NPL Report CMSC 33/03

Report to the National Measurement System Directorate, Department of Trade and Industry

From the Software Support for Metrology Programme

The comparison of algorithms for the assessment of Type A1 surface texture reference artefacts

By
A B Forbes and P M Harris
Centre for Mathematics and Scientific Computing
R K Leach
Centre for Basic, Length and Thermal Metrology

December 2003

The comparison of algorithms for the assessment of Type A1 surface texture reference artefacts

A B Forbes and P M Harris
Centre for Mathematics and Scientific Computing
R K Leach
Centre for Basic, Length and Thermal Metrology

December 2003

ABSTRACT

The dimensional assessment of a Type A1 surface texture reference artefact involves estimating the depth of a groove by fitting a pair of nominally parallel lines to measurement data and then calculating the separation between the lines. The evaluation of the uncertainty associated with this estimate is important and needs to take into account i) the statistical model for the measurement data including correlation between the measured x and z values and ii) the algorithm for determining an estimate of the depth. In this report, we analyse a range of least squares regression algorithms that can be applied to determining the geometry of these artefacts for a range of data uncertainty structures. The analysis shows that while for completely general uncertainty matrices the algorithms can produce markedly different results, for realistic data and uncertainty models, their behaviour is generally very similar. As a consequence, we can conclude it is safe to use the more straightforward approaches, so long as appropriate attention is given to the uncertainty associated with the fitted parameters. The report describes a number of analysis techniques and algorithm testing methods that have quite wide application.

This work has been funded by the National Measurement System Programmes for Length and Software Support for Metrology.

ISSN 1471–0005

Authorised by Dr Dave Rayner,
Head of the Centre for Mathematics and Scientific Computing

National Physical Laboratory,
Queens Road, Teddington, Middlesex, United Kingdom TW11 0LW

Contents

1 Introduction

ISO 5436 Part 1 [8] advocates the use of artefacts with specific surface geometries to calibrate surface texture measuring instruments. These artefacts can themselves be calibrated using a surface texture measuring instrument that has traceable metrology in the vertical and horizontal axes [9]. Many stylus-based instruments are calibrated using magnification standards that take the form of rectangular grooves in a substrate. These are useful for checking the vertical magnification factor of an instrument but do not give metrological information regarding horizontal magnification or the transmission characteristics (frequency response) of an instrument. ISO 5436 Part 1 and BS 1134 [2] refer to these artefacts as Type A1 and describe the measurement process as follows:

> A continuous straight mean line equal in length to three times the width of the groove is drawn over the groove to represent the upper level of the surface and another to represent the lower level, both lines extending symmetrically about the centre of the groove (see figure 1.1). To avoid the influence of any rounding of the corners, the upper surface on each side of the groove is to be ignored for a length equal to one-third of the width of the groove. The surface at the bottom of the groove is assessed only over the central third of its width. The portions to be used for assessment purposes are therefore those at A, B, and C in Figure 1. The depth d of the groove shall be assessed perpendicularly from the upper mean line to the mid-point of the lower mean line. A significant number, not less than five, of evenly distributed traces shall be taken.

The dimensional assessment of a Type A1 surface texture reference artefact therefore involves fitting a pair of nominally parallel lines to measurement data and then calculating their separation. The evaluation of the uncertainty associated with this estimate of the depth is important and needs to take into account i) the statistical model for the measurement data including correlation between the measured x and z values, and ii) the algorithm for determining an estimate of the separation. In this report, we analyse a range of least squares regression algorithms that can be applied to determining the geometry of these artefacts for a range of data uncertainty structures. The analysis shows that while for completely general uncertainty matrices the algorithms can produce markedly different results, for realistic data and uncertainty models, their behaviour is generally very similar. As a consequence, we can conclude it is safe to use the more straightforward approaches, so long as appropriate attention is given to the uncertainty

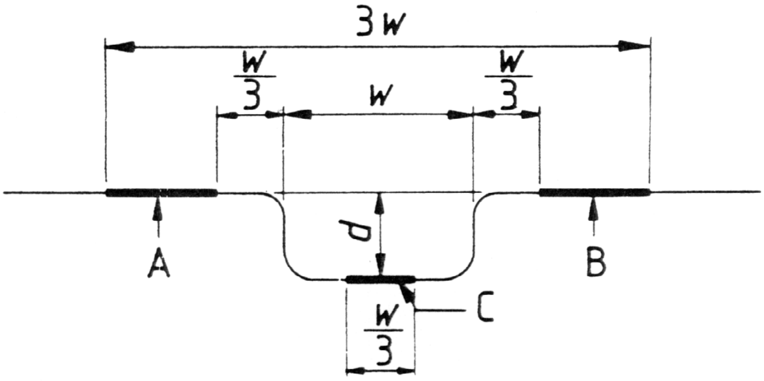

Figure 1: Measurement strategy for the assessment of a Type A1 artefact according to ISO 5346 Part 1.

associated with the estimated parameters. The report describes a number of analysis techniques and algorithm testing methods that have quite wide application.

The report is organised as follows. In section 2, we present the mathematical model associated with the measurement of a Type A1 artefact. In section 2.4, we provide a general scheme for how the model parameters, including the separation parameter of direct interest, along with the associated uncertainties, can be determined from measurement data. In section 3, we show how this general approach can be applied to the regression problem in hand and adapted to provide effective estimation methods for different uncertainty structures. The analysis of the various algorithms is presented in section 4. The analysis is further developed using experimental data in section 5. Our concluding remarks are given in section 6.

2 Mathematical model

2.1 Geometry of the artefact

The nomimal geometry of the artefact is described by two parallel lines L_k,

$$\left.\begin{array}{lcll} L_1: & z & = & a_1 + bx, \quad x \in \mathcal{C} \\ L_2: & z & = & a_2 + bx, \quad x \in \mathcal{A} \text{ or } x \in \mathcal{B} \end{array}\right\}, \tag{1}$$

where $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ correspond to the regions indicated in Figure 1, so that $\mathcal{A}$ and $\mathcal{B}$ define the ranges on the upper line segment, $\mathcal{C}$ on the lower. Given $C < A < B < D$, we define a more general model function

$$\phi(x, \mathbf{b}) = \left\{\begin{array}{ll} a_1 + bx, & x \in (A, B), \\ a_2 + bx, & \text{otherwise}, \end{array}\right. \tag{2}$$

where $\mathbf{b} = (a_1, a_2, b)^{\mathrm{T}}$. Of primary interest is the orthogonal separation d of the two lines:

$$d = \frac{a_2 - a_1}{\sqrt{1 + b^2}}.$$

2.2 Measurement data

We assume that the measurement data consists of two sets of data points $X_k = \{(x_i, z_i)^{\mathrm{T}}, i \in I_k\}$ with the kth set representing m_k data points gathered from L_k, $k = 1, 2$. The measured data is subject to random effects $\boldsymbol{\epsilon} = (\ldots, \epsilon_i, \ldots)^{\mathrm{T}}$, so that

$$\left[\begin{array}{c} \mathbf{x} \\ \mathbf{z} \end{array}\right] = \left[\begin{array}{c} \mathbf{x}^* \\ \mathbf{z}^* \end{array}\right] + \boldsymbol{\epsilon}, \quad z_i^* = \phi(x_i^*, \mathbf{b}), \quad \boldsymbol{\epsilon} \sim N(\mathbf{0}, V). \tag{3}$$

With this model we allow for a completely general correlation structure for the measurement data.

2.3 Example system: interferometric transducers

Suppose the x and z coordinates are measured using interferometric displacement transducers. The coordinates are determined from interferometric fringe counts which measure the optical displacement corrected for refractive index effects to provide the geometric displacements. For this measurement system, the following model is appropriate:

$$\left.\begin{array}{rcl} x_i & = & x_i^* + \alpha_i + \gamma_i + \omega, \\ z_i & = & z_i^* + \beta_i + \gamma_i + \omega, \\ \alpha_i & \sim & N(0, \nu_i^2), \\ \beta_i & \sim & N(0, \sigma_i^2), \\ \gamma_i & \sim & N(0, \rho_i^2), \\ \omega & \sim & N(0, \tau^2), \end{array}\right\} \qquad (4)$$

where α_i and β_i represent random effects particular to measurements x_i and z_i, respectively, γ_i a random effect common to x_i and z_i and ω a random effect common to all the measurements.

Let D be the $(3m+1) \times (3m+1)$ diagonal matrix with ν_i in the ith diagonal element, σ_i in the $(m+i)$th diagonal element, ρ_i in the $(2m+i)$th diagonal element and τ in the $(3m+1)$th diagonal element. Then D^2 is the uncertainty matrix (covariance matrix) associated with $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and ω. Let G be the $2m \times (3m+1)$ matrix

$$G = \left[\begin{array}{cccc} I_m & & I_m & \mathbf{1}_m \\ & I_m & I_m & \mathbf{1}_m \end{array}\right]$$

where I_m is the $m \times m$ identity matrix and $\mathbf{1}$ the $m \times 1$ vector of 1's. Then the data uncertainty matrix V associated with the measurements $\mathbf{x}$ and $\mathbf{z}$ is

$$V = GD^2G^{\mathrm{T}}.$$

The uncertainty matrix V_{ij} associated with the four measurements (x_i, z_i, x_j, z_j) is given by the appropriate (re-ordered) submatrix of V:

$$V_{ij} = \left[\begin{array}{cccc} \nu_i^2 + \rho_i^2 + \tau^2 & \rho_i^2 + \tau^2 & \tau^2 & \tau^2 \\ \rho_i^2 + \tau^2 & \sigma_i^2 + \rho_i^2 + \tau^2 & \tau^2 & \tau^2 \\ \tau^2 & \tau^2 & \nu_j^2 + \rho_j^2 + \tau^2 & \rho_j^2 + \tau^2 \\ \tau^2 & \tau^2 & \rho_j^2 + \tau^2 & \sigma_j^2 + \rho_j^2 + \tau^2 \end{array}\right].$$

2.4 Solution estimates of the model parameters

Estimates of the model parameters $\mathbf{b}$ associated with the model (3) can be found by solving the generalised Gauss-Markov problem [5]

$$\min_{\mathbf{b},\mathbf{x}^*} \left[\begin{array}{c} \mathbf{x} - \mathbf{x}^* \\ \mathbf{z} - \mathbf{z}^* \end{array}\right]^{\mathrm{T}} V^{-1} \left[\begin{array}{c} \mathbf{x} - \mathbf{x}^* \\ \mathbf{z} - \mathbf{z}^* \end{array}\right].$$

In the next section we consider algorithms to solve this problem for different types of structure in the uncertainty matrix V taking into account the simple form of the function $z = \phi(x, \mathbf{b})$ in (2).

3 Algorithms generalised Gauss-Markov regression

We consider first the nonlinear Gauss-Markov regression problem [5]. Let $\mathbf{f}(\mathbf{a})$ be m functions f_i of n parameters $\mathbf{a} = (a_1, \ldots, a_n)^{\mathrm{T}}$ and $V = LL^{\mathrm{T}}$ an $m \times m$ semi-positive definite matrix specified in terms of the factor L. The matrix V represents the uncertainty matrix associated with the quantities $\mathbf{f}$. The *nonlinear Gauss-Markov regression problem* is

$$\min_{\mathbf{a},\boldsymbol{\eta}} \quad \boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{\eta} \quad \text{subject to} \quad \mathbf{f}(\mathbf{a}) = L\boldsymbol{\eta}. \tag{5}$$

If L and hence V is invertible the constraints define $\boldsymbol{\eta}$ as $\boldsymbol{\eta} = L^{-1}\mathbf{f}(\mathbf{a})$ and (5) is equivalent, mathematically, to

$$\min_{\mathbf{a}} \quad \mathbf{f}(\mathbf{a})^{\mathrm{T}} V^{-1} \mathbf{f}(\mathbf{a}). \tag{6}$$

3.1 Nonlinear least squares

If $V = I$, then (5) becomes a straightforward nonlinear least squares problem

$$\min_{\mathbf{a}} \sum_{i=1}^{m} f_i^2(\mathbf{a}), \tag{7}$$

and the Gauss-Newton algorithm can be applied: given estimates $\mathbf{a}$ of the solution, we i) calculate $\mathbf{f} = \mathbf{f}(\mathbf{a})$ and corresponding $m \times n$ Jacobian matrix $J_{ij}(\mathbf{a}) = \frac{\partial f_i}{\partial a_j}$ of partial derivatives, ii) solve

$$J\mathbf{p} = -\mathbf{f}, \tag{8}$$

in the least squares sense for $\mathbf{p}$, and iii) update $\mathbf{a} := \mathbf{a} + \mathbf{p}$. These steps are repeated until the algorithm is judged to have converged. The solution $\mathbf{p}$ in (8) can be found by solving the $n \times n$ system of normal equations

$$J^{\mathrm{T}} J\mathbf{p} = -J^{\mathrm{T}}\mathbf{f}.$$

However, the formation and use of $J^{\mathrm{T}}J$ is likely to worsen any ill-conditioning in the problem. A better approach is to use a QR factorisation approach [4, 7] to factor $J = QR$ as the product of an $m \times m$ orthogonal matrix Q (so that $Q^{\mathrm{T}}Q = QQ^{\mathrm{T}} = I$, the $m \times m$ indenty matrix) with an $m \times n$ upper triangular matrix[1]

$$R = \left[\begin{array}{c} R_1 \\ \mathbf{0} \end{array} \right].$$

[1] Geometrically, we can regard the column vectors of J as vectors in $m-$dimensional space. The QR factorisation matrix produces an orthogonal matrix Q (essentially a product of rotations and reflections) such that when we apply Q^{T} to J the first column vector is transformed to lie along the first co-ordinate axis, the second is transformed to lie in the plane containing the first and second co-ordinate axes, etc.

If

$$-Q^{\mathrm{T}}\mathbf{f} = \begin{bmatrix} \hat{\mathbf{h}}_1 \\ \hat{\mathbf{h}}_2 \end{bmatrix},$$

the update step is computed by solving the upper triangular system $R_1\mathbf{p} = \hat{\mathbf{h}}_1$.

