DETECTION FROM A MULTI-CHANNEL SENSOR USING A HIERARCHICAL BAYESIAN MODEL

I. Smith and A. Ferrari

UMR 6525 H. Fizeau, Université de Nice-Sophia Antipolis, CNRS, Observatoire de la Côte d'Azur
Campus Valrose, F-06108 Nice cedex, France
isabelle.smith@unice.fr, ferrari@unice.fr

ABSTRACT

Direct imaging of exoplanets involves very low signal-to-noise ratio data, that need to be carefully acquired and processed. Multi-band devices enable the simultaneous record of images in different spectral bands. They can be used either for spectroscopy purposes or to improve detection capabilities.

This work aims at detecting a potential source, when the source moves on a random background spatially and inter-spectrally correlated. A hierarchic Bayesian model is derived to take into account correlations and their randomness, and the high dynamic range involved in potentially low signal to noise ratio data.

The point null hypothesis test is addressed using the posterior distribution of the likelihood ratio. Its percentiles are computed using a simple Markov Chain Monte Carlo method. This algorithm is illustrated using 1D simulated data of a dual band signal.

***Index Terms*—** Bayes procedure, Signal detection, Object detection, Speckle, Estimation, Astronomy

1. INTRODUCTION AND SUMMARY

Detection and spectrum measurement of exoplanets using ground-based direct imaging are current driving motivations in astronomy. These motivations are materialized by future planet finders such as the Very Large Telescope (VLT) instrument SPHERE (Spectro-Polarimetric High-Contrast Exoplanet REsearch, [1]). The expected images are characterized by an intense background -due to the orbited star- highly contrasted both temporally and spatially, and a potentially very low source intensity -the exoplanet of interest. The background, caused by the orbited star, arises from the combination of residual speckles from the turbulent atmosphere and uncorrected by the adaptive optics system, and static or quasi-static ones from the optical system aberrations. These aberrations being similar for two close wavelengths, a solution to improve estimation/detection performances is to acquire images simultaneously in different spectral bands. The VLT SPHERE project will include this facility. The source could be dectected by processing the simple difference of properly rescaled images, as proposed in [2, 3], but differential aberrations remain an important issue.

The purpose of this communication is to propose a detector based on a Bayesian approach which relies on the extension of the model used in [4] in order to take into account:

1. a spatial and inter-spectral correlation of the optical aberrations (background). See for example [5] and included references for a discussion on static and quasi-static speckles.
2. the high range of intensities of the exoplanet (source) that can be expected in the different bands, and their possible statistical dependence.

General prior distributions related to multivariate normal distributions are defined for the unknown signal parameters and hyperparameters. Such a hierarchical model has physical motivations and can be expected to make posterior inferences robust with respect to the choice of the required hyperparameters values [6]. The nuisance parameters are analytically marginalized out to obtain an analytical expression of the joint distribution of the intensities in the different channels.

In addition to the posterior distribution of the unknown source intensities, a detection test is applied. It consists in the simple (null intensity of the source in all channels) versus composite hypothesis test. We choose the Bayesian detection test proposed in [7, 8] where the authors advocate the use of the posterior distribution of the likelihood ratio rather than the common Bayes factor [9]. In order to compute some point estimates of the intensities and the percentile of the posterior of the likelihood ratio, a Markov Chain Monte Carlo (MCMC) method is used.

2. HIERARCHICAL BAYESIAN MODEL

2.1. Likelihood

The dataset is represented by N successive sets of L images, where each image is represented by a $M \times 1$ vector $\boldsymbol{i}_\ell(k)$, $k = 1 \dots N$ and $\ell = 1 \dots L$. M is the number of pixels in the image, L the number of spectral channels and N the number of time exposures.

A precise statistical model for $\boldsymbol{i}_\ell(k)$ was studied in [4] in the case of a single band. It involves a non-uniform background, a general speckle model (that relies on a correlated Gaussian modeling of the complex amplitude arising from the Central Limit Theorem) and Poisson process. It was shown to lead under a high flux assumption and after renormalization to a simple Gaussian model.

In order to increase the detection performance, it is important to account at a given time k both for spatial and inter-channel correlations. The spatial correlation arises from various phenomena like low and high orders quasi-static aberrations, and the interspectral correlation arises when two channels are acquired in wave-bands close enough for the two backgrounds to be statistically related.

