A HIGH-ORDER KALMAN FILTER FOR FOCAL PLANE CALIBRATION OF NASA'S SPACE INFRARED TELESCOPE FACILITY (SIRTF)

David S. Bayard and Bryan H. Kang

Autonomy and Control Section (345)
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109

ABSTRACT

This paper discusses the Instrument Pointing Frame (IPF) Kalman Filter, which is an algorithm developed for focal plane calibration of NASA's Space Infrared Telescope Facility (SIRTF). The IPF Kalman filter is a high-order square-root iterated linearized Kalman filter which is parametrized specifically for calibrating the telescope focal plane and aligning the science instrument arrays with respect to the telescope boresight. The most stringent calibration requirement specifies an alignment accuracy of 0.14 arcseconds, 1-sigma, radial. In order to achieve this level of accuracy, the filter carries 37 states to estimate desired frames while also correcting for systematic errors in pointing. The estimated pointing frames and calibration parameters are essential for precision telescope pointing, and end-to-end "pixels on the sky" ground pointing reconstruction.

1 INTRODUCTION

1.1 Overview

The Space Infrared Telescope Facility (SIRTF) is a space-based telescope that provides high-resolution views of the universe in the infrared range of the spectrum. SIRTF represents the fourth and final element in NASA's Great Observatory program. The new space telescope is presently scheduled for launch on a Delta II in August 2003. SIRTF has an 85 cm telescope aperture and requires the use of expendable cryogen for cooling the critical optical and detector units. It will be launched into a heliocentric orbit, slowly moving away from Earth with a small drift rate of about 0.12 AU per year [10]. Working in the infrared, SIRTF complements the range of science observations and wavelengths covered by the other three previous NASA Great Observatories (Chandra for X-ray, Hubble for visual, GRO for gamma-rays).

This paper discusses the Instrument Pointing Frame (IPF) Kalman Filter algorithm which has been designed specifically for calibrating SIRTF's telescope focal plane. The IPF Kalman filter is a high-order square-root iterated linearized Kalman filter which is parametrized for calibrating the telescope focal plane and aligning the science instrument arrays with respect to the telescope boresight. The most stringent calibration requirement specifies an alignment accuracy of 0.14 arcseconds, 1-sigma, radial [20]. In order to achieve this level of accuracy, the filter utilizes 37 states to estimate desired alignments while also correcting for expected systematic errors due to: (1) optical distortions, (2) scanning mirror scale-factor and misalignment, (3) alignment variation due to thermomechanically induced drift, and (4) gyro bias and bias-drift in all axes. The gyro scale factor and alignment parameters are not included because they are calibrated separately in-flight using a dedicated Gyro Calibration Filter (GCF) filter. The estimated pointing frames and calibration parameters support on-board precision pointing capability, in addition to end-to-end "pixels on the sky" ground pointing reconstruction efforts.

Various Kalman filtering approaches to calibrating frame alignments and focal plane distortions have been considered in the literature. Work in this area includes Murrell [11], Deutshmann and Bar-Itzhack [7], Shuster [16], Shuster and Lopes [17], Bierman and Shuster [4], Davis and Lai [6], and Pittelkau [13]. An overview of the sensor alignment literature can be found in [16] and [13].

Compared to these references, the IPF Kalman filter approach is novel in that it integrates both science and engineering parameters into a single formulation. For example, plate-scale parameters are estimated simultaneously with alignment calibration parameters. As pointed out in [17] these two problems are typically solved in separate steps and by different teams of analysts. Combining the two steps is unique to SIRTF, and results in higher accuracy estimates and improved operational efficiency. The main challenge becomes one of addressing the filter's high order. This challenge is addressed by using a modern array square-root filter formulation (cf., [9]).

An additional feature of the IPF filter is its use of polynomials to characterize time-

dependent behaviours. The gyro drift and thermomechanically induced alignment drift are characterized in this manner. Since polynomial coefficients are constant, this design allows the global re-linearization of the Kalman filter (an advantage of the approach in [6], and similar to the "iterated" Kalman filter approach discussed in [8]) while still accommodating time-varying behaviours. Such a polynomial approach was successfully used in NASA's SRTM mission [19] where the gyro bias was found to accurately fit a quadratic polynomial (in angle) over a 4 hour time span, using similar inertial grade mechanical gyros.

SIRTF is designed to carry three science payload instruments: the Infrared Array Camera (IRAC), the Infrared Spectrograph (IRS) and the Multi-band Imaging Photometer for SIRTF (MIPS). All instrument arrays and spectroscopy apertures are located in the telescope's circular focal plane which is cooled down to a temperature of 1.5 Kelvin. IRAC is designed to provide imaging and polarimetry over wavelengths of 1.8 to 27 microns. The IRS provides medium resolution spectra of astrophysical objects over wavelengths of 4 to 200 microns. The IRS instrument also includes peak-up arrays to provide real-time centroids of targeted IR objects, to facilitate their accurate transfer to the spectroscopy slits [1]. MIPS provides imaging, polarimetry and large area mapping over wavelengths of 20 to 200 microns. A main distinguishing element of the MIPS instrument is its scanning mirror, which moves along a single axis and is coordinated with spacecraft motions to facilitate science observations.

During SIRTF's 3 month In-Orbit Checkout (IOC) period, the space telescope will be commanded to perform a series of repeated carefully designed calibration maneuvers for each science instrument array. After a maneuver series is completed for a given science array, the IPF filter processes the collected attitude history data and instrument centroid data, and produces an estimate of the instrument frame along with estimates of other alignments and calibration parameters. The basic philosophy is to combine a high-order Kalman filter with carefully designed on-orbit experiment designs to achieve the overall desired calibration accuracy.

In order to meet calibration requirements, the IPF Kalman filter has several novel and important features [3]. These features include (1) A gyro pre-processor which allows gyro sensitivities to be pre-computed and stored beforehand. This completely eliminates the need for repeated and time-consuming gyro sensitivity propagation during each iteration of the filter cycle; (2) A parameter "masking" capability which allows the user to restrict the model to include only an arbitrary subset of parameters. This provides a completely flexible parametrization which allows one to match estimation model fidelity to the needs and requirements of each science array type; (3) A formulation based on a square-root iterated linearized Kalman filter which allows sequential data processing while providing good numerical conditioning [9]; (4) The flexibility to sequentially update prior estimates based on multiple data sets taken on separate days of the mission (i.e., a "multi-run" tool); (5) A sandwich-based experiment design which provides observability of all desired parameters by starting and ending on the same reference sensor [20], and which allows the same Kalman filter to be used for a multitude of different array types (cameras, spectroscopy slits, scanning instruments); (6) The ability to integrate both visible and infra-red sources in the same calibration data set; (7) The ability to process partial centroids which only contain information

along a single axis of the array (i.e., "slit mode"). This occurs, for example, when calibrating the entrance aperture of a spectroscopy slit by first scanning a source across the narrow slit width, and then subsequently along its length; (8) Operation in one of several possible "lite" modes to allow a trade-off between accuracy and robustness. For example, a completely gyroless mode can be invoked for small and/or incomplete data sets.

This paper will discuss the IPF Kalman filter parametrization, formulation and algorithm derivation. Only the most basic operation of the IPF filter will be presented. Details behind the other modes of operations are given in [3]. A simulation will demonstrate operation of the IPF filter as used for calibrating the MIPS 24 um array.

1.2 Performance Requirements

The end-to-end pointing requirement for SIRTF is 5 arcseconds absolute [20], and 1.4 arcseconds absolute for pointing reconstruction, all 1-sigma, radial. The most stringent requirement is for relative pointing, where attitude offsets accurate to 0.4 arcseconds are required to place an infra-red source at the center of an IRS spectroscopy slit. This amounts to a derived calibration requirement of 0.14 arcseconds for the IRS Short-Hi spectroscopy slit. Similar derived calibration requirements are tabulated for each of the instrument frames in Table 1.1.

There are two sets of error budgets corresponding to Coarse and Fine surveys. Coarse surveys are more relaxed and take place earlier in IOC when the telescope is still cooling. Fine surveys occur later in the IOC period after the telescope has cooled sufficiently and the telescope is fully operational.

Array/Slit	Required TPF to IPF Alignment	
	Coarse	Fine
IRS PeakUp Arrays	1	0.25
IRS PeakUp Sweetspots	1	0.14
IRS Short λ Slits	1	0.14
IRS Long λ Slits	1	0.28
IRAC Arrays	1	0.14
MIPS 24 μm	1	0.14
MIPS 70 μm small	1.12	1.1
MIPS 70 μm large	2.65	2.6
MIPS 160 μm	3.75	3.7
MIPS SED	1.15	1.1

Table 1.1: Derived Requirements for Focal Plane Survey (arcseconds, 1-sigma, radial) [20]

1.3 Experiment Design - Sandwich Maneuvers

Each Prime frame is calibrated using a series of "sandwich" type maneuvers of the type shown in Figure 1.1. Here, a source is first centroided on PCRS1, then centroided on PCRS2, then moved over to the science array for a series of centroids (taken in a grid or pattern of points), and then centroided again on PCRS1. These are called sandwich maneuvers because they start and end on the PCRS sensors while containing science centroids in between. The centroids on the science array can be arbitrarily chosen by the designer, but must result in a time-tagged list of centroid locations in pixel coordinates. Examples include a 3x3 grid (IRS Peakup arrays), a 7x3 grid (MIPS arrays), and a pattern of 5 points arranged like the 5-of-diamonds playing card (IRAC arrays). The source used for science centroiding can be different from the one used on for PCRS centroiding. Calibration maneuvers are typically repeated several times to allow statistical averaging of results, giving a corresponding reduction of errors.

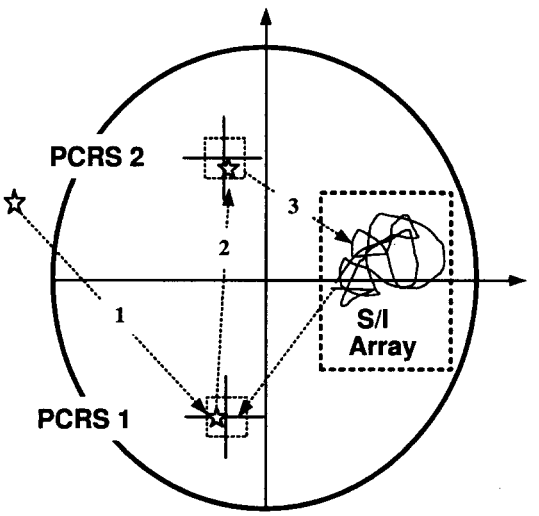

Figure 1.1: SIRTF Sandwich Maneuver for IPF Calibration

2 PRELIMINARY DEFINITIONS

2.1 Pointing-Relevant Frames

The main frames relevant to SIRTF pointing are shown in Figure 2.1 and the transformations between them are summarized in Table 2.2. The transformations A, R, T, C denote 3×3 direction cosine matrices for the indicated mappings.

