



# Best integer equivariant estimation for elliptically contoured distributions

P. J. G. Teunissen<sup>1,2,3</sup>

Received: 18 March 2020 / Accepted: 27 July 2020 / Published online: 12 August 2020  
© The Author(s) 2020

## Abstract

This contribution extends the theory of integer equivariant estimation (Teunissen in *J Geodesy* 77:402–410, 2003) by developing the principle of best integer equivariant (BIE) estimation for the class of elliptically contoured distributions. The presented theory provides new minimum mean squared error solutions to the problem of GNSS carrier-phase ambiguity resolution for a wide range of distributions. The associated BIE estimators are universally optimal in the sense that they have an accuracy which is never poorer than that of any integer estimator and any linear unbiased estimator. Next to the BIE estimator for the multivariate normal distribution, special attention is given to the BIE estimators for the contaminated normal and the multivariate  $t$ -distribution, both of which have heavier tails than the normal. Their computational formulae are presented and discussed in relation to that of the normal distribution.

**Keywords** Global navigation satellite system (GNSS) · Integer equivariant (IE) estimation · Best integer equivariant (BIE) · Best linear unbiased estimation (BLUE) · Elliptically contoured distribution (ECD) · Multivariate normal · Contaminated normal · Multivariate  $t$ -distribution · LAMBDA method

## 1 Introduction

This contribution extends the theory of integer equivariant (IE) estimation as introduced in Teunissen (2003). The theory of IE estimation provides a solution to the problem of carrier-phase ambiguity resolution, which is key to high-precision GNSS positioning and navigation. It is well known that the so-called fixed GNSS baseline estimator is superior to its ‘float’ counterpart if the integer ambiguity success rate is sufficiently close to its maximum value of one. Although this is a strong result, the necessary condition on the success rate does not make it hold for all measurement scenarios. This restriction was the motivation to search for a class of estimators that could provide a universally optimal estimator while still benefiting from the integerness of the carrier-phase

ambiguities. The solution was found in the class of IE estimators, with its best integer equivariant (BIE) estimator being best in this class in the mean squared error sense (Teunissen 2003). The class of IE estimators obeys the integer remove–restore principle and was shown to be larger than the class of integer (I) estimators as well as larger than the class of linear unbiased (LU) estimators. As a consequence, the BIE estimator has the universal property that its mean squared error (MSE) is never larger than that of any I estimator or any LU estimator. This shows that from the MSE perspective one should always prefer the use of the BIE baseline over that of the integer least-squares (ILS) baseline and best linear unbiased (BLU) baseline.

In Teunissen (2003), an explicit expression of the BIE estimator was derived that holds true when the data can be assumed to be normally distributed. In Verhagen and Teunissen (2005), the performance of this estimator was compared with the float and fixed GNSS ambiguity estimators, while in Wen et al. (2012) it was shown how to use this BIE estimator for GNSS precise point positioning (PPP). In Brack et al. (2014), a sequential approach to best integer equivariant estimation was developed and tested, while Odolinski and Teunissen (2020) analyzed the normal distribution-based

✉ P. J. G. Teunissen  
p.j.g.teunissen@tudelft.nl

<sup>1</sup> Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, The Netherlands

<sup>2</sup> GNSS Research Centre, Curtin University of Technology, Perth, Australia

<sup>3</sup> Department of Infrastructure Engineering, The University of Melbourne, Melbourne, Australia

BIE estimation for low-cost single-frequency multi-GNSS RTK positioning.

In this contribution, we will generalize BIE estimation to the larger class of elliptically contoured distributions (ECDs). Many important distributions belong to this class (Chmielewski 1981; Cabane et al. 1981). Explicit expressions of the BIE estimator will be given for the multivariate normal distribution, the contaminated normal distribution and the multivariate  $t$ -distribution. The relevance of the contaminated distribution stems from the fact that it is a finite mixture distribution particularly useful for modeling data that are thought to contain a distinct subgroup of observations and thus can be used to model experimental error or contamination. The multivariate  $t$ -distribution is another class of distributions with tails heavier than the normal (Kibria and Joarder 2006; Roth 2013).

Several studies have already indicated the occurrence of GNSS instances where working with distributions that have tails heavier than the normal would be more appropriate. In Heng et al. (2011), for instance, it is shown that GPS satellite clock errors and instantaneous UREs have heavier tails than the normal distribution for about half of the satellites. Similar findings can be found in Dins et al. (2015). Also in fusion studies of GPS and INS, Student’s  $t$ -distribution has been proposed as the more suited distribution, see e.g., (Zhu et al. 2012; Zhong et al. 2018; Wang and Zhou 2019). And similar findings can be found in studies of multi-sensor GPS fusion for personal and vehicular navigation (Dhital et al. 2013; Al Hage et al. 2019).

This contribution is organized as follows. We start with a brief review of the theory of integer equivariant estimation in Sect. 2. Here, we emphasize the difference between integer equivariant estimation and integer estimation and show that the structure of the BIE estimator is one of a weighted average over the combined spaces of real and integer numbers. In Sect. 3, we present the general expression of the BIE estimator for elliptically contoured distributions. This will be done for the mixed-integer model, the integer-only model and the real-only model. We also emphasize here how the probability density function determines the structure of the estimator. Then, in Sects. 4, 5 and 6, we derive the explicit expressions of the BIE estimator for the multivariate normal, the contaminated normal and the multivariate  $t$ -distribution. In Sect. 7, they are compared and it is shown how they can be computed efficiently. The summary and conclusions are given in Sect. 8.

We make use of the following notation:  $A^T$  is the transpose of matrix  $A$ ,  $\mathcal{R}(A)$  denotes the range space of matrix  $A$  and  $\mathcal{N}(A)$  its null space.  $\|y\|_{\Sigma}^2 = (y)^T \Sigma^{-1}(y)$  denotes the weighted squared norm of vector  $y$  and  $|\Sigma_{yy}|$  the determinant of matrix  $\Sigma_{yy}$ . A matrix  $U$  is said to be a basis matrix of a subspace  $\mathcal{V}$ , if the column vectors of  $U$  form a basis of  $\mathcal{V}$ , i.e., the columns of  $U$  are linearly independent (full rank) and

$\mathcal{R}(U) = \mathcal{V}$ . The subspace  $\mathcal{V}^\perp$  is the orthogonal complement of  $\mathcal{V}$ . We also make a frequent use of the PDF transformation rule: If  $x = Ty + t$  and  $f_y(y)$  is the probability density function (PDF) of  $y$ , then  $f_x(x) = |T|^{-1}f_y(T^{-1}(y - t))$  is the PDF of  $x$ .  $P[E]$  denotes the probability of event  $E$ ,  $\sim$  ‘distributed as,’ and  $E(\cdot)$  and  $D(\cdot)$ , the expectation and dispersion operator, respectively.

## 2 Integer equivariant estimation

The class of integer equivariant (IE) estimators was introduced in Teunissen (2003). To appreciate the differences between IE estimators and integer estimators, we first recall the three conditions that an integer (I) estimator has to fulfill (Teunissen 1999a, b):

**Definition 1** (*I-estimation*) Let  $\hat{a} \in \mathbb{R}^n$  be a real-valued estimator of the integer vector  $a \in \mathbb{Z}^n$ . Then,  $\check{a} = \mathcal{I}(\hat{a})$ ,  $\mathcal{I} : \mathbb{R}^n \mapsto \mathbb{Z}^n$ , is an integer estimator of  $a$  if the pull-in region of  $\mathcal{I}$ ,  $\mathcal{P}_z = \{x \in \mathbb{R}^n \mid z = \mathcal{I}(x)\}$ , satisfies the three conditions:

1.  $\cup_{z \in \mathbb{Z}^n} \mathcal{P}_z = \mathbb{R}^n$
2.  $\mathcal{P}_u \cap \mathcal{P}_v = \emptyset \quad \forall u \neq v \in \mathbb{Z}^n$
3.  $\mathcal{P}_z = z + \mathcal{P}_0$ .

