

Optimal Estimation of Calibration Parameters in Polarimetric Microwave Radiometers

Derek Hudson and David G. Long, *Fellow, IEEE*

Abstract—Methods for internal calibration of a certain class of microwave polarimetric radiometers are presented by Piepmeier. In that work, the calibration parameters are estimated algebraically. We demonstrate that Bayesian estimation decreases the root-mean-square error of the estimates by a factor of two. This improvement is obtained by using knowledge of the noise structure of the measurements and by utilizing all of the information provided by the measurements. Drawbacks are the increased complexity of the method and an increase in computation. We also extend the method to estimate several hardware component parameters of interest in system calibration.

Index Terms—Polarimetric radiometer, radiometer calibration, Stokes parameters.

I. INTRODUCTION

RADIOMETER calibration is the process of estimating radiometer channel gains and the internal noise (represented by an equivalent noise temperature) generated by the radiometer. Because the gains and noise temperatures can change rapidly during operation, internal calibration is performed frequently during radiometer operation. Internal calibration is accomplished by applying “known” inputs and measuring the voltage outputs of the radiometer. For the calibration scheme considered in this paper, the known inputs are internal sources (hot and cold sources of temperature T_H and T_C) rather than external targets. For this calibration scheme, the estimation of antenna gains is a separate process and, hence, is not addressed in this paper.

Internal calibration of radiometers which measure the third Stokes parameter (T_U) can be accomplished using an additional “known” input T_{CN} , which is split and fed into both the vertical and horizontal channels so that the fluctuations in the electric fields of the two channels are correlated, simulating a third Stokes parameter input (see [1, Fig. 1]). This technique is introduced in [1] for microwave radiometers which use a hybrid coupler to synthesize $\pm 45^\circ$ linear polarizations from vertical and horizontal signals. The noise-free forward model for the calibration measurements of such a radiometer is given in Section II. The algebraic method of [1] for estimating channel

Manuscript received April 10, 2007; revised November 21, 2007. Current version published October 1, 2008. This work was supported by the NASA Goddard Space Flight Center.

D. Hudson is with Brigham Young University, Provo, UT 84602 USA, and also with the NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA (e-mail: Derek.L.Hudson@nasa.gov).

D. G. Long is with the Department of Electrical and Computer Engineering, Brigham Young University, Provo, UT 84602 USA (e-mail: long@ee.byu.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2008.921636

gains and temperatures using this noise-free forward model is summarized in Section III.

In the remainder of this paper, we improve on the method of [1]. In Section IV, we expand the forward model to include the noise in the measurements and then solve the calibration problem using Bayes’ theorem rather than algebraically as was done in [1]. This results in a joint probability distribution function (pdf) for the calibration parameters.

This joint pdf itself is the most complete answer to the calibration problem and is explored in Section VI. In Section V, we compare numerical estimates extracted from this pdf with the algebraic estimates of [1]. We show that our estimates are optimal in the sense of minimizing the root-mean-square error (rmse). They are unbiased, and their rmse is approximately half the rmse of the algebraic estimates.

Section VII explores an extension. We use Bayes’ theorem to find the first ever estimates and pdfs for the hardware parameters that comprise the end-to-end radiometer channel gains. Finally, conclusions are offered in Section VIII.

II. CALIBRATION FORWARD PROBLEM

The noise-free forward model for a single cycle of the full (Case 4) polarimetric radiometer calibration algorithm described in [1] can be written as

$$\begin{bmatrix} v_{v,C} & v_{v,H} & v_{v,CH} & v_{v,CN} \\ v_{h,C} & v_{h,H} & v_{h,CH} & v_{h,CN} \\ v_{p,C} & v_{p,H} & v_{p,CH} & v_{p,CN} \\ v_{m,C} & v_{m,H} & v_{m,CH} & v_{m,CN} \end{bmatrix} = \begin{bmatrix} G_{vv} & 0 & 0 \\ 0 & G_{hh} & 0 \\ G_{pv} & G_{ph} & G_{pU} \\ G_{mv} & G_{mh} & G_{mU} \end{bmatrix} \times \begin{bmatrix} T_C + T_1 & T_H + T_1 & T_C + T_1 & T_C + \frac{1}{2}T_{CN} + T_1 \\ T_C + T_2 & T_H + T_2 & T_H + T_2 & T_C + \frac{1}{2}T_{CN} + T_2 \\ 0 & 0 & 0 & \pm T_{CN} \end{bmatrix}. \quad (1)$$

This forward model is in [1, eq. (3)], with o_v , o_h , o_p , and o_m being given in [1, eq. (8)] and with the vector $[T_v \ T_h \ T_U]^T$ being replaced by a matrix.

On the left side are the 16 voltages which are measured in one radiometer calibration cycle: For each of the four calibration looks (cold load, hot load, mixed load, and cold load plus correlated noise load), the voltage outputs of the four polarimetric channels (v , h , p , and m) are measured.

On the right side of (1) are ten calibration parameters which are unknown to some degree: eight radiometer gains G_{xx} plus two receiver noise temperatures (mostly determined by the noise figures of the first-stage amplifiers) T_1 and T_2 .

The four columns of the rightmost matrix in (1) represent the brightness temperature inputs used in the four calibration

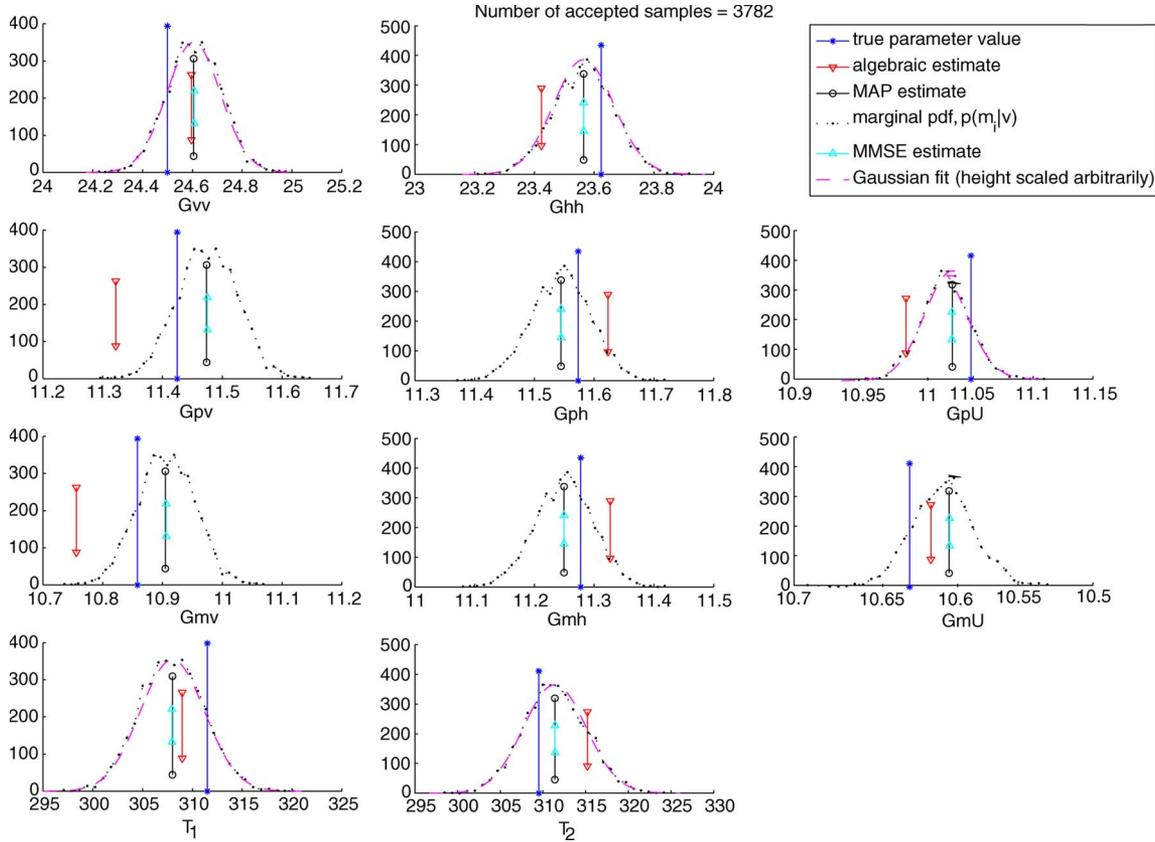


Fig. 1. Marginal pdfs for the eight radiometer gains and two noise temperatures, along with various estimates of the parameters. The height scale is arbitrary, being the number of samples in each bin. As we use more samples to generate the empirical pdfs, the empirical pdfs match the Gaussian fits more closely. The “true” parameter values used for this simulation were randomized, rather than the typical values in Table I.

subcycles. By rows, the top row has temperature inputs to the vertically polarized (v) channel, the second row has inputs to the horizontally polarized (h) channel, and the third row has third Stokes parameter inputs (representing correlation between the vertically and horizontally polarized signals T_U) (see [1, Fig. 1]).

For this paper, we pretend that T_C , T_H , and T_{CN} are perfectly known. Optimal estimation of them from thermometer measurements, on-ground calibration, and the voltages on the left side of (1) is left for future work.

Various expansions of this model are possible. An additional column, the vector $[T_v + T_1 \quad T_h + T_2 \quad T_U]^T$, could be added to the right side, and a column of corresponding voltages could also be added to the left side. This vector contains the T_v , T_h , and T_U brightnesses of the scene under observation, whose estimation is the goal of radiometry. We have left off these columns in order to simplify the problem, focusing only on the estimation of the calibration parameters. However, this paper can readily be extended to the larger problem. Other possible extensions include the following: 1) allowing for nonzero gains where there are zeros (although approximating them as zero is fairly accurate); 2) adding a small T_U term, generated by the radiometer, to the last row of the temperature matrix; 3) jointly estimating many consecutive sets of the calibration parameters, exploiting the correlation between them to improve the estimates; and 4) adapting the method expounded in this paper to other classes of radiometers by using their forward models in place of (1).

III. ALGEBRAIC ESTIMATION (PIEPMEIER’S METHOD [1])

Estimating the ten calibration parameters on the right side of (1) by the method of [1] can be summarized as follows. It is algebraic, making no attempt to model the noise in the measurements.

G_{vv} is estimated using $v_{v,C}$ and $v_{v,H}$, which is the conventional cold/hot calibration method of nonpolarimetric radiometers. From (1), the equations for these two voltages have the same form as in [1, eq. (20)], viz.,

$$\begin{bmatrix} v_{v,C} \\ v_{v,H} \end{bmatrix} = \begin{bmatrix} T_C & 1 \\ T_H & 1 \end{bmatrix} \begin{bmatrix} G_{vv} \\ G_{vv}T_1 \end{bmatrix}. \quad (2)$$

Solving for G_{vv} , the estimate is

$$\hat{G}_{vv} = \frac{v_{v,H} - v_{v,C}}{T_H - T_C} \quad (3)$$

as in [1, eq. (21)]. T_1 is estimated by solving the same equations for T_1 , yielding

$$\hat{T}_1 = \frac{-T_C v_{v,H} + T_H v_{v,C}}{v_{v,H} - v_{v,C}}. \quad (4)$$

Estimation of G_{hh} and T_2 from $v_{h,C}$ and $v_{h,H}$ is similar.

The gains G_{pv} , G_{ph} , and G_{pU} are found in [1, eq. (41)], which are reproduced here (with o_p being expanded and the minor correction that G_{pm} is replaced by G_{ph})

$$\begin{bmatrix} v_{p,C} \\ v_{p,H} \\ v_{p,CH} \\ v_{p,CN} \end{bmatrix} = \begin{bmatrix} T_C & T_C & 0 & 1 \\ T_H & T_H & 0 & 1 \\ T_C & T_H & 0 & 1 \\ T_C + \frac{1}{2}T_{CN} & T_C + \frac{1}{2}T_{CN} & \pm T_{CN} & 1 \end{bmatrix} \times \begin{bmatrix} G_{pv} \\ G_{ph} \\ G_{pU} \\ G_{pv}T_1 + G_{ph}T_2 \end{bmatrix}. \quad (5)$$

Note that this also follows from (1). To obtain G_{pv} , G_{ph} , and G_{pU} , both sides of (5) are multiplied on the left by the inverse of the 4×4 matrix in (5). G_{mv} , G_{mh} , and G_{mU} are found similarly, using the measurements $v_{m,C}$, $v_{m,H}$, $v_{m,CH}$, and $v_{m,CN}$.

