telephone: +39 051 6357 316

Adaptive Optics Systems VI, edited by Laird M. Close, Laura Schreiber, Dirk Schmidt, Proc. of SPIE Vol. 10703, 1070331 · © 2018 SPIE
CCC code: 0277-786X/18/$18 · doi: 10.1117/12.2311681

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 23 Nov 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

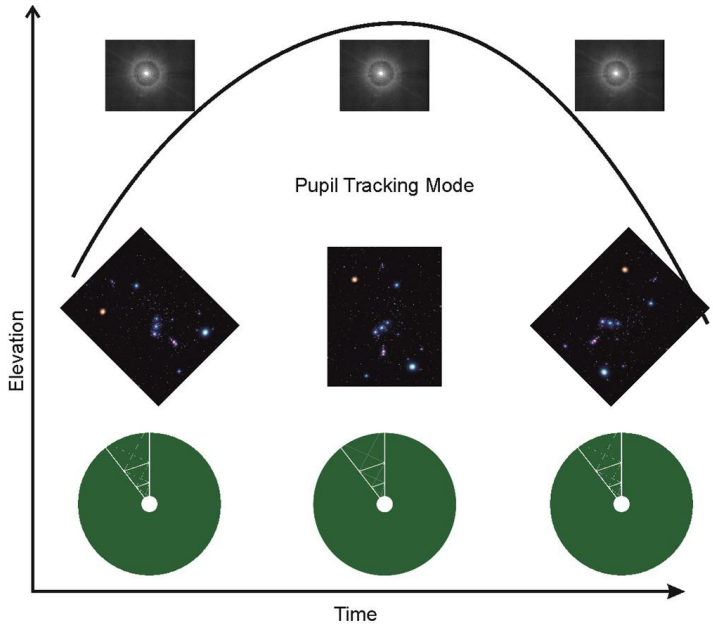


Figure 1. On Alt-Az telescope field rotates as the object crosses the sky. Typically cameras are counter-rotated, in the ADI are not and the technique takes advantage of the rotation of the object with respect to the PSF fixed to the telescope.

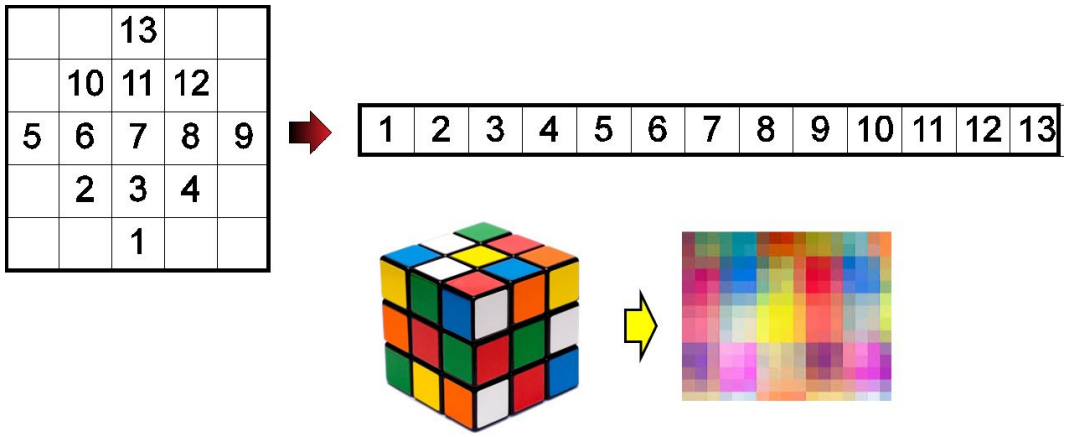


Figure 2. From each of the original frames we extract a region, and write it on a vector. The vectors are ordered on matrix.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 23 Nov 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