Each of the above three conditions states a property which is reasonable to ask of an arbitrary integer estimator. The first condition states that the pull-in regions should not leave any gaps, and the second that they should not overlap. The absence of gaps is needed in order to be able to map any ‘float’ solution  $\hat{a}$  to  $\mathbb{Z}^n$ , while the absence of overlaps is needed to guarantee that the ‘float’ solution is mapped to just one integer vector.

The third condition of the definition follows from the requirement that  $\mathcal{I}(x + z) = \mathcal{I}(x) + z, \forall x \in \mathbb{R}^n, z \in \mathbb{Z}^n$ . Also, this condition is a reasonable one to ask for. It states that when the ‘float’ solution  $\hat{a}$  is perturbed by an arbitrary integer vector  $z$ , the corresponding integer solution is perturbed by the same amount. This property thus allows one to apply the *integer remove–restore* technique,  $\mathcal{I}(\hat{a} - z) + z = \mathcal{I}(\hat{a})$ . It allows one to work with the fractional parts of  $\hat{a}$ , instead of with its complete entries.

It is this last condition which forms the guiding principle of IE estimators. Let the mean of the  $m$ -vector of observables be mixed-integer parametrized as

$$E(y) = Aa + Bb, \quad a \in \mathbb{Z}^n, b \in \mathbb{R}^p \tag{1}$$

with  $[A, B] \in \mathbb{R}^{m \times (n+p)}$  given and of full rank, and let our interest lie in estimating the linear function

$$\theta = L_a a + L_b b, \quad \theta \in \mathbb{R}^q \tag{2}$$

It then seems reasonable that the estimator should at least obey the integer remove–restore principle. For instance, when estimating integer ambiguities in case of GNSS, one would like, when adding an arbitrary number of cycles to the carrier-phase data, that the solution of the integer ambiguities be shifted by the same integer amount. For the estimator of  $\theta$ , this means that adding  $Az$  to  $y$ , for arbitrary  $z \in \mathbb{Z}^n$ , must result in a shift of  $L_a z$ . Similarly, it seems reasonable to require of the estimator that adding  $B\beta$  to  $y$ , for arbitrary  $\beta \in \mathbb{R}^p$ , results in a shift of  $L_b \beta$ . After all, we would not like the integer part of the estimator to be affected by such an addition to  $y$ . Requiring these two properties leads to the following definition of integer equivariant estimation (Teunissen 2003):

**Definition 2 (IE-estimation)** The estimator  $\hat{\theta}_{IE} = F_\theta(y)$ ,  $F_\theta : \mathbb{R}^m \mapsto \mathbb{R}^q$ , is an integer equivariant estimator of  $\theta = L_a a + L_b b \in \mathbb{R}^q$  if  $F_\theta(y + Az + B\beta) = F_\theta(y) + L_a z + L_b \beta$ ,  $\forall y \in \mathbb{R}^m, z \in \mathbb{Z}^n, \beta \in \mathbb{R}^p$ .

As the IE-class only requires the integer remove–restore property for estimating integer parameters, it encompasses the I-class, i.e.,  $I \subset IE$ . In Teunissen (2003), it was shown that the IE-class also encompasses the class of linear unbiased (LU) estimators,  $LU \subset IE$ . An important consequence of both I and LU being subsets of IE is that optimality in IE automatically carries over to I and LU. Probably, the most popular optimal estimator in the LU-class is the best linear unbiased estimator (BLUE), where ‘best’ is taken in the minimum mean squared error (MMSE) sense (Rao 1973; Koch 1999; Teunissen 2000). Using the same criterion, but now in the larger class of integer equivariant estimators, leads to the best integer equivariant (BIE) estimator.

**BIE Theorem (Teunissen 2003):** Let the vector of observables  $y \in \mathbb{R}^m$ , with PDF  $f_y(y)$ , have the mean  $E(y) = Aa + Bb$ , with  $a \in \mathbb{Z}^n$  and  $b \in \mathbb{R}^p$ . Then, the BIE estimator of the linear function  $\theta(a, b) = L_a a + L_b b$ ,  $L_a \in \mathbb{R}^{q \times n}$ ,  $L_b \in \mathbb{R}^{q \times p}$ , is given as

$$\hat{\theta}_{BIE} = \frac{\sum_{z \in \mathbb{Z}^n} \int_{\mathbb{R}^p} \theta(z, \beta) f_y(y + A(a - z) + B(b - \beta)) d\beta}{\sum_{z \in \mathbb{Z}^n} \int_{\mathbb{R}^p} f_y(y + A(a - z) + B(b - \beta)) d\beta}. \tag{3}$$

As a consequence of  $LU \subset IE$ , we have that the mean squared error (MSE) of the best linear unbiased (BLU) estimator is never better than that of a BIE estimator:

$$MSE(\hat{\theta}_{BLU}) \geq MSE(\hat{\theta}_{BIE}). \tag{4}$$

This implies that in the context of GNSS it would make sense to *always* compute the BIE baseline estimator, as its mean squared error will never be poorer than that of the ‘float’ baseline solution  $\hat{b}$ . Likewise, since  $I \subset IE$ , it follows that the BIE estimator is also MSE-superior to any integer estimator,

thus also to such popular estimators of integer least-squares (ILS), integer bootstrapping (IB) and integer rounding (IR).

The BIE estimator is a ‘weighted average.’ This can be seen if we write (3) in the compact form

$$\hat{\theta}_{BIE} = \sum_{z \in \mathbb{Z}^n} \int_{\mathbb{R}^p} w(z, \beta) \theta(z, \beta) d\beta \tag{5}$$

with weights  $w(z, \beta) = f(z, \beta) / (\sum_{z \in \mathbb{Z}^n} \int_{\mathbb{R}^p} f(z, \beta) d\beta)$ ,  $f(z, \beta) = f_y(y + A(a - z) + B(b - \beta))$ , that ‘sum up’ to unity,  $\sum_{z \in \mathbb{Z}^n} \int_{\mathbb{R}^p} w(z, \beta) d\beta = 1$ . In fact, by interpreting  $w(z, \beta)$  as a joint probability mass/density function, the BIE estimate can be interpreted as the mean of that distribution.

Would one be interested in only estimating the ambiguities, then with  $L_a = I_n$  and  $L_b = 0$ , one obtains from (5)

$$\hat{a}_{BIE} = \sum_{z \in \mathbb{Z}^n} w_z z, \text{ with } w_z = \int_{\mathbb{R}^p} w(z, \beta) d\beta. \tag{6}$$

Likewise, if one would be interested in estimating  $b$ , then with  $L_a = 0$  and  $L_b = I_p$ , one obtains from (5)

$$\hat{b}_{BIE} = \int_{\mathbb{R}^p} w(\beta) \beta d\beta, \text{ with } w(\beta) = \sum_{z \in \mathbb{Z}^n} w(z, \beta). \tag{7}$$

With both  $\hat{a}_{BIE}$  and  $\hat{b}_{BIE}$  available, one can compute the BIE estimator of  $\theta$  directly as  $\hat{\theta}_{BIE} = L_a \hat{a}_{BIE} + L_b \hat{b}_{BIE}$ . Hence, the BIE estimator of the mean  $E(y)$  is given as  $\hat{y}_{BIE} = A \hat{a}_{BIE} + B \hat{b}_{BIE}$ . This is then the expression to use, when in case of GNSS for instance, one would like to obtain the BIE solution for the pseudorange and carrier-phase data.

