Bayes One-Sample and One-Way Random Effects Analyses for 3-D Orientations with Application to Materials Science

Melissa A. Bingham[*], Stephen B. Vardeman[†] and Daniel J. Nordman[‡]

Abstract. We consider Bayes inference for a class of distributions on orientations in 3 dimensions described by 3×3 rotation matrices. Non-informative priors are identified and Metropolis-Hastings within Gibbs algorithms are used to generate samples from posterior distributions in one-sample and one-way random effects models. A simulation study investigates the performance of Bayes analyses based on non-informative priors in the one-sample case, making comparisons to quasi-likelihood inference. A second simulation study investigates the behavior of posteriors for some informative priors. Bayes one-way random effect analyses of orientation matrix data are then developed and the Bayes methods are illustrated in a materials science application.

Keywords: Bayes, credible intervals, Electron backscatter diffraction, Gibbs, Markov chain Monte Carlo, Metropolis-Hastings, one-way random effects model, orthogonal matrix, posterior density, prior distribution, UARS distribution

1   Introduction

This paper concerns Bayes inference for 3-dimensional orientations. Random observations $\mathbf{O}_1, \ldots, \mathbf{O}_n$ in the form of 3×3 orthogonal rotation matrices that preserve the right hand rule are common in many fields, including vectorcardiography (Downs 1972) and human kinetics (Rancourt, Rivest, and Asselin 2000). Such data are also important in materials science applications and arise as the output of an Electron backscatter diffraction (EBSD) machine. EBSD is used to study the microtexture of crystalline materials, such as metals (see Randle 2003). An EBSD camera, coupled with a Scanning Electron Microscope, produces diffraction patterns (when atomic planes within a target material diffract a stationary beam of electrons) which are converted to crystal orientations relative to some reference coordinate system. Each resulting observation, expressed as a 3×3 rotation matrix, gives the orientation of a cubic crystal at some scanned position on a specimen surface. As in Bingham, Nordman, and Vardeman (2009b), our motivating application involves a machine precision problem, where interest lies in quantifying the variation in orientation measurements obtained when EBSD is used repeatedly at a single location as well as the variation of measured orientations within a "grain" of

---

[*]Department of Mathematics, University of Wisconsin-La Crosse, La Crosse, WI, mailto:bingham.meli@uwlax.edu

[†]Department of Statistics and
Department of Industrial and Manufacturing Engineering, Iowa State University, Ames, IA, http://www.public.iastate.edu/~vardeman/

[‡]Department of Statistics, Iowa State University, Ames, IA, http://dnordman.public.iastate.edu/

scanned points on a metal specimen. (A grain is a homogeneous "piece" of material that produces observations which generally share a common orientation.) While EBSD precision figures are available in the materials science literature (Demirel, El-Dasher, Adams, and Rollett 2000; Wilson and Spanos 2001), the methods currently used to produce these are entirely descriptive and somewhat ad hoc in terms of how measured orientations are converted to precision statements. Further, EBSD precisions given in the literature are based on a single scan across a homogeneous specimen and do not involve repeat readings at a given location. Bingham et al. (2009b) introduced a class of distributions on orientations in 3 dimensions and developed a (quasi-likelihood-based) inference framework for the EBSD precision problem. We extend the one-sample analyses presented in Bingham et al. (2009b) by considering Bayes inference for 3-dimensional orientations in one-sample and one-way random effects models.

Bingham et al. (2009b) identified the Uniform Axis-Random Spin (UARS) class of distributions on the set of 3×3 orthogonal matrices, Ω. This class is useful for modeling the deviation of random orientations $\mathbf{O}_i \in \Omega$ from a common "central location" orientation of $\mathbf{S} \in \Omega$, here referred to as the true or principal orientation. A random rotation $\mathbf{O} \in \Omega$ from a UARS distribution with principal orientation $\mathbf{S}$ can be written as $\mathbf{O} = \mathbf{SP}$, where

$$\mathbf{P} = \mathbf{U}\mathbf{U}^T + (\mathbf{I}_{3\times 3} - \mathbf{U}\mathbf{U}^T)\cos r + \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix} \sin r \in \Omega \qquad (1)$$

comes about by rotating $\mathbf{I}_{3\times 3}$ (the 3×3 identity matrix) about an axis $\mathbf{U} = (u_1, u_2, u_3)^T \in \mathbb{R}^3$ (identified by a point uniformly distributed on the unit sphere) by a random angle $r \in (-\pi, \pi]$ (independent of $\mathbf{U}$). The angle r is assumed to follow a circular distribution on $(-\pi, \pi]$ that is symmetric about 0 with spread depending on a concentration parameter $\kappa \geq 0$. The quantity $|r|$ is sometimes referred to as a "misorientation angle" (Randle 2003), and small realizations of $|r|$ imply that a rotation $\mathbf{O}$ from a UARS model will deviate little from the location parameter $\mathbf{S}$. Since the parameter κ controls the spread of this circular distribution for r, it consequently controls the variation of a corresponding UARS observation $\mathbf{O}$ from the true direction $\mathbf{S}$. Therefore, for a particular family of circular distributions for r, the resulting subclass of UARS distributions on Ω have "location" parameter $\mathbf{S} \in \Omega$ and "spread" parameter $\kappa \in \mathbb{R}$ and will be denoted by $\mathrm{UARS}(\mathbf{S}, \kappa)$.

In the literature, the most studied distribution for rotation matrices has been the matrix Fisher (or Langevin) distribution, introduced by Downs (1972) and further investigated by other authors (Khatri and Mardia 1977; Prentice 1986; Mardia and Jupp 2000; Rancourt et al. 2000; Chikuse 2003). Jupp and Mardia (1979) studied maximum likelihood estimation for these distributions, but the inference is not simple and the usual parameterization is not so easily interpreted as that of UARS models. For estimation of a location parameter $\mathbf{S}$, Chang and Rivest (2001) and Rivest and Chang (2006) have also discussed approaches involving M- and regression-type estimators. While some of these papers have considered the large sample properties of maximum likelihood or moment estimators in some models, little consideration has been given to Bayes methods.

Our goal is to investigate Bayes one-sample inference for the parameters of the UARS($\mathbf{S}, \kappa$) distributions and to further develop Bayes one-way random effects analyses based on this model that may be used to quantify EBSD precision. The flexible two-parameter model for rotation matrices used here allows us to develop useful priors for several inference scenarios where the resulting Bayes estimators are extremely effective. For clarity in exposition and concreteness, we shall focus on one particular UARS model (that we will call the vM-UARS($\mathbf{S}, \kappa$) class of distributions), for which r follows the von Mises circular distribution with direction 0 and concentration parameter $\kappa \geq 0$ (denoted by vM($0, \kappa$)). However, we point out that the Bayes methods introduced here can also be applied to other members of the UARS class based on alternative distributions for r (for example, see Bingham, Nordman, and Vardeman (2009a) for Bayes inference for the symmetric matrix von Mises-Fisher distributions). The vM($0, \kappa$) distribution for r is unimodal on $(-\pi, \pi]$, symmetric about 0, converges to the uniform distribution on the circle as $\kappa \to 0$, and becomes approximately Normal with mean 0 and variance $1/\kappa$ as $\kappa \to \infty$ (see Mardia and Jupp 2000).

To verify that the vM-UARS($\mathbf{S}, \kappa$) model is useful for describing the measured crystal orientations, fifty observations from within a single grain of a nickel specimen were fit using both the vM-UARS($\mathbf{S}, \kappa$) and matrix Fisher distributions (see Downs 1972). For orientations $\mathbf{o}_1, \ldots, \mathbf{o}_n \in \Omega$, the value of $\mathbf{S}$ that maximizes $tr(\mathbf{S}^T \bar{\mathbf{o}})$, for $tr(\mathbf{A})$ the trace of $\mathbf{A}$ and $\bar{\mathbf{o}} = \sum_{i=1}^{n} \mathbf{o}_i/n$, is commonly used as an estimate for $\mathbf{S}$ (León et al. 2006; Khatri and Mardia 1977). The misorientation angle $|r|$ between this standard moment estimator and each observation was found, and the distributions on $|r|$ corresponding to the vM-UARS($\mathbf{S}, \kappa$) and matrix Fisher distributions were fit to the sample misorientation angles using maximum likelihood. Figure 1 shows the sample quantiles plotted against the theoretical quantiles for each distribution. The quantile plots provide evidence that vM-UARS($\mathbf{S}, \kappa$) model describes the crystal data far better than the more standard matrix Fisher model.

In Section 2, we develop Bayes analyses for the one-sample case involving the vM-UARS($\mathbf{S}, \kappa$) distribution. First, non-informative priors are identified for the parameters $\mathbf{S}$ and κ, the corresponding posterior distribution is given, and we outline a Markov chain Monte Carlo (MCMC) algorithm for generating samples from the posterior. A simulation study for the one-sample Bayes analysis with non-informative priors is presented. Then informative priors for the one-sample case are considered. Section 3 outlines Bayes inference in one-way random effects models for rotation matrices, and Section 4 applies the Bayes methods in the analysis of EBSD measurements.

