
Physically constrained causal noise models for high-contrast imaging of exoplanets

Timothy D. Gebhard

Max Planck Institute for Intelligent Systems, Max-Planck-Ring 4, 72076 Tübingen, Germany
Institute for Particle Physics & Astrophysics, ETH Zurich, 8092 Zurich, Switzerland
timothy.gebhard@tuebingen.mpg.de

Markus J. Bonse

Institute for Particle Physics & Astrophysics, ETH Zurich, 8092 Zurich, Switzerland

Sascha P. Quanz

Institute for Particle Physics & Astrophysics, ETH Zurich, 8092 Zurich, Switzerland

Bernhard Schölkopf

Max Planck Institute for Intelligent Systems, Max-Planck-Ring 4, 72076 Tübingen, Germany
Department of Computer Science, ETH Zurich, 8092 Zurich, Switzerland

Abstract

The detection of exoplanets in high-contrast imaging (HCI) data hinges on post-processing methods to remove spurious light from the host star. So far, existing methods for this task hardly utilize any of the available domain knowledge about the problem explicitly. We propose a new approach to HCI post-processing based on a modified *half-sibling regression* scheme, and show how we use this framework to combine machine learning with existing scientific domain knowledge. On three real data sets, we demonstrate that the resulting system performs clearly better (both visually and in terms of the SNR) than one of the currently leading algorithms. If further studies can confirm these results, our method could have the potential to allow significant discoveries of exoplanets both in new and archival data.

1 Introduction

Context High-contrast imaging (HCI) of extrasolar planets is a rapidly developing field in modern astrophysics [1, 2]. Today, the detection and characterization of exoplanets through HCI is pursued at all major ground-based observatories. While the current focus is on gas giant planets at large orbital separations, HCI at next-generation telescopes will yield the first-ever image of a terrestrial exoplanet around a nearby star [3]. The most crucial step of HCI post-processing is to construct an accurate model of the stellar point spread function (PSF), which we need to subtract from the data to uncover any exoplanets close to the host star. This is very challenging not only because the host star is several orders of magnitude brighter than any companions, but also because the PSF is non-static (e.g., due to the changing atmosphere and time-variable instrument performance) and contains *speckles* [4], a particular type of systematic noise that often mimics exoplanet signals. In order to differentiate between speckles and real signals, most observations employ a technique called *angular differential imaging* (ADI) [5], where the telescope is operated in pupil-stabilized mode to record a *stack* of 10^2 – 10^5 frames over a few hours (i.e., a “video” of a star and its surrounding). Due to the Earth’s rotation, the night sky (including any potential exoplanets) then appears to rotate

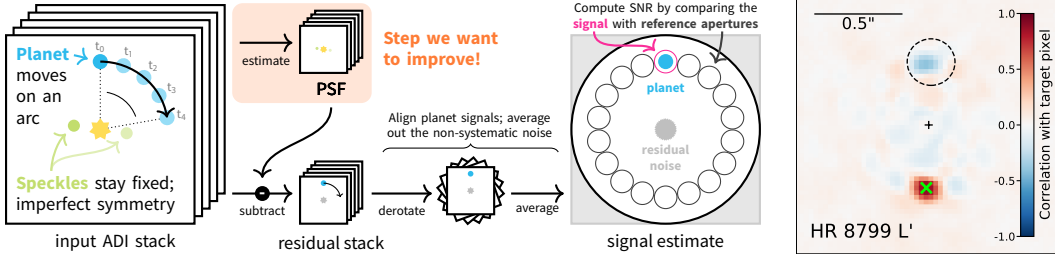


Figure 1: Finding exoplanets in HCI data requires a multi-stage post-processing pipeline. The most crucial step, however, is the estimation of the stellar PSF.

Figure 2: Example of a correlation coefficient map.

around the target star over time. The systematic noise, however, which emerges in the telescope, stays (approximately) fixed within the reference frame of the instrument. We illustrate this effect together with a standard HCI/ADI post-processing pipeline in figure 1.

Current state of the art In the past fifteen years, various algorithms have been proposed to estimate and remove the stellar PSF from ADI data, including LOCI [6], ANDROMEDA [7], LLGS [8], SODINN [9], FMMF [10], wavelets [11], PACO [12], and TRAP [13]. In practice, one of the most popular algorithms in the community is PCA-based PSF subtraction (also known as KLIP) [14, 15].¹ One notable weakness of many of these algorithms, and particularly of PCA/KLIP, is that they make only little to no use of the scientific domain knowledge that is available for the problem.

