

Papers on integer ambiguity resolution in this bundle

- De Jonge PJ, Tiberius C (1996). The LAMBDA method for integer ambiguity estimation: implementation aspects, LGR-Series, No 12. Technical report, Delft University of Technology.
- Teunissen PJG (1993). Least squares estimation of the integer GPS ambiguities. In *Invited lecture, Section IV Theory and Methodology, IAG General Meeting, Beijing*.
- Teunissen PJG (1995). The least-squares decorrelation adjustment: a method for fast GPS integer ambiguity estimation. *Journal of Geodesy*, 70: 65–82.
- Teunissen PJG (1998a). On the integer normal distribution of the GPS ambiguities. *Artificial Satellites*, 33(2): 49–64.
- Teunissen PJG (1998b). Success probability of integer GPS ambiguity rounding and bootstrapping. *Journal of Geodesy*, 72: 606–612.
- Teunissen PJG (1999). An optimality property of the integer least-squares estimator. *Journal of Geodesy*, 73(11): 587–593.
- Teunissen PJG (2003). Theory of carrier phase ambiguity resolution. *Wuhan University Journal of Natural Sciences*, 8: 471–484. 10.1007/BF02899809.
- Teunissen PJG (2010). Mixed integer estimation and validation for next generation GNSS. In W. Freeden, M. Nashed, and T. Sonar (Eds.), *Handbook of Geomathematics*, pp. 1101–1127. Springer Berlin Heidelberg.
- Verhagen S, Teunissen PJG (2006). New global navigation satellite system ambiguity resolution method compared to existing approaches. *Journal of Guidance, Control, and Dynamics*, 29(4): 981–991.

The LAMBDA method for integer ambiguity estimation: implementation aspects

Paul de Jonge and Christian Tiberius



August 1996

Contents

Preface	iii
Abstract	v
1 Introduction	1
2 Parameter estimation	3
2.1 Observation equations	3
2.2 Estimation criterion	4
2.3 Float solution	4
2.4 Integer ambiguity estimation	5
2.5 Fixed solution	5
3 Integer ambiguity estimation: transformation	7
3.1 Introduction	7
3.2 The decorrelating or Z -transformation, introduction	8
3.3 The decomposition of the variance-covariance matrix	8
3.4 Modifying the decomposition	10
3.5 The integer Gauss transformation	12
3.6 The reordering of the conditional variances	14
3.7 Putting it all together	15
3.8 Example Z -transformation	16
3.9 Concluding remarks	17
4 Integer ambiguity estimation: search	19
4.1 Introduction	19
4.2 Sequential conditional least-squares estimation	20
4.3 Computation of the bounds	21
4.4 Computation of the norm	23
4.5 The search algorithm	24
4.6 Collecting and storing the candidates	26
4.7 Alternating search around the conditional estimate	28
4.8 Shrinking the ellipsoidal region	28
4.9 The volume of the ellipsoidal region	29
4.10 Setting χ^2 , I	29
4.11 Setting χ^2 , II	30
4.12 Example ambiguity search	32

4.13 The back transformation	36
4.14 Concluding remarks	37
5 Flow diagram of the parameter estimation	39
6 Concluding remarks	41
A Inversion of a triangular matrix	43
References	45
LAMBDA bibliography	47

Preface

The development of the LAMBDA method for ambiguity resolution started with its introduction at the IAG General Meeting in Beijing in 1993 [1]. The implementation has been realized by P.J. de Jonge and C.C.J.M. Tiberius under supervision of professor P.J.G. Teunissen.

The underlying report provides a low level description of the method in its current implementation. Mrs. D. Morujão of the University of Coimbra, Portugal, and Mr. D. Charrier of SERCEL, France, are acknowledged for pointing out errors in the draft and suggesting several improvements. The second author received financial support of the Cornelis Lely foundation of Rijkswaterstaat.

Abstract

High precision relative GPS positioning is based on the very precise carrier phase measurements. In order to achieve high precision results within a short observation time span, the integer nature of the ambiguities has to be exploited. In this report the full procedure for parameter estimation based on the model of double difference GPS observations is reviewed, but the emphasis will be on the *integer estimation* of the GPS double difference ambiguities.

The LAMBDA method will be used for the integer estimation. LAMBDA stands for Least-squares AMBiguity Decorrelation Adjustment. By means of the Z -transformation, the ambiguities are decorrelated prior to the integer estimation. The integer minimization problem is then attacked by a discrete search over an ellipsoidal region, the ambiguity search ellipsoid.

The shape and orientation of the ellipsoid are governed by the variance covariance matrix of the ambiguities. The decorrelation realizes an ellipsoid that is very much sphere-like. It can be searched through very efficiently. The size of the ellipsoid can be controlled prior to the search using the volume function. The volume gives an indication of the number of candidates contained in the ellipsoid. A request for only a few candidates can be made, and this enables a straightforward implementation of the search. A limited number of candidates will be output of which one is the integer least-squares estimate for the vector of ambiguities.

The LAMBDA method provides, based on the float ambiguities and their variance covariance matrix, the integer least-squares estimate for the ambiguities. Therewith, the fixed solution can be computed. By the decorrelation, the integer estimation can be carried out very fast and efficiently. The total procedure typically takes 30 ms or less on a 486-66 MHz PC for a baseline with 12 ambiguities.

The method has been introduced in [1]. Preliminary fast positioning results using the LAMBDA method are given in [8], [9] and [10]. In this report, the implementation aspects of the method are discussed. A detailed description of the method is given, as well as the algorithms in the stylized Matlab notation of [7].

1

Introduction

High precision relative GPS positioning is based on the very precise carrier phase measurements. In order to achieve high precision results within a short observation time span, the integer nature of the ambiguities has to be exploited. In this report we will review the full procedure for parameter estimation based on the model of double difference GPS observations. The integer estimation therein is treated extensively. Note however, that integer estimation is not restricted to the double difference model; it can be applied as well in the models of un-differenced or single differenced observations with the rank-defect properly taken into account.

We will use the LAMBDA method for the integer estimation. LAMBDA stands for Least-squares AMBiguity Decorrelation Adjustment. After applying a decorrelating transformation, a sequential conditional adjustment is made upon the ambiguities. As a result, integer least-squares estimates for the ambiguities are obtained. The method has been introduced in [1]. Preliminary fast positioning results using the LAMBDA method are given in [8], [9] and [10].

The central chapters of this report are 2, 3 and 4. Chapter 2 describes the full procedure for parameter estimation. We start with the model of double difference observations, and first the float solution is computed. Based on the real valued estimate for the double difference ambiguities and the corresponding variance covariance matrix, the integer estimation is carried out. The procedure concludes with the fixed solution. Chapter 5 provides a flow diagram.

The integer minimization, which amounts to a discrete search over an ellipsoidal region, the ambiguity search ellipsoid, is treated in chapter 4. The volume of this ellipsoid, which can be computed prior to the search, plays an important role in controlling the size of the ellipsoid.

The Z -transformation is treated in chapter 3. By a systematic pair-wise decorrelation of the ambiguities, the – usually extremely elongated – ambiguity search ellipsoid is transformed towards a sphere-like shape. The Z -transformation is applied prior to the actual integer estimation. The discrete search can then be carried out very efficiently. The Z -transformation, chapter 3, and the search, chapter 4, are illustrated by means of a numerical example with three ambiguities.

2

Parameter estimation

2.1 Observation equations

In the discussion below we will set up the model of observation equations. We use a simple mathematical model appropriate for short baselines. Two types of unknowns occur: baseline coordinates and double difference ambiguities. We assume to have available only carrier phase observations. Code observations can be included as well; they are not related to the ambiguity terms.

The linearized double difference observation equations are collected in the following linear system of equations, see also section 2.1 in [2]:

$$y = Bb + Aa + e \quad (2.1)$$

where:

y the vector of observed minus computed double difference carrier phase measurements; $y \in R^m$

b the vector that contains the increments of the p baseline coordinates

a the vector of n double difference ambiguities

B the $m \times p$ design matrix for the baseline coordinates

A the $m \times n$ design matrix for the ambiguity terms

e the vector of unmodelled effects and measurement noise

The $m \times (p+n)$ design matrix $[B \ A]$ is assumed to have full rank equal $(p+n)$, i.e. enough observations have been made to determine the baseline coordinates and the double difference ambiguities. The variance-covariance matrix of the observables y is Q_y , the matrix is symmetric and positive definite.

Other unknowns, like e.g. atmospheric delay parameters can be included as well; they are then added to the vector b . In other words, all double difference ambiguities are contained in a , all other parameters in b . The double difference ambiguities are placed at the end for reasons of computational efficiency.

2.2 Estimation criterion

The least-squares principle will be used to compute estimates for the baseline coordinates and the *integer* double difference ambiguities:

$$\min_{b,a} \|y - Bb - Aa\|_{Q_y^{-1}}^2 \text{ with } b \in R^p \text{ and } a \in Z^n \quad (2.2)$$

The parameter estimation is carried out in three steps, see section 2.2 in [2]. The steps are the float solution, the integer ambiguity estimation and the fixed solution. Although model validation is important as well, we will concentrate in this report on the estimation of the unknown parameters. In particular we will treat the integer ambiguity estimation.

2.3 Float solution

The minimization (2.2) is carried out with $b \in R^p, a \in R^n$. Real valued estimates for the double difference ambiguities will be obtained.

Usually the full rank least-squares problem is solved via normal equations, see section 5.3 of [7]. The system of normal equations $Nx = r$, with N the normal matrix and r the right hand side, for (2.1) reads

$$\underbrace{\begin{bmatrix} B^*Q_y^{-1}B & B^*Q_y^{-1}A \\ A^*Q_y^{-1}B & A^*Q_y^{-1}A \end{bmatrix}}_N \underbrace{\begin{bmatrix} b \\ a \end{bmatrix}}_x = \underbrace{\begin{bmatrix} B^*Q_y^{-1}y \\ A^*Q_y^{-1}y \end{bmatrix}}_r \quad (2.3)$$

The Cholesky factorization of the normal matrix is made, see theorem 4.2.5 of [7]: $N = CC^*$, with matrix C a full rank $(p + n)$ lower triangular matrix. Algorithm 4.2.2. of [7] can be used for the Cholesky factorization. The system $CC^*x = r$ is then solved by forward- and backward substitution. The estimates and the variance-covariance matrix are:

$$\begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} \begin{bmatrix} Q_{\hat{b}} & Q_{\hat{b}\hat{a}} \\ Q_{\hat{a}\hat{b}} & Q_{\hat{a}} \end{bmatrix} \quad (2.4)$$

In the integer ambiguity estimation we will use the Cholesky factor G of the inverse of the variance-covariance matrix $Q_{\hat{a}}$. This is why the unknowns have been put in the order: first baseline coordinates and then double difference ambiguities. The Cholesky factor C is partitioned as

$$C = \begin{bmatrix} * & 0 \\ * & G \end{bmatrix} \quad (2.5)$$

with G a lower triangular matrix of dimension n , for which holds that $Q_{\hat{a}}^{-1} = GG^*$. The Cholesky factor G becomes available at no extra cost in the float solution.

2.4 Integer ambiguity estimation

The second step consists of

$$\min_a \|\hat{a} - a\|_{Q_{\hat{a}}^{-1}}^2 \text{ with } a \in Z^n \quad (2.6)$$

This minimization yields the integer least-squares estimate for the vector of ambiguities: \check{a} . The computation of the integer estimate will be treated in detail in chapters 3 and 4.

2.5 Fixed solution

The final solution, with the ambiguities fixed to their integer least-squares estimates \check{a} , reads

$$\check{b} = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} (\hat{a} - \check{a}) \quad (2.7)$$

The least-squares estimates \check{b} and \check{a} are the solution to the constrained minimization (2.2). Note that in practice \check{b} is not computed with (2.7). The final estimates for the baseline coordinates follow from the first p equations of system (2.3) in which the vector of integer estimates \check{a} has been substituted for a . Note that the Cholesky factor of matrix $B^* Q_y^{-1} B$ is already available. Estimate \check{b} is then easily obtained via forward and backward substitution. The variance-covariance matrix $Q_{\check{b}|\check{a}}$, see section 2.2 of [2], can be computed via inversion of the Cholesky factor.

3

Integer ambiguity estimation: transformation

3.1 Introduction

For the computation of the integer least-squares estimate \hat{a} , we use the LAMBDA method. It has been introduced in [1] and reviewed in [5]. The two main features of the LAMBDA method are

1. the decorrelation of the ambiguities, realized by a reparametrization, or as we call it, a Z -transformation, of the original ambiguities a to new ambiguities $z = Z^*a$, and
2. the actual integer ambiguity estimation

Through the Z -transformation the variance-covariance matrix is transformed accordingly:

$$Q_{\hat{z}} = Z^*Q_{\hat{a}}Z \quad (3.1)$$

The efficiency of the method comes from the decorrelation step and has been explained in detail by analysis of the precision and correlation of the GPS double difference ambiguities in [4]-[6].

The actual integer minimization is then made upon the transformed ambiguities. In practice the minimization (2.6) amounts to a search over grid points inside the n -dimensional ambiguity hyper-ellipsoid, defined by the variance-covariance matrix of the ambiguities

$$(\hat{z} - z)^*Q_{\hat{z}}^{-1}(\hat{z} - z) \leq \chi^2 \quad (3.2)$$

The search must yield the grid point that is nearest to the real valued estimate, with nearness measured in the metric of the variance-covariance matrix, see section 3 of [1]. Note that an n -dimensional ellipsoid, is actually an $(n - 1)$ dimensional surface in the R^n .

In this and the following chapter, the algorithmic details of the LAMBDA method will be discussed. The input basically consists of \hat{a} and $Q_{\hat{a}}$, the estimate for the vector of ambiguities and the variance-covariance matrix from the float solution.

3.2 The decorrelating or Z -transformation, introduction

To explain the concept of the decorrelating transformation we decompose the variance-covariance matrix $Q_{\hat{a}}$ into

$$Q_{\hat{a}} = L^{-*} D^{-1} L^{-1} \quad (3.3)$$

Note that this corresponds to the inverse of the LDL^* -decomposition of $Q_{\hat{a}}^{-1}$ which is easily derived from the already computed Cholesky factor (see equation (2.5)). The principle of the decorrelation is to find a matrix Z , which is an integer approximation of matrix L . If we would be able to find an integer matrix Z that fulfills the requirements in [3] and that exactly equals L , then with (3.1)

$$Q_{\hat{z}} = Z^* Q_{\hat{a}} Z = Z^* L^{-*} D^{-1} L^{-1} Z = D^{-1} \quad (3.4)$$

The transformed ambiguities \hat{z} are fully decorrelated and the integer minimization reduces to a simple rounding of the real valued estimates, see section 2.3 of [2]. In practice a complete decorrelation will not be possible due to the integer requirement.

The result of the decorrelation process is the square $n \times n$ transformation matrix Z ; for its properties we refer to [3]. The estimate \hat{z} follows from $\hat{z} = Z^* \hat{a}$. The factors of the variance-covariance matrix are updated in the decorrelation process: \tilde{D}^{-1} and \tilde{L}^{-1} . They satisfy

$$Q_{\hat{z}} = \tilde{L}^{-*} \tilde{D}^{-1} \tilde{L}^{-1} \quad (3.5)$$

The problem (2.6) has now been transformed into the following minimization

$$\min_z \| \hat{z} - z \|_{Q_{\hat{z}}^{-1}}^2 \text{ with } z \in Z^n \quad (3.6)$$

Inversion of \tilde{L}^{-1} yields \tilde{L} and inversion of \tilde{D}^{-1} yields \tilde{D} . They satisfy

$$Q_{\hat{z}}^{-1} = \tilde{L} \tilde{D} \tilde{L}^* \quad (3.7)$$

This decomposition in \tilde{L} and \tilde{D} is then used in the search, see chapter 4.

3.3 The decomposition of the variance-covariance matrix

The ambiguity decorrelating transformation can be computed in several ways depending on the matrix one starts with, and on the kind of factorization one uses. One can either use the variance-covariance matrix of the ambiguities or its inverse. The factorization can be either an LDL^* , or an L^*DL factorization. Schematic we have

$$\begin{aligned} Q_{\hat{a}}^{-1} &= L_1 D_1 L_1^* & ; & Q_{\hat{a}} = L_1^{-*} D_1^{-1} L_1^{-1} & \text{cf. [1]} \\ Q_{\hat{a}}^{-1} &= L_2^* D_2 L_2 & ; & Q_{\hat{a}} = L_2^{-1} D_2^{-1} L_2^{-*} \\ Q_{\hat{a}} &= L_3^* D_3 L_3 & ; & Q_{\hat{a}}^{-1} = L_3^{-1} D_3^{-1} L_3^{-*} \\ Q_{\hat{a}} &= L_4 D_4 L_4^* & ; & Q_{\hat{a}}^{-1} = L_4^{-*} D_4^{-1} L_4^{-1} & \text{cf. [2]} \end{aligned} \quad (3.8)$$

where:

L_i is a unit lower triangular matrix (i.e. with ones on the diagonal) and

D_i is a diagonal matrix, $D_i = \text{diag}(d_1, \dots, d_n)$

Note that $L_1 = L_3^{-1}, L_2 = L_4^{-1}$ and that $D_1 = D_3^{-1}, D_2 = D_4^{-1}$. In [1] the LDL^* factorization of the inverse of the variance-covariance matrix is used, and in [2] the LDL^* factorization of the variance-covariance matrix itself.

The implementation of the LAMBDA method described in this report is based on the description in [1]. The reason for it is that the Cholesky factor G is already available in the float solution (see equation (2.5)). The LDL^* -decomposition is easily constructed from the Cholesky factor. It holds that

$$G = L\sqrt{D} \quad (3.9)$$

Note that matrix $Q_{\hat{a}}^{-1}$ is symmetric positive definite: for the diagonal elements of D it holds that $d_i > 0 \forall i = 1, \dots, n$.

The corresponding factorization of the variance-covariance matrix reads

$$Q_{\hat{a}} = L^{-*} D^{-1} L^{-1} \quad (3.10)$$

An interpretation can be given to the elements of matrix D^{-1} , see section 5 of [1]: they are the conditional variances of the double difference ambiguities.

$$d_i^{-1} = \sigma_{\hat{a}_{i|i+1,\dots,n}}^2 \quad (3.11)$$

The inversion of the factor can be done in place, an algorithm for it can be found in chapter 6 of [11], and in appendix A of this report. The procedure proposed here is computationally more efficient than first computing the variance-covariance matrix, and then factoring it.

Instead of inverting the LDL^* -decomposition, one can compute $Q_{\hat{a}} = L^{-*} D^{-1} L^{-1}$ directly. This should be done when the factor L is not already available. We will give two algorithms for this factorization.

The first, *FMFAC5*, is an outer product formulation, and is used to explain some features of the construction of the transformation matrix Z in section 3.4.

The second algorithm, *FMFAC6*, is a ‘bordering method’ formulation (see [17]), and has as additional feature a check for positive-definiteness of the variance-covariance matrix Q . An indication for the numerical condition of matrix Q is given by the smallest collinearity number c_{min} , [ibid]. This number ranges from one to zero; if the number is less than a pre-defined constant ϵ (for the cases of 1 second time span we take $\epsilon = 10^{-9}$), the matrix is considered to be singular.

Algorithm *FMFAC5*: Given the symmetric positive definite matrix Q , an L^*DL factorization is computed by the outer product method. Matrix L may over write matrix Q ; the latter is destroyed during the computation. Only the lower triangle of Q is accessed.

```

for  $i = n : -1 : 1$ 
   $D(i, i) = Q(i, i)$ 
   $L(i, 1 : i) = Q(i, 1 : i) / \sqrt{Q(i, i)}$ 
  for  $j = 1 : i - 1$ 
     $Q(j, 1 : j) = Q(j, 1 : j) - L(i, 1 : j)L(i, j)$ 
  end
   $L(i, 1 : i) = L(i, 1 : i) / L(i, i)$ 
end

```

Algorithm *FMFACT6*: Given the symmetric positive definite matrix Q , an L^*DL factorization is computed via a backward Cholesky factorization in a bordering method formulation. Also the smallest collinearity number c_{min} is computed. Matrix L over writes matrix Q . Only the lower triangle of Q is accessed. The function *dot* computes the dot or inner product of two vectors.

```

 $L = Q$ 
 $c_{min} = 1$ 
for  $j = n : -1 : 1$ 
  for  $i = n : -1 : j + 1$ 
     $L(i, j) = (L(i, j) - dot(L(i + 1 : n, j), L(i + 1 : n, i))) / L(i, i)$ 
  end
   $t = L(j, j) - dot(L(j + 1 : n, j), L(j + 1 : n, j))$ 
   $c = t / L(j, j)$ 
  if  $c < c_{min}$ 
     $c_{min} = c$ 
  end
   $L(j, j) = \sqrt{t}$ 
end
for  $i = 1 : n$ 
   $L(i, 1 : i - 1) = L(i, 1 : i - 1) / L(i, i)$ 
   $D(i) = L(i, i)L(i, i)$ 
   $L(i, i) = 1$ 
end

```

3.4 Modifying the decomposition

Note: The construction of the transformation matrix Z is based on the variance-covariance matrix $Q_{\hat{a}}$. To simplify notation we will assume for the remainder of this chapter 3 that we have the L^*DL factorization of $Q_{\hat{a}}$ (the third option of (3.8) instead of the first), thus $d_i = \sigma_{\hat{a}_{i|i+1,\dots,n}}^2$.

The construction of the $n \times n$ Z matrix in (3.1) consists of a sequence of integer approximated Gauss transformations and permutations. Both are admissible ambiguity transformations (see e.g. [1] and [3]). The actual decorrelation is carried out by the integer Gauss transformation. If necessary the ambiguities are reordered, by a permutation, in order to allow for further decorrelation.

Before we continue, we will discuss the triangular decomposition after transformation, analogously to the decomposition discussed in [2] (formulae 47-49 at page 79).

The unit lower triangular matrix L and the diagonal matrix D are partitioned into

$$L = \begin{bmatrix} L_{11} & & \\ L_{21} & L_{22} & \\ L_{31} & L_{32} & L_{33} \end{bmatrix} \text{ and } D = \begin{bmatrix} D_{11} & & \\ & D_{22} & \\ & & D_{33} \end{bmatrix} \quad (3.12)$$

Sub-matrix L_{22} is of order 2, L_{11} of order $i - 1$, and L_{33} of order $n - i - 1$. The dimension of the other sub-matrices in L and D are determined accordingly. If we apply a two-dimensional ambiguity transformation on the i -th and the $(i + 1)$ -th ambiguity, with the block-diagonal matrix

$$Z = \begin{bmatrix} I_{i-1} & & \\ & Z_{22} & \\ & & I_{n-i-1} \end{bmatrix} \quad (3.13)$$

with Z_{22} a square and full rank 2×2 matrix, we get the new triangular decomposition $L'^* D' L'$, with

$$L' = \begin{bmatrix} L_{11} & & \\ \bar{L}_{21} & \bar{L}_{22} & \\ L_{31} & \bar{L}_{32} & L_{33} \end{bmatrix} \text{ and } D' = \begin{bmatrix} D_{11} & & \\ & \bar{D}_{22} & \\ & & D_{33} \end{bmatrix} \quad (3.14)$$

The fact that only L_{21}, L_{22}, L_{32} and D_{22} do change, can be explained by looking at the outer product form of the $L^* D L$ factorization as given in section 3.3. We have (the variance-covariance matrix is symmetric):

$$Z^* Q \hat{Q} Z = \begin{bmatrix} Q_{11} & & \\ Z_{22}^* Q_{21} & Z_{22}^* Q_{22} Z_{22} & \\ Q_{31} & Q_{32} Z_{22} & Q_{33} \end{bmatrix} = \begin{bmatrix} Q_{11} & & \\ \bar{Q}_{21} & \bar{Q}_{22} & \\ Q_{31} & \bar{Q}_{32} & Q_{33} \end{bmatrix} \quad (3.15)$$

Let us look at the updating step of the algorithm i.e. the part where an outer product is subtracted from the matrix. The parts of the outer product that are affected by Z always coincide with the parts of Q affected by Z . So the ‘reduced’ Q has the same structure as in (3.15). Since L is basically obtained from extracting rows from the reduced Q divided by the square root of the corresponding diagonal element, and D is set equal to the diagonal element, the modified decomposition will be as indicated in (3.14).

The modified factor can be related to the original one. We know that

$$L'^* D' L' = Z^* L^* D L Z$$

From this relation we easily derive for \bar{L}_{32}

$$L_{33}^* D_{33} \bar{L}_{32} = L_{33}^* D_{33} L_{32} Z_{22}$$

or,

$$\bar{L}_{32} = L_{32} Z_{22} \quad (3.16)$$

For \bar{L}_{22} holds that

$$\bar{L}_{22}^* \bar{D}_{22} \bar{L}_{22} + \bar{L}_{32}^* D_{33} \bar{L}_{32} = Z_{22}^* (L_{22}^* D_{22} L_{22} + L_{32}^* D_{33} L_{32}) Z_{22}$$

or, using (3.16)

$$\bar{L}_{22}^* \bar{D}_{22} \bar{L}_{22} = Z_{22}^* (L_{22}^* D_{22} L_{22}) Z_{22} \quad (3.17)$$

And for \bar{L}_{21} holds that

$$\bar{L}_{22}^* D_{22} \bar{L}_{21} + \bar{L}_{32}^* D_{33} L_{31} = Z_{22}^* (L_{22}^* D_{22} L_{21} + L_{32}^* D_{33} L_{31})$$

from which can be derived after substituting (3.16)

$$\bar{L}_{22}^* \bar{D}_{22} \bar{L}_{21} = Z_{22}^* L_{22}^* D_{22} L_{21}$$

Using the relation $Z_{22}^* L_{22}^* D_{22} = \bar{L}_{22}^* \bar{D}_{22} \bar{L}_{22} (L_{22} Z_{22})^{-1}$ which follows from equation (3.17) we finally get

$$\bar{L}_{21} = \bar{L}_{22} (L_{22} Z_{22})^{-1} L_{21} \quad (3.18)$$

Conclusion: once matrix Q is factored as $Q = L^* D L$, the factorization of $Z^* Q Z$ with Z defined in (3.13), can be efficiently computed from the existing factorization. In the unit lower triangular matrix L only the rows i and $i+1$ and the columns i and $i+1$ change. The modification of the diagonal matrix D is limited to the elements (i,i) and $(i+1,i+1)$. The transformation of the variance-covariance matrix can be realized by updating the factors L and D .

3.5 The integer Gauss transformation

The relations found in section 3.4 can be further simplified if we restrict the transformation (sub-)matrix to a single Gauss-transformation (see [7] par 3.2.1):

$$Z_{22} = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \quad (3.19)$$

Then

$$\bar{L}_{32} = \begin{bmatrix} l_{i+2,i} + \alpha l_{i+2,i+1} & l_{i+2,i+1} \\ l_{i+3,i} + \alpha l_{i+3,i+1} & l_{i+3,i+1} \\ \vdots & \vdots \\ l_{n,i} + \alpha l_{n,i+1} & l_{n,i+1} \end{bmatrix} \quad (3.20)$$

$$\bar{L}_{22} = \begin{bmatrix} 1 & 0 \\ l_{i+1,i} + \alpha & 1 \end{bmatrix} \text{ and } \bar{D}_{22} = \begin{bmatrix} d_i & 0 \\ 0 & d_{i+1} \end{bmatrix} \quad (3.21)$$

$$\bar{L}_{21} = L_{21} \quad (3.22)$$

We see that only the unit lower triangular matrix L changes. The diagonal matrix D is left unchanged by a single Gauss transformation. In other words: the conditional variances do not change, only the conditional covariances and thus the ‘normal’ variances change. This can intuitively be understood since the order of ambiguities was not changed.

To achieve full decorrelation we would like to choose $\alpha = -l_{i+1,i}$. In general $l_{i+1,i} \notin Z$. To meet the requirements in [3], α is approximated by $-[l_{i+1,i}]$, where $[.]$ stands for the nearest integer operator (*nint*). With this integer approximation, we can make the absolute value of any non-diagonal element of L less than or equal to .5. The integer Gauss transformation was discussed in [3] and [8].

In general, if we take for Z the unit matrix with an additional α at position (i,j) , with $i > j$, the elements that change are

$$l'_{i,j} = l_{i,j} + \alpha \quad (3.23)$$

$$l'_{k,j} = l_{k,j} + \alpha l_{k,i} \text{ for } k = i+1, \dots, n \quad (3.24)$$

The algorithm for the computation of a Z -transformation matrix, that will make the absolute value of all non-diagonal elements of L less than or equal to .5, is given as

Algorithm ZTRAN: Given the unit lower triangular factor L from the L^*DL factorization of the variance-covariance matrix $Q_{\hat{a}}$, a Z -transformation matrix is computed. Z will be lower triangular on output if set to the unit matrix on input. The computation of this matrix is performed *column wise* from *right to left*, i.e. from n to 1. Instead of decorrelating all n columns, one can apply the algorithm to only one column k , by changing the interval for i to $i = k : k$. The vector with estimates \hat{a} is transformed to $\hat{z} = Z^*\hat{a}$. The original factor L is transformed to $L' = LZ$ for which holds: $L'^*DL' = Q_{\hat{z}}$. Factor L remains lower triangular.

```

for  $i = n : -1 : 1$ 
  for  $j = i + 1 : n$ 
     $\mu = \text{nint}(L(j,i))$ 
    if  $\mu \neq 0$ 
       $L(j:n,i) = L(j:n,i) - \mu L(j:n,j)$ 
       $Z(1:n,i) = Z(1:n,i) - \mu Z(1:n,j)$ 
       $a(i) = a(i) - \mu a(j)$ 
    end
  end
end

```

In the inner loop, if $\mu \neq 0$, column j times μ is subtracted from column i , for matrix Z and factor L .

3.6 The reordering of the conditional variances

As was explained in [2] and [6] the spectrum of the conditional variances shows, in the case of a single baseline, a distinctive discontinuity when passing from the third to the fourth ambiguity. The size of this discontinuity is governed by the length of the observation time span. In [4] and [6] it was shown that it is this discontinuity in the spectrum that causes the search to be highly inefficient. Carrying out the integer estimation (2.6) will be a very time consuming task. As we saw in section 3.5, the integer Gauss transformation decorrelates the ambiguities, but it leaves the conditional variances, and thus the discontinuity, intact. Looking at the definition of conditional variance, it seems logical to change the order of the ambiguities if we want to change the spectrum. In [6], [12] and [13] this problem, and the solution for it, is explained geometrically in terms of the form and orientation of the ambiguity search space.

To change the order of the i -th and the $(i+1)$ -th ambiguity we again apply a local transformation but now using the 2-by-2 permutation matrix P for Z_{22} . With (3.16), (3.17) and (3.18) we have

$$\bar{L}_{32} = L_{32}P \quad (3.25)$$

$$\bar{L}_{22}^* \bar{D}_{22} \bar{L}_{22} = P^* L_{22}^* D_{22} L_{22} P \quad (3.26)$$

$$\bar{L}_{21} = \bar{L}_{22}(L_{22}P)^{-1}L_{21} \quad (3.27)$$

where

$$P = P^* = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

If we work out these relations we get

$$\bar{L}_{32} = \begin{bmatrix} l_{i+2,i+1} & l_{i+2,i} \\ l_{i+3,i+1} & l_{i+3,i} \\ \vdots & \vdots \\ l_{n,i+1} & l_{n,i} \end{bmatrix} \quad (3.28)$$

$$\bar{L}_{22} = \begin{bmatrix} 1 & 0 \\ l'_{i+1,i} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{l_{i+1,i}d_{i+1}}{d_i + l_{i+1,i}^2 d_{i+1}} & 1 \end{bmatrix} \quad (3.29)$$

$$\bar{D}_{22} = \begin{bmatrix} d'_i & 0 \\ 0 & d'_{i+1} \end{bmatrix} = \begin{bmatrix} d_{i+1} - \frac{l_{i+1,i}^2 d_{i+1}^2}{d_i + l_{i+1,i}^2 d_{i+1}} & 0 \\ 0 & d_i + l_{i+1,i}^2 d_{i+1} \end{bmatrix} \quad (3.30)$$

$$\bar{L}_{21} = \begin{bmatrix} -l_{i+1,i} & 1 \\ \frac{-l_{i+1,i}^2 d_{i+1}}{d_i + l_{i+1,i}^2 d_{i+1}} + 1 & \frac{l_{i+1,i} d_{i+1}}{d_i + l_{i+1,i}^2 d_{i+1}} \end{bmatrix} L_{21} \quad (3.31)$$

Simplifying equations (3.29)-(3.31) yields

$$d'_{i+1} = d_i + l_{i+1,i}^2 d_{i+1} \quad (3.32)$$

$$d'_i = \frac{d_i}{d'_{i+1}} d_{i+1} \quad (3.33)$$

$$l'_{i+1,i} = \frac{d_{i+1}}{d'_{i+1}} l_{i+1,i} \quad (3.34)$$

and

$$\bar{L}_{21} = \begin{bmatrix} -l_{i+1,i} & 1 \\ \frac{d_i}{d'_{i+1}} & l'_{i+1,i} \end{bmatrix} L_{21} \quad (3.35)$$

The factors L and D are updated, as shown in (3.14) and (3.15), as to correspond to the transformed (permuted) ambiguities.

3.7 Putting it all together

For the actual integer minimization we strive for largely decorrelated ambiguities, and furthermore to have the most precise ambiguity at position n . The reason for the latter was explained in [1]. In other words, we strive for

$$d_n \leq \dots \leq d_1 \text{ with } d_i \text{ from } D \text{ of } Q_{\hat{a}} = L^* D L$$

and therefore we interchange two conditional variances if

$$d'_{i+1} < d_{i+1} \quad (3.36)$$

Alternately we will have a decorrelation and a reordering step. We start with the last ambiguity and we try to reach the first one. At each step i we check whether the interchange of d_i and d_{i+1} will decrease the value for the latter. After each interchange we start again at the last ambiguity. The algorithm ends if during one sweep from n to 1 no further interchanges can be performed.

The algorithm requires element $l_{i+1,i}$ to be as small as possible. Therefore we take care that the absolute values of the off-diagonal elements of columns $i \dots n$ are less or equal to .5, by applying the column wise decorrelating Z -transformation of section (3.5). In algorithm *SRC1* we keep track by variable *i1* what is the last column fulfilling this criterion of ‘cleanness’. The logical ‘sw’ keeps track if there was an interchange of diagonal elements during a sweep. If during one complete sweep from n to 1, no diagonal elements were interchanged, the algorithm terminates.

Algorithm *SRC1*: Given the unit lower triangular factor L and the diagonal matrix D from the $L^* D L$ factorization (third option in (3.8)) of the variance-covariance matrix $Q_{\hat{a}}$, the Z -transformation matrix is computed. On input the matrix Z is set to the unit matrix. The vector with estimates \hat{a} is transformed to $\hat{z} = Z^* \hat{a}$ (thus using the transpose of Z). The original factor L and diagonal matrix D are transformed to L' and D' for which holds: $L'^* D' L' = Q_{\hat{z}}$.

```

i1 = n - 1 ; sw = true
while sw = true
    i = n ; sw = false
    while sw = false and i > 1
        i = i - 1
        if i ≤ i1
            apply algorithm ZTRAN to column i
        end
        δ = D(i, i) + L(i + 1, i)2D(i + 1, i + 1)           cf. eq. (3.32)
        if δ < D(i + 1, i + 1)
            λ3 = D(i + 1, i + 1)L(i + 1, i)/δ ; η = D(i, i)/δ   cf. eq. (3.34)
            D(i, i) = ηD(i + 1, i + 1)                           cf. eq. (3.33)
            D(i + 1, i + 1) = δ
            for j = 1 : i - 1
                λ1 = L(i, j) ; λ2 = L(i + 1, j)
                L(i, j) = -L(i + 1, i)λ1 + λ2                   cf. eq. (3.35)
                L(i + 1, j) = ηλ1 + λ3λ2
            end
            L(i + 1, i) = λ3
            swap columns L(i + 2 : n, i) and L(i + 2 : n, i + 1)   cf. eq. (3.28)
            swap columns Z(1 : n, i) and Z(1 : n, i + 1)
            swap elements a(i) and a(i + 1)
            i1 = i ; sw = true
        end
    end
end

```

3.8 Example *Z*-transformation

As an illustration of the *Z*-transformation we will apply it to an artificial three-dimensional example. The synthetic variance covariance matrix results from the addition of a scaled unit matrix and a rank-2 matrix with elements that are significantly larger than the scale factor of the first matrix, see also [6].

$$Q_{\hat{a}} = \begin{bmatrix} 6.290 & 5.978 & 0.544 \\ 5.978 & 6.292 & 2.340 \\ 0.544 & 2.340 & 6.288 \end{bmatrix} \quad (3.37)$$

The ambiguities have large variances and in particular a_1 and a_2 are heavily correlated.

$$\begin{aligned} \sigma_{\hat{a}_1} &= 2.508 & \rho_{\hat{a}_1 \hat{a}_2} &= 0.950 \\ \sigma_{\hat{a}_2} &= 2.508 & \rho_{\hat{a}_1 \hat{a}_3} &= 0.086 \\ \sigma_{\hat{a}_3} &= 2.508 & \rho_{\hat{a}_2 \hat{a}_3} &= 0.372 \end{aligned} \quad (3.38)$$

The matrix Z^* reads

$$Z^* = \begin{bmatrix} 1 & -1 & 0 \\ -2 & 3 & -1 \\ 3 & -3 & 1 \end{bmatrix} \quad (3.39)$$

Note that the matrix Z^* given here differs from the one obtained using algorithm *SRC1*. Reflections (change of sign) and permutations (reordering) of the ambiguities do not change the amount of correlation between the ambiguities. A permutation of the transformed ambiguities has been applied for plotting purposes (cf. figure 4.5).

Matrix Z^* has integer elements and $|Z^*| = 1$, see [3]. The variance covariance matrix of the transformed ambiguities \hat{z} reads

$$Q_{\hat{z}} = \begin{bmatrix} 0.626 & 0.230 & 0.082 \\ 0.230 & 4.476 & 0.334 \\ 0.082 & 0.334 & 1.146 \end{bmatrix} \quad (3.40)$$

and it can be seen that the ambiguities are largely decorrelated.

$$\begin{aligned} \sigma_{\hat{z}_1} &= 0.791 & \rho_{\hat{z}_1 \hat{z}_2} &= 0.137 \\ \sigma_{\hat{z}_2} &= 2.116 & \rho_{\hat{z}_1 \hat{z}_3} &= 0.097 \\ \sigma_{\hat{z}_3} &= 1.071 & \rho_{\hat{z}_2 \hat{z}_3} &= 0.147 \end{aligned} \quad (3.41)$$

3.9 Concluding remarks

In this chapter we discussed the implementation aspects of the first part of the LAMBDA method, viz. the transformation that decorrelates the ambiguities. The second part, the sequential conditional least-squares adjustment which translates into a depth-first search through the ambiguity search space, defined by the variance-covariance matrix of the ambiguities, will be the topic of chapter 4.

The transformation was materialized in matrix Z in this chapter. After the actual integer minimization, the new ambiguities \hat{z} have to be transformed back to the original ambiguities a , see section 4.13. Several alternatives are possible. They are discussed in [18].

One can also compute instead of Z , the matrix Z^{-*} directly. In algorithm *ZTRAN*, section 3.5, subtracting μ times column j from column i , yields the new column i . For computing Z^{-*} one has instead to add μ times column i to column j , to obtain the new column j .

With other alternatives, the transformation is constructed in an *implicit* manner. The fixed solution (2.7) can be computed using the transformed integer ambiguities \hat{z} directly, cf. equation (3) in [2]. Instead of starting with $Z = I$ on the input and computing matrix Z , one can insert matrix $Q_{\hat{b}\hat{a}}$ and transform it into $Q_{\hat{b}\hat{z}}$. Matrix $Q_{\hat{z}}^{-1}$ is obtained from the updated factors L and D .

4

Integer ambiguity estimation: search

4.1 Introduction

In this chapter the actual integer ambiguity estimation will be discussed. The integer estimation is also referred to as search. Based on the results of the float solution (see section 2.3) a search will be performed in order to come up with the most likely integer candidate¹ for the vector of ambiguities.

Decorrelating the ambiguities, as discussed in the previous chapter, is not a prerequisite for the integer ambiguity estimation. The search can be performed on the original ambiguities a as well, instead of on the transformed ambiguities z . The decorrelation, however, is largely beneficial for the efficiency of the search, see [4]-[6]. In the sequel the ambiguities are denoted by a , whether they represent transformed ambiguities or not.

The implementation of the search is based on the correspondence of the LDL^* -decomposition of matrix $Q_{\hat{a}}^{-1}$ (the first option of (3.8)) and the sequential conditional least-squares estimation, see section 5 of [1]. The input of this step consists of matrices L and D and the real valued estimate \hat{a} . The output is the integer least-squares estimate \check{a} .

The sequential conditional adjustment is treated in a straightforward manner in the sections 4.2 and 4.3. The implementation is given in 4.5 and 4.6. The size of the ambiguity search ellipsoid can be controlled prior to the search by means of the volume: sections 4.9 through 4.11. The search is then illustrated by means of an example with three ambiguities. As the search is made on the transformed ambiguities z , a back transformation is needed, section 4.13. It finally provides the integer least-squares estimate for the vector of original double difference ambiguities: \check{a} .

¹What is a candidate? It is a grid point that is inside or on the ambiguity search ellipsoid; it satisfies equation (4.1)

4.2 Sequential conditional least-squares estimation

As discussed in [1], no standard techniques are available for solving (2.6). A discrete search is employed instead. An ellipsoidal region in R^n is taken, on the basis of which a search is performed for the minimizer of (2.6).

$$(\hat{a} - a)^* Q_{\hat{a}}^{-1} (\hat{a} - a) \leq \chi^2 \quad (4.1)$$

For a discussion on the value for χ^2 , the constant that controls the size of the ellipsoidal region, the reader is referred to [15], [19] and sections 4.9-4.11 on the volume of the ellipsoidal region.

With the LDL^* -decomposition of matrix $Q_{\hat{a}}^{-1}$, expanding (4.1) yields

$$\sum_{i=1}^n d_i \left[(a_i - \hat{a}_i) + \sum_{j=i+1}^n l_{ji} (a_j - \hat{a}_j) \right]^2 \leq \chi^2 \quad (4.2)$$

Equation (4.2) is just an algebraic development of (4.1). In section 4.3 we will pursue on this development, as the algorithm for the integer estimation is based on (4.2).

As mentioned above, the search can also be given a statistical interpretation: the sequential conditional adjustment. The term between the square brackets is the difference of a_i and $\hat{a}_{i|i+1,\dots,n}$ and together with (3.11), (4.2) can be rewritten in

$$\sum_{i=1}^n \frac{(a_i - \hat{a}_{i|i+1,\dots,n})^2}{\sigma_{\hat{a}_{i|i+1,\dots,n}}^2} \leq \chi^2 \quad (4.3)$$

see also equation (25) of [1]. The conditional estimate $\hat{a}_{i|i+1,\dots,n}$ is the estimate for a_i conditioned on a_j with $j = i + 1, \dots, n$. The conditional estimate for ambiguity i thus reads

$$\hat{a}_{i|i+1,\dots,n} = \hat{a}_i - \sum_{j=i+1}^n l_{ji} (a_j - \hat{a}_j) \quad (4.4)$$

Equation (4.4) clearly shows that conditioning on a_j for $j = i + 1, \dots, n$ affects the estimate for a_i due to the correlation between the ambiguities. Only in case there is no correlation, $L = I$, we have

$$\hat{a}_{i|i+1,\dots,n} = \hat{a}_i \quad (4.5)$$

In the integer ambiguity estimation using the sequential conditional least-squares adjustment, the ambiguities a_j are conditioned on integers. The variance of the conditional estimator equals $\sigma_{\hat{a}_{i|i+1,\dots,n}}^2 = d_i^{-1}$, see (3.11).

4.3 Computation of the bounds

By means of a sequential conditional adjustment, the full ellipsoid will be searched through for candidates for the vector of ambiguities. From (4.2) we can construct the following bounds for ambiguity a_{i+1} ; the ambiguities a_n through a_{i+2} are already conditioned, the ambiguities a_i through a_1 are not conditioned yet.

$$\underbrace{\left[(a_{i+1} - \hat{a}_{i+1}) + \sum_{j=i+2}^n l_{j,i+1}(a_j - \hat{a}_j) \right]^2}_{\text{left}_{i+1}} \leq \underbrace{\frac{\chi^2}{d_{i+1}} - \frac{1}{d_{i+1}} \sum_{l=i+2}^n d_l \left[(a_l - \hat{a}_l) + \sum_{j=l+1}^n l_{jl}(a_j - \hat{a}_j) \right]^2}_{\text{right}_{i+1}} \quad (4.6)$$

Once ambiguity $i+1$ has been fixed to integer a_{i+1} , we compute the bounds for ambiguity i :

$$\underbrace{\left[(a_i - \hat{a}_i) + \sum_{j=i+1}^n l_{ji}(a_j - \hat{a}_j) \right]^2}_{\text{left}_i} \leq \underbrace{\frac{\chi^2}{d_i} - \frac{1}{d_i} \sum_{l=i+1}^n d_l \left[(a_l - \hat{a}_l) + \sum_{j=l+1}^n l_{jl}(a_j - \hat{a}_j) \right]^2}_{\text{right}_i} \quad (4.7)$$

Equations (4.6) and (4.7) hold for $i \in [1, n-1]$. To show that the bounds can be computed recursively, we split off the term $l = i+1$ from the summation on the right hand side of (4.7),

$$\begin{aligned} & \underbrace{\left[(a_i - \hat{a}_i) + \sum_{j=i+1}^n l_{ji}(a_j - \hat{a}_j) \right]^2}_{\text{left}_i} \leq \\ & \underbrace{\frac{d_{i+1}}{d_i} \left\{ \frac{\chi^2}{d_{i+1}} - \frac{1}{d_{i+1}} \sum_{l=i+2}^n d_l \left[(a_l - \hat{a}_l) + \sum_{j=l+1}^n l_{jl}(a_j - \hat{a}_j) \right]^2 \right\}}_{\text{right}_{i+1}} \\ & - \underbrace{\left[(a_{i+1} - \hat{a}_{i+1}) + \sum_{j=i+2}^n l_{j,i+1}(a_j - \hat{a}_j) \right]^2}_{\text{left}_{i+1}} \} \end{aligned} \quad (4.8)$$

Recognizing on the right hand side the terms $right_{i+1}$, see (4.6) and $left_{i+1}$, the equation can be simplified to

$$\underbrace{\left[(a_i - \hat{a}_i) + \sum_{j=i+1}^n l_{ji}(a_j - \hat{a}_j) \right]^2}_{left_i} \leq \underbrace{\frac{d_{i+1}}{d_i} (right_{i+1} - left_{i+1})}_{right_i} \quad (4.9)$$

This shows that, in the sequential conditional adjustment, the bounds for the ambiguities $n-1$ through 1 can be computed recursively. The recursion starts with the conditioning of a_n (substitution of $i=n$ in (4.2)):

$$\underbrace{(a_n - \hat{a}_n)^2}_{left_n} \leq \underbrace{\frac{\chi^2}{d_n}}_{right_n} \quad (4.10)$$

The interval with valid integers for ambiguity a_i follows now from elaboration on (4.9)

$$|(a_i - \hat{a}_i) + \sum_{j=i+1}^n l_{ji}(a_j - \hat{a}_j)| \leq \sqrt{right_i} \quad (4.11)$$

or

$$-\sqrt{right_i} \leq (a_i - \hat{a}_i) + \sum_{j=i+1}^n l_{ji}(a_j - \hat{a}_j) \leq \sqrt{right_i} \quad (4.12)$$

Equation (4.12) can be further developed into

$$-\sqrt{right_i} - \sum_{j=i+1}^n l_{ji}(a_j - \hat{a}_j) \leq a_i - \hat{a}_i \leq \sqrt{right_i} - \sum_{j=i+1}^n l_{ji}(a_j - \hat{a}_j) \quad (4.13)$$

and

$$\hat{a}_i - \sqrt{right_i} - \sum_{j=i+1}^n l_{ji}(a_j - \hat{a}_j) \leq a_i \leq \hat{a}_i + \sqrt{right_i} - \sum_{j=i+1}^n l_{ji}(a_j - \hat{a}_j) \quad (4.14)$$

This defines the interval for ambiguity i . It will be searched through in a straightforward manner from left to right, i.e. from the lower to the upper bound. Each valid integer will be tried, one at a time, and the adjustment proceeds with the next ambiguity a_{i-1} (the so-called depth-first search). If for a certain ambiguity a_l no valid integers can be found, one returns to the previous ambiguity a_{l+1} and takes the next valid integer for this ambiguity. Once an integer is encountered that satisfies interval (4.14) for ambiguity a_1 , a full candidate vector is found. The search terminates when all valid integers encountered, have been treated and one is back at the last ambiguity a_n . The full ellipsoid has been searched through.

To summarize: the sequential adjustment starts with a conditioning on a_n and ends with a conditioning on a_1 . In this way the bounds for the ambiguities a_n through a_1

are constructed in a recursive way. When the sequential adjustment is at ambiguity i , with the definition of the conditional estimate (4.4), the interval (4.12) can be rewritten into

$$-\sqrt{right_i} \leq (a_i - \hat{a}_{i|i+1,\dots,n}) \leq \sqrt{right_i} \quad (4.15)$$

which shows that the interval for a_i is centered at the conditional estimate $\hat{a}_{i|i+1,\dots,n}$. At this moment the integer nearest to the conditional estimate, $nint(\hat{a}_{i|i+1,\dots,n})$, is the most likely candidate for ambiguity i . From (4.3) it can be seen that the conditional variances play a decisive role in the bounds for the ambiguities. The smaller the conditional variance, the smaller the interval (4.14).

4.4 Computation of the norm

In the previous section the full ellipsoid was searched through and as a result we have available all grid points that are inside the ellipsoid. One of them, the one which yields the minimum for (2.6), is the integer least-squares estimate \check{a} .

The squared norm $t(a) = \|\hat{a} - a\|_{Q_{\hat{a}}}^2$ of a candidate can be computed by substitution of a into (4.1). It can also be computed from the bounds for ambiguity a_1 at the moment the candidate is encountered in the search. The squared norm, the left hand side of (4.2), can be rewritten into (by respectively adding and subtracting χ^2 and splitting off the first term)

$$\begin{aligned} t(a) &= \sum_{i=1}^n d_i \left[(a_i - \hat{a}_i) + \sum_{j=i+1}^n l_{ji}(a_j - \hat{a}_j) \right]^2 \\ &= \chi^2 - d_1 \underbrace{\left\{ \frac{\chi^2}{d_1} - \frac{1}{d_1} \sum_{i=2}^n d_i \left[(a_i - \hat{a}_i) + \sum_{j=i+1}^n l_{ji}(a_j - \hat{a}_j) \right]^2 \right\}}_{right_1} \\ &\quad - \underbrace{\left[(a_1 - \hat{a}_1) + \sum_{j=2}^n l_{j1}(a_j - \hat{a}_j) \right]^2}_{left_1} \} \end{aligned} \quad (4.16)$$

The squared norm $t(a)$ is less or equal to χ^2 , as the grid point is on or inside the ellipsoid.

If, at level $i = 1$, more than one candidate is available, the squared norm of successive candidate vectors a' can be computed very easily once the squared norm of one candidate a has been computed already. This is because $a'_i = a_i$ for $i = 2, \dots, n$ and

$a'_1 = a_1 + \nabla$ (where ∇ can be chosen to be any integer). So now the squared norm becomes

$$t(a') = t(a) + d_1 \{ 2\nabla [(\underbrace{a_1 - \hat{a}_1}_{dist_1}) + \underbrace{\sum_{j=2}^n l_{j1}(a_j - \hat{a}_j)}_{lef_1}] + \nabla^2 \} \quad (4.17)$$

4.5 The search algorithm

The search algorithm *FI71* is the realization of the sequential conditional estimation.

Algorithm *FI71*: Given the LDL^* -decomposition of $Q_{\hat{a}}^{-1}$, a suitable value for χ^2 , and the float ambiguity vector \hat{a} , ‘maxcan’ candidates inside the ellipsoidal search space with minimum distance to the vector of float ambiguities will be computed.

```

right(n + 1) =  $\chi^2$  ; left(n + 1) = 0 ; ende=false; ncan = 0
for i = 1 : n - 1
    dq(i) = D(i + 1, i + 1)/D(i, i)
end
dq(n) = 1/D(i, i)
i = n + 1
while ende=false
    i = i - 1
    lef(i) = dot(L(i + 1 : n, i), dist(i + 1 : n))
    right(i) = (right(i + 1) - left(i + 1)) · dq(i)
    reach =  $\sqrt{right(i)}$ 
     $\delta = a(i) - reach - lef(i)$ 
    dist(i) = JNT1( $\delta$ ) - a(i)
    if dist(i) > reach - lef(i)
        call BACKTS (n, i, end, dist, lef, left, ende)
    else
        end(i) = reach - lef(i) - 1
        left(i) = (dist(i) + lef(i))^2
    end
    if i = 1
        call COLLECTs (n, maxcan, D, lef, left, right, dist, end, ncan,
                           disall, cands, tmax, imax)
        call BACKTS (n, i, end, dist, lef, left, ende)
    end
end
```

If during the search at some level no (more) integers can be found we ‘back-track’ to a higher level (towards n). This function is performed by *BACKTS*.

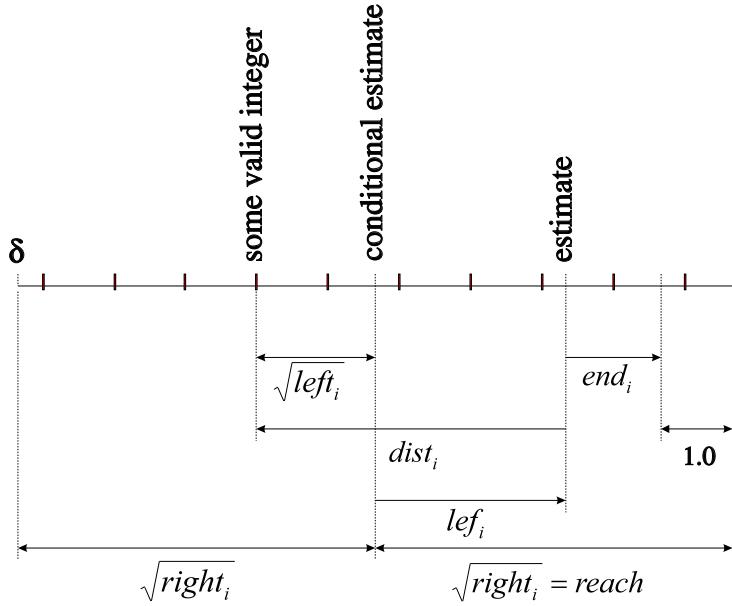


Figure 4.1: The relation between some important variables from FI71. An arrow pointing to the right indicates a positive number, an arrow pointing to the left a negative number, and a two-sided arrow indicates that the number is intrinsically positive.

Algorithm *BACKTS* ($n, i, end, dist, lef, left, ende$)

```

 $j = i + 1$ 
for  $i = j : n$ 
  if  $dist(i) \leq end(i)$ 
     $dist(i) = dist(i) + 1$ 
     $left(i) = (dist(i) + lef(i))^2$ 
    go to 10
  end
end
ende=true
10 continue
```

The performance of the algorithm could possibly be improved by recognizing that to compute $lef(i)$ it is not necessary to form every time the inner product between column i of L and vector $dist$. If we were at this level in the previous cycle, we know that only the value of the i -th ambiguity has changed; more specifically, its value has increased by 1. Consequently only $dist(i)$ has decreased by 1. This means that we can compute the new inner product from the inner product at the previous cycle by adding $L(i+1, i) \cdot 1$ to it. To incorporate this in FI71 we have to add

$iold = i$

just before the ‘while’-loop, and replace

$lef(i) = \text{dot}(L(i+1:n, i), dist(i+1:n))$

by

```

if  $iold \leq i$ 
     $lef(i) = lef(i) + L(i+1, i)$ 
else
     $lef(i) = \text{dot}(L(i+1:n, i), dist(i+1:n))$ 
end
 $iold = i$ 

```

For the function $JNT1$ ('ceil' (d) = rounding d towards $+\infty$) there are several implementations possible; use the one that works fastest on your computer.

Function $JNT1(d)$

$$JNT1(d) = \text{int}(d - \text{int}(d) - 1) + \text{int}(d) + 1$$

or

Function $JNT1(d)$

```

if  $d \leq 0$ 
     $JNT1 = \text{int}(d)$ 
else
     $JNT1 = \text{int}(d) + 1$ 
end

```

4.6 Collecting and storing the candidates

Most of the validation techniques in use nowadays, need two integer candidates, viz. the one which has minimum norm (the best), and the second best. Algorithm $COLLECTs$ takes care for the storage of the candidates.

If on level $i = 1$ a valid integer is encountered, a full candidate vector has been found. Its squared norm is computed, and the candidate and its norm are stored if, 1. the number of candidates found thus far is less than the number of candidates required (' $maxcan$ '), or 2. the squared norm is less than the maximum of the squared norms of the candidates stored thus far.

Algorithm $COLLECTs(n, maxcan, D, lef, left, right, \chi^2, dist, end, ncan, disall, cands, tmax, imax)$

$$\begin{aligned} t &= \chi^2 - (right(1) - left(1)) \cdot D(1, 1) \\ end(1) &= end(1) + 1 \end{aligned}$$

```

while  $dist(1) \leq end(1)$ 
     $n_{can} = n_{can} + 1$ 
    if  $n_{can} \leq max_{can}$ 
        call  $STOREs(n_{can}, n_{can}, imax, t, tmax, dist, cands, disall)$ 
    else
        if  $t < tmax$ 
            call  $STOREs(max_{can}, imax, imax, t, tmax, dist, cands, disall)$ 
        end
    end
     $t = t + (2 \cdot (dist(1) + lef(1)) + 1) \cdot D(1, 1)$ 
     $dist(1) = dist(1) + 1$ 
end

```

Algorithm $STOREs(ican, ipos, imax, t, tmax, dist, cands, disall)$

```

for  $i = 1 : n$ 
     $cands(i, ipos) = dist(i)$ 
end
 $disall(ipos) = t$ 
 $tmax = t$ 
 $imax = ipos$ 
for  $i = 1 : ican$ 
    if  $disall(i) > tmax$ 
         $imax = i$ 
         $tmax = disall(i)$ 
    end
end

```

When max_{can} candidates were requested and n_{can} candidates were found, then the best $\min(n_{can}, max_{can})$ candidates are stored by $STOREs$. Since the array $cands$ contains the difference between the vector of float ambiguities and the integer candidates, the integer least-squares estimate is obtained by

$$\check{a} = \hat{a} + cands(1 : n, k)$$

where k is taken such that $dist(k) \leq dist(i) \forall i = 1, \min(n_{can}, max_{can})$.

As a result of the search, the integer least-squares estimate \check{a} is obtained. $disall(k)$ contains the squared norm $t(\check{a}) = \|\hat{a} - \check{a}\|_{Q_{\hat{a}}^{-1}}^2$. Current validation procedures require, besides \check{a} , also the second best candidate \check{a}' . If requested (and present) also this second best candidate can be retrieved from the array $cands$, with $t(\check{a}') = \|\hat{a} - \check{a}'\|_{Q_{\hat{a}}^{-1}}^2$.

Note: When handling integer parameters on a computer, one must take care of their magnitude. The range for integer parameters usually is much smaller than that for real (double precision) parameters. It also is machine and/or compiler dependent.

To avoid that integer overflows occur, the magnitude of the original float ambiguities \hat{a} is reduced such that it is between -1.0 and $+1.0$. This is performed prior to the decorrelation step, and can for example be done by the modulo function:

$$\hat{a} := \text{mod}(\hat{a}, 1.0) \quad (4.18)$$

Through this *integer* translation, the center of the ellipsoid is shifted by the vector $\text{mod}(\hat{a}, 1.0)$. The integer part $\hat{a} - \text{mod}(\hat{a}, 1.0)$ is stored, and added again to the fixed ambiguity \check{a} afterwards.

4.7 Alternating search around the conditional estimate

In this and the following section, we will, as a side step, mention two implementations, that are alternatives to the straightforward search treated so far.

Instead of scanning the interval per ambiguity (4.14) from left to right for integers, one can search in an alternating way around the conditional estimate. In the conditional estimation we will start by conditioning ambiguity a_i to integer $\text{nint}(\hat{a}_{i|i+1,\dots,n})$. Next, the second nearest, third nearest, etc are tried. The candidates for a_i are chosen in a sequence, that starts with the one closest to the center of the interval (4.15) and goes alternately towards the bounds, section 2.4 in [2].

With this strategy one tries, in each step of the sequential conditioning, to stay as far as possible from the borders of the ellipsoidal region (4.1).

When all candidates contained in the ellipsoid are required, this strategy has no benefit over the straightforward search in which the candidates per interval are taken from ‘left to right’. When used in connection with the shrinking strategy, and when only one of a great number of candidates within the ellipsoid is required, the strategy of searching alternately around the conditional estimate will have benefit.

4.8 Shrinking the ellipsoidal region

The best candidate, the grid point nearest to \hat{a} , is the integer least-squares estimate for the vector of ambiguities. As we are in principle interested in only this candidate, the search can be designed to find it as quickly as possible. The sequential conditional estimation aims in the first place at constructing a complete n -dimensional vector with integers that fulfills (4.3), or in other words aims at finding a grid point that is inside the ellipsoid.

As soon as such a vector \bar{a} is found, the corresponding squared norm $\|\hat{a} - \bar{a}\|_{Q_{\hat{a}}^{-1}}^2$ is taken as a new value for χ^2 . We shrink the ellipsoidal region. The sequential conditional estimation is then continued (not started over!) in this shrunken ellipsoid, section 5 of [1]. If one, possibly after repeated shrinking, fails to find a candidate in the ellipsoid, the last found integer vector is the sought for integer least-squares estimate, $\check{a} = \bar{a}$, see section 2.4 of [2].

4.9 The volume of the ellipsoidal region

In the next three sections we will elaborate on how to control, prior to the search, the size of the ambiguity search space. We will recognize the relation between the volume of the ellipsoid and the number of candidates contained. By this relation, the straightforward search is an effective instrument in solving the minimization problem (2.6).

The volume, expressed in [cyclesⁿ], of the ellipsoidal region (4.1) is given by

$$E_n = \chi^n \sqrt{|Q_{\hat{a}}|} V_n \quad (4.19)$$

see [14]. The volume function in (4.19) is

$$V_n = \frac{2}{n} \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} \quad (4.20)$$

where Γ is the gamma function, defined as

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt \text{ for } x > 0 \quad (4.21)$$

The volume function can be computed recursively for $n \geq 3$ by

$$V_n = \frac{2\pi}{n} V_{n-2} \quad (4.22)$$

with

$$V_1 = 2 \text{ and } V_2 = \pi \quad (4.23)$$

For the determinant of the variance covariance matrix we have the following relations

$$|Q_{\hat{a}}| = \prod_{i=1}^n \lambda_i = \prod_{i=1}^n \sigma_{\hat{a}_{i|i+1}, \dots, n}^2 \quad (4.24)$$

where λ_i is the i -th eigenvalue of matrix $Q_{\hat{a}}$. The volume E_n can thus easily be computed, as the conditional variances are available from matrix D , see (3.11).

4.10 Setting χ^2 , I

The volume E_n turns out to be a good indicator for the number of candidates (grid points) contained in the ellipsoid. A volume of E_n cyclesⁿ corresponds to approximately $k = n \text{int}(E_n)$ candidates. For k less than a few, the relation is not accurate. The mismatch is caused by the discrete nature of the grid points. Centering the same ellipsoid at a different location, may result in a different number of candidates, $k \in Z$, while the volume $E_n \in R$, remains unchanged.

The value χ^2 can be taken such that a certain number of candidates will be inside the ellipsoidal region. A straightforward search can then be performed to obtain the

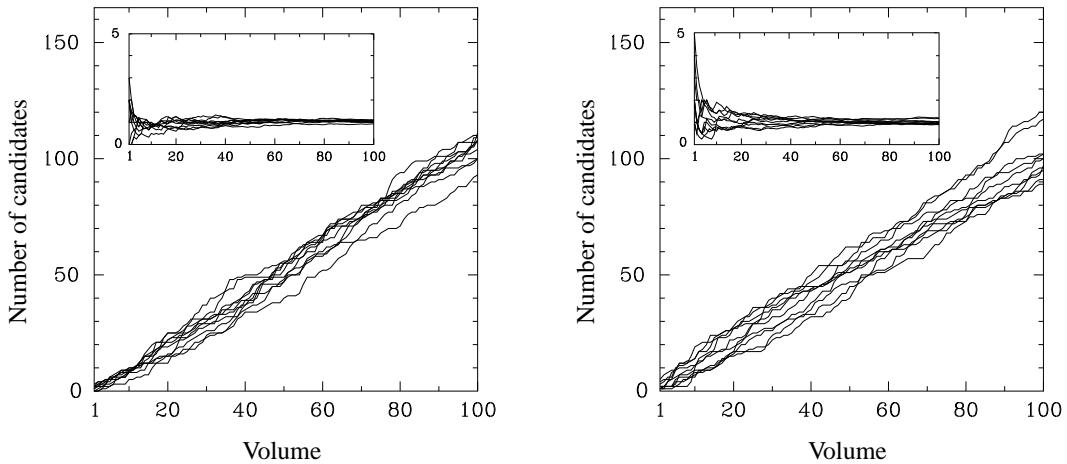


Figure 4.2: Number of candidates inside ellipsoid versus volume; left: single frequency phase data, right: dual frequency phase data

requested number of candidates. A list with the best k candidates, possibly ordered after their norms, can be set up and updated during the sequential adjustment. In this way the size of the ellipsoid is controlled prior to the search. To some extent one can already infer the quality of the integer estimator in advance.

Figure 4.2 concerns ten experiments each with two epochs (sampling interval 1 second) of single frequency data (left) and dual frequency phase data (right) to seven satellites on a 2.2 km baseline (with dual frequency this gives 12 ambiguities). The actual number of candidates contained in the ambiguity search ellipsoid is given as function of the volume of the ellipsoid [cycles¹²]. The volume ranges from 1 to 100. All lines run under 45 degrees approximately. The volume is believed to be a good indicator for the actual number of candidates in the ambiguity search ellipsoid. The inset shows the quotient of the number of candidates and the volume, i.e. the relative error of the volume as a predictor for the number of candidates.

4.11 Setting χ^2 , II

Another method to set the value for χ^2 such that at least two candidates are contained in the ellipsoidal search space, was proposed in [16], see also [19]. Since the decorrelated ambiguities have such a high precision (typically some tenths of a cycle), rounding to the nearest integer will produce a candidate with a norm close to the minimum. So setting χ^2 equal to the squared norm of this candidate will guarantee at least one and most probably not more than a few candidates. Other candidates with small norms can be found through rounding all ambiguities but one to their nearest integer, and one ambiguity to the next-nearest integer. If we have n ambiguities, this will give us n more candidates with likely small norms. Setting χ^2 to the one-but-smallest squared norm, will guarantee now at least two candidates, and most likely not more than a few.

These $n + 1$ squared norms can be computed in an efficient way using again the LDL^* -decomposition of $Q_{\hat{a}}^{-1}$. Re-arranging the first part of equation (4.16) gives

$$t(a) = \sum_{i=1}^n d_i \left[\sum_{j=i}^n l_{ji} (a_j - \hat{a}_j) \right]^2 \quad (4.25)$$

$$= \sum_{i=1}^n d_i e_i^2 \quad (4.26)$$

(Note that e_i^2 equals $left_i$.) Equation (4.17) showed an efficient way for computing the squared norm $t(a')$ for a candidate which differs ∇ in the last ambiguity a_1 from a candidate with known squared norm $t(a)$. We will now generalize this for the case that an arbitrary ambiguity k differs ∇ :

$$a_{,k} = a + c_k \nabla \quad (4.27)$$

with c_k a vector with zeros on all positions with exception of position k which has the value one:

$$c_k = (\underbrace{0, 0, \dots, 0}_{k-1}, 1, \underbrace{0, 0, \dots, 0}_{n-k})$$

The squared norm for the vector $a_{,k}$ can be written as

$$\begin{aligned} t(a_{,k}) &= \sum_{i=1}^k d_i (e_i + l_{ki} \nabla)^2 + \sum_{i=k+1}^n d_i e_i^2 \\ &= \sum_{i=1}^k d_i (e_i^2 + 2e_i l_{ki} \nabla + l_{ki}^2 \nabla^2) + \sum_{i=k+1}^n d_i e_i^2 \\ &= t(a) + \sum_{i=1}^k d_i (2e_i l_{ki} \nabla + l_{ki}^2 \nabla^2) \end{aligned} \quad (4.28)$$

Compare this result with that of equation (4.17). There we used ($dist_1 + lef_1$) instead of e_1 since these are available during the search process. The form shown here is more efficient when we have only the LDL^* -decomposition of $Q_{\hat{a}}^{-1}$, as is the case prior to the search.

We start by taking for a

$$a = (nint(\hat{a}_1), nint(\hat{a}_2), \dots, nint(\hat{a}_n)) \quad (4.29)$$

and once we have computed its squared norm $t(a)$, we can compute very efficiently another n likely small norms $t(a_{,k})$ with (4.28) by taking for ∇

$$\nabla = \begin{cases} 1 & \text{if } nint(\hat{a}_k) - \hat{a}_k < 0 \\ -1 & \text{if } nint(\hat{a}_k) - \hat{a}_k > 0 \end{cases} \quad (4.30)$$

with k from 1 to n . During the computation of the $n + 1$ norms we keep track of the smallest and one-but-smallest value. χ^2 is set to the one-but-smallest value for the norm. The ellipsoid will contain at least two candidates.

To illustrate this strategy, the following experiment was done: we computed the squared norm belonging to the vector one gets when rounding all ambiguities to their nearest integer. We also computed the n squared norms one gets when rounding all ambiguities but one to the nearest integer: the remaining one is rounded to the next-nearest integer. With χ^2 set to these norms, the volume of the ambiguity search ellipsoid was computed.

This was done for the same 10 dual frequency experiments for which the relation between volume and number of candidates was demonstrated (see figure 4.2). Per experiment the volumes were sorted after increasing order, and plotted in figure 4.3, i.e. the smallest volume is plotted at $k = 1$, the second smallest at 2, etc. At left we plotted the results for the transformed problem, at right those for the original problem. Looking at the experiments with the transformed ambiguities, one can see that the one-but-smallest volume (which guarantees at least two candidates) never exceeds the value of 10. This translates into a maximum number of candidates of approximately 10 (see figure 4.2).

For the original ambiguities the volumes are in the order $10^{10} - 10^{12}$, which shows that no acceptable volumes are obtained if the ambiguities are not decorrelated. Searching through ellipsoids with such large volumes would mean that we have to cope with an enormous number of candidates, which clearly would be detrimental for the efficiency.

For the case of the transformed ambiguities, we actually set χ^2 to the one-but-smallest squared norm. In table 4.1 one can find the epoch numbers for the 10 experiments, the resulting volume and the number of candidates inside the ellipsoid. This number is below 10 for all experiments.

The same was done for a 10.4 km baseline (measured simultaneously with the 2.2 km baseline, observing the same 7 satellites; the baselines have one station in common). Although here we find in some experiments up to 14 candidates, the procedure is still capable to come up with the two minimum candidates in an efficient way due to the almost lack of correlation between the ambiguities.

4.12 Example ambiguity search

The search in the ambiguity ellipsoid will be illustrated using the original ambiguities a of the three-dimensional example of section 3.8. Next, in order to show what is gained by the decorrelating transformation, also the search with the transformed ambiguities z will be demonstrated. This section is concluded by comparing a search with original and transformed ambiguities for a GPS baseline.

The (real valued) estimates for the ambiguities a are:

$$\begin{aligned}\hat{a}_1 &= 5.45 \\ \hat{a}_2 &= 3.10 \\ \hat{a}_3 &= 2.97\end{aligned}\tag{4.31}$$

The so-called search tree is depicted in figure 4.4 at left, and should be read from left to right. Per level, i.e. per ambiguity, the candidates encountered are represented by

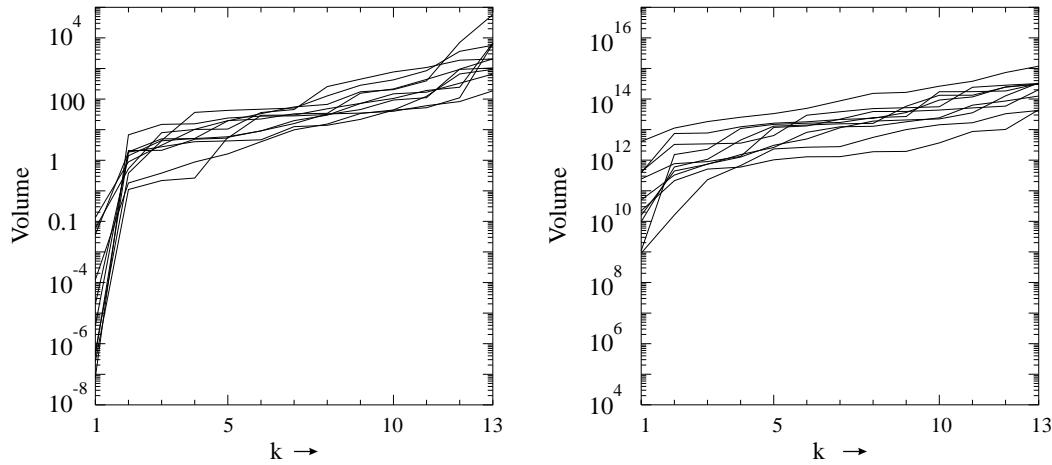


Figure 4.3: Volume of the ambiguity search space containing at least k ‘near’ candidates; left: transformed ellipsoid, right: original ellipsoid.

epochs	2.2 km		10.4 km	
	volume	number of candidates	volume	number of candidates
1 - 2	0.38	2	8.76	14
31 - 32	0.11	2	1.36	2
61 - 62	1.41	2	13.95	14
91 - 92	6.76	4	5.58	9
121 - 122	1.89	2	1.59	3
151 - 152	1.85	4	2.89	2
181 - 182	2.07	2	0.67	2
211 - 212	0.53	2	0.90	3
241 - 242	0.89	2	9.50	7
271 - 272	0.18	2	1.83	3

Table 4.1: Volume and number of candidates inside the ambiguity search ellipsoid, by setting χ^2 to the one-but-smallest squared norm

a node. A full candidate vector, i.e. a grid point inside the ellipsoid, is found when a node on level a_1 is connected with a node on level a_2 , which is connected with a node on level a_3 .

The search starts with ambiguity a_3 . The interval is centered at $\hat{a}_3 = 2.97$ and ranges from 0.462 to 5.478. Valid integers for a_3 are thus [1,2,3,4,5]. This interval is scanned from left to right. Ambiguity a_3 is conditioned on $a_3 = 1$. Equation (4.14) applied for ambiguity a_2 , yields the integers [1,2,3]. After conditioning ambiguity $a_2 = 1$, it is not possible with (4.14) applied to ambiguity a_1 , to find valid integers for ambiguity a_1 . Conditioning a_1 to an integer will result in a grid point that is outside the ambiguity search ellipsoid, no matter which integer is taken. We proceed with the second candidate for ambiguity $a_2 = 2$. For this candidate we find the integer $a_1 = 5$. As we have reached level 1, we have found a full integer vector, that satisfies (4.1), i.e. a grid point that is inside the ambiguity search ellipsoid. The vector is $(a_1 = 5, a_2 = 2, a_3 = 1)$, see also figure 4.5. The process is continued until no integers are left on any level. The search is terminated and six candidate vectors have been found. The volume of this ambiguity search ellipsoid is $E_3 = 7.3$ cycles³.

The (real valued) estimates for the ambiguities z are:

$$\begin{aligned}\hat{z}_1 &= 2.35 \\ \hat{z}_2 &= -4.57 \\ \hat{z}_3 &= 10.02\end{aligned}\tag{4.32}$$

The search tree for the transformed problem is depicted in figure 4.4 at right, the ambiguity search ellipsoid in figure 4.5 at right. In comparing the tree at left with the tree at right in figure 4.4, it can be seen that the search in the transformed problem can be performed much more efficiently. In the tree at left, there are 13 so-called ‘dead ends’, in the tree at right only 3. For a dead end, computations have to be carried out (the bounds), that do not result in a full candidate vector. It is the discontinuity in the spectrum of conditional variances that causes this so-called ‘halting’ of the search. This discontinuity is analyzed in [6] and the implications for the search are described in [4]: the first few bounds (4.14) are rather loose, the remaining ones are very tight.

The solution, the integer least-squares estimate, is $\check{a} = (5, 3, 4)$. In the transformed problem we obtain $\check{z} = (2, -5, 10)$. This solution has to be transformed by Z^{-*} and yields exactly the same solution. The squared norm is $t(\check{a}) = 0.218$.