The International Celestial Reference System (ICRS) frame serves as SIRTF's principle inertial reference frame. With a suitable relabelling, the star-tracker instrument frame serves

Transformation	Description	From	To
A	Attitude	ICRS	Body
R	Alignment	Body	TPF
T	Instrument	TPF	IPF_0
C	Scan Mirror Offset	IPF_0	IPF_Γ

Table 2.2: IPF Filter Transformations

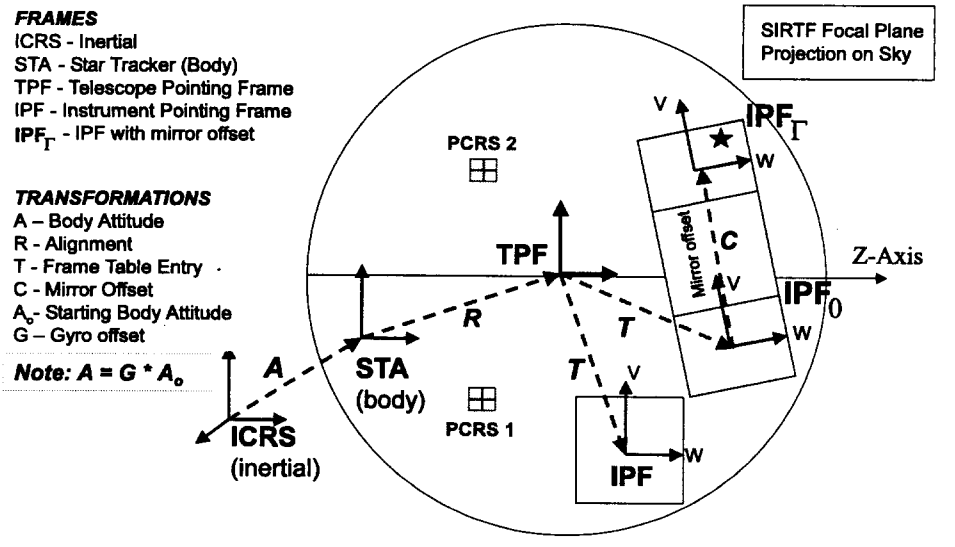

Figure 2.1: SIRTF Frames and Transformations

as the SIRTF Body frame (i.e., when spelled with its boresight as the x axis - see [3]). The mapping from ICRS to the Body Frame is denoted as the spacecraft attitude A. During a sandwich maneuver, the current attitude A is attained from a gyro offset G relative to a starting attitude A_0, in which case the attitude matrix is,

$$A = GA_0 \tag{2.1}$$

The Telescope Pointing Frame (TPF) has the telescope boresight as its x axis, and is defined rigorously in terms of the null points of the two PCRS sensors in [3]. The mapping from the Body Frame to the TPF is denoted as the alignment matrix R.

Instrument Pointing Frames (IPF) are defined by specific pixel locations in each science array and which adopt the orientation of the pixel rows and columns. The mapping from the TPF to any specified IPF is denoted as T. The IPF frames are stored in an on-board "Frame Table" as 128 values for T (stored as quaternions) and is used extensively for commanding purposes. Certain important IPF frames are denoted as **Prime Frames** (e.g., located at the center of each of the instrument arrays). Other frames are called **Inferred Frames** and are defined by a pixel offset relative to a nearby Prime frame. The nominal orientations of the science instruments and their associated Prime frames in the telescope focal plane are shown in Figure 2.2. Also shown are the associated w and v directions associated with each frame. The T matrix is the main quantity of interest for the IPF filter estimation effort.

The C matrix represents a scan mirror offset from a nominal starting position $\Gamma = 0$ to its current local mirror position $\Gamma \neq 0$. For non-MIPS instruments, the C matrix is set to identity. For MIPS, the frame defined when the mirror is at position Γ is denoted as IPF_Γ. Note that as the scan mirror moves there is an entire family of IPF_Γ frames generated as a continuous function of the variable Γ.

The attitude A is time-varying due to telescope repositioning, and R is time-varying due to thermo-mechanically induced alignment drift. The mapping T is assumed constant due to the fact that the telescope focal plane is actively cooled. The mapping C is time-varying due to a time-varying scan-mirror offset angle Γ.

2.2 Standard Coordinates

Let $u \in \mathcal{R}^3$ be a unit vector associated with a star location in the ICRS frame, i.e.,

$$u = \begin{bmatrix} \cos(\texttt{DEC}) * \cos(\texttt{RA}) \\ \cos(\texttt{DEC}) * \sin(\texttt{RA}) \\ \sin(\texttt{RA}) \end{bmatrix} \tag{2.2}$$

where $\texttt{RA}, \texttt{DEC}$ denotes the Right Ascension and Declination of the source (in radians).

Let $\ell \in \mathcal{R}^3$ denote the unit vector after a velocity aberration correction has been applied (cf., [15]),

$$\ell = \frac{u + \frac{V_{SC}}{c}}{||u + \frac{V_{SC}}{c}||} \tag{2.3}$$

where c denotes the speed of light.

Define the vector s as the resolution of ℓ in the IPF_Γ frame to give,

$$s = CTRA\ell \tag{2.4}$$

When the current attitude is the result of a gyro offset G from an initial attitude A_0 one can decompose A as,

$$A = GA_0 \tag{2.5}$$

Substituting (2.5) into (2.4) gives,

$$s = CTRGA_0\ell \tag{2.6}$$

Let the components of s be given as,

$$s = \begin{bmatrix} s_x \\ s_y \\ s_z \end{bmatrix} \tag{2.7}$$

Since s is a unit vector in the IPF_Γ frame, it can be projected into focal plane coordinates to give,

$$z = \begin{bmatrix} z_w \\ z_v \end{bmatrix} = \begin{bmatrix} s_z/s_x \\ s_y/s_x \end{bmatrix} \tag{2.8}$$

The elements of $z \in \mathcal{R}^2$ will be said to be in **Standard Coordinates**.

Let the matrices C, T, R, G be parameterized in terms of the elements of the parameter vector p_{2f} (to be defined in detail in Section 3), and let A_0 be related to an available initial attitude estimate $\widehat{A}_0$ as follows,

$$A_0 = (I - \psi^\times)\widehat{A}_0 \tag{2.9}$$

where $\psi \in \mathcal{R}^3$ denotes the initial attitude error. Then one can write (2.8) in the functional form,

$$z = h_z(p_{2f}, \psi) \tag{2.10}$$

This representation of the target source location in Standard Coordinates will be the starting point for the calibration process.

2.3 Oriented Angular Pixel (OAP) Coordinates

Typically, science centroids are obtained in units of pixels. However, calibration is more easily performed if pixel measurements are converted to units of angle (radians), and expressed with respect to an agreed upon origin and orientation. Oriented Angular Pixel (OAP) coordinates have been developed specifically for this purpose.

A pixel coordinate `(CX,CY)` (in the instrument (x, y) coordinate system) is converted to OAP coordinates using the following transformation,

$$y = \begin{bmatrix} y_w \\ y_v \end{bmatrix} = \begin{bmatrix} \mathtt{D11} & \mathtt{D12} \\ \mathtt{D21} & \mathtt{D22} \end{bmatrix} \begin{bmatrix} \mathtt{PIX2RADX} & 0 \\ 0 & \mathtt{PIX2RADY} \end{bmatrix} \begin{bmatrix} \mathtt{CX-CX0} \\ \mathtt{CY-CY0} \end{bmatrix} \tag{2.11}$$

Here, PIX2RADX,PIX2RADY are nominal plate scales, and the pixel coordinate (CX0,CY0) specifies the desired location where the Prime frame is to be embedded. The quantities D11, D12, D21, D22 are flip parameters (having values $0, -1, +1$), which specify how to map the instrument (x, y) coordinate directions into the focal plane (w, v) coordinate directions defined in Figure 2.2.

2.4 Mapping OAP to Standard Coordinates

Let $y_{true} \in \mathcal{R}^2$ be a target source as observed in OAP coordinates assuming that there is no centroiding error,

$$y_{true} = \begin{bmatrix} y_{w_true} \\ y_{v_true} \end{bmatrix} \tag{2.12}$$

Generally, y_{true} will not coincide exactly with z in (2.10) due to non-idealities in the optical system. To accommodate such non-idealities, a model which maps y_{true} in OAP coordinates to z in Standard coordinates is taken to be of the form,

$$z = \begin{bmatrix} z_w \\ z_v \end{bmatrix} = (I + M(p_1, \Gamma, y_{true})) \begin{bmatrix} y_{w_true} \\ y_{v_true} \end{bmatrix} \tag{2.13}$$

Here $M \in \mathcal{R}^{2\times 2}$ is a perturbation matrix which captures the non-idealities such as optical distortions, plate scale errors, etc. The exact form of M will be discussed in Section 3.2 as a function of the distortion parameters p_1, the scan mirror offset Γ, and the centroid measurement y_{true}.

The relation (2.13) assumes noiseless centroids. To generalize the model, a noisy centroid measurement y of the form (2.11) is introduced,

$$y = \begin{bmatrix} y_w \\ y_v \end{bmatrix} \tag{2.14}$$

The noisy centroid y is used to replace y_{true} in (2.13) according to the following relation,

$$z = \begin{bmatrix} z_w \\ z_v \end{bmatrix} = (I + M(p_1, \Gamma, y)) \begin{bmatrix} y_w \\ y_v \end{bmatrix} - \nu \tag{2.15}$$

where ν denotes the centroiding error in y. If M is small (which should always be the case), equation (2.15) is first-order equivalent to the more familiar additive noise model $y \simeq y_{true} + \nu$.

2.5 Calibration Equation

By equating (2.10) and (2.15) the following Calibration Equation is obtained,

$$\boxed{(I + M(p_1, \Gamma, y))\, y = h_z(p_{2f}, \psi) + \nu} \tag{2.16}$$

This is the main equation to be used for all SIRTF focal plane calibration. It is an end-to-end relation that maps the source location (known from a star catalog with velocity correction) to the pixel location where the source is observed on the science instrument array. Accordingly, it contains both optical distortions parameterized by p_1 and systematic pointing errors parameterized by p_{2f}. The end-to-end pointing transformations associated with the Calibration Equation (2.16) are summarized in Figure 2.3.

3 IPF FILTER PARAMETERS

3.1 Full State Description

The starting point for the Kalman filter design is the Calibration Equation (2.16). A full state vector x_f is defined as,

$$x_f = \begin{bmatrix} p_1 \\ p_{2f} \end{bmatrix} \tag{3.1}$$

where p_1 describes optical distortions and p_{2f} describes the systematic pointing errors in the Calibration Equation (2.16).

The parameters in p_1 and p_{2f} are selected such that they are constant parameters. A summary of the states p_1 and p_{2f} is given in Table 3.3 and Table 3.4. The next few subsections will be devoted to giving a detailed description of each of these parameters.