Note that both the 4×4 matrix in (5) and its inverse have a condition number (for the two-norm) of about 3000 when we use the following load temperatures typical of NASA’s near-future Aquarius radiometer: $T_C = 288$ K and $T_H = T_{CN} = 800$ K [2]. This suggests that noise in the voltage measurements can disturb the estimates significantly. This may explain why our optimal estimates of the last six gains are the most improved compared with algebraic estimates (see Table II, as will be described later). Also, note that four of the available measurements, $v_{v,CH}$, $v_{v,CN}$, $v_{h,CH}$, and $v_{h,CN}$, are not used in this method. Nevertheless, the method of Piepmeier [1] works well (evidence for this is given by the first row of numbers in Table II, as will be described later).

IV. BAYESIAN ESTIMATION

The algebraic estimation method sketched in the previous section is not unique. For example, G_{vv} could also be estimated using

$$\hat{G}_{vv} = \frac{v_{v,H} - v_{v,CH}}{T_H - T_C}. \quad (6)$$

A better estimate would be an average of (3) and (6) because the noise in each estimate is somewhat different and is therefore reduced by averaging. Information on G_{vv} is also contained in the measurement $v_{v,CN}$, and similarly with the other nine calibration parameters. How can we combine all of the information in the voltages to get the most accurate estimates of the ten calibration parameters? The answer is to approach the problem using probability theory rather than algebraically. By using Bayes’ theorem, a pdf can express all

the information on a parameter that is available from various measurements, from the noise-free forward model, and from a probabilistic description of the noise [3]. Let \mathbf{v} be the vector of measurements (voltages) in (1), and let \mathbf{m} be the vector of calibration parameters (the eight gains plus T_1 and T_2). Then, Bayes’ theorem tells us that the pdf for the parameters, given the voltages, is

$$p(\mathbf{m}|\mathbf{v}) = \frac{p(\mathbf{v}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{v})}. \quad (7)$$

The pdf $p(\mathbf{m})$ represents prior or external information on the model parameters, and similarly with $p(\mathbf{v})$. In this paper, we assume that no prior or external information is available. The absence of information can be represented by a constant pdf, so that $p(\mathbf{m})$ and $p(\mathbf{v})$ are both constants. Then, $p(\mathbf{m})/p(\mathbf{v})$ is another constant, c' , and

$$p(\mathbf{m}|\mathbf{v}) = c'p(\mathbf{v}|\mathbf{m}). \quad (8)$$

The remainder of this section finds $p(\mathbf{m}|\mathbf{v})$ explicitly.

A. Pdf for the Voltages, Given the Parameters, $p(\mathbf{v}|\mathbf{m})$

We first elaborate the probability distribution for voltages, given a set of calibration parameters, $p(\mathbf{v}|\mathbf{m})$.

1) *Noise Model:* Equation (1) is a noise-free forward model for the 16 voltages from the calibration parameters. However, the temperature inputs in the last matrix of (1) are only mean values. Actual thermal emissions fluctuate. These random fluctuations can be treated as Gaussian noise (commonly called $NE\Delta T$) that is added to each “true” (mean-value) temperature in the forward model. Therefore, the forward model with noise included is (9), shown at the bottom of the page, where the n_i terms are added noise.

The noise term n_1 represents the total fluctuation away from $T_C + T_1$ during the first calibration subcycle. It is zero-mean Gaussian noise with a standard deviation (STD) that is equal to the mean temperature divided by the root of the time bandwidth product, $(T_C + T_1)/\sqrt{B\tau_c}$, where B is the sensor bandwidth and τ_c is the calibration integration time [4, eq. (6.51)]. $n_2 - n_9$ are similar.

The noise terms $n_1 - n_6$ are all independent of one another (and of $n_7 - n_9$) for one or both of the following reasons: 1) They originate from different sources (the v -channel hot and cold inputs, as well as the amplifiers producing T_1 , are separate from the h -channel inputs and the amplifiers producing T_2 , see [1, Fig. 1]), and 2) they are realized during different calibration subcycles (i.e., are different realizations of the noise, and the rapidity of the fluctuations means that the realization

$$\begin{bmatrix} v_{v,C} & v_{v,H} & v_{v,CH} & v_{v,CN} \\ v_{h,C} & v_{h,H} & v_{h,CH} & v_{h,CN} \\ v_{p,C} & v_{p,H} & v_{p,CH} & v_{p,CN} \\ v_{m,C} & v_{m,H} & v_{m,CH} & v_{m,CN} \end{bmatrix} = \begin{bmatrix} G_{vv} & 0 & 0 \\ 0 & G_{hh} & 0 \\ G_{pv} & G_{ph} & G_{pU} \\ G_{mv} & G_{mh} & G_{mU} \end{bmatrix} \cdot \begin{bmatrix} T_C + T_1 + n_1 & T_H + T_1 + n_3 & T_C + T_1 + n_5 & T_C + T_1 + \frac{1}{2}T_{CN} + n_7 \\ T_C + T_2 + n_2 & T_H + T_2 + n_4 & T_H + T_2 + n_6 & T_C + T_2 + \frac{1}{2}T_{CN} + n_8 \\ 0 & 0 & 0 & T_{CN} + n_9 \end{bmatrix} \quad (9)$$

during one interval is independent of the realization during the next interval [5]). The noise terms n_7 – n_9 are not independent of one another because they all originate (at least in part) from the correlated calibration source; they are treated in Appendix A2.

Even though n_1 – n_6 are independent of one another, there is correlation among the voltages on the left side of (9). For example, $v_{p,C}$ and $v_{m,C}$ are both functions of n_1 and n_2 and are therefore correlated. All the correlations that exist among the voltages can be summarized in a covariance matrix C that is derived in Appendix A. As a result, the probability distribution for \mathbf{v} , given \mathbf{m} , is a 16-dimensional Gaussian pdf, with mean being given by the right-hand side of (1) denoted $\mathbf{g}(\mathbf{m})$,

$$p(\mathbf{v}|\mathbf{m}) = \frac{1}{\sqrt{(2\pi)^{16}|C|}} e^{-\frac{1}{2}(\mathbf{v}-\mathbf{g}(\mathbf{m}))^T C^{-1}(\mathbf{v}-\mathbf{g}(\mathbf{m}))}. \quad (10)$$

2) *Eigendecomposition of Singular C*: The covariance matrix C is a function of the calibration parameters \mathbf{m} . Numerical calculations using arbitrary values for \mathbf{m} show that, although C is 16×16 , its rank r is consistently only nine. From a theoretical standpoint, this corresponds to the nine noise sources on the right-hand side of (9).

Because C is not full rank, it cannot be inverted. It also has $16 - r = 7$ eigenvalues with value zero so that $|C| = 0$. How do we evaluate (10) in this situation?

Consider the eigendecomposition of C

$$C = [VDV^T] = [V_1 \ V_2] \text{diag}([\mathbf{\Lambda} \ 0]) \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \quad (11)$$

where the $r = 9$ columns of V_1 are eigenvectors with nonzero eigenvalues $\mathbf{\Lambda}$, whereas the seven columns of V_2 are eigenvectors with eigenvalues that are equal to zero, and $\text{diag}(\cdot)$ is a diagonal matrix with the elements of the argument along its diagonal. Matrix theory tells us that, because C is symmetric, $V \equiv [V_1 \ V_2]$ is unitary (because V is real, it is also an orthogonal matrix). Therefore, $V^T \mathbf{v} = \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \mathbf{v}$ is a rotation of the voltages.

Consider a partition of these rotated voltages into $V_1^T \mathbf{v}$ and $V_2^T \mathbf{v}$. The covariance of $V_2^T \mathbf{v}$ is

$$\text{Cov}(V_2^T \mathbf{v}) = \left\langle (V_2^T \mathbf{v} - \langle V_2^T \mathbf{v} \rangle) (V_2^T \mathbf{v} - \langle V_2^T \mathbf{v} \rangle)^T \right\rangle \quad (12)$$

$$= V_2^T \left\langle (\mathbf{v} - \langle \mathbf{v} \rangle) (\mathbf{v} - \langle \mathbf{v} \rangle)^T \right\rangle V_2 \quad (13)$$

$$= V_2^T \text{Cov}(\mathbf{v}) V_2 \quad (14)$$

$$= V_2^T C V_2 = [\mathbf{0}] \quad (15)$$

where the last step follows from the fact that the columns of V_2 are eigenvectors of C with eigenvalues that are equal to zero. Because the variances of the elements of $V_2^T \mathbf{v}$ (diagonal elements of $\text{Cov}(V_2^T \mathbf{v})$) are zero, $V_2^T \mathbf{v}$ is a constant vector.

That is, for a given \mathbf{m} , $V_2^T \mathbf{v}$ will be the same for any realization of the voltages.

With \mathbf{m} being given, we can find the particular $\mathbf{v} = \mathbf{g}(\mathbf{m})$ corresponding to the voltages obtained when all nine noise sources happen to be zero. Then, because $V_2^T \mathbf{v}$ is constant for a given \mathbf{m} as just shown, we can write $V_2^T \mathbf{v} = V_2^T \mathbf{g}(\mathbf{m})$ for any realization of \mathbf{v} .

This last expression is a constraint when evaluating $p(\mathbf{v}|\mathbf{m})$. It says that the pdf is concentrated on a manifold (in the 16-dimensional space of voltages) defined by $V_2^T \mathbf{v} = V_2^T \mathbf{g}(\mathbf{m})$. For any \mathbf{v} which does not satisfy this constraint (i.e., does not lie on the manifold), $p(\mathbf{v}|\mathbf{m}) = 0$.

Next, consider the rotated voltages $V_1^T \mathbf{v}$. By the same steps as in (12)–(15), we find that

$$\begin{aligned} \text{Cov}(V_1^T \mathbf{v}) &= V_1^T C V_1 = V_1^T [V_1 \ V_2] \text{diag}([\mathbf{\Lambda} \ 0]) \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} V_1 \\ &= V_1^T V_1 \text{diag}(\mathbf{\Lambda}) V_1^T V_1 = \text{diag}(\mathbf{\Lambda}). \end{aligned} \quad (16)$$

This covariance matrix is invertible, and its determinant is the product of the elements of $\mathbf{\Lambda}$. We can eliminate the singularities in (10) by reducing the dimensionality of the Gaussian pdf $p(\mathbf{v}|\mathbf{m})$ from 16 to $r = 9$, using $V_1^T \mathbf{v}$ and its nonsingular covariance matrix in place of \mathbf{v} and C (also, note that the mean of $V_1^T \mathbf{v}$ is $V_1^T \mathbf{g}(\mathbf{m})$).

In summary, to deal with the singularities in (10), we rotate and partition our measurements \mathbf{v} into measurements ($V_2^T \mathbf{v}$) which provide a constraint and those ($V_1^T \mathbf{v}$) which provide a smaller dimensional pdf, so that (10) can be rewritten as (17), shown at the bottom of the page.

