Keck Interferometer nuller instrument performance

M. M. Colavita, a E. Serabyn, a S. Ragland, b R. Millan-Gabet, c R. L. Akeson c

a Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109
b W. M. Keck Observatory, 65-1120 Mamalahoa Hwy, Kamuela, HI 96743
c NASA Exoplanet Science Institute, California Institute of Technology, 100-22, Pasadena, CA 91125

ABSTRACT
The Keck Interferometer combines the two 10 m Keck telescopes as a long baseline interferometer. It is funded by NASA as a joint development among the Jet Propulsion Laboratory, the W. M. Keck Observatory, and the NASA Exoplanet Science Institute. In February 2008, the 10 um nulling mode began a 32 night observing program with three key science teams to perform a survey of nearby stars for exozodiational dust. This program has recently concluded, and has been followed by nuller observing on a variety of science topics through the standard proposal process. We provide a review and update of the nuller implementation, and describe the data reduction process, including the calibration approach. We then review the technical performance of the instrument based on the full key science data set, including sensitivity and systematic errors. We also provide some summary data on atmospheric effects applicable to the cophasing approach.

Keywords: Interferometry, Nulling, Keck, Exozodiacal dust

1. INTRODUCTION
The Keck Interferometer (KI)1-7 combines the two 10 m Keck telescopes as a long-baseline interferometer. It is funded by NASA as a joint development among the Jet Propulsion Laboratory, the W. M. Keck Observatory, and the NASA Exoplanet Science Institute. KI supports several different observing modes, which are available to the Keck community (Caltech, UC, NASA, UH, & NOAO) through the standard proposal process.8 These include H, K, and L-band visibility-amplitude modes at a variety of spectral resolutions, and the 10 um nulling mode. See Refs. 2 and 3, this conference, for an operations and science update. Under NSF funding, Keck Observatory is continuing development of its ASTRA (ASTrometric and phase-Referenced Astronomy) project9 to add additional capability to KI. ASTRA’s self-phase-referenced high-spectral-resolution mode is currently available to the Keck community, the dual-field mode is entering the shared-risk observing phase, and the astrometric mode is under development. See Ref. 9, this conference, for an update.

The KI nuller is the topic of this paper; it was developed to measure the quantity of exozodiational dust around nearby stars. After completion of its validation program, it carried out a year-long key science program between 2008 February and 2009 January. This program used a total of 32 Keck-Keck nights, and observed 44 unique targets in support of three key science teams.

In Section 2, below, we review the instrument and highlight some of the key technical aspects. Sections 3 & 4 review the operational scenario and the data processing approach. Section 5 reviews the instrument performance obtained during the key-science program. The appendix shows some sample atmospheric data describing water-vapor turbulence strength.

2. NULLER DESCRIPTION
The nuller approach and implementation are described in several papers.10-15 Below, we provide a brief summary and note some significant aspects.
2.1 Configuration

The nuller splits the pupil from each Keck telescope into two halves, and implements a four-aperture nulling configuration as shown in Figure 1. Modified Mach-Zehnder nullers on the long 85 m baselines are used to null out the light from the central star. The nulled light from the two modified Mach-Zehnders on a short 4 m baseline is combined in a Michelson combiner. KI being a ground-based instrument, SNR limitations are mostly attributable to the thermal background, not to the shot noise from the bright star. Thus the long baseline nullers are used to accommodate the dynamic range between the bright star and the faint exozodiacal dust. Said differently, the same SNR would accrue were we to measure visibility $V$ instead of null leakage $L$, and these two quantities are simply related\textsuperscript{10,13} for typical cases as $L = (1-V)/(1+V)$. But practically, when calibration and systematic effects are considered, a leakage of, say, 0.01, can be determined more accurately from a direct measurement of the leakage itself than from the deviation from unity of a direct measurement of a visibility with nominal value 0.99.