$p_1 \in \mathcal{R}^{17}$	Description	$\delta p_1 \in \mathcal{R}^{17}$
a_{00}	Constant Plate Scales	δa_{00}
b_{00}		δb_{00}
c_{00}		δc_{00}
a_{10}	Γ Dependent Plate Scales	δa_{10}
b_{10}		δb_{10}
c_{10}		δc_{10}
d_{10}		δd_{10}
a_{20}	Γ^2 Dependent Plate Scales	δa_{20}
b_{20}		δb_{20}
c_{20}		δc_{20}
d_{20}		δd_{20}
a_{01}	Linear Plate Scales	δa_{01}
b_{01}		δb_{01}
c_{01}		δc_{01}
d_{01}		δd_{01}
e_{01}		δe_{01}
f_{01}		δf_{01}

Table 3.3: p_1 State Variables, Perturbations and Mask Vector

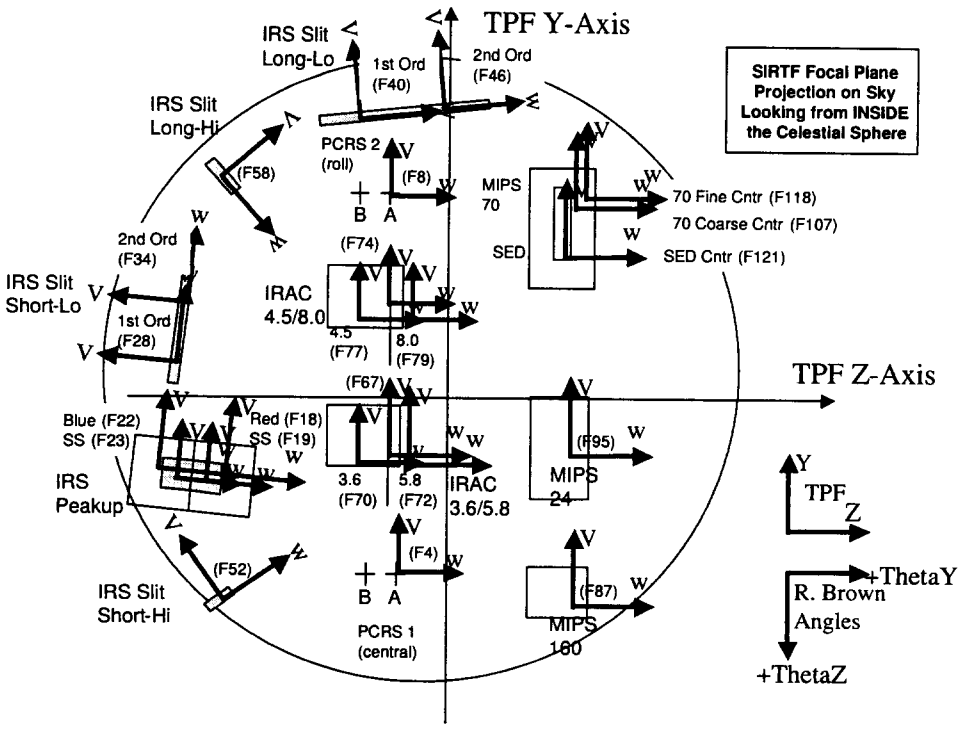

Figure 2.2: SIRTF Prime Frame Definitions

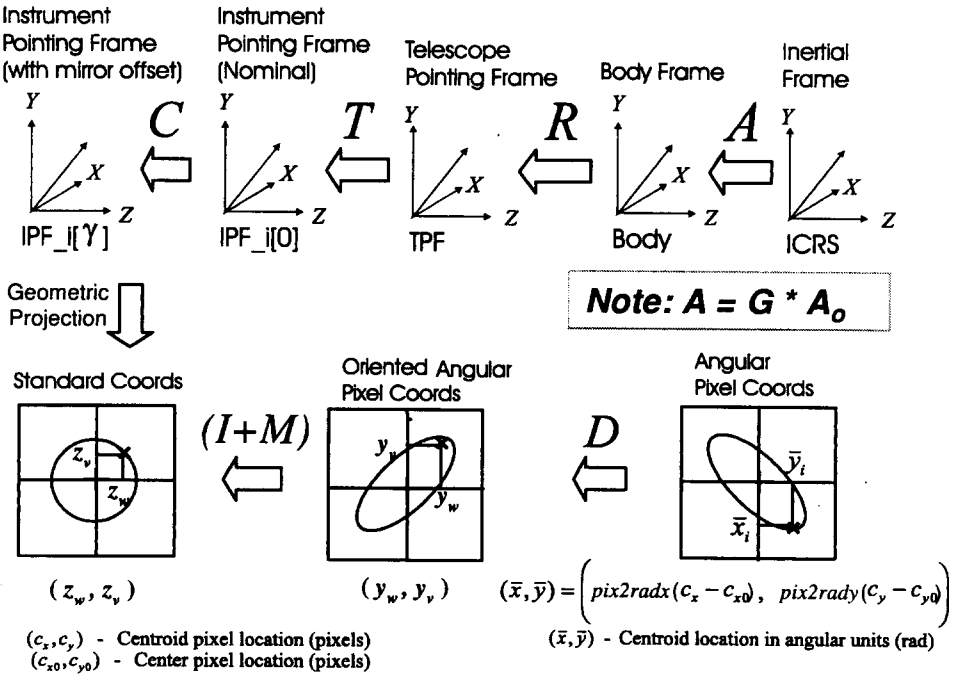

Figure 2.3: End-to-End Pointing Transformations

$p_2 \in \mathcal{R}^{24}$	Description	$\delta p_2 \in \mathcal{R}^{20}$
a_{m1}	mirror rotation axis unit vector in IPF (x)	
a_{m2}	mirror rotation axis unit vector in IPF (y)	$\delta\alpha$
a_{m3}	mirror rotation axis unit vector in IPF (z)	
β	scan mirror rotation angle scale factor	$\delta\beta$
q_{T1}	T (TPF to IPF) frame quaternion, $q_T(1)$	$\delta\theta_1$
q_{T2}	T (TPF to IPF) frame quaternion, $q_T(2)$	$\delta\theta_2$
q_{T3}	T (TPF to IPF) frame quaternion, $q_T(3)$	$\delta\theta_3$
q_{T4}	T (TPF to IPF) frame quaternion, $q_T(4)$	
q_{R1}	R Alignment quaternion, $q_R(1)$	δa_{rx}
q_{R2}	R Alignment quaternion, $q_R(2)$	δa_{ry}
q_{R3}	R Alignment quaternion, $q_R(3)$	δa_{rz}
q_{R4}	R Alignment quaternion, $q_R(4)$	
b_{rx}	Linear time varying contribution on alignment x-axis	δb_{rx}
b_{ry}	Linear time varying contribution on alignment y-axis	δb_{ry}
b_{rz}	Linear time varying contribution on alignment z-axis	δb_{rz}
c_{rx}	Quadratic time varying contribution on alignment x-axis	δc_{rx}
c_{ry}	Quadratic time varying contribution on alignment y-axis	δc_{ry}
c_{rz}	Quadratic time varying contribution on alignment z-axis	δc_{rz}
b_{gx}	Delta Gyro Bias from Nominal, x-axis	δb_{gx}
b_{gy}	Delta Gyro Bias from Nominal, y-axis	δb_{gy}
b_{gz}	Delta Gyro Bias from Nominal, z-axis	δb_{gz}
c_{gx}	Gyro Bias Drift Rate, x-axis	δc_{gx}
c_{gy}	Gyro Bias Drift Rate, x-axis	δc_{gy}
c_{gz}	Gyro Bias Drift Rate, x-axis	δc_{gz}

Table 3.4: p_2 State Variables, Perturbations and Mask Vector

3.2 Optical Distortion Parameters

The optical distortions in the calibration equation (2.16) are parameterized in terms of the matrix $M \in \mathcal{R}^{2\times 2}$ which has the form,

$$M(p_1, \Gamma, y) = M_{00} + \Gamma M_{10} + \Gamma^2 M_{20} + M_{01}(y) \tag{3.2}$$

where

$$M_{00} = \begin{bmatrix} a_{00} & c_{00} \\ c_{00} & b_{00} \end{bmatrix}; \quad M_{10} = \begin{bmatrix} a_{10} & c_{10} \\ d_{10} & b_{10} \end{bmatrix}; \quad M_{20} = \begin{bmatrix} a_{20} & c_{20} \\ d_{20} & b_{20} \end{bmatrix}; \tag{3.3}$$

$$M_{01}(y) = \begin{bmatrix} a_{01}y_w + c_{01}y_v & b_{01}y_v \\ d_{01}y_w & f_{01}y_w + e_{01}y_v \end{bmatrix}. \tag{3.4}$$

The parameter c_{00} is repeated symmetrically in M_{00} to disallow a redundant rotation with θ_1 of T (cf., [17]).

3.3 Scan Mirror Rotation Parameters

For science arrays having a scan mirror (i.e., MIPS arrays), the scan mirror rotation transformation can be defined by a direction cosine matrix C which maps the nominal IPF frame (denoted as IPF_0) to the IPF frame with a mirror offset (denoted as IPF_Γ). Mathematically, C has the form,

$$C(p_{2f}, \Gamma) = \cos(\beta\Gamma) \cdot I + (1 - \cos(\beta\Gamma))\, a_m a_m^T - \sin(\beta\Gamma)\, a_m^\times \tag{3.5}$$

Here $a_m = \begin{bmatrix} a_{m1} & a_{m2} & a_{m3} \end{bmatrix}^T$ is the scan mirror spin axis, Γ is the measured scan mirror angle (in radians), and β is the scale factor associated with measured mirror angle. The vector a_m is constrained to have unit norm, i.e.,

$$a_{m1}^2 + a_{m2}^2 + a_{m3}^2 = 1 \tag{3.6}$$

and the mirror transformation becomes the identity when the mirror is located in its nominal reference position ($\Gamma = 0$), i.e.,

$$C(p_{2f}, 0) = I. \tag{3.7}$$

For non-MIPS instruments (without a scan mirror), the condition $C = I$ is enforced.

3.4 Telescope Frame Parameters

The direction cosine matrix T transforms from TPF to IPF_0 and can be parameterized with a quaternion q_T as shown below.

$$T(q_T) = \begin{bmatrix} q_{T1}^2 - q_{T2}^2 - q_{T3}^2 + q_{T4}^2 & 2(q_{T1}q_{T2} + q_{T3}q_{T4}) & 2(q_{T1}q_{T3} - q_{T2}q_{T4}) \\ 2(q_{T1}q_{T2} - q_{T3}q_{T4}) & q_{T2}^2 - q_{T3}^2 + q_{T4}^2 - q_{T1}^2 & 2(q_{T2}q_{T3} + q_{T1}q_{T4}) \\ 2(q_{T1}q_{T3} + q_{T2}q_{T4}) & 2(q_{T2}q_{T3} - q_{T1}q_{T4}) & q_{T3}^2 + q_{T4}^2 - q_{T1}^2 - q_{T2}^2 \end{bmatrix} \tag{3.8}$$

3.5 Thermomechanical Drift Parameters

The direction cosine matrix R represents the time-varying mapping from the STA-defined Body frame to the TPF frame. The IPF filter parametrizes this misalignment in R as a quadratic function of time, i.e.,

$$R \triangleq \left(I_{3\times 3} - \left(b_r t + c_r \frac{t^2}{2} \right)^{\times} \right) R_0(q_R) \tag{3.9}$$

where,

$$b_r = \begin{bmatrix} b_{rx} \\ b_{ry} \\ b_{rz} \end{bmatrix} ; \quad c_r = \begin{bmatrix} c_{rx} \\ c_{ry} \\ c_{rz} \end{bmatrix} \tag{3.10}$$

The time $t = 0$ in (3.9) corresponds to the time tag of the first centroid of the very first sandwich maneuver. Accordingly, the quantity R_0 is the static alignment at time $t = 0$. For notational simplicity, the quaternion equivalent of the initial alignment R_0 is denoted as q_R (rather than q_{R0}).

3.6 Attitude and Gyro Parameters

The gyro offset G in (2.5) can be found by integrating the true rate $\omega \in \mathcal{R}^3$ as,

$$\dot{G} = -\omega^{\times} G \tag{3.11}$$

Since the true rate $\omega \in \mathcal{R}^3$ is not known exactly, an estimate must be generated. For computational convenience, this is done in two stages. First, the gyro pre-processor produces a nominal rate vector estimate $\omega_m^{\circ} \in \mathcal{R}^3$ according a certain construction (discussed in [3]). Second, an additive correction is applied to the nominal rate vector ω_m° to get the true rate as,

$$\omega = \omega_m^{\circ} + b_g + c_g t \tag{3.12}$$

4 FILTER MECHANIZATION

4.1 IPF Kalman Filter Architecture

The SIRTF IPF Kalman Filter algorithm is architectured as a square-root iterated linearized Kalman filter. The filter operates in block sequential form as shown in Figure 4.1. First, the nominal state estimate to be used for linearization purposes is prescribed at the beginning of the data set, corresponding to $t = 0$. For notation purposes, the start time of the jth sandwich maneuver is denoted as t_j, and the individual centroid times are denoted as T_k. Centroid data from each individual maneuver is then "stacked" into a single tall measurement vector which is used to update the Kalman filter. Accordingly, for a calibration data set having N sandwich

maneuvers, there will be N vector measurement updates. After processing the entire data set, the estimated correction is applied and the filter is re-linearized about the resulting new best state estimate. This process of re-linearization is repeated until convergence is obtained.