To account for both types of correlations without assuming that the backgrounds in two channels are too simply related, we define the $LM \times 1$ vector $\boldsymbol{x}_k^t = (\boldsymbol{i}_1(k)^t, \dots, \boldsymbol{i}_L(k)^t)$. A direct extension of the one channel model to the multi-channel case consists in assuming that, conditioned on some constant parameters, the images are independent and described by:

$$\boldsymbol{x}_k|\boldsymbol{\mu}, \Sigma, \boldsymbol{\eta}, H_k \sim \mathcal{N}(H_k\boldsymbol{\eta} + \boldsymbol{\mu}, \Sigma) \tag{1}$$

with:

- $\boldsymbol{\mu}^t = (\boldsymbol{\mu}_1^t, \ldots, \boldsymbol{\mu}_L^t)$ where $\boldsymbol{\mu}_\ell$ is the $M \times 1$ vector denoting the *unknown* background in the channel ℓ. In our case, it arises from the coronagraph, the static speckles, *etc*.
- $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_L)^t$ where η_ℓ are the *unknown* intensities of the source in the different channels. Those are the quantities of interest.
- H_k is the $LM \times L$ matrix

$$H_k = \begin{pmatrix} \boldsymbol{p}_1(k) & \mathbf{0}_{M\times 1} & \ldots & \mathbf{0}_{M\times 1} \\ \mathbf{0}_{M\times 1} & \boldsymbol{p}_2(k) & \ldots & \mathbf{0}_{M\times 1} \\ \mathbf{0}_{M\times 1} & \mathbf{0}_{M\times 1} & \ldots & \boldsymbol{p}_L(k) \end{pmatrix}$$

  where $\boldsymbol{p}_\ell(k)$ is the $M \times 1$ vector representing the *known* instrumental response of the source at time k in the channel ℓ.
- Σ is the *unknown* covariance matrix of size $LM \times LM$ accounting for spatial and inter-spectral correlations. For example, in the dual band case where $L = 2$, the main diagonal band of Σ (*ie* the terms indexed by $\{(i,i), (i, i\pm 1), (i, i\pm 2), ...\}$) account for the spatial correlation, and the middle off-diagonal (with indices $\{(i, i\pm M), (i, i\pm(M+1)), ...\}$) account for the inter-spectral correlations.

Note that when the source is moving with respect to the background, *ie* when there exists $\boldsymbol{p}_\ell(k) \neq \boldsymbol{p}_\ell(k')$, the $\boldsymbol{x}_k$ are not identically distributed. We denote the dataset as $\boldsymbol{x}^t = (\boldsymbol{x}_1^t, \ldots, \boldsymbol{x}_N^t)$.

2.2. Priors for the source intensities

Due to the high dynamics involved in possible low signal to noise ratio cases, and due to the inherent wide range of sources intensities [10], the distribution of $\boldsymbol{\eta}$ has to spread out on several order of magnitudes. This requirement appears in astronomy through the common use of a "magnitude" of objects, defined as mag $= -2.5\log(F/F_0)$ where F is the electromagnetic flux of the object, and F_0 is the flux of a reference object. Error bars on the intensities infered are also commonly given in the magnitude scale [11].

It is therefore natural to assume that $(\ln\eta_1, \ldots, \ln\eta_L)^t$ is jointly gaussian. This so-called multivariate lognormal distribution [12] describes high dynamics signals, has a positive support, is proper, and is described with few parameters that are furthermore directly related to the moments of the magnitudes. It takes the general form

$$p(\boldsymbol{\eta}|\boldsymbol{m}, B) = \frac{1}{|B|^{\frac{1}{2}}} \prod_{\ell=1}^{L} \frac{H(\eta_\ell)}{\sqrt{2\pi}\eta_\ell} \exp\left(-\frac{(\ln\boldsymbol{\eta} - \boldsymbol{m})^t B^{-1} (\ln\boldsymbol{\eta} - \boldsymbol{m})}{2}\right) \tag{2}$$

where $m = \mathsf{E}[\ln\boldsymbol{\eta}] = (\mathsf{E}[\ln\eta_1], \ldots, \mathsf{E}[\ln\eta_L])^t$, $B = \mathsf{cov}[\ln\eta]$ and $H(.)$ is the $\{0,1\}$ valued Heaviside function.