In Appendix B, we derive V_2 analytically and show that the constraint $V_2^T \mathbf{v} = V_2^T \mathbf{g}(\mathbf{m})$ reduces to

$$\begin{aligned} G_{pv} &= G_{vv} \frac{(v_{p,C} v_{h,H} - v_{h,C} v_{p,H})}{(v_{v,C} v_{h,H} - v_{h,C} v_{v,H})} \\ G_{ph} &= G_{hh} \frac{(v_{p,C} v_{v,H} - v_{v,C} v_{p,H})}{(v_{h,C} v_{v,H} - v_{v,C} v_{h,H})} \end{aligned} \quad (18)$$

$$\begin{aligned} G_{mv} &= G_{vv} \frac{(v_{m,C} v_{h,H} - v_{h,C} v_{m,H})}{(v_{v,C} v_{h,H} - v_{h,C} v_{v,H})} \\ G_{mh} &= G_{hh} \frac{(v_{m,C} v_{v,H} - v_{v,C} v_{m,H})}{(v_{h,C} v_{v,H} - v_{v,C} v_{h,H})} \end{aligned} \quad (19)$$

$$G_{mU} = G_{pU}$$

$$\frac{G_{mv} G_{hh} v_{v,CN} + G_{mh} G_{vv} v_{h,CN} - G_{vv} G_{hh} v_{m,CN}}{G_{pv} G_{hh} v_{v,CN} + G_{ph} G_{vv} v_{h,CN} - G_{vv} G_{hh} v_{p,CN}}. \quad (20)$$

$$p(\mathbf{v}|\mathbf{m}) = \begin{cases} \frac{1}{\sqrt{(2\pi)^9 |\text{diag}(\mathbf{\Lambda})|}} e^{-\frac{1}{2} [V_1^T \mathbf{v} - V_1^T \mathbf{g}(\mathbf{m})]^T \cdot \text{diag}(\mathbf{\Lambda})^{-1} \cdot [V_1^T \mathbf{v} - V_1^T \mathbf{g}(\mathbf{m})]}, & \text{if } V_2^T \mathbf{v} = V_2^T \mathbf{g}(\mathbf{m}) \\ 0, & \text{if } V_2^T \mathbf{v} \neq V_2^T \mathbf{g}(\mathbf{m}) \end{cases} \quad (17)$$

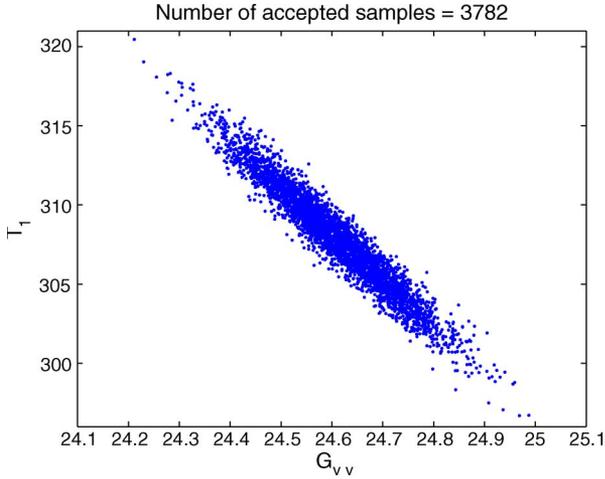


Fig. 2. Dots showing the G_{vv} and T_1 coordinates of samples of $p(\mathbf{m}|\mathbf{v})$. The density of dots illustrates the 2-D joint pdf $p(G_{vv}, T_1|\mathbf{v})$. The samples are from the same simulation that produced Fig. 1.

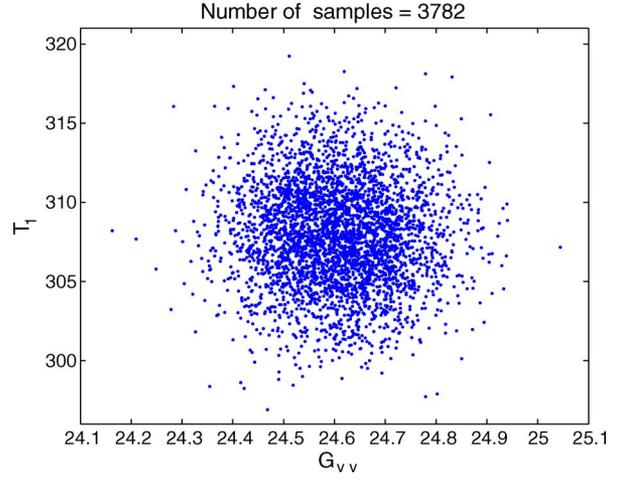


Fig. 3. Depiction of the 2-D joint pdf implied by reporting only marginal means and variances—using the means and variances of the samples in Fig. 2.

B. Posterior Pdf $p(\mathbf{m}|\mathbf{v})$

We are now ready to present a key result of this paper, $p(\mathbf{m}|\mathbf{v})$, which is the pdf for the calibration parameters, given a set of measured voltages. As shown in (8), we obtain it from (17) by simply reversing the roles of input and output and multiplying by a normalizing constant (we will not attempt to find the constant because it is not necessary for finding estimates and because the pdf can be displayed in unnormalized form). Explicitly, $p(\mathbf{m}|\mathbf{v})$ is given in (21), shown at the bottom of the page, where we show that V_1 and Λ are functions of the unknown \mathbf{m} (through C) and where c is an unknown constant.

V. MAP ESTIMATION

A. Theory

In general, the most complete answer to an estimation problem is a joint pdf on the variables of interest [6], in our case, (21). A joint pdf usually contains much more information than merely reporting numbers and STDs for the variables (this is illustrated in Section VI-B2). The increase in computing power over the last few decades enables us to begin the use of pdfs, such as $p(\mathbf{m}|\mathbf{v})$, as inputs and outputs to algorithms, rather than simple estimates and their uncertainties. It is our hope that science and engineering will move in that direction. In this section, however, we follow the tradition by reporting simple numerical estimates and their uncertainties (rmse).

From the joint pdf $p(\mathbf{m}|\mathbf{v})$, what should we extract and report as estimates of the calibration parameters? Of all possible estimates of \mathbf{m} from \mathbf{v} , the minimum-mean-square-error (mmse) estimate is the mean of $p(\mathbf{m}|\mathbf{v})$, i.e., its expected value with respect to \mathbf{m} [7].

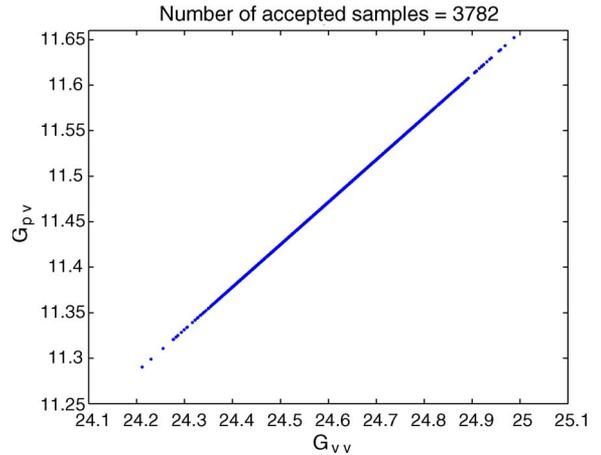


Fig. 4. Dots showing the G_{vv} and G_{pv} coordinates of samples of $p(\mathbf{m}|\mathbf{v})$. The density of dots illustrates the 2-D joint pdf $p(G_{vv}, G_{pv}|\mathbf{v})$.

Finding the mean of $p(\mathbf{m}|\mathbf{v})$ is difficult, but a shortcut is available. Marginal 1- and 2-D pdfs for one or two of the parameters are obtained by integrating $p(\mathbf{m}|\mathbf{v})$ with respect to the remaining parameters. Typical examples of these pdfs are shown in Figs. 1–4 (the generation of such figures is explained in Section VI). From such examples, it appears that $p(\mathbf{m}|\mathbf{v})$ is unimodal and symmetric in most (if not all) cases. This is confirmed by the Gaussian structure of (21). These properties signify that the mean of $p(\mathbf{m}|\mathbf{v})$ is the same as its mode (at least as a good approximation if not exactly). Finding the mode is easier than finding the mean; therefore, we use the mode as our primary estimate. We examine the properties of the mode in detail in this section and return to consider the mean in Section VI-B3.

$$p(\mathbf{m}|\mathbf{v}) = \begin{cases} \frac{c}{|\text{diag}(\Lambda(\mathbf{m}))|^{\frac{1}{2}}} e^{-\frac{1}{2} [V_1^T(\mathbf{m})\mathbf{v} - V_1^T(\mathbf{m})\mathbf{g}(\mathbf{m})]^T \cdot \text{diag}(\Lambda(\mathbf{m}))^{-1} \cdot [V_1^T(\mathbf{m})\mathbf{v} - V_1^T(\mathbf{m})\mathbf{g}(\mathbf{m})]}, & \text{if (18)–(20) are satisfied} \\ 0, & \text{otherwise} \end{cases} \quad (21)$$

TABLE I
GAIN-RELATED PARAMETERS TYPICAL OF AN L-BAND
RADIOMETER AND USED IN SIMULATIONS

Radiometer hardware parameters	
detector sensitivities, $c_v, c_h, c_p,$ and c_m	450 V/W
channel amplifier gain, G_1	1.8×10^7 W/W
gain imbalance, g	1.585
scattering parameter, s	0.7
α_e	0.934
bandwidth, B	20 MHz
Resulting gains	
G_{vv}	2.24×10^{-6} V/K
G_{hh}	3.55×10^{-6} V/K
G_{pv}	1.10×10^{-6} V/K
G_{ph}	1.81×10^{-6} V/K
G_{pU}	1.31×10^{-6} V/K
G_{mv}	1.14×10^{-6} V/K
G_{mh}	1.74×10^{-6} V/K
G_{mU}	-1.31×10^{-6} V/K

$p(\mathbf{m}|\mathbf{v})$ is often called the posterior distribution because it is the distribution for \mathbf{m} after measuring data \mathbf{v} . The mode of $p(\mathbf{m}|\mathbf{v})$ is the set of parameters \mathbf{m} , which maximizes the posterior distribution. Therefore, it is referred to as the maximum *a posteriori* (MAP) estimate. Because our $p(\mathbf{m}|\mathbf{v})$ is equal to a constant times $p(\mathbf{v}|\mathbf{m})$ (due to the absence of prior information), the MAP estimate is equivalent to the maximum-likelihood estimate, which is obtained by considering $p(\mathbf{v}|\mathbf{m})$ to be a function of \mathbf{m} and finding the \mathbf{m} which maximizes it.

Finding the value of \mathbf{m} which maximizes $p(\mathbf{m}|\mathbf{v})$ can be cast as a standard multidimensional optimization problem. It is readily accomplished by a blackbox minimization algorithm such as MATLAB's `fminsearch` (although more advanced techniques could find it with less computation). The search can be initialized using the algebraic estimate of \mathbf{m} (see Section III). To limit the search to \mathbf{m} which satisfies the constraint, we only search over the five parameters $G_{vv}, G_{hh}, G_{pU}, T_1,$ and T_2 . When the other five are needed, they are generated from the constraint equations (18)–(20).

Note that it may be possible to find a closed-form expression for the mean or the mode of $p(\mathbf{m}|\mathbf{v})$. If found, then the MAP estimate would require much less computation than is required by the brute force optimization method of the previous paragraph. However, our attempts to find a closed-form expression suggest that the task is difficult.

B. Simulation and Results

To compare MAP estimates with algebraic estimates, we simulate the estimation process. Typical values of radiometer hardware parameters defined in [1], which are obtained from [1, Table I] and [2], are shown in Table I. These are used in [1, eq. 17] to calculate typical values for the eight gains G_{xx} , which are also shown in Table I.

For typical T_1 and T_2 , we use 310 K; these, plus the eight gains G_{xx} in Table I, are considered the “true” values in our simulations (\mathbf{m}_{true}). For $T_C, T_H,$ and T_{CN} , we use 288, 800,

and 800 K, respectively. For τ_c , we use 9 ms (these $T_1, T_2, T_C, T_H, T_{CN}$, and τ_c are the anticipated values for NASA's upcoming Aquarius radiometer [2]).

