DETECTION AND CHARACTERIZATION OF EXOPLANETS USING PROJECTIONS ON KARHUNEN–LOEVE EIGENIMAGES: FORWARD MODELING

LAURENT PUEYO
Space Telescope Science Institute, 3700 San Martin Drive, Baltimore MD 21218, USA; pueyo@stsci.edu

Received 2015 August 15; accepted 2016 March 17; published 2016 June 21

ABSTRACT

A new class of high-contrast image analysis algorithms that empirically fit and subtract systematic noise has lead to recent discoveries of faint exoplanet/substellar companions and scattered light images of circumstellar disks. These methods are extremely efficient at enhancing the detectability of a faint astrophysical signal, but they generally create systematic biases in their observed properties. This paper provides a general solution for this outstanding problem. We present an analytical derivation of a linear expansion that captures the impact of astrophysical noise in current image analysis techniques. We examine the general case for which the reference images of the astrophysical scene move azimuthally and/or radially across the field of view as a result of the observation strategy. Our new method is based on perturbing the covariance matrix of the exposure image and propagating this perturbation through the data analysis algorithm. Most of the work in this paper is presented in the Principal Component Analysis framework, but it can be easily generalized to methods relying on the linear combination of images (instead of eigenmodes). Based on this linear expansion, which is obtained in the general case, we then demonstrate practical applications of this new algorithm. We first consider the spectral extraction of faint point sources in IFS data and illustrate, using public Gemini Planet Imager commissioning data, that our novel perturbation-based Forward Modeling, which we named Karhunen Loeve Image Processing (KLIP-FM), can indeed alleviate algorithmic biases. We then apply KLIP-FM to the detection of point sources and show how it decreases the rate of false negatives while keeping the rate of false positives unchanged when compared to classical KLIP. This can potentially have important consequences on the design of follow-up strategies of ongoing direct imaging surveys.

Key words: planetary systems – techniques: image processing

1. INTRODUCTION

Progress in the domain of high-contrast image analysis has spearheaded recent discoveries of faint exoplanet/substellar companions, and resulted in spectacular scattered light images of circumstellar disks around nearby stars. This progress has been mostly driven by a new class of direct imaging data analysis algorithms (Lafrenière et al. 2007; Amara & Quanz 2012; Soummer et al. 2012) that empirically fit and subtract systematic noise in coronagraph data (also called speckle noise). Speckles stems from light diffracted by the optics in the telescope and the instrument. They are a major nuisance when seeking to detect faint circumstellar point or extended sources, due to their characteristic temporal and spatial scales (respectively of the order of the exposure time and the size of the image of a point source). Modern coronagraph data analysis methods calibrate this source of noise by using local estimates of the speckles’ correlation between the science exposures and a library of noise realizations. This speckle fitting is carried out in the least-squares sense. The collection of reference images is sometimes obtained using observations of calibration stars that act as true references (Reference Differential Imaging; hereafter RDI). However, in most ground-based cases, the library of noise realizations is assembled using exposures of the source of interest in configurations for which the observer knows a priori that the location of the faint astrophysical source moves in the frame attached to the speckles. In these cases, each image can be treated as a science frame and also included in the reference stack corresponding to other exposures and/or wavelengths in the sequence. Observation strategies enabling this feature include azimuthal motion of the astrophysical signal with respect to the speckles (Angular Differential Imaging ADI; Marois et al. 2006) and radial motion (with Integral Field Spectrograph observations, Spectroscopic Spectral Differential Imaging SSDI; Sparks & Ford 2002), or they are based on the intrinsic properties of the hypothetical sources surveyed for Polarization Differential Imaging, PDI, or presence of sharp spectral feature for Spectral Differential Imaging, SDI (Biller et al. 2004).

Once a library of noise realizations, or Point Spread Functions (PSF), has been assembled according to one or more of these strategies, least-squares fitting algorithms can be finely tuned. This is achieved in a variety of ways including optimizing how the field of view is partitioned before speckle fitting (e.g., adapting the analysis to how locally one thinks the speckles are correlated), varying the selection criteria that select the “best” noise realizations from the ensemble of references, and regularizing the inverse problem. While implementations and choice of algorithmic parameters vary among authors, the consensus emerging in the community is that these methods are extremely efficient at enhancing the detectability of faint astrophysical signals, but do generally create systematic biases in their observed properties (namely photometry, spectra, astrometry of point sources and morphology, and surface brightness of circumstellar disks). Now that large surveys based on Extreme Adaptive Optics Coronagraph instruments are hitting their full stride (Beuzit et al. 2008; Hinkley et al. 2011; Macintosh et al. 2014), these biases are becoming one of the chief problems in high-contrast image analysis.

During the past few years several authors have proposed algorithmic modifications in order to mitigate such biases (Marois et al. 2010a; Pueyo et al. 2012, 2015; Fergus et al. 2014; Marois et al. 2014a). Forward Modeling in the
context of exoplanet imaging was first proposed by Marois et al. (2010a). It aims at jointly estimating the instrument response and the astrophysical signal. To do so, negative synthetic sources are injected in the raw data across the entire observing sequence. This new data set, with both positive astrophysical and negative synthetic signals, is then propagated through the reduction algorithm. Jointly minimizing the residuals in such processed images (by exploring the range of possible astrophysical properties for the synthetic negative sources) retrieves in principle the unbiased observables of the astrophysical signal. Soummer et al. (2012) suggested that carrying out a least-squares speckle subtraction using Karhunen Loeve Image Processing (KLIP, or Principal Component Analysis, PCA) provides a simple and computationally efficient framework to implement unbiased astrophysical inference that is equivalent to injecting a negative synthetic source in the raw data. However, that paper did not fully describe how to implement this Forward Modeling with KLIP (hereafter, KLIP-FM) in the most general case. Pueyo et al. (2015) revisited this problem and described how to apply KLIP-FM in the context of RDI, when the library of reference images is built using calibrator stars (with no astrophysical signal in the library). That paper then discussed how to modify the ADI/SDI problem so it mimics the RDI configuration, and thus in principle reduces biases on astrophysical estimates. That technique was used in Hinkley et al. (2013), Oppenheimer et al. (2013), and Crepp et al. (2015). In parallel, Brandt et al. (2013) and Esposito et al. (2014), for point sources and disks, respectively, discussed how the presence of an astrophysical signal in PSF libraries obtained using ADI can be accounted for as a small perturbation of the least-squares coefficients. They then showed how these small perturbations could be included in a Forward Modeling framework to self-calibrate biases on astrophysical observables a posteriori.

In the present manuscript we generalize this class of perturbation analysis to all type of observations. Our main objective is to describe the principles underlying KLIP-FM in the most general case (e.g., without the strong hypothesis previously discussed in the literature). The novelty of our method relies on an analytical expansion for the Principal Components, when astrophysical signal is present in the reference images. Because of their high technicallity, we leave both the proof of this analytical expansion and the algorithmic details regarding its implementation out of main body of the paper. Instead, Section 2 provides a high-level description of our main result and places it into the context of previously published work. We then demonstrate the advantages of our approach by applying it to two key exoplanet imaging applications: spectral characterization with an Integral Field Spectrograph (Section 3) and point source detection (Section 4). Note that we limit the scope of this paper to these two practical examples. In Section 5 we conclude by listing other science cases for which our method could be potentially beneficial. The technical background underlying our results is discussed in depth in the Appendices:

1. Appendix A provides the most general formalism for an ADI+SSDI observing sequence and lays out the formal foundations for our work.
2. Appendix B summarizes the notations Appendix A in a table format. In order to facilitate numerical implementation, it provides the dimensions of the various matrices discussed in this paper.
3. Appendix C introduces Forward Modeling in the most general case, and then discusses the specific configuration of RDL. This was already presented in Pueyo et al. (2015), but serves here to set up the stage for Appendix F.
4. Appendix D describes Forward Modeling for the astrometry and photometry of point sources using the linear algebra notations introduced in Appendices A and C. It also set up the stage for the spectral estimation algorithm described in Appendix F.
5. Appendix E contains the proof of our main result. It heavily relies on the notations introduced in Appendix A and summarized in Appendix B.
6. Appendix F describes how to take advantage of the result in Appendix E to carry out Forward Modeling for the estimation of the point source’s spectra using IFS data.

2. GENERALIZED FORWARD MODELING

2.1. Over-subtraction and Self-subtraction

We start with the notations discussed in Soummer et al. (2012), and assume the case of a target image \( T(x) \) (where \( x \) is the spatial dimension) along with a set of reference images \( R_k(x) \). The details of how \( T(x) \) and \( R_k(x) \) are chosen among some generic coronagraph sequence are not discussed here. We refer the reader to Appendix A for a thorough presentation of the parameters associated with building a target/reference library in the most general case. An orthonormal basis \( Z_k(x) \) is then obtained based on the eigenvectors of the references covariance matrix. The associated eigenvalues \( \lambda_k \) are ranked in decreasing order. They quantify how prevalent each mode \( Z_k(x) \) is in the reference stack. When \( \lambda_k \gg 1 \) the mode is present in most of the reference images, and conversely when \( \lambda_k \ll 1 \) it is absent from most references. Again, linear algebra details are given in Appendix A. When astrophysical signal \( A(x) \) is present in the target—e.g., \( T(x) = I_v(x) + A(x) \), with \( I_v(x) \) standing for the speckle noise realization in the target image to remain consistent with Soummer et al. (2012)—the resulting processed image \( P(x) \) is given by the sum of two terms \( P(x) = P_{\text{spe}}(x) + P_{\text{sig}}(x) \):

1. The residual speckles that have not been fully captured by the PCA:

\[
P_{\text{spe}}(x) = I_v(x) - \sum_{k=1}^{K_{\text{clip}}} < I_v(x), Z_k(x) >_S Z_k(x),
\]

where \( K_{\text{clip}} \) corresponds to the number of Principal Components over which the target image is projected and \( < \cdot, \cdot >_S \) stands for the \( L_2 \) inner product on the portion of the field of view over which the speckle fitting is carried out (also called the \( S \) zone).

2. The astrophysical signal, corrupted by the KLIP algorithm:

\[
P_{\text{sig}}(x) = A(x) - \sum_{k=1}^{K_{\text{clip}}} < A(x), Z_k(x) >_S Z_k(x).
\]

This latter term is the source of the biases that we seek to calibrate with Forward Modeling.

When the \( Z_k(x) \) do not depend on the astrophysical signal, then the corruption of \( A(x) \) is a linear process. It can be interpreted as confusion: namely the algorithm fits the
astrophysical signal with speckle noise. This occurs for instance in RDI. In this configuration, the \( R_{t}(x) \) were built using images of other stars and thus do not contain the astrophysical signal of interest. We call this phenomenon over-subtraction. On the other hand, when the references, and thus the \( Z_{k}(x) \), do depend on the astrophysical signal, the corruption of \( A(x) \) is a nonlinear process. Because the astrophysical signal in the reference images \( R_{k}(x) \) is added to the speckle noise, its impact on the covariance matrix, which scales as the square of the references, is quadratic. As a consequence, the Principal Components also depend quadratically on the astrophysical signal. We call this phenomenon self-subtraction. In the context of the Locally Optimized Combination of Images algorithm—LOCI, Lafrenière et al. (2007)—this effect can be interpreted as the subtraction of the astrophysical object with itself as it rotates across the field of view during an ADI sequence. In this case, we write \( Z_{k}(x) = Z_{k}^{A}(x) \) to denote the dependence of the Principal Components on the astrophysical signal.

2.2. Forward Modeling Complications due to Self-subtraction

When a true astrophysical source is present in the data, a detection algorithm is first used to discriminate the \( P_{\text{sig}}(x) \) and \( P_{\text{spc}}(x) \) components. If \( P_{\text{sig}}(x) \) is corrupted by the speckle fitting algorithm, then Forward Modeling is used in an attempt to estimate the faint source’s underlying astrophysical properties. This is often done by injecting a synthetic astrophysical source in the data \( A(x) \) and carrying out a joint minimization over both properties of this source and the speckle noise, as discussed in Section 1. This minimization can be formally written as:

\[
\begin{align*}
\min_{A} & \quad \left| \left| P(x) - A(x) + \sum_{k=1}^{K_{\text{Klip}}} < A(x), Z_{k}^{A}(x) >_{\Sigma} Z_{k}^{A}(x) \right| \right|^{2} \\
\end{align*}
\]

where \( \min_{A} \) stands for the minimization over the observable properties of the negative synthetic signal, and \( Z_{k}^{A}(x) \) for the Principal Components resulting from injecting this negative source in the observing sequence. Note that even if the discussions in this section rely on the example in Equation (3), which uses the formalism of Soummer et al. (2012), they are applicable to any least-squares speckle fitting algorithm. Appendices A to F provide this more general framework. Direct inspection of Equation (3) shows that Forward Modeling is a nonlinear optimization in which the speckle subtraction (the determination of the Principal Components in our example), is nested within an outer nonlinear loop. As a consequence, KLIP has to be carried out every time the cost function in Equation (3) is evaluated. One hopes that this two-step process breaks degeneracies, and extensive tests using low-dimensional configurations by a variety of authors have shown this to be true in most cases (see Marois et al. 2010a or Morzinski et al. 2015). However, this approach presents two main limitations. First, it becomes quickly untractable numerically when the number of astrophysical observables is large (~30 in the case of IFS data). Second, and more importantly, there is no guarantee that the optimization in Equation (3) will converge to its global minimum, for which \( \hat{A}(x) = A(x) \). Indeed, because \( Z_{k}^{A}(x) \) is a nonlinear function of the negative synthetic signal, there is no guarantee that Equation (3) is strictly convex with respect to the astrophysical properties captured in \( \hat{A}(x) \). In other words, there is no guarantee that the Forward Modeling cost function contours are always similar to the convex parabolas shown in Morzinski et al. (2015). When such pathological cases occur, the minimization can easily stall in local minima, yielding biased observables. This is an important and fundamental drawback stemming from self-subtraction. Up until now it could only be addressed using sophisticated nonlinear optimizers.

2.3. Forward Modeling for Over-subtraction

In the case of over-subtraction, the Principal Components do not depend on the astrophysical signal. As a consequence, the propagation of \( \hat{A}(x) \) through the algorithm is linear. Using the example of the KLIP algorithm, this can also be written as \( \mathcal{KLIP}[T(x) - A(x)] = P(x) - \mathcal{KLIP}[A(x)] \). In that case, the Forward Modeling cost function is truly quadratic and convex. Unbiased astrophysical observables can be readily retrieved by direct application of Equation (3). Appendices C and D describe how this can be implemented in practice, and how in the case of point sources and RDI there exist numerical algorithms that are more tractable than a brute force minimization of Equation (3).