In describing the nullers we typically state that the modulation on the short baseline is used to extract the (small) signal from the (large) thermal background. This is indeed important, as it allows detection without rapid telescope chopping. However, the short baseline plays a second key role. By changing an internal pathlength by $\lambda/2$, the long-baseline nullers can be adjusted to pass most of the light. In this case, the fringe measurement on the short baseline provides a near zero-baseline, or photometric, calibration of the fringe signal. This is analogous to the case of visibility measurements, where the raw fringe amplitude $N V$ requires normalization by the source flux $N$ to yield $V$. For the nullers, the raw leakage amplitude is $N L$, and requires normalization by $N$ to yield $L$. The short baseline measurement provides the product of the flux and the visibility on the short baseline, $N V_s$. But typical targets are unresolved on the short baseline, so that $V_s \sim 1$, and thus simple normalization of the raw leakage amplitude is possible. And to emphasize the dynamic range point, because we measure leakage directly, a 1% error in this photometric calibration only causes a 1% error in the leakage; the calibration requirements to measure this level of leakage from a visibility observation are much tighter.

![Figure 1. Four-aperture nuller schematic. The light from each Keck telescope is split at the pupil, and modified Mach-Zehnder nullers operate on the parallel long baselines. The outputs of the Mach-Zehnders are combined in the cross-combiner implemented as a conventional Michelson interferometer with fast pathlength modulation.](image-url)
A detailed description of the nuller breadboard and camera is beyond our present scope, and they are described in more detail in Refs. 11-14, and references therein. Note that optical symmetry is a key requirement to obtain deep nulls, and this applies not just to the modified Mach-Zehnder nullers, but to the entire beam train. The nuller camera incorporates spatial filtering to suppress leakage from spatial aberrations; we implement ours with pinholes, reflecting available technology at the time of construction. The nullers also incorporate a longitudinal atmospheric dispersion compensator to allow control of group delay; path delay is controlled by optical delay lines. The error budget for null leakage has a number of terms, including single-mode terms due to phase and amplitude mismatches, multi-mode terms due to residual spatial aberrations, chromatic terms due the finite optical bandwidth, and asymmetry terms due to, e.g., polarization and rotation mismatches, and beam-shear errors. Most of these terms are associated with the full system and the limitations of observing through the atmosphere, rather than being specific to the nuller back-end – in the laboratory, the nuller breadboard can achieve much deeper nulls than are possible when observing on the sky. Note that null leakage is generally a quadratic function of the various perturbations. Thus the leakage contributions from all terms simply add, rather than combining in quadrature. However, as noted above, a deep null is not required from SNR considerations. What we do require is a deep enough null that photometric and other calibration can be done at acceptable accuracy, and, that the null be as stable as possible throughout the observing process.

2.2 Cophasing

The turbulent atmosphere makes the system implementation challenging. In particular, the integration time required to obtain a high SNR directly from the 10 um science light is longer than the relevant atmospheric time scales, and thus another approach is required to provide phase coherence. The KI nullers are stabilized by two, 2 um fringe trackers that use the short-wavelength light. This facility function of the fringe trackers is largely transparent to the nullers – they just see much less atmosphere fluctuations; the only significant communications are hold commands from the fringe trackers that indicate lock state. This is not to say that the nuller is passive: it also implements fringe tracking on the long and short baselines, but at bandwidths consistent with the achievable 10 um SNR. Thus the 2 um cophasing system must provide stabilization only at frequencies greater than about 0.01 Hz; lower frequency fluctuations and drifts are controlled using the 10 um light. The 2 um system is almost completely common mode with nuller, and the ac nature of the cophasing is such that we do not require any tie metrology (although we had originally included that provision). There is still extensive metrology, but it is used for overall vibration suppression and delay line servo control, similar to the other KI modes. We should note that we do not have a 2 um fringe tracker on the short baseline. This introduces a small coherence loss due to fluctuations at frequencies above the control bandwidth of nuller’s short-baseline tracker, but which is common to the leakage and short-baseline normalization measurements. The overall 2 um stabilization system is highly optimized to stabilize the 10 um path in the presence of both atmospheric and man-made disturbances. More detail on the system design, and design considerations, is given in Ref. 17.

