# Making high accuracy null depth measurements for the LBTI Exozodi survey

Bertrand Mennesson*[a], Denis Defrère[b], Matthias Nowak[c], Philip Hinz[d], Rafael Millan-Gabet[e], Olivier Absil[b], Vanessa Bailey[d], Geoffrey Bryden[a], William Danchi[f], Grant M. Kennedy[g], Lindsay Marion[b] , Aki Roberge[f], Eugene Serabyn[a], Andy J. Skemer[d], Karl Stapelfeldt[a], Alycia J. Weinberger[h] and Mark Wyatt[g]

[a]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109; [b]Space sciences, Technologies, and Astrophysics Research (STAR) Institute, Université de Liège, 19c allée du Six Août, B-4000 Liège, Belgium; [c]LESIA/Observatoire de Paris, CNRS, UPMC, Université Paris Diderot, 5 place Jules Janssen, F-92195 Meudon, France; [d]Steward Observatory, Department of Astronomy, University of Arizona, 933 N Cherry Avenue, Tucson, AZ 85721, USA; [e] NASA Exoplanet Science Institute, California Institute of Technology, 770 South Wilson Avenue, Pasadena, CA 91125, USA; [f]NASA Goddard Space Flight Center, Exoplanets & Stellar Astrophysics Laboratory, Code 667, Greenbelt, MD 20771, USA; [g]Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK; [h]Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road NW, Washington, DC 20015, USA

**ABSTRACT**

The characterization of exozodiacal light emission is both important for the understanding of planetary systems evolution and for the preparation of future space missions aiming to characterize low mass planets in the habitable zone of nearby main sequence stars. The Large Binocular Telescope Interferometer (LBTI) exozodi survey aims at providing a ten-fold improvement over current state of the art, measuring dust emission levels down to a typical accuracy of ~12 zodis per star, for a representative ensemble of ~30+ high priority targets. Such measurements promise to yield a final accuracy of about 2 zodis on the median exozodi level of the targets sample. Reaching a 1σ measurement uncertainty of 12 zodis per star corresponds to measuring interferometric cancellation ("null") levels, i.e visibilities at the few 100 ppm uncertainty level. We discuss here the challenges posed by making such high accuracy mid-infrared visibility measurements from the ground and present the methodology we developed for achieving current best levels of 500 ppm or so. We also discuss current limitations and plans for enhanced exozodi observations over the next few years at LBTI.

**Keywords:** Exozodiacal light, nulling, mid-infrared, high contrast, interferometry, direct imaging, exoplanets
*Bertrand.Mennesson@jpl.nasa.gov; phone 1 818 354-0494

## 1. INTRODUCTION

Exozodiacal light is well known to be a double edge sword when it comes to the characterization of exoplanetary systems. On one hand, a large amount of dust emission in the habitable zone of main sequence stars is indicative of strong dynamical stirring that may be caused by planets in the system. On the other, a large amount of exozodiacal emission may hamper the detection of faint low mass exoplanets embedded in the dust. Indeed, taking the example of the solar system, the spatially integrated flux arising from zodiacal dust emission is significantly (~300x) larger than the flux of the Earth, both at visible and mid-infrared wavelengths. Observing a solar type star located 10 pc away in the visible, the exozodi flux captured in a spatial resolution element would still outshine that of a putative exo-Earth in the habitable zone of that star, at least for any telescope diameter smaller than ~10m[1]. Besides the loss of sensitivity due to increased background noise[2], a bright exozodiacal dust cloud also represents a strong challenge for the proper identification and

characterization of faint exoplanets embedded in such disks. Indeed, models of disk planet interactions predict the creation of ring structures by resonant trapping of in-spiraling dust, and the trapped grains are expected to orbit in lock with the perturbing planet[3,4,5]. These models also predict ring asymmetries, including a region of density deficit at the planet location (gap), and regions of density enhancement (clumps) trailing and leading the planet in its orbit. Such resonant clumps have actually been observed in the Solar system by the CODE/DIRBE instrument[6]. The trailing clump is very close to the Earth, and its integrated flux over the telescope PSF typically corresponds to a few percents of the Earth visible flux seen at quadrature. Exozodiacal dust clouds likely harbor similar structures. Simulations conducted in the case of high contrast direct imaging observations of an Earth analog embedded in exo-zodiacal clouds of different brightnesses[5] predict that at a level of ~10-20 zodis, these clumps could be brighter than an exo-Earth and hence constitute major sources of confusion and false positives. The exact location and strength of these exo-zodi clumps is expected to vary with planet mass, semi-major axis and outer dust characteristics[3] (e.g, density and typical grain size), and spectroscopic observations will likely allow to differentiate between bona-fide planets and bright dust clumps. However, we retain here that exo-zodi clouds at a level of ~ 10 zodis or more may generate bright enough structures to make data interpretation difficult, if not preclude the detection of exo-Earths for systems seen under high inclination. The requirement that was set from this coarse analysis, is that the LBTI be able to measure the median – i.e typical - exo-zodi level of nearby main sequence stars down to a 1σ uncertainty of ~2 zodis or less, a small fraction of the ~ 10 zodis confusion limit. **Assuming for instance that 36 stars are observed by LBTI with an *individual* measurement uncertainty of 12 zodis, the final uncertainty on the sample *median* would be of the order of that 2 zodis limit**. In comparison, the current best estimate for the median exozodi level of solar-type stars is 12 +/- 24 zodis, as derived from the Keck Interferometer observations of 20 solar-type stars between 2008 and 2011, at a typical accuracy level of 100 to 150 zodis per star[7]. This shows that a tenfold improvement over the state of the art is required of LBTI measurements to properly assess the risk of confusing exo-Earths in the habitable zone with bright exozodi clumps. We discuss hereafter some of the key challenges associated with measuring the exozodi level of individual stars down to a ~12 zodis accuracy level using ground based mid-infrared nulling interferometry.

## 2. MAIN CHALLENGES

The problem of detecting faint exozodiacal emission from the ground is notoriously difficult. The challenge resides in making high angular resolution high contrast detections, - already challenging per se -, in the presence of a considerably brighter thermal background.