Note that the above theorem holds true for any PDF of the vector of observables  $y$  might have. A closer look at (3) reveals, however, that the unknowns  $a$  and  $b$  are needed in order to compute the estimator. This dependence on  $a$  and  $b$  is present in the numerator of (3), but not in its denominator as the dependence there disappears due to the integer summation and integration. Would the dependence persist, we would not be able to compute the BIE estimator. Note, however, that it disappears if the PDF of  $y$  has the structure  $f_y(y) = f(y - Aa - Bb)$ . Fortunately, this is still true for a large class of PDFs, and in particular for the class of *elliptically contoured distributions* (ECD), which we consider next.

### 3 BIE for elliptically contoured distributions

The class of elliptically contoured distributions is defined as follows (Cabane et al. 1981; Chmielewski 1981):

**Definition 3** (*Elliptically contoured distribution (ECD)*) A random vector  $y \in \mathbb{R}^m$  is said to have an ECD, denoted as  $y \sim \text{ECD}_m(\bar{y}, \Sigma, g)$ , if its PDF is of the form

$$f_y(y) = |\Sigma_{yy}|^{-1/2} g(\|y - \bar{y}\|_{\Sigma_{yy}}^2) \tag{8}$$

where  $\bar{y} \in \mathbb{R}^m$ ,  $\Sigma_{yy} \in \mathbb{R}^{m \times m}$  is positive definite, and  $g : \mathbb{R} \mapsto [0, \infty)$  is a decreasing function that satisfies  $\int_{\mathbb{R}^m} g(y^T y) dy = 1$ .

Note that the contours of constant density of an ECD are ellipsoids,  $(y - \bar{y})^T \Sigma_{yy}^{-1} (y - \bar{y}) = \text{constant}$ , from which the ECD draws its name. Also note, since (8) is symmetric with respect to  $\bar{y}$ , that  $\bar{y}$  in (8) is the mean of  $y$ ,  $E(y) = \bar{y}$ . The positive-definite matrix  $\Sigma_{yy}$  in (8), however, is in general not the variance matrix of  $y$ . It can be shown that the variance matrix of  $y$  is a scaled version of  $\Sigma_{yy}$ ,  $D(y) = \sigma^2 \Sigma_{yy}$ . The function  $g(\cdot)$  of (8) is called the generating function of the ECD (not to be confused with its moment generating function).

Several of the properties of ECDs are similar to those of the multivariate normal distribution. For instance, ECDs also have the linearity property: If  $x = Ty + t$ , with  $|T| \neq 0$ , and  $f_y(y) = |\Sigma_{yy}|^{-1/2} g(\|y - \bar{y}\|_{\Sigma_{yy}}^2)$  is the PDF of  $y$ , then  $f_x(x) = |\Sigma_{xx}|^{-1/2} g(\|x - \bar{x}\|_{\Sigma_{xx}}^2)$  is the PDF of  $x$ , with  $\bar{x} = T\bar{y} + t$  and  $\Sigma_{xx} = T \Sigma_{yy} T^T$ . Also, marginal and conditional PDFs of ECDs are again an ECD. Several important distributions belong to the ECD-class. Such examples are the multivariate normal distribution, the contaminated normal distribution and the multivariate  $t$ -distribution.

If  $y$  has mean (1), its likelihood function reads  $f_y(y|a, b) = |\Sigma_{yy}|^{-1/2} g(\|y - Aa - Bb\|_{\Sigma_{yy}}^2)$ , from which the maximum likelihood estimators of  $a \in \mathbb{Z}^n$  and  $b \in \mathbb{R}^p$  follow as

$$\begin{aligned} (\check{a}, \check{b}) &= \arg \max_{a \in \mathbb{Z}^n, b \in \mathbb{R}^p} f_y(y|a, b) \\ &= \arg \min_{a \in \mathbb{Z}^n, b \in \mathbb{R}^p} \|y - Aa - Bb\|_{\Sigma_{yy}}^2, \end{aligned} \tag{9}$$

thus showing that the maximum likelihood estimator coincides with the (mixed) integer least-squares estimator. The minimization (9) can be further worked out if we make use of the *orthogonal decomposition* (Teunissen 1995):

$$\|y - Aa - Bb\|_{\Sigma_{yy}}^2 = \|\hat{e}\|_{\Sigma_{yy}}^2 + \|\hat{a} - a\|_{\Sigma_{\hat{a}\hat{a}}}^2 + \|\hat{b}(a) - b\|_{\Sigma_{\hat{b}\hat{b}|a}}^2 \tag{10}$$

with  $\hat{e} = y - A\hat{a} - B\hat{b}$ , the least-squares residual vector,  $\hat{a} = \Sigma_{\hat{a}\hat{a}} \bar{A}^T \Sigma_{yy}^{-1} y$  and  $\hat{b} = \Sigma_{\hat{b}\hat{b}} \bar{B}^T \Sigma_{yy}^{-1} y$ , the ‘float’ least-squares solutions of  $a$  and  $b$ , respectively, and  $\hat{b}(a) = \hat{b} - \Sigma_{\hat{b}\hat{a}} \Sigma_{\hat{a}\hat{a}}^{-1} (\hat{a} - a)$ , the conditional least-squares solution of  $b$  given  $a$ . The matrices are given as  $\Sigma_{\hat{a}\hat{a}} = (\bar{A}^T \Sigma_{yy}^{-1} \bar{A})^{-1}$ ,  $\Sigma_{\hat{b}\hat{b}} = (\bar{B}^T \Sigma_{yy}^{-1} \bar{B})^{-1}$ ,  $\Sigma_{\hat{b}\hat{b}|a} =$

$\Sigma_{\hat{b}\hat{b}} - \Sigma_{\hat{b}\hat{a}} \Sigma_{\hat{a}\hat{a}}^{-1} \Sigma_{\hat{a}\hat{b}}$ , and  $\Sigma_{\hat{b}\hat{a}} = \Sigma_{\hat{b}\hat{b}} \bar{B}^T \Sigma_{yy}^{-1} \bar{A} \Sigma_{\hat{a}\hat{a}}$ , where  $\bar{A} = (I_m - P_B)A$ ,  $\bar{B} = (I_m - P_A)B$ , with orthogonal projectors  $P_A = A(A^T \Sigma_{yy}^{-1} A)^{-1} A^T \Sigma_{yy}^{-1}$  and  $P_B = B(B^T \Sigma_{yy}^{-1} B)^{-1} B^T \Sigma_{yy}^{-1}$ , respectively. Substitution of (10) into (9) shows that

$$\check{a} = \arg \min_{z \in \mathbb{Z}^n} \|\hat{a} - z\|_{\Sigma_{\hat{a}\hat{a}}}^2, \quad \check{b} = \hat{b} - \Sigma_{\hat{b}\hat{a}} \Sigma_{\hat{a}\hat{a}}^{-1} (\hat{a} - \check{a}). \tag{11}$$

In Teunissen (1999b), it is proven that for ECDs the ILS estimator  $\check{a}$  is optimal in the sense that it has, of all estimators in the I-class, the largest probability of correct integer estimation (i.e., the largest success rate).

The above property that the maximum likelihood estimate of  $a \in \mathbb{Z}^n$  and  $b \in \mathbb{R}^p$  remains the same, irrespective the choice for  $g(\cdot)$ , does in general not carry over to best integer equivariant estimation. Hence, for different functions  $g(\cdot)$ , we will have different BIE solutions. In Sects. 4, 5 and 6, we will give the explicit BIE solutions when the distributions are normal, contaminated normal and multivariate  $t$ . First, however, we will determine the general expression for the ECD–BIE estimator. We will do so by considering the three cases:  $A = 0$  (real-valued model),  $B = 0$  (integer-valued model) and  $A \neq 0, B \neq 0$  (mixed-integer real model).