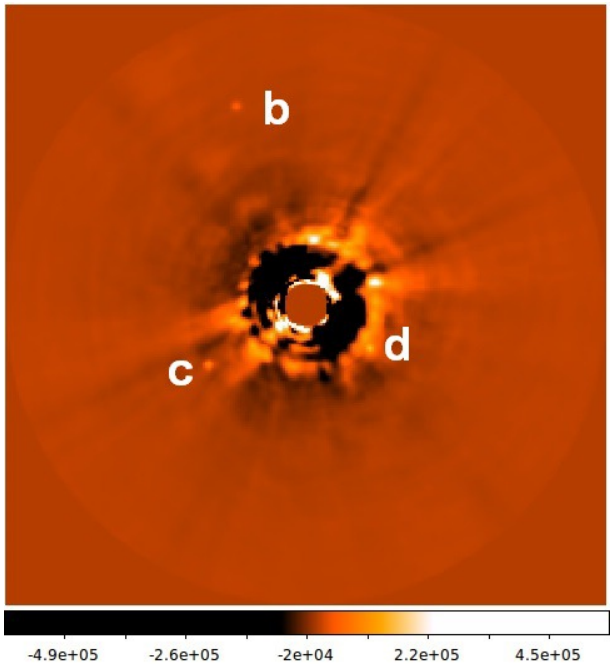


Figure 3. The HR8799 H-band image composed by the residual of the ADI technique. Three out of four planets (b, c, d) are visible. The detection of planet "d" has poor SNR, being easily confused with speckles residuals.

Where $\hat{A}$ is the median of the available images.

The subtraction of the main dazzling object to observe the faint companion was already improved by the LOCI algorithm[7] dividing images in subsections and obtaining, for each subsection, a linear combination of the reference images whose subtraction from the target image will minimize the noise. The subtraction was further improved by the Karhunen-Loève[8] (KL) image projection[9] (KLIP) algorithm. Both algorithms implement forms of Principal Component Analysis (PCA) to obtain the matrix of the residuals. We may factorize the matrix $\mathbf{A} \in \mathbb{R}^{\mathbf{m}\times\mathbf{n}}$, being m the number of useful pixels on each image and n the number of images, by using a general singular-value decomposition (SVD):

$$\mathbf{A} = \sum_{\mathbf{i=1}}^{\mathbf{n}} \sigma_\mathbf{i} \mathbf{u_i} \mathbf{v_i} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\mathbf{T} \tag{2}$$

and considering just the first k singular value we write the $\mathbf{A}$ approximation using the principal components $\mathbf{V_k} \in \mathbb{R}^{\mathbf{m}\times\mathbf{n}}$:

$$\mathbf{A_k} = \mathbf{U_k}\mathbf{\Sigma_k}\mathbf{V_k^T} \tag{3}$$

This factorization writes the matrix $\mathbf{A}$ into a matrix $\mathbf{A_k}$ with lower rank (k), projecting $\mathbf{A}$ on the new base $\mathbf{V}$. The base $\mathbf{V}$ is optimal, in the sense that the SVD is the factorization minimizing the residual $||\mathbf{R}||$:

$$\mathbf{R_k} = \mathbf{A} - \mathbf{A_k} \tag{4}$$

for every possible rank $k < n$. The residual $||\cdot||$ here is the Euclidean Norm, corresponding to the *rms*. Higher the rank used larger is the effective self-subtraction of planet companion around the parent star, see Figure 4.

In the following we take as data example the ADI sequences collected at LBT[10] during the Science Demonstration Time in the October 2011 using the PISCES camera[11] fed by the First Light Adaptive Optics (FLAO) classical AO module. The data set used is composed by $n = 1396$, 5sec, H band frames of HR8799[12] taken at LBT using PISCES+FLAO in pupil tracking mode.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 23 Nov 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