### 2.1 Extreme fringe visibility accuracy

Looking at a solar system analog, exozodiacal dust concentrates in the inner few AUs. Integrating over the whole disk, the total exozodi flux only represents a tiny fraction of the stellar flux. Indeed, the dust over star flux ratio is only $4 \times 10^{-8}$ at V band, clearly too faint to be detected by current ground-based telescopes. This ratio increases significantly in the mid-infrared, reaching $5 \times 10^{-5}$ at 11 μm for instance. While the required contrast now appears more manageable, a >20m telescope would be required to reach sufficient angular resolution to survey a significant number of stars at mid-infrared wavelengths. A more practical solution is to use a nulling interferometer, as operated previously at the Keck Interferometer and now starting science operations at LBTI.

In order to reach a 12 zodis uncertainty limit on a Sun-like star, as described above, and assuming that 50% of the exozodiacal flux is transmitted through the interferometer null fringe pattern, we get that the required null measurement accuracy is $12 \times 0.5 \times 5 \times 10^{-5}$ = 300 ppm rms at 11 μm. Using the complex source visibility (V) as a more usual observable, and noting that the astronomical null depth is simply defined as $(1-|V|)/(1+|V|)$, this corresponds to measuring the stellar fringe visibility to a record accuracy of 600 ppm rms. This is basically unheard of at any wavelength for interferometric measurements between separate telescopes, let alone in the mid-infrared.

### 2.2 Extreme background measurement accuracy

Independently of the issue of measuring high accuracy visibilities, operating in the mid-infrared from the ground implies estmating and subtracting the thermal background to a very high accuracy. Even with its small number of warm surfaces per telescope and resulting small emissivity, the thermal background measured at LBTI still contributes a very large flux.

At 11 μm, the background flux detected in the core of the telescope PSF (the so-called "photometric region" of radius 0.51 λ/D used for null measurements at LBTI[8]) is equivalent to the flux detected through the system when observing a ~100Jy star. For a 1Jy target star, on the faint end of the survey sample[9], the 300ppm null accuracy requirement then translates into a 3ppm uncertainty on the background estimation. Such accuracy has been demonstrated on ground-based telescopes using fast (> 10 Hz) chopping and nodding techniques. Fast secondary mirror chopping would however be impractical in the context of nulling interferometry between separate apertures where fast fringe tracking loops are constantly operating to maintain the optical path difference within a small fraction of the observing wavelength. There as well, some different observing methods needed to be implemented to provide very accurate background estimation.

## 3. OBSERVING STRATEGY AND BACKGROUND ESTIMATION QUALITY

The nominal observing sequence was devised to provide the best null depth estimation given the different constraints of fast background estimation and accurate visibility calibration. For a given science target, the nominal sequence is ~3.5 hours long and consists of 7 "pointings" of about 30 minutes each (including overheads), alternating between science and calibrator observations (e.g. "cal1/sci/cal2/sci/cal1/sci/cal3"), as usually done in ground based interferometry. As shown in Figure 1, each individual ~30 minute long pointing is composed of many successive observing blocks (OBs), where each OB consists of 1000– 2000 detector frames, each having an integration time of typically 10– 100 ms, depending on the brightness of the star. The pointing sequence contains typically 10 OBs recorded at null, i.e., with the beams from both apertures coherently overlapped in phase opposition, followed by one OB of photometric measurements with the beams separated on the detector, and one OB of background measurements with the beams nodded off the detector.

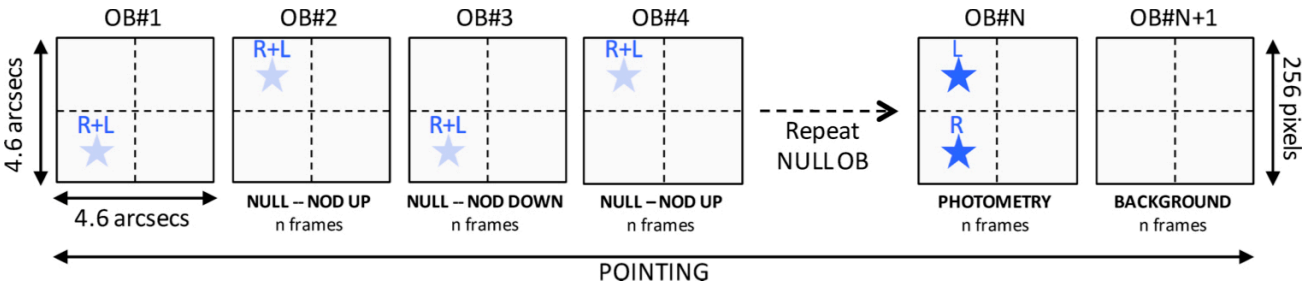


**Figure 1** (adapted from [8]): Typical observing sequence used for the LBTI-HOSTS exozodi survey. The sequence is divided in basic observing blocks (OB) consisting of one to two thousand frames. Each square represents a 256 x 256 subframe of the NOMIC detector that covers a region of approximately 4.6" x 4.6". The beams are aligned vertically in the middle of a given channel (see blue stars) to maximize the effective field of view and nodded back and forth by 2.3" (up-down to preserve the differential pathlength). The dashed lines represent the limits of different detector channels. See text for further details.

The OBs at null are acquired alternatively in two telescope nod positions separated by 2.3"on the detector. The nodding direction is set perpendicular to the interferometric baseline, allowing to re-acquire fringes very rapidly when alternating nod positions. Looking at an empty region of the sky, the background recorded in the photometric region drifts significantly over periods of seconds to minutes (example shown in Figure 2, top panel). This means that the pedestal signal that the star sits on is not only affected by white shot noise, but also exhibits low frequency drifts essentially due to atmospheric effects (e.g. variation of the line of sight water vapor absorption and emissivity as the telescope points further from zenith). Additionally, some instrumental effects are also present, such as beam walk on intervening optics with non-uniform emissivity, and significant slow drifts of the Aquarius detector bias that are already known to exist [10].