The dispersion of atmospheric water vapor is also a significant consideration. The effect is that the 2 um system is not truly common mode with the 10 um system with respect to water-vapor pathlength fluctuations. For example, a differential water column that introduces 1 um of phase delay at λ=2.2 um introduces 2.3 um of delay at λ=10 um, and except for nights of very low water vapor seeing – typically very dry nights – the pathlength fluctuations from water vapor can’t be ignored. However, the large dispersion of atmospheric water vapor allows for its specific contribution to be estimated from the difference between the 2 um phase and group delays, and we include a low bandwidth correction for water vapor turbulence in the cophasing system. Finally, again because of the water vapor dispersion, phase stabilization alone is not adequate: we must also stabilize group delay. This term can also be estimated from the 2 um dispersion measurements, and the cophasing system includes a group-delay feedforward path. The computation of the 10 um phase and group delay feedforward terms from the 2 um data at KI is described in more detail in Refs. 18 & 19; we’ll discuss some atmospheric data later in the appendix.

Figure 2 shows all of the connections of the fringe-stabilization system.
2.3 Nuller control

The direct nuller stabilization scheme has several key aspects. The first is the use of a time-multiplexed microsequence as shown in Figure 3. The nuller has only a single camera; however, it has three baselines to stabilize. The approach we have used successfully is to time interleave fringe-tracking measurements on the three baselines with measurements of the leakage at null. As shown in the figure, we track on each of the three baselines for 50 ms during each 400 ms period of the microsequence, and integrate on the null fringe for 250 ms. When a fringe tracker is not active in the sequence, it’s essentially held stationary. The multiplexing occurs much faster than the underlying servo time constants. Specifically, our typical observing configuration uses a 1.25 s coherent integration filter for computing the phase for each of the long baselines, and for each iteration of the micro-sequence we push 50 ms of data into the filter and run one cycle of the servo. Because of the fine-grained interleaving, the effective control latency is just the data age of 5 s (i.e., it takes 10 s to fill the filter) plus one half of the 400 ms microsequence length. The net effect is that the three servos appear to work in parallel.
Figure 3 also indicates the pathlength modulation which switches the long baseline nullers between null and peak during different parts of the microsequence. The microsequence uses null-peak mode to measure the phase of the long-baseline fringe; this implementation directly servo to minimize the leakage on each baseline, rather than indirectly through servoing to a calibrated phase offset (as we do during acquisition).\textsuperscript{10,11} It is similar in philosophy to Ref. 20, and as an additional benefit, provides a better SNR than a traditional implementation. The tradeoff is that the mode has a limited acquisition range, and a more traditional (“gated”) mode with a large acquisition range (and a different microsequence, not shown – see Ref. 12.) is used to find the fringes. While the SNR gain of the null-peak mode improves the science SNR, the need to transition through gated mode sets the limiting magnitude. Not shown in the figure is an underlying slow (10 s) pathlength dither. This added by the delay lines, and removed by phasor derotation in the nullers (and again in the data post-processing). The net effect is average out phase-dependent systematics.

Figure 4 illustrates all of the control paths used in the nuller’s 10-um control.
3. OPERATIONAL CONSIDERATIONS

The nuller, and the overall science observing process, would be impossible to operate reliability without a high degree of automation, and sequencing is implemented at multiple levels (we refer now to higher-level sequencing, distinct from the microsequence just discussed). At the real-time level, there are sequencers which control fringe acquisition and tracking; these are implemented on both the 2 and 10 um fringe trackers. There are also stand-alone sequencers to automate specific tasks, such as the per-star shear adjustment, and the per-cluster pointing optimization. Finally, there is an observing sequencer that sequences through all of the steps required to collect data for each star, including angle-tracker acquisition, 2 um fringe acquisition, 10 um fringe acquisition, gated-mode acquisition, and final data collection. Figure 5 shows a timeline from the end of data collection on one star to the start of data collection on a second star.

