Journal of Geodesy (2002) 76: 82–94
DOI 10.1007/s00190-001-0226-5

Journal of Geodesy

The Bayesian approach applied to GPS ambiguity resolution. A mixture model for the discrete–real ambiguities alternative

M. C. de Lacy[1], F. Sansò[1], G. Rodriguez-Caderot[2], A. J. Gil[3]

[1] Facoltà di Ingegneria di Como, Politecnico di Milano, Italy, e-mail: clara@torno.ing.unico.it; Tel.: +39-0313327521
Fax: +39-0313327519
[2] Sección Departamental Astrononomia y Geodesia, Facultad de CC. Matematicas, Universidad Complutense de Madrid, Spain
e-mail: grc@mat.vcm.es, Tel.: +34-913944588 Fax: +34-913944607
[3] Departamento Ingenieria Geodesica, Cartografica y Fotogrametria, Universidad de Jaen, Spain, e-mail: ajgil@ujaen.es
Tel.: +34-953002467 Fax: +34-953012343

Received: 29 March 2001 / Accepted: 3 September 2001

Abstract. The problem of phase ambiguity resolution in global positioning system (GPS) theory is considered. The Bayesian approach is applied to this problem and, using Monte Carlo simulation to search over the integer candidates, a practical expression for the Bayesian estimator is obtained. The analysis of the integer grid points inside the search ellipsoid and their evolution with time, while measurements are accumulated, leads to the development of a Bayesian theory based on a mathematical mixture model for the ambiguity.

Key words: Mixture Model – Bayesian Approach – Monte Carlo Simulation – Ambiguity Resolution

1 Introduction

In the mathematical model of double-difference (DD) global positioning system (GPS) observation there are two groups of unknowns: baseline coordinates and initial ambiguities. According to this model the ambiguity vector should have integer components. Then the problem is to find the correct estimate for the ambiguity with integer components and the corresponding estimate for the coordinates. A usual approach to the ambiguity resolution is first to estimate coordinates and ambiguities with no constraints, i.e. floating solution, and then to apply some 'suitable' testing procedure to decide whether the floating ambiguity vector is compatible or not with one, and only one, integer estimator; in the affirmative case the estimate of the coordinates is repeated by fixing the ambiguity to the value of the integer estimator. The methods Ratio Test, FARA (fast ambiguity resolution approach; Beutler and Frei 1990) and LAMBDA (least-squares ambiguity decorrelation adjustment; Teunissen and Kleusberg 1998) use this kind of approach. First Blewitt (1989), then Betti et al. (1993), and more recently Gundlich and Koch (submitted) proposed, as an alternative theory, the Bayesian approach whose main characteristics are the following:

1. It yields a posterior distribution for all variables, discrete and continuous, conditional to the observed quantities.
2. It takes into account the information contained in the full covariance matrix derived of the least-squares (LS) adjustment, i.e. the covariance matrix of the floating solution.
3. It does not need to resolve the ambiguity vector.
4. It provides a solution without any further adjustment.

In this paper we continue this work, producing a practical expression for the Bayesian estimator using the Monte Carlo simulation to define the search ellipsoid in which the integer candidates are contained. The Monte Carlo simulation will allow us to obtain approximated expressions for the Bayesian solution and its covariance matrix. By studying the information and the evolution with time of the position and shape of the search ellipsoid in the bias vector space, we are urged to develop a novel enlarged model where a new variable is introduced, labeling the case under analysis as an ordinary case where the bias vector has integer components, or as one of the special cases where the presence of other biases in the observations prevents us from fixing them to integer values. Of course the two cases will be discriminated only at a certain moment in time when the accumulated information becomes sufficient for that purpose.

It seems useful here to clarify that indeed the Bayesian approach is not the only consistent probabilistic set-up for the estimation of position and integer ambiguity parameters. On the contrary, the classical LS theory for normal variates has been suitably developed into a new theory for integer LS models (cf. Teunissen 1999a, b), in the framework of which a full description of the statistical behaviour of the estimators is available. In other words,

Correspondence to: M. C. de Lacy

the time when a testing procedure was used as estimation tool (i.e. the selected integer ambiguity was fixed, disappearing from the vector of the unknowns) has passed, and the frequentist approach can offer nowadays a consistent picture of the estimation problem as the Bayesian approach can in principle do. This suggests the interesting scientific question of how they compare to each other; however, this is not the aim of the present paper, which is considered by the authors as just a contribution to the development of the 'Bayesian way' which still seems to be incompletely mature. At the same time, it is because of the flexibility of Bayesian theory that we come across the idea that the alternative integer ambiguity (general real bias vector) could be *included* into the estimation model and we thought that it was useful to offer the argument to the discussion of the scientific community. Therefore in the present contribution we concentrate on the Bayesian approach applied to the ambiguity resolution problem using the Monte Carlo simulation to define a search strategy over integer candidates. The implementation of the Monte Carlo method and analysis of the first results, in comparison with the FARA method, are presented in Sect. 3. The Bayesian theory for the mathematical mixture model is explained in Sect. 4. This model is applied to two numerical examples.

2 The Bayesian approach

GPS observational models, in linearized form, can be written as

$$\mathbf{Y} = B_1\mathbf{r} + B_2\mathbf{b} + \mathbf{v} \tag{1}$$

where

- $\mathbf{Y}$ = vector of observables, in this case double differences
- $\mathbf{r}$ = continuous parameter corresponding to station position corrections, troposphere parameters, etc.
- $\mathbf{b}$ = vector of biases which in agreement with model should have integer components (in units of wavelengths)
- B_1 and B_2 = design matrices for parameters $\mathbf{r}$ and $\mathbf{b}$ derived from linearized observation equations
- $\mathbf{v}$ = vector of model noise, independent of $\mathbf{r}$ and $\mathbf{b}$.

According to Bayes' theorem, if an observation model has the observable vector $\mathbf{Y}$, the unknown parameters $(\mathbf{r}, \mathbf{b})$, the prior distribution $\tilde{p}(\mathbf{r}, \mathbf{b})$ on the parameter space and the likelihood function $L(\mathbf{Y}|\mathbf{r}, \mathbf{b})$ describing the observation process, then a new posterior distribution is generated in the parameter space with the rule (Box and Tiao 1992)

$$p(\mathbf{r}, \mathbf{b}|\mathbf{Y}) = K L(\mathbf{Y}|\mathbf{r}, \mathbf{b})\tilde{p}(\mathbf{r}, \mathbf{b}) \tag{2}$$

where the 'constant' K is such that Eq. (2) satisfies a standard normalization relation in $(\mathbf{r}, \mathbf{b})$ for fixed $\mathbf{Y}$. It is important to note that Eq. (2) says that observations $\mathbf{Y}$ modify the knowledge of the prior distribution of $(\mathbf{r}, \mathbf{b})$ into the posterior distribution $p(\mathbf{r}, \mathbf{b} \mid \mathbf{Y})$. Using the proprieties of the Bayesian theory, we interpret p as a probability function with respect to the discrete parameter $\mathbf{b}$ and as a probability density function for the continuous parameter $\mathbf{r}$. So Eq. (2) can be rewritten as

$$p(\mathbf{r}, \mathbf{b}|\mathbf{Y}) = \frac{L(\mathbf{Y}|\mathbf{r}, \mathbf{b})\tilde{p}(\mathbf{r}, \mathbf{b})}{\sum_{\mathbf{b}} \int L(\mathbf{Y}|\mathbf{r}, \mathbf{b})\tilde{p}(\mathbf{r}, \mathbf{b})\mathrm{d}(\mathbf{r})} \tag{3}$$

where the summation is over the discrete values of $\mathbf{b}$ and the integral is calculated over R^n.

From the posterior distribution we can draw a synthetic knowledge of the variables of interest, e.g. a location and a dispersion parameter. Concentrating for the moment on the location, this can be either the posterior mode (MAP) or the posterior mean, i.e.

$$\mathbf{r}_B = E[\mathbf{r}|\mathbf{Y}] = \frac{\sum_{\mathbf{b}} \int \mathbf{r} L(\mathbf{Y}|\mathbf{r}, \mathbf{b})\tilde{p}(\mathbf{r}, \mathbf{b})\mathrm{d}(\mathbf{r})}{\sum_{\mathbf{b}} \int L(\mathbf{Y}|\mathbf{r}, \mathbf{b})\tilde{p}(\mathbf{r}, \mathbf{b})\mathrm{d}(\mathbf{r})} \tag{4}$$

It is important to note the following:

1. To derive information on $\mathbf{r}$ it is not necessary to resolve the ambiguities; it is enough to sum over all possible ambiguities with proper weights directly derived from the likelihood function.
2. In order to perform any Bayesian inference it is always necessary to know the prior of $(\mathbf{r}, \mathbf{b})$. For the sake of simplicity in the subsequent computations we will assume that we have no prior information on $(\mathbf{r}, \mathbf{b})$ and that both priors are independent. Then Eq. (2) reads

   $$p(\mathbf{r}, \mathbf{b}|\mathbf{Y}) = KL(\mathbf{Y}|\mathbf{r}, \mathbf{b})\tilde{p}(\mathbf{r})\tilde{p}(\mathbf{b}) \tag{5}$$

   where

   $$\begin{aligned} &\tilde{p}(\mathbf{r}) = c \quad \text{improper} \\ &\tilde{p}(\mathbf{b}) = c\sum_{\mathbf{b}_I} \delta(\mathbf{b} - \mathbf{b}_I) \quad \text{improper} \end{aligned} \tag{6}$$

   where c is a constant, $\mathbf{b}_I$ is the integer ambiguity vector and δ represents the Dirac distribution.

