

# Generation of three-dimensional random rotations in fitting and matching problems

Michael Habeck

Received: 9 January 2008 / Accepted: 23 April 2009 / Published online: 8 May 2009  
© The Author(s) 2009. This article is published with open access at Springerlink.com

**Abstract** An algorithm is developed to generate random rotations in three-dimensional space that follow a probability distribution arising in fitting and matching problems. The rotation matrices are orthogonally transformed into an optimal basis and then parameterized using Euler angles. The conditional distributions of the three Euler angles have a very simple form: the two azimuthal angles can be decoupled by sampling their sum and difference from a von Mises distribution; the cosine of the polar angle is exponentially distributed and thus straightforward to generate. Simulation results are shown and demonstrate the effectiveness of the method. The algorithm is compared to other methods for generating random rotations such as a random walk Metropolis scheme and a Gibbs sampling algorithm recently introduced by Green and Mardia. Finally, the algorithm is applied to a probabilistic version of the Procrustes problem of fitting two point sets and applied in the context of protein structure superposition.

**Keywords** Markov chain Monte Carlo · Random rotation · Euler angles · Von Mises distribution · Procrustes problem · Nearest rotation matrix

## 1 Introduction

In various probabilistic inference tasks, one faces the problem of generating a rotation matrix  $R$  from a probability density function:

---

M. Habeck (✉)

Department of Empirical Inference, Max-Planck-Institute for Biological Cybernetics,  
Spemannstr. 38, 72076 Tübingen, Germany  
e-mail: michael.habeck@tuebingen.mpg.de

M. Habeck

Department of Protein Evolution, Max-Planck-Institute for Developmental Biology,  
Spemannstr. 35, 72076 Tübingen, Germany

$$p(R|A) dR = \text{etr}(A^T R) dR \quad (1)$$

where  $A$  is some  $3 \times 3$  matrix and  $\text{etr}\{A\} = \exp\{\text{tr } A\}$ ;  $dR$  is the invariant (prior) measure on  $SO(3)$ , the group of all three-dimensional rotation matrices. When superimposing two rigid bodies or point sets by successive rotation and translation, the posterior probability density function of the rotation matrix has the above form. Recently, Green and Mardia derived such a posterior distribution in the context of protein structure alignment (Green and Mardia 2006). The same distribution occurs in a probabilistic solution of the Procrustes problem (Theobald and Wuttke 2006). There are also applications in computer vision and robotics: Bayesian three-dimensional pose estimation, for example, involves a distribution of the same functional form (see e.g. Choi et al. 2002).

The aim of a truly Bayesian analysis is to not only provide parameter estimates but also quantify their uncertainty (MacKay 2003). The best method to obtain such uncertainties when an analytical treatment is infeasible, is to generate random samples from the full posterior distribution. In recent years, a number of Markov chain Monte algorithms have been developed to make this approach practical (e.g. Chen et al 2002; MacKay 2003).

In their article, Green and Mardia (2006) present an algorithm for sampling three-dimensional rotation matrices from the above distribution (1). However, their sampling algorithm appears to be suboptimal in two respects: the parameters (Euler angles) are coupled, and the polar angle cannot be drawn from a known distribution but has to be generated using a random walk Metropolis algorithm.

In this article, a Gibbs sampling algorithm is developed to generate rotation matrices from probability distribution (1) without resorting to a Metropolis algorithm. The rotation matrices are orthogonally transformed into an optimal basis and then parameterized using Euler angles. The resulting distribution of the Euler angles is much easier to simulate. Log-probability plots and an autocorrelation analysis indicate that the algorithm converges very rapidly when compared to a random walk Metropolis scheme and to the algorithm of Green and Mardia. An application in structural bioinformatics is discussed.

## 2 Preliminaries

A rotation matrix is column and row orthogonal, i.e.  $R^T R = R R^T = I$ , where  $I$  is the identity matrix. Furthermore the determinant of  $R$  is one:  $|R| = 1$ .

### 2.1 Maximally probable rotation

We first discuss the problem of maximizing the probability density function (Eq. 1) over the space of all rotations in  $n$ -dimensional space. This problem can be viewed as a matrix nearness problem: let  $\|A\|_F = \sqrt{\sum_{ij} A_{ij}^2}$  denote the Frobenius or Euclidean norm for arbitrary  $n \times m$  matrices  $A$ . We have:  $\|A\|_F^2 = \text{tr}(A^T A)$ . Then  $\|A - B\|_F$  is the distance between two matrices  $A$  and  $B$ .