The differences between original and transformed ambiguities are even much larger for real GPS examples. In figure 4.6 we give the number of valid integers per ambiguity, encountered during the full search. They are given for both the original and transformed ambiguities. The data is from a 2.2 km baseline with dual frequency data to seven satellites. Two epochs of data were taken, with one second in between.

When a baseline (three coordinate unknowns) is observed for a short time span, three conditional variances, of the original ambiguities, are very large and the remaining nine ones are very small. From figure 4.6 we see that at level $i = 10$ (the third ambiguity in the search) there are over 3.10^8 candidates. After having proceeded to ambiguity a_1 there are only 2 full candidate vectors left, which implies that there are very many dead ends. The volume of the ambiguity search ellipsoid is $E_{12} = 2.8$

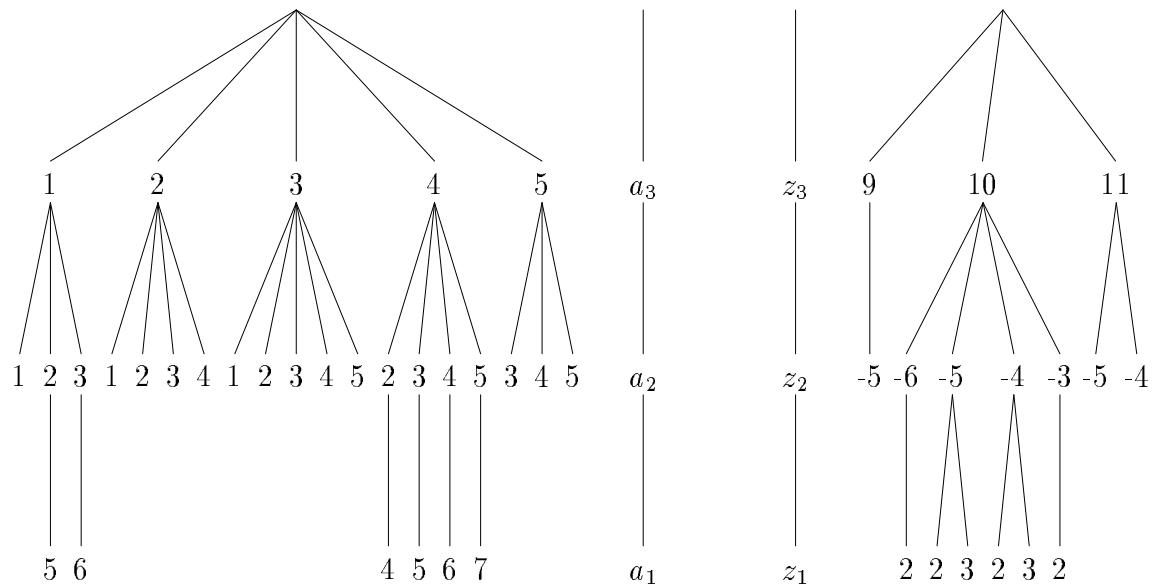


Figure 4.4: The search tree for the original problem (left) and for the transformed problem (right).

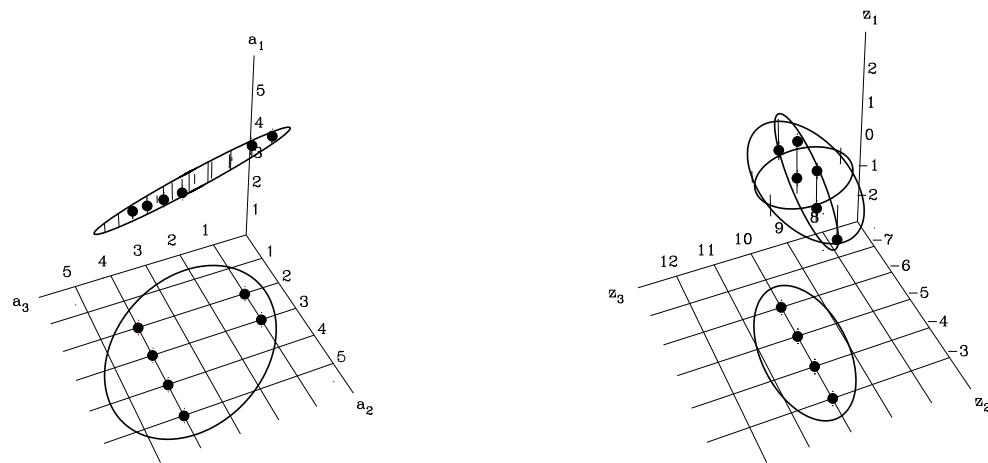


Figure 4.5: The ellipsoid in 3-D and its perpendicular projection onto the 2-3 plane for the original problem (left) and for the transformed problem (right).

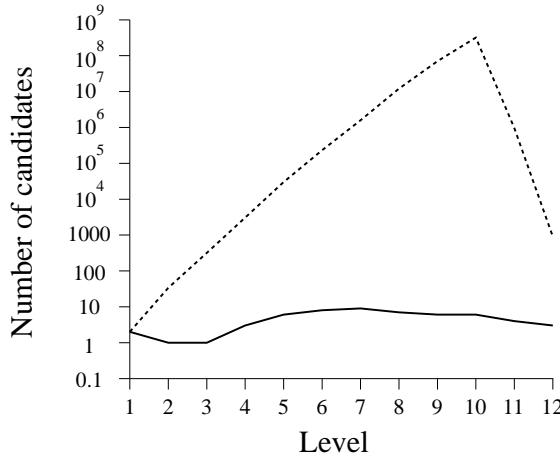


Figure 4.6: Number of candidates per level; dashed line: before transformation, solid line: after transformation.

cycles¹². By the transformation, the spectrum of conditional variances is flattened, and so is the graph of the number of candidates per level: the search on the transformed ambiguities can be performed very efficiently, there are only a few dead ends left.

4.13 The back transformation

Once the transformed integer minimization problem (3.6) has been solved, we have to back transform the integer estimate \check{z} , in order to obtain the integer least-squares estimate \check{a} . The relation reads

$$Z^* \check{a} = \check{z}$$

The inverse of the matrix Z^* does not need to be computed explicitly. Matrix Z^* has only integer elements, has full rank, with $|Z^*| = \pm 1$, and is square and usually dense, as the transformation is truly a multi-channel transformation [3]. Therefore an *LU*-factorization can be made by Gaussian elimination with partial pivoting, see chapter 3 of [7] and chapter 1 of [11] (the LINPACK routines DGEFA and DGESL). Then the integer least-squares estimate \check{a} is obtained via forward and backward substitution. Note that the intermediate result – after the forward substitution – is in general *not* an integer vector. The forward and backward substitution can be repeated for any other candidate, e.g. the second best \check{a}' .

Alternatives to the back transformation, computing matrix Z^{-*} directly or computing matrix $Q_{\check{b}\check{z}}$, are suggested in section 3.9.

4.14 Concluding remarks

The integer ambiguities are estimated in a sequential conditional least-squares adjustment. The ambiguities are constrained to integers, in a sequential manner from ambiguity a_n to a_1 . In case the search is performed on the original ambiguities a , they should be ordered according to their conditional precision [1]:

$$\sigma_{\hat{a}_{1|2,\dots,n}}^2 \geq, \dots, \geq \sigma_{\hat{a}_n}^2 \quad (4.33)$$

The sequential adjustment should start with the most precise ambiguity. The decorrelating Z -transformation typically flattens the spectrum of conditional variances: the variances are equalized. This makes that the sequential conditional adjustment can be carried out very efficiently.

5

Flow diagram of the parameter estimation

Procedure for solving the linear system (2.2)

Float solution:

- normal equations $Nx = r$ (equation 2.3)
- Cholesky factorization of normal matrix, $N = CC^*$, e.g. algorithm 4.2.2. [7]
- forward substitution $C(C^*x) = r$, algorithm 3.1.1. [7]
- backward substitution $C^*x = C^{-1}r$, algorithm 3.1.2. [7]

Integer ambiguity estimation:

- split Cholesky factor $G, G = L\sqrt{D} \rightarrow L$ and D (equation 3.9)
- inversion of $D \rightarrow D^{-1}$
- inversion of $L \rightarrow L^{-1}$, LTINV
- construction of matrix Z , transformation of $D^{-1} \rightarrow \tilde{D}^{-1}$, $L^{-1} \rightarrow \tilde{L}^{-1}$ and $\hat{a} \rightarrow \check{z}$, SRC1
- inversion of $\tilde{D}^{-1} \rightarrow \tilde{D}$
- inversion of $\tilde{L}^{-1} \rightarrow \tilde{L}$, LTINV
- integer minimization, FI71 $\rightarrow \check{z}$ (equation 3.6)
- transposition of $Z \rightarrow Z^*$
- LU -factorization of Z^* , algorithm 3.4.1. [7]
- forward substitution, algorithm 3.1.1. [7]
- backward substitution, algorithm 3.1.2. [7] $\rightarrow \check{a}$

Fixed solution:

- forward substitution, algorithm 3.1.1. [7]
- backward substitution, algorithm 3.1.2. [7] $\rightarrow \check{b}$

6

Concluding remarks

The LAMBDA method for integer estimation of the GPS double difference ambiguities consists of firstly a decorrelation of the ambiguities and secondly a sequential conditional adjustment of the ambiguities. In this report the implementation aspects have been discussed. The integer minimization problem is attacked by a discrete search over an ellipsoidal region, the ambiguity search ellipsoid.

The shape and orientation of the ellipsoid are governed by the variance covariance matrix of the ambiguities. The decorrelation realizes an ellipsoid that is very much sphere-like. It can be searched through very efficiently. The size of the ellipsoid can be controlled prior to the search using the volume function, using one of the two strategies presented. The volume gives an indication of the number of candidates contained in the ellipsoid. Making a request for a few candidates and carrying out the search on the transformed ambiguities, enables a straightforward implementation of the search. A limited number of candidates will be gathered of which one is the integer least-squares estimate for the vector of ambiguities.

For future information concerning the LAMBDA method, we refer to our WWW page:

<http://www.geo.tudelft.nl/mgp/lambda>

A

Inversion of a triangular matrix

The construction of the Z-transformation matrix is based on the variance-covariance matrix $Q_{\hat{a}}$. With the first option of (3.8) the inverse of the unit lower triangular matrix L has to be computed. The inversion of diagonal matrix D is trivial.

Algorithm *LTINV*: Given the full rank lower triangular matrix L , this algorithm computes the inverse $X = L^{-1}$. The relation between matrix L and its inverse X reads $LX = I$. For element (i, j) of I we have

$$\sum_{k=j}^i L(i, k)X(k, j) = \begin{cases} 0 & i > j \\ 1 & i = j \end{cases}$$

From this follows that for $i > j$

$$X(i, j) = \frac{\sum_{k=j}^{i-1} L(i, k)X(k, j)}{L(i, i)}$$

and for $i = j$

$$X(i, i) = \frac{1}{L(i, i)}$$

The algorithm reads

```

for  $i = 1 : n$ 
  for  $j = 1 : i - 1$ 
     $X(i, j) = -dot(L(i, j : i - 1), X(j : i - 1, j)) / L(i, i)$ 
  end
   $X(i, i) = 1 / L(i, i)$ 
end

```

The inverse is computed row wise and in-place; the inverse may overwrite matrix L . Function *dot* computes the dot or inner product of two vectors.

References

- [1] Teunissen, P.J.G. (1993). *Least-squares estimation of the integer GPS ambiguities*. Invited lecture, Section IV Theory and Methodology, IAG General Meeting, Beijing, China, August, also in *Delft Geodetic Computing Centre LGR series*, No. 6, 16 pp.
- [2] Teunissen, P.J.G. (1995). The least-squares ambiguity decorrelation adjustment: A method for fast GPS integer ambiguity estimation. *Journal of Geodesy*, Vol. 70, No. 1-2, pp. 65-82.
- [3] Teunissen, P.J.G. (1995). The invertible GPS ambiguity transformations. *Manuscripta Geodaetica*. Vol. 20, No. 6, pp. 489-497.
- [4] Teunissen, P.J.G. (1995). On the GPS double difference ambiguities and their partial search spaces. *Proceedings IAG Symposium No. 114, Geodetic Theory Today*, Springer Verlag, pp. 39-48, III Hotine-Marussi symposium on Mathematical Geodesy, L'Aquila, Italy, May 29-June 3, 1994
- [5] Teunissen, P.J.G., and C.C.J.M. Tiberius (1994). Integer least-squares estimation of the GPS phase ambiguities. *Proceedings of International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation KIS'94*, Banff, Canada, August 30-September 2, pp. 221-231.
- [6] Teunissen, P.J.G., P.J. de Jonge, and C.C.J.M. Tiberius (1994). On the spectrum of the GPS DD-Ambiguities. *Proceedings of ION GPS-94, 7th International Technical Meeting of the Satellite Division of the Institute of Navigation*, Salt Lake City, UT, September 20-23, pp. 115-124.
- [7] Golub, G.H., and C.F. van Loan (1989). *Matrix computations*. Second edition. The Johns Hopkins University Press, Baltimore, Maryland, USA.
- [8] Teunissen, P.J.G. (1994). A new method for fast carrier phase ambiguity estimation. *Proceedings IEEE Position, Location and Navigation Symposium PLANS'94*. Las Vegas, NV, April 11-15, pp. 562-573.
- [9] Jonge de, P.J., and C.C.J.M. Tiberius (1994). A new GPS ambiguity estimation method based on integer least squares. *Proceedings of the 3rd International Symposium on Differential Satellite Navigation Systems DSNS'94*. London, UK, April 18-22, Paper No. 73, 9 pp.

- [10] Tiberius, C.C.J.M., and P.J. de Jonge (1995). Fast positioning using the LAMBDA-method. *Proceedings of the 4th International Symposium on Differential Satellite Navigation Systems DSNS'95*. Bergen, Norway. April 24-28, Paper No. 30, 8 pp.
- [11] Dongarra, J.J., J.R. Bunch, C.B. Moler, and G.W. Stewart (1979). *LINPACK Users' Guide*. SIAM Publications, Philadelphia, USA.
- [12] Teunissen, P.J.G., P.J. de Jonge, and C.C.J.M. Tiberius (1995). A new way to fix carrier-phase ambiguities. *GPS World*, Vol. 6, No. 4, April 1995, pp. 58-61.
- [13] Teunissen, P.J.G. (1996). GPS carrier phase ambiguity fixing concepts. Chapter 8 of Lecture Notes in Earth Sciences, No. 60, *GPS for Geodesy*, International School, Delft, The Netherlands, March 26-April 1, 1995, A. Kleusberg and P.J.G. Teunissen, (Eds.) Springer Verlag.
- [14] Apostol, T.M. (1969). *Multi-variable calculus and linear algebra, with applications to differential equations and probability*. Calculus Vol. 2, Second edition, Wiley, New York.
- [15] Jonge de, P.J., and C.C.J.M. Tiberius (1996). Integer ambiguity estimation with the LAMBDA method. *Proceedings IAG Symposium No. 115, GPS trends in precise terrestrial, airborne and spaceborne applications*, XXI General Assembly of IUGG, July 2-14, Boulder, CO, July 2-14, 1995, G. Beutler et.al., (Eds.), Springer Verlag, pp. 280-284.
- [16] Teunissen, P.J.G., P.J. de Jonge, and C.C.J.M Tiberius (1995). The least-squares ambiguity decorrelation adjustment: its performance on short GPS baselines, submitted to *Journal of Geodesy*.
- [17] Marel, van der, H. (1989). Algorithmic aspects of solving large systems of equations (1989). *Proceedings of ISPRS intercommission working group III/IV, Tutorial on mathematical aspects of data analysis*, Pisa, Italy, June 1-2, pp. 49-103.
- [18] Jonge de P.J., C.C.J.M. Tiberius, and P.J.G. Teunissen (1996). Computational aspects of the LAMBDA method for GPS ambiguity resolution. Poster to be presented at *ION GPS-96, 9th International Technical Meeting of the Satellite Division of the Institute of Navigation*, Kansas City, Missouri, Sept. 17-20.
- [19] Teunissen, P.J.G., P.J. de Jonge, and C.C.J.M. Tiberius (1996). The volume of the GPS ambiguity search space and its relevance for integer ambiguity resolution. Paper to be presented at *ION GPS-96, 9th International Technical Meeting of the Satellite Division of the Institute of Navigation*, Kansas City, Missouri, Sept. 17-20.

LAMBDA bibliography

- Jonge de, P.J., and C.C.J.M. Tiberius (1994). A new GPS ambiguity estimation method based on integer least squares. *Proceedings of the 3rd International Symposium on Differential Satellite Navigation Systems DSNS'94*, London, UK, April 18-22, Paper No. 73, 9 pp.
- Jonge de, P.J., and C.C.J.M. Tiberius (1996). Integer ambiguity estimation with the LAMBDA method. *Proceedings IAG Symposium No. 115, GPS trends in precise terrestrial, airborne and spaceborne applications*, XXI General Assembly of IUGG, July 2-14, Boulder, CO, July 2-14, 1995, G. Beutler et.al., (Eds.), Springer Verlag, pp. 280-284.
- Jonge de, P.J., and C.C.J.M. Tiberius (1996). The LAMBDA method for integer ambiguity estimation: implementation aspects. *Delft Geodetic Computing Centre LGR series*, No. 12, (this report)
- Jonge de P.J., C.C.J.M. Tiberius, and P.J.G. Teunissen (1996). Computational aspects of the LAMBDA method for GPS ambiguity resolution. Paper to be presented at *ION GPS-96, 9th International Technical Meeting of the Satellite Division of the Institute of Navigation*, Kansas City, Missouri, Sept. 17-20.
- Teunissen, P.J.G. (1993). *Least-squares estimation of the integer GPS ambiguities*. Invited lecture, Section IV Theory and Methodology, IAG General Meeting, Beijing, China, August, also in *Delft Geodetic Computing Centre LGR series*, No. 6, 16 pp.
- Teunissen, P.J.G. (1994). A new method for fast carrier phase ambiguity estimation. *Proceedings IEEE Position, Location and Navigation Symposium PLANS'94*, Las Vegas, NV, April 11-15, pp. 562-573.
- Teunissen, P.J.G. (1995). On the GPS double difference ambiguities and their partial search spaces. *Proceedings IAG Symposium No. 114, Geodetic Theory Today*, Springer Verlag, pp. 39-48, III Hotine-Marussi symposium on Mathematical Geodesy, L'Aquila, Italy, May 29-June 3, 1994.
- Teunissen, P.J.G. (1995). The invertible GPS ambiguity transformations. *Manuscripta Geodaetica*, Vol. 20, No. 6, pp. 489-497.
- Teunissen, P.J.G. (1995). The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. *Journal of Geodesy*, Vol. 70, No. 1-2, pp. 65-82.

- Teunissen, P.J.G. (1996). GPS carrier phase ambiguity fixing concepts. Chapter 8 of Lecture Notes in Earth Sciences, No. 60, *GPS for Geodesy*, International School, Delft, The Netherlands, March 26-April 1, 1995, A. Kleusberg and P.J.G. Teunissen, (Eds.) Springer Verlag.
- Teunissen, P.J.G. (1996). Size and shape of L1/L2 ambiguity search space. *Proceedings IAG Symposium No. 115, GPS trends in precise terrestrial, airborne and spaceborne applications*, XXI General Assembly of IUGG, Boulder, CO, July 2-14, 1995, G. Beutler et.al., (Eds.), Springer Verlag, pp. 275-279.
- Teunissen, P.J.G. (1996). Rank defect integer least-squares with applications to GPS. *Bulletino di Geodesia e Scienze Affini*, Vol. 55, No. 3, pp. 225-238.
- Teunissen, P.J.G. (1996). An analytical study of ambiguity decorrelation using dual frequency code and carrier phase. *Journal of Geodesy*, Vol. 70, No. 8, pp. 515-528.
- Teunissen, P.J.G. (1996). On the geometry of the ambiguity search space with and without ionosphere. *Zeitschrift für Vermessungswesen*, Vol. 121, No. 7, pp. 332-341.
- Teunissen, P.J.G., P.J. de Jonge, and C.C.J.M. Tiberius (1994). On the spectrum of the GPS DD-ambiguities. *Proceedings of ION GPS-94, 7th International Technical Meeting of the Satellite Division of the Institute of Navigation*, Salt Lake City, UT, September 20-23, pp. 115-124.
- Teunissen, P.J.G., P.J. de Jonge, and C.C.J.M. Tiberius (1995). A new way to fix carrier-phase ambiguities. *GPS World*, Vol. 6, No. 4, pp. 58-61.
- Teunissen, P.J.G., P.J. de Jonge, and C.C.J.M. Tiberius (1995). The LAMBDA method for fast GPS surveying. *Proceedings of International Symposium 'GPS technology applications'*, Bucharest, Rumania, September 26-29, pp. 203-210.
- Teunissen, P.J.G., P.J. de Jonge, and C.C.J.M. Tiberius (1995). The least-squares ambiguity decorrelation adjustment: its performance on short GPS baselines, submitted to *Journal of Geodesy*.
- Teunissen, P.J.G., P.J. de Jonge, and C.C.J.M. Tiberius (1996). The volume of the GPS ambiguity search space and its relevance for integer ambiguity resolution. Paper to be presented at *ION GPS-96, 9th International Technical Meeting of the Satellite Division of the Institute of Navigation*, Kansas City, Missouri, Sept. 17-20.
- Teunissen, P.J.G., and C.C.J.M. Tiberius (1994). Integer least-squares estimation of the GPS phase ambiguities. *Proceedings of International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation KIS'94*, Banff, Canada, August 30-September 2, pp. 221-231.

Tiberius, C.C.J.M., and P.J. de Jonge (1995). Fast positioning using the LAMBDA-method. *Proceedings of the 4th International Symposium on Differential Satellite Navigation Systems DSNS'95*, Bergen, Norway, April 24-28, Paper 30, 8 pp.

Tiberius, C.C.J.M., and P.J. de Jonge (1995). Introduction to GPS Surveying (Part 4). *GIM International Journal for Geomatics*, Vol. 9, No. 10, pp. 61-67.

Least-Squares Estimation of the Integer GPS Ambiguities

P.J.G. Teunissen

Delft Geodetic Computing Centre (LGR)
Department of the Geodetic Engineering
Delft University of Technology
Thijsseweg 11, 2629 JA DELFT
The Netherlands

Abstract

The Global Positioning System (GPS) double-difference carrier-phase data are biased by an integer number of cycles. In this contribution a new method is introduced that enables very fast integer least-squares estimation of the ambiguities. The method makes use of an ambiguity transformation that allows one to reformulate the original ambiguity estimation problem as a new problem that is much easier to solve. The transformation aims at decorrelating the least-squares ambiguities and is based on an integer approximation of the conditional least-squares transformation. And through a flattening of the typical discontinuity in the GPS-spectrum of conditional variances of the ambiguities, the transformation returns new ambiguities that show a dramatic improvement in precision in comparison with the original double-difference ambiguities.

Contents

1	Introduction
2	Least-Squares
3	Integer Least-Squares
4	Ellipsoidal Planes of Support
5	Conditional Least-Squares
6	The Integer GPS Ambiguity Transformation
6.1	The idea of reparametrization
6.2	The admissible ambiguity transformations
6.3	On the choice of reparametrization in 2D
6.4	Bounding the triangular factor
6.5	Flattening the spectrum of conditional variances
7	Summary and concluding remarks
8	References

This contribution is based on the invited lecture given by the author, for Section IV "Theory and Methodology", at the General Meeting of the International Association of Geodesy, Beijing, China, August 1993.

1 Introduction

Nonlinear optimization, nonlinear least-squares and densities of nonlinear estimators are a trilogy of problems that are intimately related in the framework of estimation or adjustment of geodetic data. The description of physical phenomena generally proceeds through models in which a mapping, A , is defined, from a set of parameters, N , to a set of experimental outcomes, M . M is supposed to contain the image of the map A . The usual assumptions are that the map A is sufficiently smooth and that both the spaces M and N are continuous subspaces of R^n and R^m respectively. And indeed many of our geodetic estimation problems can be described as such. But not all! In particular it may happen that we know a priori that some of the parameters of interest have a discrete-like nature and are therefore not allowed to range through the whole of R . If this happens to be the case, standard gradient-like algorithms, such as for instance the Gauss-Newton method as used for solving smooth nonlinear least-squares problems fail to hold. Therefore alternative methods need to be devised for solving estimation problems in which N is of a discrete-like nature.

An important estimation problem where the parameters fail to range through the whole of R occurs in the field of GPS phase-data processing. The GPS double-difference phase ambiguities are known to be integer valued. And it is this a priori information that one would like to incorporate in the estimation procedure so as to improve the precision of the result. This is a non-trivial problem if one aims at numerical efficiency. And indeed, this topic has been a rich source of GPS-research over the last decade or so, see e.g. [2-8], [10], [17]. Starting from rather simple but time-consuming integer rounding schemes, the methods have evolved into complex and very efficient search algorithms. Nevertheless, at present times, it is still opportune to seek ways for improving the efficiency of the various search methods. This is in particular true for the real-time applications of GPS. But, to a certain extent, this is also true for some typical static applications of GPS. Because, if we are really able to significantly reduce the computational effort for estimating the integer ambiguities, it may also become worthwhile to tackle problems that have dimensions higher than the one's considered so far. And in particular this could open the way of treating all ambiguities of the various baselines in a GPS-network simultaneously.

The present study was initiated by the desire to obtain a better understanding of existing ambiguity search algorithms and through this, to obtain a better grip on the nature of the intrinsic difficulties that are associated with the problem of GPS-ambiguity fixing. And in particular, improve the computational speed of integer estimation. This would then, hopefully, also lead to ways of improving existing methods of GPS-ambiguity fixing. Based on our numerical experiences sofar, we believe to have succeeded in formulating such an improved method of GPS-ambiguity fixing. Our proposal, which in part is based on existing ideas, is presented in section 5 and section 6, and is summarized in section 7.

Our proposal is based on a one-to-one transformation from the original double-difference ambiguities to a new set of integer ambiguities. And the essence of the method is that this reparametrization enables one to obtain new ambiguities which have a smaller variance and that are less correlated. The idea of transforming the double-difference ambiguities is of course not completely new. It can be recognized in the transformation from the L_1 - and L_2 - ambiguities to the wide-lane ambiguity. This transformation, however, is not one-to-one. Also this transformation is enacted at the level of a single channel and therefore does not take into account the presence of the receiver-satellite geometry. The idea of transforming the double-difference ambiguities can also be recognized to some extent in the customary practice of choosing that reference satellite which has a favourable influence on the precision of the double-difference ambiguities. In fact, as will be shown in the sequel, the one-to-one transformation from one set of double-difference ambiguities to another set having a different satellite as reference, belongs to the class of admissible transformations that will be considered in the sequel.

In order to properly judge the significance of the present contribution, it is important that one distinguishes between the following two problems of GPS-ambiguity fixing. First one has the problem of finding ways, preferably efficient ways, for fixing the ambiguities. This in fact is the *estimation* problem. Secondly, one has the problem of validating the fixed values of the ambiguities. This is the *testing* problem. And for a proper evaluation of the validity of the fixed ambiguities one will need the probability densities of the corresponding estimators; reference is made to the discussions in e.g. [2] and [13]. Although the procedures for validating the fixed ambiguities which are currently in use in practice, appear to work satisfactory, it is the author's opinion that a sound theoretical basis for these validation

procedure is still lacking. As a case in point, consider the customary practice of ambiguity validation. Usually the fixed ambiguities are validated by testing the ratio of the quadratic forms of residuals belonging to the most-likely and second most-likely integer candidates. This ratio is then tested against a critical value computed from an F-distribution. But this is incorrect, because of the stochastic dependency that exists between the two quadratic forms. Despite the importance of proper validation procedures, the present contribution only addresses the first problem and not the second. Hence, we will only be concerned with the problem of finding a numerically efficient way for estimating the integer ambiguities. There is therefore no harm in stressing, if the data is contaminated with unmodelled effects, that our method, efficient as it may be, might still come up with the wrong ambiguities.

Although the material of this contribution is intended for solving GPS-ambiguity fixing problems, we have tried, for those who are not too familiar with the typical intricacies of GPS, to refrain as much as possible from the use of standard GPS-terminology. The presentation is therefore cast in the framework of adjustment-theory and the problem of GPS-ambiguity fixing based on the least-squares criterium is formulated as an integer least-squares problem. Also detailed derivations of results are avoided in the paper. They will be taken up in a future presentation [15]. Instead we keep to the basic ideas involved and try to motivate the main results by appealing to intuition and stochastic or geometric interpretations.

This paper is organized as follows. In section 2 we briefly review the standard linear and nonlinear least-squares problem, and emphasize the (differential) geometric interpretation that can be given to least-squares problems [14].

In section 3 we introduce and define the integer least-squares problem. Again the geometric interpretation is emphasized. For reason of simplicity we assume the map $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ to be linear. This also allows an easier access to the more intricate details of the integer least-squares problem. It is shown how an integer least-squares problem can be decomposed into parts. This is shown both for the hybrid as well as for the non-hybrid case. The decomposition is based on the theorem of Pythagoras and allows one to solve the problem in two steps. The first step is rather straightforward and amounts to an ordinary least-squares adjustment. The second step comprises the minimization of a non-homogeneous quadratic form over

the set of integers. And it is with this second step that the intricacy of the problem manifests itself. The stepped-wise approach agrees with the approach that is usually taken in case of GPS-ambiguity fixing. Since the minimizer of the original integer least-squares problem is shown to be identical to the integer minimizer of the quadratic form, the remaining part of the sequel will focus on finding ways of solving for this integer minimizer.

The sections 4, 5 and 6 are closely related. In each one of these three sections a concept for solving integer least-squares problems is presented. The concepts of section 4 and section 5 have already been in use, in one form or another, for fixing GPS-ambiguities. The reason for including these two concepts is not only because they are of importance in their own right and that they reveal clearly the intricate nature of integer least-squares problems, but also because they pave the way for a proper discussion of the material of section 6.

The first concept is reviewed in section 4. It is based on the idea that an ellipsoidal region can be described by using the infinite set of ellipsoidal planes of support. This approach is very similar to the use of simultaneous confidence intervals in statistics for multiple comparisons [11]. Within the context of GPS-ambiguity fixing the method of Frei is based on it [6]. In this section it is shown how a finite subset of the infinite ellipsoidal planes of support can be used for selecting integer candidates from which then the sought for integer minimizer is chosen.

The second concept, which is based on the completion of squares of a quadratic form, is reviewed in section 5. This concept makes use of a triangular decomposition of the positive-definite matrix that describes the ellipsoidal region. This allows one then to come up with bounds for the integer candidates that are sharper in general than the bounds derived in the previous section. It is shown how these bounds can be employed for the formulation of an efficient search algorithm. Within the context of GPS-ambiguity fixing, examples of approaches that, in one form or another, make use of a triangular decomposition, are [2], [5] and [17]. The method of Wübbena [17], resembles the approach of the present section most. The method of Euler/Landau, [5], however, still uses an a priori chosen rectangular search window. That is, the triangular decomposition is not used for selecting integer candidates, but instead, it is used after a candidate set has been selected, for efficiently eliminating candidates.

Our proposal for efficiently solving the estimation problem of GPS-ambiguity fixing is presented in section 6. First the idea of reparametrization is introduced. It is argued that the integer least-squares problem becomes easier to solve if the reparametrization can achieve a scaling and rotation of the ellipsoidal region that will result in a region which has its principal axes (almost) parallel to the grid axes. The objective is thus, to decorrelate the least-squares ambiguities and to diagonalize their variance-covariance matrix. For GPS, this is in particular of relevance when only short timespan carrier-phase data is used. The ambiguities will then be extremely correlated, their confidence ellipsoid will be extremely elongated, and the spectrum of conditional variances of the ambiguities will then show a large discontinuity. In fact, it is the discontinuity in the spectrum of conditional variances, that prohibits an efficient search for the integer least-squares ambiguities. Although the aim is to decorrelate the ambiguities, true diagonality of the variance-covariance matrix will be difficult to reach. This is due to the fact that only a particular class of ambiguity transformations is admissible. They need to be integer and volume-preserving. Based on an integer approximation of the conditional least-squares transformation, a decorrelating two-dimensional transformation, which is both integer and volume-preserving, is introduced in subsection 6.3. It returns ambiguities with an improved precision and guarantees that the correlation coefficient stays sufficiently bounded. To tackle the n -dimensional problem, the two-dimensional transformation is repeatedly applied to pairs of conditional least-squares estimates of the ambiguities. This approach has been motivated by the presence of the typical discontinuity in the GPS spectrum of ambiguity conditional variances and is based on ideas from [9]. The success of our method is largely due to the presence of the discontinuity. And this discontinuity in the GPS-spectrum of ambiguity conditional variances also stipulates the relevance of satellite-redundancy and the use of dual-frequency data. By removing the discontinuity with the ambiguity transformation, the spectrum is flattened and lowered, and transformed ambiguities are obtained that show a dramatic improvement in precision. As a result the search for the transformed integer least-squares ambiguities can be performed in a highly efficient manner.

2 Least squares

The problem of least-squares can be formulated as the minimization problem:

$$(1) \quad \min_x \|y - A(x)\|^2,$$

where: $A: R^n \rightarrow R^m$ ($m \geq n$); $\|\cdot\|^2 = (\cdot)^T Q_y^{-1}(\cdot)$ and Q_y is positive-definite. For varying values of $x \in R^n$, $A(x)$ traces (locally) an n -dimensional surface or manifold embedded in R^m . With the metric Q_y^{-1} of R^m , the scalar $\|y - A(x)\|$ therefore equals the distance from the datapoint y to the point $A(x)$ on the manifold. Hence, the problem of (1) corresponds to the problem of finding that point on the manifold, say $\hat{y} = A(\hat{x})$, which has least distance to y .

There are two conditions that $\hat{y} = A(\hat{x})$ needs to satisfy in order for it to be a (local) solution of (1). The first condition states that the least-squares residual vector $\hat{e} = y - \hat{y}$ should be orthogonal to the tangentspace of the manifold at the solution \hat{y} . And the second condition states that the datapoint y should lie within a hypersphere with centre \hat{y} and a radius equal to the largest principal normal curvature corresponding with the normal direction of the least-squares residual vector. Both these conditions are necessary and sufficient. And both these conditions are intrinsic in the sense that they are invariant to a change of variables or a reparametrization.

If the map $A(\cdot)$ is nonlinear (which happens to be the case in the majority of geodetic applications), the corresponding manifold traced by $A(x)$ is curved, and then generally no direct methods exist for solving (1) (there are exceptions). In this case one has to fall back on using iterative solution techniques. These iteration methods, such as the Gauss-Newton method, are usually based on repetitively solving linear or linearized least-squares problems.

If the map $A(\cdot)$ is linear, the corresponding manifold traced by $A(x)$ is flat. In this linear case the absence of curvature allows one to solve the minimization problem (1) explicitly. The corresponding linear least-squares estimates are given by the well-known formulae:

$$(2) \quad \begin{aligned} \hat{y} &= P_A y, \quad \hat{e} = P_A^\perp y, \\ \hat{x} &= A^{-1} P_A y, \quad \|\hat{e}\|^2 = \|P_A^\perp y\|^2, \end{aligned}$$

where: matrix P_A is the orthogonal projector that projects onto the range space of A and along its orthogonal complement; $P_A^\perp = I - P_A$; and A^{-1} is an (arbitrary) inverse of A . The estimates \hat{y} , \hat{e} and $\|\hat{e}\|^2$ are all unique, and the estimate \hat{x} is unique if and only if the linear map A has full rank n .

In the remaining of the sequel we will assume the map A to

be linear. This does of course restrict the focus of our discussion somewhat. But it is of importance to first understand the linear case before the nonlinear case can be tackled.

3 Integer least squares

A least-squares problem is said to be of the integer-type if the parameters are constrained to integer values. The problem of integer least-squares can therefore be formulated as the minimization problem:

$$(3) \quad \min_x \|y - Ax\|^2, \quad x \in Z^n.$$

Compare with (1) and note that R^n has been replaced by Z^n . In order to describe the integer least-squares problem geometrically, consider the standard grid of coordinate lines in R^n . This standard grid is mapped by A into a new, but non-standard oblique grid in the data space R^m . This new grid is superimposed on the flat manifold spanned by the columnvectors of A . The set of gridpoints of this grid equals the set that follows when Z^n is mapped under A . Hence, it is from this mapped set of gridpoints in the manifold that points should be chosen as possible candidates for solving (3). In fact, the point of this mapped set of gridpoints that has the least distance to the given datapoint $y \in R^m$, minimizes the constrained objective function of (3). In the linear case there are at most 2^n such points; in the nonlinear case however, there may be even more. Once such a point has been found, the corresponding parameter vector may be obtained through an inverse mapping of A . And this solution is unique if A has full rank n . Note however that in some cases the solution may still be unique even if A is not of full rank. This happens if the set $\{x \in R^n | x - z = u, Au = 0\}$, where z is an integer minimizer of (3), has only one point in common with Z^n , namely z . In the following we will assume A to be of full rank.

The integer least-squares problem (3) may be decomposed in two parts. In order to show this we apply Pythagoras to get the following decomposition of the objective function of (3):

$$(4) \quad \|y - Ax\|^2 = \|P_A y - Ax\|^2 + \|P_A^\perp y\|^2.$$

The second term on the right-hand side equals the squared-norm of the least-squares residual vector \hat{e} . Since this term is independent of the parameters, the minimizer of (3) is identical to the minimizer of

$$(5) \quad \min_x \|P_A y - Ax\|^2, \quad x \in Z^n.$$

Hence, the two problems (3) and (5) are equivalent in the sense that the parameter vector that minimizes (3) also minimizes (5), and vice versa. From this follows that the solution of (3) can be obtained in two steps. In the first step an ordinary least-squares estimation is performed. This amounts to replacing y in (3) by its least-squares estimate $P_A y$. This then gives (5), which is solved in the second step. Once the minimizer of (5) has been found, the minimum value of the original objective function is obtained by adding the squared-norm of the least-squares residual vector to the minimum value of the objective function of (5).

In the above discussed integer least-squares problem the complete parameter vector x was constrained to lie in Z^n . It may also happen however that only some, but not all, of the parameters are constrained to integer values. In that case we are in a *hybrid* situation where, with $x = (x_1, x_2)$, we have $x_1 \in R^{n_1}$ and $x_2 \in Z^{n_2}$. The integer least-squares problem (3) is then replaced by

$$(6) \quad \min_{x_1, x_2} \|y - A_1 x_1 - A_2 x_2\|^2, \quad x_1 \in R^{n_1}, x_2 \in Z^{n_2}.$$

But also this integer least-squares problem can be decomposed into parts. First we introduce a reparametrization through the one-to-one transformation

$$(7) \quad \begin{bmatrix} \bar{x}_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} I_{n_1} & (A_1^* Q_y^{-1} A_1)^{-1} A_1^* Q_y^{-1} A_2 \\ O & I_{n_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

With this reparametrization we may now replace (6) by

$$(8) \quad \min_{\bar{x}_1, x_2} \|y - A_1 \bar{x}_1 - \bar{A}_2 x_2\|^2, \quad \bar{x}_1 \in R^{n_1}, x_2 \in Z^{n_2},$$

where $\bar{A}_2 = P_{A_1^\perp} A_2$. Analogous to (4), we can then decompose the objective function of (8) as

$$(9) \quad \|y - A_1 \bar{x}_1 - \bar{A}_2 x_2\|^2 = \|P_A(y - A_1 \bar{x}_1 - \bar{A}_2 x_2)\|^2 + \|P_A^\perp y\|^2.$$

Since the second term on the right-hand side is independent of the parameters we only need to consider the first term for the minimization. With $P_A = P_{A_1^\perp} P_{\bar{A}_2}$, this first term can be further decomposed as

$$(10) \quad \|P_A(y - A_1 \bar{x}_1 - \bar{A}_2 x_2)\|^2 = \|P_{A_1}(y - A_1 \bar{x}_1)\|^2 + \|P_{\bar{A}_2}(y - \bar{A}_2 x_2)\|^2.$$

Hence, it follows from (9) and (10), that (8) may be written as

$$(11) \quad \begin{aligned} & \min_{\bar{x}_1 \in R^{n_1}, x_2 \in Z^{n_2}} \|y - A_1 \bar{x}_1 - \bar{A}_2 x_2\|^2 = \|P_A^\perp y\|^2 + \\ & + \min_{\bar{x}_1 \in R^{n_1}} \|P_{A_1}(y - A_1 \bar{x}_1)\|^2 + \min_{x_2 \in Z^{n_2}} \|P_{\bar{A}_2}(y - \bar{A}_2 x_2)\|^2. \end{aligned}$$

From this decomposition follows then how one can proceed to obtain the x_1 - and x_2 -minimizers of (6). First the x_2 -minimizer is obtained from solving the integer least-squares problem

$$(12) \quad \min_{x_2 \in Z^{n_2}} \|P_{\bar{A}_2}(y - \bar{A}_2 x_2)\|^2.$$

Hence, in the hybrid case, (12) takes over the role of (5). Once the x_2 -minimizer is known, the x_1 -minimizer can be computed as follows.

Note that since the second term on the right-hand side of (11) equals zero, the corresponding minimizer is given as $\hat{x}_1 = (A_1^* Q_y^{-1} A_1)^{-1} A_1^* Q_y^{-1} y$. This estimate together with the x_2 -minimizer allows one then to compute the x_1 -minimizer through the use of transformation (7). In the remaining of the sequel we will assume for reasons of simplicity to have a non-hybrid integer least-squares problem. Hence, we will consider (3) instead of (6).

Decomposition (4) implies that as far as the minimizer of (3) is concerned we may as well start from the minimization problem (5). If \hat{x} is the least-squares estimate of x , then $P_A y = A \hat{x}$ and (5) may be written as

$$(13) \quad \min_x \|\hat{x} - x\|_{Q_x}^2, \quad x \in Z^n,$$

where $\|\cdot\|_{Q_x}^2 = (\cdot)^* Q_x^{-1} (\cdot)$ and $Q_x^{-1} = A^* Q_y^{-1} A$.

As was pointed out already in the previous section, the minimization problem (13) may not have a unique solution. Although it is very unlikely that (13) has more than one solution, it is possible in principle that (13) has up to 2^n different minimizers. Still however, we will assume in the present sequel that (13) has one and only one solution. Our motivation for this assumption stems from the way the integer least-squares problem is applied in the context of GPS-ambiguity fixing. If x stands for the vector of double-difference ambiguities, a non-unique solution for (13) will namely imply that a reliable fixing of the ambiguities is not feasible. In the remaining of the sequel we will focus on solving (13).

Unfortunately there are in general no standard techniques available for solving (13) as they are available for solving ordinary least-squares problems. As a consequence one has to resort to methods that in one way or another make use of

a discrete search strategy for finding the minimizer of (13). The idea is to use the objective function of (13) for introducing an ellipsoidal region in R^n , on the basis of which a search is performed for the minimizer of (13). The ellipsoidal region is given by

$$(14) \quad (x-\hat{x})^* Q_{\hat{x}}^{-1} (x-\hat{x}) \leq \chi^2.$$

Through the selection of the positive constant χ^2 the size of the ellipsoidal region can be controlled. However, already with the selection of χ^2 care has to be exercised. A small value for χ^2 may result in an ellipsoidal region that fails to contain the minimizer of (13). And a too large value for χ^2 may result in a region for which the search for the minimizer becomes too time-consuming. Unfortunately it is difficult to formulate a data-independent criterion for selecting χ^2 , that ensures that the region contains one or more gridpoints. This is due to the fact that the ellipsoidal region is centered at $\hat{x} \in R^n$ and not centered at a gridpoint of Z^n . Would the latter be the case, then the volume of the ellipsoid could be used to set a reference value for χ^2 . It can namely be shown that for the case $\hat{x} \in Z^n$, the ellipsoidal region (14) would at least contain one gridpoint other than \hat{x} if its volume is larger than or equal to 2^n (which is the volume of the cube $|x_i| \leq 1$, $i=1,\dots,n$). In our case an alternative approach has to be taken. One approach would be to round all the individual coordinates of \hat{x} to their nearest integer, substitute the so obtained vector for x in the left-hand side of (14) and then take χ^2 to be equal to the function value of the quadratic form. This approach at least ensures that (14) contains minimally one gridpoint. However, the so obtained value for χ^2 may also be overly conservative. And this may occur especially when the ellipsoidal region is extremely elongated (which is typically the case with GPS when the observational timespan is short). In the context of selecting χ^2 , it is of interest to note that our numerical experiments indicate that the *volume* of the ellipsoidal region (14) gives a good approximation to the number of integer vectors that lie within the ellipsoidal region. This suggests that one could use the volume of the ellipsoidal region as indicator to decide whether or not the scalar χ^2 should be scaled down or scaled up.

An alternative and from a statistical testing point of view more appealing approach would be to rely on and to make use of the statistical distribution of the least-squares estimator of x . If the observables are normally distributed with mean Ax and variance matrix Q_y , then the quadratic form of (14) has a central Chi-square distribution with $m-n$ degrees of

freedom (it is a central F -distribution if $Q_{\hat{x}}$ has been scaled with the a posteriori variance factor). As reference value for χ^2 one may now choose χ^2 to be equal to the α -percentage point of the Chi-square (or F -) distribution. With this choice for χ^2 one is of course not certain that (14) indeed contains a gridpoint. But the choice does ensure that (14) contains a grid point with probability $1-\alpha$. This gridpoint is the mean of the least-squares estimator of x .

The above shows that one should give some consideration to the way χ^2 is selected. If the only objective is to solve the minimization problem (13), then χ^2 should be chosen in a way that guarantees that (14) contains the minimizer. If however the objective is also to statistically validate the minimizer, then the approach based on the α -percentage point of the Chi-square (or F -) distribution can be used. Because, if α is chosen small enough and (14) still fails to contain a grid point, then the minimizer of (13) can be considered to be invalidated.

From now on it will be assumed that a value for χ^2 has been selected. With this value for χ^2 the ellipsoidal region (14) is then taken as the point of departure for developing a search strategy to obtain the minimizer of (13). Different search strategies are possible and some of them have in fact been proposed already in the GPS-literature. In the following we will review two of such concepts that already have been in use for GPS ambiguity fixing. They are based on using the planes of support of an ellipsoid and on completing a quadratic equation to squares. These two concepts are reviewed in sections 4 and 5.

4 Ellipsoidal planes of support

One way of finding the minimizer of (13) is to identify first the set of gridpoints that satisfy the inequality (14) and then to pick that gridpoint that gives the smallest function value for the quadratic form. However, the quadratic form of (14) can not be used as such to identify the set of candidate gridpoints. The idea is therefore to replace inequality (14) with an equivalent description that is based on using the planes of support of the ellipsoid. This equivalence can be constructed as follows: Let a be an arbitrary vector of R^n and let $x-\hat{x}$ be orthogonally projected onto a . The orthogonal projection (where orthogonality is measured with respect to the metric $Q_{\hat{x}}$) of $x-\hat{x}$ onto a is then given as: $a(a^* Q_{\hat{x}}^{-1} a)^{-1} a^* Q_{\hat{x}}^{-1} (x-\hat{x})$. And the square of the length of this vector reads: $[a^* Q_{\hat{x}}^{-1} (x-\hat{x})]^2 / (a^* Q_{\hat{x}}^{-1} a)$. Now, since the length of the orthogonal projection of a vector onto an arbitrary direction is always less than or equal to the length of the vector

itself, we have

$$(x-\hat{x})^* Q_{\hat{x}}^{-1} (x-\hat{x}) = \max_{a \in R^n} \frac{[a^* Q_{\hat{x}}^{-1} (x-\hat{x})]^2}{a^* Q_{\hat{x}}^{-1} a}.$$

From this follows then, when a is replaced by $Q_{\hat{x}} c$, the equivalence

$$(15) \quad (x-\hat{x})^* Q_{\hat{x}}^{-1} (x-\hat{x}) \leq \chi^2 \Leftrightarrow \frac{[c^* (x-\hat{x})]^2}{c^* Q_{\hat{x}} c} \leq \chi^2, \forall c \in R^n.$$

Both type of inequalities describe the same ellipsoidal region. In the second type we recognize $c^*(x-\hat{x}) = \pm(c^* Q_{\hat{x}} c)^{1/2} \chi$, which is the pair of parallel planes of support of the ellipsoid having vector c as normal. The above equivalence therefore states that the ellipsoidal region coincides with the region that follows from taking all intersections of the areas between each pair of ellipsoidal planes of support. Hence, in order to find the candidate gridpoints that satisfy (14) we may as well make use of the ellipsoidal planes of support.

When working with the above equivalence for our purposes, there are however two restrictions that need to be appreciated. First of all, the above equivalence only holds for the *infinite* set of planes of support. But for all practical purposes one can only work with a finite set. Working with a finite set implies however that the region bounded by the planes of support will be larger in size than the original ellipsoidal region. Of course, one could think of minimizing the increase in size by choosing an appropriate set of normal vectors c . For instance, if the normal vectors c are chosen to be in the direction of the major and minor axes of the ellipsoidal region, then the resulting region will fit the ellipsoid best. But here is where the second restriction comes into play. One simply has no complete freedom in choosing the planes of support. Their normals c should namely be chosen such that the resulting interval $[c^*(x-\hat{x})]^2 \leq c^* Q_{\hat{x}} c \chi^2$ can indeed be used for selecting candidate grid points. Hence, the normals c can not be chosen arbitrarily.

The simplest approach to the above would be to submit oneselfs to this situation and to choose the normals to be parallel to the grid axes. When the normals are chosen as $c_i = (0, \dots, 1, 0, \dots, 0)^*$, with the 1 as the i th-coordinate, the region bounded by the planes of support becomes

$$(16) \quad (x_i - \hat{x}_i)^2 \leq \sigma_{\hat{x}_i}^2 \chi^2, \quad i=1, \dots, n,$$

where $\sigma_{\hat{x}_i}^2$ is the variance of the least-squares estimator of x_i . The intervals of (16) can be used to select candidate gridpoints from which then the minimizer of (13) can be chosen.

Although the approach based on (16) is certainly a valuable one, it will be clear that it can become quite time-consuming when the region defined by (16) is significantly larger than the original ellipsoidal region. And this will definitely be the case when the ellipsoid is both elongated and rotated with respect to the grid axes. Of course, one can reduce the size of the region by introducing additional planes of support. For instance, in addition to $c_i = (0, \dots, 1, 0, \dots, 0)^*$, admissible choices for the normals are the sum or differences of c_i and c_j , $i, j=1, \dots, n$, $i \neq j$. And depending on the elongation and orientation of the ellipsoid, this may indeed significantly reduce the size of the region enclosed by the planes of support. The problem remains however that this way of including additional planes of support is somewhat ad hoc and need not necessarily lead to a significant reduction in size of the region.

One conclusion that can be drawn from the above discussion is that the efficacy of the method depends to a large extent on the elongation and orientation of the ellipsoid. This observation has been the motivation for developing the method that will be discussed in section 6. First however we will review another interesting concept that already has been in use, in one form or another, for GPS ambiguity fixing. This concept is based on completing the square of a quadratic equation, or phrased in statistical terms, it is based on a sequential conditional least-squares adjustment.

5 Conditional Least-Squares

When use is made of the planes of support as previously discussed, all bounds (such as $\sigma_{\hat{x}_i}^2 \chi^2$ in (16)) are set prior to the actual search. That is, the setting of the bounds is independent of the search process. One may wonder however whether it is not possible to keep adjusting these bounds during the search process. For instance, let x be partitioned as $x = (x_1^*, x_2^*)^*$ and assume that already candidate integers have been found for the elements of x_1 . Then clearly it is possible to formulate new bounds for the elements of x_2 that are sharper in general than the original bounds in the previous section. And as it turns out, this idea can be implemented very efficiently when use is made of completing the square of a quadratic equation.

In order to describe the method we will start with the two-dimensional case first. Let the two-dimensional ellipsoidal region be given as

$$(17) \quad ax_1^2 + 2bx_1 x_2 - cx_2^2 \leq \chi^2,$$

where $a>0$, $c>0$, $ac-b^2>0$. For the moment we simply assume the ellipse to be centered at the origin. When we use the approach of the previous section and apply (16), the two bounds for x_1 and x_2 follow as

$$(18) \quad x_1^2 \leq \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \chi^2 = \chi^2 / (a - b^2/c),$$

and

$$(19) \quad x_2^2 \leq \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \chi^2 = \chi^2 / (c - b^2/a).$$

Bound (18) is the sharpest possible bound for x_1 when nothing is known about x_2 . Similarly, bound (19) is the sharpest possible bound for x_2 when nothing is known about x_1 . But each of these bounds can be improved once the other parameter is known. This can be seen when (17) is completed to a sum of squares. Completing the square of (17) gives

$$(20) \quad a(x_1 - \frac{b}{a}x_2)^2 + (c - \frac{b^2}{a})x_2^2 \leq \chi^2.$$

And from this follows that one can bound x_1 as

$$(21) \quad (x_1 - \frac{b}{a}x_2)^2 \leq \frac{c-b^2/a}{a} \left[\frac{\chi^2}{c-b^2/a} - x_2^2 \right].$$

This shows, since $a \geq a-b^2/c$ and $x_2^2 \geq 0$, that except for the trivial case $b=0$, the bound of (21) is always sharper than the bound of (18). Advantage can therefore be gained from replacing (18) by (21). And this gain may be considerable when the ellipse is elongated and rotated with respect to the grid axes. Since the bound of (21) depends on x_2 , first (19) should be used to come up with an integer candidate for x_2 . This step is then followed by (21) for determining an integer candidate for x_1 . Once a pair of integer candidates has been found, a new and smaller reference value for χ^2 can be computed. This enables us then to shrink the ellipsoidal region and to perform a renewed search for integer candidates. In this way one can efficiently scan the ellipsoidal region to locate the sought for minimizer of (13).

In order to generalize the above to the multi-dimensional case, we first note that (20) can be written in vector-matrix form as

$$(22) \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ b/a & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & (c-b^2/a) \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ b/a & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \chi^2.$$

This shows that completing the square corresponds to a triangular decomposition. Also note that with

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}^{-1},$$

$$\begin{bmatrix} 1 & 0 \\ b/a & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\sigma_{12}\sigma_2^{-2} \\ 0 & 1 \end{bmatrix}, \text{ and}$$

$$\begin{bmatrix} a & 0 \\ 0 & c-b^2/a \end{bmatrix} = \begin{bmatrix} (\sigma_1^2 - \sigma_{12}^2/\sigma_2^2)^{-1} & 0 \\ 0 & 1/\sigma_2^2 \end{bmatrix},$$

the inequality of (20) can be written as

$$(23) \quad (\hat{x}_{1|2}/\sigma_{1|2})^2 + (x_2/\sigma_2)^2 \leq \chi^2,$$

and the two inequalities (19) and (21) can be written as

$$(24) \quad x_2^2 \leq \sigma_2^2 \chi^2 \text{ and } \hat{x}_{1|2}^2 \leq \frac{\sigma_{1|2}^2}{\sigma_2^2} (\sigma_2^2 \chi^2 - x_2^2),$$

in which we recognize the conditional least-squares estimate $\hat{x}_{1|2} = x_1 - \sigma_{12}\sigma_2^{-2}x_2$ and its variance $\sigma_{1|2}^2 = \sigma_1^2 - \sigma_{12}^2/\sigma_2^2$. This shows that the triangular decomposition corresponds to a conditional least-squares adjustment [1], [12]. To generalize we therefore will use a *sequential conditional least-squares* adjustment for the multi-dimensional case.

A sequential conditional least-squares adjustment (or an LDU-decomposition of $Q_{\hat{x}}^{-1}$), starting with a conditioning on x_n and ending with a conditioning on x_1 , will then in analogy to (23) give for the multi-dimensional case,

$$(25) \quad \sum_{i=1}^n (x_i - \hat{x}_{i|i-1,\dots,n})^2 / \sigma_{i|i-1,\dots,n}^2 \leq \chi^2.$$

From this and in analogy to (24), we can construct the following bounds

$$(26) \quad (x_i - \hat{x}_{i|i-1,\dots,n})^2 \leq \lambda_{\hat{x}_i} \sigma_{i|i-1,\dots,n}^2 \chi^2, \quad i = 1, \dots, n,$$

with

$$(27) \quad \lambda_{\hat{x}_i} = \left[1 - \sum_{j=i+1}^n \frac{(x_j - \hat{x}_{j|i-1,\dots,n})^2}{\sigma_{j|i-1,\dots,n}^2 \chi^2} \right]$$

Here we have made use of the notation $\hat{x}_{i|i-1,\dots,n}$ to denote the least-squares estimate of x_i conditioned on fixing x_j , $j = i+1, \dots, n$.

Compare (26) with (16). Since clearly $0 \leq \lambda_i \leq 1$, and since a conditional variance is always smaller than or equal to its

unconditional counterpart, $\sigma_{i|i-1,\dots,n}^2 \leq \sigma_{\hat{x}_i}^2$, it follows that the bounds of (26) (or (23)) are always sharper than or at least as sharp as the bounds of (16).

Also note the regularity in the above bounds. The bound for x_i is equal to the gap in the previous bound times the ratio $\sigma_{i|i-1,\dots,n}^2 / \sigma_{i-1|i-2,\dots,n}^2$. The above set of inequalities can now be used as follows for solving (13). First a candidate integer is determined for x_n using the first bound. This candidate integer is then used to compute the second bound, from which a candidate integer for x_{n-1} is determined. This process is continued up to the point that a complete vector x with candidate integer coordinates is constructed. With this vector one can then shrink the ellipsoidal search region and perform a renewed search within the shrunken ellipsoidal region. Repeated application of the above steps will then finally lead to the minimizer of (13). Note that the renewed search need not necessarily commence with x_n . If for instance it is known that due to the shrinking of χ^2 , the integers $x_n, x_{n-1}, \dots, x_{k-1}$ are the only candidates, the renewed search may commence with x_k .

It may happen that the above procedure halts before a complete candidate vector x has been found. This occurs when the size of the bound is such that no candidate integer lies within the interval. If this happens to be the case, one should return to the previous bound and increase (or decrease) the previously found candidate integer by one. From there on one can then continue again.

It can be deduced that the bounds of (26) have the *tendency* to become smaller as the index i gets smaller. Clearly λ_i gets smaller as the index i gets smaller. But generally also $\sigma_{i|i+1,\dots,n}^2$ has the tendency to get smaller as the index i gets smaller. The more constraints are included the smaller the conditional variance gets. This tendency can also be explained if we interpret $\sigma_{i|i+1,\dots,n}^2$ geometrically. It can be shown that

$$\sigma_{i|i+1,\dots,n}^2 = \|P_{A_{(i)}}^\perp a_i\|^2 = \|a_i\|^2 \sin^2 \alpha_i$$

where: a_i is the i th-column vector of matrix A , $A_{(i)}$ is the matrix that follows from taking the first $(i-1)$ -number of column vectors from A and α_i is the angle between vector a_i and the range space of matrix $A_{(i)}$. Now, the angle α_i will have the tendency to become smaller the larger the dimension of the range space of $A_{(i)}$ gets. In fact the angle will be zero when the dimension of the range space equals n . Therefore, when the lengths of the column vectors of matrix A are approximately constant, also $\sigma_{i|i+1,\dots,n}^2$ will have the tendency

to decrease when the index i gets larger.

The above suggests that in order to reduce the potential of halting, it may be worthwhile to order the elements of \hat{x} according to their (conditional) precision. Because, even if the bound of (26) is small at a certain level $i=l$, halting will not take place as long as $\hat{x}_{l+1,\dots,n}$ is sufficiently pushed towards an integer value. But this requires that the previously chosen integers $x_i, i=l+1,\dots,n$, are indeed coordinates of an integer vector x that lies within the ellipsoidal region. And the probability that this will be the case is higher the better the precision of these elements is.

The problem that the search for the integer candidate vector halts, is a serious one in case of GPS carrier-phase processing. For a single baseline model, it can namely be shown, see [15], that the spectrum of conditional variances of the ambiguities, $\sigma_{i|i-1,\dots,n}^2$ for $i=n,\dots,1$, has a large discontinuity when passing from $\sigma_{n-2|n-1,n}^2$ to $\sigma_{n-3|n-2,n-1,n}^2$. In fact, one can show that $\sigma_{j|j-1,\dots,n}^2 < \sigma_{i|i-1,\dots,n}^2$ for $j=1,\dots,n-3$ and $i=n-2,n-1,n$. This implies, since the first three bounds of (26) will be rather loose, that quite a number of integer-triples satisfy these bounds. But, this on its turn implies, when we start working with the fourth bound, which is very tight due to the steep decrease in value of the conditional variances, that we have a high likelihood of not being able to find an integer candidate that satisfies this fourth bound. The potential of halting is therefore very significant when one passes from the third to the fourth bound. As a consequence a large number of trials are required, before one is able to move on to the next bound. And it is this inefficiency, that will be tackled by our method proposed in the next section. The method that will be introduced in the next section, overcomes the problem of halting, through a flattening and a lowering of the level of the GPS spectrum of ambiguity conditional variances.

6 The integer GPS ambiguity transformation

6.1 The idea of reparametrization

In the previous two sections we have dealt with two ways of solving the integer least-squares problem (5) (or (13)). First the use of the ellipsoidal planes of support was discussed. But as was pointed out, the bounds that follow from using the ellipsoidal planes of support can be rather conservative, in particular when the ellipsoid is elongated and rotated with respect to the grid axes. Moreover, these bounds are fixed from the outset. This observation then led

to the idea to introduce adjustable bounds, bounds that are made dependent on the stage of progress of the search process. These bounds were obtained through a sequential conditional least-squares adjustment, which resulted in the introduction of the conditional least-squares estimates $\hat{x}_{ii=1,\dots,n}$, $i=1,\dots,n$. And it was shown that these bounds are indeed much less conservative. Up to this point however, we have been working solely on the basis of representation (14). But one may wonder whether it is not possible to obtain a further improvement in the search process, if one can replace (14) with an alternative but equivalent representation. And this indeed turns out to be the case. A new idea of the present section is therefore to *reparametrize* the integer least-squares problem such that an equivalent formulation is obtained, but one that is much easier and hence much faster to solve. In order to understand what our reparametrization should achieve, we first pause for a moment to present two ways of visualizing the integer least-squares problem. We will start with the data space point of view.

Assume that $Q_y = I_m$ (if this is not the case one simply has to replace in the following, A by $Q_y^{-1/2}A$ and P_Ay by $Q_y^{-1/2}P_Ay$), and let $G_y = \{y \in R^m | y = Ax, x \in Z^n\}$ be the set of gridpoints that is generated by the column vectors of A . Then (5) amounts to finding that element of G_y that has the least (cartesian) distance to the least-squares estimate P_Ay . In general this is a nontrivial problem to solve. The intricacy of the problem stems from the fact that although the metric is standard ($Q_y = I_m$), the grid G_y is not (the columnvectors of A are oblique in general). The problem becomes trivial however if in addition to the metric being standard, also the grid is standard (or at least orthogonal). Because if this happens to be the case, then the columnvectors of A are mutual orthogonal and (5) can simply be solved as follows. First the consistent system of equations

$$P_Ay = Ax$$

is solved for x , giving the least-squares estimate \hat{x} . And then the minimizer of (5) is obtained from a simple rounding of the individual elements of \hat{x} to the nearest integer.

The same conclusion is reached if we visualize the integer least-squares problem from a parameter space point of view. But contrary to the data space point of view we now have a non-standard metric with a standard grid. In formulation (13), x ranges namely through the standard set of gridpoints of R^n , which is Z^n , whereas distance is now measured with a non-standard metric, namely Q_x^{-1} . But as with the data space point of view we again observe that (13) becomes a trivial problem once both the metric and grid are standard. The

conclusion reads therefore that the integer least-squares problem (5) or ((13)) can simply be solved by means of rounding, if the columnvectors of A are mutual orthogonal, or if the matrix Q_x is diagonal. In order to see what happens to the methods of section 4 and 5, when Q_x is diagonal, consider the following. If Q_x is diagonal, the orientation of the ellipsoid is such that its major and minor axes are parallel to the grid axes. And in that case the n -dimensional rectangular box defined by (16) indeed fits the ellipsoid best. In that case the conditional variances also reduce to ordinary variances, and the left-hand sides of (23) become identical to those of (16). The bounds of (23) remain however sharper than those of (16).

Since the situation where Q_x is diagonal is the best one can hope for in any integer least-squares problem, we will try to find ways to come as close as possible to this ideal situation. This in short, is the essence of the method of this section.

6.2. The admissible ambiguity transformations

Let Z be an n -by- n matrix of full rank and define

$$(28) \quad z = Z^*x, \hat{z} = Z^*\hat{x}, Q_{\hat{z}} = Z^*Q_xZ.$$

Then

$$(29) \quad (x - \hat{x})^*Q_x^{-1}(x - \hat{x}) = (z - \hat{z})^*Q_{\hat{z}}^{-1}(z - \hat{z}).$$

The variance-covariance matrix $Q_{\hat{z}}$ is clearly diagonal if matrix Z contains the eigenvectors of Q_x . Unfortunately however, this choice for matrix Z is not admissible in case of our *integer* least-squares problem. Because, if this choice for Z would be used, the vector in Z^n that follows from rounding the coordinates of \hat{z} to their nearest integer would in general fail to produce a vector x in Z^n . In other words, if $z \in Z^n$ then generally $x = Z^{-1}z \notin Z^n$. This dilemma points out that only a restricted class of transformations qualifies for reparametrizing the integer least-squares problem. Fortunately, this class of transformations can easily be characterized [16]. They need to be *volume preserving* and have elements which are *integers*. Typical examples of matrices that fall in this class are the identity-matrix and the permutation matrices. Note that we already made use of the permutation matrices in the previous section when ordering the elements of \hat{x} according to their (conditional) precision. Also within the context of GPS double-difference ambiguity fixing, one in fact already has been using transformations that are volume preserving and have integer elements. This is the case when one changes from one set of double-

difference ambiguities to another set having a different satellite as reference. This can be seen as follows. Take as an example the situation that five satellites are available. The single-difference ambiguity related to satellite i is denoted as a_i , and the corresponding double-difference ambiguity having satellite j as reference is denoted as $a_i^{(j)} = a_i - a_j$. The regular transformation from $a_i^{(1)}$ to $a_i^{(2)}$ reads then

$$\begin{pmatrix} a_1^{(2)} \\ a_3^{(2)} \\ a_4^{(2)} \\ a_5^{(2)} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_2^{(1)} \\ a_3^{(1)} \\ a_4^{(1)} \\ a_5^{(1)} \end{pmatrix}.$$

The transformation matrix in this expression has integer elements and its determinant equals -1. Hence, it follows that this matrix is indeed a member of the class of admissible ambiguity transformations. In a similar way one can show that all transformations that change the reference satellite of the double-difference ambiguities belong to the above mentioned class of admissible transformations.

By choosing a matrix Z from the above mentioned class, we can now replace our original integer least-squares problem (13) by the equivalent, but *reparametrized* integer least-squares problem

$$(30) \quad \min_z (z - \hat{z})^T Q_{\hat{z}}^{-1} (z - \hat{z}), \quad z \in \mathbb{Z}^n.$$

And once the minimizer of (30) has been found, the minimizer of (13) can be recovered from invoking $x = Z^{-1}z$.

It will be clear that because of the restrictions on Z , no true diagonality of $Q_{\hat{z}}$ can be hoped for. This leaves us with two questions. Firstly, how to measure the non-diagonality of $Q_{\hat{z}}$, and secondly, how to choose matrix Z so as to obtain near-diagonality. In order to answer the first question, we note that the variance-covariance matrix $Q_{\hat{z}}$ is diagonal if and only if its *correlation* matrix $R_{\hat{z}}$ equals the identity matrix. That is, $Q_{\hat{z}}$ is diagonal if and only if all elements of \hat{z} are fully decorrelated. In the two-dimensional case the determinant of $R_{\hat{z}}$ is related to the correlation coefficient as: $\det(R_{\hat{z}}) = 1 - \rho_{\hat{z}}^2$. This shows that for the two-dimensional case, $Q_{\hat{z}}$ is diagonal if and only if $\det(R_{\hat{z}}) = 1$. But, it can be shown that this also holds true for dimensions higher than two [15]. We therefore introduce as measure of diagonality of $Q_{\hat{z}}$, the scalar

$$(31) \quad r_{\hat{z}} = \det(R_{\hat{z}})^{\frac{1}{2}} \quad (0 \leq r_{\hat{z}} \leq 1).$$

Since the scalar $r_{\hat{z}}$ measures the decorrelation between the

elements of \hat{z} , it will be referred to as the *decorrelation number* of \hat{z} . The elements of \hat{z} are fully decorrelated when $r_{\hat{z}}$ equals one, and they are poorly decorrelated when $r_{\hat{z}}$ is close to zero. It follows from the triangular decomposition of $Q_{\hat{z}}$, that $r_{\hat{z}}$ is related to the spectrum of *conditional* and *unconditional* standard deviations as

$$(32) \quad r_{\hat{z}} = \prod_{i=1}^n \frac{\sigma_{i|i-1,\dots,n}}{\sigma_{\hat{z}_i}}.$$

From the fact that the nominator in this expression is independent of Z , since $\det(Q_{\hat{z}}) = \det(Z^T Q_{\hat{x}} Z) = \det(Q_{\hat{x}})$, follows, that the elements of \hat{z} are less correlated than those of \hat{x} , $r_{\hat{z}} > r_{\hat{x}}$, when $\sigma_{\hat{x}_i} > \sigma_{\hat{z}_i}$. Hence, the variance-covariance matrix $Q_{\hat{z}}$ is less *non-diagonal* than $Q_{\hat{x}}$, when its *diagonal* elements are smaller than those of $Q_{\hat{x}}$. The gain in decorrelation can be measured by the ratio

$$(33) \quad r_{\hat{z}}/r_{\hat{x}} = \prod_{i=1}^n \frac{\sigma_{\hat{x}_i}}{\sigma_{\hat{z}_i}}.$$

This gain can be given the following geometrical interpretation. The volume of the n -dimensional rectangular box (16) that encloses the ellipsoidal region (14) is given as $2^n \chi^n \Pi \sigma_{\hat{x}_i}$. Similarly, the n -dimensional box that encloses the ellipsoidal region defined by $Q_{\hat{z}}$, has volume $2^n \chi^n \Pi \sigma_{\hat{z}_i}$. This shows, that it is the relative decrease in volume of the n -dimensional box, that directly measures the gain in decorrelation. A maximum decrease in volume is achieved when $\prod_{i=1}^n \sigma_{\hat{z}_i} = \prod_{i=1}^n \sigma_{\hat{x}_i}$, in which case $r_{\hat{z}} = 1$.

6.3 On the choice of reparametrization in 2D

In order to answer the question as to how to construct matrix Z , we first consider the problem in two dimensions. Let \hat{x} and $Q_{\hat{x}}$ be given as

$$(34) \quad \hat{x} = \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} \text{ and } Q_{\hat{x}} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}.$$

We now need to come up with a matrix Z , which has integer entries, which is volume-preserving, and which allows us to decorrelate the two elements of \hat{x} . We already know from section 5, see equation (23), that a complete decorrelation is obtained, when \hat{x}_1 is replaced by its corresponding *conditional least-squares estimate* $\hat{x}_{1|2}$. The transformation that achieves this, is given as

$$(35) \quad \begin{pmatrix} \hat{x}_{1|2} \\ \hat{x}_2 \end{pmatrix} = \begin{pmatrix} 1 & -\sigma_{12}\sigma_2^{-2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix}$$

For reasons of convenience, we have assumed the expectations of \hat{x}_1 and \hat{x}_2 to be simply zero for the moment. Note, that transformation (35) not only decorrelates, but, in line with the correspondence between linear least-squares estimation and best linear unbiased estimation, also returns $\hat{x}_{1|2}$ as the element which has the best precision of all linear unbiased functions of \hat{x}_1 and \hat{x}_2 . Instead of using (35), we can of course also interchange the role of the two entries of \hat{x} and use the transformation

$$(36) \quad \begin{pmatrix} \hat{x}_1 \\ \hat{x}_{2|1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\sigma_{21}\sigma_1^{-2} & 1 \end{pmatrix} \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix}.$$

Both of the above transformations fully decorrelate. Also note, that both transformations are volume-preserving. Hence, the only condition that is still not satisfied, is the condition that the entries of the transformation need to be integer. In order to repair this situation, we *approximate* the above transformations by replacing $\sigma_{21}\sigma_1^{-2}$ by $[\sigma_{21}\sigma_1^{-2}]$, or $\sigma_{12}\sigma_2^{-2}$ by $[\sigma_{12}\sigma_2^{-2}]$, where $[.]$ stands for "rounding to the nearest integer". The volume-preserving property is retained by this integer-approximation. The decorrelation-property is however, not retained. But still, one can show, although the integer-approximation does not allow for a complete decorrelation, that it allows one to improve the precision of the elements and that it allows one to bound the correlation between the elements when the two transformations are used in an alternating fashion. Based on (35) and (36), the idea is therefore to use the following two type of transformations

$$(37) \quad Z_1^* = \begin{pmatrix} 1 & z_{12} \\ 0 & 1 \end{pmatrix} \text{ and } Z_2^* = \begin{pmatrix} 1 & 0 \\ z_{21} & 1 \end{pmatrix},$$

in which z_{12} and z_{21} are appropriately chosen integers. The two type of transformations are applied in such a way, that they replace the element with the poorest precision with one that has an improved precision. Thus, when $\sigma_2^2 \leq \sigma_1^2$, we start with Z_1^* and we choose the scalar z_{12} as $z_{12} = -[\sigma_{12}\sigma_2^2]$. This gives

$$(38) \quad Z_1 Q_{\hat{x}} Z_1^* = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \text{ with } \sigma_1^2 \leq \sigma_2^2.$$

Then, if the precision of the first element is still not better than that of the second, $\sigma_2^2 \leq \sigma_1^2$, we stop, else we continue with Z_2^* and choose z_{21} as $z_{21} = -[\sigma_{21}\sigma_1^2]$. This gives then

$$(39) \quad Z_2 Z_1^* Q_{\hat{x}} Z_1 Z_2 = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \text{ with } \sigma_2^2 \leq \sigma_1^2.$$

Then, if the precision of the second element is still not better than that of the first, $\sigma_1^2 \leq \sigma_2^2$, we stop, else we continue again with Z_1^* and choose z_{12} as $z_{12} = -[\sigma_{12}\sigma_2^2]$. This whole process of alternatingly using Z_1^* and Z_2^* finally stops when one fails to improve the precision of the elements. And when this happens, the correlation coefficient is bounded as $\rho_{\hat{z}}^2 \leq \frac{1}{4}$, since then both of the inequalities, $|\sigma_{12}\sigma_2^2| \leq \frac{1}{2}$ and $|\sigma_{21}\sigma_1^2| \leq \frac{1}{2}$, are satisfied.

Geometrically, the above sequence of transformations can be given the following useful interpretation. Imagine the confidence-ellipse of \hat{x} . The first transformation Z_1^* then pushes the two *vertical* tangents of the ellipse from the $\pm\sigma_1\chi$ level towards the $\pm\sigma_1\chi$ level, while at the same time keeping fixed the volume (area) of the ellipse and the location of the two horizontal tangents of the ellipse. The second transformation Z_2^* then pushes the two *horizontal* tangents of the ellipse from the $\pm\sigma_2\chi$ level towards the $\pm\sigma_2\chi$ level, while at the same time keeping fixed the volume of the ellipse and the location of the two vertical tangents. And this process is continued until the next transformation reduces to the trivial identity. Since the volume of the ellipse is kept constant at all times, whereas the volume of the enclosing rectangular box is reduced in each step, it follows that not only the decorrelation number gets improved, but also that the shape of the ellipse is forced to become more sphere-like.

Once the above sequence of transformations that make up Z^* has been applied, we have $\rho_{\hat{z}}^2 \leq \frac{1}{4}$, which implies for the decorrelation number that

$$(40) \quad r_{\hat{z}}^2 \geq 3/4.$$

This is a very significant result, since we know that the original double-difference ambiguities are extremely correlated when based on short timespan carrier-phase data.

From this bound also follows, together with $\sigma_{\hat{z}_2}^2 \leq \sigma_{\hat{z}_1}^2$ and $\sigma_{\hat{z}_{1|2}}^2 = r_{\hat{z}}^2 \sigma_{\hat{z}_1}^2$, that

$$(41) \quad \sigma_{\hat{z}_{1|2}}^2 \geq \frac{3}{4} \sigma_{\hat{z}_1}^2.$$

Hence, the transformation Z^* guarantees that the transformed conditional variance $\sigma_{\hat{z}_{1|2}}^2$ will never be much smaller than

the variance $\sigma_{\hat{x}_2}^2$. But this implies, that the transformation removes to a large extent any discontinuity that might be present in the original variances, $\sigma_{\hat{x}_{1|2}}^2 \ll \sigma_{\hat{x}_2}^2$. And as we observed earlier in section 5, this is precisely the situation that we are confronted with in case of GPS carrier-phase processing.