In order to obtain a practical expression for Eq. (4) we rewrite the model of our estimation problem as

$$\mathbf{Y} = B_1\mathbf{r} + B_2\mathbf{b} + \mathbf{v} = A\mathbf{x} + \mathbf{v} \tag{7}$$

where
A is the design matrix and $\mathbf{x} = \binom{\mathbf{r}}{\mathbf{b}}$ is the parameter vector to be estimated. We will soon need the symbols $\hat{\mathbf{x}} = \binom{\hat{\mathbf{r}}}{\hat{\mathbf{b}}}$, the floating solution of the coordinates and ambiguities computed from a standard LS adjustment, and

$$C_{\mathbf{xx}} = \sigma_0^2 \begin{pmatrix} N_{GG} & N_{Gb} \\ N_{bG} & N_{bb} \end{pmatrix}^{-1}$$

the variance–covariance matrix of this adjustment.

Assuming $\mathbf{Y} \approx N(\mathbf{y}, C_{YY})$ to be a normal variate and taking into account Eq. (6), Eq. (5) can be expressed as

$$p(\mathbf{r}, \mathbf{b}|\mathbf{Y}) \propto \mathrm{e}^{-\frac{1}{2}(\mathbf{Y}-A\mathbf{x})^T C_{\mathbf{xx}}^{-1}(\mathbf{Y}-A\mathbf{x})} \sum_{\mathbf{b}_I} \delta(\mathbf{b} - \mathbf{b}_I) \tag{8}$$

and, exploiting the ordinary decomposition of the quadratic form at exponent, Eq. (8) is equivalent to

$$p(\mathbf{r}, \mathbf{b}|\mathbf{Y}) \propto \mathrm{e}^{-\frac{1}{2}(\mathbf{Y}-\hat{\mathbf{Y}})^T C_{\mathbf{xx}}^{-1}(\mathbf{Y}-\hat{\mathbf{Y}})} \cdot \mathrm{e}^{-\frac{1}{2}(\mathbf{x}-\hat{\mathbf{x}})^T A^T C_{\mathbf{xx}}^{-1} A(\mathbf{x}-\hat{\mathbf{x}})} \sum_{\mathbf{b}_I} \delta(\mathbf{b} - \mathbf{b}_I) \tag{9}$$

Let us introduce the shifted variables

$$\boldsymbol{\delta}\mathbf{r} = \mathbf{r} - \hat{\mathbf{r}}$$

$$\boldsymbol{\beta} = \mathbf{b} - \hat{\mathbf{b}}$$

$$\boldsymbol{\beta}_I = \mathbf{b}_I - \hat{\mathbf{b}}$$

$$\boldsymbol{\beta} - \boldsymbol{\beta}_I = \mathbf{b} - \mathbf{b}_I$$

Then, keeping only the part depending on the sufficient statistic $(\hat{\mathbf{r}}, \hat{\mathbf{b}})$, the other part is absorbed into the constant as is customary for functions of $\mathbf{Y}$ only in Bayesian theory, and Eq. (9) reads

$$\begin{aligned} p(\mathbf{r}, \mathbf{b}|\mathbf{Y}) &\propto \sum_{\mathbf{b}_I} \mathrm{e}^{-\frac{1}{2\sigma_0^2}\left[(\mathbf{r}-\hat{\mathbf{r}})^T (\mathbf{b}_I-\hat{\mathbf{b}})^T\right] \begin{pmatrix} N_{GG} & N_{Gb} \\ N_{bG} & N_{bb} \end{pmatrix} \begin{pmatrix} \mathbf{r}-\hat{\mathbf{r}} \\ \mathbf{b}_I-\hat{\mathbf{b}} \end{pmatrix}} \\ &\quad \times \delta(\mathbf{b} - \mathbf{b}_I) \\ &= \sum_{\boldsymbol{\beta}_I} \mathrm{e}^{-\frac{1}{2\sigma_0^2}\Psi(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}_I)} \cdot \delta(\mathbf{b} - \mathbf{b}_I) \end{aligned} \tag{10}$$

where

$$\Psi(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}_I) = \left[(\mathbf{r}-\hat{\mathbf{r}})^T (\mathbf{b}_I-\hat{\mathbf{b}})^T\right] \begin{pmatrix} N_{GG} & N_{Gb} \\ N_{bG} & N_{bb} \end{pmatrix} \begin{pmatrix} \mathbf{r}-\hat{\mathbf{r}} \\ \mathbf{b}_I-\hat{\mathbf{b}} \end{pmatrix} \tag{11}$$

As it is shown in Appendix 1, we obtain the Bayesian estimator

$$\boldsymbol{\delta}\mathbf{r}_B = E\{\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}\} = -N_{GG}^{-1} N_{Gb} \frac{\sum_{\boldsymbol{\beta}_I} \boldsymbol{\beta}_I \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I}}{\sum_{\boldsymbol{\beta}_I} \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I}} \tag{12}$$

It is important to stress the following.

1. N_{GG}^{-1}, N_{Gb} are known since they have been calculated in the LS adjustment to obtain the floating solution and they need to be used only once in Eq. (12).
2. In order to calculate Eq. (12) it is necessary to know

$$\boldsymbol{\mu}_{\boldsymbol{\beta}_I} = \sum_{\boldsymbol{\beta}_I} \boldsymbol{\beta}_I p(\boldsymbol{\beta}_I) \tag{13}$$

where

$$p(\boldsymbol{\beta}_I) = p(\boldsymbol{\beta}_I|\mathbf{Y}) \propto \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I} \tag{14}$$

So, the problem is now reduced from the numerical point of view to just that of averaging over integer candidates, weighting them with their corresponding probability. Indeed Eq. (13) is an infinite sum and cannot be computed exactly, but we have to truncate the summation; this is a quite substantial but unavoidable drawback intrinsic to the problem. As a matter of fact, also in the frequentist approach a complete description of the estimator would imply projecting a full normal distribution in the space of the vector $\mathbf{b}$ onto the lattice of points with integer coordinates. Given the shape of the distribution $p(\boldsymbol{\beta}_I)$ it is only natural to try to use a summation set or characteristic ellipsoidal region

$$\mathscr{E} = \left\{\boldsymbol{\beta}_I = \mathbf{b}_I - \hat{\mathbf{b}}, \mathbf{b}_I \text{ integer}; \boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I \leq \sigma_0^2 \chi^2_{\dim\boldsymbol{\beta}}\right\} \tag{15}$$

We will call O (for outer region) the complementary of this region in lattice of $\boldsymbol{\beta}$.

When N_r is small enough, there are many integer grid points in $\mathscr{E}$; in this case we propose that a Monte Carlo simulation method be used to compute Eq. (13) and so Eq. (12). To this aim we draw a sample of values $\mathbf{b}_i$ distributed as $\mathbf{b}_i \approx N(\hat{\mathbf{b}}, \sigma_0^2 N_r^{-1})$, by exploiting the relation (Jennings 1977, Press et al. 1992; Sheldon 1997)

$$\mathbf{b}_i = \hat{\mathbf{b}} + T^T \mathbf{z}_i \tag{16}$$

where
$\hat{\mathbf{b}}$ is the vector of floating ambiguities
T is the Cholesky factor of $C_{\boldsymbol{\beta\beta}} = \sigma_0^2 N_r^{-1} = T^T T$
$\mathbf{z}$ is a vector of independent random numbers distributed as a normal N(0, I).

Rounding every vector $\mathbf{b}_i$ to the vector of the nearest integer components, $\mathbf{b}_i^I$, and using the strong law of large numbers, Eq. (13) can be approximated as

$$\boldsymbol{\mu}_{\boldsymbol{\beta}_I} \approx \frac{1}{k} \sum_{i=1}^{k} (\mathbf{b}_i^I - \hat{\mathbf{b}}) \tag{17}$$

This approximation is introduced into Eq. (12) in order to obtain the Bayesian correction to be applied to the floating solution.

The same line of thought can be used to compute an approximation to the covariance matrix of the Bayesian solution. The covariance matrix of the distribution of Eq. (3) is given by

$$C_{\boldsymbol{\delta}\mathbf{r}_B\boldsymbol{\delta}\mathbf{r}_B} = \mathrm{Cov}(\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}) = E\{(\boldsymbol{\delta}\mathbf{r}\boldsymbol{\delta}\mathbf{r}^T - \boldsymbol{\mu}_{\delta r|Y}\boldsymbol{\mu}_{\delta r|Y}^T)|\mathbf{Y}\} \tag{18}$$

where $\boldsymbol{\mu}_{\delta r|Y}$ is the mean of the $\boldsymbol{\delta}\mathbf{r}$ estimations conditional to the observations and, as it is proved in Appendix 1, the exact expression of this covariance matrix is

$$C_{\boldsymbol{\delta}\mathbf{r}_B\boldsymbol{\delta}\mathbf{r}_B} = \sigma_0^2 N_{GG}^{-1} + N_{GG}^{-1} N_{Gb} C_{\boldsymbol{\beta}_I\boldsymbol{\beta}_I} N_{bG} N_{GG}^{-1} \tag{19}$$

We can again apply the Monte Carlo method to obtain the approximated covariance matrix of $\boldsymbol{\beta}_I$

$$C_{\boldsymbol{\beta}_I\boldsymbol{\beta}_I} = C_{\mathbf{b}_I-\hat{\mathbf{b}},\mathbf{b}_I-\hat{\mathbf{b}}} \approx \frac{1}{k} \sum_{\mathbf{b}_i} (\mathbf{b}_i^I - \hat{\mathbf{b}})(\mathbf{b}_i^I - \hat{\mathbf{b}})^T - \boldsymbol{\mu}_{\boldsymbol{\beta}}\boldsymbol{\mu}_{\boldsymbol{\beta}}^T \tag{20}$$

where μ_β is given by Eq. (17); the result is then substituted back in Eq. (19).

