Pushing ELT's sensitivity through PCA background subtraction

Hélène Rousseau, Denis Defrère, Steve Ertel, Virginie Faramaz, Kevin Wagner

To cite this version:

Hélène Rousseau, Denis Defrère, Steve Ertel, Virginie Faramaz, Kevin Wagner. Pushing ELT's sensitivity through PCA background subtraction. Adaptive Optics for Extremely Large Telescopes 7th Edition, ONERA, Jun 2023, Avignon, France. 10.13009/AO4ELT7-2023-055 . hal-04402890

HAL Id: hal-04402890
https://hal.science/hal-04402890

Submitted on 18 Jan 2024

Pushing ELT's sensitivity through PCA background subtraction

Hélène Rousseau[a,b,c], Steve Ertel[a,c], Denis Defrère[d], Virginie Faramaz[c], and Kevin Wagner[c]

[a]Large Binocular Telescope Observatory, University of Arizona, 933 N Cherry Ave., Tucson, AZ 85721-0065, USA
[b]AGO Department, University of Liège, Allée du 6 août, 19C B-4000 Liège 1,Belgium
[c]Department of Astronomy and Steward Observatory, University of Arizona, 933 N Cherry Ave., Tucson, AZ 85721-0065, USA
[d]Institute of Astronomy, KU Leuven, Celestijnenlaan 200D, 3001 Leuven, Belgium

ABSTRACT

The thermal infrared is an important wavelength range for a wide variety of astronomical observations. At these wavelengths, the large ground-based telescopes such as the ELTs reach higher angular resolution than space-based telescopes. However, their sensitivity is limited by the high thermal background due to both photon noise and imperfect removal of background structures from the sky and warm telescope optics. Developing more effective methods for the removal of spatially and temporally variable background structure is paramount for unlocking the full potential of existing and future large and extremely large ground-based telescopes operating at thermal-infrared wavelength. As recent studies have shown, the Principal-Component-Analysis (PCA) method can significantly improve the background subtraction compared to the more common method of subtracting the mean image from dedicated background exposure. We developed and optimized background subtraction routine based on PCA. We used imaging data taken in nulling-interferometric mode from the Large Binocular Telescope Interferometer (LBTI) in the N band (11 μm). The LBTI connects the two aperture of the LBT to form the largest single-mount telescope in the word. This make the LBTI the best place to pioneer data reduction methods and procedures for future ELTs. We present a comparison of classical (mean) and PCA background subtraction for both aperture photometry (mean retrieval and RMS) and high contrast imaging (contrast curves). The PCA background subtraction allows for improvement factors of two to three for both observing techniques. This will

---

Further author information:
H.R.: E-mail: hrousseau@lbti.org

allow for improving the sensitivity of suitable, existing data by the same factors. More importantly, it will reduce by a factor five to ten the extremely valuable integration time required to detect faint sources with future ELTs!

Keywords: Thermal background, PCA, background subtraction, data reduction

1. INTRODUCTION

The LBT, located in mount Graham in Arizona, possesses two 8-meters primary mirrors, on a single mount. The LBTI [10, 9] is an instrument which makes use of the 23-meters baseline of the LBT to turn it in the only currently available ELT-like facility. As such, the LBTI can be used as a pathfinder for the future ELTs [7, 8]. In particular the LBTI is optimized for observation over a wavelength range from the L to the N-band, which will also be covered by the ELTs. As the thermal background in these wavelength is really high, it limits our sensitivity. This high background will also limit the future ELTs' sensitivity. Thus, improving the thermal background removal is critical to unlock their full potential. We here present our efforts of using the LBTI data to develop a method, allowing for a better thermal background removal, in particular in the N-band.

Future ELTs, as well as future space mission, will devote a large amount of time to directly image exoplanets. In particular, to detect and characterize small, rocky planets, in the habitable zone of their hosts stars. In order to reach this objective, both ELTs and space mission will have to overcome important challenges. One of them is the exozodiacal light, which can easily outshine Earth-like exoplanets. This light originates from clouds or disks of dust located in the habitable zone of the hosts stars. Detecting those clouds or disks and constraining their amount of dust is necessary to make sure detecting rocky planets in the habitable zone is feasible and within which amount of time. Detecting and characterizing this exozodiacal dust was the goal of the HOSTS survey.