For invertible V, we use the factorisation $V = LL^{\mathrm{T}}$ and set

$$\tilde{\mathbf{f}}(\mathbf{a}) = L^{-1}\mathbf{f}(\mathbf{a}).$$

The Gauss-Newton algorithm can now be applied to minimise $\tilde{\mathbf{f}}(\mathbf{a})^{\mathrm{T}}\tilde{\mathbf{f}}(\mathbf{a})$ so that at each iteration we are required to solve, in the least squares sense,

$$L^{-1}J\mathbf{p} = -L^{-1}\mathbf{f}, \tag{9}$$

where $J_{ij}(\mathbf{a}) = \frac{\partial f_i}{\partial a_j}$, as before. If L is well-conditioned this approach represents a satisfactory approach to solving (5). Otherwise, the formation of L^{-1} and $\tilde{J} = L^{-1}J$ can lead to numerical instability. (If L is singular this approach cannot be applied at all.) However, the linear least squares problem (9) is equivalent to

$$\min_{\boldsymbol{\eta},\mathbf{p}} \boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{\eta} \quad \text{subject to } J\mathbf{p} + L\boldsymbol{\eta} = -\mathbf{f}. \tag{10}$$

This problem can be solved stably using the *generalised QR factorisation* (GQR) [1, 5, 10]. This approach also uses the QR factorisation of J and the RQ factorisation of the $m \times m$ matrix $Q^{\mathrm{T}}L$ so that

$$Q^{\mathrm{T}}L = TU,$$

where

$$T = \begin{bmatrix} T_{11} & T_{12} \\ & T_{22} \end{bmatrix},$$

is $m \times m$ upper triangular and U is $m \times m$ orthogonal. We multiply the equation $-\mathbf{f} = J\mathbf{p} + L\boldsymbol{\eta}$ by Q^{T} and partition the result as

$$\begin{bmatrix} \hat{\mathbf{h}}_1 \\ \hat{\mathbf{h}}_2 \end{bmatrix} = \begin{bmatrix} R_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} T_{11} & T_{12} \\ & T_{22} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\eta}}_1 \\ \hat{\boldsymbol{\eta}}_2 \end{bmatrix}, \tag{11}$$

where

$$\begin{bmatrix} \hat{\mathbf{h}}_1 \\ \hat{\mathbf{h}}_2 \end{bmatrix} = -Q^{\mathrm{T}}\mathbf{f}, \quad \begin{bmatrix} \hat{\boldsymbol{\eta}}_1 \\ \hat{\boldsymbol{\eta}}_2 \end{bmatrix} = U\boldsymbol{\eta}.$$

The bottom set of equations in (11) defines $\hat{\boldsymbol{\eta}}_2$ as the solution of

$$T_{22}\hat{\boldsymbol{\eta}}_2 = \hat{\mathbf{h}}_2,$$

but we are free to assign $\hat{\boldsymbol{\eta}}_1$. We minimise $\boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{\eta} = \hat{\boldsymbol{\eta}}^{\mathrm{T}}\hat{\boldsymbol{\eta}}$ by setting $\hat{\boldsymbol{\eta}}_1 = \mathbf{0}$. The update step $\mathbf{p}$ therefore solves

$$R_1\mathbf{p} = \hat{\mathbf{h}}_1 - T_{12}\hat{\boldsymbol{\eta}}_2, \quad T_{22}\hat{\boldsymbol{\eta}}_2 = \hat{\mathbf{h}}_2.$$

The solution $\boldsymbol{\eta}$ is given by

$$\boldsymbol{\eta} = U^{\mathrm{T}} \left[\begin{array}{c} \mathbf{0} \\ \hat{\boldsymbol{\eta}}_2 \end{array} \right].$$

A modified Gauss-Newton algorithm can therefore be applied to solve (5) which at each iteration solves (10) to determine the update step $\mathbf{p}$. This algorithm can be also be applied in the case where L is not full rank.

3.2 Evaluation of the uncertainty matrix

For a nonlinear least squares problem (7), the uncertainty matrix $V_{\mathbf{a}}$ associated with the fitted parameters $\mathbf{a}$ is given by

$$V_{\mathbf{a}} = (J^{\mathrm{T}}J)^{-1},$$

where J is the Jacobian matrix of partial derivatives evaluated at the solution.

If $J = QR$, then $V_{\mathbf{a}} = (R^{\mathrm{T}}R)^{-1}$. If a generalised QR factorisation approach is used to solve (6), then

$$V_{\mathbf{a}} = KK^{\mathrm{T}},$$

where K solves $R_1K = T_{11}$ [5]. If the input uncertainty matrix V is known only approximately, we can set

$$V_{\mathbf{a}} = \hat{\sigma}^2(J^{\mathrm{T}}J)^{-1},$$

where is $\hat{\sigma}$ is an estimate of the standard deviation of the residuals. One such estimate is given by

$$\hat{\sigma}^2 = \mathbf{f}^{\mathrm{T}}\mathbf{f}/(m-n). \tag{12}$$

The quantity $\hat{\sigma}$ can be regarded as a scale correction to input uncertainty arrived at following the least squares analysis of the data ([3]). We refer to $\hat{\sigma}$ as a *posterior* estimate of the the standard deviation.

3.3 Generalised Gauss-Markov regression with curves

We now consider Gauss-Markov regression in the context of curve fitting of which (3) is a particular example. Let $z = \phi(x, \mathbf{b})$ be a curve specified by parameters $\mathbf{b}$ and suppose we have data $X = \{(x_i, z_i)\}_{i=1}^m$ nominally lying on such a curve but subject to random effects with general uncertainty matrix $V = LL^{\mathrm{T}}$. The generalised Gauss-Markov regression problem is

$$\min_{\boldsymbol{\eta},\mathbf{x}^*,\mathbf{b}} \boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{\eta} \text{ subject to } \left[\begin{array}{c} \mathbf{x} \\ \mathbf{z} \end{array}\right] = \left[\begin{array}{c} \mathbf{x}^* \\ \mathbf{z}^* \end{array}\right] + L\boldsymbol{\eta}, \qquad (13)$$

where

$$z_i^* = \phi(x_i^*, \mathbf{b}), \quad i = 1, \ldots, m.$$

We use the qualification *generalised* to emphasise the fact that we allow for the possibility of measurements of x as well as those of z being subject to random effects. Setting

$$\mathbf{a} = \left[\begin{array}{c} \mathbf{x}^* \\ \mathbf{b} \end{array}\right], \quad \mathbf{f}(\mathbf{a}) = \left[\begin{array}{c} \mathbf{x} - \mathbf{x}^* \\ \mathbf{z} - \mathbf{z}^* \end{array}\right],$$

the problem of finding the best-fit curve is formulated as a nonlinear Gauss-Markov regression problem (5). The Jacobian matrix associated with $\mathbf{f}(\mathbf{a})$ is the $2m \times (m + n)$ matrix J with

$$-J = \left[\begin{array}{cc} I & \mathbf{0} \\ J_x & J_{\mathbf{b}}^* \end{array}\right], \qquad (14)$$

where $J_{\mathbf{b}}^*$ is the $m \times n$ matrix with $J_{\mathbf{b}}^*(i,j) = \partial\phi(x_i^*, \mathbf{b})/\partial b_j$ and J_x is the diagonal matrix with $J_x(i,i) = \partial\phi(x_i^*, \mathbf{b})/\partial x$.

We now consider how different types of structure in the uncertainty matrix V lead to generally simpler estimation methods.

3.4 Ordinary linear regression

In this case, the measurements z_i of the heights are subject to uncorrelated random effects but the locations x_i are assumed to be known accurately. In this situation,

$$V = \left[\begin{array}{cc} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & V_{22} \end{array}\right],$$

where V_{22} is the diagonal matrix with $\sigma_i^2 > 0$ in the ith diagonal element. The corresponding optimisation problem is

$$\min_{\mathbf{b}} \sum_{i=1}^m w_i^2 (z_i - \phi(x_i, \mathbf{b}))^2,$$

with $w_i = 1/\sigma_i$ and this problem can be solved directly using a QR factorisation approach.

3.5 Ordinary Guass-Markov linear regression

In this case, the heights z_i are subject to correlated random effects with covariance matrix $V_{22} = L_{22}L_{22}^{\mathrm{T}}$ but the locations x_i are assumed to be known accurately. In this situation,

$$V = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & V_{22} \end{bmatrix},$$

and the corresponding optimisation problem is

$$\min_{\mathbf{b},\boldsymbol{\eta}} \; \boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{\eta} \;\text{ subject to }\; \mathbf{z} = \mathbf{z}^* + L_{22}\boldsymbol{\eta}.$$

If $\phi(x, \mathbf{b})$ is linear in $\mathbf{b}$ so that $\mathbf{z}^* = \mathbf{z}_0 + A\mathbf{b}$, then this optimisation problem can be solved directly using a GQR factorisation approach.

3.6 Generalised distance regression

In this case, the heights z_i and locations x_i are subject to correlated random effects, but the ith set of measurements are independent of the jth, $i \neq j$. In this situation, the observations can be ordered such that V is composed of diagonal 2×2 blocks $V_i = L_iL_i^{\mathrm{T}}$ and the regression problem can be formulated as

$$\min_{\mathbf{a},\{\boldsymbol{\eta}_i\}} \; \sum_{i=1}^{m} \boldsymbol{\eta}_i^{\mathrm{T}}\boldsymbol{\eta}_i$$

subject to the constraints

$$\begin{bmatrix} x_i - x_i^* \\ z_i - \phi(x_i^*, \mathbf{b}) \end{bmatrix} = L_i\boldsymbol{\eta}_i, \quad i = 1, \ldots, m.$$

This problem can be solved efficiently using a separation of variables approach in which the x_i^* are determined from the solutions of the corresponding footpoint problems:

$$\min_{\boldsymbol{\eta}_i, x_i^*} \; \boldsymbol{\eta}_i^{\mathrm{T}}\boldsymbol{\eta}_i$$

subject to the constraints

$$\begin{bmatrix} x_i - x_i^* \\ z_i - \phi(x_i^*, \mathbf{b}) \end{bmatrix} = L_i\boldsymbol{\eta}_i.$$

For the case of model (2), for fixed $\mathbf{b}$ the constraints are linear in the parameter x_i^* and so the footpoint problem can be solved directly using the

GQR algorithm. Setting $\mathbf{n}_i = (-\dot{\phi}_i, 1)^{\mathrm{T}}/(1+\dot{\phi}_i^2)^{1/2}$, $\dot{\phi}_i = \partial\phi/\partial x$ evaluated at x_i^*, and $s_i = \|L_i^{\mathrm{T}}\mathbf{n}_i\|$, then

$$d_i(\mathbf{b}) = \frac{1}{s_i}\mathbf{n}_i^{\mathrm{T}} \left[\begin{array}{c} x_i - x_i^* \\ z_i - z_i^* \end{array} \right], \tag{15}$$

and

$$\frac{\partial d_i}{\partial b_j} = \frac{1}{s_i}\mathbf{n}_i^{\mathrm{T}} \left[\begin{array}{c} 0 \\ -\frac{\partial \phi}{\partial b_j}(x_i^*) \end{array} \right]. \tag{16}$$

The best-fit curve is found from solving

$$\min_{\mathbf{b}} \sum_{i=1}^{m} d_i^2(\mathbf{b})$$

using the Gauss-Newton algorithm with d_i and its derivatives calculated according to (15) and (16). We note that at no point in the calculation is L_i^{-1} required. In fact, this approach can be used effectively even if L_i is rank deficient.

3.7 Example system: interferometric transducers

We refer back to the model for interferometric transducers described in section 2.3. If $\nu = \rho = \tau = 0$, only the z_i measurements are subject to random effects and the resulting fitting problem is an ordinary linear least squares regression problem. If $\tau = 0$, the fitting problem is a generalised distance regression problem. If $\tau > 0$, the problem is a generalised Gauss-Markov regression problem.

4 Analysis of algorithm behaviour

In this section, we look at the behaviour of different parameter estimation methods for simulated data. Given an uncertainty matrix V for the measurement data, we know that the generalised Gauss Markov (GGM) estimator has optimal behaviour in the sense that it uses the input uncertainty matrix V to 'weight' the input information in the most appropriate way. However its implementation, even for the simple models considered here, requires optimisation components considerably more involved than those used for ordinarly least squares regression. We are therefore interested in comparing the behaviour of different estimation algorithms relative to the generalised Gauss Markov approach.

4.1 Monte Carlo simulation data generation

We generate test data according to the model (13) as follows.

Suppose the $2m \times 2m$ data uncertainty matrix $V = LL^{\mathrm{T}}$ has been specified.

i Fix end points $C < A < B < D$, parameters $\mathbf{b}^\sharp$ and abscissae $\mathbf{x}^\sharp = (x_1^\sharp, \ldots, x_m^\sharp)^{\mathrm{T}}$, $C \le x_i^\sharp \le D$.

ii Generate heights $\mathbf{z}^\sharp$ corresponding to the nominal geometry so that $z_i^\sharp = \phi(x_i^\sharp, \mathbf{b}^\sharp)$, $i = 1, \ldots, m$.

iii Evaluate the $2m \times (m+3)$ Jacobian matrix J for parameters $\mathbf{x}^\sharp$ and $\mathbf{b}^\sharp$ and form the generalised QR factorisation for the pair $[J,\ L]$:

$$J = Q \begin{bmatrix} R_1 \\ \mathbf{0} \end{bmatrix}, \quad Q^{\mathrm{T}} L = \begin{bmatrix} T_{11} & T_{12} \\ & T_{22} \end{bmatrix} Z.$$

iv Evaluate the $(m+3) \times (m+3)$ matrix

$$V_{\mathbf{a}}^\sharp = KK^{\mathrm{T}}, \quad R_1 K = T_{11}.$$

$V_{\mathbf{a}}^\sharp$ is the uncertainty matrix associated with the fitted parameters $\mathbf{x}^\sharp$ and $\mathbf{b}^\sharp$ for the GGM estimator and the lower right 3×3 submatrix of $V_{\mathbf{a}}^\sharp$ is the uncertainty matrix $V_{\mathbf{b}}^\sharp$ associated with the parameters $\mathbf{b}$.

v Generate a $2m \times 1$ vector $\boldsymbol{\eta}$ whose elements are drawn from a standard normal distribution, i.e., $\eta_i \sim N(0,1)$ and set $\boldsymbol{\epsilon} = L\boldsymbol{\eta}$. Then $\boldsymbol{\epsilon} \sim N(\mathbf{0}, V)$.

vi Set

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{x}^\sharp \\ \mathbf{z}^\sharp \end{bmatrix} + \boldsymbol{\epsilon}.$$

Steps v and vi can be repeated N times to generate data sets $X_q = \{(x_{i,q}, z_{i,q})\}$, $q = 1, \ldots, N$, such that the GGM estimates $\mathbf{a}_q$ of the parameters associated with X_q are approximately distributed as

$$\mathbf{a}_q \sim N(\mathbf{0}, V_{\mathbf{a}}^{\sharp}), \quad \mathbf{b}_q \sim N(\mathbf{0}, V_{\mathbf{b}}^{\sharp}).$$

4.2 Generation of data uncertainty matrices

4.2.1 General covariance structure

We first consider the case in which the measurement errors have a completely general covariance structure. We assume that we have m data points so that V is $(2m)\times(2m)$. Any covariance matrix V has an eigenvalue decomposition of the form

$$V = US^2U^{\mathrm{T}},$$

where S is a diagonal matrix with $s_{ii} \geq 0$ and U is an orthogonal matrix. The eigenvalues of V are the squares of the diagonal elements of S and the eigenvectors are the columns of U. The jth eigenvalue is the variance of the linear combination of the random variables specified by the jth eigenvector. A zero eigenvalue corresponds to perfect correlation in a such a linear combination. By choosing an orthogonal matrix at random and eigenvalues in a specified range, we can generate covariance matrices with a range of properties. In the experiments reported on here, we set

$$s_{ii} = e^{-c(i-1)/(2m)}, \quad i = 1, \ldots, 2m,$$

where $c \geq 0$ is an input parameter that controls the amount of correlation: the larger the value of c, the more correlation. We note that such a V can be decomposed as $V = LL^{\mathrm{T}}$ where $L = US$.

Figures 2–4 graph the square roots of the diagonal elements of the data uncertainty matrix V generated with correlation parameters $c = 2$, 4 and 6, respectively, while Figures 5–7 provide histograms of the off-diagonal elements of the associated correlation matrices. For this dataset, $m = 75$. Each colour corresponds to a different column of the correlation matrix. As c increases, the correlation coefficients tend to become larger in absolute value.