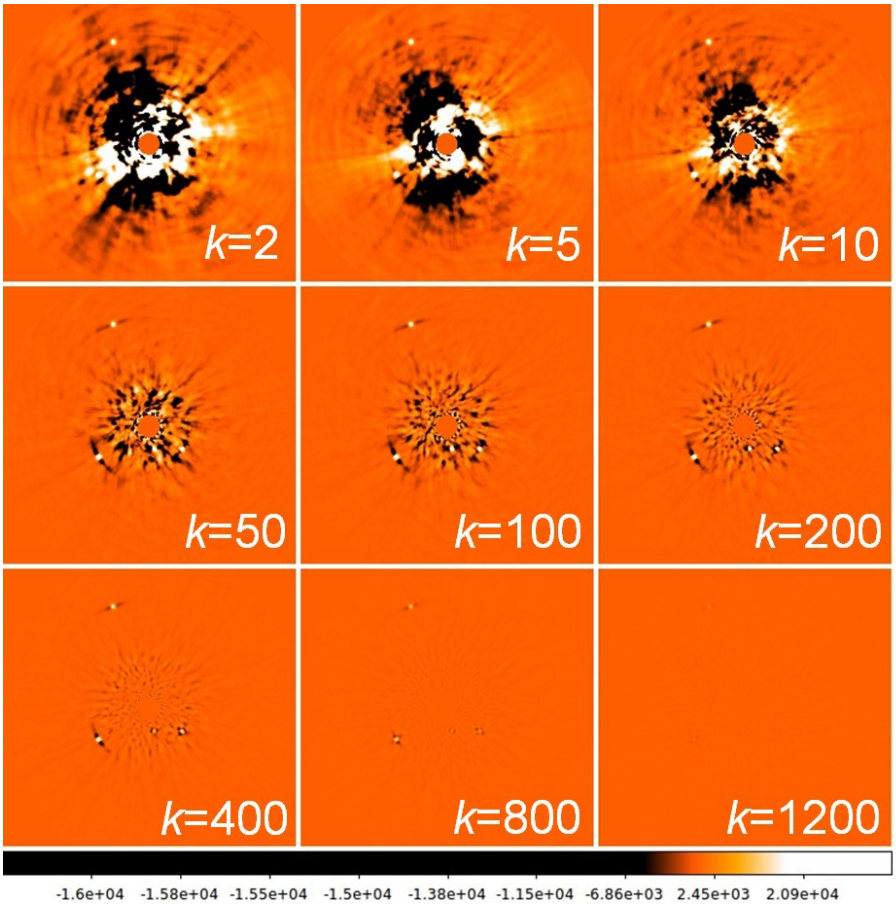


Figure 4. Increasing the rank of the matrix $\mathbf{A_k}$ we see that greatly improve speckles residuals. However, at the same time the flux of the companions is reduced more and more.

3. NONNEGATIVE MATRIX FACTORIZATION: IMPROVING PHOTOMETRY AND DETECTION

To detect a faint planet around a star we need to remove the main component image. A stable and known Point Spread Function (PSF) is mandatory to disentangle what is noise with respect to the true signals. Actually ground-based adaptive optics assisted telescopes or space telescopes produce pretty stable PSF. The Angular Differential Imaging (ADI) aims improving contrast in SCAO imaging o from Space Telescopes. In particular its better achievements have been pursued on the imaging of exoplanet on NIR camera feed by an Extreme AO module.

However, in both cases long lasting- and slow evolving- speckles due to the system optical path distortion decreases the Signal to Noise Ratio (SNR) and generates false positives much more than the pure Poisson photon noise.

Take for example the case of KLIP, the methods foresees to build a KL-base for the data projection starting from a data set different than the images to be reduced: in this way aiming to reduce the effect of the self subtraction. However the faint speckles bed below the PSF is different since gravity and thermoelastic flexure are different. The use of the same science data set to generate the projecting space greatly improves speckles subtraction, however may fail in the detection (take the case of LOCI were the position of the companion should be known in advance) since is still dubious if low SNR objects in the reconstructed are real objects or true detections.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 23 Nov 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