2 One-sample Bayes Analyses for the vM-UARS(S, κ) Distribution

2.1 One-Sample Model

We begin by identifying potentially *non-informative* prior distributions for the parameters of a vM-UARS($\mathbf{S}, \kappa$) distribution. For the location parameter $\mathbf{S}$ we use a prior
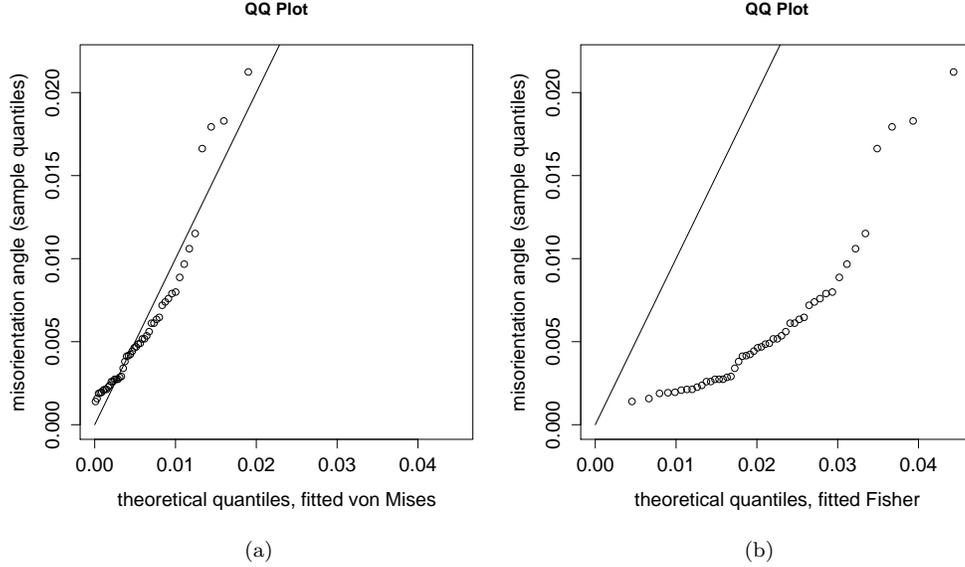
Figure 1: Sample quantiles of misorientation angles vs. theoretical quantiles for absolute spins using fitted (a) von-Mises and (b) Fisher distributions for fifty measured crystal orientations

distribution uniform on Ω. The invariant Haar measure acts as a "uniform distribution" and dominating measure for other distributions on Ω (see Downs 1972), so we adopt a prior distribution on $\mathbf{S}$ specified by density

$$h_1(\mathbf{S}) = 1, \ \mathbf{S} \in \Omega \tag{2}$$

with respect to the Haar measure. This prior is equivalent to a UARS prior model for $\mathbf{S}$ where we set $\mathbf{S} \equiv \mathbf{P}$ in (1) for r having a density on $[-\pi, \pi]$ proportional to $1 - \cos r$ (Miles 1965); see Downs (1972) for other characterizations of the Haar measure. For the spread parameter κ, we consider a Jeffreys prior. Since κ is the concentration parameter from the von Mises circular distribution with density $v(r|\kappa) = [2\pi I_0(\kappa)]^{-1} \exp[\kappa \cos(r)], \ r \in (-\pi, \pi]$, for $I_i(\kappa)$ the modified Bessel function of order i, we have

$$-\mathrm{E}\left(\frac{d^2}{d^2\kappa} \log(v(r|\kappa))\right) = \frac{I_0(\kappa)^2 - \frac{1}{\kappa} I_0(\kappa) I_1(\kappa) - I_1(\kappa)^2}{I_0(\kappa)^2}.$$

Therefore, the Jeffreys prior is specified by the (improper) density

$$h_2(\kappa) = \frac{\sqrt{I_0(\kappa)^2 - \frac{1}{\kappa} I_0(\kappa) I_1(\kappa) - I_1(\kappa)^2}}{I_0(\kappa)}, \ \kappa \in [0, \infty) \tag{3}$$

displayed in Figure 2. The function $h_2(\kappa) \to 1/\sqrt{2}$ as $\kappa \to 0$ and $\kappa h_2(\kappa) \to 1/\sqrt{2}$ as $\kappa \to \infty$.
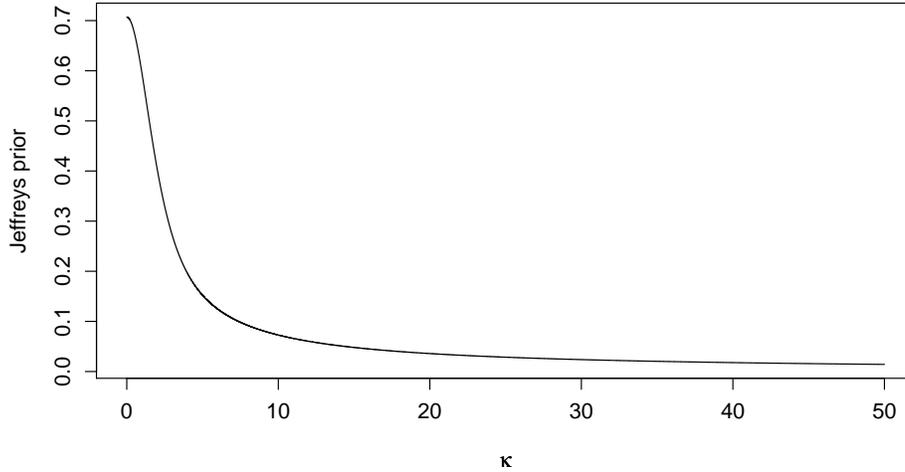
Figure 2: Graph of the Jeffreys prior, $h_2(\kappa)$, against κ

Now, the vM-UARS$(\mathbf{S}, \kappa)$ distribution has density

$$f(\mathbf{o}|\mathbf{S}, \kappa) = \frac{2}{3 - tr(\mathbf{S}^T\mathbf{o})} \left[I_0(\kappa)\right]^{-1} \exp\left[\frac{\kappa}{2}(tr(\mathbf{S}^T\mathbf{o}) - 1)\right], \quad \mathbf{o} \in \Omega, \qquad (4)$$

with respect to the Haar measure (see Bingham et al. 2009b), giving likelihood

$$L(\mathbf{S}, \kappa) = \frac{2^n \exp\left[\dfrac{\kappa}{2} \displaystyle\sum_{i=1}^{n} \left(tr(\mathbf{S}^T\mathbf{o}_i) - 1\right)\right]}{\left[I_0(\kappa)\right]^n \displaystyle\prod_{i=1}^{n} \left(3 - tr(\mathbf{S}^T\mathbf{o}_i)\right)}. \qquad (5)$$

Multiplying (2), (3), and (5) the posterior density is proportional to

$$g(\mathbf{S}, \kappa) = \frac{\exp\left[\dfrac{\kappa}{2} \displaystyle\sum_{i=1}^{n} \left(tr(\mathbf{S}^T\mathbf{o}_i) - 1\right)\right] \sqrt{I_0(\kappa)^2 - \frac{1}{\kappa}I_0(\kappa)I_1(\kappa) - I_1(\kappa)^2}}{\left[I_0(\kappa)\right]^{n+1} \displaystyle\prod_{i=1}^{n} \left(3 - tr(\mathbf{S}^T\mathbf{o}_i)\right)}. \qquad (6)$$

(We observe that this posterior density and likelihood (4) are unbounded, with singularities at each $\mathbf{o}_i$.) Although the Jeffreys prior for κ is improper, the posterior is proper and a proof of this is provided in the Appendix.

To simulate values from the posterior distribution we use a Metropolis-Hastings within Gibbs algorithm. Suppose that after $j - 1$ iterations of the algorithm one has parameters $\mathbf{S}^{j-1}$ and κ^{j-1}. For the parameter $\mathbf{S}$, we obtain a candidate for $\mathbf{S}^j$ as $\mathbf{S}^{j*} \sim$ vM-UARS$(\mathbf{S}^{j-1}, \rho)$, where ρ is a tuning parameter that can be adjusted to make the algorithm efficient. We note that this choice of proposal for $\mathbf{S}$ is symmetric in the sense that $f(\mathbf{S}'|\mathbf{S}, \rho) = f(\mathbf{S}|\mathbf{S}', \rho)$ for f in (4). For the concentration parameter κ, we take $\log(\kappa^{j*}) \sim \mathrm{N}(\log(\kappa^{j-1}), \sigma^2)$, i.e. κ^{j*} is log-normal with parameters $\log(\kappa^{j-1})$ and (the tuning parameter) σ^2. If $t(x|\mu, \sigma^2)$ represents the log-normal density, then $\dfrac{t(\kappa'|\log(\kappa), \sigma^2)}{t(\kappa|\log(\kappa'), \sigma^2)} = \dfrac{\kappa}{\kappa'}$.