Scientific domain knowledge For instance, the only real assumption that goes into PCA/KLIP is that the systematic noise accounts for most of the variance in the data. We do, however, know much more about the problem. For example: (1) We know the expected spatial size of the planet signal and its temporal behavior (it moves on a circular arc with known opening angle around the star), which is determined by the known sky rotation, parametrized by the *parallactic angle* $\prime(t)$. (2) We have a good understanding of the causal structure of the data-generating process: our data are a (potentially clipped) sum of the signal, the systematic noise (e.g., speckles), as well as stochastic noise (e.g., read-out noise). (3) The theoretically expected structure of the stellar PSF has been studied extensively in the literature [4, 19–31]. One particular result is that, under certain circumstances, the speckle pattern is expected to be approximately (anti-)symmetric across the origin, meaning that if there is a speckle at position $(x; y)$, we should also see an (anti-)speckle at position $(-x; -y)$, where $(0;0)$ is the location of the star. In figure 2, we present empirical evidence for this, which we have obtained by the following experiment: For a given pixel P at position $(x; y)$ (indicated by the green cross) in the speckle-dominated regime close to the star, we compute the correlation coefficient (along the time axis) with all other spatial pixels. As shown in figure 2, the region around $(-x; -y)$ (indicated by the dashed circle) is clearly anti-correlated with P , which can be interpreted as evidence that speckles also exhibit some degree of symmetry in practice. (4) Additional meta-information, such as the observing conditions, are available and provide information about the temporal variation of the systematic noise.

Objectives Inspired by the availability (and current under-utilization) of this rich body of scientific domain knowledge, we develop a strategy that seeks to incorporate explicitly this information into a machine learning-based approach for post-processing high-contrast imaging data.

2 Method

Idea We propose a modified version of *half-sibling regression* (HSR) [32], taking inspiration also from the *CPM difference imaging* approach of Wang et al. [33]. Our method works as follows: To denoise the time series of a given (spatial) pixel Y , we choose a set of predictor pixels $\mathbf{X} =$

¹Note: While the limitations of pure PCA for denoising tasks are known and possible extensions (e.g., for taking into account known properties of the noise distribution) have been proposed in other fields such as neuroscience (see, e.g., [16] or [17]), these innovations have not yet been adopted in the field of high-contrast imaging. Instead, “vanilla” PCA is still commonly used to detect new exoplanets; see [18] for a recent example.

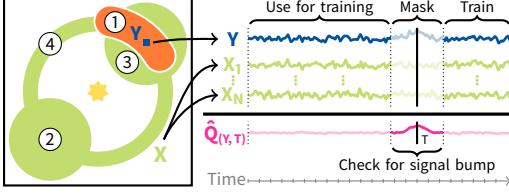


Figure 3: In step 1, we select the predictors \mathbf{X} for a target Y , train the HSR models, and store candidates.

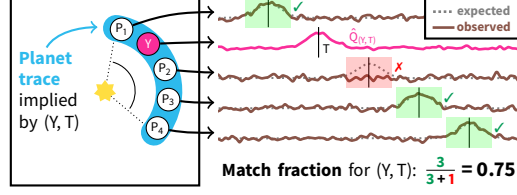


Figure 4: In step 2, we run a consistency test for every candidate $(Y; T)$ and compute its *match fraction*.

$fX_1; \dots; X_Ng$ and train (along the time axis, i.e., using different time steps as samples) a machine learning model m to predict Y from \mathbf{X} . The constraint here is that \mathbf{X} must be chosen such that, under the assumption that Y does contain planet signal, \mathbf{X} does *not* contain planet signal. Due to the above additivity assumption, our estimate for the planet signal in Y is then given by the residual time series $\hat{Q} = Y - m(\mathbf{X})$. The intuition here is that, because we choose \mathbf{X} to be *causally independent* of the planet signal, the prediction $m(\mathbf{X})$ can not contain any planet signal, but only the systematic noise. Since a planet’s position is not known a priori, we have to loop over a spatial region of interest (ROI), treat every pixel in it as Y , and denoise its time series accordingly. This step can easily be parallelized. Once we have assembled the full residual stack, we obtain the final signal estimate by derotating each frame in the residual stack by its respective parallactic angle before taking the mean along the time axis (cf. figure 1).