Let us now consider square matrices, i.e.  $m = n$ . The rotation matrix that is nearest to  $A$  can be found by minimizing the distance between  $A$  and  $R$  under the constraint that  $R$  must be a rotation matrix. In this case, the matrix distance is:

$$\|A - R\|_F^2 = \|A\|_F^2 + n - 2\text{tr}(A^T R), \quad R \in SO(n).$$

That is, solving the matrix nearness problem of finding the rotation matrix that is closest to some arbitrary  $n \times n$  matrix  $A$  is equivalent to maximizing the overlap or scalar product  $\text{tr}(A^T R)$ , which is the logarithm of the probability density function (Eq. 1). Therefore, maximization of the probability  $p(R|A)$  is equivalent to solving the nearest rotation matrix problem. This matrix nearness problem can be solved by calculating the singular value decomposition of  $A$ .

A singular value decomposition of  $A$  yields the factorization  $A = U\Lambda V^T$  where  $U$  and  $V$  are orthogonal  $n \times n$  matrices and  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  is the diagonal matrix of singular values with  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ . The nearest orthogonal matrix is  $R_{\max} = UV^T$  (Higham 1989). If  $|UV^T| = -1$ , the nearest orthogonal matrix is not a pure rotation but also involves a reflection. The nearest rotation matrix is then found by replacing  $U$  with  $U(I - 2e_n e_n^T)$  and  $\lambda_n$  with  $-\lambda_n$  where  $e_n$  is the basis vector pointing into the direction that corresponds to the smallest singular value. If for the original matrices  $|UV^T| = 1$ , the maximal overlap between  $A$  and a rotation matrix is  $\max\{\text{tr}(A^T R)\} = \sum_i \lambda_i$ . This value is diminished by  $2\lambda_n$ , if the nearest orthogonal matrix involves a reflection. In the following, we can therefore assume that  $|U| = 1$  and  $|V| = 1$  and  $\lambda_1 \geq \lambda_2 \geq \dots \geq |\lambda_n| \geq 0$ .

Due to the outlined equivalence between the matrix distance and the overlap, the probability distribution Eq. (1) could also be written

$$p(R|A) = \exp\left\{-\frac{1}{2}\|A - R\|_F^2\right\}.$$

Therefore  $p(R|A)$  can be viewed as an analog of a Gaussian distribution for rotation matrices.

### 2.2 Reparameterization

We parameterize the rotation matrix  $R$  in terms of deviations from the maximally probable rotation  $R_{\max}$ . We choose the parameterization  $R = USV^T$  where  $S$  again is a rotation matrix. This parameterization guarantees that  $R$  is a rotation matrix if  $S$  is one. For the special case  $S = I$ , the rotation  $R$  is identical to the maximally probable rotation; the more  $S$  differs from  $I$ , the more  $R$  will deviate from the maximally probable rotation.

To substitute the rotation matrix  $R$  with the rotation matrix  $S$  using the described orthogonal transformation, we need to calculate the Jacobian of this parameter transformation, i.e. we have to calculate the determinant of the matrix  $\partial \text{vec}[R]/\partial \text{vec}[S]$  where  $\text{vec}[A]$  is the  $n^2$ -dimensional vector obtained by stacking the columns of a  $n \times n$

matrix  $A$ . We have (Horn and Johnson 1991):

$$(V \otimes U) \text{vec}[S] = \text{vec}[R]$$

where  $\otimes$  denotes the Kronecker product. It is now straightforward to see that

$$\frac{\partial \text{vec}[R]}{\partial \text{vec}[S]} = V \otimes U,$$

and the Jacobian becomes

$$\left| \frac{\partial \text{vec}[S]}{\partial \text{vec}[R]} \right| = |V \otimes U| = |V|^n |U|^n = 1.$$

Using the new parameterization we have:

$$p(R|A)dR = p(S|\Lambda)dS = \exp\left(\sum_i \lambda_i S_{ii}\right) dS \tag{2}$$

We consider the probability  $p(S|\Lambda)$  as a standardized form of the more general distribution (Eq. 1). Once we are able to generate  $S$  from  $p(S|\Lambda)$ , we obtain a sample from  $p(R|A)$  by letting  $R = USV^T$ .

### 2.3 Euler angles

So far, the previous remarks are valid in spaces of arbitrary dimension  $n$ . In practical applications, the three-dimensional case deserves special attention because it occurs in the widespread problem of fitting, matching or superimposing two rigid bodies represented by a set of corresponding objects such as landmarks, atoms, and vertices.