Thus, for observations $\mathbf{o}_1, \ldots, \mathbf{o}_n \in \Omega$, beginning with some starting values $\mathbf{S}^0$ and κ^0, our Metropolis-Hastings within Gibbs algorithm for $j = 1, 2, \ldots$ is then:

1. Generate $\mathbf{S}^{j*} \sim$ vM-UARS$(\mathbf{S}^{j-1}, \rho)$ as a proposal for $\mathbf{S}^j$.

2. Compute $r_j^1 = \dfrac{g(\mathbf{S}^{j*}, \kappa^{j-1})}{g(\mathbf{S}^{j-1}, \kappa^{j-1})}$ for g in (6) and generate $W_j^1 \sim$ Bernoulli$(\min(1, r_j^1))$.

3. Take $\mathbf{S}^j = W_j^1 \mathbf{S}^{j*} + (1 - W_j^1)\mathbf{S}^{j-1}$.

4. Generate $\log(\kappa^{j*}) \sim \mathrm{N}(\log(\kappa^{j-1}), \sigma^2)$, with κ^{j*} as a candidate for κ^j.

5. Compute $r_j^2 = \dfrac{g(\mathbf{S}^j, \kappa^{j*})\kappa^{j*}}{g(\mathbf{S}^j, \kappa^{j-1})\kappa^{j-1}}$ for g in (6) and generate
   $W_j^2 \sim$ Bernoulli$(\min(1, r_j^2))$.

6. Take $\kappa^j = W_j^2 \kappa^{j*} + (1 - W_j^2)\kappa^{j-1}$.

Next, we report a simulation study in which the above algorithm was used to perform one-sample Bayes analyses for the vM-UARS$(\mathbf{S}, \kappa)$ distribution for various values of κ and sample size, n. We compare frequentist coverage probabilities and sizes of confidence/credible regions obtained from the Bayes analyses to the ones produced by a quasi-likelihood method of Bingham et al. (2009b). This particular comparison is of interest because the quasi-likelihood method produces an estimator of the location parameter $\mathbf{S}$ which is common in the literature (explained in the next section). Additionally, the quasi-likelihood method requires distributional approximations which are potentially poor in finite samples and the Bayes analysis offers a potentially superior alternative. We emphasize that while the Bayes analyses outlined here are for the vM-UARS$(\mathbf{S}, \kappa)$ distribution, they are easily modified for use with other members of the UARS class (see Bingham et al. (2009a) for one such example).

2.2 Simulation Study

We simulated data from the vM-UARS$(\mathbf{S}, \kappa)$ distribution for various combinations of κ and n and then used the algorithm outlined above to generate samples from the

posterior distribution. The values used for κ were 1, 5, 20, and 500 and sizes of the data sets were $n = 10, 30$, and 100. (Recall that a vM-UARS$(\mathbf{S}, \kappa)$ observation $\mathbf{O} = \mathbf{SP}$ can be simulated from the construction (1) involving a von Mises angle r and independent uniform vector $\mathbf{U} \in \mathbb{R}^3$. As $\kappa \to \infty$, the von Mises circular distribution is approximately Normal with variance $1/\kappa$ so that the distribution of r is essentially Normal with mean 0 and variance .002 when $\kappa = 500$.) The parameter $\mathbf{S}$ was held constant at

$$\mathbf{S} = \begin{pmatrix} -0.491493 & 0.804619 & -0.333202 \\ -0.562824 & -0.001501 & 0.826576 \\ 0.664578 & 0.593790 & 0.453596 \end{pmatrix}. \tag{7}$$

Symmetry arguments show this detail to be irrelevant since inference for any other principal direction $\mathbf{T}$ can be done by rotating inference results for $\mathbf{S}$ by $\mathbf{R}$, where $\mathbf{T} = \mathbf{RS}$.

For each (n, κ) combination we simulated 1000 data sets from the vM-UARS$(\mathbf{S}, \kappa)$ distribution. Based on each of these data sets we generated a sample of size 20000 from the posterior distribution in (6) (taken after a burn-in of 5000 iterations) using the algorithm with starting values $\mathbf{S}^0$ and κ^0 set at the true parameters. (Inspection of various starting values indicated that the choice of starting value did not affect the output of posterior simulations after the indicated burn-in period.) Further, for each (n, κ) combination, the tuning parameters ρ and σ were set to give Metropolis-Hastings jumping rates near 40% (see Gelman, Carlin, Stern, and Rubin 2004, sec. 11.10). The values of the tuning parameters used are given in Table 1.

Table 1: Values of tuning parameters ρ and σ

	$n = 10$		$n = 30$		$n = 100$	
	ρ	σ	ρ	σ	ρ	σ
$\kappa = 1$	25	2	50	1.5	1000	1
$\kappa = 5$	100	1.5	1000	1	10000	0.5
$\kappa = 20$	1000	1.5	1000	1	50000	0.5
$\kappa = 500$	1000	1	10000	0.5	100000	0.5

For each vM-UARS$(\mathbf{S}, \kappa)$ data set simulated for a given (n, κ) pair we obtained point and set estimates for $\mathbf{S}$ and κ using both Bayes methods and the quasi-likelihood method of Bingham et al. (2009b). For the latter, with iid observations $\mathbf{o}_1, \ldots, \mathbf{o}_n \in \Omega$ from the vM-UARS$(\mathbf{S}, \kappa)$ distribution, point estimates were obtained by maximizing (the quasi-likelihood)

$$Q(\mathbf{S}, \kappa) = [I_0(\kappa)]^{-n} \prod_{i=1}^n \exp\left[\frac{\kappa}{2}(tr(\mathbf{S}^T \mathbf{o}_i) - 1)\right]. \tag{8}$$

Bingham et al. (2009b) note that $\hat{\mathbf{S}}$ (the maximum quasi-likelihood estimate for the parameter $\mathbf{S}$) can be obtained by maximizing $tr(\mathbf{S}^T \bar{\mathbf{o}})$, where $\bar{\mathbf{o}} = \sum_{i=1}^n \mathbf{o}_i/n$. Thus, $\hat{\mathbf{S}}$ corresponds to a commonly used estimate of "location" for 3-dimensional rotation data,

and is the moment estimator for the modal rotation of the Cayley distribution in 3 dimensions (León et al. 2006, p. 421) and the mean direction for the matrix von Mises-Fisher distribution in 3 dimensions (Khatri and Mardia 1977, p. 96). Both the likelihood (5) and posterior density (6) have singularities in $\mathbf{S}$ at each observation $\mathbf{o}_i$ and have no maximum. Since the technique of maximizing $tr(\mathbf{S}^T\bar{\mathbf{o}})$ is a commonly used and sensible approach for estimating a "mean orientation" $\mathbf{S}$ from a set of orientations $\mathbf{o}_1, \ldots, \mathbf{o}_n$, we use it to obtain a Bayes point estimate for $\mathbf{S}$ based on the 20000 orientations simulated from the posterior. A Bayes point estimate for κ can be simply obtained by taking the mean of the κ values from the simulated posterior sample.

We compare the point estimators for κ under the two inference approaches by examining the mean squared error (MSE) obtained for different (n, κ) pairs. To compare point estimators for $\mathbf{S}$, the misorientation angle (intrinsically in $[0, \pi]$) between an estimate $\hat{\mathbf{S}}$ and the true $\mathbf{S}$ was found; see (11) in the Appendix for more details. These misorientation angles were then averaged for the 1000 data sets for each (n, κ) pair. The MSEs (for κ) and average misorientation angles (for $\mathbf{S}$) are given in Table 2. We see that the Bayes point estimates for κ result in smaller errors than the maximum quasi-likelihood estimates for small n, but that the two estimation techniques provide similar results as n increases. When estimating $\mathbf{S}$, the Bayes point estimates give a smaller average misorientation angle than the quasi-likelihood estimates for all cases, with the relative advantage of the Bayes method increasing (substantially) with n.