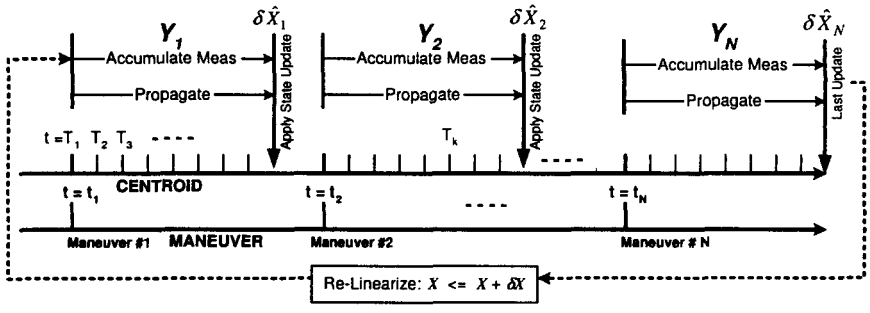

Figure 4.1: Square-Root, Iterated and Linearized Kalman Filtering Process

4.1.1 Time Update

Corresponding to the full state vector x_f defined in (3.1), an incremental state vector is given as,

$$\delta x = \begin{bmatrix} \delta p_1 \\ \delta p_2 \end{bmatrix} \in \mathcal{R}^{37} \tag{4.1}$$

where the quantities $p_1, \delta p_1, p_{2f}, \delta p_2$ have been defined in Table 3.3 and Table 3.3. It is noted that $\delta p_2 \in \mathcal{R}^{20}$ is of smaller dimension than $p_{2f} \in \mathcal{R}^{24}$ due to constraints on the parameters.

Since the IPF Kalman filter is parameterized by constant coefficients, the incremental state propagation equation can be written as,

$$\delta \dot{x} = \begin{bmatrix} \dot{\delta p_1} \\ \dot{\delta p_2} \end{bmatrix} = 0. \tag{4.2}$$

Accordingly, the discrete form of the incremental filter propagation equation for the mean and the square-root covariance can be written as,

$$\delta \hat{x}_{j+1|j} = \delta \hat{x}_{j|j} \tag{4.3}$$

$$P^{\frac{1}{2}}_{j+1|j} = P^{\frac{1}{2}}_{j|j}. \tag{4.4}$$

where $j+1|j$ signifies the predicted value at the start of the $j+1$'th maneuver, given measurements from the past j maneuvers.

4.1.2 Measurement Update

The Calibration Equation (2.16) will serve as a measurement equation for the Kalman filter by rearranging it into the form,

$$y = h(p_1, p_{2f}, \psi, y) + v \tag{4.5}$$

where,

$$h(p_1, p_{2f}, \psi, y) = -M(p_1, y)y + h_z(p_{2f}, \psi) \tag{4.6}$$

Note that the measurement y is used on both sides of (4.5), which requires a slight abuse of Kalman filter conventions (motivated by the implicit form of y in (2.16)). Note also that the notational dependence of M on Γ has been dropped for convenience.

Equation (4.5) can be linearized to obtain the desired Kalman filter update equation. To this effect, a prediction $\widehat{h}$ of h is constructed using the nominal state estimates $\hat{p}_1$ and $\hat{p}_{2f}$ as follows,

$$\widehat{h} = h(\widehat{p}_1, \widehat{p}_{2f}, 0, y) \tag{4.7}$$

Subtracting (4.7) from (4.5) gives the desired incremental measurement equation as,

$$\begin{aligned}
\delta y &\triangleq y - h\left(\hat{p}_1, \hat{p}_{2f}, 0, y\right) && (4.8)\\
&= \mathcal{K}_1 \delta p_1 + \mathcal{K}_2 \delta p_2 + H_\psi \psi + \nu && (4.9)\\
&= \begin{bmatrix} \mathcal{K}_1 & \mathcal{K}_2 \end{bmatrix} \begin{bmatrix} \delta p_1 \\ \delta p_2 \end{bmatrix} + H_\psi \psi + \nu && (4.10)\\
&= H \delta x + n && (4.11)
\end{aligned}$$

where,

$$\begin{aligned}
H &\triangleq \begin{bmatrix} \mathcal{K}_1 & \mathcal{K}_2 \end{bmatrix} && (4.12)\\
\mathcal{K}_1 &\triangleq \left.\frac{\partial h}{\partial p_1}\right|_{\hat{p}_1, \hat{p}_{2f}} = \left.\frac{\partial}{\partial p_1}\left(-M\left(p_1, y\right) y\right)\right|_{\hat{p}_1, \hat{p}_{2f}} && (4.13)\\
\mathcal{K}_2 &\triangleq \left.\frac{\partial h_z}{\partial (\delta p_2)}\right|_{\hat{p}_1, \hat{p}_{2f}} && (4.14)\\
H_\psi &\triangleq \left.\frac{\partial h_z}{\partial \psi}\right|_{\hat{p}_1, \hat{p}_{2f}} && (4.15)\\
\delta x &\triangleq \begin{bmatrix} \delta p_1 \\ \delta p_2 \end{bmatrix} && (4.16)\\
n &\triangleq H_\psi \psi + \nu && (4.17)
\end{aligned}$$

Equation (4.11) is the desired incremental measurement relation. The perturbation parameter vector $\delta p_1 \in \mathcal{R}^{17}$ is defined in Table 3.3, and $\delta p_2 \in \mathcal{R}^{20}$ is defined in Table 3.4.

It is emphasized that the perturbation $\delta p_2 \in \mathcal{R}^{20}$ has a smaller dimension than the vector it perturbs $p_{2f} \in \mathcal{R}^{24}$. This is due to the many constraints that exist in the problem. For

example, quaternions in p_{2f} have 4 elements, but only 3 of them are independent and have been carried in the perturbation δp_2.

In general, the derivative of a function $f(\xi_f)$ with respect to a constrained parameter vector ξ_f, will be defined in terms of only the independent parameters $\delta\xi$ (where $\delta\xi$ can be of lower dimension than ξ_f) by the matrix $\mathcal{K}$ that satisfies,

$$f(\xi_f) = f(\hat{\xi}_f) + \mathcal{K}\delta\xi + \mathcal{O}(||\delta\xi||^2) \tag{4.18}$$

Because in this general case the variable ξ may only exist in its incremental form $\delta\xi$, the following notation will be used throughout the remainder of the report,

$$\mathcal{K} \triangleq \left.\frac{\partial f}{\partial(\delta\xi)}\right|_{\hat{\xi}_f} \tag{4.19}$$

For example, this notation has been used in (4.14) for describing $\mathcal{K}_2$ which involves a derivative with respect to δp_2.

The measurement noise n in (4.17) is a combination of both attitude error ψ and centroiding error ν which gives a non-diagonal structure to the measurement noise covariance. In particular, assuming ψ and ν are statistically independent,

$$R \triangleq E[nn^T] = E[\,(H_\psi\delta\psi + \nu)\,(H_\psi\delta\psi + \nu)^T\,] \tag{4.20}$$

$$= H_\psi P_\psi H_\psi^T + \Xi \tag{4.21}$$

where $P_\psi = Cov[\psi]$ is the attitude error covariance and $\Xi = Cov[\nu]$ is the centroiding error covariance.

Let the following factorized matrices be defined,

$$R \triangleq R^{\frac{1}{2}} R^{\frac{T}{2}} \tag{4.22}$$

$$P_\psi \triangleq P_\psi^{\frac{1}{2}} P_\psi^{\frac{T}{2}} \tag{4.23}$$

$$\Xi \triangleq \Xi^{\frac{1}{2}} \Xi^{\frac{T}{2}} \tag{4.24}$$

Then (4.24) can be rewritten as,

$$R^{\frac{1}{2}} R^{\frac{T}{2}} = H_\psi P_\psi^{\frac{1}{2}} P_\psi^{\frac{T}{2}} H_\psi^T + \Xi^{\frac{1}{2}} \Xi^{\frac{T}{2}} \tag{4.25}$$

Recognizing that this has the general form $CC^T = AA^T + BB^T$, the square-root factor $R^{\frac{1}{2}}$ can be determined by using a QR factorization (cf., [3]).

Equation (4.11) provides an incremental measurement δy from a single centroid measurement. For filtering purposes, all of the centroids associated with the j'th maneuver are stacked into a single measurement vector as follows,

$$\delta\tilde{y}_j = \tilde{H}_j \delta x + \tilde{n}_j \tag{4.26}$$

where,

$$\delta\tilde{y}_j \triangleq \begin{bmatrix} \delta y_1 \\ \vdots \\ \delta y_{m_j} \end{bmatrix}; \quad \tilde{n}_j \triangleq \begin{bmatrix} n_1 \\ \vdots \\ n_{m_j} \end{bmatrix} \tag{4.27}$$

$$\tilde{H}_j \triangleq \begin{bmatrix} H_1 \\ \vdots \\ H_{m_j} \end{bmatrix}; \quad \tilde{R}_j^{\frac{1}{2}} = diag\{R_1^{\frac{1}{2}}, ..., R_{m_j}^{\frac{1}{2}}\} \tag{4.28}$$

where m_j is the number of centroids in the j'th maneuver.

Using the stacked measurement equation (4.26), the Kalman Filter gain and the square-root covariance update can be obtained. Specifically, given $\tilde{H}_j$, $P_{j|j-1}^{\frac{1}{2}}$ and $\tilde{R}_j^{\frac{1}{2}}$ consider the unitary triangularization of the following matrix,

$$\begin{bmatrix} \tilde{R}_j^{\frac{1}{2}} & \tilde{H}_j P_{j|j-1}^{\frac{1}{2}} \\ 0 & P_{j|j-1}^{\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix} \Theta \tag{4.29}$$

where Θ^T is an orthogonal matrix (i.e., $\Theta^T\Theta = I$). The factorization in (4.29) can be performed using a QR factorization [3]. Extracting X, Y and Z, one can compute the Kalman Filter Gain K and square-root covariance update equation as (cf., [3][9]),

$$K_j = YX^{-1} \tag{4.30}$$

$$P_{j|j-1}^{\frac{1}{2}} = Z \tag{4.31}$$

Note that the inverse in (4.30) is simple to compute and well-conditioned because it is a square-root covariance in block diagonal form. Given the Kalman gain K_j in (4.30) the incremental state update equation after the j'th maneuver is,

$$\delta\hat{x}_{j|j} = \delta\hat{x}_{j|j-1} + K_j\ \delta\tilde{y}_j \tag{4.32}$$

4.2 Sensitivity Equations for p_1 Parameters

The sensitivity $\mathcal{K}_1$ is defined from (4.13) as,

$$\mathcal{K}_1 = \frac{\partial}{\partial p_1}\left(-M\left(p_1, y\right)y\right)\Big|_{\hat{p}_1, \hat{p}_{2f}} \tag{4.33}$$

Using Kronecker identities (cf., [3],[12]) gives,

$$\begin{aligned} -M\left(p_1, y\right)y &= -(M_{00} + \Gamma M_{10} + \Gamma^2 M_{20} + M_{01}\left(y\right))y && (4.34) \\ &= -\left(y^T \otimes I\right) Vec\left(M\right) && (4.35) \\ &= -\left(y^T \otimes I\right)\left[\ S_{00} \quad \Gamma S_{10} \quad \Gamma^2 S_{20} \quad f_u\left(y\right)\ \right] p_1 && (4.36) \end{aligned}$$

where,

$$S_{00} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad S_{10} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad S_{20} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{4.37}$$

$$f_u(y) = \begin{bmatrix} y_w & 0 & y_v & 0 & 0 & 0 \\ 0 & 0 & 0 & y_w & 0 & 0 \\ 0 & y_v & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & y_v & y_w \end{bmatrix} \tag{4.38}$$

$$y^T \otimes I = \begin{bmatrix} y_w & 0 & y_v & 0 \\ 0 & y_w & 0 & y_v \end{bmatrix} \tag{4.39}$$

Substituting (4.36) into (4.33) gives the desired expression,

$$\mathcal{K}_1 = -\left(y^T \otimes I\right) \left[\begin{array}{cccc} S_{00} & \Gamma S_{10} & \Gamma^2 S_{20} & f_u(y) \end{array} \right] \tag{4.40}$$

4.3 Sensitivity Equations for p_2 and ψ Parameters

This section will use equations (2.6)-(2.8) to calculate the sensitivities $\mathcal{K}_2$ and H_ψ defined in (4.14) and (4.15)(), respectively.