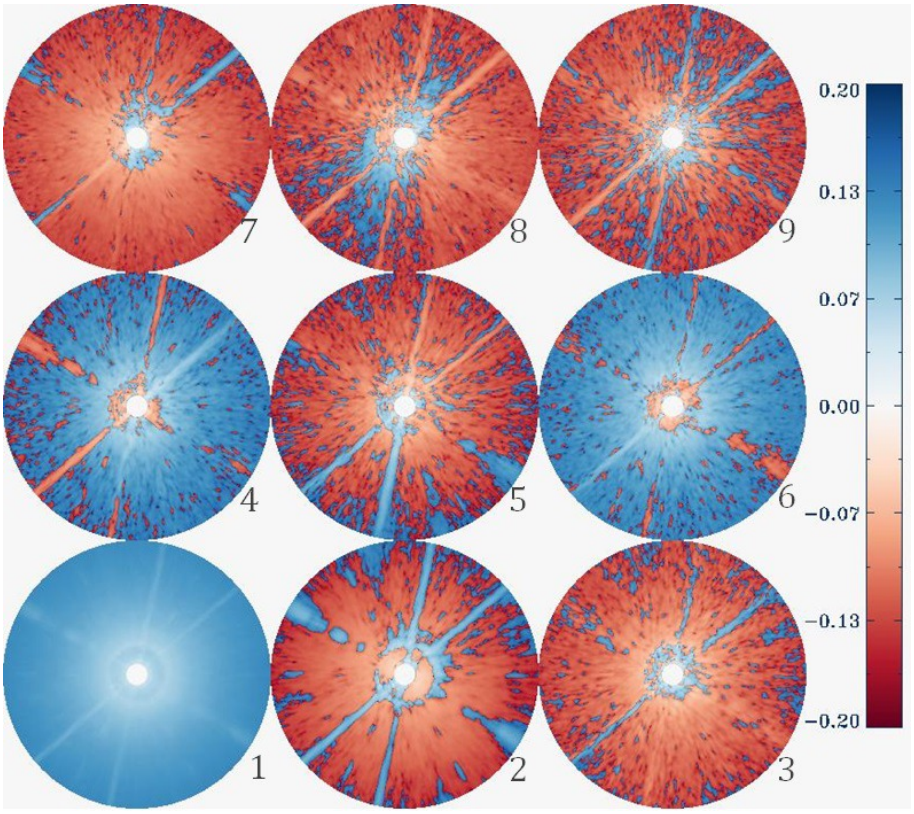


Figure 5. The projection space in the case of PCA is composed by positive and negative values. Only the first principal component, the "1" has physical meaning being the average of all the frames.

Let's consider the SVD: actually the projecting base $\mathbf{V}$ has both positive and negative numbers and, with the exception of the first eigenmode corresponding to the average of the PSF, it has not physical meaning, see Figure 5

Our idea is to build a projecting base that is more similar to the original data set and that possibly includes a-priori information such as telescope pupil, AO telemetry, seeing value, deformable mirror (DM) characteristics etc.. Actually we foresee to write the image as it was a long exposure PSF, as the linear combination of a limited number of instantaneous PSFs, as it is in the actual image formation process.

The Nonnegative Matrix Factorization (NMF) uses only positive components, mimicking the actual composition of the long exposure images. The NMF is an approximate iterative method with less demanding computer power with respect to the PCA-like approaches, as LOCI. The PCA approach returns the best subtraction of the main star: however it is paying that in terms of self-subtraction of the faint companion that is more and more severe increasing the number of degrees of freedom used.

The NMF problem consists in the minimization of the functional

$$f(\mathbf{W},\mathbf{H}) = \frac{1}{2}||\mathbf{A}-\mathbf{WH}||^{\mathbf{2}}_{\mathbf{F}}, \mathbf{W} \geq \mathbf{0}, \mathbf{H} \geq \mathbf{0}. \tag{5}$$

Where $||\cdot||_F$ is the Froebenius norm. Actually, this problem is not convex both in $\mathbf{W}$ and $\mathbf{H}$, existing multiple couples of matrices $\mathbf{W} \in \mathbb{R}^{m\times k}$ and $\mathbf{H} \in \mathbb{R}^{k\times n}$ that solve it. Several iterative algorithms are in use to solve the problem above, such as the class of Multiplicative update algorithms [13, Lee and Seung (2011)], that works as gradient descent algorithms, and the Alternate Least Square (ALS) that we also tried, obtaining better results.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 23 Nov 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