2.3. Hyperparameters priors

The nuisance parameters of the likelihood and the prior are generally unknown or subject to uncertainties. This issue can be adressed by assuming they are random, described by some underlying distribution and finally marginalized out, leading to more robust inferences [6]. The integration is in general numerically too time-consuming, but can be performed analytically using conjugate distributions.

Both the likelihood (1) and the prior (2) are related to a Normal distribution, the first with a time varying component and the second through a log-function. The Normal - inverse Wishart distribution, conjugate for the mean and covariance of a Normal distribution [13], is still conjugate for the nuisance parameters involved in both cases.

2.3.1. Background

We consider a Normal - inverse Wishart distribution for $(\boldsymbol{\mu}, \Sigma)$:

$$\boldsymbol{\mu}|\Sigma, \boldsymbol{\mu}_0, \lambda_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \lambda_0\Sigma) \tag{3}$$

$$\Sigma|\Sigma_0, \nu \sim \mathcal{W}_{LM}^{-1}(\Sigma_0, \nu) \tag{4}$$

where Σ_0 is assumed to be definite positive, and satisfies $\mathsf{E}[\Sigma] = (\nu - LM - 1)^{-1}\Sigma_0$.

The present case differs from the standard one [13] because of the motion of the source: although deterministic, $H_k\boldsymbol{\eta}$ that contributes to the mean value is time-varying and involves the variable of interest. Eqs. (3,4) imply:

$$\begin{aligned} p(\boldsymbol{\mu}, \Sigma|\boldsymbol{\mu}_0, \lambda_0, \nu) &= (2\pi)^{-LM/2} |\lambda_0\Sigma|^{-1/2} \\ &\times \mathrm{etr}\left(-\frac{1}{2\lambda_0}\Sigma^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^t\right) \\ &\times \frac{|\Sigma_0|^{\nu/2}}{2^{\frac{\nu+LM}{2}}\Gamma_{LM}(\nu/2)} |\Sigma|^{-\frac{\nu+LM+1}{2}} \mathrm{etr}\left(-\frac{1}{2}\Sigma^{-1}\Sigma_0\right) \end{aligned} \tag{5}$$

where $\mathrm{etr}(M) = \exp(\mathrm{trace}(M))$. Combining this equation with the likelihood (1) we obtain:

$$\begin{aligned} p(\boldsymbol{x}, \boldsymbol{\mu}, \Sigma|\boldsymbol{\eta}, \boldsymbol{\mu}_0, \lambda_0, \nu) &= \frac{(2\pi)^{-\frac{LM(N+1)}{2}}}{2^{\frac{\nu LM}{2}} \lambda_0^{\frac{LM}{2}} \Gamma_{LM}(\nu/2)} |\Sigma_0|^{\nu/2} \\ &\times |\Sigma|^{-\frac{\nu+N+LM+2}{2}} \mathrm{etr}\left(-\frac{1}{2}\Sigma^{-1}(\Sigma_0 + A(\boldsymbol{\mu}))\right) \end{aligned} \tag{6}$$

where $A(\boldsymbol{\mu})$ is defined by:

$$A(\boldsymbol{\mu}) = \sum_k (\boldsymbol{\mu} - \boldsymbol{x}_k + H_k\boldsymbol{\eta})(\boldsymbol{\mu} - \boldsymbol{x}_k + H_k\boldsymbol{\eta})^t + \frac{1}{\lambda_0}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^t$$

The matrix $A(\boldsymbol{\mu})$ can be rewritten as a "quadratic form" in $\boldsymbol{\mu}$:

$$A(\boldsymbol{\mu}) = (N + \lambda_0^{-1})(\boldsymbol{\mu} - \boldsymbol{h}(\boldsymbol{\eta}))(\boldsymbol{\mu} - \boldsymbol{h}(\boldsymbol{\eta}))^t + C_x(\boldsymbol{\eta}) \tag{7}$$

$$\text{with} \quad \boldsymbol{h}(\boldsymbol{\eta}) = \frac{\lambda_0}{N\lambda_0 + 1} \sum_k (\boldsymbol{x}_k - H_k\boldsymbol{\eta}) + \frac{1}{N\lambda_0 + 1}\boldsymbol{\mu}_0 \tag{8}$$

$$\begin{aligned} \text{and} \quad C_x(\boldsymbol{\eta}) &= \sum_k (\boldsymbol{x}_k - H_k\boldsymbol{\eta})(\boldsymbol{x}_k - H_k\boldsymbol{\eta})^t + \frac{1}{\lambda_0}\boldsymbol{\mu}_0\boldsymbol{\mu}_0^t \\ &- (N + \lambda_0^{-1})\boldsymbol{h}(\boldsymbol{\eta})\boldsymbol{h}(\boldsymbol{\eta})^t \end{aligned} \tag{9}$$