We parameterize  $S$  in Euler angles using a  $zyz$  representation:

$$S(\alpha, \beta, \gamma) = R_z(\gamma)R_y(\beta)R_z(\alpha) \tag{3}$$

with azimuthal angles  $\alpha, \gamma \in [0, 2\pi]$  and polar angle  $\beta \in [0, \pi]$ ;  $R_z$  and  $R_y$  are rotations about the  $z$  and the  $y$  axis, respectively. Explicitly we have:

$$S(\alpha, \beta, \gamma) = \begin{pmatrix} \cos \alpha \cos \beta \cos \gamma - \sin \alpha \sin \gamma & \sin \alpha \cos \beta \cos \gamma + \cos \alpha \sin \gamma & -\sin \beta \cos \gamma \\ -\cos \alpha \cos \beta \sin \gamma - \sin \alpha \cos \gamma & -\sin \alpha \cos \beta \sin \gamma + \cos \alpha \cos \gamma & \sin \beta \sin \gamma \\ \cos \alpha \sin \beta & \sin \alpha \sin \beta & \cos \beta \end{pmatrix}. \tag{4}$$

Above, we found that  $dR = dS$ . The invariant measure of the rotation group parameterized in Euler angles is (Miles 1965):

$$dS = \frac{1}{8\pi^2} \sin \beta \, d\alpha \, d\beta \, d\gamma. \tag{5}$$

This is the natural prior distribution over elements of the rotation group (Miles 1965; Hartigan 1964).

Using the above results, we find that the task of generating a rotation from probability distribution (1) boils down to generating Euler angles from:

$$f(\alpha, \beta, \gamma) = \text{etr}\{\mathbf{\Lambda}S(\alpha, \beta, \gamma)\} \sin \beta. \tag{6}$$

We use a Gibbs sampling scheme (Geman and Geman 1984), for this task: we successively generate the azimuthal Euler angles from  $f(\alpha, \gamma|\beta)$  and the polar angle from  $f(\beta|\alpha, \gamma)$ .

### 2.4 Conditional distribution of the azimuthal angles

To calculate the distribution of the azimuthal angles  $\alpha$  and  $\gamma$  conditioned on the polar angle  $\beta$  we have to evaluate  $\text{tr}(\mathbf{\Lambda}S)$  in the Euler parameterization. We obtain:

$$\text{tr}(\mathbf{\Lambda}S) = (\lambda_1 \cos \beta + \lambda_2) \cos \alpha \cos \gamma - (\lambda_1 + \lambda_2 \cos \beta) \sin \alpha \sin \gamma + \lambda_3 \cos \beta. \tag{7}$$

Since  $f(\alpha, \beta, \gamma)$  depends on  $\alpha$  and  $\gamma$  only through the products

$$\cos \alpha \cos \gamma = [\cos(\alpha + \gamma) + \cos(\alpha - \gamma)]/2$$

and

$$-\sin \alpha \sin \gamma = [\cos(\alpha + \gamma) - \cos(\alpha - \gamma)]/2$$

it is convenient to work with new angles

$$\Phi = \alpha + \gamma, \quad \Psi = \alpha - \gamma.$$

The Jacobian of the parameter transformation is

$$\left| \frac{\partial(\Phi, \Psi)}{\partial(\alpha, \gamma)} \right| = 2$$

and the domain  $[0, 2\pi] \times [0, 2\pi]$  transforms to  $\Omega = \cup_{i=1}^4 \Omega_i$  where

$$\begin{aligned} \Omega_1 &= \{(\Phi, \Psi) \mid 0 \leq \Phi \leq 2\pi, 0 \leq \Psi \leq \Phi\}, \\ \Omega_2 &= \{(\Phi, \Psi) \mid 2\pi \leq \Phi \leq 4\pi, 0 \leq \Psi \leq 4\pi - \Phi\}, \\ \Omega_3 &= \{(\Phi, \Psi) \mid 0 \leq \Phi \leq 2\pi, -\Phi \leq \Psi \leq 0\}, \\ \Omega_4 &= \{(\Phi, \Psi) \mid 2\pi \leq \Phi \leq 4\pi, -4\pi + \Phi \leq \Psi \leq 0\}. \end{aligned}$$