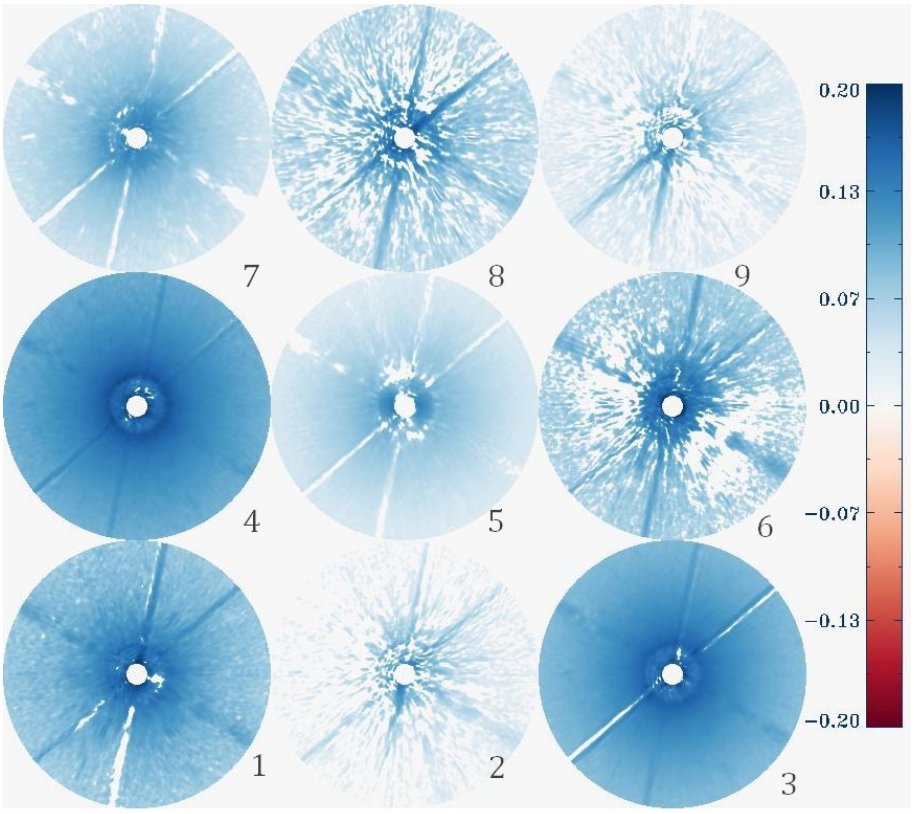


Figure 6. The projection space, **H**, in the case of NMF is composed by positive only values. All the elements have a physical interpretation, see text.

4. ALTERNATE LEAST SQUARE

The Alternate Least Square algorithm already introduced in [5, Lee and Seung (1999)] exploits the fact that nevertheless the NMF problem is not convex in both **W** and **H**, it is either in **W** or in **H**. So, fixing one of the two, it is possible to find the unique one that solves the problem. The idea is to use an alternate solution in **W** and in**H** that respects the nonnegativity condition. Starting from a random and positive **W** we solve for **H**:

$$\mathbf{W}^T\mathbf{W}\mathbf{H} = \mathbf{W}^T\mathbf{A} \tag{6}$$

writing **H** as:

$$\mathbf{H} = \left(\mathbf{W}^T\mathbf{W}\right)^{-1}\mathbf{W}^T\mathbf{A} \tag{7}$$

Then we set all the negative elements of **H** to zero and we solve for $\mathbf{W}^T$:

$$\mathbf{H}^T\mathbf{H}\mathbf{W}^T = \mathbf{H}\mathbf{A}^T \tag{8}$$

writing $\mathbf{W}^T$ as:

$$\mathbf{W}^T = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}\mathbf{A}^T \tag{9}$$

and setting also in this case all negative elements of **W** to zero. Repeating these operation up to convergence.