Finally, three points are worthy of comment:

1. The Bayesian solution depends in general on the variable σ_0^2. When this is not known the theory has to be modified as shown, for instance, in Betti et al. (1993), although the direct use of the floating estimate σ_0^2 in the previous computations seems to supply a suitable solution.
2. The Bayesian approach in its pure form considers as outcome of the 'estimation' procedure the posterior distribution of the parameters of interest (Box and Tiao 1992). It is only to the extent of synthesizing the information of this distribution that one is entitled to reduce it to the knowledge of a few parameters. It is customary in Bayesian literature to use the MAP point together with the average of second derivatives of the logarithm of the posterior (information matrix) with respect to parameters. Here we have chosen the more traditional posterior mean and covariance because we are aware that the posterior of $\mathbf{\delta r}$ will be multimodal and, moreover, the high-probability regions around the maximo may still contain a sensible probability, if the observations time is short (Teunissen 1999a; Gundlich 2001). In this situation we feel that mean and covariance are still more descriptive of the coarse, overall spread of probability, because of the universal validity of the Tchebychef theorem.
3. The Monte Carlo approach presented here produces two different types of approximations. One is classical for Monte Carlo methods, yielding, for example in Eq. (17) an error covariance

$$\begin{aligned} \boldsymbol{\eta} &= \frac{1}{k}\sum_i \boldsymbol{\beta}_i^I - \sum \boldsymbol{\beta}_I p(\boldsymbol{\beta}_I) \\ C_{\boldsymbol{\eta\eta}} &\approx \frac{\sigma_0^I N_r^{-1}}{k} \end{aligned} \tag{21}$$

Since typically the diagonal of $\sigma_0^I N_r^{-1}$ is between 1 and 0.01 cycles ($\approx$ 20 cm to 2 mm), as soon as k is larger than 10^3 the approximation can be considered as effective. The other approximation is by the mechanism of sampling a normal and then shifting the sample vector $\mathbf{b}_i$ to the nearest neighbour $\mathbf{b}_i^I$. This assigns to the final lattice point $\mathbf{b}_i^I$ the same probability of the normal distribution $N(\hat{\mathbf{b}}, \sigma_0^2 N_r^{-1})$ integrated over the square block of points which are nearest to it. This approximation is more effective if the characteristic $\mathscr{E}$ is large with respect to the above blocks (see the discussion in the next section) and possible in the spherical shape. For this reason we expect the Monte Carlo approach to work much better if first the space $\mathbf{b}$ is Z-transformed according to the LAMBDA concept (Teunissen and Kleusberg 1998), thus performing a rounding of the critical regions. We intend to perform that in future. Finally, despite the general statement of Eq. (21) on the error covariance and its approach to zero, one could be afraid that the Monte Carlo sampling would be too slow in filling the regions of high probability and, more generally, one could ask how fast this sampling is filling the integer lattice in $\boldsymbol{\beta}$ space. This question is discussed in Appendix 2.

3 Implementation and limits of the Monte Carlo method

The Monte Carlo method has been implemented in Bamba software (Betti et al. 1996) and applied to the baseline Herreros–Cijancos belonging to a non-permanent GPS network with geodynamical purposes in Eastern Granada in the south of Spain (Gil et al. 2001). The test baseline is about 8.5 km long, the height difference between Herreros and Cijancos stations is close to 600 m, and measurements were taken using the same receiver and antenna, SR9500 with antenna AT302.

Data processing was performed using Bernese 4.0 (Rothacher et al. 1996) and Bamba software. Only L1 frequency observations from 16 h 15 m to 16 h 45 m with 15^s of sample rate were used. In this observation period the satellite constellation was formed by the satellites PRN 3, 17, 21, 22, 23 and 31. Satellite 23 was adopted as reference satellite because it had the highest signal/noise rate. Herrero has been adopted as reference station. CODE ephemerides from the Astronomical Institute of the University of Berne were included for the computation. The Saastamoinen model with a standard atmosphere has been used for the tropospheric refraction. All observations having a mask lower than 20^o were discarded.

After obtaining the floating solution, we simulated a sample of 5000 ambiguities $\mathbf{b}_i$ distributed as $\mathbf{b}_i \approx N(\hat{\mathbf{b}}, \sigma_0^2 N_r^{-1})$ with dim $\mathbf{b} = 5$ and rounded every ambiguity to the nearest integer defining the discrete search space. The results of the Monte Carlo approximation have been compared with the FARA method. The choice of using the Bernese software, in which the FARA method is implemented, was made because this is in wide use and rigorous in that it computes and exploits the full covariance information; moreover, this was the software available to us. The authors are aware that more work should be done in order to have a systematic comparison with available methods. As for the choice of the Monte Carlo samples: we had the feeling, supported by the reasons explained in Sect. 2 , that any number above 1000 would do. In addition, we have seen that estimating with 4000 or 5000 samples was giving practically the same result, which we took as sign of success in the convergence of the estimators.

The final ambiguities calculated by the FARA and Monte Carlo methods are shown in Tables 1 and 2, respectively. In Table 1 can be seen a discontinuous behaviour in the ambiguity, showing jumps between integer and floating ambiguities, that does introduce a discontinuous behaviour in the coordinate estimator. In Table 2 it is noted that the Monte Carlo method is conservative and it tends little by little to one 'particular' integer. It is interesting to observe that the ambiguity in general does not coincide with an integer value because it is the average of a discrete distribution. Nevertheless,

Table 1. Final DD ambiguities calculated by FARA (cycles)

Minutes	Sat. 3	Sat. 17	Sat. 21	Sat. 22	Sat. 31
4	34 316	−9 960	27 813	31 048	52 799
5	34 316	−9 960	27 813	31 048	52 799
6	34 316	−9 960	27 813	31 048	52 799
7	34 316	−9 960	27 813	31 048	52 799
8	34 316	−9 960	27 813	31 048	52 799
9	34 316	−9 960	27 813	31 048	52 799
10	34 316.97	−9 959.65	27 813.14	31 048.39	52 800.49
11	34 317.24	−9 959.57	27 813.15	31 048.57	52 800.76
12	34 317.40	−9 959.52	27 813.13	31 048.68	52 800.90
13	34 317.26	−9 959.47	27 813.02	31 048.53	52 800.72
14	34 317.11	−9 959.44	27 812.91	31 048.40	52 800.47
15	34 316.87	−9 959.49	27 812.84	31 048.29	52 800.11
16	34 316.59	−9 959.54	27 812.78	31 048.09	52 799.75
17	34 316	−9 960	27 813	31 048	52 799
18	34 316	−9 960	27 813	31 048	52 799
19	34 316	−9 960	27 813	31 048	52 799
20	34 316	−9 960	27 813	31 048	52 799
21	34 316	−9 960	27 813	31 048	52 799
22	34 316	−9 960	27 813	31 048	52 799
23	34 316	−9 960	27 813	31 048	52 799
24	34 316	−9 960	27 813	31 048	52 799
25	34 316	−9 960	27 813	31 048	52 799
26	34 316	−9 960	27 813	31 048	52 799
27	34 315.58	−9 959.85	27 812.73	31 047.50	52 798.50
28	34 315.55	−9 959.86	27 812.73	31 047.53	52 798.47
29	34 315.51	−9 959.87	27 812.74	31 047.52	52 798.43
30	34 315.49	−9 959.88	27 812.74	31 047.51	52 798.39

Table 2. Final DD ambiguities calculated by the Monte Carlo method (cycles)

Minutes	Sat. 3	Sat. 17	Sat. 21	Sat. 22	Sat. 31
4	34 317.372	−9 959.025	27 813.196	31 048.29	52 799.690
5	34 315.909	−9 959.155	27 812.876	31 047.49	52 797.493
6	34 316.299	−9 959.006	27 813.973	31 047.588	52 798.193
7	34 316.72	−9 958.999	27 812.973	31 047.641	52 798.861
8	34 317.061	−9 958.999	27 812.996	31 047.873	52 799.459
9	34 317.452	−9 9658.999	27 813	31 048	52 799.986
10	34 317.994	−9 958.974	27 813	31 048.117	52 800.607
11	34 318.157	−9 959.744	27 813	31 048.54	52 800.970
12	34 318.469	−9 958.368	27 813	31 048.891	52 801.047
13	34 318.254	−9 958.068	27 813	31 048.657	52 800.994
14	34 318.050	−9 958.005	27 813	31 048.23	52 800.908
15	34 318.002	−9 958.046	27 813	31 048.023	52 800.327
16	34 317.983	−9 958.335	27 813	31 048	52 800.005
17	34 317.278	−9 958.840	27 812.998	31 048	52 799.829
18	34 317.004	−9 958.986	27 812.986	31 047.988	52 799.122
19	34 317	−9 958.999	27 812.98	31 047.95	52 799.013
20	34 317	−9 958.999	27 812.966	31 047.832	52 799.004
21	34 317	−9 959	27 812.972	31 047.782	52 799
22	34 317	−9 959	27 812.981	31 047.743	52 799
23	34 317	−9 959	27 812.992	31 047.755	52 799
24	34 317	−9 959	27 812.996	31 047.747	52 798.999
25	34 316.999	−9 959	27 813	31 047.747	52 798.994
26	34 316.999	−9 959	27 813	31 047.78	52 798.987
27	34 316.99	−9 959	27 813	31 047.564	52 798.957
28	34 316.974	−9 959	27 813	31 047.462	52 798.926
29	34 316.915	−9 959	27 813	31 047.294	52 798.835
30	34 316.828	−9 959	27 813	31 047.208	52 798.654

we expect that when the bias vector **b** is likely to attain integer components the corresponding ambiguity will drift close to the correct solution, while it will remain far from grid knots if the true ambiguity is not really integer (Betti et al. 1993). In order to check this fact, the integer grid points generated by the Monte Carlo simulation have been analysed after 4, 9 and 17 minutes. In Fig. 1 (four-minute observation period) it can be seen there are some candidates for each component of the ambiguity vector and no vector has probability close to one. In Fig. 2 (after nine-minute observation period), for satellites 21 and 22 we find the integer value with prob-