The inequality constraints on the allowed values of  $\Phi$  and  $\Psi$  are cumbersome to work with. A more convenient parameterization is to introduce two angular variables  $\varphi, \psi \in [0, 2\pi]$  with associated binary variables  $v, w \in \{0, 1\}$  such that

$$\Phi = \varphi + 2\pi v, \quad \Psi = \psi + 2\pi w.$$

The four domains  $\Omega_i$  then correspond to:

$$\begin{aligned} 0 \leq \varphi \leq 2\pi, \quad 0 \leq \psi \leq \varphi, \quad v = 0, \quad w = 0; \\ 0 \leq \varphi \leq 2\pi, \quad \varphi \leq \psi \leq 2\pi, \quad v = 1, \quad w = 1; \\ 0 \leq \varphi \leq 2\pi, \quad 0 \leq 2\pi - \psi \leq \varphi, \quad v = 0, \quad w = 1; \\ 0 \leq \varphi \leq 2\pi, \quad \varphi \leq 2\pi - \psi \leq 2\pi, \quad v = 1, \quad w = 0. \end{aligned}$$

The original Euler angles are obtained by letting

$$\alpha = \frac{\varphi + \psi}{2} + \pi(v - w), \quad \gamma = \frac{\varphi - \psi}{2} + \pi(v + w).$$

Due to the periodicity of  $f(\alpha, \beta, \gamma)$  in  $\alpha$  and  $\gamma$ , we can combine both cases where  $v = w$  and obtain:

$$(\varphi, \psi) \in [0, 2\pi] \times [0, 2\pi], \quad \alpha = \frac{\varphi + \psi}{2}, \quad \gamma = \frac{\varphi - \psi}{2}.$$

Similarly, we can combine the cases where  $v + w = 1$ :

$$(\varphi, \psi) \in [0, 2\pi] \times [0, 2\pi], \quad \alpha = \frac{\varphi + \psi}{2} + \pi, \quad \gamma = \frac{\varphi - \psi}{2} + \pi.$$

Thus we can use the unconstrained angles  $\varphi, \psi$  defined on  $[0, 2\pi] \times [0, 2\pi]$  and a binary variable  $u = |v - w| \in \{0, 1\}$  such that the azimuthal angles are

$$\alpha = \frac{\varphi + \psi}{2} + \pi u, \quad \gamma = \frac{\varphi - \psi}{2} + \pi u. \quad (8)$$

After replacing  $\alpha$  and  $\gamma$  in  $\text{tr}(\mathbf{\Lambda}S)$  (Eq. 7) with  $\varphi$  and  $\psi$ , we observe that the conditional probability of the new angles is the product of two von Mises distributions:

$$f(\varphi, \psi | \beta) = \exp\{\kappa_\varphi(\beta) \cos \varphi\} \exp\{\kappa_\psi(\beta) \cos(\psi - \pi)\} \quad (9)$$

where

$$\kappa_\varphi(\beta) = \cos^2(\beta/2)(\lambda_1 + \lambda_2) \geq 0 \quad (10)$$

and

$$\kappa_\psi(\beta) = \sin^2(\beta/2)(\lambda_1 - \lambda_2) \geq 0. \quad (11)$$

That is, the sum and the difference of the azimuthal Euler angles  $\alpha$  and  $\gamma$  are conditionally independent. The sum of the two angles,  $\varphi$ , follows a von Mises distribution with mean angle 0 and shape parameter  $\kappa_\varphi$  (Eq. 10). The difference angle  $\psi$  follows a

von Mises distribution with mean angle  $\pi$  and shape parameter  $\kappa_\psi$  (Eq. 11). The generation of a random sample from a von Mises distribution is straightforward (Devroye 1986).

### 2.5 Conditional distribution of the polar angle

The conditional distribution of the polar angle  $\beta$  is proportional to:

$$f(\beta|\alpha, \gamma) = \sin \beta \exp \{ \kappa_\beta \cos \beta/2 \} \tag{12}$$

where  $\beta \in [0, \pi]$  and in the  $\varphi/\psi$  parameterization

$$\kappa_\beta = (\lambda_1 + \lambda_2) \cos \varphi + (\lambda_1 - \lambda_2) \cos \psi + 2\lambda_3. \tag{13}$$

Distribution (12) is not a von Mises distribution because it is defined on  $[0, \pi]$  and has an additional  $\sin \beta$  factor.

However, the generation of a random sample from this distribution is straightforward. Let us substitute  $\cos \beta$  with  $x \in [-1, 1]$ , then:

$$f(x|\alpha, \gamma) = \exp\{\kappa_\beta x/2\}$$

with distribution function:

$$F(x|\alpha, \gamma) = \int_{-1}^x dt f(t|\alpha, \gamma) \bigg/ \int_{-1}^1 dt f(t|\alpha, \gamma) = \frac{\exp(\kappa_\beta x/2) - \exp(-\kappa_\beta/2)}{\exp(\kappa_\beta/2) - \exp(-\kappa_\beta/2)}.$$

We generate a random sample  $x$  by generating a uniformly distributed random number  $r$  in  $[0, 1]$  and then obtain  $x$  by inversion of  $r = F(x|\alpha, \gamma)$ . By letting  $\beta = \arccos x$  one obtains a random sample of the polar angle.