6.4 Bounding the triangular factor

In order to obtain a higher-dimensional version of Z^* , we first try to find a generalization of Z_1^* . In two dimensions, \hat{x} is transformed by Z_1^* as

$$(42) \quad \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} = \begin{pmatrix} 1 & -[\sigma_{12}\sigma_2^{-2}] \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix}$$

If we substitute the inverse of (35) into the right-hand side of (42) and apply the error propagation law, we get

$$\begin{pmatrix} \sigma_{\hat{x}_1}^2 & \sigma_{\hat{x}_1 \hat{x}_2} \\ \sigma_{\hat{x}_2 \hat{x}_1} & \sigma_{\hat{x}_2}^2 \end{pmatrix} \begin{pmatrix} 1 & \varepsilon \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sigma_{\hat{x}_{1|2}}^2 & 0 \\ 0 & \sigma_{\hat{x}_2}^2 \end{pmatrix} \begin{pmatrix} 1 & \varepsilon \\ 0 & 1 \end{pmatrix} \text{ with } |\varepsilon| \leq \frac{1}{2}.$$

(43)

It is the inverse of the unique *LDU*-decomposition of the inverse of the variance-covariance matrix of \hat{x}_1 and \hat{x}_2 . This result illustrates once again that Z_1^* tries to diagonalize the variance-covariance matrix, by bounding its triangular factor.

In order to generalize (43) to dimensions higher than two, we start from the inverse of the *LDU*-decomposition of $Q_{\hat{x}}^{-1}$. It reads $Q_{\hat{x}}^{-1} = U^{-1}D^{-1}L^{-1}$. Note that like L , L^{-1} is lower triangular having one's on the main diagonal. Also note that if the elements of L would be integer, then so would be the elements of L^{-1} . In fact, matrix L , being integer and volume-preserving, would then be the perfect candidate to diagonalize $Q_{\hat{x}}$. One would then be able to truly diagonalize $Q_{\hat{x}}$ in just one step. This observation suggests, even though the entries of L will be non-integer in general, that we choose the n -by- n matrix Z_1 as a lower triangular matrix, with integer entries and with one's on the main diagonal. Now, in order to make $Z_1^* Q_{\hat{x}} Z_1$ approximately diagonal, one could in first instance think, in analogy with the two-dimensional case, of setting Z_1 equal to L after all its elements have been rounded to their nearest integer. Unfortunately, this approach fails for the higher-dimensional case. It can namely not guarantee that *all* the non-diagonal entries of $L^{-1}Z_1$ get sufficiently bounded. Fortunately, one can do better than this by means of sweeping integer-multiples of the rows of L^{-1} . Matrix Z_1^* can then be constructed from subtracting suitable integer

multiples of the last ($n-i$) rows of L^{-1} from row i of L^{-1} , for $i=1,\dots,(n-1)$. Using this matrix Z_1^* , one obtains

$$(44) \quad Z_1^* Q_{\hat{x}} Z_1 = (Z_1^* U^{-1}) D^{-1} (L^{-1} Z_1),$$

in which, in analogy with (43), the absolute values of all non-diagonal elements of $L^{-1}Z_1$ are guaranteed to be bounded by a half. This implies, when the non-diagonal elements of L^{-1} are larger than a half in absolute value, that the diagonal entries of $Z_1^* Q_{\hat{x}} Z_1$ will be smaller than those of $Q_{\hat{x}}$. Hence, the decorrelation number will undergo an improvement through Z_1^* . But note, since (44) is the inverse of the *unique LDU*-decomposition of the inverse of $Z_1^* Q_{\hat{x}} Z_1$, that all the conditional variances stay invariant under the transformation Z_1^* . For the GPS-ambiguities this implies, not only that the variance σ_n^2 remains large, but also that the discontinuity in the spectrum of conditional variances, which is so distinctive of GPS carrier-phase processing, stays untouched. Hence, one should not expect too much from the *single* transformation Z_1^* .

6.5 Flattening the spectrum of conditional variances

In the two-dimensional case, the two type of transformations Z_1^* and Z_2^* of (37) are used in an alternating fashion in their construction of Z^* . Instead of Z_1^* and Z_2^* , one could also say that only the type Z_1^* is used, but then each time followed by a permutation of the two elements in the vector to be transformed. This suggests for the n -dimensional case, that we perform a re-ordering of the n -elements after each time that the transformation Z_1^* of the previous subsection is applied. The difficulty that we are faced with is however, what type of re-ordering to choose? That is, in dimensions higher than two, different reordering schemes are possible, all of which reduce to a simple interchange when applied to the two-dimensional case. Hence, no unambiguous generalization of the two-dimensional case seems to exist.

Fortunately, in case of GPS carrier-phase processing, already the two-dimensional scheme based on a pairwise re-ordering, allows us to obtain results that show a dramatic improvement over the original ambiguities. In order to make this clear, we first need to consider the spectrum of conditional variances. For the GPS single baseline model, based on carrier-phase data only, we have

$$(45) \quad \sigma_{\hat{x}_{j|i-1,n}}^2 \ll \sigma_{\hat{x}_{i+1,n}}^2 \text{ for } j=1,\dots,n-3; \quad i=n-2,n-1,n.$$

And it is this large discontinuity in the spectrum of

conditional variances, that forms a hindrance for the efficient search for the integer least-squares estimates. Our aim in constructing transformation Z^* should therefore at least be, to remove the discontinuity from the spectrum of conditional variances. And a very important consequence of such a flattening of the spectrum is, that when $n>3$, the three large variances in the spectrum get reduced by a very significant amount. The volume-preserving property of Z^* implies namely, that the product of conditional variances remains unaffected by the transformation. Hence, by flattening the spectrum, the presence of the very small conditional variances automatically implies, that the three large variances in the spectrum have to get much smaller.

This observation now also stipulates the significance of *satellite redundancy* and *dual frequency* data. When both are absent, we have $n=3$. In that case, the absence of very small conditional variances prohibits us from "pulling down" the large variances in the spectrum. In case of satellite redundancy and/or dual frequency data however, we have $n>3$. Now the presence of the very small conditional variances does allow us to bring the large variances in the spectrum down to much smaller values. And the larger $n-3$ is, the more we are able to bring the flattened spectrum to a lower level.

Thus, in case of GPS carrier-phase processing, a dramatic improvement can be realized, if we would be able to remove the discontinuity and enforce the spectrum of conditional variances to become much flatter. We know from subsection 6.3, see (41), that this is precisely what the transformation Z^* does for the two-dimensional case. But this suggests for the n -dimensional case, that a steep decrease in value between two consecutive conditional variances, $\sigma_{\hat{x}_{i-1|n}}^2$ and $\sigma_{\hat{x}_{i|n}}^2$, can be removed when the two-dimensional transformation Z^* is applied to the $(i-1)$ th and i th least-squares estimates both of which are conditioned on the last $(n-i)$ estimates: $\hat{x}_{i-1|i-1..n}$ and $\hat{x}_{i|i-1..n}$. Thus, instead of applying the two-dimensional transformation of subsection 6.3 to the unconditional least-squares estimates, the idea is to apply it to the *conditional* least-squares estimates. This would then give, in analogy of (41),

$$(46) \quad \sigma_{\hat{x}_{i-1|i..n}}^2 \geq \frac{3}{4} \sigma_{\hat{x}_{i|i-1..n}}^2.$$

Note that the other conditional variances remain unaffected by the transformation. This is simply a consequence of the fact that we are transforming *conditional* least-squares estimates. Result (46) implies for $i=n-2$, that we are able to

close the large gap between the $(n-3)$ th and $(n-2)$ th conditional variance. Ofcourse, after the transformation has been applied, other, but smaller, discontinuities emerge. For instance, if the transformation has been applied for $i=n-2$, then $\sigma_{\hat{x}_{i-1|i..n}}^2 < \sigma_{\hat{x}_{i-1..n}}^2$ and $\sigma_{\hat{x}_{i-1..n}}^2 < \sigma_{\hat{x}_{i-2..n}}^2$. But, also they can be removed by applying the two-dimensional transformation. In fact, one can continue in this way and flatten the complete spectrum of conditional variances.

In summary, the proposed method thus flattens the n -dimensional spectrum of conditional variances through a repeated application of the two-dimensional transformation Z^* to the conditional least-squares estimates. And the larger $n-3$ is, the lower the level of the transformed spectrum. As a result the n th ambiguity shows a dramatic improvement in precision, $\sigma_{\hat{x}_n}^2 << \sigma_{\hat{x}}^2$. And this can also be assured for the remaining ambiguities, when the low level of the transformed spectrum is combined with a bounding of the triangular factor, as given in the previous subsection. Thus the proposed method of ambiguity-transformation, returns less correlated ambiguities with a significantly improved precision and allows for a very efficient search for the transformed integer least-squares ambiguities. Our numerical experiments indicate for instance, that with dual-frequency carrier-phase data, based on two epochs of data, with a one second sampling interval, standard deviations of the transformed ambiguities are obtained that usually are well below the one cycle level.

7 Summary and concluding remarks

In this contribution a new method was introduced for computing the integer least-squares estimates of the GPS ambiguities. It was shown in section 3 that this problem can be reduced to the integer least-squares problem

$$(47) \quad \min_x (\hat{x}-x)^T Q_{\hat{x}}^{-1} (\hat{x}-x), \quad x \in Z^n.$$

In case of GPS however, when short timespan carrier-phase data is used, the elements of \hat{x} , being the least-squares estimates of the double-difference ambiguities, are extremely correlated. Also the confidence ellipsoid will then be extremely elongated. It is not uncommon for instance, to have an elongation in the order of 10^4 , when the data is based on two epochs, one second apart (if length of minor axis 1 cm, then length of major axis 100 mtr.). The amount of correlation between the ambiguities and therefore the non-diagonality of $Q_{\hat{x}}$ can be measured by the determinant $r_{\hat{x}}^2$

of the correlation matrix. When $r_{\hat{x}}$ equals one, $Q_{\hat{x}}$ is diagonal, and when $r_{\hat{x}}$ is close to zero, then $Q_{\hat{x}}$ is far from diagonal. The scalar $r_{\hat{x}}^2$ reads in terms of the conditional variances and unconditional variances, as

$$(48) \quad r_{\hat{x}}^2 = \prod_{i=1}^n \frac{\sigma_{\hat{x}_{i+1:n}}^2}{\sigma_{\hat{x}_i}^2}.$$

And because of the large discontinuity in the spectrum of conditional variances of the ambiguities, $r_{\hat{x}}$ is usually extremely small. For instance, with dual-frequency data and a satellite redundancy of only one, $r_{\hat{x}}$ can be in the order of 10^{-19} , when the data is based on two epochs of data, one second apart. As a result of this extreme non-diagonality of $Q_{\hat{x}}$, the efficiency in solving the above integer least-squares problem is severely hindered. The idea is therefore to reparametrize the above integer least-squares problem, such that an equivalent formulation is obtained, but one that is much easier to solve. By introducing the reparametrization

$$(49) \quad \hat{z} - z = Z^*(\hat{x} - x),$$

with Z^* being integer and volume-preserving, we obtain the equivalent minimization problem

$$(50) \quad \min_z (\hat{z} - z)^* Q_{\hat{z}}^{-1} (\hat{z} - z), \quad z \in \mathbb{Z}^n,$$

with the new variance-covariance matrix $Q_{\hat{z}} = Z^* Q_x Z$. The transformed integer least-squares problem becomes trivial, when the new variance-covariance matrix $Q_{\hat{z}}$ is diagonal. That is, when $r_{\hat{z}}$ equals one. The idea is therefore to come up with a matrix Z^* that allows $r_{\hat{z}}$ to be close to one. Based on integer approximating the conditional least-squares transformation, the construction of such a matrix was given in subsection 6.3 for the two-dimensional case. It returns ambiguities with an improved precision and guarantees, because of $r_{\hat{z}}^2 \geq \frac{3}{4}$, that

$$(51) \quad \sigma_{z_{1p}}^2 \geq \frac{3}{4} \sigma_{z_2}^2.$$

From this it followed, that one can remove the discontinuity in the spectrum of conditional variances of the original ambiguities, by means of a repeated application of the two-dimensional transformation to the sequential conditional least-squares estimates of the ambiguities. As a result, the method returns: less correlated ambiguities (for instance an improvement from the above given $r_{\hat{x}} \approx 10^{-19}$ to $r_{\hat{z}} \approx 0.5$); significantly more precise ambiguities (standard deviations of the transformed ambiguities that are well below the one cycle level are not uncommon, when dual-frequency carrier-phase data is used, based on two epochs of data, with a one second

interval); and it allows one to perform the search for the transformed integer least-squares ambiguities, based on

$$(52) \quad (z_i - \hat{z}_{i|j-1, \dots, n})^2 \leq \lambda_{\hat{z}_i} \sigma_{\hat{z}_{i+1:n}}^2 \chi^2,$$

in a highly efficient manner.

To conclude, we finish with a few remarks on some untouched GPS issues. As it was pointed out in the introduction, the proposed method is directed towards solving the estimation problem of GPS-ambiguity fixing. But still, it also bears some relation to the validation step. This is particularly true, when the validation step is based on a comparison of the most-likely and the second most-likely integer minimizer. With some minor changes in the search, the method namely also allows an efficient computation of the second most-likely integer minimizer. Also, since the method is significantly faster than existing methods for GPS-ambiguity fixing, the gain in efficiency leaves us room for handling higher dimensional integer least-squares problems. This may in particular be useful in a GPS-network approach. And finally, we would like to point out that the given ambiguity transformation is completely determined by the variance-covariance matrix of the ambiguities. Even the a posteriori variance-factor need not be known. This stipulates that actual measurements are not needed to perform the transformation to near-diagonality. Hence, the necessary computations can be done in principle at the designing stage, prior the actual measurement stage.

8 References

1. Baarda, W. (1967): *Statistical Concepts in Geodesy*, Netherlands Geodetic Commission, Publication on Geodesy, New Series, Vol. 2, No. 4.
2. Blewitt, G. (1989): Carrier Phase Ambiguity Resolution for the Global Positioning System Applied to Geodetic Baselines up to 2000 km. *Journal of Geophysical Research*, Vol. 94, No. B8, pp. 10.187-10.203.
3. Cocard, M. and A. Geiger (1992): Systematic Search for all Possible Widelanes. Paper presented at the *Sixth International Geodetic Symposium on Satellite Positioning*, Columbus, Ohio, March 17-20, 1992.
4. Counselman, C.C. and S.A. Gourevitch (1981): Miniature Interferometer Terminals for Earth Surveying: Ambiguity and Multipath with Global Positioning System. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. GE-19, No. 4, pp. 244-252.
5. Euler, H.J. and H. Landau (1992): Fast GPS ambiguity

- resolution on-the-fly for real-time applications. Paper presented at the *Sixth International Geodetic Symposium on Satellite Positioning*, Columbus, Ohio, March 17-20, 1992.
6. Frei, E. (1991): *Rapid Differential Positioning with the Global Positioning System*. Schweizerischen Geodätischen Kommission, Band 44.
 7. Hatch, R. (1991): Instantaneous Ambiguity Resolution. Proceedings of IAG *International Symposium 107 on Kinematic Systems in Geodesy, Surveying and Remote Sensing*, Sept. 10-13, 1990, Springer Verlag, New York, pp. 299-308.
 8. Kleusberg, A. (1990): A Review of Kinematic and Static GPS Surveying Procedures. Proceedings of the *Second International Symposium on Precise Positioning with the Global Positioning System*, Ottawa, Canada, September 3-7 1990, pp. 1102-1113.
 9. Lenstra, H.W. (1981): *Integer Programming with a Fixed Number of Variables*. University of Amsterdam, Dept. of Mathematics, report 81-03.
 10. Remondi, B.W. (1986): Performing Centimeter-Level Surveys in Seconds with GPS Carrier Phase: Initial Results. *Journal of Navigation*, Volume III, The Institute of Navigation.
 11. Scheffé, H. (1956): *The Analysis of Variance*. John Wiley and Sons.
 12. Tienstra, J.M. (1956): *Theory of the Adjustment of Normally Distributed Observations*, Argus, Amsterdam.
 13. Teunissen, P.J.G. (1990): GPS op afstand bekijken. In: *Een halve eeuw in de goede richting*. Lustrumboek Snellius 1985-1990, pp. 215-233.
 14. Teunissen, P.J.G. (1990): Nonlinear Least Squares. *Manuscripta Geodaetica*, 15, pp. 137-150.
 15. Teunissen, P.J.G. (1993): Integer Least-Squares: Fast Estimation of the Carrier Phase Ambiguities. *First International Symposium on the Mathematical and Physical Foundations of Geodesy*, September 7-9 1993, Invited paper (in print), Stuttgart.
 16. Teunissen, P.J.G. (1993): *The Invertible GPS Ambiguity Transformations*, Delft Geodetic Computing Centre (LGR), 9 p.
 17. Wübbena, G. (1991): *Zur Modellierung von GPS-Beobachtungen für die hochgenaue Positionsbestimmung*, Hannover, 1991.

The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation

P. J. G. Teunissen

Delft Geodetic Computing Centre (LGR), Faculty of Geodetic Engineering, Delft University of Technology, Thijsseweg 11, 2629 JA Delft, The Netherlands

Received 7 July 1994; Accepted 14 November 1994

Abstract

The GPS double difference carrier phase measurements are ambiguous by an unknown integer number of cycles. High precision relative GPS positioning based on short observational timespan data, is possible, when reliable estimates of the integer double difference ambiguities can be determined in an efficient manner. In this contribution a new method is introduced that enables very fast integer least-squares estimation of the ambiguities. The method makes use of an ambiguity transformation that allows one to reformulate the original ambiguity estimation problem as a new problem that is much easier to solve. The transformation aims at decorrelating the least-squares ambiguities and is based on an integer approximation of the conditional least-squares transformation. This least-squares ambiguity decorrelation approach, flattens the typical discontinuity in the GPS-spectrum of ambiguity conditional variances and returns new ambiguities that show a dramatic improvement in correlation and precision. As a result, the search for the transformed integer least-squares ambiguities can be performed in a highly efficient manner.

Contents

1. Introduction
2. The GPS Ambiguity Estimation Problem
 - 2.1 The Carrier-Phase Observation Equation
 - 2.2 Integer Least-Squares Estimation
 - 2.3 Sequential Conditional Least-Squares Estimation
 - 2.4 Search for the Integer Least-Squares Ambiguities
 - 2.5 The GPS Spectrum of Ambiguity Conditional Variances
3. The Reparametrized Ambiguity Estimation Problem
 - 3.1 The Class of Admissible Ambiguity Transformations
 - 3.2 A 2D-Decorrelating Ambiguity Transformation

3.3 Flattening the Spectrum of Ambiguity Conditional Variances

3.4 On the Spectrum and the Gain in Baseline Precision

4. Summary and Concluding Remarks
5. References

1. Introduction

High precision relative GPS positioning is based on the very precise carrier phase measurements. A prerequisite for obtaining high precision relative positioning results, is that the double-difference carrier phase ambiguities become sufficiently separable from the baseline coordinates. Different approaches are in use and have been proposed to ensure a sufficient separability between these two group of parameters [1-8]. In particular, the approaches that explicitly aim at resolving the integer-values of the double-difference ambiguities have been very successful. Once the integer ambiguities are fixed, the carrier phase measurements will start to act as if they were high-precision pseudorange measurements, thus allowing for a baseline solution with a comparable high precision. However, the fixing of the integer ambiguities is a non-trivial problem, in particular if one aims at numerical efficiency. This topic has therefore been a rich source of GPS-research over the last decade or so. Starting from rather simple but timeconsuming integer rounding schemes, the methods have evolved into complex and effective search algorithms [9-17]. Nevertheless, at present times, it is still expedient to seek ways for improving the efficiency of the various search methods. This is in particular true for real-time applications of GPS. But to a certain extent, this is also true for some typical static applications of GPS. If we are namely really able to significantly reduce the computational effort for estimating the integer ambiguities, it may also become worthwhile to tackle problems that have dimensions higher than the one's considered sofar. For instance, it may become much easier then to simultaneously

estimate all integer double-difference ambiguities when adjusting a GPS network [18].

In this contribution a new method will be presented for the integer least-squares estimation of the double-difference carrier phase ambiguities. The method consists of two steps. In the first step, an ambiguity transformation Z^* is constructed that allows one to reparametrize the original double-difference ambiguities, such that new ambiguities are obtained, having certain desirable properties. Hence, the original integer ambiguity vector a , its real-valued least-squares estimate \hat{a} and corresponding variance-covariance matrix $Q_{\hat{a}}$, are transformed as

$$(1) \quad z = Z^*a, \hat{z} = Z^*\hat{a}, Q_{\hat{z}} = Z^*Q_{\hat{a}}Z.$$

The ambiguity transformation Z^* is required to be integer and volume preserving [19]. The kernelletter z is used to denote the transformed ambiguities. The idea of transforming the double-difference ambiguities is of course not completely new. Certain linear combinations of the GPS-observables already play a prominent role in the problem of ambiguity fixing. In particular "wide-laning" techniques have proven to be very successful. Well-known examples are the narrow-lane, the wide-lane and extra wide-lane combinations [20-22]. But also other wide-lane combinations have been studied [23]. At present however, the various integer linear combinations that are considered, are restricted to the single-channel dual-frequency case. They therefore require the use of dual-frequency carrier phase data. Moreover, since they enact at the level of a single channel only, they do not allow one to take care of the receiver-satellite geometry. The above ambiguity transformation Z^* however, does not require the use of dual-frequency carrier phase data per se and also allows one to take care of the slowly changing receiver-satellite geometry, a geometry which is so emphatically present in the ambiguity variance-covariance matrix $Q_{\hat{a}}$.

Once the transformed ambiguities have been obtained, the actual search for the integer least-squares ambiguities is performed in the second step. The search is based on a sequential conditional least-squares adjustment of \hat{z} and it derives its efficiency from the properties of the transformed ambiguities. In terms of the conditional least-squares ambiguity estimates, the transformed ambiguity searchspace is described by the inequality

$$(2) \quad \sum_{i=1}^n (\hat{z}_{i|l} - z_i)^2 / \sigma_{\hat{z}|l,i|l} \leq \chi^2,$$

in which $\hat{z}_{i|l}$ denotes the least-squares estimate of the i th ambiguity conditioned on the first $(i-1)$ number of

ambiguities and $\sigma_{\hat{z}|l,i|l}$ denotes its variance. χ^2 is an appropriately chosen positive constant that ensures that the ambiguity searchspace indeed contains the sought for integer least-squares ambiguities. It is the sum-of-squares structure in the above inequality that allows one to formulate bounds for the individual ambiguities, thus enabling that a search for the transformed integer least-squares ambiguity \hat{z} can be performed. Although the recovery of a sum-of-squares structure is not new in itself, there are important differences in how it is interpreted and used for the search [13],[16,17],[24]. Once the integer estimate \hat{z} has been determined, the corresponding baseline solution \check{b} can be recovered as

$$(3) \quad \check{b} = \hat{b} - Q_{\hat{b}\hat{z}} Q_{\hat{z}}^{-1}(\hat{z} - \check{z}),$$

in which \check{b} denotes the non-fixed least-squares baseline solution.

The basic idea that lies at the root of the method - both in the construction of Z^* , as in the formulation of the search - is, that integer least-squares ambiguity estimation becomes trivial once all the least-squares ambiguities are fully decorrelated. Their confidence ellipsoid would then be aligned with the grid axes and the sought for integer least-squares ambiguities would then simply follow from a rounding of the real-valued least-squares estimates to their nearest integer. The ambiguities would be fully decorrelated when the conditional least-squares estimates are identical to their unconditional counterparts, i.e. when $\hat{a}_{i|l} \equiv \hat{a}_i$ holds true for all i . In case of GPS however, the least-squares ambiguities are usually highly correlated and their confidence ellipsoid is usually extremely elongated. This is particularly true in case of short observational timespans and in the absence of precise P-code data. As a consequence of the intrinsic structure of the ambiguity variance-covariance matrix, the spectrum of ambiguity conditional variances, $\sigma_{\hat{a}|l,i|l}$ $i = 1, \dots, n$, generally shows a large discontinuity when passing from the third to the fourth conditional variance. But this implies, when the above inequality is used in the original parametrization, that the search would suffer from the fact that the bounds for the first three ambiguities are rather loose, whereas the remaining bounds are extremely tight. The essence of the method is therefore to aim at constructing a decorrelating ambiguity transformation Z^* , that removes the discontinuity from the spectrum. In two dimensions this is achieved by basing the ambiguity transformation on an integer approximation of the fully decorrelating conditional least-squares transformation. Transformations of this type are also known as Gauss-transformations and in the non-integer case they are considered to be the basic tools for zeroing

entries in matrices [25]. The n -dimensional case is tackled through a repeated use of the two-dimensional decorrelating ambiguity transformation. But instead of applying it to the unconditional ambiguities, it is now applied to pairs of conditional least-squares ambiguities. This approach has been motivated by the presence of the typical discontinuity in the GPS spectrum of ambiguity conditional variances and is based on ideas from [26]. The success of our method is largely due to the presence of this discontinuity and it stipulates the relevance of satellite redundancy and the use of dual-frequency data. Once the spectrum of conditional variances has been flattened, less correlated and very precise ambiguities \hat{z}_i are returned, thus allowing that the search for the transformed integer least-squares ambiguities can indeed be performed in a highly efficient manner.

In order to properly judge the significance of the present contribution, it is important that one distinguishes between the following two problems of GPS-ambiguity fixing:

1. The ambiguity *estimation* problem, and
2. The ambiguity *validation* problem.

The present contribution only addresses the first problem and not the second. The second problem, which depends on the outcome of the first, is concerned with the validation of the estimated integer ambiguities. Although the procedures for validating the estimated ambiguities which are currently in use in practice, appear to work satisfactory, it is the author's opinion that there is still some room for improving the theoretical basis of these validation procedures. For a proper statistical evaluation of both the estimated ambiguities as well as the corresponding baseline solution, it would be very helpful indeed if one has at one's disposal the probability densities of the corresponding integer estimators. This however, is a non-trivial problem, but reference is made to the discussions in [13], [27-29]. Despite the importance of proper validation procedures, the present contribution is only concerned with the integer ambiguity estimation problem. There is therefore no harm in stressing, that our method, efficient as it may be, might still come up with the wrong integer ambiguities if the data are contaminated with unmodelled effects.

2. The GPS Ambiguity Estimation Problem

In this chapter the integer least-squares ambiguity estimation problem is formulated and discussed. It shows how the search can be based on a sequential conditional least-squares

adjustment of the ambiguities and it explains why in case of GPS such a search, when performed for the original ambiguities, suffers from being inefficient. As such, chapter two identifies the dilemma with which one is confronted and shows along which lines solutions can be sought. In section 2.1, first the basic observation equation of the carrier phase observable is given. We restrict our attention to carrier phase data to accentuate that our method is in principle independent of the use of code data or even the use of dual-frequency data. But the method can easily accommodate code data as well, although with Anti-Spoofing turned on, very precise code data will probably no longer be available. The double-difference version of the carrier phase observation equation forms the basis of the model that will be used for the integer least-squares ambiguity estimation. The concept of integer least-squares estimation is discussed in section 2.2. In section 2.3 a sequential conditional least-squares estimation of the ambiguities is performed in order to recover a sum-of-squares representation for the ambiguity search space. The sum-of-squares structure is used to formulate sharp bounds for the individual ambiguities. The search procedure based on these bounds, is discussed in section 2.4. In section 2.5 the spectrum of ambiguity conditional variances is studied. It is shown, when short observational timespan carrier phase data are used, that the special structure of the variance-covariance matrix of the ambiguities usually results in a spectrum having a large discontinuity. As a consequence, the search for the original ambiguities suffers from the potential problem of halting.

2.1 The Carrier-Phase Observation Equation

The GPS observables are (P or C/A) code-derived pseudorange measurements and carrier phase measurements. These data can be available on both of the two frequencies L_1 and L_2 . In the following we will restrict our attention to the carrier phase measurement Φ . The carrier phase measurement can be represented as

$$(4) \quad \Phi = \|r-R\| + c(dt-dT) + \lambda N + \varepsilon ,$$

where:

- r : is the unknown receiver antenna position vector at signal reception time,
- R : is the given satellite antenna position vector at signal transmission time,
- c : is the speed of light in vacuum,
- cdt : is the receiver clock range offset,
- cdT : is the satellite clock range offset,
- λ : is the carrier wavelength of the signal (L_1 or L_2),

and

N : is the carrier phase ambiguity (L_1 or L_2).

The term ϵ represents carrier phase measurement noise and biases such as satellite ephemeris errors, tropospheric and ionospheric delays, and ranging errors caused by multipath. In the processing of phase data it is not uncommon to difference the carrier phase measurements between satellites and between receivers to eliminate the satellite and receiver clock offsets. This gives the double difference carrier phase observation equation

$$(5) \quad DD\Phi = DD\|r-R\| + \lambda DDN + DD\epsilon ,$$

where DD stands for the double-difference operator. In this equation the ambiguity-term DDN is known to be *integer*-valued. For reasons of simplicity we will restrict ourselves in the following to the two-receiver situation. This however, has no effect on the general applicability of the proposed method. The method is namely independent of the number of receivers used, and is therefore also applicable in situations where more than two receivers are used. This is for instance the case when the GPS-data are simultaneously adjusted for in a network-mode.

Linearization of the observation equations with respect to the unknown parameters, and a collection of these linearized equations into a linear system of equations, gives

$$(6) \quad y = Aa + Bb + e ,$$

where:

- y : is the vector of observed minus computed double difference carrier phase measurements,
- a : is the vector of unknown integer double difference ambiguities,
- b : is the vector that contains the increments of the unknown baseline components,
- A, B : are the design matrices for ambiguity terms and baseline components, and
- e : is the vector of unmodelled errors.

This system of observation equations is taken as point of departure for computing estimates of the unknown parameters a and b . Our estimation criterion will be based on the principle of least-squares. From a statistical viewpoint this choice is motivated by the fact that, in the absence of modelling errors, properly weighted linear least-squares estimators are identical to unbiased minimum variance estimators. Furthermore, these estimators are also maximum

likelihood estimators if the assumption of normality holds for the phase observables. In the following we will assume that the bias-terms in e are either corrected for or sufficiently small to be neglected.

2.2 Integer Least-Squares Estimation

The least-squares criterion for solving the linear(ized) system of observation equations (6) reads

$$(7) \quad \min_{a,b} \|y - Aa - Bb\|_{Q_y}^2 \text{ with } a \in Z^n, b \in R^3 ,$$

where $\|\cdot\|_{Q_y}^2 = (\cdot)^* Q_y^{-1}(\cdot)$, Q_y is the variance-covariance matrix of the double-difference carrier phase observables, Z^n is the n -dimensional space of integers and R^3 is the 3-dimensional space of reals. The minimizers of (7) will be denoted as respectively $\check{a} \in Z^n$ and $\check{b} \in R^3$. Note that (7) is a *constrained* least-squares problem. This is due to the presence of the *integer*-constraint $a \in Z^n$. The minimization problem (7) will therefore be referred to as an *integer* least-squares problem.

The quadratic objective function of the above integer least-squares problem can be decomposed into a sum of three squares,

$$(8) \quad \|y - Aa - Bb\|_{Q_y}^2 = \|\hat{e}\|_{Q_y}^2 + \|\hat{b}|a - b\|_{Q_{\hat{b}|a}}^2 + \|\hat{a} - a\|_{Q_a}^2 ,$$

where: \hat{a} is the real-valued, unconstrained least-squares ambiguity vector, having $Q_{\hat{a}}$ as variance-covariance matrix; $\hat{b}|a$ is the *conditional* least-squares baseline vector, conditioned on a , having $Q_{\hat{b}|a}$ as variance-covariance matrix; and \hat{e} is the unconstrained least-squares residual vector. From the above decomposition follows, that the last two squares vanish identically, if the objective function (8) would be minimized as function of $a \in R^n$ and $b \in R^3$. Hence, the minimizers would then be given by $\hat{a} \in R^n$ and $\hat{b} \in R^3$, and the minimum of the objective function would be given by the squared norm of the least-squares residual vector \hat{e} . In our case however, the objective function needs to be minimized as function of $a \in Z^n$ and $b \in R^3$. In that case, only the second square vanishes identically and the minimizers are given by $\check{a} \in Z^n$ and $\check{b} = \hat{b}|a \in R^3$. The corresponding minimum of the objective function is then given by $\|\hat{e}\|_{Q_y}^2 + \|\hat{a} - \check{a}\|_{Q_a}^2$.

The above shows that the integer least-squares problem (7) may be solved in two steps. The first step consists then of solving (7) with Z^n replaced by R^n . Hence, in the first step the integer-constraint is removed, reducing the problem to an ordinary unconstrained least-squares problem. As a result of this first step, real-valued estimates for both the ambiguities,

\hat{a} , and the baseline components, \hat{b} , are obtained, together with their corresponding variance-covariance matrices:

$$(9) \quad \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}, \begin{pmatrix} Q_{\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}} \end{pmatrix}.$$

This result forms then the input for the second step. In the second step one first solves for the vector of *integer* least-squares estimates of the ambiguities, \check{a} . It follows from solving the integer least-squares problem

$$(10) \quad \min_a (\hat{a} - a) Q_{\hat{a}}^{-1} (\hat{a} - a) \text{ with } a \in Z^n.$$

Once the integer least-squares ambiguity vector $\check{a} \in Z^n$ has been obtained, the residual $(\hat{a} - \check{a})$ is used to adjust the unconstrained baseline solution \hat{b} , to get $\check{b} = \hat{b}| \check{a}$. As a result, the final baseline solution is obtained as

$$(11) \quad \check{b} = \hat{b}| \check{a} = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} (\hat{a} - \check{a}).$$

The above step-wise approach agrees conceptually with the procedure that is usually followed in practice, when ambiguity-fixing is included in the baseline computations. The estimates \hat{a} and \hat{b} are sometimes referred to as the "float-solution", and the estimates \check{a} and \check{b} , being the solution of (7), are then referred to as the "fixed-solution".

Note that the minimization problem (10) may not have a unique solution. Although it is very unlikely that it has more than one solution, it is possible in principle that (10) has up to 2^n different integer minimizers. Still however, we will assume in the present sequel that (10) has one and only one solution. Our motivation for this assumption is based on the fact, that a reliable fixing of the ambiguities will not be considered feasible whenever the solution of (10) is non-unique.

The variance-covariance matrix of $\hat{b}|a$ is given as $Q_{\hat{b}|a} = Q_{\hat{b}} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} Q_{\hat{a}\hat{b}}$. In practice, this variance-covariance matrix is usually used to describe the precision of the final baseline solution \check{b} (approaches may differ, however, in the way this matrix is scaled). Here however, some remarks of caution are in order [27]. The structure of (11) shows, that \check{b} can be interpreted as a *conditional* least-squares estimate. That is, starting from the result of the first step, (9), \check{b} can be interpreted as the least-squares estimate of b , conditioned on the integer ambiguity-vector \check{a} . And consistent with this interpretation, the precision of \check{b} may indeed be described by the variance-covariance matrix $Q_{\hat{b}|a}$. When doing so, one should recognize however, that $Q_{\hat{b}|a}$ is interpreted as a

conditional variance-covariance matrix. Hence, in the light of predicting empirical outcomes of future experiments, the following meaning should be given to the formal variance-covariance matrix $Q_{\hat{b}|a}$. When the measurement experiment is repeated a sufficient number of times under similar circumstances, the matrix $Q_{\hat{b}|a}$ describes the spread one can expect in the various baseline solutions, when all baseline solutions are based on the *same* value for a . But this shows, since in practice each baseline determination will have its own integer estimate for the ambiguity vector $a \in Z^n$, that strictly speaking, $Q_{\hat{b}|a}$ does *not* describe what it should describe. In other words, $Q_{\hat{b}|a}$ is the variance-covariance matrix of $\hat{b}|a$, but not necessarily that of $\check{b} = \hat{b}| \check{a}$. The fact that the least-squares estimator of the ambiguity vector is integer, *does not* imply that it is nonstochastic. What is needed therefore, is the *unconditional* variance-covariance matrix of \check{b} . Hence, in order to obtain the theoretically correct variance-covariance matrix of \check{b} , the stochasticity of the integer estimator of the ambiguity vector should be taken into account when applying the error propagation law to (11). This is a nontrivial problem (the probability density function of \check{a} is of the discrete type) and one that has not yet been solved satisfactorily from a theoretical point of view. Fortunately, the practical relevance of this problem may be minor, in particular when a sound validation procedure has been used for the validation of \check{a} . One of the features of a proper validation procedure should namely be to verify whether or not sufficient probability mass is located at a single gridpoint of Z^n . And when this can be assured to a sufficient degree, the influence of the stochasticity of \check{a} on \check{b} will be small and matrix $Q_{\hat{b}|a}$ can be taken as a sufficient realistic measure for the precision of \check{b} .

It will be clear that $Q_{\hat{b}|a} < Q_{\hat{b}}$. When short observational timespans are used, we in fact have $Q_{\hat{b}|a} \ll Q_{\hat{b}}$. This can be explained as follows. Since GPS satellites are in very high altitude orbits, their relative positions with respect to the receiver change slowly, which implies in case of short observational timespans, that the ambiguities - when treated as being real-valued - become very poorly separable from the baseline coordinates. As a result, the precision with which the baseline can be estimated will be rather poor. However, when one explicitly aims at resolving for the integer-values of the ambiguities and assumes that their discrete probability density function is sufficiently peaked, the high-precision carrier-phase observables will start to act as if they were high-precision pseudo-range observables. As a result, the baseline coordinates become estimable with a comparable high precision and $Q_{\hat{b}} \equiv Q_{\hat{b}|a} \ll Q_{\hat{b}}$ holds true. The sole purpose of "ambiguity-fixing" is thus, to be able, via the

inclusion of the integer-constraint $a \in \mathbb{Z}^n$, to obtain a drastic improvement in the precision of the baseline solution. In this way, long observational timespans can be avoided, which otherwise would have been needed if the ambiguities were treated as being real-valued.

In the remaining of the sequel we will not be concerned directly with the baseline solution. Instead we will concentrate our attention on solving the ambiguity integer least-squares problem (10). And it is with the minimization of this constrained objective function, that the intricacy of the integer ambiguity estimation problem manifests itself.

2.3 Sequential Conditional Least-Squares Estimation

Due to the presence of the integer-constraint $a \in \mathbb{Z}^n$, there are unfortunately in general no standard techniques available for solving (10) as they are available for solving ordinary least-squares problems. As a consequence one has to resort to methods that in one way or another make use of a discrete search strategy for finding the integer minimizer of (10). However, before we start thinking of setting up such a search strategy, it helps if we ask ourselves the question what the structure of (10) must be in order to be able to apply the simplest of all integer estimation methods. Clearly, the simplest integer estimation method is "rounding to the nearest integer". And when applied to (10), this method will give the correct answer provided that the ambiguity variance-covariance matrix $Q_{\hat{a}}$ is diagonal, i.e. when all least-squares ambiguities are fully decorrelated. A diagonal $Q_{\hat{a}}$ implies namely that (10) reduces to the minimization of a sum of independent squares,

$$(12) \quad \underset{a_1, \dots, a_n \in \mathbb{Z}}{\text{minimize}} \sum_{i=1}^n (\hat{a}_i - a_i)^2 / \sigma_{\hat{a}(i,i)},$$

where $\sigma_{\hat{a}(i,i)}$ denotes the variance of the i th least-squares ambiguity. Hence, in that case we can work with n separate scalar integer least-squares problems. And the integer minimizers of each of these individual squares are then simply given by the integers nearest to \hat{a}_i . The conclusion reads therefore, that the ambiguity integer least-squares problem becomes trivial, when all least-squares ambiguities are fully decorrelated.

Unfortunately, in reality, the least-squares ambiguities are usually highly correlated and the variance-covariance matrix $Q_{\hat{a}}$ is far from being diagonal. Still however, it is possible to recover a sum-of-squares structure of the objective function if we diagonalize $Q_{\hat{a}}$. Not every diagonalization works however. What is needed in addition, is, that the diagonalization realizes, like in (12), that the individual

ambiguities can be assigned to the individual squares in the total sum-of-squares. This for instance, rules out a diagonalization based on an eigenvalue decomposition of the ambiguity variance-covariance matrix. In the same spirit of decomposition (8), we will therefore apply a conditional least-squares decomposition to the ambiguities. And this will be done on an ambiguity-by-ambiguity basis. Hence, a *sequential conditional* least-squares estimation will be applied. This implies that the first least-squares ambiguity \hat{a}_1 remains unchanged. The second least-squares ambiguity however, is replaced by its least-squares estimate, conditioned on the first ambiguity a_1 . This gives $\hat{a}_{2|1} = \hat{a}_2 - \sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-1} (\hat{a}_1 - a_1)$. Note that $\hat{a}_{2|1}$ is uncorrelated with a_1 . The third least-squares ambiguity is replaced by its least-squares estimate conditioned on the first two ambiguities a_1 and a_2 . This gives $\hat{a}_{3|2,1} = \hat{a}_3 - \sigma_{\hat{a}(3,1)} \sigma_{\hat{a}(1,1)}^{-1} (\hat{a}_1 - a_1) - \sigma_{\hat{a}(3,2|1)} \sigma_{\hat{a}(2|1,2|1)}^{-1} (\hat{a}_{2|1} - a_2)$. Note that $\hat{a}_{3|2,1}$ is uncorrelated with both $\hat{a}_{2|1}$ and \hat{a}_1 . By continuing in this way, we obtain for the i th step, using the shorthand notation $\hat{a}_{j|J}$ for $\hat{a}_{j|(j-1), \dots, 1}$,

$$(13) \quad \hat{a}_{i|I} = \hat{a}_i - \sum_{j=1}^{i-1} \sigma_{\hat{a}(i,j|J)} \sigma_{\hat{a}(j|J,j|J)}^{-1} (\hat{a}_{j|J} - a_j).$$

And $\hat{a}_{i|I}$ is uncorrelated with all $\hat{a}_{j|J}$ for $j=1, \dots, (i-1)$. It now follows from (13) that the ambiguity difference $(\hat{a}_i - a_i)$ can be written in terms of the differences $(\hat{a}_{j|J} - a_j)$, $j=1, \dots, i$, as $(\hat{a}_i - a_i) = (\hat{a}_{i|I} - a_i) + \sum_{j=1}^{i-1} \sigma_{\hat{a}(i,j|J)} \sigma_{\hat{a}(j|J,j|J)}^{-1} (\hat{a}_{j|J} - a_j)$. Hence, when this is written out in vector-matrix form, using the notation $\hat{d} = (\hat{a}_1, \hat{a}_{2|1}, \dots, \hat{a}_{n|N})^*$, and the error propagation law is applied, it follows, because of the fact that the conditional least-squares ambiguities are mutually uncorrelated, that

$$(14) \quad (\hat{a} - a) = L(\hat{d} - a) \text{ and } Q_{\hat{a}} = LDL^*,$$

where: $D = \text{diag}(\dots, \sigma_{\hat{a}(i|I,i|I)}, \dots)$ and $(L)_{ij} = 0$ for $1 \leq i < j \leq n$, and $(L)_{ii} = 1$ for $i=j$, and $(L)_{ij} = \sigma_{\hat{a}(i,j|J)} \sigma_{\hat{a}(j|J,j|J)}^{-1}$ for $1 \leq j < i \leq n$. The above matrix decomposition is well-known and is usually referred to as the LDL^* -decomposition [25]. With our "re-discovery" of the LDL^* -decomposition, we now can give a clear statistical interpretation to each of the entries of the lower triangular matrix L and to each of the entries of the diagonal matrix D . This interpretation will also be of help for the construction of the ambiguity transformation in the next chapter. Substitution of (14) into the objective function of (10) gives the desired sum-of-squares structure, and allows us to rewrite the integer least-squares problem as

$$(15) \quad \underset{a_1, \dots, a_n \in \mathbb{Z}}{\text{minimize}} \sum_{i=1}^n (\hat{a}_{i|I} - a_i)^2 / \sigma_{\hat{a}(i|I,i|I)}.$$

Note that the sum-of-squares of (15) reduces to that of (12), when all least-squares ambiguities would be fully decorrelated. Also note, that in case of (15), a simple

"rounding to the nearest integer" will now not necessarily give the correct integer minimizer to (10). This is due to the fact that $\hat{a}_{i|I}$ depends on a_j for $j=1,\dots,(i-1)$. But, the sum-of-squares structure of (15) still allows one to set up sharp bounds for the individual ambiguities. This is shown in the next section.

2.4 Search for the Integer Least-Squares Ambiguities

In order to solve (15) we will first restrict the solution space by replacing the space of integers, Z^n , by a smaller subset that can be enumerated. The idea is to use the objective function of (15) for introducing an ellipsoidal region in R^n , on the basis of which the search can be performed. This ellipsoidal ambiguity search space is determined by

$$(16) \quad \sum_{i=1}^n (\hat{a}_{i|I} - a_i)^2 / \sigma_{\hat{a}(i|I,i|I)}^{-1} \leq \chi^2.$$

This ellipsoidal region is centred at $\hat{a} \in R^n$, its orientation and elongation are governed by the ambiguity variance-covariance matrix Q_a , and its size can be controlled through the selection of the positive constant χ^2 . It will be assumed that the positive constant χ^2 has been chosen such that the region at least contains the sought for integer minimizer of (15) [24].

In order to discuss our search for the integer least-squares ambiguities, we will first consider the two-dimensional case. For $n=2$, the inequality (16) reduces to

$$(17) \quad (\hat{a}_1 - a_1)^2 / \sigma_{\hat{a}(1,1)}^{-1} + (\hat{a}_{2|1} - a_2)^2 / \sigma_{\hat{a}(2|1,2|1)}^{-1} \leq \chi^2.$$

This ellipsoidal region is shown in figure 1. Also shown in the figure are the rectangular box that encloses the ellipse and the line passing through the centre of the ellipse, (\hat{a}_1, \hat{a}_2) , having $(1, \sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-1})$ as direction vector. This line intersects the ellipse at the two points where the normal of the ellipse is directed along the a_1 -axis. Note that the point $(a_1, \hat{a}_{2|1})$ moves along this line when a_1 is varied. With (17) we are now in the position to construct the following two bounds for the two ambiguities a_1 and a_2 ,

$$(18) \quad \begin{cases} (\hat{a}_1 - a_1)^2 \leq \sigma_{\hat{a}(1,1)}^{-1} \chi^2 \\ (\hat{a}_{2|1} - a_2)^2 \leq \sigma_{\hat{a}(2|1,2|1)}^{-1} \lambda(a_1) \chi^2, \end{cases}$$

with $\lambda(a_1) = 1 - (\hat{a}_1 - a_1)^2 / \chi^2 \sigma_{\hat{a}(1,1)}^{-1}$. The corresponding intervals and their lengths are also shown in figure 1.

Our search for the integer least-squares ambiguities now proceeds as follows. First one selects an integer ambiguity

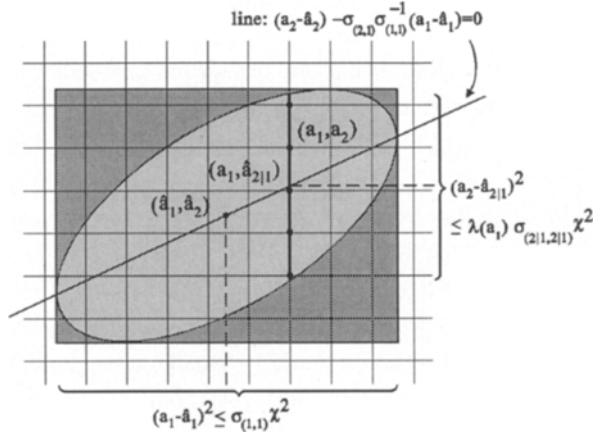


Figure 1: Ambiguity search space and bounds.

a_1 that satisfies the first bound of (18). It will be assumed that the two ambiguities are ordered such that one starts with the most precise least-squares ambiguity, i.e. that $\sigma_{\hat{a}(1,1)} \leq \sigma_{\hat{a}(2,2)}$. Then, based on this chosen integer ambiguity value a_1 , the conditional least-squares estimate $\hat{a}_{2|1}$ and scalar $\lambda(a_1)$ are computed. These values are then used to select an integer ambiguity a_2 that satisfies the second bound of (18). Since we aim at finding the integer minimizer, it is natural to choose the integer candidates in such a way that the individual squares in the sum-of-squares (17) are made as small as possible. This implies that a_2 should always be chosen as the integer nearest to $\hat{a}_{2|1}$. But remember that $\hat{a}_{2|1}$ depends on a_1 . For a_1 one first chooses the integer nearest to \hat{a}_1 . If one then fails to find an integer a_2 that satisfies the second bound, one restarts and chooses for a_1 the second nearest integer to \hat{a}_1 , and so on. Note that in this way, one is roughly following the direction of the line $(a_1, \hat{a}_{2|1})$, working with a_1 along the a_1 -axis from the inside of the ellipse, in an alternating fashion, towards the bounds of the ellipse. And this is continued until an admissible integer-pair (a_1, a_2) is found, i.e. until a gridpoint is found that lies inside the ellipse. Then a shrinking of the ellipse is applied, by applying an appropriate downscaling of χ^2 , after which one continues with the next and following nearest integers to \hat{a}_1 . This process is continued until one fails to find an admissible integer for a_1 . The last found integer-pair is then the sought for integer least-squares solution.

Generalizing the above to the n -dimensional case, results in the following bounds for the individual ambiguities,

$$(19) \quad (\hat{a}_{i|I} - a_i)^2 \leq \sigma_{\hat{a}(i|I,i|I)}^{-1} \lambda(a_i) \chi^2, \quad i=1,\dots,n,$$

with $\lambda(a_i) = (1 - \sum_{j=1}^{i-1} (\hat{a}_{j|J} - a_j)^2 / \chi^2 \sigma_{\hat{a}(j|J,j|J)}^{-1})$. These bounds are now used, quite analogous to the two dimensional case, for the search of the integer least-squares ambiguities. Note that

the bounds of (19) are sharper than those that would result when using the ellipsoidal planes of support for bounding the individual ambiguities. In that case, one would have: $(\hat{a}_i - a_i)^2 \leq \sigma_{\hat{a}(i,i)} \chi^2$ for $i=1,\dots,n$. And clearly, since $0 \leq \lambda(a_i) \leq 1$ and since a conditional variance is always smaller or at least as small as its unconditional counterpart, it follows that the bounds of (19) are always sharper or at least as sharp as $\sigma_{\hat{a}(i,i)} \chi^2$.

To conclude this section, we briefly remark on the practical implementation of the sequential conditional least-squares estimation. In the previous section, see (14), the correspondence was shown between the LDL^* -decomposition of $Q_{\hat{a}}$ and the sequential conditional least-squares estimation. It will be clear that the LDL^* -decomposition is easily constructed from the triangular Cholesky-decomposition. Many geodetic least-squares adjustment software-packages make use of the Cholesky-decomposition, but it is usually the Cholesky-decomposition of the normal matrix, $Q_{\hat{a}}^{-1}$, that is standard available and not the Cholesky-decomposition of $Q_{\hat{a}}$. This however, does not pose any difficulty. From the Cholesky-decomposition of the normal matrix, the LDL^* -decomposition of $Q_{\hat{a}}^{-1}$ is easily constructed. Inversion gives then $Q_{\hat{a}} = L^{-*} D^{-1} L^{-1}$. This decomposition of $Q_{\hat{a}}$ is again unique. The difference with the decomposition of $Q_{\hat{a}}$ in (14) is however, that matrix L^{-*} in $Q_{\hat{a}} = L^{-*} D^{-1} L^{-1}$ is upper-triangular and not lower-triangular. Hence, this decomposition would follow if one would apply a "backward" triangular decomposition to $Q_{\hat{a}}$. That is, instead of starting with a_1 , one then starts with a_n . The conclusion reads therefore, that when one starts from a triangular decomposition of the normal matrix $Q_{\hat{a}}^{-1}$, bounds for the ambiguities can be constructed as easily as those of (19). The only difference would then be, that these bounds would correspond to a sequential conditional least-squares estimation that commences with a conditioning on the last ambiguity and stops with a conditioning on the first ambiguity.

2.5 The GPS Spectrum of Ambiguity Conditional Variances

In the previous section it was shown how we perform the search for the integer least-squares ambiguities. But as it was pointed out, it may happen that the search halts before a complete integer ambiguity candidate vector has been found. This occurs when the size of the bound is such that no candidate integer lies within the interval. The problem that the search for the integer least-squares ambiguities halts is a

very serious one in case of GPS carrier-phase processing, in particular when short observational timespans are used. As it will be explained, this problem is intrinsically related to the structure of the GPS carrier-phase model of observation equations and the chosen parametrization.

For a single baseline model, it can be shown that the spectrum of ambiguity conditional variances, $\sigma_{\hat{a}(i|I,i|I)}$ for $i=1,\dots,n$, has a large discontinuity when passing from the third conditional variance $\sigma_{\hat{a}(3|2,1 ; 3|2,1)}$ to the fourth conditional variance $\sigma_{\hat{a}(4|3,2,1 ; 4|3,2,1)}$. In order to explain this behaviour of the spectrum of conditional variances, the following example considers a synthetic 2-by-2 ambiguity variance-covariance matrix. The structure of this matrix is chosen such, that it resembles the structure of the actual n -by- n ambiguity variance-covariance matrices.

Example 1

Let the variance-covariance matrix of the two least-squares ambiguities \hat{a}_1 and \hat{a}_2 be given as

$$(20) \quad \begin{pmatrix} \sigma_{\hat{a}(1,1)} & \sigma_{\hat{a}(1,2)} \\ \sigma_{\hat{a}(2,1)} & \sigma_{\hat{a}(2,2)} \end{pmatrix} = \sigma^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}^T.$$

It will be assumed that

$$(21) \quad \sigma^2 \ll \beta_1^2, \beta_2^2; \beta_1^2 \equiv \beta_2^2.$$

Note that the above 2-by-2 matrix is given as the sum of a scaled rank-2 matrix and a rank-1 matrix. And because of (21), the entries of the rank-2 matrix are very much smaller than the entries of the rank-1 matrix.

First we consider the correlation between the two ambiguities \hat{a}_1 and \hat{a}_2 . It follows from (20) that the square of the correlation coefficient is given as

$$(22) \quad \rho^2 = ((1 + \sigma^2/\beta_1^2)(1 + \sigma^2/\beta_2^2))^{-1}.$$

Together with (21) this shows that $\rho^2 \approx 1$. Hence, the two ambiguities are very heavily correlated. As a consequence of this extreme correlation, one will observe a large discontinuity in the conditional variances. To show this, consider the variance $\sigma_{\hat{a}(1,1)}$ and the conditional variance $\sigma_{\hat{a}(2|1,2|1)}$. It follows from (20) that

$$(23) \quad \sigma_{\hat{a}(1,1)} = \sigma^2 + \beta_1^2; \sigma_{\hat{a}(2|1,2|1)} = \sigma^2 + \beta_2^2 \frac{\sigma^2/\beta_1^2}{1 + \sigma^2/\beta_1^2}.$$

Together with (21) this shows that $\sigma_{\hat{a}(2|1,2|1)} \ll \sigma_{\hat{a}(1,1)}$.

Hence, there is a tremendous drop in value when one goes from the variance of the first ambiguity to the conditional variance of the second ambiguity. With β_1^2 sufficiently large, we approximately have $\sigma_{\hat{a}(1,1)} \approx \beta_1^2$ and $\sigma_{\hat{a}(2|1,2|1)} \approx 2\sigma^2$. The important implication of this result for the search of the integer least-squares ambiguities, is the following. When $\sigma_{\hat{a}(1,1)}$ is large and $\sigma_{\hat{a}(2|1,2|1)}$ is extremely small, the problem of search-halting will be significant. A large $\sigma_{\hat{a}(1,1)}$ implies namely, that the first bound of (18) will be rather loose. Quite a number of integers will therefore satisfy this first bound. This on its turn implies, when we go to the second bound, which is very tight due to $\sigma_{\hat{a}(2|1,2|1)} \ll \sigma_{\hat{a}(1,1)}$, that we have a high likelihood of not being able to find an integer that satisfies this second bound. The potential of halting is therefore very significant when one goes from the first to the second bound. As a consequence a large number of trials are required, before one is able to find a candidate integer-pair. \square

The above two-dimensional example has shown, how the search is affected when the conditional variances show a large discontinuity. This effect is also typically experienced in the actual search for the integer least-squares ambiguities. The structure of (20) resembles namely the structure of the actual ambiguity variance-covariance matrices. In case of the single baseline model, the actual n -by- n ambiguity variance-covariance matrices can also be written as the sum of two matrices. The first matrix in this sum is then of rank n , but contains very small entries because of the high precision with which the carrier-phases can be observed. The second matrix in the sum will be of rank 3, and contains entries which are relatively large due to the poor precision with which the baseline can be estimated when the ambiguities are still estimated as reals. As a result of this structure, the actual spectrum of ambiguity conditional variances will generally show a large discontinuity after the third variance:

$$(24) \quad \sigma_{\hat{a}(j|I,j|J)} \ll \sigma_{\hat{a}(i|I,i|I)} \quad i=1,2,3, \quad j=4,\dots,n.$$

This implies, since the first three bounds of (19) will be rather loose and the remaining bounds will be very tight due to the discontinuity, that the potential of halting will be significant when one passes from the third bound to the fourth bound. The following example gives a representative illustration of the characteristics of the spectrum of ambiguity conditional variances.

Example 2

This example is based on a 7 satellite configuration using dual frequency carrier-phase data for an observational

timespan of two seconds. Figure 2 shows the spectrum of conditional standard deviations expressed in cycles. Note the logarithmic scale along the vertical axis.

Since 7 satellites were observed on both frequencies, we have twelve double-difference ambiguities and therefore also

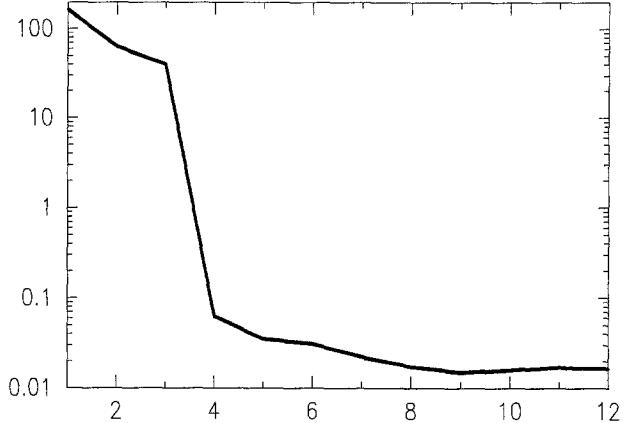


Figure 2: A GPS spectrum of conditional standard deviations.

twelve conditional standard deviations. The figure clearly shows the tremendous drop in value when passing from the third to the fourth standard deviation. There are three large conditional standard deviations and nine extremely small ones. The nine small conditional standard deviations are due to the presence of satellite redundancy and the fact that dual frequency carrier-phase data is observed. And it is because of the large values of the first three conditional standard deviations and the extreme small values for the remaining conditional standard deviations, that the search for the integer least-squares ambiguities will be rather timeconsuming. \square

It will be clear that the efficiency of the search for the integer least-squares ambiguities could be improved considerably, if we would be able to eliminate the discontinuity in the spectrum and lower the values of the large conditional standard deviations. In the next chapter it will be shown how this can be achieved through a reparametrization of the ambiguities.

3. The Reparametrized Ambiguity Estimation Problem

This chapter is devoted to the elimination of the potential problem of halting. The original integer least-squares problem is reparametrized such that an equivalent formulation is obtained, but one that is easier to solve. As our search procedure of the previous chapter, also the reparametrizing ambiguity transformation makes use of the

sequential conditional least-squares estimation of the ambiguities. Although the aim is to fully decorrelate the ambiguities, true diagonality of the variance-covariance matrix will be difficult to reach. In section 3.1 it is shown that this is due to the fact that only a particular class of ambiguity transformations is admissible. Being restricted to this class, section 3.2 shows to what extent the ambiguities can be decorrelated for the two-dimensional case. Based on an integer approximation of the conditional least-squares transformation, a decorrelating two-dimensional ambiguity transformation is introduced. With this transformation it is guaranteed that the square of the correlation coefficient of the transformed ambiguities is less than or equal to 1/4. In section 3.3 the n -dimensional problem is discussed. The discontinuity in the GPS-spectrum of conditional variances is flattened through a repeated use of the two-dimensional ambiguity transformation. By removing the discontinuity with the ambiguity transformation, transformed ambiguities are obtained that show a dramatic improvement in both precision and correlation. As a result the search for the transformed ambiguities can be performed in a highly efficient manner. In section 3.4 a measure for the gain in baseline precision due to ambiguity fixing is introduced. It is shown how this gain can be computed directly from the spectrum of conditional variances.

3.1 The Class of Admissible Ambiguity Transformations

Certain linear combinations of the GPS-observables play a prominent role in the problem of ambiguity fixing. For the purpose of ambiguity fixing one usually aims at those integer linear combinations that produce a phase observable which has a relatively long wavelength, a relatively low noise behaviour and a reasonable small ionospheric delay [20]. And these properties are indeed very beneficial to the ambiguity fixing process. At present however, all integer linear combinations that are considered, are restricted to the single-channel dual-frequency case. But this implies, that the relative receiver-satellite geometry, which is so emphatically present in the ambiguity variance-covariance matrix, is not taken into account in the current linear combinations. This observation suggests therefore, that it is of interest to consider multi-channel integer linear combinations, instead of only the current single-channel integer linear combinations.

In [19] the admissible class of multi-channel integer linear combinations of the ambiguities was identified. It was shown that a full-rank n -by- n transformation Z^* of the original ambiguity vector a to a new ambiguity vector z , giving $z =$

Z^*a , is admissible if and only if Z^* is *volume preserving* and has entries which all are *integer*. Typical examples of transformations that fall in this class are the identity matrix and the permutation matrices. But also all ambiguity transformations that change the choice of reference satellite in the double-difference ambiguities, are admissible. And on the single-channel level, the same holds for the transformation from the L_1 and L_2 ambiguity to the L_1 ambiguity and the wide-lane ambiguity. Note however, that the one-to-one single-channel transformation from the L_1 and L_2 ambiguity to the narrow-lane and wide-lane ambiguity is not admissible.

With the class of admissible ambiguity transformations identified, we can now try to use members from this class to aid the ambiguity fixing process. Let Z^* be an ambiguity transformation, which is used to transform the ambiguities as

$$(25) \quad z = Z^*a, \quad \hat{z} = Z^*\hat{a}, \quad Q_{\hat{z}} = Z^*Q_a Z.$$

The ambiguity integer least-squares problem (10) would then transform accordingly into the equivalent minimization problem

$$(26) \quad \min_z (\hat{z} - z)^* Q_{\hat{z}}^{-1} (\hat{z} - z) \text{ with } z \in Z^n.$$

Similarly, the original ambiguity searchspace (16) would transform into the new ambiguity searchspace

$$(27) \quad \sum_{i=1}^n (\hat{z}_{i|I} - z_i)^2 / \sigma_{\hat{z}(i|I, i|I)} \leq \chi^2.$$

Note that this transformed searchspace not only has a volume which is identical to the volume of the original searchspace, but it also has, as it should be, the same number of candidate gridpoints. Based on the transformed searchspace (27), the search for the integer minimizer \check{z} of (26), can now be performed in exactly the same way as it has been described in section 2.4 for the original searchspace. And once the integer least-squares ambiguity \check{z} has been found, the integer minimizer \check{a} of (10) can be recovered from invoking $\check{a} = Z^{-1}\check{z}$. The final baseline solution \check{b} follows then from (11). Alternatively, one could also use

$$(28) \quad \check{b} = \hat{b} - Q_{\hat{b}\check{z}} Q_{\hat{z}}^{-1} (\hat{z} - \check{z})$$

to obtain the final baseline solution.

In order to have any use for the ambiguity transformation Z^* , we should aim at finding a transformation that makes the transformed integer least-squares problem (26) easier to solve than the original problem (10). Clearly, the ideal situation would be, to have a transformation Z^* that allows for a full

decorrelation of the ambiguities. In that case, $Q_{\hat{z}}$ is diagonal and (26) can simply be solved by rounding the entries of \hat{z} to their nearest integer. Unfortunately however, the restrictions on Z^* do generally not allow for a complete diagonalization of the variance-covariance matrix. For instance, the choice where Z contains the (normalized) eigenvalues of $Q_{\hat{a}}$ is not allowed. Although this transformation preserves volume, it generally does not have entries which all are integer. Also a diagonalization based on $Z^* = L^{-1}$, with L being the triangular factor of $Q_{\hat{a}}$, is not admissible. Again, although L is volume-preserving, its non-zero off-diagonal entries will generally not be integer. This shows that in terms of diagonality, one will have to be content with a somewhat less perfect result. Nevertheless a decrease in correlation, although not complete, will already be very helpful, since it would close the existing gap between $\sigma_{\hat{a}(1,1)}$ and $\sigma_{\hat{a}(2,1,2|1)}$ of example 1. In the next section it will be shown to what extent the ambiguities can be decorrelated in the two-dimensional case.

3.2 A 2D-Decorrelating Ambiguity Transformation

In order to answer the question as to how to construct the ambiguity transformation Z^* , we first consider the problem in two dimensions. Let the ambiguities and their variance-covariance matrix be given as

$$(29) \quad \hat{a} = \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} \text{ and } Q_{\hat{a}} = \begin{pmatrix} \sigma_{\hat{a}(1,1)} & \sigma_{\hat{a}(1,2)} \\ \sigma_{\hat{a}(2,1)} & \sigma_{\hat{a}(2,2)} \end{pmatrix}.$$

From section 2.3 we already know, that the *conditional least-squares* based transformation returns least-squares ambiguities that are fully decorrelated. When (13) is written in vector-matrix form, we obtain for the two-dimensional case, the transformation

$$(30) \quad \begin{pmatrix} \hat{a}_1 \\ \hat{a}_{2|1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-1} & 1 \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix}.$$

Since we are studying the effect of transformations on $Q_{\hat{a}}$, we have for reasons of convenience skipped the elements a_1 and a_2 in the above transformation. Note, that this transformation not only decorrelates, but, in line with the correspondence between linear least-squares estimation and best linear unbiased estimation, also returns $\hat{a}_{2|1}$ as the element which has the best precision of all linear unbiased functions of \hat{a}_1 and \hat{a}_2 . Also note, that the above transformation is volume-preserving. Hence, the only condition that prohibits the above transformation from becoming an admissible ambiguity transformation is, that not

all entries of the transformation are integer. But, this shortcoming is easily repaired. We simply approximate the above transformation by replacing $\sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-1}$ by $[\sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-1}]$, where $[.]$ stands for "rounding to the nearest integer". This gives

$$(31) \quad \begin{pmatrix} \hat{a}_1 \\ \hat{a}_{2'} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -[\sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-1}] & 1 \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix}.$$

The volume-preserving property is retained by this integer approximation. But the full decorrelation property is of course not retained. What remains to be verified therefore, is whether the above transformation still allows us to reduce the correlation between the ambiguities. In order to verify this, we first assume, as was done in our search algorithm of section 2.4, that the two ambiguities are ordered such that $\sigma_{\hat{a}(1,1)} \leq \sigma_{\hat{a}(2,2)}$. This implies that $|\sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-1}| \geq |\sigma_{\hat{a}(1,2)} \sigma_{\hat{a}(2,2)}^{-1}|$. We also assume that $\sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-1} \notin (-1/2, +1/2]$. Because, otherwise we would have $[\sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-1}] = 0$, which would mean that the ambiguity transformation (31) reduces to the trivial identity transformation. Also an interchange of the two ambiguities would then not help, because if $[\sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-1}] = 0$, then also $[\sigma_{\hat{a}(1,2)} \sigma_{\hat{a}(2,2)}^{-1}] = 0$, since $|\sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-1}| \geq |\sigma_{\hat{a}(1,2)} \sigma_{\hat{a}(2,2)}^{-1}|$.

It follows from the volume-preserving property of the ambiguity transformation that the determinant of the variance-covariance matrix of the original ambiguities \hat{a}_1 and \hat{a}_2 is identical to the determinant of the variance-covariance matrix of the transformed ambiguities \hat{a}_1 and $\hat{a}_{2'}$. We therefore have the following equality of determinants

$$(32) \quad \sigma_{\hat{a}(1,1)} \sigma_{\hat{a}(2,2)} (1 - \rho^2) = \sigma_{\hat{a}(1,1)} \sigma_{\hat{a}(2,2)} (1 - \rho'^2).$$

From this equality we conclude that the correlation decreases whenever the variance of the transformed ambiguity $\hat{a}_{2'}$ is smaller than the variance of the original ambiguity \hat{a}_2 . Hence, by decreasing the variance we automatically reach a decrease in correlation, $\rho'^2 < \rho^2$, and vice versa. Note, that this coupling between variance and correlation is a direct consequence of the volume-preserving property of the ambiguity transformation. We already know that the conditional least-squares estimate $\hat{a}_{2|1}$ has the smallest variance attainable. What remains to be shown is therefore whether its integer approximation $\hat{a}_{2'}$ also has a variance that is smaller than the variance of \hat{a}_2 . In order to show this, we express $(\hat{a}_1, \hat{a}_{2'})$ in terms of $(\hat{a}_1, \hat{a}_{2|1})$. With (31) and the inverse of (30), this gives

$$(33) \quad \begin{pmatrix} \hat{a}_1 \\ \hat{a}_{2'} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \epsilon & 1 \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_{2|1} \end{pmatrix}, \text{ with } |\epsilon| \leq 1/2.$$

Application of the error propagation law gives then

$$(34) \quad \begin{cases} \sigma_{\hat{a}(1,2)} = \epsilon \sigma_{\hat{a}(1,1)} \\ \sigma_{\hat{a}(2',2')} = \epsilon^2 \sigma_{\hat{a}(1,1)} + (1 - \rho^2) \sigma_{\hat{a}(2,2)}. \end{cases}$$

For the second equation we may also write $\sigma_{\hat{a}(2',2')} = \sigma_{\hat{a}(2,2)} - (\sigma_{\hat{a}(2,2)}^2 \sigma_{\hat{a}(1,1)}^{-2} - \epsilon^2) \sigma_{\hat{a}(1,1)}$. This shows, since $\sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-2} \geq 1/4$ and $\epsilon^2 \leq 1/4$ that $\sigma_{\hat{a}(2',2')} \leq \sigma_{\hat{a}(2,2)}$, with the equality only when $\sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-1} = -1/2$. The conclusion reads therefore, that the ambiguity transformation (31) indeed allows us to decrease the correlation between the two ambiguities.

So far, we have only dealt with one of the two ambiguities, namely a_2 . But now that \hat{a}_2 has been transformed into $\hat{a}_{2'}$, we could think of interchanging the role of the two ambiguities and proceed with transforming the first ambiguity \hat{a}_1 into $\hat{a}_{1'}$. The corresponding ambiguity transformation reads then

$$(35) \quad \begin{pmatrix} \hat{a}_{1'} \\ \hat{a}_{2'} \end{pmatrix} = \begin{pmatrix} 1 & -[\sigma_{\hat{a}(1,2)} \sigma_{\hat{a}(2',2')}^{-1}] \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_{2'} \end{pmatrix}.$$

This transformation makes only sense however, when $\sigma_{\hat{a}(1,2)} \sigma_{\hat{a}(2',2')}^{-1} \notin (-1/2, +1/2]$. It follows from the first equation of (34) that $\sigma_{\hat{a}(1,2)} \sigma_{\hat{a}(2',2')}^{-1} = \epsilon \sigma_{\hat{a}(1,1)} \sigma_{\hat{a}(2',2')}^{-1}$. We therefore have $\sigma_{\hat{a}(1,2)} \sigma_{\hat{a}(2',2')}^{-1} \in (-1/2, +1/2]$ and thus $[\sigma_{\hat{a}(1,2)} \sigma_{\hat{a}(2',2')}^{-1}] = 0$ whenever $\sigma_{\hat{a}(2',2')} \geq \sigma_{\hat{a}(1,1)}$. This shows that it only makes sense to continue when $\sigma_{\hat{a}(2',2')} < \sigma_{\hat{a}(1,1)}$. And when one continues, one again will be able to obtain a further decrease in the correlation between the two ambiguities. In fact, one can proceed in this way and construct a concatenated form of ambiguity transformations Z^* , in which each individual transformation contributes to a lessening of the correlation. The last ambiguity transformation in this string of transformations will then reduce to the trivial identity transformation. And when this happens, we would have for the transformed ambiguities, denoted as \hat{z}_1 and \hat{z}_2 , that $[\sigma_{\hat{z}(1,1)} \sigma_{\hat{z}(1,1)}^{-1}] = 0$ and $[\sigma_{\hat{z}(1,2)} \sigma_{\hat{z}(2,2)}^{-1}] = 0$, or that $|\sigma_{\hat{z}(1,1)} \sigma_{\hat{z}(1,1)}^{-1}| \leq 1/2$ and $|\sigma_{\hat{z}(1,2)} \sigma_{\hat{z}(2,2)}^{-1}| \leq 1/2$, from which it follows that

$$(36) \quad \rho_{\hat{z}}^2 = \frac{\sigma_{\hat{z}(1,2)}^2}{\sigma_{\hat{z}(1,1)} \sigma_{\hat{z}(2,2)}} \leq 1/4.$$

This is an important result, because it shows that we can transform the original ambiguities, which are usually highly correlated, into new ambiguities which are much less correlated and which also, because of the volume-preserving property, see (32), are more precise. From the above bound it follows, together with $\sigma_{\hat{z}(2|1,2|1)} = (1 - \rho_{\hat{z}}^2) \sigma_{\hat{z}(2,2)}$ and $\sigma_{\hat{z}(1,1)}$

$\leq \sigma_{\hat{z}(2,2)}$ that

$$(37) \quad \sigma_{\hat{z}(2|1,2|1)} \geq \frac{3}{4} \sigma_{\hat{z}(1,1)}.$$

Hence, the ambiguity transformation Z^* guarantees that the transformed conditional variance $\sigma_{\hat{z}(2|1,2|1)}$ will never be much smaller than the variance $\sigma_{\hat{z}(1,1)}$. But this implies, that Z^* indeed removes to a large extent any discontinuity present in the original conditional variances, $\sigma_{\hat{a}(2|1,2|1)} << \sigma_{\hat{a}(1,1)}$.

Geometrically, the above constructed sequence of ambiguity transformations can be given the following useful interpretation (see figure 3). Imagine the original two-

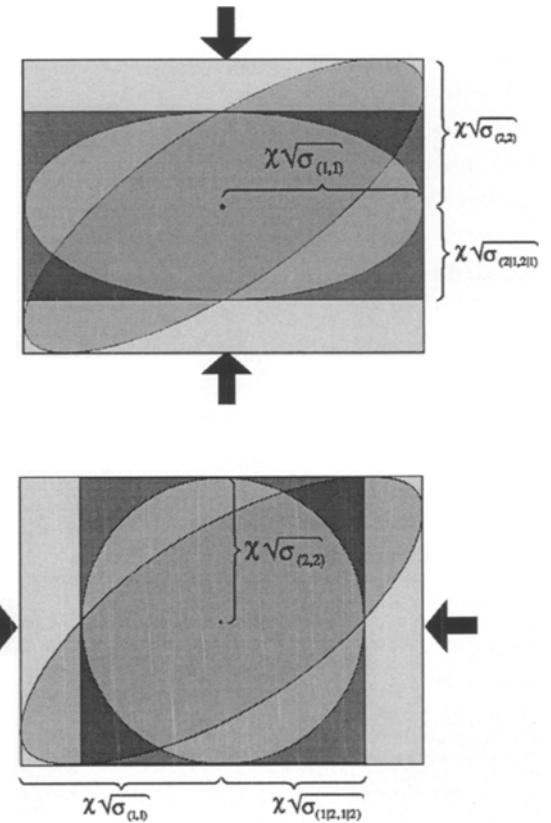


Figure 3: Decorrelating ambiguities by pushing tangents.

dimensional ambiguity searchspace centred at \hat{a} . A full decorrelation between the two ambiguities can be realized, if we push the two horizontal tangents of the ellipse from the $\pm(\sigma_{\hat{a}(2,2)} \chi^2)^{1/2}$ level towards the $\pm(\sigma_{\hat{a}(2|1,2|1)} \chi^2)^{1/2}$ level, while at the same time keeping fixed the volume (area) of the ellipse and the location of the two vertical tangents. Alternatively, one can also achieve a full decorrelation, if instead of the two horizontal tangents, the two vertical tangents are pushed from the $\pm(\sigma_{\hat{a}(1,1)} \chi^2)^{1/2}$ level towards the $\pm(\sigma_{\hat{a}(1|2,1|2)} \chi^2)^{1/2}$ level. Unfortunately however, both these transformations are not admissible, since it is not guaranteed that they result in integer ambiguities. Therefore, instead of using these two

transformations, we make use of their integer approximations. The first transformation (31) then pushes the two *horizontal* tangents of the ellipse from the $\pm(\sigma_{\hat{a}(2,2)}\chi^2)^{1/2}$ level towards the $\pm(\sigma_{\hat{a}(2',2')}\chi^2)^{1/2}$ level, while at the same time keeping fixed the volume of the ellipse and the location of the two vertical tangents of the ellipse. The second transformation (35) then pushes the two *vertical* tangents from the $\pm(\sigma_{\hat{a}(1,1)}\chi^2)^{1/2}$ level towards the $\pm(\sigma_{\hat{a}(1',1')}\chi^2)^{1/2}$ level, while at the same time keeping fixed the volume of the ellipse and the location of the two horizontal tangents. And this process is continued until the next transformation reduces to the trivial identity transformation. Since the volume of the ellipse is kept constant at all times, whereas the volume of the enclosing box is reduced in each step, it also follows that the ellipse is forced to become more sphere-like.