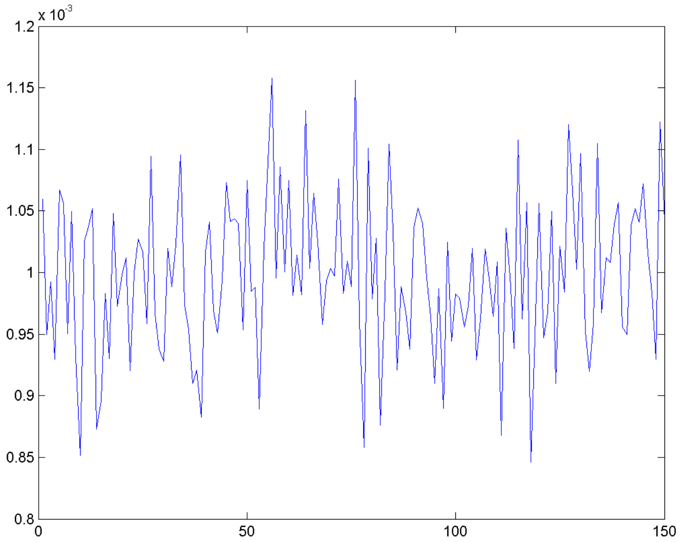

Figure 2: Square root of the diagonal elements of data uncertainty matrix V generated with $c = 2$. The units in the vertical axis are arbitrary.

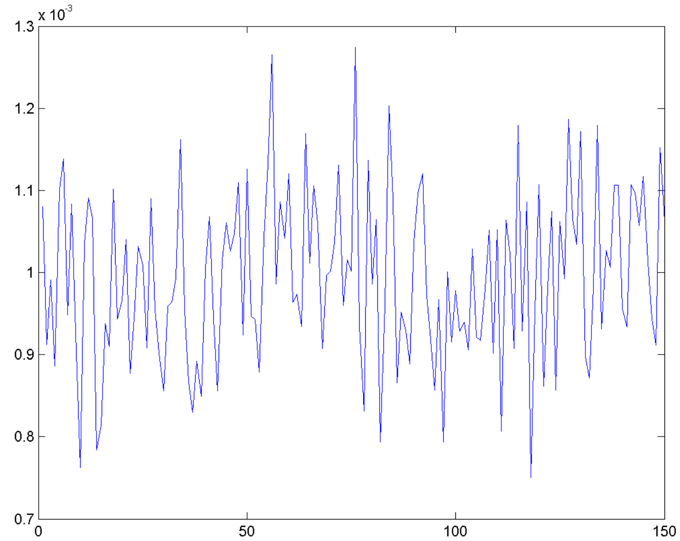

Figure 3: As Figure 2 but generated with $c = 4$.

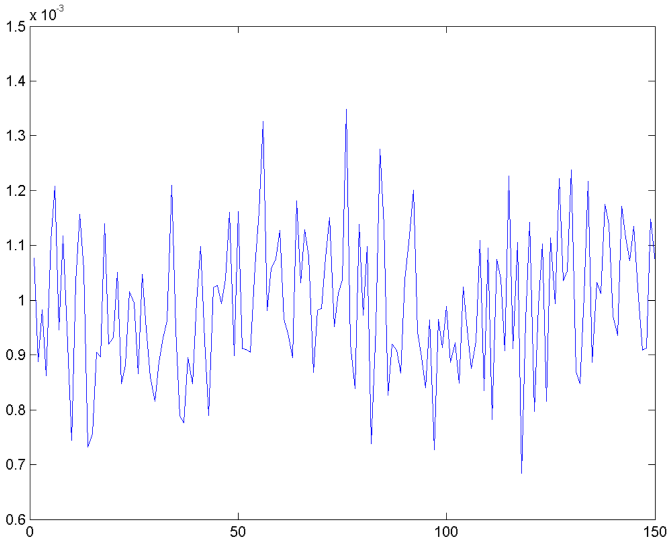

Figure 4: As Figure 2 but generated with $c = 6$.

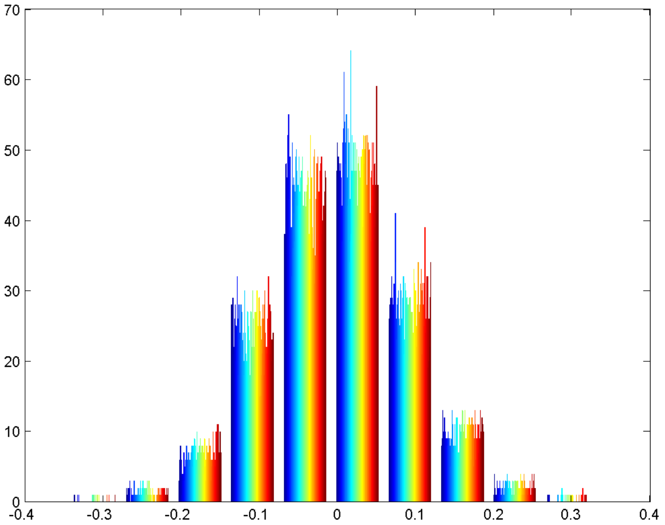

Figure 5: Histogram of the off-diagonal elements of the correlation matrix associated with the data uncertainty matrix V generated with $c = 2$. Each colour corresponds to a different column of the correlation matrix. The vertical axis gives is the number of elements in each column that fall into the range specified on the horizontal axis. (Each column has $2m = 150$ elements.)

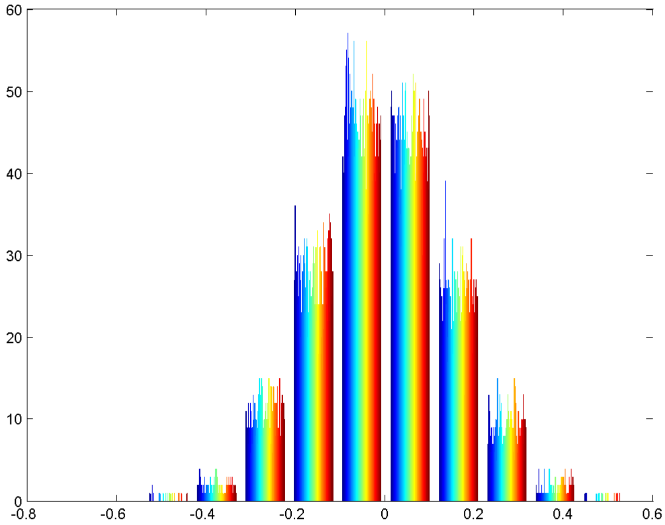

Figure 6: As Figure 5 but with $c = 4$.

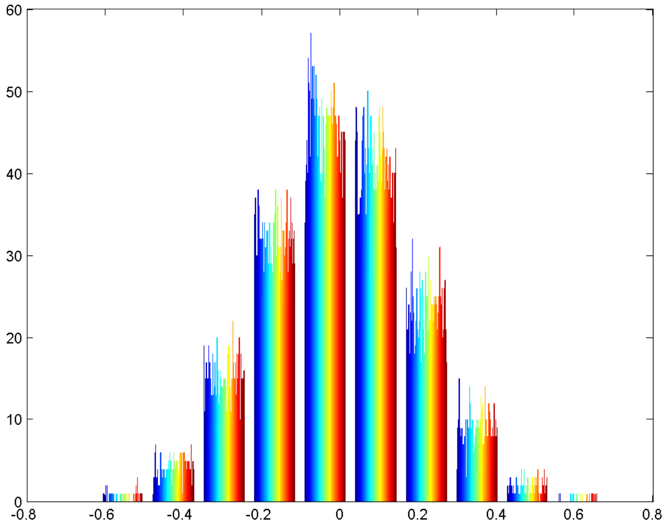

Figure 7: As Figure 5 but with $c = 6$.

4.2.2 Estimation algorithms for general uncertainty matrices

We apply the following algorithms to data generated for general V.

LLS linear least squares, in which we solve

$$\min_{\mathbf{b}} \sum_i (z_i - \phi(x_i, \mathbf{b}))^2.$$

WLS weighted linear least squares, in which we solve

$$\min_{\mathbf{b}} \sum_i w_i^2 (z_i - \phi(x_i, \mathbf{b}))^2,$$

with $w_i = 1/\sigma_i$ where σ_i is the standard uncertainty associated with z_i, i.e., $\sigma_i^2 = V_{m+i,m+i}$.

ODR orthogonal distance regression, in which we solve

$$\min_{\mathbf{x}^*, \mathbf{b}} \sum_i \{(x_i - x_i^*)^2 + (z_i - \phi(x_i^*, \mathbf{b}))^2\}.$$

WDR weighted distance regression, in which we solve

$$\min_{\mathbf{x}^*, \mathbf{b}} \sum_i \{v_i^2 (x_i - x_i^*)^2 + w_i^2 (z_i - \phi(x_i^*, \mathbf{b}))^2\}.$$

with w_i as above and $v_i = 1/\nu_i$ where ν_i is the standard uncertainty associated with x_i, i.e., $\nu_i^2 = V_{ii}$.

GDR general distance regression, in which we solve

$$\min_{\mathbf{a}, \{\boldsymbol{\eta}_i\}} \sum_{i=1}^{m} \boldsymbol{\eta}_i^{\mathrm{T}} \boldsymbol{\eta}_i$$

subject to the constraints

$$\left[\begin{array}{c} x_i - x_i^* \\ z_i - \phi(x_i^*, \mathbf{b}) \end{array} \right] = L_i \boldsymbol{\eta}_i, \quad i = 1, \ldots, m,$$

where $V_i = L_i L_i^{\mathrm{T}}$ is the uncertainty matrix associated with the measurements x_i and z_i, i.e.,

$$V_i = \left[\begin{array}{cc} V_{i,i} & V_{i,m+i} \\ V_{m+i,i} & V_{m+i,m+i} \end{array} \right].$$

GM Gauss-Markov regression in which we solve

$$\min_{\boldsymbol{\eta},\mathbf{b}} \boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{\eta} \text{ subject to } \mathbf{z} = A\mathbf{b} + L_{22}\boldsymbol{\eta},$$

where the ith row of the $m \times 3$ matrix A is

$$A(i,1:3) = \left\{ \begin{array}{lll} (1,0,x_i), & x_i & \in & (A,B), \\ (0,1,x_i), & x_i & \notin & (A,B), \end{array} \right. \tag{17}$$

and L_{22} is an $m \times m$ matrix such that $V_{22} = L_{22}L_{22}^{\mathrm{T}}$. Here, V_{22} is the covariance matrix associated with $\mathbf{z}$, i.e.,

$$V_{22} = V(m+1:2m, m+1:2m).$$

GGM generalised Gauss-Markov regression in which we solve

$$\min_{\boldsymbol{\eta},\mathbf{x}^*,\mathbf{b}} \boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{\eta} \text{ subject to } \begin{bmatrix} \mathbf{x} \\ \mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{x}^* \\ \mathbf{z}^* \end{bmatrix} + L\boldsymbol{\eta},$$

where $z_i^* = \phi(x_i^*, \mathbf{b})$. We note that $\mathbf{z}^* = A^*\mathbf{b}$ where A^* is the $m \times 3$ matrix

$$A^*(i,1:3) = \left\{ \begin{array}{lll} (1,0,x_i^*), & x_i^* & \in & (A,B), \\ (0,1,x_i^*), & x_i^* & \notin & (A,B). \end{array} \right.$$

Of these algorithms, GGM is expected to perform best as it reflects accurately the covariance structure in the data.

All these algorithms are forms of least squares regression and the evaluation of the uncertainty matrix associated with the fitted parameters can be performed following the general approach outlined in section 3.2.

4.2.3 Covariance matrices associated with the interferometric model

The covariance matrix associated with the interferometric model (4) is determined by the three m-vectors $\boldsymbol{\nu}$, $\boldsymbol{\sigma}$ and $\boldsymbol{\rho}$ and the scalar τ that specify the diagonal matrix D in $V = GD^2G^{\mathrm{T}}$. In the simulations reported on below, we first set three scalars ν_0, σ_0 and ρ_0, and then set

$$\nu_i = \nu_0(1+\delta_i^\nu), \quad \sigma_i = \sigma_0(1+\delta_i^\sigma), \quad \rho_i = \nu_0(1+\delta_i^\rho), \tag{18}$$

where δ_i^ν, δ_i^σ, $\delta_i^\rho \in [0,\ 1]$ are drawn from a rectangular (uniform) distribution.

If the $(3m+1) \times 2m$ matrix DG^{T} has QR factorisation $DG^{\mathrm{T}} = QR$, and R_1 is the submatrix of R comprising the first $2m$ rows, then $V = LL^{\mathrm{T}}$ where $L = R_1^{\mathrm{T}}$.

4.2.4 Estimation algorithms for interferometric model data uncertainty matrices

In order to analyse the algorithm behaviour for uncertainty matrices that arise in practice, we apply the same set of algorithms to data generated for a data uncertainty matrix V associated with the interferometric model as those described in section 4.2.3, but with the following modifications:

WLS weighted linear least squares, defined in terms of weights $w_i = 1/\sigma_i$ where σ_i is as in (18).

WDR weighted distance regression, defined in terms of weights v_i and w_i with w_i as above and $v_i = 1/\nu_i$ where ν_i is defined in (18).

GDR general distance regression, using the ith covariance matrix

$$V_i = \left[\begin{array}{cc} \nu_i^2 + \rho_i^2 & \rho_i^2 \\ \rho_i & \sigma_i^2 + \rho_i^2 \end{array} \right].$$

with ν_i, σ_i and ρ_i defined by (18).

These three algorithms differ from their general covariance counterparts in that they correspond to the assumptions that $\nu_i = \rho_i = \tau = 0$, $\rho_i = \tau = 0$, and $\tau = 0$, respectively. For the general covariance matrix case they were defined in terms of the corresponding elements of V.

4.3 Monte Carlo simulation results

The following graphs record the normalised estimates $(d - d^\sharp)/u_{GGM}$ where

$$d = (a_2 - a_1)/(1 + b^2)^{1/2},$$

is the estimated depth, $d^\sharp$ is the depth determined by $\mathbf{b}^\sharp$ and u_{GGM} is the standard uncertainty associated with d for the estimate determined by the GGM algorithm. We therefore expect approximately 57% of the results associated with GGM to fall within ±1. Each simulation run involved 100 data sets ($N = 100$).

4.3.1 General covariance matrix

Nine simulation runs were made on data generated as described in section 4.2.1 with $\mathbf{b}^\sharp = (0, 2, 1)^{\mathrm{T}}$, $(0, 2, -1)^{\mathrm{T}}$ and $(0, 2, 0.001)^{\mathrm{T}}$ and $c = 2$, 4 and 6. Graphs 8–16 show:

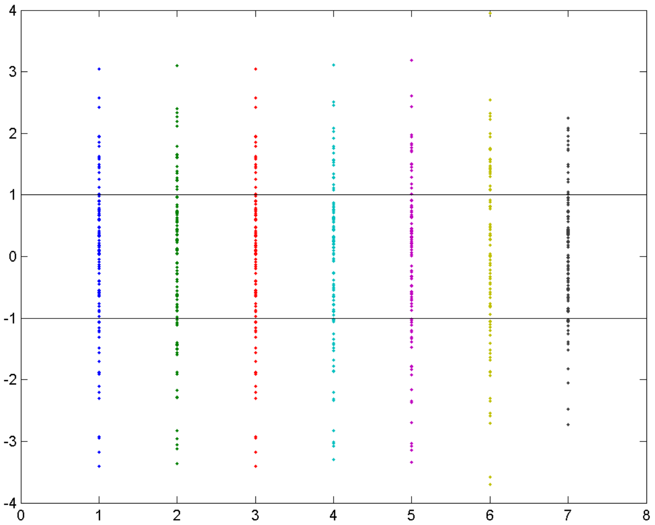

Figure 8: Normalised estimates $(d - d^\sharp)/u_{GGM}$ calculated by 1) LLS, 2) WLS, 3) ODR, 4) WDR, 5) GDR, 6) GM and 7) GGM algorithms for 100 simulated data sets with $\mathbf{b}^\sharp = (0, 2, 1)^{\mathrm{T}}$ and $c = 2$. The units on the vertical axis have been normalised so that the standard uncertainty ($k = 1$) associated with the GGM estimate of the separation d is 1.

