Michelson Science Center

KI Visibility Amplitude (V^2) Data Reduction Guide

Rachel Akeson and Rafael Millan-Gabet

Michelson Science Center

`version 1.0`

March 24 2005

Contents

1 Introduction

1.1 Purpose of this document

This document describes the visibility data reduction for the Keck Interferometer (KI). Raw interferometer data (called Level 0) is processed into internally calibrated and time-averaged visibility data (called Level 1) by a software package called `Kvis`. This manual details the algorithms used so that the user can understand the analysis method. `Kvis` is run on all KI visibility science data as part of the standard pipeline processing and the reduced Level 1 data are made available to the users via a web interface. See our web documentation for an overview of the data flow and formats [1].

It is not the purpose of this document to introduce the reader to the basic concepts of optical interferometry. The reader is referred to Section 5 for a list of references. It is also not the purpose of this document to describe the external calibration of interferometer data. External calibration is accomplished by making use of the software packages `wbCalib` and `nbCalib` and is described elsewhere [2].

1.2 Questions and comments

Questions and comments should be sent to msc-help@ipac.caltech.edu.

2 Data flow and processing

The beam combination for each baseline produces two outputs which are measured by a fringe tracker. One output is wide-band (WB) and is sometimes referred to as the "white light" channel. The second output is dispersed and directed to an array of pixels, which are referred to as the spectrometer channels. Visibilities are calculated for the wide-band pixel, the summation of the spectrometer channels and for individual spectrometer channels.

2.1 Data acquisition

Once the fringe tracker engine finds the main fringe and locks, science data are acquired according to the following sequence:

1. fringe data for 25 sec..

2. K1 only flux data for 5 sec.

3. K2 only flux data for 5 sec.

4. repeat steps $1 - 3$ five times.

5. dark data for 25 sec.

6. foreground data (on-source, off-fringe) for 25 sec.

---

[1] available at: `http://msc.caltech.edu/KISupport/dataOverview.html`

[2] see the relevant sections in the KI support page: `http://msc.caltech.edu/KISupport/`

2.2 Visibility estimator

The fringe detector uses a four bin algorithm with path length modulation, see Colavita (1999) for a detailed description of this method. The detector array is read 5 times, once before the modulation starts and 4 times after each quarter-wave of modulation. These reads are called z, a, b, c and d. Each of these reads consists of a number of sub-reads to reduce the effective read-noise. The basic output of the fringe detector, which is called a sample (Section 2.3), is the array counts for each channel and these values are read in by `Kvis`.

For each channel the integrated flux in each bin is calculated as:

$$
\begin{aligned}
A = (a - z)/n_s \quad & B = (b - a)/n_s \\
C = (c - b)/n_s \quad & D = (d - c)/n_s,
\end{aligned}
\tag{1}
$$

where n_s is the number of non-destructive reads of the detector. The fringe quadratures are then given by:

$$
\begin{aligned}
X &= C - A \\
Y &= D - B
\end{aligned}
\tag{2}
$$

and the total flux is:

$$
N = A + B + C + D.
\tag{3}
$$

A measure of the energy is given by:

$$
NUM = X^2 + Y^2.
\tag{4}
$$

From these values we calculate the squared visibility and fringe phase:

$$
V^2 = \frac{\pi^2}{2} \frac{NUM}{N^2}
\tag{5}
$$

$$
\phi = \arctan \frac{Y}{X}.
\tag{6}
$$

2.3 Time terminology

The data processing in `Kvis` involves the following basic time units:

Sample This is the time over which one complete measurement of a fringe is made (a complete set of *zabcd* data including multiple reads if applicable) and is the level at which fringe data is recorded in the archive. A typical value is 5 msec, corresponding to a fringe tracker rate of 200 Hz.

Frame Configurable length of time over which samples are combined. Spectrometer phases are referenced to the average WB phase over this time. The visibility (V^2) is computed using incoherent and coherent estimates for each channel. This time scale affects the coherent V^2 and generally should not be longer than 0.1 sec for KI data. This parameter (`frameTime`) is specified in the configuration file `ini.params` (see Section 3) and has a default value of 0.1 sec.

Block Configurable length of time over which frames are combined. The squared visibility (V^2) is calculated coherently and incoherently for the WB channel, a weighted average of the spectrometer channels, and individually for each spectrometer channel. This parameter (`blockTime`) is specified in the configuration file `ini.params` and has a default value of 5 sec.