Table 2: MSE for κ and average misorientation angles for $\mathbf{S}$ for point estimation, using (non-informative prior) Bayes and quasi-likelihood inference

(n, κ)	κ		$\mathbf{S}$	
	Bayes	Q-L	Bayes	Q-L
$(10, 1)$	0.3662	0.4824	0.2091	0.3244
$(30, 1)$	0.1021	0.1090	0.0678	0.1825
$(100, 1)$	0.0283	0.0288	0.0203	0.0966
$(10, 5)$	10.77	14.88	0.0780	0.1350
$(30, 5)$	1.85	2.11	0.0271	0.0770
$(100, 5)$	0.493	0.512	0.0085	0.0415
$(10, 20)$	259.9	341.4	0.0366	0.0666
$(30, 20)$	34.4	39.3	0.0134	0.0375
$(100, 20)$	8.7	9.2	0.0040	0.0209
$(10, 500)$	154500	215400	0.0072	0.0132
$(30, 500)$	24400	28130	0.0026	0.0075
$(100, 500)$	5481	5717	0.0008	0.0041

We will next compare Bayes and quasi-likelihood set estimation. Bingham et al. (2009b, sec. 4.2, Propositions 2 and 3) derived the asymptotic distributions of a quasi-likelihood ratio test statistic and a Wald test statistic for the vM-UARS$(\mathbf{S}, \kappa)$ distribution and developed confidence regions for κ and $\mathbf{S}$ through inversion of such tests. Our aim here is to compare the coverage probabilities and sizes of sets obtained using quasi-likelihood ratio and Wald methods to those of credible sets from Bayes analyses. For
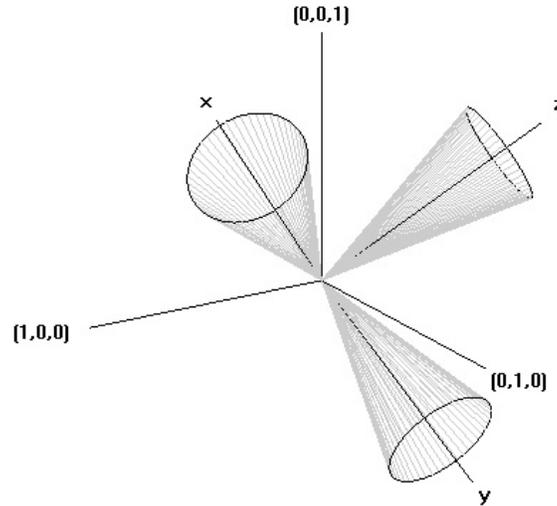
Figure 3: Graphical display of a confidence or credible region for the parameter **S**, with x, y, and z representing the orientation of a corresponding point estimate for **S**

concreteness, we consider 95% confidence and credible sets. To get a posterior credible interval for κ, let $\Delta_\kappa = \{\kappa^1, \kappa^2, \ldots, \kappa^N\}$ represents the set of $N = 20000$ values for κ simulated from the posterior. With $\kappa_{.025} =$ the .025 quantile of Δ_κ and $\kappa_{.975} =$ the .975 quantile of Δ_κ, the interval $[\kappa_{.025}, \kappa_{.975}]$ represents a 95% credible interval for κ. To construct a credible set for **S**, suppose that $\mathbf{S}_B$ represents the Bayes point estimate for **S** discussed above and that $\Delta_S = \{\mathbf{S}^1, \mathbf{S}^2, \ldots, \mathbf{S}^N\}$ represents the set of values for **S** simulated from the posterior. For $i = 1, \ldots, N$, we find the (non-negative) angle between each of the coordinate axes rotated from their reference direction by $\mathbf{S}_B$ with the corresponding axis rotated by $\mathbf{S}^i$ and let δ^i represent the maximum of these three angles. With $\delta_{.95} =$ the .95 quantile of $\{\delta^1, \ldots, \delta^N\}$, cones of constant angle $\delta_{.95}$ around the coordinate axis rotated by $\mathbf{S}_B$ create a region representing 95% of sets of directions of coordinate axes rotated by values from Δ_S. Bingham et al. (2009b) use conic regions centered at coordinate axes rotated by the maximum quasi-likelihood estimate for **S** to construct Wald and quasi-likelihood ratio confidence regions. Thus, a confidence or credible region for **S** can be graphically displayed as in Figure 3 and the notion of "size" of these regions can be reduced to the size of the angle between the center and edge of the cones.

After creating confidence and credible regions for each of the 1000 vM-UARS$(\mathbf{S}, \kappa)$ data sets generated for a given (n, κ) combination, we checked if the regions for κ and **S** contained the true parameter values. Corresponding estimated frequentist coverage probabilities for the inference methods are given in Table 3. The probabilities in the table verify that, with the exception of the small n Wald intervals for κ, all regions are holding their nominal frequentist coverage rates. The fact that the Wald intervals for κ cover the true parameter value less often than nominal when n is small is in accord

with the results in Bingham et al. (2009b). (It is observed there that convergence to the limiting distribution for the Wald test statistic is slower than the corresponding convergence of the quasi-likelihood ratio test statistic.)

Table 3: Coverage rates (%) for $\mathbf{S}$ and κ for (non-informative prior) Bayes, quasi-likelihood ratio test, and Wald techniques for various (n, κ)

	Bayes		Q-LRT		Wald	
(n, κ)	$\mathbf{S}$	κ	$\mathbf{S}$	κ	$\mathbf{S}$	κ
$(10, 1)$	94.7	97.2	97.0	94.8	94.6	88.7
$(30, 1)$	95.6	94.4	98.2	94.8	94.9	93.4
$(100, 1)$	95.2	95.5	98.9	95.5	94.6	94.1
$(10, 5)$	96.1	95.8	94.1	94.4	95.6	83.5
$(30, 5)$	94.2	94.9	95.6	94.9	95.7	91.1
$(100, 5)$	94.7	94.1	96.1	95.3	95.7	93.5
$(300, 5)$	94.0	95.4	95.8	95.8	95.4	95.8
$(1000, 5)$	93.6	94.8	95.8	94.9	95.1	94.4
$(10, 20)$	93.8	94.5	93.7	94.0	96.5	82.5
$(30, 20)$	94.4	95.7	95.3	95.5	95.9	91.7
$(100, 20)$	94.8	94.6	94.0	94.6	94.0	93.4
$(10, 500)$	94.7	95.0	93.3	926	95.8	80.3
$(30, 500)$	95.5	96.6	94.8	95.2	95.3	88.6
$(100, 500)$	94.9	94.6	95.2	94.2	95.5	94.2

Recognizing that both Bayes and quasi-likelihood sets produce about the "right" coverage probabilities, we compare the "sizes" of the regions obtained from the different methods. For each (n, κ) combination, we have 1000 confidence intervals for κ and 1000 confidence sets for $\mathbf{S}$ from inversion of the quasi-likelihood ratio tests and Wald tests. Similarly, we have 1000 posterior credible regions for each parameter.

Table 4 gives the median widths for the 1000 intervals for κ. From the table we see that the posterior credible intervals for κ outperform both types of quasi-likelihood intervals. While the median width of the quasi-likelihood ratio intervals is relatively close to the that of the posterior credible intervals, the median width for the Wald intervals is much larger than that of the credible intervals for small sample sizes. Thus, Bayes method captures κ in narrower intervals, still while holding the desired coverage probabilities.

We note that instead of using $[\kappa_{.025}, \kappa_{.975}]$ as a 95% credible interval for κ, we could have used shortest length/highest posterior density (HPD) intervals. For all cases, the HPD intervals gave median widths close to those presented in Table 4. Figure 4 displays fitted densities for posterior values of κ generated using a data set of size 30 with $\kappa = 5$ and a data set of size 30 with $\kappa = 500$. Since the densities are unimodal and fairly symmmetric, the two methods for obtaining 95% credible intervals for κ can be expected to give similar results. The width of the $[\kappa_{.025}, \kappa_{.975}]$ interval corrsponding to Figure 4a ($\kappa = 5$) is 3.58, while the HPD interval is 3.55. These widths are 603 and

596, respectively, for the case of $\kappa = 500$ presented in Figure 4b. Because the HPD 95% intervals are similar to the intervals constructed as $[\kappa_{.025}, \kappa_{.975}]$, we will from here on use the latter for κ.

Table 4: Median width of (non-informative prior) Bayes credible intervals, quasi-likelihood ratio intervals, and Wald intervals for κ for various (n, κ)

(n, κ)	Bayes	Q-LRT	Wald
$(10, 1)$	2.052	2.274	6.991
$(30, 1)$	1.204	1.231	1.372
$(100, 1)$	0.660	0.665	0.683
$(10, 5)$	8.646	9.940	42.520
$(30, 5)$	4.745	4.979	6.695
$(100, 5)$	2.566	2.599	2.818
$(300, 5)$	1.489	1.497	1.537
$(1000, 5)$	0.815	0.817	0.824
$(10, 20)$	36.32	42.09	177.90
$(30, 20)$	20.44	21.36	28.52
$(100, 20)$	11.02	11.22	12.13
$(10, 500)$	944.9	1075.0	4543.0
$(30, 500)$	515.3	536.4	716.0
$(100, 500)$	279.1	283.3	306.2

Again, the size of a confidence or credible region for $\mathbf{S}$ is characterized by the angle of the conic regions. The median angles for the 1000 regions produced by each method are given in Table 5. The credible regions are perhaps even surprisingly smaller than both the quasi-likelihood ratio and Wald sets for $\mathbf{S}$, and the Bayes median angles seem to decrease at rate $1/n$ while the quasi-likelihood median angles seem to decrease at the "usual" $1/\sqrt{n}$ rate. (The $(300, 5)$ and $(1000, 5)$ lines of Tables 3, 4, and 5 were added to further study this phenomenon and bear out this observation.) For $\kappa = 5$, the Bayes median angles are approximately $1.46/n$ while the Q-LRT median angles are approximately $.749/\sqrt{n}$ and Wald median angles are approximately $.729/\sqrt{n}$. Our considered belief regarding the origin of this phenomenon is that it can be traced exclusively to the non-regularity of the likelihood (and posterior) in this problem. The likelihood (5) has singularities at all observations (a fact which drove us to initially consider quasi-likelihood), while the quasi-likelihood (8) is smooth and its asymptotics "standard." But it is not unheard of to be able to get a better convergence rate than $1/\sqrt{n}$ with an appropriate method in a non-regular problem. (For a more complete analysis of this kind of rate issue for Bayes methods, albeit in a simpler related context, see Nordman, Vardeman, and Bingham (2009).)