From the aforementioned parameters, we next generate simulated voltages. As discussed in Section IV-A2, we can find $V_2^T \mathbf{v}$ as $V_2^T \mathbf{g}(\mathbf{m}_{\text{true}})$. We also need realizations of $V_1^T \mathbf{v}$. Because $V_1^T \mathbf{v}$ is a linear combination of \mathbf{v} , it is Gaussian as is \mathbf{v} . We know that $V_1^T \mathbf{v}$ has a mean of $V_1^T \mathbf{g}(\mathbf{m})$ and a covariance matrix of $\text{diag}(\lambda)$, so we can readily generate random realizations of it by adding the mean to zero-mean Gaussian noise with that covariance. Then, with $V_2^T \mathbf{v}$ and $V_1^T \mathbf{v}$ in hand, we can form \mathbf{v} by derotating (left multiplying by V) because

$$V \begin{bmatrix} V_1^T \mathbf{v} \\ V_2^T \mathbf{v} \end{bmatrix} = V(V^T \mathbf{v}) = (VV^T) \mathbf{v} = \mathbf{v}. \quad (22)$$

With the simulated voltages, we then use the method of Piepmeier [1], as summarized in Section III, to algebraically estimate the ten calibration parameters. We repeat this for 10^6 different realizations of the voltages (the means stay the same, but the noise changes). The bias of these 10^6 estimates is computed as the difference between their mean and \mathbf{m}_{true} . The STD of these 10^6 estimates is also computed. Finally, the rmse of these 10^6 estimates is computed as the root sum square of the bias and STD (this is equivalent to $\text{rmse} = \sqrt{\langle (\hat{\mathbf{m}} - \mathbf{m}_{\text{true}})^2 \rangle}$, where the averaging is over 10^6 realizations of noise).

This process is repeated for MAP estimates, and the results are compared. For both estimation methods, and for all ten parameters, the bias is less than 0.01% of the true parameter values. The STD is therefore the same as the rmse, to four significant digits. In the first two rows of Table II, we report the rmse for each method and for each of the ten parameters, as percentages of the true parameter values. In the third row, we report the factor by which the rmse of MAP estimates is lower than the rmse of algebraic estimates.

To summarize our results with a single number, we take the average of these ten improvement factors, which is 2.04. That is, the rmse of MAP estimates is about two times smaller than the rmse of algebraic estimates.

To establish the accuracy of this number, the entire procedure of the last four paragraphs is repeated 100 times to provide 100 estimates of the average improvement factor. The average of these 100 numbers is 2.041, with an STD of 0.001.

Note that this improvement is independent of the available integration time τ_c . For example, if τ_c is quadrupled (equivalent to averaging four contiguous calibration looks), then the STD (and hence rmse) of both methods drops by a factor of two. The improvement factors reported in Table II remain the same. We verified this for seven sets of 10^6 estimates.

The cost for the increased accuracy of MAP estimates is an increase in computation. On average, a MAP estimate requires 40 000 times more computation than an algebraic estimate. This could be reduced if pains were taken to increase the efficiency of the search algorithm or of the eigendecomposition. Even so, this is still quite tractable: a 2.4-GHz workstation finds the MAP estimate of all ten parameters in about 0.06 s.

1) *Improvement as a Function of \mathbf{m}_{true}* : The improvement in accuracy of MAP estimates (over algebraic estimates) is a weak function of \mathbf{m}_{true} . If we repeat the earlier procedure for 100 different values of \mathbf{m}_{true} 's, chosen randomly within

TABLE II
RMSE RESULTS FOR 10^6 ESTIMATES FROM VOLTAGES GENERATED BY TYPICAL \mathbf{m}_{true} .
RMSE IS GIVEN AS A PERCENT OF THE TRUE PARAMETER VALUE

Parameter	G_{vv}	G_{hh}	G_{pv}	G_{ph}	G_{pU}	G_{mv}	G_{mh}	G_{mU}	T_1	T_2
RMSE of algebraic estimates (%)	0.58	0.58	1.33	0.63	0.78	1.24	0.63	0.59	1.39	1.39
RMSE of MAP estimates (%)	0.44	0.43	0.44	0.43	0.21	0.44	0.43	0.21	1.05	1.18
Improvement Factor (row 1 / row 2)	1.3	1.4	3.0	1.5	3.7	2.8	1.5	2.8	1.3	1.2

expected ranges,¹ we find that the rmse of MAP estimates is consistently the same as the values in Table II. The rmse of algebraic estimates is slightly lower than the values in Table II, resulting in an average improvement factor that ranges from 1.86 to 2.03, with the mean being 1.90.

VI. SAMPLING THE POSTERIOR PDF

We now turn to exploration of the more complete answer to the calibration problem, the posterior pdf $p(\mathbf{m}|\mathbf{v})$.

If samples of $p(\mathbf{m}|\mathbf{v})$ are available, they can be used to visualize the posterior pdf, to report it, or to calculate its mean \mathbf{m}_{mmse} . In this section, we describe how to generate such samples and then use them for the aforementioned purposes.

A. Sampling $p(\mathbf{m}|\mathbf{v})$ by the Rejection Method

Samples of the posterior pdf $p(\mathbf{m}|\mathbf{v})$ can be generated by the well-known rejection method [6], [7]. First, samples of G_{vv} , G_{hh} , and G_{pU} are proposed from independent uniform distributions. These distributions are centered on an initial guess, such as the MAP estimate.

With the proposed G_{vv} , G_{hh} , and G_{pU} coordinates, we next find the proposed G_{ph} , G_{mh} , G_{pv} , G_{mv} , and G_{mU} coordinates from the constraint equations (18)–(20). Also, T_1 and T_2 coordinates are proposed from independent uniform distributions centered on the initial guess. Each set of ten proposed coordinates now comprises one proposed sample of $p(\mathbf{m}|\mathbf{v})$, which we denote by \mathbf{m}_{prop} .

Because the constraint equations were used, each \mathbf{m}_{prop} is a sample from a uniform distribution over a region of the constraint manifold. Let the constant value of this uniform distribution be denoted by k . In order to correctly generate samples of $p(\mathbf{m}|\mathbf{v})$, we must accept each \mathbf{m}_{prop} with probability P where

$$P = \frac{p(\mathbf{m}_{\text{prop}}|\mathbf{v})/k}{\max_{\mathbf{m}}(p(\mathbf{m}|\mathbf{v})/k)} = \frac{p(\mathbf{m}_{\text{prop}}|\mathbf{v})}{\max_{\mathbf{m}}(p(\mathbf{m}|\mathbf{v}))}. \quad (23)$$

The numerator of (23) is readily calculated using (21), with V_1 and \mathbf{A} being found from a numerical eigendecomposition of C and C being calculated from \mathbf{m}_{prop} . The denominator of (23) is the peak value of $p(\mathbf{m}|\mathbf{v})$. The search for this peak is

¹ $c_v, c_h, c_p,$ and c_m are chosen independently from normal distributions with mean of 450 and STD of 17 mV/mW. G_1 is chosen from a normal distribution with mean of $1.8 * 10^7$ and STD of $2.7 * 10^6$ W/W; $10 \log_{10}(g)$ is chosen from a normal distribution with mean of zero and STD of 1 dB; s is chosen from a normal distribution with mean of $1/\sqrt{2}$ and STD of $0.02/\sqrt{2}$; α_e is chosen from a normal distribution with mean of 0.934 and STD of 0.01; and T_1 and T_2 are chosen independently from normal distributions with mean of 310 and STD of 1 K. Also, T_C is chosen from a normal distribution with mean of 288 and STD of 0.5 K, whereas T_H and T_{CN} are chosen independently from normal distributions with mean of 800 and STD of 2 K.

the same search that finds the MAP estimate of \mathbf{m} —i.e., the denominator is simply $p(\mathbf{m}_{\text{MAP}}|\mathbf{v})$.²

B. Using the Samples

1) *Marginal Posterior Pdfs:* Once we have a number of samples of $p(\mathbf{m}|\mathbf{v})$, we can immediately obtain plots of the marginal probability distribution for the i th parameter by simply binning the i th coordinate values of the samples (we believe that this follows from [6]). Some marginal pdfs obtained in this manner are shown in Fig. 1.

Fig. 1 shows that the marginal pdfs are symmetric. This verifies the claim made earlier that our MAP estimates (modes of marginal pdfs) are the same as mmse estimates (means of marginal pdfs). As further proof, when mmse estimates (see hereafter) are made, they have the same error statistics as MAP estimates, as reported in Table II and Section V-B.

Fig. 1 also shows that the marginal pdfs are Gaussian (or at least very nearly so). The Gaussians that are plotted have STD from the second line of Table II. Hence, each marginal pdf can be completely characterized by its mode (the MAP estimate) and STD (= rmse of MAP estimates, given in Table II).

2) *Additional Information in Joint Pdfs:* Joint pdfs can convey many times more information than marginal pdfs. For example, consider G_{vv} and T_1 . The 2-D joint pdf for these two parameters (made from the same samples as Fig. 1) is shown in Fig. 2. This joint pdf contains significant correlation information. For example, it shows that there is a fair chance that $G_{vv} \approx 24.5$ and $T_1 \approx 311$ but almost no chance that $G_{vv} \approx 24.5$ and $T_1 \approx 304$. If we had reported only the 1-D marginal pdfs in Fig. 1, both possibilities would have appeared equally likely.

The current practice of reporting only the mean and variance for each parameter is equivalent, in effect, to reporting independent 1-D marginal Gaussian pdfs for the parameters [6]. The only joint pdf that can logically be reconstructed from marginal pdfs is the product of the marginal pdfs (this follows from [6]). An example of such a reconstruction is shown in Fig. 3. All correlation information is lost, as well as any non-Gaussian characteristics of the posterior pdf.

The situation is even more pronounced for the parameters whose estimation is 100% correlated, due to the constraint equations. For example, the 2-D joint pdf for G_{vv} and G_{pv} is shown in Fig. 4. The pdf is completely concentrated along a 1-D line in the 2-D space. If only means and variances were reported, the appearance would be similar to Fig. 3.

We have demonstrated that the most complete answer to an estimation problem is a joint pdf, but how can a 10-D posterior pdf be reported? We can report the equation for it, such as (21).

²Numerical note: The sampling process is much faster when all the G_{xx} (or voltages) are scaled by an appropriate factor. A factor of 10^7 was used in generating the figures.

A numerical alternative is to simply report a large number of samples because most, if not all, calculations done with a posterior pdf can be done using these samples [6]. A very succinct alternative is available when a multidimensional pdf is sufficiently Gaussian, as in the present case. In such cases, all of the information can be conveyed by a vector of means and a covariance matrix.

3) *MMSE Estimation*: Another use for the samples of a distribution is in calculating its mean. For example, to find the i th coordinate of the mean, we simply average the i th coordinate of the samples, which converges to the mean as the number of samples increases [6]. The means found by this method are shown in Fig. 1, from which it is also seen that MAP and mmse estimates extracted from $p(\mathbf{m}|\mathbf{v})$ are equivalent.

VII. INFORMATION ON HARDWARE PARAMETERS

As a tangential but useful extension of the previous results in this paper, we can obtain information on the radiometer hardware parameters that comprise the eight gains G_{xx} . The definition of these gains in terms of hardware parameters, found by comparing (1) and [1, eq. 17], is reproduced in (24), shown at the bottom of the page. By simply replacing the eight gains G_{xx} ³ in (21) with their component definitions on the right side of (24), we immediately obtain the joint posterior pdf for the eight hardware parameters G_1 , G_2 , α_e , c_v , c_h , c_p , c_m , and s , plus T_1 and T_2 (the bandwidth B could also be considered a parameter, but for this paper, we treat it as a known constant; the parameter k is Boltzmann's constant). For brevity, we simply summarize our discoveries about this pdf as follows.