The Hunt for Observable Signature of Terrestrial Systems (HOSTS [6, 5, 2, 4]) surveyed 38 nearby star to detect exozodiacal dust round them, and to constrain the amount of dust in those systems. The observations were perform in the N-band (11 μm) with the Nulling-Optimized Mid-Infrared Camera (NOMIC, [11]) of the LBTI between September 2016 and May 2018. Those observations were carried out using nulling interferometry. This technique allows for both the small angular resolution, coming from interferometry, and an effective starlight suppression. The light from the on-axis source is suppressed by destructive interferences while off-axis sources are producing constructive interferences. Thus the starlight is effectively removed while keeping a small inner working angle, allowing to detect close-in faint objects such as planets or disks. After the successful completion of HOSTS, the LBTI is now executing targeted observations to characterize the detected systems (PI: S. Ertel, [6]).

One important result from the HOSTS survey was to put a constraint on the median amount of dust present in the systems it surveyed. This amount of dust is described by zodis, where one zodi is the equivalent of the amount of dust in our own system. The HOSTS survey has found the median zodi upper limit to be 9 zodis within a 1σ [5]. However with a 95% interval of confidence, this upper limit became 27 zodis [5]. It has been shown that the maximum zodi level allowing for a Earth-like exoplanet detection is 10 zodis [1]. If the 1σ upper limit demonstrates that the exozodiacal light should not prevent rocky exoplanets detection, a precision, higher by a factor 2 to 3 would be needed to make such conclusion with the 95% interval of confidence.

The study of the performances of the sensitivity of the HOSTS survey [3] has shown, that the imperfect thermal background removal was the major limitation. The bias in the reduced data is in particular responsible for this limitation. Reducing the bias by a factor 2 to 3 will similarly improve the sensitivity ([3, 5]). To reach this goal, we introduced a new PCA-based background subtraction based on algorithm previously developed for coronographic data background subtraction in the L and M bands [12].

In Section 2, we briefly describe the method of our background subtraction. In Section 3, we present our results with the PCA background subtraction compared to the classical Mean background subtraction, for both aperture photometry (AP) and high contrast imaging (HCI). We discuss the impact and the limitations of our method in Section 4.

2. METHOD

The constraints on the upper limit of exozodiacal dust around nearby stars, set by the HOSTS survey, were derived from data using a classical mean background subtraction. It has been shown that the main limitation of the sensitivity for those data were the bias left after the imperfect background subtraction [3]. Thus we developed a new PCA-based method to reduce this bias to a minimum. The PCA approach is a statistical method which allows one to reconstruct the data as a basis of eigenvectors corresponding to the level of variance in the data. It thus determines the dominant features of this set of data. The number of principal components used in this analysis determines the strength of the features that the analysis will consider, from the most dominant to the less significant ones. Thus, it determines the level of detail the analysis will be able to reconstruct. In the following we will compare the performances of the classical, mean background subtraction and the PCA background subtraction.

The data used in this work were obtained during the commissioning time (February 2014 - November 2015) of the HOSTS survey (September 2016 - May 2018). We use in particular two datasets, one for HCI and one for AP. For HCI, we used a dataset of β Leo, which is composed of 16800 frames with 60ms exposure time with an offset every 1000 frames. For AP, the dataset is a background-only dataset, composed of 24000 frames of 45ms exposure time, without any sky offsets. For this dataset we recreated artificial groups using the 1000 frames per group model from the dataset of β Leo.

In our data, the star can only take two positions: in the top left quadrant and in the bottom left quadrant. When preparing the images for the background subtraction, for both methods, the right part of the images is removed, the remaining top and bottom quadrants are separated and sorted. For both quadrants we alternate between groups of source exposures and of background exposures. After this first step, we correct all the images for bad pixels, and subtract their mean value. We refers to those steps as the image preparation, which is identical for both mean and background subtraction. Similarly, we use the same librairies to build the corrections for both Mean and PCA background subtraction.

In the case of the mean background subtraction approach, we compute the mean of a sequence of dedicated background exposures and then subtract it from every on-source exposure. The high-contrast analysis, to search for circumstellar emission, is then performed without further treatment of the background. The AP analysis is performed using a background annulus to estimate the background under the photometric aperture. The photometry estimated from the annulus is then subtracted from the photometry in the region of interest. The resulting photometry is then used for further analysis. This approach is extensively described in [3].