Integration Complete observation of a source, including all the internal calibrations discussed below. As described in Section 2.1, a complete integration contains 125 sec of fringe data.

2.4 Internal calibrations

Several internal calibration measurements are made during normal observations. Some are made at the beginning of the observing night or after a change of the fringe tracker parameters (DARK, BRIGHT) and others are taken once for every source (DARK, FOREGROUND, RATIO1 and RATIO2) [3]. BRIGHT measurements are on an internal light source, FOREGROUND are on-source but off-fringe and the interleaved RATIO1/RATIO2 are on-source, off-fringe with either Keck II or Keck I shuttered internally.

The detector bias terms are measured and removed from the data on the sample time scale. The bias terms for X, Y and N (called BX, BY and BN) are calculated from the either the DARK or BACKGROUND data using equations 2–3. The X, Y, N values for each data sample are corrected by subtracting the bias terms and the corrected value of NUM is:

$$NUM^* = (X - BX)^2 + (Y - BY)^2. \tag{7}$$

Two additional bias terms come from the photon noise and the read noise. The variance of the read noise term, Brn, is calculated from the average value of NUM^* during the DARK or BACKGROUND observations. The variance of the photon term is kN, where k is the number of detected counts per photoelectron. A high-flux calibration observation (BRIGHT) is used to measure k which is given by:

$$k = (NUM^* - Brn)/\hat{N}, \tag{8}$$

where the notation $\hat{N}$ is used to denote bias corrected values. The final corrected value for the energy measure can now be calculated as:

$$\widehat{NUM} = NUM^* - k\hat{N} - Brn. \tag{9}$$

Control of when to calibrate the data and which data are used to calculate the bias values is described in Sec. 3.2. Note that the value of k currently used is set by internal measurements rather than nightly measurements due to the detector characteristics.

The RATIO1 and RATIO2 measurements are separately averaged over the 25 sec for each integration, and a single flux ratio value (R) is formed for the integration, in order to compute a ratio correction as [4]:

$$RC = \frac{(1+R)^2}{4R}, \tag{10}$$

---

[3] DARK measurements are shuttered internally, BACKGROUND measurements (off-source) are also possible but not currently implemented, as in the near-infrared both are approximately equivalent

[4] see our memo on this topic at: `http://msc.caltech.edu/KISupport/dataMemos/`

The ratio correction is calculated for the wideband channel, each spectral channel, and the summed spectrometer channel, and is output with the visibility data (but not applied to the measured visibilities at the reduction step; applying the ratio correction is an option of the external calibration programs wb/nbCalib).

2.5 Dewarping

If the path-length modulation does not match the wavelength of the channel, a correction must be applied to the fringe quadratures. For a stroke length in radians of s, where $s = 2\pi$ is a matched stroke, the dewarped values are given by:

$$
\begin{aligned}
X_c &= \frac{s}{2\pi}\left(\frac{X+Y}{\beta} + \frac{X-Y}{\alpha}\right) \\
Y_c &= \frac{s}{2\pi}\left(\frac{X+Y}{\beta} - \frac{X-Y}{\alpha}\right) \\
N_c &= N - \frac{\gamma}{2\alpha}(Y-X) \\
NUM_c &= NUM - 2GkN_c - \frac{\alpha\pi}{s}k(Y_c - X_c),
\end{aligned} \tag{11}
$$

where X, Y, N are the measured quantities and the constants are given by:

$$
\begin{aligned}
\alpha &= 2\sin(s/4) - \sin(s/2) \\
\beta &= 1 - \cos(s/2) \\
\gamma &= 2\sin(s/2) \\
G &= \left(\frac{s}{2\pi}\right)^2\left(\frac{1}{\alpha^2} + \frac{1}{\beta^2}\right).
\end{aligned} \tag{12}
$$

The second term in equation 11 for NUM takes the place of the kN term in equation 9. This dewarping correction is applied to all data (fringe tracking and calibration) before the bias corrections discussed in Sec. 2.4 are applied. This results in a unbiased estimate of the visibility. Whether or not the dewarping correction is applied is controlled by the `doDewarp` parameter (Section 3).