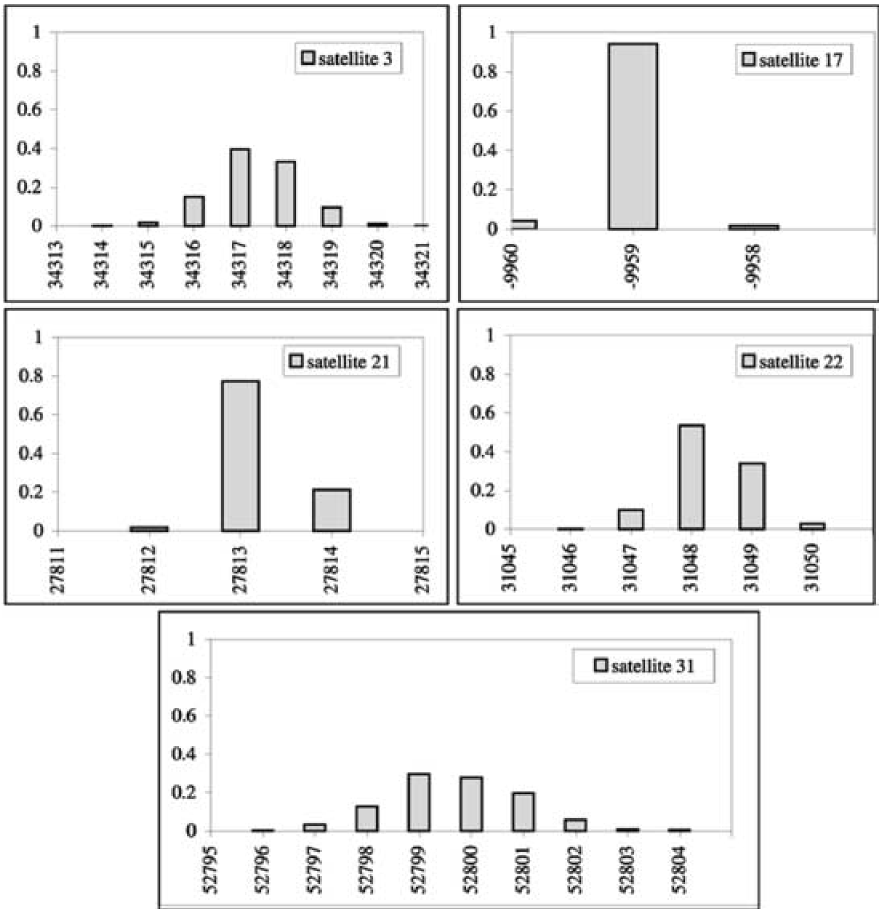

Fig. 1 Integers selected by Monte Carlo method after four minutes of observation

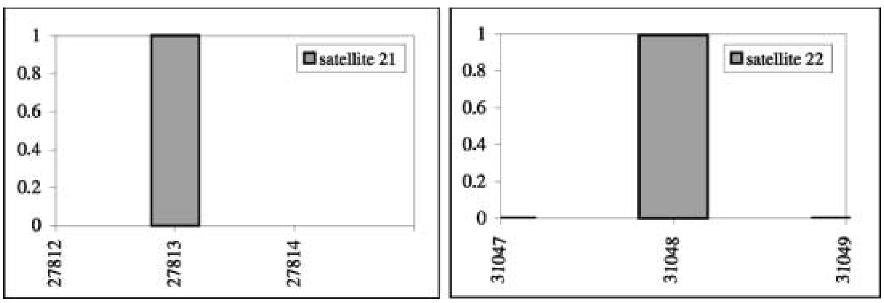

Fig. 2 Integers selected by Monte Carlo method after nine minutes of observation

ability equal to one. But when we analyse the situation after 17 minutes of measure before rounding to the nearest integer (Fig. 3), we find that for satellites 17 and 21 the ambiguities generated by the Monte Carlo method are far from the integer grid points but the values shown in Table 2 tend to −9959 and 27 813.

These results indicate that, within a short observation span, we expect that there are some candidates inside the search space. The shape of the search ellipsoid is illustrated in Fig. 4a and the corresponding normal distribution in Fig. 5a. In this case the mathematical model with integer ambiguities works because it is not really stringent, the Monte Carlo simulation works, and the approximation to the nearest integer works as well. After some minutes the shape of the ellipsoid changes and so the normal distribution changes (Fig. 4b and 5b); it may still contain more than one integer grid point but some of them can be very close to the bounds of the search ellipsoid. In this situation, the mathematical model of integer ambiguities is still acceptable but the Monte Carlo simulation does not work. In fact, our Monte Carlo approximation is based on the approximation $i = NN_\delta(x)$ with x distributed as a normal $N(\mu, \sigma^2) \equiv g(x \mid \mu, \sigma^2)$ and $p(i) \cong g(x \mid \mu, \sigma^2) \cdot \delta$, and where δ is the grid step size. $NN_\delta(x)$ means the nearest neighbour of x within the grid and $g(x \mid \mu, \sigma^2)$ is a normal probability density. Indeed, such an approximation holds reasonably when $\delta < \sigma$ but it becomes too coarse when $\delta > \sigma$.

Because of the elongated shape of the ellipsoid it could happen that a few minutes later no integer grid point is inside the search ellipsoid (Figs. 4c and 5c).

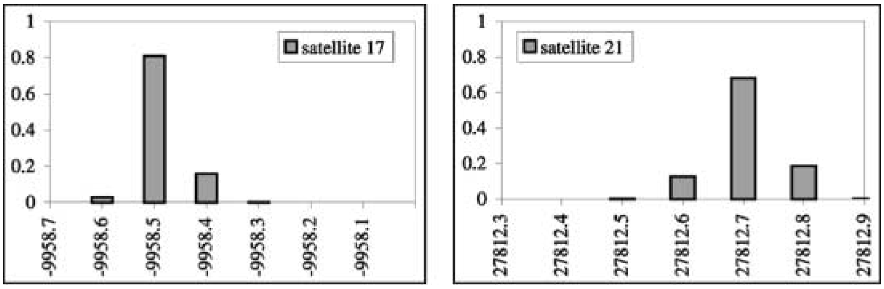

Fig. 3 Integers selected by Monte Carlo method after 17 minutes of observation

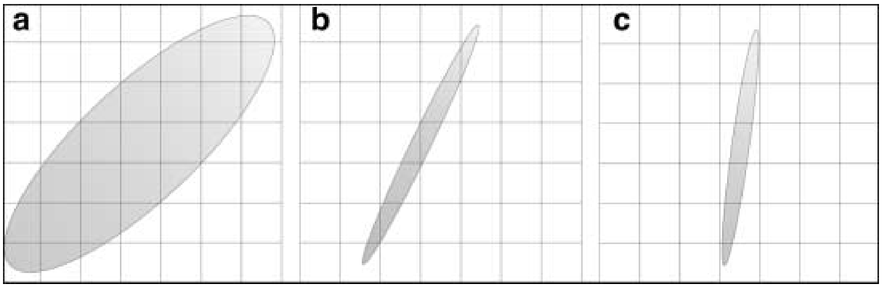

Fig. 4 Three different shapes of the ambiguity search ellipsoid

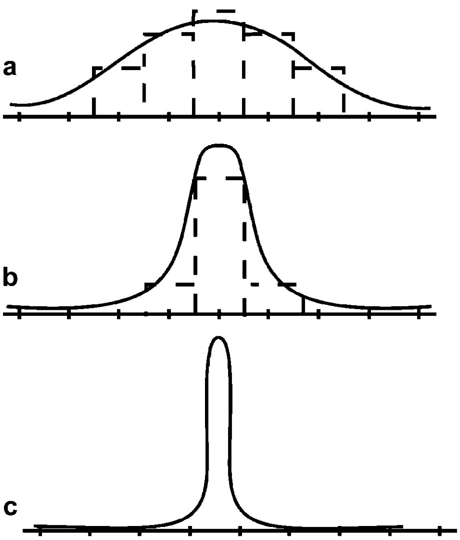

Fig. 5. Three different shapes of the normal distribution

Now the mathematical model with integer ambiguities is no longer tenable and we have to admit that our data suggest that it is very unlikely that $\mathbf{b}$ can coincide with any knot in the grid.

Reconsidering this behaviour, we find that it is from the beginning that we have to build a model admitting two possible alternatives, one where $\mathbf{b}$ is an integer vector and another one where $\mathbf{b}$ has no integer constraints, i.e. $\mathbf{b}$ is generated by other effects than the simple initial ambiguity. It will then be the accumulation of information with time and the corresponding evolution of the characteristic ellipsoid that will tell us which of the two alternatives is more probable. This mixed model will be developed in the next paragraph.

4 A mathematical mixed model for the integer–real ambiguities alternative

According to the discussion of the last section, let us try to design a new model larger than Eq. (1) where, among the parameters, beyond $(\mathbf{r}, \mathbf{b})$ we introduce

$$\omega = \begin{cases} I & \text{if } \mathbf{b} \text{ is integer, } \mathbf{b} = \mathbf{b}_I \\ F & \text{if } \mathbf{b} \text{ is real, } \mathbf{b} \in R^n \end{cases} \tag{22}$$

ω has the meaning of a discrete label variate and we assume it to be a priori independent from $\mathbf{r}$.