Based on the variance-covariance matrix of example 1, the following example illustrates how the ambiguity transformation Z^* succeeds in decorrelating the two ambiguities.

Example 3

When the values

$$(38) \quad \sigma = 0.2, \beta_1 = 5.0 \text{ and } \beta_2 = 6.0$$

are substituted into (20), we obtain

$$(39) \quad Q_{\hat{a}} = \begin{pmatrix} 25.04 & 30 \\ 30 & 36.04 \end{pmatrix}.$$

From this it follows that the two ambiguities are highly correlated and that the with (39) corresponding ambiguity search space is very elongated. The correlation coefficient $\rho_{\hat{a}}$ and elongation $e_{\hat{a}}$ read

$$(40) \quad \rho_{\hat{a}} = 0.999 \text{ and } e_{\hat{a}} = 39.064.$$

Elongation is measured as the ratio of the lengths of the major and minor semi-axes of the ambiguity search ellipse. If we now apply our method of ambiguity decorrelation, the corresponding ambiguity transformation becomes

$$(41) \quad Z^* = \begin{pmatrix} 6 & -5 \\ -1 & 1 \end{pmatrix}.$$

The variance-covariance matrix of the transformed ambiguities reads therefore

$$(42) \quad Q_{\hat{z}} = \begin{pmatrix} 2.44 & -0.44 \\ -0.44 & 1.08 \end{pmatrix}.$$

This result shows, that the transformed ambiguities are indeed less correlated, more precise, and also have an ambiguity search space that is less elongated:

$$(43) \quad \rho_{\hat{z}} = -0.271 \text{ and } e_{\hat{z}} = 1.645.$$

And because of this result, the gap in the conditional variances has been largely removed: $\sigma_{\hat{a}(1,1)} = 25.04, \sigma_{\hat{a}(2|1,2|1)} = 0.098$ versus $\sigma_{\hat{z}(1,1)} = 2.44, \sigma_{\hat{z}(2|1,2|1)} = 1.00$. Also note that the determinant of the variance-covariance matrix remained invariant under the ambiguity transformation. Both $Q_{\hat{a}}$ and $Q_{\hat{z}}$ have a determinant equal to 2.442. This invariance is due to the volume-preserving property of (41). \square

3.3 Flattening the Spectrum of Ambiguity Conditional Variances

As it was shown in section 2.5, it is the large discontinuity in the spectrum of ambiguity conditional variances, that forms a hindrance for the efficient search of the integer least-squares ambiguities. A flattened spectrum will therefore be very beneficial indeed for our search. Having a flattened spectrum automatically implies, when $n > 3$, that the three large variances in the spectrum get significantly smaller. This can be seen as follows. Since the ambiguity transformation is volume-preserving, the determinant of the ambiguity variance-covariance matrix remains invariant under the transformation. Hence, since the determinant of the lower triangular matrix L in (14) equals one, it follows that also the determinant of the diagonal matrix D and therefore the product of all conditional variances remains invariant under the transformation. Compare with (32) of section 3.2. But this implies, that by flattening the spectrum, the very small conditional variances in the spectrum automatically allow us to significantly decrease the three large conditional variances. And as a result of this, our search for the integer least-squares ambiguities would then already commence with a very tight bound, thus assuring with a very high likelihood that the first chosen integer candidate is indeed already the one sought.

The fact that the very small conditional variances in the spectrum allow us to significantly decrease the three large conditional variances, now also makes quite clear what role is played by *satellite redundancy* and *dual frequency* data. When both are absent, we have $n=3$. In that case, the

absence of very small conditional variances prohibits us to a high degree from "pulling down" the large variances in the spectrum. In case of satellite redundancy and/or dual frequency data however, we have $n > 3$. Now the presence of very small conditional variances does allow us to bring the large variances down to much smaller values. And the larger $n-3$ is, the more we are able to bring the flattened spectrum to a lower level.

In the previous section it was shown how to decorrelate the two ambiguities, thereby removing the gap between $\sigma_{\hat{a}(1,1)}$ and $\sigma_{\hat{a}(2|1,2|1)}$. The two-dimensional ambiguity transformation was constructed from a sequence of transformations of the following two types:

$$(44) \quad Z_1^* = \begin{pmatrix} 1 & 0 \\ z_{21} & 1 \end{pmatrix} \text{ and } Z_2^* = \begin{pmatrix} 1 & z_{12} \\ 0 & 1 \end{pmatrix},$$

in which z_{21} and z_{12} are appropriately chosen integers. To generalize this to the n -dimensional case, we first need to generalize these type of transformations accordingly. Here however, we are faced with a difficulty, since one can think of different generalizations all of which reduce to one of the above when $n=2$. One rather straightforward generalization follows however, when one considers the bounding of the triangular factor that is achieved by Z_1^* for $z_{21} = -[\sigma_{\hat{a}(2,1)} \sigma_{\hat{a}(1,1)}^{-1}]$. From an application of the error propagation law to (33) follows namely

$$\begin{pmatrix} \sigma_{\hat{a}(1,1)} & \sigma_{\hat{a}(1,2)} \\ \sigma_{\hat{a}(2',1)} & \sigma_{\hat{a}(2',2)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \epsilon & 1 \end{pmatrix} \begin{pmatrix} \sigma_{\hat{a}(1,1)} & 0 \\ 0 & \sigma_{\hat{a}(2|1,2|1)} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \epsilon & 1 \end{pmatrix},$$

(45)

with $|\epsilon| \leq 1/2$. In order to generalize this result to dimensions higher than two, we start from $Q_a = LDL^*$. Note that if the entries of L would be integer, then so would the entries of its inverse L^{-1} be. In fact, the inverse L^{-1} , being integer and volume-preserving, would then be the perfect candidate to diagonalize Q_a . One would then be able to fully decorrelate all ambiguities in just one step. This observation suggests, even though the entries of L will be non-integer in general, that we choose the n -by- n matrix Z_1^* as a lower triangular matrix, with integer entries and with one's on its diagonal. Matrix Z_1^* can then be constructed from subtracting suitable integer multiples of the first $(i-1)$ rows of L from row i of L for $i=n, \dots, 2$. Using this matrix Z_1^* , one obtains the triangular decomposition $Z_1^* Q_a Z_1 = (Z_1^* L) D (Z_1^* L)^*$, in which, in analogy with (45), the absolute values of all non-diagonal entries of $Z_1^* L$ are guaranteed to be bounded by a half. This bounding of the triangular factor $Z_1^* L$ implies, when the non-

diagonal entries of L are larger than a half in absolute value, that the precision of all but the first transformed ambiguity will be improved. And this improvement will be more noteworthy, the more the non-diagonal entries of L are decreased in size.

Although the above bounding of the triangular factor does allow for some decorrelation and some improvement of precision, one should note however, that all conditional variances remain invariant under the transformation Z_1^* . Hence, not only the variance $\sigma_{\hat{a}(1,1)}$ remains large, but the complete spectrum of ambiguity conditional variances, including its discontinuity, remains untouched. One should therefore not expect too much from the *single* transformation Z_1^* . This suggests, in analogy with the two-dimensional case, that one tries, in one way or another, to set up a *sequence* of n -by- n transformations. But here, the problem of dimension takes its revenge in the sense that one can try to pursue many different alternatives. Fortunately, there is no need to follow this line of thought, since excellent results can already be obtained when we keep ourselves to the two-dimensional ambiguity transformation of the previous section and apply it to the n -dimensional case as well.

Transformations of the type (44) are known as Gauss-transformations and they are considered to be the basic tools for zeroing entries in matrices [25]. In our case, due to the integer nature of z_{12} and z_{21} , they will be used to decrease the conditional correlations instead of zeroing them, thereby trying to flatten the spectrum of ambiguity conditional variances. In order to see how we can construct the overall n -dimensional ambiguity transformation on the basis of the two-dimensional transformation of the previous section, consider the discontinuity in the spectrum. The discontinuity is located at the two neighbouring conditional variances $\sigma_{\hat{a}(i|I,i|I)}$ and $\sigma_{\hat{a}(i+1|I+1,i+1|I+1)}$ for $i=3$. Hence, if we let $\hat{a}_{i|I}$ and $\hat{a}_{i+1|I}$, for $i=3$, play the role of our two ambiguities \hat{a}_1 and \hat{a}_2 of the previous section, we should be able to remove this discontinuity from the spectrum by using the decorrelating two-dimensional ambiguity transformation of the previous section. Thus, instead of applying the two-dimensional ambiguity transformation to the unconditional least-squares ambiguities, it is applied to the *conditional* least-squares ambiguities. This would then give, in analogy to (37),

$$(46) \quad \sigma_{\hat{a}(i+1|I+1,i+1|I+1)} \geq \frac{3}{4} \sigma_{\hat{a}(i|I,i|I)},$$

for $i=3$. Note, that as a consequence of the fact that we are transforming conditional least-squares ambiguities, the other conditional variances remain unaffected by this

transformation. Now, in order to construct and apply the two-dimensional transformation, we need the variance-covariance matrix of $\hat{a}_{i|l}$ and $\hat{a}_{i+1|l}$. In addition, we also need to know how the transformation affects the variance-covariance matrix $Q_{\hat{a}}$. We will therefore again make use of the sequential conditional least-squares decomposition (14).

To consider what happens to (14) when the two-dimensional ambiguity transformation is applied to \hat{a}_i and \hat{a}_{i+1} , we partition the lower triangular matrix L and diagonal matrix D of $Q_{\hat{a}} = LDL^*$ as

$$(47) \quad L = \begin{pmatrix} L_{11} & & \\ L_{21} & L_{22} & \\ L_{31} & L_{32} & L_{33} \end{pmatrix} \text{ and } D = \begin{pmatrix} D_{11} & & \\ & D_{22} & \\ & & D_{33} \end{pmatrix},$$

in which L_{11} is of order $(i-1)$, L_{22} is of order 2 and L_{33} is of order $(n-i)$. The dimensions of the other submatrices in L and D are determined accordingly. Note that $(L_{21}D_{11}L_{21}^* + L_{22}D_{22}L_{22}^*)$ is the variance-covariance matrix of the least-squares ambiguities \hat{a}_i and \hat{a}_{i+1} , whereas $L_{22}D_{22}L_{22}^*$ is the variance-covariance matrix of the conditional least-squares ambiguities $\hat{a}_{i|l}$ and $\hat{a}_{i+1|l}$. It is this last matrix that is now used for the construction of the two-dimensional ambiguity transformation. Hence, the conditional variance-covariance matrix $L_{22}D_{22}L_{22}^*$ now plays the role of the variance-covariance matrix in (29). The two-dimensional ambiguity transformation will be denoted as Z_{22}^* . With Z_{22}^* determined from $L_{22}D_{22}L_{22}^*$, we can Z_{22}^* -transform the unconditional least-squares ambiguities \hat{a}_i and \hat{a}_{i+1} , while leaving the other least-squares ambiguities unchanged. As a result we obtain the transformed ambiguity vector $\hat{a}' = (\hat{a}_1, \dots, \hat{a}_{i-1}, \hat{z}_i, \hat{z}_{i+1}, \hat{a}_{i+2}, \dots, \hat{a}_n)^*$, with corresponding variance-covariance matrix $Q_{\hat{a}'}$. This matrix will have a new triangular-decomposition $Q_{\hat{a}'} = L'D'L'$, with

$$(48) \quad L' = \begin{pmatrix} L_{11} & & \\ \bar{L}_{21} & \bar{L}_{22} & \\ L_{31} & \bar{L}_{32} & L_{33} \end{pmatrix} \text{ and } D' = \begin{pmatrix} D_{11} & & \\ & \bar{D}_{22} & \\ & & D_{33} \end{pmatrix},$$

and where

$$(49) \quad \begin{cases} \bar{L}_{21} = Z_{22}^* L_{21} \\ \bar{L}_{22} \bar{D}_{22} \bar{L}_{22}^* = Z_{22}^* (L_{22} D_{22} L_{22}^*) Z_{22} \\ \bar{L}_{32} = L_{32} (Z_{22}^* L_{22})^{-1} L_{22}. \end{cases}$$

Note that the transformation of \hat{a}_i and \hat{a}_{i+1} with Z_{22}^* , only has an effect on the i th and $(i+1)$ th row and column of L and D .

Hence, the recovery of the triangular decomposition after the two-dimensional transformation has been applied, is rather straightforward. This is also understandable, considering the properties of sequential conditional least-squares estimation. The two two-dimensional matrices \bar{L}_{22} and \bar{D}_{22} follow from the unique triangular decomposition of $Z_{22}^* (L_{22} D_{22} L_{22}^*) Z_{22}$. Matrix \bar{L}_{22} is lower triangular, having ones on its diagonal and $\sigma_{\hat{z}(i+1,l|l)} \sigma_{\hat{z}(i,l|l)}^{-1}$ as its off-diagonal entry. This off-diagonal entry is in absolute value less than or equal to 1/2. Compare with the triangular factor of (45). The diagonal matrix \bar{D}_{22} has the conditional variances $\sigma_{\hat{z}(i,l|l)}$ and $\sigma_{\hat{z}(i+1,l+1,i+1|l+1)}$ as its entries and they satisfy the inequality (46).

The above result implies for $i=3$, that we are able to close the large gap between the third and fourth conditional variance in the spectrum. But of course, after the transformation has been applied, other, but smaller discontinuities emerge. For instance, if the transformation has been applied for $i=3$, then $\sigma_{\hat{z}(3|2,1;3|2,1)} < \sigma_{\hat{a}(2|1,2|1)}$ and $\sigma_{\hat{a}(5|4,..;1;5|4,..)} < \sigma_{\hat{z}(4|3,2,1;4|3,2,1)}$. But also they can be removed by applying the two-dimensional transformation. The idea is therefore to continue applying the transformation to pairs of neighbouring ambiguities until the complete spectrum of conditional variances is flattened and (46) holds true for all i . Once this has been completed, the n -dimensional ambiguity transformation Z^* is known and the original least-squares ambiguity vector \hat{a} can be transformed as $\hat{z} = Z^* \hat{a}$. Its variance-covariance matrix is given as $Q_{\hat{z}} = Z^* Q_{\hat{a}} Z$ and the entries of the diagonal matrix \bar{D} of its triangular decomposition $Q_{\hat{z}} = \bar{L} \bar{D} \bar{L}^*$ will then contain the transformed spectrum of conditional variances. The following example illustrates how the ambiguity transformation Z^* succeeds in flattening the spectrum of conditional variances.

Example 4

This example is based on the same data as example 2. Figure 4 shows both the conditional standard deviations of the original least-squares ambiguities \hat{a}_i , as well as the conditional standard deviations of the transformed least-squares ambiguities \hat{z}_i . We clearly observe a dramatic improvement in the spectrum. The discontinuity has been removed and the transformed conditional standard deviations are all of about the same small order. The dramatic decrease in value of the first three conditional standard deviations and their levelling with the remaining conditional standard deviations, now implies that our search for the integer least-squares ambiguities already commences with very tight bounds, thus assuring with a very high likelihood that the first chosen integer candidates are indeed the coordinates of

the sought for integer least-squares ambiguity vector $\check{\zeta}$.

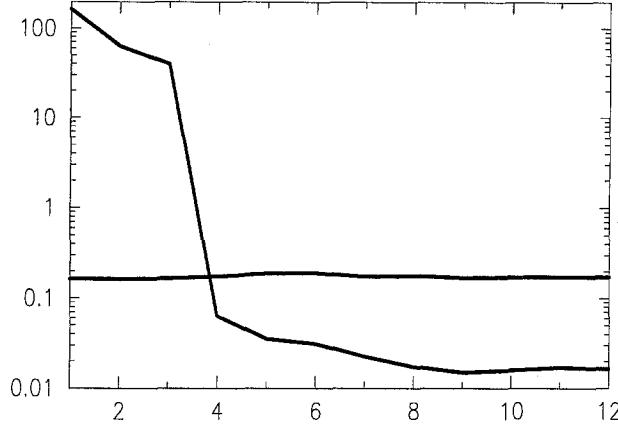


Figure 4: The original and transformed spectrum of conditional standard deviations.

Together with a bounding of the triangular factor as discussed above, the very low level of the transformed spectrum also allows us to assure that all unconditional variances of the transformed ambiguities are of the same small order. \square

To conclude this chapter, the following section will show how the spectrum of conditional variances can also be used to determine the gain in baseline precision that is achieved through the fixing of the ambiguities.

3.4. On the Spectrum and the Gain in Baseline Precision.

As it was remarked in section 2.2, the purpose of ambiguity fixing is to improve the precision of the baseline solution. This suggests that the precision of the least-squares ambiguities can be seen as an indicator of the amount of improvement one can achieve in the baseline precision. This indicator-property will be proven by showing the relation that exists between the determinants of Q_a , Q_b and $Q_{b|a}$.

For the conditional variance-covariance matrices $Q_{b|a}$ and $Q_{a|b}$ we may write: $Q_{b|a} = (I - Q_{ba}Q_a^{-1} \cdot Q_{ab}Q_b^{-1})Q_b$ and $Q_{a|b} = (I - Q_{ab}Q_b^{-1} \cdot Q_{ba}Q_a^{-1})Q_a$. From this follows, together with the determinant identity $\det(I - Q_{ba}Q_a^{-1} \cdot Q_{ab}Q_b^{-1}) = \det(I - Q_{ab}Q_b^{-1} \cdot Q_{ba}Q_a^{-1})$, that

$$(50) \quad \frac{\det(Q_b)}{\det(Q_{b|a})} = \frac{\det(Q_a)}{\det(Q_{a|b})}.$$

This relation is generally valid and it is independent of whether one is dealing with a single baseline model or a model where several baselines are adjusted for simultaneously. Note that $\det(Q_a)$ equals the product of all

ambiguity conditional variances $\sigma_{\check{\zeta}_{(i)|I,i|D}}$ for $i = 1, \dots, n$. Also note that one may replace the dependence on the original ambiguities in (50) by the transformed ambiguities.

The variance-covariance matrix $Q_{a|b}$ in (50) describes the precision of the least-squares double-difference ambiguities conditioned on a fixing of the baseline(s). Hence, $Q_{a|b}$ does not depend on the relative receiver-satellite geometry. Its determinant is therefore easily determined. For a single baseline model, where an s -number of satellites are tracked over a k -number of epochs, we have for the single-frequency and dual-frequency case

$$(51) \quad \begin{cases} L_1: \det(Q_{a|b}) = s(\sigma_1^2/k\lambda_1^2)^{s-1} \\ L_1, L_2: \det(Q_{a|b}) = s(\sigma_1^2/k\lambda_1^2)^{s-1}s(\sigma_2^2/k\lambda_2^2)^{s-1}, \end{cases}$$

in which σ_1^2 and σ_2^2 are the a-priori variances of the single-differenced carrier-phase observables on L_1 and L_2 . This result can now be used in combination with (50) to show how the gain in baseline precision is governed by the spectrum of conditional variances. For instance, for the single-frequency single baseline case we get

$$(52) \quad \frac{\det(Q_b)}{\det(Q_{b|a})} = [\prod_{i=1}^n (k\lambda_i^2 \sigma_{\check{\zeta}_{(i)|I,i|D}}/\sigma_1^2)]/(n+1).$$

Note that the square-root of the left-hand side of this expression equals the volume of the three-dimensional standard ellipsoid of the "floated" baseline, divided by the volume of the three-dimensional standard ellipsoid of the "fixed" baseline. Also note, since the left-hand side of (52) is independent of the chosen reference satellite in the double-difference ambiguities, that also the right-hand side must be independent of this choice. Hence, although the individual conditional variances are dependent on the choice of reference satellite, their product is not.

4. Summary and Concluding Remarks

In this contribution a method for the fast estimation of the integer least-squares ambiguities has been introduced. The basic idea that lies at the root of the method is that integer ambiguity estimation becomes trivial once the ambiguities are decorrelated. The approach followed is therefore to aim at decorrelating the least-squares ambiguities, while retaining their intrinsic integer nature. The method consists of two steps, a transformation step and a search step, both of which rely on the results of a sequential conditional least-squares ambiguity adjustment. The transformation-step starts from extending current usage of single-channel integer ambiguity

linear combinations to invertible multi-channel linear combinations. The importance of this extension is, that it provides significant leeway to influence the dependence of the ambiguity variance-covariance matrix on the designmatrix containing the receiver-satellite geometry. For short observational timespans, this dependence has been identified as a discontinuity in the spectrum of ambiguity conditional variances. As a consequence a direct search for the original integer least-squares ambiguities will be rather timeconsuming. Based on an integer approximation of a fully decorrelating transformation, transformed conditional least-squares ambiguities were recovered having an almost flattened spectrum. As a result, the search for the transformed integer least-squares ambiguities already commences with tight bounds, thus assuring that the solution can be found in a highly efficient manner.

As it was stressed in the introduction, our method is concerned with the *estimation* problem and not with the *validation* problem. Hence, a proper validation should still be used after the integer least-squares ambiguities have been identified. And since the quadratic forms which are usually used for validation purposes are invariant to our reparametrization (e.g. $\|\hat{a} - \check{a}\|_{Q_a}^2 = \|\hat{z} - \check{z}\|_{Q_z}^2$), any of the existing validation procedures can be used. Also note, that one can use, if so desired, ones own search procedure. But instead of applying it to the original ambiguities, it should then be applied to our transformed ambiguities.

We would also like to remark that the ambiguity transformation used, is completely determined by the variance-covariance matrix of the ambiguities. Even the a posteriori variance-factor need not be known. This stipulates that actual measurements are not needed for the transformation-step. Hence, the necessary computations for this step can be done in principle at the designing stage, prior to the actual measurement stage.

In sections 2.1 and 2.2 we have followed the customary practice of working with the double-difference version of the carrier-phase observation equations. This however, is not really necessary. One could as well work with the undifferenced version of the carrier-phase observation equations, as long as the original set of undifferenced noninteger ambiguities is reparametrized into a set of integer ambiguities plus a remaining set of noninteger parameters. In fact, this is precisely the approach used in our numerical software implementation. Similar approaches have been advocated in [13] and [30]. In sections 2.1 and 2.2 we have, for reasons of simplicity, also restricted ourselves to a

parametrization in which only the ambiguities and baseline components appear as parameters. But if needed and when estimable, additional parameters (e.g. different types of delays) can be included without affecting the principle of the approach. On the observational side, the method is also independent in principle, on whether code data are used or not, and on whether dual-frequency data are used or not. And when dual-frequency data are used, one could also apply the method to other type of ambiguities, such as the wide-lane or narrow-lane ambiguities.

Acknowledgement

The author acknowledges the assistance of ir. P.J. de Jonge and ir. C.C.J.M. Tiberius for preparing the numerical examples.

5. References

- [1] Hatch, R. (1986): Dynamic Differential GPS at the Centimeter Level. *Proceedings 4th International Geod. Symp. on Satellite Positioning*, Austin, Texas, 28 April - 2 May, pp. 1287-1298.
- [2] Hofmann-Wellenhof, B., B.W. Remondi (1988): The Antenna Exchange: one Aspect of High-Precision Kinematic Survey. Presented at the International GPS Workshop, *GPS Techniques Applied to Geodesy and Surveying*, Darmstadt, FRG, 10-13 April.
- [3] Remondi, B.W. (1984): *Using the Global Positioning System (GPS) Phase Observable for Relative Geodesy: Modelling, Processing, and Results*, Ph.D. dissertation, NOAA, Rockville, 360 pp..
- [4] Seeber, G. G. Wübbena (1989): Kinematic Positioning with Carrier Phases and "On the Way" Ambiguity Solution. *Proceedings 5th Int. Geod. Symp. on Satellite Positioning*. Las Cruces, New Mexico, March 1989.
- [5] Frei, E., G. Beutler (1990): Rapid Static Positioning Based on the Fast Ambiguity Resolution Approach FARA: Theory and First Results. *Manuscripta Geodaetica*, Vol. 15, No. 6, 1990.
- [6] Blewitt, G.(1990): A New Tool for Dense Monitoring of Crustal Strain: GPS Rapid Static Surveying. *EOS Transaction American Geophysical Union*, Vol. 71, No. 17, p. 483.
- [7] Euler, H.-J., C. Goad (1990): On Optimal Filtering of GPS Dual Frequency Observations without using Orbit Information. *Bulletin Geodesique*, Vol. 65, pp. 130-143.
- [8] Kleusberg, A. (1990): A Review of Kinematic and

- Static GPS Surveying Procedures. *Proceedings of the Second International Symposium on Precise Positioning with the global Positioning system*, Ottawa, Canada, September 3-7 1990, pp. 1102-1113.
- [9] Counselman, C.C., S.A. Gourevitch (1981): Miniature Interferometer Terminals for Earth Surveying: Ambiguity and Multipath with Global Positioning System. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. GE-19, No. 4, pp. 244-252.
- [10] Hatch, R. (1989): Ambiguity Resolution in the Fast Lane. Proceedings *ION GPS-89*, Colorado Springs, CO, 27-29 September, pp. 45-50.
- [11] Remondi, B.W. (1986): Performing Centimeter-Level Surveys in Seconds with GPS Carrier Phase: Initial Results. *Journal of Navigation*, Volume III, The Institute of Navigation.
- [12] Remondi, B.W. (1991): Pseudo-Kinematic GPS Results Using the Ambiguity Function Method, *Journal of Navigation*, Vol. 38, No. 1, pp. 17-36.
- [13] Blewitt, G. (1989): Carrier Phase Ambiguity Resolution for the Global Positioning System Applied to Geodetic Baselines up to 2000 km. *Journal of Geophysical Research*, Vol. 94, No. B8, pp. 10.187-10.203.
- [14] Hatch, R. (1991): Instantaneous Ambiguity Resolution. *Proceedings of IAG International Symposium 107 on Kinematic Systems in Geodesy, Surveying and Remote Sensing*, Sept. 10-13, 1990, Springer Verlag, New York, pp. 299-308.
- [15] Frei, E. (1991): *Rapid Differential Positioning with the Global Positioning System*. In: *Geodetic and Geophysical Studies in Switzerland*, Vol. 44.
- [16] Wübbena, G. (1991): *Zur Modellierung von GPS-Beobachtungen für die hochgenaue Positionsbestimmung*, Hannover, 1991.
- [17] Euler, H.-J., H. Landau (1992): Fast GPS Ambiguity Resolution On-The-Fly for Real-Time Applications. *Proceedings 6th Int. Geod. Symp. on Satellite Positioning*. Columbus, Ohio, 17-20 March 1992, pp. 650-729.
- [18] Mervart, L., G. Beutler, M. Rothacher, U. Wild (1994): Ambiguity Resolution Strategies using the Results of the International GPS Geodynamics Service (IGS), *Bulletin Geodesique*, 68: 29-38.
- [19] Teunissen, P.J.G. (1993): *The Invertible GPS Ambiguity Transformations*. Delft Geodetic Computing Centre (LGR), 9 p.
- [20] Wübbena, G. (1989): The GPS Adjustment Software Package - GEONAP - Concepts and Models. *Proceedings 5th Int. Geod. Symp. on Satellite Positioning*. Las Cruces, New Mexico, 13-17 March 1989, pp. 452-461.
- [21] Allison, T. (1991): Multi-Observable Processing Techniques for Precise Relative Positioning. *Proceedings ION GPS-91*. Albuquerque, New Mexico, 11-13 September 1991, pp. 715-725.
- [22] Goad, C. (1992) Robust Techniques for Determining GPS Phase Ambiguities. *Proceedings 6th Int. Geod. Symp. on Satellite Positioning*. Columbus, Ohio, 17-20 March 1992, pp. 245-254.
- [23] Cocard, C., A. Geiger (1992): Systematic Search for all Possible Widelanes. *Proceedings 6th Int. Geod. Symp. on Satellite Positioning*. Columbus, Ohio, 17-20 March 1992.
- [24] Teunissen, P.J.G. (1993): *Least-Squares Estimation of the Integer GPS Ambiguities*. Delft Geodetic Computing Centre (LGR), 16 p. Invited Lecture, Section IV Theory and Methodology. IAG General Meeting, Beijing, China, August 1993. Also in LGR-Series no. 6.
- [25] Golub, G.H. and C.F. Van Loan (1986): *Matrix Computations*. North Oxford Academic.
- [26] Lenstra, H.W. (1981): *Integer Programming with a Fixed Number of Variables*. University of Amsterdam, Dept. of Mathematics, report 81-03.
- [27] Teunissen, P.J.G. (1990): GPS op afstand bekijken (in Dutch). In: *Een halve eeuw in de goede richting*. Lustrumboek Snellius 1950-1990, pp. 215-233.
- [28] Euler, H.-J., B. Schaffrin (1991): On a Measure for the Discernability between Different Ambiguity Resolutions in the Static-Kinematic GPS-mode. *Proceedings of IAG International Symposium 107 on Kinematic Systems in Geodesy, Surveying and Remote Sensing*, Sept. 10-13, 1990, Springer Verlag, New York, pp. 285-295.
- [29] Betti, B., M. Crespi, and F. Sanso (1993): A Geometric Illustration of Ambiguity Resolution in GPS Theory and a Bayesian Approach, *Manuscripta Geodaetica* 18: 317-330.
- [30] Goad, C.C. (1985): Precise Relative Position Determination Using Global Positioning System Carrier Phase Measurements in a Nondifference Mode. *Proceedings of First Int. Symposium on Precise Positioning with the Global Positioning System*, April 15-19, 1985, Rockville, Maryland, Vol. 1, pp. 347-356, U.S. Department of Commerce, National Geodetic Information Center, NOAA.

ON THE INTEGER NORMAL DISTRIBUTION OF THE GPS AMBIGUITIES

P.J.G. Teunissen

Delft Geodetic Computing Centre (LGR)

Faculty of Geodesy

Delft University of Technology

Thijsseweg 11

2629 JA Delft, The Netherlands

Fax: ++ 31 15 278 3711

e-mail: P.J.G.Teunissen@geo.tudelft.nl

ABSTRACT

Carrier phase ambiguity resolution is the key to fast and high precision GPS kinematic positioning. Critical in the application of ambiguity resolution is the quality of the computed integer ambiguities. Unsuccessful ambiguity resolution, when passed unnoticed, will too often lead to unacceptable errors in the positioning results. In order to describe the quality of the integer ambiguities, their distributional properties need to be known. This contribution introduces the probability mass function of the integer least-squares ambiguities. This integer normal distribution is needed in order to infer objectively whether or not ambiguity resolution can be expected to be successful. Some of its properties are discussed. Attention is given in particular to the probability of correct integer estimation. Various diagnostic measures are presented for evaluating this probability.

Keywords: ambiguity resolution, integer normal distribution, probability of correct integer ambiguity estimation

1 Introduction

GPS ambiguity resolution is the process of resolving the unknown cycle ambiguities of the double-difference (DD) carrier phase data as integers [Hofmann-Wellenhof *et al.*, 1997], [Kleusberg and Teunissen, 1996], [Leick, 1995], [Strang and Borre, 1997]. Once resolved, one usually keeps the ambiguities fixed at their computed integer estimates. That is, all the results that depend on the ambiguity resolution process are usually evaluated as if the integer ambiguities were deterministic constants. From a *theoretical* point of view this is not correct. The estimated ambiguities, although integer, are still stochastic variates. They have been computed from the data and since the vector of observables is assumed to be random, also the integer ambiguity

estimator is a random vector. Conceptually we may write $\check{a} = F(y)$, where \check{a} denotes the integer ambiguity vector, y the vector of observables, and $F(\cdot)$ the mapping from the continuous vector of observables to the integer vector of ambiguities. Thus when y is random, \check{a} is random as well. As far as their randomness is concerned, the marked difference between y and \check{a} is that the probability distribution of y is continuous, whereas that of \check{a} is of the discrete type. Thus \check{a} has a probability *mass* function attached to it. This point was emphasized in [Teunissen, 1990] and highlighted again in [Teunissen, 1997a] as one of the pitfalls in the more classical approaches to ambiguity resolution.

Although theoretically not correct, it is possible that a treatment of the ambiguities as if they were deterministic constants is acceptable from a *practical* point of view. But in order to be able to judge whether this is feasible or not, one first needs to know more about the stochastic characteristics of the integer ambiguities. That is, one needs to know the probability mass function of the ambiguities and in addition, have means available of evaluating this distribution.

In this contribution we will assume that the data are normally distributed and we will use the integer least-squares criterion as estimation principle. This combination assures that a maximum likelihood solution is obtained. In addition, the solution is admissible and minimax as well. But other principles can also be used. One of the earliest is the ambiguity function method [Counselman and Gourevitch, 1981]. A Bayesian approach is used in [Betti *et al.*, 1993], and [Blewitt, 1989] and [Dong and Bock, 1989] used the bootstrapped estimator. The distribution of the bootstrapped estimator was studied in [Teunissen, 1998a].

This contribution is organized as follows. In Sect. 2 a brief review is given of the conceptual steps in solving the integer least-squares problem. In Sect. 3 we present the probability mass function of the integer least-squares ambiguities. This distribution is coined the integer normal distribution. In Sect. 4 we discuss some properties of this distribution. We show that it is symmetric and that its maximum equals the probability of correct integer estimation. We also show that the integer least-squares ambiguities are unbiased and that their precision bounds the probability of correct integer estimation. In Sect. 5 finally, we present some approaches for evaluating the integer normal distribution. Although an exact evaluation is rather difficult, approximations and bounds are given for the probability of correct integer estimation. Some of these approaches were already discussed in [Teunissen, 1997a-c]. The first approach is based on simulating the probability mass function. This is possible, since the shape of distribution is independent of the unknown integer ambiguities. For the second approach we make use of the bootstrapped ambiguity estimator. Its probability of correct integer estimation is easily computed. In the third approach we use eigenvalues to bound the variance matrix of the integer least-squares ambiguities. Finally in the last approach, the minimum grid point distance, as measured by the metric of the ambiguity variance matrix, is used to obtain bounds on the region of integration.

2 Integer least-squares estimation

As our point of the departure we will take the following system of linear(ized) observation equations

$$y = Aa + Bb + e \quad (1)$$

where y is the given data vector, a and b are the unknown parameter vectors and e is the noise vector. In principle all the GPS models can be cast in this frame of observation equations. The data vector will then usually consist of the 'observed minus computed' single- or dual-frequency double-differenced (DD) phase and/or pseudo range (code) observations, accumulated over all observation epochs. The entries of the n -vector a are the DD carrier phase ambiguities, expressed in units of cycles rather than range. They are known to be integer valued. The entries of the p -vector b consist of the remaining unknown parameters, such as for instance baseline components (coordinates) and possibly atmospheric delay parameters (troposphere, ionosphere).

When using the least-squares principle, the above system of observation equations can be solved by means of the minimization problem

$$\min_{a,b} (y - Aa - Bb)^T Q_y^{-1} (y - Aa - Bb) , a \in Z^n , b \in R^p \quad (2)$$

with Q_y the variance-covariance matrix of the observables, Z^n the n -dimensional space of integers and R^p the p -dimensional space of real numbers. This type of least-squares problem was first introduced in [Teunissen, 1993] and has been coined with the term '*integer least-squares*'. It is a nonstandard least-squares problem due to the integer constraints $a \in Z^n$.

Conceptually one can divide the computation of the solution to (2) into three different steps. In the first step one simply disregards the integer constraints on the ambiguities and performs a standard least-squares adjustment. As a result one obtains the (real-valued) least-squares estimates of a and b , together with their variance-covariance matrix

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} , \begin{bmatrix} Q_{\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}} \end{bmatrix} \quad (3)$$

This solution is often referred to as the 'float' solution. In the second step the 'float' ambiguity estimate \hat{a} and its variance-covariance matrix are used to compute the corresponding integer ambiguity estimate. This implies solving the minimization problem

$$\min_{a \in Z^n} (\hat{a} - a)^T Q_{\hat{a}}^{-1} (\hat{a} - a) \quad (4)$$

Its solution will be denoted as \check{a} . Finally in the third step, the integer ambiguities are used to correct the 'float' estimate \hat{b} . As a result one obtains the 'fixed' solution

$$\check{b} = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} (\hat{a} - \check{a}) \quad (5)$$

Thus in summary, the 'float' solutions \hat{a} and \hat{b} follow from solving (2) *without* the integer constraints, while the 'fixed' solutions \check{a} and \check{b} follow from solving (2) *with* the integer constraints included.

From a computational point of view, the most difficult part in the above three steps is the solution of (4). The difficulty lies in the fact that most GPS least-squares ambiguities are highly correlated. This is due to the short observation time spans used and the fact, that in case of GPS, the relative receiver-satellite geometry changes only slowly with time. As a consequence of the high correlation, the search needed to solve (4) becomes very time consuming. To remedy this situation, the least-squares ambiguity decorrelation adjustment (LAMBDA) was introduced

[Teunissen, 1993], [Jonge de and Tiberius, 1996]. By means of the decorrelation process of this method the original DD ambiguities are transformed to new ambiguities which have the property of being far more precise than the original ambiguities, see also the textbooks [Kleusberg and Teunissen, 1996], [Hofmann-Wellenhof et al., 1997], [Strang and Borre, 1997]. Examples of practical results can be found in e.g. [Tiberius and de Jonge, 1995], [Jonge de and Tiberius, 1996], [Jonkman, 1998].

The goal of the present contribution is not to discuss the intricacies of the above computational steps, but instead to present some of the distributional properties of the integer least-squares estimator \hat{a} . This is in particular of relevance for GPS ambiguity resolution. It is namely only through the distributional properties of \hat{a} that one will be able to objectively decide whether or not a successful 'fixing' of the integer ambiguities is likely to happen.

3 The distribution of the integer least-squares ambiguities

It will be assumed that our vector of observables is normally (Gaussian) distributed with mean $E\{y\} = Aa + Bb$ and dispersion $D\{y\} = Q_y$. As a result the (real-valued) least-squares ambiguities will be normally distributed too, with mean a and variance-covariance matrix $Q_{\hat{a}}$. Thus

$$y \sim N(Aa + Bb, Q_y) \implies \hat{a} \sim N(a, Q_{\hat{a}}) \quad (6)$$

Note that the mean of the real-valued least-squares estimator is an integer vector. The multivariate normal probability density function of \hat{a} reads therefore

$$p_{\hat{a}}(a) = \frac{1}{\sqrt{\det(Q_{\hat{a}})}(2\pi)^{\frac{1}{2}n}} \exp\left\{-\frac{1}{2} \|\hat{a} - a\|_{Q_{\hat{a}}}^2\right\} \quad (7)$$

with the squared weighted norm $\|\cdot\|_{Q_{\hat{a}}}^2 = (\cdot)^T Q_{\hat{a}}^{-1}(\cdot)$ and $a \in Z^n$.

Since the *integer* least-squares ambiguities follow from solving (4), they are functions of the stochastic *real valued* least-squares ambiguities and therefore stochastic variates themselves as well. Thus both \hat{a} and \check{a} are random vectors, although their distributions differ. The distribution of \hat{a} is a continuous one, whereas the distribution of \check{a} is of the discrete type. The mapping from the continuous random vector \hat{a} to the discrete random vector \check{a} is namely a many-to-one map. To see this, consider the following subset of R^n

$$S_z = \{x \in R^n \mid \|x - z\|_{Q_{\hat{a}}}^2 \leq \|x - u\|_{Q_{\hat{a}}}^2, \forall u \in Z^n\} \quad (8)$$

It contains all values of \hat{a} which are mapped to the single integer grid point $z \in Z^n$ when solving the integer least-squares problem (4). Thus the integer least-squares solution equals z when \hat{a} lies in S_z and vice versa. Hence

$$\hat{a} \in S_z \iff \check{a} = z \quad (9)$$

Thus the subset S_z acts as a *pull-in-region* for \hat{a} . That is, whenever \hat{a} lies in the subset S_z , it is pulled to z , being the centre grid point of the set. These pull-in-regions were studied by [Jonkman, 1998] for the geometry-free GPS model. Note that each integer grid point $z \in Z^n$ has such a pull-in-region assigned to it. Also note that these subsets are disjoint and that they

together cover R^n , i.e. $S_{z_i} \cap S_{z_j} = \{0\}$ for $i \neq j$ and $R^n = \cup_{z \in Z^n} S_z$. With this information it is now possible to formulate the probability *mass* function of \check{a} . It reads

$$P(\check{a} = z) = P(\hat{a} \in S_z) = \int_{S_z} \frac{1}{\sqrt{\det(Q_{\hat{a}})}(2\pi)^{\frac{1}{2}n}} \exp\left\{-\frac{1}{2} \|x - a\|_{Q_{\hat{a}}}^2\right\} dx, \quad a, z \in Z^n \quad (10)$$

The discrete distribution of the integer least-squares ambiguities follows thus from mapping the volume of the normal distribution over the subsets S_z to each of their centre grid points z . Since this distribution does not appear to have a name yet, see e.g. [Johnson *et al.*, 1994], we will coin it the *integer normal distribution*. In the following section we will present some of its properties.

4 Properties of the integer normal distribution

In this section we will focus on a few properties of the integer normal distribution. They are related to the mode and shape of the distribution.

4.1 The probability of correct integer estimation is largest

Of the infinite number of probability masses of (10), there is one which is particularly of importance for GPS ambiguity resolution. It is $P(\check{a} = a)$, the probability of correct integer estimation. This is the probability that \check{a} coincides with the true but unknown integer mean a . In order for the integer least-squares principle to make sense, the least we can ask of this principle is that the resulting probability of correct integer estimation is always larger than any of the probabilities of wrong integer estimation. We will now show that this indeed holds true. Thus

$$\max_{z \in Z^n} P(\check{a} = z) = P(\check{a} = a) \quad (11)$$

This will be proven by showing that

$$\int_{S_a} c \exp\left\{-\frac{1}{2} \|x - a\|_{Q_{\hat{a}}}^2\right\} dx \geq \int_{S_a} c \exp\left\{-\frac{1}{2} \|x - z\|_{Q_{\hat{a}}}^2\right\} dx = \int_{S_z} c \exp\left\{-\frac{1}{2} \|x - a\|_{Q_{\hat{a}}}^2\right\} dx$$

where c is the proportionality factor of the multivariate normal distribution. Note that the left-hand side equals $P(\check{a} = a)$ and the right-hand side $P(\check{a} = z)$. The inequality follows by noting that $\hat{a} \in S_a$ implies $\exp\left\{-\frac{1}{2} \|\hat{a} - a\|_{Q_{\hat{a}}}^2\right\} \geq \exp\left\{-\frac{1}{2} \|\hat{a} - z\|_{Q_{\hat{a}}}^2\right\}$, $\forall z \neq a$. In order to prove the above integral equality, we apply the general transformation formula for integrals [Fleming, 1977]. The following change of variable transformation is applied to the second integral, $T : x = -y + a + z$. Note that the absolute value of the Jacobian equals one and that the region of integration transforms from S_a to S_z , since $T^{-1}(S_a) = \{y \in R^n \mid \|y - z\|_{Q_{\hat{a}}}^2 \leq \|y - a - z + u\|_{Q_{\hat{a}}}^2, \forall u \in Z^n\} = S_z$. This shows that the above integral equality holds true.

Although it is of course comforting to know that the probability of correct integer estimation is the largest of all nonzero probabilities, it is not sufficient for ambiguity resolution to be successful. For that to be the case, one still has to evaluate the probability of correct integer estimation and check whether its value is sufficiently close to one. This issue will be taken up in Sect. 5.

4.2 The integer normal distribution is symmetric

Apart from knowing the maximum of the probability mass function, it is also of interest to know how the probability masses are distributed about this maximum. This concerns the shape of the distribution. We know that the multivariate normal distribution (7) is symmetric about a . We will now show that the corresponding integer normal distribution is symmetric about a as well. We have

$$P(\check{a} = a - z) = P(\check{a} = a + z), \forall z \in Z^n \quad (12)$$

In order to show this, we have to prove

$$\int_{S_{a-z}} c \exp\left\{-\frac{1}{2} \|x - a\|_{Q_{\hat{a}}}^2\right\} dx = \int_{S_{a+z}} c \exp\left\{-\frac{1}{2} \|x - a\|_{Q_{\hat{a}}}^2\right\} dx$$

This result follows from applying the change of variable transformation, $T : x = 2a - y$ and noting that $T^{-1}(S_{a-z}) = \{y \in R^n \mid \|y - a - z\|_{Q_{\hat{a}}}^2 \leq \|y - 2a + u\|_{Q_{\hat{a}}}^2, \forall u \in Z^n\} = S_{a+z}$.

4.3 The 'float' and 'fixed' solutions are unbiased

A third property of the integer least-squares estimator is its unbiasedness. We are thus in the happy situation that not only the real-valued least-squares ambiguities are unbiased, but their integer least-squares counterparts as well. Hence,

$$E\{\check{a}\} = E\{\hat{a}\} = a \quad (13)$$

This property of unbiasedness is a direct consequence of the symmetry of the distribution. To see this, recall the definition of the expectation of \check{a} . It reads $E\{\check{a}\} = \sum z P(\check{a} = z)$, with the sum taken over all grid points of Z^n . This may also be written as $E\{\check{a}\} = \sum (z + a) P(\check{a} = a + z)$ and as $E\{\check{a}\} = \sum (a - z) P(\check{a} = a - z)$. Taking the sum of these last two expressions and noting that $P(\check{a} = a + z) = P(\check{a} = a - z)$, gives $2E\{\check{a}\} = \sum 2a P(\check{a} = a + z) = 2a$ from which the unbiasedness follows.

From the unbiasedness of the integer least-squares ambiguities it also follows that the 'fixed' solution \check{b} is unbiased. For the 'float' solution we have $E\{\hat{a}\} = a$ and $E\{\hat{b}\} = b$. This together with $E\{\check{a}\} = a$, shows that the expectation of (5) is given as

$$E\{\check{b}\} = E\{\hat{b}\} = b \quad (14)$$

With the above results we have proven that the inclusion of the integer constraints does not introduce any biases when using the least-squares principle. That is, the 'fixed' solutions are unbiased whenever the 'float' solutions are. But biases may still be introduced of course, when models are used that are misspecified (e.g. due to cycle slips or outliers in the data).

Finally we note that the property of unbiasedness is not restricted to the integer least-squares estimator. Certain other integer estimators can be shown to be unbiased as well. Such a class of unbiased integer estimators was introduced in [Teunissen, 1998b].

4.4 Precision and probability of the integer ambiguities

Although it is comforting to know that the integer least-squares ambiguities are unbiased, this is not enough for ambiguity resolution to be successful. For that to be the case, we also need a sufficiently small variability of the integer least-squares ambiguities about their integer means. We will now show how the ambiguity precision is related to the probability of correct integer estimation.

By definition, the variance-covariance matrix of the integer least-squares ambiguities is given as

$$Q_{\check{a}} = \sum_{z \in Z^n} (z - a)(z - a)^T P(\check{a} = z) \quad (15)$$

where, as before, a denotes the integer mean of \check{a} . The sum is taken over all grid points in Z^n . Note the absence of $P(\check{a} = a)$ in the above expression. This shows that the expected variability of \check{a} is due to the probabilities of wrong integer estimation. Hence, the integer ambiguities will have a poor precision if the probabilities of wrong integer estimation are not negligible. The reverse of this statement is true as well. That is, if the variance of the integer ambiguities is sufficiently small, then the probability of correct integer estimation is sufficiently large. To see this, consider the j th diagonal entry of (15). It reads

$$c_j^T Q_{\check{a}} c_j = \sum_{z \in Z^n} [c_j^T (z - a)]^2 P(\check{a} = z)$$

where c_j denotes the j th canonical unit vector. The left-hand side equals the variance of the j th integer ambiguity. In the sum on the right-hand-side, the contribution for $z = a$ is absent and the non-zero minimum of $[c_j^T (z - a)]^2$ equals 1. This gives the inequality $\sigma_{\check{a}_j}^2 \geq \sum_{z \in Z^n \setminus \{a\}} P(\check{a} = z)$, or $P(\check{a} = a) \geq 1 - \sigma_{\check{a}_j}^2$. Since such an inequality holds for all diagonal entries of the ambiguity variance matrix, we finally get

$$\boxed{P(\check{a} = a) \geq 1 - \frac{1}{n} \text{trace } Q_{\check{a}}} \quad (16)$$

This shows that the probability of correct integer estimation is bounded from below by one minus the average variance of the integer least-squares ambiguities. Hence, smaller variances will push the probability of correct integer estimation closer to one. In fact, this probability equals one already when one of the variances of the integer least-squares ambiguities vanishes.

5 Evaluation of the integer normal distribution

In this section we will present different approaches for evaluating the integer normal distribution. Particular attention will be given to the probability of correct integer estimation. Although these approaches differ in the way they try to approximate this probability, they all make use of the variance matrix of the least-squares ambiguities. The first approach is based on simulating the probability of correct integer estimation. The second approach uses the probability of correct integer estimation of a less optimal integer estimator, namely the bootstrapped estimator. The third approach uses bounds on the ambiguity variance-covariance matrix to obtain corresponding bounds on the probability of correct integer estimation. Finally, in the last approach such bounds are obtained by bounding the region of integration.

5.1 Simulating the probability mass function

In general it is very difficult to evaluate the integer normal distribution (10) exactly. This is due to the rather complicated geometry of the integration region S_z . The method of simulation can however be used to obtain approximations of the probabilities $P(\check{a} = z)$. This goes as follows. We know that the 'float' solution is distributed as $\hat{a} \sim N(a, Q_{\hat{a}})$. We also know that the integer normal distribution is symmetric about the mean a . Hence, in order to obtain the required probability masses we may shift the distribution over a and restrict our attention to $N(0, Q_{\hat{a}})$, draw samples from it and use these samples to obtain the corresponding integer samples by means of solving (4). Repeating this procedure a sufficient number of times, allows us then to built up the required frequency table. The probability of correct integer estimation is then given as $P(\check{a} = 0)$.

Thus first one starts generating, using a random generator, n independent samples from the univariate standard normal distribution, say s_1, \dots, s_n from $N(0, 1)$. These samples are then collected in the vector $s = (s_1, \dots, s_n)^T$ and transformed by means of $\hat{a} = Gs$, where matrix G equals the Cholesky factor of the ambiguity variance-covariance matrix $Q_{\hat{a}}$, i.e. $Q_{\hat{a}} = GG^T$. Hence, \hat{a} is now a sample from $N(0, Q_{\hat{a}})$. Using this sample to solve (4) results in the corresponding integer least-squares sample. By repeating this process an N -number of times, one obtains a collection of N integer vectors. Of this collection one can now infer how often a particular grid point, say z , is visited. This gives the frequency N_z . An approximation to the required probability masses follows then from the relative frequencies. Thus

$$P(\check{a} = z) \approx \frac{N_z}{N} \quad (17)$$

Successful ambiguity resolution can now be expected feasible when the probability $P(\check{a} = 0)$ is sufficiently close to one. Note that this procedure requires that problem (4) has to be solved N -times. For large N , this becomes a very time consuming task if not an efficient search is in place for solving the integer least-squares problem. This shows that the simulation should not be based on the original DD ambiguities, but instead on the transformed ambiguities obtained by means of the LAMBDA method [Teunissen, 1993], [Jonge de and Tiberius, 1996]. The integer normal distribution of the transformed ambiguities differs of course from the one of the DD ambiguities. For the DD ambiguities for instance, the nonzero probabilities $P(\check{a} = z)$, with $z \neq 0$, will be more spread out. However, since there is a one-to-one correspondence between the two distributions, the probability of correct integer estimation will be the same for both distributions.

In order to get an idea of how large N should be taken in the simulation, we consider the probability that N_0 out of N integer vectors equal the zero vector. If the N samples are drawn independently from the normal distribution $N(0, Q_{\hat{a}})$, then this probability is governed by the binomial distribution and is given as

$$P(N_0) = \frac{N!}{(N - N_0)!N_0!} P_0^{N_0} (1 - P_0)^{N - N_0}$$

where we made use of the abbreviation $P_0 = P(\check{a} = 0)$. The mean (expectation) and variance

(dispersion) of the relative frequency N_0/N follow therefore as

$$E\{N_0/N\} = P_0 \quad \text{and} \quad D\{N_0/N\} = P_0(1 - P_0)/N$$

Note that the first expression is in fact the motivation for using the relative frequency as an estimator for P_0 , the probability of correct integer estimation. The second expression gives the precision of this estimator. It depends on both P_0 and N .

Using the above mean and variance we may now apply the Chebyshev inequality to obtain an upperbound on the probability that the relative frequency N_0/N differs more than ϵ from P_0 . The corresponding Chebyshev inequality reads

$$P(|\frac{N_0}{N} - P_0| \geq \epsilon) \leq \frac{P_0(1 - P_0)}{N\epsilon^2} \quad (18)$$

The required number of samples N can be obtained by setting both ϵ and the upperbound to a small enough value. For instance, when the probability of correct integer estimation equals $P_0 = 1 - 10^{-3}$, an upperbound of one percent and a deviation of $\epsilon = 10^{-3}$ leads to a required number of samples of $N = 10^5$. This shows that in general a large number of samples are needed to get a sufficiently precise estimate of the probability of correct integer estimation. Instead of using the above Chebyshev inequality, one may also use the Gaussian approximation for the binomial distribution to obtain an estimate of the required number of samples, when N is large. This will usually give a somewhat less conservative estimate of N .

Instead of using a random generator to obtain samples of the integer normal distribution, one may of course apply the same idea to actual 'real-world' experiments. In that case the experiment will consist of repeatedly estimating the integer ambiguities, while keeping count of the success rates. Such a study was performed in [Jonkman, 1998] for the geometry-free model and in [Tiberius and de Jonge, 1995] for the geometry-based GPS model.

5.2 Probability of the bootstrapped estimator

Instead of using simulation, one may also try to formulate bounds on the probability of correct integer least-squares estimation. One such bound is obtained if we consider the probability of a less optimal integer estimator, the bootstrapped estimator.

As it was remarked earlier, it is difficult in general to evaluate the integer normal distribution (10) exactly. The evaluation becomes relatively simple though when \hat{a} is a scalar or when $Q_{\hat{a}}$ is diagonal. In the scalar case the integer least-squares estimator coincides with the operation 'round to the nearest integer'. For this case the integer normal distribution takes the form

$$P(\check{a} = i) = \int_{(i-a)-\frac{1}{2}}^{(i-a)+\frac{1}{2}} \frac{1}{\sigma_{\hat{a}}\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x/\sigma_{\hat{a}})^2\right\} dx \quad (19)$$

where i ranges over the set of integers. For the purpose of ambiguity resolution we are particularly interested in the probability of correct integer estimation. This probability is given as $P(\check{a} = a) = P(|\hat{a} - a| \leq \frac{1}{2})$, which can be evaluated by means of (19). The probability of correct integer estimation becomes then

$$P(\check{a} = a) = 2\Phi\left(\frac{1}{2\sigma_{\hat{a}}}\right) - 1 \quad \text{with} \quad \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z^2\right\} dz \quad (20)$$

In the multivariate case one can still use the 'rounding operation' to obtain the integer least-squares solution, provided the variance-covariance matrix is diagonal. In that case the problem decouples into n scalar problems of the above type. Hence, the probability of correct integer estimation becomes then

$$P(\check{a} = a) = \prod_{i=1}^n [2\Phi\left(\frac{1}{2\sigma_{\hat{a}_i}}\right) - 1] \quad (21)$$

In the actual practice of GPS, the ambiguity variance-covariance matrix is of course fully populated and therefore nondiagonal. Expression (21) does therefore not apply. However, a similar expression can be obtained if we consider the so-called bootstrapped estimator. This integer estimator follows from a sequential conditional least-squares adjustment and is computed as follows. If n ambiguities are available, one starts with the first ambiguity \hat{a}_1 , and rounds its value to the nearest integer. Having obtained the integer value of this first ambiguity, the real-valued estimates of all remaining ambiguities are then corrected by virtue of their correlation with the first ambiguity. Then the second, but now corrected, real-valued ambiguity estimate is rounded to its nearest integer. Having obtained the integer value of the second ambiguity, the real-valued estimates of all remaining $n - 2$ ambiguities are then again corrected, but now by virtue of their correlation with the second ambiguity. This process is continued until all ambiguities are taken care of. Since the components of this bootstrapped estimator are conditionally independent, it follows that its probability of correct integer estimation takes the form of (21), but now with the unconditional standard deviations replaced by their sequential conditional counterparts. Thus if we denote the bootstrapped estimator as \check{a}_B , we have

$$P(\check{a}_B = a) = \prod_{i=1}^n [2\Phi\left(\frac{1}{2\sigma_{\hat{a}_{i|(i-1), \dots, 1}}}\right) - 1] \quad (22)$$

where $\sigma_{\hat{a}_{i|(i-1), \dots, 1}}$ denotes the *conditional* standard deviation. It can be shown that this probability provides an upperbound on the probability of correct integer estimation based on the simple 'componentwise rounding' mechanism, i.e. $P(\check{a}_B = a) \geq P(\cap_{i=1}^n | \hat{a}_i - a_i | \leq \frac{1}{2})$, see [Teunissen, 1998a]. The bootstrapped solution will thus more often lead to the correct integer ambiguities than the solution based on the 'componentwise rounding'. In fact, when (22) is sufficiently close to one, the bootstrapped estimator becomes a viable alternative to the integer least-squares estimator. Note however that (22) depends on the conditional standard deviations of the individual ambiguities. Since the first three conditional standard deviations of the DD ambiguities are known to be large in general [Teunissen, 1996], the probability $P(\check{a}_B = a)$ can be expected to be small when using DD ambiguities. This shows that one should first apply the decorrelation process of the LAMBDA method, before using (22) to evaluate the probability of correct integer estimation.

5.3 Bounding the ambiguity variance matrix

Another way to get a grip on the probability of correct integer estimation, while at the same time avoiding the complicated integration of (10), is to make use of scaled unit matrices as variance matrices. We know that the computation of the probability of correct integer estimation becomes straightforward when the ambiguity variance matrix is diagonal. That is, when the least-squares

ambiguities are fully decorrelated. This suggests that we bound the actual ambiguity variance matrix from above and below by diagonal matrices, and compute the probability of correct integer estimation that would belong to these diagonal matrices. The simplest way of bounding the actual ambiguity variance matrix from above and below, is to make use of its maximum and minimum eigenvalue. This gives

$$\lambda_{\min} I_n \leq Q_{\hat{a}} \leq \lambda_{\max} I_n$$

Using these bounds, the corresponding bounds for the probability of correct integer estimation reads

$$[2\Phi(\frac{1}{2\sqrt{\lambda_{\max}}}) - 1]^n \leq P(\check{a} = a) \leq [2\Phi(\frac{1}{2\sqrt{\lambda_{\min}}}) - 1]^n \quad (23)$$

Note that the two bounds coincide when the two extreme eigenvalues coincide. This is the case when the ambiguity variance matrix itself is a scaled unit matrix. In the actual practice of GPS this will not happen. In fact, the two extreme eigenvalues will differ considerably when the variance matrix of the DD ambiguities is used. In that case the above two bounds would become too loose to be useful. When using the decorrelated ambiguities as produced by the LAMBDA method, the elongation of the ambiguity search space is considerably reduced and the ratio of the two extreme eigenvalues is pushed towards its minimum of one. Hence, the above bounds are much sharper when using the eigenvalues of the transformed ambiguity variance matrix, than when using the eigenvalues of the original DD ambiguity variance matrix.

5.4 Bounding the region of integration

Another way to get a grip on the probability of correct integer estimation is to replace the original region of integration S_a by a subset $L_a \subset S_a$ and by an enclosing set $U_a \supset S_a$. In that case the probability of correct integer estimation lies in the interval

$$P(\hat{a} \in L_a) \leq P(\check{a} = a) = P(\hat{a} \in S_a) \leq P(\hat{a} \in U_a) \quad (24)$$

Both regions of integration should of course be chosen such that the corresponding probabilities are easily evaluated in practice. The probability $P(\hat{a} \in L_a)$ can then be used to infer whether ambiguity resolution can be expected to be successful, while the probability $P(\hat{a} \in U_a)$ will show when one can expect successful ambiguity resolution to fail. Thus there is enough confidence that ambiguity resolution will be successful when the lowerbound is sufficiently close to one, while no such confidence exists when the upperbound turns out to be too small.

In the following we will use the minimum grid point distance, or minimum norm, to bound the region of integration [Teunissen, 1997b]. The smallest distance between two integer grid points, as measured in the metric of the ambiguity variance matrix, will correspond with the situation that one has the greatest difficulty in discriminating between two grid points. Hence, if this distance is large enough one can expect ambiguity resolution to be successful. On the other hand, successful ambiguity resolution will become problematic if this distance is too small.

5.4.1 The probability $P(\hat{a} \in U_a)$

Let us first consider the geometry of the pull-in-region S_a . Since

$$\| \hat{a} - a \|_{Q_{\hat{a}}}^2 \leq \| \hat{a} - z \|_{Q_{\hat{a}}}^2 \iff (z - a)^T Q_{\hat{a}}^{-1} (\hat{a} - a) \leq \frac{1}{2} \| z - a \|_{Q_{\hat{a}}}^2, \forall z \in Z^n$$

it follows that

$$S_a = \{ \hat{a} \in R^n \mid |w_i| \leq \frac{1}{2} \| c_i \|_{Q_{\hat{a}}}, \forall c_i \in Z^n \} \text{ with } w_i = \frac{c_i^T Q_{\hat{a}}^{-1} (\hat{a} - a)}{\sqrt{(c_i)^T Q_{\hat{a}}^{-1} (c_i)}} \quad (25)$$

Note that w_i is the well-known w -test statistic for testing one-dimensional alternative hypotheses [Baarda, 1968], [Teunissen, 1985]. It is distributed as $w_i \sim N(0, 1)$. Geometrically, w_i can be interpreted as an orthogonal projector which projects $(\hat{a} - a)$ onto the direction vector c_i . Hence, S_a is the intersection of the *banded* subsets $\{ \hat{a} \in R^n \mid |w_i| \leq \frac{1}{2} \| c_i \|_{Q_{\hat{a}}} \}$, all centred at a and each having a width $\| c_i \|_{Q_{\hat{a}}}$. Since any *finite* intersection of these banded subsets encloses S_a , the following choice for U_a is suggested

$$U_a = \{ \hat{a} \in R^n \mid |w_i| \leq \frac{1}{2} \| c_i \|_{Q_{\hat{a}}}, i = 1, \dots, p \} \supset S_a \quad (26)$$

Thus U_a is taken as the intersection of p such bands. Note that the choice for p and for the grid vectors c_i is still left open. The simplest choice would be $p = 1$. In that case the probability is easily evaluated due to the standard normal distribution of w_1 . For c_1 one could still take any one of the grid vectors. The best choice in this case would be the nonzero grid vector having shortest length. The probability reads then

$$P(\hat{a} \in U_a) = 2\Phi\left(\frac{1}{2} \min_{z \in Z^n \setminus \{0\}} \|z\|_{Q_{\hat{a}}}\right) - 1 \quad (27)$$

This probability can now be used as upperbound for the probability of correct integer estimation. If a sharper upperbound is needed, one will have to choose p larger than 1. The situation becomes then more complicated due to the fact that the w_i are correlated. In order to tackle this case, we first define the p -vector

$$v = (v_1, \dots, v_p)^T \text{ with } v_i = \frac{w_i}{\|c_i\|_{Q_{\hat{a}}}} \quad (28)$$

Then $U_a = \{ \hat{a} \in R^n \mid \cap_{i=1}^p |v_i| \leq \frac{1}{2} \}$. The probability $P(\hat{a} \in U_a)$ equals therefore the probability that the 'componentwise rounding' of the vector $v = (v_1, \dots, v_p)^T$ produces the zero vector. Hence, we can now make use of the bootstrapped results of the previous subsection. That is, $P(\check{a} \in U_a)$ and thus also $P(\check{a} = a)$ will be bounded from above by the probability that the 'sequential rounding' of the entries of v produces the zero vector. Hence,

$$P(\check{a} = a) \leq \prod_{i=1}^p \left[2\Phi\left(\frac{1}{2\sigma_{v_i|(i-1), \dots, 1}}\right) - 1 \right] \quad (29)$$

where $\sigma_{v_i|(i-1),\dots,1}$ denotes the conditional standard deviation of v_i . These sequential conditional standard deviations of v follow from applying an LDL^T -decomposition to the variance-covariance matrix of v . The entries of this variance-covariance matrix Q_v are given as

$$\sigma_{v_i v_j} = \frac{c_i^T Q_{\hat{a}}^{-1} c_j}{\|c_i\|_{Q_{\hat{a}}}^2 \|c_j\|_{Q_{\hat{a}}}^2} \quad (30)$$

After applying the LDL^T -decomposition to this variance-covariance matrix, the sequential conditional variances follow as the diagonal entries of the diagonal matrix D . Note that the variance-covariance matrix needs to be of full rank in order to avoid that some of the conditional variances of v become zero. This implies that the grid vectors c_i , $i = 1, \dots, p \leq n$, need to be linear independent.

5.4.2 The probability $P(\hat{a} \in L_a)$

As with the choice $U_a \supset S_a$, also a choice for the subset $L_a \subset S_a$ can be made by considering the geometry of the pull-in-region S_a . Note that the widths of the bands that make up the intersection of (25) vary in length. This shows that a subset of S_a can be obtained by replacing all these varying widths by one single width, namely the smallest one. Hence, $\{\hat{a} \in R^n \mid |w_i| \leq \frac{1}{2} \min_{z \in Z^n \setminus \{0\}} \|z\|_{Q_{\hat{a}}}, \forall z \in Z^n\} \subset S_a$. Moreover, since $|w_i| = \frac{1}{2} \min_{z \in Z^n \setminus \{0\}} \|z\|_{Q_{\hat{a}}}$ describes a pair of opposite *planes of support* of the ellipsoid $\|\hat{a} - a\|_{Q_{\hat{a}}}^2 = (\frac{1}{2} \min_{z \in Z^n \setminus \{0\}} \|z\|_{Q_{\hat{a}}})^2$ [Teunissen, 1996], it follows that the ellipsoid

$$L_a = \{\hat{a} \in R^n \mid \|\hat{a} - a\|_{Q_{\hat{a}}}^2 \leq (\frac{1}{2} \min_{z \in Z^n \setminus \{0\}} \|z\|_{Q_{\hat{a}}})^2\} \quad (31)$$

is also a subset of S_a . We therefore have the lowerbound

$$P(\check{a} = a) \geq P(\chi^2(n, 0) \leq (\frac{1}{2} \min_{z \in Z^n \setminus \{0\}} \|z\|_{Q_{\hat{a}}})^2) \quad (32)$$

with $\chi^2(n, 0)$ the central Chi-square distribution with n degrees of freedom. These probabilities can be computed using [Johnson et al., 1994]

$$P(\chi^2(n, 0) \leq x) = \begin{cases} 1 - e^{-x/2} \sum_{i=0}^{\frac{1}{2}(n-2)} \left(\frac{x}{2}\right)^i / i! & n = \text{even} \\ (2\Phi(\sqrt{x}) - 1) - e^{-x/2} \sum_{i=0}^{\frac{1}{2}(n-3)} \left(\frac{x}{2}\right)^{i+\frac{1}{2}} / , (i + \frac{3}{2}) & n = \text{odd} \end{cases}$$

Note that for $n = 1$, we have $\chi^2(1, 0) = N^2(0, 1)$ and $\min_{z \in Z^n \setminus \{0\}} \|z\|_{Q_{\hat{a}}}^2 = 1/\sigma_{\hat{a}}^2$. This shows that for the one-dimensional case, (32) reduces to (20). The upperbound (29) and lowerbound (32) are then also identical.

Note that the minimum grid point distance is used in both the upperbound (27) and lowerbound (32). Computation of this integer least-squares problem can be avoided if one settles for less sharp bounds, using eigenvalues and/or (conditional) variances. The following four bounds can then be used instead,

$$\begin{cases} \frac{1}{\lambda_{max}} & \leq \min_{z \in Z^n \setminus \{0\}} \|z\|_{Q_{\hat{a}}}^2 \leq \frac{1}{4\sigma_{\hat{a}}^2} (e + \frac{1}{e})^2 \text{ with } e^2 = \frac{\lambda_{max}}{\lambda_{min}}, i = 1, \dots, n \\ \frac{1}{\sigma_{\hat{a}_{i|I}}^2} & \leq \min_{z \in Z^n \setminus \{0\}} \|z\|_{Q_{\hat{a}}}^2 \leq \frac{1}{\sigma_{\hat{a}_{n|N}}^2} \text{ when } z_i \neq 0, z_j = 0 \text{ for } j < i \end{cases} \quad (33)$$

with $\sigma_{\hat{a}_i}^2$ and $\sigma_{\hat{a}_{i|I}}^2$ being the i th unconditional and conditional ambiguity variance respectively, and with λ_{min} and λ_{max} being the two extreme eigenvalues of the ambiguity variance matrix. Note, although the minimum norm is invariant for the choice of ambiguity parametrization, that the above bounds depend on the type of ambiguities used.

The first lowerbound follows from the definition of the maximum eigenvalue. The second lowerbound, sharper than the first, follows when the square of the grid point distance is given a sum-of-squares form using the *LDU*-decomposition and noting that this sum is never smaller than the first nonzero entry in the sum. To use this lowerbound, one still needs to check whether $z_i \neq 0, z_j = 0$ for $j < i$. This can be avoided by taking as lowerbound the smallest reciprocal (conditional) ambiguity variance. For the unconditional variances this can be understood as follows. Since $z \neq 0$, at least one of its entries is nonzero, say z_i . Thus the reciprocal value of $\sigma_{\hat{a}_i}^2$ can be taken as lowerbound. But this shows that certainly the minimum of all reciprocal ambiguity variances can be used as lowerbound. For the conditional variances a similar reasoning applies. The first upperbound follows from applying Kantovorich inequality, see e.g. [Rao, 1973], with z chosen as the canonical unit vector having 1 as its i th entry. Finally, the second upperbound follows from taking the length of $z = (1, 0, \dots, 0)^T$ as upperbound.

6 Summary

In this contribution we introduced the discrete distribution of the integer least-squares ambiguities and called it the *integer normal distribution*. This probability mass function was shown to be symmetric and centred at the integer mean of the real-valued least-squares ambiguity vector \hat{a} . It was also shown that the 'fixed' solutions \check{a} and \check{b} are unbiased. Hence, no biases are introduced when collapsing the pull-in-regions S_z to their respective grid points. The maximum of the integer normal distribution coincides with the probability of correct integer estimation. Since this maximum is particularly of relevance for GPS ambiguity resolution, we presented different approaches for evaluating this maximum. It was pointed out that an exact evaluation is rather difficult in general, due to the complicated geometry of the pull-in-regions. The first approach was based on reconstructing the probability mass function by means of a simulation. The second approach made use of the integer bootstrapped estimator. Due to the sequential conditioning on which this estimator is based, the probability of correct integer estimation is easily evaluated for the bootstrapped estimator. The third approach was based on bounding the ambiguity variance matrix. Finally, the last approach was based on geometric bounds for the pull-in-regions, where use was made of the minimum grid point distance.

When comparing the different approaches, a few remarks can be made. The advantage of the first approach, although rather computational intensive, is that the approximations of the probabilities can be obtained in principle with any desired level of accuracy. This is not the case with the approaches that use bounds, although they are often much easier to evaluate. Here one generally depends on the precision of the least-squares ambiguities. The more precise the ambiguities are, the sharper the bounds generally become. The only influence one can exercise on the bounds is by using a proper ambiguity parametrization. That is, the noninvariant bounds can be made sharper by using decorrelated ambiguities instead of DD ambiguities. Note however that not all bounds converge to the exact probability of correct integer estimation in case the

ambiguity variance matrix becomes a diagonal matrix. This is only the case for the second and third approach, but not for the last approach.

We emphasize in conclusion that the integer normal distribution is completely specified by the variance-covariance matrix of the real-valued least-squares ambiguities. Hence, all diagnostic measures presented can be computed once this matrix is known. This implies that it is possible to compute the probability of correct integer estimation before the actual measurements are carried out. Only the design matrix and variance-covariance matrix of the GPS observables need to be known. The theory presented in this contribution can therefore be used to analyse different measurement scenarios as to their strength in resolving the integer ambiguities successfully.

7 References

- [1] Baarda, W. (1968): *A testing procedure for use in geodetic networks*, Netherlands Geodetic Commission, Publications on Geodesy, New Series, Vol. 2, No. 5.
- [2] Betti, B., M. Crespi, F. Sanso (1993): A geometric illustration of ambiguity resolution in GPS theory and a Bayesian approach. *Manuscripta Geodaetica*, 18:317-330.
- [3] Blewitt, G. (1989): Carrier-phase ambiguity resolution for the Global Positioning System applied to geodetic baselines up to 2000 km. *Journal of Geophysical Research*, Vol. 94, No. B8, pp. 10.187- 10.302.
- [4] Dong, D., Y. Bock (1989): Global Positioning System network analysis with phase ambiguity resolution applied to crustal deformation studies in California. *Journal of Geophysical Research*, 94, pp. 3949-3966.
- [5] Counselman, C.C., S.A. Gourevitch (1981): Miniature interferometer terminals for earth surveying: ambiguity and multipath with the Global Positioning System. *IEEE Transactions on Geoscience and Remote Sensing*, GE-19(4): 244-252.
- [6] Fleming, W. (1977): *Functions of Several Variables*, Springer Verlag.
- [7] Hofmann-Wellenhof, B., H. Lichtenegger, J. Collins (1997): *Global Positioning System: Theory and Practice*. 4th edition. Springer Verlag.
- [8] Jonge de, P.J., C.C.J.M. Tiberius (1996): Integer ambiguity estimation with the LAMBDA method. *Proceedings IAG Symposium No. 115, GPS trends in precise terrestrial, airborne and spaceborne applications*, G. Beutler et al. (eds.), Springer Verlag, pp. 280-284.
- [9] Jonge de P.J., C.C.J.M. Tiberius (1996): The LAMBDA method for integer ambiguity estimation: implementation aspects. Publications of the Delft Computing Centre, *LGR-Series* No. 12.
- [10] Johnson N.L., S. Kotz, N. Balakrishnan (1994): *Continuous univariate distributions*, volume 1, second edition, John Wiley and Sons.
- [11] Jonkman, N.F. (1998): Integer ambiguity estimation without the receiver-satellite geometry. Publications of the Delft Geodetic Computing Centre, *LGR-Series*, No. 18.

- [12] Kleusberg, A., P.J.G. Teunissen (eds) (1996): *GPS for Geodesy*, Lecture Notes in Earth Sciences, Vol. 60, Springer Verlag.
- [13] Leick, A. (1995): *GPS Satellite Surveying*. 2nd edition, John Wiley, New York.
- [14] Rao, C.R. (1973): *Linear Statistical Inference and its Applications*, 2nd edition, John Wiley and Sons.
- [15] Strang, G., K. Borre (1997): *Linear Algebra, Geodesy, and GPS*, Wellesley-Cambridge Press.
- [16] Teunissen, P.J.G. (1985): Quality control in geodetic networks. Chapter 17 in *Optimization and Design of Geodetic Networks*, E. Grafarend and F. Sanso (eds), Springer Verlag.
- [17] Teunissen, P.J.G. (1990): *GPS op afstand bekijken*. In: Een halve eeuw in de goede richting, Lustrumboek Snellius pp. 215-233.
- [18] Teunissen, P.J.G. (1993): Least-squares estimation of the integer GPS ambiguities. Invited Lecture, Section IV Theory and Methodology, IAG General Meeting, Beijing, China, August 1993. Also in: *LGR Series*, No. 6, Delft Geodetic Computing Centre.
- [19] Teunissen, P.J.G. (1996): GPS carrier phase ambiguity fixing concepts. Chapter 8 in: *GPS for Geodesy*. A. Kleusberg and P.J.G. Teunissen (eds). Lecture Notes in Earth Sciences, Vol. 60, Springer Verlag, pp. 263-335.
- [20] Teunissen, P.J.G. (1997a): Some remarks on GPS ambiguity resolution. *Artificial Satellites*, Vol. 32, No. 3, pp. 119-130.
- [21] Teunissen, P.J.G. (1997b): A canonical theory for short GPS baselines. Part IV: precision versus reliability. *Journal of Geodesy*, 71:513-525.
- [22] Teunissen, P.J.G. (1997c): Closed form expressions for the volume of the GPS ambiguity search spaces. *Artificial Satellites*, Vol. 31, No. 1, pp. 5-20.
- [23] Teunissen, P.J.G. (1998a): Probability of integer GPS ambiguity rounding. *Journal of Geodesy*, in print.
- [24] Teunissen, P.J.G. (1998b): A class of unbiased integer GPS ambiguity estimators. *Artificial Satellites*, Vol. 33, No. 1.
- [25] Tiberius, C.C.J.M., P.J. de Jonge (1995): Fast positioning using the LAMBDA method. *Proceedings DSNS-95*, paper 30, 8p.