First, the transformed constraint equations leave the following six hardware parameters unconstrained: G_1 , G_2 , α_e , c_v , T_1 , and T_2 . Next, the constraint equations dictate that the parameter s is completely determined by the voltages

$$s = \sqrt{\frac{AD + \sqrt{ADBC}}{AD - BC}} \quad (25)$$

where

$$\begin{aligned} A &\equiv v_{p,C}v_{h,H} - v_{h,C}v_{p,H} & D &\equiv v_{m,C}v_{v,H} - v_{v,C}v_{m,H} \\ B &\equiv v_{p,C}v_{v,H} - v_{v,C}v_{p,H} & C &\equiv v_{m,C}v_{h,H} - v_{h,C}v_{m,H}. \end{aligned} \quad (26)$$

There is no uncertainty in this estimate—it is not affected by $NE\Delta T$ (in simulation, therefore, this estimate is the exact value of the true s ; however, in real practice, the uncertainties not captured by our model in (9), such as imperfect knowledge of T_C and T_H , will cause this estimate to have some error). The

³ G_{xx} are found both in \mathbf{m} and in the constraint equations.

constraint also dictates that c_h , c_p , and c_m are constrained as follows:

$$c_h = c_v \sqrt{\frac{AC}{BD}} \quad (27)$$

$$c_p = c_v \frac{\sqrt{A}}{E\sqrt{D}} (\sqrt{AD} - \sqrt{BC}) \quad (28)$$

$$c_m = -c_v \frac{\sqrt{C}}{E\sqrt{B}} (\sqrt{AD} - \sqrt{BC}) \quad (29)$$

where $E \equiv v_{v,C}v_{h,H} - v_{h,C}v_{v,H}$. One interpretation of these equations is that the ratio of any two c_x is perfectly resolved, which is equal to a function of the voltages.

Pdfs (2-D) for the unconstrained hardware parameters, which are obtained using the technique of Section VI with the necessary modifications, are shown in Fig. 5. Numerical simulation verifies that s is perfectly resolved; therefore, pdfs for s are omitted. Pdfs for c_h , c_p , and c_m are not shown either, because such pdfs are simply multiples of the pdfs involving c_v , due to (27)–(29).

Intuition suggests that the ability to perfectly resolve s must be compensated by a lack of ability to resolve other parameters. This is indeed the case: As shown by Fig. 5, c_v , G_1 , and G_2 cannot be separately resolved. We can only resolve their pairwise products (for example, $c_v G_1$ is a constant times G_{vv} , which was well resolved in the earlier sections of this paper). Marginal pdfs for c_v , G_1 , and G_2 are essentially uniform.⁴

Our final observation is that α_e is resolved well. The average rmse of MAP estimates of α_e is 0.33%, whether using the typical values given in Table I or the randomized values for the true hardware parameters. T_1 and T_2 are resolved as accurately as before (see Table II and Fig. 1).

VIII. CONCLUSION

In this paper, we have demonstrated optimal estimation of calibration parameters in polarimetric microwave radiometers which use hybrid coupler-based correlators to measure the third Stokes parameter, such as NASA's upcoming Aquarius radiometer. By exploiting statistical knowledge of measurement noise using Bayesian estimation, the rmse is reduced by a factor of two compared to estimation without such knowledge. Most of the principles that we have employed are well known in estimation theory; however, this paper is their first published application to microwave radiometer calibration. Many extensions of this paper can be made by expanding the forward model to include other unknowns and/or other measurements.

⁴These uniform marginal pdfs are only limited by prior knowledge. However, if we have tighter prior knowledge on one parameter, for example, on c_v , then good resolution of a product such as $c_v G_1$ can cause tighter bounds on the other parameter G_1 . This effect can be seen in several subplots of Fig. 5.

$$\begin{bmatrix} G_{vv} & 0 & 0 \\ 0 & G_{hh} & 0 \\ G_{pv} & G_{ph} & G_{pU} \\ G_{mv} & G_{mh} & G_{mU} \end{bmatrix} \equiv kB \begin{bmatrix} c_v G_1 & 0 & 0 \\ 0 & c_h G_2 & 0 \\ c_p s^2 G_1 & c_p (1 - s^2) G_2 & c_p s \sqrt{1 - s^2} \alpha_e \sqrt{G_1 G_2} \\ c_m (1 - s^2) G_1 & c_m s^2 G_2 & -c_m s \sqrt{1 - s^2} \alpha_e \sqrt{G_1 G_2} \end{bmatrix} \quad (24)$$

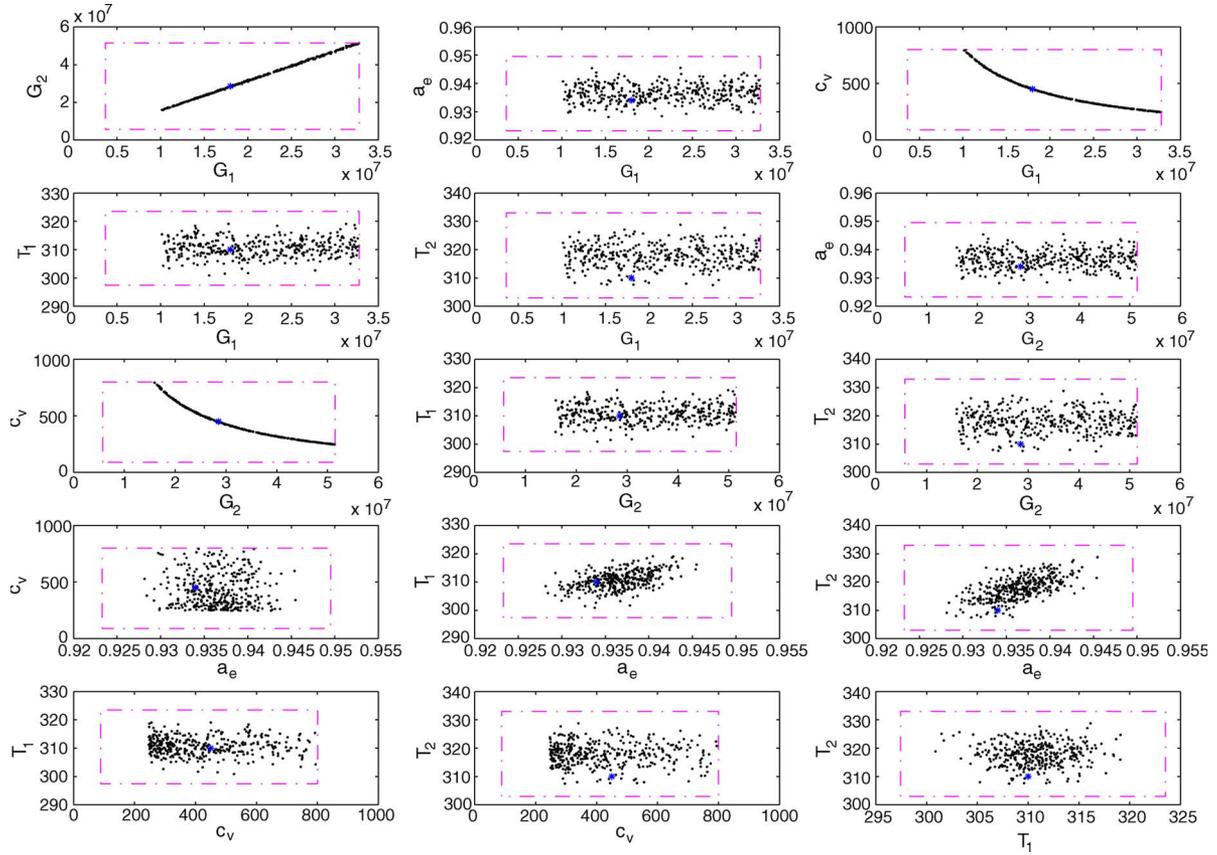


Fig. 5. Two-dimensional pdfs for the four unconstrained radiometer hardware parameters and two noise temperatures. The density of dots illustrates the pdfs, whereas asterisks show the location of the true parameter values. The bounding boxes show the bounds of the uniform pdfs from which samples were proposed. These plots illustrate that it is only possible to obtain useful estimates of the hardware parameters α_e , T_1 , and T_2 (in addition to s , which is not illustrated).

This paper has also illustrated the fact that much more information can be conveyed by a posterior probability distribution for a set of parameters than by simple estimates comprised only of marginal means and variances. Finally, we have demonstrated the acquisition of information on the eight hardware parameters that comprise the overall channel gains in the class of radiometer which is analyzed in this paper. Two of these hardware parameters can be accurately estimated from calibration measurements; for the other six, only products or ratios of pairs of hardware parameters can be resolved.

We note that the forward model (1) is based on a number of stated assumptions—see Section II—and our simulations incorporate those assumptions. In future work, we hope to assess the validity of these assumptions and the consequent predictions of this paper by testing the proposed method on real radiometer data. Future work can also assess the improvement in scene brightness temperature accuracy that can be achieved by using the methodology of this paper rather than algebraic estimation.

APPENDIX A

A. Derivation of Covariance Matrix C

Using the information stated in Section IV-A1 about the nine noise variables, we derive here the variances and covariances of the voltages in (9). Each column of voltages is independent of the voltages in the other columns, so we proceed column by column.

1) *First Three Columns:* Consider the first column of voltages in (9). There are just two noise terms, and they are independent of one another because they are from different sources. Therefore, $v_{v,C}$ and $v_{h,C}$ are independent. All other relationships have nonzero correlation

$$\begin{aligned} \text{Cov}(v_{v,C}, v_{p,C}) &= \text{Cov}(G_{vv}n_1, G_{pv}n_1) \\ &= E(G_{vv}G_{pv}n_1^2) \\ &= G_{vv}G_{pv} \frac{(T_C + T_1)^2}{B\tau_c}. \end{aligned} \quad (30)$$

Similarly we have

$$\text{Cov}(v_{h,C}, v_{p,C}) = G_{hh}G_{ph} \frac{(T_C + T_2)^2}{B\tau_c} \quad (31)$$

$$\text{Cov}(v_{v,C}, v_{m,C}) = G_{vv}G_{mv} \frac{(T_C + T_1)^2}{B\tau_c} \quad (32)$$

$$\text{Cov}(v_{h,C}, v_{m,C}) = G_{hh}G_{mh} \frac{(T_C + T_2)^2}{B\tau_c}. \quad (33)$$

For the final covariance for this column

$$\begin{aligned} \text{Cov}(v_{p,C}, v_{m,C}) &= E((G_{pv}n_1 + G_{ph}n_2)(G_{mv}n_1 + G_{mh}n_2)) \\ &= E(G_{pv}G_{mv}n_1^2 + G_{ph}G_{mh}n_2^2) \\ &= G_{pv}G_{mv} \frac{(T_C + T_1)^2}{B\tau_c} \\ &\quad + G_{ph}G_{mh} \frac{(T_C + T_2)^2}{B\tau_c}. \end{aligned} \quad (34)$$

By inspection, the four variances (diagonal terms) are $G_{vv}^2((T_C + T_1)^2/B\tau_c)$, $G_{hh}^2((T_C + T_2)^2/B\tau_c)$, $G_{pv}^2((T_C + T_1)^2/B\tau_c) + G_{ph}^2((T_C + T_2)^2/B\tau_c)$, and $G_{mv}^2((T_C + T_1)^2/B\tau_c) + G_{mh}^2((T_C + T_2)^2/B\tau_c)$.

The variances and covariances of the second column of voltages in (9) are the same as those of the first column, except replacing $(T_C + T_1)$ and $(T_C + T_2)$ with $(T_H + T_1)$ and $(T_H + T_2)$, respectively. The variances and covariances of the third column are the same as those of the first one, except replacing $(T_C + T_2)$ with $(T_H + T_2)$.

2) *Fourth Column:* The variances and covariances of the fourth column are quite different because T_{CN} is a component of all three inputs. We first rewrite the calibration inputs (T_C for both the vertical and horizontal channels, $T_{CN}/2$, T_1 , and T_2) in terms of electric fields.

The v -channel cold load emits an electric field, which we denote by c_1 . Its second moment (defined as $\langle c_1^2 \rangle$, where $\langle \cdot \rangle$ is the ensemble average) is T_C (here and hereafter, we ignore the constant that converts the product of two electric fields to a brightness temperature). The v -channel low-noise-amplifier (LNA) noise is another source, whose equivalent electric field (referred to the input of the first LNA so that it is on the same level as c_1) is denoted by r_1 . Its second moment is T_1 . Similarly, the h -channel cold load outputs c_2 , with second moment T_C , and the h -channel LNA noise is r_2 , whose second moment is T_2 .