Choice of predictors For a given position Y , we select our predictors as illustrated in the left panel of figure 3 (remember that we need to loop over all possible positions Y in our ROI). Our choice is motivated by our knowledge about the temporal movement of a planet signal, as well as the expected structure of the speckle pattern. Region ① (in orange) is the *exclusion region*, consisting of pixels which may contain signal if Y at *some point in time* contains a planet. These pixels (which we can compute using our knowledge of the signal shape and the parallactic angles) must not be used as predictors, lest we run the risk of “explaining away” the signal that we are after in the first place. The actual predictors \mathbf{X} (in green) consist of three parts: Region ② is chosen symmetrically opposite of Y because we know from theory that if there is a speckle at Y , there should also be an (anti-)speckle at ②, meaning the pixels there should be good predictors for the systematic noise. Region ③ is chosen to capture any “local” effects around Y due to the instrument, and the annulus ④ is used because we know that the systematic noise significantly depends on the radial variable. This specific selection of predictors works well, but other choices are still part of our ongoing research.

Besides the ADI data itself, we also have access to meta-information about the observing conditions, such as wind speed or atmospheric turbulence. These quantities are guaranteed to be causally independent from the true planet signal affecting the pixel measurements, but may contain information about the systematic noise. The HSR framework allows us to include these data in the form of additional predictors, which, to the best of our knowledge, is something no other approach has explored so far.

Learning models Due to the flexibility of the HSR framework, we can use virtually *any* type of regression model to learn m . For simplicity, we choose ridge regression with generalized cross-validation (RidgeCV in `sklearn`). We can now learn such a model m using the full time series for \mathbf{X} and Y , and call this our *default HSR model*. However, if there is a strong planet signal in Y (which our predictors cannot explain as we have chosen them to be causally independent of the signal), the fit can be poor. Therefore, we employ the following *signal masking* approach consisting of two steps.

In step 1, illustrated in figure 3, we first define a grid of possible planet positions in time. Our domain knowledge allows us to compute the *expected signal form* for a given target pixel Y and a time T on the grid. For every such tuple $(Y; T)$, the shape of the expected signal implies a temporal interval (where the planet signal is non-zero) that we mask out when training the model $m_{(Y;T)}$. Once trained, we apply $m_{(Y;T)}$ to the full predictor time series and use the model’s prediction to compute the residual time series $\hat{Q}_{(Y;T)} = Y - m_{(Y;T)}(\mathbf{X}_{(Y;T)})$. Note that we write $\mathbf{X}_{(Y;T)}$ for the predictors of the model $m_{(Y;T)}$ because the exclusion region (and thus the predictors) depends both on the spatial position Y and the assumed time T at which the signal reaches its peak in Y . Next, we use a simple heuristic to check if $\hat{Q}_{(Y;T)}$ contains a “signal bump” at T that matches the

expected planet signal shape. If this is the case, we store $(Y; T)$ as a *candidate*. Finally, we prune our list of candidates and only keep the best (i.e., highest bump) candidate for each Y .

In step 2, illustrated in figure 4, we perform a *consistency test* on all candidates from step 1. Each candidate $(Y; T)$ implies a hypothesis about the planet’s exact trajectory (i.e., which pixels in the stack will be contain planet signal, and when). Therefore, we can select other spatial positions along this implied signal path and test if their residuals also show a signal bump at the expected time. For each candidate, we compute the *match fraction*, that is, the fraction of test positions along the implied planet path that show such a consistent behavior. We expect that only candidates due to an actual planet signal will yield a high match fraction.

Finally, we choose a threshold for the match fraction and assemble the residual stack in the following way: For positions Y with a match fraction above the threshold, we use the residuals obtained using the signal masking-based model. For all other positions, we use the residuals from the default HSR model. The signal estimate is then computed from the residual stack in the usual way (cf. figure 1). Because this last step is relatively cheap computationally, we run it multiple times for different threshold values and choose the result with the highest SNR.

3 Experimental evaluation and results

Data sets We showcase the performance of our improved HSR approach by applying it to three publicly available HCI data sets from the Very Large Telescope (VLT) that are known to contain exoplanets (for details, see table 1 in the appendix). To preprocess the raw data, we use a standard pipeline based on PynPoint [15, 34]. Our analysis focuses on the L^0 ($\lambda = 3.80 \mu\text{m}$) and M^0 ($\lambda = 4.78 \mu\text{m}$) wavelength bands because hundreds of archival data sets are readily available and next-generation HCI instruments for the VLT (ERIS [35]) and the ELT (METIS [36]) will be operating in this regime.