Figure 1 illustrates how in this configuration KLIP-FM yields unbiased photometry. This result was obtained using Hubble Space Telescope-NICMOS data and the KLIP algorithm when injecting a synthetic point source of known flux. The left panel shows the reduced images for four values of \( K_{\text{Klip}} \) (the number of Principal Components used for the data analysis) and illustrates how the detectability of the point source changes with this parameter. When \( K_{\text{Klip}} \) is too small, the point source is not detected. It only becomes apparent for larger values \( K_{\text{Klip}} = 50 \), albeit with some residual spatially correlated speckle noise in the image (e.g., \( P_{\text{spc}}(x) \sim 0 \)). This noise obviously contaminates the astrophysical observables. When \( K_{\text{Klip}} = 200 - 400 \), the residual noise disappears (\( P_{\text{spc}}(x) \sim 0 \)) but the point source has been significantly over-subtracted. The right panel of Figure 1 illustrates over-subtraction increasing with \( K_{\text{Klip}} \) when Forward Modeling is not used. In this case, only the numerator of Equation (22) is taken into account. This corresponds to a matched filter or to cross-correlating the reduced image of a point source, including corrugations due to the data analysis algorithm, with the uncorrugated instrument PSF. Without Forward Modeling and for large \( K_{\text{Klip}} \) then the photometric estimate is wrong by a factor of three. However, when using Forward Modeling (e.g., Equation (22)) we find that the injected photometry is retrieved at the ~5% level for \( K_{\text{Klip}} \) that is large enough. This corresponds to the regime for which the residual speckle noise is sufficiently well behaved. Unfortunately, most modern high-contrast instruments often privilege strategies combining ADI + SSDI. In this case, the reference images contain astrophysical signals and \( \mathcal{KLIP}[T(x) - \hat{A}(x)] = P(x) - \mathcal{KLIP}[\hat{A}(x)] \); as a consequence, the method outlined in Appendix C, which relies on only considering over-subtraction, is most often not applicable.
2.4. Propagation of Astrophysical Signal through KLIP

In this paper we introduce an analytical expansion that quantifies the propagation of the astrophysical signal through KLIP, even the presence of self-subtraction. Moreover we show that when the astrophysical signal is small, this expansion only depends on $A(x)$ in a linear fashion. The proof of this result is described in Appendix E, along with the linear algebra formalism necessary to implement it in computer calculations. We do not provide these technical details here, and we only focus on the implications of this expansion. Moreover, instead of discussing the most general framework of Appendix E, we give the example of a faint point source detected in IFS data (as in Appendix F). The spectrum of this point source is $s$. If the source is faint enough with respect to the speckles, then the Principal Components associated with any reference stack picked within a most general ADI+SSDI observing sequence, can simply be written as:

$$Y_k(x) = Z_k^A(x) = Z_k(x) + f^T \Delta Z_k(x),$$

(4)

where $\Delta Z_k(x)$ is a matrix that only depends on the $Z_k(x)$, $\Lambda_k$ eigenpair and on the instrument PSF; it does not depend on the point source’s spectrum. In other words, in our example of a faint source this analytical expansion captures the propagation of astrophysical signal through a least-squares speckle fitting algorithm (KLIP here) in a linear fashion: $KLIP[I_0(x) + A(x)] = KLIP[I_0(x)] + f^T \Delta KLIP[I_0(x), PSF(x)]$. The actual expression of this expansion is given in Appendix E Equations (53) and (55). While this result is presented in the context of PCA-based algorithms, it can also be applied to algorithms that rely on linear combinations of images (e.g., LOCI). This is in virtue of the direct equivalence between LOCI and KLIP discussed by Savransky (2015). The three main terms in this expansion of $Z_k^A(x)$ have already been discussed in the literature in the context of LOCI. We describe them qualitatively here:

1. the unperturbed Principal Components $Z_k(x)$ that capture the correlations of the instrument PSF. These are normalized such that $||Z_k(x)|| = 1$ and are responsible for over-subtraction.
2. the perturbation to the Principal Components that captures the direct self-subtraction associated with the presence of an astrophysical source at various parallactic angles and wavelengths in the observing sequence. If $s$ is the brightness of the astrophysical source, then this term scales as $s/\sqrt{\lambda_i}$. In the case of LOCI, this term can be modeled by multiplying synthetic images of the astrophysical source at various parallactic angles and wavelengths by their corresponding LOCI coefficients. This is the term that Esposito et al. (2014) correct in the case of disk imaging.
3. the perturbation to the Principal Components that captures the indirect self-subtraction associated with correlations, albeit small ones, between the astrophysical signal and the speckles. This term scales as $s/\Lambda_k$. In the case of LOCI+ADI this term can be quantified by conducting a perturbation analysis of the LOCI coefficients, as introduced by Brandt et al. (2013).

Because the unperturbed eigenvalues are ordered by decreasing magnitude, we identify three regimes of astrophysical biases in high-contrast data analysis:

1. when $K_{KLIP}$ is small, over-subtraction dominates. Provided that the astrophysical source can be detected, the solution described in Section 2.3 can be applied.
2. when $K_{\text{Klip}}$ has an intermediate value, direct self-subtraction dominates the biases. Provided that the astrophysical source can be detected, methods based on linear combinations of images (e.g., LOCI), along with the method described in Esposito et al. (2014), are best suited.

3. when $K_{\text{Klip}}$ is large, which might be the only recourse for very faint astrophysical sources, indirect self-subtraction dominates the biases. In this configuration one can take advantage of our expansion to predict the influence of an injected synthetic negative source of spectrum $\tilde{f}$:

$$Z^2_f(x) = Y_f(x) + \tilde{f}^T \Delta Y_f(x)$$

In other words, here we have applied our analytical expansion to propagate to the synthetic negative source through the data analysis algorithm: $K \text{LTP}[T(x) - A(x)] = K \text{LTP}[T(x)] - \tilde{f}^T \Delta K \text{LTP}[T(x), \text{PSF}(x)]$. Substituting this expression for $Z^2_f(x)$ into Equation (3) yields a quadratic Forward Modeling cost function. This ensures that the Forward Modeling optimization will converge toward the global minimum (e.g., no pathological biases). Moreover, because each evaluation of Equation (3) is calculated only via a simple matrix multiplication (see Appendix F for details), Forward Modeling then becomes tractable even with highly dimensional astrophysical observables (such as IFS data).

This latter case is of course the most interesting one, which we discuss when presenting practical applications of Equation (4) in Sections 3 and 4.

2.5. Validity of This Expansion

We now study the validity of our linear approximation. The mathematical rationale associated with this aspect is described Appendix E. We in particular direct the reader toward Equation (42), which can be used priori to decide whether or not Equation (4) is valid. We illustrate the various regimes of this approximation using public Gemini Planet Imager J-band data on Beta Pictoris, obtained in 2013 December as part of GPI commissioning activities. Details about the observations and scientific implications are presented in Bonnefoy et al. (2014). In this paper we do not discuss the exoplanet in this system and instead inject synthetic planets at other locations in the GPI field of view. We chose this data for our numerical examples in this paper because GPI raw J-band data is dominated by speckles (when compared to the results reported in Ingraham et al. 2014; Chilcote et al. 2015) and because Beta Pictoris is the brightest star with GPI public commissioning data in this filter. Figures 2–3 illustrate how the linear model described by Equation (4) fares when compared with propagating numerically (without any approximation) a synthetic source through the KLIP algorithm. We carried out this test using target images from a single GPI exposure, without limiting the ensemble of potential references (e.g., the target image is chosen for a given $t_o$ but the references are picked among all $t_1 \ldots t_{N_{ref}}$). There is no loss of generality associated with using a single exposure to illustrate the validity of Equation (4), because it can easily be generalized by derotation and summation and over all exposures of an ADI sequence. We start with Figure 2, which addresses the case of a point source that cannot be detected in raw IFS data. The top panel compares numerical KLIP data with the linear model for non-aggressive parameters. A detailed description of algorithm parameters is given in Appendix A. For the sake of our discussions here (and for the remainder of the paper) the main consideration to remember is that “aggressive” corresponds to parameters that are tuned to reduce speckles very efficiently, and thus reveal the faintest underlying point sources. In the case of the top row of Figure 2, while the model fares very well, the injected point source (at the same location as in the bottom panel) is barely seen by eye and cannot be distinguished from local speckles. On the other hand, when using the same geometry but more aggressive settings, the faint point source is detected. The cross sections in the bottom panel of Figure 2 illustrate how the PSF morphology of the injected source is very much altered by KLIP. This results in biases on the astrophysical estimates when the inference is carried out on these reduced images and in the absence of Forward Modeling. However, the results from numerical KLIP and the linear model are in very good agreement, even with an aggressively selected PSF library, the cross sections corresponding to the reduced data and the linear model are completely indistinguishable. This demonstrates that the analytical expansion in Equation (4) does indeed capture with high fidelity the degradation of the astrophysical signal due to the speckle noise fitting algorithm. As a consequence, one can in principle predict this degradation prior to any measurements and use our analytical model for unbiased astrophysical inference.

On the other hand, Figure 3 illustrates two cases for which the linear approximation in Equation (4) does not hold. Indeed, as predicted in Equation (42) the linear approximation is not valid for point sources brighter (top panel of Figure 3) or as bright as (bottom panel) the local speckles when using relatively aggressive KLIP parameters. Fortunately, Figure 4 shows that simply changing the KLIP parameters to less aggressive settings yields better agreement between the actual reduced KLIP image and the linear model in both configurations (point source brighter and as bright as speckles). In this case, the cross sections corresponding to the reduced data and the linear model are much closer one to another on Figure 4 than on Figure 3 (albeit not matching perfectly). These cases are somewhat of limited interest because they operate in configurations for which the point source can be detected in raw IFS data (either in a single slice or in an IFS cube where it would stand immobile when compared to the speckles). For those brightnesses, aggressive KLIP might not be needed. This illustrates the limitations of the perturbation method presented in Section 2.4 and how its applicability can be extended to bright objects, provided that the least-squares PSF subtraction parameters are chosen to be non-aggressive.

3. EXAMPLE 1: IFS SPECTROSCOPY OF POINT SOURCES

3.1. Forward Modeling with Astrophysical Signal in the PSF Library

Up until this point, our discussion, and in particular the analytical perturbation of Principal Components presented in Section 2, was general and could be applied to extended objects. We now consider a more specific example, in which we show that Equation (4) can be used to estimate the spectrum of faint point sources in IFS data. Because of the high dimensionality ($N_\lambda$) of the astrophysical observables potentially affected by
self-subtraction, this problem is often considered to be one of the most challenging in coronagraph data analysis. The injection of fake point sources and/or the direct minimization of Equation (3) can be made tractable in the cases of RDI or ADI (Marois et al. 2010a; Mazoyer et al. 2014). However, the presence of astrophysical signal at other wavelengths in the reference library (e.g., when using SSDI) renders the spectral estimation problem very degenerate. These degeneracies, along with the large number of unknown astrophysical quantities are an important obstacle to the spectral characterization of the fainter substellar companions discovered using modern high-contrast instruments (see Marois et al. 2010a; Pueyo et al. 2012 for examples).

The linear expansion in Equation (4) can alleviate this problem entirely. Indeed, the negative synthetic source can be propagated through the algorithm a priori, and thus inference can occur without having to compute multiple times the costly matrix inversion associated with KLIP. Moreover, under the assumption that Equation (4) holds (see discussion in Section 2.5 and in Appendix E), it does capture the actual degradation of the astrophysical signal due to over- and self-subtraction in a linear fashion. When neglecting the higher order terms in \( \tilde{f} \) in \( \beta(\mathbf{x}) \), \( Z_k^A(\mathbf{x}) \geq Z_k^A(\mathbf{x}) \) (e.g., for \( \tilde{f} \sim f \) is small), the Forward Modeling cost function in Equation (3) becomes quadratic. This ensures that there are no pathological cases for which KLIP-FM converges to a local minimum. In Appendix F we describe in greater detail how to carry out astrophysical inference by injecting Equation (4) into Equation (3). There, we show that the spectral extraction consists of (1) running the KLIP algorithm, (2) building a wavelength and pixel dependent model of the point source propagated through the KLIP algorithm, (3) based on this model, building a \( N_x \times N_y \) matrix whose diagonal terms capture over-subtraction,
while off-diagonal terms capture self-subtraction, and (4) invert this matrix to retrieve the spectrum of the detected point source. We insist that while our method is mathematically equivalent to injecting a negative synthetic source into the data, we do not carry out the multidimensional minimization in Equation (3) “as is.” Instead, we rely on the formalism of Appendices E and F to built a linear model of the corruption of the astrophysical signal and then invert this model. Our algorithm thus does not feature a series of iterations to find the global minimum of Equation (3). Of course, this method is not perfectly suited to quantify the stochastic uncertainties associated with the extracted spectrum. However the “frequentist” approach presented in Appendix F has the benefit of being computationally cheap: in this paper, we limit our examples to this simple inversion algorithm. As discussed in Section 3.3, Forward Modeling can also be carried out in the Bayesian sense: in this case, Equation (3) (along with adequate weighting to capture the properties of the residual noise) becomes a likelihood function in which the contribution of the negative synthetic source is accounted for according to Equation (4).

3.2. Results with Synthetic Point Sources

In order to test the KLIP-FM IFS spectral estimation algorithm described in Appendix F, we inject point sources of known spectra in the GPI J-band Beta Pictoris public data set, and illustrate how this method can alleviate SSDI algorithmic biases. To do so, we study the three regimes illustrated in Figures 2–3, using two types of underlying spectra for the synthetic sources: a flat spectrum (in units of contrast: same point source to star flux ratio as a function of wavelength) and a sharp triangular spectrum. These two cases can be considered, respectively, as the most and least challenging in terms of high-contrast IFS spectral estimation. Indeed, the former is particularly difficult because the presence of the point source at other wavelengths in the reference images yields a “local derivative” of the spectrum after KLIP or LOCI, see Marois et al. (2014a). As a consequence, retrieving a flat
spectrum might be difficult using the simple inversion described above. On the other hand, a sharp triangular spectrum, with a significant fraction of the bandpass for which the point source’s flux is small, might be intuitively more amenable to this type of analysis. In this section we again illustrate the performances of KLIP-FM in various configurations by using only one data cube for the target image.

Figure 5 illustrates how the estimated spectrum varies as a function of the number of Principal Components ($K_{\text{Klip}}$) when the injected source has a sharp spectral feature whose maximum flux is brighter (left column) and as bright (right column) as the speckles. The top row shows the estimated spectrum when using aperture photometry and the bottom row when using KLIP-FM. The solid thick line represents the injected spectrum and each thin dashed line is an estimated spectrum that corresponds to $K_{\text{Klip}} = 1 \ldots N_{\text{Corr}}$ (where $N_{\text{Corr}}$ is the number of images in the PSF library). It was generated using non-aggressive parameters close to the ones in Figure 4.

Because indirect self-subtraction, which scales as $1/\lambda_k$, worsens when increasing the number of KLIP modes, one expects the estimated spectrum after KLIP in the absence of Forward Modeling to vary as $K_{\text{Klip}}$ increases. This is exactly the behavior that the top two panels of Figure 5 (and subsequent figures) exhibit. On the other hand, self-subtraction is accounted for with Forward Modeling and thus the estimated spectrum should in principle not depend on $K_{\text{Klip}}$ (provided that the residual speckle noise is small enough and the linear approximation is valid). Again, this is exactly what happens in the bottom panels of Figure 5: KLIP-FM reduces the sensitivity of the estimated spectrum to $K_{\text{Klip}}$ and brings the estimated spectrum closer to the injected one. However, in cases for which the point source is at least as bright as the speckles, Forward Modeling is not absolutely necessary because aperture photometry after KLIP with non-aggressive reductions still yields an estimate of the spectrum within $\sim 10\%$ of the injected signal (top panels). This is because over-subtraction does not

Figure 4. Example of configurations for which Equation (4) holds for bright point sources. Two leftmost panels show the raw and KLIP processed images in the spectral channel centered around 1.26 µm of the GPI J-band filter. The middle panel shows a model of the KLIP processed image, predicted by Equation (55) based on the Principal Components calculated in the absence of the synthetic source in the reference library and the model of the point source moving azimuthally/radially across the PSF library. The two rightmost panels then show the horizontal and vertical cross section of both the actual and model-based KLIP images at the location of the injected synthetic source. On these cross sections, both the data and mode overlap. When compared to Figure 3, this example illustrates that simply changing the KLIP parameters to less aggressive settings yields better agreement between the actual reduced KLIP image and the linear model, both in the brighter than and as bright as speckles configurations.
depend on $\Lambda_k$, and direct self subtraction only scales as $1/\sqrt{\Lambda_k}$.
In this context KLIP-FM is a tool to reduce uncertainties. This is of course only true when using algorithm parameters chosen so that Equation (4) holds. Indeed, when the point source is relatively bright and the speckle fitting is aggressive, Equation (4) does not hold and biases still remain after KLIP-FM when comparing the estimated and injected spectra. This is illustrated on Figure 6: in spite of a somewhat reduced sensitivity to $K_{KLip}$, a significant offset remains between injected and extracted spectra.