The correlated calibration source (shown in [1, Fig. 1]) emits an electric field n with second moment T_{CN} . When the energy from this source is split between the vertical and horizontal channels, the electric field in each channel is then $n/\sqrt{2}$, whose second moment is $T_{CN}/2$.

The five electric fields just described (c_1 , r_1 , c_2 , r_2 , and n) are independent of one another because of their distinct origins. They are all zero-mean normal random variables.

The voltages in the fourth column on the left side of (1) are found by summing these electric fields, squaring, integrating, and multiplying by a channel gain

$$v_{v,CN} = G_{vv}I \quad (35)$$

$$v_{h,CN} = G_{hh}J \quad (36)$$

$$v_{p,CN} = G_{pv}I + G_{ph}J + G_{pU}K \quad (37)$$

$$v_{m,CN} = G_{mv}I + G_{mh}J + G_{mU}K \quad (38)$$

where

$$I \equiv \frac{1}{\tau_c} \int_0^{\tau_c} \left(c_1 + \frac{n}{\sqrt{2}} + r_1 \right)^2 dt \quad (39)$$

$$J \equiv \frac{1}{\tau_c} \int_0^{\tau_c} \left(c_2 + \frac{n}{\sqrt{2}} + r_2 \right)^2 dt \quad (40)$$

$$K \equiv \frac{1}{\tau_c} \int_0^{\tau_c} n^2 dt. \quad (41)$$

Note that K arises from the correlation of the inputs to the hybrid coupler.

Because all these voltages are expressed in terms of I , J , and K , all the variances and covariances can also be expressed

in terms of the variances and covariances of I , J , and K . As shown in [5], I , J , and K can be rewritten as sums of independent samples

$$I = \frac{1}{N_c} \sum_{i=1}^{N_c} \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 \quad (42)$$

where $N_c = 2B\tau_c$, and similarly with J and K . In the remainder of this section, we use (42) to find the means, variances, and covariances of I , J , and K and then of $v_{v,CN}$, $v_{h,CN}$, $v_{p,CN}$, and $v_{m,CN}$. These derivations are similar to those in [8, Appendix].

Means and variances of I , J , and K : First, using the independence of each of the N_c samples from one another and of the five electric fields from one another, the means of I , J , and K are

$$\begin{aligned} \langle I \rangle &= \left\langle \left(c_1 + n/\sqrt{2} + r_1 \right)^2 \right\rangle \\ &= \left\langle c_1^2 + n^2/2 + r_1^2 + 2c_1n/\sqrt{2} + 2c_1r_1 + 2nr_1/\sqrt{2} \right\rangle \\ &= T_C + T_{CN}/2 + T_1 \equiv TT1 \end{aligned} \quad (43)$$

$$\begin{aligned} \langle J \rangle &= \left\langle \left(c_2 + n/\sqrt{2} + r_2 \right)^2 \right\rangle \\ &= T_C + T_{CN}/2 + T_2 \equiv TT2 \end{aligned} \quad (44)$$

$$\langle K \rangle = \langle n^2 \rangle = T_{CN}. \quad (45)$$

These means coincide with the final column of the temperature matrix in (1), verifying our formulation of the problem in terms of electric fields. To determine variance

$$\begin{aligned} \text{Var}(I) &= \left\langle \frac{1}{N_c} \sum_{i=1}^{N_c} \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 \right. \\ &\quad \times \left. \frac{1}{N_c} \sum_{j=1}^{N_c} \left(c_{1,j} + \frac{n_j}{\sqrt{2}} + r_{1,j} \right)^2 \right\rangle - TT1^2 \quad (46) \\ &= \frac{1}{N_c^2} \left\langle \sum_{i=1}^{N_c} \sum_{j=1}^{N_c} \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 \right. \\ &\quad \times \left. \left(c_{1,j} + \frac{n_j}{\sqrt{2}} + r_{1,j} \right)^2 \right\rangle - TT1^2. \quad (47) \end{aligned}$$

Separating the expected value operation into terms for which $i \neq j$ and for which $i = j$

$$\begin{aligned} \text{Var}(I) &= \frac{1}{N_c^2} \left(\sum_{i=1}^{N_c} \sum_{j=1(\neq i)}^{N_c} \left\langle \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 \right. \right. \\ &\quad \times \left. \left. \left(c_{1,j} + \frac{n_j}{\sqrt{2}} + r_{1,j} \right)^2 \right\rangle \right. \\ &\quad \left. + \sum_{i=1}^{N_c} \left\langle \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^4 \right\rangle \right) - TT1^2. \end{aligned} \quad (48)$$

Using the independence of samples i and j , the independence of c_1 , n , and r_1 from everything but themselves, and the known

fourth moment of zero-mean normal random variables

$$\begin{aligned} \text{Var}(I) = & \frac{1}{N_c^2} \left(\sum_{i=1}^{N_c} \left\langle \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 \right\rangle \right. \\ & \times \sum_{j=1(\neq i)}^{N_c} \left\langle \left(c_{1,j} + \frac{n_j}{\sqrt{2}} + r_{1,j} \right)^2 \right\rangle \\ & \left. + \sum_{i=1}^{N_c} 3 \left\langle \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 \right\rangle^2 \right) - TT1^2. \end{aligned} \quad (49)$$

Then, using (43)

$$\begin{aligned} \text{Var}(I) = & \frac{1}{N_c^2} (N_c TT1(N_c - 1) TT1 + 3N_c TT1^2) - TT1^2 \\ = & \frac{2}{N_c} TT1^2 \end{aligned} \quad (50)$$

and we finally arrive at the ensemble variance

$$\text{Var}(I) = \frac{TT1^2}{B\tau_c} \quad \text{and similarly,} \quad \text{Var}(J) = \frac{TT2^2}{B\tau_c}. \quad (51)$$

Also

$$\begin{aligned} \text{Var}(K) = & \left\langle \frac{1}{N_c} \sum_{i=1}^{N_c} n_i^2 \frac{1}{N_c} \sum_{j=1}^{N_c} n_j^2 \right\rangle - T_{CN}^2 \\ = & \frac{1}{N_c^2} \left\langle \sum_{i=1}^{N_c} \sum_{j=1}^{N_c} n_i^2 n_j^2 \right\rangle - T_{CN}^2 \end{aligned} \quad (52)$$

$$= \frac{1}{N_c^2} \left(\sum_{i=1}^{N_c} \sum_{j=1(\neq i)}^{N_c} \langle n_i^2 n_j^2 \rangle + \sum_{i=1}^{N_c} \langle n_i^4 \rangle \right) - T_{CN}^2 \quad (53)$$

$$= \frac{1}{N_c^2} \left(\sum_{i=1}^{N_c} \langle n_i^2 \rangle \sum_{j=1(\neq i)}^{N_c} \langle n_j^2 \rangle + \sum_{i=1}^{N_c} 3 \langle n_i^2 \rangle^2 \right) - T_{CN}^2 \quad (54)$$

$$\begin{aligned} = & \frac{1}{N_c^2} (N_c(N_c - 1) T_{CN}^2 + 3N_c T_{CN}^2) - T_{CN}^2 \\ = & \frac{T_{CN}^2}{B\tau_c}. \end{aligned} \quad (55)$$

Covariances of I, J, and K: Derivation of the covariances of I, J, and K is similar

$$\begin{aligned} \text{Cov}(I, J) = & \left\langle \frac{1}{N_c} \sum_{i=1}^{N_c} \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 \frac{1}{N_c} \right. \\ & \times \sum_{j=1}^{N_c} \left(c_{2,j} + \frac{n_j}{\sqrt{2}} + r_{2,j} \right)^2 \left. \right\rangle - TT1 TT2 \end{aligned} \quad (56)$$

$$\begin{aligned} = & \frac{1}{N_c^2} \left\langle \sum_{i=1}^{N_c} \sum_{j=1}^{N_c} \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 \right. \\ & \times \left. \left(c_{2,j} + \frac{n_j}{\sqrt{2}} + r_{2,j} \right)^2 \right\rangle - TT1 TT2. \end{aligned} \quad (57)$$

The expected value operation separates into terms for which $i \neq j$ and for which $i = j$

$$\begin{aligned} & \sum_{i=1}^{N_c} \sum_{j=1(\neq i)}^{N_c} \left\langle \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 \left(c_{2,j} + \frac{n_j}{\sqrt{2}} + r_{2,j} \right)^2 \right\rangle \\ & + \sum_{i=1}^{N_c} \left\langle \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 \left(c_{2,i} + \frac{n_i}{\sqrt{2}} + r_{2,i} \right)^2 \right\rangle. \end{aligned} \quad (58)$$

Using the independence of samples i and j , the independence of c_1, c_2, r_1 , and r_2 from everything but themselves, and the known fourth moment of zero-mean normal random variables, this becomes

$$\begin{aligned} & N_c TT1(N_c - 1) TT2 + N_c \left[TT1 TT2 + \frac{T_{CN}^2}{2} \right] \\ = & N_c^2 TT1 TT2 + N_c \frac{T_{CN}^2}{2}. \end{aligned} \quad (59)$$

Putting this back into (57)

$$\text{Cov}(I, J) = \frac{T_{CN}^2}{4B\tau_c}. \quad (60)$$

Similarly we have

$$\begin{aligned} \text{Cov}(I, K) = & \left\langle \frac{1}{N_c} \sum_{i=1}^{N_c} \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 \frac{1}{N_c} \sum_{j=1}^{N_c} n_j^2 \right\rangle \\ & - TT1 T_{CN} \end{aligned} \quad (61)$$

$$\begin{aligned} = & \frac{1}{N_c^2} \left\langle \sum_{i=1}^{N_c} \sum_{j=1}^{N_c} \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 n_j^2 \right\rangle \\ & - TT1 T_{CN} \end{aligned} \quad (62)$$

$$\begin{aligned} = & \frac{1}{N_c^2} \left(\sum_{i=1}^{N_c} \left\langle \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 \right\rangle \sum_{j=1(\neq i)}^{N_c} \langle n_j^2 \rangle \right. \\ & \left. + \sum_{i=1}^{N_c} \left\langle \left(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i} \right)^2 n_i^2 \right\rangle \right) \\ & - TT1 T_{CN} \end{aligned}$$

$$\begin{aligned} = & \frac{1}{N_c^2} \left(N_c TT1(N_c - 1) T_{CN} + N_c \right. \\ & \times \left. \left[T_{CN} T_{CN} + \frac{\langle n^4 \rangle}{2} + T_1 T_{CN} \right] \right) \\ & - TT1 T_{CN} \end{aligned} \quad (63)$$

$$\begin{aligned} = & \frac{1}{N_c^2} \left(-N_c TT1 T_{CN} + N_c T_{CN} \right. \\ & \times \left. \left[T_C + \frac{3}{2} T_{CN} + T_1 \right] \right) \\ = & -\frac{TT1 T_{CN}}{N_c} + \frac{(TT1 + T_{CN}) T_{CN}}{N_c} \end{aligned} \quad (64)$$

leaving

$$\text{Cov}(I, K) = \frac{T_{CN}^2}{2B\tau_c} \quad \text{and similarly} \quad \text{Cov}(J, K) = \frac{T_{CN}^2}{2B\tau_c}. \quad (65)$$

Variances of the voltages: Now that the means, variances, and covariances of I , J , and K are known, we can find the means, variances, and covariances of the voltages in (66)–(70), shown at the bottom of the page.

$v_{p,CN}$ is identical to $v_{m,CN}$ —same realizations of noise—except for multiplication by different G_{xx} . Therefore, variances and covariances of the first will be identical to those of the second if we simply replace G_{px} with G_{mx} , as indicated in (71), shown at the bottom of the page.