1. The behaviour of algorithms LLS, WLS, ODR, WDR and GDR are similar with respect to variation in their estimates of the separation parameter d.

2. The behaviour of algorithm GM depends on the value of the slope parameter b. For b small, it gives similar estimates to that supplied by the GGM algorithm. For $b = \pm 1$, its estimates show the largest variation.

3. As c increases, corresponding to a greater degree of correlation, the performance of the GGM algorithm becomes better relative to the others.

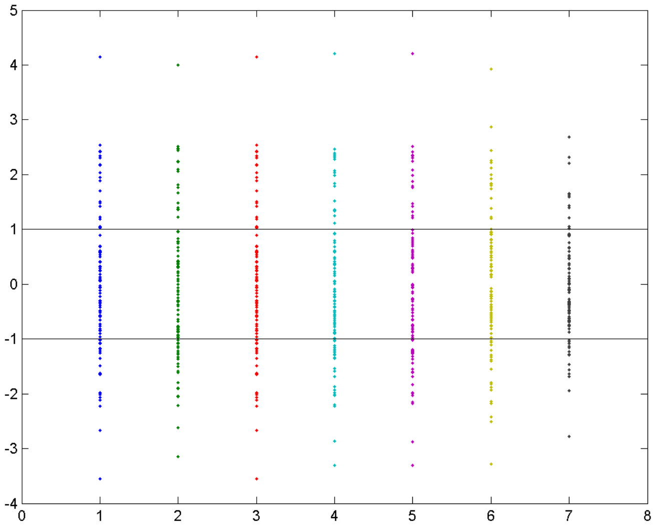

Figure 9: As Figure 8 but with $\mathbf{b}^\sharp = (0, 2, -1)^\mathrm{T}$.

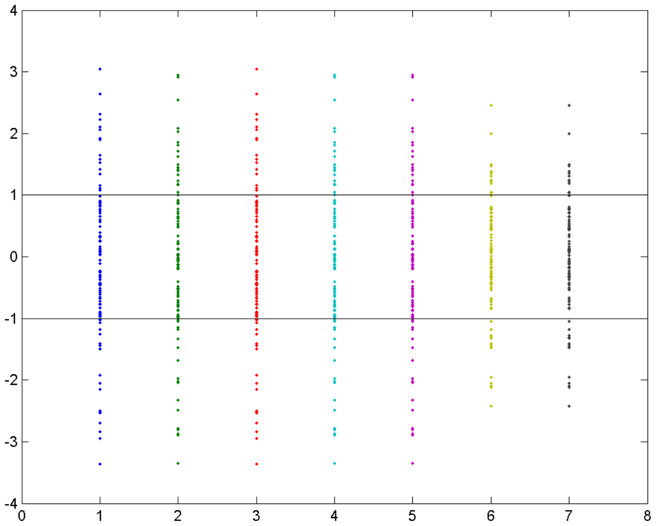

Figure 10: As Figure 8 but with $\mathbf{b}^\sharp = (0, 2, 0.001)^\mathrm{T}$.

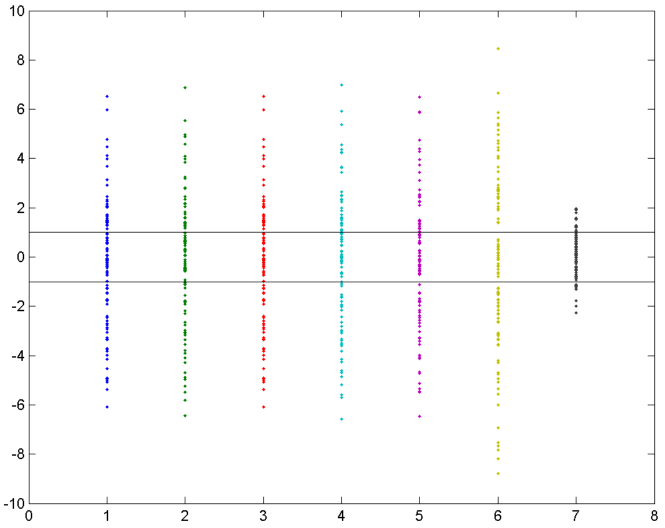

Figure 11: As Figure 8 but with $\mathbf{b}^\sharp = (0, 2, 1)^\mathrm{T}$ and $c = 4$.

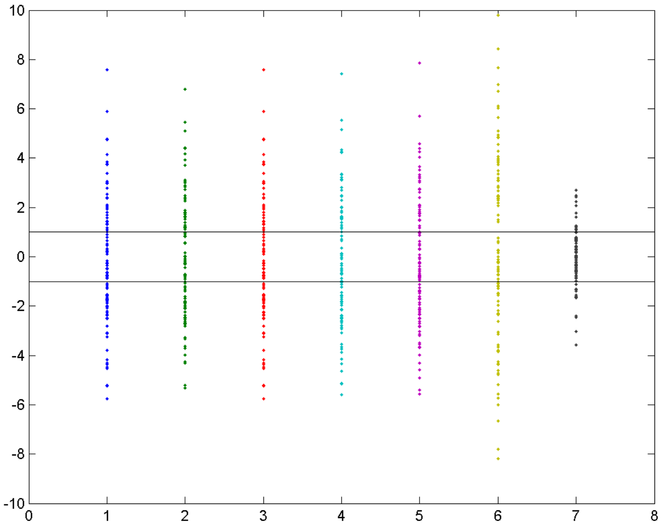

Figure 12: As Figure 8 but with $\mathbf{b}^\sharp = (0, 2, -1)^\mathrm{T}$ and $c = 4$.

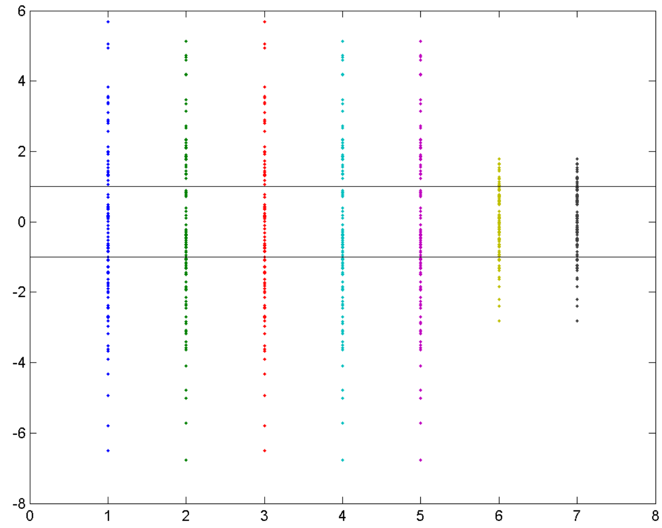

Figure 13: As Figure 8 but with $\mathbf{b}^\sharp = (0, 2, 0.001)^\mathrm{T}$ and $c = 4$.

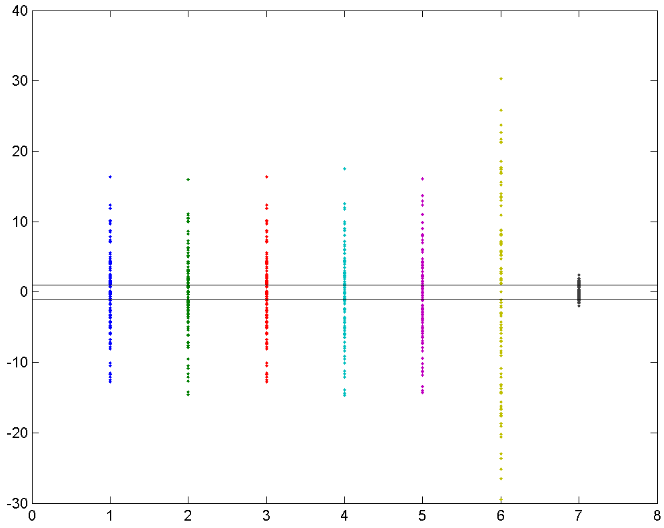

Figure 14: As Figure 8 but with $\mathbf{b}^\sharp = (0, 2, 1)^\mathrm{T}$ and $c = 6$.

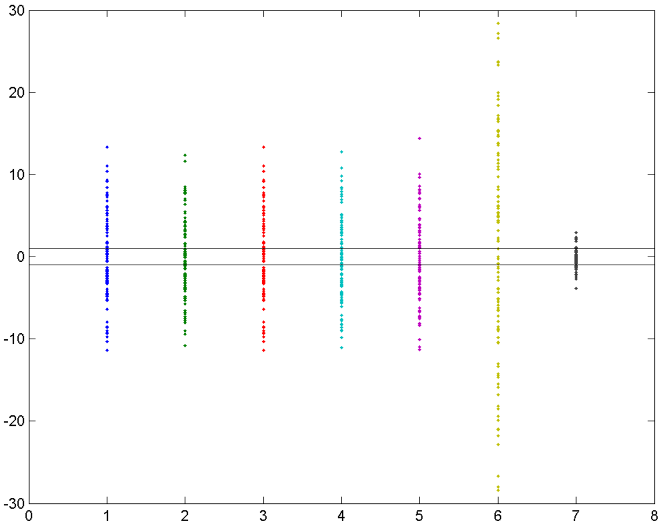

Figure 15: As Figure 8 but with $\mathbf{b}^\sharp = (0, 2, -1)^{\mathrm{T}}$ and $c = 6$.

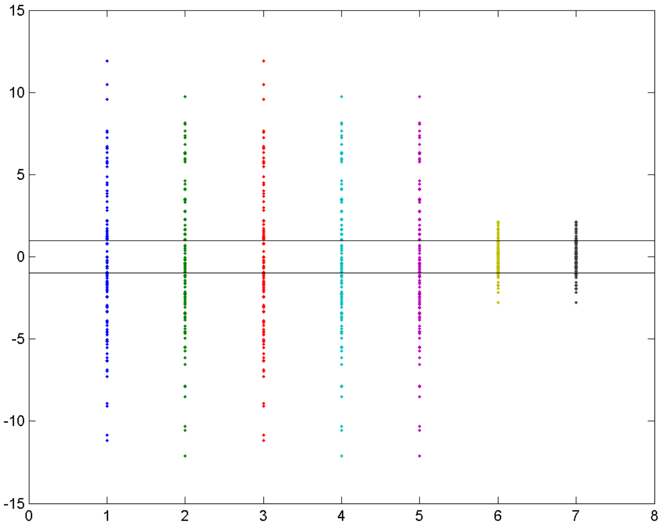

Figure 16: As Figure 8 but with $\mathbf{b}^\sharp = (0, 2, 0.001)^{\mathrm{T}}$ and $c = 6$.

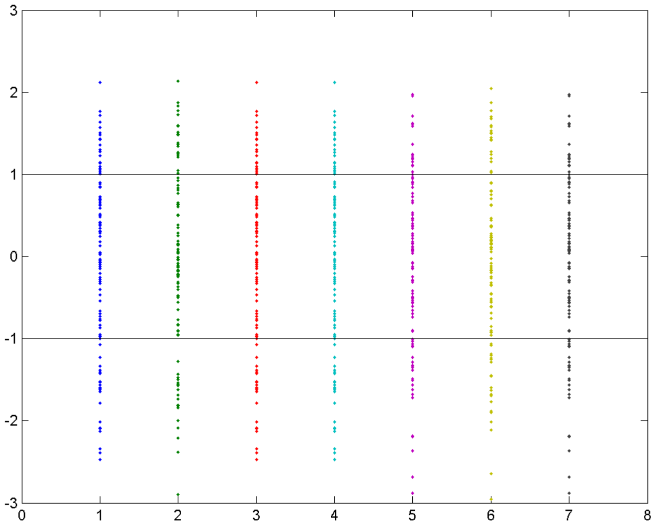

Figure 17: Normalised estimates $(d - d^\sharp)/u_{GGM}$ calculated by 1) LLS, 2) WLS, 3) ODR, 4) WDR, 5) GDR, 6) GM and 7) GGM algorithms for 100 simulated data sets with $\mathbf{b}^\sharp = (0, 2, 1)^{\mathrm{T}}$ and $(\nu_0, \sigma_0, \rho_0, \tau) = (0.001, 0.001, 0.001, 0.001)$. The units on the vertical axis have been normalised so that the standard uncertainty ($k = 1$) associated with the GGM estimate of the separation d is 1.

4.3.2 Interferometric model

Six simulation runs were made for data generated as described in section 4.2.3 with $\mathbf{b}^\sharp = (0, 2, 1)^{\mathrm{T}}$, $(0, 2, -1)^{\mathrm{T}}$ and $(0, 2, 0.001)^{\mathrm{T}}$ (as above) and

$$(\nu_0, \sigma_0, \rho_0, \tau) = (0.001, 0.001, 0.001, 0.001)$$

and

$$(\nu_0, \sigma_0, \rho_0, \tau) = (0.002, 0.001, 0.005, 0.003).$$

Graphs 17–22 show that all the algorithms have a similar behaviour for this type of data, at least in terms of their estimates of the separation parameter d.

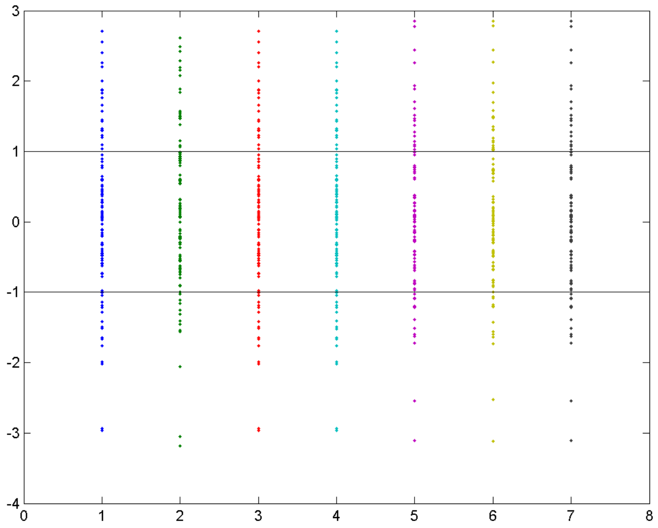

Figure 18: As Figure 17 but with $\mathbf{b} = (0, 2, -1)^{\mathrm{T}}$.