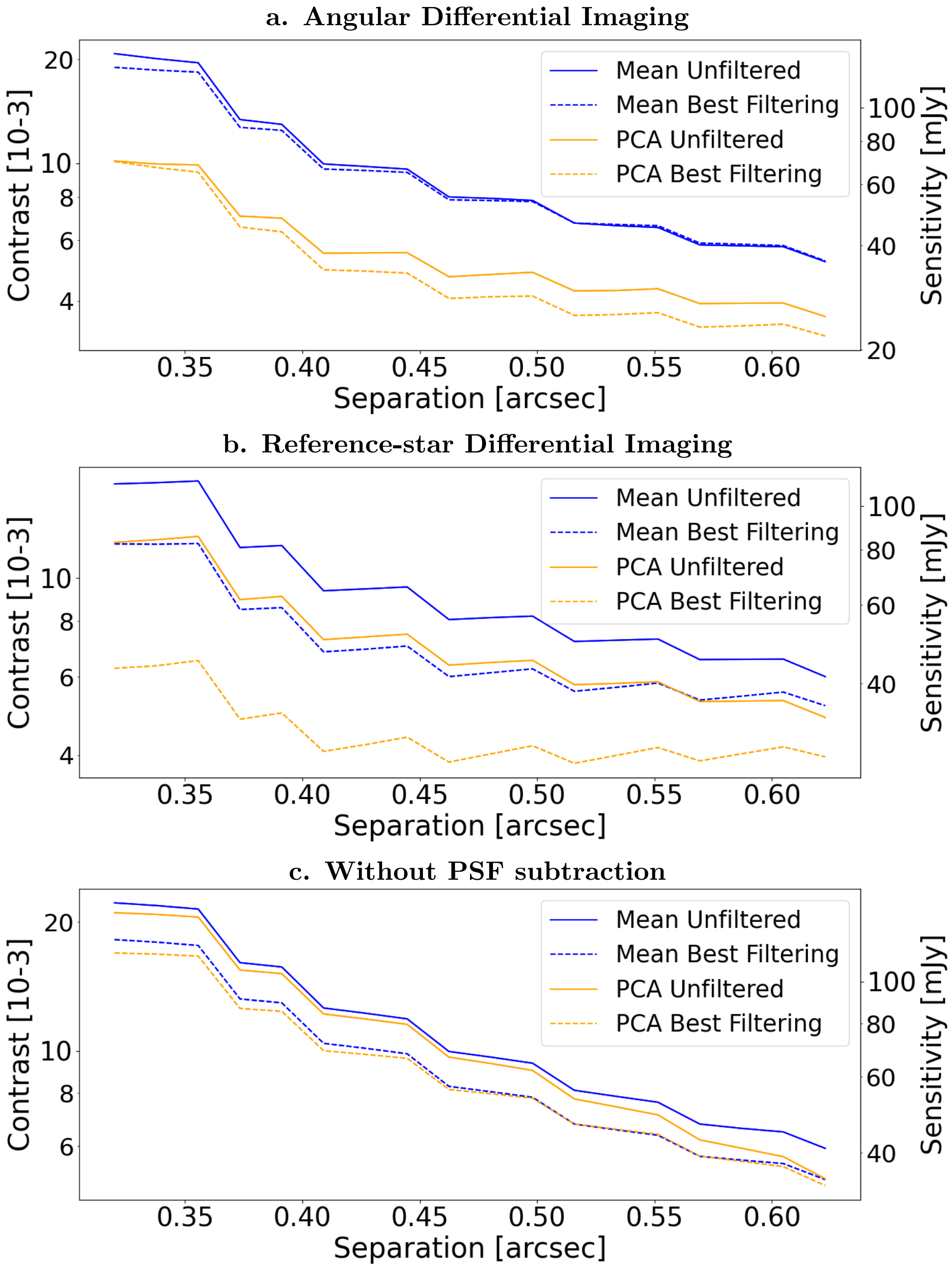
With our new method, we perform PCA on background exposures. This determines the principal components, onto which the science images are projected. In order to avoid source over-subtraction, a mask on the star needs to be introduced during this step. To limit the impact due to the mask and its intrinsic loss of information, another step is added before the background subtraction: the pre-subtraction. This steps is equivalent to the mean background subtraction. Indeed we build the correction on the background images only, without any adaptation to the image to correct. With the PCA background subtraction, an optimal correction is computed for, and subtracted from, each on-source science image. This technique allows for an additional background removal to the classical subtraction for high-contrast analysis. In the case of AP, it replaces the background annulus. PCA is indeed capable of reconstructing the background at the position of the star, which constitutes a more sophisticated estimate than the background annulus.