Success probability of integer GPS ambiguity rounding and bootstrapping

P. J. G. Teunissen

Delft Geodetic Computing Centre (LGR), Faculty of Geodesy, Delft University of Technology, Thijsseweg 11, 2629 JA, The Netherlands
E-mail: P.J.G. Teunissen@geo.tudelft.nl

Received: 24 March 1998 / Accepted: 8 June 1998

Abstract. Global Positioning System ambiguity resolution is usually based on the integer least-squares principle (Teunissen 1993). Solution of the integer least-squares problem requires both the execution of a search process and an ambiguity decorrelation step to enhance the efficiency of this search. Instead of opting for the integer least-squares principle, one might also want to consider less optimal integer solutions, such as those obtained through rounding or sequential rounding. Although these solutions are less optimal, they do have one advantage over the integer least-squares solution: they do not require a search and can therefore be computed directly. However, in order to be confident that these less optimal solutions are still good enough for the application at hand, one requires diagnostic measures to predict their rate of success. These measures of confidence are presented and it is shown how they can be computed and evaluated.

Key words. GPS ambiguity · Probability · Integer rounding · Integer bootstrapping

1 Introduction

Global Positioning System (GPS) ambiguity resolution, which is the process of resolving the unknown cycle ambiguities of the double-difference (DD) carrier phase data as integers, consists of an estimation part and a validation part. The purpose of the validation part is to infer, given the data set at hand, whether or not the actual integer solution is sufficiently consistent with the GPS model used. The purpose of the estimation part, on the other hand, is to come up with a procedure for obtaining an integer solution and to show which quality this solution is likely to have. In other words, for the estimation part one has to decide which integer estima-

tor to use, as well as provide a description of its qualitative properties.

Although GPS research in the last decade or so has resulted in a variety of different methods and proposals for estimating integer ambiguities, there is a growing consensus that the use of the integer least-squares principle is the preferred approach for integer ambiguity estimation. There are two main reasons for this. First, the integer least-squares estimator is optimal in the sense that it maps the largest probability mass onto the integer vector of ambiguity means. Second, it has been demonstrated that by means of a least-squares ambiguity decorrelation adjustment, it is now possible to compute the integer least-squares solution rigorously and efficiently. The pitfalls associated with some of the more classical approaches can thus be avoided. For a discussion of these pitfalls, see Teunissen (1997a).

When the ambiguity variance-covariance matrix is diagonal, the integer least-squares estimator reduces to the estimator that corresponds to a simple componentwise rounding of the real-valued least-squares ambiguity vector. This is not the case, however, when the ambiguity variance-covariance matrix is non-diagonal. Nevertheless, in this case one could still decide to stick to the simple rounding mechanism in order to obtain an integer solution. Although the integer estimator so obtained will be less optimal, it might still be ‘optimal enough’ for the particular application at hand. This requires, however, that one is also able to predict the success rate of this simple integer estimator. Hence, in order to evaluate the quality of this estimator, we need to be able to evaluate its probability of correct integer estimation. It is the purpose of this contribution to present these measures of confidence for two simple integer estimators, the one that corresponds to a componentwise rounding and the one that corresponds to a sequential rounding of the individual ambiguities.

In Sect. 2 we first give a brief overview of the estimation steps involved in ambiguity resolution. In this section we also illustrate how to set the size of the ambiguity search space. The integer vector obtained by

means of either simple rounding or a sequential rounding is particularly suited for this, provided the ambiguity decorrelation process precedes the rounding step. In Sect. 3, we present diagnostic measures to evaluate whether or not it makes sense to use the simple rounding mechanism in order to obtain the integer ambiguity vector. In Sect. 4 this is also done for the integer ambiguity vector obtained through a sequential rounding. We also prove that the probability of correct integer estimation based on a sequential rounding is larger than, or at least as large as, the probability of correct integer estimation based on a straightforward rounding. Finally, we give an upper bound for the probability of correct integer estimation which is invariant for the whole class of admissible ambiguity transformations. It is based on the determinant of the ambiguity variance–covariance matrix, for which closed-form expressions are available for a variety of different GPS models.

2 The ambiguity search space

2.1 Integer least-squares estimation

In principle, all the GPS models can be cast in the following conceptual frame of linear(ized) observation equations:

$$y = Aa + Bb + e \quad (1)$$

where y is the given data vector, a and b are the unknown parameter vectors of order n and p , respectively, and e is the noise vector of order m . The matrices A and B are the corresponding design matrices of order $m \times n$ and $m \times p$ respectively. The data vector y will usually consist of the ‘observed minus computed’ single- or dual-frequency DD phase and/or pseudo range (code) observations, accumulated over all observation epochs. The entries of vector a are then the DD carrier phase ambiguities, expressed in units of cycles rather than range. They are known to be integers. The entries of vector b will then consist of the remaining unknown parameters such as, for example, baseline components (coordinates) and possibly atmospheric delay parameters (troposphere, ionosphere). Note that we have followed the customary practice of using the DD version of the code and carrier phase observation equations. This is, however, not strictly needed. As an alternative one can also work with undifferenced or single-differenced observations (de Jonge 1998). In that case the non-integer ambiguities will have to be reparametrized so as to obtain integer ambiguities again.

When using the least-squares principle, the above system of observation equations can be solved by means of the minimization problem

$$\begin{aligned} & \min_{a,b} (y - Aa - Bb)^T Q_y^{-1} (y - Aa - Bb), \\ & a = \text{integer}, \quad b = \text{real} \end{aligned} \quad (2)$$

with Q_y the variance–covariance matrix of the observables. This is a non-standard least-squares problem, due

to the integer constraints on the ambiguities. This type of least-squares problem was first introduced by Teunissen (1993) and has been coined ‘integer least-squares’. Conceptually, one can divide the computation of Eq. (2) into three different steps. In the first step one simply disregards the integer constraints on the ambiguities and performs a standard least-squares adjustment. As a result one obtains the (real-valued) least-squares estimates of a and b , together with their variance–covariance matrix

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}; \quad \begin{bmatrix} Q_{\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}} \end{bmatrix} \quad (3)$$

This solution is often referred to as the ‘float’ solution. If we assume that the noise vector e of Eq. (1) is normally distributed with zero mean and variance–covariance matrix Q_y , then \hat{a} is normally distributed with integer mean a and variance–covariance matrix $Q_{\hat{a}}$. Thus

$$\hat{a} \sim N(a, Q_{\hat{a}}), \quad a = \text{integer} \quad (4)$$

In the second step the ‘float’ ambiguity estimate \hat{a} and its variance–covariance matrix are used to compute the corresponding integer ambiguity estimate. This implies that one has to solve the minimization problem

$$\min_a (\hat{a} - a)^T Q_{\hat{a}}^{-1} (\hat{a} - a), \quad a = \text{integer} \quad (5)$$

Its solution will be denoted as \check{a} . Since the mapping from \hat{a} to \check{a} is nonlinear, the distribution of the integer least-squares estimator will no longer be normal (Gaussian). In fact, due to the mapping involved, it generally becomes very difficult to compute the exact distribution. It can be shown, however, that it is an admissible and min–max estimator. Approximations or bounds can be obtained for its probability of correct integer estimation. Examples can be found in works by Teunissen et al. (1996) and Teunissen (1997b, c) and also in Sect. 4 of the present paper.

Once the integer ambiguities are computed, they are finally used in the third step to correct the ‘float’ estimate of b . As a result one obtains the ‘fixed’ solution

$$\check{b} = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} (\hat{a} - \check{a}) \quad (6)$$

The unbiasedness of the integer least-squares estimator was proven by Teunissen (1998). This implies that, since the ‘float’ solutions are also unbiased

$$E\{\check{a}\} = E\{\hat{a}\} = a \quad \text{and} \quad E\{\check{b}\} = E\{\hat{b}\} = b$$

with $E\{\cdot\}$ the expectation operator. However, the ‘fixed’ solution will not be normally distributed, but it is approximately normal when the probability of correct integer estimation is sufficiently close to one. In that case, the ‘fixed’ solution also inherits the very high precision that corresponds with the case that the integer ambiguities are non-stochastic. In fact, this is the whole principle on which GPS ambiguity resolution is based.

From a computational point of view, the most difficult part in the above three steps is the solution of Eq. (5). It requires the minimization of a quadratic form

over the whole n -dimensional space of integers. In order to tackle this problem, we first replace the whole space of integers by a smaller set of integers. This is the so-called ambiguity search space. It is in this local space that the search for the integer least-squares solution is performed. In the case of the GPS, however, this search is seriously hindered by the fact that the (real-valued) least-squares ambiguities are usually highly correlated. To remedy this situation, the least-squares ambiguity decorrelation adjustment (LAMBDA) was introduced (Teunissen 1993); for details of its performance see, e.g. de Jonge and Tiberius (1996) Boon and Ambrosius (1997), Jonkman (1998), Teunissen et al. (1997). By using this method, the original DD ambiguities are replaced by a set of transformed ambiguities which have the property of being very precise and largely decorrelated; see also the textbooks of Kleusberg and Teunissen (1996), Hofmann-Wellenhof et al. (1997) and Strang and Borre (1997). The ambiguity decorrelation process has the very beneficial effect of moulding the search space such that it transforms from a highly elongated ellipsoid to one that corresponds more closely to a sphere.

2.2 Setting the size of the ambiguity search space

The ambiguity search space is defined as the set of gridpoints a that satisfy

$$(\hat{a} - a)^T Q_{\hat{a}}^{-1} (\hat{a} - a) \leq \chi^2 \quad (7)$$

The positive constant χ^2 is chosen such that it guarantees that the search space contains the solution sought. This can be achieved by computing the constant as

$$\chi^2 = (\hat{a} - a^0)^T Q_{\hat{a}}^{-1} (\hat{a} - a^0)$$

in which a^0 is an arbitrary integer vector. However, in order to avoid the search space containing an abundance of unnecessary grid points, it helps if the constant can be chosen such that the size of the search space becomes small, while at the same time it remains guaranteed that it contains the integer least-squares solution sought. For that purpose a^0 should be a sufficiently good approximation to the integer least-squares solution. It may occasionally even happen that this approximation coincides with the integer least-squares solution. Of course, if this happens, a further search is no longer needed. A simple check as to whether the approximate solution coincides with the integer least-squares solution can be performed as follows. Using the cosine-rule-based decomposition

$$\begin{aligned} \|\hat{a} - a^0 + \nabla\|^2 &= \|\hat{a} - a^0\|^2 + \|\nabla\|^2 \\ &\quad + 2\|\hat{a} - a^0\|\|\nabla\| \cos \alpha \end{aligned}$$

with α the angle between $\hat{a} - a^0$ and the integer vector ∇ and the norm taken with respect to the metric defined by the ambiguity variance-covariance matrix, it follows that a^0 will coincide with the integer least-squares solution when the sum of the last two terms is larger than or equal to zero for any non-zero integer vector ∇ .

This will certainly be the case when $\|\nabla\| \geq 2\|\hat{a} - a^0\|$ and thus also when $1/\sqrt{\lambda_{\max}} \geq 2\|\hat{a} - a^0\|$, where λ_{\max} equals the largest eigenvalue of the ambiguity variance-covariance matrix. Hence, a sufficient check to see whether a^0 equals the integer least-squares solution is to check whether the reciprocal of the square root of the largest eigenvalue is larger than or equal to twice the distance between \hat{a} and a^0 . In order to obtain looser upper bounds, one may also replace the largest eigenvalue by the largest variance of \hat{a}_i or by the largest sequential conditional variance.

There are two easy methods of computing an approximate integer solution. The first involves applying a componentwise rounding to the entries of \hat{a} and the second a sequential rounding to the entries of \hat{a} . Results that show how well these approximate solutions perform in setting the size of the search space can be found in Teunissen et al. (1996) and de Jonge and Tiberius (1996). As is shown in these studies, the approximate solutions perform far better when they are preceded by the ambiguity decorrelation process. This also holds true for their respective probabilities of correct integer estimation. These probabilities will be developed in the following sections.

3 Integer ambiguity rounding

The simplest method of obtaining an approximation to the integer least-squares solution is to apply a componentwise rounding scheme to the entries of \hat{a} . This approximation may be used to set the size of the search space, but it can also be used as an integer estimator in its own right. Although it will be less optimal than the integer least-squares estimator, for particular applications at hand it may still perform sufficiently well. Diagnostics to evaluate whether or not this is the case are presented in this section.

3.1 The scalar case

Let us start with the simplest case that the ‘float’ solution \hat{a} is a scalar. Thus we assume that there is only one single ambiguity. If we denote ‘rounding to the nearest integer’ by $[\cdot]$, the integer nearest to \hat{a} reads

$$\check{a}_R = [\hat{a}] \quad (8)$$

where the subscript ‘ R ’ refers to the fact that the integer is obtained through rounding. The probability mass function of this integer estimator is given as

$$P(\check{a}_R = i) = \int_{(i-a)-\frac{1}{2}}^{(i-a)+\frac{1}{2}} \frac{1}{\sqrt{2\pi}\sigma_{\hat{a}}} \exp\left\{-\frac{1}{2}\hat{a}^2/\sigma_{\hat{a}}^2\right\} d\hat{a} \quad (9)$$

where i ranges over the set of integers and $\sigma_{\hat{a}}$ denotes the standard deviation of \hat{a} . Hence, the probabilities are obtained through an integration of the normal distribution over intervals of length 1 which are all centred at an integer. Since the normal distribution is symmetric about its mean and since the mean of \hat{a} is an integer, it

immediately follows that the above integer estimator is unbiased, i.e. its expectation equals the integer a .

In order to infer the quality of the integer estimator for the purpose of ambiguity resolution, we require the probability of the event that Eq. (8) coincides with the true but unknown integer ambiguity a . This probability reads as

$$P(\check{a}_R = a) = P(|\hat{a} - a| \leq \frac{1}{2}) \quad (10)$$

This probability can be evaluated by means of Eq. (9). The probability of rounding to the correct integer value then becomes

$$P(\check{a}_R = a) = 2\Phi\left(\frac{1}{2\sigma_{\hat{a}}}\right) - 1 \quad (11)$$

with

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}z^2) dz$$

Note that the probability of rounding to the correct integer value increases as the standard deviation of \hat{a} becomes smaller. This is also what one would expect. By evaluating Eq. (11) and checking whether the probability is sufficiently close to 1, one can now make the decision as to whether or not to treat the rounded ambiguity as being deterministic.

Although the computation of Eq. (11) by means of the function $\Phi(x)$ is straightforward, it is also helpful if we can approximate the probability by means of simple analytical expressions. There are several such approximations, one of which is

$$\sqrt{1 - \exp\{-\frac{1}{2}x^2\}} \leq P(\check{a}_R = a) \leq \sqrt{1 - \exp\{-x^2\}}$$

with $x = 1/2\sigma_{\hat{a}}$; see e.g. Johnson et al. (1994).

3.2 The vectorial case

We now assume that \hat{a} is a vector of order n , normally distributed with integer mean a and variance–covariance matrix $Q_{\hat{a}}$. Componentwise rounding to the nearest integer gives the integer ambiguity vector

$$\check{a}_R = ([\hat{a}_1], \dots, [\hat{a}_n])^T \quad (12)$$

The probability of rounding to the correct integer ambiguity vector a now reads

$$P(\check{a}_R = a) = P(\bigcap_{i=1}^n \{|\hat{a}_i - a_i| \leq \frac{1}{2}\}) \quad (13)$$

It is the probability that \hat{a} lies in the n -dimensional cube, centred at a and having sides of length 1. This probability is easy to evaluate when the ambiguities are fully decorrelated, i.e. when the ambiguity variance matrix is diagonal. In that case the problem decouples into n independent scalar problems of the type of Eq. (10). The probability Eq. (13) then equals the product of n probabilities of the type of Eq. (11). The exact evaluation of the probability Eq. (13) becomes very difficult when the ambiguity variance–covariance matrix

is non-diagonal. Unfortunately, this is the case in actual practice. However, although it is difficult to evaluate Eq. (13) exactly in the correlated case, it is possible to formulate a lower bound for it. This lower bound is given by the probability corresponding to the decorrelated case. Thus we have

$$\prod_{i=1}^n \left(2\Phi\left(\frac{1}{2\sigma_{\hat{a}_i}}\right) - 1\right) \leq P(\check{a}_R = a) \quad (14)$$

This lower bound can now be used to check whether the simple estimator ‘round to the nearest integer’ guarantees sufficient success of obtaining the correct integer ambiguity vector. One simply has to evaluate the lower bound and check whether it is sufficiently close to 1.

Note that the lower bound is only dependent on the diagonal entries of the ambiguity variance–covariance matrix. Hence, this lower bound is not invariant for the class of admissible ambiguity transformations (Teunissen 1995). Since the precision of the individual DD ambiguities is usually rather poor, the lower bound of Eq. (14) will usually be rather loose when applied to the DD ambiguities. However, the lower bound becomes much sharper when it is applied to ambiguities which are almost decorrelated. This shows that, for an actual application, the above lower bound should be evaluated for the decorrelated ambiguities obtained through the LAMBDA method. Since the transformed ambiguities obtained by this method are far more precise than the original DD ambiguities, the lower bound becomes sharper due to its increase in value.

4 Integer ambiguity bootstrapping

Another easy method of obtaining an approximation to the integer least-squares solution is to apply a sequential rounding scheme to the entries of \hat{a} . This approximation may also be used as an integer estimator in its own right. In this section we will present its probability of correct integer estimation and show how it is related to the results of the previous section.

4.1 Probability of correct integer estimation

The integer estimation scheme of componentwise rounding does not take the ambiguity correlations into account. It simply treats the multivariate integer estimation problem as if it were a problem consisting of n independent scalar estimation problems. A method which does take some of the correlation into account is the ‘sequential integer rounding’ method, which also is referred to as the ‘integer bootstrapping’ method, see e.g. Blewitt (1989) and Dong and Bock (1989). This method is a generalization of the ‘integer rounding’ method, and it goes as follows. If n ambiguities are available, one starts with the first ambiguity \hat{a}_1 and, as before, rounds its value to the nearest integer. Having obtained the integer value of this first ambiguity, the real-valued estimates of all remaining ambiguities are

then corrected by virtue of their correlation with the first ambiguity. Then the second, but now corrected, real-valued ambiguity estimate is rounded to its nearest integer. Having obtained the integer value of the second ambiguity, the real-valued estimates of all remaining $n - 2$ ambiguities are then again corrected, but now by virtue of their correlation with the second ambiguity. This process is continued until all ambiguities are taken care of. In essence this ‘bootstrapping’ technique boils down to the use of a sequential conditional least-squares adjustment (Teunissen 1993, 1996), with a conditioning on the integer ambiguity values obtained in the previous steps. The components of the integer ambiguity vector so obtained therefore read

$$\begin{aligned} [\hat{a}_1] &= [\hat{a}_1] \\ [\hat{a}_{2|1}] &= [\hat{a}_2 - \sigma_{\hat{a}_2 \hat{a}_1} \sigma_{\hat{a}_1}^{-2} (\hat{a}_1 - [\hat{a}_1])] \\ \vdots &\quad \vdots \\ [\hat{a}_{n|N}] &= [\hat{a}_n - \sum_{i=1}^{n-1} \sigma_{\hat{a}_n \hat{a}_{i|I}} \sigma_{\hat{a}_{i|I}}^{-2} (\hat{a}_{i|I} - [\hat{a}_{i|I}])] \end{aligned} \quad (15)$$

where the shorthand notation $\hat{a}_{i|I}$ stands for the i th least-squares ambiguity obtained through a conditioning on the previous $I = \{1, \dots, (i-1)\}$ sequentially rounded ambiguities. The integer bootstrapped solution therefore reads

$$\check{a}_B = ([\hat{a}_1], \dots, [\hat{a}_{n|N}])^T \quad (16)$$

Note that this integer estimator is a generalization of the previous one [Eq. (12)]. The two integer estimators are identical in the case that the variance–covariance matrix is diagonal and they differ when this matrix is non-diagonal. In the non-diagonal case, however, the bootstrapped estimator still makes use of the simple ‘integer rounding’ operation, but now by taking the (sequential) correlations into account as well. The bootstrapped probability of correct integer estimation is given as

$$P(\check{a}_B = a) = P(\cap_{i=1}^n \{|\hat{a}_{i|I} - a_i| \leq \frac{1}{2}\}) \quad (17)$$

In contrast with Eq. (13), this bootstrapped probability has a useful property in that it can be computed exactly in a rather straightforward manner. This follows from applying the chain rule of conditional probabilities

$$\begin{aligned} P(\check{a}_B = a) &= \prod_{i=1}^n P([\hat{a}_{i|I}] = a_i | [\hat{a}_1] = a_1, \\ &\quad \dots, [\hat{a}_{(i-1)|(I-1)}] = a_{i-1}) \end{aligned} \quad (18)$$

Each of the probabilities in this chain is of the type of Eq. (11), but now with the ambiguity standard deviations replaced by the sequential conditional standard deviations. Hence, the bootstrapped probability of obtaining the correct integer ambiguity vector reads as

$$P(\check{a}_B = a) = \prod_{i=1}^n (2\Phi\left(\frac{1}{2\sigma_{\hat{a}_{i|I}}}\right) - 1) \quad (19)$$

This probability can now be used to infer whether or not it makes sense to use the integer bootstrapped solution.

Note, that since each of the individual probabilities in the product of Eq. (19) is smaller than one, the overall probability has the tendency to get smaller as the dimension n increases. Hence, for a particular application it could well be that the overall probability is too small, while the product of the first $j < n$ terms is still large enough to render successful fixing of the first j ambiguities possible. In that case Eq. (19) can be used to study the success rate of a partial fixing of the ambiguities. Also note that the probability is not invariant against a reparametrization of the ambiguities. In fact, the bootstrapped solution and its corresponding probability are not even invariant against a reordering or permutation of the ambiguities. Hence, as was the case with the previous integer estimator, it also makes sense for the bootstrapped estimator to first transform the ambiguities to a new set of more precise ambiguities before commencing with the bootstrapping.

4.2 The probability of bootstrapping and rounding

The relative performance of the ‘integer bootstrapping’ estimator and the ‘integer rounding’ estimator can be evaluated if we compare their respective probabilities of correct integer estimation. Since the bootstrapped solution takes part of the correlation structure into account, whereas the solution obtained through rounding does not, one will be inclined to think that the bootstrapped probability of correct integer estimation is larger than or at least as large as the probability that corresponds with rounding. Hence, the conjecture is that the inequality

$$P(\check{a}_R = a) \leq P(\check{a}_B = a) \quad (20)$$

holds true. In order to prove this inequality we will start with the probability of correct rounding. It reads

$$\begin{aligned} P(\check{a}_R = a) &= \int_R \frac{1}{(2\pi)^{\frac{1}{2}n} \sqrt{\det(Q_{\hat{a}})}} \\ &\quad \times \exp\left\{-\frac{1}{2}(\hat{a} - a)^T Q_{\hat{a}}^{-1}(\hat{a} - a)\right\} d\hat{a} \end{aligned} \quad (21)$$

with the region of integration

$$R = \left\{ (\hat{a}_1, \dots, \hat{a}_n)^T \in \mathbf{R}^n \mid |\hat{a}_i - a_i| \leq \frac{1}{2}, i = 1, \dots, n \right\}$$

This region is a cube having all side lengths equal to 1. The above integral is parametrized in terms of \hat{a}_i . In order to obtain a link with the bootstrapped probability of correct integer estimation, we will reparametrize the integral in terms of $\hat{a}_{i|I}$. The corresponding transformation reads

$$\hat{a}_i = \hat{a}_{i|I} + \sum_{j=1}^{i-1} \sigma_{\hat{a}_i \hat{a}_{j|J}} \sigma_{\hat{a}_{j|J}}^{-2} (\hat{a}_{j|J} - a_j) \quad \text{for } i = 1, \dots, n \quad (22)$$

In order to apply this transformation to the integral of Eq.(21), we need the general transformation formula for integrals. This reads (Fleming, 1977)

$$\int_R f(y) dy = \int_{g^{-1}(R)} f(g(x)) |\det(\partial_x g(x))| dx$$

By letting Eq. (22) play the role of $y = g(x)$, we obtain, noting that the matrix of transformation (22) is unit triangular and therefore that the determinant of the Jacobian equals one

$$\begin{aligned} P(\check{a}_R = a) &= \int_{g^{-1}(R)} \frac{1}{(2\pi)^{\frac{1}{2}n} \prod_{i=1}^n \sigma_{\hat{a}_{i|I}}} \\ &\quad \times \prod_{i=1}^n \exp \left\{ -\frac{1}{2} \left(\frac{\hat{a}_{i|I} - a_i}{\sigma_{\hat{a}_{i|I}}} \right)^2 \right\} \prod_{i=1}^n d\hat{a}_{i|I} \end{aligned} \quad (23)$$

with the region of integration

$$g^{-1}(R) = \{(\hat{a}_1, \hat{a}_{2|1}, \dots, \hat{a}_{n|N})^T \in \mathbf{R}^n | S_i, i = 1, \dots, n\}$$

where

$$S_i = \left\{ \left| (\hat{a}_{i|I} - a_i) + \sum_{j=1}^{i-1} \sigma_{\hat{a}_i \hat{a}_{j|J}} \sigma_{\hat{a}_{j|J}}^{-2} (\hat{a}_{j|J} - a_j) \right| \leq \frac{1}{2} \right\}$$

This region of integration is now no longer a cube. However, due to the product form of the integral and the sequential form of the intervals in $g^{-1}(R)$, it follows that when we write the integral as

$$\begin{aligned} P(\check{a}_R = a) &= \int_{g^{-1}(R)|S_n} F_n \prod_{i=1}^{n-1} \frac{1}{\sqrt{2\pi} \sigma_{\hat{a}_{i|I}}} \\ &\quad \times \exp \left\{ -\frac{1}{2} \left(\frac{\hat{a}_{i|I} - a_i}{\sigma_{\hat{a}_{i|I}}} \right)^2 \right\} \prod_{i=1}^{n-1} d\hat{a}_{i|I} \\ F_n &= \int_{S_n} \frac{1}{\sqrt{2\pi} \sigma_{\hat{a}_{n|N}}} \exp \left\{ -\frac{1}{2} \left(\frac{\hat{a}_{n|N} - a_n}{\sigma_{\hat{a}_{n|N}}} \right)^2 \right\} d\hat{a}_{n|N} \end{aligned}$$

the dependence on $\hat{a}_{n|N}$ is completely captured in F_n . We are now in a position to formulate an upper bound. Note that the interval S_n is not centred at the mean a_n . Hence, this integral of the standard normal distribution will increase in value if the interval S_n is replaced by the interval $B_n = \{|\hat{a}_{n|N} - a_n| \leq 1/2\}$. This shows that

$$F_n \leq \left(2\Phi \left(\frac{1}{2\sigma_{\hat{a}_{n|N}}} \right) - 1 \right)$$

By repeating these steps for the other variables also, we finally obtain the upper bound

$$P(\check{a}_R = a) \leq \prod_{i=1}^n \left(2\Phi \left(\frac{1}{2\sigma_{\hat{a}_{i|I}}} \right) - 1 \right)$$

which concludes the proof of Eq. (20). Note that when the upper bound Eq. (20), is combined with the lower bound, Eq. (14), we have an easy-to-compute interval that completely bounds the probability of correct integer

rounding. To reduce the length of the interval, the bounds should be evaluated with the transformed and decorrelated ambiguities.

4.3 An invariant upper bound

We have already observed that the bootstrapped probability [Eq. (19)] depends on the chosen ambiguity parametrization. Hence, by using more precise transformed ambiguities instead of the DD ambiguities, one can increase the probability of correct integer estimation. Although this probability can be computed exactly for every chosen parametrization, it would still be helpful if we could come up with an upper bound which is easy to compute and at the same time invariant for the choice of parametrization. Such an upper bound can be found as follows. Note that Eq. (17) equals the integral of a normal distribution over an n -dimensional box. When the side lengths are taken equal to the reciprocal value of the sequential conditional standard deviations, the normal distribution takes its standard form of having zero mean and a unit variance–covariance matrix. This implies, if we vary the side lengths but constrain the volume of the box to be constant, that this probability reaches its maximum for a cube. Since this cube must have the same volume as the original box, its side lengths are all equal to the geometric average of the reciprocal sequential conditional standard deviations. Hence, we have

$$\begin{aligned} P \left(\bigcap_{i=1}^n \left\{ \frac{|\hat{a}_{i|I} - a_i|}{\sigma_{\hat{a}_{i|I}}} \leq \frac{1}{2\sigma_{\hat{a}_{i|I}}} \right\} \right) \\ \leq P \left(\bigcap_{i=1}^n \left\{ \frac{|\hat{a}_{i|I} - a_i|}{\sigma_{\hat{a}_{i|I}}} \leq \frac{1}{2} \sqrt[n]{\prod_{i=1}^n \frac{1}{\sigma_{\hat{a}_{i|I}}}} \right\} \right) \end{aligned}$$

In this expression we recognize on the right-hand side the ADOP (Ambiguity Dilution of Precision) which was introduced by Teunissen (1997b) as

$$\text{ADOP} = \sqrt{\det Q_a^{\frac{1}{n}}} \quad (\text{cycle})$$

It measures the average precision of the ambiguities and it is invariant for the class of admissible ambiguity transformations (Teunissen and Odijk 1997). The above upper bound can thus be formulated in a compact way as

$$P(\check{a}_B = a) \leq \left(2\Phi \left(\frac{1}{2 \text{ADOP}} \right) - 1 \right)^n \quad (24)$$

Since the ADOP is invariant for all admissible ambiguity transformations, this also holds true for the above upper bound. The upper bound can now be used, independently of which parametrization is used, to check whether it makes sense to consider the bootstrapping method as a viable option for integer estimation. The bootstrapping method should not be used when the upper bound is too small. Closed form expressions for the determinant of the ambiguity variance–covariance matrix are given by Teunissen (1997c). By means of

these closed form expressions it becomes obvious in what way the various factors such as number of satellites tracked, number of observation epochs used, type of observables used and amount of change in relative receiver–satellite geometry, contribute to the above bound.

5 Conclusions

In this contribution we have studied two simple alternatives of the integer least-squares estimator. Although they are less optimal, they have the advantage that no search at all is needed for their actual computation. However, before the choice is made to opt for one of these two integer estimators, one should make sure that their probability of correct integer estimation is sufficiently close to one. This is especially of importance for GPS ambiguity resolution, where in all subsequent processing steps the integer ambiguity solution is treated as non-stochastic.

The first integer estimator considered was the one that corresponds to a componentwise rounding of the ‘float’ solution. Although it is difficult to compute its probability of correct integer estimation exactly, it was shown that this probability is bounded as

$$\prod_{i=1}^n \left(2\Phi\left(\frac{1}{2\sigma_{\hat{a}_i}}\right) - 1 \right) \leq P(\check{a}_R = a) \leq \prod_{i=1}^n \left(2\Phi\left(\frac{1}{2\sigma_{\hat{a}_{ijl}}}\right) - 1 \right)$$

Since the probability (as well as its two bounds) is dependent on the choice of ambiguity parametrization, one should apply this estimator only after the ambiguity decorrelation process has been applied.

The second integer estimator considered was the bootstrapped one. It follows from a sequential rounding of the entries of the ‘float’ solution. Since the variance–covariance matrix of the sequential conditional least-squares solution is diagonal, the probability of correct integer estimation of the bootstrapped solution can be computed exactly. It was shown that

$$\prod_{i=1}^n \left(2\Phi\left(\frac{1}{2\sigma_{\hat{a}_{ijl}}}\right) - 1 \right) = P(\check{a}_B = a) \leq \left(2\Phi\left(\frac{1}{2\text{ADOP}}\right) - 1 \right)^n$$

Hence, with the bootstrapped solution one has a better chance of obtaining the correct integer ambiguities. Here, the remark that the evaluation should be based on the reparametrized and decorrelated ambiguities again applies. Although the probability of correct integer estimation can be computed exactly for the bootstrapped solution, the above given upper bound is still

very useful. First, it is invariant for the choice of ambiguity parametrization, and second, simple closed-form expressions are available for evaluating the ADOP of the ambiguity variance–covariance matrix.

References

- Blewitt G (1989) Carrier-phase ambiguity resolution for the Global Positioning System applied to geodetic baselines up to 2000 km. *J Geophys Res* 94 (B8): 10.187–10.302
- Boon F, Ambrosius B (1997) Results of real-time application of the LAMBDA method in GPS based aircraft landings. *Proc. KIS97*, pp 339–345
- de Jonge PJ (1998) A processing strategy for the application of the GPS in networks. *Netherlands Geodetic Commission, Publications on Geodesy*, No. 46.
- de Jonge PJ, Tiberius CCJM (1996) The LAMBDA method for integer ambiguity estimation: implementation aspects LGR-Series no. 12. Delft Geodetic Computing Centre, 49 pp
- Dong D, Bock Y (1989) Global Positioning System network analysis with phase ambiguity resolution applied to crustal deformation studies in California. *J Geophys Res* 94: 3949–3966
- Fleming W (1977) Functions of Several Variables. Springer, Berlin Heidelberg New York
- Hofmann-Wellenhof B, Lichtenegger H, Collins J (1997) Global positioning system: theory and practice. 4th edn. Springer, Berlin Heidelberg New York
- Johnson NL, Kotz S, Balakrishnan N (1994) Continuous univariate distributions, vol 1, 2nd edn. Wiley, New York
- Jonkman NF (1998) Integer GPS ambiguity estimation without the receiver–satellite geometry. LGR-Series no. 18. Delft Geodetic Computing Centre, 95 pp
- Kleusberg A, Teunissen PJG (eds) (1996) GPS for Geodesy. Lecture Notes in Earth Sciences, vol 60. Springer, Berlin Heidelberg New York
- Strang G, Borre K (1997) Linear algebra, geodesy, and GPS. Wellesley-Cambridge Press, Cambridge, Massachusetts
- Teunissen PJG (1993) Least-squares estimation of the integer GPS ambiguities. Invited Lecture, Sect IV theory and methodology, IAG general meeting, Beijing. Also in: LGR-Series no. 6. Delft Geodetic Computing Centre
- Teunissen PJG (1995) The invertible GPS ambiguity transformations. *Manusci Geod* 20: 489–497
- Teunissen PJG (1996) GPS carrier phase ambiguity fixing concepts. In: Kleusberg A, Teunissen PJG (eds) GPS for geodesy. Lecture Notes in Earth Sciences, vol 60. Springer, Berlin Heidelberg New York, pp 263–335
- Teunissen PJG (1997a) Some remarks on GPS ambiguity resolution. *Artificial Satellites* 32(3), pp 119–130
- Teunissen PJG (1997b) A canonical theory for short GPS baselines, Parts I–IV. *J Geod* 71: 320–336, 389–401, 486–501, 513–525
- Teunissen PJG (1997c) Closed form expressions for the volume of the GPS ambiguity search spaces. *Artif Satellites* 32(1): 5–20
- Teunissen PJG (1998) A class of unbiased integer GPS ambiguity estimators. *Artif Satellites* 33(1), pp 3–10
- Teunissen PJG, Odijk D (1997) Ambiguity dilution of precision: definition, properties and application. *Proc ION GPS-97*, vol 1, pp 891–899
- Teunissen PJG, de Jonge PJ, Tiberius CCJM (1996) The volume of the GPS ambiguity search space and its relevance for integer ambiguity resolution. *Proc ION GPS-96*, pp 889–898
- Teunissen PJG, de Jonge PJ, Tiberius CCJM (1997) The performance of the LAMBDA method for fast GPS ambiguity resolution. *Navigation* 44(3): 373–383

An optimality property of the integer least-squares estimator

P. J. G. Teunissen

Department of Mathematical Geodesy and Positioning, Faculty of Civil Engineering and Geosciences,
 Delft University of Technology, Thijsseweg 11, 2629 JA Delft, The Netherlands
 e-mail: P.J.G.Teunissen@geo.tudelft.nl; Fax: +31 15 278 3711

Received: 11 January 1999 / Accepted: 9 July 1999

Abstract. A probabilistic justification is given for using the integer least-squares (LS) estimator. The class of admissible integer estimators is introduced and classical adjustment theory is extended by proving that the integer LS estimator is best in the sense of maximizing the probability of correct integer estimation. For global positioning system ambiguity resolution, this implies that the success rate of any other integer estimator of the carrier phase ambiguities will be smaller than or at the most equal to the ambiguity success rate of the integer LS estimator. The success rates of any one of these estimators may therefore be used to provide lower bounds for the LS success rate. This is particularly useful in case of the bootstrapped estimator.

Key words. Integer LS · GPS ambiguity resolution · Ambiguity success rate

1 Introduction

Ambiguity resolution applies to a great variety of global positioning system (GPS) models currently in use. These range from single-baseline models used for kinematic positioning to multi-baseline models used as a tool for studying geodynamic phenomena. An overview of these and other GPS models, together with their application in surveying, navigation and geodesy, can be found in textbooks such as those of Leick (1995), Parkinson and Spilker (1996), Hofmann-Wellenhof et al. (1997), Strang and Borre (1997) and Teunissen and Kleusberg (1998). Despite the differences in application of the various GPS models, it is important to understand that their ambiguity resolution problems are intrinsically the same. That is, the GPS models on which ambiguity resolution is based can all be cast in the following conceptual frame of linear(ized) observation equations

$$y = Aa + Bb + e \quad (1)$$

where y is the given GPS data vector of order m , a and b are the unknown parameter vectors respectively of order n and o , and e is the noise vector. The matrices A and B are the corresponding design matrices. The data vector y will usually consist of the ‘observed minus computed’ single- or dual-frequency double-difference (DD) phase and/or pseudorange (code) observations accumulated over all observation epochs. The entries of vector a are then the DD carrier phase ambiguities, expressed in units of cycles rather than range. They are known to be integers, $a \in \mathbb{Z}^n$. The entries of the vector b will consist of the remaining unknown parameters, such as for instance baseline components (coordinates) and possibly atmospheric delay parameters (troposphere, ionosphere). They are known to be real-valued, $b \in \mathbb{R}^o$.

The procedure which is usually followed for solving the GPS model of Eq. (1) can be divided into three steps (for more details we refer to e.g. Teunissen 1993 or de Jonge and Tiberius 1996). In the *first* step we simply disregard the integer constraints $a \in \mathbb{Z}^n$ on the ambiguities and perform a standard adjustment. As a result we obtain the (real-valued) estimates of a and b , together with their variance-covariance matrix

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}, \quad \begin{bmatrix} Q_{\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}} \end{bmatrix} \quad (2)$$

This solution is referred to as the ‘float’ solution. In the *second* step the ‘float’ ambiguity estimate \hat{a} is used to compute the corresponding integer ambiguity estimate \check{a} . This implies that a mapping $F : \mathbb{R}^n \mapsto \mathbb{Z}^n$, from the n -dimensional space of real numbers to the n -dimensional space of integers, is introduced such that

$$\check{a} = F(\hat{a}) \quad (3)$$

Once the integer ambiguities are computed, they are used in the *third* step to finally correct the float estimate of b . As a result we obtain the ‘fixed’ solution

$$\check{b} = \hat{b} - Q_{\hat{b}\hat{a}}Q_{\hat{a}}^{-1}(\hat{a} - \check{a}) \quad (4)$$

The ambiguity residual ($\hat{a} - \check{a}$) is thus used to adjust the float solution so as to obtain the fixed solution.

It is of course not enough to compute the fixed solution and be done with it. We can always compute such a solution, whether it is of good quality or not. We therefore still need to address the question whether we have enough confidence in the computed integer ambiguity solution. After all, unsuccessful ambiguity resolution, when passed unnoticed, will all too often lead to unacceptable errors in the positioning results. We therefore need to have a way of knowing how often we can expect the computed ambiguity solution to coincide with the correct, but unknown, solution. Is this 9 out of 10 times, 99 out of a 100, or a higher percentage? It will certainly never equal 100%. After all, the integer ambiguities are computed from the data: they are therefore subject to uncertainty just like the data are.

In order to obtain such a description, we require the probability distribution of the integer ambiguities (Teunissen 1997). This distribution will be a probability mass function, due to the integer nature of the ambiguities. Of this probability mass function, the probability of correct integer ambiguity estimation is of particular interest. This probability will be denoted as $P(\check{a} = a)$. It describes the frequency with which one can expect to have a successful ambiguity resolution. It equals the expected ambiguity success rate. This probability depends on three contributing factors: the functional model (the observation equations), the stochastic model (the distribution and precision of the observables) and the chosen method of integer ambiguity estimation. Changes in any one of these will affect the success rate.

In this contribution the choice of integer ambiguity estimator will be considered. In general, we would like to have the highest success rate possible. We therefore would like to know which integer ambiguity estimator maximizes $P(\check{a} = a)$. For this purpose we first need to introduce a class of candidate integer estimators. Such a class of admissible integer estimators is introduced in Sect. 2. In constructing this class, we are led by practical considerations such as the following: the estimator should map any float solution to a unique integer solution and when the float solution is perturbed by an integer amount, the integer solution should be perturbed by the same integer amount. In Sect. 3 we give three examples of integer estimators which belong to this class of admissible estimators. They are the ‘rounding’ estimator, the ‘bootstrapped’ estimator and the integer least-squares (LS) estimator.

Section 4 contains the main result of this contribution. It is proven that the integer LS estimator has the largest ambiguity success rate of all admissible estimators. The integer LS estimator is therefore the best estimator in the sense of maximizing the probability of correct integer estimation. Any other integer ambiguity estimator, such as for instance the ‘rounding’ estimator or the ‘bootstrapped’ estimator, will have a smaller success rate. GPS ambiguity resolution will therefore be less successful when integer estimators other than the LS estimator are used.

2 A class of integer estimators

There are many ways of computing an integer ambiguity vector \check{a} from its real-valued counterpart \hat{a} . To each such method belongs a mapping $F : R^n \mapsto Z^n$ from the n -dimensional space of real numbers to the n -dimensional space of integers. Once this map has been defined, the integer ambiguity vector follows from its real-valued counterpart as $\check{a} = F(\hat{a})$. Due to the discrete nature of Z^n , the map F will not be one-to-one, but instead a many-to-one map. This implies that different real-valued ambiguity vectors may be mapped to the same integer vector. We can therefore assign a subset $S_z \subset R^n$ to each integer vector $z \in Z^n$:

$$S_z = \{x \in R^n \mid z = F(x)\}, \quad z \in Z^n \quad (5)$$

The subset S_z contains all real-valued ambiguity vectors that will be mapped by F to the same integer vector $z \in Z^n$. This subset is referred to as the *pull-in region* of z (Jonkman 1998; Teunissen 1998a). It is the region in which all ambiguity float solutions are pulled to the same fixed ambiguity vector z .

Having defined the pull-in regions, we are now in a position to give an explicit expression for the corresponding integer ambiguity estimator. It reads

$$\check{a} = \sum_{z \in Z^n} z s_z(\hat{a}) \text{ with } s_z(\hat{a}) = \begin{cases} 1 & \text{if } \hat{a} \in S_z \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Since the pull-in regions define the integer estimator completely, we can define a class of integer estimators by listing properties of these pull-in regions. In this section we introduce three properties for which it seems reasonable that they are possessed by the pull-in regions.

It seems reasonable to ask of the pull-in regions that their union covers the n -dimensional space completely

$$\bigcup_{z \in Z^n} S_z = R^n \quad (7)$$

Otherwise we would have gaps, in which case not every $\hat{a} \in R^n$ could be assigned to a corresponding integer ambiguity vector.

Another property that we require of the pull-in-regions is that any two distinct regions should not have an overlap. Otherwise we could end up in a situation where a float solution $\hat{a} \in R^n$ cannot be assigned uniquely to a single integer vector. For the interior points of two distinct pull-in regions we therefore require

$$S_{z_1} \bigcap S_{z_2} = \emptyset, \quad \forall z_1, z_2 \in Z^n, \quad z_1 \neq z_2 \quad (8)$$

We allow the pull-in regions to have common boundaries, however. This is permitted if we assume zero probability that \hat{a} lies on one of the boundaries. This will be the case when the probability density function (pdf) of \hat{a} is continuous.

The third and last property asked is that the integer map F possesses the property that $F(x+z) = F(x) + z$, $\forall x \in R^n$, $z \in Z^n$. This property is also a reasonable one to request. It states that when the float

solution is moved by an integer amount z , the corresponding integer solution is moved by the same integer amount. This property allows one to use the ‘integer remove–restore’ technique: $F(\hat{a} - z) + z = F(\hat{a})$. It therefore allows us to work with the fractional parts of the entries of \hat{a} , instead of with its complete entries, which may sometimes be large numbers.

The integer remove–restore property implies that

$$\begin{aligned} S_{z_1+z_2} &= \{x \in R^n \mid z_1 + z_2 = F(x)\} \\ &= \{x \in R^n \mid z_1 = F(x) - z_2 = F(x - z_2)\} \\ &= \{x \in R^n \mid z_1 = F(y), x = y + z_2\} \\ &= S_{z_1} + z_2, \quad \forall z_1, z_2 \in Z^n. \end{aligned}$$

Hence, it means that the pull-in regions are translated copies of one another. This third property may therefore also be stated as

$$S_z = z + S_0, \quad \forall z \in Z^n \quad (9)$$

with S_0 being the pull-in region of the origin of Z^n .

Integer ambiguity estimators that possess all three of the above stated properties form a class. This class will be referred to as the class of admissible integer ambiguity estimators. It is defined as follows.

Definition. The integer estimator $\check{a} = \sum_{z \in Z^n} z s_z(\hat{a})$ is said to be *admissible* if

1. $\bigcup_{z \in Z^n} S_z = R^n$
2. $S_{z_1} \cap S_{z_2} = \emptyset, \quad \forall z_1, z_2 \in Z^n, z_1 \neq z_2$
3. $S_z = z + S_0, \quad \forall z \in Z^n$

Various integer estimators exist that belong to this class. As the definition shows, one way of constructing admissible estimators is to choose a subset S_0 such that its translated copies cover R^n without gaps and overlaps. In two dimensions this can be achieved, for instance, by choosing S_0 as the unit square centred at the origin.

3 Examples of admissible estimators

In this section three different admissible integer estimators are considered. All three of them have been in use, in one way or another, for GPS ambiguity resolution. They are the ‘rounding’ estimator, the ‘bootstrapped’ estimator and the LS estimator.

3.1 Integer rounding

The simplest way to obtain an integer vector from the real-valued float solution is to round each of the entries of \hat{a} to its nearest integer. The corresponding integer estimator reads therefore

$$\check{a}_R = ([\hat{a}_1], \dots, [\hat{a}_n])^T \quad (10)$$

where ‘[.]’ denotes rounding to the nearest integer. This estimator is clearly admissible. The first two conditions

of the definition are satisfied, since – apart from ties in rounding – any float solution $\hat{a} \in R^n$ gets mapped to a unique integer vector. The third condition is also satisfied since rounding admits the integer remove–restore technique, that is, $[x - z] + z = [x], \forall x \in R, z \in Z$.

Since componentwise rounding implies that each real-valued ambiguity estimate $\hat{a}_i, i = 1, \dots, n$, is mapped to its nearest integer, the absolute value of the difference between the two is at most $1/2$. The pull-in regions $S_{R,z}$ that belong to this integer estimator are therefore given as

$$S_{R,z} = \bigcap_{i=1}^n \left\{ x \in R^n \mid |x_i - z_i| \leq \frac{1}{2} \right\}, \quad \forall z \in Z^n \quad (11)$$

They are n -dimensional cubes, centred at $z \in Z^n$, all having sides of length one.

3.2 Integer bootstrapping

Another relatively simple integer ambiguity estimator is the bootstrapped estimator (Blewitt 1989; Dong and Bock 1989). The bootstrapped estimator can be seen as a generalization of the previous estimator. It still makes use of integer rounding, but it also takes some of the correlation between the ambiguities into account. The bootstrapped estimator follows from a sequential conditional LS adjustment and it is computed as follows. If n ambiguities are available, we start with the first ambiguity \hat{a}_1 , and round its value to the nearest integer. Having obtained the integer value of this first ambiguity, the real-valued estimates of all remaining ambiguities are then corrected by virtue of their correlation with the first ambiguity. Then the second, but now corrected, real-valued ambiguity estimate is rounded to its nearest integer. Having obtained the integer value of the second ambiguity, the real-valued estimates of all remaining $n - 2$ ambiguities are then again corrected, but now by virtue of their correlation with the second ambiguity. This process is continued until all ambiguities are considered. The components of the bootstrapped estimator \check{a}_B are given as

$$\begin{aligned} \check{a}_{B,1} &= [\hat{a}_1] \\ \check{a}_{B,2} &= [\hat{a}_{2|1}] = [\hat{a}_2 - \sigma_{\hat{a}_2 \hat{a}_1} \sigma_{\hat{a}_1}^{-2} (\hat{a}_1 - \check{a}_{B,1})] \\ &\vdots \\ \check{a}_{B,n} &= [\hat{a}_{n|N}] = \left[\hat{a}_n - \sum_{i=1}^{n-1} \sigma_{\hat{a}_n \hat{a}_{i|I}} \sigma_{\hat{a}_{i|I}}^{-2} (\hat{a}_{i|I} - \check{a}_{B,i}) \right] \end{aligned} \quad (12)$$

where the shorthand notation $\hat{a}_{i|I}$ stands for the i th LS ambiguity obtained through a conditioning on the previous $I = \{1, \dots, (i-1)\}$ sequentially rounded ambiguities.

The bootstrapped estimator is admissible. The first two conditions of the definition are satisfied, since – apart from ties in rounding – any float solution gets mapped to a unique integer ambiguity vector. The integer remove–restore technique again applies. To see this, let \check{a}'_B be the bootstrapped estimator which corre-

sponds with $\hat{a}' = \hat{a} - z$. It then follows from Eq. (12) that $\check{a}_B = \check{a}'_B + z$.

The real-valued sequential conditional LS solution can be obtained by means of the triangular decomposition of the ambiguity variance–covariance matrix. Let the LDU decomposition of the variance–covariance matrix be given as $Q_{\hat{a}} = LDL^T$, with L a unit lower triangular matrix and D a diagonal matrix. Then $(\hat{a} - z) = L(\hat{a}^c - z)$, where \hat{a}^c denotes the conditional LS solution obtained from a sequential conditioning on the entries of z . The variance–covariance matrix of \hat{a}^c is given by the diagonal matrix D . This shows, when a componentwise rounding is applied to \hat{a}^c , that z is the integer solution of the bootstrapped method. Thus \check{a}_B satisfies $[L^{-1}(\hat{a} - \check{a}_B)] = 0$. Hence, if c_i denotes the i th canonical unit vector having a 1 as its i th entry, the pull-in regions $S_{B,z}$ that belong to the bootstrapped estimator follow as

$$S_{B,z} = \cap_{i=1}^n \left\{ x \in R^n \mid |c_i^T L^{-1}(x - z)| \leq \frac{1}{2} \right\}, \quad \forall z \in Z^n \quad (13)$$

Note that these subsets reduce to the ones of Eq. (11) when L becomes diagonal. This is the case when the ambiguity variance–covariance matrix is diagonal. In that case the two integer estimators \check{a}_R and \check{a}_B are identical.

3.3 Integer LS

The integer LS estimator is defined as

$$\check{a}_{LS} = \arg \min_{z \in Z^n} \| \hat{a} - z \|_{Q_{\hat{a}}}^2 \quad (14)$$

where $\| \cdot \|_{Q_{\hat{a}}}^2 = (\cdot)^T Q_{\hat{a}}^{-1}(\cdot)$. This ambiguity estimator was introduced for the first time in Teunissen (1993). This estimator is also admissible. Apart from boundary ties, it produces a unique integer vector for any float solution $\hat{a} \in R^n$. And since $\check{a}_{LS} = \arg \min_{z \in Z^n} \| \hat{a} - u - z \|_{Q_{\hat{a}}}^2 + u$ holds true for any integer u , the integer remove–restore technique again applies.

It follows from Eq. (14) that the float solutions $\hat{a} \in R^n$ which are mapped to the same integer vector \check{a}_{LS} are those that lie closer to this integer vector than to any other integer vector $z \in Z^n$. This shows that the LS pull-in regions $S_{LS,z}$ consist of intersecting half-spaces, each one of which is bounded by the plane orthogonal to $(c - z)$, $c \in Z^n$ and passing through the mid-point $\frac{1}{2}(z + c)$. Here, orthogonality is taken with respect to the metric as defined by the ambiguity variance–covariance matrix. Since \hat{a} lies in one of these half-spaces when the length of the orthogonal projection of $(\hat{a} - z)$ onto $(c - z)$ is less than or equal to half the distance between c and z , it follows that

$$S_{LS,z} = \cap_{c \in Z^n} \left\{ x \in R^n \mid |w_c(x)| \leq \frac{1}{2} \| c \|_{Q_{\hat{a}}} \right\}, \quad \forall z \in Z^n \quad (15)$$

with

$$w_c(x) = \frac{c^T Q_{\hat{a}}^{-1}(x - z)}{\sqrt{c^T Q_{\hat{a}}^{-1} c}}$$

Note that $(c - z)$ has been replaced by c in Eq. (15). This is permitted since the intersection is taken with respect to all $c \in Z^n$. Also note that w_c is an example of the well-known w -test statistic for testing one-dimensional alternative hypotheses (Baarda 1968; Teunissen 1985). The absolute values of w_c are thus required to be no larger than the ‘critical values’ $\frac{1}{2} \| c \|_{Q_{\hat{a}}}$.

In our comparison of \check{a}_R and \check{a}_B , we noted that the two estimators became identical in the case that the unit triangular matrix L reduced to the identity matrix. The same holds true in the case of \check{a}_{LS} . Hence, all three estimators become identical in the case that the ambiguity variance–covariance matrix is diagonal. This condition can be relaxed, however, when comparing \check{a}_B with \check{a}_{LS} . These two estimators will already have become identical when all matrix entries of L are integer. This is the case when L is an admissible ambiguity transformation (Teunissen 1995). Thus, $\check{a}_{LS} = \check{a}_B = L[L^{-1}\hat{a}]$.

4 Maximizing the ambiguity success rate

In this section the main result of this contribution is discussed. So far, we have introduced a class of admissible integer estimators and discussed some of its members. We thus have now a variety of reasonable integer estimators available. The question which arises next is which of these estimators to choose? Does an estimator exist which one can single out as being the ‘best’? And how do we want to define the qualification ‘best’? The approach that will be followed here is a probabilistic one. That is, we will use the probability distribution of the integer estimator in order to decide which estimator to choose. Since the integer estimator \check{a} is by definition of the discrete type, its distribution will be a probability mass function (pmf). It will be denoted as $P(\check{a} = z)$, with $z \in Z^n$. In order to determine this distribution, we first need the pdf of \hat{a} . The pdf of \hat{a} will be denoted as $p_a(x)$, with $x \in R^n$. The subscript is used to show that the pdf still depends on the unknown parameter vector $a \in Z^n$.

The pmf of \check{a} can now be obtained as follows. Since the integer estimator is defined as

$$\check{a} = z \Leftrightarrow \hat{a} \in S_z \quad (16)$$

it follows that $P(\check{a} = z) = P(\hat{a} \in S_z)$. The pmf of \check{a} therefore follows as

$$P(\check{a} = z) = \int_{S_z} p_a(x) dx, \quad \forall z \in Z^n \quad (17)$$

The probability that \check{a} coincides with z is therefore given by the integral of the pdf $p_a(x)$ over the pull-in region $S_z \subset R^n$.

The pmf of \check{a} can now be used to study various properties of the integer estimator. In Teunissen (1998b), for instance, the first moments of the integer ambiguity estimators were studied. It was shown that \check{a} is an *unbiased* estimator if $p_a(x)$ is symmetric about a and S_z reflection-symmetric about z . All three estimators \check{a}_R , \check{a}_B and \check{a}_{LS} are therefore unbiased in the case that the pdf possesses the necessary property of symmetry. This is the case, for instance, when $p_a(x)$ is a member of the family of multivariate normal distributions. Note that the reflection-symmetric property of the pull-in region is not necessary for an integer estimator to be admissible. Hence, the unbiased estimators are admissible, but admissible estimators need not be unbiased.

Having the problem of GPS ambiguity resolution in mind, we focus our attention in this contribution to the chance of successful ambiguity resolution. That is, we consider the probability of correct integer estimation. This is given as

$$P(\check{a} = a) = \text{probability of correct integer estimation} \quad (18)$$

and it describes the reliability of ambiguity resolution in terms of its expected success rate. Since unsuccessful ambiguity resolution, when passed unnoticed, will all too often lead to unacceptable errors in the positioning results, we desire high success rates and therefore a large value for $P(\check{a} = a)$. It is therefore not only of theoretical interest, but also of practical interest, to know which integer estimator maximizes the ambiguity success rate. It will be proven, for a general family of pdfs, that of all admissible estimators it is the *integer LS estimator* which maximizes Eq. (18). The pdfs that we will consider all belong to the family of elliptically contoured distributions. They are defined as follows (Chmielewsky 1981).

Definition. The random vector $\hat{a} \in R^n$ is said to have an *elliptically contoured distribution* if its pdf is of the form

$$p_a(x) = \sqrt{\det(Q_{\hat{a}}^{-1})} G(\|x - a\|_{Q_{\hat{a}}}^2) \quad (19)$$

where $G : R \mapsto [0, \infty)$ is decreasing and $Q_{\hat{a}}$ is positive-definite.

Several important distributions belong to this family. The multivariate normal distribution can be shown to be a member of this family by choosing

$$G(x) = (2\pi)^{-\frac{n}{2}} \exp(-\frac{1}{2}x^T x), \quad x \in R$$

Another member is the multivariate *t*-distribution.

Since we can formulate the LS pull-in regions for all members of the family of elliptically contoured distributions as $S_{LS,z} = \{x \in R^n \mid p_z(x) \geq p_u(x), \forall u \in Z^n\}$, it follows that

$$p_a(x) \geq \sum_{z \in Z^n} s_z(x)p_z(x), \quad \forall x \in S_{LS,a} \quad (20)$$

with the indicator function

$$s_z(x) = \begin{cases} 1 & x \in S_z \\ 0 & \text{otherwise} \end{cases}$$

where S_z are the pull-in regions of an arbitrary admissible integer estimator. When taking the integral of Eq. (20) over $S_{LS,a}$, we obtain

$$\int_{S_{LS,a}} p_a(x)dx \geq \sum_{z \in Z^n} \int_{S_{LS,a} \cap S_z} p_z(x)dx \quad (21)$$

We now apply the change of variable $y = x + a - z$ and obtain the following replacements: $p_z(x) \rightarrow p_z(y - a + z) = p_a(y)$, $S_{LS,a} \rightarrow S_{LS,2a-z}$ and $S_z \rightarrow S_a$. Hence

$$\int_{S_{LS,a}} p_a(x)dx \geq \sum_{z \in Z^n} \int_{S_{LS,2a-z} \cap S_a} p_a(y)dy = \int_{S_a} p_a(y)dy \quad (22)$$

where the last equality is a consequence of $\cup_{z \in Z^n} S_{LS,2a-z} = R^n$. On the left side of Eq. (22) we recognize the probability of correct integer estimation of the LS estimator and on the right side the probability of correct integer estimation of any arbitrary admissible integer estimator. This concludes the proof that the integer LS estimator indeed maximizes the ambiguity success rate. This result is summarized in the following theorem.

Theorem. Let the integer LS estimator be given as

$$\check{a}_{LS} = \arg \min_{z \in Z^n} \| \hat{a} - z \|_{Q_{\hat{a}}}^2$$

and the pdf of \hat{a} as

$$p_a(x) = \sqrt{\det(Q_{\hat{a}}^{-1})} G(\|x - a\|_{Q_{\hat{a}}}^2)$$

where $G : R \mapsto [0, \infty)$ is decreasing and $Q_{\hat{a}}$ is positive-definite. Then

$$P(\check{a}_{LS} = a) \geq P(\check{a} = a) \quad (23)$$

for any admissible estimator \check{a} .

With this theorem and a result of Teunissen (1998c) we are now also in a position to order the three admissible estimators \check{a}_R , \check{a}_B and \check{a}_{LS} in terms of their success rates. From the theorem it follows that \check{a}_{LS} is better than both \check{a}_R and \check{a}_B , and in Teunissen (1998c) it was shown that \check{a}_B is again better than \check{a}_R . We thus have the ordering

$$P(\check{a}_R = a) \leq P(\check{a}_B = a) \leq P(\check{a}_{LS} = a) \quad (24)$$

A very useful application of this result is that it shows how one can *lower-bound* the probability of correct integer LS estimation. This is particularly useful when $P(\check{a}_B = a)$ is used as lower bound. This probability can be computed exactly and rather easily in case that the pdf $p_a(x)$ is normal. As was shown in Teunissen (1997), it can be computed as

$$P(\check{a}_B = a) = \prod_{i=1}^n \left(2\Phi\left(\frac{1}{2\sigma_{\hat{a}_{i|I}}}\right) - 1 \right) \quad (25)$$

with

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right) dy$$

Such an easy way of evaluating the success rate is usually not possible in case of the LS estimator.

When using Eq. (25) as the lower bound for $P(\check{a}_{LS} = a)$, there is one important point to be recognized. Since Eq. (25) depends only on the sequential conditional standard deviations of \hat{a} , it is generally not invariant for the class of admissible ambiguity transformations (Teunissen 1995). On the other hand, the probability of correct integer LS estimation is invariant for the class of admissible ambiguity transformations. The lack of invariance in $P(\check{a}_B = a)$ implies that we still have some degrees of freedom for improving this probability so as to make it a sharper lower bound of $P(\check{a}_{LS} = a)$.

The lower bound is usually particularly poor when applied to the DD ambiguities. The lower bound becomes much sharper, however, when it is applied to ambiguities which are almost decorrelated. Such ambiguities can be obtained by means of the decorrelating ambiguity transformation of the LAMBDA method (see e.g. Teunissen 1993; de Jonge and Tiberius 1996). Since the transformed ambiguities obtained by this method are far more precise than the original DD ambiguities, the lower bound becomes sharper due to its increase in value. Although other type of lower bounds for $P(\check{a}_{LS} = a)$ can be given, it is our experience that $P(\check{a}_B = a)$, when applied to the decorrelated ambiguities, is usually the best lower bound we obtain and very sharp indeed. Finally, note that the additional computations required for evaluating the lower bound are minimal. Since the LAMBDA method is already used for efficiently solving the integer LS problem, the sequential conditional standard deviations of the transformed ambiguities are available at no extra cost.

5 Summary

In this contribution a probabilistic justification for using the integer LS estimator has been given. It was shown that when the pdf of the ambiguity float solution is a member of the elliptically contoured distributions, the integer LS estimator will have the largest success rate of all admissible ambiguity estimators. The class of admissible estimators was defined by means of the following three properties of their pull-in regions:

1. $\bigcup_{z \in Z^n} S_z = R^n$
2. $S_{z_1} \cap S_{z_2} = \emptyset, \forall z_1, z_2 \in Z^n, z_1 \neq z_2$
3. $S_z = z + S_0, \forall z \in Z^n$

The first condition states that the pull-in regions should not leave any gaps, the second that they should not overlap and the third that the integer estimators should admit the ‘integer remove–restore’ technique. Various admissible estimators exist, three of which are as follows:

1. $\check{a}_R = ([\hat{a}_1], \dots, [\hat{a}_n])^T$
2. $\check{a}_B = ([\hat{a}_1], \dots, [\hat{a}_{n|N}])^T$
3. $\check{a}_{LS} = \arg \min_{z \in Z^n} \| \hat{a} - z \|_{Q_{\hat{a}}}^2$

Their pull-in regions are shaped as the n -dimensional versions of a square, a parallelogram and a convex polygon respectively. The theorem given in this contribution shows that

$$P(\check{a}_{LS} = a) \geq P(\check{a} = a) \text{ for all admissible } \check{a}$$

In maximizing the success of GPS ambiguity resolution we are thus better off when using the integer LS estimator than any other admissible ambiguity estimator. As a direct consequence of the theorem we have

$$P(\check{a}_{LS} = a) \geq P(\check{a}_B = a) = \prod_{i=1}^n \left(2\Phi\left(\frac{1}{2\sigma_{\hat{a}_{i|I}}}\right) - 1 \right)$$

where the last equality follows from Teunissen (1997). This shows how the easily computed success rate of the bootstrapped estimator can be used as the lower bound.

References

- Baarda W (1968) A testing procedure for use in geodetic networks, Publ geodesy, New series, vol 2(5). Netherlands Geodetic Commission, Delft
- Blewitt G (1989) Carrier-phase ambiguity resolution for the Global Positioning System applied to geodetic baselines up to 2000 km. J Geophys Res 94 (B8) 10,187–10,302
- Chmielewsky MA (1981) Elliptically symmetric distributions: a review and bibliography. Int Statist Rev 49: 67–74
- de Jonge PJ, Tiberius CCJM (1996) The LAMBDA method for integer ambiguity estimation: implementation aspects. Delft Geodetic Computing Centre LGR Series, No 12, Delft University of Technology
- Dong D, Bock Y (1989) Global positioning system network analysis with phase ambiguity resolution applied to crustal deformation studies in California. J Geophys Res 94: 3949–3966
- Hofmann-Wellenhof B, Lichtenegger H, Collins J (1997) Global positioning system: theory and practice, 4th edn. Springer, Berlin Heidelberg New York
- Jonkman NF (1998) The geometry-free approach to integer GPS ambiguity estimation. Proc ION GPS-98. This paper presents a part of the author’s MSc thesis (1996) Integer GPS-ambiguity estimation without the receiver-satellite geometry, Delft University of Technology
- Leick A (1995) GPS Satellite Surveying, 2nd edn. John Wiley, New York
- Parkinson B, Spilker JJ (eds) (1996) GPS: Theory and applications, vols 1 and 2. AIAA, Washington, DC
- Strang G, Borre K (1997) Linear algebra, geodesy, and GPS. Wellesley-Cambridge Press
- Teunissen PJG (1985) Quality control in geodetic networks. In: Graffarend E, Samso F (eds) Optimization and design of geodetic networks. Springer, Berlin Heidelberg New York
- Teunissen PJG (1993) Least-squares estimation of the integer GPS ambiguities. Invited lecture, Sect IV Theory and methodology, IAG General Meeting, Beijing, August. Also in LGR Series, No 6, Delft Geodetic Computing Centre
- Teunissen PJG (1995) The invertible GPS ambiguity transformations. Manuscripta Geod 20: 489–497

- Teunissen PJG (1997) Some remarks on GPS ambiguity resolution. *Artific Sat* 32(3): 119–130
- Teunissen PJG (1998a) On the integer normal distribution of the GPS ambiguities. *Artific Sat* 33(2): 49–64
- Teunissen PJG (1998b) A class of unbiased integer GPS ambiguity estimators. *Artific Sat* 33(1): 3–10
- Teunissen PJG (1998c) Success probability of integer GPS ambiguity rounding and bootstrapping. *J Geod* 72: 606–612
- Teunissen PJG, Kleusberg A (eds) (1998) *GPS for Geodesy*, 2nd edn. Springer, Berlin Heidelberg New York

Theory of Carrier Phase Ambiguity Resolution

Teunissen P J G

Department of Mathematical Geodesy and Positioning
Delft University of Technology, Thiusseweg 11, 2629 JA Delft, The Netherlands

Abstract. Carrier phase ambiguity resolution is the key to high precision Global Navigation Satellite System (GNSS) positioning and navigation. It applies to a great variety of current and future models of GPS, modernized GPS and Galileo. A proper handling of carrier phase ambiguity resolution requires a proper understanding of the underlying theory of integer inference. In this contribution a brief review is given of the probabilistic theory of integer ambiguity estimation. We describe the concept of ambiguity pull-in regions, introduce the class of admissible integer estimators, determine their probability mass functions and show how their variability affect the uncertainty in the so-called 'fixed' baseline solution. The theory is worked out in more detail for integer least-squares and integer bootstrapping. It is shown that the integer least-squares principle maximizes the probability of correct integer estimation. Sharp and easy-to-compute bounds are given for both the ambiguity success rate and the baseline's probability of concentration. Finally the probability density function of the ambiguity residuals is determined. This allows one for the first time to formulate rigorous tests for the integerness of the parameters.

Keywords: GNSS, ambiguity resolution, integer least-squares

0 Introduction

Global Navigation Satellite System (GNSS) ambiguity resolution is the process of resolving the unknown cycle ambiguities of double difference (DD) carrier phase data as integers. It is the key to high precision GNSS positioning and navigation. The availability of a theory of integer inference is therefore a prerequisite for a proper handling and understanding of the various intricate aspects of carrier phase ambiguity resolution. Although the theory of integer inference as a whole is unfortunately still far from mature, significant progress has been made in the last decade in the area of integer ambiguity estimation. This holds true for the computational aspects of integer estimation as well as for the corresponding probabilistic aspects. We now have a theoretical framework available which enables one to define integer estimators unequivocally, to compare their performance, to single-out optimal estimators, to judge the probabilistic consequences for the so-called 'fixed' baseline solution and to design measurement set-ups in accordance to specifications. In addition the framework has also enabled one to identify pitfalls in some of the earlier proposed methods of ambiguity resolution. It is the purpose of the current invited contribution to give a brief review of this probabilistic framework of integer estimation. The focus will be on integer least-squares estimation and some closely related integer estimation principles. The presentation will be non-Bayesian throughout. For a Bayesian approach to ambiguity resolution we refer to e.g. [1], [12], [13]. The presentation will also not touch upon the theory of integer validation, a theory which unfortunately is still in its infancy. But some first results, as reported in [33] and [37], will be given.

Biography: Teunissen P J G, male, Professor, Ph.D., research direction: GNSS data processing strategies for medium scaled networks with an emphasis on ambiguity resolution.

E-mail: P.J.G.Teunissen@citg.tudelft.nl

The practical importance of carrier phase ambiguity resolution becomes clear when one realizes the great variety of current and future GNSS models to which it applies. These models may differ greatly in complexity and diversity. They range from single-baseline models used for kinematic positioning to multi-baseline models used as a tool for studying geodynamic phenomena. The models may or may not have the relative receiver-satellite geometry included. They may also be discriminated as to whether the slave receiver(s) are stationary or in motion. When in motion, one solves for one or more trajectories, since with the receiver-satellite geometry included, one will have new coordinate unknowns for each epoch. One may also discriminate between the models as to whether or not the differential atmospheric delays (ionosphere and troposphere) are included as unknowns. In the case of sufficiently short baselines they are usually excluded.

Apart from the current Global Positioning System (GPS) models, carrier phase ambiguity resolution also applies to the future modernized GPS and the future European Galileo GNSS. An overview of GNSS models, together with their applications in surveying, navigation, geodesy and geophysics, can be found in textbooks such as [16], [19], [20], [21], [24] and [25].

In the present contribution we emphasize the probabilistic aspects of integer ambiguity estimation. This contribution is organized as follows. In section 1 we introduce a general class of integer ambiguity estimators, determine their probability mass functions and show how their variability affect the uncertainty in the computed GNSS baselines. This theory is worked out in sections 2 and 3 for two of the most important integer ambiguity estimators. We refrain from giving proofs of the theorems and corollaries. For these proofs we refer to the referenced literature. In section 2 we discuss the properties of integer bootstrapping and in section 3 those of integer least-squares. The properties of these two estimators are com-

pared. It is shown that integer least-squares maximizes the probability of correct integer estimation. We also give sharp and easy-to-compute bounds for the ambiguity success rate. In the final section 4 we present the joint pdf of the 'float' and 'fixed' ambiguities and from it determine the pdf of the ambiguity residuals.

1 Integer Ambiguity Resolution

1.1 The GNSS model

As our point of departure we will take the following system of linear(ized) observation equations

$$y = Aa + Bb + e \quad (1)$$

where y is the given GNSS data vector of order m , a and b are the unknown parameter vectors respectively of order n and p , and where e is the noise vector. In principle all the GNSS models can be cast in this frame of observation equations. The data vector y will usually consist of the 'observed minus computed' single- or dual-frequency double-difference (DD) phase and/or pseudorange (code) observations accumulated over all observation epochs. The entries of vector a are then the DD carrier phase ambiguities, expressed in units of cycles rather than range. They are known to be *integers*, $a \in \mathbb{Z}^n$. The entries of the vector b will consist of the remaining unknown parameters, such as for instance baseline components (coordinates) and possibly atmospheric delay parameters (troposphere, ionosphere). They are known to be real-valued, $b \in \mathbb{R}^p$.

The procedure which is usually followed for solving the GNSS model (1), can be divided into three steps. In the *first* step one simply disregards the integer constraints $a \in \mathbb{Z}^n$ on the ambiguities and performs a standard least-squares adjustment. As a result one obtains the (real-valued) estimates of a and b , together with their variance-covariance (vc-) matrix

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}, \begin{bmatrix} Q_{\hat{a}\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}\hat{b}} \end{bmatrix} \quad (2)$$

This solution is referred to as the 'float' solution. In the *second* step the 'float' ambiguity estimate \hat{a} is used to compute the corresponding integer ambiguity estimate \check{a} . This implies that a mapping $S : \mathbb{R}^n \mapsto \mathbb{Z}^n$, from the n -dimensional space of reals to the n -dimensional space of integers, is introduced such that

$$\check{a} = S(\hat{a}) \quad (3)$$

Once the integer ambiguities are computed, they are used in the *third* step to finally correct the 'float' estimate of b . As a result one obtains the 'fixed' solution

$$\check{b} = \hat{b} - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{-1}(\hat{a} - \check{a}) \quad (4)$$

In the present review we will primarily focus our attention on the probabilistic properties of (3) and (4).

1.2 Admissible integer estimation

There are many ways of computing an integer ambiguity vector \check{a} from its real-valued counterpart \hat{a} . To each such method belongs a mapping $S : \mathbb{R}^n \mapsto \mathbb{Z}^n$ from the n -dimensional

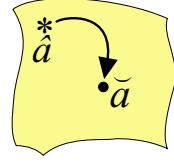


Fig. 1 An ambiguity pull-in region of $z = \check{a}$.

space of real numbers to the n -dimensional space of integers. Due to the discrete nature of \mathbb{Z}^n , the map S will not be one-to-one, but instead a many-to-one map. This implies that different real-valued ambiguity vectors will be mapped to the same integer vector. One can therefore assign a subset $S_z \subset \mathbb{R}^n$ to each integer vector $z \in \mathbb{Z}^n$:

$$S_z = \{x \in \mathbb{R}^n \mid z = S(x)\}, \quad z \in \mathbb{Z}^n \quad (5)$$

The subset S_z contains all real-valued ambiguity vectors that will be mapped by S to the same integer vector $z \in \mathbb{Z}^n$. This subset is referred to as the *pull-in region* of z (see figure 1). It is the region in which all ambiguity 'float' solutions are pulled to the same 'fixed' ambiguity vector z .

Using the pull-in regions, one can give an explicit expression for the corresponding integer ambiguity estimator. It reads

$$\check{a} = \sum_{z \in \mathbb{Z}^n} z s_z(\hat{a}) \quad (6)$$

with the indicator function:

$$s_z(\hat{a}) = \begin{cases} 1 & \text{if } \hat{a} \in S_z \\ 0 & \text{otherwise.} \end{cases}$$

Since the pull-in regions define the integer estimator completely, one can define classes of integer estimators by imposing various conditions on the pull-in regions. One such class is referred to as the class of admissible integer estimators. This class was introduced in [30] and it is defined as follows.

Definition 1

The integer estimator $\check{a} = \sum_{z \in \mathbb{Z}^n} z s_z(\hat{a})$ is said to be *admissible* if

- (i) $\bigcup_{z \in \mathbb{Z}^n} S_z = \mathbb{R}^n$
- (ii) $\text{Int}(S_{z_1}) \cap \text{Int}(S_{z_2}) = \emptyset, \quad \forall z_1, z_2 \in \mathbb{Z}^n, z_1 \neq z_2$
- (iii) $S_z = z + S_0, \quad \forall z \in \mathbb{Z}^n$

This definition is motivated as follows (see figure 2). Each one of the above three conditions describes a property of which it seems reasonable that it is possessed by an arbitrary integer ambiguity 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 $\hat{a} \in \mathbb{R}^n$ to \mathbb{Z}^n , while the absence of overlaps is needed to guarantee that the 'float' solution is mapped to just one integer vector. Note that we allow the pull-in regions to have common boundaries. This is permitted if we assume to have zero probability that \hat{a} lies on one of the boundaries. This will be the case when the probability density function (pdf) of \hat{a} is continuous.