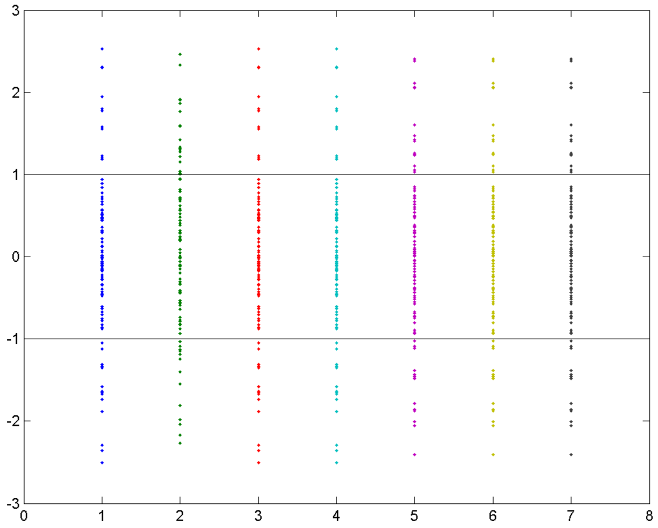

Figure 19: As Figure 17 but with $\mathbf{b} = (0, 2, 0.001)^{\mathrm{T}}$.

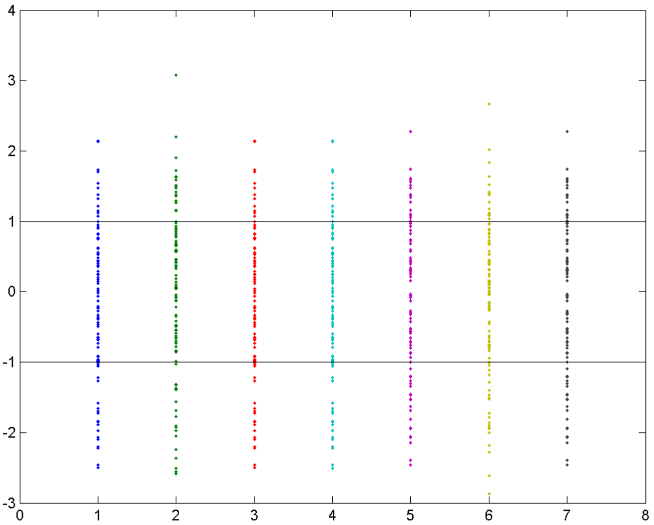

Figure 20: As Figure 17 but with $(\nu_0, \sigma_0, \rho_0, \tau) = (0.002, 0.001, 0.005, 0.003)$.

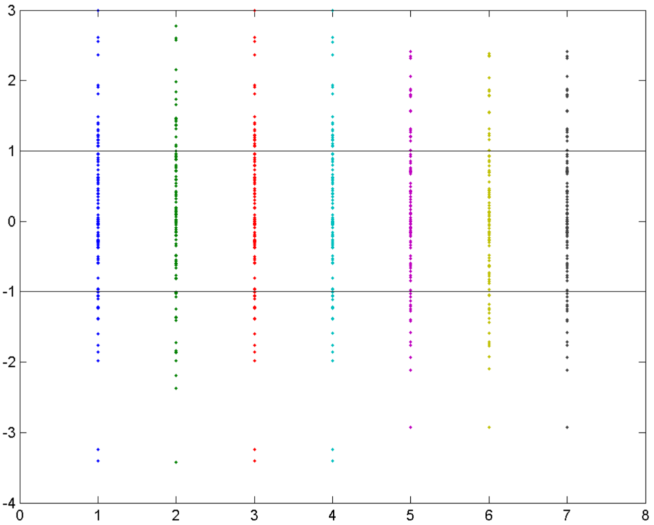

Figure 21: As Figure 17 but with $\mathbf{b} = (0, 2, -1)^{\mathrm{T}}$ and $(\nu_0, \sigma_0, \rho_0, \tau) = (0.002, 0.001, 0.005, 0.003)$.

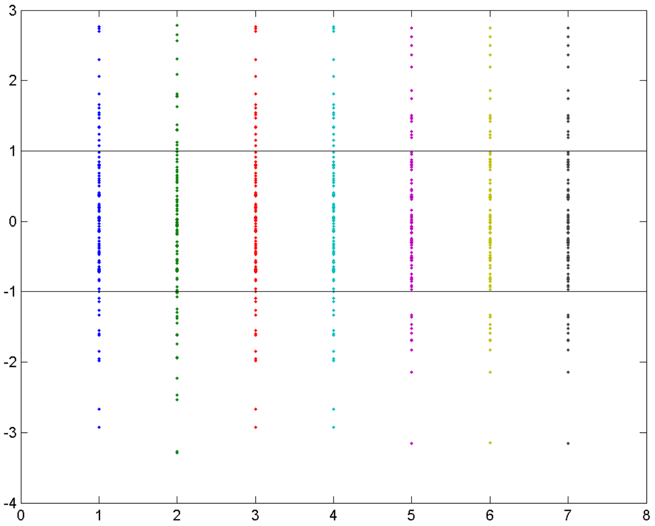

Figure 22: As Figure 17 but with $\mathbf{b} = (0, 2, 0.001)^{\mathrm{T}}$ and $(\nu_0, \sigma_0, \rho_0, \tau) = (0.002, 0.001, 0.005, 0.003)$.

4.4 Null-space data generation scheme

In order to compare more directly the approximate algorithms with the GGM algorithm we can apply the approximate algorithms on data for which the GGM solution is known.

We wish to generate test data for the problem of generalised Gauss-Markov regression with curves (13). We assume that the uncertainty matrix V is full rank. A first order condition for $\mathbf{b}$ and $\mathbf{x}^*$ to define a local minimum is that

$$J^{\mathrm{T}} V^{-1} \left[\begin{array}{c} \mathbf{x} - \mathbf{x}^* \\ \mathbf{z} - \mathbf{z}^* \end{array} \right] = \mathbf{0},$$

where, as before, $z_i^* = \phi(x_i^*, \mathbf{b})$ and J is the Jacobian matrix defined in (14). Therefore, if

$$J^{\mathrm{T}} \boldsymbol{\delta} = \mathbf{0},$$

and

$$\left[\begin{array}{c} \mathbf{x} \\ \mathbf{z} \end{array} \right] = \left[\begin{array}{c} \mathbf{x}^* \\ \mathbf{z}^* \end{array} \right] + V \boldsymbol{\delta},$$

then these optimality conditions are satisfied. Below we describe a data generation scheme based on the analysis above that uses the generalised QR factorisation to improve the efficiency and numerical stability.

Suppose $2m \times 2m$ uncertainty matrix $V = LL^{\mathrm{T}}$ has been specified. The simulation data is generated according to the following scheme.

i Fix end points $C < A < B < D$, parameters $\mathbf{b}^\sharp$ and abscissae $\mathbf{x}^\sharp = (x_1^\sharp, \ldots, x_m^\sharp)^{\mathrm{T}}$, $C \le x_i^\sharp \le D$.

ii Generate heights $\mathbf{z}^\sharp$ corresponding to the nominal geometry so that $z_i^\sharp = \phi(x_i^\sharp, \mathbf{b}^\sharp)$, $i = 1, \ldots, m$.

iii Evaluate $2m \times (m+3)$ Jacobian matrix J for parameters $\mathbf{x}^\sharp$ and $\mathbf{b}^\sharp$ and form the generalised QR factorisation for the pair $[J,\ L]$:

$$J = Q \left[\begin{array}{c} R_1 \\ \mathbf{0} \end{array} \right], \quad Q^{\mathrm{T}} L = \left[\begin{array}{cc} T_{11} & T_{12} \\ & T_{22} \end{array} \right] U.$$

iv Evaluate the $(m+3) \times (m+3)$ matrix

$$V_{\mathbf{a}}^\sharp = KK^{\mathrm{T}}, \quad R_1 K = T_{11}.$$

v Generate at random a $m-3$ vector $\boldsymbol{\zeta}$ and normalise it so that $\boldsymbol{\zeta}^{\mathrm{T}} \boldsymbol{\zeta} = m-3$. Set

$$\boldsymbol{\epsilon} = Q \left[\begin{array}{c} T_{12} \\ T_{22} \end{array} \right] \boldsymbol{\zeta}, \quad \boldsymbol{\eta} = U^{\mathrm{T}} \left[\begin{array}{c} \mathbf{0} \\ \boldsymbol{\zeta} \end{array} \right].$$

vi Set

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{x}^\sharp \\ \mathbf{z}^\sharp \end{bmatrix} + \boldsymbol{\epsilon}.$$

Then $\mathbf{b}^\sharp$ is the GGM estimate associated with the data $\mathbf{x}$ and $\mathbf{z}$ and $V_{\mathbf{a}}^\sharp$ is the uncertainty matrix associated with the fitted parameters $\mathbf{x}^\sharp$ and $\mathbf{b}^\sharp$ for the GGM estimator. The lower right 3×3 submatrix of $V_{\mathbf{a}}^\sharp$ is the required uncertainty matrix $V^\sharp$ associated with the parameters $\mathbf{b}$. From $\mathbf{b}^\sharp$ and $V_{\mathbf{a}}^\sharp$ we can calculate the separation $d^\sharp$ and its associated uncertainty $u^\sharp$.

Steps v and vi can be repeated any number of times if required.

With this data generation scheme we can compare approximate estimate methods with the GGM estimator to see if their performance is acceptable.

4.5 Null space simulations

4.5.1 General covariance matrix

Nine simulation runs were made on null-space data for covariance matrices generated as in section 4.2.1 with $\mathbf{b}^\sharp = (0, 2, 1)^{\mathrm{T}}$, $(0, 2, -1)^{\mathrm{T}}$ and $(0, 2, 0.001)^{\mathrm{T}}$ and $c = 2$, 4 and 6. Graphs 23–31 show:

1. As for the Monte Carlo simulation data, as c increases the variations in the parameter estimates by the approximate algorithms increase.
2. For slope $b = 0.001$, the GM estimate is close to the GGM solution $d^\sharp$, relative to the uncertainty in $d^\sharp$.

We can explain the last property as follows. Suppose $\mathbf{b}$ and $\mathbf{x}^*$ solve the GGM problem for data $\mathbf{x}$ and $\mathbf{z}$ and uncertainty matrix V. Then there is a $\boldsymbol{\delta}$ such that

$$\begin{bmatrix} I & bI \\ \mathbf{0} & (A^*)^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \end{bmatrix} = \mathbf{0}, \quad \begin{bmatrix} \mathbf{x} - \mathbf{x}^* \\ \mathbf{z} - \mathbf{z}^* \end{bmatrix} = \begin{bmatrix} V_{11}\boldsymbol{\delta}_1 + V_{12}\boldsymbol{\delta}_2 \\ V_{21}\boldsymbol{\delta}_1 + V_{22}\boldsymbol{\delta}_2 \end{bmatrix}, \tag{19}$$

where A^* is the observation matrix associated with $\mathbf{x}^*$. Similarly, for $\mathbf{b}$ to be a solution of the GM problem for data $\mathbf{x}$ and uncertainty matrix V_{22}, we must have

$$A^{\mathrm{T}} V_{22}^{-1}(\mathbf{z} - \mathbf{z}^*) = \mathbf{0},$$

where A is the observation matrix associated with $\mathbf{x}$. We write $A = A^* + \Delta$ where $\|\Delta\| = O(\|\mathbf{x} - \mathbf{x}^*\|) = O(\|\boldsymbol{\delta}\|)$, i.e., the difference between A and A^*

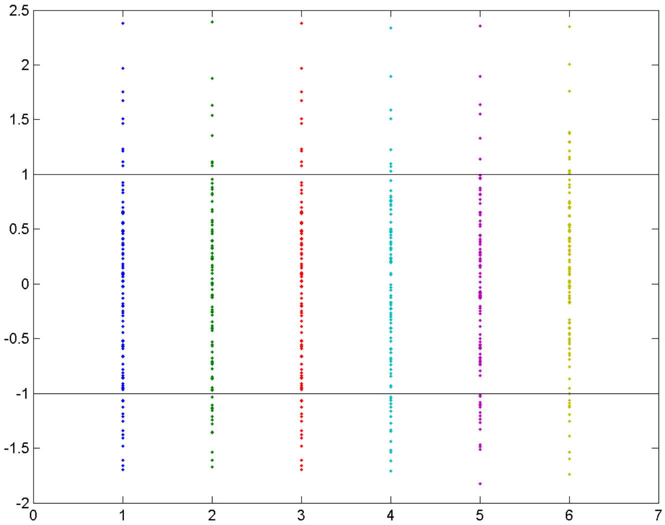

Figure 23: Normalised estimates $(d - d^\sharp)/u_{GGM}$ calculated by 1) LLS, 2) WLS, 3) ODR, 4) WDR, 5) GDR and 6) GM algorithms for 100 simulated data sets with $\mathbf{b}^\sharp = (0, 2, 1)^{\mathrm{T}}$ and $c = 2$. The units on the vertical axis have been normalised so that the standard uncertainty ($k = 1$) associated with the GGM estimate of the separation d is 1.

is of the same order of magnitude as that of $\mathbf{x} - \mathbf{x}^*$ and $\boldsymbol{\delta}$. We have from (19) that $\boldsymbol{\delta}_1 = -b\boldsymbol{\delta}_2$ (and is zero if b is zero) and

$$\mathbf{z} - \mathbf{z}^* = V_{22}\boldsymbol{\delta}_2 - bV_{21}\boldsymbol{\delta}_2,$$

so that

$$\begin{aligned} A^{\mathrm{T}}V_{22}^{-1}(\mathbf{z} - \mathbf{z}^*) &= A^{\mathrm{T}}(\boldsymbol{\delta}_2 - bV_{22}^{-1}V_{12}\boldsymbol{\delta}_2), \\ &= \Delta^{\mathrm{T}}(\boldsymbol{\delta}_2 - bV_{22}^{-1}V_{12}\boldsymbol{\delta}_2) - b(A^*)^{\mathrm{T}}V_{22}^{-1}V_{12}\boldsymbol{\delta}_2. \end{aligned}$$

The first term is $O(\|\boldsymbol{\delta}\|^2)$ while the second term is $O(b\|\boldsymbol{\delta}\|)$. From this, we expect that the difference between the GM and GGM to vary linearly with the size of the noise in the data for slopes significantly different from zero, but quadratically if $b = 0$.

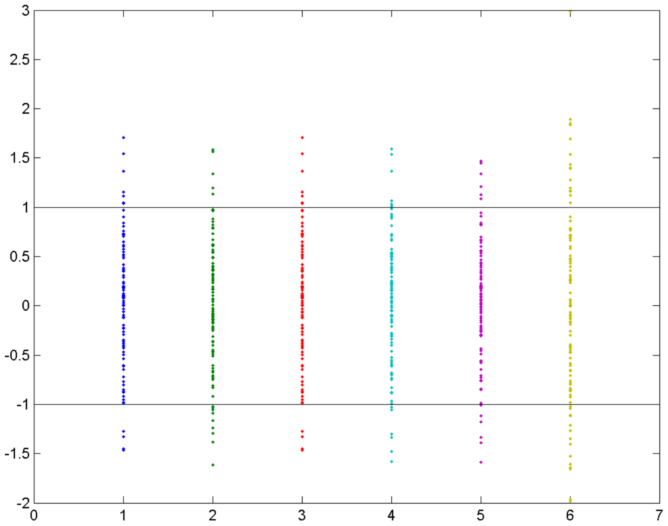

Figure 24: As Figure 23 but with $\mathbf{b} = (0, 2, 1)^{\mathrm{T}}$.