Substitution of (7) into (6) shows that the prior (5) is conjugate to (1). Integration with respect to $\boldsymbol{\mu}$ leads to:

$$\begin{aligned} p(\boldsymbol{x}, \Sigma|\boldsymbol{\eta}) &= \frac{(2\pi)^{-\frac{LM(N-1)}{2}}}{2^{\frac{\nu LM}{2}} (N\lambda_0 + 1)^{\frac{LM}{2}} \Gamma_{LM}(\nu/2)} |\Sigma_0|^{\nu/2} \\ &|\Sigma|^{-\frac{\nu+N+LM+1}{2}} \mathrm{etr}\left(-\frac{1}{2}\Sigma^{-1}(\Sigma_0 + C_x(\boldsymbol{\eta}))\right) \end{aligned}$$

and integration with respect to Σ finally gives:

$$p(\boldsymbol{x}|\boldsymbol{\eta}) = \frac{2^{\frac{LM}{2}}\pi^{-\frac{LM(N-1)}{2}}}{(N\lambda_0 + 1)^{\frac{LM}{2}}} \frac{\Gamma_{LM}(\frac{\nu+N}{2})}{\Gamma_{LM}(\frac{\nu}{2})} \frac{|\Sigma_0|^{\frac{\nu}{2}}}{|\Sigma_0 + C_x(\boldsymbol{\eta})|^{\frac{\nu+N}{2}}} \tag{10}$$

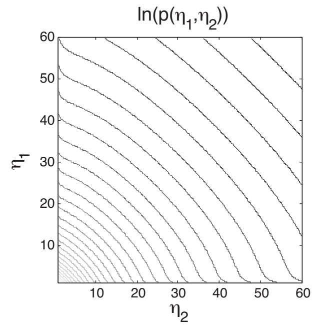

Fig. 1. Contour levels of the prior distribution, shown in a log view.

2.3.2. *Intensities*

The same way, we consider a Normal - inverse Wishart distribution for the hyperparameters (m, B) which is conjugate to the prior (2):

$$\boldsymbol{m}|\boldsymbol{m}_0, c_0, B \sim \mathcal{N}(\boldsymbol{m}_0, c_0 B) \tag{11}$$

$$B|B_0, \omega \sim \mathcal{W}_L^{-1}(B_0, \omega) \tag{12}$$

A similar calculation as for the likelihood is performed. The marginalized prior is finally given by:

$$p(\boldsymbol{\eta}) = \frac{(\pi(c_0+1))^{-\frac{L}{2}}}{\prod_{\ell=1}^{L} \eta_\ell} \frac{\Gamma_L\left(\frac{\omega+1}{2}\right)}{\Gamma_L\left(\frac{\omega}{2}\right)} \frac{|B_0|^{\frac{\omega}{2}} \prod_{\ell=1}^{L} H(\eta_\ell)}{\left|B_0 + \frac{(\ln\boldsymbol{\eta} - \boldsymbol{m}_0)(\ln\boldsymbol{\eta} - \boldsymbol{m}_0)^t}{c_0+1}\right|^{\frac{\omega+1}{2}}} \tag{13}$$

An example is given in Fig. 1 with $L = 2$, $\boldsymbol{m}_0 = (4, 4)^t$, $c_0 = 0.99$, $B_0 = (\omega - 3)\mathsf{E}[B]$ where we fix $\mathsf{E}[B_{1,1}] = \mathsf{E}[B_{2,2}] = \sigma^2$ and $\mathsf{E}[B_{1,2}] = \rho\sigma^2$ with $\sigma^2 = 10$ and $\rho = 0.8$. The choice of ω can be made upon the variance on the diagonal terms of B: $\mathsf{var}(B_{i,i}) = \frac{2(B_0)_{i,i}^2}{(\omega-3)^2(\omega-5)^2}$. We choose $\omega = 4$. The choice of this variables set was motivated by a rather conservative view with little correlation assumed on η_1 and η_2 and a quite large uncertainty about the different parameters. In practice, parameters can be chosen to fit the prior of physical models such as [10].