### 2.6 Marginal distribution of the polar angle

The marginal probability distribution of the polar angle  $\beta$  can now be calculated straightforwardly up to a normalization constant:

$$\begin{aligned} f(\beta) &= \int d\varphi d\psi \sin \beta \exp \{ \lambda_3 \cos \beta + \kappa_\varphi(\beta) \cos \varphi + \kappa_\psi(\beta) \cos(\psi - \pi) \} \\ &= 4\pi^2 \sin \beta e^{\lambda_3 \cos \beta} I_0 \left[ (\lambda_1 + \lambda_2) \cos^2 \left( \frac{\beta}{2} \right) \right] I_0 \left[ (\lambda_1 - \lambda_2) \sin^2 \left( \frac{\beta}{2} \right) \right] \end{aligned} \tag{14}$$

where  $I_0$  is the zeroth order modified Bessel function of the first kind.

### 3 Algorithm

Our algorithm to sample rotation matrices from  $p(R|A)$  (Eq. 1) works as follows:

1. Optimal basis: calculate the singular value decomposition of  $A$ , scale the last column of  $U$  such that the determinant of  $UV^T$  is one; replace the smallest singular value  $\lambda_3$  with  $-\lambda_3$  if necessary.
2. Gibbs sampling: generate random Euler angles  $\alpha, \beta, \gamma$  by iterating through
  - (a) sample  $\alpha$  and  $\gamma$ :

$$\varphi \sim \text{vonMises}(0, \kappa_\varphi)$$

$$\psi \sim \text{vonMises}(\pi, \kappa_\psi)$$

$$u \sim \text{Bernoulli}(1/2)$$

where  $\kappa_\varphi$  and  $\kappa_\psi$  depend on the previous value of  $\beta$  (initial value  $\beta = 0$  in the first iteration) and are defined in Eqs. (10) and (11), respectively. Then  $\alpha = (\varphi + \psi)/2 + \pi u$  and  $\gamma = (\varphi - \psi)/2 + \pi u$ .<sup>1</sup>

- (b) sample  $\beta$ :

- $\kappa_\beta = (\lambda_1 + \lambda_2) \cos \varphi + (\lambda_1 - \lambda_2) \cos \psi + 2\lambda_3$
- generate a uniformly distributed random number  $r$ :

$$r \sim \text{Uniform}(0, 1)$$

- calculate  $x$ :

$$x = 1 + 2 \log\{r + (1 - r)e^{-\kappa_\beta}\}/\kappa_\beta$$

- calculate  $\beta$ :

$$\beta = \arccos x$$

3. Output: obtain an approximate sample from distribution (1) by letting  $R = US(\alpha, \beta, \gamma)V^T$  with  $S(\alpha, \beta, \gamma)$  defined in Eq. (4).

### 4 Results

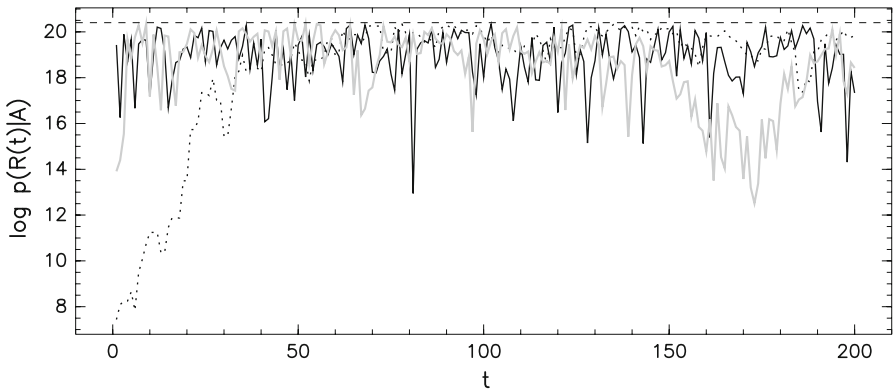
#### 4.1 Performance and comparison to other algorithms

First we test our algorithm for a simple case where  $A$  is chosen randomly:

$$A = 10^{-1} \begin{pmatrix} 85 & 78 & 43 \\ 11 & 39 & 64 \\ 41 & 60 & 48 \end{pmatrix}$$

<sup>1</sup> Instead of shifting  $\alpha$  and  $\gamma$  we could flip the sign of the polar angle  $\beta$  and obtain the same rotation.