Assume that current estimates $\hat{C}, \hat{T}, \hat{R}, \hat{G}, \hat{A}_0$ are sufficiently close to C, T, R, G, A_0 so that one can define the perturbations $c_k, \theta_k, \phi_k, g_k, \psi_k$ according to the following relationships,

$$C = (I - c_k^\times)\hat{C} \tag{4.41}$$

$$T = (I - \theta_k^\times)\hat{T} \tag{4.42}$$

$$R = (I - \phi_k^\times)\hat{R} \tag{4.43}$$

$$G = (I - g_k^\times)\hat{G} \tag{4.44}$$

$$A_0 = (I - \psi_k^\times)\hat{A}_0 \tag{4.45}$$

Substituting (4.41)-(4.45) into (2.6) and rearranging gives (cf., [3]),

$$s \triangleq CTRGA_0\ell \tag{4.46}$$

$$\simeq \left(I - c_k^\times\right)\hat{C}\left(I - \theta_k^\times\right)\hat{T}\left(I - \phi_k^\times\right)\hat{R}\left(I - g_k^\times\right)\hat{G}\left(I - \psi_k^\times\right)\hat{A}_0\ell \tag{4.47}$$

$$\simeq \left[I - \left(c_k + \hat{C}\theta_k + \hat{C}\hat{T}\phi_k + \hat{C}\hat{T}\hat{R}g_k + \hat{C}\hat{T}\hat{R}\hat{G}\psi_k\right)^\times\right]\hat{C}\hat{T}\hat{R}\hat{G}\hat{A}_0\ell \tag{4.48}$$

$$= \left(I - \eta^\times\right)\hat{N}\ell. \tag{4.49}$$

where,

$$\eta = c_k + \hat{C}\theta_k + \hat{C}\hat{T}\phi_k + \hat{C}\hat{T}\hat{R}g_k + \hat{C}\hat{T}\hat{R}\hat{G}\psi_k \tag{4.50}$$

$$\hat{N} = \hat{C}\hat{T}\hat{R}\hat{G}\hat{A}_0 \tag{4.51}$$

It is seen from (4.50) that the perturbation η is of the form,

$$\eta = \begin{bmatrix} I & \hat{C} & \hat{C}\hat{T} & \hat{C}\hat{T}\hat{R} \end{bmatrix} \begin{bmatrix} c_k \\ \theta_k \\ \phi_k \\ g_k \end{bmatrix} + \hat{C}\hat{T}\hat{R}\hat{G}\psi_k \tag{4.52}$$

$$= H_\eta \lambda_k + L_\psi \psi_k. \tag{4.53}$$

where,

$$\lambda \triangleq \begin{bmatrix} c_k^T & \theta_k{}^T & \phi_k{}^T & g_k{}^T \end{bmatrix}^T \in \mathcal{R}^{12} \tag{4.54}$$

$$H_\eta \triangleq \begin{bmatrix} I & \hat{C} & \hat{C}\hat{T} & \hat{C}\hat{T}\hat{R} \end{bmatrix} \tag{4.55}$$

$$L_\psi \triangleq \hat{C}\hat{T}\hat{R}\hat{G} \tag{4.56}$$

Given the above construction, the desired sensitivities can be written in the form,

$$\mathcal{K}_2 = \frac{\partial h_z}{\partial(\delta p_2)} = \left[\frac{\partial z}{\partial s}\right]\left[\frac{\partial s}{\partial \eta}\right]\left[\frac{\partial \eta}{\partial \lambda}\right]\left[\frac{\partial \lambda}{\partial(\delta p_2)}\right] = H_z H_s H_\eta H_\lambda \tag{4.57}$$

$$H_\psi = \frac{\partial h_z}{\partial \psi} = \left[\frac{\partial z}{\partial s}\right]\left[\frac{\partial s}{\partial \eta}\right]\left[\frac{\partial \eta}{\partial \psi}\right] = H_z H_s L_\psi \tag{4.58}$$

where evaluation on $\hat{p}_1, \hat{p}_{2f}$ and $\psi = 0$ is implied. The specific terms in these expressions for $\mathcal{K}_2$ and H_ψ are derived in the Appendix.

5 SIMULATION EXAMPLE

A simulation example is given to demonstrate the calibration of the MIPS 24 um array. Simulation is performed using FLUTE (FiLter Unit Test Environment) which is a tool specifically designed to unit test the IPF filter [2]. For this test, a 10 percent plate scale error is intentionally introduced to aid error visualization.

The experiment design calls for the commanding of a 7 by 3 grid of observations shown in Figure 5.1, where each row of 3 uses a separate sandwich maneuver. At each point in a row the spacecraft attitude is fixed, and there are six scan mirror offsets which move the source

3 positions up and 3 positions down in 25 arcsecond increments (some of which fall off the array and are not recorded).

The effect of the large plate scale error is seen in a-priori predictions shown in Figure 5.2. After filter convergence, the centroids are attitude corrected on a maneuver-by-maneuver basis, and the resulting centroids are plotted in Figure 5.3. The a-posteriori RMS centroid error is approximately .4 arcseconds (1-sigma, radial) which agrees well with the actual centroiding error used in the simulation having a value of .35 arcseconds. The estimated Prime frame and Inferred frames are depicted by the small window symbols (some are obstructed by the data). The innovations of the filter is shown in Figure 5.4 (w-axis only, the v-axis is similar). As expected the 1-sigma bound on the residuals approaches a steady-state value, and the innovations are well behaved with respect to this bound.

The sizes of the corrections applied for re-linearization (on a parameter by parameter basis) are shown as a function of the iteration number in Figure 5.5. It is seen that the filter has converged in approximately 8 iterations (i.e., re-linearizations).

After the filter has converged with respect to the nonlinear iterations, the 1-sigma bounds are plotted as a function of maneuver number in Figure 5.6. It is seen that a large amount of information is in the first 5 maneuvers, after which the errors decrease at a slower rate. This type of plot provides a useful covariance analysis for determining how many calibration maneuvers are needed.

The contribution of the thermomechanical drift to the final error is shown in Figure 5.7 as fitted to a quadratic, and the contribution of the gyro drift is shown in Figure 5.8, as fitted to a straight line.

The 1-sigma a-posteriori bounds are shown in Figure 5.10 along with the actual errors (this plot is possible because the true parameter values are known from simulation). The symbol key for this plot is given in Figure 5.9. As expected the actual errors lie within a few sigma of their a-posteriori bounds. The 1-sigma errors in three rotation angles $(\theta_1, \theta_2, \theta_3)$ associated with the IPF frame T are given by $(31, .035, .043)$ arcseconds, respectively. This gives a radial error of .055 arcseconds which meets the requirement of 0.14.

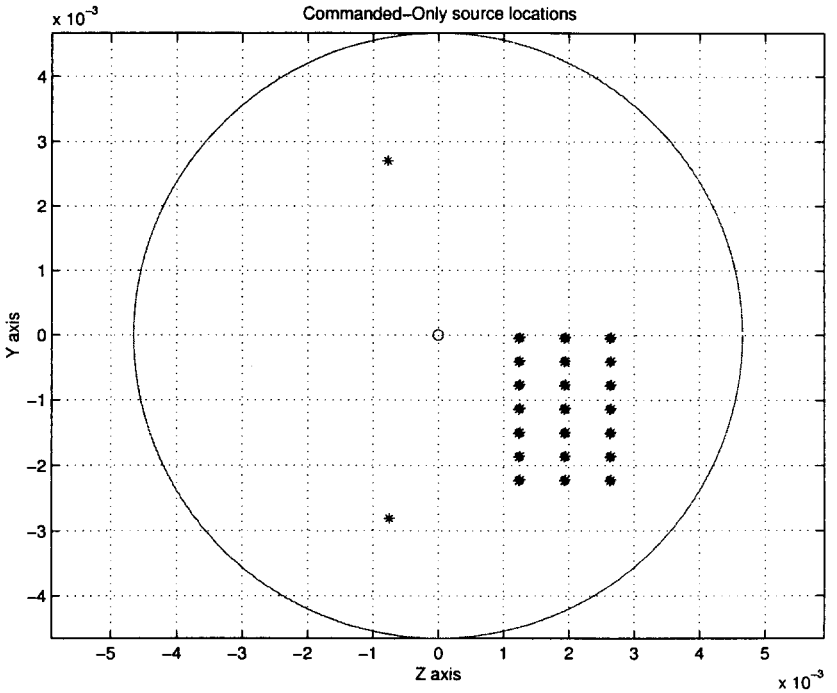

Figure 5.1: Commanded source locations

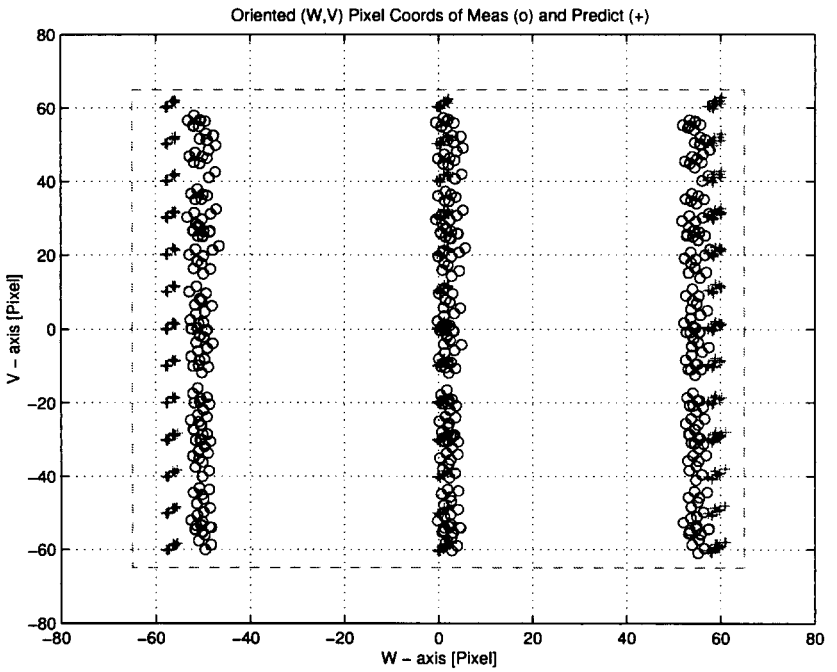

Figure 5.2: Measurements and predicts (a-priori)