2.6 Time and spectral averaging

As previously described, `Kvis` combines groups of samples into frames. The samples are filtered through a buffer to eliminate the data immediately before a transition from lock to semi-lock in the fringe tracker. The size of this buffer corresponds to the averaging used by the fringe tracker when determining the lock state. Each good sample is dewarped and the bias corrections are applied. For the wide-band channel, an average phase is calculated for the frame. Coherent estimators for each frame are produced for each channel by derotating the fringe quadratures with respect to the WB average phase (ϕ_{WB}). For the spectrometer channels, ϕ_{WB} is scaled by the ratio of the WB and channel wavelengths.

$$
\begin{aligned}
\hat{X}_{i\,coh} &= \langle \hat{X}_i \cos(\phi_{WB}\lambda_{WB}/\lambda_i) - \hat{Y}_i \sin(\phi_{WB}\lambda_{WB}/\lambda_i)\rangle \\
\hat{Y}_{i\,coh} &= \langle \hat{X}_i \sin(\phi_{WB}\lambda_{WB}/\lambda_i) + \hat{Y}_i \cos(\phi_{WB}\lambda_{WB}/\lambda_i)\rangle.
\end{aligned} \tag{13}
$$

$$
\tag{14}
$$

The coherent value for NUM is given by:

$$NUMC = (\hat{X}_{i\ coh})^2 + (\hat{X}_{i\ coh})^2 - (k\hat{N} + Brn)/N_{samps} \tag{15}$$

where N_{samps} is the number of samples in the frame and the incoherent value by:

$$NUMI = \langle \widehat{NUM} \rangle, \tag{16}$$

where the average is calculated from the sample values.

In combining frames to form a block the values of $NUMI, NUMC$ and N are simply averaged. The coherent and incoherent values of V^2 for each channel are given by equation 5. The spectral line channels are combined to form a composite spectrometer visibility:

$$V_{spec}^2 = \frac{\sum_i V_i^2 W_i}{\sum_i W_i} \tag{17}$$

where the weights W_i are set by the user (see Section 3).

3 Parameters set by the user

Parameters controlling the averaging times and the internal calibrations can be set by the user, and all correspond to parameter values specified in the configuration file `ini.params` (see Table 1). The KI pipeline processing uses the same default values for all data. A memo describing how the default values of these parameters were determined using KI observations is available at `http://msc.caltech.edu/KISupport/dataMemos/`.

3.1 Grouping parameters

The `frameTime` and `blockTime` parameters can be set to any value. The minimum filling fractions (`minFrameFill, minBlockFill`) specify the minimum fraction of good data within the specified frame or block time. This calculation is done in terms of the number of good items in the next lower processing level. For example, `minFrameFill = 0.5` means that one half of the samples within the frame must be good. At the block level, all frames are counted equally towards the `minBlockFill`.

3.2 Calibration parameters

Which data are used to calculate the bias values is controlled by the `calType` parameter.

calType $= 0$ The bias values are read from a file `calData.input` which must be in the working directory. The file format is 1 row per channel, with the WB values first. The bias values are given in the order BX, BY, BN, Brn, k. A file in the correct format called `calData.input` can be produced by using the `saveCalData` parameter. Using this parameter will destroy any previous versions of the file.

calType $= 1$ The bias values are calculated from averages of all the DARK and BRIGHT data present in the archive data file. The entire file will be searched, not just the time range specified on the command line.

calType $= 2$ Bias values for BX, BY, BN, Brn are calculated from the average of each BACKGROUND instance and for k using an average of all the BRIGHT data. For each block, the set of bias values is chosen from the BACKGROUND data closest in time.

grouping			
frameTime	Length of frame in seconds		
	Allowed values:	real	
minFrameFill	Minimum fraction of frame with good data		
	Allowed values:	0 to 1	
blockTime	Length of block in seconds		
	Allowed values:	real	
minBlockFill	Minimum fraction of block with good data		
	Allowed values:	0 to 1	
processing			
specAvgType	Sets weights used in spectrometer channel averaging		
	Allowed values:	0	weights $= 1$
		1	weights $\propto 1/\sigma_{V_i^2}^2$
doCal	Allowed values:	0	no internal calibration applied
		1	internal calibration applied
calType	Controls which calibration data is used		
	Allowed values:	0	data read from file `calData.input`
		1	use DARK,BRIGHT data
		2	use BRIGHT and closest BACKGROUND data
		3	use BRIGHT and median BACKGROUND data
		4	use BRIGHT, BACKGROUND and FOREGROUND data
		5	use BRIGHT, BACKGROUND and FOREGROUND data
saveCalData	Controls writing of calibration data to file		
	Allowed values:	0	no file written
		1	write data to file `calData.input`
doDewarp	Allowed values:	0	no dewarping calibration applied
		1	dewarping calibration applied
flags			
doPrintRaw	prints all samples (WB only) regardless of data type		
doPrintSample	prints calibrated samples, one line per channel, grouped by frame		
doPrintFrame	prints calibrated frames, one line per channel		
doPrintBlock	prints WB and averaged spectrometer channels		
doPrintSpec	prints individual spec channel blocks		
doPrintAnc	prints ancillary telemetry		
	Allowed values:	0	do not print
		1	print