In the case of HCI, the filtering is performed before the PSF subtraction. In this case, we remove all images for which the photometry at the position of the star is higher than a given threshold. In the case of AP, we compute the photometry for each image in an aperture varying from 8 to 32 pixels in radius. We then remove the outlier groups of images which produce significantly higher RMS or different mean retrieval.

3. RESULTS

3.1 High-contrast imaging

In this section we present a comparison of performances between mean and PCA background subtractions, in the context of HCI. To estimate the improvement achieved by our new PCA method, we produced contrast curves for both PSF subtracted datasets, using Angular Differential Imaging (ADI) or Reference-star Differential Imaging (RDI), and non-subtracted datasets. The PSF subtraction step (if used) is identical for both Mean background subtraction and PCA background subtraction. In each case, and for each background subtraction method, we selected the best contrast-curve. This selection is made among different principal component number for the PSF subtraction (ranging from 1 to 10), different filtering levels (ranging from 100% to 20% of the total photometry) and in the case of PCA background subtraction, different principal component number for the background subtraction (ranging from 1 to 30). In Figure 1, we present our results for the three processing techniques.



a. Angular Differential Imaging

b. Reference-star Differential Imaging

c. Without PSF subtraction

Figure 1. Best contrast curves obtained for ADI (panel a), RDI (panel b), and without PSF subtraction (panel c). We present the best constrast curves for mean background subtraction in blue in the unfiltered (straight lines) and best filtering (dashed lines) cases. The best contrast curves for PCA background subtraction are presented in orange in the unfiltered (straight lines) and best filtering (dashed lines) cases.

As we can see in Figure 1, for ADI and RDI, the PCA background subtraction provides better results. The improvement factors between Mean and PCA background subtraction ranges between 1.5 to 2 for ADI and RDI. This improvement is valid for all distances from the star probed by our contrast curves. This behavior comes from the nulling-interferometric nature of our data, in which the starlight is effectively suppress. As seen by shallow decrease of the contrast toward smaller separations, the separation probed by our data fall into the contrast-limited regime compared to background-photon noise. However, the background bias dominates the final contrast as seen by the improvement brought by PCA background subtraction. In the case without PSF subtraction, since the starlight is not removed as in the case of ADI and RDI, the contrast dominate and we observe only a slight improvement by using PCA background subtraction.

3.2 Aperture Photometry

We present here a comparison of performances for Mean and PCA background subtraction with AP. We distinguish two cases: the general photometry case where the size of the aperture, the mask and the source match; and the extended source case where the size of the aperture can vary to probe different area of the source. In this last case, the mask matches the source extension. The aperture, however, can be smaller.

To achieved a realistic measurement of the improvement brought by PCA background subtraction compared to mean background subtraction, we computed, for each group of a thousand images, the RMS and the mean retrieval on both datasets. With this method, we obtained 24 measurements for both datasets. We present in Figure 2 panel a, the improvement factors obtained when using PCA background subtraction instead of Mean background subtraction in the general case photometry. We studied different size of aperture and matching mask size. In panel b of Figure 2 we present the results for the extended source case. In particular, we used for here an aperture of 8 pixels in radius and let the mask size vary from 17 to 32 pixels in radius

As we can see in both cases, the RMS (left panels) does not significantly change with the new background subtraction. However, the mean retrieval is significantly improved, in other words the bias is significantly reduced. In both cases, we obtain an improvement factor of 2 to 3 for most apertures and mask sizes. As we discussed earlier, the major limiting factor in HOSTS data being the imperfect thermal background removal and the resulting bias, such improvement would directly transfer to a similar sensitivity improvement. However, it is important to note that those results are only representative of the background limited regime.