Figure 4. Nuller control flow in null-peak sum-diff mode as used for data collection. In the figure, PD and GD refer to phase delay and group delay, and the suffixes O, I, and OI refer to data from outer ports only, inner ports only, and the coherent phasor combination of the outer and inner ports. For simplicity, the nuller control pickoffs are shown at the nuller block outputs; in reality, as there is only one camera, they are taken from the left side of the drawing during the appropriate beat of the microsequence, i.e., PD_{OI} uses light from all four camera ports. While not explicit in the drawing, unwrapping is generally disabled in this mode. Null-peak mode is only entered after transitioning through gated mode, which uses a different control strategy to provide a larger acquisition range. As final note, the control bandwidth for the three internal vernier delays is very low. These delays are initialized nightly on a bright star in “separate” mode, and subsequent data is taken in sum-diff mode. See Ref. 11 for more discussion of gated-mode and sum-diff control.
It’s worth noting that the daytime preparation is critical to nighttime success, and there are also sequencers that are used for daytime beam-train alignment (especially the coudé alignment) and preparation. We also use checklists to good effect – there is a lot to keep track of. However, the checklists serve an additional purpose: there are a number of applications of feedforward in the system whose effect is either only sensed at low bandwidth in the 10 um system, or only seen in the final data reduction. Thus run preparation must address both items that would cause obvious problems, e.g., like not finding fringes, as well as subtle problems that would affect the data quality but that wouldn’t cause obvious problems, and could be mistaken for, e.g., bad seeing.

4. DATA REDUCTION

The nuller data reduction is discussed in detail in Ref. 10. In principle, the processing is simple: the null leakage is just the ratio of the fringe amplitude on the short baseline with both long baseline nullers at null, to the fringe amplitude with them at peak. In practice, it’s a bit more complicated: for example, we must correct for the chromaticity of the phase step from null to peak. In addition, proper data gating is important to achieve consistent performance, e.g., to ensure that all systems are in the proper state and that all servos have settled. We do this objectively, with a set of criteria for system state and servo residuals that we apply consistently to all science data.

Nuller observing interleaves measurements of calibrator stars to measure the system leakage, analogous to the interleaving of calibrators for $V^5$ measurements to measure the system visibility. To establish the calibrator sizes, we use, in addition to archival data, the fringe visibilities measured by the 2 um cophasing system. Ref. 10 includes a table of 118 calibrator star diameters determined as part of the key science observing program. For each nuller target, we typically included multiple target scans with bracketing calibrators. We refer to such a set of science and a calibrator scans as a “cluster,” and this was our main unit for science planning; each cluster typically takes several hours to complete. Each cluster was measured at a particular long delay line setting, and was preceded by a nuller pointing optimization using a bright star; more detail is giving in Refs. 10 & 11.

5. KEY SCIENCE PROGRAM SUMMARY

The nuller exozodiacal dust key science project was carried out between 2008 February and 2009 January (Keck semesters 2008A and 2008B). The overall project objective was to observe nearby main sequence stars that could be potential targets for future planet-finding missions, as well as to observe targets which were known to have debris disks.
The project used 32 Keck-Keck night over those two semesters, distributed as 28 full nights and 8 half nights, and arranged into 8 observing runs. Three key science teams, competitively selected in 2007 November, provided the targets. In total, of the 46 targets submitted by the three teams, 44 were observed as part of the key science project.

The overall data set consisted of 92 clusters on 44 unique targets. 70% of these clusters had two or more target scans, along with associated calibrator scans. The overall data set included 700 individual scans (target or calibrator) with a typical microsequence duration of 500 s.

Figure 6 shows the achieved photometric performance; more discussion is given in Ref. 10. The system achieves an SNR on the long baseline at 2 Jy of 18.5 for phase and 42 for group delay for underlying coherent integration times of 1.25 s and 6.25 s, respectively (as discussed above, the clock time to fill these filters is 8× longer because of the microsequence). This is for the final null-peak configuration; the SNR is a factor of two lower while transitioning through gated mode.1 Because of non-linear computations (e.g. arctangent) in the gated-mode phase calculation, the SNR must be >>1 to work well: our practical limiting magnitude is ~1.5 Jy. In principle, as we’re phase referenced, we can go fainter with longer underlying filters; in practice, the acquisition times become excessively long because of the low 10 um servo bandwidth. As it is, the 10-90% rise time for the ADC loop (the slowest loop in the system) is 80 s with the filter lengths given above. Thus settling after initial acquisition takes typically one or two minutes, and is part of the timeline in Figure 5.