Table 1: Brief descriptions of all parameters in the `ini.params` file.

calType = 3   Bias values for BX, BY, BN, Brn are calculated from the average of each BACKGROUND instance and for k using an average of all the BRIGHT data. For each block, a set of 3 bias value sets is chosen based on closest time and each bias value is determined separately as the median value from those 3 sets. This is the default for KI data.

calType = 4   As for calType = 3, except BX and BY are calculated from the FORE-GROUND observations.

calType = 5   As for calType = 3, except BX, BY, Brn are calculated from the FORE-GROUND observations. This method results in the removal of an estimated scintillation term from the data.

By default, for `calType` = 1 to 5 the entire input data file is searched for calibration data, regardless of the start and stop times specified. This loop through the data can take many minutes for data files lasting for several hours.

If no BRIGHT data are present, the value for k is set to its default value, which is 0.26 counts/photo-electron (May 2003).

4 L1 data output

4.1 L1 files and contents

KI users download the results of the data reduction described above, in the form of a set of L1 data files (available in ASCII and FITS formats). The contents of each L1 file is described in detail in `http://msc.caltech.edu/KISupport/v2/KIV2dataProducts.html`.

4.2 Browsing your L1 data (`KvisPlot`

Your L1 data constitutes the main input to the data calibration programs (wbCalib and nbCalib) that will return L2 data ready for astrophysical interpretation. However, if you wish to also browse your L1 data, you may use the program `KvisPlot`, distributed with the wb/nbCalib packages. `KvisPlot` simply takes a L1 data file as input and produces a series of plots representing the most relevant quantities. For a summary of `KvisPlot` useage and its options, please type `KvisPlot -help` once you have installed the wb/nbCalib package.

4.3 L1 data Quality Assesment

Included in the L1 data set, are files which contain detailed information on the data quality, relative to the ensemble of KI observations. Please refer to the QA description document at `http://msc.caltech.edu/KISupport/v2/QAoverview.html` for detailed information on the QA metrics used and contents of the QA files.

In order to also facilitate a rapid assesment of data quality by our KI users, the detailed QA attributes are used to assing a "grade" to each piece of L1 data, which may be used to filter the inputs that go to the calibration step.

Each grade is a three digit field where the digits run from 0 to 3 with the following meanings:

0 : good data with no known issues or problems

1 : fair data which may have minor issues but will still calibrate normally

2 : poor data (weather or instrument) that will need very careful calibration

3 : data severely comprised by weather or instrument problems, should only be used with extreme caution

Each digit represents a given time span according to our usual nomenclature:

1st digit : this block

2nd digit : this integration

3rd digit : this night

Thus, the 2nd digit is the same for all blocks in an integration, and the 3rd digit will be the same for all blocks in a night. Note that the grades are be normalized, that is, there is no effort to make 1/4 of the grades be 0.

5   References

- General optical and infrared interferometry

  - Shao, M. and Colavita, M.M. 1992, *ARAA*, 30, 457.
  - Quirrenbach, A. 2001, *ARAA*, 39, 353.
  - Principles of Long Baseline Stellar Interferometry, ed. P. Lawson, Course notes from the 1999 Michelson summer school, available at `http://olbin.jpl.nasa.gov/iss1999/`. Paper copies also available by request to Peter.R.Lawson@jpl.nasa.gov.

- Visibility data reduction

  - Colavita, M.M. 1999, *PASP*, 111, 111.

- Technical desciption of the Keck Interferometer

  - Colavita, M.M. and Wizinowich, P., 2003, SPIE vol. 4838, p. 79