Despite the fact that the prior (13) tends to ∞ as $\boldsymbol{\eta} \to \boldsymbol{0}$, this prior can be shown to be proper using standard calculations on improper integrals. Note that the standard scale invariant distribution $\prod_{\ell=1}^{L} \eta_\ell^{-1} H(\eta_\ell)$ also tends to ∞ as $\boldsymbol{\eta} \to \boldsymbol{0}$ but is improper, which could be inconvenient for example if the Bayes factor was retained for the detection [6, 9].

The final posterior distribution is computed from (10,13):

$$p(\boldsymbol{\eta}|\boldsymbol{x}) \propto p(\boldsymbol{x}|\boldsymbol{\eta})p(\boldsymbol{\eta}) \tag{14}$$

3. DETECTION: POINT NULL HYPOTHESIS TEST

The detection test consists in the comparison of the simple hypothesis $\boldsymbol{\eta} = \boldsymbol{0}$ and the more general composite hypothesis, namely:

$$H_0 : \boldsymbol{\eta} = \boldsymbol{0} \qquad H_1 : \eta_\ell \geq 0,\ \forall \ell \tag{15}$$

A common approach consists in using the usual Bayes factor [9]. In our case the null hypothesis is supported by a set of measure 0. This case has been extensively discussed, both on theoretical grounds and case studies: see [7, 9, 14] and included references. In order not to stress the sensitivity of the detector with respect to the choice of the prior, we retain the approach used in [8]. It consists in rejecting H_0 if the probability that "the likelihood ratio (LR) of H_0 over H_1 is less than ζ" is greater than some p value, where the likelihood is given under both hypotheses by Eq. (10):

$$H_0 \text{ rejected if} \quad \Pr\{LR(\boldsymbol{\eta}) < \zeta|\boldsymbol{x}\} > p \tag{16}$$

$$\text{where } LR(\boldsymbol{\eta}) = \frac{p(\boldsymbol{x}|\boldsymbol{\eta} = \boldsymbol{0})}{p(\boldsymbol{x}|\boldsymbol{\eta})} \tag{17}$$

In practice, this can be done by sampling the posterior distribution (14) as $\boldsymbol{\eta}^{[k]} \sim p(\boldsymbol{\eta}|\boldsymbol{x})$ using a MCMC sampling method and computing the empirical posterior cumulative distribution of $LR(\boldsymbol{\eta}^{[k]})$ defined by $F_{LR|\boldsymbol{x}}(\zeta) = \Pr\{LR(\boldsymbol{\eta}) < \zeta|\boldsymbol{x}\}$.

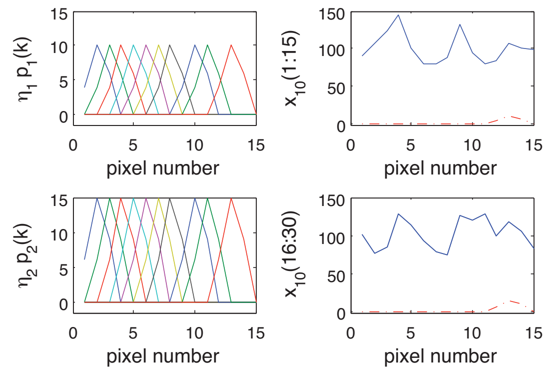

Fig. 2. Up: band $\ell = 1$. Down: band $\ell = 2$. Left: source profile $\eta_\ell \boldsymbol{p}_\ell(k)$ superimposed for all $k = 1, ..., 10$. Right, continuous line: data $\boldsymbol{x}_k$ for $k = 10$ (first half of the vector $\boldsymbol{x}_{10}$ for $\ell = 1$ and second half for $\ell = 2$). Right, dashed line: $\eta_\ell \boldsymbol{p}_\ell(10)$.

4. SIMULATIONS

The hierarchical model derived in section 2 is illustrated using a simple 1D dataset in the dual band case $L = 2$, simulated from the model (1) with parameters $M = 15$, $N = 10$, $\boldsymbol{\eta} = (20, 30)^t$.