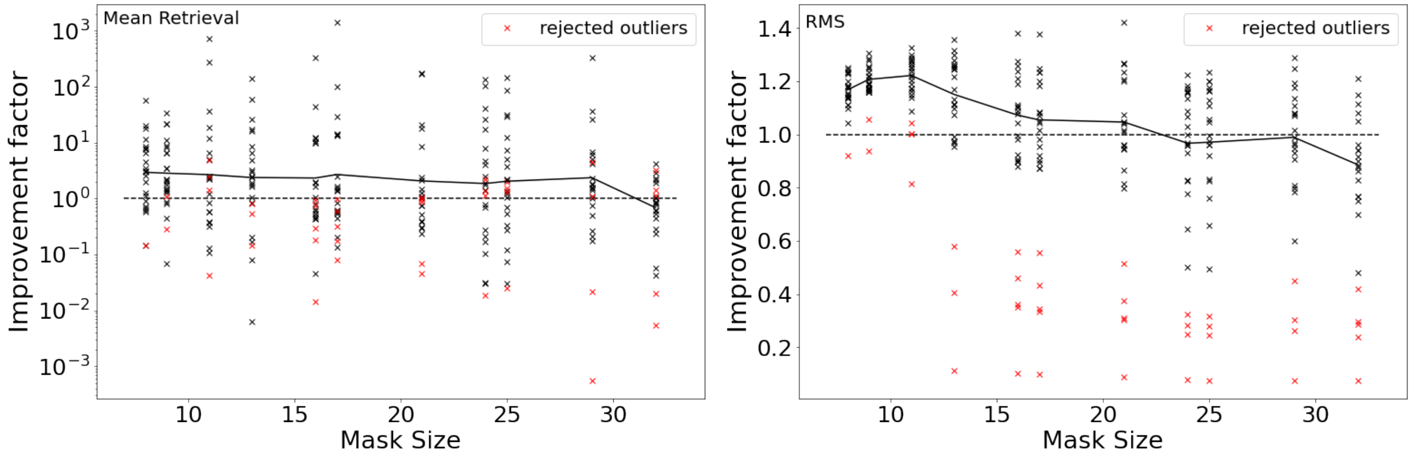
4. IMPACT AND LIMITATIONS

We aim to apply our new PCA-based background subtraction method to the whole HOSTS survey (38 stars) to improve the data sensitivity. As the imperfect thermal background subtraction is the major limiting factor of the HOSTS survey sensitivity, a factor 2 to 3 improved bias will also result in a factor 2 to 3 improved sensitivity. This improved precision will allow us to confirm detections with higher precision on the amount of dust present in the system, to rule out false positives, to improve upper limit or even add potential new detections. Altogether, this will allow us to put much stronger constraints on the median zodi level around nearby stars. Those constraints will inform future direct imaging missions aiming at the detection of rocky planets in the habitable zone of their hosts stars, both in space (Habitable World Observatory and the Large Interferometer for Exoplanets) or on the ground with the ELTs. In particular constraints on the median zodi level will inform mission designs in terms of observing time needed for an Earth-like planet to be detected and characterized.

We tested our method on nulling-interferometric data, however, this method is not limited to this type of data and can be applied to any data with a nodding sequence. Our improvement factors will remain valid when in the background-limited regime. ELTs will thus directly benefits from this method, for a wide range of observations in the infrared and in particular in the N-band. Indeed, an improved sensitivity by a factor 2 to 3 will result in a reduction of the integration time needed by a factor 4 to 9. This would greatly improve the efficiency and scientific return of the ELTs which observing time will be, in addition, strongly overbooked.

We are aware, however, that our method currently presents some limitations. In particular, with large mask sizes we observe strong variations of the photometry within one group of images. We are currently developing an additional temporal PCA-based correction to mitigate those effects. We expect from this additional correction an additional slight improvement of the bias and a more significant improvement of the RMS.