The posterior distribution of the parameters will now be given by

$$p(\mathbf{r}, \mathbf{b}, \omega|\mathbf{Y}) = CL(\mathbf{Y}|\mathbf{r}, \mathbf{b}, \omega)\tilde{p}(\mathbf{r})\tilde{p}(\mathbf{b}, \omega) \tag{23}$$

C being a constant such that Eq. (23) is a probability distribution in $(\mathbf{r}, \mathbf{b}, \omega)$ for fixed $\mathbf{Y}$. In order to describe the a priori relation between ω and $\mathbf{b}$ we will use the improper (prior) distribution

$$\tilde{p}(\mathbf{b}, \omega) = \tilde{p}(\mathbf{b}|\omega)\tilde{p}(\omega) \propto \begin{cases} \alpha \sum_{\mathbf{b}_I} \delta(\mathbf{b} - \mathbf{b}_I) & \text{if } \omega = I \\ (1 - \alpha) & \text{if } \omega = F \end{cases} \tag{24}$$

where α becomes a parameter controlling the prior probability between the two alternatives I, F. Therefore α is what is called in Bayesian literature a hyperparameter (Box and Tiao 1992), and more correctly one should write in Eq. (24)

$$\tilde{p}(\mathbf{b}, \omega, \alpha) = \tilde{p}(\mathbf{b}|\omega, \alpha)\tilde{p}(\omega|\alpha)\tilde{p}(\alpha) \tag{25}$$

In this section, for the sake of simplicity we will simply assume that α is a variable 'known from experience' and we will fix it to $\alpha = 90\%$; only at the end of the paragraph will we show how to deal with it in the completely opposite hypothesis based on the use of a non-informative prior for α. With this remark in mind, the posterior distribution then becomes

$$p(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}, \omega|\mathbf{Y}) \propto \mathrm{e}^{-\frac{1}{2\sigma_0^2}(\boldsymbol{\delta}\mathbf{r}+N_{GG}^{-1}N_{Gb}\boldsymbol{\beta})^T N_{GG}(\boldsymbol{\delta}\mathbf{r}+N_{GG}^{-1}N_{Gb}\boldsymbol{\beta})} \times \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}^T N_r \boldsymbol{\beta}} \cdot \begin{cases} \alpha \sum_{\boldsymbol{\beta}_I} \delta(\boldsymbol{\beta} - \boldsymbol{\beta}_I) & \text{if} \quad \omega = I \\ (1-\alpha) & \text{if} \quad \omega = F \end{cases} \tag{26}$$

From Eq. (26) all the necessary marginal distributions can be derived; in particular we find

$$p(\omega|\mathbf{Y}) = \begin{cases} P_I = \frac{\alpha}{K} \cdot \sum_{\boldsymbol{\beta}_I} \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I} & \omega = I \\ 1 - P_I = \frac{1-\alpha}{K} \cdot \frac{(2\pi)^{\frac{m}{2}}\sigma_0^m}{\sqrt{|N_r|}} & \omega = F \end{cases} \tag{27}$$

where $m = \dim \boldsymbol{\beta}$ and

$$K = \alpha \sum_{\boldsymbol{\beta}} \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I} + (1-\alpha) \cdot \frac{(2\pi)^{\frac{m}{2}}\sigma_0^m}{\sqrt{|N_r|}} \tag{28}$$

We know that

$$p(\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}, \omega = I) \propto C \sum_{\boldsymbol{\beta}_I} \mathrm{e}^{-\frac{1}{2\sigma_0^2}(\boldsymbol{\delta}\mathbf{r}+N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I)^T N_{GG}(\boldsymbol{\delta}\mathbf{r}+N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I)} \times \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I} \tag{29}$$

and

$$p(\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}, \omega = F) \propto C\mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\delta}\mathbf{r}^T (N_{GG} - N_{Gb}N_{bb}^{-1}N_{bG})\boldsymbol{\delta}\mathbf{r}} \tag{30}$$

The derivations can be found in Appendix 3. From the above relations, and considering that

$$p(\boldsymbol{\delta}\mathbf{r}|\mathbf{y}) = P_I p(\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}, I) + (1 - P_I)p(\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}, F) \tag{31}$$

one is then able to derive the Bayes posterior mean

$$\boldsymbol{\delta}\mathbf{r}_B = E\{\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}\} = P_I E\{\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}, \omega = I\} + (1 - P_I)E\{\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}, \omega = F\} = P_I \boldsymbol{\delta}\mathbf{r}_{BI} \tag{32}$$

where $\boldsymbol{\delta}\mathbf{r}_{BI}$ is exactly what we have found in Eq. (12), while

$$\boldsymbol{\delta}\mathbf{r}_{BF} = E\{\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}, \omega = F\} = 0 \tag{33}$$

because without the lattice constraint, i.e. when $\omega = f$, the best estimate is already the floating without any correction.

Reasoning in a similar way (see Appendix 3) we can derive the posterior covariance, which turns out to be

$$C_{\boldsymbol{\delta}\mathbf{r}_B\boldsymbol{\delta}\mathbf{r}_B} = P_I C_I + (1 - P_I)\sigma_0^2 N_{GG}^{-1} + (1 - P_I) \times (\sigma_0^2 N_{GG}^{-1}N_{Gb}N_r^{-1}N_{bG}N_{GG}^{-1} - P_I \boldsymbol{\delta}\mathbf{r}_{BI}\boldsymbol{\delta}\mathbf{r}_{BI}^T) \tag{34}$$

where C_I is the covariance matrix [Eq. (19)] corresponding to the integer case already discussed.

Therefore, the mixture model should refine the Monte Carlo method in the following way:

1. The floating solution and its covariance matrix will be derived from the LS adjustment.
2. A sample of ambiguities normally distributed will be generated by the Monte Carlo method and every value will be rounded to the nearest integer defining the ambiguity search space.
3. All possible integer combinations will be made to compute P_I, and then summing over all these combinations we will be able to obtain the exact expression for Eqs. (32) and (34).

Summarizing, it is interesting to stress that, after all, the application of this model can follow any floating solution, with in addition an indicator of what has to be the search space, which in the present work is derived from an application of the Monte Carlo method.

This procedure, based on the mathematical mixture model following the steps 1, 2, and 3, has been applied to two examples. The first one is the baseline Cijancos–Romeral belonging to the non-permanent GPS network already mentioned in Sect. 3. Its length is about 5 km and we have processed the GPS measurements in the same manner as explained in Sect. 3 for the baseline Herreros–Cijancos. The second example is the proper baseline Herreros–Cijancos. In both cases we have built the mixed model starting from the floating solution of the Bernese software because we wanted to investigate with our new tool the behaviour of the ambiguities over time.

The results corresponding to the first test baseline in terms of P_I are shown in Table 3. We can observe, accordingly to FARA (Table 4), a regular evolution of

Table 3. Mixed-model analysis of Cijancos–Romeral basis

Minutes	P_I	$1-P_I$
1	0.968	0.032
2	0.948	0.052
3	0.001	0.999
4	0.135	0.865
5	0.999	0.001
6	0.999	0.001
7	0.999	0.001
8	0.999	0.001
9	0.999	0.001
10	0.999	0.001
11	0.999	0.001
12	0.999	0.001
13	0.999	0.001
14	0.999	0.001
15	0.999	0.001
16	0.999	0.001
17	0.999	0.001
18	0.999	0.001
19	0.999	0.001
20	0.999	0.001

Table 4. Final DD ambiguities (cycles) calculated by FARA for Cijancos–Romeral basis

Minutes	Ambiguity vector
1	(−12 522.42, 32 556.57, 23 872.49, −23 660.5, −2283.66)
2	(927.41, 8680.77, −23 871.79, −10 211.88, −26 155.7, −24 958.44)
3	(924.75, 8671.01, −23 871.43, −10 216.78, −26 155.86, −24 961.14)
4	(924, −19 091, −23 871, −10 222, −26 156, −24 963)
5	(924, 8671, −23 871, −10 222, −26 156, −24 963)
6	(924, 8671, −23 871, −10 222, −26 156, −24 963)
7	(924, 8671, −23 871, −10 222, −26 156, −24 963)
8	(924, 8671, −23 871, −10 222, −26 156, −24 963)
9	(924, 8671, −23 871, −10 222, −26 156, −24 963)
10	(924, 8671, −23 871, −10 222, −26 156, −24 963)
11	(924, 8671, −23 871, −10 222, −26 156, −24 963)
12	(924, 8671, −23 871, −10 222, −26 156, −24 963)
13	(924, 8671, −23 871, −10 222, −26 156, −24 963)
14	(924, 8671, −23 871, −10 222, −26 156, −24 963)
15	(924, 8671, −23 871, −10 222, −26 156, −24 963)
16	(924, 8671, −23 871, −10 222, −26 156, −24 963)
17	(924, 8671, −23 871, −10 222, −26 156, −24 963)
18	(924, 8671, −23 871, −10 222, −26 156, −24 963)
19	(924, 8671, −23 871, −10 222, −26 156, −24 963)
20	(924, 8671, −23 871, −10 222, −26 156, −24 963)

the ambiguity estimate with time. Only at the beginning P_I does an irregular behaviour due to a new satellite arise after 1.30 minutes of observation. In contrast to this example, we can analyse the erratic behaviour of the ambiguity vector in the Herreros–Cijancos baseline (Table 1). The results provided by the mixed model in terms of P_I are shown in Table 5. These indicate that for a short observation period there are some integer candidates inside the search space of the ambiguities and in this case P_I is high enough to justify the belief of an integer bias. When observations are accumulated, the shape of the characteristic ellipsoid $\mathscr{E}$ changes and P_I decreases, closing to zero. This means that $\hat{\mathbf{b}}$ is not centered perfectly on an integer knot and so no integer grid point is inside the search space. Later, only one integer grid point appears in the characteristic ellipsoid. In a situation like this, the Bayesian correction $\boldsymbol{\delta}\mathbf{r}_B$ given by Eq. (32) will compensate for this phenomenon.

It is interesting to note that finding a basis for which we were sure that one and only one (over a time span of 20 minutes) integer bias could be chosen was not immediate; this brings us back to the question of the 'a priori' value of α. We would like to close the section by outlining the procedure to adopt in the case that α is considered as a hyperparameter, for instance with a non-informative prior.

To this aim we first observe that all the formulas worked out until now, and in particular from Eq. (27) to Eq. (34), still hold true as they are with the only proviso that all distributions $P(\cdot|\mathbf{Y})$ and averages $E\{\cdot|\mathbf{Y}\}$ be substituted by $P(\cdot|\mathbf{Y},\alpha)$ and $E\{\cdot|\mathbf{Y},\alpha\}$. Accordingly, we assume that

$$\tilde{p}(\alpha) = 1, \quad 0 \le \alpha \le 1 \tag{35}$$

Table 5. Mixed-model analysis of Herreros–Cijancos basis

Minutes	P_I	$1-P_I$
4	0.999	0.001
5	0.999	0.001
6	0.999	0.001
7	0.999	0.001
8	0.999	0.001
9	0.031	0.969
10	0	1
11	0	1
12	0	1
13	0	1
14	0	1
15	0	1
16	0.996	0.004
17	0.999	0.001
18	0.999	0.001
19	0.999	0.001
20	0.999	0.001

is the non-informative prior for α, we have only to integrate the various quantities as functions of α, to obtain their posterior marginal distribution or mean, conditional to $\mathbf{Y}$ only.