In order to correctly estimate and subtract this mid-IR background, which varies both spatially and temporally across the detector, a 2-step process is used (see illustrative sequence shown in Figure 2):

- The background in the photometric region is first estimated instantaneously by measuring the total signal detected in two semi-circular regions located on each side of the photometric region (Figure 3). The difference between this estimate and the actual background signal detected in the photometric region is shown in the middle panel of figure 2. The estimation error is much more stable over time than the raw background signal of the top panel, but it still exhibits a large offset that slowly drifts over time.
- A second stage of correction is then necessary. It uses nodding of the telescopes to periodically measure and reset the residual background offset between the photometric region and the background estimation region. The bottom panel of figure 2 shows the final measurement error that would be committed in the best case scenario, nodding at the detector frame rate of ~30Hz. In that case, the background estimation would essentially be perfect, just limited by the total background shot noise.

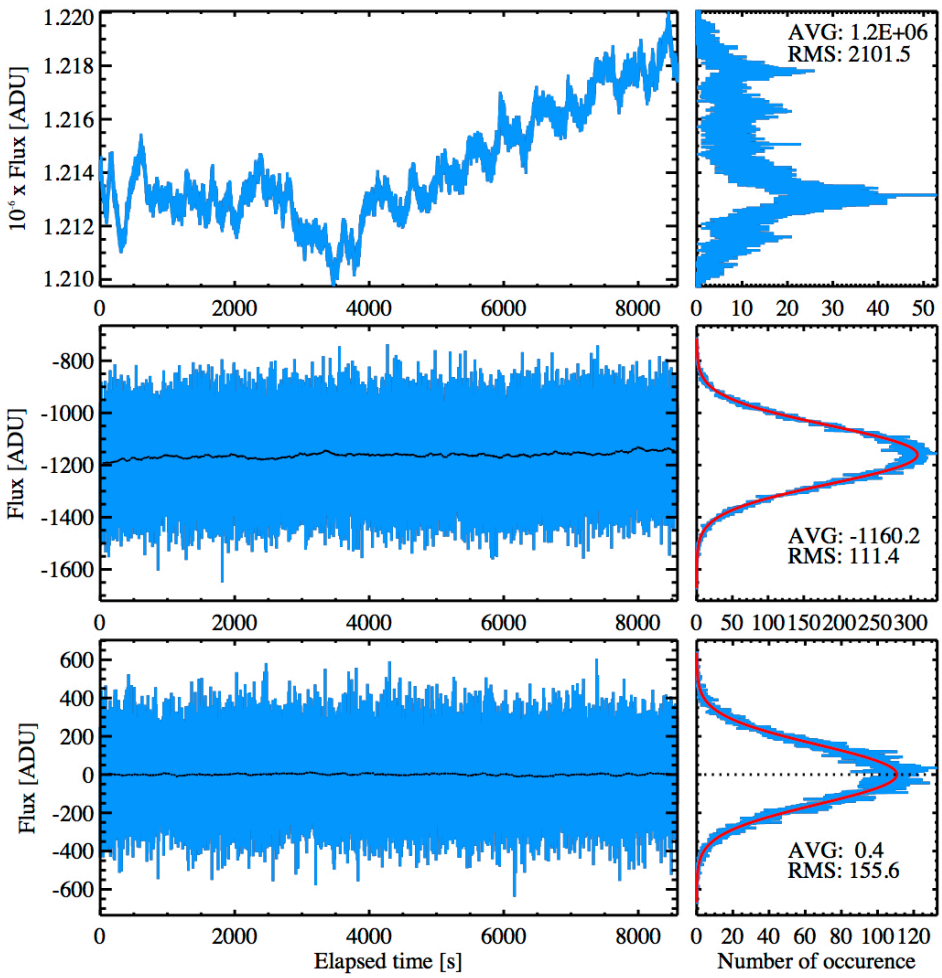


Figure 2 (adapted from[8]): Top: example of on-sky raw thermal background measurements obtained in the N′ band with the telescope pointing at an empty region of the sky and covering approximately 15 degrees of elevation change during the whole duration of the 2.4 hour long sequence. The left panel shows the flux integrated over a photometric aperture of 8 pixels in radius while the right panel shows the corresponding distribution. Middle: same measurements after subtraction of simultaneous background measurements (left). The corresponding distribution (right) is now Gaussian and shows a relatively large offset. The black line represents a running average of 100 s to better show the low-frequency drift due to slowly changing instrumental background between the 2 regions. Bottom: same measurements after subtraction of simultaneous background measurements and nod subtraction (left). For this example, nod subtraction has been performed at the maximum frequency (i.e., using adjacent frames). The corresponding distribution (right) is now Gaussian and centered on 0. These data have been obtained using an integration time of 28 ms (on 2014 May 14).

In practice, the telescopes are only nodded every ~60s (at then end of each OB), and current on-sky tests using 30 ms detector exposures show that the background measurement error is of the order of ~3 ADU rms per nod. Assuming 10 nods per individual 30 minute pointing, the final uncertainty on the background estimation is then ~1 ADU rms. Given the total background flux of ~2 $10^5$ ADU measured in the photometric region, this corresponds to a 5ppm rms background estimation error per pointing, or ~3ppm for 3 independent pointings. This corresponds precisely to the *total* error acceptable for null measurements at the 300ppm level on a 1Jy star. This also says that for all stars fainter than 2Jy, the current background estimation uncertainty represents at least 50% of the overall tolerable error.

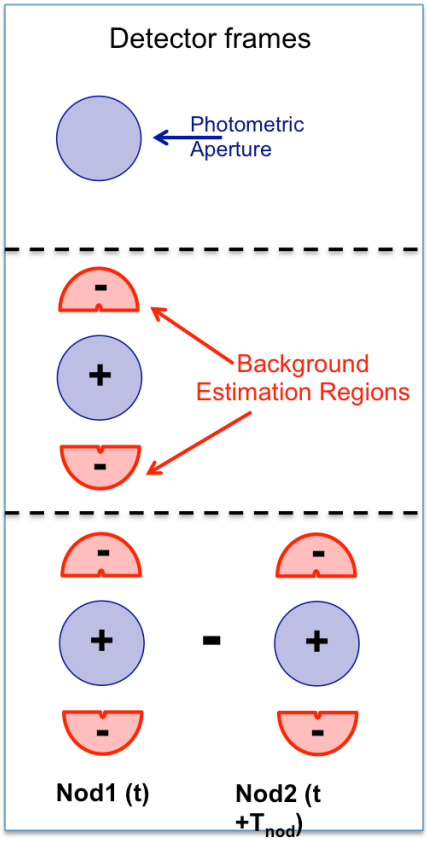


Figure 3: Principle of background estimation. The background signal recorded at time t in the photometric aperture (starlight PSF core, top panel) is estimated by measuring the signal recorded at the same time in a nearby "background estimation region" composed of two distinct half-circle apertures (middle panel). Finally, the remaining -time variable- background difference between the photometric and background estimation regions is measured and subtracted periodically (every $T_{nod}$ ~ 60s) by nodding the telescope, i.e moving the star off that region of the array (bottom panel).