a. General photometry case



b. Extended source case

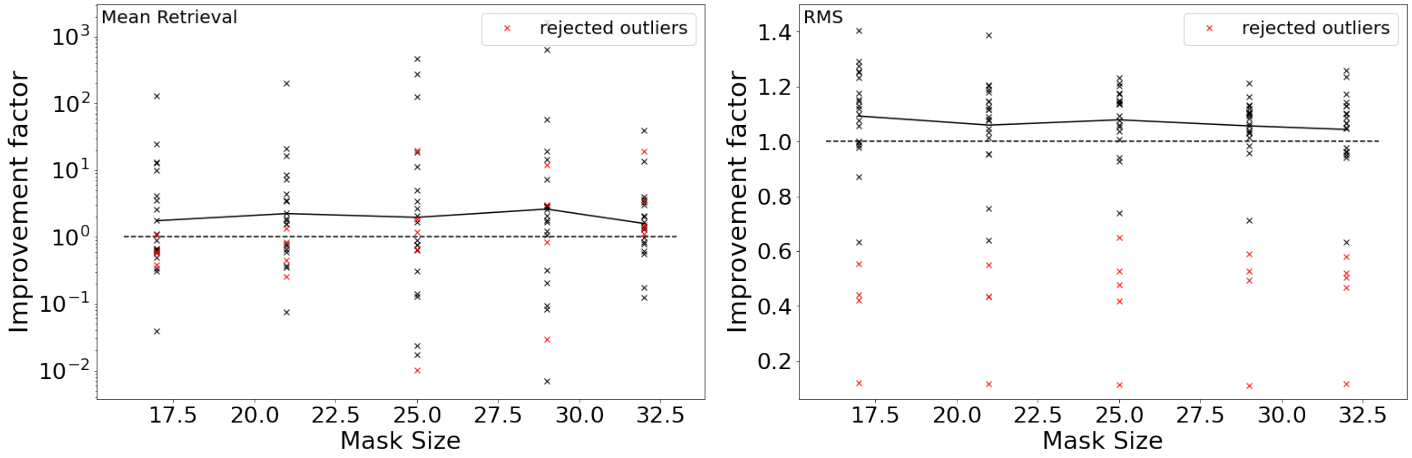


Figure 2. Improvement factors between Mean and PCA background subtraction for the mean retrieval (left panels) and the RMS (right panels). The geometric mean of those improvement factor is used to determine the overall improvement factor and is shown by the black straight line. The outliers, shown in red, are not used to compute the overall improvement factors.

5. CONCLUSIONS AND PROSPECTS

This study shows that PCA thermal background subtraction can achieve significant improvement over mean background subtraction for both AP and HCI in the mid-infrared (N band). For the latter, we have demonstrated that a PCA background subtraction can improve the reachable contrast, when in the regime limited by the background bias, by a factor 1.5 to 2.

For aperture photometry, we have shown in particular that, without degrading the photometric precision, we can reach an improvement factor from 2 to 3 on the accuracy of the mean retrieval. Imperfect thermal-background subtraction has been shown to be a major sensitivity limitation of the HOSTS survey [3]. This limitation is mainly due to the bias on the individual, calibrated null measurements rather than their errors bars. With a factor 2 to 3 improvement on the accuracy of the mean retrieval over most of the range of apertures, we expect at least a factor 2 improvement for those biases, and thus on the sensitivity. In future work we intend to apply this new method to the whole HOSTS survey target sample and re-analyze the data with the new PCA background subtraction. From this factor 2 to 3 improvement on the mean retrieval, we expect to both put stronger constraints on the already detected exozodis, but also to detect a few more among the HOSTS target sample. Furthermore, we intend to perfect this method with the introduction of a temporal correction, in addition to the spatial correction described in this work. This reanalyze of the HOSTS survey will strongly benefit future direct imaging mission.

The approach presented in this work can be applied to a wide variety of existing data sets and future observations, as it only requires background observations that are regularly interleaved with the science observations (e.g., through nodding). Similarly, future datasets with those characteristics such as data taken by JWST, large ground-based telescopes and future ELTs, would also strongly benefit from this approach. A factor 2 to 3 improvement in sensitivity in thermal-infrared observations will make such observations 4 to 9 times less time consuming, and hence greatly improve the science return of these observatories.

ACKNOWLEDGMENTS

The Large Binocular Telescope Interferometer is funded by the National Aeronautics and Space Administration as part of its Exoplanet Exploration Program. The LBT is an international collaboration among institutions in the United States, Italy, and Germany. LBT Corporation partners are The University of Arizona on behalf of the Arizona university system; Instituto Nazionale di Astrofisica, Italy; LBT Beteiligungsgesellschaft, Germany, representing the Max-Planck Society, the Astrophysical Institute Potsdam, and Heidelberg University; The Ohio State University, and The Research Corporation, on behalf of The University of Notre Dame, University of Minnesota and University of Virginia. Observations have benefited from the use of ALTA Center (alta.arcetri.inaf.it) forecasts performed with the Astro-Meso-Nh model. Initialization data of the ALTA automatic forecast system come from the General Circulation Model (HRES) of the European Centre for Medium Range Weather Forecasts. D.D. has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement CoG - 866070).