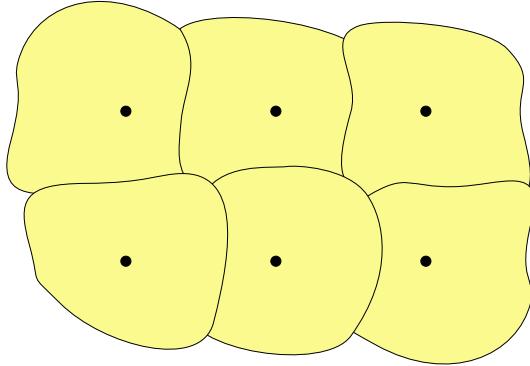


Fig. 2 Pull-in regions that cover R^n without gaps and overlaps.

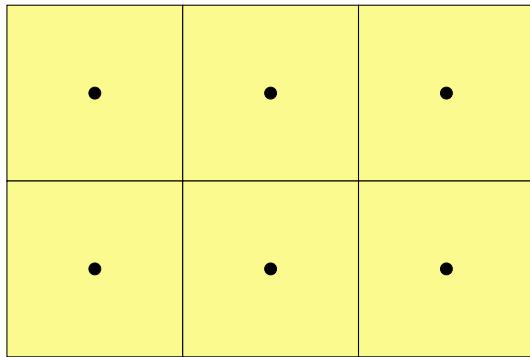


Fig. 3 An example of integer translational invariant pull-in regions that cover R^n without gaps and overlaps.

The third and last condition of the definition follows from the requirement that $S(x+z) = S(x) + z, \forall x \in R^n, z \in Z^n$ (see figure 3). Also this condition is a reasonable one to ask for. It states that when the 'float' solution is perturbed by $z \in Z^n$, the corresponding integer solution is perturbed by the same amount. This property allows one to apply the *integer remove-restore* technique: $S(\hat{a} - z) + z = S(\hat{a})$. It therefore allows one to work with the fractional parts of the entries of \hat{a} , instead of with its complete entries.

With the division of R^n into mutually exclusive pull-in regions, we are now in the position to consider the distribution of \check{a} . This distribution is of the *discrete* type and it will be denoted as $P(\check{a} = z)$. It is a probability mass function, having zero masses at non-grid points and nonzero masses at some or all grid points. If we denote the *continuous* probability density function of \hat{a} as $p_{\hat{a}}(x)$, the distribution of \check{a} follows as

$$P(\check{a} = z) = \int_{S_z} p_{\hat{a}}(x) dx, \quad z \in Z^n \quad (7)$$

This expression holds for any distribution the 'float' ambiguities \hat{a} might have. In most GNSS applications however, one assumes the vector of observables y to be normally distributed. The estimator \hat{a} is therefore normally distributed too, with mean $a \in Z^n$ and vc-matrix $Q_{\hat{a}}$. Its probability density function reads

$$p_{\hat{a}}(x) = \frac{1}{\sqrt{\det(Q_{\hat{a}})}(2\pi)^{\frac{1}{2}n}} \exp\left\{-\frac{1}{2} \|x - a\|_{Q_{\hat{a}}}^2\right\} \quad (8)$$

with the squared weighted norm $\|\cdot\|_{Q_{\hat{a}}}^2 = (\cdot)^T Q_{\hat{a}}^{-1}(\cdot)$. Note that $P(\check{a} = a)$ equals the probability of *correct* integer am-

biguity estimation. It describes the expected success rate of GNSS ambiguity resolution.

1.3 The baseline solution

We are now in the position to determine the pdf of the 'fixed' baseline estimator (4). In order to determine this pdf, one needs to propagate the uncertainty of the 'float' solution, \hat{a} and \hat{b} , as well as the uncertainty of the integer solution \check{a} through (4). Should one neglect the random character of the integer solution and therefore consider the ambiguity vector \check{a} as deterministic and equal to, say, z , then the pdf of \check{b} would equal the conditional baseline distribution

$$p_{\check{b}|\hat{a}}(x | z) = \frac{1}{\sqrt{\det Q_{\check{b}|\hat{a}}}(2\pi)^{\frac{1}{2}p}} \exp\left\{-\frac{1}{2} \|x - b(z)\|_{Q_{\check{b}|\hat{a}}}^2\right\} \quad (9)$$

with conditional mean $b(z) = b - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1}(a - z)$, conditional variance matrix $Q_{\check{b}|\hat{a}} = Q_{\check{b}} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} Q_{\hat{a}\hat{b}}$ and $\|\cdot\|_{Q_{\check{b}|\hat{a}}}^2 = (\cdot)^T Q_{\check{b}|\hat{a}}^{-1}(\cdot)$. However, since \check{a} is random and not deterministic, the conditional baseline distribution will give a too optimistic description of the quality of the 'fixed' baseline. To get a correct description of the 'fixed' baseline's pdf, the integer ambiguity's pmf needs to be considered. As the following theorem shows this results in a baseline distribution, which generally will be multi-modal.

Theorem 1 (Pdf of the 'fixed' baseline)

Let the 'float' solution, \hat{a} and \hat{b} , be normally distributed with mean $a \in Z^n$ and mean $b \in R^p$, and vc-matrix (2), let \check{a} be an admissible integer estimator and let the 'fixed' baseline \check{b} be given as in (4). The pdf of \check{b} reads then

$$p_{\check{b}}(x) = \sum_{z \in Z^n} p_{\check{b}|\hat{a}}(x | z) P(\check{a} = z) \quad (10)$$

Note that, although the model (1) is linear and the observables normally distributed, the distribution of the 'fixed' baseline is not normal, but multi-modal (see figure 4). As the theorem shows, the 'fixed' baseline distribution equals an infinite sum of weighted conditional baseline distributions. These conditional baseline distributions $p_{\check{b}|\hat{a}}(x | z)$ are shifted versions of one another. The size and direction of the shift is governed by $Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} z, z \in Z^n$. Each of the conditional baseline distributions in the infinite sum is down-weighted. These weights are given by the probability masses of the distribution of the integer ambiguity estimator \check{a} . This shows that the dependence of the 'fixed' baseline distribution on the choice of integer estimator is only felt through the weights $P(\check{a} = z)$.

1.4 On the quality of the 'fixed' baseline

In order to describe the quality of the 'fixed' baseline, one would like to know how close one can expect the baseline estimate \check{b} to be to the unknown, but true baseline value b . As a measure of confidence, we take

$$P(\check{b} \in R) = \int_R p_{\check{b}}(x) dx \quad \text{with } R \subset R^p \quad (11)$$

But in order to evaluate this integral, we first need to make a choice about the shape and location of the subset R . Since it is

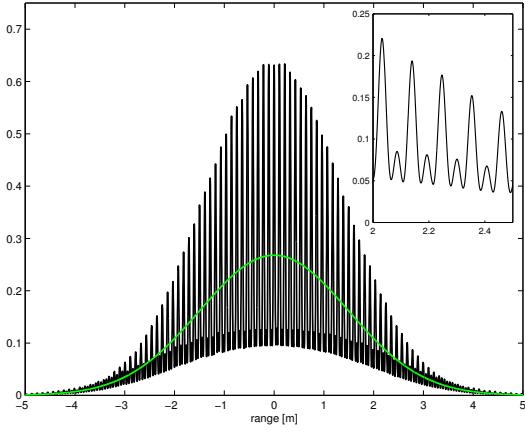


Fig. 4 An example of the multi-modal pdf of \check{b} and corresponding uni-modal pdf of \hat{b} .

common practice in GNSS positioning to use the vc-matrix of the conditional baseline estimator as a measure of precision for the 'fixed' baseline, the vc-matrix $Q_{\hat{b}|\hat{a}}$ will be used to define the shape of the confidence region. For its location, we choose the confidence region to be centered at b . After all, we would like to know by how much the baseline estimate \check{b} can be expected to differ from the true, but unknown baseline value b . That is, one would like (11) to be a measure of the baseline's probability of concentration about b .

With these choices on shape and location, the region R takes the form

$$R = \{x \in R^p \mid (x - b)^T Q_{\hat{b}|\hat{a}}^{-1}(x - b) \leq \beta^2\} \quad (12)$$

The size of the region can be varied by varying β . The following theorem shows how the baseline's probability of concentration (11) can be evaluated as a weighted sum of probabilities of noncentral Chi-square distributions.

Theorem 2 (*The 'fixed' baseline's probability of concentration*)

Let \check{b} be the 'fixed' baseline estimator, let R be defined as in (12), and let $\chi^2(p, \lambda_z)$ denote the noncentral Chi-square distribution with p degrees of freedom and non-centrality parameter λ_z . Then

$$P(\check{b} \in R) = \sum_{z \in Z^n} P(\chi^2(p, \lambda_z) \leq \beta^2) P(\check{a} = z) \quad (13)$$

with

$$\lambda_z = \| \nabla \check{b}_z \|_{Q_{\hat{b}|\hat{a}}}^2 \quad \text{and} \quad \nabla \check{b}_z = Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1}(z - a)$$

This result shows that the probability of the 'fixed' baseline lying inside the ellipsoidal region R centered at b equals an infinite sum of probability products. If one considers the two probabilities of these products separately, two effects are observed. First the probabilistic effect of shifting the conditional baseline estimator away from b and secondly the probabilistic effect of the peakedness or non-peakedness of the ambiguity pmf. The second effect is related to the expected performance of ambiguity resolution, while the first effect has to do with the sensitivity of the baseline for changes in the values of the integer ambiguities. This effect is measured by

the non-centrality parameter λ_z . Since the tail of a noncentral Chi-square distribution becomes heavier when the non-centrality parameter increases, while the degrees of freedom remain fixed, $P(\chi^2(p, \lambda_z) \leq \beta^2)$ gets smaller when λ_z gets larger.

The two probabilities in the product reach their maximum values when $z = a$. The following corollary shows how these two maxima can be used to lower bound and to upper bound the probability $P(\check{b} \in R)$. Such bounds are of importance for practical purposes, since it is difficult in general to evaluate (13) exactly.

Corollary 1 (*Lower and upper bounds*)

Let \check{b} be the 'fixed' baseline estimator and let R be defined as in (12). Then

$$P(\hat{b}_{|\hat{a}=a} \in R) P(\check{a} = a) \leq P(\check{b} \in R) \leq P(\hat{b}_{|\hat{a}=a} \in R) \quad (14)$$

with

$$P(\hat{b}_{|\hat{a}=a} \in R) = P(\chi^2(p, 0) \leq \beta^2)$$

Note that the two bounds relate the probability of the 'fixed' baseline estimator to that of the conditional estimator and the ambiguity success rate. The above bounds become tight when the ambiguity success rate approaches one. This shows, although the probability of the conditional estimator always overestimates the probability of the 'fixed' baseline estimator, that the two probabilities are close for large values of the success rate. This implies that in case of GNSS ambiguity resolution, one should first evaluate the success rate $P(\check{a} = a)$ and make sure that its value is close enough to one, before making any inferences on the basis of the distribution of the conditional baseline estimator. In other words, the (uni-modal) distribution of the conditional estimator is a good approximation to the (multi-modal) distribution of the bootstrapped baseline estimator, when the success rate is sufficiently close to one.

2 Integer Bootstrapping

2.1 The bootstrapped estimator

The distributional results presented so far hold for any admissible ambiguity estimator. The simplest way to obtain an integer vector from the real-valued 'float' solution is to round each of the entries of \hat{a} to its nearest integer. The corresponding integer estimator reads therefore

$$\check{a}_R = ([\hat{a}_1], \dots, [\hat{a}_n])^T \quad (15)$$

where '[.]' denotes rounding to the nearest integer. The pull-in region of this integer estimator equals the multivariate version of the unit square (see figure 3).

Another relatively simple integer ambiguity estimator is the bootstrapped estimator. The bootstrapped estimator can be seen as a generalization of the previous estimator, [29] and [32]. It still makes use of integer rounding, but it also takes some of the correlation between the ambiguities into account. The bootstrapped estimator follows from a sequential conditional least-squares adjustment and it is computed as follows. If n ambiguities are available, one starts with the first ambiguity \hat{a}_1 , and rounds its value to the nearest integer. Having obtained the integer value of this

first ambiguity, the real-valued estimates of all remaining ambiguities are then corrected by virtue of their correlation with the first ambiguity. Then the second, but now corrected, real-valued ambiguity estimate is rounded to its nearest integer. Having obtained the integer value of the second ambiguity, the real-valued estimates of all remaining $n - 2$ ambiguities are then again corrected, but now by virtue of their correlation with the second ambiguity. This process is continued until all ambiguities are considered. We thus have the following definition.

Definition 2 (Integer bootstrapping)

Let $\hat{a} = (\hat{a}_1, \dots, \hat{a}_n)^T \in R^n$ be the ambiguity 'float' solution and let $\check{a}_B = (\check{a}_{B,1}, \dots, \check{a}_{B,n})^T \in Z^n$ denote the corresponding integer bootstrapped solution. The entries of the bootstrapped ambiguity estimator are then defined as

$$\begin{aligned}\check{a}_{B,1} &= [\hat{a}_1] \\ \check{a}_{B,2} &= [\hat{a}_{2|1}] = [\hat{a}_2 - \sigma_{21}\sigma_1^{-2}(\hat{a}_1 - \check{a}_{B,1})] \\ &\vdots \\ \check{a}_{B,n} &= [\hat{a}_{n|N}] = [\hat{a}_n - \sum_{j=1}^{n-1} \sigma_{n,j|J}\sigma_{j|J}^{-2}(\hat{a}_{j|J} - \check{a}_{B,j})]\end{aligned}\tag{16}$$

where ' $[.]$ ' denotes the operation of rounding to the nearest integer, and $\sigma_{i,j|J}$ denotes the covariance between \hat{a}_i and $\hat{a}_{j|J}$, and $\sigma_{j|J}^2$ is the variance of $\hat{a}_{j|J}$. The shorthand notation $\hat{a}_{i|I}$ stands for the i th least-squares ambiguity obtained through a conditioning on the previous $I = \{1, \dots, (i-1)\}$ sequentially rounded ambiguities.

Note that the bootstrapped estimator is not unique. Changing the order in which the ambiguities appear in vector \hat{a} will already produce a different bootstrapped estimator. Although the principle of bootstrapping remains the same, every choice of ambiguity parameterization has its own bootstrapped estimator.

2.2 The bootstrapped pull-in regions

The pull-in regions for rounding are unit cubes centered at integer grid points. For bootstrapping the shape of the pull-in regions will depend on the vc-matrix of the ambiguities. They will coincide with the unit cubes only in case the vc-matrix is a diagonal matrix. Bootstrapping reduces namely to rounding in the absence of any correlation between the ambiguities. The following theorem gives a description of the bootstrapped pull-in regions in the general case.

Theorem 3 (Bootstrapped pull-in regions)

The pull-in regions of the bootstrapped ambiguity estimator $\check{a}_B = (\check{a}_{B,1}, \dots, \check{a}_{B,n})^T \in Z^n$ are given as

$$\begin{aligned}S_{B,z} &= \{x \in R^n \mid |c_i^T L^{-1}(x - z)| \leq \frac{1}{2}, \\ &\quad i = 1, \dots, n\}, \quad \forall z \in Z^n\end{aligned}\tag{17}$$

where L denotes the unique unit lower triangular matrix of the ambiguity vc-matrix' decomposition $Q_{\hat{a}} = LDL^T$ and c_i denotes the i th canonical unit vector having a 1 as its i th entry and zeros otherwise.

That the bootstrapped estimator is indeed admissible, can now be seen as follows. The first two conditions of Definition 1 are easily verified using the definition of the bootstrapped estimator. Since every real-valued vector \hat{a} will be mapped by the bootstrapped estimator to an integer vector, the pull-in regions $S_{B,z}$ cover R^n without any gaps. There is also no overlap between the pull-in regions, since - apart from boundary ties - any real-valued vector \hat{a} is mapped to not more than one integer vector. To verify the last condition of Definition 1, we make use of (17). From

$$\begin{aligned}S_{B,z} &= \{x \in R^n \mid |c_i^T L^{-1}(x - z)| \leq \frac{1}{2}, \\ &\quad i = 1, \dots, n\} \\ &= \{x \in R^n \mid |c_i^T L^{-1}y| \leq \frac{1}{2}, \\ &\quad x = y + z, \quad i = 1, \dots, n\} \\ &= S_{B,0} + z\end{aligned}$$

it follows that all bootstrapped pull-in regions are translated copies of $S_{B,0}$. All pull-in regions have therefore the same shape and the same volume. Their volumes all equal 1. This can be shown by transforming $S_{B,0}$ to the unit cube centered at the origin. Consider the linear transformation $y = L^{-1}x$. Then

$$L^{-1}(S_{B,0}) = \{y \in R^n \mid |c_i^T y| \leq \frac{1}{2}, \quad i = 1, \dots, n\}$$

equals the unit cube centered at the origin. Since the determinant of the unit lower triangular matrix L^{-1} equals one and since the volume of the unit cube equals one, it follows that the volume of $S_{B,0}$ must equal one as well. To infer the shape of the bootstrapped pull-in region, we consider the two-dimensional case first. Let the lower triangular matrix L be given as

$$L = \begin{bmatrix} 1 & 0 \\ l & 1 \end{bmatrix}$$

Then

$$\begin{aligned}S_{B,0} &= \{x \in R^2 \mid |c_i^T L^{-1}x| \leq \frac{1}{2}, \quad i = 1, 2\} \\ &= \{x \in R^2 \mid |x_1| \leq \frac{1}{2}, \quad |x_2 - lx_1| \leq \frac{1}{2}\}\end{aligned}$$

which shows that the two-dimensional pull-in region equals a parallelogram (see figure 5). Its region is bounded by the two vertical lines $x_1 = 1/2$ and $x_1 = -1/2$, and the two parallel slopes $x_2 = lx_1 + 1/2$ and $x_2 = lx_1 - 1/2$. The direction of the slope is governed by $l = \sigma_{21}\sigma_1^{-2}$. Hence, in the absence of correlation between the two ambiguities, the parallelogram reduces to the unit square. In higher dimensions the above construction of the pull-in region can be continued. In three dimensions for instance, the intersection of the pull-in region with the x_1x_2 -plane remains a parallelogram, while along the third axis the pull-in region becomes bounded by two parallel planes.

2.3 The bootstrapped pmf

Since the integer bootstrapped estimator is defined as $\check{a}_B = z \iff \hat{a} \in S_{B,z}$, it follows that $P(\check{a}_B = z) = P(\hat{a} \in S_{B,z})$.

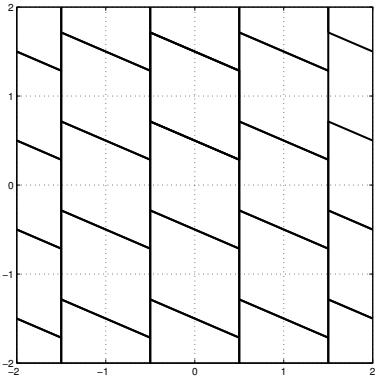


Fig. 5 The two-dimensional pull-in regions of integer bootstrapping.

The pmf of \hat{a}_B follows therefore as

$$P(\hat{a}_B = z) = \int_{S_{B,z}} p_{\hat{a}}(x) dx, \quad z \in Z^n \quad (18)$$

Hence, the probability that \hat{a}_B coincides with z is given by the integral of the pdf $p_{\hat{a}}(x)$ over the bootstrapped pull-in region $S_{B,z} \subset R^n$. The above expression holds for any distribution the 'float' ambiguities \hat{a} might have. In most GNSS applications however, one usually assumes the vector of observables y to be normally distributed. For that case the following theorem gives an exact expression of the bootstrapped pmf (see figure 6).

Theorem 4 (The integer bootstrapped pmf)

Let \hat{a} be distributed as $N(a, Q_{\hat{a}})$, $a \in Z^n$, and let \hat{a}_B be the corresponding integer bootstrapped estimator. Then

$$\begin{aligned} P(\hat{a}_B = z) &= \prod_{i=1}^n [\Phi\left(\frac{1 - 2l_i^T(a - z)}{2\sigma_{\hat{a}_{i|I}}}\right) \\ &\quad + \Phi\left(\frac{1 + 2l_i^T(a - z)}{2\sigma_{\hat{a}_{i|I}}}\right) - 1], \quad z \in Z^n \end{aligned} \quad (19)$$

with

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}v^2\right\} dv$$

and with l_i the i th column vector of the unit lower triangular matrix L^{-T} and $\sigma_{\hat{a}_{i|I}}^2$ the variance of the i th least-squares ambiguity obtained through a conditioning on the previous $I = \{1, \dots, (i-1)\}$ ambiguities.

The bootstrapped pmf equals a product of univariate pmf's and is therefore easy to compute. Note that the bootstrapped pmf is completely governed by the ambiguity vc-matrix $Q_{\hat{a}}$. The pmf follows once the triangular factor L and the diagonal matrix D of the decomposition $Q_{\hat{a}} = LDL^T$ are given. The above result also shows that the bootstrapped pmf is symmetric about the mean of \hat{a} . This implies that the bootstrapped estimator \hat{a}_B is an unbiased estimator of $a \in Z^n$. Since the 'float' solutions, \hat{a} and \hat{b} , are unbiased too, it follows from taking the expectation of (4) that the bootstrapped baseline is also unbiased.

For the purpose of predicting the success of ambiguity resolution, the probability of correct integer estimation is

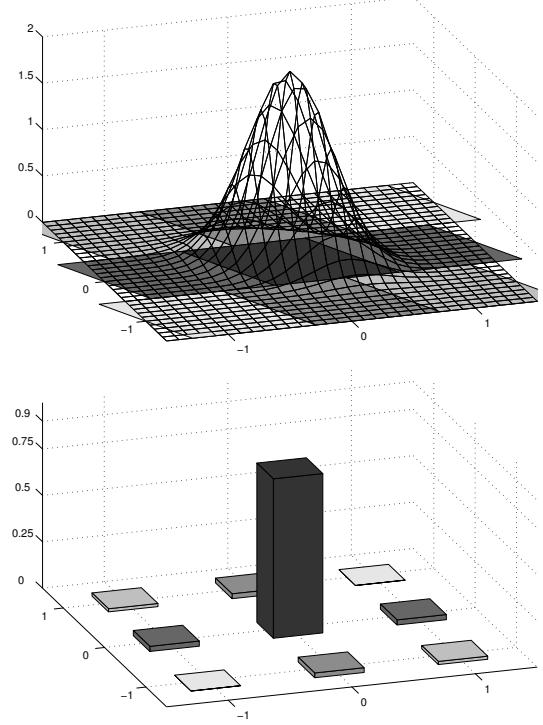


Fig. 6 (top) The two-dimensional pdf of the 'float' solution \hat{a} situated over the bootstrapped pull-in regions; (bottom) the two-dimensional pmf of the integer bootstrapped solution \hat{a}_B .

of particular interest. For the bootstrapped estimator this success rate is given in the following corollary.

Corollary 2 (The bootstrapped success rate)

The bootstrapped probability of correct integer estimation (the success rate) is given as

$$P(\hat{a}_B = a) = \prod_{i=1}^n [2\Phi\left(\frac{1}{2\sigma_{\hat{a}_{i|I}}}\right) - 1] \quad (20)$$

The method of integer bootstrapping is easy to implement and it does not need, as opposed to the method of integer least-squares (see next section), an integer search for computing the sought for integer solution. However, as it was mentioned earlier, the outcome of bootstrapping depends on the chosen ambiguity parameterization. Bootstrapping of DD ambiguities, for instance, will produce an integer solution which generally differs from the integer solution obtained from bootstrapping of reparameterized ambiguities. Since this dependency also holds true for the bootstrapped pmf, one still has some important degrees of freedom left for improving (20).

In order to improve the bootstrapped success rate, one should work with decorrelated ambiguities instead of with the original ambiguities. The method of bootstrapping performs relatively poor, for instance, when applied to the DD ambiguities. This is due to the usually high correlation between the DD ambiguities. Bootstrapping should therefore be used in combination with the decorrelating Z -transformation of the LAMBDA method, see [26] and [27]. This transformation decorrelates the ambiguities further than the best reordering would achieve and thereby reduces the values of the sequential conditional variances. By reducing the values of the se-

quential conditional variances, the bootstrapped success rate gets enlarged.

It may however happen that it is simply not possible to resolve the complete vector of ambiguities with sufficient probability. As an alternative of resolving the complete vector of ambiguities, one might then consider resolving only a subset of the ambiguities. The idea of partial ambiguity resolution is based on the fact that the success rate will generally increase when fewer integer constraints are imposed. However, in order to apply partial ambiguity resolution, one first will have to determine which subset of ambiguities to choose. It will be clear that this decision should be based on the precision of the 'float' ambiguities. The more precise the ambiguities, the larger the ambiguity success rate. It is at this point where the decorrelation step of the LAMBDA method and the bootstrapping principle can be applied. Once the transformed and decorrelated ambiguity vc-matrix is obtained, the construction of the subset proceeds in a sequential fashion. One first starts with the most precise ambiguity, say \hat{z}_1 , and computes its success rate $P(\check{a}_1 = z_1)$. If this success rate is large enough, one continues and determines the most precise pair of ambiguities, say (\hat{z}_1, \hat{z}_2) . If their success rate is still large enough, one continues again by trying to extend the set. This procedure continues until one reaches a point where the corresponding success rate becomes unacceptably small. When this point is reached, one can expect that the previously identified ambiguities can be resolved successfully.

Once the subset for partial ambiguity resolution has been identified, one still needs to determine what this will do to improve the baseline estimator. After all, being able to successfully resolve the ambiguities does not necessarily mean that the 'fixed' solution is significantly better than the 'float' solution. The theory presented in the previous sections provides the necessary tools for performing such an evaluation.

3 Integer Least-squares

3.1 The ILS estimator

When using the least-squares principle, the GNSS model can be solved by means of the minimization problem

$$\min_{a,b} \|y - Aa - Bb\|_{Q_y}^2, a \in Z^n, b \in R^p \quad (21)$$

with Q_y the vc-matrix of the GNSS observables. This type of least-squares problem was first introduced in [26] and has been coined with the term '*integer least-squares*'. It is a non-standard least-squares problem due to the integer constraints $a \in Z^n$, see [27] and [15]. The solution of (21) is consistent with the three solution steps of section 1. This can be seen as follows. It follows from the orthogonal decomposition

$$\begin{aligned} \|y - Aa - Bb\|_{Q_y}^2 &= \|\hat{e}\|_{Q_y}^2 + \|\hat{a} - a\|_{Q_{\hat{a}}}^2 \\ &\quad + \|\hat{b}(a) - b\|_{Q_{\hat{b}|a}}^2 \end{aligned} \quad (22)$$

with $\hat{e} = y - A\hat{a} - B\hat{b}$ and $\hat{b}(a) = \hat{b} - Q_{\hat{b}\hat{a}}Q_{\hat{a}}^{-1}(\hat{a} - a)$, that the sought for minimum is obtained when the second term on the right-hand side is minimized for $a \in Z^n$ and the last term is set to zero. The integer least-squares (ILS) estimator of the ambiguities is therefore defined as follows.

Definition 3 (Integer least-squares)

Let $\hat{a} = (\hat{a}_1, \dots, \hat{a}_n)^T \in R^n$ be the ambiguity 'float' solution and let $\check{a}_{LS} \in Z^n$ denote the corresponding integer least-squares solution. Then

$$\check{a}_{LS} = \arg \min_{z \in Z^n} \|\hat{a} - z\|_{Q_{\hat{a}}}^2 \quad (23)$$

In contrast to integer rounding and integer bootstrapping, an integer search is needed to compute \check{a}_{LS} . Although we will refrain from discussing the computational intricacies of ILS estimation, the conceptual steps of the computational procedure will be described briefly. The ILS procedure is mechanized in the GNSS LAMBDA (Least-squares AMBiguity Decorrelation Adjustment) method, which is currently one of the most applied methods for GNSS carrier phase ambiguity resolution. For more information on the LAMBDA method, we refer to e.g. [6], [26] and [27] or to the textbooks [16], [24], [25]. Practical results obtained with it can be found, for example, in [3], [4], [5], [7], [8], [14], [18], [22], [34], [35].

The main steps as implemented in the LAMBDA method are as follows. One starts by defining the ambiguity search space

$$\Omega_a = \{a \in Z^n \mid (\hat{a} - a)^T Q_{\hat{a}}^{-1}(\hat{a} - a) \leq \chi^2\} \quad (24)$$

with χ^2 a to be chosen positive constant. The boundary of this search space is ellipsoidal. It is centered at \hat{a} , its shape is governed by the vc-matrix $Q_{\hat{a}}$ and its size is determined by χ^2 . In case of GNSS, the search space is usually extremely elongated, due to the high correlations between the ambiguities. Since this extreme elongation usually hinders the computational efficiency of the search, the search space is first transformed to a more spherical shape,

$$\Omega_z = \{z \in Z^n \mid (\hat{z} - z)^T Q_{\hat{z}}^{-1}(\hat{z} - z) \leq \chi^2\} \quad (25)$$

using the admissible ambiguity transformations $\hat{z} = Z^T \hat{a}$, $Q_{\hat{z}} = Z^T Q_{\hat{a}} Z$. Ambiguity transformations Z are said to be admissible when both Z and its inverse Z^{-1} have integer entries. Such matrices preserve the integer nature of the ambiguities. In order for the transformed search space to become more spherical, the volume-preserving Z -transformation is constructed as a transformation that decorrelates the ambiguities as much as possible. Using the triangular decomposition of $Q_{\hat{z}}$, the left-hand side of the quadratic inequality in (25) is then written as a sum-of-squares:

$$\sum_{i=1}^n \frac{(\hat{z}_{i|I} - z_i)^2}{\sigma_{i|I}^2} \leq \chi^2 \quad (26)$$

On the left-hand side one recognizes the conditional least-squares estimator $\hat{z}_{i|I}$, which follows when the conditioning takes place on the integers z_1, z_2, \dots, z_{i-1} . Using the sum-of-squares structure, one can finally set up the n intervals which are used for the search. These sequential intervals are given as

$$\begin{aligned} (\hat{z}_1 - z_1)^2 &\leq \sigma_1^2 \chi^2 \\ (\hat{z}_{2|1} - z_2)^2 &\leq \sigma_{2|1}^2 \left(\chi^2 - \frac{(\hat{z}_1 - z_1)^2}{\sigma_1^2} \right) \\ &\vdots \end{aligned} \quad (27)$$

In order for the search to be efficient, one not only would like the vc-matrix $Q_{\hat{z}}$ to be as close as possible to a diagonal matrix, but also that the search space does not contain too many integer grid points. This requires the choice of a small value for χ^2 , but one that still guarantees that the search space contains at least one integer grid point. Since the bootstrapped estimator is so easy to compute and at the same time gives a good approximation to the ILS estimator (see section 3.4), the bootstrapped solution is an excellent candidate for setting the size of the ambiguity search space. Following the decorrelation step $\hat{z} = Z^T \hat{a}$, the LAMBDA-method therefore uses, as one of its options, the bootstrapped solution \check{z}_B for setting the size of the ambiguity search space as

$$\chi^2 = (\hat{z} - \check{z}_B)^T Q_{\hat{z}}^{-1} (\hat{z} - \check{z}_B) \quad (28)$$

In this way one can work with a very small search space and still guarantee that the sought for integer least-squares solution is contained in it.

3.2 The ILS pull-in region

The pull-in regions of integer rounding are unit cubes, while those of integer bootstrapping are multivariate versions of parallelograms. To determine the ILS pull-in regions we need to know the set of 'float' solutions $\hat{a} \in R^n$ that are mapped to the same integer vector $z \in Z^n$. This set is described by all $x \in R^n$ that satisfy $z = \arg \min_{u \in Z^n} \|x - u\|_{Q_{\hat{a}}}^2$. The ILS pull-in region that belongs to the integer vector z follows therefore as

$$S_{LS,z} = \{x \in R^n \mid \|x - z\|_{Q_{\hat{a}}}^2 \leq \|x - u\|_{Q_{\hat{a}}}^2, \forall u \in Z^n\} \quad (29)$$

It consists of all those points which are closer to z than to any other integer point in R^n . The metric used for measuring these distances is determined by the vc-matrix $Q_{\hat{a}}$. Based on (29), one can give a representation of the ILS pull-in regions that resembles the representation of the bootstrapped pull-in regions. This representation reads as follows.

Theorem 5 (ILS pull-in regions)

The pull-in regions of the ILS ambiguity estimator $\check{a}_{LS} \in Z^n$ are given as

$$\begin{aligned} S_{LS,z} &= \bigcap_{c_i \in Z^n} \{x \in R^n \mid |c_i^T Q_{\hat{a}}^{-1}(x - z)| \\ &\leq \frac{1}{2} \|c_i\|_{Q_{\hat{a}}}^2\}, \quad \forall z \in Z^n \end{aligned} \quad (30)$$

This shows that the ILS pull-in regions are constructed from intersecting half-spaces. One can also show that at most $2^n - 1$ pairs of such half spaces are needed for constructing the pull-in region. The ILS pull-in regions are convex, symmetric sets of volume 1, which satisfy the conditions of Definition 1. The ILS estimator is therefore admissible. The ILS pull-in regions are hexagons in the two-dimensional case (see figure 7).

3.3 Maximizing the success rate

Although various integer estimators exist which are admissible, some may be better than others. Having the problem of GNSS ambiguity resolution in mind, one is particularly

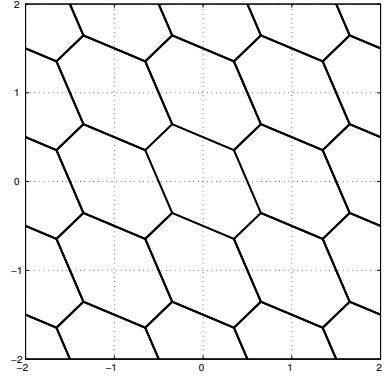


Fig. 7 The two-dimensional ILS pull-in regions.

interested in the estimator which maximizes the probability of correct integer estimation. This probability equals $P(\check{a} = a)$, but it will differ for different ambiguity estimators. The following theorem, due to [31], shows that the ILS estimator maximizes the probability of correct integer estimation.

Theorem 6 (ILS is optimal)

Let the pdf of the 'float' solution \hat{a} be given as

$$p_a(x) = \sqrt{\det(Q_{\hat{a}}^{-1})} G(\|x - a\|_{Q_{\hat{a}}}^2) \quad (31)$$

where $G : R \mapsto [0, \infty)$ is decreasing and $Q_{\hat{a}}$ is positive-definite. Then

$$P(\check{a}_{LS} = a) \geq P(\check{a} = a) \quad (32)$$

for any admissible estimator \check{a} .

This theorem gives a probabilistic justification for using the ILS estimator. For GNSS ambiguity resolution it shows, that one is better off using the ILS estimator than any other admissible integer estimator. The family of distributions defined in (31), is known as the family of elliptically contoured distributions. Several important distributions belong to this family. The multivariate normal distribution can be shown to be a member of this family by choosing $G(x) = (2\pi)^{-\frac{n}{2}} \exp -\frac{1}{2}x^T x, x \in R$. Another member is the multivariate *t*-distribution.

As a direct consequence of the above theorem we have the following corollary.

Corollary 3 (The effect of the weight matrix)

Let Σ be any positive-definite matrix of order n and define

$$\check{a}_{\Sigma} = \arg \min_{z \in Z^n} \|\hat{a} - z\|_{\Sigma}^2 \quad (33)$$

Then \check{a}_{Σ} is admissible and

$$P(\check{a}_{LS} = a) \geq P(\check{a}_{\Sigma} = a) \quad (34)$$

In order to prove the corollary, we only need to show that \check{a}_{Σ} is admissible. Once this has been established, the stated result (34) follows from theorem 6. The admissibility can be shown as follows. The first two conditions of Definition 1 are satisfied, since the ILS-map produces - apart from boundary ties - a unique integer vector for any 'float' solution $\hat{a} \in R^n$.

And since $\check{a}_\Sigma = \arg \min_{z \in Z^n} \| \hat{a} - u - z \|_\Sigma^2 + u$ holds true for any integer $u \in Z^n$, also the integer remove-restore technique applies.

As the corollary shows, a proper choice of the data weight matrix is also of importance for ambiguity resolution. The choice of weights is optimal when the weight matrix equals the inverse of the ambiguity vc-matrix. A too optimistic precision description or a too pessimistic precision description, will both result in a less than optimal ambiguity success rate. In the case of GNSS, the observation equations (the functional model) are sufficiently known and well documented. However, the same can not yet be said of the vc-matrix of the GNSS data. In the many GNSS textbooks available, we will usually find only a few comments, if any, on this vc-matrix. Examples of studies that have been reported in the literature are: [9], [10], [11], and [17], who studied the elevation dependence of the observation variances; [18] and [36], who considered time correlation and cross correlation; and [23], [2] and [28], who considered the inclusion of stochastic ionospheric constraints.

3.4 Bounding the ILS success rate

A very useful application of theorem 6 is that it shows how one can *lower bound* the ILS probability of correct integer estimation. This is particularly useful since the ILS success rate is usually difficult to compute. This is due to the rather complicated geometry of the ILS pull-in region. The bootstrapped success rate is a good candidate for the ILS success rates' lower bound. The bootstrapped success rate is easy to compute and it becomes a sharp lower bound when applied to the decorrelated ambiguities $\hat{z} = Z^T \hat{a}$. In fact, at present, the bootstrapped success rate is the sharpest available lower bound of the ILS success rate.

Apart from having a lower bound, it is also useful to have an upper bound available. For obtaining an upper bound one can make use of the *geometric mean* of the ambiguity conditional variances. This geometric mean is referred to as the Ambiguity Dilution of Precision (ADOP) and it is given as

$$\text{ADOP} = \sqrt{\det Q_{\hat{a}}}^{\frac{1}{n}} \quad (\text{cycles}) \quad (35)$$

Note that this scalar measure of the ambiguity precision is invariant for the admissible volume preserving ambiguity transformations. With the ADOP one can obtain an upper bound by making use of the fact that the probability content of the ILS pull-in region $S_{LS,a}$ would be maximal if its shape would coincide with that of the ambiguity search space, while its volume would still be constrained to 1. We have the following bounds for the ILS success rate.

Theorem 7 (Bounds on the ILS success rate)

The ILS success rate $P(\check{a}_{LS} = a)$ is bounded from below and from above as

$$P(\check{z}_B = z) \leq P(\check{a}_{LS} = a) \leq P\left(\chi^2(n, 0) \leq \frac{c_n}{\text{ADOP}^2}\right) \quad (36)$$

with: $c_n = \left(\frac{n}{2}\Gamma\left(\frac{n}{2}\right)\right)^{2/n}/\pi$

4 The PDF of the Ambiguity Residuals

Having determined the various probabilistic properties of the 'fixed' ambiguities and 'fixed' baseline solution, we are now in a position to draw our attention to the ambiguity residuals. In order to determine their pdf, we first need the joint distribution of the 'float' ambiguities, \hat{a} , and the 'fixed' ambiguities, \check{a} . The joint distribution of \hat{a} and \check{a} will be denoted as $p_{\hat{a},\check{a}}(x, z)$. We have the following result.

Theorem 8 (*The joint distribution of 'float' and 'fixed' ambiguities*)

Let $p_{\hat{a}}(x)$ be the pdf of \hat{a} and let the integer ambiguity estimator be defined as $\check{a} = \sum_{z \in Z^n} z s_z(\hat{a})$, with $s_z(x)$ the indicator function of the pull-in region $S_z \subset R^n$, $z \in Z^n$. The joint distribution of \hat{a} and \check{a} is then given as

$$p_{\hat{a},\check{a}}(x, z) = p_{\hat{a}}(x) s_z(x), \quad x \in R^n, z \in Z^n \quad (37)$$

For the proof we refer to [33].

The joint distribution (37) can be used to determine the distribution of functions of \hat{a} and \check{a} . An important example of such a function is the *ambiguity residual*. We define it as

$$\check{\epsilon} = \hat{a} - \check{a} \quad (38)$$

We will determine the distribution of $\check{\epsilon}$ in two steps. We first determine the joint distribution of $\check{\epsilon}$ and \check{a} and then determine the marginal distribution $p_{\check{\epsilon}}(x)$ by means of summation. In deriving the joint distribution, we make use of the following transformation law for probability density functions. Let two random vectors u and v be related as $v = Tu + t$, with T and t known, and matrix T invertible. The pdf of v can then be expressed in the pdf of u as $p_v(v) = |\det T^{-1}| p_u(T^{-1}(v - t))$.

By using the invertible transformation

$$\begin{bmatrix} \check{\epsilon} \\ \check{a} \end{bmatrix} = \begin{bmatrix} I_n & -I_n \\ 0 & I_n \end{bmatrix} \begin{bmatrix} \hat{a} \\ \check{a} \end{bmatrix}$$

we can express the joint distribution of $\check{\epsilon}$ and \check{a} in the joint distribution of \hat{a} and \check{a} . This gives $p_{\check{\epsilon},\check{a}}(x, z) = p_{\hat{a},\check{a}}(x + z, z)$. If we now make use of (37) we obtain $p_{\check{\epsilon},\check{a}}(x, z) = p_{\hat{a}}(x + z) s_0(x)$, $x \in R^n, z \in Z^n$. The pdf of the ambiguity residuals follows then from summing this joint distribution over all integers,

$$p_{\check{\epsilon}}(x) = \sum_{z \in Z^n} p_{\hat{a}}(x + z) s_0(x), \quad x \in R^n, z \in Z^n \quad (39)$$

In order to show how the distribution of the ambiguity residuals is constructed from the distribution of the 'float' ambiguities, a visualization of the steps involved is given in figure 8 for the one-dimensional case ($n = 1$). This figure shows the four distributions, $p_{\hat{a}}(x)$ (top left), $p_{\hat{a},\check{a}}(x, z)$ (top right), $p_{\check{\epsilon},\check{a}}(x, z)$ (bottom left) and $p_{\check{\epsilon}}(x)$ (bottom right). For each integer $z \in Z^n$ the joint distribution $p_{\hat{a},\check{a}}(x, z)$ is composed of slices from the marginal distribution $p_{\hat{a}}(x)$ located at (z, z) . Translating these slices parallel to the x -axis to the line $x = 0$ gives the joint pdf $p_{\check{\epsilon},\check{a}}(x, z)$. A further translation along the z -axis to the origin then finally provides $p_{\check{\epsilon}}(x)$. The distribution of the ambiguity residuals is clearly non-Gaussian. We have $p_{\check{\epsilon}}(x) = 0$ for all $x \notin S_0$. This implies that the norm of

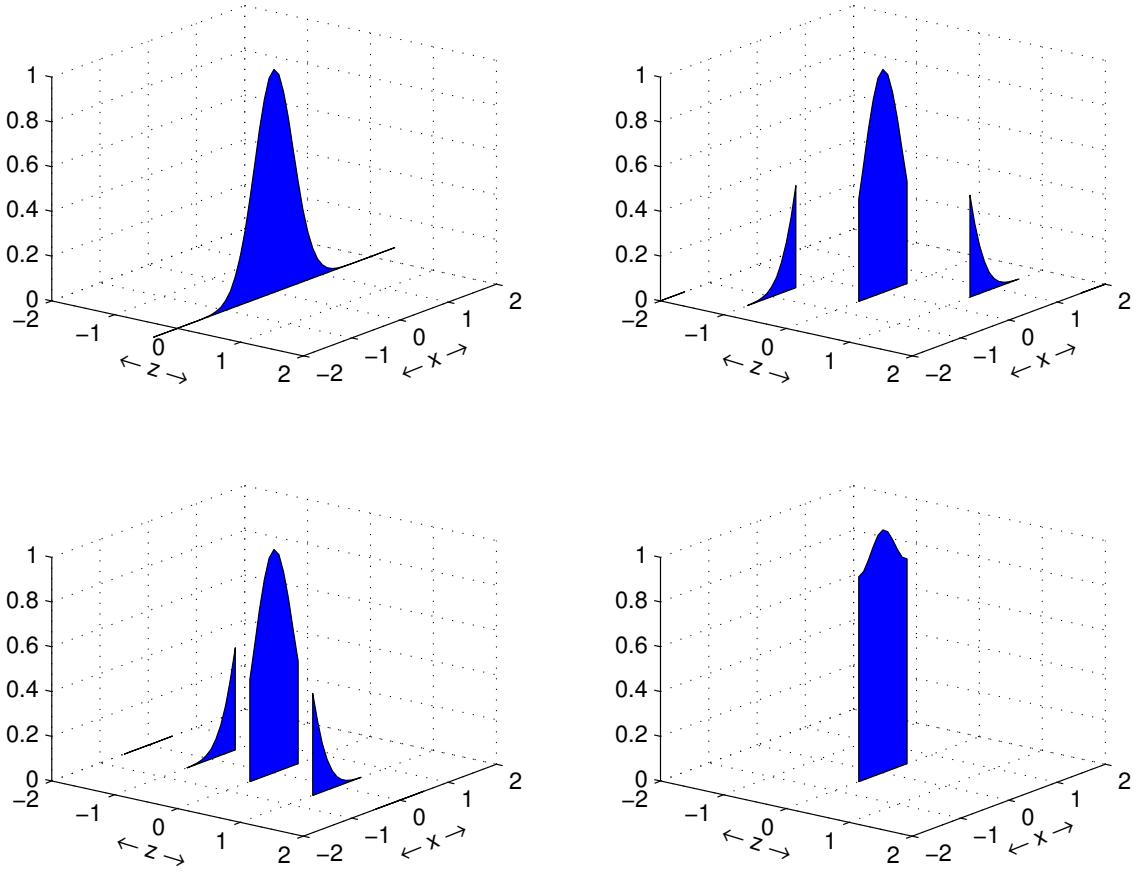


Fig. 8 The construction of $p_\epsilon(x)$ from $p_{\hat{a}}(x)$. (top left) pdf $p_{\hat{a}}(x)$, (top right) joint pdf $p_{\hat{a},\hat{a}}(x,z)$, (bottom left) joint pdf $p_{\epsilon,\hat{a}}(x,z)$, (bottom right) pdf $p_\epsilon(x)$.

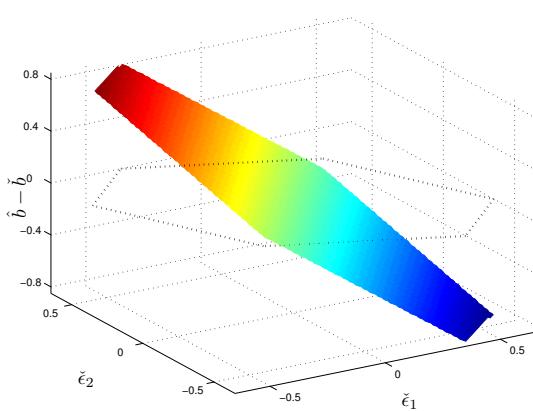


Fig. 9 An example showing the bounded influence of the ambiguity residuals on the difference between the 'float' and 'fixed' baseline solution.

the vector of ambiguity residuals is *always* bounded irrespective the values taken by the ambiguity 'float' solution \hat{a} . For the one-dimensional case we have $\check{\epsilon} \in [-1/2, +1/2]$. For the higher dimensional case the bound depends on the shape of the pull-in region and therefore on the type of integer estimator chosen. The fact that the ambiguity residuals are bounded has an important implication. It implies that the difference between the 'float' baseline solution \hat{b} and the 'fixed' baseline solution \bar{b} is also bounded, irrespective the values taken by the 'float' ambiguities (see figure 9).

The distribution $p_\epsilon(x)$ can have different shapes. There are two extreme cases between which one can discriminate, namely the uniform distribution and the impulse function distribution. In order to understand these two extreme cases, consider what happens when the precision of the 'float' ambiguities, and thereby the peakedness of $p_{\hat{a}}(x)$, is varied. When the 'float' distribution $p_{\hat{a}}(x)$ gets more peaked, which happens when the 'float' ambiguities get more precise, the pdf $p_\epsilon(x)$ gets more peaked as well. However, since all the probability mass of $p_\epsilon(x)$ is located within the pull-in region S_0 , the peakedness of $p_\epsilon(x)$ will only start to manifest itself when $p_{\hat{a}}(x)$ is sufficiently peaked in relation to the size of the pull-in region. When this is not the case, the distribution of the ambiguity residuals will remain flat and therefore be close to $s_0(x)$, which is the uniform distribution for the pull-in region S_0 . The 'float' distribution $p_{\hat{a}}(x)$ may be considered peaked in relation to the size of the pull-in region, when most of its probability mass is located within S_0 . This happens when the ambiguity success rate is sufficiently close to one, in which case the two distributions will also not differ by much. Further improvement of the precision of the ambiguities will then in the limit produce an impulse function for both $p_{\hat{a}}(x)$ and $p_\epsilon(x)$ (see figure 10 for the one-dimensional case). The distribution of the ambiguity residuals is symmetric and independent of the unknown integer ambiguity vector $a \in \mathbb{Z}^n$. It gets its symmetry inherited from the 'float' distribution $p_{\hat{a}}(x)$, while the independence of $a \in \mathbb{Z}^n$ follows from $\sum_{z \in \mathbb{Z}^n} p_{\hat{a}}(x + a + z) = \sum_{z \in \mathbb{Z}^n} p_{\hat{a}}(x + z)$. The point of symmetry of the distribution is the origin. This implies that

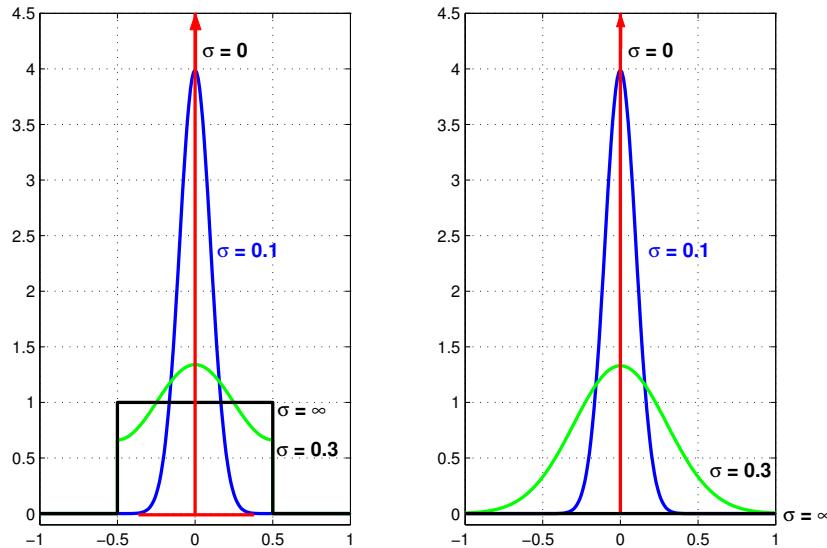


Fig. 10 One dimensional pdf for different values of the ambiguity standard deviation (in cycles): (left) pdf of ambiguity residual, (right) pdf of corresponding 'float' ambiguity.

the mean of the ambiguity residual equals zero,

$$E\{\tilde{\epsilon}\} = 0 \quad (40)$$

This result combined with the fact the pdf of $\tilde{\epsilon}$ is completely known once the precision of the 'float' ambiguities is given and once the choice of integer ambiguity estimator is made, allows one for the first time to formulate rigorous tests for the integerness of the parameters.

Acknowledgement

The author acknowledges the help of Sandra Verhagen and Peter Joosten in preparing (figures and formatting) the present contribution.

References

- [1] Betti, B., M. Crespi, and F. Sanso (1993): A geometric illustration of ambiguity resolution in GPS theory and a Bayesian approach. *Manuscr Geod*, 18: 317-330.
- [2] Bock, Y. (1998): Medium distance GPS measurements. In: Teunissen, P.J.G., and A. Kleusberg (Eds), *GPS for Geodesy*, 2nd edition, Springer Verlag.
- [3] Boon, F., B. Ambrosius (1997): Results of real-time applications of the LAMBDA method in GPS based aircraft landings. *Proceedings KIS97*, pp. 339-345.
- [4] Boon, F., P.J. de Jonge, C.C.J.M. Tiberius (1997): Precise aircraft positioning by fast ambiguity resolution using improved troposphere modelling. *Proceedings ION GPS-97*, Vol. 2, pp. 1877-1884.
- [5] Cox, D.B. and J.D.W. Brading (1999): Integration of LAMBDA ambiguity resolution with Kalman filter for relative navigation of spacecraft. *Proceedings ION NTM-99*, pp. 739-745.
- [6] de Jonge P.J., C.C.J.M. Tiberius (1996a): The LAMBDA method for integer ambiguity estimation: implementation aspects. Publications of the Delft Computing Centre, *LGR-Series No. 12*.
- [7] de Jonge, P.J., C.C.J.M. Tiberius (1996b): Integer estimation with the LAMBDA method. *Proceedings IAG Symposium No. 115, 'GPS trends in terrestrial, airborne and spaceborne applications'*, G. Beutler et al. (eds), Springer Verlag, pp. 280-284.
- [8] de Jonge, P.J., C.C.J.M. Tiberius, P.J.G. Teunissen (1996): Computational aspects of the LAMBDA method for GPS ambiguity resolution. *Proceedings ION GPS-96*, pp. 935-944.
- [9] Euler, H.J., C. Goad (1991): On optimal filtering of GPS dual frequency observations without using orbit information. *Bull Geod*, 65: 130- 143.
- [10] Gerdan, G.P. (1995): A comparison of four methods of weighting double difference pseudo range measurements. *Trans Tasman Surv*, 1(1): 60-66.
- [11] Gianniu, M. (1996): Genauigkeitssteigerung bei kurzzeit-statischen und kinematischen Satellitenmessungen bis hin zu Echtzeitanwendung. PhD-thesis, DGK, Reihe C, no. 458, Mnchen.
- [12] Gundlich, B., K-R. Koch (2002): Confidence regions for GPS baselines by Bayesian statistics. *Journal of Geodesy*, 76: 55-62.
- [13] Gundlich, B., P.J.G. Teunissen (2002): Multiple Models: Fixed, Switching and Interacting. *Proceedings V Hotine-Marussi Symposium*, Matera, Italy.
- [14] Han, S. (1995): Ambiguity resolution techniques using integer least-squares estimation for rapid static or kinematic positioning. *Symposium Satellite Navigation Technology: 1995 and beyond*, Brisbane, Australia, 10 p.

- [15] Hassibi, A., S. Boyd (1998): Integer parameter estimation in linear models with applications to GPS. *IEEE Transactions on Signal Processing*, vol. 46, no. 11, pp. 2938-2952.
- [16] Hofmann-Wellenhof, B., H. Lichtenegger, J. Collins (2001): *Global Positioning System: Theory and Practice*. 5th edition. Springer Verlag.
- [17] Jin, X.X. and C.D. de Jong (1996): Relationship between satellite elevation and precision of GPS code measurements. *J Navig*, 49: 253-265.
- [18] Jonkman, N.F. (1998): Integer GPS ambiguity estimation without the receiver-satellite geometry. Publications of the Delft Geodetic Computing Centre, *LGR-Series*, No. 18.
- [19] Leick, A. (1995): *GPS Satellite Surveying*. 2nd edition, John Wiley, New York.
- [20] Misra, P., P. Enge (2001): *Global Positioning System: Signals, Measurements, and Performance*. Ganga-Jamuna Press.
- [21] Parkinson, B., J.J. Spilker (eds) (1996): *GPS: Theory and Applications*, Vols 1 and 2, AIAA, Washington DC.
- [22] Peng, H.M., F.R. Chang, L.S. Wang (1999): Attitude determination using GPS carrier phase and compass data. *Proceedings ION NTM-99*, pp. 727-732.
- [23] Schaffrin, B., Y. Bock (1988): A unified scheme for processing GPS dual-band observations. *Bull Geod*, 62: 142-160.
- [24] Strang, G., K. Borre (1997): *Linear Algebra, Geodesy, and GPS*, Wellesley-Cambridge Press.
- [25] Teunissen, P.J.G., A. Kleusberg (eds) (1998): *GPS for Geodesy*, 2nd enlarged edition, Springer Verlag.
- [26] Teunissen, P.J.G. (1993): Least-squares estimation of the integer GPS ambiguities. Invited Lecture, Section IV Theory and Methodology, IAG General Meeting, Beijing, China, August 1993. Also in: *LGR Series*, No. 6, Delft Geodetic Computing Centre.
- [27] Teunissen, P.J.G. (1995): The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. *Journal of Geodesy*, 70: 65-82.
- [28] Teunissen, P.J.G. (1998a): The ionosphere-weighted GPS baseline precision in canonical form. *Journal of Geodesy*, 72: 107-117.
- [29] Teunissen, P.J.G. (1998b): Success probability of integer GPS ambiguity rounding and bootstrapping. *Journal of Geodesy*, 72: 606-612.
- [30] Teunissen, P.J.G. (1999a): The probability distribution of the GPS baseline for a class of integer ambiguity estimators. *Journal of Geodesy*, 73: 275-284.
- [31] Teunissen, P.J.G. (1999b): An optimality property of the integer least-squares estimator. *Journal of Geodesy*, 73: 587-593.
- [32] Teunissen, P.J.G. (2001): The probability distribution of the ambiguity bootstrapped GNSS baseline. *Journal of Geodesy*, 75: 267-275.
- [33] Teunissen, P.J.G. (2002): The parameter distributions of the integer GPS model. *Journal of Geodesy*, 76: 41-48.
- [34] Tiberius, C.C.J.M., P.J. de Jonge (1995): Fast positioning using the LAMBDA method. *Proceedings DSNS-95*, paper 30, 8p.
- [35] Tiberius, C.C.J.M., P.J.G. Teunissen, P.J. de Jonge (1997): Kinematic GPS: performance and quality control. *Proceedings KIS97*, pp. 289-299.
- [36] Tiberius, C.C.J.M. (1998): Recursive data processing for kinematic GPS surveying. *Publ Geodesy*, no.45, Netherlands Geodetic Commission, Delft.
- [37] Verhagen, S., P.J.G. Teunissen (2002): PDF evaluation of the ambiguity residuals. *Proceedings V Hotine-Marussi Symposium*, Matera, Italy.

37 Mixed Integer Estimation and Validation for Next Generation GNSS

Peter J.G. Teunissen^{1,2}

¹Department of Spatial Sciences, Curtin University of Technology, Perth, Australia

²Department of Earth Observation and Space Systems,
Delft University of Technology, Delft, The Netherlands

1	<i>Introduction</i>	1103
1.1	Next Generation GNSS.....	1103
1.1.1	Background.....	1103
1.1.2	Benefits of GNSS.....	1103
1.1.3	Theory.....	1104
1.2	Mixed Integer Model.....	1104
1.3	Chapter Overview.....	1106
2	<i>Principles of Integer Inference</i>	1107
2.1	Integer Estimation.....	1107
2.1.1	Pull-In Regions.....	1107
2.1.2	PMF and Success-Rate.....	1108
2.1.3	Optimal Integer Estimation.....	1108
2.2	Integer Aperture Estimation.....	1109
2.2.1	Aperture Pull-In Regions.....	1109
2.2.2	Probability Distribution and Successful Fix-Rate.....	1110
2.2.3	Optimal Integer Aperture Estimation.....	1111
2.3	Integer Equivariant Estimation.....	1111
2.3.1	A Larger Class of Estimators.....	1111
2.3.2	Best Integer Equivariant Estimation.....	1112
3	<i>Three Popular Integer Estimators</i>	1113
3.1	Integer Rounding.....	1113
3.1.1	Scalar and Vectorial Rounding.....	1113
3.1.2	Rounding Success-Rate.....	1114
3.1.3	Z-Transformations.....	1115
3.2	Integer Bootstrapping.....	1115
3.2.1	The Bootstrapping Principle.....	1115
3.2.2	The Bootstrapped PMF and Success-Rate.....	1117
3.2.3	Z-Transformations.....	1118
3.3	Integer Least-Squares.....	1119
3.3.1	Mixed Integer Least-Squares.....	1119

3.3.2	The ILS Search.....	1120
3.3.3	The Least-Squares PMF and Success-Rate.....	1121
4	<i>Baseline Quality and Model Validation</i>	1122
4.1	Fixed and Float Baseline.....	1122
4.2	Cross-Validation of Mixed Integer Model.....	1123
5	<i>Conclusions</i>	1125
	<i>References</i>	1125

Abstract The coming decade will bring a proliferation of Global Navigation Satellite Systems (GNSS) that are likely to revolutionize society in the same way as the mobile phone has done. The promise of a broader multi-frequency, multi-signal GNSS “system of systems” has the potential of enabling a much wider range of demanding applications compared to the current GPS-only situation. In order to achieve the highest accuracies one must exploit the unique properties of the received carrier signals. These properties include the multi-satellite system tracking, the mm-level measurement precision, the frequency diversity, and the integer ambiguities of the carrier phases. Successful exploitation of these properties results in an accuracy improvement of the estimated GNSS parameters of two orders of magnitude. The theory that underpins this ultraprecise GNSS parameter estimation and validation is the theory of integer inference. This theory is the topic of the present chapter.

1 Introduction

1.1 Next Generation GNSS

1.1.1 Background

Global navigation satellite systems (GNSSs) involve satellites, ground stations, and user receiver equipment and software to determine positions anywhere around the world at any time. The global positioning system (GPS) from the United States is the best-known and currently fully operational GNSS. Fueling growth during the next decade will be the next generation GNSSs that are currently being deployed and developed. Current and prospective providers of GNSS systems are the United States, Russia, the European Union, China, Japan, and India. The United States is modernizing its dual-frequency GPS. A third civil frequency will be added, with expected 24-satellite full constellation capability (FOC) around 2015. Russia is revitalizing its GLONASS system, from a current only partially functioning system to 24-satellite FOC reached by 2010. The European Union is developing a complete new multi-frequency GNSS, called Galileo, which is currently in orbit validation phase and which will have its 30-satellite FOC by 2012. China is developing its own 30-satellite GNSS, called Compass, of which the first satellite was launched in April 2007. Finally, India and Japan are developing GNSS augmentation systems. India’s 7-satellite IRNSS (Indian Regional Navigational Satellite System) is expected operational in 2012 and Japan will soon launch its first of three QZSS (Quasi-Zenith Satellite System) satellites. QZSS is designed to increase the number of satellites available at high-elevation angles over Japan.

1.1.2 Benefits of GNSS

The promise of a broader and more diverse system of GNSSs has enormous potential for improving the accuracy, integrity, and efficiency of positioning worldwide, the importance of which can hardly be overstated.

The availability of many more satellites and signals creates exciting opportunities to extend current GPS applications and to enable new applications in areas where the GPS-only situation has been a hindrance to market growth. Extending the operational range of precise carrier-phase GNSS, currently restricted to about 15 km, will allow instantaneous cm-level accuracy

positioning at remote locations on land and offshore. Improved integrity will service various industries having high marginal costs (e.g., mining, agriculture, machine-guided construction). Single-frequency tracking of more satellites will create opportunities for the low-cost receiver market of precise real-time location devices, such as handheld or in moving vehicles. In addition, environmental- and spaceborne-GNSS will benefit enormously from tracking multiple satellites on multiple frequencies. Environmental GNSS, in general, benefits from denser atmospheric profiling, while short-term weather prediction in particular benefits from a reduction in the latency of GNSS integrated water vapor estimates. The benefits for spaceborne GNSS are highly accurate orbit determinations of Earth-orbiting space platforms, possibly even in real-time, thus offering increased spacecraft autonomy, simplification of spacecraft operations, and support to rapid delivery of end-user data products such as atmospheric profiles from occultation or synthetic aperture radar images for deformation monitoring. An overview of this great variety of GNSS models and their applications can be found in textbooks like Parkinson and Spilker (1996), Strang and Borre (1997), Teunissen and Kleusberg (1998), Farrell and Barth (1999), Leick (2004), Misra and Enge (2006), and Hofmann-Wellenhoff et al. (2008).

1.1.3 Theory

Several key issues need to be addressed in order to achieve the fullest exploitation of the opportunities created by future GNSSs. The highest possible accuracies can only be achieved if one is able to exploit the unique properties of the received carrier signals. These properties include the mm-level precision with which the carrier-phases can be tracked, the frequency-diversity of the carriers, and the knowledge that certain functions of the carriers are integer-valued. The process of exploiting these properties is known as integer ambiguity resolution (IAR). IAR improves the precision of the estimated GNSS model parameters by at least two orders of magnitude. For positioning, successful IAR effectively transforms the estimated fractional carrier-phases into ultraprecise receiver-satellite ranges, thus making high-precision (cm- to mm-level) positioning possible. As a beneficial by-product, it also improves other GNSS model parameters, such as atmospheric parameters, and it enables reduction of GNSS parameter-estimation space, sometimes up to 50%, thus simplifying computations considerably and accelerating the time to position. However, the success of IAR depends on the strength of the underlying GNSS model. The weaker the model, the more data needs to be accumulated before IAR can be successful and the longer it therefore takes before one can profit from the ultraprecise carrier signals. Clearly, the aim is to have short times-to-convergence, preferably zero, thereby enabling truly instantaneous GNSS positioning.

The theory that underpins ultraprecise GNSS parameter estimation is the theory of integer inference. This theory of estimation and validation is the topic of the present chapter. Although a large part of the theory has been developed since the 1990s for GPS, the theory has a much wider range of applicability.

1.2 Mixed Integer Model

Central in the theory of integer inference is the mixed integer model. To introduce this model, we use the GNSS pseudorange and carrier-phase observables as leading example. If we denote the j -frequency pseudorange and carrier-phase for the $r-s$ receiver-satellite combination at

epoch t as $p_{r,j}^s(t)$ and $\phi_{r,j}^s(t)$, respectively, then their observation equations can be formulated as

$$\begin{aligned} p_{r,j}^s(t) &= \rho_r^s(t) + T_r^s(t) + \mu_j I_r^s(t) + c dt_r^s(t) + e_{r,j}^{s,p}(t) \\ \phi_{r,j}^s(t) &= \rho_r^s(t) + T_r^s(t) - \mu_j I_r^s(t) + c \delta t_r^s(t) + \lambda_j M_{r,j}^s + e_{r,j}^{s,\phi}(t) \end{aligned} \quad (1)$$

where ρ_r^s is the receiver-satellite range, T_r^s is the tropospheric delay, I_r^s is the ionospheric delay, dt_r^s and δt_r^s are the pseudorange and carrier-phase receiver-satellite clock biases, $M_{r,j}^s$ is the time-invariant carrier-phase ambiguity, c is the speed of light, λ_j is the j -frequency wave length, $\mu_j = (\lambda_j/\lambda_1)^2$, and $e_{r,j}^{s,p}$ and $e_{r,j}^{s,\phi}$ are the remaining error terms, respectively. The real-valued carrier-phase ambiguity $M_{r,j}^s = \phi_{r,j}(t_0) + \phi_j^s(t_0) + N_{r,j}^s$ is the sum of the initial receiver-satellite phases and the integer ambiguity $N_{r,j}^s$.

Through differencing of the observation equations one can eliminate the initial phases and the clock biases. The so-called double differenced (DD) observation equations then take the form

$$\begin{aligned} p_{qr,j}^{ts}(t) &= \rho_{qr}^{ts}(t) + T_{qr}^{ts}(t) + \mu_j I_{qr}^{ts}(t) + e_{qr,j}^{ts,p}(t) \\ p_{qr,j}^{ts}(t) &= \rho_{qr}^{ts}(t) + T_{qr}^{ts}(t) - \mu_j I_{qr}^{ts}(t) + \lambda_j N_{qr,j}^{ts} + e_{qr,j}^{ts,\phi}(t) \end{aligned} \quad (2)$$

where $p_{qr,j}^{ts}(t) = [p_{r,j}^s(t) - p_{r,j}^t(t)] - [p_{q,j}^s(t) - p_{q,j}^t(t)]$, with a similar notation for the other DD variates. The DD tropospheric slant delays are usually further reduced to a single DD vertical delay T_{qr}^{vert} by means of mapping functions. Furthermore, the need for having the ionospheric delays present depends very much on the baseline length between receivers. These delays can usually be neglected for distances less than 15 km.

If we assume the error terms $e_{qr,j}^{ts,p}(t)$ and $e_{qr,j}^{ts,\phi}(t)$ in Eq. 2 to be zero-mean random variables, the observation equations can be used to set up a linear model in which some of the unknown parameters are reals and others are integer ($N_{qr,j}^{ts}$). Such a model is an example of a mixed integer linear model.

The observation equations of Eq. 2 are parametrized in the DD ranges $\rho_{qr}^{ts}(t)$. These ranges depend on the receiver positions and on the satellite positions. Assuming the satellite orbits known, these ranges are usually further linearized with respect to the unknown receiver coordinates. As a result one obtains linearized equations that are parametrized in the between-receiver baseline vector increments. Such a model is an example of a mixed integer *linearized* model. These linearized GNSS models can usually be treated as if they are linear, since the nonlinearities are small.

We now define the general form of a mixed integer linear model.

Definition 1 (Mixed integer linear model)

Let (A, B) be a given $m \times (n + p)$ matrix of full rank and let Q_{yy} be a given $m \times m$ positive definite matrix. Then

$$y \sim N(Aa + Bb, Q_{yy}), \quad a \in \mathbb{Z}^n, \quad b \in \mathbb{R}^p \quad (3)$$

will be referred to as the mixed integer linear model.

The notation “~” is used to describe “distributed as.” In a GNSS context, the m -vector y contains the pseudorange and carrier-phase observables, the n -vector a the integer DD ambiguities, and the real-valued p -vector b the remaining unknown parameters, such as baseline components (coordinates) and possibly atmospheric delay parameters (troposphere, ionosphere). As in most GNSS applications, the underlying distribution of the above mixed integer model is

assumed to be a multivariate normal distribution. Various results in the following sections are also valid, however, for general distributions.

1.3 Chapter Overview

The mixed integer model (Eq. 3) is usually solved and validated in a number of steps. We now briefly present the contributions of this chapter in relation to these steps.

In the first step, the integer nature of a is discarded. The parameters a and b are estimated using least-squares (LS) estimation, which in the present case is equivalent to using maximum likelihood (ML) or best linear unbiased estimation (BLUE). As a result one obtains the so-called float solution,

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} \sim N \left(\begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} Q_{\hat{a}\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}\hat{b}} \end{bmatrix} \right) \quad (4)$$

In this first step, one usually also tests the data and GNSS model for possible model misspecifications, e.g., outliers, cycle slips, or other modeling errors. This can be done with the standard theory of hypothesis testing (Baarda 1968; Koch 1999; Teunissen 2006).

In the second step, a mapping $\mathcal{I} : \mathbb{R}^n \mapsto \mathbb{Z}^n$ is introduced that takes the integer constraints $a \in \mathbb{Z}^n$ into account,

$$\check{a} = \mathcal{I}(\hat{a}) \quad (5)$$

There are many ways in which the mapping \mathcal{I} can be defined. In [Sect. 2](#), we introduce three different classes of such estimators. They are the class of integer estimators (I), the class of integer aperture estimators (IA), and the class of integer equivariant estimators (IE). These three classes were introduced in Teunissen (1999a, 2003a, b). They are subsets of one another and related as

$$I \subset IA \subset IE \quad (6)$$

Each class consists of a multitude of estimators. For each class we present the optimal estimator. As optimality criterion we either use the maximization of the probability of correct integer estimation or the minimization of the mean squared error.

Since estimators from the class of integer estimators are most often used, we present in [Sect. 3](#) an analysis of the properties of the three most popular integer estimators. They are the estimators of integer rounding (IR), integer bootstrapping (IB), and integer least-squares (ILS). Special attention is given to computational issues and to their success-rates, i.e., the probabilities of correct integer estimation. It is shown that the performances of these three integer estimators are related as

$$P(\check{a}_{\text{IR}} = a) \leq P(\check{a}_{\text{IB}} = a) \leq P(\check{a}_{\text{ILS}} = a) \quad (7)$$

Knowing that a is integer, strengthens the model and allows one, in principle, to re-evaluate the validation of the model as compared to the first step validation. However, the standard theory of hypothesis testing is then not applicable anymore. This validation problem will be addressed in [Sect. 4](#), where we also present a cross-validation method for the mixed integer model.

In the final step, once \check{a} of Eq. 5 is computed and accepted, the float estimator \hat{b} is readjusted to obtain the so-called fixed estimator

$$\check{b} = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}\hat{a}}^{-1} (\hat{a} - \check{a}) \quad (8)$$

Whether or not \check{b} is an improvement over \hat{b} depends in a large part on the probabilistic properties of \check{a} . We therefore present in [Sect. 4](#) the probability density function (PDF) of \check{b} and show how it is influenced by the probability mass function (PMF) of \check{a} .

2 Principles of Integer Inference

In this section, we present three different classes of integer parameter estimators. They are the integer estimators, the integer aperture estimators, and the integer equivariant estimators. Within each class we determine the optimal estimator.

2.1 Integer Estimation

2.1.1 Pull-In Regions

We start with the requirement that the estimator \check{a} needs to be integer, $\check{a} = \mathcal{I}(\hat{a}) \in \mathbb{Z}^n$. Then $\mathcal{I} : \mathbb{R}^n \mapsto \mathbb{Z}^n$ is a many-to-one map, instead of a one-to-one map. Different real-valued vectors will be mapped to one and the same integer vector. One can therefore assign a subset, say $\mathcal{P}_z \subset \mathbb{R}^n$, to each integer vector $z \in \mathbb{Z}^n$,

$$\mathcal{P}_z = \{x \in \mathbb{R}^n \mid z = \mathcal{I}(x)\}, \quad z \in \mathbb{Z}^n \quad (9)$$

This subset is referred to as the *pull-in region* of z . It is the region in which all vectors are pulled to the same integer vector z .

The concept of pull-in regions can be used to define integer estimators.

Definition 2 (Integer estimators)

The mapping $\check{a} = \mathcal{I}(\hat{a})$ is said to be an integer estimator if its pull-in regions satisfy

- (1) $\bigcup_{z \in \mathbb{Z}^n} \mathcal{P}_z = \mathbb{R}^n$
- (2) $\text{Int}(\mathcal{P}_u) \cap \text{Int}(\mathcal{P}_v) = \emptyset, \quad \forall u, v \in \mathbb{Z}^n, \quad u \neq v$
- (3) $\mathcal{P}_z = z + \mathcal{P}_0, \quad \forall z \in \mathbb{Z}^n$

According to this definition an integer estimator is completely specified once the pull-in region \mathcal{P}_0 is given. The following explicit expression can be given for an integer estimator,

$$\check{a} = \sum_{z \in \mathbb{Z}^n} z p_z(\hat{a}) \quad (10)$$

with the indicator function, $p_z(x)$, defined as

$$p_z(x) = \begin{cases} 1 & \text{if } x \in \mathcal{P}_z \\ 0 & \text{if } x \notin \mathcal{P}_z \end{cases}$$

Note, since $\sum_{z \in \mathbb{Z}^n} p_z(x) = 1, \forall x \in \mathbb{R}^n$, that the $p_z(\hat{a})$ can be interpreted as weights. The integer estimator \check{a} is therefore equal to a weighted sum of integer vectors with binary weights. Examples of I-estimators are integer rounding, integer bootstrapping, and integer least-squares. Their pull-in regions are the multivariate versions of a square, a parallelogram, and a hexagon. The properties of these popular integer estimators will be further detailed in [Sect. 3](#).

2.1.2 PMF and Success-Rate

The outcome of an integer estimator should only be used if one has enough confidence in its solution. To evaluate one's confidence in \check{a} , one needs its PMF. The PMF of \check{a} is obtained by integrating the PDF of \hat{a} , $f_{\hat{a}}(x|a)$, over the pull-in regions $\mathcal{P}_z \subset \mathbb{R}^n$,

$$P(\check{a} = z) = P(\hat{a} \in \mathcal{P}_z) = \int_{\mathcal{P}_z} f_{\hat{a}}(x|a) dx, \quad z \in \mathbb{Z}^n \quad (11)$$

In case $\hat{a} \sim N(a, Q_{\hat{a}\hat{a}})$, the PDF is given as $f_{\hat{a}}(x|a) = C \exp\left\{-\frac{1}{2} \|x - a\|_{Q_{\hat{a}\hat{a}}}^2\right\}$, where C is a normalizing constant and $\|\cdot\|_M^2 = (\cdot)^T M^{-1}(\cdot)$.

The PMF of \check{a} depends, of course, on the pull-in regions \mathcal{P}_z and therefore on the chosen integer estimator. Since various integer estimators exist, some may be better than others. Having the problem of GNSS ambiguity resolution in mind, one is particularly interested in the probability of correct integer estimation. This probability is referred to as the *success-rate*, $P_s = P(\check{a} = a)$. Its complement is referred to as the *fail-rate*, $P_f = 1 - P_s$. The success-rate is computed as

$$P_s = \int_{\mathcal{P}_a} f_{\hat{a}}(x|a) dx = \int_{\mathcal{P}_0} f_{\hat{a}}(x + a|a) dx \quad (12)$$

This shows, if the PDF has the translational property $f_{\hat{a}}(x + a|a) = f_{\hat{a}}(x|0)$, that the success-rate can be computed without knowledge of the unknown integer vector $a \in \mathbb{Z}^n$. Obviously, this is the case for the PDF of the multivariate normal distribution.