Such considerations do not apply to point sources that are fainter than the speckles, for which Equation (4) is always valid. We discuss results obtained in this configuration using both flat (left columns) and peaky (right columns) spectra along with non-aggressive (Figure 7) and aggressive (Figure 8) KLIP settings. In the case of non-aggressive subtractions, the post-KLIP aperture photometry spectra are significantly biased by residual speckle noise. Because KLIP-FM still operates under the assumption that the residual speckle noise is well behaved (e.g., $P_{spe}(x) \sim 0$), which is clearly not applicable to this case, Forward Modeling does still yield biases in both the flat and peaky spectra configurations (Figure 7). This ought to be expected for such algorithm settings that do not seek to reach the absolute best speckle least-squares fitting, yielding the type of correlated residuals illustrated in the top panel of Figure 2. On the contrary, Figure 8, which was obtained using aggressive settings this time, shows that the bright portion of a peaky spectrum becomes very close to the injected spectrum when using KLIP-FM. The spectral fidelity in this case is almost as good as Figure 5, even though the injected point source is 15 times fainter. As predicted above, results using a flat spectrum exhibit somewhat lesser fidelity when compared to the ground truth: in particular both the aggressive and non-aggressive settings seem to yield similar biases. However, these biases are much smaller in the case of KLIP-FM than without using Forward Modeling. Even in this worse-case scenario of a flat spectrum point source only detectable at the $\sim 1\sigma$ level (see images in the bottom panel of Figure 2), residual biases are at the 20%–30% level, which is well below the statistical uncertainties that are expected for the low-significance detections we simulated here.

3.3. Residual Biases and Statistical Uncertainties

Figures 5–8 clearly demonstrate how KLIP-FM reduces the systematic biases associated with spectral extraction of faint point sources in IFS coronagraph data. However, even if the final estimated spectrum is a much weaker function of $K_{KLip}$ (which we use as a proxy to establish the ability of KLIP-FM to correct for over- and self-subtraction), deviations from the injected spectrum are noticeable in the case of fainter point sources. We investigated
this feature by carrying out the same analysis as in Figures 5–8, except that in a first step we set the flux of the injected point source to zero to quantify the residual speckle floor. For illustration, the estimated spectrum obtained under this null hypothesis is given in the top panel of Figure 9. Subtracting this “residual speckle noise flux” to the KLIP-FM spectral estimate yields the bottom panel of Figure 9. We indeed obtain a bias-free estimated spectrum in the bright channels of the spectrum. We confirm by eye inspection that the bluer end of the spectrum corresponds to non-detections, which explains the remaining offset. We find a similar outcome when repeating this test for all the configurations shown on Figures 5–8. The test on Figure 9 illustrates that with KLIP-FM, spectral estimation is not limited by over- and self-subtraction. By and large the post-KLIP-FM biases stem from the residual speckles in the reduced images.

Of course in practice this null test cannot be carried out and the estimated spectrum, along with its associated uncertainties, will be affected by poorly subtracted speckles. When using a full ADI sequence this will be alleviated by co-adding cubes over time, in virtue of the central limit theorem as described in Marois et al. (2008a). In that respect, our non-ADI single cube test is somewhat of a pessimistic configuration. If the observing strategy does not include ADI, the brightness of the residual speckles can also be minimized by adjusting KLIP parameters (within the range for which the linear approximation remains valid, as discussed in Section 2.5). Regardless of these adjustments, the estimation of confidence intervals associated with the, now unbiased, KLIP-FM extracted spectrum is of critical importance for astrophysical inference. Most often, these confidence intervals are calculated by injecting synthetic point sources at various positions in the coronagraph field of view. Errors bars associated with this process include the contribution of both the possible algorithmic biases and the residual post-processed speckles. Using KLIP-FM makes the former term negligible. Correlations between spectral channels (associated with the latter term) can also be estimated based on residual noise statistics at other locations than the one of the detected point source. All these approaches yield realistic confidence intervals under the assumption that the noise properties are spatially uniform across the field of view (or at least over all azimuths at a given angular separation). When using a ground-based Adaptive Optics system, this assumption is not always true due to signatures of the wind direction in the coronagraph PSF. Our Forward Modeling approach alleviates this assumption and enables the estimation of confidence intervals based the contribution of the residual speckles at the location of the point source. This can be achieved by estimating spatial and spectral co-variances at positions where astrophysical signal is absent and introducing them into Equation (3), so it becomes a true likelihood function in the Bayesian sense (see Greco & Brandt (2016) for details on how the spectral correlation can be included). This represents

Figure 6. Comparison between KLIP+aperture photometry and KLIP-FM with a source featuring sharp spectral features and with flux brighter than and as bright as the speckles. Aggressive reductions. Top left: KLIP+aperture photometry with a source brighter than speckles. Top right: KLIP+aperture photometry with a source as bright as speckles. Bottom left: KLIP-FM with a source brighter than speckles. Bottom right: KLIP-FM with a source as bright as speckles. The solid tick line represents the injected spectrum and each thin dashed line is an estimated spectrum corresponding to $K_{\text{KLIP}}$. This figure serves as a cautionary tale in cases for which the point source can be identified in the raw data. The downward arrow indicates the variations of the estimated spectrum as a function of $K_{\text{KLIP}}$. As predicted by our analytical expansion, self-subtraction scales at $1/L_\lambda$ and gets more and more severe in the absence of Forward Modeling as $K_{\text{KLIP}}$ increases. With KLIP-FM, this sensitivity is greatly reduced, however some significant biases remain. This was predicted in Figure 3, where we found that in such configurations the linear model-based KLIP deviates from numerical KLIP. This issue can simply be solved by using less aggressive parameters; see Figure 5.
significant progress when compared with the present state, in particular for data sets that feature significant residuals associated with atmospheric wind. However a full end to end demonstration of this approach is beyond the scope of this paper, and we leave this analysis to an upcoming publication (J. Wang et al. 2016, in preparation).

4. EXAMPLE 2: DETECTABILITY OF FAINT POINT SOURCES IN IFS DATA

4.1. Detection Threshold and Completeness

We now tackle the actual detection problem, which is the decision process that chooses whether or not to trigger a detection alarm and take action accordingly (in the case of a first epoch this action consists of carrying out confirmation observations). This problem has been extensively discussed by Caucci et al. (2007), Marois et al. (2008a), Mugnier et al. (2009), Ygouf et al. (2013), Mawet et al. (2014), Wahhaj et al. (2015), Cantalloube et al. (2015), and Gomez Gonzalez et al. (2016). Here we revisit these results in the context of KLIP-FM. A detection algorithm can be seen as an observer\(^1\) that estimates the probability that flux in a given set of pixels originates from an astrophysical signal rather than scattered starlight (speckles) or other sources of noise. Because the characteristic scales of speckle noise mimic the presence of a planet for most classes of observers, a preliminary routine aimed at calibrating this noise is necessary. In this paper we discuss algorithms based on least-squares PSF fitting for this denoising step. After this is carried out, statistical inference regarding the presence of a certain class of point sources, or lack thereof, then occurs using the chosen observer. If a given combination of pixels (as defined by the observer) is above a given threshold (often chosen to minimize the false positive rate), then an alarm is triggered and follow-up observations are carried out. On the other hand, if no alarm is triggered, the range of astrophysical objects that are absent from the data (e.g., completeness) is then quantified to inform the statistical distribution of such object across the ensemble of stars observed (see work by Vigan et al. 2012; Nielsen et al. 2013; Brandt et al. 2014 for recent exoplanet surveys). In this section we describe how, for a given set of chosen algorithm parameters, KLIP-FM can keep the false positive rate similar to the one obtained without Forward Modeling while significantly increasing completeness.

\(^{1}\) Here “observer” refers to an image analysis algorithm such as the Hotelling Observer described in Caucci et al. (2007), not an individual collecting astronomical data with a telescope.

4.2. Maximizing True Positives while Minimizing False Negatives

Our goal is to quantify the efficiency of KLIP-FM when applied to the detection problem. To do so we compare two types of observers:

1. De-noising with KLIP and then aperture photometry (denoted KLIP+ApPhot).
2. De-noising with KLIP and then inversion of Equation (70) (KLIP-FM).

We depart from the common practice in the high-contrast imaging community that consists of comparing the local signal-to-noise ratio \( S/N \) of images obtained using various algorithms. Instead we follow the prescription described in Caucci et al. (2007) and recently revisited by Gomez Gonzalez et al. (2016). Using such metrics is now becoming common practice in high-contrast imaging. Note that this approach was also pointed out by Wahhaj et al. (2015), who demonstrated that the rigorous way to assess the efficiency of various joint denoising and detection methods is to study them under the paradigm of minimizing the False Positive Fraction \( \text{FPF} \) while maximizing the True Positive Fraction \( \text{TPF} \). For the sake of brevity we do not recall the formal definition of these quantities, and refer the reader to the excellent presentations in Mawet et al. (2014) and Wahhaj et al. (2015). In practice the FPF and TPF can be estimated using numerical simulations as follows:

1. We choose a hypothesis for the underlying population of astrophysical signal we want to test: this includes the brightness of the point sources, their separation from the star, and their underlying spectrum.
2. We generate a series of \( 2 \times N_{\text{synthetic}} \) data sets: \( N_{\text{synthetic}} \) have a signal as prescribed in the previous step and the other \( N_{\text{synthetic}} \) do not have signal (e.g., their brightness has been set to zero). This latter data set serves as a null hypothesis test in the assessment of the data analysis algorithm.
3. We propagate these data sets through each denoising + observer algorithm whose performance we want to assess.
4. Based on these simulations, we build the empirical Probability Density Function (PDF) of the scalar metrics given by each observer under both the signal present and absent hypotheses. This yields the histograms shown in Figures 10 and 11.
5. The FPF captures the probability that, for a given threshold, the observer will classify an event as an astrophysical detection while it is actually stemming from noise realizations. It is thus calculated as the area under the curve of the “no point source” histogram, from the threshold to \( +\infty \).
6. The TPF measures completeness (i.e., the probability that the astrophysical signal will be classified as such and not as noise). It is thus calculated as the area under the curve of the “point source” histogram, from the threshold to $+\infty$.

7. We then move the threshold from the left to the right of each histogram and compute the FPF and TPF at each threshold value. This yields the Receiver Operating Characteristic (ROC), which is parametric curve describing $\text{TPF} = \text{roc}(\text{FPF})$. This ROC can then be used to compare denoising+observer methods.

As described extensively in the Imaging Science literature (Caucci et al. 2007 and references therein), having the ROC follow a straight line between $(0, 0)$ and $(1, 1)$ implies that TPF and FPF are always equal for all values of threshold: the observer is no better than a coin toss. On the other hand, having the ROC follow a perfect elbow from $(0, 0)$ to $(0, 1)$ to $(1, 1)$ implies that there exists an optimal value for the threshold for which the “no point source” and “point source” histograms do not overlap at all, thus enabling the recovery of all possible astrophysical signals without any false positive. In this case the observer is ideal. In this framework the area integrated under the ROC curve (AUC) is the figure of merit that quantifies the performances of a given denoising+observer combination. We thus ran a series of numerical tests to compare the performance of KLIP+Aperture Photometry and KLIP-FM under this metric.

Figures 10 and 11 show the histograms of observed counts in the case of synthetic companions as bright as the speckles. This figure was generated using $N_{\text{s}} = 50$ injections of true positives, at a separation of $0^{\prime\prime}4$ and random azimuthal positions.

In all cases, we find that using Forward Modeling in conjunction with KLIP fares better than simply using aperture photometry after KLIP. KLIP-FM shifts to the right the “point source present” histogram while only slightly changing the tail of the histogram associated with no signal. This reduces the area of the “confusion zone,” where the two histograms overlap.
conjunction with KLIP fares better than simply using aperture photometry after this algorithm. A closer look at Figures 10 and 11 illustrates how KLIP-FM does shift to the right the histogram of counts when an astrophysical signal is present, while only slightly changing the tail of the histogram associated with an absent signal. This reduces the area of the “confusion zone” where the two histograms overlap.

Figure 11. Histograms of the observed counts in the case of synthetic companions as bright as the speckles. This figure was generated using a $N_{\text{synth}} = 50$ injection of true positives at separation of $0''4$ and random azimuthal positions. Left: the observer measures the maximum of the counts in an aperture the size of the FWHM of a PSF in the “blue” channels at the base of the spectral feature in both the KLIP+ApPhot and the KLIP-FM cases. Right: the observer measures the maximum of the counts in an aperture the size of the FWHM of a PSF in the “red” channels at the peak of the spectrum in both the KLIP+ApPhot and the KLIP-FM cases. In both cases, we find that using Forward Modeling in conjunction with KLIP fares better than simply using aperture photometry. KLIP-FM shifts to the right the histogram of counts associated with point source signal injected in the data, while only slightly changing the tail of the histogram associated with no signal. This reduces the area of the “confusion zone” where the two histograms overlap.

Finally we illustrate how this procedure can also be used in a more systematic manner for “planet search.” Instead of building a model for hypothetical point sources scattered across the field of view, we present a point-wise implementation of KLIP-FM that inverts Equation (70) at each point of the astrophysical scene, one at a time. Figure 13 was generated using this method over a portion of the GPI field of view, with a synthetic companion that is five times fainter than the local speckles, aggressive PSF parameters, and we only used one GPI data cube. The leftmost column of each panel Figure 13 shows the images from the KLIP algorithm, and the next column shows the flux across wavelengths obtained with point-wise KLIP-FM. The two right columns show the horizontal and vertical cross section of both images at the location of the injected point source. This figure highlights the four advantages of point-wise KLIP-FM:

1. In the low flux channels, denoted as $\lambda = 1.199, 1.232, 1.291$, and $1.324 \mu m$, the KLIP image features hints of faint flux at the location of the injected point source, but the single wavelength detection is much more convincing in the left column with point-wise KLIP-FM. This is in part due to the convolution by the instrument PSF, which would occur regardless of Forward Modeling. However, this is also due to the fact that inverting Equation (70) does shift to the right the signal present histogram (as shown in Figure 10), thus helping to discriminate faint astrophysical signals from residual speckle noise.

2. In all channels the centroid of the signal after KLIP is very much affected by the over/self-subtraction. On the other hand, with point-wise KLIP-FM, the maximum of the retrieved signal is at the location of the injected source in the channels for which the residual speckle noise is uncorrelated. This is highly beneficial when using detection metrics that rely on the stability of the point source location as a function of wavelength. It also has significant advantages for astrometry.
3. In all channels the point-wise KLIP-FM counts correspond to the unbiased spectrum of the point source (see Section 4).

4. For all locations corresponding to non-detections, over/self-subtraction have also been corrected; such images can be readily used to derive detection limits.