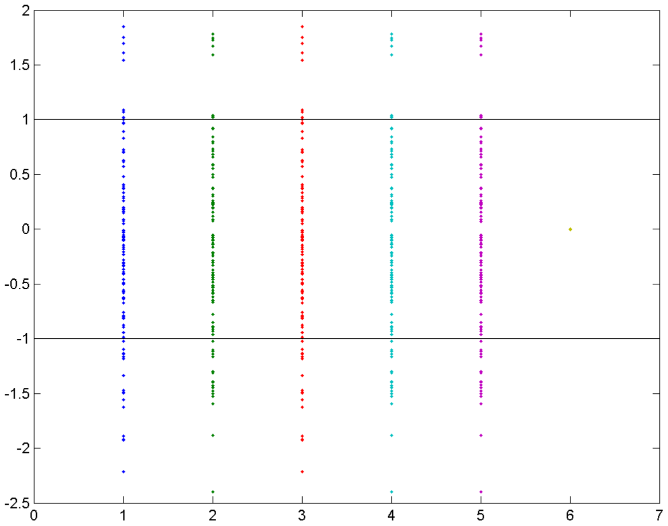

Figure 25: As Figure 23 but with $\mathbf{b} = (0, 2, 0.001)^{\mathrm{T}}$. The results for the GM algorithm are essentially identical to those for the GGM algorithm.

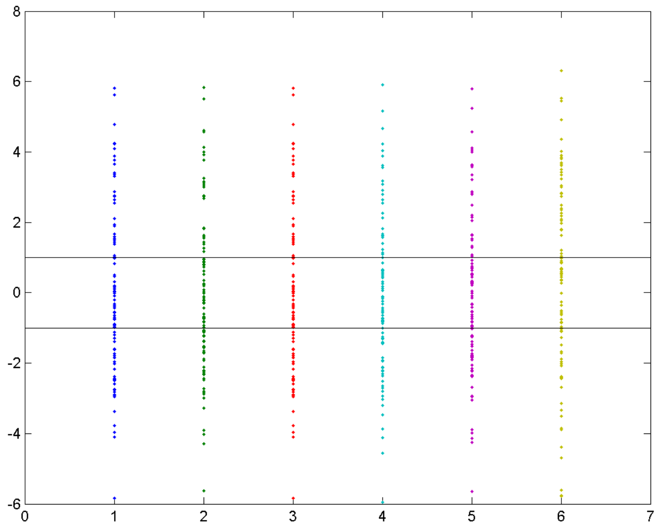

Figure 26: As Figure 23 but with $c = 4$.

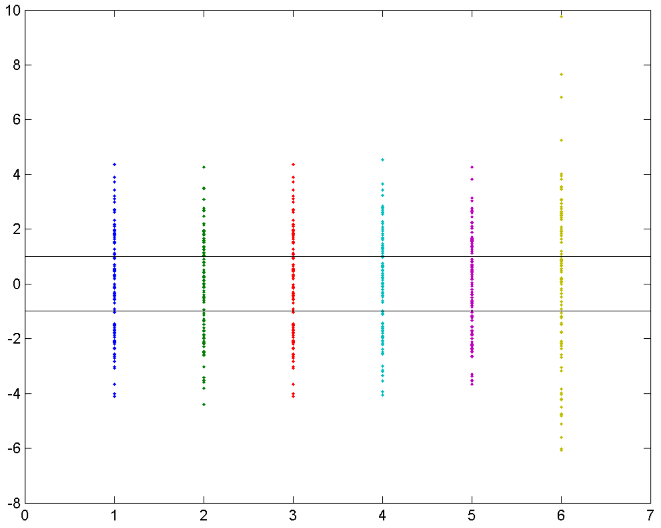

Figure 27: As Figure 23 but with $\mathbf{b} = (0, 2, -1)^{\mathrm{T}}$ and $c = 4$.

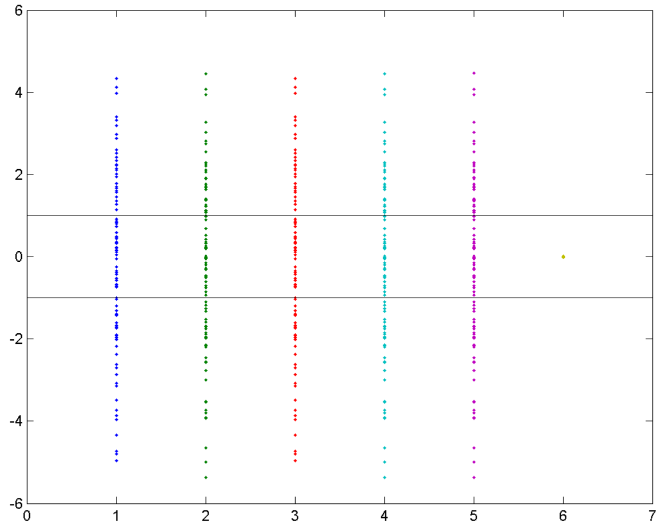

Figure 28: As Figure 23 but with $\mathbf{b} = (0, 2, 0.001)^{\mathrm{T}}$ and $c = 4$. The results for the GM algorithm are essentially identical to those for the GGM algorithm.

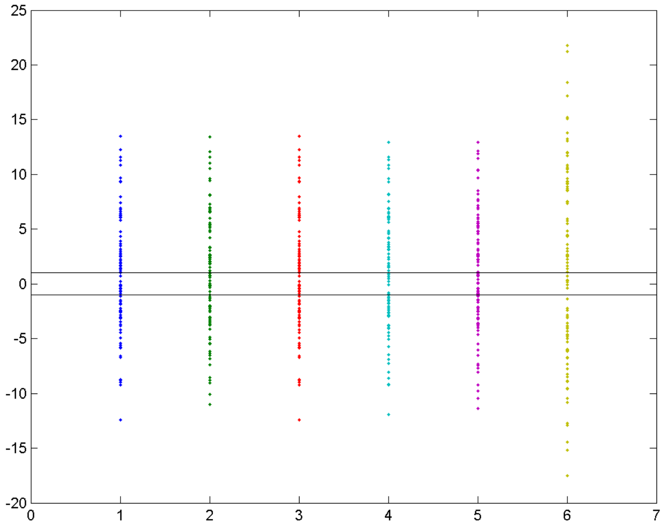

Figure 29: As Figure 23 but with $c = 6$.

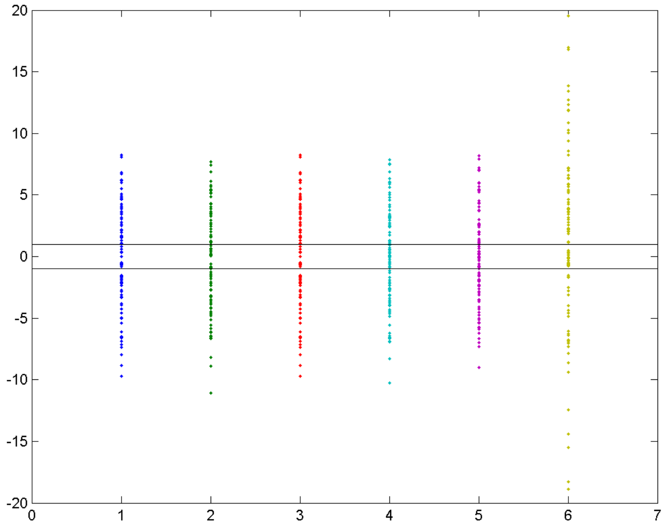

Figure 30: As Figure 23 but with $\mathbf{b} = (0, 2, -1)^{\mathrm{T}}$ and $c = 6$.

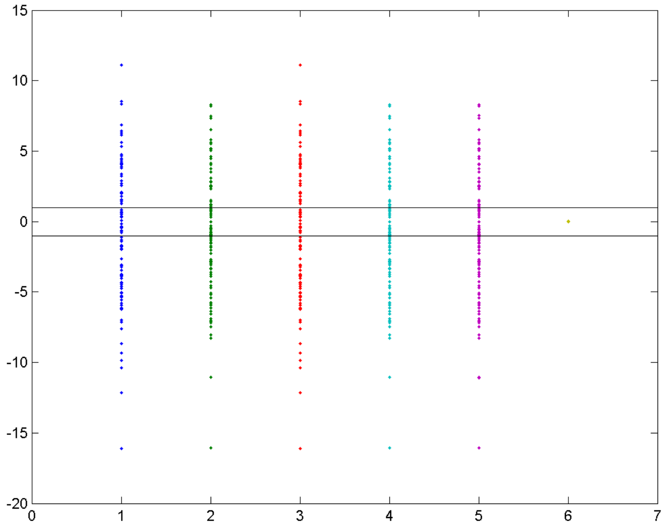

Figure 31: As Figure 23 but with $\mathbf{b} = (0, 2, 0.001)^{\mathrm{T}}$ and $c = 6$. The results for the GM algorithm are essentially identical to those for the GGM algorithm.

4.5.2 Interferometric model

Six simulation runs were made for null space data corresponding to covariance matrices generated as described in section 4.2.3 with $\mathbf{b}^\sharp = (0, 2, 1)^{\mathrm{T}}$, $(0, 2, -1)^{\mathrm{T}}$ and $(0, 2, 0.001)^{\mathrm{T}}$ (as above) and

$$(\nu_0, \sigma_0, \rho_0, \tau) = (0.001, 0.001, 0.001, 0.001)$$

and

$$(\nu_0, \sigma_0, \rho_0, \tau) = (0.002, 0.001, 0.005, 0.003).$$

Graphs 32–37 show:

1. The GDR estimator provides accurate estimates relative to the GGM estimator for all the data sets.
2. The WDR estimator provides accurate estimates relative to the GGM estimator for slope $b = 1$.
3. The GM estimator provides accurate estimates relative to the GGM estimator for slope $b = 0.001$.

The third observation has already been explained above. To explain the first, we note that the GDR estimator is solving the generalised Gauss-Markov problem for the data $\mathbf{x}$ and $\mathbf{z}$ but uncertainty matrix $V_{GDR} = V - \tau^2\mathbf{e}\mathbf{e}^{\mathrm{T}}$ where $\mathbf{e}$ is the $2m \times 1$ vector with 1 in each element. For this type of uncertainty matrix, we can rewrite (19) as

$$\begin{bmatrix} I & bI \\ \mathbf{0} & (A^*)^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \end{bmatrix} = \mathbf{0}, \quad \begin{bmatrix} \mathbf{x} - \mathbf{x}^* \\ \mathbf{z} - \mathbf{z}^* \end{bmatrix} = (V_{GDR} + \tau^2\mathbf{e}\mathbf{e}^{\mathrm{T}})\boldsymbol{\delta}. \qquad (20)$$

From the definition of A^* if $\mathbf{n}$ is a vector such $(A^*)^{\mathrm{T}}\mathbf{n} = \mathbf{0}$, then necessarily $\sum_i n_i = 0$. In particular the elements of $\boldsymbol{\delta}_2$ must sum to zero. Since $\boldsymbol{\delta}_1 = -b\boldsymbol{\delta}_2$, the elements of $\boldsymbol{\delta}_1$ must also sum to zero so that $\mathbf{e}^{\mathrm{T}}\boldsymbol{\delta} = 0$. Applying this to (20), we see that a solution of the GGM problem for uncertainty matrix

$$V = V_{GDR} + \tau^2\mathbf{e}\mathbf{e}^{\mathrm{T}}$$

must also be a solution for uncertainty matrix V_{GDR}.

To explain the second point, we note that if $b = 1$ then $\boldsymbol{\delta}_1 = -\boldsymbol{\delta}_2$ and

$$\begin{bmatrix} \nu_i^2 + \rho_i^2 & \rho_i^2 \\ \rho_i^2 & \sigma_i^2 + \rho_i \end{bmatrix} \begin{bmatrix} \delta_i \\ -\delta_i \end{bmatrix} = \begin{bmatrix} \delta_i\nu_i^2 \\ -\delta_i\sigma_i^2 \end{bmatrix} = \begin{bmatrix} \nu_i^2 & \\ & \sigma_i^2 \end{bmatrix} \begin{bmatrix} \delta_i \\ -\delta_i \end{bmatrix}.$$

If V_{WDR} is the diagonal matrix with ν_i^2 in the ith diagonal element and σ_i^2 in the $(m+i)$th, and $\boldsymbol{\delta}$ is in the nullspace of the Jacobian matrix J then $V\boldsymbol{\delta} = V_{GRD}\boldsymbol{\delta} = V_{WDR}\boldsymbol{\delta}$. This means that for the case $b = 1$ the WDR, GDR and GGM solutions coincide.

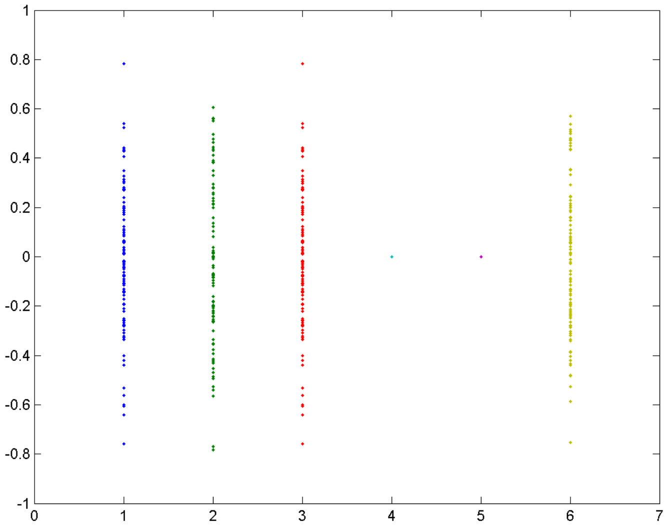

Figure 32: Normalised estimates $(d - d^{\sharp})/u_{GGM}$ calculated by 1) LLS, 2) WLS, 3) ODR, 4) WDR, 5) GDR and 6) GM algorithms for 100 simulated nullspace data sets with $\mathbf{b}^{\sharp} = (0, 2, 1)^{\mathrm{T}}$ and $(\nu_0, \sigma_0, \rho_0, \tau) = (0.001, 0.001, 0.001, 0.001)$. The units on the vertical axis have been normalised so that the standard uncertainty ($k = 1$) associated with the GGM estimate of the separation d is 1.

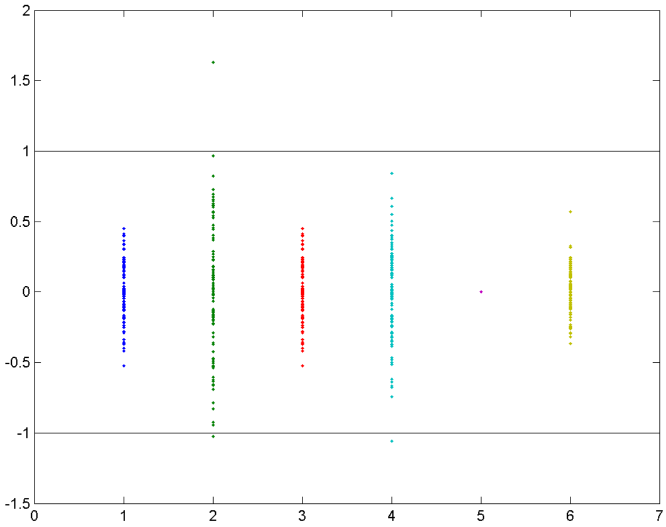

Figure 33: As Figure 32 but with $\mathbf{b}^{\sharp} = (0, 2, -1)^{\mathrm{T}}$.