Experiments We apply three variants of our algorithm to our data sets: just the default HSR, HSR with signal masking (SM), and HSR with signal masking and using the observing conditions as additional predictors (SM+OC). For a full list of all observing conditions that we used including a short description, see table 2 in the appendix. As an additional pre-processing step, we median-combine blocks of frames to create data sets with an effective integration time of 6.5 s. This has proven beneficial in preliminary experiments, and generally also improves the results obtained with PCA/KLIP. For the signal masking, we use a grid with 50 temporal positions in step 1, and evaluate the match fraction using 20 test positions in step 2. The final signal estimates are then compared to the best result obtained using PCA/KLIP, both visually and quantitatively using the signal-to-noise ratio (SNR) as defined in Mawet et al. [37].

Results We show exemplary results for the Beta Pictoris L^0 data set in figure 5. Already visually, our proposed HSR algorithm—when used in combination with signal masking—achieves a better separation between the signal and background than the PCA baseline. Quantitatively, we find that the achieved SNR is significantly higher for HSR than for PCA (up to a factor of 4 for Beta Pictoris L^0). We also notice that adding the observing conditions as additional predictors yields a substantial performance improvement. Similar results (SNR improvements of a factor 2–3; adding OC as predictors consistently improves performance) are also found in our other two data sets; see figure 6 in the appendix for a full comparison.

Note of caution While these results are very encouraging, we would like to point out that the SNR alone is not yet a sufficient metric to fully characterize the performance of an HCI post-processing algorithm. More extensive studies—for example, in the form of performance maps [38] or contrast curves—are needed to assess more thoroughly our method’s ability to make new detections.

4 Discussion and outlook

We have outlined a new algorithm for post-processing HCI data for exoplanet science and described our strategy for integrating existing scientific domain knowledge into a flexible machine learning approach. Our preliminary results are very encouraging and indicate that HSR could constitute a significant improvement over existing community standards. If further studies (see above) can confirm these results, HSR could potentially enable new discoveries in hundreds of archival or new data sets.

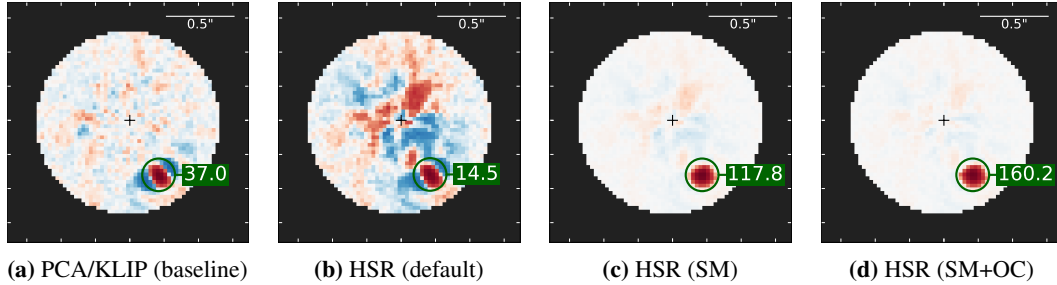


Figure 5: Exemplary results (i.e., signal estimates) for Beta Pictoris L^0 . The labels (in green) indicate the respective SNR. All images are oriented such that up = North. The color scale is a symmetric logscale from $-V_{\max}$ to V_{\max} , where V_{\max} is 1.1 times the amplitude of the signal (i.e., all plots use a different absolute scale).

We see the following steps for future research: (1) Run additional experiments to characterize the method’s performance more comprehensively (e.g., compute contrast curves), (2) study more extensively different base models for the HSR, in particular non-linear models, (3) investigate how to treat models for spatially close pixels not completely independently, (4) study the influence of the choice of predictors, (5) test if the signal estimate obtained with HSR can, in contrast to the PCA estimate, be used to directly measure the relative brightness of an exoplanet, and (6) extend our method to multi-wavelength data (i.e., from an integral field spectrograph), which would allow us to also incorporate additional domain knowledge about the wavelength-dependent behavior of speckles.

Broader impact statement

The authors are not aware of any immediate ethical or societal implications of this research. Astrophysically, the development of new advanced post-processing algorithms for HCI data may be impactful in several ways. In the short and medium-term, such algorithms can be used to find new extrasolar planets in archival data or improve existing limits. In the long run, the architecture of these methods may also influence the design of new instruments and facilities. For instance, if further studies can confirm that incorporating the observing conditions into the denoising process improves the performance, this could affect decisions about which and how much additional meta-information about HCI observations is recorded.

Acknowledgments and Disclosure of Funding

The authors thank Tomas Stolker for his help in preparing the data sets. T.D.G. acknowledges partial funding through the Max Planck ETH Center for Learning Systems. Part of this work has been carried out within the framework of the NCCR PlanetS supported by the Swiss National Science Foundation. Finally, the authors also thank the anonymous reviewers for their helpful feedback.