For instance, Eq. (27) gives us $P_I = P(\omega = I|\mathbf{Y},\alpha)$ and we can write

$$P_I = p(\omega = I|\mathbf{Y}) = \int_0^1 P(\omega = I|\mathbf{Y},\alpha)\tilde{p}(\alpha)\mathrm{d}\alpha \tag{36}$$

Accordingly, we write Eq. (32) as

$$\boldsymbol{\delta}\mathbf{r}_B(\alpha) = P_I(\alpha)\boldsymbol{\delta}\mathbf{r}_{BI} \tag{37}$$

and the final result is

$$\boldsymbol{\delta}\mathbf{r}_B = \int_0^1 P_I(\alpha)\ \boldsymbol{\delta}\mathbf{r}_{BI} = P_I\boldsymbol{\delta}\mathbf{r}_{BI} \tag{38}$$

In a similar way, Eq. (27) holds true in the same form, but with P_I given by Eq. (36). Therefore the only exercise to perform is to compute the integral of Eq. (36) from Eq. (27), taking into account that K is a function of α too; the explicit result is

$$P_I = \frac{q}{q-1}\left\{1 - \frac{1}{q-1}\log q\right\} \tag{39}$$

where

$$q = \sum_{\boldsymbol{\beta}_I} \frac{\sqrt{|N_r|}}{(2\pi)^{\frac{m}{2}}\sigma_0^m} \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I} \tag{40}$$

5 Conclusions

In this paper we first returned to the Bayesian approach of GPS baseline estimation (Blewitt 1989; Betti et al. 1993), showing how the Monte Carlo method provides a possible practical tool to define a finite set of knots of the lattice of integer bias vectors on which a reasonable approximation of the infinite series, inherent in the theory, can be obtained. The technique proposed is not free of the further approximations and we feel that improvements are possible and deserve further research.

In performing this work we realized that there were cases in which by forming **b** in any integer knot we were in fact jumping into an area of extremely low probability. Indeed, if we say that the integer ambiguity model is correct, any estimation method will give us back pure integer estimates; a frequentist approach will choose one particular integer set, a Bayesian approach will give us a posterior distribution with a single spike practically equal to 1. However, if we suspect that the model may be wrong then we can build, for example in the Bayesian framework, a more general model where according to the data we can choose from two alternatives: **b** is integer ($\omega = I$) or **b** is real ($\omega = F$). This is the mathematical mixture model worked out in Sect. 4. A justification of this mixed model is in that we can imagine that a number of effects not perfectly accounted for by the model of Eq. (1) enter as systematic factors in v, which however is modelled as a pure random noise. As we know, this introduces biases in the LS estimation of parameters, so that $\hat{\mathbf{b}}$ might not be perfectly entered in an integer knot. In Bayesian terms, we can say that restricting the prior distribution of **b** to the integer lattice is not consistent with the data.

Of course, if $\sigma_0^2 N_r^{-1}$ is still large (a characteristic ellipsoid $\mathscr{E}$ includes several knots of the lattice), we are not able to discriminate between the two cases, but when more observations are accumulated with time it is likely that that only one knot, $\mathbf{b}_I$, will have a high posterior probability under both models $\omega = I$ and $\omega = F$; in frequentist terms we could say that only $\mathbf{b}_I$ falls in the characteristic ellipsoid. If this situation continues while $\sigma_0^2 N_r^{-1}$ shrinks, we can claim that $\omega = I$ is plausible, if on the contrary at a certain moment the posterior probability of $\mathbf{b}_I$ drops (i.e. if it slips out of the ellipsoid $\mathscr{E}$) we can say that a significant bias with respect to the integer ambiguity model is increasing. The Bayesian approach can compensate automatically for this phenomenon and in addition provides a good index, namely P_I, to monitor which of the two models is prevailing with time; when $P_I \to 0$ it is the floating model which better interprets the data, and when $P_I \to 1$ the opposite is true.

Appendix 1

We want to prove how, from the posterior distribution of Eq. (10), we can derive the posterior mean and its covariance matrix. Let us consider the quadratic form of Eq. (11) written as

$$\Psi(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}_I) = \boldsymbol{\delta}\mathbf{r}^T N_{GG}\boldsymbol{\delta}\mathbf{r} + 2\boldsymbol{\delta}\mathbf{r}^T N_{Gb}\boldsymbol{\beta}_I + \boldsymbol{\beta}_I^T N_{bb}\boldsymbol{\beta}_I \tag{A1}$$

After some manipulations, Eq. (A1) can be expressed as

$$\begin{aligned}\Psi(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}_I) &= \boldsymbol{\delta}\mathbf{r}^T N_{GG}\boldsymbol{\delta}\mathbf{r} + 2\boldsymbol{\delta}\mathbf{r}^T N_{GG}N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I \\ &\quad + \boldsymbol{\beta}_I^T N_{bG}N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I \\ &\quad + \boldsymbol{\beta}_I^T (N_{bb} - N_{bG}N_{GG}^{-1}N_{Gb})\boldsymbol{\beta}_I\end{aligned} \tag{A2}$$

Calling $N_r = N_{bb} - N_{bG}N_{GG}^{-1}N_{Gb}$, Eq. (A2) can be written as

$$\begin{aligned}\Psi(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}_I) &= (\boldsymbol{\delta}\mathbf{r} + N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I)^T N_{GG}(\boldsymbol{\delta}\mathbf{r} + N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I) \\ &\quad + \boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I\end{aligned} \tag{A3}$$

Therefore the posterior distribution will be given by

$$\begin{aligned}p(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}|\mathbf{Y}) &\propto \sum_{\boldsymbol{\beta}_I} \boldsymbol{\delta}(\boldsymbol{\beta} - \boldsymbol{\beta}_I) \cdot \mathrm{e}^{-\frac{1}{2\sigma_0^2}(\boldsymbol{\delta}\mathbf{r} + N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I)^T N_{GG}(\boldsymbol{\delta}\mathbf{r} + N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I)} \\ &\quad \times \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I}\end{aligned} \tag{A4}$$

and from Eq. (A4) we can derive

$$p(\boldsymbol{\delta}\mathbf{r}|\boldsymbol{\beta} = \boldsymbol{\beta}_I, \mathbf{Y}) \propto \mathrm{e}^{-\frac{1}{2\sigma_0^2}(\boldsymbol{\delta}\mathbf{r} + N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I)^T N_{GG}(\boldsymbol{\delta}\mathbf{r} + N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I)} \tag{A5}$$

$$p(\boldsymbol{\beta} = \boldsymbol{\beta}_I|\mathbf{Y}) \propto \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I} \tag{A6}$$

From Eqs. (A6), (A4) and (4) the expression of the Bayesian estimator becomes

$$\begin{aligned}\boldsymbol{\delta}\mathbf{r}_B &= E\{\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}\} = E_{\boldsymbol{\beta}_I}\{E_{\boldsymbol{\delta}\mathbf{r}}\{\boldsymbol{\delta}\mathbf{r}|\boldsymbol{\beta}_I, \mathbf{Y}\}\} \\ &= -\frac{\sum_{\boldsymbol{\beta}_I} N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I \cdot \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I}}{\sum_{\boldsymbol{\beta}_I} \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I}} \\ &= -N_{GG}^{-1}N_{Gb}E\{\boldsymbol{\beta}_I\}\end{aligned} \tag{A7}$$

which does coincide with Eq. (12), as it was to be proved.

The covariance matrix of the distribution of Eq. (3) is given by

$$C_{\boldsymbol{\delta}\mathbf{r}_B\boldsymbol{\delta}\mathbf{r}_B} = \mathrm{Cov}(\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}) = E\{(\boldsymbol{\delta}\mathbf{r}\boldsymbol{\delta}\mathbf{r}^T - \boldsymbol{\mu}_{\delta r|Y}\boldsymbol{\mu}_{\delta r|Y}^T)|\mathbf{Y}\} \tag{A8}$$

where $\boldsymbol{\mu}_{\delta r|Y}$ is the mean of the $\boldsymbol{\delta}\mathbf{r}$ estimations conditional to the observations.

From Eq. (12), Eq. (A8) becomes

$$C_{\boldsymbol{\delta}\mathbf{r}_B\boldsymbol{\delta}\mathbf{r}_B} = E\{\boldsymbol{\delta}\mathbf{r}\boldsymbol{\delta}\mathbf{r}^T|\mathbf{Y}\} - N_{GG}^{-1}N_{Gb}\boldsymbol{\mu}_{\boldsymbol{\beta}_I}\boldsymbol{\mu}_{\boldsymbol{\beta}_I}^T N_{bG}N_{GG}^{-1} \tag{A9}$$

and

$$\begin{aligned}C_{\boldsymbol{\delta}\mathbf{r}_B\boldsymbol{\delta}\mathbf{r}_B} &= E_{\boldsymbol{\beta}_I}\{E_{\boldsymbol{\delta}\mathbf{r}}\{\boldsymbol{\delta}\mathbf{r}\boldsymbol{\delta}\mathbf{r}^T|\boldsymbol{\beta}_I, \mathbf{Y}\}\} - N_{GG}^{-1}N_{Gb}\boldsymbol{\mu}_{\boldsymbol{\beta}_I}\boldsymbol{\mu}_{\boldsymbol{\beta}_I}^T N_{bG}N_{GG}^{-1} \\ &= E_{\boldsymbol{\beta}_I}\{\mathrm{cov}(\boldsymbol{\delta}\mathbf{r}|\boldsymbol{\beta}_I, \mathbf{Y}) + N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I\boldsymbol{\beta}_I^T N_{bG}N_{GG}^{-1}|\mathbf{Y}\} \\ &\quad - N_{GG}^{-1}N_{Gb}\boldsymbol{\mu}_{\boldsymbol{\beta}_I}\boldsymbol{\mu}_{\boldsymbol{\beta}_I}^T N_{bG}N_{GG}^{-1}\end{aligned} \tag{A10}$$