In spite of all these advantages, point-wise KLIP-FM, as implemented to generate Figure 13, presents one major drawback: it is painstakingly slow and memory hungry (generating Figure 13 required a day of computations on a standard macbook pro laptop, for only one GPI data cube and a small fraction of the field of view). Carrying out the forward model construction and inversion “as is” for an entire field of view and over an entire observing sequence (e.g., following the procedure outlined in Appendices A–F for every pixel in the field of view of a coronagraph imager) is thus prohibitively expensive computationally. However, simplifications exist: in particular the exceedingly large intermediate matrices of Appendices E and F do not need to be evaluated in all cases, and one can show that temporary variables of lower dimensionally can be used instead. The details of the linear algebra associated with these simplifications is beyond the scope of this paper and will be presented in the future (J. B. Rufio et al. 2016, in preparation). Here we simply use our computationally inefficient implementation point-wise KLIP-FM to illustrate in Figure 13 that this algorithm has the potential to become a very powerful tool for exoplanet detection via direct imaging.

5. CONCLUSION AND PERSPECTIVES

In this paper we introduced a linear expansion that captures the impact of over/self-subtraction in high-contrast imaging data. This is done in the most the general case for which the reference images of the astrophysical scene move azimuthally and/or radially across the field of view (ADI and/or SSDI). This method is based on perturbing the covariance matrix underlying any least-squares speckles problem and propagating this perturbation through the data analysis algorithm. Most of the work in this paper has been presented in the PCA framework, but it can be easily generalized to methods relying on linear combinations of images (instead of eigenmodes). Based on this linear expansion, we then demonstrated how this new algorithm could be used in practice. We first considered the case of the spectral extraction of faint point sources in IFS data (under the ADI+SSDI observation strategy) and illustrated, using public Gemini Planet Imager commissioning data, that our novel perturbation-based Forward Modeling can alleviate algorithmic biases. We then applied KLIP-FM to the detection of point sources and showed how it decreases the rate of false negatives while keeping the rate of false positives unchanged, when compared to classical KLIP.

Beyond these two examples, our analytical result is broadly applicable to a wide range of high-contrast science:

1. **Planet detection**: should the point-wise KLIP-FM described in Section 4.3 be improved upon so it can be implemented in an efficient manner, it will facilitate the detection of the faintest end of the point sources buried in the residual speckles of ongoing surveys. Note, however, that this gain will only occur if astronomers relax their standards to trigger follow-up observations. Indeed, for very faint planets if the threshold is solely based on a FPF...
then the TPF will be close to 0% regardless of whether or not Forward Modeling is used. This is illustrated in Figure 12 for the faintest example considered in this paper. However, if a FPF of 20% is tolerated, then Forward Modeling will bring the completeness, or TPF, up 70%. Without Forward Modeling, increasing the allowable FPF to 20% only brings completeness up to 20%, which is still no better than a coin toss. This demonstrates that loosening detection threshold might be beneficial, now that we are equipped with a tool that can greatly increase completeness at only a modest cost in observing efficiency (in our example, one in five follow-up observations is triggered based on a bright speckle instead of a true astrophysical point source). Given the paucity of currently directly imaged exoplanets, we argue that the experimental design of ongoing surveys should consider such an option.

2. **Planet detection**: in this manuscript we only considered observers that integrate flux over an aperture (with and without Forward Modeling). More sophisticated observers such as the ones presented in Kasdin & Braems (2006) and Caucci et al. (2007) could be used. Moreover, the linear model developed here could also be used in a Bayesian framework. It was recently shown that a simple model of dual-image ADI subtraction in such a framework was an effective method for the detection of faint sources (Cantalloube et al. 2015). Because of the analytical derivation presented herein, similar work can now be conducted using more sophisticated PCA-based denoising algorithms.

3. **Detection limits in IFS data**: In the case of non-detections, completeness is estimated for each hypothesis regarding the potential astrophysical signal that was not observed. In the case of broadband imaging with RDI or ADI, this ensemble of astrophysical hypothesis is of low dimensionality because the observables are simply separation and integrated brightness. In the case of IFS observations, this dimensionality significantly increases. Moreover, when using SDI or SSDI, the ROC, and thus the completeness, varies as a function of the hypothesized underlying spectrum. This significantly complicates the population statistics for this type of observation, and is one of the outstanding problems for the statistical analysis of ongoing large high-contrast surveys with IFS (Beuzit et al. 2008; Hinkley et al. 2008; Macintosh et al. 2014). In Appendix F, we briefly describe how KLIP-FM could be used to address this issue, but leave out numerical examples for future work.
4. Astrometry: point-wise Forward Modeling can be carried out at the subpixel level around a first guess for the location of a detected point source. This in principle should yield high precision astrometry, even when using ADI and/or SSDI. We will demonstrate the performance of this method in a future paper (J. Wang et al. 2016, in preparation).

5. Retrieval of physical properties of planets: as discussed in Section 3, because of its simplicity Equation (67) is amenable to be used within a Bayesian framework to evaluate correlated uncertainties associated with the spectrum of a faint point source. One could also directly fit the physical model “in the data” using Equation (67) as the cost function in retrieval codes, such as the one recently presented in Line et al. (2015).

6. Disk imagery and characterization: the biggest practical difference between high-contrast disk imagery and point source detection resides in the choice of optimal least-squares reduction parameters. In this paper we presented a Forward Modeling that is parameter independent, and as a consequence all our discussions are in principle applicable to disk imaging and characterization. Note however that in practice Forward Modeling with disks is complicated by the fact that Equation (10) cannot be simplified using a simple PSF as the astrophysical model: every hypothetical disk morphology must be explored.

The amount of work required to robustly devise these algorithmic improvements and thoroughly test them goes beyond the scope of what can be achieved by a single individual. It is our hope that the community will conduct the investigations in data analysis development outlined here in a collaborative manner and include promising advances in publicly available tools, such as Wang et al. (2015).

We thank the two anonymous referees for their feedback and extremely important suggestions, which made this paper significantly more readable. This manuscript also benefited from critical inputs from the entire Gemini Planet Imager team. In particular we thank Dimitry Savransky, who carefully went over the linear algebra presented in the Appendices; Abhijit Rajan, Kim Ward-Duong, and Marshall Perrin, who provided insightful suggestions on the writing and presentation; and Christian Marois and Kate Morzinski, who provided over the linear algebra presented in the Appendices; Abhijit Rajan, Kim Ward-Duong, and Marshall Perrin, who provided insightful suggestions on the writing and presentation; and Christian Marois and Kate Morzinski, who provided important suggestions regarding the context of the paper. The vast majority of the results in Section 4 stemmed from initial discussions during the Exoplanet Imaging Workshop whose findings are presented in Lawson et al. (2012). Finally, this paper would have never been written without Jason Wang and Jean Baptiste Ruffio, who were able to confirm the findings presented herein using a separate implementation of KLIP-FM and made their code available to the community (Wang et al. 2015).

APPENDIX

Reminder of the structure of the Appendices:

1. Appendix A provides the most general formalism for an ADI+SSDI observing sequence and lays out the formal foundations for our work.

2. Appendix B summarizes the notations in Appendix A in a table format. In order to facilitate numerical implementation, it provides the dimensions of the various matrices discussed in this paper.

3. Appendix C introduces the formalism underlying Forward Modeling in the most general case, and then discusses the specific configuration of RDI. This was already presented in Pueyo et al. (2015), but serves to set up the stage for Appendix F.

4. Appendix D describes how to carry out Forward Modeling for the astrometry and photometry of point sources for RDI within the framework of the linear algebra notations introduced in Appendices A and C.

5. Appendix E contains the proof of our novel analytical expansion. It heavily relies on the notations introduced in Appendix A and contains the key innovation of this paper.

6. Appendix F describes how to take advantage of the results in Appendix E to carry out Forward Modeling to estimate the spectrum of point sources in IFS data.

APPENDIX A

NARRATIVE EXPLANATION OF THE VARIOUS NOTATIONS

In this appendix we provide a detailed description of our mathematical notations regarding the most general case of a high-contrast observing sequence and a generic implementation of the KLIP algorithm (e.g., we put Table I into words). Note that all the material in this appendix has already been discussed in the literature, but is revisited here to provide a rigorous framework for the latter introduction of the KLIP-FM algorithm. An illustration of the algorithm parameters discussed here is given in Figure 14.

A.1. Observing Sequence

We consider the general case of an observing sequence with an Integral Field Spectrograph (IFS) and in the presence of field rotation (ADI). An image, $I_\lambda(x)$ at the wavelength $\lambda$ and at time $t$ (parallactic angle $\theta_l$) within the observing sequence, can be written as:

$$I_\lambda(x) = S_{\psi_{\lambda}} \left( \frac{x}{\lambda} \right) + \epsilon a_0 A_\lambda (R_0 [x])$$

(6)

where:

1. $S_{\psi_{\lambda}}$ is the focal plane intensity associated with speckles (integrated over the narrow bandpass around $\lambda$ at time $t$) that results from a random realization $\psi_{\lambda}$ of the telescope + instrument wavefront.

2. $x$ are the 2D coordinates across the field of view.

3. $\epsilon$ is equal to zero if there is no astrophysical signal. If an object is present, it is equal to the integrated photometry of the astrophysical source over the entire bandpass of the IFS.

4. $a = [a_1, a_2, \ldots, a_N]$ is the normalized spectrum of the object. Namely, if the actual spectrum of the object at the resolution of the IFS is $f = [f_1, f_2, \ldots, f_N]$ then

$$\epsilon = \sum_{p=1}^{N} f_p a_p = f_p / \epsilon.$$  

5. $\epsilon a_0 A_\lambda (x)$ is the image at $\lambda$ of the astrophysical source, at the spatial resolution of the instrument, rotated north up.

6. $R_0$ corresponds to the 2D rotation matrix—with respect to the stellar location—associated with the azimuthal motion of the astrophysical source across the ADI observing sequence. $\theta_l$ corresponds to the parallactic angle (direction of north in the images), which varies across an ADI sequence.
7. throughout the paper $x$ corresponds to the 2D coordinates across the field of view. In the linear algebra formalism discussed in the appendices, and for practical implementations, these two dimensions can be collapsed onto one. For instance if $x$ describes all the possible pixel coordinates across the field of view (of size $N_{pix} \times N_{pix}$), then $x$ is a $1 \times N_{pix}^2$ array.

PCA-based reduction algorithms use a well-chosen library of images to build an empirical model of the speckle noise realization associated with each target image within the observing sequence. Each empirical model is then subtracted from its corresponding target image to increase the S/N of potential astrophysical sources. Without a loss of generality, we choose the target image at wavelength $\lambda_0$ at the exposure starting at $t_0$ (parallactic angle $\theta_0$).

$$T(x) = I_{\lambda_0,t_0}(x) = S_{\psi_{\lambda_0,t_0}}(x) + \epsilon A_{\lambda_0}(R_{\psi_0}[x]),$$

(7)

The corresponding reference library is then assembled by choosing among all other possible images with $(\lambda, t) \neq (\lambda_0, t_0)$. This captures the most general configuration discussed in this paper. Of course there exist observation scenarios for which it greatly simplifies:

1. when using RDI (a PSF library built using images of other sources) and under the assumption that the library has been built to be “signal free” (see for instance Choquet et al. 2014), the atmospheric signal is only present in the target image $T = S_{\psi_{\lambda_0,t_0}}(x) + \epsilon A_{\lambda_0}(x)$. In this case $\epsilon = 0$ in all images in the reference in the PSF library and thus $R_k = S_{\psi_{\lambda_0,t_0}}$. These are the shorthanded notations described in Soummer et al. (2012) (here the state of the telescope+instrument $\psi$ does not depend on time and wavelength, it is simply indexed over the ensemble of reference stars).

2. when using ADI with non-IFS data (or when using IFS data that excludes images at other wavelengths from the PSF library), we can drop the wavelength dependence for both the spatial scaling of the speckle noise and the brightness of the potential astronomical objects. Then Equation (6) reduces to:

$$I_{\lambda}(x) = S_{\psi_{\lambda_0,t_0}}(x) + \epsilon A_{\lambda_0}(R_{\psi_0}[x]),$$

(8)

where $a$ is now a scalar instead of a vector.

A.2. Reference PSF Library

A.2.1. Spatial Rescaling and Image Plane Motion of A Point Source

The first step in least-squares speckle fitting is to build for each target $T(x) = I_{\lambda_0,t_0}(x)$ its corresponding ensemble of reference PSFs — $R_{\lambda,t}(x)$—by choosing among all other possible images with $(\lambda, t) \neq (\lambda_0, t_0)$. In the most general case (e.g., when using SSSI and ADI) all IFS slices are first spatially rescaled to $\lambda_0$, so that the characteristic scale of the speckle noise in the references matches the noise in the target image. Thus, the reference images are:

$$R_{\lambda,t}(x) = I_{\psi_0,\lambda_0}(x) = S_{\psi_{\lambda_0,t_0}}(x) + \epsilon A_{\lambda_0}(R_{\psi_0}[x]),$$

(9)

In the case of ADI only, this rescaling is not necessary and the wavelength dependence can be dropped. The flux normalized astrophysical scene seen by the instrument can be written as the convolution of the sky—$\text{Sky}(x)$—by the instrument PSF:

$$A_\lambda(x) = \int \text{Sky}_\lambda(u)PSF_{\lambda,t}(u-x)du,$$

(10)

where the wavelength and field dependence of the coronagraphic PSF are captured in the subscripts of $PSF_{\lambda,t}$. Using these notations and neglecting PSF field dependence, the motion of the astrophysical signal at given field point $x_S$ associated with wavelength scaling and field rotation $-A_\lambda(R_{\psi_0}[x_S/\lambda_0])$—is captured by $PSF_\lambda(u - R_\psi[x_S/\lambda_0])$ with:

$$R_\psi[x_S/\lambda_0] = x_S + \delta(x_S,\lambda)$$

(11)

$$\delta(x_S,\lambda) = ||x_S||\Lambda/\lambda_0(\cos(\theta)n + \sin(\theta)e).$$

(12)

where $n, e$ are the unit vectors pointing north and east. $\delta(x_S,\lambda)$ is a 2D vector that relates the position in the field of view of a hypothetical point source in the science image of interest—at $(t_0, \lambda_0)$—to its position in each of the spatially rescaled reference images (at $(t_0, \lambda_0) = (t, \lambda)$).

This motion of the astrophysical scene with respect to the speckle noise across the instrument field of view is key to building PSF libraries for which the signal in the reference PSFs is not located at $x_S$ (thus enabling local empirical fitting of the speckles only, with “minimal contamination from the signal”).

A.2.2. Reference Selection Criteria

More formally, we build such a collection of reference images by ensuring that $\delta(x_S,\lambda)$ is large enough so that there is no (or little) astrophysical signal at $x_S$ in the PSF library. We write these references as $R = \{R_k(x), k = 1 ... N_R\}$. This library is constructed over subsections of the image, or subtraction zones $S$, centered on $x_S$. Here we parameterize these zone in polar coordinates: radial extent $dr$ (or $N_r$ annuli across the field of view) or azimuthal extent $d\phi$ (or $N_{\phi}$ sectors per annulus). Note that in this paper we do not follow the method described in Lafreniere et al. (2007), which splits the geometry of the problem between optimization ($O$) and subtraction zones ($S$), and we solely focus on the case for which $O = S$. However, in principle, the KLIP-FM formalism is also applicable when $S$ is a subregion of $O$. We write $u_{S,S}$ and $u_{S}\delta$ as the radial and tangential unit vectors in the direction of $x_S$. Whether or not an image is included in this library is then decided using a combination of the following criteria:

1. $k$ such that $(\delta(x_S,\lambda)_S - \delta(x_S,\lambda)_S)_S > N_0 * \text{FWHM}(PSF_{\lambda_0})$ to account for the minimal motion of a source due to field rotation. FWHM($PSF_{\lambda_0}$) is the Full Width at Half Maximum of the instrument’s PSF at wavelength $\lambda_0$.