This research has made use of the following Python packages: `astropy` [39, 40], `bottleneck`, `context-timer`, `coverage.py`, `deepldiff`, `flake8`, `h5py`, `joblib`, `jupyter` [41], `matplotlib` [42], `mypy`, `numpy` [43], `pandas` [44, 45], `photutils` [46], `pytest`, `scikit-learn` [47], `scipy` [48], `seaborn` [49], `tqdm`.

References

- [1] S. P. Quanz et al. (2010). “First results from very large telescope NACO apodizing phase plate: 4 μm images of the exoplanet Pictoris b.” *The Astrophysical Journal*, 722(1): L49–L53. DOI: [10.1088/2041-8205/722/1/L49](https://doi.org/10.1088/2041-8205/722/1/L49).
- [2] B. P. Bowler (2016). “Imaging Extrasolar Giant Planets.” *Publications of the Astronomical Society of the Pacific*, 128(968): 102001. DOI: [10.1088/1538-3873/128/968/102001](https://doi.org/10.1088/1538-3873/128/968/102001).
- [3] S. P. Quanz et al. (2015). “Direct detection of exoplanets in the 3–10 μm range with E-ELT/METIS.” *International Journal of Astrobiology*, 14(2): 279–289. DOI: [10.1017/S1473550414000135](https://doi.org/10.1017/S1473550414000135).

- [4] E. E. Bloemhof et al. (2001). “Behavior of Remnant Speckles in an Adaptively Corrected Imaging System.” *The Astrophysical Journal*, 558(1): L71–L74. DOI: [10.1086/323494](https://doi.org/10.1086/323494).
- [5] C. Marois et al. (2006). “Angular Differential Imaging: A Powerful High-Contrast Imaging Technique.” *The Astrophysical Journal*, 641(1): 556–564. DOI: [10.1086/500401](https://doi.org/10.1086/500401).
- [6] D. Lafrenière et al. (2007). “A New Algorithm for Point-Spread Function Subtraction in High-Contrast Imaging: A Demonstration with Angular Differential Imaging.” *The Astrophysical Journal*, 660(1): 770–780. DOI: [10.1086/513180](https://doi.org/10.1086/513180).
- [7] F. Cantalloube et al. (2015). “Direct exoplanet detection and characterization using the ANDROMEDA method: Performance on VLT/NaCo data.” *Astronomy & Astrophysics*, 582: A89. DOI: [10.1051/0004-6361/201425571](https://doi.org/10.1051/0004-6361/201425571).
- [8] C. A. Gomez Gonzalez et al. (2016). “Low-rank plus sparse decomposition for exoplanet detection in direct-imaging ADI sequences.” *Astronomy & Astrophysics*, 589: A54. DOI: [10.1051/0004-6361/201527387](https://doi.org/10.1051/0004-6361/201527387).
- [9] C. A. Gomez Gonzalez et al. (2018). “Supervised detection of exoplanets in high-contrast imaging sequences.” *Astronomy & Astrophysics*, 613: A71. DOI: [10.1051/0004-6361/201731961](https://doi.org/10.1051/0004-6361/201731961).
- [10] J.-B. Ruffio et al. (2017). “Improving and Assessing Planet Sensitivity of the GPI Exoplanet Survey with a Forward Model Matched Filter.” *The Astrophysical Journal*, 842(1): 14. DOI: [10.3847/1538-4357/aa72dd](https://doi.org/10.3847/1538-4357/aa72dd).
- [11] M. J. Bonse et al. (2018). “Wavelet based speckle suppression for exoplanet imaging — Application of a de-noising technique in the time domain.” [arXiv:1804.05063](https://arxiv.org/abs/1804.05063).
- [12] O. Flasseur et al. (2018). “Exoplanet detection in angular differential imaging by statistical learning of the nonstationary patch covariances.” *Astronomy & Astrophysics*, 618: A138. DOI: [10.1051/0004-6361/201832745](https://doi.org/10.1051/0004-6361/201832745).
- [13] M. Samland et al. (2020). “TRAP: A temporal systematics model for improved direct detection of exoplanets at small angular separations.” [arXiv:2011.12311](https://arxiv.org/abs/2011.12311).
- [14] R. Soummer et al. (2012). “Detection And Characterization of Exoplanets And Disks Using Projections On Karhunen-Loève Eigenimages.” *The Astrophysical Journal*, 755(2): L28. DOI: [10.1088/2041-8205/755/2/L28](https://doi.org/10.1088/2041-8205/755/2/L28).
- [15] A. Amara et al. (2012). “PYNPOINT: an image processing package for finding exoplanets.” *Monthly Notices of the Royal Astronomical Society*, 427(2): 948–955. DOI: [10.1111/j.1365-2966.2012.21918.x](https://doi.org/10.1111/j.1365-2966.2012.21918.x).
- [16] J. V. Manjón et al. (2015). “MRI noise estimation and denoising using non-local PCA.” *Medical Image Analysis*, 22(1): 35–47. DOI: [10.1016/j.media.2015.01.004](https://doi.org/10.1016/j.media.2015.01.004).
- [17] P.-L. Bazin et al. (2019). “Denoising High-Field Multi-Dimensional MRI With Local Complex PCA.” *Frontiers in Neuroscience*, 13. DOI: [10.3389/fnins.2019.01066](https://doi.org/10.3389/fnins.2019.01066).
- [18] A. J. Bohn et al. (2020). “Two Directly Imaged, Wide-orbit Giant Planets around the Young, Solar Analog TYC 8998-760-1.” *The Astrophysical Journal*, 898(1): L16. DOI: [10.3847/2041-8213/aba27e](https://doi.org/10.3847/2041-8213/aba27e).
- [19] E. E. Bloemhof (2002). “Statistics of remnant speckles in an adaptively corrected imaging system.” In: *Adaptive Optics Systems and Technology II*. Edited by R. K. Tyson et al. SPIE, 2002. DOI: [10.1117/12.454811](https://doi.org/10.1117/12.454811).
- [20] E. E. Bloemhof (2002). “Suppression of Speckle Noise by Speckle Pinning in Adaptive Optics.” *The Astrophysical Journal*, 582(1): L59–L62. DOI: [10.1086/346100](https://doi.org/10.1086/346100).
- [21] E. E. Bloemhof (2003). “Speckles in a highly corrected adaptive optics system.” In: *Astronomical Adaptive Optics Systems and Applications*. Edited by R. K. Tyson et al. SPIE, 2003. DOI: [10.1117/12.507241](https://doi.org/10.1117/12.507241).
- [22] E. E. Bloemhof (2004). “Anomalous intensity of pinned speckles at high adaptive correction.” *Optics Letters*, 29(2): 159. DOI: [10.1364/ol.29.000159](https://doi.org/10.1364/ol.29.000159).
- [23] E. E. Bloemhof (2004). “Remnant Speckles in a Highly Corrected Coronagraph.” *The Astrophysical Journal*, 610(1): L69–L72. DOI: [10.1086/423172](https://doi.org/10.1086/423172).