So in sum, the Bayes methods based on non-informative priors for both κ and $\mathbf{S}$ in the vM-UARS$(\mathbf{S}, \kappa)$ one-sample problem are preferable to the quasi-likelihood methods of Bingham et al. (2009b). This is weakly true for inference about κ and strongly true for inference on $\mathbf{S}$.
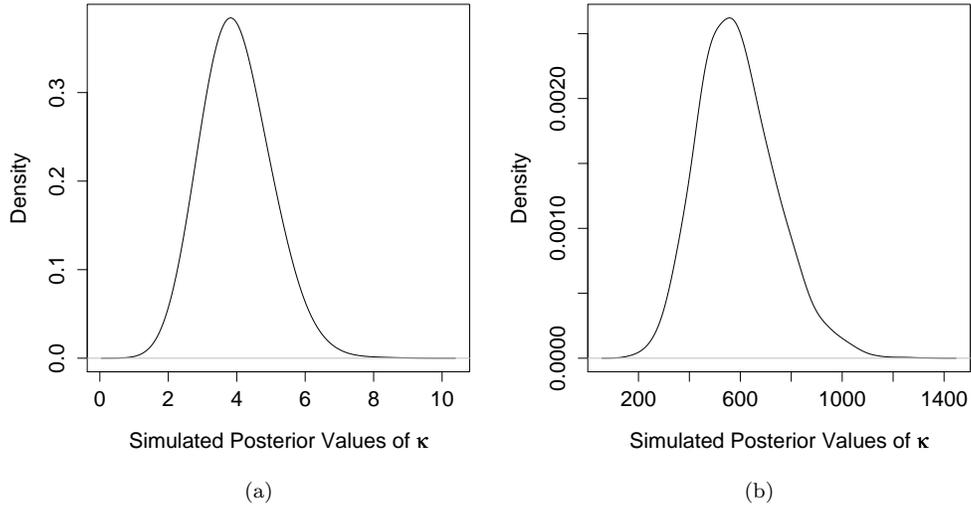
Figure 4: Fitted densities for posterior samples of κ generated using a data set of size 30 with (a) $\kappa = 5$ and (b) $\kappa = 500$

Table 5: Median angle of (non-informative prior) Bayes credible regions, quasi-likelihood ratio regions, and Wald regions for **S** for various (n, κ)

(n, κ)	Bayes	Q-LRT	Wald
$(10, 1)$	0.3857	0.6215	0.5439
$(30, 1)$	0.1264	0.3641	0.3104
$(100, 1)$	0.0374	0.1960	0.1665
$(10, 5)$	0.1474	0.2513	0.2312
$(30, 5)$	0.0495	0.1334	0.1321
$(100, 5)$	0.0152	0.0723	0.0714
$(300, 5)$	0.0048	0.0432	0.0426
$(1000, 5)$	0.0014	0.0237	0.0233
$(10, 20)$	0.0683	0.1177	0.1091
$(30, 20)$	0.0247	0.0654	0.0642
$(100, 20)$	0.0074	0.0352	0.0351
$(10, 500)$	0.0138	0.0234	0.0215
$(30, 500)$	0.0050	0.0130	0.0126
$(100, 500)$	0.0015	0.0070	0.0070

2.3   Informative Priors

So far we have considered Bayes inference for the vM-UARS$(\mathbf{S}, \kappa)$ distribution based on non-informative priors for the parameters κ and $\mathbf{S}$. We now consider the use of informative priors. For the parameter $\mathbf{S}$ we propose a vM-UARS$(\mathbf{M}, \nu)$ prior. (A member of *any* UARS class might be equally easily and naturally used as a prior.) Recall that for large κ, the von Mises circular distribution for r in the vM-UARS$(\mathbf{S}, \kappa)$ construction is approximately normal with mean 0 and variance $1/\kappa$. The inverse gamma distribution serves as a conjugate prior for a normal variance. This naturally suggests a Gamma(α, β) prior for the parameter κ. We can use the obvious variant of the Metropolis-Hastings within Gibbs algorithm outlined in Section 2.1 to simulate values from the posterior using these priors.

We next report a numerical "thought exercise" where we imagine losing half the data but place a greater inference burden on certain prior models selected from this prior class. This is intended to give a context to our claim of "informative." Having separately demonstrated the efficacy of non-informative Bayes inference, we now aim to illustrate that appropriate "informative" priors provide the same quality of inference as non-informative priors given twice the data. This is one way of providing evidence that this prior class can honestly and usefully encode information.

A simulation study for (n, κ) pairs $(30, 5)$, $(100, 5)$, $(30, 20)$, and $(100, 20)$ was done, with $\mathbf{S}$ again fixed at (7). For each (n, κ) pair, 1000 vM-UARS$(\mathbf{S}, \kappa)$ data sets were generated. First, median sizes of Bayes credible regions for $\mathbf{S}$ and κ were found for the case of non-informative priors (2) and (3) exactly in the manner discussed in Section 2.2 (based on a posterior sample of size 20000). Then, to investigate the degree to which "independent vM-UARS and gamma priors" can represent (at least empirically-based) prior knowledge about $(\mathbf{S}, \kappa)$, we did the following. We divided each of the individual data sets originally of size n into two subsets, each of size $n/2$. Based on the first $n/2$ observations and the non-informative prior and MCMC algorithm of Section 2.1, we generated a posterior sample of size 20000 exactly as described in that section. We then fit a vM-UARS$(\mathbf{M}, \nu)$ distribution to the marginal posterior of $\mathbf{S}$ and a Gamma(α, β) distribution to the marginal posterior of κ. The product of these fitted forms for marginal posteriors was then used as a prior distribution for $(\mathbf{S}, \kappa)$ using the last $n/2$ observations. We then simulated 20000 values from the fitted posterior using the informative priors and half-size data sets. Credible regions for $\mathbf{S}$ and κ were found as in Section 2.2.

Table 6 gives the approximate frequentist coverage probabilities and Table 7 gives the median sizes for the 95% credible regions obtained using first the non-informative prior and sample size n, and second, the informative prior based on the product of fitted marginal posteriors (based on $n/2$ prior observations and $n/2$ observations). The estimated coverage probabilities in Table 6 verify that all credible regions are holding nominal coverage rates. Table 7 shows that there is little difference in the median sizes of the credible regions obtained under the two methods. (Notice also that the regions for $\mathbf{S}$ still appear to decrease in size at a rate proportional to $1/n$.) The comparison here gives evidence that our proposal for informative priors, namely a UARS$(\mathbf{M}, \nu)$ proir for $\mathbf{S}$ and a Gamma prior for κ, is an effective and appropriate coding of prior information.

Table 6: Coverage rates (%) for $\mathbf{S}$ and κ for various (n, κ) pairs

	Non-informative Prior (sample size n)		(Fitted) Informative Prior (sample size $n/2$)	
(n, κ)	κ	$\mathbf{S}$	κ	$\mathbf{S}$
$(30, 5)$	95.1	94.4	95.0	93.3
$(100, 5)$	93.9	94.5	93.7	93.6
$(30, 20)$	94.5	93.2	94.4	93.5
$(100, 20)$	94.6	94.1	94.3	93.4

Table 7: Median sizes of 95% Bayes credible regions for κ and $\mathbf{S}$

	Non-informative Prior (sample size n)		(Fitted) Informative Prior (sample size $n/2$)	
(n, κ)	κ	$\mathbf{S}$	κ	$\mathbf{S}$
$(30, 5)$	4.782	0.0511	4.780	0.0546
$(100, 5)$	2.597	0.0156	2.579	0.0168
$(30, 20)$	20.52	0.0247	20.63	0.0263
$(100, 20)$	11.00	0.0075	10.97	0.0080

3 One-Way Random Effects Bayes Analyses for the vM-UARS($\mathbf{S}, \kappa$) Distribution

With numerical evidence indicating that Bayes methods perform well for the one-sample case, we next develop Bayes analyses for a 3-d rotation version of a one-way random effects model. This allows effective study of two sources of variation arising between and among groups of rotation matrix observations. Suppose that for $i = 1, \ldots, r$ and $k = 1, \ldots, m_i$,

$$\mathbf{O}_{ik} = \mathbf{P}_i \mathbf{Q}_{ik} \in \Omega$$

for $\mathbf{P}_i \sim^{iid}$ vM-UARS($\mathbf{S}, \kappa$) independent of $\mathbf{Q}_{ik} \sim^{iid}$ vM-UARS($\mathbf{I}_{3 \times 3}, \tau$). Thus, we have r groups with m_i observations in the ith group, where κ represents the size of the between-group variation and τ represents the size of the within-group variation (with larger values indicating less variation). We will use Bayes methods to estimate these parameters as well as the overall "location" parameter $\mathbf{S}$.