2. k such that $\delta(x_S,\lambda)_S > N_{\delta S} * \text{FWHM}(A_{\lambda_0})$ to account for the minimal outward motion of a source due to speckle chromaticity.

3. $k$ such that $\delta(x_S,\lambda)_S < N_{\delta S} \times \text{FWHM}(A_{\lambda_0})$ to account for the minimal inward motion of a source due to speckle chromaticity. Very often $N_{\delta S} = N_{\delta \lambda}$. However, as
explained in Marois et al. (2014a), it can be very beneficial to use different values when seeking to detect faint companions with sharp spectral features. In this case, $\mathcal{N}_{\lambda}^{\delta}$ can be chosen based on the hypothetical underlying sharp spectral feature of the hypothetical astrophysical signal.

4. $k$ such that the reference belongs to the $\mathcal{N}_{\text{Corr}}$ images with the largest correlation with the target. Note that we adopt the following notation for correlations for the remainder of the paper: $\langle R_k, T \rangle_S = \int_S R_k(x)T(x)dx$.

Note that in the case of point sources, when $\text{Sky}_k(\mu) = a_kP_{\text{source}}(\mu - u_{\text{source}})$, with $P_{\text{source}}(0, 0) = 1$ and zero otherwise, the first three selection criteria above directly relate to the flux contamination across wavelengths and rotation angle. For all the examples in this manuscript, we simplify the reference selection by using $N_\delta^{-\lambda} = N_\delta^{\lambda} = N_\delta^0$. We also introduce the following shorthand notations:

1. $E_\text{obs}(x)$ is the overall ensemble of PSFs in the observing sequence. When folding the two-dimensional spatial variable $x$ into a line vector it can be seen as a matrix with $N_{\text{Exp}} \times N_l$ lines and $N_{\text{Pix}}$ columns ($N_{\text{Exp}}$ is the number of exposures in the observing sequence and $N_{\text{Pix}}$ the number of pixels in the $S$ zones). Note that $x$ corresponds to 2D coordinates that are folded into one dimension for practical reasons. That is, if the $S$ zone was the entire field of view (of dimension $N_{\text{fov}} \times N_{\text{fov}}$ pixels), then one row entry of $E_\text{obs}(x)$ would be of dimension $1 \times N_{\text{fov}}^2$. The same applies to $R(x)$.

2. $R_{\lambda_0, \theta_0, (N_{\text{Exp}}N_lN_{\text{Pix}}, N_{\text{Exp}}N_{\text{Pix}}, N_{\text{Exp}}N_{\text{Corr}}, N_{\text{Corr}}, N_{\text{Pix}}, N_{\text{fov}})}(x)$ is the ensemble of reference PSFs chosen to analyze a target image at $(\lambda_0, \theta_0)$. When folding $x$ into a line vector, it becomes a matrix with $N_\text{Exp}$ lines and $N_{\text{Pix}}$ columns ($N_\text{Exp}$ is the number of frames selected in the PSF library). For clarity, we drop the dependence on algorithm parameters and write this matrix as $R(x)$.

3. $\text{Sel}_{\lambda_0, \theta_0, (N_{\text{Exp}}N_lN_{\text{Pix}}, N_{\text{Exp}}N_{\text{Corr}}, N_{\text{Corr}}, N_{\text{Pix}}, N_{\text{fov}})}$ is an $N_\text{Exp} \times N_l$ selection matrix whose entries are defined by $S[i,j] = 1$ if $E_{\text{obs}}[j](x)$ is the $i$th entry in the reference library and 0 otherwise. Or, more succinctly: $R(x) = \text{Sel} E_{\text{obs}}(x)$, \hspace{1cm} (13)

where again we write $\text{Sel}$, dropping the dependence on $S$, $\lambda_0$, $\theta_0$ and $(N_{\text{Exp}}, N_l, N_{\text{Corr}}, N_{\text{Pix}}, N_{\text{fov}})$.

A.2.3. Principal Component Analysis

Once the reference library corresponding to a given target image has been assembled, the PCA is carried out as follows:

1. Zero mean $\bar{T}(x)$ and $R_k(x)$ over $S$.

2. Calculate the Karhunen–Loève transform of $R(x)$:

$$Z_k(x) = \frac{1}{\sqrt{\lambda_k}} \sum_{m=1}^{N_l} v_k[m] R_m(x),$$ \hspace{1cm} (14)

where the vectors $V_k = [v_k[1], \ldots, v_k[N_l]]$ are the eigenvectors of the $N_l$-dimensional covariance matrix of the reference library $C_{RB} = \langle R(x), R(x) \rangle_S = \langle R(x) R(x)^T \rangle$, and correspond to its eigenvalues $\{\lambda_k\}_{k=1, \ldots, N_l}$.

3. Choose a cutoff, $K_{\text{Klip}}$, for the number of modes the target image will be projected on.

4. Project the target image on the Principal Components and subtract this projected speckle noise model from the target image:

$$\mathcal{KLIP}_R[T(x)] = P(x) = T(x) - \sum_{k=1}^{K_{\text{Klip}}} \langle T(x), Z_k(x) \rangle_S Z_k(x).$$ \hspace{1cm} (15)

This algorithm was outlined in Amara & Quanz (2012) and Soummer et al. (2012). Soummer et al. (2012) suggested that such a formalism could serve as a foundation to calibrate potential systematic errors on the astrophysical observables due to the reduction algorithms, however they did not delve into the details of such a procedure. The present manuscript addresses this outstanding point.

APPENDIX B

TABLE SUMMARIZING OUR MATHEMATICAL NOTATIONS

Here we summarize the various notations introduced in the discussions in the appendices (along with the dimensions of each variable). It is our hope that this summary will help the reader through the more technical arguments of this manuscript, along with helping interested parties implement the KLIP-FM algorithm.

APPENDIX C

FORWARD MODELING IN THE CASE OF RDI

In this appendix we provide a detailed description of our mathematical notations and Forward Modeling implementation in the case of RDI, in configurations for which there is no astrophysical signal in the PSF library. Most of the material in this appendix has already been discussed in Pueyo et al. (2015), albeit in somewhat less detail. We revisit it here to provide a rigorous context for the latter introduction of the KLIP-FM algorithm in the general case of ADI and/or SSDI.

C.1. Basic Principle of Forward Modeling

Forward Modeling in the context of exoplanet imaging was first proposed by Marois et al. (2010a) and aims at jointly estimating the instrument response and the astrophysical signal. To do so, negative synthetic sources are injected in the raw data across the entire observing sequence. This new data set, with both positive astrophysical and negative synthetic signals, is then propagated through the reduction algorithm. Jointly minimizing the residuals in such processed images (by exploring the range of possible astrophysical properties for the synthetic negative sources) retrieves in principle the properties of the astrophysical signal. We call $\hat{A}$ the ensemble of estimated astrophysical observables $\hat{A} = \{\hat{\alpha}, \hat{\alpha}_{\lambda}, \hat{\alpha}_{\lambda_{\gamma}}, \hat{A}_{\lambda}(x), \ldots, \hat{\alpha}_{\lambda_{\gamma}}(x)\}$. These are the quantities corresponding to the synthetic negative astrophysical signal injected in the data, while $A$ are the quantities corresponding to the actual signal. This notation covers the most general case (i.e., resolved source, not centered on the star, and whose morphology changes with wavelength). Of course in practice one never faces such a challenge and the dimensionality of astrophysical estimates is much smaller. We can then write the Forward Modeling problem at a given wavelength $\lambda_0$ and time $t_0$ as the following minimization:

$$\min_{\hat{A}} \| \mathcal{LSQ}_{\mathcal{R}_\lambda}[\hat{A}] [T(x) - \hat{\alpha}_{\lambda_{\gamma}}(R_{\theta_0}(x))] \|_\mathcal{F}^2,$$ \hspace{1cm} (16)
Table 1
Description of the Mathematical Notations in This Appendix

| Symbol       | Expression | Dimensions | Comments |
|--------------|------------|------------|----------|
| \(x\)        |            | 1 x 2      | Coordinates in focal plane |
| \(x_S\)      |            | 1 x 2      | Center of PCA subtraction zone |
| \(\delta^{(\text{AD}-\lambda)t}x_S\) |            | 1 x 2      | Radial and azimuthal motion of an hypothetical source located at \(x_S\) over an ADI+SSDI observing sequence |
| \(\theta_t\) |            | 1 x 2      | ADI field rotation corresponding to the exposure at time \(t\) |
| \(dr = \frac{1}{N_t}\) |            | 1 x 1      | Radial extent of the local \(S\) zone of the field of view over which the speckles' least-squares fitting occurs |
| \(d\theta = \frac{1}{N_0}\) |            | 1 x 1      | Azimuthal extent of the \(S\) zone |
| \(\delta x_q\) |            | 1 x 1      | Azimuthal displacement of an astrophysical source at the \(q\)th exposure |
| \(N_{00}\)   |            | 1 x 1      | ADI exclusion criterion |
| \(N_{0\lambda}\) |            | 1 x 1      | SDI exclusion criterion (inward) |
| \(N_{0\lambda}^+\) |            | 1 x 1      | SDI exclusion criterion (outward) |
| \(N_{\text{Corr}}\) |            | 1 x 1      | Number of most correlated references kept in PSF library |
| \(N_{\Omega}\) |            | 1 x 1      | Number of references in PSF library, in this paper we use \(N_{\text{Corr}} = N_{\Omega}\) |
| \(N_{\text{pix}}\) |            | 1 x 1      | Number of pixels in the \(S\) zone |

Coordinates and Algorithm Parameters

| Symbol       | Expression | Dimensions | Comments |
|--------------|------------|------------|----------|
| \(x_\lambda\) |            | 1 x 2      | Location of an astrophysical point source |
| \(x_{\text{syn}}\) |            | 1 x 2      | Location of a synthetic negative point source underlying Forward Modeling |
| \(x_{\lambda}\) |            | 1 x 2      | Estimated location of an astrophysical point source |
| \(\epsilon\)    |            | 1 x 1      | Photometry of an astrophysical source |
| \(\bar{\epsilon}\) |            | 1 x 1      | Photometry of a synthetic negative point source underlying Forward Modeling |
| \(\bar{\epsilon}\) |            | 1 x 1      | Estimated photometry of an astrophysical source |
| \(f\)        | \(f_i\) \[
\begin{bmatrix}
    f_1 \\
    f_2 \\
    \vdots \\
    f_{N_{\lambda}}
\end{bmatrix}
\] | \(N_{\lambda} \times 1\) | Spectrum of an astrophysical point source |
| \(f_{\lambda}\) |            | \(N_{\lambda} \times 1\) | Normalized spectrum of an astrophysical source |
| \(f_{\text{syn}}\) |            | \(N_{\lambda} \times 1\) | Spectrum of a synthetic negative point source underlying Forward Modeling |
| \(\tilde{a}\)  | \(\tilde{a}\) \[
\begin{bmatrix}
    \tilde{a}_{\lambda 1} \\
    \vdots \\
    \tilde{a}_{\lambda N_{\lambda}}
\end{bmatrix}
\] | \(N_{\lambda} \times 1\) | Normalized spectrum of a synthetic negative source underlying Forward Modeling |
| \(\tilde{f}\)   | \(\tilde{f}\) \[
\begin{bmatrix}
    f_1 \\
    f_2 \\
    \vdots \\
    \tilde{f}_{N_{\lambda}}
\end{bmatrix}
\] | \(N_{\lambda} \times 1\) | Estimated spectrum of an astrophysical source |
| \(\tilde{\tilde{a}}\) | \(\tilde{\tilde{a}}\) \[
\begin{bmatrix}
    \tilde{a}_{\lambda 1} \\
    \vdots \\
    \tilde{a}_{\lambda N_{\lambda}}
\end{bmatrix}
\] | \(N_{\lambda} \times 1\) | Estimated normalized spectrum of an astrophysical source |
| \(\tilde{a}_{\lambda}\) | \(\tilde{a}_{\lambda}\) \[
\begin{bmatrix}
    a_{\lambda 1} \\
    \vdots \\
    a_{\lambda N_{\lambda}}
\end{bmatrix}
\] | \(N_{\lambda} \times N_{\lambda}\) | Vector of a normalized flux of the astrophysical source corresponding to the signal contained in each \(N_{\lambda}\) reference image |
| \(a\)          | \(a\) \[
\begin{bmatrix}
    a_{\lambda 1} & \cdots & 0 & \cdots & 0 \\
    \vdots & \ddots & \vdots & \ddots & \vdots \\
    0 & \cdots & a_{\lambda N_{\lambda}} & \cdots & 0 \\
    \vdots & \ddots & \vdots & \ddots & \vdots \\
    0 & \cdots & 0 & \cdots & a_{\lambda N_{\lambda}}
\end{bmatrix}
\] | \(N_{\lambda} \times N_{\lambda}\) | Matrix whose diagonal elements have been populated by each one of the \(N_{\lambda}\) entries of \(\tilde{a}\) |