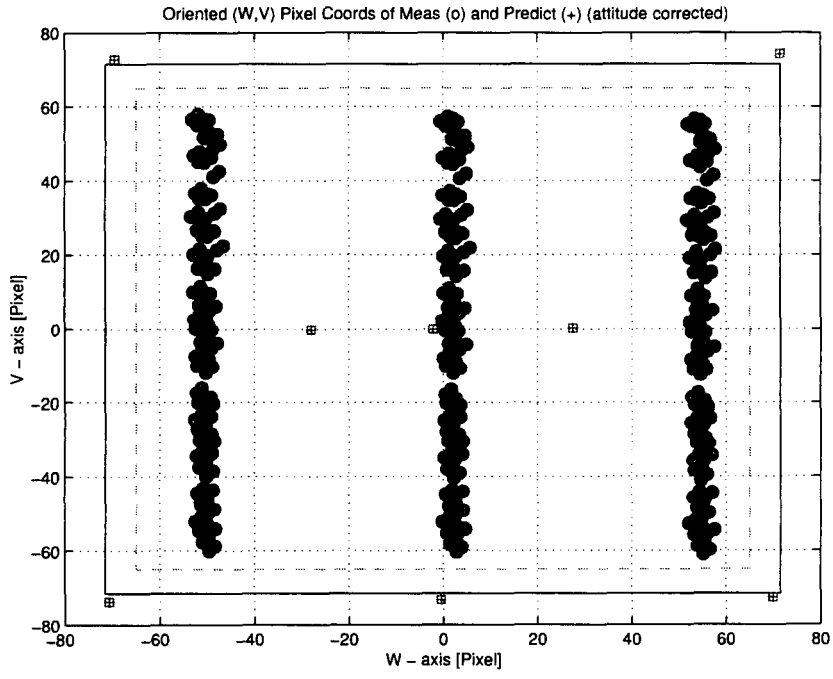

Figure 5.3: Measurements and predicts (a-posteriori)