- The source profile $\boldsymbol{p}_\ell(k)$ is identical for all ℓ and k, aside from a global motion with almost constant speed. The profile has a spatial support of 3 pixels equal to $(0.2, 0.5, 0.3)^t$,
- the background is uniform and identical for both bands, with $\mu(i) = 100\ \forall i = 1, ..., 2M$,
- Σ is Toeplitz with first column vector $\boldsymbol{v} = 500\boldsymbol{u}$ with $u(1) = 1$, $u(2) = 0.3$, $u(3) = 0.05$, $u(M+1) = 0.6$, $u(M+2) = 0.2$ and all other components are 0.

Fig. 2 illustrates a realization of such a dataset. To account for "imperfect" assumptions, the following likelihood hyperparameters are chosen for Eq. (10):

- $\mu_0(i) = 500\ \forall i = 1, ..., 2M$: the background is assumed to be uniform but with a wrong constant value,
- $\lambda_0 = 2$ and $\nu = 2M$: reasonable weight is given to the data,
- $\Sigma_0 = 4I$: no correlation is assumed, neither spatially nor between the two bands,

The consecutive likelihood function is shown on the left image of Fig. 3. Its product with the prior shown in a log view on Fig. 1 gives the posterior distribution shown on the right image of Fig. 3.

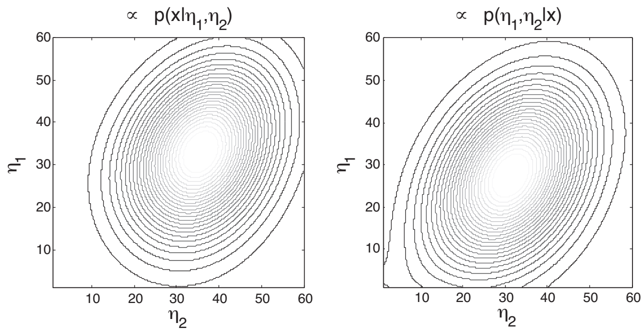

Fig. 3. Left: Contour levels of the likelihood (10). Right: corresponding posterior distribution under the prior of Fig. 1.

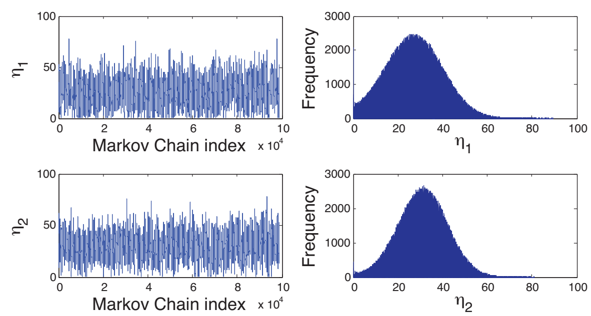

Fig. 4. Up: η_1 samples. Down: η_2 samples. Left: piece of the Markov chain. Right: corresponding histograms from the whole chain.

Due to the complexity of this posterior distribution $p(\eta_1, \eta_2|\boldsymbol{x})$ given in Eq. (14), appropriate simulation methods such as MCMC methods [15] are required for estimation and detection. Unfortunately, because of the presence of a determinant that cannot be simplified, the posterior cannot be re-expressed using conditional distributions, as necessary for the Gibbs sampler [15]. In our case, the slice sampling algorithm which has satisfactory convergence properties will be used. Fig. 4 shows a part of the Markov Chain, computed from the posterior distribution shown in the previous figure.

The chain can be first used to make inferences about the source intensities: the Maximum A Posteriori is $\hat{\boldsymbol{\eta}}_{MAP} = (27.5, 31.4)^t$, the Posterior Mean is $\hat{\boldsymbol{\eta}}_{PM} = (27.3, 31.1)^t$ (with true values $\boldsymbol{\eta} = (20, 30)^t$) and the Standard Deviation of the marginal posterior distribution of each intensity is $(12.2, 10.8)^t$, which is quite large but expected from the conservative hyperparameters chosen.

The chain is then used for the detection test (16). Fig. 5 shows the empirical posterior cumulative distribution of $LR(\boldsymbol{\eta})$. It shows that the probability that $p(\boldsymbol{x}|H_0) < 100\ p(\boldsymbol{x}|H_1, \boldsymbol{\eta})$ is slightly above 0.9. The H_0 hypothesis can be reasonably rejected.

This algorithm has been applied on several datasets generated using very broad ranges of parameters, and under hyperparameters more or less in adequation with the data. All the cases considered gave satisfactory results with no totally counterintuitive conclusion.