## 4. ESTIMATING THE ASTRONOMICAL NULL DEPTH

### 4.1 "What is the null?"

As stated above, the basic observable is the astronomical null depth N, which can be written as N= (1-|V|) / (1+|V|), where V is the usual complex source visibility measured at the interferometric resolving baseline (fixed at 14.4m at the LBTI), and over some spectral bandpass. For all results discussed here the bandpass is the defined by LBTI's N' band which has an effective wavelength of 11.1 μm and a FWHM of 2.27 μm[8].

### 4.2 Basic principle of visibility (and null) self-calibration

What is measured however is not directly the astronomical null depth, but a times series of interferometric signals. For the LBTI nuller, these are recorded in the vicinity of the central dark fringe with a typical OPD fluctuation of 300 to 500 nm rms. Let's first use a simplistic monochromatic simulation, where both telescope beams have the same constant intensity $I_0$, and where signal variations are caused by OPD and background fluctuations. In other words, the observed interferometric signal is simulated as:

$$I(t) = 2I_0 * [1 - |V| * \cos(\Delta\phi(t))] + B(t) \qquad [1]$$

where the phasing error $\Delta\phi(t)$ represents the offset from the central dark fringe, and B(t) is the instantaneous background signal, after subtraction of its best instantaneous estimate, as explained in section 3. The top left panel of Figure 4 shows an illustrative time series obtained for a simulated dataset of $10^4$ points (i.e $10^4$ individual detector reads), where the signal of Equation (1) was divided by the expected constructive signal $2I_0*(1+V)$. The top right panel shows the histogram of this normalized signal occurrences. In this particular simulation, the source visibility V is set to 0.996, i.e the astronomical null depth N is 0.002. The random phasing error $\Delta\phi(t)$ is drawn from a Gaussian variable of mean value $\mu$=0.2 radian and standard deviation $\sigma$=0.3 radian, both typical of LBTI fringe tracking performance around 11 μm. The residual background estimation error is also assumed to be Gaussian, with a zero mean and a standard deviation set to 0.06 $I_0$. Given this type of observed data, the goal is to infer the true astronomical null depth N.

Even in this simple illustrative case, ensemble characteristics such as the mean or median of the $I_{obs}$ data provide completely wrong estimates of the astrophysical null, respectively 0.033 and 0.022. Additionally, these estimation errors strongly depend on the phase setting and rms jitter conditions for a given OB, i.e the exact values of $\mu$ and $\sigma$, and will generally not calibrate out between science and calibrator observations. The average of the lowest 5% observed nulls also provides a poor estimate in these background dominated conditions, with a value of -0.03. It would calibrate out with good accuracy if the science and calibrator stars have the same mid-infrared flux[11], but it only uses a small part of the data, and is still slightly sensitive to the phase tracking conditions. A better estimator, in this particular case, would be the mode (highest occurrence null signal) of the histogram, which is 0.011. But even that estimator does not hold, unless stars are infinitely bright. In fact, none of these raw estimators is even close to the actual astrophysical null value of 0.002.

Our approach, known as "Null Self Calibration (NSC)"[12-14] or "Visibility Self Calibration (VSC)"[15], is based instead on fitting the *whole* observed histogram of occurrences. Since we do not know a priori the temporal signal evolution, we rather use its statistical distribution, where the data order is irrelevant. For the toy data simulated here, and assuming that the background statistical fluctuations are well captured by data recorded in the complementary nod, there are only 3 free parameters left in the observed distribution of interferometric signals: the Gaussian phase error parameters ($\mu$,$\sigma$) and the object's visibility V. In the –numerical- self calibration method currently used for reducing LBTI data, a model sequence of temporal signals I'(t, $\mu$, $\sigma$, V) is generated for all possible values of the 3 free parameters. The resulting histogram is computed for each triplet ($\mu$, $\sigma$, V) and compared to the observed histogram. The middle panels of Figure 4 show the resulting histograms obtained when trying different values for these parameters. The parameters values are adjusted so that the model distribution matches the observed one, minimizing a goodness of fit Pearson $\chi^2$ test, as shown in the bottom panels of figure 4. For each visibility value tried, the $\chi^2$ is minimized on the other 2 parameters ($\mu$,$\sigma$). The main output of this fitting procedure is of course the object's visibility V, or equivalently its astronomical null depth. The best fit value is found around an astrophysical null of 0.002, which is the correct input value.

### 4.3 Null Self Calibration (NSC) applied to LBTI on-sky data

As we just showed, the proper estimation of the underlying astronomical null depth using NSC requires an accurate model of the interferometric signal statistical fluctuations. Only then can we separate variable instrumental effects from constant astrophysical properties by fitting the observed signal distribution. In reality the detected interferometric signal is more complex than presented in section 4.2. A first complication comes from the fact that individual beam signals have different intensities that fluctuate over time. A better model of the observed signal is then:

$$I(t) = I_1(t) + I_2(t) - 2|V|.\sqrt{I_1(t).I_2(t)}.\cos(\Delta\phi(t)) + B(t) \qquad [2]$$