**Fig. 1** Log–probability traces  $\log p(R|A) = \text{tr}[A^T R(t)]$  for three sampling algorithms: results obtained with the algorithm outlined in this article are shown as *black line*, random walk Metropolis scheme: *dotted black line*, algorithm of Green and Mardia: *thick grey line*. The *dashed horizontal line* indicates the maximum of the log–probability

with singular values 16.17, 4.80, and 0.57. Since the determinants of the orthogonal matrices  $U$  and  $V$  have opposite signs, we have to flip the sign of the last singular value and of the last column of one orthogonal matrix to obtain a valid rotation matrix. The maximum of the logarithm of  $p(R|A)$  is  $\text{tr } \Lambda = 20.4$ .

We generated 10,000 rotation matrices using our algorithm, a random walk Metropolis scheme and the algorithm outlined in Green and Mardia (2006). In the Metropolis scheme and in the algorithm of Green and Mardia, the rotation matrix  $R$  is parameterized in Euler angles without transforming it into the optimal basis. In the Metropolis scheme, a single proposal step consists in perturbing the Euler angles with an increment that is uniformly chosen from the interval  $[-0.2, 0.2]$  which was tuned so as to obtain an acceptance rate of 70 %.

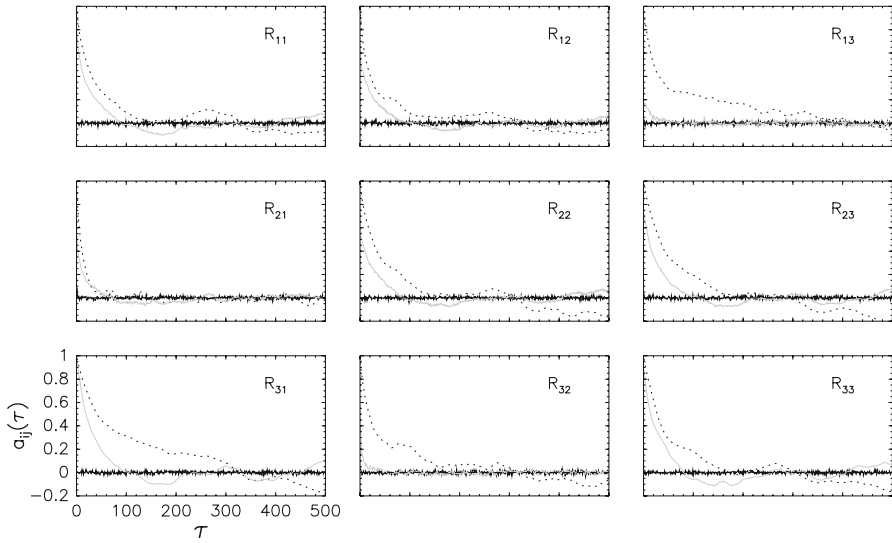
Figure 1 shows traces of the first 200 sampled log-probability values  $\text{tr}(A^T R)$  obtained with the three algorithms. It is evident from this figure that the new algorithm converges almost instantaneously and outperforms the other algorithms.

To further investigate the sampling properties, an autocorrelation analysis was carried out. For each element  $R_{ij}$  of the rotation matrix, an autocorrelation curve  $a_{ij}(\tau)$  was calculated over the samples. Figure 2 shows  $a_{ij}(\tau)$  for the three sampling algorithms investigated here. Again, the new algorithm’s performance is superior.

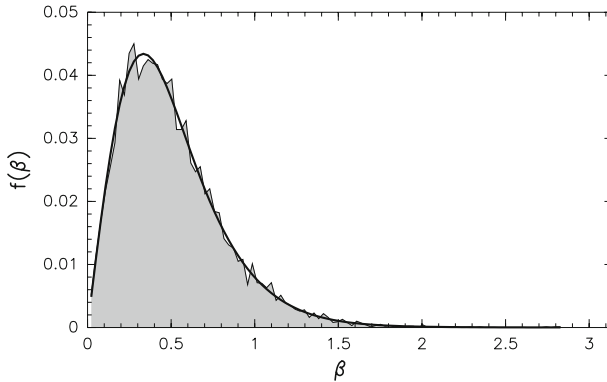
As another indicator we investigate the marginal distribution of the polar angle  $f(\beta)$  (Eq. 14). Figure 3 shows this distribution and the histogram of the polar angle sampled with our algorithm. Again, the very good agreement between the theoretical and simulated distribution highlights the excellent performance of the new algorithm.