5. CONCLUSION

The hierarchical bayesian model presented in this communication opens large possibilities for detection and estimation purpose by

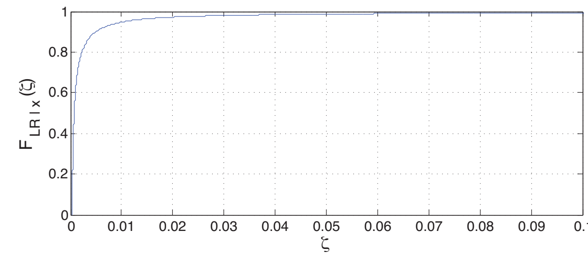

Fig. 5. Posterior probability that H_0 is rejected if $LR(\boldsymbol{\eta}) < \zeta$.

- taking into account some correlation in the data, without having to make too precise estimates neither of the covariance matrix nor of the background,
- modeling in a flexible form the prior of $\boldsymbol{\eta}$ using its statistical properties.

The next step consists in generalizing the likelihood by allowing a time correlation, in order to take into account the presence of quasi-static speckles. For continuous and symmetry issues, the use of independent and redundant blocks of data should be avoided. The treatment of this issue therefore requires a quite different approach.

6. REFERENCES

[1] J.-L. Beuzit, M. Feldt, K. Dohlen, et al., "SPHERE: A 'Planet Finder' Instrument for the VLT," *The Messenger*, vol. 125, pp. 29, 2006.

[2] L.M. Mugnier, J.-F. Sauvage, T. Fusco, and G. Rousset, "Multi-channel planet detection algorithm for angular differential imaging," in *Optical Society of America*, B. Ellerbroek, Ed., 2007, Adaptive Optics: Analysis and Methods.

[3] A. Boccaletti, D. Mouillet, et al., "Analysis of ground-based differential imager performance," in *Proceedings of the International Astronomical Union*. 2005, pp. 519–524, Cambridge University Press.

[4] I. Smith, A. Ferrari, and M. Carbillet, "Detection of a moving source in speckle noise. application to exoplanet detection," *IEEE Transactions on Signal Processing*, 2008, submitted.

[5] R. Soummer, A. Ferrari, C. Aime, and L. Jolissaint, "Speckle noise and dynamic range in coronagraphic images," *The Astrophysical Journal*, vol. 669, pp. 642–656, nov 2007.

[6] Christian P. Robert, *The Bayesian Choice*, Springer, February 2001.

[7] A. P. Dempster, "The direct use of likelihood for significance testing," *Statistics and Computing*, vol. 7, no. 4, pp. 247–252, 1997.

[8] Murray Aitkin, "The calibration of p-values, posterior bayes factors and the aic from the posterior distribution of the likelihood," *Statistics and Computing*, vol. 7, no. 4, pp. 253–261, 1997.

[9] R.E. Kass and A.E. Raftery, "Bayes factors," *Journal of the American Statistical Association*, vol. 90, no. 430, pp. 773–795, 1995.

[10] C. Moutou et al., "Science case for VLT-Planet Finder," in *IAU Colloq. 200: Direct Imaging of Exoplanets: Science and Techniques*, C. Aime and F. Vakili, Eds., 2006, pp. 159–164.

[11] F. Wesley, J.J. Kavelaars, et al., "The Kuiper belt luminosity function from m(R)=21 to 26," *Icarus*, vol. 195, no. 2, pp. 827–843, 2008.

[12] Samuel Kotz, N. Balakrishnan, and Norman L. Johnson, *Continuous Multivariate Distributions, Volume 1, Models and Applications, 2nd Edition*, Wiley-Interscience, 2000.

[13] J. M. Dickey, D. V. Lindley, and S. J. Press, "Bayesian estimation of the dispersion matrix of a multivariate normal distribution," *Communications in Statistics - Theory and Methods*, vol. 14, pp. 1019–1034, 1985.

[14] Murray Aitkin, Richard J. Boys, and Tom Chadwick, "Bayesian point null hypothesis testing via the posterior likelihood ratio," *Statistics and Computing*, vol. 15, no. 3, pp. 217–230, 2005.

[15] Christian P. Robert and George Casella, *Monte Carlo Statistical Methods*, Springer-Verlag, 1999.