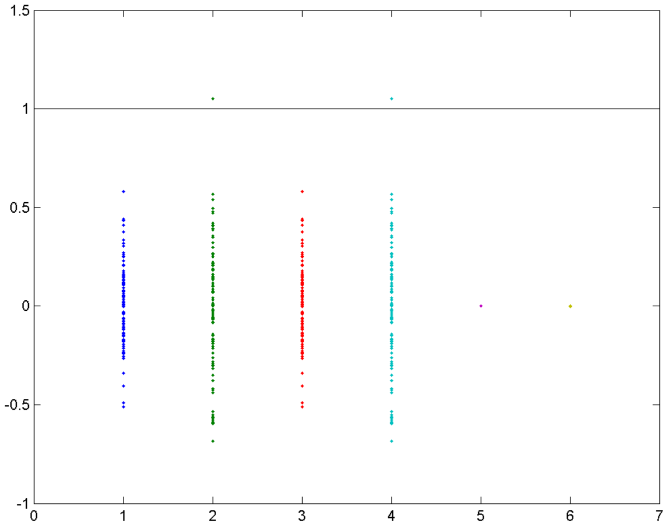

Figure 34: As Figure 32 but with $\mathbf{b}^\sharp = (0, 2, 0.001)^\mathrm{T}$.

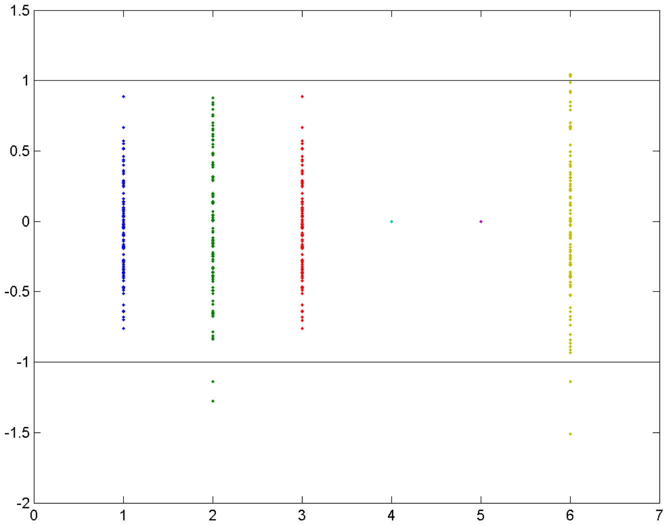

Figure 35: As Figure 32 but with $(\nu_0, \sigma_0, \rho_0, \tau) = (0.002, 0.001, 0.005, 0.003)$.

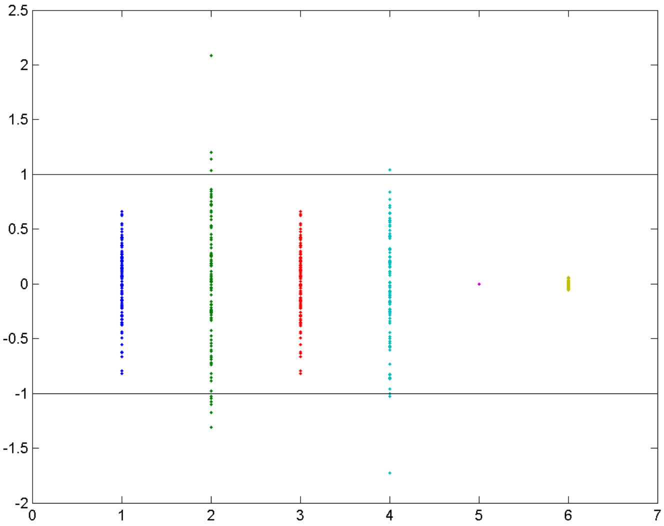

Figure 36: As Figure 32 but with $\mathbf{b}^\sharp = (0, 2, -1)^{\mathrm{T}}$ and $(\nu_0, \sigma_0, \rho_0, \tau) = (0.002, 0.001, 0.005, 0.003)$.

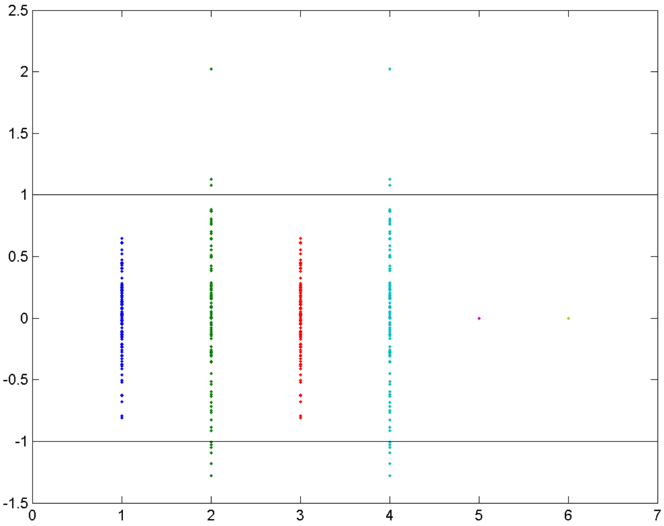

Figure 37: As Figure 32 but with $\mathbf{b}^\sharp = (0, 2, 0.001)^{\mathrm{T}}$ and $(\nu_0, \sigma_0, \rho_0, \tau) = (0.002, 0.001, 0.005, 0.003)$.

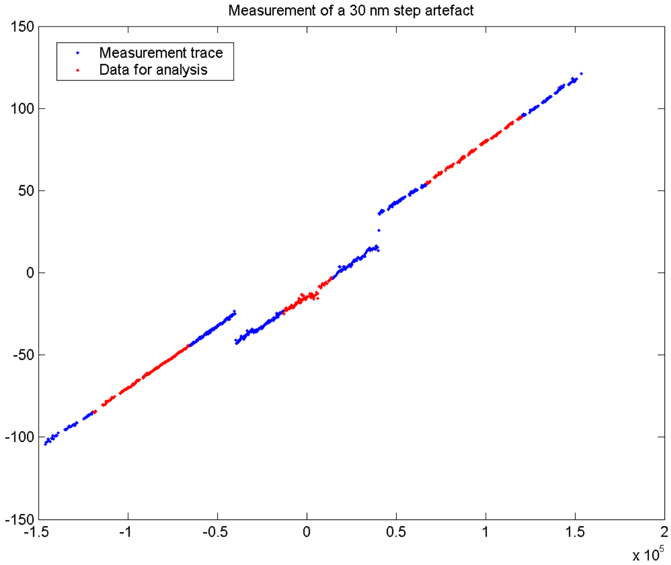

Figure 38: Data measured from an artefact with a nominal step of 30 nm. The vertical axis is height and the horizontal axis represents location along x axis. The units are in nanometres.

5 Analysis of measurement data

In this section we apply the algorithms to real measurement data. We look at three data sets arising from measurements of a 30 nm, 300 nm and 3000 nm step reference artefacts. The approximate slopes corresponding to these data sets are 7.5×10^{-4}, -2.1×10^{-5} and 8.0×10^{-4}.

5.1 Example 1: 30 nm step artefact

Figure 38 graphs a section of the measurement trace. The data points selected for the assessment of the artefact are shown in red.

We have applied the seven algorithms already considered assuming an interferometric error structure with a) $\nu_i = 1$, $\sigma_i = 1$, $\rho_i = 1$ and $\tau = 1$ and b) $\nu_i = 2$, $\sigma_i = 1$, $\rho_i = 5$ and $\tau = 3$. (The units are in nanometres.) Figure 39 graphs the residuals associated with a fit of the model to the data using the GGM algorithm. The estimates for the separation d produced by the seven algorithms for the two uncertainty matrices are given in Table 1. For case a), the standard uncertainty associated with the estimate of the depth, based on the input data uncertainty matrix, is approximately 0.13 nm for the first four algorithms and 0.18 nm for the last three. If we calculate the uncertainty

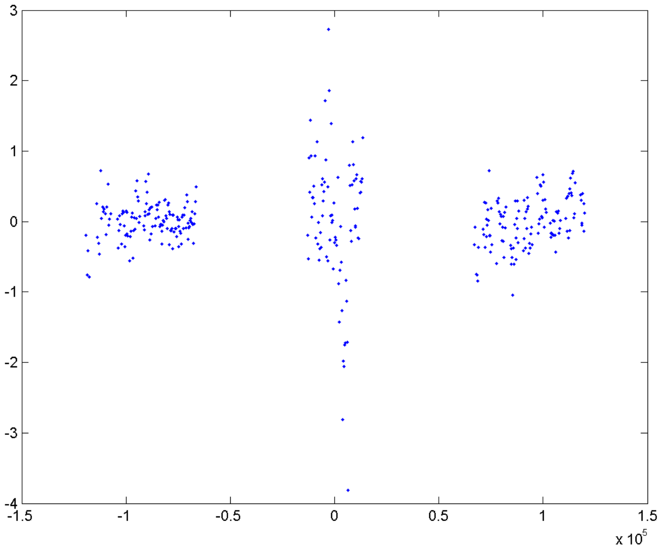

Figure 39: Residuals associated with the 30 nm step artefact data. The units for the vertical axis are nanometres.

based on the posterior estimate of σ as in (12) all seven algorithms give the same value of 0.10 nm. For case b) the corresponding numbers are 0.12, 0.64 and 0.10. Relative to the uncertainty in d, the estimates of d agree to 1 part in 10^7. In terms of the estimate of the depth the algorithms have essentially the same behaviour. It therefore follows that the estimates of the uncertainty associated with the depth should be essentially the same. While the uncertainty estimates based on the input uncertainty matrices differ significantly, those using the posterior estimate $\hat{\sigma}$ agree. The effect of an inappropriate combination of uncertainty matrix and solution algorithm is compensated for by the use of the posterior estimate.

5.2 Example 2: 300 nm step artefact

Figure 40 graphs a section of the measurement trace. The data points selected for the assessment of the artefact are shown in red.

We have applied a similar analysis to this data set. Figure 41 graphs the residuals associated with a GGM fit of the model to the data. The estimates for the separation d produced by the seven algorithms for the two uncertainty matrices are given in Table 2. For case a), the standard uncertainty associated with estimate of the depth, based on the input data uncertainty matrix, is approximately 0.26 nm for the first four algorithms

	a)	b)
LLS	19.53334409194924	19.53334409194924
WLS	19.53334409194924	19.53334409194924
ODR	19.53334409198467	19.53334409198467
WDR	19.53334409198467	19.53334409209101
GDR	19.53334406827659	19.53334404636478
GM	19.53334409194921	19.53334409194921
GGM	19.53334406827659	19.53334404636478

Table 1: Estimates of the separation parameter for a 30 nm step reference artefact associated with an interferometric error structure with a) $(\nu_i, \sigma_i, \rho_i, \tau) = (1, 1, 1, 1)$ and b) $(\nu_i, \sigma_i, \rho_i, \tau) = (2, 1, 5, 3)$.

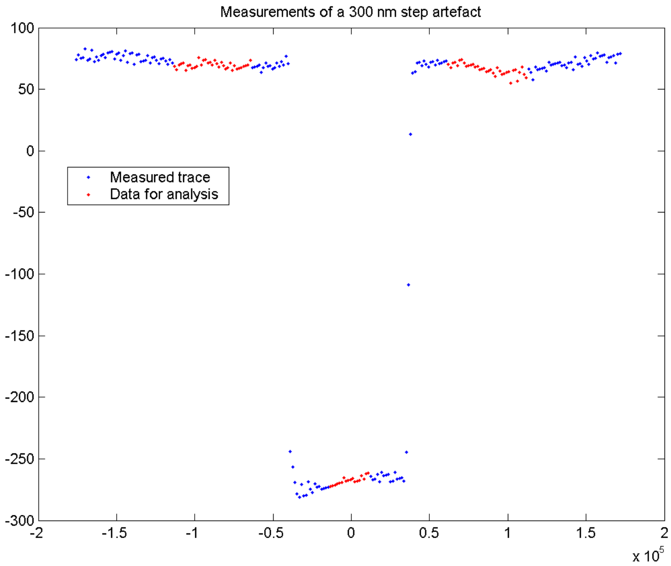

Figure 40: Data measured from an artefact with a nominal step of 300 nm. The vertical axis is height and the horizontal axis represents location along x axis. The units are in nanometres.

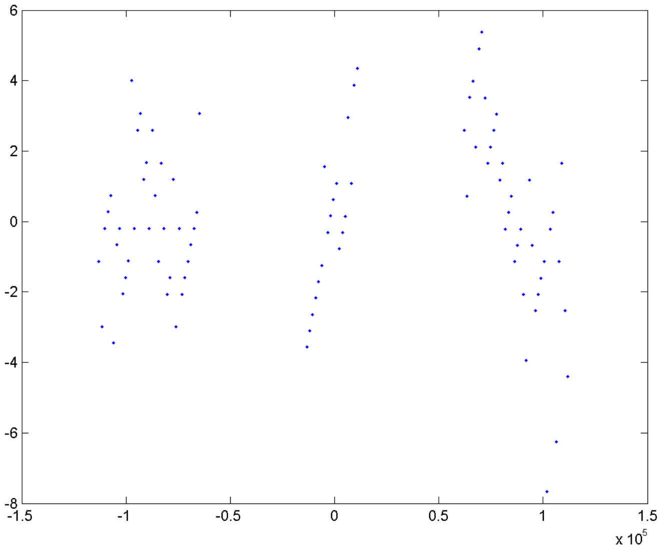

Figure 41: Residuals associated with the 300 nm step artefact data. The units for the vertical axis are nanometres.

and 0.37 nm for the last three. If we calculate the uncertainty based on the posterior estimate of σ all seven algorithms give the same estimate of 0.90 nm. For case b) the corresponding numbers are 0.26, 1.35 and 0.90. Relative to the uncertainty in d, the estimates of d agree to 1 part in 10^5.

5.3 Example 3: 3000 nm step artefact

Figure 42 graphs a section of the measurement trace. The data points selected for the assessment of the artefact are shown in red.

	a)	b)
LLS	335.458216718040	335.458216718040
WLS	335.458216718040	335.458216718040
ODR	335.458216718098	335.458216718098
WDR	335.458216718098	335.458216718272
GDR	335.458218082135	335.458219341201
GM	335.458216718040	335.458216718040
GGM	335.458218082135	335.458219341201

Table 2: Estimates of the separation parameter for a 300 nm step reference artefact associated with an interferometric error structure with a) $(\nu_i, \sigma_i, \rho_i, \tau) = (1, 1, 1, 1)$ and b) $(\nu_i, \sigma_i, \rho_i, \tau) = (2, 1, 5, 3)$.