**ECD–BIE Theorem** Let  $y \sim \text{ECD}_m(Aa + Bb, \Sigma_{yy}, g)$  have the PDF  $f_y(y) = |\Sigma_{yy}|^{-1/2} g(\|y - Aa - Bb\|_{\Sigma_{yy}}^2)$ ,  $a \in \mathbb{Z}^n$ ,  $b \in \mathbb{R}^p$ . Then, the BIE estimator of  $\theta = L_a a + L_b b$  is given as  $\hat{\theta}_{\text{BIE}} = L_a \hat{a}_{\text{BIE}} + L_b \hat{b}_{\text{BIE}}$ , with

$$\hat{a}_{\text{BIE}} = \sum_{z \in \mathbb{Z}^n} \frac{h(z)}{\sum_{u \in \mathbb{Z}^n} h(u)} z \tag{12}$$

where

$$\begin{aligned} \text{if } A \neq 0, B \neq 0 : & h(z) = \int_0^\infty g(c_z + r^2) r^{p-1} dr \\ & \hat{b}_{\text{BIE}} = \hat{b} - \Sigma_{\hat{b}\hat{a}} \Sigma_{\hat{a}\hat{a}}^{-1} (\hat{a} - \hat{a}_{\text{BIE}}) \\ \text{if } A \neq 0, B = 0 : & h(z) = g(c_z) \\ & L_b = 0 \\ \text{if } A = 0, B \neq 0 : & L_a = 0 \\ & \hat{b}_{\text{BIE}} = (B^T \Sigma_{yy}^{-1} B)^{-1} B^T \Sigma_{yy}^{-1} y \end{aligned} \tag{13}$$

and  $c_z = \|\hat{e}\|_{\Sigma_{yy}}^2 + \|\hat{a} - z\|_{\Sigma_{\hat{a}\hat{a}}}^2$ .

**Proof** For the proof, see the ‘‘Appendix’’. □

Note in case the model has no integer parameters ( $A = 0$ ), the BIE estimator of the real-valued parameter vector is identical to its BLUE. This is a consequence of the fact that the class of linear unbiased estimators is a subset of the IE-class, while both estimators have the MMSE property. Also note that in case of a mixed model ( $A \neq 0, B \neq 0$ ), the BIE estimator  $\hat{b}_{\text{BIE}}$  has the same structure as  $\check{b}$  of (11). By replacing  $\check{a}$  in the expression of  $\check{b}$  by  $\hat{a}_{\text{BIE}}$ , one obtains  $\hat{b}_{\text{BIE}}$ .

An important difference between the two type of estimators is, however, that in the (mixed) ILS case the mapping of  $y$  to  $\check{a}$  and  $\check{b}$  does not depend on  $g(\cdot)$ , whereas in the BIE case it does. This dependence on  $g(\cdot)$  reveals itself in the function  $h(z)$  (cf. 13). In general, it depends, through  $c_z$ , on both  $\|\hat{e}\|_{\Sigma_{yy}}^2$  and  $\|\hat{a} - z\|_{\Sigma_{\hat{a}\hat{a}}}^2$ . Hence,  $h(z)$  will then not only be smaller when  $\hat{a}$  is further away from  $z$ , but also when the norm of the least-squares residual vector gets larger.

The above theorem also shows that the complexity of the BIE estimator of  $a \in \mathbb{Z}^n$  differs depending on whether the model contains additional real-valued parameters or not. Both cases are relevant for GNSS. The model without real-valued parameters occurs, for instance, in the geometry-fixed case, when the data of a short GNSS baseline with known coordinates are analyzed for the purpose of stochastic model estimation.

As the estimator  $\hat{a}_{\text{BIE}}$  is simpler to compute when the model only contains integer parameters, one would perhaps be inclined, since real-valued parameters are easily eliminated by means of a linear transformation of the vector of observables, to always opt for using case ( $B = 0$ ) instead of case ( $B \neq 0$ ). Such would, however, be wrong. Although the class of elliptically contoured distributions is closed under linear transformations, the transformed ECD will generally not be a simple scaled version of the original. The following lemma makes this clear:

**Lemma 1** (Linear function of ECD) *Let  $y$  have the PDF  $f_y(y) = |\Sigma_{yy}|^{-1/2} g(\|y - Aa - Bb\|_{\Sigma_{yy}}^2)$  and  $y_c = C^T y$ , with basis matrix  $C \in \mathcal{R}(B)^\perp$  (i.e.,  $C^T B = 0$ ). Then,*

$$y_c \sim |\Sigma_{y_c y_c}|^{-1/2} g_c \left( \|y_c - C^T Aa\|_{\Sigma_{y_c y_c}}^2 \right) \tag{14}$$

with  $g_c(x) = |\Sigma_{y_b y_b}|^{-1/2} \int_{\mathbb{R}^p} g(x + \|y_b - \bar{y}_b\|_{\Sigma_{y_b y_b}}^2) dy_b$ ,  $\bar{y}_b = \Sigma_{y_b y_b} B^T \Sigma_{yy}^{-1} Aa + b$ ,  $\Sigma_{y_b y_b} = (B^T \Sigma_{yy}^{-1} B)^{-1}$ , and  $\Sigma_{y_c y_c} = C^T \Sigma_{yy} C$ .

**Proof** For proof, see the ‘‘Appendix’’. □

This result shows that although the PDF of  $y_c = C^T y$  is again an ECD, its generating function  $g_c(\cdot)$  differs from the original  $g(\cdot)$  as it is now an integrated version of it. It thus depends on  $g(\cdot)$  whether or not one would be allowed, in the computation of  $\hat{a}_{\text{BIE}}$ , to still work with  $g(\cdot)$  when eliminating the real-valued parameters from the model. As we will see in the next sections, this is the case with the normal distribution, but not in general.

### 4 BIE for multivariate normal distribution

The elliptically contoured PDF  $f_y(y) = |\Sigma_{yy}|^{-1/2} g(\|y - Aa - Bb\|_{\Sigma_{yy}}^2)$  becomes that of a multivariate normal distribu-

tion  $N_m(Aa + Bb, \Sigma_{yy})$ , when the generating function  $g(x)$  is chosen as

$$g(x) = (2\pi)^{-m/2} e^{-\frac{1}{2}x}. \tag{15}$$

Thus, in this case,  $\Sigma_{yy}$  is the variance matrix of  $y$ . By now using (15) with (13), one can obtain the weights for the BIE estimator in case  $y$  is normally distributed.

**Corollary 1** (BIE weights for normal distribution) *Let  $y \sim N_m(Aa + Bb, \Sigma_{yy})$ ,  $a \in \mathbb{Z}^n$ ,  $b \in \mathbb{R}^p$ . Then, the BIE weights of (12) follow using*

$$h(z) \propto \exp\{-\frac{1}{2}\|\hat{a} - z\|_{\Sigma_{\hat{a}\hat{a}}}^2\}. \tag{16}$$

**Proof** By substituting (15) into (13), one obtains for the case  $B \neq 0$ :  $h(z) = (2\pi)^{-m/2} \exp\{-\frac{1}{2}\|\hat{e}\|_{\Sigma_{yy}}^2\} \exp\{-\frac{1}{2}\|\hat{a} - z\|_{\Sigma_{\hat{a}\hat{a}}}^2\} \int_0^\infty \exp\{-\frac{1}{2}r^2\} r^{p-1} dr$ , of which only the  $z$ -dependent part remains when substituted into the weights of (12). The same result is obtained for the case  $B = 0$ . □

Note that where the weights of the general ECD–BIE expression (cf. 13) depend on  $c_z$  and thus on  $\|\hat{e}\|_{\Sigma_{yy}}^2$ , that this dependence on the least-squares residual vector  $\hat{e}$  is absent in case of a normal distribution.