#### 4.2 Application to a Procrustes problem: protein structure superposition

In the orthogonal Procrustes problem, one minimizes  $\|Y - XR^T\|_F$  subject to the constraint that  $R$  must be an element of  $SO(3)$  (Higham 1989);  $X$  and  $Y$  are two  $n \times 3$



**Fig. 2** Autocorrelation analysis: for all three simulations, the sample autocorrelation curves of all nine elements of the rotation matrix were calculated and are plotted (matrix element indicated as plot label). Autocorrelation curves obtained with the algorithm outlined in this article are shown as *black lines*, random walk Metropolis scheme: *dotted black line*, algorithm of Green and Mardia: *thick grey line*



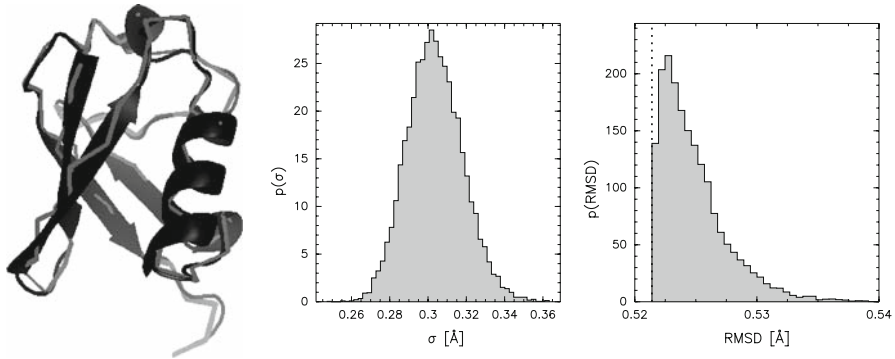
**Fig. 3** Marginal distribution of the polar angle  $f(\beta)$ . The *thick black curve* is the analytical curve given in Eq. (14), the *grey histogram* is the distribution obtained from the sampled polar angles

centered configuration matrices whose rows  $x_i$  and  $y_i$ , respectively, are the positions of some objects (atoms, landmarks, etc.) in three-dimensional space.

Let us first derive a probabilistic formulation of the Procrustes problem (Theobald and Wuttke 2006). The generative model is:

$$y_i = Rx_i + e_i, \quad e_i \sim \text{Normal}(0, \sigma^2 I)$$

where the error vector  $e_i$  is isotropically Gaussian distributed with standard deviation  $\sigma$ . The likelihood function is:



**Fig. 4** Analysis of the Ubiquitin structures 1UBQ and 1D3Z: The *left panel* shows the crystal structure 1UBQ as *black ribbon* onto which the NMR structure 1D3Z was superimposed. The different orientations are visualized as an ensemble of structures (*gray lines*). The *middle panel* shows the distribution of the standard deviation  $\sigma$ . The *right panel* shows the distribution of the RMSD values (the *dotted line* indicates the optimal RMSD value of 0.5214 Å)

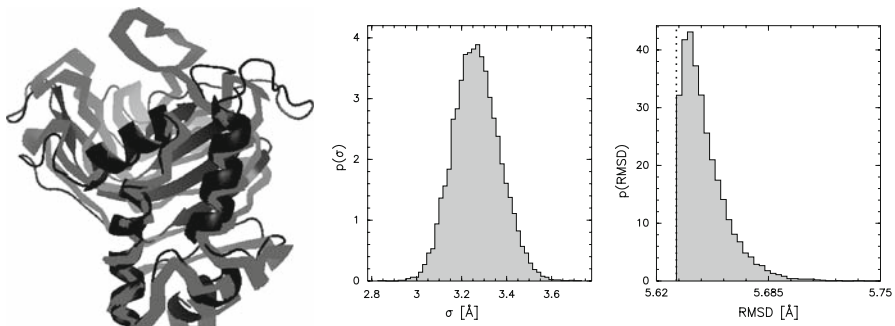
$$\begin{aligned}
 L(R, \sigma) &= \prod_i \text{Normal}(y_i - Rx_i, \sigma^2 I) \\
 &= (2\pi\sigma^2)^{-3n/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \|y_i - Rx_i\|^2 \right\} \\
 &= (2\pi\sigma^2)^{-3n/2} \exp \left\{ -\frac{1}{2\sigma^2} \|Y - XR^T\|_F^2 \right\}
 \end{aligned}$$

and the conditional posterior probability of the rotation matrix  $p(R|\sigma)$  is of the form (1) with  $A = Y^T X / \sigma^2$ . The conditional posterior distribution of  $\sigma^{-2}$  is a Gamma distribution:  $\sigma^{-2} \sim \text{Gamma}(3n/2, \|Y - XR^T\|_F^2 / 2)$ .