Figure 7 shows the formal error on the null leakage per scan vs. flux; the formal error shown is the formal error on the mean of the null leakage computed from dividing the typical 500 s scan into 10 segments. On the fainter science targets, the expected 1/SNR behavior is seen; the coefficient of the fit, 0.25% per scan at 2 Jy, matches that expected from the SNR reported by the real-time system in the previous figure. For bright stars, systematics come into play, and the formal error exhibits a floor level. However, while the formal error is interesting, we really care about the external (“true”) error, i.e., the formal error only tells us about the fluctuations about the mean leakage – it doesn’t tell us if the mean leakage is correct.

---

1 We partially accommodate the lower gated-mode SNR by dynamically increasing the phase filter length during this mode.
We can get closer to determining the external error by examining the scatter of the calibrated scans. So, for each cluster with more than one calibrated scan, we can compute the (weighted) rms, and plot this vs. flux. While this is a fairly noisy plot, given that we’re computing the rms from only several points per cluster, we can do a median fit to estimate the floor level for bright stars: this yields for our data approximately 0.2% per calibrated scan. This is a step closer to a proper floor estimate, incorporating the calibration process. Better still is to look at the scatter between repeated clusters on the same object. We were able to do this on one bright object, and it yielded 0.2% rms per cluster, and we adopt this value as our recommended error floor for bright stars, i.e., no matter what the formal error propagation or intra-cluster errors say, the final cluster error cannot be better than this value (or to be conservative, you can add this value in quadrature with the propagated error).

Ref. 10 goes through this in more detail, and also includes the flux dependence and a discussion of the error behavior of the spectral channels. All of the previous discussion pertains to our broadband 8–9 um channel; this synthetic channel is the coherent combination of the short-wavelength channels of our 8–13 um spectrometer. Because of diffraction, material absorption, pinhole mode matching, and other factors, the system responsivity is strongly peaked toward the blue, and the 8–9 um channel has 2/3 of the SNR of the weighted full-band synthetic WL channel used by the real-time system. For brighter stars, the data in the individual spectral channels can also be used for science. These have different systematic behavior, in particular, the error floor increases with wavelength, attributable, we believe, to residual correlation between the two aperture halves from telescope top-end thermal emission that diffracts more strongly into the beam train at long wavelengths (presumably, the short wavelength tail of this effect accounts for some of error floor of the 8–9 um channel).

However, for most stars, we need the SNR provided by the synthetic white-light channel. Incorporating the WL systematic term, above, with the overall photometric performance, our summary 1-sigma performance for the key science program for a 2 Jy source is approx. 0.3% rms per cluster.

ACKNOWLEDGMENTS

The Keck Interferometer is funded by the National Aeronautics and Space Administration (NASA). Observations presented here were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and NASA. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. Part of this work was performed at the Jet Propulsion

Figure 7. Formal error per scan vs. flux from Ref. 10. For faint stars, the behavior is as 1/SNR. For bright stars, there is evidence of a systematic noise floor (see text).
Laboratory, California Institute of Technology, and at the NASA Exoplanet Science Institute, California Institute of Technology, under contract with NASA.

REFERENCES


APPENDIX

The 700 scans of the key science program provide an interesting data set with which to examine basic atmospheric parameters. While the dry air pathlength fluctuations with an interferometer have been measured by several authors, the water-vapor fluctuations measured at infrared wavelengths have not been well examined. Figure 8, from Ref. 19, shows the mean dry air and water vapor power spectra computed from the K-band pathlength and group delay collected along with the N-band data. Cross-spectral techniques were used to measure down to very low levels of water vapor turbulence (which was measured from the K-band group delay following Ref. 18). The dry air spectrum is as expected, with near Kolmogorov shape and plausible total power for this baseline and expected seeing. The water vapor spectrum has a similar shape, but somewhat surprisingly has a lower level than expected based on archival radio data, likely due to different contributions from surface layer seeing. This lower level for the water vapor seeing, and also its lower
effective wind speed, makes its compensation easier. As shown in Ref. 19, the strength of the observed water vapor seeing varies more than the strength of the dry air seeing.

Figure 8. Power spectra of 2 um phase delay attributable to dry air (upper traces) and water vapor (lower traces) from Ref. 19, from data taken during the nuller key science program.