Applying such as algorithm to our HR8799 problem, we find for the desired k^{th} rank the pair of **WH** minimizing the problem, see Figure 6

Using this procedure we notice that a few problematic points raised:

- the result depends on the initial randomization of the matrix **W**;

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 23 Nov 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

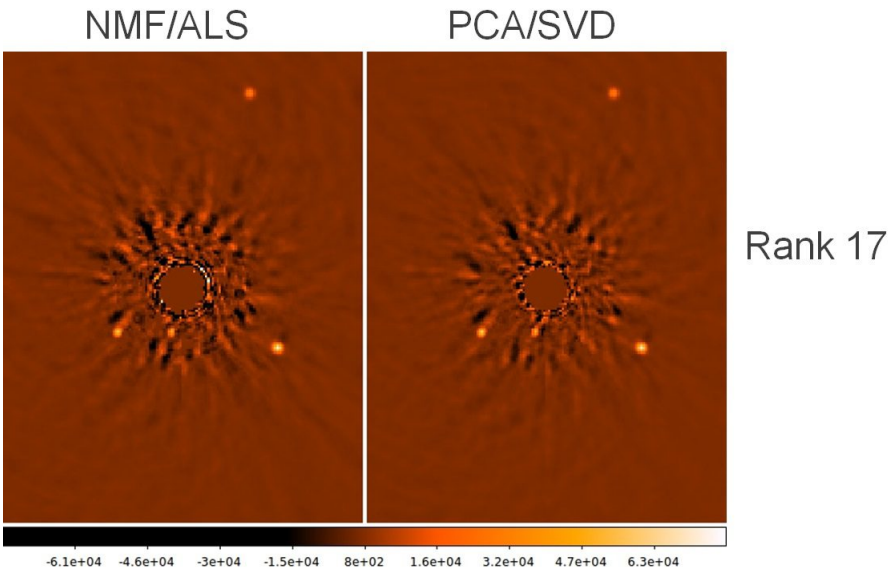


Figure 7. In the frame above we show the result of NMF (right panel) and PCA-truncated SVD (left panel). The SNR of the companion are very similar, we notice a small difference in the residual speckles pattern.

- we miss a convergence criterion to avoid local minimum;
- locking phenomenon: once one element reaches the zero it will not change anymore;

The first two points are on a strict relation: the fact that the NMF is not convex both on $\mathbf{W}$ and $\mathbf{H}$, and only on $\mathbf{W}$ or $\mathbf{H}$ implies the convergence on local minimum, that depends on the starting point.

For those two problems solution exists, take for example [14, Langville et al., 2014]: we tested positively a version of the SVD-centroid initialization taking as initial guess for the $\mathbf{H}$ the positive elements of the SVD decomposition term $\mathbf{V} \in \mathbb{R}^{n \times k}$.

We applied the ALS with the SVD-centroid initialization for rank up to the 35^{th}, while for higher rank we noticed that the solution was not converging or converging on lower ranks with higher norm.

The projection space on which the NMF algorithm converges is a true "space of the PSF" containing the variation of the PSF on the data set: actually we found (see Figure 6) in the $\mathbf{H}$ the main PSF components, such as the average PSF on frame 4, error in the centering of PSF (vibrations) on frame 3, the PSF halo and uncorrelated residuals on frame 7, spider-diffraction and residual speckles on the 2, 6 and 8, the diffraction pattern and residual within the control radius on the 5.

Looking into the results we found that residuals of the same rank for SVD and NMF produces very similar results, see it on Figure 7.

5. FURTHER WORK

Once we demonstrated the ability of the NMF to expand the images on a linear combinations of a positive space, we can include the a-priori information. A number of solutions are possible, we are converging on using auxiliary matrices orthogonal, $\mathbf{Q}$, or invertible $\mathbf{P}$ in order to write the NMF problem as:

$$\mathbf{A}_k = \mathbf{W}\mathbf{Q}\mathbf{Q}^T\mathbf{H} \tag{10}$$

in the case of an orthogonal matrix $\mathbf{Q}$, or

$$\mathbf{A}_k = \mathbf{W}\mathbf{P}\mathbf{P}^{-1}\mathbf{H} \tag{11}$$

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 23 Nov 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

in the case of the invertible $\mathbf{P}$. The iterative algorithm as the ALS, in this case, would solve for $\mathbf{WQ}$ and $\mathbf{Q}^T\mathbf{H}$ instead of the simple $\mathbf{W}$ and $\mathbf{H}$. A different solution is the inclusion of penalty terms[15] to enforce constraints through regularization parameters.