[Equation 12](#) is very important for GNSS applications. It allows the GNSS user to evaluate (often even before the actual measurements are taken) whether or not the strength of the underlying GNSS model is such that one can expect successful integer ambiguity resolution. The evaluation of the multivariate integral of [Eq. 12](#) can generally be done through Monte Carlo integration (Robert and Casella 1999). For some important integer estimators we also have easy-to-compute expressions and/or sharp (lower and upper) bounds of their success-rates available (cf. [Sect. 3](#)).

2.1.3 Optimal Integer Estimation

Since the success-rate depends on the pull-in region and therefore on the chosen integer estimator, it is of importance to know which integer estimator maximizes the probability of correct integer estimation ([Teunissen 1999b](#)).

Theorem 1 (Optimal integer estimation)

Let $f_{\hat{a}}(x|a)$ be the PDF of \hat{a} and let the integer maximum likelihood (IML) estimator

$$\check{a}_{\text{IML}} = \arg \max_{z \in \mathbb{Z}^n} f_{\hat{a}}(\hat{a}|z) \quad (13)$$

be an integer estimator. Then

$$P(\check{a}_{\text{IML}} = a) \geq P(\check{a} = a) \quad (14)$$

for any integer estimator \check{a} .

This result shows that of all integer estimators, the IML-estimator has the largest success-rate. The theorem holds true for an arbitrary PDF of \hat{a} . In case $\hat{a} \sim N(a, Q_{\hat{a}\hat{a}})$, the optimal I-estimator is the integer least-squares (ILS) estimator

$$\check{a}_{\text{ILS}} = \arg \min_{a \in \mathbb{Z}^n} \| \hat{a} - a \|_{Q_{\hat{a}\hat{a}}}^2 \quad (15)$$

The above theorem therefore gives a probabilistic justification for using the ILS estimator when the PDF is Gaussian. We will have more to say about the ILS-estimator in [Sect. 3.3](#).

2.2 Integer Aperture Estimation

2.2.1 Aperture Pull-In Regions

The outcome of an integer (I) estimator is always an integer, whether the fail-rate is large or small. Since the user has no direct control over this fail-rate (other than strengthening the underlying model a priori), the user has no direct influence on the confidence of its integer solution.

To give the user control over the fail-rate, the class of integer aperture (IA) estimators was introduced. This class is larger than the I-class. The IA-class is defined by dropping one of the three conditions in Definition 2, namely the condition that the pull-in regions should cover \mathbb{R}^n completely. We thus allow the IA pull-in regions to have gaps.

Definition 3 (Integer aperture estimators)

Let $\Omega \subset \mathbb{R}^n$ and $\Omega_z = \Omega \cap \mathcal{P}_z$, where \mathcal{P}_z is a pull-in region of an arbitrary I-estimator. Then the estimator

$$\check{a} = \begin{cases} \hat{a} & \text{if } \hat{a} \notin \Omega \\ z & \text{if } \hat{a} \in \Omega_z \end{cases} \quad (16)$$

is said to be an integer aperture estimator if its pull-in regions satisfy

- (1) $\bigcup_{z \in \mathbb{Z}^n} \Omega_z = \Omega \subset \mathbb{R}^n$
- (2) $\text{Int}(\Omega_u) \cap \text{Int}(\Omega_v) = \emptyset, \quad \forall u, v \in \mathbb{Z}^n, \quad u \neq v$
- (3) $\Omega_z = z + \Omega_0, \quad \forall z \in \mathbb{Z}^n$

If we compare with Definition 2, we note that the role of the complete space \mathbb{R}^n has been replaced by the subset $\Omega \subset \mathbb{R}^n$. It is easily verified from the above conditions that Ω is z -translational invariant, $\Omega = \Omega + z, \forall z \in \mathbb{Z}^n$. Also note that the I-class is a subset of the IA-class. Thus I-estimators are IA-estimators, but the converse is not necessarily true.

An IA-estimator maps the float solution \hat{a} to the integer vector z if $\hat{a} \in \Omega_z$ and it maps the float solution to itself if $\hat{a} \notin \Omega$. An IA-estimator can therefore be expressed explicitly as

$$\check{a}_{\text{IA}} = \hat{a} + \sum_{z \in \mathbb{Z}^n} (z - \hat{a}) \omega_z(\hat{a}) \quad (17)$$

with $\omega_z(x)$ the indicator function of Ω_z .

Note that the IA-estimator is completely determined once $\Omega_0 \subset \mathcal{P}_0$ is given. Thus Ω_0 plays the same role for the IA-estimators as \mathcal{P}_0 does for the I-estimators. By changing the size and shape of Ω_0 one changes the outcome of the IA-estimator. The subset Ω_0 can therefore be seen as an adjustable pull-in region with two limiting cases. The limiting case in which Ω_0 is empty

and the limiting case when Ω_0 equals \mathcal{P}_0 . In the first case the IA-estimator becomes identical to the float solution \hat{a} , and in the second case the IA-estimator becomes identical to an I-estimator. The subset Ω_0 therefore determines the *aperture* of the pull-in region.

2.2.2 Probability Distribution and Successful Fix-Rate

In order to evaluate the performance of an IA-estimator, the following three outcomes need to be distinguished: $\check{a} \in \Omega_a$ for success (correct integer estimation), $\check{a} \in \Omega \setminus \Omega_a$ for failure (incorrect integer estimation), and $\check{a} \notin \Omega$ for undecided (a not estimated as integer). The corresponding probabilities of success (s), failure (f), and undecided (u) are given as

$$\begin{aligned} P_s &= P(\check{a}_{\text{IA}} = a) = \int_{\Omega_a} f_{\check{a}}(x|a) dx \\ P_f &= \sum_{z \in \mathbb{Z}^n \setminus \{a\}} P(\check{a}_{\text{IA}} = z) = \sum_{z \in \mathbb{Z}^n \setminus \{a\}} \int_{\Omega_z} f_{\check{a}}(x|a) dx \\ P_u &= 1 - P_s - P_f = 1 - \int_{\Omega_0} f_{\check{a}}(x|a) dx \end{aligned} \quad (18)$$

where

$$f_{\check{a}}(x|a) = \sum_{z \in \mathbb{Z}^n} f_{\check{a}}(x + z|a) p_0(z) \quad (19)$$

is the PDF of the residual $\check{\delta} = \hat{a} - \check{a}$ (Teunissen 2002, Verhagen and Teunissen 2005).

Since $\Omega_a \subset \mathcal{P}_a$, it follows that the success-rate of an IA-estimator will never be larger than that of the corresponding I-estimator. So what have we gained? What we have gained is that we can now control the fail-rate and in particular control the successful fix-rate, i.e., the probability of successful fixing. Note that the complement of the undecided probability, $1 - P_u = P_s + P_f$, is the fix-probability, i.e., the probability that the outcome of the IA-estimator is integer. The probability of successful fixing is therefore given by the ratio

$$P_{sf} = \frac{P_s}{P_s + P_f} \quad (20)$$

To have confidence in the integer outcomes of IA-estimation, a user would like to have P_{sf} close to 1. This can be achieved by setting the fail-rate P_f at a small-enough level. Thus the user chooses the level of fail-rate he finds acceptable and then determines the size of the aperture pull-in region that corresponds with this fail-rate level. With such a setting, the user has the guarantee that the fail-rate of his IA-estimator will never become unacceptably large.

As with I-estimation, the user can choose from a whole class of IA-estimators simply by using different shape-definitions for the aperture pull-in regions. Various examples have been given in Verhagen and Teunissen (2006). The ILS-estimator combined with the popular GNSS *Ratio-Test* (Leick 2004), is such an example. Unfortunately one can still find various incorrect interpretations of the Ratio-Test in the literature. So is its use often motivated by stating that it determines whether the ILS-solution is true or false. This is not correct. Also the current ways of choosing the tolerance value τ are ad hoc or based on false theoretical grounds. Often a fixed value of $\frac{1}{2}$ or $\frac{1}{3}$ is used. However, as shown in Verhagen and Teunissen (2006), Teunissen and Verhagen (2009), instead of using a fixed τ -value, one should use the *fixed fail-rate* approach. From the fixed fail-rate, one can then compute the variable τ -value (it varies with varying strength of the underlying GNSS model).

2.2.3 Optimal Integer Aperture Estimation

So far we considered IA-estimation with a priori chosen aperture pull-in shapes. Now we determine which of the IA-estimators performs best. As the optimal IA-estimator we choose the one which maximizes the success-rate subject to a given fail-rate. The optimal IA-estimator is given by the following theorem (Teunissen 2005).

Theorem 2 (Optimal integer aperture estimation)

Let $f_{\hat{a}}(x|a)$ and $f_{\delta}(x|a)$ be the PDFs of \hat{a} and $\delta = \hat{a} - \check{a}_{\text{IML}}$, respectively, and let P_s and P_f be the success-rate and the fail-rate of the IA-estimator. Then the solution to

$$\max_{\Omega_0 \subset \mathcal{P}_0} P_s \text{ subject to given } P_f \quad (21)$$

is given by the aperture pull-in region

$$\Omega_0 = \{x \in \mathcal{P}_0 \mid f_{\delta}(x|a) \leq \lambda f_{\hat{a}}(x+a|a)\} \quad (22)$$

where

$$\mathcal{P}_0 = \{x \in \mathbb{R}^n \mid 0 = \arg \max_{z \in \mathbb{Z}^n} f_{\hat{a}}(x|z)\}$$

and with the aperture parameter λ chosen so as to satisfy the a priori fixed fail-rate P_f .

The steps in computing the optimal IA-estimator are therefore as follows: (1) Compute the optimal I-estimator $\hat{a} = \arg \max_{z \in \mathbb{Z}^n} f_{\hat{a}}(\hat{a}|z)$; (2) Determine the aperture parameter λ from the user-defined fail-rate P_f . Ways of doing this are discussed in Verhagen (2005a,b) and Teunissen and Verhagen (2009); (3) Check whether $\delta = \hat{a} - \check{a}$ lies in Ω_0 . If $\delta \in \Omega_0$, then \check{a} is the outcome of the optimal IA-estimator, otherwise the outcome is \hat{a} .

2.3 Integer Equivariant Estimation

2.3.1 A Larger Class of Estimators

The class of IA-estimators includes the class of I-estimators. Now we introduce an even larger class. This larger class is obtained by dropping another condition of Definition 2. Since we would at least like to retain the integer remove–restore property, we keep the condition that the estimators must be z -translational invariant. Such estimators will be called *integer equivariant* (IE) estimators.

Definition 4 (Integer equivariant estimators)

The estimator $\hat{\theta}_{\text{IE}} = F_\theta(\hat{a})$, with $F_\theta : \mathbb{R}^n \mapsto \mathbb{R}$, is said to be an integer equivariant estimator of the linear function $\theta = l^T a$ if

$$F_\theta(x+z) = F_\theta(x) + l^T z, \quad \forall x \in \mathbb{R}^n, \quad z \in \mathbb{Z}^n \quad (23)$$

It will be clear that I-estimators and IA-estimators are also IE-estimators. The converse, however, is not necessarily true.

The class of IE-estimators is also larger than the class of linear unbiased estimators, assuming that the float solution is unbiased. Let $F_\theta^T \hat{a}$, for some $F_\theta \in \mathbb{R}^n$, be the linear estimator of $\theta = l^T a$. For it to be unbiased one needs, using $E\{\hat{a}\} = a$, that $F_\theta^T E\{\hat{a}\} = l^T a$, $\forall a \in \mathbb{R}^n$ holds true,

or that $F_\theta = I$. But this is equivalent to stating that $F_\theta^T(\hat{a} + a) = F_\theta^T\hat{a} + l^T a$, $\forall \hat{a} \in \mathbb{R}^n, a \in \mathbb{R}^n$. Comparison with (23) shows that the condition of linear unbiasedness is more restrictive than the condition of integer equivariance. The class of linear unbiased estimators is therefore a subset of the IE-class. This implies that a “best” IE-estimator must be at least as good as the BLUE \hat{a} . Afterall the float solution \hat{a} is an IE-estimator as well.

2.3.2 Best Integer Equivariant Estimation

We denote the best integer equivariant (BIE) estimator of θ as $\hat{\theta}_{\text{BIE}}$ and use the mean squared error (MSE) as our criterion of “best.” The BIE-estimator of $\theta = l^T a$ is therefore defined as

$$\hat{\theta}_{\text{BIE}} = \arg \min_{F_\theta \in \text{IE}} \text{E}\{(F_\theta(\hat{a}) - \theta)^2\} \quad (24)$$

in which IE stands for the class of IE-estimators. The minimization is thus taken over all integer equivariant functions that satisfy the condition of Definition 4. Thus the BIE-estimator is the optimal IE-estimator in the MSE-sense.

The reason for choosing the MSE-criterion is twofold. First, it is a well-known probabilistic criterion for measuring the closeness of an estimator to its target value, in our case $\theta = l^T a$. Second, the MSE-criterion is also often used as measure for the quality of the float solution itself. It should be kept in mind however, that the MSE-criterion is a weaker criterion than the probabilistic criterion used in the previous two sections.

The BIE-estimator is given by the following theorem (Teunissen 2003b).

Theorem 3 (Best integer equivariant estimation)

Let $f_{\hat{a}}(x|a)$ be the PDF of \hat{a} and let $\hat{\theta}_{\text{BIE}}$ be the best integer equivariant estimator of $\theta = l^T a$. Then $\hat{\theta}_{\text{BIE}} = l^T \hat{a}_{\text{BIE}}$, where

$$\hat{a}_{\text{BIE}} = \sum_{z \in \mathbb{Z}^n} z w_z(\hat{a}) \quad (25)$$

with the weighting functions $w_z(x)$ given as

$$w_z(x) = \frac{f_{\hat{a}}(x + a - z|a)}{\sum_{u \in \mathbb{Z}^n} f_{\hat{a}}(x + a - u|a)} \quad (26)$$

As the I-estimator, the BIE-estimator is also a weighted sum of all integer vectors in \mathbb{Z}^n . In the present case, however, the weights are not binary. They vary between 0 and 1, and their values are determined by the float solution and its PDF. As a consequence the BIE-estimator will be real-valued in general, instead of integer-valued.

An important consequence of the above theorem is that the BIE-estimator is always better than or at least as good as any integer estimator as well as any linear unbiased estimator. Afterall the class of integer estimators and the class of linear unbiased estimators are both subsets of the class of IE-estimators. The nonlinear BIE-estimator is therefore also better than the best linear unbiased estimator (BLUE),

$$\text{MSE}(\hat{\theta}_{\text{BIE}}) \leq \text{MSE}(\hat{\theta}_{\text{BLUE}}) \quad (27)$$

The BLUE is the minimum variance estimator of the class of linear unbiased estimators and it is given by the well-known Gauss–Markov theorem. The two estimators $\hat{\theta}_{\text{BIE}}$ and $\hat{\theta}_{\text{BLUE}}$ therefore both minimize the mean squared error, albeit within a different class.

The above theorem holds true for any PDF the float solution \hat{a} might have. In the Gaussian case $\hat{a} \sim N(a, Q_{\hat{a}\hat{a}})$, the weighting-function of Eq. 26 becomes

$$w_z(x) = \frac{\exp\left\{-\frac{1}{2} \|x - z\|_{Q_{\hat{a}\hat{a}}}^2\right\}}{\sum_{u \in \mathbb{Z}^n} \exp\left\{-\frac{1}{2} \|x - u\|_{Q_{\hat{a}\hat{a}}}^2\right\}} \quad (28)$$

Since the space of integers \mathbb{Z}^n can be seen as a certain discretized version of the space of real numbers \mathbb{R}^n , one would expect if the integer grid size gets smaller in relation to the size and extend of the PDF, that the difference between the two estimators, \hat{a}_{BIE} and \hat{a} , gets smaller as well. Similarly, if the PDF gets more peaked in relation to the integer grid size, one would expect that the BIE-estimator \hat{a}_{BIE} tends to an integer estimator. This is made precise in the following lemma.

Lemma 1 (Limits of the BIE-estimator)

(i) If we replace $\sum_{z \in \mathbb{Z}^n}$ by $\int_{\mathbb{R}^n} dz$ in Eqs. 25 and 28, then

$$\hat{a}_{BIE} = \hat{a}$$

(ii) Let $\hat{a} \sim N(a, Q_{\hat{a}\hat{a}} = \sigma^2 G)$. Then

$$\lim_{\sigma \rightarrow 0} \hat{a}_{BIE} = \check{a}_{ILS}$$

A probabilistic performance comparison between \hat{a} , \check{a}_{ILS} , and \hat{a}_{BIE} can be found in Verhagen and Teunissen (2005).

It is interesting to observe that the above given Gaussian-based expression for \hat{a}_{BIE} is identical to its Bayesian counterpart as given in Betti et al. (1993), Gundlich and Koch (2002), and Gundlich and Teunissen (2004). This is not quite true for the general case however. Still, this Gaussian equivalence nicely bridges the gap between the current theory of integer inference and the Bayesian approach.

3 Three Popular Integer Estimators

In this section, we discuss integer rounding, integer bootstrapping, and integer least-squares, with special attention to computational issues and the success-rates. We assume $\hat{a} \sim N(a \in \mathbb{Z}^n, Q_{\hat{a}\hat{a}})$.

3.1 Integer Rounding

3.1.1 Scalar and Vectorial Rounding

The simplest integer estimator is “rounding to the nearest integer.” In the scalar case, its pull-in regions (intervals) are given as

$$\mathcal{R}_z = \{x \in \mathbb{R} \mid |x - z| \leq 1/2\}, \quad z \in \mathbb{Z} \quad (29)$$

Any outcome of $\hat{a} \sim N(a \in \mathbb{Z}, \sigma_a^2)$, that satisfies $|\hat{a} - z| \leq 1/2$, will thus be pulled to the integer z . We denote the rounding estimator as \check{a}_R and the operation of integer rounding as $[.]$. Thus $\check{a}_R = [\hat{a}]$ and $\check{a}_R = z$ if $\hat{a} \in \mathcal{R}_z$.

The PMF of $\check{a}_R = \lceil \hat{a} \rceil$ is given as

$$P(\check{a}_R = z) = \Phi\left(\frac{1 - 2(a - z)}{2\sigma_{\hat{a}}}\right) + \Phi\left(\frac{1 + 2(a - z)}{2\sigma_{\hat{a}}}\right) - 1, \quad z \in \mathbb{Z} \quad (30)$$

where $\Phi(x)$ denotes the normal distribution function, $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}v^2\right\} dv$.

Note that the PMF is symmetric about a . Thus integer rounding provides for an unbiased integer estimator, $E(\check{a}_R) = a \in \mathbb{Z}$. Also note that the PMF becomes more peaked when $\sigma_{\hat{a}}$ gets smaller.

For GNSS ambiguity resolution, the success-rate is of particular importance. The success-rate of integer rounding follows from Eq. 30 by setting z equal to a ,

$$P(\check{a}_R = a) = 2\Phi\left(\frac{1}{2\sigma_{\hat{a}}}\right) - 1 \quad (31)$$

Thus the smaller the standard deviation, the larger the success-rate. A success-rate better than 0.99 requires $\sigma_{\hat{a}} \leq 0.15$ cycle.

Scalar rounding is easily generalized to the vectorial case. It is defined as the component-wise rounding of $\hat{a} = (\hat{a}_1, \dots, \hat{a}_n)^T$, $\check{a}_R = (\lceil \hat{a}_1 \rceil, \lceil \hat{a}_2 \rceil, \dots, \lceil \hat{a}_n \rceil)^T$. The pull-in regions of vectorial rounding are the multivariate versions of the scalar pull-in intervals,

$$\mathcal{R}_z = \left\{ x \in \mathbb{R}^n \mid \left| c_i^T (x - z) \right| \leq 1/2, \quad i = 1, \dots, n \right\}, \quad z \in \mathbb{Z}^n \quad (32)$$

where c_i denotes the unit vector having a 1 as its i th entry and 0s otherwise. Thus the pull-in regions of rounding are unit-squares in 2D, unit-cubes in 3D, etc.

3.1.2 Rounding Success-Rate

To determine the joint PMF of the components of \check{a}_R , we have to integrate the PDF of $\hat{a} \sim N(a, Q_{\hat{a}\hat{a}})$ over the pull-in regions \mathcal{R}_z . These n -fold integrals are unfortunately difficult to evaluate, unless the variance matrix $Q_{\hat{a}\hat{a}}$ is diagonal, in which case the components of \check{a}_R are independent and their joint PMF follows as the product of the univariate PMFs of the components. The corresponding success-rate is then given by the n -fold product of the univariate success-rates.

In case of GNSS, the ambiguity variance matrix will usually be fully populated, meaning that one will have to resort to methods of Monte Carlo simulation for computing the joint PMF. In case of the success-rate, one can alternatively make use of the following bounds (Teunissen 1998a).

Theorem 4 (Rounding success-rate bounds)

Let $\hat{a} \sim N(a \in \mathbb{Z}^n, Q_{\hat{a}\hat{a}})$. Then the rounding success-rate can be bounded from below and from above as

$$\prod_{i=1}^n \left[2\Phi\left(\frac{1}{2\sigma_{\hat{a}_i}}\right) - 1 \right] \leq P(\check{a}_R = a) \leq \left[2\Phi\left(\frac{1}{2\sigma_{\max}}\right) - 1 \right] \quad (33)$$

where $\sigma_{\max} = \max_{i=1, \dots, n} \sigma_{\hat{a}_i}$.

These easy-to-compute bounds are very useful for determining the expected success of GNSS ambiguity rounding. The upper bound is useful to quickly decide against such ambiguity

resolution. It shows that ambiguity resolution based on vectorial rounding can not be expected successful, if already one of the scalar rounding success-rates is too low.

The lower bound is useful to quickly decide in favor of vectorial rounding. If the lower bound is sufficiently close to 1, one can be confident that vectorial rounding will produce the correct integer ambiguity vector. Note that this requires each of the individual probabilities in the product of the lower bound to be sufficiently close to 1.

3.1.3 Z-Transformations

Although \check{a}_R is easy to compute, the rounding estimator suffers from a lack of invariance against integer reparametrizations or the so-called *Z-transformations*. A matrix is called a *Z-transformation* if it is one-to-one (i.e., invertible) and integer (Teunissen 1995a). Such transformations leave the integer nature of the parameters in tact.

By saying that the rounding estimator lacks *Z*-invariance, we mean that if the float solution is *Z*-transformed, the integer solution does not transform accordingly. That is, rounding and transforming do not commute,

$$\check{z}_R \neq Z\check{a}_R \text{ if } \hat{z} = Z\hat{a} \quad (34)$$

Only in case *Z* is a permutation matrix, $Z = \Pi$, do we have $\check{z}_R = \Pi\check{a}_R$. In this case, the transformation is a simple reordering of the ambiguities.

Also the success-rate lacks *Z*-invariance. Since the pull-in regions of rounding remain unaffected by the *Z*-transformation, while the distribution of the float solution changes to $\hat{z} \sim N(z = Za, Q_{zz} = ZQ_{\hat{a}\hat{a}}Z^T)$, we have, in general,

$$P(\check{z}_R = z) \neq P(\check{a}_R = a) \quad (35)$$

This lack of invariance implies that integer rounding is not optimal in the vectorial case. The lack of invariance does not occur in the scalar case, since multiplication by ± 1 is then the only admissible *Z*-transformation.

Does the mentioned lack of invariance mean that rounding is unfit for GNSS integer ambiguity resolution? No, by no means. Integer rounding is a valid ambiguity estimator, since it obeys the principle of integer equivariance, and it is an attractive estimator, because of its computational simplicity. Whether or not it can be successfully applied in any concrete situation, depends solely on the value of its success-rate for that particular situation. What the lack of invariance shows, is the nonoptimality of rounding. Despite being nonoptimal, rounding can achieve high success-rates, provided the underlying GNSS model is of sufficient strength and provided the proper ambiguity parametrization is chosen. In [Sect. 3.3.2](#), we come back to this issue and see how we can use the existing degrees of freedom of integer parametrization to our advantage.

3.2 Integer Bootstrapping

3.2.1 The Bootstrapping Principle

Integer bootstrapping is a generalization of integer rounding; it combines integer rounding with sequential conditional least-squares estimation and as such takes some of the correlation between the components of the float solution into account. The method goes as follows.

If $\hat{a} = (\hat{a}_1, \dots, \hat{a}_n)^T$, one starts with \hat{a}_1 and as before rounds its value to the nearest integer. Having obtained the integer of the first component, the real-valued estimates of all remaining components are then corrected by virtue of their correlation with \hat{a}_1 . Then the second, but now corrected, real-valued component is rounded to its nearest integer. Having obtained the integer value of this second component, the real-valued estimates of all remaining $n - 2$ components are then again corrected by virtue of their correlation with the second component. This process is continued until all n components are taken care of. We have the following definition.

Definition 5 (Integer bootstrapping)

Let $\hat{a} = (\hat{a}_1, \dots, \hat{a}_n)^T \in \mathbb{R}^n$ be the float solution and let $\check{a}_B = (\check{a}_{B,1}, \dots, \check{a}_{B,n})^T \in \mathbb{Z}^n$ denote the corresponding integer bootstrapped solution. Then

$$\begin{aligned}\check{a}_{B,1} &= \lceil \hat{a}_1 \rceil \\ \check{a}_{B,2} &= \lceil \hat{a}_{2|1} \rceil = \lceil \hat{a}_2 - \sigma_{21}\sigma_1^{-2}(\hat{a}_1 - \check{a}_{B,1}) \rceil \\ &\vdots \\ \check{a}_{B,n} &= \lceil \hat{a}_{n|N} \rceil = \left\lceil \hat{a}_n - \sum_{j=1}^{n-1} \sigma_{n,j|J}\sigma_{j|J}^{-2}(\hat{a}_{j|J} - \check{a}_{B,j}) \right\rceil\end{aligned}\tag{36}$$

where $\hat{a}_{i|I}$ is the least-squares estimator of a_i conditioned on the values of the previous $I = \{1, \dots, (i-1)\}$ sequentially rounded components, $\sigma_{i,j|J}$ is the covariance between \hat{a}_i and $\hat{a}_{j|J}$, and $\sigma_{j|J}^2$ is the variance of $\hat{a}_{j|J}$. For $i = 1$, $\hat{a}_{i|I} = \hat{a}_1$.

As the definition shows, the bootstrapped estimator can be seen as a generalization of integer rounding. The bootstrapped estimator reduces to integer rounding in case correlations are absent, i.e., in case the variance matrix $Q_{\hat{a}\hat{a}}$ is diagonal.

The bootstrapped estimator combines sequential conditional least-squares estimation with integer rounding. If we replace the “least-squares estimation” part by “linear estimation,” we can construct a whole class of sequential integer estimators. This class is defined as follows.

Definition 6 (Sequential integer estimation)

Let $\hat{a} = (\hat{a}_1, \dots, \hat{a}_n)^T \in \mathbb{R}^n$ be the float solution. Then $\check{a} = (\check{a}_1, \dots, \check{a}_n)^T \in \mathbb{Z}^n$ is a sequential integer estimator of $E(\hat{a}) = a \in \mathbb{Z}^n$ if $\check{a}_i = \lceil \hat{a}_i + \sum_{j=1}^{i-1} r_{ij}(\hat{a}_j - \check{a}_j) \rceil$, $i = 1, \dots, n$, or, in vector-matrix form, if

$$\check{a} = \lceil \hat{a} + (R - I_n)(\hat{a} - \check{a}) \rceil\tag{37}$$

with R a unit lower triangular matrix.

By showing how the bootstrapped estimator can be computed from using the triangular factorization of the variance matrix $Q_{\hat{a}\hat{a}}$, it becomes immediately clear that the bootstrapped estimator \check{a}_B is indeed a member of the class of sequential integer estimators. We have the following result (Teunissen 2007a).

Theorem 5 (Bootstrapping and the triangular decomposition)

Let $\hat{a} \in \mathbb{R}^n$ be the float solution and let the unit lower triangular decomposition of its variance matrix be given as $Q_{\hat{a}\hat{a}} = LDL^T$. The entries of L and D are then given as

$$(L)_{ij} = \begin{cases} 0 & \text{for } 1 \leq i < j \leq n \\ 1 & \text{for } i = j \\ \sigma_{i,j|J}\sigma_{j|J}^{-2} & \text{for } 1 \leq j < i \leq n \end{cases} \quad \text{and } D = \text{diag}(\dots, \sigma_{j|J}^2, \dots)\tag{38}$$

and the bootstrapped estimator $\check{a}_B \in \mathbb{Z}^n$ of Eq. 36 can be expressed as

$$\check{a}_B = \lceil \hat{a} + (L^{-1} - I_n)(\hat{a} - \check{a}_B) \rceil \quad (39)$$

Comparing Eq. 39 with Eq. 37 shows that the bootstrapped estimator is indeed a sequential integer estimator.

Note that in the construction of the bootstrapped estimator, the triangular factor L of $Q_{\hat{a}\hat{a}} = LDL^T$ is used, but not the diagonal matrix D . Thus bootstrapping takes only part of the information of the variance matrix into account. Although the diagonal matrix D is not used in the bootstrapped mapping itself, it determines—as we will see below—the bootstrapped success-rate.

3.2.2 The Bootstrapped PMF and Success-Rate

To determine the bootstrapped PMF, we first need to determine the bootstrapped pull-in regions. The bootstrapped pull-in regions are given as

$$\mathcal{B}_z = \left\{ x \in \mathbb{R}^n \mid \left| c_i^T L^{-1}(x - z) \right| \leq 1/2, i = 1, \dots, n \right\}, \quad \forall z \in \mathbb{Z}^n \quad (40)$$

where c_i denotes the unit vector having a 1 as its i th entry and 0s otherwise.

To prove Eq. 40, let $\tilde{a} = z + L^{-1}(\hat{a} - z) = \hat{a} + (L^{-1} - I_n)(\hat{a} - z)$. Then according to Eq. 39, $\check{a}_B = z$ if $\lceil \tilde{a} \rceil = z$. Thus $\check{a}_B = z$ if $\lceil \tilde{a} - z \rceil = 0$, which implies that all components of $L^{-1}(\hat{a} - z)$ have to be less than or equal to $\frac{1}{2}$ in absolute value. It is easily verified that the \mathcal{B}_z satisfy the conditions of Definition 2. In Xu (2006), however, it is incorrectly claimed that Teunissen's bootstrapped estimator Eq. 36 is not a genuine integer estimator.

The bootstrapped PMF follows from integrating the multivariate normal distribution over the bootstrapped pull-in regions. In contrast to the multivariate integral for integer rounding, the multivariate integral for bootstrapping can be simplified considerably. As shown by the following theorem, the bootstrapped PMF can be expressed as a product of univariate integrals (Teunissen 1998a).

Theorem 6 (Bootstrapped PMF)

Let $\hat{a} \sim N(a \in \mathbb{Z}^n, Q_{\hat{a}\hat{a}})$ and let \check{a}_B be the bootstrapped estimator of a . Then

$$P(\check{a}_B = z) = \prod_{i=1}^n \left[\Phi \left(\frac{1 - 2l_i^T(a - z)}{2\sigma_{\hat{a}_{i|I}}} \right) + \Phi \left(\frac{1 + 2l_i^T(a - z)}{2\sigma_{\hat{a}_{i|I}}} \right) - 1 \right], \quad \forall z \in \mathbb{Z}^n \quad (41)$$

where l_i is the i th column vector of the unit upper triangular matrix L^{-T} .

The bootstrapped PMF is symmetric about the mean of \hat{a} . This implies that the bootstrapped estimator \check{a}_B is an unbiased estimator of $a \in \mathbb{Z}^n$.

As a direct consequence of the above theorem, we have an exact and easy-to-compute expression for the bootstrapped success-rate.

Corollary 1 (Bootstrapped success-rate)

Let $\hat{a} \sim N(a \in \mathbb{Z}^n, Q_{\hat{a}\hat{a}})$. Then the bootstrapped success-rate is given as

$$P(\check{a}_B = a) = \prod_{i=1}^n \left[2\Phi \left(\frac{1}{2\sigma_{\hat{a}_{i|I}}} \right) - 1 \right] \quad (42)$$

This is an important result for GNSS applications as it provides a very simple way of evaluating the bootstrapped success-rate.

The following result shows what happens to the success-rate if the computed bootstrapped estimator is based on a too optimistic or a too pessimistic description of the float precision (Teunissen 2007a).

Theorem 7 (Use of wrong weight matrix)

Let $\hat{a} \sim N(a \in \mathbb{Z}^n, Q)$ and let \check{a}_B^Σ be the bootstrapped estimator constructed on the basis of the positive definite matrix Σ . Then

$$P(\check{a}_B^\Sigma = a) \leq P(\check{a}_B^Q = a) \quad (43)$$

with strict inequality if the unit triangular factors of Σ and Q differ.

The success-rate will thus get smaller if a wrong weight matrix is used. This theorem has two important consequences. First, since it was shown that the bootstrapped estimator is a member of the class of sequential integer estimators, the above result directly implies that the bootstrapped estimator is the optimal estimator within this restricted class.

Second, since integer rounding is also a sequential integer estimator (i.e., take $R = I_n$ in Eq. 37), it follows that the rounding success-rate will never be larger than the bootstrapped success-rate. Thus bootstrapping is a better integer estimator than rounding.

3.2.3 Z-Transformations

Like rounding, bootstrapping also suffers from a lack of Z -invariance. Similarly to Eqs. 34 and 35, we have for bootstrapping,

$$\check{z}_B \neq Z\check{a}_B \text{ and } P(\check{z}_B = z) \neq P(\check{a}_B = a) \text{ if } \hat{z} = Z\hat{a} \quad (44)$$

Thus the success-rate of bootstrapping (and of rounding as well) changes if a different parametrization is used, say $z = Za \in \mathbb{Z}^n$ instead of $a \in \mathbb{Z}^n$. It is thus of importance, when using bootstrapping (or rounding), that a proper parametrization is used. Bootstrapping performs relatively poor, for instance, when applied to the GNSS DD ambiguities for RTK positioning. This is due to the usually high correlation between such DD ambiguities. Bootstrapping should therefore be used in combination with a decorrelating Z -transformation. For computational details on how such Z -transformations can be constructed, we refer to Teunissen (1993, 1995b), de Jonge and Tiberius (1996a), de Jonge et al. (1996, and the references cited therein). Also see Liu et al. (1999), Grafarend (2000), Xu (2001), Joosten and Tiberius (2002), Svendsen (2006).

Despite the fact that we have an exact and easy-to-compute formula for the bootstrapped success-rate (cf. Eq. 42), an easy-to-compute upper bound of it would still be useful if it would be Z -invariant. Such an upper bound can be constructed when use is made of the Z -invariant *Ambiguity Dilution of Precision* (ADOP) (Teunissen 2000).

Theorem 8 (Bootstrapped success-rate invariant upper bound)

Let $\hat{a} \sim N(a \in \mathbb{Z}^n, Q_{\hat{a}\hat{a}})$, $\hat{z} = Z\hat{a}$, and $ADOP = \det(Q_{\hat{a}\hat{a}})^{\frac{1}{2n}}$. Then

$$P(\check{z}_B = z) \leq \left[2\Phi\left(\frac{1}{2ADOP}\right) - 1 \right]^n \quad (45)$$

for any admissible Z -transformation.

Thus if the upper bound is too small, we can immediately conclude, for *any* parametrization, that bootstrapping nor rounding will be successful.

3.3 Integer Least-Squares

3.3.1 Mixed Integer Least-Squares

Application of the least-squares principle to the mixed integer model (3) gives the minimization problem

$$\min_{a,b} \|y - Aa - Bb\|_{Q_{yy}}^2, \quad a \in \mathbb{Z}^n, \quad b \in \mathbb{R}^P \quad (46)$$

where Q_{yy} is the variance matrix of y . This type of least-squares problem was first formulated in Teunissen (1993) and was coined a (mixed) integer least-squares (ILS) problem. It is a nonstandard least-squares problem due to the integer constraints $a \in \mathbb{Z}^n$.

To solve Eq. 46, we start from the orthogonal decomposition

$$\|y - Aa - Bb\|_{Q_{yy}}^2 = \|\hat{e}\|_{Q_{yy}}^2 + \|\hat{a} - a\|_{Q_{\hat{a}\hat{a}}}^2 + \|\hat{b}(a) - b\|_{Q_{b(a)b(a)}}^2 \quad (47)$$

where $\hat{e} = y - A\hat{a} - B\hat{b}$, with \hat{a} and \hat{b} the unconstrained least-squares estimators of a and b , respectively, $\hat{b}(a) = \hat{b} - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{-1}(\hat{a} - a)$, and $Q_{b(a)b(a)} = Q_{\hat{b}\hat{b}} - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{-1}Q_{\hat{a}\hat{b}}$. Note that the first term on the right-hand side of Eq. 47 is constant and that the third term can be made zero for any a by setting $b = \hat{b}(a)$. Hence, the mixed integer minimizers of Eq. 46 are given as

$$\check{a}_{\text{ILS}} = \arg \min_{z \in \mathbb{Z}^n} \|\hat{a} - z\|_{Q_{\hat{a}\hat{a}}}^2 \quad \text{and} \quad \check{b} = \hat{b}(\check{a}_{\text{ILS}}) = \hat{b} - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{-1}(\hat{a} - \check{a}_{\text{ILS}}) \quad (48)$$

In contrast to rounding and bootstrapping, the ILS-principle is Z -invariant. We have

$$\check{z}_{\text{ILS}} = Z\check{a}_{\text{ILS}} \quad \text{and} \quad \check{b} = \hat{b} - Q_{\hat{b}\hat{z}}Q_{\hat{z}\hat{z}}^{-1}(\hat{z} - \check{z}_{\text{ILS}}) \quad (49)$$

Application of the ILS-principle to $Z\hat{a}$ gives therefore the same result as Z times the result of applying the ILS-principle to \hat{a} . Also \check{b} is invariant for the integer reparametrization.

We have seen that the 2D pull-in regions of rounding and bootstrapping are squares and parallelograms, respectively. It follows that those of ILS are hexagons. The ILS pull-in region of $z \in \mathbb{Z}^n$ consists by definition of all those points that are closer to z than to any other integer vector in \mathbb{R}^n , $\mathcal{L}_z = \{x \in \mathbb{R}^n \mid \|x - z\|_{Q_{\hat{a}\hat{a}}}^2 \leq \|x - u\|_{Q_{\hat{a}\hat{a}}}^2, \forall u \in \mathbb{Z}^n\}, z \in \mathbb{Z}^n$. By rewriting the inequality, we obtain a representation that more closely resembles the ones of rounding, \mathcal{R}_z , and bootstrapping, \mathcal{B}_z , (see Eqs. 32 and 40),

$$\mathcal{L}_z = \left\{ x \in \mathbb{R}^n \mid \left| u^T Q_{\hat{a}\hat{a}}^{-1}(x - z) \right| \leq 1/2 \|u\|_{Q_{\hat{a}\hat{a}}}^2, \forall u \in \mathbb{Z}^n \right\}, \quad z \in \mathbb{Z}^n \quad (50)$$

This shows that the ILS pull-in regions are constructed from intersecting half-spaces. One can show that at most $2^n - 1$ pairs of such half spaces are needed for constructing the pull-in region. It is easily verified that the ILS pull-in regions are convex, symmetric sets that satisfy the conditions of Definition 2. Note that $\mathcal{L}_z = \mathcal{R}_z$ when $Q_{\hat{a}\hat{a}}$ is diagonal.

3.3.2 The ILS Search

In contrast to rounding and bootstrapping, an integer search is needed to compute an ILS solution. The search space is defined as

$$\Psi_a = \{a \in \mathbb{Z}^n \mid \| \hat{a} - a \|_{Q_{\hat{a}\hat{a}}}^2 \leq \chi^2\} \quad (51)$$

where χ^2 is a to be chosen positive constant. This ellipsoidal search space is centred at \hat{a} , its elongation is governed by $Q_{\hat{a}\hat{a}}$ and its size is determined by χ^2 . In case of GNSS, the search space is usually extremely elongated due to the high correlations between the carrier phase ambiguities. Since this extreme elongation hinders the computational efficiency of the search, the search space is first transformed to a more spherical shape by means of a decorrelating Z-transformation,

$$\Psi_z = \{z \in \mathbb{Z}^n \mid \| \hat{z} - z \|_{Q_{\hat{z}\hat{z}}}^2 \leq \chi^2\} \quad (52)$$

where $\hat{z} = Z\hat{a}$ and $Q_{\hat{z}\hat{z}} = ZQ_{\hat{a}\hat{a}}Z^T$.

In order for the search to be efficient, one would like the search space to be small such that it contains not too many integer vectors. This requires the choice of a small value for χ^2 , but one that still guarantees that the search space contains at least one integer vector. After all, Ψ_z has to be nonempty to guaranty that it contains \check{z}_{ILS} . Since the easy-to-compute (decorrelated) bootstrapped estimator gives a good approximation to the ILS estimator (cf. Theorem 9), \check{z}_B is an excellent candidate for setting the size of the search space,

$$\chi^2 = \| \hat{z} - \check{z}_B \|_{Q_{\hat{z}\hat{z}}}^2 \quad (53)$$

In this way one can work with a very small search space and still guarantee that the sought for ILS solution is contained in it. If the rounding success-rate is sufficiently high, one may also use \check{z}_R instead of \check{z}_B .

For the actual search, the quadratic form $\| \hat{z} - z \|_{Q_{\hat{z}\hat{z}}}^2$ is first written as a sum-of-squares. This is achieved by using the triangular decomposition $Q_{\hat{z}\hat{z}} = LDL^T$ (cf. Theorem 5),

$$\sum_{i=1}^n \frac{(\hat{z}_{i|I} - z_i)^2}{\sigma_{i|I}^2} \leq \chi^2 \quad (54)$$

This sum-of-squares structure can now be used to set up the n intervals that are used for the search. These sequential intervals are given as

$$\begin{aligned} (\hat{z}_1 - z_1)^2 &\leq \sigma_1^2 \chi^2 \\ (\hat{z}_{2|1} - z_2)^2 &\leq \sigma_{2|1}^2 \left(\chi^2 - \frac{(\hat{z}_1 - z_1)^2}{\sigma_1^2} \right) \\ &\vdots \\ (\hat{z}_{n|(n-1),\dots,1} - z_n)^2 &\leq \sigma_{n|(n-1),\dots,1}^2 \left(\chi^2 - \sum_{i=1}^{n-1} \frac{(\hat{z}_{i|I} - z_i)^2}{\sigma_{i|I}^2} \right) \end{aligned} \quad (55)$$

To search for all integer vectors that are contained in Ψ_z , one can now proceed as follows. First, collect all integers z_1 that are contained in the first interval. Then for each of these integers, one computes the corresponding length and center point of the second interval, followed by collecting all integers z_2 that lie inside this second interval. By proceeding in this way to the last interval, one finally ends up with the set of integer vectors that lie inside Ψ_z . From this set one

then picks the ILS solution as the integer vector that returns the smallest value for $\|\hat{z} - z\|_{Q_{zz}}^2$. Various refinements on this search are possible (see, e.g., Teunissen 1995b; de Jonge and Tiberius 1996a; de Jonge et al. 1996; de Jonge 1998; Chang et al. 2005).

To understand why the decorrelating Z -transformation is necessary to improve the efficiency of the search, consider the structure of the above given sequential intervals and assume that they are formulated for the DD ambiguities of a single-baseline GNSS-RTK model. The DD ambiguity sequential conditional variances will then show a large discontinuity when going from the third to the fourth ambiguity. The RTK DD ambiguities are namely poorly estimable, i.e., have large variances, unless already three of them are assumed known, since with three DD ambiguities known, the baseline and remaining ambiguities can be estimated with a very high precision. This discontinuity of the DD ambiguity sequential conditional variances implies that σ_1^2 , $\sigma_{2|1}^2$, and $\sigma_{3|2,1}^2$ are large, while the remaining variances $\sigma_{i|1}^2$, $i = 4, \dots, n$ are very small. Thus the first three bounds of Eq. 55 are rather loose, while those of the remaining ($n - 3$) inequalities are very tight. As a consequence one will experience “search halting.” Of many of the collected integer candidates that satisfy the first three inequalities of Eq. 55, one will not be able to find corresponding integers that satisfy the remaining inequalities. This inefficiency in the search is eliminated when using the Z -transformed ambiguities instead of the DD ambiguities. The decorrelating Z -transformation eliminates the discontinuity and, by virtue of the fact that the product of the sequential variances remains invariant (volume is preserved), also reduces the large values of the first three conditional variances. For the construction of decorrelating Z -transformations, see Teunissen (1995b) and the references cited therein.

The ILS procedure is mechanized in the GNSS LAMBDA (Least-squares AMBiguity Decorrelation Adjustment) method, which is currently one of the most applied methods for GNSS carrier phase ambiguity resolution. For more information on the LAMBDA method, we refer to Teunissen (1993, 1995b), de Jonge and Tiberius (1996a), de Jonge et al. (1996), and Chang et al. (2005) or to the textbooks Strang and Borre (1997), Teunissen and Kleusberg (1998), Leick (2004), Misra and Enge (2006), and Hofmann-Wellenhof et al. (2008).

The LAMBDA method is in use for a variety of different applications. Examples of such applications are baseline- and network positioning (Tiberius and de Jonge 1995; de Jonge and Tiberius 1996b; Boon and Ambrosius 1997; Odijk 2002; Teunissen et al. 1997; Tsai et al. 2007), satellite formation flying (Cox and Brading 1999; Wu and Bar-Sever 2006, Buist et al. 2008), InSAR and VLBI (Kampes and Hanssen 2004; Hobiger et al. 2009), GNSS attitude determination (Park and Teunissen 2003; Dai et al. 2004; Monikes et al. 2005; Giorgi et al. 2008), and next generation GNSS (Eissfeller et al. 2002; Wu et al. 2004; Ji et al. 2007).

3.3.3 The Least-Squares PMF and Success-Rate

The ILS PMF is given as

$$P(\check{a}_{\text{ILS}} = z) = \int_{\mathcal{L}_z} f_a(x|a) dx \quad (56)$$

To obtain the ILS success-rate, set $z = a$. Due to the complicated geometry of the ILS pull-in regions, methods of Monte Carlo simulation are needed to evaluate the multivariate integral Eq. 56. For $\hat{a} \sim N(a, Q_{\hat{a}\hat{a}})$, the ILS success-rate is given by the integral

$$P(\check{a}_{\text{ILS}} = a) = \int_{\mathcal{L}_0} C \exp \left\{ -\frac{1}{2} \|x\|_{Q_{\hat{a}\hat{a}}}^2 \right\} dx \quad (57)$$

As a consequence of Theorem 1, this success-rate is the largest of all integer estimators.

Note that a is not needed for the computation of the success-rate. Thus one may simulate as if \hat{a} has a zero mean. Also note that the ILS success-rate is Z -invariant, $P(\check{z}_{\text{ILS}} = Z\hat{a}) = P(\check{a}_{\text{ILS}} = a)$. This property can be used to ones advantage when simulating. Since the simulation requires the repeated computation of an ILS solution, one is much better of doing this for a largely decorrelated $\check{z} = Z\hat{a}$, than for the original \hat{a} .

The first step is to use a random generator to generate n independent samples from the univariate standard normal distribution $N(0, 1)$, and then collect these in an n -vector s . This vector is transformed as Gs , with G equal to the Cholesky factor of $Q_{\check{z}\check{z}} = GG^T$. The result is a sample Gs from $N(0, Q_{\check{z}\check{z}})$, and this sample is used as input for the ILS estimator. If the output of this estimator equals the null vector, then it is correct, otherwise it is incorrect. This simulation process can be repeated N number of times, and one can count how many times the null vector is obtained as a solution, say N_s times, and how often the outcome equals a nonzero integer vector, say N_f times. The approximations of the success-rate and fail-rate follow then as,

$$P_s \approx \frac{N_s}{N} \quad \text{and} \quad P_f \approx \frac{N_f}{N} \quad (58)$$

Instead of using simulation, one may also consider using bounds on the success-rate. The following theorem gives sharp lower- and upper bounds on the ILS success-rate.

Theorem 9 (ILS success-rate bounds)

Let $\hat{a} \sim N(a \in \mathbb{Z}^n, Q_{\hat{a}\hat{a}})$, $\check{z} = Z\hat{a}$, and $c_n = (\frac{n}{2}\Gamma(\frac{n}{2}))^{2/n}/\pi$, with $\Gamma(x)$ the gamma function. Then

$$P(\check{z}_B = z) \leq P(\check{a}_{\text{ILS}} = a) \leq P\left(\chi^2(n, 0) \leq \frac{c_n}{\text{ADOP}^2}\right) \quad (59)$$

for any admissible Z -transformation.

The lower bound is due to Teunissen (1998, 1999b). The upper bound was first given in Hassibi and Boyd (1998), albeit without proof. A proof is given in Teunissen (2000). The above lower bound (after decorrelation) is currently the sharpest lower bound available for the ILS success-rate. A study on the performances of the various bounds can be found in Thomsen (2000) and Verhagen (2003, 2005a,b).

4 Baseline Quality and Model Validation

4.1 Fixed and Float Baseline

The estimation of $a \in \mathbb{Z}^n$ is usually not a goal in itself. The integer constraints are usually included so as to get a better estimator for the baseline, i.e., the real-valued parameter vector $b \in \mathbb{R}^p$. To study the impact of the integer constraints, we need the distribution of the fixed baseline \check{b} . It is given by the following theorem (Teunissen 1999a).

Theorem 10 (Fixed baseline PDF)

Let \hat{a}, \check{b} be distributed as in Eq. 4 and let \check{a} be any integer estimator of a . Then the PDF of the fixed baseline estimator, $\check{b} = \hat{b} - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{-1}(\hat{a} - \check{a})$, is given as

$$f_{\check{b}}(x) = \sum_{z \in \mathbb{Z}^n} f_{\hat{b}(z)}(x)P(\check{a} = z) \quad (60)$$

where $f_{\hat{b}(z)}(x)$ is the PDF of $\hat{b}(z) \sim N(b - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{-1}(a - z), Q_{\hat{b}\hat{b}} - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{-1}Q_{\hat{a}\hat{b}})$ and $P(\check{a} = z)$ is the PMF of \check{a} .

This result shows that the fixed baseline distribution is *multi modal*, it equals an infinite sum of weighted conditional baseline distributions. These conditional baseline distributions are shifted versions of one another. The size and direction of the shift are governed by $Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{-1}z$, $z \in \mathbb{Z}^n$. Each of the conditional baseline distributions in the sum is downweighted as z gets further apart from a . The weights are the masses of \check{a} 's PMF.

Knowing the fixed baseline PDF allows one to study its quality by means of the probability that \check{b} lies in a certain region $C_b \subset \mathbb{R}^p$. In general it will be difficult to evaluate such probability exactly. For practical purposes it is therefore of importance to have bounds available for the probability $P(\check{b} \in C_b)$. We will assume C_b to be convex and symmetric about $E(\hat{b}) = E(\check{b}) = b$. In that case the integral of $f_{\hat{b}(z)}(x)$ over C_b reaches its maximum for $z = a$. This shows that $P(\hat{b}(z = a) \in C_b)$ can be taken as upper bound. A lowerbound is also easily found. Since the entries in the sum of (60) are all nonnegative, any finite sum of nonzero entries can be used to obtain a lower bound. The more nonzero entries are used in this finite sum, the sharper this lower bound becomes. As a result we have obtained the following corollary.

Corollary 2 (Fixed baseline coverage probability)

Let $C_b \subset \mathbb{R}^p$ be any convex set symmetric about b . Then the fixed baseline coverage probability $P(\check{b} \in C_b)$ can be bounded from above and below as

$$P(\hat{b}(a) \in C_b)P(\check{a} = a) \leq P(\check{b} \in C_b) \leq P(\hat{b}(a) \in C_b) \quad (61)$$

Note that these bounds become tight when the success-rate approaches 1. This shows, although the probability of the conditional estimator always overestimates the probability of the fixed baseline estimator, that the two probabilities are close for large success-rates. In other words, the (unimodal) distribution of the conditional estimator is a good approximation to the (multimodal) distribution of the fixed baseline estimator, when the success-rate is sufficiently close to 1.

The tolerance value to be chosen for $P(\check{a} = a)$ depends on the error one is willing to accept. If one accepts a *relative error* of $0 \leq [P(\hat{b}(a) \in C_b) - P(\check{b} \in C_b)]/P(\check{b} \in C_b) \leq 10^{-\epsilon}$, then a success-rate of $P(\check{a} = a) = \frac{1}{1+10^{-\epsilon}}$ is required.

4.2 Cross-Validation of Mixed Integer Model

Without the integer constraints $a \in \mathbb{Z}^n$, the mixed integer model (3) reduces to a standard linear model. For such models, validation can be executed using the standard theory of hypothesis testing. In our case, a is integer and we may consider the following null- and alternative hypothesis

$$\begin{aligned} H_0 : y &\sim N(Aa + Bb, Q_{yy}), \quad a \in \mathbb{Z}^n, \quad b \in \mathbb{R}^p \\ H_a : y &\sim N(Aa + Bb + Cc, Q_{yy}), \quad a \in \mathbb{Z}^n, \quad (b^T, c^T)^T \in \mathbb{R}^p \times \mathbb{R}^q \end{aligned} \quad (62)$$

where the additional term Cc , with matrix C known and vector c unknown, models under the alternative hypothesis the supposed modeling errors, e.g., outliers, slips, instrumental biases, etc. (Teunissen and Kleusberg 1998, Chaps. 5 and 8).

Knowing that a is integer, strengthens the model and allows one to reevaluate the model. This is the situation one will have when testing the validity of a phase-based GNSS model with integer ambiguities. Instead of validating it as a linear model, one should think of validating it as a mixed integer linear model, as was emphasized in Teunissen and Verhagen (2008). Many, however, still apply the standard theory of hypothesis testing to Eq. 62. Stating that a is known to be integer, is however not the same as stating that a is a known integer. In order to be able to validate the mixed integer model, one must be able to answer questions like (1) What are the appropriate test statistics for testing H_0 against H_a ? (2) How are these test statistics distributed under H_0 against H_a ? (3) What are the appropriate acceptance and rejection regions?

Here we present a solution based on the idea of cross-validation. First note that the observation equations of Eq. 62 can always be one-to-one transformed to a set in which only a part of the transformed observables is used under H_a . For this transformed case, the C -matrix of H_a will be of the form $C = (I_q, 0)^T$, which implies that the corresponding y_0 of $y = (y_0^T, y_1^T)^T$ will not contribute to the parameter solution. We may therefore use the solution of H_a to predict the extra observable(s) of H_0 , namely y_0 . An unlikely prediction error is then reason for rejecting H_0 on the basis of H_a .

To keep things simple, we assume in the following that the necessary one-to-one transformation has been applied already. This allows us to work with $y = (y_0^T, y_1^T)^T$ and use y_1 to predict y_0 . Let $F = (F_0^T, F_1^T)^T = (A, B)$, $x = (a^T, b^T)^T$ and let $\hat{x}_{(1)}$ be the float estimator of x based on y_1 . Then the float prediction error is given as $\hat{e}_0 = y_0 - \hat{y}_0$, with the predictor $\hat{y}_0 = F_0\hat{x}_{(1)} + Q_{y_0 y_1} Q_{y_1 y_1}^{-1}(y_1 - F_1\hat{x}_{(1)})$ (see, e.g., Grafarend 1976). This prediction error is distributed under H_0 as $\hat{e}_0 \sim N(0, Q_{\hat{e}_0 \hat{e}_0})$. To test H_0 , one can check whether or not the outcome of \hat{e}_0 lies in the tails of $N(0, Q_{\hat{e}_0 \hat{e}_0})$. If it does, H_0 is rejected on the basis of H_a . This is how the tests are executed for the standard linear model. A typical example is Baarda's *data snooping* (1968) for outlier detection, where each individual observation plays on its turn the role of y_0 .

For our mixed integer model, we should not use the float prediction error $\hat{e}_0 = y_0 - \hat{y}_0$, but instead the fixed prediction error $\check{e}_0 = y_0 - \check{y}_0$, where the predictor \check{y}_0 has taken the integerness of a into account (Teunissen 2007b).

Theorem 11 (Fixed prediction error PDF)

Let $y = (y_0^T, y_1^T)^T$ and let $\hat{e}_0 = y_0 - \hat{y}_0$ be the float prediction error. Then for any integer estimator \check{a} , the fixed prediction error is given as $\check{e}_0 = \hat{e}_0 - Q_{\hat{e}_0 \check{a}} Q_{\check{a} \check{a}}^{-1}(\check{a} - \check{a})$ and its PDF under H_0 as

$$f_{\check{e}_0}(x) = \sum_{z \in \mathbb{Z}^n} f_{\hat{e}_0(z)}(x) P(\check{a} = z) \quad (63)$$

where $f_{\hat{e}_0(z)}(x)$ is the PDF of $\hat{e}_0(z) \sim N(-Q_{\hat{e}_0 \check{a}} Q_{\check{a} \check{a}}^{-1}(a - z), Q_{\hat{e}_0 \hat{e}_0} - Q_{\hat{e}_0 \check{a}} Q_{\check{a} \check{a}}^{-1} Q_{\check{a} \hat{e}_0})$ and $P(\check{a} = z)$ is the PMF of \check{a} .

This result can now be used to test whether or not it is likely that the outcome of \check{e}_0 is a sample from $f_{\check{e}_0}(x)$. But before one can execute the test, we need to know the acceptance and rejection regions. Determining these regions is made difficult by the multi-modality of $f_{\check{e}_0}(x)$. Let $\mathcal{A} \subset \mathbb{R}^q$ be the acceptance region with coverage probability $P(\check{e}_0 \in \mathcal{A}|H_0) = 1 - \alpha$. Since we want the rejection to be rare when the underlying model is correct, the false alarm probability α is chosen as a small value. But since there are an infinite number of subsets that can produce this false alarm probability, we still need to determine a way of defining a proper \mathcal{A} . It seems reasonable to define the optimal subset as the one which has the smallest volume. In that case the probability $1 - \alpha$ would be the most concentrated. This acceptance region is given as

$\mathcal{A} = \{x \in \mathbb{R}^q | f_{\hat{\epsilon}_0}(x) \geq \lambda\}$, where λ is chosen so as to satisfy the given probability constraint. The outcome of the test leads then to rejection of H_0 if $\hat{\epsilon}_0 \notin \mathcal{A}$. Note, due to the multi-modality of $f_{\hat{\epsilon}_0}(x)$, that the acceptance region may consist of a number of disconnected regions. This is a consequence of having the integrerness of a taken into account.

5 Conclusions

This chapter presented a brief review of the theory of integer inference as originally developed for GPS. The theory has however a much wider range of applicability than only GPS. It also applies directly to the next generation GNSSs, such as modernized GPS, Galileo, or Compass, as well as to VLBI, InSAR, and acoustic interferometry, to name a few examples.

The mixed integer model was introduced as the model that underlies (interferometric) measurement systems observing fractional carrier phases. Three different classes of estimators were introduced for the mixed integer model. They are the class of integer estimators (I), the class of integer aperture estimators (IA), and the class of integer equivariant estimators (IE). Each class consists of a multitude of estimators. For each class the optimal estimator was presented. As optimality criterion we used the maximization of the probability of correct integer estimation or the minimization of the mean squared error.

For the I-class, we presented in addition the properties of the three most popular integer estimators. They are the estimators of integer rounding (IR), integer bootstrapping (IB), and integer least-squares (ILS). Special attention was given to computational and statistical issues. The distributional properties of the I-class were given and it was shown how they impact the quality of the corresponding real-valued mixed model parameter estimators. It was also shown that the increase in strength of the mixed integer linear model over that of the standard linear model, allows one to reevaluate the validation of the model. A cross-validation method for the mixed integer model was presented together with the necessary multimodal distribution.

Acknowledgments

The author is a recipient of the Australian Research Council Federation Fellowship (project number FF0883188). This support is gratefully acknowledged.

References

- Baarda W (1968) A testing procedure for use in geodetic networks. In: Publications in geodesy, 2(5). Netherlands Geodetic Commission, Delft
- Betti B, Crespi M, Sanso F (1993) A geometric illustration of ambiguity resolution in GPS theory and a Bayesian approach. *Manuscr Geod* 18:317–330
- Boon F, Ambrosius B (1997) Results of real-time applications of the LAMBDA method in GPS based aircraft landings. In: Proceedings of the KIS97, Banff, Canada, pp 339–345
- Buist PJ, Teunissen PJG, Giorgi G, Verhagen S (2008) Instantaneous multi-baseline ambiguity resolution with constraints. In: Proceedings of the international symposium on GPS/GNSS, Tokyo, Japan, pp 862–871
- Chang XW, Yang X, Zhou T (2005) MLAMBDA: a modified LAMBDA method for integer least-squares estimation. *J Geod* 79:552–565
- Cox DB, Brading JDW (1999) Integration of LAMBDA ambiguity resolution with Kalman filter for relative navigation of spacecraft. In:

- Proceedings of the ION NTM 99, San Diego, pp 739–745
- Dai L, Ling KV, Nagarajan N (2004) Real-time attitude determination for microsatellite by LAMBDA method combined with Kalman filtering. In: Proceedings of the AIAA-ICSS 2004, Montery, vol 3118, pp 1–8
- de Jonge PJ (1998) A processing strategy for the application of the GPS in networks. In: Publications on geodesy, vol 46. Netherlands Geodetic Commission, Delft
- de Jonge PJ, Tiberius CCJM (1996a) The LAMBDA method for integer ambiguity estimation: implementation aspects. In: LGR-Series No. 12. Publications of the Delft Computing Centre, Delft
- de Jonge PJ, Tiberius CCJM (1996b) Integer estimation with the LAMBDA method. In: Beutler G et al (eds) Proceedings of the IAG symposium, vol 115, GPS trends in terrestrial, airborne and spaceborne applications, Boulder. Springer, Berlin pp 280–284
- de Jonge PJ, Tiberius CCJM, Teunissen PJG (1996) Computational aspects of the LAMBDA method for GPS ambiguity resolution. In: Proceedings of the ION GPS-96, Kansas City, pp 935–944
- Eissfeller B, Tiberius CCJM, Pany T, Biberger R, Schueler T, Heinrichs G (2002) Instantaneous ambiguity resolution for GPS/Galileo RTK positioning. *J Gyros Navigat* 38(3):71–91
- Farell JA, Barth M (1999) The global positioning system and inertial navigation. McGraw-Hill, New York
- Giorgi G, Teunissen PJG, Buist P (2008) A search and shrink approach for the constrained LAMBDA method: experimental results. In: Proceedings of the international GPS/GNSS symposium, Tokyo, Japan, pp 797–806
- Grafarend EW (1976) Geodetic applications of stochastic processes. *Phys Earth Planet Interiors* 12:151–179
- Grafarend EW (2000) Mixed integer-real valued adjustment problems: GPS initial cycle ambiguity resolution by means of the LLL algorithm. *GPS Solut* 4(2):31–44
- Gundlich B, Koch KR (2002) Confidence regions for GPS baselines by Bayesian statistics. *J Geod* 76:55–62
- Gundlich B, Teunissen PJG (2004) Multiple models: fixed, switching and interacting. In: Sanso F (ed) Proceedings of the 5th Hotine-Marussi symposium. IAG symposia series, vol 127. Springer, Berlin, pp 129–136
- Hobiger T, Sekido M, Koyama Y, Kondo T (2009) Integer phase ambiguity estimation in next generation very long baseline interferometry. *Adv Space Res* 43(1):187–192
- Hofmann-Wellenhof B, Lichtenegger H, Wasle E (2008) GNSS global navigation systems; GPS, GLONASS, Galileo and More. Springer, Wien
- Ji S, Chen W, Zhao C, Ding X, Chen Y (2007) Single epoch ambiguity resolution for Galileo with the CAR and LAMBDA methods. *GPS Solut* 11(4):259–268
- Joosten P, Tiberius CCJM (2002) LAMBDA: FAQs. *GPS Solut* 6:109–114
- Kampes BM, Hanssen RF (2004) Ambiguity resolution for permanent scatterer interferometry. *IEEE Trans Geosci Remote Sens* 42(11):2446–2453
- Koch KR (1999) Parameter estimation and hypothesis testing in linear models, 2nd edn. Springer, Berlin
- Leick A (2004) GPS satellite surveying, 3rd edn. Wiley, New York
- Liu LT, Hsu HT, Zhu YZ, Ou JK (1999) A new approach to GPS ambiguity decorrelation. *J Geod* 73:478–490
- Misra P, Enge P (2006) Global positioning system: signals, measurements, and performance, 2nd edn. Ganga-Jamuna, Lincoln
- Moenikes R, Wendel J, Trommer GF (2005) A modified LAMBDA method for ambiguity resolution in the presence of position domain constraints. In: Proceedings of the ION GNSS 2005, Long Beach, pp 81–87
- Odijk D (2002) Fast precise GPS positioning in the presence of ionospheric delays. Publications on geodesy, vol 52. Netherlands Geodetic Commission, Delft
- Park C, Teunissen PJG (2003) A new carrier phase ambiguity estimation for GNSS attitude determination systems. In: Proceedings of the international GPS/GNSS symposium, Tokyo, Japan, pp 1–8
- Parkinson B, Spilker JJ (1996) GPS: theory and applications, vols 1 and 2. AIAA, Washington, DC
- Robert CP, Casella G (1999) Monte Carlo statistical methods. Springer Texts in Statistics. Springer, New York
- Strang G, Borre K (1997) Linear algebra, geodesy, and GPS. Wellesley-Cambridge, Wellesley
- Svendsen JGG (2006) Some properties of decorrelation techniques in the ambiguity space. *GPS Solut* 10:40–44
- Teunissen PJG (1993) Least-squares estimation of the integer GPS ambiguities. In: Invited Lecture, Section IV theory and methodology, IAG general meeting, Beijing, China, August 1993. Also in: LGR Series, No. 6. Delft Geodetic Computing Centre, Delft
- Teunissen PJG (1995a) The invertible GPS ambiguity transformations. Manus Geod 20:489–497

- Teunissen PJG (1995b) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. *J Geod* 70: 65–82
- Teunissen PJG (1998) Success probability of integer GPS ambiguity rounding and bootstrapping. *J Geod* 72:606–612
- Teunissen PJG (1999a) The probability distribution of the GPS baseline for a class of integer ambiguity estimators. *J Geod* 70:65–82
- Teunissen PJG (1999b) An optimality property of the integer least-squares estimator. *J Geod* 73: 587–593
- Teunissen PJG (2000) ADOP based upper bounds for the bootstrapped and the least-squares ambiguity success-rates. *Artifi Satel* 35(4):171–179
- Teunissen PJG (2002) The parameter distributions of the integer GPS model. *J Geod* 76:41–48
- Teunissen PJG (2003a) Integer aperture GNSS ambiguity resolution. *Artif Satel* 38(3):79–88
- Teunissen PJG (2003b) Theory of integer equivariant estimation with applications to GNSS. *J Geod* 77:402–410
- Teunissen PJG (2005) GNSS ambiguity resolution with optimally controlled failure-rate. *Artif Satel* 40(4):219–227
- Teunissen PJG (2006) Testing theory: an introduction, 2nd edn. VSSD, Delft
- Teunissen PJG (2007a) Influence of ambiguity precision on the success rate of GNSS ambiguity bootstrapping. *J Geod* 81:351–358
- Teunissen PJG (2007b) Least-squares prediction in linear models with integer unknowns. *J Geod* 81:565–579
- Teunissen PJG, Kleusberg A (1998) GPS for geodesy, 2nd enlarged edn. Springer, Berlin
- Teunissen PJG, Verhagen S (2008) GNSS carrier phase ambiguity resolution: challenges and open problems. In: IAG symposia series, vol 133. Springer, Berlin, pp 785–792
- Teunissen PJG, Verhagen S (2009) The GNSS ambiguity ratio-test revisited: a better way of using it. *Surv Rev* 41:138–151
- Teunissen PJG, de Jonge PJ, Tiberius CCJM (1997) The least-squares ambiguity decorrelation adjustment: its performance on short GPS baselines and short observation time spans. *J Geod* 71:589–602
- Thomsen H (2000) Evaluation of upper and lower bounds on the sucess probability. In: Proceedings of the ION GPS 2000, Salt Lake City, pp 183–188
- Tiberius CCJM, de Jonge PJ (1995) Fast positioning using the LAMBDA method. In: Proceedings of the DSNS-95, Bergen, Norway, paper 30, 8 p
- Tsai YF, Juang JC (2007) Ambiguity resolution validation based on LAMBDA and eigendecomposition. In: Proceedings of the 63rd AM-ION GPS, Portland, pp 299–304
- Verhagen S (2003) On the approximation of the integer least-squares success-rate: which lower or upper bound to use? *J Global Position Syst* 2(2):117–124
- Verhagen S (2005a) The GNSS integer ambiguities: estimation and validation. In: Publications on geodesy, vol 58. Netherlands Geodetic Commission, Delft
- Verhagen S (2005b) On the reliability of integer ambiguity resolution. *Navigation* 52(2):99–110
- Verhagen S, Teunissen PJG (2004) PDF evaluation of the ambiguity residuals. In: Sanso F (ed) Proceedings of the 5th Hotine-Marussi symposium. IAG symposia series, vol 127. Springer, Berlin, pp 120–129
- Verhagen S, Teunissen PJG (2005) Performance comparison of the BIE estimator with the float and fixed GNSS ambiguity estimator. In: Sanso F (ed) IAG symposia series, vol 128. Springer, Berlin, pp 428–433
- Verhagen S, Teunissen PJG (2006) New global navigation satellite system ambiguity resolution method compared to existing approaches. *J Guid Control Dynam* 29(4):981–991
- Wu SC, Bar-Sever YE (2006) Real-time sub-cm differential orbit determination of two low-earth orbiters with GPS bias-fixing. In: Proceedings of the ION GNSS 2006, Fort Worth, pp 2515–2522
- Wu F, Kubo N, Yasuda A (2004) Performance evaluation of GPS augmentation using quasi zenith satellite system. *IEEE Trans Aerosp Electron Syst* 40(4):1249–1260
- Xu P (2001) Random simulation and GPS decorrelation. *J Geod* 75:408–423
- Xu P (2006) Voronoi cells, probabilistic bounds and hypothesis testing in mixed integer linear models. *IEEE Trans Inform Theory* 51(7):3122–3138



New Global Navigation Satellite System Ambiguity Resolution Method Compared to Existing Approaches

Sandra Verhagen* and Peter J. G. Teunissen†
Delft University of Technology, 2629 HS Delft, The Netherlands

Integer carrier phase ambiguity resolution is the process of resolving unknown cycle ambiguities of double-differenced carrier phase data as integers, and it is a prerequisite for rapid and high-precision global navigation satellite system positioning and navigation. Besides integer estimation, integer ambiguity resolution also involves validation of the integer estimates. In this contribution a new ambiguity resolution method is presented, based on the class of integer aperture estimators, which for the first time reveals an overall approach to the combined problem of integer estimation and validation. Furthermore, it is shown how the different discrimination tests that are currently in use in practice can be cast into the framework of the new approach.

Introduction

THE principle of positioning with a global navigation satellite system (GNSS) is based on determining the distance of at least four GNSS satellites to a receiver. For that purpose two types of distance measurements are employed: the code-based pseudoranges and the carrier-phase measurements. The code observation is based on a binary code, which is modulated on the signal carrier. Using only the code measurements allows positioning with an accuracy of several meters. The phase measurement equals the difference between the phase of the receiver-generated carrier signal at reception time and the phase of the carrier signal generated in the satellite at transmission time. It can be measured to within millimeters. However, an integer number of full cycles is unknown since only the fractional phase can be measured. The distance from receiver to satellite, l , is thus

$$l = \lambda\phi + \lambda N + e \quad (1)$$

with ϕ the phase measurement, λ the wavelength of the carrier signal, N the real-valued ambiguity, and e containing all noise and biases, such as atmospheric delays, multipath, and clock errors. Currently, GPS transmits on two frequencies, with wavelengths of 19.0 and 24.4 cm. In the future a third frequency will be available, with a wavelength of 25.5 cm. The future Galileo system will transmit on four frequencies with wavelengths of 19.0, 23.4, 24.8, and 25.5 cm.

To allow positioning with subcentimeter accuracy, it is thus required to use the very precise carrier-phase measurements, and hence to resolve the unknown integer ambiguities. Furthermore, the principle of differential GNSS is used. Differential GNSS involves the use of linear combinations of the GNSS observations from different satellites and receivers, so that common errors are eliminated or mitigated, such as clock errors and atmospheric delays. Besides the user receiver, one or more reference receivers in known locations are required for that purpose. The GNSS models may differ greatly in complexity and diversity. They range from single-baseline models used for kinematic positioning to multibaseline models used as a tool for studying geodynamic phenomena. The models may or may not have the relative receiver–satellite geometry included. They may

also be discriminated as to whether the slave receiver(s) is stationary or in motion, or whether the differential atmospheric delays (ionosphere and troposphere) are included as unknowns. A more detailed description of the GNSS observations and an overview of models can be found in textbooks such as Refs. 1–5.

GNSS ambiguity resolution is the fundamental problem to be addressed in this contribution. Its importance stems from the fact that it is the key to rapid and precise GNSS relative positioning. Once the integer ambiguities are resolved one can take full advantage of the highly precise carrier phase measurements. GNSS ambiguity resolution has important applications in surveying, navigation, geodesy, and geophysics. The applications become more demanding in terms of precision, reliability, availability, and integrity all the time. This is possible thanks to advances in hardware and algorithms. Examples of some highly demanding applications are precision approaches and landing of aircraft,^{6,7} attitude determination,^{8–11} and formation flying of satellites.^{12,13}

To explain the problem of integer ambiguity resolution, we will start with the general GNSS model. Any linear(ized) GNSS relative position model can be cast in the following system of linear observation equations:

$$E\{y\} = Aa + Bb, \quad a \in \mathbb{Z}^n, \quad b \in \mathbb{R}^p \quad (2)$$

with $E\{\cdot\}$ the mathematical expectation operator, y the random m -vector of observables, a the n -vector of unknown integer parameters, and b the p -vector of unknown real-valued parameters. The data vector y will then usually consist of the “observed minus computed” single- or multifrequency double-difference phase and/or pseudorange (code) observations accumulated over all observation epochs. The entries of the integer vector a are the double differences of the real-valued carrier phase ambiguities N , expressed in units of cycles. The entries of the real-valued vector b will consist of the remaining unknown parameters, such as baseline components (coordinates) and atmospheric delay parameters (troposphere, ionosphere).

It is the goal of integer ambiguity resolution to exploit the integrality of the ambiguity vector a when estimating the parameters of interest, which are usually the components of b . In the literature on GNSS ambiguity resolution one can recognize two different approaches to this problem, the Bayesian approach and the non-Bayesian approach. In the Bayesian approach, not only the vector of observables y is assumed random, but also the two vectors of unknown parameters, a and b . The improper priors of a and b are then assumed independent and proportional to an integer-centered pulse train and a constant, respectively. The Bayes estimate of b is then taken as the conditional mean of the posterior probability density function (PDF) of b . Examples of Bayesian analyses of GNSS ambiguity resolution can be found in Refs. 14–18. How the Bayesian approach compares to the non-Bayesian approach of integer equivariant minimum variance estimation is shown in Ref. 19.

Received 3 February 2005; revision received 10 October 2005; accepted for publication 12 October 2005. Copyright © 2005 by Sandra Verhagen. Published by the American Institute of Aeronautics and Astronautics, Inc., with permission. Copies of this paper may be made for personal or internal use, on condition that the copier pay the \$10.00 per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923; include the code 0731-5090/06 \$10.00 in correspondence with the CCC.

*Postdoctoral Researcher, Delft Institute of Earth Observation and Space Systems, Kluyverweg 1; A.A.Verhagen@TUDelft.nl.

†Professor, Delft Institute of Earth Observation and Space Systems.

In the present contribution we will use a non-Bayesian approach throughout. Thus only y is considered to be a random vector. The two parameter vectors a and b are unknown and nonrandom. The procedure for solving this GNSS model is then usually divided into three steps. In the first step the integer constraints $a \in \mathbb{Z}^n$ are discarded and a and b are estimated by the method of least squares. As a result, one obtains the so-called “float” solution \hat{a} and \hat{b} . This solution is real-valued and it is given as $\hat{a} = (\bar{A}^T Q_y^{-1} \bar{A})^{-1} \bar{A}^T Q_y^{-1} y$ and $\hat{b} = (\bar{B}^T Q_y^{-1} \bar{B})^{-1} \bar{B}^T Q_y^{-1} y$, respectively, with $\bar{A} = P_B^\perp A$ and $\bar{B} = P_A^\perp B$, where $P_A^\perp = I_m - A(A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1}$, $P_B^\perp = I_m - B(B^T Q_y^{-1} B)^{-1} B^T Q_y^{-1}$, and Q_y is the variance matrix of y .