Astrophysical quantities

Vectors and Matrices in Data Space

| Symbol       | Expression | Dimensions | Comments |
|--------------|------------|------------|----------|
| \(S_{\lambda}(x)\) | \(S_{\lambda}(x)\) \[
\begin{bmatrix}
    S_{\lambda}(x) \\
    S_{\lambda}(x) \\
    \vdots \\
    S_{\lambda}(x)
\end{bmatrix}
\] | \(1 \times N_{\text{pix}}\) | Scattered starlight (speckles) at wavelength \(\lambda\) in the zone \(S\) corresponding to the state of the instrument at time \(t\) |
| \(S(x)\)      | \(S(x)\) \[
\begin{bmatrix}
    S_{\lambda}(x) \\
    S_{\lambda}(x) \\
    \vdots \\
    S_{\lambda}(x)
\end{bmatrix}
\] | \(N_{\lambda} \times N_{\text{pix}}\) | Matrix of the concatenated speckles realizations kept in the reference PSF library |
| \(S\)         | \(S\) \[
\begin{bmatrix}
    S_{\lambda}(x) \\
    S_{\lambda}(x) \\
    \vdots \\
    S_{\lambda}(x)
\end{bmatrix}
\] | \(N_{\lambda} \times N_{\text{pix}}\) | Short-handed notation for \(S(x)\) |
| \(A_{\lambda}(R_{\theta}(x))\) | \(A_{\lambda}(R_{\theta}(x))\) \[
\begin{bmatrix}
    a_{\lambda}(x) \\
    a_{\lambda}(x) \\
    \vdots \\
    a_{\lambda}(x)
\end{bmatrix}
\] | \(1 \times N_{\text{pix}}\) | Astrophysical image at wavelength \(\lambda\) that has been rotated by \(\theta_t\) due to ADI field rotation |
| \(T(x)\)      | \(T(x)\) \[
\begin{bmatrix}
    T(x) \\
    T(x) \\
    \vdots \\
    T(x)
\end{bmatrix}
\] | \(1 \times N_{\text{pix}}\) | Target image at \((\lambda_0, \theta_0)\) |
| \(R_{\lambda}(x)\) | \(R_{\lambda}(x)\) \[
\begin{bmatrix}
    R_{\lambda}(x) \\
    R_{\lambda}(x) \\
    \vdots \\
    R_{\lambda}(x)
\end{bmatrix}
\] | \(1 \times N_{\text{pix}}\) | \(k\)th image in the reference PSF library for the target image |
Table 1 (Continued)

| Symbol | Expression | Dimensions | Comments |
|--------|------------|------------|----------|
| $R(x)$ | \[ \begin{bmatrix} R_{\lambda_0,0}(x) \\ \vdots \\ R_{\lambda_g,0}(x) \end{bmatrix} \] | $N_R \times N_{\mathrm{pix}}$ | Matrix of the concatenated references in the PSF library |
| $P_{\lambda,t}(x)$ | $S + a \lambda_t$ | $1 \times N_{\mathrm{pix}}$ | Processed image at wavelength $\lambda$ and time $t$ |
| $R$ | $A_{\lambda}(x) = \begin{bmatrix} A_{\lambda_0}(R_{\lambda_0}(x) \cdot \frac{\lambda_0}{\lambda}) \\ \vdots \\ A_{\lambda_g}(R_{\lambda_g}(x) \cdot \frac{\lambda_g}{\lambda}) \end{bmatrix}$ | $N_R \times N_{\mathrm{pix}}$ | Short-handed notation for $R_{\lambda}(x)$ |
| $A_{\lambda}$ | Matrix of the concatenated $N_R$ astrophysical images in the reference library, at wavelength $\lambda$ and $\theta_t$ due to ADI field rotation and rescaled to wavelength $\lambda_0$ |
| $Z_k(x)$ | \[ \begin{bmatrix} Z_1(x) \\ \vdots \\ Z_{N_R}(x) \end{bmatrix} \] | $N_R \times N_{\mathrm{pix}}$ | Matrix of the concatenated KL modes associated with speckles |
| $\Delta Z_k(x)$ | \[ \begin{bmatrix} \Delta Z_k^1(x) \\ \vdots \\ \Delta Z_k^{N_R}(x) \end{bmatrix} \] | $N_R \times N_{\mathrm{pix}}$ | Perturbation of the speckles' KL modes decomposed as a function of wavelength |
| $Y_k(x)$ | $Y_k(x) = Z_k(x) + \epsilon \Delta Z_k(x)$ | $1 \times N_{\mathrm{pix}}$ | KL modes of the Instrument PSF perturbed by astrophysical signal = KL modes of the actual data |
| $\Delta Y_k(x)$ | \[ \begin{bmatrix} \Delta Y_k^1(x) \\ \vdots \\ \Delta Y_k^{N_R}(x) \end{bmatrix} \] | $N_R \times N_{\mathrm{pix}}$ | Perturbation of the speckles' KL modes decomposed as a function of wavelength |
| $F_{\lambda,t}(x)$ | see text, too ugly | $N_R \times N_{\mathrm{pix}}$ | Model of the astrophysical source propagated through the data analysis algorithm at $(\lambda, t)$, decomposed as a function of wavelength |

Eigenvalues, Eigenvectors, Covariance Matrices

| Symbol | Expression | Dimensions | Comments |
|--------|------------|------------|----------|
| $C_{SS}$ | \[ SS^T \] | $N_R \times N_R$ | Covariance of the speckles |
| $\lambda_k$ | \[ v_k[1] \\ \vdots \\ v_k[N_R] \] | $N_R \times 1$ | $k$th eigenvalue of $C_{SS}$ |
| $V$ | \[ \begin{bmatrix} v_1[1] & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & v_k[m] \end{bmatrix} \] | $N_R \times N_R$ | Elements of the $k$th eigenvector of $C_{SS}$ arranged on the diagonal of a square matrix |
| $C_{RR}$ | \[ RR^T \] | $N_R \times N_R$ | Covariance matrix of the reference library |
| $C_{AS}$ | \[ a \lambda \Sigma_S S + \lambda \Sigma_S \Sigma^T \] | $N_R \times N_R$ | Cross term between speckles and an astrophysical signal |
Table 1 (Continued)

| Symbol | Expression | Dimensions | Comments |
|--------|------------|------------|----------|
| $\Gamma_k$ | $u_k[1]$ | $N_R$ | $k$th eigenvector of $C_{RR}$ as defined by the perturbation analysis of this paper |
| $u_k$ | $\ldots$ | $u_k[N_R]\ldots$ | |

Miscellaneous linear algebra

| Symbol | Expression | Dimensions | Comments |
|--------|------------|------------|----------|
| $I_{NR}$ | $[1 \ldots 0 \ldots 0]$ | $N_R \times N_R$ | Identity matrix of size $N_R$ |
| $L$ | $[I_{N_e} \ldots I_{N_e}]$ | $N_e \times N_e$ | Rectangular matrix that relates each slice of each exposure to the its wavelength |
| $Sel$ | $S[i, j] = 1$ if $E_{p(j),q(j)}$ | $N_R \times N_e$ | Selection matrix that relates each slice of each exposure to its position in the reference library (note it depends on algorithm parameters $(\mathbf{S}_{N_{Exp},N_{Exp}},N_{Exp},N_{Exp})$) |
| $Sel_{\text{A}}$ | $L \text{Sel}^p$ | $N_e \times N_R$ | Selection matrix that relates each slice in the reference library to its wavelength |

where $LSQ_{R(A,\tilde{A})}$ describes a most general and generic least-squares speckle fitting algorithm (using reference images that depend both on the astrophysical observables, $A$, and their synthetic negative counterparts, $\tilde{A}$), $F$ is a “fit” region of the field of view that does not necessarily correspond to the $S$ zone. In the context of this paper, we assume that $LSQ_{R(A,\tilde{A})}$ corresponds to $KLIIP_{R(A,\tilde{A})}$ described in Appendix A.2.2, with reference images chosen according to the rules described in Appendix A.2.3. Note that in practice, images from a sequence of exposures (and/or wavelengths) are co-added before the signal is estimated, and as a consequence Equation (16) is only representative of realistic cases up to one or two summations. However, for the sake of clarity we present our work without these summations and only discuss them when outlining practical implementations for spectral extraction in Appendix F.

C.2. RDI of a Point Source

We review here applications of Equation (16) to the case of RDI. In this configuration the signal can be over-subtracted, that is, the image of a point source can be fitted using some combination of instrument noise realizations (e.g., it can be over-subtracted by a speckle at the same exact location in the reference library). In this case the reference library does not depend on the astrophysical ($A$) or the synthetic negative ($\tilde{A}$) signals. The least-squares speckles fitting algorithm can then be written as an operator on any arbitrary image $I(x)$: $KLIIP_{R(A,\tilde{A})}$ = $KLIP_{R(A,\tilde{A})} \{ I(x) \} = I(x) - \sum_{k=1}^{K_{\text{Max}}} \langle I(x), Z_k(x) \rangle Z_k(x)$. Here the simplification $R(A,\tilde{A}) = R$ is key because it alleviates all issues associated with self-subtraction: only over-subtraction remains. Assuming that the morphology of the PSF is known (e.g., $A(x) = A(x) = \text{PSF}(x)$) and that is not field dependent, then there are only three scalar unknowns: the photometry over the bandwidth of interest $\epsilon$ and the location of the point source (astrometry) in the scene $x_\Lambda$. We call $(\tilde{\epsilon}, x_\Lambda)$, the values of $(\tilde{\epsilon}, x_\Lambda)$ that actually minimize Equation (16). Under these assumptions the Forward Modeling problem, Equation (16) reduces to the following minimization:

$$
(\tilde{\epsilon}, x_\Lambda) = \arg \min_{(\tilde{\epsilon}, x_\Lambda)} \|LSQ_{R}[T(x)] - (\tilde{\epsilon}A(x) - x_\Lambda) - \sum_{k=1}^{K_{\text{Max}}} \langle Z_k(x), Z_k(x) \rangle Z_k(x) \|_F
$$

(17)

where Equation (17) means that because the PSF library (and thus the Karuhnen–Loeve modes) does not contain any astrophysical signal, estimating the astrophysical observables of a detected point source can occur in processed image space. This implies that one does not need to reprocess the data for every evaluation of the Forward Modeling cost function (e.g., every value of $(\tilde{\epsilon}, x_\Lambda)$ that is hypothesized while iterating to find the global minimum of Equation (17)). In particular, the CPU intensive matrix inversion associated with the determination of the Principal Components is only carried out once. Moreover Equation (17) also provides insights regarding the signal estimation algorithm. Indeed, the quantity that is being minimized can be decomposed into three terms:

1. A noise term that represents the remaining speckle noise that has not been captured by the PCA:

$$
P_{\text{spe}}(x) = S_{\text{spe}}(x) - \sum_{k=1}^{K_{\text{Max}}} \langle S_{\text{spe}}(x), Z_k(x) \rangle Z_k(x).
$$

(18)

Algorithms such as LOCI or KLIP are aimed at changing the PDF of this noise from spatially correlated + Rician, in the raw data (Soummer & Aime 2007a), to spatially uncorrelated + Gaussian (Marois et al. 2008a; with the smallest standard deviation).

2. A term capturing the difference between the astrophysical signal $\epsilon A(x - x_\Lambda)$ and the negative synthetic source,
\[ \hat{\xi} A(x - \bar{x}_A); \text{ both are propagated through the PCA:} \]
\[ (\xi A(x - \bar{x}_A) - \hat{\xi} A(x - \bar{x}_A)) + \sum_{k=1}^{K_{\text{KLIP}}} (\xi A(x - \bar{x}_A) - \hat{\xi} A(x - \bar{x}_A)) \]
\[ (\xi A(x - \bar{x}_A), Z_A(x))_S Z_k(x). \]  
(19)

Plugging these expressions into Equation (17) yields:
\[ (\hat{\xi}, \bar{x}_A) = \arg \min_{(\hat{\xi}, \bar{x}_A)} \| P(x) - \hat{\xi} G(x - \bar{x}_A) \|_S^2, \]  
(20)

where \( \hat{\xi} G(x - \bar{x}_A) \) corresponds to the negative synthetic source propagated through KLIP. We have thus established that Forward Modeling with RDI simply consists of inverting the “transfer function” of the algorithm: namely solving for \( (\hat{\xi}, \bar{x}_A) \) in the least-squares sense. When the algorithm parameters have been well chosen, the noise term \( P_{\text{spec}}(x) \) is indeed zero mean, Gaussian, and does not feature spatial correlations; it can be substituted into Equation (25) as:
\[ \begin{align*}
\hat{f}_0 &= \arg \min_{f_0} \| b(x) - f_0^T G(x - \bar{x}_0) \| \cr
\bar{x}_0 &= \arg \min_{x_0} \| b(x) - \bar{\gamma}_0^T G(x - \bar{x}_0) \| \end{align*} \]  

Thus when seeking to minimize a quadratic cost function of a functional form similar to Equation (24) the spatial offset \( x_0 \) can be first calculated by maximizing the cross-correlation in Equation (29), assuming a first guess for the spectrum. Based on this value of \( x_0 \), one can update \( \bar{\gamma}_0 \) using Equation (28) and iterate until astrometry and photometry have converged.

**APPENDIX D**

**LINEAR ALGEBRA NOTATIONS UNDERLYING ASTROMETRY AND PHOTOMETRY OF A POINT SOURCE WITH KLIP-FM**

This Appendix details the linear algebra associated with the derivation of Equations (22) and (21). It also sets up the stage for the mechanics underlying the spectral estimation algorithm discussed in Appendix F.

**D.1. General Case**

We consider the general case of minimizing the following Forward Modeling cost function:
\[ \arg \min_{(f_0, x_0)} \| b(x) - f_0^T G(x - x_0) \| \]  
(24)

where \( f_0 \) is a \( N_x \times 1 \) column vector, \( G(x - x_0) \) is a \( N_x \times N_{\text{pix}} \) matrix, and \( b(x) \) is a \( 1 \times N_{\text{pix}} \) line vector. This is the most general case for point sources whose position, spectrum, and astrometry are being estimated. Thus, the pair \( (\hat{f}_0, \bar{x}_0) \) is a solution to this least-squares minimization problem if and only if:
\[ \hat{f}_0 = \arg \min_{f_0} \| b(x) - f_0^T G(x - \bar{x}_0) \| \]  
(25)

and
\[ \bar{x}_0 = \arg \min_{x_0} \| b(x) - \bar{\gamma}_0^T G(x - x_0) \| \]  
(26)

Using the shorthanded notation \( G_{\bar{x}_0} = G(x - \bar{x}_0) \), we can write Equation (25) as:
\[ \hat{f}_0 = \arg \min_{f_0} \| b b^T + f_0^T G_{\bar{x}_0} G_{\bar{x}_0}^T f_0 - 2 f_0^T G_{\bar{x}_0} b^T \]. \]  
(27)

Taking the first derivative with respect to \( f_0 \) yields:
\[ G_{\bar{x}_0} G_{\bar{x}_0}^T \hat{f}_0 = G_{\bar{x}_0} b^T \]  
(28)

which can be substituted into Equation (26) and yield after simplification:
\[ \bar{x}_0 = \arg \max_{x_0} \bar{\gamma}_0^T G(x - x_0) b^T \]. \]  
(29)

C.3. Calculating Correlations with Fourier Transforms

Note that in practice, this algorithm can be implemented very efficiently at the subpixel precision level using Fourier Transforms. Indeed, the correlation between the two images \( I_1(x) \) and \( I_2(x) \) over \( \mathcal{S} \) can be written as:
\[ (I_1(x), I_2(x - \bar{x}_A))_S = \text{MFT} \left[ \text{FFT}[m_S I_1(x)] \text{FFT}[m_S I_2(x)] \right] (\bar{x}_A) \]  
(23)

where \( m_S \) denotes a mask over the image that is zero everywhere but in \( \mathcal{S} \), FFT is a fast Fourier Transform, and MFT is the Matrix Fourier Transform discussed in Soummer et al. (2007b), calculated over a subpixel grid of \( \bar{x}_A \) centered around an initial guess of the position of the detected point source. As a consequence, the entire 2D grid of correlations necessary to solve for the location of the point source can be calculated without resorting to CPU expensive loops.
APPENDIX E

ANALYTICAL PROPAGATION OF THE ASTROPHYSICAL SIGNAL THROUGH A PCA

This appendix describes the derivation of the main theoretical result of this paper.

E.1. Linear Expansion of The Covariance Matrix

The expression of a target image and its associated references are:

\[
T(x) = S_{\psi_{\lambda_0}}(x) + c a A_{\lambda_0}(R_0[x]) \]  

(36)

\[
R_{\lambda_0}(x) = S_{\psi_{\lambda_0}}(x) + c a A_{\lambda_0}(R_0[x]) \]  

(37)

where the references have been chosen among all images in the observing sequence according to the parameters \(N_0, N_{\psi_0}, N_{\lambda_0})\). We drop this dependence for simplicity and use the following shorthand notations:

1. \(R(x)\) is the \(N_R\) by \(N_{Pix}\) matrix whose \(kth\) line entry is \(R_k(x) = R_{\lambda_k}(x)\). Note that each line entry of \(R(x)\) is zero mean over \(S\).

2. \(S(x)\) is the \(N_R\) by \(N_{Pix}\) matrix whose \(kth\) line entry is \(S_k(x) = S_{\lambda_k}(x)\). Note that each line entry of \(S(x)\) is zero mean over \(S\).

3. \(a\) the \(N_R\) diagonal matrix whose \(kth\) diagonal entry is \(a_{\lambda_k}\) (the normalized astrophysical flux at the wavelength corresponding to the \(kth\) reference).