### 5 BIE for contaminated normal distribution

The contaminated normal distribution is a *mixture* of two normal distributions having the same mean but proportional variance matrices. A mixture of two distributions is the distribution of a random vector  $y$  of which the sample is created from the realizations of two other random vectors  $x_1$  and  $x_2$  as follows: First, one of the two random vectors is selected by chance according to the two given probabilities of selection, say  $\epsilon$  for  $x_1$  and  $1 - \epsilon$  ( $0 \leq \epsilon \leq 1$ ) for  $x_2$ , and then the sample of the selected random vector is realized.

The PDF of the so-created random vector  $y$  can be expressed as a convex combination of the density functions  $f_{x_1}(y)$  and  $f_{x_2}(y)$  of the two random vectors:  $f_y(y) = (1 - \epsilon)f_{x_1}(y) + \epsilon f_{x_2}(y)$ . Thus, if  $f_{x_1}(y) = |\Sigma_{yy}|^{-1/2} g(\|y - \bar{y}\|_{\Sigma_{yy}}^2)$  and  $f_{x_2}(y) = |\delta \Sigma_{yy}|^{-1/2} g(\|y - \bar{y}\|_{\delta \Sigma_{yy}}^2)$  ( $\delta > 1$ ) are two ECDs with same mean but proportional variance matrices, then  $f_y(y)$  is a contaminated ECD.

The contaminated ECD is again an ECD, with generating function  $(1 - \epsilon)g(x) + \epsilon\delta^{-m/2}g(\frac{x}{\delta})$ . Thus, with  $g(x) = (2\pi)^{-m/2} \exp\{-\frac{1}{2}x\}$  being the generating function for the multivariate normal distribution, the generating function for the contaminated normal distribution is given as

$$g(x) = (2\pi)^{-\frac{m}{2}} \left[ (1 - \epsilon)e^{-\frac{1}{2}x} + \epsilon\delta^{-\frac{m}{2}} e^{-\frac{1}{2\delta}x} \right]. \tag{17}$$

Usually  $\delta$  is chosen large and  $\epsilon$  small. The idea is that the main distribution  $N_m(\bar{y}, \Sigma_{yy})$  is slightly ‘contaminated’ by the wider distribution  $N_m(\bar{y}, \delta \Sigma_{yy})$ . This results in a distribution with heavier tails than the main one.

By substituting (17) into (13), one obtains the weights for the contaminated normal BIE estimator.

**Corollary 2** (BIE weights for contaminated normal) *Let  $y \sim (1 - \epsilon)N_m(Aa + Bb, \Sigma_{yy}) + \epsilon N_m(Aa + Bb, \delta \Sigma_{yy})$ ,  $a \in \mathbb{Z}^n$ ,  $b \in \mathbb{R}^p$ ,  $0 \leq \epsilon \leq 1$ ,  $\delta > 1$ . Then, the BIE weights of (12) follow using*

$$h(z) \propto k(z) \times \exp\{-\frac{1}{2} \|\hat{a} - z\|_{\Sigma_{\hat{a}\hat{a}}}^2\} \tag{18}$$

with

$$k(z) = \left[ 1 + \frac{\delta^{-\frac{m-p}{2}} \epsilon / (1 - \epsilon)}{\exp\{-\frac{1}{2} c_z (\delta - 1) / \delta\}} \right] \tag{19}$$

in which  $c_z = \|\hat{e}\|_{\Sigma_{yy}}^2 + \|\hat{a} - z\|_{\Sigma_{\hat{a}\hat{a}}}^2$ .

**Proof** By substituting (17) into (13), recognizing that  $\int \exp\{-\frac{1}{2\delta} r^2\} r^{p-1} dr = \delta^{p/2} \int \exp\{-\frac{1}{2} r^2\} r^{p-1} dr$ , the result follows.  $\square$

Compare (18) with (16). Note hereby, as  $k(z)$  depends on  $c_z$ , that the weights of the contaminated normal BIE estimator depend on the least-squares residual, this in contrast to the normal BIE estimator. This dependence gets less, the closer  $\delta$  is chosen to one and in the limit we have  $\lim_{\delta \rightarrow 1} k(z) = 1 + \epsilon / (1 - \epsilon)$ , which as a constant cancels in the weights  $w_z$ . Also note the dependence of  $k(z)$  on the dimension of the real-valued vector  $b \in \mathbb{R}^p$ . When all else remains the same,  $k(z)$  gets larger for larger  $p$ . The case  $B = 0$  corresponds with  $p = 0$ .

### 6 BIE for multivariate $t$ -distribution

The random vector  $y \in \mathbb{R}^m$  is said to have a multivariate  $t$ -distribution with  $d > 2$  degrees of freedom, denoted as  $y \sim T_m(\bar{y}, \Sigma_{yy}, d)$ , if its PDF is given as (Zellner 1973; Kibria and Joarder 2006; Roth 2013)

$$f_y(y) = \frac{\Gamma(\frac{m+d}{2})}{(d\pi)^{\frac{d}{2}} \Gamma(\frac{d}{2}) |\Sigma_{yy}|^{1/2}} \left[ 1 + \frac{\|y - \bar{y}\|_{\Sigma_{yy}}^2}{d} \right]^{-\frac{m+d}{2}} \tag{20}$$

in which  $\Gamma(\cdot)$  denotes the gamma-function. The mean of  $y$  is  $E(y) = \bar{y}$ , and its variance matrix is  $D(y) = \frac{d}{d-2} \Sigma_{yy}$ .

It is noted that  $T_1(0, 1, d)$  is Student’s  $t$ -distribution with  $d$  degrees of freedom (Gosset 1908; Koch 1999). The  $t$ -distribution has heavier tails than the normal distribution, but in the limit  $d \rightarrow \infty$  converges to the standard normal distribution  $N_1(0, 1)$ . The analogy in the multivariate case is that  $T_m(0, \Sigma_{yy}, d)$  converges to  $N_m(0, \Sigma_{yy})$  as  $d \rightarrow \infty$ .

Also, the multivariate  $t$ -distribution is an elliptically contoured distribution. Its generating function is given as

$$g(x) = \frac{\Gamma(\frac{m+d}{2})}{(d\pi)^{\frac{d}{2}} \Gamma(\frac{d}{2})} \left[ 1 + \frac{x}{d} \right]^{-\frac{m+d}{2}}. \tag{21}$$

By substituting (21) into (13), one obtains the weights for the multivariate  $t$  BIE estimator.

**Corollary 3** (BIE weights for multivariate  $t$ ) *Let  $y \sim T_m(Aa + Bb, \Sigma_{yy}, d)$ ,  $a \in \mathbb{Z}^n$ ,  $b \in \mathbb{R}^p$ . Then, the BIE weights of (12) follow using*

$$h(z) \propto \left[ 1 + \frac{c_z}{d} \right]^{-\frac{m+d}{2} + p} \tag{22}$$

in which  $c_z = \|\hat{e}\|_{\Sigma_{yy}}^2 + \|\hat{a} - z\|_{\Sigma_{\hat{a}\hat{a}}}^2$ .