- [24] E. E. Bloemhof (2004). “Speckle noise in highly corrected coronagraphs.” In: *Advanced Wavefront Control: Methods, Devices, and Applications II*. Edited by J. D. Gonglewski et al. SPIE, 2004. DOI: [10.1117/12.560370](https://doi.org/10.1117/12.560370).
- [25] E. E. Bloemhof (2004). “Static point-spread function correction dominating higher-order speckle terms at high adaptive correction.” *Optics Letters*, 29(20): 2333. DOI: [10.1364/ol.29.002333](https://doi.org/10.1364/ol.29.002333).
- [26] E. E. Bloemhof (2006). “Suppression of speckles at high adaptive correction using speckle symmetry.” In: *Instruments, Methods, and Missions for Astrobiology IX*. Edited by R. B. Hoover et al. SPIE, 2006. DOI: [10.1117/12.681477](https://doi.org/10.1117/12.681477).
- [27] E. E. Bloemhof (2007). “Feasibility of symmetry-based speckle noise reduction for faint companion detection.” *Optics Express*, 15(8): 4705. DOI: [10.1364/oe.15.004705](https://doi.org/10.1364/oe.15.004705).
- [28] A. Boccaletti et al. (2002). “Speckle Symmetry with High-Contrast Coronagraphs.” *Publications of the Astronomical Society of the Pacific*, 114(792): 132–136. DOI: [10.1086/338914](https://doi.org/10.1086/338914).
- [29] A. Sivaramakrishnan et al. (2002). “Speckle Decorrelation and Dynamic Range in Speckle Noise-limited Imaging.” *The Astrophysical Journal*, 581(1): L59–L62. DOI: [10.1086/345826](https://doi.org/10.1086/345826).
- [30] M. D. Perrin et al. (2003). “The Structure of High Strehl Ratio Point-Spread Functions.” *The Astrophysical Journal*, 596(1): 702–712. DOI: [10.1086/377689](https://doi.org/10.1086/377689).
- [31] E. N. Ribak et al. (2008). “Fainter and closer: finding planets by symmetry breaking.” *Optics Express*, 16(20): 15553. DOI: [10.1364/oe.16.015553](https://doi.org/10.1364/oe.16.015553).
- [32] B. Schölkopf et al. (2016). “Modeling confounding by half-sibling regression.” *Proceedings of the National Academy of Sciences*, 113(27): 7391–7398. DOI: [10.1073/pnas.1511656113](https://doi.org/10.1073/pnas.1511656113).
- [33] D. Wang et al. (2017). “A pixel-level model for event discovery in time-domain imaging.” [arXiv:1710.02428](https://arxiv.org/abs/1710.02428).
- [34] T. Stolker et al. (2019). “PynPoint: a modular pipeline architecture for processing and analysis of high-contrast imaging data.” *Astronomy & Astrophysics*, 621: A59. DOI: [10.1051/0004-6361/201834136](https://doi.org/10.1051/0004-6361/201834136).
- [35] R. Davies et al. (2018). “ERIS: revitalising an adaptive optics instrument for the VLT.” In: *Ground-based and Airborne Instrumentation for Astronomy VII*. Edited by H. Takami et al. SPIE, 2018. DOI: [10.1117/12.2311480](https://doi.org/10.1117/12.2311480).
- [36] B. R. Brandl et al. (2016). “Status of the mid-infrared E-ELT imager and spectrograph METIS.” In: *Ground-based and Airborne Instrumentation for Astronomy VI*. Edited by C. J. Evans et al. SPIE, 2016. DOI: [10.1117/12.2233974](https://doi.org/10.1117/12.2233974).
- [37] D. Mawet et al. (2014). “Fundamental Limitations Of High Contrast Imaging Set By Small Sample Statistics.” *The Astrophysical Journal*, 792(2): 97. DOI: [10.1088/0004-637x/792/2/97](https://doi.org/10.1088/0004-637x/792/2/97).
- [38] R. Jensen-Clem et al. (2017). “A New Standard for Assessing the Performance of High Contrast Imaging Systems.” *The Astronomical Journal*, 155(1): 19. DOI: [10.3847/1538-3881/aa97e4](https://doi.org/10.3847/1538-3881/aa97e4).
- [39] Astropy Collaboration et al. (2013). “Astropy: A community Python package for astronomy.” *Astronomy & Astrophysics*, 558: A33. DOI: [10.1051/0004-6361/201322068](https://doi.org/10.1051/0004-6361/201322068).
- [40] Astropy Collaboration et al. (2018). “The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package.” *The Astronomical Journal*, 156(3): 123. DOI: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f).
- [41] T. Kluyver et al. (2016). “Jupyter Notebooks – a publishing format for reproducible computational workflows.” In: *Positioning and Power in Academic Publishing: Players, Agents and Agendas*. Edited by F. Loizides et al. IOS Press, 2016, pages 87–90.
- [42] J. D. Hunter (2007). “Matplotlib: A 2D graphics environment.” *Computing in Science & Engineering*, 9(3): 90–95. DOI: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55).
- [43] C. R. Harris et al. (2020). “Array programming with NumPy.” *Nature*, 585(7825): 357–362. DOI: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2).
- [44] W. McKinney (2010). “Data Structures for Statistical Computing in Python.” In: *Proceedings of the 9th Python in Science Conference*. SciPy, 2010. DOI: [10.25080/majora-92bf1922-00a](https://doi.org/10.25080/majora-92bf1922-00a).