First, we can write

$$f(\mathbf{o}_{ik}, \mathbf{p}_i | \mathbf{S}, \kappa, \tau) = f(\mathbf{o}_{ik} | \mathbf{p}_i, \tau) f(\mathbf{p}_i | \mathbf{S}, \kappa),$$

for f in (4). Thus, for the one-way random effects case we have a joint density for the

observable $\mathbf{o}_{ik}$ and unobservable $\mathbf{p}_i$

$$f(\mathbf{o}_{11}, \ldots, \mathbf{o}_{rm_r}, \mathbf{p}_1, \ldots, \mathbf{p}_r | \mathbf{S}, \kappa, \tau) = \prod_{i=1}^{r} \left(\left[\prod_{k=1}^{m_i} f(\mathbf{o}_{ik} | \mathbf{p}_i, \tau) \right] f(\mathbf{p}_i | \mathbf{S}, \kappa) \right)$$

$$= \prod_{i=1}^{r} \left(\frac{2^{(m_i+1)} \exp\left[\frac{\tau}{2} \sum_{k=1}^{m_i} \left(tr(\mathbf{p}_i^T \mathbf{o}_{ik}) - 1 \right) + \frac{\kappa}{2} \left(tr(\mathbf{S}^T \mathbf{p}_i) - 1 \right) \right]}{I_0(\tau)^{m_i} I_0(\kappa) \left(3 - tr(\mathbf{S}^T \mathbf{p}_i) \right) \prod_{k=1}^{n_i} \left[3 - tr(\mathbf{p}_i^T \mathbf{o}_{ik}) \right]} \right).$$

$$(9)$$

Now, we again place a uniform (Haar) prior on $\mathbf{S}$ and independent Jeffreys priors on κ and τ. We then have posterior density for the unobservable $\mathbf{p}_i$ and parameters proportional to

$$G(\mathbf{S}, \kappa, \tau, \mathbf{p}_1, \ldots, \mathbf{p}_r) = f(\mathbf{p}_1, \ldots, \mathbf{p}_r | \mathbf{S}, \kappa, \tau) h_2(\tau) h_2(\kappa) \qquad (10)$$

for f given in (9) with $\mathbf{o}_{ik}$ fixed and dependence upon them suppressed and h_2 in (3). For observations $\mathbf{o}_{ik}$, $i = 1, \ldots, r$ and $k = 1, \ldots, m_i$, beginning with starting values $\mathbf{S}^0$, κ^0, τ^0, and $\{\mathbf{p}_1^0, \ldots, \mathbf{p}_r^0\}$ we implement a Metropolis-Hastings within Gibbs algorithm for $j = 1, 2, \ldots$ as follows:

1. Generate $\mathbf{S}^{j*} \sim$ vM-UARS($\mathbf{S}^{j-1}, \rho_1$) as a proposal for $\mathbf{S}^j$.

2. Compute $r_j^1 = \dfrac{G(\mathbf{S}^{j*}, \kappa^{j-1}, \tau^{j-1}, \mathbf{p}_1^{j-1}, \ldots, \mathbf{p}_r^{j-1})}{G(\mathbf{S}^{j-1}, \kappa^{j-1}, \tau^{j-1}, \mathbf{p}_1^{j-1}, \ldots, \mathbf{p}_r^{j-1})}$ for G in (10) and generate $W_j^1 \sim$ Bernoulli($\min(1, r_j^1)$).

3. Take $\mathbf{S}^j = W_j^1 \mathbf{S}^{j*} + (1 - W_j^1) \mathbf{S}^{j-1}$.

4. Generate $\log(\kappa^{j*}) \sim$ N($\log(\kappa^{j-1})$, σ_1^2), with κ^{j*} as a proposal for κ^j.

5. Compute $r_j^2 = \dfrac{G(\mathbf{S}^j, \kappa^{j*}, \tau^{j-1}, \mathbf{p}_1^{j-1}, \ldots, \mathbf{p}_r^{j-1}) \kappa^{j*}}{G(\mathbf{S}^j, \kappa^{j-1}, \tau^{j-1}, \mathbf{p}_1^{j-1}, \ldots, \mathbf{p}_r^{j-1}) \kappa^{j-1}}$ for G in (10) and generate $W_j^2 \sim$ Bernoulli($\min(1, r_j^2)$).

6. Take $\kappa^j = W_j^2 \kappa^{j*} + (1 - W_j^2) \kappa^{j-1}$.

7. Generate $\log(\tau^{j*}) \sim$ N($\log(\tau^{j-1})$, σ_2^2), with τ^{j*} as a proposal for τ^j.

8. Compute $r_j^3 = \dfrac{G(\mathbf{S}^j, \kappa^j, \tau^{j*}, \mathbf{p}_1^{j-1}, \ldots, \mathbf{p}_r^{j-1}) \tau^{j*}}{G(\mathbf{S}^j, \kappa^j, \tau^{j-1}, \mathbf{p}_1^{j-1}, \ldots, \mathbf{p}_r^{j-1}) \tau^{j-1}}$ for G in (10) and generate $W_j^3 \sim$ Bernoulli($\min(1, r_j^3)$).

9. Take $\tau^j = W_j^3 \tau^{j*} + (1 - W_j^3) \tau^{j-1}$.

10. For $k = 1, \ldots, r$

    (a) Generate $\mathbf{p}_k^{j*} \sim$ vM-UARS($\mathbf{p}_k^{j-1}, \rho_2$) as a proposal for $\mathbf{p}_k^j$.

(b) Compute $q_j^k = \dfrac{G(\mathbf{S}^j, \kappa^j, \tau^j, \mathbf{p}_1^j, \ldots, \mathbf{p}_{k-1}^j, \mathbf{p}_k^{j*}, \mathbf{p}_{k+1}^{j-1}, \ldots, \mathbf{p}_r^{j-1})}{G(\mathbf{S}^j, \kappa^j, \tau^j, \mathbf{p}_1^j, \ldots, \mathbf{p}_{k-1}^j, \mathbf{p}_k^{j-1}, \mathbf{p}_{k+1}^{j-1}, \ldots, \mathbf{p}_r^{j-1})}$ for G in (10)
and generate $V_j^k \sim \text{Bernoulli}(\min(1, q_j^k))$.

(c) Take $\mathbf{p}_k^j = V_j^k \mathbf{p}_k^{j*} + (1 - V_j^k) \mathbf{p}_k^{j-1}$.

(Again, we emphasize that while here we concentrate on the von Mises case, this methodology is easily modified and applied to other UARS subclasses.)

To check the efficacy of the one-way random effects method above, we simulated 1000 independent data sets for a case where $m_i = 30$ for $i = 1, \ldots, 30$, $\tau = 20$, $\kappa = 20$, and $\mathbf{S}$ is as in Section 2.2. (Because the Metropolis-Hastings algorithm for the one-way random effects case requires much more computing time than for the one-sample case, we could not perform a more extensive simulation study.) Using each of the 1000 data sets, we generated a sample of size 8000 from the posterior density in (10) after a burn-in of 2000 iterations. As in the one-sample case, starting values had no effect on the nature of posterior samples for such a burn-in period. For each of the 1000 cases, we obtained Bayes point estimates and 95% posterior credible sets. The point estimates and credible sets for $\mathbf{S}$ were computed as for the one-sample situation, and estimates and intervals for κ and τ were both computed as for κ in the one-sample case.

Checking the 95% credible regions for each parameter, we found the true values were captured at rates of 93.9%, 95.3%, and 94.2% for $\mathbf{S}$, κ, and τ, respectively. So, the frequentist coverage rates are close to the credible levels. We also found the median size of the credible regions for each parameter. For $\mathbf{S}$ the median conic angle was 0.0241, for κ the median credible interval width was 20.52, and for τ the median credible interval width was 3.652. Thus, we are able to capture the indicator of within group variation τ in smaller intervals than the indicator of between group variation κ. Now, we apply the methods from this section, as well as the one-sample methods from the previous section, to some real EBSD data.