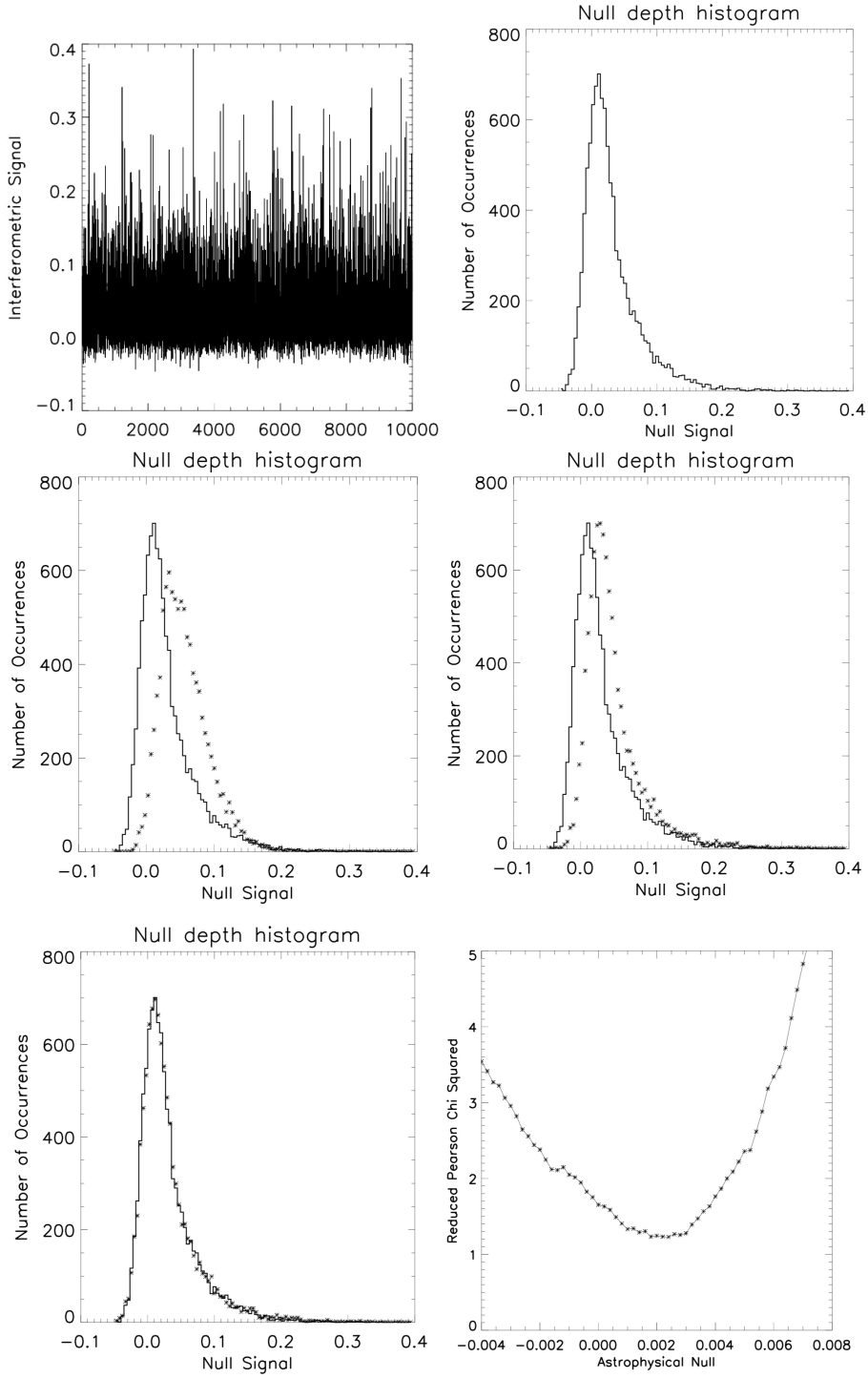


Figure 4: Top panels. Left: simulated time series of null signals affected by OPD fluctuations and background noise; right: corresponding histogram, showing numbers of occurrences per null signal recorded. Middle panels compare different model distributions (asterisks) with the observed one. Left: histogram obtained with wrong model parameter values for (V, μ, σ). Note that both the shape and position of the distribution are affected. Right: distribution obtained for the right OPD jitter parameters (μ,σ), but the wrong visibility, resulting in a lateral shift between the model and observed distributions. Bottom panels. Left: observed distribution and best fit distribution, essentially identical. Right: Reduced Pearson $\chi^2$ measuring the goodness of fit between the observed distribution and a set of model distributions with different source visibilities (astrophysical null values). See text for details.

where $I_{1,2}(t)$ are the instantaneous beam intensities. These individual photometric signals are measured at the end of each pointing, when the 2 beams are spatially separated on the detector. However, these measurements are affected by the large background shot noise, and their fluctuations are hence not representative of the actual stellar signal variations per beam. For this reason, and also because of the high Strehl stability provided by the AO system in the mid-infrared, only the mean measured values of $I_{1,2}$ are used to generate the model signal I'(t, μ, σ, V).

A second - more serious- complication comes from finite bandwidth effects, both temporal and spectral. Indeed, the detector exposure time is at least 30ms, which allows for rapid phase fluctuations within each exposure. These are caused by residual instrument vibrations and atmospheric piston, and are characterized by an rms $\sigma\phi_t(t)$, which is itself time-variable. Similarly, for each individual exposure centered around time t, there is a spectral phase rms variation $\sigma\phi_\lambda(t)$ across the N' band filter, and a time variable phase value at zero group delay. Under not too stringent assumptions, it is possible to show that after integrating over the finite wavelength range and detector exposure time, the signal recorded per frame is well modeled by:

$$I(t) \cong I_1(t) + I_2(t) - 2|V|.\left(1 - \frac{\sigma^2_{\varphi_t}(t)}{2} - \frac{\sigma^2_{\varphi_\lambda}(t)}{2}\right).\sqrt{I_1(t).I_2(t)}.\cos\big(\Delta\phi(t)\big) + B(t) \qquad [3]$$

where $I_{1,2}$ (t) and B(t) are now integrated over wavelength and time within each exposure centered around time t. As stated above, $I_{1,2}$ are assumed to be constant over time in our model. |V| is now the modulus of the broad-band complex source visibility. $\Delta\phi(t)$ is the residual phase offset averaged over all wavelengths and time during a given exposure. Its temporal variations are now caused by fluctuations of the OPD, represented as a random Gaussian variable (μ, σ) in the NSC model signal, and of the phase at zero group delay. In the case of a source with a top hat spectral profile centered at wavenumber $\sigma_0$ with FWHM Δσ, and assuming that the only effect of water vapor dispersion is to change the zero group delay phase $\phi_0(t)$, we get:

$$I(t) \cong I_1 + I_2 - 2|V|.\left(1 - \frac{\sigma^2_{\varphi_t}(t)}{2}\right).\sqrt{I_1.I_2}.\cos\big(\phi_0(t) + 2\pi.\sigma_0.OPD(t)\big).sinc(\pi.OPD(t).\Delta\sigma) + B(t) \quad [4]$$