- [45] J. Reback et al. (2020). *pandas-dev/pandas: Pandas 1.1.2*. 2020. DOI: [10.5281/zenodo.3509134](https://doi.org/10.5281/zenodo.3509134).
- [46] L. Bradley et al. (2020). *astropy/photutils: 1.0.1*. 2020. DOI: [10.5281/zenodo.596036](https://doi.org/10.5281/zenodo.596036).
- [47] F. Pedregosa et al. (2011). “Scikit-learn: Machine Learning in Python.” *Journal of Machine Learning Research*, 12: 2825–2830.
- [48] P. Virtanen et al. (2020). “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python.” *Nature Methods*, 17: 261–272. DOI: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2).
- [49] M. Waskom et al. (2020). *mwaskom/seaborn: v0.11.0 (September 2020)*. 2020. DOI: [10.5281/zenodo.592845](https://doi.org/10.5281/zenodo.592845).
- [50] O. Absil et al. (2013). “Searching for companions down to 2 AU from κ Pictoris using the L^0 -band AGPM coronagraph on VLT/NACO.” *Astronomy & Astrophysics*, 559: L12. DOI: [10.1051/0004-6361/201322748](https://doi.org/10.1051/0004-6361/201322748).
- [51] M. Bonnefoy et al. (2013). “The near-infrared spectral energy distribution of κ Pictoris b.” *Astronomy & Astrophysics*, 555: A107. DOI: [10.1051/0004-6361/201220838](https://doi.org/10.1051/0004-6361/201220838).