4   Application to EBSD Data

Although EBSD precision has received attention in the materials science literature, published precision estimates are based on descriptive methods applied to a single scan on a homogeneous specimen (i.e. a single crystal/grain). Using inference for the vM-UARS$(\mathbf{S}, \kappa)$ class we can quantify EBSD precision by examining data comprised of multiple scans of a single grain providing both single-site and between site measures of variability in orientation. Bingham et al. (2009b) analyzed data from a sample of TSL (an EBSD company) calibration standard high-Iron-concentration nickel. The nickel specimen had a surface area $40\mu\text{m} \times 40\mu\text{m}$ and the same area was scanned 14 times with at least 4000 measurements per run. Using the quasi-likelihood inference of Bingham et al. (2009b), we fit the vM-UARS$(\mathbf{S}, \kappa)$ distribution to two sets of observations. Data set 1 contains 14 repeat observations from one location in the scans and data set 2 contains 50 observations from different locations appearing to be within the same grain on a single scan. (The data here are highly concentrated around their principal direction,

requiring use of the normal approximation to the von Mises distribution.) We also applied the (non-informative prior) one-sample Bayes methods from Section 2.1 to the same two data sets by simulating a sample of size 20000 from the posterior given in (6). Densities were fit to the simulated posterior values of κ for each of the data sets and are presented in Figure 5. Estimates obtained using both Bayes and quasi-likelihood inference are presented in Table 8 and are similar for the two estimation approaches.
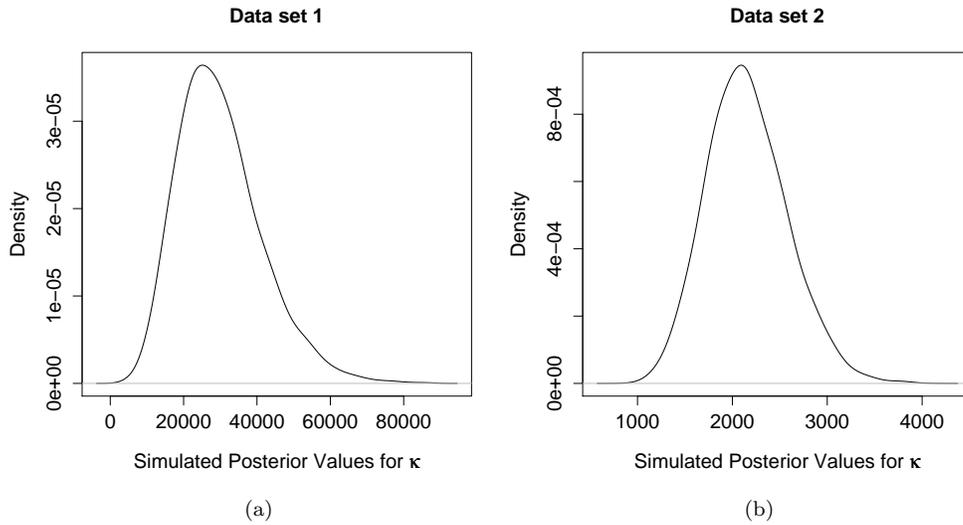
Data set 1 **Data set 2**

Figure 5: Fitted densities for posterior samples of κ for (a) the single location data (Data set 1) and (b) data from within a grain on a single scan (Data set 2)

Table 8: Estimates of the parameters κ and **S** (given in Euler angle form; see Bingham et al. (2009b)) for the two EBSD data sets using maximum quasi-likelihood and (non-informative prior) Bayes methods

		κ	(α, β, γ)
Data set 1:	MQL	33599.2	(5.8532, 0.9291, 4.2265)
	Bayes	30774.0	(5.8534, 0.9292, 4.2259)
Data set 2:	MQL	2197.0	(5.8599, 0.9252, 4.2191)
	Bayes	2156.5	(5.8567, 0.9259, 4.2225)

We next put the Bayes estimates for κ into the context of EBSD precision. With $\hat{\kappa} = 30774.0$, the spin angles r (see Section 1) for the repeated nickel data are approximately normal with mean 0 and standard deviation $1/\sqrt{30774.0} = 0.0057$ radians, or $0.3266°$, and the fitted distribution on the misorientation angles $|r|$ places 99% probability in the interval $(0°, 0.8413°)$. As in Bingham et al. (2009b), we report this as an EBSD misorientation precision of $0.8413°$. For the measurements within a grain on a single

scan, r is approximately normal with mean 0 and standard deviation $1/\sqrt{2156.5} = 0.0215$ radians, or $1.234°$. Thus, the fitted distribution on the misorientation angles $|r|$ places 99% probability in the interval $(0°, 3.179°)$, for a within-grain misorientation precision of $3.179°$.

We can also provide 95% confidence and credible intervals for κ for both data sets. Confidence intervals obtained using the inference of Bingham et al. (2009b) are in Table 9 along with the posterior credible intervals we obtained for κ using the (non-informative prior) Bayes approach of Section 2.1. For both data sets, the width of the Bayes credible interval for κ is less than that of the Wald and quasi-likelihood ratio intervals. The intervals for data set 2 are located entirely to the left of those for data set 1, which is expected since the intervals for the second data set capture both single-site variation and (location-to-location) variation between sites. (In fact, based on the posterior samples displayed in Figure 5, only 0.115% of the simulated κ values from data set 2 overlap the range of simulated κ values for data set 1.)

Table 9: 95% quasi-likelihood ratio test and Wald confidence intervals and 95% credible intervals for κ for the two EBSD data sets

	Data set 1	Data set 2
Q-LRT	(14437, 64979)	(1444, 3174)
Wald	(19301, 129623)	(1578, 3613)
Bayes	(11961, 58327)	(1411, 3056)

Using the (non-informative prior) Bayes analysis for one-way random effects developed in Section 3, we can look simultaneously at the two types of EBSD variation. We used a sample of 14 repeat observations from each of 50 positions on the scans that all appear to be within a single grain. Thus, $m_i = 14$ for each $i = 1, \ldots, 50$. Using the algorithm outlined in Section 3, we obtained a sample of size 8000 from the posterior in (10). Fitted densities for the posterior values of κ and τ are presented in Figure 6. Based on this sample, Bayes estimates are $\kappa_B = 2652$, $\tau_B = 5474$, and

$$\mathbf{S}_B = \left(\begin{array}{rrr} -0.6449 & -0.2866 & -0.7084 \\ 0.6874 & -0.6228 & -0.3738 \\ -0.3341 & -0.7280 & 0.5987 \end{array} \right).$$

The 95% posterior credible intervals are $(1717, 3797)$ and $(4899, 6066)$ for κ and τ, respectively, with conic regions of angle 0.0034 representing a 95% credible region for $\mathbf{S}$. Notice that the interval for τ represents less variability in orientation than that for κ, so the between group variation is greater than the within group variation. The Bayes estimates for κ and τ correspond to estimated EBSD misorientation precisions of $2.866°$ and $1.995°$, respectively.

Because the estimates for κ and τ are much larger here than the parameters used in the single simulation in Section 3, we also conducted a simulation using κ_B, τ_B, and $\mathbf{S}_B$ as the true parameter values with $m_i = 14$ for $i = 1, \ldots, 50$. Again, 1000 data sets were generated using these parameter values and then the Metropolis-Hastings within
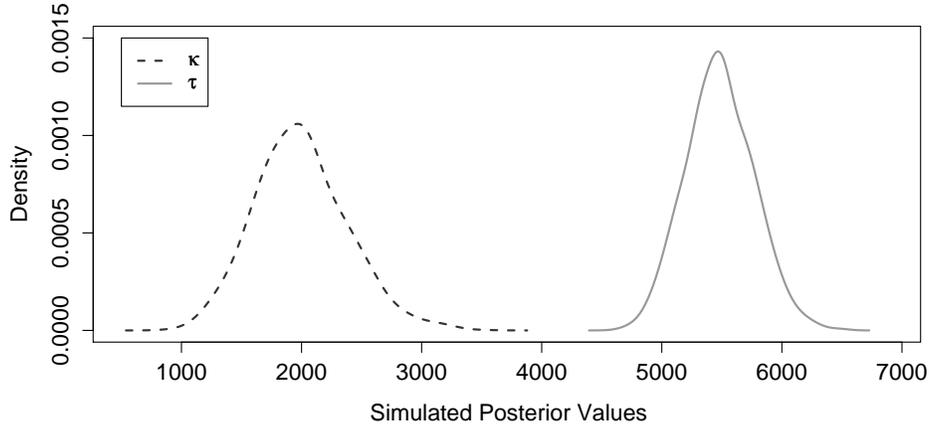
Figure 6: Fitted densities for posterior samples of κ and τ in the one-way random effects case

Gibbs algorithm was used to get a sample of size 8000 from the posterior for each of the 1000 data sets. Median 95% credible region sizes were calculated as before, yielding median width of 2072 for κ, 1147 for τ, and median angle of 0.0013 for $\mathbf{S}$. Coverage rates were 93.3%, 96.2%, and 95.4% for $\mathbf{S}$, κ, and τ, respectively. This is at least anecdotal evidence that Bayes one-way random effects analyses also behave sensibly for large values of κ and τ.