Since we can put

$$C_{\boldsymbol{\beta}_I\boldsymbol{\beta}_I} = E\{(\boldsymbol{\beta}_I\boldsymbol{\beta}_I^T - \boldsymbol{\mu}_{\boldsymbol{\beta}_I}\boldsymbol{\mu}_{\beta_I}^T)|\mathbf{Y}\} \tag{A11}$$

we derive the exact formula for the sought covariance

$$C_{\delta r_B \delta r_B} = \sigma_0^2 N_{GG}^{-1} + N_{GG}^{-1} N_{Gb} C_{\beta_I \beta_I} N_{bG} N_{GG}^{-1} \quad \text{(A12)}$$

Appendix 2

Let us consider the splitting of Eq. (13) into a finite sum and a rest

$$\boldsymbol{\mu}_{\beta_I} = \sum_{\beta_I \in \mathscr{E}} \boldsymbol{\beta}_I p(\boldsymbol{\beta}_I) + \boldsymbol{\epsilon} \quad \text{(A13)}$$

where the finite sum can be considered as an approximation of $\boldsymbol{\mu}_{\beta_I}$ and

$$\boldsymbol{\epsilon} = \sum_{\beta_I \in O} \boldsymbol{\beta}_I p(\boldsymbol{\beta}_I) \quad \text{(A14)}$$

We can substitute this term in Eq. (13) with a finite sum over random $\boldsymbol{\beta}_I$ and these reach a maximum norm (in the natural covariance metric $\sigma_0^{-2} N_r$) which defines the maximal $\mathscr{E}^M$ including all the visited knots. The corresponding error for the outer (non-visited) knots is

$$\boldsymbol{\epsilon} = \sum_{\beta_I \in O^M} \boldsymbol{\beta}_I p(\boldsymbol{\beta}_I) \quad \text{(A15)}$$

and we would like to relate this to the dimensions k of the sample. We will do that by considering a kind of relative error coefficient

$$RE^2 = \frac{\boldsymbol{\epsilon}^T N_r \boldsymbol{\epsilon}}{E\{\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I\}} \quad \text{(A16)}$$

where the factor σ_0^2 is cancelled between numerator and denominator. We have, by the Schwartz inequality

$$\begin{aligned} RE^2 &= \frac{\sum_{\beta_I, \beta_J \in O^M} \boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_J \cdot p(\boldsymbol{\beta}_I) \cdot p(\boldsymbol{\beta}_J)}{\sum_{\beta_I} \boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I \cdot p(\boldsymbol{\beta}_I)} \\ &\leq \frac{\left[\sum_{\beta_I \in O^M} \sqrt{\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I} \cdot p(\boldsymbol{\beta}_I)\right]^2}{\sum_{\beta_I} \boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I \cdot p(\boldsymbol{\beta}_I)} \\ &\leq P\{\boldsymbol{\beta}_I \in O^M\} \frac{\sum_{\beta_I \in O^M} \boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I \cdot p(\boldsymbol{\beta}_I)}{\sum_{\beta_I} \boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I \cdot p(\boldsymbol{\beta}_I)}, \end{aligned} \quad \text{(A17)}$$

Let us explicitly note that in Eq. (A17) the $\sum_{\beta_I}$ in the denominator runs over all the knots of the lattice in $\boldsymbol{\beta}_I$ space.

Now, for the sake of finding an approximate value, we substitute the summation with integrals and we note that we come out with averages over the $\chi_m^2 (m = \dim \boldsymbol{\beta})$ distribution in both the denominator and numerator, although this is truncated to the value

$$\xi^M = \frac{1}{\sigma_0^2} \boldsymbol{\beta}_I^{MT} N_r \boldsymbol{\beta}_I^M (\mathscr{E}^M = \{\boldsymbol{\beta}; \boldsymbol{\beta}^T N_r \boldsymbol{\beta} \leq \sigma_0^2 \xi^M\}) \quad \text{(A18)}$$

Namely, we find

$$\begin{aligned} \frac{\sum_{\beta_I \in O^M} \boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I \cdot p(\boldsymbol{\beta}_I)}{\sum_{\beta_I} \boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I \cdot p(\boldsymbol{\beta}_I)} &\approx \frac{\int_{O^M} \boldsymbol{\beta}^T N_r \boldsymbol{\beta} \cdot e^{-\frac{1}{2\sigma_0^2} \boldsymbol{\beta}^T N_r \boldsymbol{\beta}} d_m \beta}{\int_{R^M} \boldsymbol{\beta}^T N_r \boldsymbol{\beta} \cdot e^{-\frac{1}{2\sigma_0^2} \boldsymbol{\beta}^T N_r \boldsymbol{\beta}} d_m \beta} \\ &= \frac{\int_{\xi^M}^{+\infty} \xi \xi^{\frac{m}{2}-1} e^{\frac{-1}{2}\xi} d\xi}{\int_0^{+\infty} \xi \xi^{\frac{m}{2}-1} e^{\frac{-1}{2}\xi} d\xi} = p\{\chi_{m+2}^2 \geq \xi^M\} \end{aligned} \quad \text{(A19)}$$

For the same reason we can use the approximation

$$p\{\boldsymbol{\beta} \in O^M\} \approx p\{\chi_m^2 \geq \xi^M\} \quad \text{(A20)}$$

Moreover, since χ_m^2 and χ_{m+2}^2 are not that different, from Eq. (A17), summarizing, we can write

$$RE \leq p\{\chi_m^2 \geq \xi^M\} \quad \text{(A21)}$$

Now, Eq. (A21) is an inequality between two positive random variables, so in order to give a numerical appreciation of it we could take the average of the right-hand side, namely

$$E\{p(\chi_m^2 \geq \xi^M)\} = E\{1 - F_m(\xi^M)\} \quad \text{(A22)}$$

where F_m is the distribution function of χ_m^2 while ξ^M is a random variable

$$\xi^M = \text{Max}_{i=1,\ldots,k} \xi_i, \; \xi_i \approx IID(\chi_m^2) \quad \text{(A23)}$$

with density

$$f_M(\xi) = k F_m^{k-1} f_m(\xi) \; (\chi_m^2 \approx f_m(\xi) = F'_m(\xi)) \quad \text{(A24)}$$

Accordingly, Eq. (A22) gives

$$\begin{aligned} E\{1 - F_m(\xi^M)\} &= \int_o^{+\infty} [1 - F_m(\xi)] k F_m^{k-1}(\xi) f_m(\xi) d\xi \\ &= (1 - \frac{k}{k+1}) = \frac{1}{k+1} \end{aligned} \quad \text{(A25)}$$

Relations (A21), (A22) and (A25) indicate that

$$E\{RE\} \leq \frac{1}{k+1} \quad \text{(A26)}$$

Basically Eq. (A26) indicates that the error of the Monte Carlo method in estimating $\boldsymbol{\mu}_\beta$ is due much more to the inaccuracy of the empirical estimates of the probabilities inside the encompassing ellipsoid $\mathscr{E}^M$ which are of $O(\frac{1}{\sqrt{k}})$, than to the completely neglected external zone O^M.

Appendix 3

In this appendix we derive the marginal distributions of Eqs. (27), (29) and (30) and the covariance matrix of Eq. (34) from Eq. (26).

$$\begin{aligned} p(\boldsymbol{\beta}, \omega | \mathbf{Y}) &= \int p(\boldsymbol{\delta r}, \boldsymbol{\beta}, \omega | \mathbf{Y}) d(\delta r) \propto e^{-\frac{1}{2\sigma_0^2} \boldsymbol{\beta}^T N_r \boldsymbol{\beta}} \\ &\times \begin{cases} \alpha \sum_{\beta_I} \delta(\boldsymbol{\beta} - \boldsymbol{\beta}_I) & \text{if } \omega = I \\ (1 - \alpha) & \text{if } \omega = F \end{cases} \end{aligned} \quad \text{(A27)}$$

$$p(\omega|\mathbf{Y}) = \int \mathrm{d}\beta \int p(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}, \omega|\mathbf{Y})\mathrm{d}(\delta r)$$
$$\propto \begin{cases} \alpha \sum_{\boldsymbol{\beta}_I} \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I} & \text{if } \omega = I \\ (1-\alpha) \cdot \frac{(2\pi)^{\frac{m}{2}}\sigma_0^m}{\sqrt{|N_r|}} & \text{if } \omega = F \end{cases} \tag{A28}$$

where $m = \dim\boldsymbol{\beta}$. Notice that in Eq. (A28) use has been made of the relation

$$\int \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}^T N_r \boldsymbol{\beta}} \delta(\boldsymbol{\beta} - \boldsymbol{\beta}_I)\mathrm{d}\beta = \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I} \tag{A29}$$

To normalize Eq. (A28) we can write

$$K = \alpha \sum_{\boldsymbol{\beta}} \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I} + (1-\alpha) \cdot \frac{(2\pi)^{\frac{m}{2}}\sigma_0^m}{\sqrt{|N_r|}} \tag{A30}$$

and we then have Eq. (27)

$$p(\omega|\mathbf{Y}) = \begin{cases} \alpha \frac{\sum_{\boldsymbol{\beta}_I} \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I}}{K} & = P_I \\ (1-\alpha) \cdot \frac{(2\pi)^{\frac{m}{2}}\sigma_0^m}{K\sqrt{|N_r|}} & = 1 - P_I \end{cases} \tag{A31}$$

In particular, the posterior probability for the case in which the bias vector is due only to the ambiguity terms, and the information available is sufficient to find their integer values, is

$$P_I = \alpha \frac{\sum_{\boldsymbol{\beta}_I} \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I}}{K} \tag{A32}$$

Therefore

$$p(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}|\mathbf{Y}) = \sum_{\omega} p(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}, \omega|\mathbf{Y}) = \sum_{\omega} p(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}|\omega, \mathbf{Y}) \cdot p(\omega|\mathbf{Y})$$
$$= p(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}|\mathbf{Y}, I)P_I + p(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}|\mathbf{Y}, F)(1 - P_I) \tag{A33}$$

As can be seen, Eq. (A33) depends on a combination of the posterior distribution of the integer case and the posterior distribution of the floating case. In the integer case, the posterior distribution of the parameter $\boldsymbol{\delta}\mathbf{r}$ is the posterior marginal of

$$p(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}|\mathbf{Y}, I) = p(\boldsymbol{\delta}\mathbf{r}|\boldsymbol{\beta}, \mathbf{Y}, I)p(\boldsymbol{\beta}|\mathbf{Y}, I)$$
$$= C\mathrm{e}^{-\frac{1}{2\sigma_0^2}(\boldsymbol{\delta}\mathbf{r} + N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I)^T N_{GG}(\boldsymbol{\delta}\mathbf{r} + N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}_I)}$$
$$\times \frac{\sum_{\boldsymbol{\beta}_I} \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I} \delta(\boldsymbol{\beta} - \boldsymbol{\beta}_I)}{\sum_{\boldsymbol{\beta}_I} \mathrm{e}^{\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I}} \tag{A34}$$

i.e. it coincides with Eq. (A4). Moreover, the integral over $\boldsymbol{\beta}$ of Eq. (A34) gives Eq. (29).