The typical temporal phase rms per frame $\sigma\phi_t(t)$ is ~0.2 radian at 11 μm. It varies over time and is computed for each frame using contemporaneous K-band 1kHz fringe tracker data. It is actually injected in the model following equation (3) above. Taking this time-variable phase jitter term into account significantly improves the histogram fit quality (lower Pearson $\chi^2$). It also significantly improves the repeatability and accuracy of derived astronomical null depths for a given target. As for the spectral phase dispersion term, it consists both in a static offset and in a time variable term due to differential water vapor fluctuations effects above the 2 telescopes. (Conversely to Keck, the common mount approach of LBTI limits such effects considerably, but there are still some differential water vapor seeing effects between the 2 beams). While there are on-going efforts to estimate the phase at zero group delay fluctuations and intra-band dispersion terms for each frame using simultaneous phases estimates at H and K bands[15], and some extrapolation to the N' band, these terms are currently not included in the model signal I'(t, μ, σ, V) used in NSC fitting (i.e. $\phi_0(t) = 0$).

Figure 5 shows some of the best nulling measurements obtained to date at LBTI using this improved model with the NSC approach. They were obtained during a 3.5 hour sequence of observations of β Leo, including interleaved observations of different calibrators[8]. While the observing efficiency at the time did not allow to record more than 8 OBs per pointing, and only 2 science target pointings instead of the nominal 3, a null excess was still clearly detected around β Leo, at a level of $4.8 \times 10^{-3}$ +/- $5 \times 10^{-4}$. These are the most accurate null measurements made so far in the mid-infrared. Note that this 500 ppm precision corresponds to only 7-10 zodis in the particular case of β Leo (an A3V star), and would correspond to 15-20 zodis for a G2V star at 10pc, depending on the exact exozodi model used.

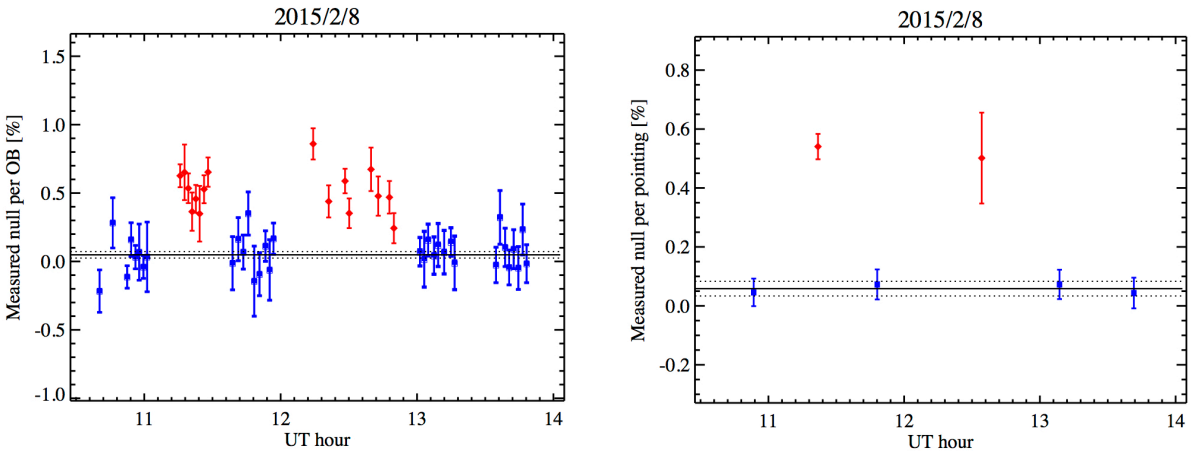


Figure 5 (adapted from[8]): Left: astronomical null depth measurements per OB as a function of UT time obtained on 2015 February 8. The blue squares show the calibrator measurements, while the red diamonds represent the β Leo measurements. The estimated instrumental null floor is represented by the solid black line and the corresponding 1σ uncertainty by the dotted lines. Right: corresponding null depth measurements per pointing (same notation). The longer time spent to acquire the second β Leo pointing increased the background bias and hence the dispersion of the null depth measurements per OB in the left-hand plot. This effect results in a larger systematic error for the second β Leo pointing and explains the larger error bar in the right-hand plot.

Figure 6 shows the improvement in astronomical null depth measurement uncertainty since the LBTI nuller commissioning started in February 2013. The best on-sky performance so far is 500 ppm, as demonstrated on a few stars including the 2 calibrated pointings of β Leo presented above. Further improvements are expected, especially through improved observing efficiency and active dispersion correction. Some further possible improvements are detailed next.