4. \(A_{\lambda_k}(x)\) is the \(N_R\) by \(N_{Pix}\) matrix whose \(kth\) line entry is \(A_{\lambda_k}(x) = A_{\lambda_k}(R_0[x])\). Note that each line entry of \(A_{\lambda_k}(x)\) is zero mean over \(S\). In the case of a point source and neglecting the field dependence of the PSF, this can be further simplified as:

\[
A_{\lambda_k}(R_0[x]) \approx \text{PSF}_{\lambda_k}(x - x_S - \delta(\lambda_k - \lambda_0)x_S) \]  

(38)

In this framework the reference library can be written in a matrix form:

\[
R(x) = S(x) + c a A_{\lambda}(x) \]  

(39)

When using PCA-based algorithms the next step is to calculate the Karhune Loeve transform of this ensemble of references (e.g., Equation (14)). Our goal is to evaluate how the signal in the references—e.g., \(a A_{\lambda}(x)\)—propagates through this Principal Component decomposition. To do so, we write the covariance matrix in the presence of an astrophysical signal as the sum of the speckles covariance and the cross-term between the astrophysical signal and the speckle noise:

\[
C_{RR} = R(x) R(x)^T \]  

\[
C_{RR} = S(x) S(x)^T + \epsilon a A_{\lambda}(x) S(x)^T + \epsilon^2 a A_{\lambda}(x) A_{\lambda}(x)^T a^T + O(\epsilon^2) \]  

Note that here we dropped the \(1/\sqrt{N_{\text{corr}}-1}\) factor in front of the covariance matrix. Thus the presence of an astrophysical signal in the PSF library becomes a quadratic (scaling as \(\epsilon^2\)) perturbation of the reference’s covariance matrix: this non-linearity, associated with the eigenmodes truncation, is the source of the self-subtraction biases for astrophysical estimates (over-subtraction occurs even when \(\epsilon = 0\) in the references).
However, when

$$e^2 a A_k(x) A_k(x)^T a^T \ll \epsilon C_{A,S}$$ (41)

is true for each entry in these matrix, then the dependence on the astrophysical signal becomes linear and can be modeled a posteriori in a tractable fashion. This argument is the crux of the analysis presented in this paper. This inequality is true when:

$$\epsilon \ll \max \left( \frac{A[C_{A,S}]}{A[a A_k(x) A_k(x)^T a^T]} \right)$$ (42)

where $A[M]$ denotes the operator that calculates the eigenvalue of a matrix $M$. As a consequence, the linear approximation holds either when the astrophysical signal $\epsilon$ is small with respect to the local speckles or when the algorithm parameters are chosen so that the correlations of astrophysical images across the PSF library $-a A_k(x) A_k(x)^T a^T$—have much smaller eigenvalues than the cross-term between the astrophysical signal and the speckle noise $-C_{A,S}$. This latter case occurs when the algorithm parameters are chosen not to be aggressive (e.g., $N_{\phi}, N_{\phi}^+, N_{\phi}^- \ll$ larger than the FWHM of a point source PSF). Note, however, that while this condition is necessary (e.g., when it is true the linear approximation holds), it is not sufficient; there exist cases where this inequality is not strictly true but where the linear model holds. In practice it is preferable to use a metric based on numerical evaluation of the eigenmodes or eigenvalues, such as the one introduced in 15. The case of small $\epsilon$ is of most interest in the framework of high-contrast imaging. Indeed when a source is brighter than the speckles, a PCA-based algorithm might not be necessary (or can be tuned so that $N_{\phi}, N_{\phi}^+, N_{\phi}^- \ll$ is large enough), and thus the sophistications presented in this manuscript can be circumvented.

E.2. Perturbed Eigenpair of The Covariance Matrix

We remind the reader that $V_k = [v_k[1] \ldots v_k[N]]$ are the eigenvectors of $C_{SS}$ and $\Lambda_k$ its eigenvalues (see discussion associated with Equation (14)). Under the linear approximation described by Equation (40), we now propagate the astrophysical signal in the PSF library through the calculation of eigenvalues/vectors of the covariance matrix. This identity is a standard linear algebra result, however because it is the cornerstone of KLIP-FM, we recall here the main steps of its derivation. We seek to express the perturbed eigenpair $\Gamma_k, U_k$ of $C_{SS} + \epsilon C_{A,S}$ as:

$$\Gamma_k = \Lambda_k + \epsilon \delta \Lambda_k$$

$$U_k = V_k + \epsilon \sum_{p=1}^{N_S} C_{k,p} V_p$$ (43)

(44)

This implies that:

$$(C_{SS} + \epsilon C_{A,S}) U_k = \Gamma_k U_k$$

$$= (\Lambda_k + \epsilon \delta \Lambda_k) (V_k + \epsilon \sum_{p=1}^{N_S} C_{k,p} V_p)$$

$$= \epsilon \delta \Lambda_k V_k + \epsilon \Lambda_k \sum_{p=1}^{N_S} C_{k,p} V_p + O(\epsilon^2)$$

$$= \epsilon \delta \Lambda_k V_k + \epsilon \Lambda_k \sum_{p=1}^{N_S} C_{k,p} V_p + O(\epsilon^2)$$ (45)

where we have used $C_{SS} V_k = \Lambda_k V_k$. We neglect the terms of order $\epsilon^2$, left multiply Equation (45) by $V_k^T$, and use the fact that $V_k^T V_q = \delta_{p,q}$ (e.g., the eigenbasis $\{V_k\}_{k=1}^{N_S}$ is orthonormal). We find:

$$\delta \Lambda_k = V_k^T C_{A,S} V_k$$ (46)

Similarly, left multiplying Equation (45) by $V_j^T, j \neq k$ yields:

$$V_j^T C_{A,S} V_k + c_{k,j} \Lambda_k = c_{k,j} \Lambda_k$$

$$c_{k,j} = \frac{V_j^T C_{A,S} V_k}{\Lambda_k - \Lambda_j}$$ (48)
Finally, forcing the normalization of \( \{ U_k \}_{k=1}^{N_k} \) yields \( a_{k,k} = 0 \). We find that the eigenpair of \( C_{RR} \) can be written as a small perturbation of eigenvalues/vectors of the signal-free reference covariance matrix:

\[
\Gamma_k = \Lambda_k + \epsilon \, V_k^T C_{A,S} V_k
\]

\[
U_k = V_k + \epsilon \sum_{j=1,j\neq k}^{N_k} \frac{V_j^T C_{A,S} V_k}{\Lambda_k - \Lambda_j} V_j
\]

Note that in this framework the perturbation to both \( \Lambda_k \) and \( V_k \) is a linear function of the astrophysical signal. The validity of Equation (50) depends both on the validity of the inequality in Equation (42) and on the magnitude of the terms in \( \epsilon^2 \).

### E.3. Validity of The Eigenpair Expansion

Figure 15 shows how this approximation fares when changing the brightness \( \epsilon \) of a synthetic point source injected in IFS coronagraph data and varying algorithm parameters. To do so we used the same public Gemini Planet Imager J-band data on the source Beta Pictoris used in the body of the manuscript. Our proxy to quantify the fidelity of the approximation is the largest of all possible \( N_{\text{Corr}} \) relative differences between the true eigenvalues of \( C_{RR} \) and the ones calculated using Equation (50). We find that the linear approximation holds very well for small values of \( \epsilon \) and making the algorithm parameters less aggressive by increasing \( N_b \) or reducing \( N_{\text{Corr}} \) results in a wider range of \( \epsilon \) for which the approximation is well behaved. It is also important to note that this metric is somewhat conservative. Indeed, in most cases we found that for \( K_{\text{Klip}} \) smaller than 20% of the dimensionality of the reference library, the approximation holds very well even with bright point sources. Indeed when using \( K_{\text{Klip}} \geq 0.2 \times N_{\text{Corr}} \) we find that in most cases, even with a point source brighter than the local speckles, our eigenvalues-based metric remains below 10%.

### E.4. From Perturbed Eigenpair to Perturbed of The Principal Components

Next, we propagate the linear expansion in Equation (50) into the Principal Components. This is achieved by plugging Equation (50) into Equation (14):

\[
Y_k(x) = \frac{1}{\sqrt{\Lambda_k}} \sum_{m=1}^{N_k} u_k[m] R_m(x)
\]

\[
Y_k(x) = \frac{1}{\sqrt{\Lambda_k} + \epsilon \, V_k^T C_{A,S} V_k}
\]

\[
\times \sum_{m=1}^{N_k} v_k[m] + \epsilon \sum_{j=1,j\neq k}^{N_k} \frac{V_j^T C_{A,S} V_k}{\Lambda_k - \Lambda_j} v_j[m]
\]

\[
\times (S_m(x) + \alpha a_{m,a} \, a_{m}(x)).
\]

This expression can be simplified by replacing the definition of the unperturbed Principal Components—e.g., Equation (14)—into Equation (52), in order to express the \( Y_k(x) \) as a function of the \( Z_k(x) \). This finally yields:

\[
Y_k(x) = Z_k(x) + \epsilon \Delta Z_k(x)
\]

\[
\Delta Z_k(x) = -\frac{1}{2\Lambda_k} V_k^T C_{A,S} V_k \, Z_k(x)
\]

\[
+ \sum_{j=1,j\neq k}^{N_k} \sqrt{\frac{\Lambda_j}{\Lambda_k - \Lambda_j}} \, V_j^T C_{A,S} V_k \, Z_j(x)
\]

\[
+ \frac{1}{\sqrt{\Lambda_k}} V_k^T a A_k(x),
\]

where again we have neglected the terms of \( O(\epsilon^2) \). Equation (53) captures how the presence of an astrophysical signal in the reference library translates into a small perturbation of the Principal Components. When neglecting the \( O(\epsilon^2) \) terms, this perturbation is linear with respect the signal’s photometry \( \epsilon \) and/or its spectrum \( f \). Here we identify the three terms discussed in Section 2.4:

1. **Over-subtraction**, scaling at \( \sim 1 \) \( Z_k(x) \)
2. **Direct self-subtraction**, scaling as \( \sim \epsilon / \sqrt{\Lambda_k} \)
3. **Indirect self-subtraction**, scaling as \( \sim \epsilon / \Lambda_k \)

\[
\epsilon \left( -\frac{1}{2\Lambda_k} V_k^T C_{A,S} V_k \, Z_k(x) + \sum_{j=1,j\neq k}^{N_k} \sqrt{\frac{\Lambda_j}{\Lambda_k - \Lambda_j}} \, V_j^T C_{A,S} V_k \, Z_j(x) + \frac{1}{\sqrt{\Lambda_k}} V_k^T a A_k(x) \right).
\]

### E.5. Linearity of The Perturbed Principal Components when Using IFS Data

One of the main features of Equation (53) is that the presence of the astrophysical signal in the reference library is captured in a linear fashion. This has important consequences when using IFS data, because the high dimensionality of the astrophysical unknowns makes it very difficult to carry out the Forward Modeling minimization in Equation (16) “as is” for SSDI observations. Here we rewrite the expression of \( \Delta Z_k(x) \) in a way that highlights this linear dependence on \( a = [a_1 \ldots a_{N_A}] \). To do so, we use the following linear algebra identity, which is true for any eigenvector \( V_i \) of the signal-less covariance matrix:

\[
V_i^T a A_i(x) = a^T \text{Sel}_i V_i A_i(x).
\]

This stems from:

1. defining \( L \) as the \( N_a \times (N_a N_{\text{Exp}}) \) rectangular matrix that relates each slice of each exposure to its wavelength \(-L = [\mathbb{1}_{N_a} \ldots \mathbb{1}_{N_a}] \). We then write \( \text{Sel}_i = L \, \text{Sel}^T \).
2. recognizing that \( \alpha^T \text{Sel}_k \) is an \( N_R \) dimensional vector whose \( k \)th entry is \( a_{\lambda_k} \) (the normalized astrophysical flux at the wavelength corresponding to the \( k \)th reference).
3. recognizing that for any \( N_R \) dimensional vector \( b: b \, V_i = V_i^T b \), where \( V_i \) and \( b \) denote the matrices whose diagonal elements are populated with the \( N_R \) entries of the vectors \( V_i \) and \( b \). Applying this identity to \( \alpha^T \text{Sel}_k V_i \) yields Equation (54).

We then use Equation (54) to simplify Equation (53). Some linear algebra manipulations finally yield the main theoretical
result of this manuscript:

$$\epsilon \Delta Z_k(x) = c \alpha^T \Delta Z_k^\delta(x) = f^T \Delta Z_k^\delta(x)$$

$$\Delta Z_k^\delta(x) = \frac{\text{Sel}_k}{\sqrt{\Lambda_k}} \left[ -1 \sqrt{\Delta} V_k A_{\delta}(x) S(x)^T V_k Z_k(x) ight. \\
+ V_k A_{\delta}(x) \sum_{j=1, j \neq k}^{N_k} \frac{\Lambda_j}{\Lambda_k} (V_k A_{\delta}(x) S(x)^T V_j) \\
+ V_j A_{\delta}(x) S(x)^T V_k Z_k(x)) \right]$$

(55)

While Equation (55) can seem daunting at first, we emphasize that for a given unperturbed basis set $Z_k(x)$ and a given model of the image of the astrophysical source, then $\Delta Z_k^{(\delta)}(x)$ only needs to be pre-computed once and stored. Then the propagation of hypothetical sources of any arbitrary spectra $f$ through the least-squares speckle subtraction algorithm can be evaluated using a simple matrix multiplication. This has considerable advantages for statistical inference in the case of non-detections and for the estimation of astrophysical observables when faint substellar companions are detected.

6.6. Generalization to Perturbed LOCI Coefficients

While this manuscript used the formalism laid out in Soummer et al. (2012), we can also use the invariance between Equation (18) and Equation (29) in Savransky (2015) under the assumption of full rank correlation matrices, to apply our formalism to the case of LOCI. In the absence of astrophysical signal in the PSF library, the processed image using LOCI can be written as:

$$\mathcal{P}_{LOCI}(I(x)) = T(x)^T - S(x)^T \epsilon_{LOCI}$$

where the LOCI coefficients are given by Equation (18) and Equation (29) in Savransky (2015). Note that here we have transposed all quantities in Savransky (2015) to follow the array conventions in Soummer et al. (2012).

$$\epsilon_{LOCI} = V^T \Lambda^{-1} V S^T (x)^T$$

(57)

We use the following notations for the perturbed problem:

1. $V$, $\delta V$ are, respectively, the matrices whose column entries are $V_k$, $\delta V_k$.
2. $S, A_{\delta}$ are as defined in the body of the manuscript.
3. $\Lambda^{-1}, \delta \Lambda^{-1}$ are, respectively, the diagonal matrices whose diagonal entries are $1/\Lambda_k, 1/\delta \Lambda_k$. When using an eigenvalue truncation for LOCI, and only keeping the first $K_{\text{Klip}}$ modes, then the last $K_{\text{Klip}}$ diagonal terms of both matrices are set to zero.

In the presence of an astrophysical signal the reduced image becomes:

$$\mathcal{P}_{LOCI}(I(x))^T = T(x)^T - S(x)^T \epsilon_{LOCI} \\
- \alpha A_{\delta}(x)^T \epsilon_{LOCI} - \epsilon S(x)^T \delta \epsilon_{LOCI}$$

(58)

and we can use the formalism of the present manuscript to write $\delta \epsilon_{LOCI}$

$$\delta \epsilon_{LOCI} = (V^T \Lambda^{-1} \delta V + \delta V^T \Lambda^{-1} V + \delta V^T \\
\times \Lambda^{-1} V S^T \Lambda^{-1} V A_{\delta}^T - V^T \delta \Lambda^{-1} V S^T) T(x)^T$$

(59)

where $\delta V$ and $\delta \Lambda^{-1}$ can be calculated using Equation (50). The formalism developed herein is also applicable to LOCI implementation of least-squares speckle fitting algorithms.