6. CONCLUSIONS

We demonstrated the ability of the NMF to solve astronomical ADI problems. The projection space on which the NMF algorithm converge is actually a "space of the PSF" containing the variation of the PSF on the data set: +we are able to decompose the original data set into a linear combination of PSF elements. Moreover we explored initialization method and propose way to include a-priori constrains.

ACKNOWLEDGMENTS

The authors collaboration was established in the framework of "Progetto Premiale" Adaptive Optics Made in Italy OU 5 Operative Unit 5 - "Analysis of adaptive Optics images". The science data were obtained using LBT and the PISCES camera during the Science Demonstration Time in the October 2011.

REFERENCES

[1] Sanders, G. H., "The Thirty Meter Telescope (TMT): An International Observatory," *Journal of Astrophysics and Astronomy* **34**, 81–86 (2013).

[2] Gilmozzi, R. and Spyromilio, J., "The European Extremely Large Telescope (E-ELT)," *The Messenger* **127**, 11 (2007).

[3] Johns, M., McCarthy, P., Raybould, K., Bouchez, A., Farahani, A., Filgueira, J., Jacoby, G., Shectman, S., and Sheehan, M., "Giant Magellan Telescope: overview," in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **8444** (2012).

[4] Paatero, P. and Tapper, U., "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics* **5**(2), 111–126 (1994).

[5] Lee, D. D. and Seung, H. S., "Learning the parts of objects by nonnegative matrix factorization," *Nature* **401**, 788–791 (1999).

[6] Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., and Nadeau, D., "Angular Differential Imaging: A Powerful High-Contrast Imaging Technique," ApJ **641**, 556–564 (2006).

[7] Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., and Artigau, É., "A New Algorithm for Point-Spread Function Subtraction in High-Contrast Imaging: A Demonstration with Angular Differential Imaging," ApJ **660**, 770–780 (2007).

[8] Karhunen, K., "Über lineare methoden in der wahrscheinlichkeitsrechnung," *Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys.* **37**, 1–79 (1947).

[9] Soummer, R., Pueyo, L., and Larkin, J., "Detection and Characterization of Exoplanets and Disks Using Projections on Karhunen-Loève Eigenimages," ApJ **755**, L28 (2012).

[10] Hill, J. M., "The Large Binocular Telescope," *Applied Optics* **49**, 115–122 (2010).

[11] McCarthy, Jr., D. W., Ge, J., Hinz, J. L., Finn, R. A., and de Jong, R. S., "PISCES: A Wide-Field, 1-2.5 μm Camera for Large-Aperture Telescopes," PASP **113**, 353–361 (2001).

[12] Esposito, S., Mesa, D., Skemer, A., Arcidiacono, C., Claudi, R. U., Desidera, S., Gratton, R., Mannucci, F., Marzari, F., Masciadri, E., Close, L., Hinz, P., Kulesa, C., McCarthy, D., Males, J., Agapito, G., Argomedo, J., Boutsia, K., Briguglio, R., Brusa, G., Busoni, L., Cresci, G., Fini, L., Fontana, A., Guerra, J. C., Hill, J. M., Miller, D., Paris, D., Pinna, E., Puglisi, A., Quiros-Pacheco, F., Riccardi, A., Stefanini, P., Testa, V., Xompero, M., and Woodward, C., "LBT observations of the HR 8799 planetary system. First detection of HR 8799e in H band," A&A **549**, A52 (2013).

[13] Lee, D. D. and Seung, H. S., "Algorithms for non-negative matrix factorization," in [*Advances in Neural Information Processing Systems 13*], Leen, T. K., Dietterich, T. G., and Tresp, V., eds., 556–562, MIT Press (2001).

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 23 Nov 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

[14] Langville, A. N., Meyer, C. D., Albright, R., Cox, J., and Duling, D., "Algorithms, initializations, and convergence for the nonnegative matrix factorization," *CoRR* **abs/1407.7299** (2014).

[15] Berry, M. W., Browne, M., Langville, A. N., Pauca, V. P., and Plemmons, R. J., "Algorithms and applications for approximate nonnegative matrix factorization," *Computational Statistics & Data Analysis* **52**(1), 155–173 (2007).

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 23 Nov 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use