Then, in the second step, the float solution \hat{a} is further adjusted to take the integerness of the ambiguities into account. This is done by mapping \hat{a} to an integer vector \check{a} ,

$$\check{a} = S(\hat{a}) \quad (3)$$

with $S : \mathbb{R}^n \mapsto \mathbb{Z}^n$. This estimator is then used in the final step to adjust the float estimator \hat{b} . As a result one obtains the so-called “fixed” estimator of b as

$$\check{b} = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} (\hat{a} - \check{a}) \quad (4)$$

in which $Q_{\hat{a}}$ denotes the variance matrix of \hat{a} and $Q_{\hat{b}\hat{a}}$ denotes the covariance matrix of \hat{b} and \hat{a} . The matrix product $Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1}$, when expressed in A and B , reads $Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} = (\bar{B}^T Q_y^{-1} \bar{B})^{-1} \bar{B}^T Q_y^{-1} \bar{A}$.

The above three-step procedure is still ambiguous in the sense that it leaves room for choosing the n -dimensional map S . Different choices for S will lead to different ambiguity estimators and thus also to different baseline estimators \check{b} . Examples of integer estimators that are in use in practice are the ones based on integer rounding, integer bootstrapping, and integer least squares.

Next to the integer estimation step, integer validation also plays a crucial role in the process of ambiguity resolution. After all, even when one uses an optimal, or close to optimal, integer ambiguity estimator, one can still come up with an unacceptable integer solution. Unfortunately, however, there does not yet exist a rigorous probabilistic theory for the validation of integer ambiguities. Various validation procedures have been proposed in the literature. Some seem to have good performance; others, however, can be shown to perform poorly, whereas still others perform poorly in some cases and well in other cases. One of the earliest and most popular ways of validating the integer ambiguity solution is to make use of the so-called ratio test. The test statistic of the ratio test is defined as the ratio of the squared norm of the “best” ambiguity residual vector and the squared norm of the “second-best” ambiguity residual vector. The computed integer ambiguity solution is then rejected in favor of the float solution when this ratio exceeds a certain user-defined threshold. Also, the so-called F -ratio test, the difference test, and the projector test are in use in practice. They also are defined so that the test statistics depend on the best and second-best ambiguity estimators.

In this contribution we will study the properties and underlying concept of these tests. It will be shown that the procedure underlying the tests can be given a firm theoretical footing, but one that differs significantly from the ones described in the literature. For the practical user of the tests this has the important consequence that the tests have to be evaluated differently than thought so far.

The firm theoretical basis that can be given to the tests is made possible by the recently introduced theory of integer aperture (IA) inference.²⁰ The required ingredients of this theory will be briefly described. It will be shown that the procedure underlying the ratio test, the difference test, and the projector test are all members from the class of integer aperture estimators. This allows us 1) to quantify and qualify the acceptance region of the tests, 2) to give an exact and overall probabilistic evaluation of the combined integer estimation and validation solution, and 3) to show the user how he/she needs to compute the critical value of the discrimination test at hand. But it will also be shown how the optimal integer aperture estimator can be defined, so that the probability of correct integer estimation is

maximized, and at the same time the failure rate will not exceed a user-defined threshold.

The outline of this contribution is as follows. First the theory of integer estimation is reviewed. Then the theory of integer aperture estimation is introduced. The following sections are then devoted to the various tests that are currently in use in practice. For each of the tests it will be proven that the underlying procedures are indeed members from the class of integer aperture estimators, and it will be shown how the corresponding acceptance regions are constructed. Finally, the properties of the tests will be evaluated based on simulations. The results will also be used to compare the performance of the discrimination tests with that of the optimal integer aperture estimator.

Integer Ambiguity Resolution

Integer Estimation

We start by introducing the class of integer estimators. Because $S : \mathbb{R}^n \mapsto \mathbb{Z}^n$ in Eq. (3) is a many-to-one map, different real-valued vectors will be mapped by S to the same integer vector. Therefore, a subset $S_z \subset \mathbb{R}^n$ can be assigned to each integer vector $z \in \mathbb{Z}^n$:

$$S_z = \{x \in \mathbb{R}^n \mid z = S(x), z \in \mathbb{Z}^n\} \quad (5)$$

This subset S_z contains all real-valued float ambiguity vectors that will be mapped to the same integer vector z , and it is called the pull-in region of z .^{21,22} One can now define the class of integer estimators through the conditions that these pull-in regions have to fulfill. We require the pull-in regions to satisfy the following three conditions:

- 1) $\bigcup_{z \in \mathbb{Z}^n} S_z = \mathbb{R}^n$
 - 2) $\text{Int}(S_u) \cap \text{Int}(S_z) = \emptyset, \quad \forall u, z \in \mathbb{Z}^n, \quad u \neq z$
 - 3) $S_z = z + S_0, \quad \forall z \in \mathbb{Z}^n$
- (6)

where “Int” denotes the interior of the subset. Hence, the pull-in regions are required to be translationally invariant and to cover \mathbb{R}^n without gaps and overlaps. In Ref. 23 the motivation for this definition is given. It follows that every integer estimator \check{a} can now be expressed as

$$\check{a} = \sum_{z \in \mathbb{Z}^n} z s_z(\hat{a}) \quad (7)$$

with the indicator function, $s_z(x)$, defined as

$$s_z(x) = \begin{cases} 1 & \text{if } x \in S_z \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

Examples of estimators that belong to this class of integer estimators are integer rounding, integer bootstrapping, and integer least squares (ILS). Their pull-in regions are the multivariate versions of a square, a parallelogram, and a hexagon, respectively. The pull-in regions always have a volume equal to one.

It will be clear that the outcome of an integer estimator should only be used if one has enough confidence in the solution. Hence, one needs to evaluate the distribution of

$$\check{a} = \sum_{z \in \mathbb{Z}^n} z s_z(\hat{a})$$

This distribution, denoted as $P(\check{a} = z)$, is a probability mass function with zero masses at noninteger points, and nonzero masses at some or all integer points. With the division of \mathbb{R}^n into mutually exclusive pull-in regions, the distribution of \check{a} follows as²⁴

$$P(\check{a} = z) = P(\hat{a} \in S_z) = \int_{S_z} f_{\hat{a}}(x) dx, \quad z \in \mathbb{Z}^n \quad (9)$$

where $f_{\hat{a}}(x)$ is the continuous probability density function (PDF) of the float ambiguity vector \hat{a} . In most GNSS applications the data are assumed to be normally distributed, and thus the estimator \check{a} will

be normally distributed too with mean $a \in \mathbb{Z}^n$ and variance matrix $Q_{\hat{a}}$. The PDF is then given by

$$f_{\hat{a}}(x) = C \exp\left\{-\frac{1}{2}\|x - a\|_{Q_{\hat{a}}}^2\right\} \quad (10)$$

where C is a normalization constant.

From the probability mass function of \hat{a} , one is particularly interested in the probability of correct integer estimation, the so-called success rate of ambiguity resolution. The success rate $P_s = P(\hat{a} = a)$ follows from Eq. (9) as

$$P_s = P(\hat{a} = a) = \int_{S_a} f_{\hat{a}}(x) dx = \int_{S_0} f_{\hat{a}}(x + a) dx \quad (11)$$

Note that the success rate can be computed without knowledge of a for any distribution of \hat{a} that is translated over a when the mean a is subtracted from the stochastic variable \hat{a} , that, $f_{\hat{a}}(x) = f_{\hat{a}-a}(x-a)$. Obviously, this is the case for the normal distribution of Eq. (10).

Since the success rate depends on the pull-in region and therefore on the chosen integer estimator, it is of importance to know which integer estimator maximizes the probability of correct integer estimation. For an arbitrary PDF the solution is given in Ref. 25. For elliptically contoured distributions, such as the multivariate normal distribution, the ILS-estimator can be shown to be the optimal estimator.²³ The ILS estimator is given as

$$\check{a}_{\text{LS}} = \arg \min_{z \in \mathbb{Z}^n} \|\hat{a} - z\|_{Q_{\hat{a}}}^2 \quad (12)$$

with the squared norm $\|\cdot\|_Q^2 = (\cdot)^T Q^{-1}(\cdot)$. Note, since

$$\begin{aligned} \|y - Aa - Bb\|_Q^2 &= \|y - A\hat{a} - B\hat{b}\|_Q^2 + \|\hat{a} - a\|_{Q_{\hat{a}}}^2 \\ &\quad + \|\hat{b} - b - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1}(\hat{a} - a)\|_{(B^T Q_y^{-1} B)^{-1}}^2 \end{aligned}$$

that \check{a}_{LS} and $\check{b}_{\text{LS}} = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1}(\hat{a} - \check{a}_{\text{LS}})$ is the solution of $\min_{a \in \mathbb{Z}^n, b \in \mathbb{R}^n} \|y - Aa - Bb\|_Q^2$ (Ref. 26).

In contrast to integer rounding and integer bootstrapping, an integer search is needed to compute \check{a}_{LS} . The ILS-procedure is mechanized in the LAMBDA (least-squares ambiguity decorrelation adjustment) method,^{26–28} which is currently one of the most applied methods for GNSS carrier phase ambiguity resolution. Practical results obtained with it can be found, for example, in Refs. 9, 11, 13, 21, and 29–41. Further studies on the decorrelation step of the LAMBDA method can be found in Refs. 42–46.

To evaluate the success rate of ILS-estimation, we need to integrate the PDF over the ILS pull-in region; cf. Eq. (11). The ILS pull-in region that belongs to the integer vector z is given as

$$\begin{aligned} S_z &= \left\{ x \in \mathbb{R}^n \mid \|x - z\|_{Q_{\hat{a}}}^2 \leq \|x - u\|_{Q_{\hat{a}}}^2, \forall u \in \mathbb{Z}^n \right\} \\ &= \bigcap_{c \in \mathbb{Z}^n} \left\{ x \in \mathbb{R}^n \mid |c^T Q_{\hat{a}}^{-1}(x - z)| \leq \frac{1}{2}\|c\|_{Q_{\hat{a}}}^2 \right\} \quad (13) \end{aligned}$$

Hence, the ILS pull-in regions are constructed as intersecting half-spaces, which are bounded by the plane orthogonal to $(u - z)$, $u \in \mathbb{Z}^n$, and passing through $\frac{1}{2}(u + z)$. Because of the complexity of the geometry of the ILS pull-in region, it is not possible to give an exact evaluation of the ILS success rate. One has to resort to Monte Carlo methods or to the use of lower and upper bounds. Fortunately sharp bounds are available for evaluating the ILS success rate. An overview and evaluation of lower and upper bounds is given in Ref. 47.

Integer Validation in Practice

It is important to realize that the success rate as introduced here is a diagnostic measure for the expected performance of the integer estimator, but not of its actual performance. Thus even if the success rate is considered large enough, it can still happen that an actual outcome of the integer estimator is erroneous. This is why in practice so-called discrimination tests are used. Of course, one can never be sure whether the integer outcome is correct or not. But what one can do is to single out integer outcomes that are suspect. This is the whole purpose of the discrimination tests. Several such tests

have been proposed in the literature and are currently in use in practice.^{38,48–55}

All these tests compare in one way or another the ILS solution \check{a}_{LS} with a so-called “second best” integer solution \check{a}' . Based on this comparison, the outcome of the discrimination tests is either to accept the integer solution \check{a}_{LS} , or to reject it in favor of the float solution \hat{a} . Examples of discrimination tests are the ratio test, the difference test, and the projector test.

Unfortunately, all discrimination tests proposed in the literature until now lack a sound theoretical basis. They have been introduced in an ad hoc way, their evaluation is based on the incorrect assumption that \check{a}_{LS} and \check{a}' are nonrandom, they use critical values that are chosen as fixed values, and their performance is treated separately from the performance of the integer estimator. An overview of how these tests are used in practice and their pitfalls is given in Ref. 56. In this contribution it will be shown that it is possible to give a theoretical foundation to these tests. It will be shown that the tests all fit in the framework of the new class of integer aperture estimators. As a result, the probabilistic evaluation of integer aperture estimation makes it possible to give an overall evaluation of both the integer estimation part and the integer discrimination part.

Integer Aperture Estimation

The idea behind IA estimation is to distinguish between the following three situations: *success* if the integer ambiguity is estimated correctly, *failure* if the integer ambiguity is estimated incorrectly, and *undecided* if the real-valued float solution is maintained. These three possible outcomes can be brought together in one estimator if one drops the condition that there may be no gaps between the pull-in regions. To achieve this, we now take, instead of the whole space \mathbb{R}^n , a subset $\Omega \subset \mathbb{R}^n$ as the region for which \hat{a} is mapped to an integer. Thus \hat{a} is mapped to an integer if $\hat{a} \in \Omega$ and the float solution is maintained if $\hat{a} \in \mathbb{R}^n \setminus \Omega$. The region Ω is called the aperture space. To ensure that any integer-perturbed version of \hat{a} gets mapped to an integer whenever \hat{a} itself is mapped to an integer, the aperture space is required to be translationally invariant, $\Omega = \Omega + z$, $\forall z \in \mathbb{Z}^n$. To determine to which integer the float solution is mapped, we introduce

$$\Omega_z = \Omega \cap S_z, \quad \forall z \in \mathbb{Z}^n \quad (14)$$

where S_z is a pull-in region satisfying the conditions of Eq. (6). Then

- 1) $\bigcap_{z \in \mathbb{Z}^n} \Omega_z = \Omega$
- 2) $\text{Int}(\Omega_u) \cap \text{Int}(\Omega_z) = \emptyset, \quad \forall u, z \in \mathbb{Z}^n, \quad u \neq z$
- 3) $\Omega_z = z + \Omega_0, \quad \forall z \in \mathbb{Z}^n$

This shows that the subsets $\Omega_z \subset S_z$ satisfy conditions similar to those satisfied by S_z , with \mathbb{R}^n replaced by $\Omega \subset \mathbb{R}^n$. The IA estimator can now be defined as follows. The IA estimator maps the float solution \hat{a} to the integer vector z when $\hat{a} \in \Omega_z$ and it maps the float solution to itself when $\hat{a} \notin \Omega$. Any IA estimator, denoted as \bar{a} , can then be expressed as

$$\bar{a} = \sum_{z \in \mathbb{Z}^n} z \omega_z(\hat{a}) + \hat{a} \left(1 - \sum_{z \in \mathbb{Z}^n} \omega_z(\hat{a}) \right) \quad (16)$$

with the indicator function, $\omega_z(x)$, defined as

$$\omega_z(x) = \begin{cases} 1 & \text{if } x \in \Omega_z \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

Compare Eq. (16) with Eq. (7). Note that an IA estimator is completely defined once Ω_0 is chosen. This subset can be seen as an adjustable pull-in region with two limiting cases: the limiting case in which Ω_0 is empty and the limiting case in which Ω_0 equals S_0 . In the first case the IA estimator becomes identical to the float estimator \hat{a} , and in the second case the IA estimator becomes identical to an integer estimator. The subset Ω_0 therefore determines the *aperture* of the pull-in region.

To be able to evaluate an IA estimator, we need to distinguish between the following three cases:

- | | |
|---|--|
| $\hat{a} \in \Omega_a$ | success: correct integer estimation |
| $\hat{a} \in \Omega \setminus \{\Omega_a\}$ | failure: incorrect integer estimation |
| $\hat{a} \notin \Omega$ | undecided: ambiguity not fixed to an integer |

where $\Omega \setminus \Omega_a$ means that Ω_a is deleted from the set Ω , with a being the unknown integer ambiguity vector.

The corresponding probabilities of success (s), failure (f), and undecidedness (u) are given by

$$\begin{aligned} P_s &= P(\bar{a} = a) = \int_{\Omega_a} f_{\bar{a}}(x) dx \\ P_f &= \sum_{z \in \mathbb{Z}^n \setminus \{a\}} \int_{\Omega_z} f_{\bar{a}}(x) dx = \int_{\Omega_0} f_{\bar{e}}(x) dx - \int_{\Omega_a} f_{\bar{a}}(x) dx \\ P_u &= 1 - P_s - P_f = 1 - \int_{\Omega_0} f_{\bar{e}}(x) dx \end{aligned} \quad (18)$$

The first two probabilities are referred to as success rate and failure rate, respectively. The expression for the failure rate is obtained by using the probability density function of the ambiguity residuals $\bar{e} = \hat{a} - \bar{a}$ (Ref. 24):

$$f_{\bar{e}}(x) = \sum_{z \in \mathbb{Z}^n} f_{\bar{a}}(x + z) s_0(x) \quad (19)$$

with $s_0(x)$ the indicator function of Eq. (8), and \bar{e} the estimator of $e = \hat{a} - a$ with $f_{\bar{e}}(x) = f_{\bar{a}}(x + a)$. Evaluation of the PDF in Eq. (19) involves an infinite sum. In Ref. 57 it is demonstrated how a finite integer set can be chosen for which this sum converges with an approximation error on the order of 10^{-8} within reasonable computation times.

As mentioned earlier, for a user it is especially important that the failure rate be below a certain limit. The approach of integer aperture estimation now allows us to choose a threshold for the failure rate, and then determine the size of the aperture pull-in regions such that the failure rate will be equal to or below this threshold. Once the failure rate P_f has been chosen, the various other probabilities can be computed. For a user, knowledge of the unconditional success rate $P_s = P(\bar{a} = a)$ and the conditional success rate $P(\bar{a} = a | \hat{a} \in \Omega)$ is particularly important. The unconditional success rate describes the probability of a correct outcome, whereas the conditional success rate describes the probability that an integer outcome is correct. Because $P_s + P_f$ equals the probability of an integer outcome, the conditional success rate is given by the ratio $P_s / (P_s + P_f)$. Hence, the conditional success rate will be close to one when the failure rate is chosen close to zero. This implies that one can have very high confidence in the correctness of the integer outcomes of the integer aperture estimator, even for modest values of the unconditional success rate. If one used an integer estimator instead of an integer aperture estimator, then such a high confidence could only be reached once the success rate of the integer estimator was close to one; cf. Eq. (11).

Application of this fixed failure rate approach means that implicitly the ambiguity estimate is validated using a sound criterion. This is an important result, because it means that the integer aperture estimation method for the first time reveals an overall approach to the problem of integer ambiguity estimation and validation, which allows probabilistic evaluation of the final solution.

Perhaps superfluously, we note that these properties hold true on the average, but not necessarily for the individual outcomes of a single experiment. This is inherent in any probabilistic evaluation of random phenomena. Thus an outcome $\bar{a} = \hat{a}$, for instance, is no guarantee that the corresponding rejected outcome \bar{a} is further apart from the correct integer a than \hat{a} is from a . The probability, however, that the estimator \bar{a} is further apart from a than the estimator \hat{a} is from a is less than the failure rate and therefore small.

So far, we have not specified the aperture pull-in regions completely. There are still several options left with respect to the choice of the shape of the aperture pull-in regions. In the following sections it will be shown that the procedures underlying the discrimination tests are all examples of integer aperture estimators. For the different tests it will be shown how the shape of the aperture pull-in regions is determined.

Optimal Integer Aperture Estimation

It is also possible to define an *optimal* integer aperture estimator by requiring that the success rate will be maximized for a given failure rate; that is, the optimization problem is given by

$$\max_{\Omega_0 = \Omega \cap S_0} P_s \quad \text{subject to} \quad P_f = \beta \quad (20)$$

where β is the chosen value for the fixed failure rate. For an arbitrary PDF of \hat{a} the solution to the optimization problem is presented in Ref. 58. For elliptically contoured distributions, the optimal integer aperture estimator can be regarded as a combination of ILS-estimation and the following test:

$$\text{Accept } \check{a}_{\text{LS}} \text{ iff: } \frac{f_{\bar{e}}(\check{e}_{\text{LS}})}{f_{\bar{e}}(\check{e}_{\text{LS}})} \leq \mu, \quad \mu > 1 \quad (21)$$

with $\check{e}_{\text{LS}} = \hat{a} - \check{a}_{\text{LS}}$. Note that for the computation of $f_{\bar{e}}(\check{e}_{\text{LS}})$ and $f_{\bar{e}}(\check{e}_{\text{LS}})$ knowledge of a is not needed for any distribution for which $f_{\bar{a}}(x) = f_{\bar{a}-a}(x-a)$. If for instance the normal distribution is used, $f_{\bar{e}}(\check{e}_{\text{LS}}) = f_{\bar{a}}(\check{e}_{\text{LS}} + a)$ can be computed with Eq. (10), and then a is eliminated. Equation (19) is used for the computation of $f_{\bar{e}}(\check{e}_{\text{LS}})$. Because of the infinite sum over all integers, again knowledge of the true integer a is not required. Furthermore note that, because $\check{e}_{\text{LS}} \in S_0$, the factor $s_0(\check{e}_{\text{LS}})$, which is needed for the evaluation of $f_{\bar{e}}(\check{e}_{\text{LS}})$ with Eq. (19), will always come out to be 1. Hence, with the normal distribution for the float ambiguity estimator, the ratio in Eq. (21) becomes

$$\frac{f_{\bar{e}}(\check{e}_{\text{LS}})}{f_{\bar{e}}(\check{e}_{\text{LS}})} = \frac{\sum_{z \in \mathbb{Z}^n} \exp\left\{-\frac{1}{2} \|\check{e}_{\text{LS}} + z\|_{Q_{\bar{a}}}^2\right\}}{\exp\left\{-\frac{1}{2} \|\check{e}_{\text{LS}}\|_{Q_{\bar{a}}}^2\right\}}$$

We note that the above optimal integer aperture estimator can also be interpreted from two different alternative viewpoints. The first one is a Bayesian viewpoint in which all parameters are considered random, with a diffuse prior assumption for the integer parameters. In this context, the ratio on the left-hand side of the inequality in Eq. (21) is equal to the reciprocal marginal a posteriori probability of \check{a}_{LS} being the true integer vector.^{15,16} The second alternative viewpoint is provided by the non-Bayesian penalized ambiguity estimator introduced in Ref. 59. With this estimator the user is given the possibility of assigning penalties to each of the possible outcomes of the estimator. The average of all penalties is minimized by the penalized ambiguity estimator. When penalties are assigned to the three cases, success, failure, and undecided, the structure of the penalized ambiguity estimator is the same as that of the optimal integer aperture estimator. The parameter μ of Eq. (21) will then be a function of the assigned penalties.

The optimal test of Eq. (21) has the highest possible success rate, unconditional as well as conditional, for a given failure rate. The critical value μ is referred to as the aperture parameter, because it determines the size of the aperture pull-in regions; it is completely determined by the constraint on the failure rate, $P_f = \beta$. The higher the failure rate, the higher the value of μ , and hence the larger the aperture of Ω_0 . The aperture parameter μ cannot be smaller than one, because the numerator of Eq. (21) is always larger than the denominator. If the failure rate is required to be smaller than the success rate, which is a reasonable condition, it can be shown that μ must then be smaller than 2, because $P_f \leq (\mu - 1)P_s$ (Ref. 58).

The computational steps involved in computing the optimal integer aperture estimator are now as follows. First compute the integer least-squares solution, \check{a}_{LS} . Then form the ambiguity residual, $\check{e}_{\text{LS}} = \hat{a} - \check{a}_{\text{LS}}$, and check whether $\check{e}_{\text{LS}} \in \Omega_0$; cf. the test of Eq. (21). If

this is the case then the outcome of the optimal estimator is \check{a}_{LS} , otherwise the outcome is \hat{a} . For computational efficiency it is advisable to compute \check{a}_{LS} with the LAMBDA method and use the LAMBDA-transformed ambiguities also for the evaluation of $\check{x}_{LS} \in \Omega_0$.

In the next and following sections we will consider four discrimination tests that are currently in use and show that each of the underlying procedures is an example of an IA-estimator. The tests considered are the ratio test, the F -ratio test, the difference test, and the projector test.

Ratio Test

A very popular discrimination test is the one introduced in Ref. 50. It is given by

$$\text{Accept } \check{a}_{LS} \text{ iff: } \frac{\|\hat{a} - \check{a}_2\|_{Q_{\hat{a}}}^2}{\|\hat{a} - \check{a}_{LS}\|_{Q_{\hat{a}}}^2} = \frac{R_2}{R_1} \geq c \quad (22)$$

where the notation R_i is used for the squared norm of ambiguity residuals of the best ($i = 1$) and second-best ($i = 2$) integer solution, \check{a}_{LS} and \check{a}_2 respectively, as measured by the squared norm of the ambiguity residual vector.

A problem with this approach is that the determination of the critical value is not straightforward. In Ref. 50 the test statistic is derived by applying the classical theory of hypothesis testing. However, to determine the critical value, incorrect assumptions are made on the probability distributions of the parameters involved. Other approaches have also been proposed in the literature, all based on choosing a fixed critical value for test (22), without a theoretical basis. See Ref. 60 for more detailed comments on the approaches proposed in the literature.

Another point of criticism of the ratio test is that the combined integer estimation and validation solution lacks an overall probabilistic evaluation.

But, despite this criticism, integer validation based on the ratio test is often reported to work satisfactorily in practice.

Integer Aperture Estimation with the Ratio Test

It will be shown that the procedure underlying the ratio test is a member of the class of integer aperture estimators. In the sequel the inverse of the test statistic of Eq. (22) will be used:

$$\text{Accept } \check{a}_{LS} \text{ iff: } R_1/R_2 \leq \mu, \quad 0 < \mu \leq 1 \quad (23)$$

The critical value is denoted as μ . For the following to be true, the condition $0 < \mu \leq 1$ must be fulfilled. The acceptance region or aperture space is then given as

$$\Omega = \{x \in \mathbb{R}^n \mid \|x - \check{x}_{LS}\|_{Q_{\hat{a}}}^2 \leq \mu \|x - \check{x}_2\|_{Q_{\hat{a}}}^2, 0 < \mu \leq 1\} \quad (24)$$

with \check{x}_{LS} and \check{x}_2 the best and second-best estimators of x , respectively. Let $\Omega_z = \Omega \cap S_z$; that is, Ω_z is the intersection of Ω with the ILS pull-in region as defined in Eq. (13). Then all conditions of Eq. (15) are fulfilled, because

$$\left\{ \begin{array}{l} \Omega_0 = \{x \in \mathbb{R}^n \mid \|x\|_{Q_{\hat{a}}}^2 \leq \mu \|x - z\|_{Q_{\hat{a}}}^2, \forall z \in \mathbb{Z}^n \setminus \{0\}\} \\ \Omega_z = \Omega_0 + z, \quad \forall z \in \mathbb{Z}^n \\ \Omega = \bigcup_{z \in \mathbb{Z}^n} \Omega_z \end{array} \right. \quad (25)$$

The proof is as follows:

$$\begin{aligned} \Omega_z &= \Omega \cap S_z \\ &= \{x \in S_z \mid \|x - \check{x}_{LS}\|_{Q_{\hat{a}}}^2 \leq \mu \|x - \check{x}_2\|_{Q_{\hat{a}}}^2, 0 < \mu \leq 1\} \\ &= \{x \in S_z \mid \|x - \check{x}_{LS}\|_{Q_{\hat{a}}}^2 \leq \mu \|x - u\|_{Q_{\hat{a}}}^2, \forall u \in \mathbb{Z}^n \setminus \{\check{x}_{LS}\}, \\ &\quad 0 < \mu \leq 1\} \\ &= \{x \in S_z \mid \|x - z\|_{Q_{\hat{a}}}^2 \leq \mu \|x - u\|_{Q_{\hat{a}}}^2, \forall u \in \mathbb{Z}^n \setminus \{z\}, 0 < \mu \leq 1\} \end{aligned}$$

$$\begin{aligned} &= \{x \in \mathbb{R}^n \mid \|x - z\|_{Q_{\hat{a}}}^2 \leq \mu \|x - u\|_{Q_{\hat{a}}}^2, \forall u \in \mathbb{Z}^n \setminus \{z\}, \\ &\quad 0 < \mu \leq 1\} \\ &= \{x \in \mathbb{R}^n \mid \|y\|_{Q_{\hat{a}}}^2 \leq \mu \|y - v\|_{Q_{\hat{a}}}^2, \forall v \in \mathbb{Z}^n \setminus \{0\}, \\ &\quad x = y + z, 0 < \mu \leq 1\} \\ &= \Omega \cap S_0 + z \\ &= \Omega_0 + z \end{aligned}$$

The first two equalities follow from the definition of the ratio test, and the third from $\|\hat{x} - \check{x}_{LS}\|_{Q_{\hat{a}}}^2 \leq \|\hat{x} - \check{x}_2\|_{Q_{\hat{a}}}^2 \leq \|\hat{x} - u\|_{Q_{\hat{a}}}^2, \forall u \in \mathbb{Z}^n \setminus \{\check{x}_{LS}\}$. The fourth equality follows because $\check{x}_{LS} = z$ is equivalent to $x \in S_z$. The fifth equality follows from the fact that $0 < \mu \leq 1$. The last equalities follow from a change of variables, and the definition of $\Omega_0 = \Omega \cap S_0$. Finally, note that

$$\bigcup_{z \in \mathbb{Z}^n} \Omega_z = \bigcup_{z \in \mathbb{Z}^n} (\Omega \cap S_z) = \Omega \cap \left(\bigcup_{z \in \mathbb{Z}^n} S_z \right)$$

This ends the proof of Eq. (25).

The acceptance region of the ratio test consists thus of an infinite number of regions, each one of which is an integer-translated copy of $\Omega_0 \subset S_0$. The acceptance region plays the role of the aperture space, and μ plays the role of the aperture parameter, because it controls the size of the aperture pull-in regions.

It has now been shown that indeed there is a theoretical basis for the ratio test, because the underlying procedure is an integer aperture estimator and there is a sound criterion available for choosing the critical value, or aperture parameter, by means of the fixed-failure-rate approach described at the end of the section on integer aperture estimation.

Ratio Test Aperture Pull-in Region

To understand how the size and shape of the ratio test aperture pull-in region is governed by the aperture parameter and the variance matrix of the float solution, we will now reveal its geometry. From the definition of the aperture pull-in region, Ω_0 , the following can be derived:

$$\begin{aligned} \Omega_0 : \|x\|_{Q_{\hat{a}}}^2 &\leq \mu \|x - z\|_{Q_{\hat{a}}}^2, \quad \forall z \in \mathbb{Z}^n \setminus \{0\}, \quad 0 < \mu \leq 1 \\ \iff \|(1 - \mu)/\mu)x + z\|_{Q_{\hat{a}}}^2 &\leq (1/\mu)\|z\|_{Q_{\hat{a}}}^2, \quad \forall z \in \mathbb{Z}^n \setminus \{0\} \\ \iff \|x + [\mu/(1 - \mu)]z\|_{Q_{\hat{a}}}^2 &\leq [\mu/(1 - \mu)^2]\|z\|_{Q_{\hat{a}}}^2 \\ \forall z \in \mathbb{Z}^n \setminus \{0\} \end{aligned} \quad (26)$$

This shows that the aperture pull-in region is equal to the intersection of all ellipsoids with centers $-[\mu/(1 - \mu)]z$ and radius $[\sqrt{\mu/(1 - \mu)}]\|z\|_{Q_{\hat{a}}}$.

In fact, the intersection region is only determined by the adjacent integers, because $\forall u \in \mathbb{Z}^n$ not being adjacent, Eq. (26) is always fulfilled $\forall x \in S_0$. It follows from Eqs. (24) and (25) that for all $x \in S_0$ on the boundary of Ω_0 ,

$$\|x\|_{Q_{\hat{a}}}^2 = \mu \|x - z\|_{Q_{\hat{a}}}^2, \quad z = \arg \min_{z \in \mathbb{Z}^n \setminus \{0\}} \|x - z\|_{Q_{\hat{a}}}^2$$

So z is the second closest integer and must therefore be adjacent. Hence, the following is always true:

$$\|x\|_{Q_{\hat{a}}}^2 = \mu \|x - z\|_{Q_{\hat{a}}}^2 \leq \mu \|x - u\|_{Q_{\hat{a}}}^2, \quad \forall u \in \mathbb{Z}^n \setminus \{0, z\}$$

and thus the integers u do not have any influence on the boundary of the aperture pull-in region.

This means that the integer least-squares pull-in region S_0 can be split up into sectors, all having another integer z_2 as second closest; that is, $\|x\|_{Q_{\hat{a}}}^2 \leq \|x - z_2\|_{Q_{\hat{a}}}^2 \leq \|x - u\|_{Q_{\hat{a}}}^2, \forall u \in \mathbb{Z}^n \setminus \{0\}$. Within a sector, the aperture pull-in region is then equal to the intersection of the sector with the ellipsoid with center $-[\mu/(1 - \mu)]z_2$ and radius

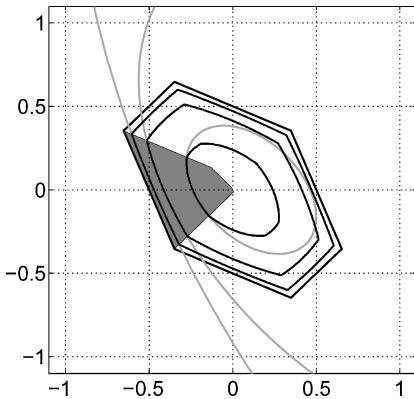


Fig. 1 Geometry of ratio test aperture pull-in region. For all x that fall into the gray region, the second closest integer is $z_2 = [-1 \ 0]^T$. The boundary of the aperture pull-in region within this gray region is then equal to the ellipsoid with center $-[\mu/(1-\mu)]z_2$ and radius $[\sqrt{\mu/(1-\mu)}]\|z_2\|_{Q_{\hat{a}}}$. Examples are shown for μ , 0.1, 0.5, and 0.8, respectively. Ellipsoids are gray; aperture pull-in regions black.

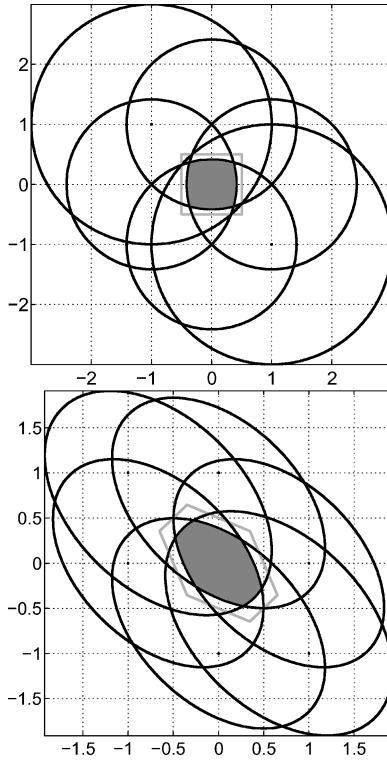


Fig. 2 Geometry of ratio test aperture pull-in region (gray); examples for two variance matrices. Top: identity matrix, $\mu = 0.5$; bottom: Q_1 , $\mu = 0.4$.

$[\sqrt{\mu/(1-\mu)}]\|z_2\|_{Q_{\hat{a}}}$. This is illustrated for the two-dimensional case in Fig. 1, with

$$Q_{\hat{a}} = \begin{bmatrix} 0.0865 & -0.0364 \\ -0.0364 & 0.0847 \end{bmatrix} = Q_1 \quad (27)$$

It can be seen that especially for larger μ , the shape of the aperture pull-in region starts to resemble that of the ILS pull-in region. The reason is that for larger μ the radius of the ellipsoid increases, and the center of the ellipsoid is further away, but in the direction of $[z_2, -z_2]$, and the ellipsoid has the same orientation as the ILS pull-in region.

Figure 2 shows two two-dimensional examples of the geometry of the aperture pull-in regions. For a diagonal matrix, the ILS pull-in region becomes square and there are only four adjacent integers. It

can be seen in the left panel that indeed only these four integers determine the shape of the aperture pull-in region. With the covariance matrix Q_1 there are six adjacent integers determining the shape.

F-Ratio Test

The *F*-ratio test also considers the residuals of the best and second-best integer solution, but additionally takes into account the residuals of the complete float solution, $\hat{e} = y - A\hat{a} - B\hat{b}$. It is defined as follows:

$$\text{Accept } \check{a}_{\text{LS}} \text{ iff: } \frac{\|\hat{e}\|_{Q_y}^2 + \|\hat{a} - \check{a}_2\|_{Q_{\hat{a}}}^2}{\|\hat{e}\|_{Q_y}^2 + \|\hat{a} - \check{a}_{\text{LS}}\|_{Q_{\hat{a}}}^2} \geq c \quad (28)$$

in which $c > 1$ is the chosen critical value. Note that the choice of the critical value is not trivial. In practice, it is often assumed that the test statistic has an *F*-distribution, so that the critical value can be based on a choice of the level of significance. However, this is not correct because the terms in the numerator and denominator are not independent and not χ^2 distributed.⁶¹ Still, the test in Eq. (28) is often used and seems to work satisfactorily; see, for examples, Refs. 48 and 51. In Ref. 54 it is proposed to use the test statistic with a critical value of $c = 2$.

Integer Aperture Estimation with the F-Ratio Test

With the ratio test, the solution is always accepted or rejected for a certain value of the ambiguity residuals and for a certain μ . With the *F*-ratio test in Eq. (28), however, this is not the case. The test statistic namely depends on the residuals of the complete float solution and on the ambiguity residuals. Using the inverse ratio of Eq. (28), the test becomes

$$\text{Accept } \check{a}_{\text{LS}} \text{ iff: } \frac{\|\hat{e}\|_{Q_y}^2 + \|\hat{a} - \check{a}_{\text{LS}}\|_{Q_{\hat{a}}}^2}{\|\hat{e}\|_{Q_y}^2 + \|\hat{a} - \check{a}_2\|_{Q_{\hat{a}}}^2} \leq \mu, \quad 0 < \mu \leq 1 \quad (29)$$

It can be shown that the procedure underlying this test is a member of the class of IA estimators. An integer perturbation of the observations $y + Az$, with $z \in \mathbb{Z}^n$, does not propagate into the least-squares residuals, and \hat{a} , \check{a}_{LS} , and \check{a}_2 all undergo the same change, so that both $\hat{a} - \check{a}_{\text{LS}}$ and $\hat{a} - \check{a}_2$ are invariant under the integer perturbation.

F-Ratio Test Aperture Pull-in Region

Similarly to the ratio test, it can be shown what the geometry of the aperture pull-in regions of the *F*-ratio test IA estimator is:

$$\begin{aligned} \Omega_0 : \|x\|_{Q_{\hat{a}}}^2 + \|\hat{e}\|_{Q_y}^2 &\leq \mu(\|x - z\|_{Q_{\hat{a}}}^2 + \|\hat{e}\|_{Q_y}^2) \\ \forall z \in \mathbb{Z}^n \setminus \{0\}, \quad 0 < \mu \leq 1 \\ \iff \|x\|_{Q_{\hat{a}}}^2 &\leq \mu\|x - z\|_{Q_{\hat{a}}}^2 + (\mu - 1)\|\hat{e}\|_{Q_y}^2 \\ \iff \|x + [\mu/(1-\mu)z]\|_{Q_{\hat{a}}}^2 &\leq [\mu/(1-\mu)^2]\|z\|_{Q_{\hat{a}}}^2 - \|\hat{e}\|_{Q_y}^2 \\ \forall z \in \mathbb{Z}^n \setminus \{0\} \end{aligned} \quad (30)$$

So Ω_0 is equal to the intersection of all ellipsoids with centers $-[\mu/(1-\mu)]z$ and radius equal to the square root of the term on the right-hand side of the last inequality in Eq. (30). The shape of the aperture pull-in region is identical to that of the ratio test IA estimator. Obviously, the difference with the ratio test IA estimator is that the radii depend on the sample value of \hat{e} . For the same value of μ they will be smaller than or equal to the radii of the ellipsoids determining the ratio test IA aperture pull-in region.

Difference Test

The ratio test statistic is defined as the ratio of the quadratic forms R_1 and R_2 . Another approach would be to look at the difference³⁸ such that the test becomes

$$\text{Accept } \check{a}_{\text{LS}} \text{ iff: } R_2 - R_1 \geq c \quad (31)$$

It will be clear that again the problem with this test is the choice of the critical value c : only an empirically determined value can be used. Note that in contrast to the ratio test (22), with this test the critical value depends on the variance factor of unit weight, σ^2 , if $Q_{\hat{a}} = \sigma^2 G_{\hat{a}}$.

Integer Aperture Estimation with the Difference Test

Similarly to the ratio test, it can be shown that the procedure underlying the difference test (31) belongs to the class of IA estimators. The acceptance region of this test is given as follows:

$$\Omega = \{x \in \mathbb{R}^n \mid \|x - \check{x}_{LS}\|_{Q_{\hat{a}}}^2 \leq \|x - \check{x}_2\|_{Q_{\hat{a}}}^2 - \mu\} \quad (32)$$

Let $\Omega_z = \Omega \cap S_z$. Then

$$\begin{cases} \Omega_0 = \{x \in \mathbb{R}^n \mid \|x\|_{Q_{\hat{a}}}^2 \leq \|x - z\|_{Q_{\hat{a}}}^2 - \mu, \forall z \in \mathbb{Z}^n \setminus \{0\}\} \\ \Omega_z = \Omega_0 + z, \quad \forall z \in \mathbb{Z}^n \\ \Omega = \bigcup_{z \in \mathbb{Z}^n} \Omega_z \end{cases} \quad (33)$$

The proof is as follows:

$$\begin{aligned} \Omega_z &= \Omega \cap S_z \\ &= \{x \in S_z \mid \|x - \check{x}_{LS}\|_{Q_{\hat{a}}}^2 \leq \|x - \check{x}_2\|_{Q_{\hat{a}}}^2 - \mu\} \\ &= \{x \in S_z \mid \|x - \check{x}_{LS}\|_{Q_{\hat{a}}}^2 \leq \|x - u\|_{Q_{\hat{a}}}^2 - \mu, \forall u \in \mathbb{Z}^n \setminus \{\check{x}_{LS}\}\} \\ &= \{x \in S_z \mid \|x - z\|_{Q_{\hat{a}}}^2 \leq \|x - u\|_{Q_{\hat{a}}}^2 - \mu, \forall u \in \mathbb{Z}^n \setminus \{z\}\} \\ &= \{x \in \mathbb{R}^n \mid \|x - z\|_{Q_{\hat{a}}}^2 \leq \|x - u\|_{Q_{\hat{a}}}^2 - \mu, \forall u \in \mathbb{Z}^n \setminus \{z\}\} \\ &= \{x \in \mathbb{R}^n \mid \|y\|_{Q_{\hat{a}}}^2 \leq \|y - v\|_{Q_{\hat{a}}}^2 - \mu, \forall v \in \mathbb{Z}^n \setminus \{0\}, x = y + z\} \\ &= \Omega \cap S_0 + z \\ &= \Omega_0 + z \end{aligned}$$

The first two equalities follow from the definition of the difference test, and the third from $\|\hat{x} - \check{x}_{LS}\|_{Q_{\hat{a}}}^2 \leq \|\hat{x} - \check{x}_2\|_{Q_{\hat{a}}}^2 \leq \|\hat{x} - u\|_{Q_{\hat{a}}}^2, \forall u \in \mathbb{Z}^n \setminus \{\check{x}_{LS}\}$. The fourth equality follows since $\check{x}_{LS} = z$ is equivalent to $x \in S_z$. The fifth equality follows from the fact that $\mu \geq 0$. The last equalities follow from a change of variables and the definition of $\Omega_0 = \Omega \cap S_0$. Finally, note that

$$\bigcup_{z \in \mathbb{Z}^n} \Omega_z = \bigcup_{z \in \mathbb{Z}^n} (\Omega \cap S_z) = \Omega \cap \left(\bigcup_{z \in \mathbb{Z}^n} S_z \right)$$

This ends the proof of Eq. (33).

It has been shown that the acceptance region of the difference test consists of an infinite number of integer-translated copies of a subset of the ILS pull-in region S_0 . And thus, the difference test procedure is an IA estimator with aperture parameter μ .

Difference Test Aperture Pull-in Region

In a way similar to that of the ratio test aperture pull-in region, the geometry of the aperture pull-in region of the difference test can be revealed. From the definition of Ω_0 it follows that

$$\begin{aligned} \Omega_0 : \|x\|_{Q_{\hat{a}}}^2 &\leq \|x - z\|_{Q_{\hat{a}}}^2 - \mu, \quad \forall z \in \mathbb{Z}^n \setminus \{0\}, \quad \mu \geq 0 \\ \iff x^T Q_{\hat{a}}^{-1} z &\leq \frac{1}{2} (\|z\|_{Q_{\hat{a}}}^2 - \mu) \\ \iff \frac{z^T Q_{\hat{a}}^{-1} x}{\|z\|_{Q_{\hat{a}}}^2} &< \frac{\|z\|_{Q_{\hat{a}}}^2 - \mu}{2\|z\|_{Q_{\hat{a}}}^2} \end{aligned} \quad (34)$$

On the left-hand side of Eq. (34) we recognize the orthogonal projection of x onto the direction z . This shows that the aperture pull-in region Ω_0 is determined by intersecting half-spaces that are

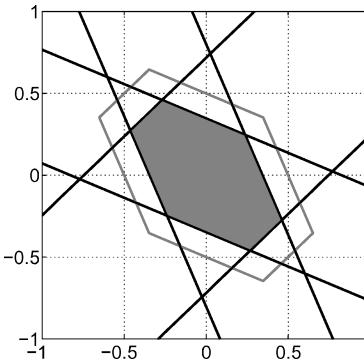


Fig. 3 Geometry of difference test aperture pull-in region (gray) for Q_1 , $\mu = 0.4$.

bounded by planes orthogonal to z and passing through the points $\frac{1}{2}[1 - (\mu/\|z\|_{Q_{\hat{a}}}^2)]z$.

The geometry of Ω_0 is thus very similar to that of the ILS pull-in region S_0 , which is also determined by half-spaces bounded by the planes orthogonal to z ; see Eq. (13). The difference is that these planes pass through the midpoint $\frac{1}{2}z$, whereas in the case of the difference test this point depends on the distance $\|z\|_{Q_{\hat{a}}}^2$ and on μ . This implies that the difference test aperture pull-in region is a down-scaled version of the ILS pull-in region, but that the scaling depends on the direction.

Figure 3 shows a two-dimensional example of the geometry of the aperture pull-in region. The black lines are the planes orthogonal to z and passing through the point $\frac{1}{2}[1 - (\mu/\|z\|_{Q_{\hat{a}}}^2)]z$.

Projector Test

The projector test is defined as^{52,55}

$$\text{Accept } \check{a}_{LS} \text{ iff: } \left| \frac{(\check{a}_2 - \check{a}_{LS})^T Q_{\hat{a}}^{-1} (\hat{a} - \check{a}_{LS})}{\|\check{a}_2 - \check{a}_{LS}\|_{Q_{\hat{a}}}} \right| \leq \mu \quad (35)$$

This test is based on the erroneous assumption that the classical theory of hypothesis testing in linear models is applicable to the current situation. The test is referred to as the projector test because the term on the left-hand side of Eq. (35) equals a projector. It projects $\hat{a} - \check{a}_{LS}$ orthogonally on the direction of $\check{a}_2 - \check{a}_{LS}$, in the metric of $Q_{\hat{a}}$. Furthermore, note that the following is true:

$$\left| \frac{(\check{a}_2 - \check{a}_{LS})^T Q_{\hat{a}}^{-1} (\hat{a} - \check{a}_{LS})}{\|\check{a}_2 - \check{a}_{LS}\|_{Q_{\hat{a}}}} \right| \leq \frac{1}{2} \|\check{a}_2 - \check{a}_{LS}\|_{Q_{\hat{a}}} \quad (36)$$

This inequality implies that the fixed solution will always be accepted if $\mu \geq \frac{1}{2} \|\check{a}_2 - \check{a}_{LS}\|_{Q_{\hat{a}}}$.

Integer Aperture Estimation with the Projector Test

The acceptance region of the projector test is given as

$$\Omega = \{x \in \mathbb{R}^n \mid (\check{x}_{LS} - \check{x}_2)^T Q_{\hat{a}}^{-1} (x - \check{x}_{LS}) \leq \mu \|\check{x}_{LS} - \check{x}_2\|_{Q_{\hat{a}}}^2\} \quad (37)$$

In a way similar to that for the difference test it can be proven that

$$\begin{cases} \Omega_0 = \{x \in S_0 \mid c^T Q_{\hat{a}}^{-1} x \leq \mu \|c\|_{Q_{\hat{a}}}, \quad c = \arg \min_{z \in \mathbb{Z}^n \setminus \{0\}} \|x - z\|_{Q_{\hat{a}}}\} \\ \Omega_z = \Omega_0 + z, \quad \forall z \in \mathbb{Z}^n \\ \Omega = \bigcup_{z \in \mathbb{Z}^n} \Omega_z \end{cases} \quad (38)$$

And thus the procedure underlying the projector test also belongs to the class of IA estimators.

Projector Test Aperture Pull-in Region

From Eq. (38) it follows that Ω_0 is bounded by the planes orthogonal to c and passing through μc , and these planes themselves are bounded by the condition that $c = \arg \min_{z \in \mathbb{Z}^n \setminus \{0\}} \|x - z\|_{Q_{\hat{a}}}$ for all x on the plane.

Figure 4 shows a two-dimensional example of the geometry of the aperture pull-in region Ω_0 . The black lines are the planes orthogonal

to c and passing through μc . For the geometry of Ω_0 , these planes are bounded by the condition that $c = \arg \min_{z \in \mathbb{Z}^n \setminus \{0\}} \|x - z\|_{Q_a}^2$ for all $x \in \Omega_0$. Therefore, sectors within the ILS pull-in region are also shown as alternating gray and white regions, with the sectors containing all x with a certain integer c as second closest integer. The region Ω_0 follows as the region bounded by the intersection of the black lines and the sectors, and by the boundaries of the sectors. This results in a strange nonconvex shape of the aperture pull-in regions in the direction of the vertices of the ILS pull-in region.

Evaluation of the Tests

To illustrate the principle and the properties of the discrimination tests and the optimal test, examples will be given here. These examples are also used to illustrate the differences between the tests.

Simulations were carried out to generate 500,000 samples of the float range/baseline parameters and ambiguities for different models. The first step is to use a random generator to generate n independent samples from the univariate standard normal distribution $N(0, 1)$, and then collect these in a vector x . This vector is transformed by means of $\hat{a} = Gx$, with G equal to the Cholesky factor $Q_a = GG^T$. The result is a sample \hat{a} from $N(0, Q_a)$, and this sample is used as input for integer least-squares estimation. If the output of this estimator equals the null vector, then it is correct; otherwise it is incorrect. This process can be repeated N times, and one can count how many times the null vector is obtained as solution, say N_s times, and how often the outcome equals a nonzero integer vector, say N_f times. The approximations of the success rate and failure rate then follow as

$$P_s = N_s/N, \quad P_f = N_f/N$$

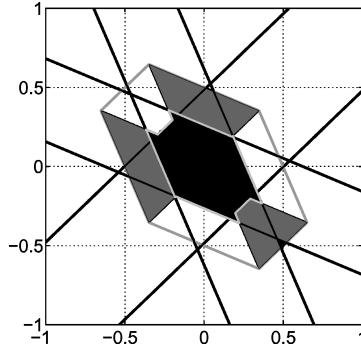


Fig. 4 Geometry of projector test aperture pull-in region (black) for Q_1 , $\mu = 1$.

To get good approximations, the number of samples N must be chosen sufficiently large; see Ref. 22.

We will start with a two-dimensional example. The following variance matrices Q_a are used:

$$Q_1 = \begin{bmatrix} 0.0865 & -0.0364 \\ -0.0364 & 0.0847 \end{bmatrix}, \quad Q_2 = Q_1 + \begin{bmatrix} 0.8 & 0 \\ 0 & 0 \end{bmatrix} \quad (39)$$

The first variance matrix corresponds to a dual-frequency GPS model for one satellite-receiver pair.

Figure 5 shows examples of the aperture pull-in regions for the different tests as obtained with two different failure rates. The aperture pull-in regions of the F -ratio test are not shown, because for the same failure rate the size of the corresponding regions still depends on the sample. Recall that the shape is identical to that of the ratio test aperture pull-in region.

The geometry of the regions obtained with the different discrimination tests was revealed in the previous sections, obviously resulting in quite different shapes. The shape of the optimal aperture pull-in region is determined by the probability distributions of the float ambiguities and the ambiguity residuals.

Especially with Q_1 , the regions for the ratio test and optimal test are very similar. This is an explanation for why the performance of the ratio test is good in many practical situations. The regions are especially different in the direction of the vertices of the ILS pull-in region. That is because in those directions there are two second-best integer vectors such that $\|\hat{a} - \check{a}_2\|_{Q_a}^2 = \|\hat{a} - \check{a}_3\|_{Q_a}^2$, but the squared norm on the right-hand side is not considered by the ratio test, whereas it does affect the size of the optimal test statistic. If the aperture parameter is chosen so that the optimal IA pull-in region just touches the ILS pull-in region, this results in the largest difference between the ratio test and optimal test. The reason is that for a large aperture parameter the optimal IA pull-in region is cut off by the ILS pull-in region. This is not the case for the ratio test pull-in region, because it starts to resemble the shape of the ILS pull-in region for large μ .

For Q_2 , the difference between the tests seems to be much more significant. However, it should be noted that for this example there is a very low probability that \hat{a} will fall into the region where the two aperture pull-in regions do not overlap, so that the performance of the two tests is still quite similar.

With respect to the difference test, conclusions similar to those for the ratio test can be drawn, although the difference from the optimal test is somewhat more significant. This is due to the shape of the aperture pull-in regions. Since these regions are bounded by planes, the pull-in regions are different not only from the optimal

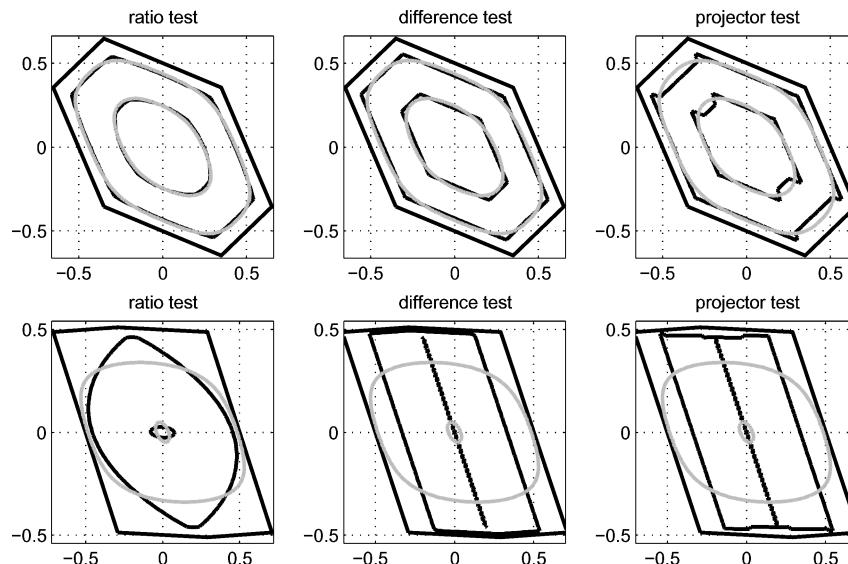


Fig. 5 Examples of two-dimensional aperture pull-in regions of the discrimination tests (black) and optimal test (gray) for two different failure rates. Top: Q_1 , failure rates of 0.005 and 0.06; Bottom: Q_2 , failure rates of 0.005 and 0.4.

IA pull-in regions in the direction of the vertices of the ILS pull-in region, but also in the direction of the adjacent integers.

The difference between the aperture pull-in regions of the optimal test and the projector test is much clearer than for the other two tests, due to the nonconvex shape of the region corresponding to the projector test.

To get an idea of the actual performance of the tests in practice, simulations are used for several geometry-based, dual-frequency GPS models. The GPS constellation is based on the Yuma almanac for GPS week 184 and a cut-off elevation of 15 deg. Undifferenced standard deviations of 30 cm and 3 mm are used for the code and phase observations, respectively. The GPS model is set up for a single epoch for two different times, at which four and six satellites were visible, respectively. The baseline length was chosen to be short to medium length by varying the ionospheric standard deviation σ_I . For the simulation 500,000 samples are used.

Table 1 shows the success rates as obtained with the various integer aperture estimators. Two different failure rates are used with each of the models. The results for Q_1 , corresponding to a geometry-free model for two satellites, are also included. By definition the optimal integer aperture estimator will result in the highest success rates, but it turns out that the discrimination tests result in comparable success rates for most examples. The success rates obtained with the ratio test and the difference test are especially close to those obtained with optimal IA estimation.

Table 2 shows the percentage of decisions identical to the optimal test, based on a comparison of the decisions for each generated sample. This table confirms that indeed the ratio test and difference test perform close to optimally. The *F*-ratio test and the projector test clearly perform less close to optimally.

These results are as could be expected from the comparison of the aperture pull-in regions for the two-dimensional examples.

Table 1 Success rates obtained with our fixed failure rate approach for ratio test (R), *F*-ratio test (F-R), difference test (D), projector test (P), and optimal test (O)^a

No. SV	σ_I [cm]	P_f	Success rates				
			R	F-R	D	P	O
2	0	0.005	0.369	0.318	0.365	0.363	0.369
		0.025	0.637	0.615	0.636	0.633	0.637
4	0	0.005	0.328	0.297	0.320	0.312	0.330
		0.025	0.550	0.536	0.543	0.532	0.552
4	1	0.005	0.026	0.021	0.025	0.025	0.026
		0.025	0.082	0.073	0.079	0.079	0.083
6	1	0.005	0.976	0.976	0.976	0.910	0.976
		0.010	0.987	0.987	0.987	0.978	0.987
6	3	0.005	0.025	0.023	0.027	0.027	0.031
		0.025	0.089	0.084	0.091	0.083	0.098

^aThe first two columns show the number of satellites (No. SV) and ionospheric standard deviation (σ_I) respectively.

Table 2 Percentage of decisions identical to optimal test for ratio test (R), *F*-ratio test (F-R), difference test (D), and projector test (P) if our fixed failure rate approach is used^a

No. SV	σ_I [cm]	P_f	% identical			
			R	F-R	D	P
2	0	0.005	98.8	87.1	97.2	96.7
		0.025	99.2	92.6	97.6	96.5
4	0	0.005	97.8	91.0	95.6	93.9
		0.025	97.9	93.4	95.0	92.6
4	1	0.005	99.6	97.3	98.6	98.7
		0.025	98.6	93.7	96.0	96.1
6	1	0.005	99.6	99.6	99.9	92.0
		0.010	99.9	99.9	99.9	98.7
6	3	0.005	97.0	96.6	97.9	97.6
		0.025	93.8	92.7	94.9	92.7

^aThe first two columns show the number of satellites (No. SV) and ionospheric standard deviation (σ_I) respectively.

Table 3 Critical values obtained with our fixed failure rate approach for ratio test (R), *F*-ratio test (F-R), difference test (D), projector test (P), and optimal test (O)^a

No. SV	σ_I [cm]	P_f	Critical values				
			R	F-R	D	P	O
2	0	0.005	0.106	0.160	7.803	0.888	1.031
		0.025	0.318	0.390	4.379	1.343	1.151
4	0	0.005	0.345	0.449	7.915	1.498	1.040
		0.025	0.551	0.640	4.837	1.848	1.167
4	1	0.005	0.187	0.343	5.670	0.772	1.253
		0.025	0.335	0.501	4.216	1.001	1.445
6	1	0.005	0.854	0.891	2.878	2.690	1.250
		0.010	0.977	0.983	0.440	3.499	1.844
6	3	0.005	0.392	0.609	6.095	0.476	1.206
		0.025	0.542	0.717	4.576	0.835	1.379

^aThe first two columns show the number of satellites (No. SV) and ionospheric standard deviation (σ_I) respectively.

Currently, in practice, the discrimination tests are used in combination with a fixed critical value. The results in Table 3 show that it is not desirable to do so, because it may result in tests that are either too conservative or too optimistic. From the table it follows that to obtain the same failure rate for differently models, the corresponding critical values must be chosen quite differently. For the ratio test a critical value of $\frac{1}{3}$ or 0.5 is often used.^{2,53} The critical values obtained with the fixed failure rates in Table 3 are much larger or much smaller in most cases. Hence, the corresponding failure rates are either smaller or larger, depending on the model at hand. This is also true for the *F*-ratio test, for which a critical value of 0.5 is mostly used.⁵⁴ If the projector test statistic is assumed to have a (truncated) normal distribution, and the critical value is chosen based on a level of significance of 0.05, the corresponding failure rates vary between 0.000 for the model with six satellites and $\sigma_I = 1$ cm, and 0.487 for the model with four satellites and $\sigma_I = 1$ cm. The critical values of the difference test obtained for the model with six satellites and $\sigma_I = 1$ cm are very different from those obtained for the other models. Hence, the difference test applied with a fixed critical value will also result in different failure rates. Obviously, the approach of integer aperture estimation with a fixed failure rate is to be preferred above the traditional approaches.

Conclusions

In this contribution an overall approach is presented for the combined problem of integer estimation and validation. This approach, which is based on the theory of integer aperture estimation, for the first time includes an overall probabilistic evaluation of GNSS ambiguity resolution. With the freedom to set the size and shape of the aperture pull-in region, integer aperture estimation allows the user to get control over the failure rate of ambiguity resolution. It has been shown that the various integer discrimination tests that are in use in practice for integer validation are all members of the same class of integer aperture estimators. This common theoretical basis makes a rigorous comparison of their performances possible. By revealing the geometry of the aperture pull-in regions, further insight was gained into the driving mechanisms of the different discrimination tests.

We also presented the optimal integer aperture estimator. This is the estimator that maximizes the success rate for a user-given failure rate. It was given for an arbitrary PDF of the float solution. For the case of a multivariate normal distribution, the optimal estimator was compared to the discrimination tests, both in a qualitative and in a quantitative manner. In this comparison we identified some pitfalls and erroneous assumptions on which these tests were based in the literature. It also followed from the comparison that the ratio test and the difference test could both give a performance close to optimal. However, this performance is valid when our fixed failure rate approach is used, but not if the classical approach of using a fixed critical value is used.

The fixed-failure-rate approach implies that the aperture parameter, which governs the size of the pull-in region, is automatically

adapted to the strength of the underlying model. If the strength of the underlying model increases, for example, due to the use of more data or more precise data, the size of the aperture pull-in region will automatically increase, thus resulting in more frequent acceptance of the integer solution. Hence, with this approach the time to first fix will be shorter, while at the same time it is guaranteed that the probability of incorrect fixing is below a user-defined threshold.

References

- ¹Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J., *Global Positioning System: Theory and Practice*, 5th ed., Springer-Verlag, Berlin, 2001.
- ²Leick, A., *GPS Satellite Surveying*, 3rd ed., Wiley, New York, 2003.
- ³Parkinson, B. W., and Spilker, J. J. (eds.), *Global Positioning System: Theory and Applications*, Vols. 1 and 2, Progress in Aeronautics and Astronautics, Vol. 164, AIAA, Washington, DC, 1996.
- ⁴Strang, G., and Borre, K., *Linear Algebra, Geodesy, and GPS*, Wellesley-Cambridge Press, Wellesley, MA, 1997.
- ⁵Teunissen, P. J. G., and Kleusberg, A., *GPS for Geodesy*, 2nd ed., Springer-Verlag, Berlin, 1998.
- ⁶Kim, U. S., De Lorenzo, D. S., Akos, D., Gautier, J., Enge, P., and Orr, J., "Precise Phase Calibration of a Controlled Reception Pattern GPS Antenna for JPALS," *Proceedings of IEEE PLANS'04*, Inst. of Electrical and Electronics Engineers, New York, 2004, pp. 478–485.
- ⁷De Lorenzo, D. S., Alban, S., Gautier, J., Enge, P., and Akos, D., "GPS Attitude Determination for a JPALS Testbed: Integer Initialization and Testing," *Proceedings of IEEE PLANS'04*, Inst. of Electrical and Electronics Engineers, New York, 2004, pp. 762–770.
- ⁸Crassidis, J. L., Markley, F. L., and Lightsey, E. G., "Global Positioning System Integer Ambiguity Resolution Without Attitude Knowledge," *Journal of Guidance, Control, and Dynamics*, Vol. 22, No. 2, 1999, pp. 212–218.
- ⁹Peng, H. M., Chang, F. R., and Wang, L. S., "Attitude Determination Using GPS Carrier Phase and Compass Data," *Proceedings of ION NTM-1999*, The Inst. of Navigation, Fairfax, VA, 1999, pp. 727–732.
- ¹⁰Bell, T., "Global Positioning System-Based Attitude Determination and the Orthogonal Procrustes Problem," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 5, 2003, pp. 820–822.
- ¹¹Park, C. S., and Teunissen, P. J. G., "A New Carrier Phase Ambiguity Estimation for GNSS Attitude Determination Systems," *Proceedings of International Symposium on GPS/GNSS*, Japan GPS Council, Tokyo, 2003, pp. 249–255.
- ¹²Leung, S., and Montenbruck, O., "Real-Time Navigation of Formation Flying Spacecraft Using GPS Measurements," *Journal of Guidance, Control, and Dynamics*, Vol. 28, No. 2, 2005, pp. 226–235.
- ¹³Kroes, R., Montenbruck, O., Bertiger, W., and Visser, P., "Precise GRACE Baseline Determination Using GPS," *GPS Solutions*, Vol. 9, No. 1, 2005, pp. 21–31.
- ¹⁴Betti, B., Crespi, M., and Sansò, F., "A Geometric Illustration of Ambiguity Resolution in GPS Theory and a Bayesian Approach," *Manuscripta Geodaetica*, Vol. 18, No. 6, 1993, pp. 317–330.
- ¹⁵Gundlich, B., "Statistische Untersuchung Ganzzahliger und Reellweger Unbekannter Parameter im GPS-Modell," Ph.D. Dissertations, DGK, Reihe C, No. 549, Munich, 2002.
- ¹⁶Gundlich, B., and Koch, K. R., "Confidence Regions for GPS Baselines by Bayesian Statistics," *Journal of Geodesy*, Vol. 76, No. 1, 2002, pp. 55–62.
- ¹⁷Gundlich, B., and Teunissen, P. J. G., "Multiple Models: Fixed, Switching and Interacting," *V. Hotine-Marussi Symposium on Mathematical Geodesy*, edited by F. Sansò, IAG Symposia, No. 127, Springer-Verlag, Berlin, 2004.
- ¹⁸De Lacy, M. C., Sansò, F., Rodriguez-Caderot, G., and Gil, A. J., "The Bayesian Approach Applied to GPS Ambiguity Resolution. A Mixture Model for the Discrete–Real Ambiguities Alternative," *Journal of Geodesy*, Vol. 76, No. 2, 2002, pp. 82–94.
- ¹⁹Teunissen, P. J. G., "Theory of Integer Equivariant Estimation with Application to GNSS," *Journal of Geodesy*, Vol. 77, No. 7–8, 2003, pp. 402–410.
- ²⁰Teunissen, P. J. G., "Integer Aperture GNSS Ambiguity Resolution," *Artificial Satellites*, Vol. 38, No. 3, 2003, pp. 79–88.
- ²¹Jonkman, N. F., "Integer GPS Ambiguity Estimation Without the Receiver–Satellite Geometry," LGR Series, No. 18, Delft Geodetic Computing Centre, Delft Univ. of Technology, 1998.
- ²²Teunissen, P. J. G., "On the Integer Normal Distribution of the GPS Ambiguities," *Artificial Satellites*, Vol. 33, No. 2, 1998, pp. 49–64.
- ²³Teunissen, P. J. G., "An Optimality Property of the Integer Least-Squares Estimator," *Journal of Geodesy*, Vol. 73, No. 11, 1999, pp. 587–593.
- ²⁴Teunissen, P. J. G., "The Parameter Distributions of the Integer GPS Model," *Journal of Geodesy*, Vol. 76, No. 1, 2002, pp. 41–48.
- ²⁵Teunissen, P. J. G., "A Theorem on Maximizing the Probability of Correct Integer Estimation," *Artificial Satellites*, Vol. 34, No. 1, 1999, pp. 3–9.
- ²⁶Teunissen, P. J. G., "Least Squares Estimation of the Integer GPS Ambiguities," Invited lecture, Section IV, *Theory and Methodology, IAG General Meeting*, International Association of Geodesy, Budapest, 1993.
- ²⁷Teunissen, P. J. G., "The Least-Squares Ambiguity Decorrelation Adjustment: A Method for Fast GPS Integer Ambiguity Estimation," *Journal of Geodesy*, Vol. 70, No. 1–2, 1995, pp. 65–82.
- ²⁸De Jonge, P. J., and Tiberius, C. C. J. M., "The LAMBDA Method for Integer Ambiguity Estimation: Implementation Aspects," LGR Series, No. 12, Delft Geodetic Computing Centre, Delft Univ. of Technology, 1996.
- ²⁹Boon, F., and Ambrosius, B., "Results of Real-Time Applications of the LAMBDA Method in GPS-Based Aircraft Landings," *Proceedings of the International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation*, Dept. of Geomatics Engineering, Univ. of Calgary, Calgary, Canada, 1997, pp. 339–345.
- ³⁰Boon, F., De Jonge, P. J., and Tiberius, C. C. J. M., "Precise Aircraft Positioning by Fast Ambiguity Resolution Using Improved Troposphere Modelling," *Proceedings of ION GPS-1997*, The Inst. of Navigation, Fairfax, VA, 1997, pp. 1877–1884.
- ³¹Chaitin-Chatelin, F., Dallakyan, and Frayssé, V., "On the GPS Carrier Phase Ambiguity Resolution. The LAMBDA Method: An Analysis of Speed, Efficiency and Numerical Robustness," TR IR/PA/99/23, Cerfacs, 1999.
- ³²Cox, D. B., and Brading, J. D. W., "Integration of LAMBDA Ambiguity Resolution with Kalman Filter for Relative Navigation of Spacecraft," *Navigation*, Vol. 47, No. 3, 2000, pp. 205–210.
- ³³De Jonge, P. J., and Tiberius, C. C. J. M., "Integer Ambiguity Estimation with the LAMBDA Method," *IAG Symposia No. 115, GPS Trends in Terrestrial, Airborne and Spaceborne Applications*, edited by G. Beutler, Springer-Verlag, New York, 1996, pp. 280–284.
- ³⁴Han, S., "Ambiguity Resolution Techniques Using Integer Least-Squares Estimation for Rapid Static or Kinematic Positioning," *Proceedings of Symposium Satellite Navigation Technology and Beyond*, Paper 34, Space Centre for Satellite Navigation, Queensland Univ. of Technology, Brisbane, Australia, 1995, p. 10.
- ³⁵Kiamehr, R., "Comparison of Ambiguity Resolution Using LAMBDA and KTH Method," *Proceedings of 2003 International Symposium on GPS/GNSS*, Japan GPS Council, Tokyo, 2003.
- ³⁶Lee, H. K., Wang, J., and Rizos, C., "An Integer Ambiguity Resolution Procedure for GPS/Pseudolite/INS Integration," *Journal of Geodesy*, Vol. 79, No. 4–5, 2005, pp. 242–255.
- ³⁷Milbert, D., "Influence of Pseudorange Accuracy on Phase Ambiguity Resolution in Various GPS Modernization Scenarios," *Navigation*, Vol. 52, No. 1, 2005, pp. 29–38.
- ³⁸Tiberius, C. C. J. M., and De Jonge, P. J., "Fast Positioning Using the LAMBDA Method," *Proceedings of DSNS'95*, Paper 30, The Nordic Inst. of Navigation, Oslo, 1995.
- ³⁹Tiberius, C. C. J. M., Teunissen, P. J. G., and De Jonge, P. J., "Kinematic GPS: Performance and Quality Control," *Proceedings of KIS'97*, Dept. of Geomatics Engineering Univ. of Calgary, Calgary, Canada, 1997, pp. 289–299.
- ⁴⁰Weiming, T., "Ambiguity Resolution of Single Epoch Single Frequency Data with Baseline Length Constraints Using LAMBDA Method," *Proceedings of 2004 International Symposium on GPS/GNSS*, Univ. of New South Wales, Sydney, Australia, 2004.
- ⁴¹Zhou, Y., and Liu, J., "Another Form for LAMBDA Method," *Geospatial Information Science*, Vol. 6, No. 3, 2003, pp. 66–70.
- ⁴²De Jonge, P. J., Tiberius, C. C. J. M., and Teunissen, P. J. G., "Computational Aspects of the LAMBDA Method for GPS Ambiguity Resolution," *Proceedings of ION GPS-1996*, The Inst. of Navigation, Fairfax, VA, 1996, pp. 935–944.
- ⁴³Xu, P., "Random simulation and GPS Decorrelation," *Journal of Geodesy*, Vol. 75, No. 7, 2001, pp. 408–423.
- ⁴⁴Svendsen, J. G. G., "Some Properties of Decorrelation Techniques in the Ambiguity Space," *GPS Solutions*, [online Journal], Vol. 10, No. 1, 2006, pp. 40–44, DOI: 10.1007/s10291-005-0004-6, 2005.
- ⁴⁵Lou, L., and Grafarend, E., "GPS Integer Ambiguity Resolution by Various Decorrelation Methods," *Zeitschrift für Vermessungswesen*, Vol. 128, No. 3, 2003, pp. 203–210.
- ⁴⁶Chang, X. W., Yang, X., and Zhou, T., "A Modified LAMBDA Method for Integer Least-Squares Estimation," *Journal of Geodesy*, Vol. 79, No. 9, 2005, pp. 552–565.
- ⁴⁷Verhagen, S., "On the Reliability of Integer Ambiguity Resolution," *Navigation*, Vol. 52, No. 2, 2005, pp. 99–110.
- ⁴⁸Abidin, H. A., "Computational and Geometrical Aspects of On-the-Fly Ambiguity Resolution," Ph.D. Dissertation, TR 104, Univ. of New Brunswick, 1993.
- ⁴⁹Chen, Y., "An Approach to Validate the Resolved Ambiguities in GPS Rapid Positioning," *Proceedings of the International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation*, Dept. of Geomatics Engineering, Univ. of Calgary, Calgary, Canada, 1997, pp. 301–304.

- ⁵⁰Euler, H. J., and Schaffrin, B., "On a Measure for the Discernibility Between Different Ambiguity Solutions in the Static–Kinematic GPS-Mode," IAG Symposia No. 107, *Kinematic Systems in Geodesy, Surveying, and Remote Sensing*, Springer-Verlag, New York, 1991, pp. 285–295.
- ⁵¹Frei, E., and Beutler, G., "Rapid Static Positioning Based on the Fast Ambiguity Resolution Approach FARA: Theory and First Results," *Manuscripta Geodaetica*, Vol. 15, No. 6, 1990, pp. 325–356.
- ⁵²Han, S., "Quality Control Issues Relating to Instantaneous Ambiguity Resolution for Real-Time GPS Kinematic Positioning," *Journal of Geodesy*, Vol. 71, No. 6, 1997, pp. 351–361.
- ⁵³Han, S., and Rizos, C., "Validation and Rejection Criteria for Integer Least-Squares Estimation," *Survey Review*, Vol. 33, No. 260, 1996, pp. 375–382.
- ⁵⁴Landau, H., and Euler, H. J., "On-the-Fly Ambiguity Resolution for Precise Differential Positioning," *Proceedings of ION GPS-1992*, The Inst. of Navigation, Fairfax, VA, 1992, pp. 607–613.
- ⁵⁵Wang, J., Stewart, M. P., and Tsakiri, M., "A Discrimination Test Procedure for Ambiguity Resolution On-the-Fly," *Journal of Geodesy*, Vol. 72, No. 11, 1998, pp. 644–653.
- ⁵⁶Verhagen, S., "Integer Ambiguity Validation: An Open Problem?" *GPS Solutions*, Vol. 8, No. 1, 2004, pp. 36–43.
- ⁵⁷Verhagen, S., and Teunissen, P. J. G., "On the Probability Density Function of the GNSS Ambiguity Residuals," *GPS Solutions* [online journal], Vol. 10, No. 1, 2006, pp. 21–28, DOI: 10.1007/s10291-005-0148-4, 2005.
- ⁵⁸Teunissen, P. J. G., "GNSS Ambiguity Resolution with Optimally Controlled Failure-Rate," *Artificial Satellites*, Vol. 40, No. 4, 2005, pp. 219–227.
- ⁵⁹Teunissen, P. J. G., "Penalized GNSS Ambiguity Resolution," *Journal of Geodesy*, Vol. 78, No. 4–5, 2004, pp. 235–244.
- ⁶⁰Teunissen, P. J. G., and Verhagen, S., "On the Foundation of the Popular Ratio Test for GNSS Ambiguity Resolution," *Proceedings of ION GNSS-2004*, The Inst. of Navigation, Fairfax, VA, 2004, pp. 2529–2540.
- ⁶¹Teunissen, P. J. G., "GPS Carrier Phase Ambiguity Fixing Concepts," *GPS for Geodesy*, edited by P. J. G. Teunissen and A. Kleusberg, Springer-Verlag, Berlin, 1998, Chap. 8.