A probabilistic comparison of two protein structures  $X$  and  $Y$  is based on a Gibbs sampler which successively draws a rotation matrix  $R$  and a standard deviation  $\sigma$  from their conditional posterior distributions. The traditional measure to compare protein structures is the root mean square deviation (RMSD) of  $C\alpha$  positions (Kabsch 1976). The RMSD is defined as

$$\text{RMSD} = \min_{R \in SO(3)} \|Y - XR^T\|_F / \sqrt{n}.$$

We apply the probabilistic Procrustes analysis to two pairs of protein structures, one with low RMSD and one with high RMSD value. The structures were downloaded from the Protein Data Bank (PDB) (Berman et al. 2000). The first pair is a structure of the protein Ubiquitin obtained with X-ray crystallography (PDB accession code: 1UBQ) and with nuclear magnetic resonance (NMR) spectroscopy (PDB code: 1D3Z). The second pair is the apo and holo form of the hemophore HasA (PDB codes 1YBJ and 1B2V) obtained with NMR and with X-ray crystallography, respectively. For both pairs, 10,000 rotations and standard deviations were generated using the outlined Gibbs sampler. The results are shown in Figs. 4 and 5.



**Fig. 5** Analysis of the HasA structures 1YBJ and 1B2V: same panels as in Fig. 4. Now, the apo form of HasA 1YBJ is shown as ribbon, the ensemble corresponds to the holo form. The optimal RMSD value is 5.63 Å

Figure 4 shows a comparison of the Ubiquitin structures. The crystal and the NMR structure match quite well, which is reflected in a small variability of the sampled orientations (indicated by the tight structure ensemble). As a consequence, the standard deviation is estimated to be rather small and exhibits a narrow posterior distribution.

In contrast, the HasA structures match less well due to a large conformational change in one of the loops (cf. Fig. 5). As a consequence, the  $\sigma$  distribution is much broader and centered at a value which is a magnitude larger than for the Ubiquitin structures. The same holds for the distribution of RMSD values.

## 5 Conclusion

This article outlines an algorithm for generating three-dimensional rotations from probability distributions of the form  $\text{etr}\{A^T R\}$  (Eq. 1). The algorithm relies on an efficient parameterization of the rotations: the basis system is chosen such that the rotation matrix with maximal probability corresponds to the identity matrix. Deviations from the optimal rotation are parameterized in Euler angles. The main advantage of the new algorithm when compared to existing methods is that the conditional distributions of the Euler angles are maximally decoupled and of a standard form for which random number generators exist. Therefore, it is not necessary to resort to a Metropolis algorithm. All checks are indicative of a superior performance when compared to other sampling algorithms.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. *Nucleic Acids Res* 28:235–242

- Chen MH, Shao QM, Ibrahim JG (2002) Monte Carlo Methods in Bayesian Computation. Springer Verlag, Inc., New York
- Choi KN, Carcassoni M, Hancock ER (2002) Recovering facial pose with the EM algorithm. *Pattern Recognit* 35:2073–2093
- Devroye L (1986) Non-uniform random variate generation. Springer, New York
- Geman S, Geman D (1984) Stochastic relaxation, gibbs distributions, and the Bayesian restoration of images. *IEEE Trans PAMI* 6(6):721–741
- Green PJ, Mardia KV (2006) Bayesian alignment using hierarchical models, with applications in protein bioinformatics. *Biometrika* 93:235–254
- Hartigan J (1964) Invariant prior distributions. *Ann Math Stat* 35:836–845
- Higham NJ (1989) Matrix nearness problems and applications. In: Gover MJC, Barnett S (eds) Applications of matrix theory. Oxford University Press, Oxford pp 1–27
- Horn RA, Johnson CR (1991) Topics in matrix analysis. Cambridge University Press, Cambridge
- Kabsch W (1976) A solution for the best rotation to relate two sets of vectors. *Acta Cryst A* 32:922–923
- MacKay DJC (2003) Information theory, inference, and learning algorithms. Cambridge University Press, Cambridge
- Miles RE (1965) On random rotations in  $R^3$ . *Biometrika* 52:636–639
- Theobald DL, Wuttke DS (2006) Empirical Bayes hierarchical models for regularizing maximum likelihood estimation in the matrix Gaussian Procrustes problem. *Proc Natl Acad Sci USA* 103:18,521–18,527