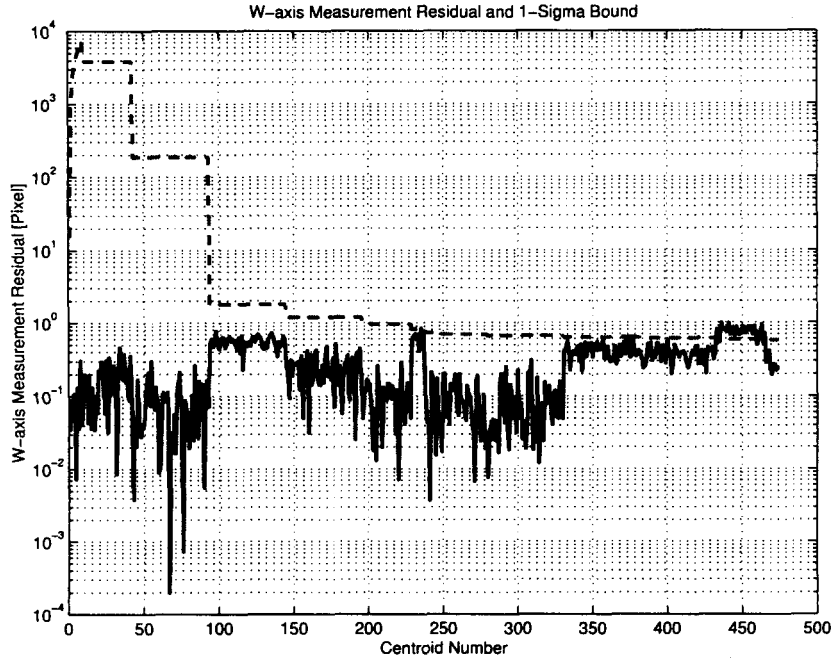

Figure 5.4: W-axis filter residual with 1-sigma bound

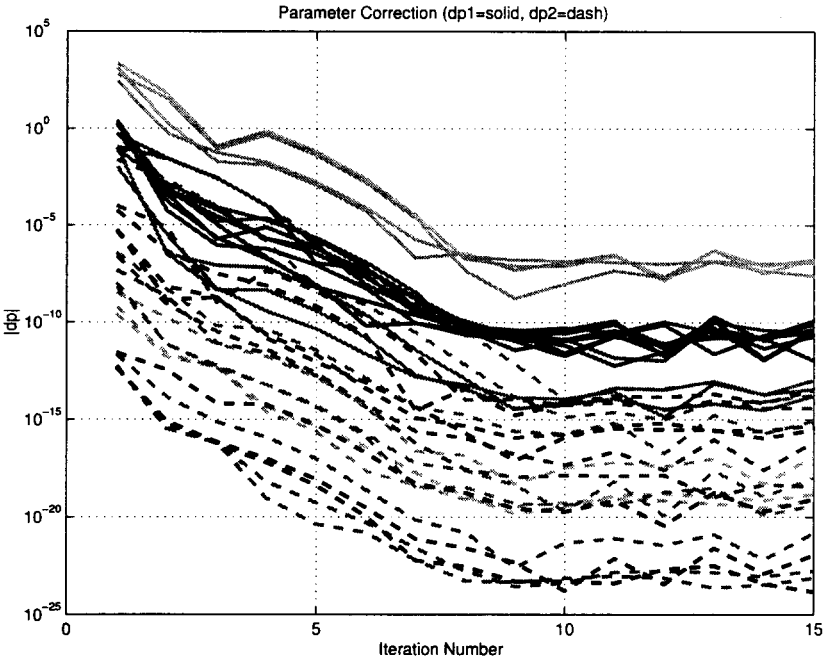

Figure 5.5: Correction size versus iteration number

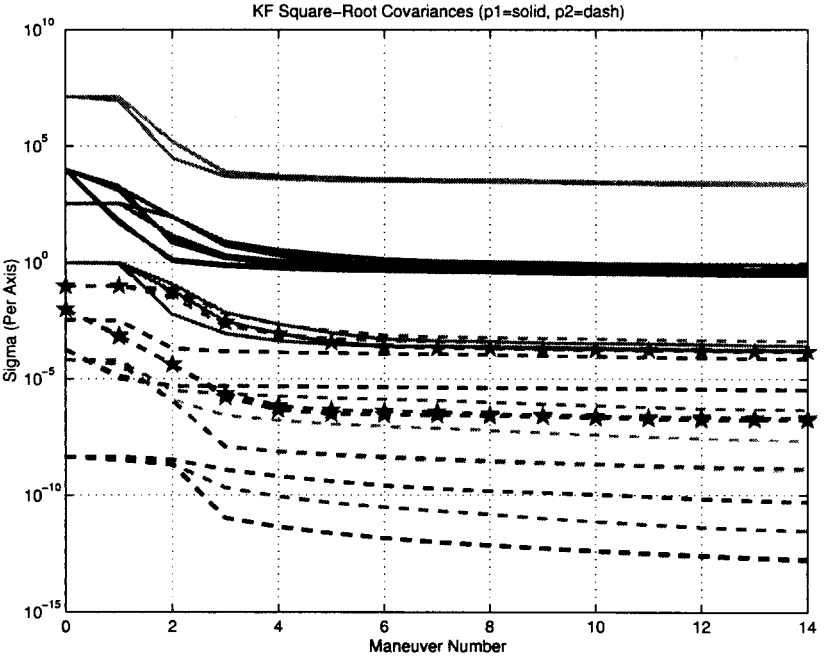

Figure 5.6: Filter convergence versus maneuver number

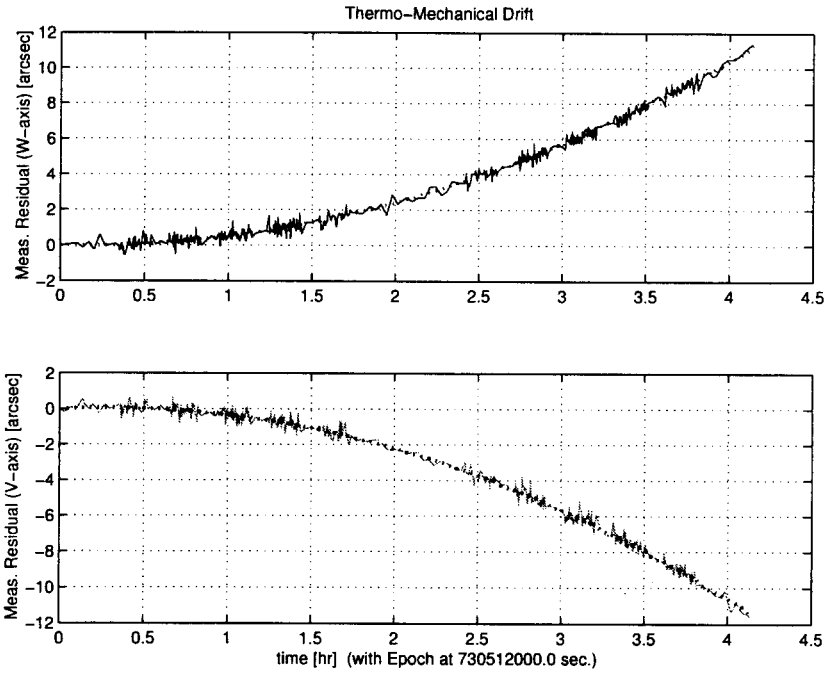

Figure 5.7: Estimate of thermomechanical drift

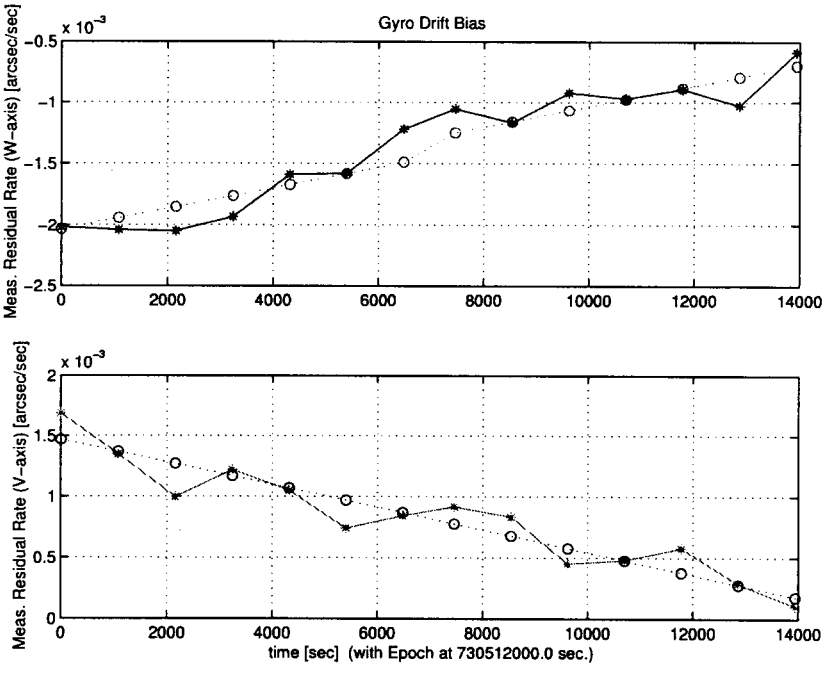

Figure 5.8: Estimate of gyro drift

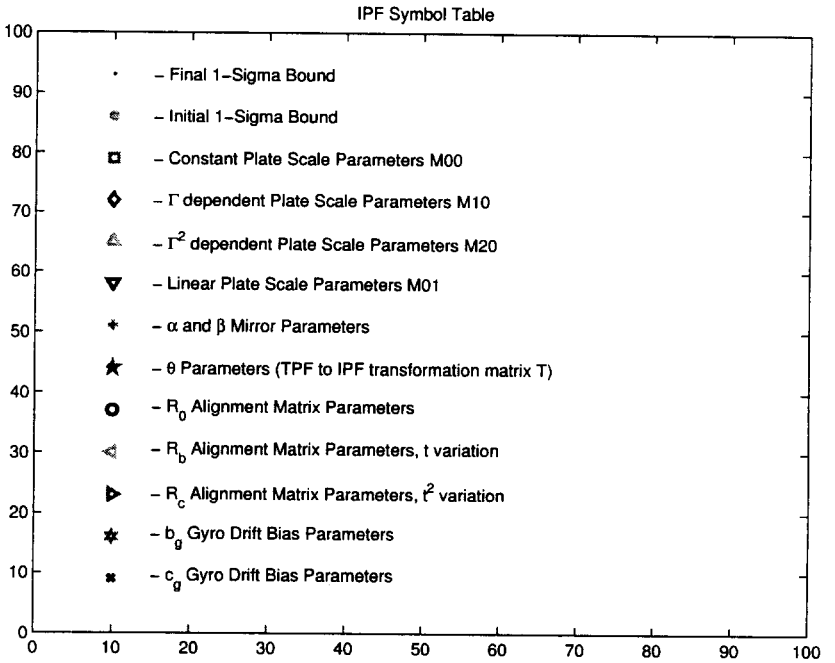

Figure 5.9: IPF Symbol Table

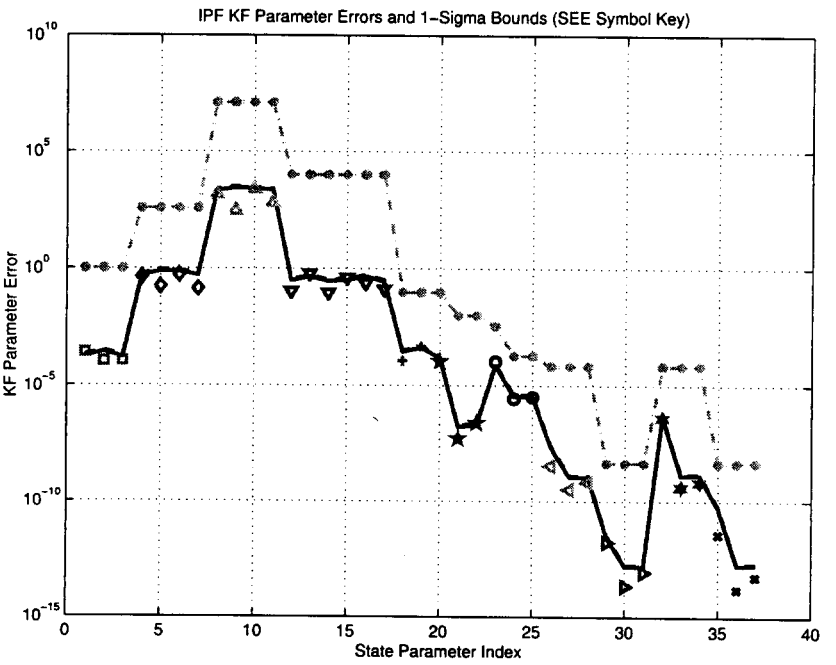

Figure 5.10: Parameter error 1-sigma plot

A APPENDIX: Sensitivity Calculations

A.1 $\frac{\partial z}{\partial s}$ Derivation

Starting with equation (2.8), the partial derivative of z with respect to s can be calculated as,

$$H_z = \frac{\partial z}{\partial s} = \begin{bmatrix} -\frac{\hat{s}_z}{\hat{s}_x^2} & 0 & \frac{1}{\hat{s}_x} \\ -\frac{\hat{s}_y}{\hat{s}_x^2} & \frac{1}{\hat{s}_x} & 0 \end{bmatrix} \tag{A.1}$$

where,

$$\hat{s} = \hat{N}\ell = \hat{C}\hat{T}\hat{R}\hat{G}\hat{A}_0\ell \tag{A.2}$$

A.2 $\frac{\partial s}{\partial \eta}$ Derivation

Equation (4.49) can be rearranged to give,

$$s = \left(I - \eta^{\times}\right)\hat{N}\ell = \hat{N}\ell - \eta^{\times}\hat{N}\ell \tag{A.3}$$

$$= \hat{N}\ell + \left(\hat{N}\ell\right)^{\times}\eta \tag{A.4}$$

where the vector cross product property has been used ($-a \times b = b \times a$). Taking the partial derivative of (A.4) with respect to η gives,

$$H_s = \frac{\partial s}{\partial \eta} = \left(\hat{N}\ell\right)^{\times} = \left(\hat{C}\hat{T}\hat{R}\hat{G}\hat{A}_0\ell\right)^{\times} \tag{A.5}$$

A.3 $\frac{\partial \eta}{\partial \lambda}$ Derivation

Starting with equation (4.53), the partial derivative of η with respect to λ can be calculated as,

$$H_\eta = \frac{\partial \eta}{\partial \lambda} = \begin{bmatrix} I & \hat{C} & \hat{C}\hat{T} & \hat{C}\hat{T}\hat{R} \end{bmatrix}. \tag{A.6}$$

Equation (A.6) is the most general expression. However, since sensitivities with respect to c_k are only needed for the MIPS instruments (which have scan mirrors), and since sensitivities with respect to θ_k are not needed for the PCRS sensors (which define the TPF frame), the actual value of H_η will be calculated differently depending on which array the specified centroid was taken on, i.e.,

$$H_\eta = \begin{cases} [I, \hat{C}, \hat{C}\hat{T}, \hat{C}\hat{T}\hat{R}] & \text{for MIPS} \\ [0, I, \hat{C}\hat{T}, \hat{C}\hat{T}\hat{R}] & \text{for non-MIPS} \\ [0, 0, \hat{C}\hat{T}, \hat{C}\hat{T}\hat{R}] & \text{for PCRS} \end{cases} \tag{A.7}$$

A.4 $\frac{\partial\lambda}{\partial(\delta p_2)}$ Derivation

Given $\lambda = \begin{bmatrix} c_k & \theta_k & \phi_k & g_k \end{bmatrix}^T$, it is convenient to decompose the sensitivity of λ with respect to δp_2 into the following matrix,

$$H_\lambda = \frac{\partial\lambda}{\partial(\delta p_2)} = \begin{bmatrix} H_c & 0 & 0 & 0 \\ 0 & H_\theta & 0 & 0 \\ 0 & 0 & H_\phi & 0 \\ 0 & 0 & 0 & H_g \end{bmatrix} \in \mathcal{R}^{12\times 20} \tag{A.8}$$

where,

$$H_c = \left[\frac{\partial c_k}{\partial(\delta\alpha)} \vdots \frac{\partial c_k}{\partial(\delta\beta)}\right] \in \mathcal{R}^{3\times 2} \tag{A.9}$$

$$H_\theta = \frac{\partial\theta_k}{\partial(\delta\theta)} \in \mathcal{R}^{3\times 3} \tag{A.10}$$

$$H_\phi = \left[\frac{\partial\phi_k}{\partial(\delta a_r)} \vdots \frac{\partial\phi_k}{\partial(\delta b_r)} \vdots \frac{\partial\phi_k}{\partial(\delta c_r)}\right] \in \mathcal{R}^{3\times 9} \tag{A.11}$$

$$H_g = \left[\frac{\partial g_k}{\partial(\delta b_g)} \vdots \frac{\partial g_k}{\partial(\delta c_g)}\right] \in \mathcal{R}^{3\times 6} \tag{A.12}$$

The quantities $H_c, H_\theta, H_\phi, H_g$ will be calculated next.

A.4.1 Scan Mirror Axis Sensitivity H_c

Consider the perturbation on the scan mirror rotation given in (4.41) as,

$$C = \left(I - c_k^\times\right)\hat{C} \tag{A.13}$$

Based on a result found in [3] (cf., Angle-Axis Perturbation), the perturbation c_k in (A.13) can be written as,

$$c_k = \left[\sin\hat{\beta}\Gamma_k \cdot I - \left(1 - \cos\hat{\beta}\Gamma_k\right)\hat{a}^\times\right]\delta a + \hat{a}\Gamma_k\delta\beta \tag{A.14}$$

where Γ_k is Γ at the k'th centroid time T_k; $\hat{\beta}$ is the nominal scale factor; and $\hat{a} \in \mathcal{R}^3$ is the nominal mirror axis. In general, the misalignment on the scan mirror axis $\delta a \in \mathcal{R}^3$ has two degrees of freedom corresponding to in-plane and out-of-plane errors. However, the out-of-plane error can be ignored because it manifests itself as a frame misalignment, and is estimated as part of T. The remaining degree of freedom is the in-plane misalignment which can be parametrized as,

$$\delta a = h_a\left(\hat{a}\right)\delta\alpha \tag{A.15}$$

where,

$$h_a(\hat{a}) = \begin{bmatrix} 0 \\ \hat{a}_3 \\ -\hat{a}_2 \end{bmatrix} \tag{A.16}$$

Substituting (A.15) into (A.14) gives,

$$c_k = \left[\left(\sin\hat{\beta}\Gamma_k \cdot I - \left(1 - \cos\hat{\beta}\Gamma_k\right)\hat{a}^\times\right) h_a(\hat{a}) \ \vdots \ \hat{a}\Gamma_k \right] \begin{bmatrix} \delta\alpha \\ \delta\beta \end{bmatrix} \tag{A.17}$$

Consequently, the desired partial derivative of c_k with respect to $\delta\alpha, \delta\beta$ is,

$$H_c = \left[\left(\sin\hat{\beta}\Gamma_k \cdot I - \left(1 - \cos\hat{\beta}\Gamma_k\right)\hat{a}^\times\right) h_a(\hat{a}) \ \vdots \ \hat{a}\Gamma_k \right] \tag{A.18}$$

A.4.2 IPF_0 Alignment Sensitivity H_θ

It is seen from (4.42) that the variables θ_k and $\delta\theta$ represent the same physical perturbation of T. Consequently, the partial $\frac{\partial\theta_k}{\partial(\delta\theta)}$ is a 3×3 identity matrix, i.e.,

$$H_\theta = I_{3\times3} \tag{A.19}$$

A.4.3 TPF Alignment Sensitivity H_ϕ

The direction cosine matrix R is the mapping from the STA-defined Body frame to the TPF frame. It is parametrized in (3.9) as the following quadratic function of time,

$$R(q_R, b_r, c_r) \triangleq \left(I_{3\times3} - \left(b_r t + \frac{c_r t^2}{2} \right)^\times \right) R_0(q_R) \tag{A.20}$$

where R_0 (and its equivalent quaternion q_R) corresponds to the initial alignment at time $t = 0$, and b_r, c_r are parameters associated with the time-varying alignment drift.