**Proof** We need to solve  $h(z) = \int_0^\infty g(c_z + r^2) r^{p-1} dr$  for  $g(x) \propto [1 + \frac{x}{d}]^{-\frac{m+d}{2}}$ . Since  $\int_0^\infty [1 + (ax)^2]^{-s} x^{n-1} dx = \frac{\Gamma(\frac{n}{2}) \Gamma(s - \frac{n}{2})}{2a^{2n} \Gamma(s)}$  (Gradshteyn and Ryzhik 2007), it follows from  $h(z) \propto [1 + \frac{c_z}{d}]^{-\frac{m+d}{2}} \int_0^\infty [1 + (\frac{r}{c_z+d})^2]^{-\frac{m+d}{2}} r^{p-1} dr$  that  $h(z) \propto [1 + \frac{c_z}{d}]^{-\frac{m+d}{2}} [c_z + d]^p$ , from which the result (22) follows.  $\square$

Note, as with the contaminated normal distribution, that the BIE weights for the multivariate  $t$ -distribution depend on the least-squares residual vector  $\hat{e}$ . Also note, when all else remains the same, that  $h(z)$  gets larger for larger  $p$ , i.e., if the underlying model gets weaker. The case  $B = 0$  corresponds with  $p = 0$ .

As the  $t$ -distribution converges to the normal distribution for  $d \rightarrow \infty$ , one would expect  $h(z)$  of (22) to converge to (16) for increasing degrees of freedom. And indeed, since  $\lim_{n \rightarrow \infty} [1 + \frac{x}{n}]^n = e^x$ , it follows that

$$\lim_{d \rightarrow \infty} \left[ 1 + \frac{c_z}{d} \right]^{-\frac{m+d}{2} + p} = \exp\{-\frac{1}{2} c_z\} \propto \exp\{-\frac{1}{2} \|\hat{a} - z\|_{\Sigma_{\hat{a}\hat{a}}}^2\}$$

which is the  $h(z)$  of (16).

### 7 On the computation of the BIE estimators

As (12) cannot be computed in practice due to the occurrence of infinite sums, we need to replace the sum over all integers

by a sum over a finite set. It seems reasonable to neglect the contributions to the sum, when  $h(z)$  is sufficiently small. As  $h(z)$  gets smaller, the larger the (weighted) distance between  $z$  and  $\hat{a}$  is (cf. 12), we define the integer set as

$$\Omega_{\hat{a}}^\lambda = \{z \in \mathbb{Z}^n \mid \|\hat{a} - z\|_{\Sigma_{\hat{a}\hat{a}}}^2 \leq \lambda^2\} \tag{23}$$

and approximate  $\hat{a}_{\text{BIE}}$  by

$$\hat{a}_{\text{BIE}}^\lambda = \sum_{z \in \Omega_{\hat{a}}^\lambda} z \frac{h(z)}{\sum_{u \in \Omega_{\hat{a}}^\lambda} h(u)}. \tag{24}$$

This will be a good approximation of  $\hat{a}_{\text{BIE}}$  if the size of the integer set  $\Omega_{\hat{a}}^\lambda$ , and thus  $\lambda$  can be properly chosen. To find such way of choosing  $\lambda$ , we first note that the approximation  $\hat{a}_{\text{BIE}}^\lambda$  can also be seen as an *exact* BIE estimator in its own right, namely when the PDF is given by the following *truncated* version of  $f_y(y)$ ,

$$f_y^\lambda(y) = \frac{\delta_{\hat{a}}^\lambda(\hat{a}) f_y(y)}{\int_{E_{\hat{a}}^\lambda} f_y(y) dy} \tag{25}$$

with  $\delta_{\hat{a}}^\lambda(x)$  being the indicator function of the ellipsoidal region  $E_{\hat{a}}^\lambda = \{x \in \mathbb{R}^n \mid \|x - \hat{a}\|_{\Sigma_{\hat{a}\hat{a}}}^2 \leq \lambda^2\}$ . Thus, by using (25) as PDF instead of (8), one will obtain instead of (12), with its infinite sums, the BIE estimator (24), having finite sums, since  $\sum_{z \in \mathbb{Z}^n} \delta_z^\lambda(\hat{a}) h(z) = \sum_{z \in \Omega_{\hat{a}}^\lambda} h(z)$ .

As (25) produces (24) as BIE estimator, one will have a good approximation of (12) when  $f_y^\lambda(y)$  is a good approximation to the original PDF  $f_y(y)$ , which will be the case when the denominator of (25) is close enough to one, and thus

$$\int_{E_{\hat{a}}^\lambda} f_y(y) dy = 1 - \alpha \tag{26}$$

for  $\alpha$  small enough. By applying the appropriate change-of-variable rule, one will recognize the integral of (26) as the probability of  $\hat{a}$  residing in  $E_{\hat{a}}^\lambda$ . Hence, the proper size of the integer set  $\Omega_{\hat{a}}^\lambda$  in (24) is found by choosing  $\lambda$  to satisfy

$$P[\hat{a} \in E_{\hat{a}}^\lambda] = P[\|\hat{a} - a\|_{\Sigma_{\hat{a}\hat{a}}}^2 \leq \lambda^2] = 1 - \alpha. \tag{27}$$

The following lemma shows how this probability can be computed for the three different ECDs that we have considered.

**Lemma 2** (On the size of the integer set  $\Omega_{\hat{a}}^\lambda$ ): *Let  $y$ , with mean  $\bar{y} = Aa + Bb$ ,  $a \in \mathbb{Z}^n, b \in \mathbb{R}^p$ , be distributed as (i)  $y \sim N_m(\bar{y}, \Sigma_{yy})$ , (ii)  $y \sim (1 - \epsilon)N_m(\bar{y}, \Sigma_{yy}) + \epsilon N_m(\bar{y}, \delta \Sigma_{yy})$ , and (iii)  $y \sim T_m(\bar{y}, \Sigma_{yy}, d)$ , respectively. Then, the probability  $P[\|\hat{a} - a\|_{\Sigma_{yy}}^2 \leq \lambda^2] = 1 - \alpha$  is given*

by

- (i)  $P[\chi^2(n) \leq \lambda^2] = 1 - \alpha$
- (ii)  $(1 - \epsilon)P[\chi^2(n) \leq \lambda^2] + \epsilon P[\chi^2(n) \leq \frac{1}{\delta} \lambda^2] = 1 - \alpha$
- (iii)  $P[F(n, d) \leq \lambda^2] = 1 - \alpha$

in which  $\chi^2(n)$  and  $F(n, d)$  are the central chi-squared and central  $F$ -distribution with  $n$  and  $n, d$  degrees of freedom, respectively.

**Proof** As  $\hat{a} = \Sigma_{\hat{a}\hat{a}} \bar{A}^T \Sigma_{yy}^{-1} y$  is linear in  $y$ , it follows that (i)  $\hat{a} \sim N_n(a, \Sigma_{\hat{a}\hat{a}})$ , (ii)  $\hat{a} \sim (1 - \epsilon)N_n(a, \Sigma_{\hat{a}\hat{a}}) + \epsilon N_n(a, \delta \Sigma_{\hat{a}\hat{a}})$ , and (iii)  $\hat{a} \sim T_n(a, \Sigma_{\hat{a}\hat{a}}, d)$ , respectively. Therefore, (i)  $\|\hat{a} - a\|_{\Sigma_{yy}}^2 \sim \chi^2(n)$  and (iii)  $\|\hat{a} - a\|_{\Sigma_{yy}}^2 \sim F(n, d)$  (Kibria and Joarder 2006), from which the result follows.  $\square$

Once the choice for the size of the integer set  $\Omega_{\hat{a}}^\lambda$  has been made, the integer vectors contained in it have to be collected, so as to be able to compute (24). This collection can be done efficiently with the LAMBDA method (Teunissen 1995; De Jonge et al. 1996), as also demonstrated in Verhagen and Teunissen (2005) and (Odolinski and Teunissen 2020).