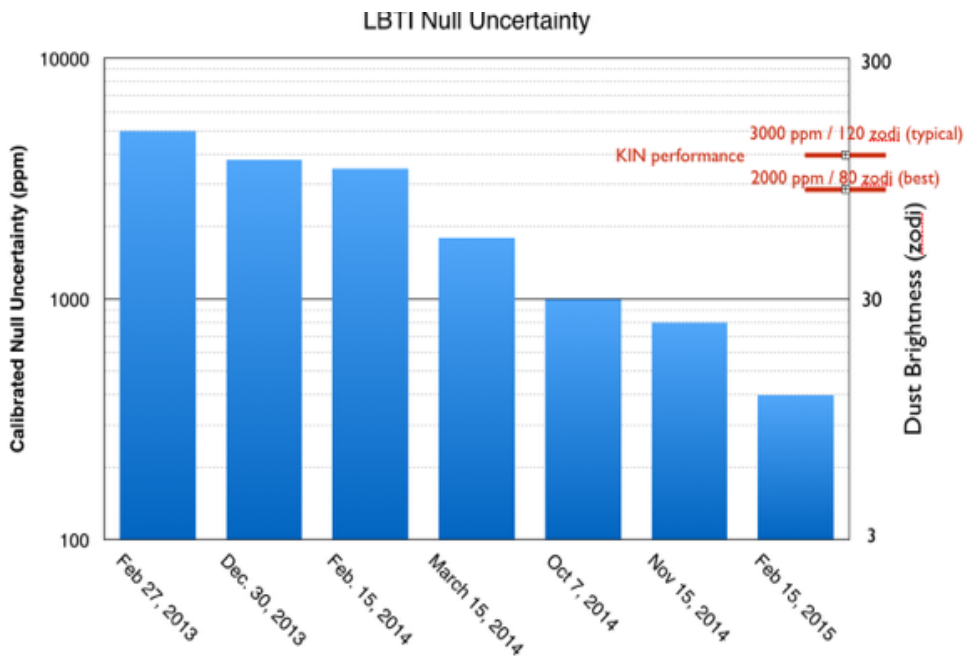


Figure 6: Evolution of on-sky performance over time during LBTI nuller commissioning. The null uncertainty went from ~5000ppm in early 2013, down to 500 ppm in early 2015, through a combination of instrumental improvements and the use of the NSC data reduction technique since the fall of 2014. The Keck Interferometer performance level (typically 2000 to 3000 ppm around a wavelength of 8.5 μm) is also indicated as a reference.

### 4.4 Further improvements to null estimation

There are at least 4 main improvements worth considering to break the barrier of 500ppm accuracy: the first two on the instrument side, the last two on the modeling side.

- Background estimation. As shown in section 3, the background estimation error is currently a large source of uncertainty, representing most of the 300ppm rms error budget for sources fainter than 2 Jy. It possibly even causes some biases in the null estimation, since NSC can not correct for any systematic error in estimating the background pedestal level. More background estimation tests are required on long sequences recorded on empty regions of the sky, still tracking a star to get all of the relevant effects. In this ground truth case where the actual background is known, different nodding /chopping strategies should be explored and compared, using different background reference regions and flat fielding techniques. Part of the background spatio-temporal variations also come from detector characteristics, basically the well known excess low frequency noise issue of the NOMIC mid-IR Si:As hybrid Raytheon Aquarius array[10], which translates in fast changes in the bias signal and which is yet to be mitigated.

- Variable chromatic dispersion. As stated above, this error term is currently not folded in the NSC model shown in Equation 3&4, and any change in the intra-band dispersion between star and calibrator observations will not cancel out. In other words, if the water vapor seeing conditions or any other dispersing effects change between the two pointings, this will cause a calibration error. The current work on water vapor dispersion will allow to send an offset command to the N' band OPD control system, basically telling the servo-system that the phase error value measured at K-band should be slightly corrected before being applied to the mid-infrared. This will reduce the apparent phase induced null fluctuations seen at N' band and increase the effective time spent at null. This latter effect is already captured by the NSC method, which sees it as an increased OPD jitter. The intra-band dispersion will not be actively corrected by a dispersion compensator (as was done at Keck). But it will be computed from the water vapor measurements at H and K, and this estimate should be fed into the NSC model of Equation 4 for highest accuracy results.

- Departure from single-mode assumption. We implicitly assumed everywhere here that the stellar interferometric signal is truly "single-mode", meaning that it is only affected by beam intensity fluctuations and by a single (piston) phase offset between the beams. This would strictly be true if the beams were injected into a common single-mode waveguide prior to detection, as done in many stellar interferometers. This is not the case at LBTI. However, the interferometric signal is integrated over the central core of the stellar PSF (the so-called photometric aperture), and any higher order differential wavefront aberrations between the 2 beams will be partially filtered out. These differential aberrations should also be small given the high Strehl expected in the mid-infrared. A possible improvement to the current model would be to include a differential tip-tilt (tt) visibility degradation term ($\sim 1 - k^2.tt^2$), which could either be fitted blindly like the other 3 parameters, or directly derived from the tip-tilt sensor.

- Distribution fitting approach. So far, the NSC fitting we have been using is based on minimizing a goodness of fit Pearson $\chi^2$ test. This approach does not lend itself easily to the computation of error bars, even more in the case of the numerical NSC[13], for which a random OPD sequence is generated, adding some "modeling uncertainty". Additionally, minimizing the Pearson $\chi^2$ only provides a maximum likelihood estimator in some particular cases. Because of these limitations, we recently made two major improvements:

- The first improvement is to compute an analytical probability distribution for the interferometric signal described by equation 3. This analytical probability density function, noted PDF ($\mu$, $\sigma$, V) only depends on the 3 usual parameters.

- The second improvement consists in computing the likelihood of the data for a given set of PDF model parameters using the multinomial distribution, as appropriate when dealing with histograms of observed occurrences. If M is the total number of interferometric signal measurements around null (typically a few 1000 per OB), if the number of observed occurrences in the k-th bin is $Y_k$ and the model probability is $PDF_k$ (V,$\mu$, $\sigma$), the likelihood function is given by:

$$L(Y_k,/(V,\mu,\sigma)) = M! \prod_k PDF_k(V,\mu,\sigma)^{Y_k}/Y_k! \qquad (4)$$

If the interferometric signal values range from $I_{min}$ to $I_{max}$, the number of bins between these values is set to ~ $\sqrt{M}$. As shown in Figure 7, in order to compute the astrophysical null depth uncertainty, we then use the statistical properties of the quantity -2ln(R), where R is the profile likelihood function defined as:

$$R(V) = \frac{Max_{\mu,\sigma}\left(L(V,\mu,\sigma)\right)}{Max(L)} \qquad (5)$$

These 2 modifications have only been tested on a few dataset, - as shown in figure 7- , but have not yet been implemented on full data sequences. They are expected to provide more accurate error bars and better behaved estimators.