Covariances of the voltages: The covariances of $v_{v,CN}$ with the other voltages are found as follows:

$$\begin{aligned} \text{Cov}(v_{v,CN}, v_{h,CN}) &= G_{vv}G_{hh}\text{Cov}(I, J) \\ &= G_{vv}G_{hh}\frac{T_{CN}^2}{4B\tau_c} \end{aligned} \quad (72)$$

$$\begin{aligned} \text{Cov}(v_{v,CN}, v_{p,CN}) &= \text{Cov}(G_{vv}I, G_{pv}I + G_{ph}J + G_{pU}K) \\ &= G_{vv}[G_{pv}\text{Var}(I) + G_{ph}\text{Cov}(I, J) + G_{pU}\text{Cov}(I, K)] \end{aligned} \quad (73)$$

$$= G_{vv}\left[G_{pv}\frac{TT1^2}{B\tau_c} + G_{ph}\frac{T_{CN}^2}{4B\tau_c} + G_{pU}\frac{T_{CN}^2}{2B\tau_c}\right] \quad (74)$$

$$= G_{vv}\frac{4G_{pv}TT1^2 + G_{ph}T_{CN}^2 + 2G_{pU}T_{CN}^2}{4B\tau_c} \quad (75)$$

and similarly

$$\begin{aligned} \text{Cov}(v_{v,CN}, v_{m,CN}) &= G_{vv}\frac{4G_{mv}TT1^2 + G_{mh}T_{CN}^2 + 2G_{mU}T_{CN}^2}{4B\tau_c}. \end{aligned} \quad (76)$$

Likewise, the covariances of $v_{h,CN}$ with the $v_{p,CN}$ and $v_{m,CN}$ are

$$\begin{aligned} \text{Cov}(v_{h,CN}, v_{p,CN}) &= \text{Cov}(G_{hh}J, G_{pv}I + G_{ph}J + G_{pU}K) \end{aligned} \quad (77)$$

$$= G_{hh}[G_{pv}\text{Cov}(I, J) + G_{ph}\text{Var}(J) + G_{pU}\text{Cov}(J, K)] \quad (78)$$

$$= G_{hh}\left[G_{pv}\frac{T_{CN}^2}{4B\tau_c} + G_{ph}\frac{TT2^2}{B\tau_c} + G_{pU}\frac{T_{CN}^2}{2B\tau_c}\right] \quad (79)$$

$$= G_{hh}\frac{G_{pv}T_{CN}^2 + 4G_{ph}TT2^2 + 2G_{pU}T_{CN}^2}{4B\tau_c} \quad (80)$$

$$\begin{aligned} \text{Cov}(v_{h,CN}, v_{m,CN}) &= G_{hh}\frac{G_{mv}T_{CN}^2 + 4G_{mh}TT2^2 + 2G_{mU}T_{CN}^2}{4B\tau_c}. \end{aligned} \quad (81)$$

The final covariance is longer

$$\begin{aligned} \text{Cov}(v_{p,CN}, v_{m,CN}) &= \text{Cov}(G_{pv}I + G_{ph}J + G_{pU}K, G_{mv}I + G_{mh}J + G_{mU}K) \end{aligned} \quad (82)$$

$$\begin{aligned} &= G_{pv}G_{mv}\text{Var}(I) + G_{ph}G_{mh}\text{Var}(J) + G_{pU}G_{mU}\text{Var}(K) \\ &\quad + (G_{pv}G_{mh} + G_{ph}G_{mv})\text{Cov}(I, J) \\ &\quad + (G_{pv}G_{mU} + G_{pU}G_{mv})\text{Cov}(I, K) \\ &\quad + (G_{ph}G_{mU} + G_{pU}G_{mh})\text{Cov}(J, K) \end{aligned} \quad (83)$$

$$\text{Var}(v_{v,CN}) = G_{vv}^2\text{Var}(I) = G_{vv}^2\frac{TT1^2}{B\tau_c} \quad (84)$$

$$\text{Var}(v_{h,CN}) = G_{hh}^2\text{Var}(J) = G_{hh}^2\frac{TT2^2}{B\tau_c} \quad (85)$$

$$\begin{aligned} \text{Var}(v_{p,CN}) &= \text{Var}(G_{pv}I + G_{ph}J + G_{pU}K) \\ &= \text{Var}(G_{pv}I) + \text{Var}(G_{ph}J) + \text{Var}(G_{pU}K) + 2[G_{pv}G_{ph}\text{Cov}(I, J) + G_{pv}G_{pU}\text{Cov}(I, K) + G_{ph}G_{pU}\text{Cov}(J, K)] \end{aligned} \quad (86)$$

$$= G_{pv}^2\frac{TT1^2}{B\tau_c} + G_{ph}^2\frac{TT2^2}{B\tau_c} + G_{pU}^2\frac{T_{CN}^2}{B\tau_c} + 2G_{pv}G_{ph}\frac{T_{CN}^2}{4B\tau_c} + 2G_{pv}G_{pU}\frac{T_{CN}^2}{2B\tau_c} + 2G_{ph}G_{pU}\frac{T_{CN}^2}{2B\tau_c} \quad (87)$$

$$= \frac{G_{pv}^2TT1^2 + G_{ph}^2TT2^2 + (G_{pU}^2 + G_{pv}G_{ph}/2 + G_{pv}G_{pU} + G_{ph}G_{pU})T_{CN}^2}{B\tau_c} \quad (88)$$

$$\text{Var}(v_{m,CN}) = \frac{G_{mv}^2TT1^2 + G_{mh}^2TT2^2 + (G_{mU}^2 + G_{mv}G_{mh}/2 + G_{mv}G_{mU} + G_{mh}G_{mU})T_{CN}^2}{B\tau_c} \quad (89)$$

which expands and then simplifies to

$$\begin{aligned} & \text{Cov}(v_{p,CN}, v_{m,CN}) \\ &= \frac{1}{4B\tau_c} \left[4G_{pv}G_{mv}TT1^2 + 4G_{ph}G_{mh}TT2^2 \right. \\ & \quad + (4G_{pU}G_{mU} + G_{pv}G_{mh} + G_{ph}G_{mv} \\ & \quad + 2G_{pv}G_{mU} + 2G_{pU}G_{mv} + 2G_{ph}G_{mU} \\ & \quad \left. + 2G_{pU}G_{mh})T_{CN}^2 \right]. \end{aligned} \quad (85)$$

3) *Summary: Entire C:* To recapitulate

$$\mathbf{C} = \begin{bmatrix} C_C & 0 & 0 & 0 \\ 0 & C_H & 0 & 0 \\ 0 & 0 & C_{CH} & 0 \\ 0 & 0 & 0 & C_{CN} \end{bmatrix} \quad (86)$$

where C_C , C_H , C_{CH} , and C_{CN} are given in (87)–(90), shown at the bottom of the page, (note that the matrices in these

equations are all symmetric; to make them fit better, only the upper triangular elements are given), where

$$\begin{aligned} P &\equiv G_{pv}^2 TT1^2 + G_{ph}^2 TT2^2 \\ & \quad + (G_{pU}^2 + G_{pv}G_{ph}/2 + G_{pv}G_{pU} + G_{ph}G_{pU}) T_{CN}^2 \\ M &\equiv G_{mv}^2 TT1^2 + G_{mh}^2 TT2^2 \\ & \quad + (G_{mU}^2 + G_{mv}G_{mh}/2 + G_{mv}G_{mU} + G_{mh}G_{mU}) T_{CN}^2 \\ Y &\equiv G_{pv}G_{mv}TT1^2 + G_{ph}G_{mh}TT2^2 \\ & \quad + \left(G_{pU}G_{mU} + \frac{G_{pv}G_{mh}}{4} + \frac{G_{ph}G_{mv}}{4} + \frac{G_{pv}G_{mU}}{2} \right. \\ & \quad \left. + \frac{G_{pU}G_{mv}}{2} + \frac{G_{ph}G_{mU}}{2} + \frac{G_{pU}G_{mh}}{2} \right) T_{CN}^2. \end{aligned}$$

B. Derivation of the Constraint Equations

1) *Finding V_2 (Eigenvectors of C for which $\lambda = 0$):* Due to the block diagonal structure of C , it has four eigenvectors of the form $[a \ b \ c \ d \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$, where $[a \ b \ c \ d]^T$ is an eigenvector of C_C ; four of the form $[0 \ 0 \ 0 \ 0 \ a \ b \ c \ d \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$, where $[a \ b \ c \ d]^T$ is an eigenvector of C_H ; four of the form $[0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ a \ b \ c \ d \ 0 \ 0 \ 0 \ 0]^T$, where $[a \ b \ c \ d]^T$ is an eigenvector of C_{CH} ; and four of the form $[0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ a \ b \ c \ d]^T$, where $[a \ b \ c \ d]^T$ is an eigenvector of C_{CN} . Each of C_C , C_H , and C_{CH} has two eigenvalues (λ) that are equal to zero, whereas C_{CN} has only one. This is easily confirmed numerically; theoretically, it is

$$C_C = \frac{1}{B\tau_c} \begin{bmatrix} G_{vv}^2(T_C + T_1)^2 & 0 & G_{vv}G_{pv}(T_C + T_1)^2 & G_{vv}G_{mv}(T_C + T_1)^2 \\ & G_{hh}^2(T_C + T_2)^2 & G_{hh}G_{ph}(T_C + T_2)^2 & G_{hh}G_{mh}(T_C + T_2)^2 \\ & & G_{pv}^2(T_C + T_1)^2 + G_{ph}^2(T_C + T_2)^2 & G_{pv}G_{mv}(T_C + T_1)^2 \\ & & & +G_{ph}G_{mh}(T_C + T_2)^2 \\ & & & G_{mv}^2(T_C + T_1)^2 \\ & & & +G_{mh}^2(T_C + T_2)^2 \end{bmatrix} \quad (87)$$

$$C_H = \frac{1}{B\tau_c} \begin{bmatrix} G_{vv}^2(T_H + T_1)^2 & 0 & G_{vv}G_{pv}(T_H + T_1)^2 & G_{vv}G_{mv}(T_H + T_1)^2 \\ & G_{hh}^2(T_H + T_2)^2 & G_{hh}G_{ph}(T_H + T_2)^2 & G_{hh}G_{mh}(T_H + T_2)^2 \\ & & G_{pv}^2(T_H + T_1)^2 + G_{ph}^2(T_H + T_2)^2 & G_{pv}G_{mv}(T_H + T_1)^2 \\ & & & +G_{ph}G_{mh}(T_H + T_2)^2 \\ & & & G_{mv}^2(T_H + T_1)^2 \\ & & & +G_{mh}^2(T_H + T_2)^2 \end{bmatrix} \quad (88)$$

$$C_{CH} = \frac{1}{B\tau_c} \begin{bmatrix} G_{vv}^2(T_C + T_1)^2 & 0 & G_{vv}G_{pv}(T_C + T_1)^2 & G_{vv}G_{mv}(T_C + T_1)^2 \\ & G_{hh}^2(T_H + T_2)^2 & G_{hh}G_{ph}(T_H + T_2)^2 & G_{hh}G_{mh}(T_H + T_2)^2 \\ & & G_{pv}^2(T_C + T_1)^2 + G_{ph}^2(T_H + T_2)^2 & G_{pv}G_{mv}(T_C + T_1)^2 \\ & & & +G_{ph}G_{mh}(T_H + T_2)^2 \\ & & & G_{mv}^2(T_C + T_1)^2 \\ & & & +G_{mh}^2(T_H + T_2)^2 \end{bmatrix} \quad (89)$$

$$C_{CN} = \frac{1}{B\tau_c} \begin{bmatrix} G_{vv}^2 TT1^2 & G_{vv}G_{hh} \frac{T_{CN}^2}{4} & G_{vv} \frac{4G_{pv}TT1^2 + G_{ph}T_{CN}^2 + 2G_{pU}T_{CN}^2}{4} & G_{vv} \frac{4G_{mv}TT1^2 + G_{mh}T_{CN}^2 + 2G_{mU}T_{CN}^2}{4} \\ & G_{hh}^2 TT2^2 & G_{hh} \frac{G_{pv}T_{CN}^2 + 4G_{ph}TT2^2 + 2G_{pU}T_{CN}^2}{4} & G_{hh} \frac{G_{mv}T_{CN}^2 + 4G_{mh}TT2^2 + 2G_{mU}T_{CN}^2}{4} \\ & & P & Y \\ & & & M \end{bmatrix} \quad (90)$$

because the first three columns of (9) have two noise sources each, whereas the last one has three.