## 8 Summary and conclusions

In this contribution, the theory of integer equivariant (IE) estimation (Teunissen 2003) was extended to include the family of elliptically contoured distributions. As the class of IE estimators includes both the class of integer (I) estimators and class of linear unbiased (LU) estimators, any optimality in the IE-class automatically carries over to the I-class and the LU-class. Hence, if  $\check{b}$  is an arbitrary I estimator and  $\hat{b}$  an arbitrary LU estimator, then the best integer equivariant (BIE) estimator  $\hat{b}_{\text{BIE}}$ , which is optimal in the minimum mean squared error sense, will have a mean squared error (MSE) that satisfies

$$\text{MSE}(\hat{b}_{\text{BIE}}) \leq \text{MSE}(\hat{b}) \text{ and } \text{MSE}(\hat{b}_{\text{BIE}}) \leq \text{MSE}(\check{b}). \tag{28}$$

In the context of GNSS, this implies that the MSE of the baseline BIE estimator is always less than, or at the most equal to, that of both its ‘float’ and ‘fixed’ counterparts. This shows that from the MSE perspective one should always prefer the use of the BIE baseline over that of the best linear unbiased (BLU) baseline and integer least-squares (ILS) baseline.

In contrast to the BLU estimator and the ILS estimator, the expression for the BIE estimator depends on the probability density function (PDF) of the vector of observables. Different PDFs give different mappings from the vector of observables  $y$  to the baseline  $b$ . In this contribution, we developed the expressions of the BIE estimator for the family of elliptically contoured distributions. For the mixed-integer model  $E(y) =$

$Aa + Bb$ ,  $a \in \mathbb{Z}^n$ ,  $b \in \mathbb{R}^p$ , the BIE estimators were shown to be given as

$$\hat{b}_{\text{BIE}} = \hat{b} - \Sigma_{\hat{b}\hat{a}} \Sigma_{\hat{a}\hat{a}}^{-1} (\hat{a} - \hat{a}_{\text{BIE}}) \tag{29}$$

with

$$\hat{a}_{\text{BIE}} = \sum_{z \in \mathbb{Z}^n} z \frac{\int_0^\infty g(c_z + r^2) r^{p-1} dr}{\sum_{u \in \mathbb{Z}^n} \int_0^\infty g(c_u + r^2) r^{p-1} dr} \tag{30}$$

in which the choice of the multivariate PDF is felt through the generating function  $g(\cdot)$ .

Important examples of elliptically contoured distributions are the multivariate normal distribution, the contaminated normal distribution and the multivariate  $t$ -distribution. By means of their generating functions, which were shown to be given as

$$g(x) \propto \begin{cases} e^{-\frac{1}{2}x} & \text{(normal)} \\ (1 - \epsilon)e^{-\frac{1}{2}x} + \epsilon\delta^{-\frac{m}{2}} e^{-\frac{1}{2\delta}x} & \text{(contaminated)} \\ \left[1 + \frac{x}{d}\right]^{-\frac{m+d}{2}} & \text{(\textit{t}-distribution)} \end{cases} \tag{31}$$

we provided the explicit formulae of their BIE estimators. For each of these distributions, it is now thus possible to compute the GNSS baseline estimator such that it will have the smallest possible mean squared error of all integer equivariant baseline estimators.

**Author contributions** PT developed the theory and wrote the paper.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix

**Proof of ECD–BIE Theorem** First, we consider the case  $A \neq 0$ ,  $B \neq 0$  (the case  $A \neq 0$ ,  $B = 0$  goes along similar lines) and then the case  $A = 0$ . We start from the general expression

for the BIE estimator of  $b$ , which follows from (3) as

$$\hat{b}_{\text{BIE}} = \frac{\sum_{z \in \mathbb{Z}^n} \int_{\mathbb{R}^p} \beta f_y(y + A(a - z) + B(b - \beta)) d\beta}{\sum_{z \in \mathbb{Z}^n} \int_{\mathbb{R}^p} f_y(y + A(a - z) + B(b - \beta)) d\beta} \tag{32}$$

Substitution of the ECD–PDF (8), with  $\bar{y} = Aa + Bb$ , while making use of the orthogonal decomposition (10), gives

$$\hat{b}_{\text{BIE}} = \frac{\sum_{z \in \mathbb{Z}^n} \int_{\mathbb{R}^p} \beta g(c_z + \|\hat{b}(z) - \beta\|_{\Sigma_{\hat{b}\hat{b}|a}}^2) d\beta}{\sum_{z \in \mathbb{Z}^n} \int_{\mathbb{R}^p} g(c_z + \|\hat{b}(z) - \beta\|_{\Sigma_{\hat{b}\hat{b}|a}}^2) d\beta} \tag{33}$$

with  $c_z = \|\hat{e}\|_{\Sigma_{yy}}^2 + \|\hat{a} - z\|_{\Sigma_{\hat{a}\hat{a}}}^2$ . By noting that the function  $g(c_z + \|\beta - \hat{b}(z)\|_{\Sigma_{\hat{b}\hat{b}|a}}^2)$ , as function of  $\beta$ , is symmetric with respect to  $\hat{b}(z)$ , we can make use of the property

$$\int_{\mathbb{R}^p} [\beta - \hat{b}(z)] g(c_z + \|\beta - \hat{b}(z)\|_{\Sigma_{\hat{b}\hat{b}|a}}^2) d\beta = 0 \tag{34}$$

and rewrite the numerator of (33) to obtain

$$\hat{b}_{\text{BIE}} = \sum_{z \in \mathbb{Z}^n} \frac{h(z)}{\sum_{z \in \mathbb{Z}^n} h(z)} \hat{b}(z) \tag{35}$$

where

$$h(z) \propto \int_{\mathbb{R}^p} g(c_z + \|\hat{b}(z) - \beta\|_{\Sigma_{\hat{b}\hat{b}|a}}^2) d\beta. \tag{36}$$

This leaves us to further simplify the integral. A first simplification is reached if we reparametrize  $\beta$  in  $v$  as  $\beta = G^T v + \hat{b}(z)$ , with  $G$  satisfying  $\Sigma_{\hat{b}\hat{b}|a} = G^T G$ , e.g., being its Cholesky factor. Then,  $\|\beta - \hat{b}(z)\|_{\Sigma_{\hat{b}\hat{b}|a}}^2 = v^T v$ , from which, with an application of the change-of-variable rule for integrals and recognizing that  $|\Sigma_{\hat{b}\hat{b}|a}|^{1/2} = |G|$ , follows

$$h(z) \propto |\Sigma_{\hat{b}\hat{b}|a}|^{1/2} \int_{\mathbb{R}^p} g(c_z + v^T v) dv. \tag{37}$$

Note that  $v$  only appears in the integral through its squared Euclidean norm. This suggests that we apply a *polar coordinate* transformation as a further change of variables in the integral. We therefore reparametrize  $v$  in the scalar radius  $r > 0$  and the vector of angles  $\alpha = (\alpha_1, \dots, \alpha_{p-1})^T$  as  $v = ru(\alpha)$ , where the components of the unit-vector  $u(\alpha)$  are given as  $u_i(\alpha) = \cos(\alpha_i) \prod_{j=0}^{i-1} \sin(\alpha_j)$ , with  $\sin(\alpha_0) = \cos(\alpha_p) = 1$ , and where  $0 \leq \alpha_j \leq \pi$  for  $j = 1, \dots, p - 2$ ,  $0 \leq \alpha_{p-1} < 2\pi$ . As the Jacobian of this transformation is given as  $J(r, \alpha) = r^{p-1} \prod_{i=2}^{p-1} (\sin(\alpha_{i-1}))^{p-i}$  (Mardia



et al. 1989), it follows from an application of the change-of-variable rule that

$$\int_{\mathbb{R}^p} g(c_z + v^T v)dv = S(p) \int_0^\infty r^{p-1} g(c_z + r^2)dr \quad (38)$$

where  $S(p) = \int J(1, \alpha)d\alpha$  is the surface area of the  $p$ -dimensional unit sphere. Combining (37) and (38) with (35) concludes the proof for the case  $A \neq 0$ .