In the floating case, the posterior distribution is

$$p(\boldsymbol{\delta}\mathbf{r}, \boldsymbol{\beta}|\mathbf{Y}, F) = p(\boldsymbol{\delta}\mathbf{r}|\boldsymbol{\beta}, \mathbf{Y}, F) \cdot p(\boldsymbol{\beta}|\mathbf{Y}, F)$$
$$\propto \mathrm{e}^{-\frac{1}{2\sigma_0^2}(\boldsymbol{\delta}\mathbf{r} + N_{GG}^{-1}N_{Gb}\boldsymbol{\beta})^T N_{GG}(\boldsymbol{\delta}\mathbf{r} + N_{GG}^{-1}N_{Gb}\boldsymbol{\beta})} \cdot \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}^T N_r \boldsymbol{\beta}} \tag{A35}$$

which again coincides with the standard floating distribution. By integrating over $\boldsymbol{\beta}$ we obtain Eq. (30), although we will avoid the explicit use of this distribution. Then the Bayesian estimator of the coordinate vector for the integer case can be written as

$$\boldsymbol{\delta}\mathbf{r}_{BI} = E\{\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}, I\} = E_{\boldsymbol{\beta}}\{E_{\boldsymbol{\delta}\mathbf{r}}\{\boldsymbol{\delta}\mathbf{r}|\boldsymbol{\beta}, \mathbf{Y}, I\}|I\}$$
$$= -N_{GG}^{-1}N_{Gb}\frac{\sum_{\boldsymbol{\beta}_I} \boldsymbol{\beta}_I \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I}}{\sum_{\boldsymbol{\beta}_I} \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}_I^T N_r \boldsymbol{\beta}_I}} \tag{A36}$$

exactly as before in Eq. (12); for the floating case we have

$$\boldsymbol{\delta}\mathbf{r}_{BF} = E\{\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}, F\} = E_{\boldsymbol{\beta}}\{E_{\boldsymbol{\delta}\mathbf{r}}\{\boldsymbol{\delta}\mathbf{r}|\boldsymbol{\beta}, \mathbf{Y}, F\}|F\}$$
$$= -E_{\boldsymbol{\beta}}\{N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}|F\} = -N_{GG}^{-1}N_{Gb}E\{\boldsymbol{\beta}|F\} = 0 \tag{A37}$$

because

$$p(\boldsymbol{\beta}|\mathbf{Y}, F) \propto \mathrm{e}^{-\frac{1}{2\sigma_0^2}\boldsymbol{\beta}^T N_r \boldsymbol{\beta}} \approx N(0, \sigma_0^2 N_r^{-1}) \tag{A38}$$

Therefore, in the mixed model the Bayesian posterior mean of $\boldsymbol{\delta}\mathbf{r}$ is expressed by

$$\boldsymbol{\delta}\mathbf{r}_B = P_I E\{\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}, I\} + (1 - P_I)E\{\boldsymbol{\delta}\mathbf{r}|\mathbf{Y}, F\} = P_I \boldsymbol{\delta}\mathbf{r}_{BI} \tag{A39}$$

which coincides with Eq. (32), and which means that when the probability of obtaining the integer ambiguity is small the Bayesian estimator moves close to the floating estimator derived from the LS adjustment; when it is high it goes towards the pure integer estimator. Moreover, in this mixture model the posterior covariance matrix is given by the expression

$$C_{\boldsymbol{\delta}\mathbf{r}_B\boldsymbol{\delta}\mathbf{r}_B} = E\{\boldsymbol{\delta}\mathbf{r}\boldsymbol{\delta}\mathbf{r}^T|\mathbf{Y}\} - \boldsymbol{\delta}\mathbf{r}_B\boldsymbol{\delta}\mathbf{r}_B^T \tag{A40}$$

where

$$E\{\boldsymbol{\delta}\mathbf{r}\boldsymbol{\delta}\mathbf{r}^T|\mathbf{Y}\} = P_I E\{\boldsymbol{\delta}\mathbf{r}\boldsymbol{\delta}\mathbf{r}^T|\mathbf{Y}, I\} + (1 - P_I)E\{\boldsymbol{\delta}\mathbf{r}\boldsymbol{\delta}\mathbf{r}^T|\mathbf{Y}, F\} \tag{A41}$$

with

$$E\{\boldsymbol{\delta}\mathbf{r}\boldsymbol{\delta}\mathbf{r}^T|\mathbf{Y}, I\} = C_I + \boldsymbol{\delta}\mathbf{r}_{BI}\boldsymbol{\delta}\mathbf{r}_{BI}^T \tag{A42}$$

In the last formula C_I is the covariance matrix for the integer case and its expression is equal to Eq. (A12).

As for the floating case, we find

$$E\{\boldsymbol{\delta}\mathbf{r}\boldsymbol{\delta}\mathbf{r}^T|\mathbf{Y}, F\} = E_{\boldsymbol{\beta}}\{E_{\delta r}\{\boldsymbol{\delta}\mathbf{r}\boldsymbol{\delta}\mathbf{r}^T|\mathbf{Y}, \boldsymbol{\beta}, F\}|F\}$$
$$= E_{\boldsymbol{\beta}}\{\sigma_0^2 N_{GG}^{-1} + N_{GG}^{-1}N_{Gb}\boldsymbol{\beta}\boldsymbol{\beta}^T N_{bG}N_{GG}^{-1}|F\}$$
$$= \sigma_0^2(N_{GG}^{-1} + N_{GG}^{-1}N_{Gb}N_r^{-1}N_{bG}N_{GG}^{-1}) \tag{A43}$$

Combining Eqs. (A42) and (A43) and substituting into Eqs. (A41) and (A40), the covariance matrix in the mixture model is given by the somewhat more complicated formula

$$
\begin{aligned}
C &= P_I C_I + P_I \boldsymbol{\delta}\mathbf{r}_{BI}\boldsymbol{\delta}\mathbf{r}_{BI}^T + (1-P_I)\sigma_0^2 N_{GG}^{-1} \\
&\quad + (1-P_I)\sigma_0^2 N_{GG}^{-1} N_{Gb} N_r^{-1} N_{bG} N_{GG}^{-1} - P_I^2 \boldsymbol{\delta}\mathbf{r}_{BI}\boldsymbol{\delta}\mathbf{r}_{BI}^T \\
&= P_I C_I + (1-P_I)\sigma_0^2 N_{GG}^{-1} + (1-P_I) \\
&\quad \times (\sigma_0^2 N_{GG}^{-1} N_{Gb} N_r^{-1} N_{bG} N_{GG}^{-1} - P_I \boldsymbol{\delta}\mathbf{r}_{BI}\boldsymbol{\delta}\mathbf{r}_{BI}^T) \qquad \text{(A44)}
\end{aligned}
$$

which coincides with Eq. (34).

References

Betti B, Crespi M, Sansò F (1993) A geometric illustration of ambiguity resolution in GPS theory and a Bayesian Approach. Manuscr Geod 18: 317–330

Betti B, Carpino M, Crespi M, Marana B (1996) State of the art of the implementation of the new GPS data processing software BAMBA. Reports on Surveying and Geodesy, in memory of Prof. A. Gubellini and Prof. G. Folloni. Unguendoli M (ed) Edizioni Nautilus. Bolonia. pp. 237–264

Beutler G, Frei E (1990) Rapid static positioning based on the fast ambiguity resolution approach "FARA": theory and first results. Manuscr Geod 15(6): 325–356

Blewitt GP (1989) Carrier phase ambiguity resolution for the global positioning system applied to geodetic baselines up to 2000 km. J Geophys Res 94: 187–203

Box GEP, Tiao GC (1992) Bayesian inference in statistical analysis. John Wiley, New York

Gil AJ, Rodriguez-Caderot G, Lacy MC, Ruiz A (2001) Red GPS no Permanente para el control de Deformaciones en el sector Central de la cordillera bética. In: La Cuenca de Granada. (ed) CSIC. University of Granada (Spain), pp. 198–208

Gundich B, Koch KR (submitted) Confidence regions for GPS baselines by Bayesian statistics. J Geod

Jennings A (1977) Matrix computation for engineers and scientists. John Wiley, London

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) Numerical recipes in Fortran. The art of scientific computing, 2nd edn. Cambridge University Press, Cambridge

Rothacher M, Mervant L, Beutler G, Brockmann E, Fankhauser S, Gurtner W, Johnson J, Schaer S, Springer T, Weber R (1996) BERNESE GPS software version 4.0.

Sheldon MR (1997) Introduction to probability models, 6th edn. Academic Press, University of California, Berkeley

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

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

Teunissen PJG, Kleusberg A (1998) GPS for geodesy, 2nd edn. Springer, Berlin Heidelberg New York