References

[1] D. Defrère et al. "Direct imaging of exoEarths embedded in clumpy debris disks". In: *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*. Ed. by Mark C. Clampin et al. Vol. 8442. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Sept. 2012, 84420M, p. 84420M. DOI: 10.1117/12.926324.

[2] D. Defrère et al. "First-light LBT Nulling Interferometric Observations: Warm Exozodiacal Dust Resolved within a Few AU of η Crv". In: 799.1, 42 (Jan. 2015), p. 42. DOI: 10.1088/0004-637X/799/1/42. arXiv: 1501.04144 [astro-ph.EP].

[3] D. Defrère et al. "NULLING DATA REDUCTION AND ON-SKY PERFORMANCE OF THE LARGE BINOCULAR TELESCOPE INTERFEROMETER". In: *ApJ* 824.2 (June 14, 2016), p. 66. ISSN: 1538-4357. DOI: 10.3847/0004-637X/824/2/66. URL: https://iopscience.iop.org/article/10.3847/0004-637X/824/2/66 (visited on 11/16/2022).

[4] D. Defrère et al. "The HOSTS Survey: Evidence for an Extended Dust Disk and Constraints on the Presence of Giant Planets in the Habitable Zone of Leo". In: *AJ* 161.4 (Apr. 1, 2021), p. 186. ISSN: 0004-6256, 1538-3881. DOI: 10.3847/1538-3881/abe3ff. URL: https://iopscience.iop.org/article/10.3847/1538-3881/abe3ff (visited on 11/16/2022).

[5] S. Ertel et al. "The HOSTS Survey for Exozodiacal Dust: Observational Results from the Complete Survey". In: *AJ* 159.4 (Mar. 31, 2020), p. 177. ISSN: 1538-3881. DOI: 10.3847/1538-3881/ab7817. URL: https://iopscience.iop.org/article/10.3847/1538-3881/ab7817 (visited on 01/28/2023).

[6] S. Ertel et al. "The HOSTS Survey—Exozodiacal Dust Measurements for 30 Stars". In: *AJ* 155.5 (Apr. 17, 2018), p. 194. ISSN: 1538-3881. DOI: 10.3847/1538-3881/aab717. URL: https://iopscience.iop.org/article/10.3847/1538-3881/aab717 (visited on 01/28/2023).

[7] S. Ertel et al. "The Large Binocular Telescope Interferometer as an ELT pathfinder". To be published in these same AO4ELT7 proceedings. 2023.

[8] Steve Ertel et al. "Imaging nearby, habitable-zone planets with the Large Binocular Telescope Interferometer". In: *Optical and Infrared Interferometry and Imaging VIII*. Ed. by Antoine Mérand, Stephanie Sallum, and Joel Sanchez-Bermudez. Vol. 12183. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Aug. 2022, 1218302, p. 1218302. DOI: 10.1117/12.2635333.

[9] Steve Ertel et al. “Overview and prospects of the LBTI beyond the completed HOSTS survey”. In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 11446. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Dec. 2020, 1144607, p. 1144607. DOI: 10.1117/12.2561849.

[10] P. M. Hinz et al. “Overview of LBTI: a multipurpose facility for high spatial resolution observations”. In: *Optical and Infrared Interferometry and Imaging V*. Optical and Infrared Interferometry and Imaging V. Vol. 9907. SPIE, Aug. 4, 2016, pp. 8–21. DOI: 10.1117/12.2233795. URL: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9907/990704/Overview-of-LBTI--a-multipurpose-facility-for-high-spatial/10.1117/12.2233795.full (visited on 11/30/2022).

[11] William F. Hoffmann et al. “Operation and performance of the mid-infrared camera, NOMIC, on the Large Binocular Telescope”. In: *Ground-based and Airborne Instrumentation for Astronomy V*. Ed. by Suzanne K. Ramsay, Ian S. McLean, and Hideki Takami. Vol. 9147. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. July 2014, 91471O, 91471O. DOI: 10.1117/12.2057252.

[12] Silvan Hunziker et al. “A PCA-based approach for subtracting thermal background emission in high-contrast imaging data”. In: *A&A* 611 (Mar. 2018), A23. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201731428. arXiv: 1706.10069[astro-ph]. URL: http://arxiv.org/abs/1706.10069 (visited on 11/16/2022).