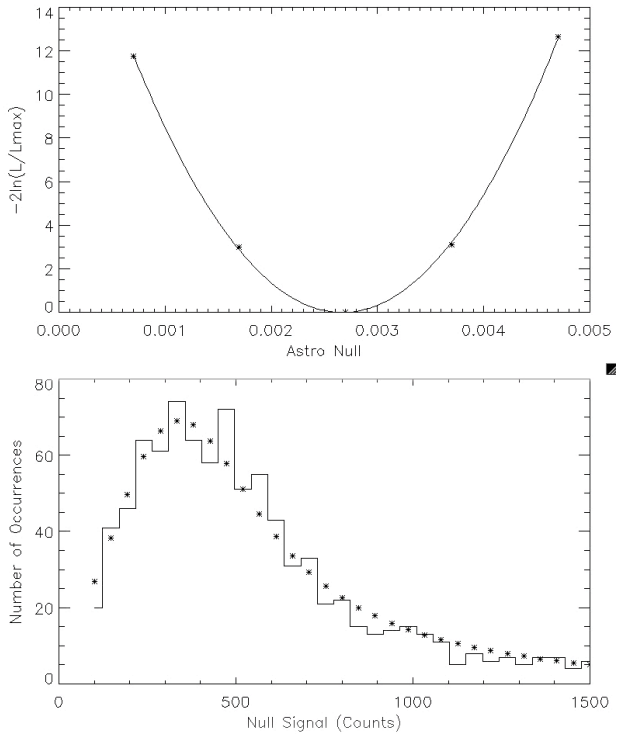


Figure 7: Example of data reduction results for one calibrator OB of the February 8 2015 LBTI data. The top panel shows the quantity $X^2$= -2ln(R(V)), where R is the profile likelihood function[7] of the data given model parameters (equation 4 & 5), and the astrophysical null is N=(1-V)/(1+V). For a large enough sample of null measurements (such as 1000 in this case), $X^2$ follows a $\chi^2$ distribution with 1 degree of freedom, and the 1σ uncertainty range corresponds to the null values for which the ordinate is equal to 1. The resulting null depth estimate for this OB is 0.0027 +/- 0.0005. The bottom panel shows the observed distribution (histogram) and in asterisks the number of occurrences predicted by the best fit (maximum likelihood) Probability Density Function (Nowak & Mennesson, in preparation).

## 5. CONCLUSIONS

During the scientific commissioning phase of the LBTI-HOSTS mid-infrared exozodi survey, astronomical null depth measurements have been demonstrated at the 500 ppm rms uncertainty level on the 7Jy star β Leo which showed the smallest exozodiacal excess detected to date. A number of instrumental improvements have occurred since these data were taken, providing better observing efficiency, vibration and dispersion control, and giving reasonable hope that routine measurements at the 300 to 500 ppm accuracy level should finally be within sight for the average target of the LBTI sample, representing a tenfold improvement over previous studies. Such measurements are extremely important for the preparation of future missions aiming at the direct imaging and spectroscopy of faint rocky planets in the habitable zone of nearby main sequence stars. Indeed, bright exozodiacal dust clumps could be confused with bona-fide planets and warrant lengthy spectroscopic observations to be correctly identified. The determination of exozodi levels is crucial

both from a statistical point of view, - with the potential of determining the median level of the class down to a few zodis -, and for selecting individual targets of future missions. These statistical and per source aspects justify observing as many stars as possible, a few dozen targets looking like the bare minimum sample size to draw any meaningful conclusions.

Finally, the LBTI formidable data reduction challenge, measuring interferometric visibilities at accuracy levels significantly better than 0.1% in the mid-infrared, is now within reach. This is in a large part thanks to the introduction of the null self-calibration (NSC) technique, which was previously applied to near-infrared data[11] and that we further adapted to the LBTI mid-infrared nulling data. Interestingly, the NSC / VSC method works equally well to determine astronomical null depths or visibilities, and should perform extremely well on any interferometers using single-mode recombination and fringe tracking to a fraction of the observing wavelength. Fast read-outs and high spectral resolution would also help minimize finite bandwidth effects, especially for long baseline interferometry, and get the most accurate fringe visibility measurements using NSC.

## ACKNOWLEDGEMENTS

Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

## REFERENCES

[1] Roberge, A. et al., "The exozodiacal dust problem for direct observations of Exo-Earths", PASP, 124, 918, 799 (2012)
[2] Stark, C. et al., "Maximizing the ExoEarth candidate yield form a future direct imaging mission, ApJ, 795, 122 (2014)
[3] Stark, C. and Kuchner, M. J., "The detectability of exo-Earths and exo super-Earths via resonant signatures in exo-zodiacal clouds", ApJ, 686, 637 (2008)
[4] Stark, C., 'The transit light curve of an exozodiacal dust cloud", AJ, 142, 123 (2011)
[5] Defrere, D. et al. 2012, "Direct imaging of exo-Earths embedded in clumpy debris disks", SPIE Astronomical Telescopes and Instrumentation Conference Series (2012)
[6] Reach, W.T., "Observational confirmation of a circumsolar dust ring by the COBE satellite", Nature, 374, 521 (1995)
[7] Mennesson, B. et al., "Constraining the exozodiacal luminosity function of main sequence stars: complete results from the Keck Nuller mid-infrared surveys", ApJ, 797, 119 (2014)
[8] Defrere, D. et al., "Nulling Data Reduction and On-sky Performance of the Large Binocular Telescope Interferometer", ApJ, 824, 66 (2016)
[9] Weinberger, A. et al., "Target Selection for the LBTI Exozodi Key Science Program", ApJS, 216, 24 (2015)
[10] Hoffmann, W. et al., "Operation and performance of the mid-infrared camera, NOMIC, on the Large Binocular Telescope". SPIE Astronomical Telescopes and Instrumentation Conference Series, 9147 (2014)
[11] Defrere, D. et al., "", "First-light LBT nulling interferometric observations: warm exozodiacal dust resolved within a few AU of η CrV", ApJ, 799, 42 (2015)
[12] Mennesson, B. et al., "New Constraints on Companions and Dust within a Few AU of Vega", ApJL, 736, L14 (2011)
[13] Mennesson, B. et al., "High-contrast Stellar Observations within the Diffraction Limit at the Palomar Hale Telescope", ApJ, 743, 178 (2011)
[14] Hanot C., et al., "Improving Interferometric Null Depth Measurements using Statistical Distributions: Theory and First Results with the Palomar Fiber Nuller", ApJ, 729, 110 (2011)
[15] Mennesson, B. et al., "Expanding the CHARA/FLUOR hot disk survey", Journal of Astronomical Instrumentation, Vol. 2 (2013)
[16] Defrere, D. et al., "Simultaneous Water Vapor and Dry Air Optical Path Length Measurements and Compensation with the Large Binocular Telescope Interferometer", SPIE Astronomical Telescopes and Instrumentation Conference Series, 9907 (2016)