Let R_0, b_r, c_r be perturbed about their current nominal estimates $\hat{R}_0, \hat{b}_r, \hat{c}_r$ by the perturbations $\delta a_r, \delta b_r, \delta c_r \in \mathcal{R}^3$ according to,

$$R_0(q_R) = (I - \delta a_r^\times)\hat{R}_0 \tag{A.21}$$

$$b_r = \hat{b}_r + \delta b_r \tag{A.22}$$

$$c_r = \hat{c}_r + \delta c_r \tag{A.23}$$

Substituting the perturbations (A.21)-(A.23) into (A.20) and rearranging gives (to first order),

$$R(q_R, b_r, c_r) = \left(I - \left(\left(\hat{b}_r t + \frac{\hat{c}_r t^2}{2} \right) + \left(\delta b_r t + \frac{\delta c_r t^2}{2} \right) \right)^\times \right) \left(I - \delta a_r^\times \right) \hat{R}_0 \tag{A.24}$$

$$= \left[I - \left(\delta a_r + \delta b_r t + \frac{\delta c_r t^2}{2}\right)^\times\right]\left[I - \left(\hat{b}_r t + \frac{\hat{c}_r t^2}{2}\right)^\times\right]\hat{R}_0 \tag{A.25}$$

$$\triangleq \left[I - \phi(t)^\times\right] R(\hat{q}_R, \hat{b}_r, \hat{c}_r) \tag{A.26}$$

where,

$$R(\hat{q}_R, \hat{b}_r, \hat{c}_r) = \left[I - \left(\hat{b}_r t + \frac{\hat{c}_r t^2}{2}\right)^\times\right]\hat{R}_0 \tag{A.27}$$

$$\phi(t) = \begin{bmatrix} I & t \cdot I & \frac{t^2}{2} \cdot I \end{bmatrix} \begin{bmatrix} \delta a_r \\ \delta b_r \\ \delta c_r \end{bmatrix} \tag{A.28}$$

The variable ϕ_k is defined by evaluating $\phi(t)$ at the k centroid time T_k to give,

$$\phi_k = \begin{bmatrix} I & T_k \cdot I & \frac{T_k^2}{2} \cdot I \end{bmatrix} \begin{bmatrix} \delta a_r \\ \delta b_r \\ \delta c_r \end{bmatrix} \tag{A.29}$$

Hence, the desired sensitivity H_ϕ is,

$$H_\phi = \begin{bmatrix} I & T_k \cdot I & \frac{T_k^2}{2} \cdot I \end{bmatrix} \tag{A.30}$$

It is noted that after the perturbations $\delta a_r, \delta b_r, \delta c_r$ are estimated, they are applied to update the nominal parameters as follows,

$$\hat{R}_0 \leftarrow (I - \delta\hat{a}_r^\times)\hat{R}_0 \tag{A.31}$$

$$\hat{b}_r \leftarrow \hat{b}_r + \delta\hat{b}_r \tag{A.32}$$

$$\hat{c}_r \leftarrow \hat{c}_r + \delta\hat{c}_r \tag{A.33}$$

In this manner the estimate of the initial alignment $\hat{R}_0$ (and its equivalent quaternion $\hat{q}_R$) is kept as a large angle quantity, while the time-varying drift is kept as a small angle. This is appropriate since time-variations are expected to be only a few arcseconds in size.

A.4.4 Gyro Attitude Offset Sensitivity H_g

The true gyro offset G at centroid time T_k can be written as,

$$G(T_k) = \left(I - \gamma(T_k)^\times\right) G^\circ(T_k) \tag{A.34}$$

where G° is a nominal gyro offset provided by the gyro pre-processor. The quantity G° is computed by the gyro pre-processor by using a certain nominal rate estimate ω_m°. (e.g., ω_m° is taken as the first on-board correction, or some other reasonable estimate [3]).

Clearly the quantity γ is not known. However, given that the true rate ω is related to the approximate rate ω_m° according to,

$$\omega = \omega_m^\circ + b_g + c_g \tag{A.35}$$

(by definition of b_g and c_g in (3.12)) then as shown in [3], the quantity γ can be parametrized linearly in b_g, c_b as,

$$\gamma\left(T_k\right) = H_g\left(T_k\right)\left[\begin{array}{c} b_g \\ c_g \end{array}\right] \tag{A.36}$$

$$H_g\left(T_k\right) = \left[\begin{array}{cc} \Lambda_b\left(T_k\right) & \Lambda_c\left(T_k\right) \end{array}\right] \tag{A.37}$$

where the quantities $\Lambda_b(T_k)$ and $\Lambda_c(T_k)$ are obtained by integrating the matrix differential equations,

$$\left(\dot{\Lambda}_b = -\left(\omega_m^\circ\right)^\times \Lambda_b + I\right)\Big|_{t_j}^{T_k(j)} \quad with \;\; I.C. \quad \Lambda_b\left(t_j\right) = 0 \tag{A.38}$$

$$\left(\dot{\Lambda}_c = -\left(\omega_m^\circ\right)^\times \Lambda_c + t \cdot I\right)\Big|_{t_j}^{T_k(j)} \quad with \;\; I.C. \quad \Lambda_c\left(t_j\right) = 0 \tag{A.39}$$

Here, the notation $(\cdot)|_a^b$ denotes integration over the time interval $t \in [a, b]$; t_j denotes the starting time of the j'th maneuver; and T_k denotes the k'th centroid time.

Since the b_g, c_g parameters are not known, perturbations $\delta b_g, \delta c_g$ are defined such that,

$$b_g = \hat{b}_g + \delta b_g \tag{A.40}$$

$$c_g = \hat{c}_g + \delta c_g \tag{A.41}$$

Using the current nominal estimates $\hat{b}_g, \hat{c}_g$, an estimate $\hat{\gamma}$ of γ at the centroid time T_k is defined by,

$$\hat{\gamma}\left(T_k\right) \triangleq H_g\left(T_k\right)\left[\begin{array}{c} \hat{b}_g \\ \hat{c}_g \end{array}\right] \tag{A.42}$$

A corresponding estimate of G is defined by,

$$\hat{G}\left(T_k\right) \triangleq \left(I - \hat{\gamma}\left(T_k\right)^\times\right) G^\circ\left(T_k\right) \tag{A.43}$$

This estimate is used in the measurement equation to form the prediction. Combining (A.34) and (A.43) yields (to first order),

$$G(T_k) = \left(I - \left(\gamma(T_k) - \hat{\gamma}(T_k)\right)^\times\right)\hat{G} \tag{A.44}$$

By comparing (A.44) with the definition of g_k in (4.44) it is clear that the relation between g_k and γ is given (to first order) as,

$$g_k \triangleq \gamma(T_k) - \hat{\gamma}(T_k) = H_g\left(T_k\right)\left[\begin{array}{c} \delta b_g \\ \delta c_g \end{array}\right] \tag{A.45}$$

where (A.36) has been subtracted from (A.42) to get the last relation. It is seen from (A.45) that H_g is the desired sensitivity function and its formula is given by (A.37).

For computational savings, the gyro pre-processor computes the quantities $\{G^\circ(T_k), H_g(T_k)\}$ only once (by numerical integration of (A.38) and (A.38)) and stores the results. These quantities are then used in the prediction equation (A.43) and the sensitivity equation (A.45) at each centroid time T_k and during each filter cycle. The gyro pre-processor is discussed in more detail in [3].

B CONCLUSIONS

The IPF Kalman filter algorithm has been described, which has been specifically designed for calibrating the SIRTF telescope focal plane. The main novelty of the filter design lies in its ability to handle a large variety of array types (cameras, arrays with scanning mirrors, spectroscopy slits, etc.) in a single formulation, and in its combining both science array and attitude parameters in the same filter. The IPF filter has many features which are not covered here due to space limitations, but are described in the algorithm document [3]. The IPF filter will be executed over 60 times on separate data sets during SIRTF's In-Orbit Checkout period, where it will be used to calibrate over 1200 hundred parameters and over 60 different Prime and Inferred frames. These frames and calibration parameters will play a key role in supporting SIRTF's in-flight precision pointing capability, and for ground pointing reconstruction efforts.

Acknowledgements

This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

[1] D.S. Bayard, "Spacecraft pointing control criteria for high resolution spectroscopy," IEEE Transactions on Aerospace and Electronic Systems, Vol. 35, No. 2, pp. 637-644, April 1999.

[2] D.S. Bayard, B.H. Kang, *SIRTF Instrument Pointing Frame (IPF) Kalman Filter Unit Test Report.* JPL Internal Document, D-24810, October 31, 2002.

[3] D.S. Bayard and B.H. Kang, *SIRTF Instrument Pointing Frame (IPF) Kalman Filter Algorithm*, JPL Internal Document, D-24809, April 30, 2002.

[4] G.J. Bierman and M.D. Schuster, "Spacecraft alignment estimation," 27th IEEE Conference on Decision and Control, Austin, TX, Dec. 7-9, 1988.

[5] R.J. Brown, *Focal Surface to Object Space Field of View Conversion.* Systems Engineering Report SER No. S20447-OPT-051, Ball Aerospace& Technologies Corp., November 23, 1999.

[6] S. Davis and J. Lai, "Attitude sensor calibration for the Ocean Topography Experiment Satellite (TOPEX)," SPIE Space Guidance, Control and Tracking Conference, Vol. 1949, Society of Photo-Optical Instrumentation Engineers, pp. 80-91, 1993.

[7] J. Deutschmann and I.Y. Bar-Itzhack, "Extended Kalman Filter for attitude estimation of the Earth Radiation Budget Satellite," Proc. Flight Mechanics/Estimation Theory Symposium, NASA Goddard Space Flight Center, Greenbelt, Maryland, pp. 319-332, May 10-11, 1989.

[8] A. Gelb (Ed.), *Applied Optimal Estimation.* The MIT Press, Cambridge, Massachusetts, 1984.

[9] T. Kailath, A.H. Sayed, B. Hassibi, *Linear Estimation.* Prentice Hall,Upper Saddle River, New Jersey, 1999.

[10] T. Kia, D.S. Bayard, F. Tolivar, "A precision pointing control system for the Space Infrared Telescope Facility (SIRTF)," American Astronautical Society, AAS97-067, 1997.

[11] J.W. Murrell, "Precision attitude determination for multimission spacecraft," Paper No. 78-1248, Proc. AIAA Guidance and Control Conference, Palo Alto, California, August 7-9, 1978.

[12] H. Neudecker, "Some theorems on matrix differentiation with special reference to Kronecker matrix products," J. of Amer. Stat. Assoc., Vol. 64, pp.953-963, 1969

[13] M. Pittelkau, "Kalman filtering for spacecraft system alignment calibration," J. Guidance, Control and Dynamics, Vol. 24, No. 6, November-December 2001.

[14] J. Schwenker, *Focal Surface to Object Space Field of View Conversion, Review and Simplified Math Model.* Systems Engineering Report SER No. S20447-OPT-057, Ball Aerospace & Technologies Corp., September 9, 2000.

[15] P.K. Seidelmann, *Explanatory Supplement to the Astronomical Almanac.* University Science Books, Sausalito, California, 1992.

[16] M.D. Shuster, "Inflight Estimation of Spacecraft Sensor Alignment," Advances in the Astronautical Sciences, 72, 1990.

[17] M.D. Shuster and R.V.F. Lopes, "Parameter interference in distortion and alignment calibration, AAS Paper AAS 94-186, Feb. 1994.

[18] R.P. Wishner, J.A. Tabaczynski, M. Athans, "A Comparison of Three Non-Linear Filters," Automatica, Vol. 5, pp.487-496, Pergamon Press, 1969.

[19] E. Wong, W. Breckenridge, D. Boussalis, P. Brugarolas, D.S. Bayard, J. Spanos, G. Singh, "Post-Flight Attitude Reconstruction for the Shuttle Radar Topography Mission," Proc. AAS/AIAA Astrodynamics Specialists Conference, Paper AAS 01-314, Quebec City, Quebec, Canada, July 30, 2001.

[20] *Space Infrared Telescope Facility In-Orbit Checkout and Science Verification Mission Plan.* JPL Internal Document, D-22622, 674-FE-301 Version 1.4, February 8, 2002.