For the case  $A = 0$ , we have

$$\hat{b}_{BIE} = \frac{\int_{\mathbb{R}^p} \beta f_y(y + B(b - \beta))d\beta}{\int_{\mathbb{R}^p} f_y(y + B(b - \beta))d\beta} \quad (39)$$

which, with  $\|y - B\beta\|_{\Sigma_{yy}}^2 = \|\hat{e}\|_{\Sigma_{yy}}^2 + \|\hat{b} - \beta\|_{\Sigma_{bb}}^2$ ,  $\hat{b} = (B^T \Sigma_{yy}^{-1} B)^{-1} B^T \Sigma_{yy}^{-1} y$ , can be written as

$$\hat{b}_{BIE} = \frac{\int_{\mathbb{R}^p} \beta g\left(\|\hat{e}\|_{\Sigma_{yy}}^2 + \|\hat{b} - \beta\|_{\Sigma_{bb}}^2\right) d\beta}{\int_{\mathbb{R}^p} g\left(\|\hat{e}\|_{\Sigma_{yy}}^2 + \|\hat{b} - \beta\|_{\Sigma_{bb}}^2\right) d\beta} = \hat{b} \quad (40)$$

which proves the stated result. □

**Proof of Lemma 1 (Linear function of ECD):** Define the one-to-one transformation  $\tilde{y} = Ty$ , with  $\tilde{y} = [y_b^T, y_c^T]^T$ ,  $T = [B^{+T}, C]^T$ ,  $B^+ = (B^T \Sigma_{yy}^{-1} B)^{-1} B^T \Sigma_{yy}^{-1}$  and  $C$  a basis matrix of  $\mathcal{R}(B)^\perp$ . With  $f_y(y) = |\Sigma_{yy}|^{-1/2} g(\|y - Aa - Bb\|_{\Sigma_{yy}}^2)$  being the PDF of  $y$ , an application of the PDF transformation rule gives then

$$f_{y_b, y_c}(y_b, y_c) = \frac{g\left(\|y_b - \bar{y}_b\|_{\Sigma_{y_b y_b}}^2 + \|y_c - \bar{y}_c\|_{\Sigma_{y_c y_c}}^2\right)}{|\Sigma_{y_b y_b}|^{1/2} |\Sigma_{y_c y_c}|^{1/2}} \quad (41)$$

with  $\bar{y}_b = B^+ Aa + b$ ,  $\bar{y}_c = C^T Aa$ ,  $\Sigma_{y_b y_b} = (B^T \Sigma_{yy}^{-1} B)^{-1}$ ,  $\Sigma_{y_c y_c} = C^T \Sigma_{yy} C$ . Hence, the PDF of  $y_c = C^T y$  follows from integrating  $y_b$  out of  $f_{y_b, y_c}(y_b, y_c)$  as

$$f_{y_c}(y_c) = |\Sigma_{y_c y_c}|^{-1/2} g_c\left(\|y_c - \bar{y}_c\|_{\Sigma_{y_c y_c}}^2\right) \quad (42)$$

with

$$g_c(x) = |\Sigma_{y_b y_b}|^{-1/2} \int_{\mathbb{R}^p} g\left(x + \|y_b - \bar{y}_b\|_{\Sigma_{y_b y_b}}^2\right) dy_b. \quad (43)$$

This concludes the proof of the lemma. □

## References

Al Hage J, Xu P, Bonnifait P (2019) Student’s  $t$  information filter with adaptive degree of freedom for multi-sensor fusion. In: 22nd international conference on information fusion, Ottawa, Canada

Brack A, Henkel P, Gunther C (2014) Sequential best integer-equivariant estimation for GNSS. *Navigation* 61(2):149–158

Chmielewski MA (1981) Elliptically symmetric distributions: a review and bibliography. *Int Stat Rev* 49:67–74

Cabane S, Huang S, Simons G (1981) On the theory of elliptically contoured distributions. *J Multivar Anal* 11:368–385

De Jonge PJ, Tiberius CCJM (1996) The LAMBDA method for integer ambiguity estimation: implementation aspects. LGR-Series Publications of the Delft Geodetic Computing Centre No. 12

Dhital A, Bancroft JB, Lachapelle G (2013) A new approach for improving reliability of personal navigation devices under harsh GNSS signal conditions. *Sensors* 13:15221–15241. <https://doi.org/10.3390/s131115221>

Dins A, Ping Y, Schipper B (2015): Statistical characterization of BeiDou and GPS SIS errors in the Asian region. In: IEEE/AIAA 34th digital avionics systems conference (DASC)

Gosset W (Student, 1908) The probable error of a mean. *Biometrika* 6(1):1–25

Gradshteyn IS, Ryzhik IM (2007) Table of integrals, series, and products, 7th edn. Academic Press, Boca Raton

Heng L, Gao GX, Walter T, Enge P (2011) Statistical characterization of GPS signal-in-space errors. In: Institute of Navigation - International Technical Meeting 2011. ITM, pp 312–319

Kibria BMG, Joarder AH (2006) A short review of the multivariate  $t$ -distribution. *J Stat Res* 40(1):59–72

Koch KR (1999) Parameter estimation and hypothesis testing in linear models. Springer, Berlin

Mardia KV, Kent JT, Bibby JM (1989) Multivariate Analysis. Academic Press, Boca Raton

Odolinski R, Teunissen PJG (2020) On the best integer equivariant estimator for low-cost single-frequency multi-GNSS RTK positioning. In: Proceedings of the 2020 international technical meeting of the institute of navigation, San Diego, California, January 2020, pp 499–508

Rao CR (1973) Linear statistical inference and its applications, 2nd edn. Wiley, Hoboken

Roth M (2013) On the multivariate  $t$ -distribution. Report No. LiTH-ISY-R-3059. Department of Electrical Engineering. Linköpings universitet, Sweden

Teunissen PJG (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. *J Geodesy* 70:65–82

Teunissen PJG (1999a) The probability distribution of the GPS baseline for a class of integer ambiguity estimators. *J Geodesy* 73:275–284

Teunissen PJG (1999b) An optimality property of the integer least-squares estimator. *J Geodesy* 73:587–593

Teunissen PJG (2000) Adjustment theory: an introduction. Delft University Press, Delft

Teunissen PJG (2003) Theory of integer equivariant estimation with application to GNSS. *J Geodesy* 77:402–410

Verhagen S, Teunissen PJG (2005) Performance comparison of the BIE estimator with the float and fixed GNSS ambiguity estimators. A Window on the Future of Geodesy, International Association of Geodesy Symposia, Springer, Berlin Heidelberg, vol 128, pp 428–433

Wang Z, Zhou W (2019) Robust linear filter with parameter estimation under Student  $t$  measurement distribution. *Circuits Syst Signal Process* 38:2445–2470

- Wen Z, Henkel P, Guenther C, Brack A (2012) Best integer equivariant estimation for precise point positioning. In: Proceedings ELMAR-2012
- Zellner A (1973) Bayesian and non-Bayesian analysis of the regression model with multivariate Student- $t$  error terms. *Journal of the American Statistical Association* 71(354):400–405
- Zhong M, Xu X, Xu X (2018) A novel robust Kalman filter for SINS/GPS integration. In: Integrated communications, navigation, surveillance conference (ICNS) <https://doi.org/10.1109/ICNSURV.2018.8384892>
- Zhu H, Leung H, He Z (2012) A variational bayesian approach to robust sensor fusion based on Student  $t$ -distribution. *Inf Sci* 221(2013):201–214