APPENDIX F

KLIP-FM WHEN AN ASTROPHYSICAL SIGNAL IS PRESENT IN THE REFERENCE LIBRARY: CASE OF IFS OBSERVATIONS

Here we describe the mathematical formalism that takes advantage of Equation (55) to generalize the RDI Forward Modeling concepts discussed in Appendix D to the case of IFS spectroscopy of faint point sources with ADI+SSD. We assume that the location of the point source is known and seek to estimate its spectrum. We leave the astrometric estimation to further investigation.

F.1. Forward Modeling Cost Function in The Presence of Astrophysical Signal in The Reference Library

We start with Equation (16). Because of the presence of an astrophysical signal in the PSF library, the data analysis operator $\mathcal{L}S_{\mathcal{Q}_{(A,A)}}$ cannot be simplified as straightforwardly as in the case of RDI. Moreover, in the most general case the dependence on $A$ and $\tilde{A}$ is nonlinear. Fortunately in the case of small perturbations (both for the actual signal and for its synthetic negative counterpart) the linear approximation in Equation (55) holds and the data analysis operator can be simplified as follows:

$$\mathcal{L}S_{\mathcal{Q}_{(A,A)}}[I(x)] = I(x) - \sum_{k=1}^{K_{\text{Klip}}} \langle I(x), Z_k(x) \rangle \\
+ f^T \Delta Z_k^{\delta}(x) S(Z_k(x) + f^T \Delta Z_k^{\delta}(x))$$

(60)

$$\mathcal{L}S_{\mathcal{Q}_{(A,A)}}[I(x)] = I(x) - \sum_{k=1}^{K_{\text{Klip}}} \langle I(x), Y_k(x) \rangle_S Y_k(x)$$

(61)

$$\mathcal{L}S_{\mathcal{Q}_{(A,A)}}[I(x)] = I(x) - \sum_{k=1}^{K_{\text{Klip}}} \langle I(x), Y_k(x) \rangle \\
- f^T \Delta Y_k^{\delta}(x) S(Y_k(x) - f^T \Delta Y_k^{\delta}(x))$$

(62)

Of course in reality $f$ and $\Delta Z_k^{\delta}(x)$ are unknown. We can apply the same treatment derived for the actual astrophysical signal in Equation (55) to the negative synthetic source. Plugging Equation (62) into Equation (16) yields the single wavelength Forward Modeling cost function at $\lambda_0$:

$$\|T_{\lambda_0}(x) - \tilde{f}_{\lambda_0} A_{\lambda_0}(x) - \sum_{k=1}^{K_{\text{Klip}}} \langle I_{\lambda_0}(x), \tilde{f}_{\lambda_0} A_{\lambda_0}(x) \rangle \|_2^2$$

$$\|Y_{\lambda_0}(x) - \tilde{f} \Delta Y_{\lambda_0}^{\lambda_0}(x) \|_2^2$$

(63)

Adding this cost function for all wavelengths finally yields the spectral extraction cost function over the full range of wavelengths.
Note that in the most rigorous case, each wavelength should be weighted by its noise and a possible correlation between the wavelengths ought to be included. This is critical when deriving a confidence interval. However, the applications presented in this paper are only focused on biases associated with the most likely estimated spectrum. We leave further sophistications related to the calculations of confidence intervals to a future communication (J. Wang et al. 2016, in preparation). In Section 3, we show that even using this simple approach yields unbiased estimated spectra (albeit without confidence intervals). Here we emphasize that the Principal Components considered are the sum of two terms:

1. the principal components of the data itself, $Y^{l_h}_k(x)$, which contain (or do not contain) perturbations due to the presence of a hypothetical point source. Unfortunately the contribution of these perturbations cannot be evaluated a priori.
2. a term corresponding to the perturbation of $Y^{l_h}_k(x)$ due to the injection of the synthetic injected negative point source $-\hat{f} \Delta Y_k(x)$. This synthetic source is used for Forward Modeling purposes. Even if this term corresponds to the injection of a non-physical source for the purpose of astrophysical inference, its impact can still be quantified using the result in Equation (55).

### F.1.1. Linearized Problem

We now simplify Equation (64) to reduce the problem to the functional form described in Appendix D (Equation (24), which is more amenable to astrophysical inference). We use a fitting region $F$ whose size is independent of the size of the search area $S$. Assuming that the point source has been detected but its position is only known at the $\sim 1-2$ pixels level, this fitting region can be an aperture of the size of theFWHM of the PSF centered around this rough position. This is what we use in practice in the examples in this paper. We introduce the following notations:

$$P_{\lambda_x}(x) = T_{\lambda_x}(x) - \sum_{k=1}^{K_{\text{lin}}} \langle T_{\lambda_x}(x), Y^{l_h}_k(x) \rangle Y^{l_h}_k(x)$$

$$F_{\lambda_x}(x) = \hat{f}_k \left( A_{\lambda_x}(x) - \sum_{k=1}^{K_{\text{lin}}} \langle A_{\lambda_x}(x), Y^{l_h}_k(x) \rangle Y^{l_h}_k(x) \right)$$

$$= \hat{f}_k \sum_{k=1}^{K_{\text{lin}}} \left( \langle T_{\lambda_x}(x), Y^{l_h}_k(x) \rangle \Delta Y^{l_h}_k(x) \right)$$

$$\left. \right. + \langle T_{\lambda_x}(x), \Delta Y^{l_h}_k(x) \rangle Y^{l_h}_k(x).$$

This latter term contains all the contributions of the synthetic negative point source that are linear in $\hat{f}$ (e.g., neglecting the terms in $O(\hat{f}^2)$ in the Forward Modeling cost function). This expression for $F_{\lambda_x}(x)$ captures both over-subtraction and self-subtraction. Here again, $F_{\lambda_x}(x)$ can be written as a simple matrix multiplication:

$$F_{\lambda_x}(x) = \hat{f} F_{\lambda_x}(x)$$

with $F_{\lambda_x}$ being a $N_\lambda \times N_{\text{pix}}$ matrix whose $q$th line entry is defined by:

$$F_{\lambda_x}[q](x) = \delta_{p,q} \left( A_{\lambda_x}(x) - \sum_{k=1}^{K_{\text{lin}}} \langle A_{\lambda_x}(x), Y^{l_h}_k(x) \rangle Y^{l_h}_k(x) \right)$$

$$- \sum_{k=1}^{K_{\text{lin}}} \langle T_{\lambda_x}(x), Y^{l_h}_k(x) \rangle \Delta Y^{l_h}_k[q](x)$$

$$+ \langle T_{\lambda_x}(x), \Delta Y^{l_h}_k[q](x) \rangle Y^{l_h}_k(x),$$

where $\Delta Y^{l_h}_k[q](x)$ is the $q$th line entry in $\Delta Y^{l_h}_k(x)$. This yields the following simplified form for the Forward Modeling cost function:

$$\hat{f} = \arg \min_{\hat{f}} \sum_{p=1}^{N_\lambda} \| P_{\lambda_x}(x) - \hat{f} F_{\lambda_x}(x) \|^2_{\hat{f}}.$$  

#### F.1.2. Spectral Extraction Algorithm

Equation (67) can seem daunting at first but it can be easily implemented following the steps below:

1. Once a faint point source has been identified, carry out a KLIP reduction using geometric parameters that keep the point source at the center of the subtraction $S$ zone. If combining ADI and SSDI, derotate each exposure and sum over time to obtain:

$$P_{\lambda_x}(x) = \sum_i P_{\lambda_x,i}(R_{-\theta_i}[x]).$$

2. Assuming that the location of the point source and the off-axis PSF of the instrument are known, build a model of the motion of the point source across wavelengths and exposures $A_{\lambda_x}(R_{-\theta_i} x_{\lambda_x}).$

3. For each $Z_k(x)$, $V_k$, $\lambda_x$, $S_{\text{el}}$ associated with the reduction of each exposure at each wavelength, use this model in conjunction with Equations (55) and (66) to calculate $\Delta Y^{l_h}_k(x)$ and $F_{\lambda_x}(x)$. When using ADI, derotation and summation over time are necessary:

$$F_{\lambda_x}(x) = \sum_i F_{\lambda_x,i}(R_{-\theta_i}[x]).$$

These three steps only consist of basic linear algebra operations (associated with some image rotations over a small subregion $F$ of the image, usually ranging from a PSF FWHM to the entire $S$ zone). Once they have been carried out, the spectrum of the point source can be retrieved using any quadratic optimization algorithm to minimize Equation (67). Here we only consider the most simple route and recognize that Equation (67) is similar to Equation (25) (up to a summation) and thus the estimated spectrum can be obtained by solving the
following inverse problem:

$$\sum_{p=1}^{N_s} F_{s_p}(x) F_{s_p}(x)^T \frac{f}{\sum_{p=1}^{N_s} F_{s_p}(x) P_{s_p}(x)^T} = \sum_{p=1}^{N_s} F_{s_p}(x) P_{s_p}(x)^T. \quad (70)$$

### F.1.3. Detection Limits in IFS Data

Here we show that the algorithm discussed above can also be used to obtain detection limits that vary as a function of the hypothetical nature of the point sources that have not been detected in IFS data. Quantifying completeness as a function of the hypothetical nature of the point sources that have not been used to obtain detection limits that vary as a function of the astrophysical scenario. It does involves a computational overhead of $N_s$ extra image rotations when compared to the analysis of an entire data set simply using KLIP. However, this overhead can be mitigated by only carrying out these rotations locally in regions surrounding the injected planets. Moreover, most implementations of PCA-based methods calculate reduced images for a range of $K_{\text{Klip}}$ all the way to $N_{\text{Corr}}$. Because the linear model in Equation (71) is valid for all $K_{\text{Klip}}$ in the case of synthetic planets at the detection limit, it is sufficient to apply only the procedure described above for one value of $K_{\text{Klip}}$.

### REFERENCES

- Amara, A., & Quanz, S. P. 2012, MNRA, 427, 948
- Beuzit, J.-L., Feldt, M., Dohlen, K., et al. 2008, Proc. SPIE, 7014, 701418
- Biller, B. A., Close, L., Lenzen, R., et al. 2004, Proc. SPIE, 5409, 389
- Bonnefoy, M., Marleau, G.-D., Galicher, R., et al. 2014, A&A, 567, L9
- Brandt, T. D., McElwain, M. W., Turner, E. L., et al. 2013, ApJ, 764, 183
- Brandt, T. D.; McElwain, M. W., Turner, E. L., et al. 2014, ApJ, 794, 159
- Cantalloube, F., Mouillet, D., Mugnier, L. M., et al. 2015, A&A, 582, A89
- Cauchi, L., Barret, H. H., Devanye, N., & Rodríguez, J. J. 2007, JOSAA, 24, 13
- Chilcote, J., Barman, T., Fitzgerald, M. P., et al. 2015, ApJL, 798, L3
- Choquet, É., Pueyo, L., Hagan, J. B., et al. 2014, Proc. SPIE, 9143, 57
- Cornia, A., Mugnier, L. M., Carbillet, M., et al. 2010, Proc. SPIE, 7736, 77361
- Crepp, J. R., Rice, E. L., Veicht, A., et al. 2015, ApJL, 798, L63
- Carrie, T., Fukagawa, M., Thalmann, C., Matsunara, S., & Plavchan, P. 2012a, ApJL, 755, L34
- Fergus, R., Hogg, D. W., Oppenheimer, B., Brenner, D., & Pueyo, L., 2014, ApJ, 794, 161
- Gomez Gonzalez, C. A., Abisil, O., Abisil, P.-A., et al. 2016, arXiv:1602.08381
- Greco, J., & Brandt, T. 2016, ApJ, 815, 40 (tAxy:1602.00591)
- Hinkley, S., Oppenheimer, B. R., Brenner, D., et al. 2008, Proc. SPIE, 7015, 701519
- Hinkley, S., Oppenheimer, B. R., Zimmerman, N., et al. 2011, PASP, 123, 74
- Hinkley, S., Pueyo, L., Faherty, J. K., et al. 2013, ApJ, 779, 153
- Ingraham, P., Marley, M. S., Sauvage, D., et al. 2014, ApJL, 794, L15
- Kasdin, N. J., & Braems, I. 2006, ApJ, 646, 1260
- Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., & Artigau, É. 2007, ApJ, 660, 770
- Lawson, P. R., Poyneer, L., Barnett, H., et al. 2012, Proc. SPIE, 8847, 22
- Line, M. R., Teske, J., Burningham, B., Fortney, J. J., & Marley, M. S., 2015, ApJ, 807, 183
- Macintosh, B., Graham, J. R., Ingraham, P., et al. 2014, PNAS, 111, 12661
- Marois, C., Correia, C., Galicher, R., et al. 2014a, Proc. SPIE, 9148, 0
- Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., & Nadeau, D. 2006, ApJ, 641, 556
- Marois, C., Lafrenière, D., Macintosh, B., & Doyon, R. 2008a, ApJ, 673, 647
- Marois, C., Macintosh, B., & Vérán, J. 2010a, Proc. SPIE, 7736, 77361J
- Marois, C., Zuckerman, B., Konopacky, Q. M., Macintosh, B., & Barman, T. 2010b, Natur, 468, 1080
- Mawet, D., Milli, J., Wahhaj, Z., et al. 2014, ApJ, 792, 97
- Mazoyer, J., Boccaletti, A., Augereau, J.-C., et al., 2014, A&A, 569, A29
- Morzinski, K. M., Males, J. R., Skemer, A. J., et al. 2015, ApJ, 815, 108
- Mugnier, L. M., Cornia, A., Sauvage, J.-F., et al. 2009, JOSAA, 26, 1326
- Nielsen, E. L., Liu, M. C., Wahhaj, Z., et al. 2017, ApJ, 776, 43
- Oppenheimer, B. R., Baranec, C., Beichman, C., et al. 2013, ApJ, 768, 24
- Pueyo, L., Crepp, J. R., Vasisht, G., et al. 2012, ApJS, 199, 6
- Pueyo, L., Soummer, R., Hofmann, J., et al. 2015, ApJ, 803, 31
- Soummer, R., Ferrari, A., Aime, C., & Jolissaint, L. 2007a, ApJ, 669, 642
- Soummer, R., Perrin, M. D., Pueyo, L., et al. 2014, ApJL, 786, L23
- Soummer, R., Pueyo, L., & Larkin, J. 2012, ApJL, 755, L28
- Soummer, R., Pueyo, L., Sivaramakrishnan, A., & Vanderbei, R. J. 2007b, OEExpr, 15, 15935
- Sparks, W. B., & Ford, H. C. 2002, ApJ, 578, 543
- Vigan, A., Patience, J., Marois, C., et al. 2012, A&A, 544, A9
- Wahhaj, Z., Cieza, L. A., Mawet, D., et al. 2015, A&A, 581, 24
- Wang, J., Ruffio, J.-B., De Rosa, R. J., et al. 2015, pyKLIPl::PSF: Subtraction for Exoplanets and Disks, Astrophysics Source Code Library, ascl:1506.001
- Ygouf, M., Mugnier, L. M., Mouillet, D., Fusco, T., & Beuzit, J.-L. 2013, A&A, 551, A138