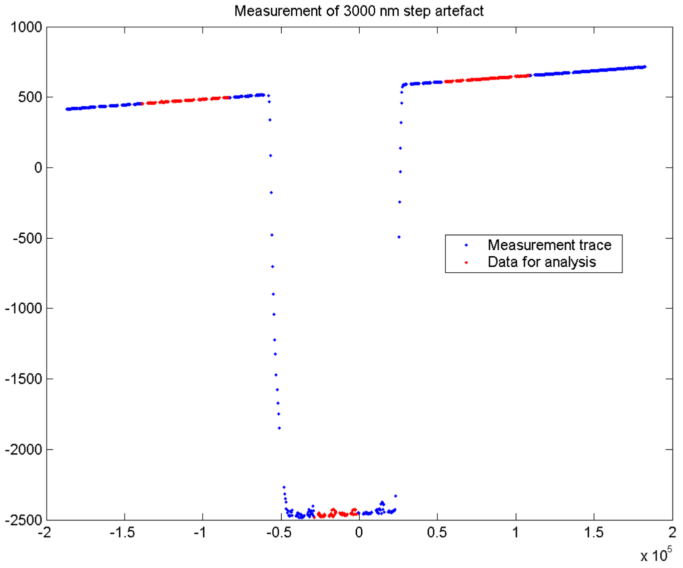

Figure 42: Data measured from an artefact with a nominal step of 3000 nm. The vertical axis is height and the horizontal axis represents location along x axis. The units are in nanometres.

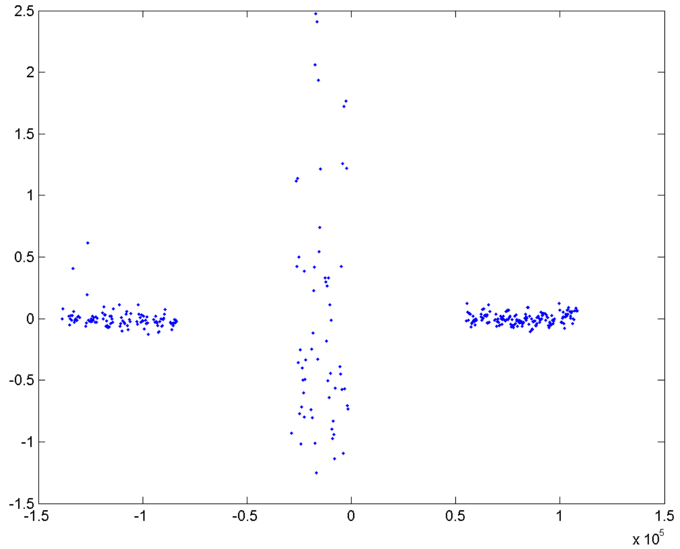

Figure 43: Residuals associated with the 3000 nm step artefact data. The units for the vertical axis are nanometres.

	a)	b)
LLS	3012.17406014788	3012.17406014788
WLS	3012.17406014788	3012.17406014788
ODR	3012.17406008297	3012.17406008297
WDR	3012.17406008297	3012.17406008297
GDR	3012.17410040646	3012.17410040646
GM	3012.17406014788	3012.17406014788
GGM	3012.17410040646	3012.17410040646

Table 3: Estimates of the separation parameter for a 3000 nm step reference artefact associated with an interferometric error structure with a) $(\nu_i, \sigma_i, \rho_i, \tau) = (1, 1, 1, 1)$ and b) $(\nu_i, \sigma_i, \rho_i, \tau) = (2, 1, 5, 3)$.

We have applied a similar analysis to this data set as the previous two. Figure 43 graphs the residuals associated with a GGM fit of the model to the data. The estimates for the separation d produced by the seven algorithms for the two uncertainty matrices are given in Table 3. For case a), the standard uncertainty associated with the depth, based on the input data uncertainty matrix, is approximately 0.14 nm for the first four algorithms and 0.20 nm for the last three. If we calculate the uncertainty based on the posterior estimate of σ all seven algorithms give the same estimate of 0.83 nm. For case b) the corresponding numbers are 0.14, 0.72 and 0.83. Relative to the uncertainty in d, the estimates of d agree to 1 part in 10^4.

5.4 Effect of form error on uncertainty evaluation

From the results for all three data sets we see that all seven algorithms have a similar behaviour in terms of the estimate of the separation. In fact there is no practical difference between them. In terms of the value of the standard uncertainty associated with the separation, the estimates based on the posterior estimate of the standard deviation of the residual errors are very close to each other for all seven algorithms, while those that depend on the input uncertainty matrix can differ significantly, as expected.

The posterior estimate of the standard deviation is based on the assumption that the residuals are drawn, approximately at least, from a normal distribution. In practice, there is a very significant contribution to the residuals due to the fact that the artefacts do not have perfect geometry. Imperfect geometry introduces two complications. The first is that use of the posterior estimate could well lead to a significant overestimate of the measurement uncertainty. Repeating a measurement run, executing the same measurement strategy, will lead to variation in the separation significantly less than that estimated from the standard deviation of the

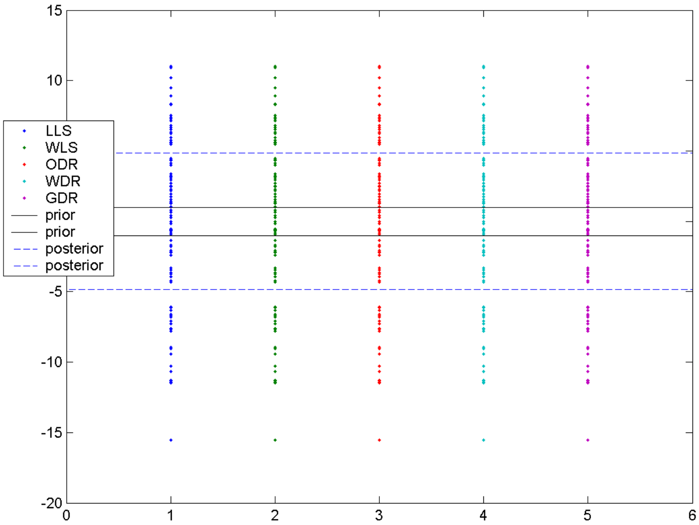

Figure 44: Estimated separations calculated by the 1) LLS, 2) WLS, 3) ODR, 4) WDR and 5) GDR algorithms for random subsets of the 3000 nm step artefact data, each with approximately half the number of data points. The solid, black horizontal lines are ±σ_p, where σ_p is the prior estimate of the standard deviation of the residuals, and the dashed, blue lines are ±σ_P, where σ_P is the posterior estimate of the standard deviation of the residuals. The units for the vertical axis are nanometres.

residuals. The second complication is that the estimate of the separation depends significantly on the measurement strategy.

To investigate this latter behaviour we have repeated the analysis of the 3000 nm step data by defining 100 subsets X_q of the orginal data set, each with approximately half the number of data points selected at random from the complete data set. We applied the LLS, WLS, ODR, WDR and GDR algorithms to these data sets and compared the estimated separations with values of the uncertainty based on the prior or posterior estimate of the standard deviation of the residuals. The results are shown in Figure 44. They show that the prior estimate significantly underestimates the variation in the separation, while the posterior estimate is in line with the observed variation.

5.5 Valid uncertainty analysis for the linear least squares estimator

The results from the analysis of the step data suggest that the estimate of the separation provided by the LLS algorithm is fit for purpose but that the usual value of the uncertainty based on prior information is likely to be invalid. The posterior estimate is also invalid since it is contaminated by form error effects. However, we know that for small slopes, the GM and GGM estimators have very similar behaviour and that the uncertainty associated with the $\mathbf{z}$ measurements can be safely ignored. We propose therefore a model of the form

$$z_i = \phi(x_i, \mathbf{b}) + \epsilon_i, \quad i = 1, \ldots, m, \quad \boldsymbol{\epsilon} \sim N(0, V_{22}).$$

If A is the matrix defined in (17), the LLS estimate of the parameters is given by

$$\mathbf{b}_{LLS} = A^{\dagger}\mathbf{z},$$

where $A^{\dagger}$ is the $3 \times m$ Moore-Penrose inverse of A:

$$A^{\dagger} = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}.$$

We note that if A has QR factorisation,

$$A = [Q_1 \; Q_2] \left[\begin{array}{c} R_1 \\ \mathbf{0} \end{array} \right],$$

then $A^{\dagger} = R_1^{-1}Q_1^{\mathrm{T}}$. The uncertainty matrix associated with $\mathbf{b}_{LLS}$ is therefore

$$V_{\mathbf{b}_{LLS}} = A^{\dagger}V_{22}\left(A^{\dagger}\right)^{\mathrm{T}},$$

and we can use this matrix to evaluate the uncertainty associated with the separation obtained from the LLS algorithm. We note that if $V_{22} = L_{22}L_{22}^{\mathrm{T}}$ and

$$Q^{\mathrm{T}}V_{22} = \left[\begin{array}{cc} T_{11} & T_{12} \\ & T_{22} \end{array} \right] U$$

completes the generalised QR factorisation of the pair $[A, \; L_{22}]$, then

$$V_{\mathbf{b}_{LLS}} = K_{LLS}K_{LLS}^{T}, \quad R_1K_{LLS} = [T_{11} \; T_{12}].$$

We can compare this with the uncertainty matrix $V_{\mathbf{b}_{GM}}$ associated the GM estimate:

$$V_{\mathbf{b}_{GM}} = K_{GM}K_{GM}^{T}, \quad R_1K_{GM} = T_{11}.$$

Comparing these two uncertainty matrices we see that any increased uncertainty in the LLS estimate over the GM estimate is due to T_{12}.

We have run a number of Monte Carlo simulations, using exactly the same data generation scheme as described in sections 4.1 and 4.2.1, to compare the estimate V of covariance matrix associated with the fitted parameters $\mathbf{b}_{LLS}$, with the sample covariance V_{MC} of the parameters $\mathbf{b}_q$ associated with the data sets X_q. Each data set corresponded to parameters $\mathbf{b} = (0, 2, 0.001)^{\mathrm{T}}$, and three runs, each of 5000 trials, were made with $c = 2$, 4 and 6. We also compare the estimated uncertainty u associated with the separation parameter with the standard deviation u_{MC} of the sampled deviations. These results are given in the Tables 4–6 and show good agreement.

We use the same approach to evaluating the uncertainty u_{LLS} associated with the LLS estimate of the separation for the experimental data sets. Since we have found the LLS estimate of the separation to be very close to the GGM estimate, we expect the uncertainties to be close. Table 7 below provides a comparison of these uncertainties for each of the three artefacts and assuming an interferometric error structure with a) $\nu_i = 1$, $\sigma_i = 1$, $\rho_i = 1$ and $\tau = 1$ and b) $\nu_i = 2$, $\sigma_i = 1$, $\rho_i = 5$ and $\tau = 3$. In all cases the differences are smaller that 1 part in 10^3.

$c = 2$	V	
3.8549	-0.2312	-0.0179
-0.2312	1.7121	0.0161
-0.0179	0.0161	0.0167
	V_{MC}	
3.9048	-0.1588	-0.0214
-0.1588	1.7024	0.0162
-0.0214	0.0162	0.0166
u	2.4555e-004	
u_{MC}	2.4341e-004	

Table 4: Estimates of the covariance matrix V and standard uncertainty u of the fitted parameters $\mathbf{b}$ and their counterparts V_{MC} and u_{MC} determined from 5000 Monte Carlo simulations. Data generated for the case $c = 2$.

$c = 4$	V	
3.5343	-0.3381	-0.0214
-0.3381	1.4253	0.0186
-0.0214	0.0186	0.0151
	V_{MC}	
3.6851	-0.2655	-0.0239
-0.2655	1.4161	0.0181
-0.0239	0.0181	0.0153
u	2.3740e-004	
u_{MC}	2.3732e-004	

Table 5: As Table 4 but with $c = 4$.

$c = 6$	V	
3.2855	-0.3823	-0.0213
-0.3823	1.2632	0.0178
-0.0213	0.0178	0.0132
	V_{MC}	
3.3718	-0.3303	-0.0224
-0.3303	1.2468	0.0171
-0.0224	0.0171	0.0135
u	2.3050e-004	
u_{MC}	2.2976e-004	

Table 6: As Table 4 but with $c = 6$.

a)	u_{GGM}	u_{LLS}
30 nm	0.1783	0.1784
300 nm	0.3732	0.3732
3000 nm	0.2015	0.2016
b)	u_{GGM}	u_{LLS}
30 nm	0.6427	0.6431
300 nm	1.3457	1.3456
3000 nm	0.7262	0.7268

Table 7: Estimates u_{GGM} and u_{LLS} of uncertainty in the separation parameter determined by the LLS and GGM algorithms on simulated data associated with an interferometric error structure with a) $(\nu_i, \sigma_i, \rho_i, \tau) = (1, 1, 1, 1)$ and b) $(\nu_i, \sigma_i, \rho_i, \tau) = (2, 1, 5, 3)$.

6 Concluding remarks

In this report, we have been concerned with algorithms for assessing the geometry of Type A1 reference artefacts for a range of uncertainty structures in the measurement data. In particular, we have included the most general case in which there is general correlation between all the measurement data. We have examined the behaviour of a range of algorithms on simulated data with the different algorithms incorporating different degrees of approximation in the uncertainty structure. The generalised Gauss-Markov algorithm (GGM) is seen to perform best. This is to be expected since it uses all the uncertainty information appropriately. The behaviour of the other algorithms depends on the degree of approximation and, for the nonlinear algorithms, the values of the parameters. In particular, for measurements of artefacts for which the slope of the fitted geometry is small, say, less than 0.005 in absolute value, the uncertainty in the x measurements can be safely ignored, leading to simpler estimation algorithms.

Applied to real measurement data, we have seen that the behaviour of all the algorithms was very similar in terms of the estimation of the separation parameters. However, the evaluation of the uncertainty matrix associated with the fitted parameters has to be based on the true data uncertainty matrix rather than the approximate uncertainty matrix implicitly associated with the algorithm. The use of the posterior estimate of the standard deviation of the residuals is likely to give invalid results if the form error is significant relative to the uncertainty associated with the height measurements.

References

[1] A. Björck. *Numerical Methods for Least Squares Problems.* SIAM, Philadelphia, 1996.

[2] British Standards Institution, London. *BS 1134 Part 1: Assessment of surface texture: methods and instrumentation.*, 1988.

[3] M. G. Cox, A. B. Forbes, J. Flowers, and P. M. Harris. Least squares adjustment in the presence of discrepant data. In *International Conference on Advanced Mathematical and Computational Tools in Metrology VI, Torino, September, 2003.* World Scientific, Singapore. *To appear.*

[4] M. G. Cox, A. B. Forbes, and P. M. Harris. Software Support for Metrology Best Practice Guide 4: Modelling Discrete Data. Technical report, National Physical Laboratory, Teddington, 2000.

[5] M. G. Cox, A. B. Forbes, P. M. Harris, and I. M. Smith. Classification and solution of regression problems for calibration. Technical Report CMSC 24/03, National Physical Laboratory, May 2003.

[6] A. B. Forbes, P. M. Harris, and I. M. Smith. Generalised Gauss-Markov Regression. In J. Levesley, I. Anderson, and J. C. Mason, editors, *Algorithms for Approximation IV*, pages 270–277. University of Huddersfield, 2002.

[7] G. H. Golub and C. F. Van Loan. *Matrix Computations.* John Hopkins University Press, Baltimore, third edition, 1996.

[8] International Organisation for Standards. *ISO 5436: Geometrical product specifications (GPS) - Surface texture: Profile method; Measurement standards*, 1999.

[9] R. K. Leach. Calibration, traceability and uncertainty issues in surface texture metrology. Technical Report CLM 7, National Physical Laboratory, Teddington, 2000.

[10] SIAM, Philadelphia. *The LAPACK User's Guide*, third edition, 1999.