A Appendix

Table 1: Details of the three HCI data sets that we used for our experiments. All data were obtained with the NACO instrument at the VLT observatory and are publicly available from the ESO archive.

Target star	Filter	Date	Stack size ⁰	Coronagraph	DIT (s) ¹	Δ' (") ²	ESO Program ID	Original reference
Beta Pictoris	L^0	2013-02-01	(29 681, 65, 65)	AGPM	0.200	83.3	60.A-9800(J)	Absil et al. [50]
Beta Pictoris	M^0	2012-11-26	(52 122, 73, 73)	—	0.065	51.8	090.C-0653(D)	Bonnefoy et al. [51]
HR 8799	L^0	2011-09-01	(21 043, 165, 165)	AGPM	0.200	32.5	087.C-0450(B)	Previously unpublished.

⁰ Format: (number of frames, frame width in pixel, frame height in pixel). ¹ Detector integration time per frame ² Field rotation of data set

Table 2: Overview of the observing conditions used as additional predictors for the HSR model in section 3. The values were obtained by the Astronomical Site Monitor (ASM) at Paranal and are directly available from the raw FITS files of the observations. For values were only a start and end value for each file was available we used a linear interpolation to obtain one value per frame.

Parameter name	Description
AIR_MASS	Air mass relative to zenith (unitless).
AIR_PRESSURE	Observatory ambient air pressure (in hPa).
AVERAGE_COHERENCE_TIME	Average coherence time τ_0 (in s).
M1_TEMPERATURE	Superficial temperature of the primary mirror M1 (in °C).
OBSERVATORY_TEMPERATURE	Observatory ambient temperature (in °C).
RELATIVE_HUMIDITY	Observatory ambient relative humidity (in %).
SEEING	Observatory Seeing (in arcsec).
WIND_SPEED	Observatory ambient wind speed (in m/s).
COS_WIND_DIRECTION	Cosine of the observatory ambient wind direction (unitless).
SIN_WIND_DIRECTION	Sine of the observatory ambient wind direction (unitless).

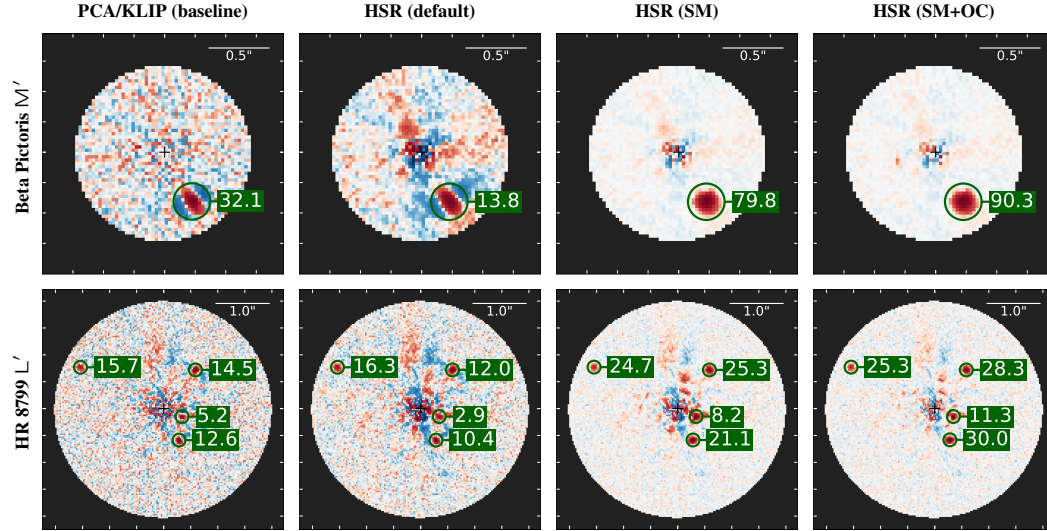


Figure 6: Additional plots for the main experiment from section 3 showing the results for the data Beta Pictoris M^0 and HR 8799 L^0 . Again, all images are oriented such that up = North, labels indicate the respective SNR, and each plot uses a symmetric logscale. We find our results from figure 5 confirmed: HSR in combination with signal masking is significantly better than PCA, both visually and in terms of SNR. Furthermore, adding the observing conditions as additional predictors improves the SNR even further.