Eigenvectors of C_C , C_H , and C_{CH} for which $\lambda = 0$: C_C , C_H , and C_{CH} can all be written in the abbreviated forms

$$\frac{1}{B\tau_c} \begin{bmatrix} v^2T & 0 & vpT & vmT \\ 0 & h^2U & hqU & hnU \\ vpT & hqU & p^2T + q^2U & pmT + qnU \\ vmT & hnU & pmT + qnU & m^2T + n^2U \end{bmatrix} \quad (91)$$

where the only difference between C_C , C_H , and C_{CH} is whether T and U are defined using T_C or T_H . The eigenvectors of this matrix are found by using the defining equation of an eigenvector with $\lambda = 0$,

$$\begin{bmatrix} v^2T & 0 & vpT & vmT \\ 0 & h^2U & hqU & hnU \\ vpT & hqU & p^2T + q^2U & pmT + qnU \\ vmT & hnU & pmT + qnU & m^2T + n^2U \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (92)$$

From the first two of the four equations in (92)

$$a = -\frac{cp + dm}{v} \quad b = -\frac{cq + dn}{h}. \quad (93)$$

Using these to substitute for a and b in the third and fourth equations in (92), these equations reduce to

$$0c + 0d = 0 \quad 0c + 0d = 0. \quad (94)$$

Any c and d will satisfy these equations. We choose simple but nontrivial values: $c = 1$, $d = 0$ and $c = 0$, $d = 1$. Then, using (93), the eigenvectors are

$$\begin{bmatrix} -p/v \\ -q/h \\ 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} -m/v \\ -n/h \\ 0 \\ 1 \end{bmatrix}. \quad (95)$$

Eigenvectors can be scaled arbitrarily. Scaling by $-vh$ and undoing the abbreviations, the eigenvectors with $\lambda = 0$ are

$$\begin{bmatrix} G_{hh}G_{pv} \\ G_{vv}G_{ph} \\ -G_{vv}G_{hh} \\ 0 \end{bmatrix} \quad \begin{bmatrix} G_{hh}G_{mv} \\ G_{vv}G_{mh} \\ 0 \\ -G_{vv}G_{hh} \end{bmatrix}. \quad (96)$$

Note that these are exactly the same for C_C , C_H , and C_{CH} because they do not depend on T and U .

Eigenvector of C_{CN} for which $\lambda = 0$: A process similar to that of the previous subsection, although more tedious, leads to

$$\text{Eigenvector of } C_{CN} \text{ with } \lambda = 0 \text{ is } \begin{bmatrix} \frac{G_{pv}G_{mU} - G_{pU}G_{mv}}{G_{vv}} \\ \frac{G_{ph}G_{mU} - G_{pU}G_{mh}}{G_{vv}} \\ -G_{mU} \\ G_{pU} \end{bmatrix}. \quad (97)$$

2) *Forming Constraint Equations From V_2 :* For fixed \mathbf{v} and unknown \mathbf{m} , the constraint is

$$V_2^T(\mathbf{m})\mathbf{v} = V_2^T(\mathbf{m})\mathbf{g}(\mathbf{m}) \quad (98)$$

where we explicitly show that V_2 is formed from the unknown \mathbf{m} (through C).

First six constraint equations: Having found the eigenvectors that form V_2 in Appendix A.2.1, we can expand the first six of the seven equations in (98) and simplify them to obtain

$$G_{hh}G_{pv}v_{v,C} + G_{vv}G_{ph}v_{h,C} = G_{vv}G_{hh}v_{p,C} \quad (99)$$

$$G_{hh}G_{pv}v_{v,H} + G_{vv}G_{ph}v_{h,H} = G_{vv}G_{hh}v_{p,H} \quad (100)$$

$$G_{hh}G_{pv}v_{v,CH} + G_{vv}G_{ph}v_{h,CH} = G_{vv}G_{hh}v_{p,CH} \quad (101)$$

$$G_{hh}G_{mv}v_{v,C} + G_{vv}G_{mh}v_{h,C} = G_{vv}G_{hh}v_{m,C} \quad (102)$$

$$G_{hh}G_{mv}v_{v,H} + G_{vv}G_{mh}v_{h,H} = G_{vv}G_{hh}v_{m,H} \quad (103)$$

$$G_{hh}G_{mv}v_{v,CH} + G_{vv}G_{mh}v_{h,CH} = G_{vv}G_{hh}v_{m,CH}. \quad (104)$$

Multiplying (99) by $v_{v,H}/v_{v,C}$ and subtracting (100) leaves

$$G_{ph}(v_{h,C}v_{v,H}/v_{v,C} - v_{h,H}) = G_{hh}(v_{p,C}v_{v,H}/v_{v,C} - v_{p,H}). \quad (105)$$

Similarly, from (99) and (101), we obtain

$$G_{ph}(v_{h,C}v_{v,CH}/v_{v,C} - v_{h,CH}) = G_{hh}(v_{p,C}v_{v,CH}/v_{v,C} - v_{p,CH}). \quad (106)$$

Solve (106) for

$$G_{hh} = G_{ph} \frac{(v_{h,C}v_{v,CH}/v_{v,C} - v_{h,CH})}{(v_{p,C}v_{v,CH}/v_{v,C} - v_{p,CH})} \quad (107)$$

then substitute it into (105) to obtain

$$G_{ph}(v_{h,C}v_{v,H}/v_{v,C} - v_{h,H}) = G_{ph} \frac{(v_{h,C}v_{v,CH}/v_{v,C} - v_{h,CH})}{(v_{p,C}v_{v,CH}/v_{v,C} - v_{p,CH})} (v_{p,C}v_{v,H}/v_{v,C} - v_{p,H}). \quad (108)$$

This last equation cannot be solved for G_{ph} —it merely reveals redundancy in the data and that (105) and (106) are redundant constraint equations.

Similarly, from (102) and (103), we obtain

$$G_{mh}(v_{h,C}v_{v,H}/v_{v,C} - v_{h,H}) = G_{hh}(v_{m,C}v_{v,H}/v_{v,C} - v_{m,H}) \quad (109)$$

or an expression that is numerically equivalent from (102) and (104).

A similar procedure produces constraint equations for G_{pv} and G_{mv} in terms of G_{vv} . In summary, we end up with the four constraint equations

$$G_{pv} = G_{vv} \frac{(v_{p,C}v_{v,H} - v_{h,C}v_{p,H})}{(v_{v,C}v_{v,H} - v_{h,C}v_{v,H})} \quad (110)$$

$$G_{ph} = G_{hh} \frac{(v_{p,C}v_{v,H} - v_{v,C}v_{p,H})}{(v_{h,C}v_{v,H} - v_{v,C}v_{h,H})}$$

$$G_{mv} = G_{vv} \frac{(v_{m,C}v_{v,H} - v_{h,C}v_{m,H})}{(v_{v,C}v_{v,H} - v_{h,C}v_{v,H})}$$

$$G_{mh} = G_{hh} \frac{(v_{m,C}v_{v,H} - v_{v,C}v_{m,H})}{(v_{h,C}v_{v,H} - v_{v,C}v_{h,H})} \quad (111)$$

or two alternative (but numerically equivalent) sets obtained by finding, for example, the equation relating G_{ph} and G_{hh} from

$$\begin{aligned}
& \frac{G_{pv}G_{mU} - G_{pU}G_{mv}}{G_{vv}}v_{v,CN} + \frac{G_{ph}G_{mU} - G_{pU}G_{mh}}{G_{hh}}v_{h,CN} - G_{mU}v_{p,CN} + G_{pU}v_{m,CN} \\
& = \frac{G_{pv}G_{mU} - G_{pU}G_{mv}}{G_{vv}}G_{vv}TT1 + \frac{G_{ph}G_{mU} - G_{pU}G_{mh}}{G_{hh}}G_{hh}TT2 \\
& \quad - G_{mU}(G_{pv}TT1 + G_{ph}TT2 + G_{pU}TCN) + G_{pU}(G_{mv}TT1 + G_{mh}TT2 + G_{mU}TCN)
\end{aligned} \tag{112}$$

(99) and (101) or from (100) and (101). Thus, of (99)–(101), any one is completely redundant, and the same goes for (102)–(104). Hence, rather than six equations in six unknowns, we only have four equations in six unknowns.

Last constraint equation: Using the last eigenvector with $\lambda = 0$, as derived in Appendix B1, the final constraint equation is expressed in (112), shown at the top of the page. The entire right side cancels itself out. Then, multiplying both sides by $G_{vv}G_{hh}$

$$\begin{aligned}
& (G_{pv}G_{mU} - G_{pU}G_{mv})G_{hh}v_{v,CN} + \\
& (G_{ph}G_{mU} - G_{pU}G_{mh})G_{vv}v_{h,CN} - \\
& G_{mU}G_{vv}G_{hh}v_{p,CN} + G_{pU}G_{vv}G_{hh}v_{m,CN} = 0. \tag{113}
\end{aligned}$$

Solving this for

$$\begin{aligned}
G_{mU} & = G_{pU} \\
& \cdot \frac{G_{mv}G_{hh}v_{v,CN} + G_{mh}G_{vv}v_{h,CN} - G_{vv}G_{hh}v_{m,CN}}{G_{pv}G_{hh}v_{v,CN} + G_{ph}G_{vv}v_{h,CN} - G_{vv}G_{hh}v_{p,CN}}.
\end{aligned} \tag{114}$$

REFERENCES

- [1] J. R. Piepmeier, "Calibration of passive microwave polarimeters that use hybrid coupler-based correlators," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 2, pp. 391–400, Feb. 2004.
- [2] J. R. Piepmeier, 2006. private communication.
- [3] E. T. Jaynes, *Probability Theory: The Logic of Science*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
- [4] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive*, vol. 1. Norwood, MA: Artech House, 1981.
- [5] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive*, vol. 2. Norwood, MA: Artech House, 1986, ch. 7, pp. 476–494.
- [6] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*. Philadelphia, PA: SIAM, 2005, pp. 44–49, 173, and 179–180. [Online]. Available: <http://www.ipgp.jussieu.fr/%7Etarantola/Files/Professional/SIAM/index.html>
- [7] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991, p. 175, 229.
- [8] D. L. Hudson, J. R. Piepmeier, and D. G. Long, "Polarization rotation correction in radiometry: An error analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3212–3223, Oct. 2007.



Derek Hudson received the B.S. and M.S. degrees in electrical engineering from Brigham Young University (BYU), Provo, UT, in 2003, where he is currently working toward the Ph.D. degree.

He has been with the Microwave Earth Remote Sensing (MERS) research group, BYU, since 1999. He is also currently with the NASA Goddard Space Flight Center, Greenbelt, MD. His research experience is in the application of inverse problem theory and in radar hardware design.



David G. Long (S'80–M'82–SM'98–F'07) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1989.

From 1983 to 1990, he was with the Jet Propulsion Laboratory (JPL), National Aeronautics and Space Administration (NASA), where he developed advanced radar remote sensing systems. While at JPL he was the Project Engineer on the NASA Scatterometer (NSCAT) project, which flew from 1996 to 1997. He also managed the SCANSAT project, the precursor to SeaWinds, which was launched in 1999 and 2002. He is currently a Professor with the Department of Electrical and Computer Engineering, Brigham Young University, Provo, UT, where he teaches upper-division and graduate courses in communications, microwave remote sensing, radar, and signal processing, and is the Director of the BYU Center for Remote Sensing. He is the Principal Investigator on several NASA-sponsored research projects in remote sensing. He is the author or coauthor of more than 275 publications in signal processing and radar scatterometry. His research interests include microwave remote sensing, radar theory, space-based sensing, estimation theory, signal processing, and mesoscale atmospheric dynamics.

Dr. Long is an Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. He has received the NASA Certificate of Recognition several times.