5   Conclusion

As we have shown, one-sample Bayes analyses for the vM-UARS$(\mathbf{S}, \kappa)$ models are effective and provide an alternative to the quasi-likelihood inference introduced in Bingham et al. (2009b). By using Bayes methods and MCMC, we are also able to extend our analyses for the vM-UARS$(\mathbf{S}, \kappa)$ models to a one-way random effects version, which provides a useful new form of inference for 3-d rotation data. Although only the von Mises version of the UARS class of distributions was examined here, the Bayes methods are easily extended to other members of the class (see, for example, Bingham et al. 2009a).

The Bayes methods introduced here can also be extended to allow for a variety of useful new analyses for 3-d rotation data. We have shown how to include random effects into analyses for orientation data and can just as easily consider fixed/regression effects. For example, for a known parametric function $\mathbf{S}(\boldsymbol{\theta}, \mathbf{x})$ taking values in Ω and data $(\mathbf{x}_i, \mathbf{o}_i)$, $i = 1, \ldots, n$, we can consider the model $\mathbf{O}_i \sim^{ind} \mathrm{UARS}(\mathbf{S}(\boldsymbol{\theta}, \mathbf{x}_i), \kappa)$. By placing a prior on $\boldsymbol{\theta}$ and using MCMC we can simulate values from a posterior for $(\boldsymbol{\theta}, \kappa)$.

This type of analysis could, for example, be used to model noise in the observation of the spinning of an object at an unknown rate about an unknown axis. Bayes methods could also be developed for "random walks" and AR(1) processes in orientations and clustering for 3-d orientation data, resulting in new types of inference for 3-d rotation data.

Appendix: Proof of Propriety of the Posterior

Recall the posterior kernel (6). Let H denote the Haar measure on the space Ω of 3-d rotations (Miles 1965; Downs 1972) and λ denote the Lebesgue measure on $\mathbb{R}_+ \equiv (0, \infty)$ (i.e., measures defined with respect to the standard σ-algebras on these sets).

Theorem

Theorem 5.1. *Conditional on rotation matrix-valued data $\mathbf{o}_1, \ldots, \mathbf{o}_n$, with $n > 1$ and $\mathbf{o}_i \neq \mathbf{o}_j$ for $i \neq j$, the posterior kernel (6) is proper over $\Omega \times \mathbb{R}_+$ with respect to $H \times \lambda$.*

Proof of Theorem 1. Between a pair $(\mathbf{S}, \mathbf{O})$ of rotation matrices, we may define the absolute misorientation angle $r \equiv r_{mis}(\mathbf{S}, \mathbf{O}) \in [0, \pi]$ and misorientation axis $\mathbf{U} \equiv \mathbf{U}_{mis}(\mathbf{S}, \mathbf{O}) \in \mathbb{R}^3$, $\|\mathbf{U}\|^2 = 1$, as the unique solution of

$$
\begin{aligned}
\mathbf{S}^T\mathbf{O} &= \exp[\Phi(r\mathbf{U})] = \mathbf{I}_{3\times3} + \Phi(r\mathbf{U})) + \frac{1}{2!}[\Phi(r\mathbf{U}))]^2 + \cdots \qquad (11)\\
&= \cos r \cdot \mathbf{I}_{3\times3} + \sin r \cdot \Phi(\mathbf{U}) + (1 - \cos r) \cdot \mathbf{U}\mathbf{U}^T
\end{aligned}
$$

where, given any $U \equiv (u_1, u_2, u_3)^T \in \mathbb{R}^3$, $\Phi(U)$ is the skew symmetric matrix defined as

$$
\Phi(U) = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix};
$$

in switching pair order, $r_{mis}(\mathbf{O}, \mathbf{S}) = r_{mis}(\mathbf{S}, \mathbf{O})$ and $\mathbf{U}_{mis}(\mathbf{O}, \mathbf{S}) = -\mathbf{U}_{mis}(\mathbf{S}, \mathbf{O})$ hold. Hence, $3 - tr(\mathbf{S}^T\mathbf{O}) = 2(1 - \cos r) \geq 0$ follows, where $3 - tr(\mathbf{S}^T\mathbf{O}) = 0$ if and only if $r = 0$ (i.e., if $\mathbf{S}^T\mathbf{O} = \mathbf{I}_{3\times3}$).

By defining $a(\mathbf{S}) \equiv 1/\prod_{i=1}^{n} a_i(\mathbf{S})$ and $b(\mathbf{S}) \equiv \sum_{i=1}^{n} a_i(\mathbf{S})$ for $a_i(\mathbf{S}) \equiv [3 - tr(\mathbf{S}^T\mathbf{o}_i)]/2 \geq 0$, $i = 1, \ldots, n$, we algebraically rewrite the posterior kernel (6) as

$$
g(\mathbf{S}, \kappa) = a(\mathbf{S})\kappa^{n/2}\exp[-\kappa b(\mathbf{S})]c(\kappa),
$$

where

$$
c(\kappa) \equiv \left(\frac{\kappa^{-1/2}\exp[\kappa]}{2I_0(\kappa)} \right)^n \sqrt{1 - \frac{1}{\kappa}\frac{I_1(\kappa)}{I_0(\kappa)} - \frac{I_1(\kappa)^2}{I_0(\kappa)^2}}.
$$

Since $\kappa^{-1/2}\exp[\kappa]/I_j(\kappa) \to (2\pi)^{1/2}$ as $\kappa \to \infty$ for $j = 0$ or 1, by basic properties of Bessel functions, we can fix $\kappa_0 \in \mathbb{R}_+$ such that $c(\kappa) \leq 2^{(n+1)/2}$ for $\kappa > \kappa_0$. It suffices to show that the posterior kernel is integrable over regions $\Omega \times (\kappa_0, \infty)$ and $\Omega \times (0, \kappa_0]$.

In the following, let C generically denote a positive constant, not depending on any $\mathbf{S} \in \Omega$, $\kappa > 0$, or given data $\mathbf{o}_1, \ldots, \mathbf{o}_n$. Using the bound on $c(\kappa)$ along with $b(\mathbf{S}) > 0$ for any $\mathbf{S} \in \Omega$ (since the observations $\mathbf{o}_i$ are distinct and $n > 1$), we integrate over $\Omega \times (\kappa_0, \infty)$ as

$$
\begin{aligned}
\int_{\Omega \times (\kappa_0, \infty)} g(\mathbf{S}, \kappa) d(H \times \lambda) &\leq C \int_\Omega a(\mathbf{S}) \left(\int_0^\infty \kappa^{n/2} \exp[-\kappa b(\mathbf{S})] d\lambda \right) dH \\
&\leq C \int_\Omega \frac{a(\mathbf{S})}{[b(\mathbf{S})]^\alpha} dH
\end{aligned}
$$

with $\alpha \equiv (n+2)/2$. Next let $\theta = \min_{1 \leq i < j \leq n} r_{mis}(\mathbf{o}_i, \mathbf{o}_j) > 0$ denote the minimum absolute misorientation angle over all pairs $(\mathbf{o}_i, \mathbf{o}_j)$, and define disjoint subregions $\Omega_i \equiv \{\mathbf{S} \in \Omega : r_{mis}(\mathbf{S}, \mathbf{o}_i) < \theta/2\}$ of Ω, $i = 1, \ldots, n$ (disjointness is verified using $\mathbf{o}_i^T \mathbf{o}_j = (\mathbf{S}^T \mathbf{o}_i)^T (\mathbf{S}^T \mathbf{o}_j)$ and (11)). Note that, for any $1 \leq i \leq n$, we have $a_i(\mathbf{S}) \geq \delta > 0$ for $\delta \equiv 1 - \cos(\theta/2)$ whenever $\mathbf{S} \in \Omega \setminus \Omega_i$ (i.e., if $\mathbf{S} \in \Omega_j$, $i \neq j$). This feature enables a bound

$$
\begin{aligned}
\int_\Omega \frac{a(\mathbf{S})}{[b(\mathbf{S})]^\alpha} dH &\leq \int_{\Omega \setminus \cup_{i=1}^n \Omega_i} \delta^{-n-\alpha} dH + \sum_{i=1}^n \int_{\Omega_i} \frac{\delta^{-(n-1)-\alpha}}{a_i(\mathbf{S})} dH \\
&\leq C + C \sum_{i=1}^n \int_\Omega \frac{1}{a_i(\mathbf{o}_i^T \mathbf{S})} dH
\end{aligned}
$$

using the fact that the Haar measure H integrates to 1 over Ω and is invariant (i.e., if $\mathbf{S} \in \Omega$ is distributed according to H, so is $\mathbf{O}^T \mathbf{S}$ for any fixed $\mathbf{O} \in \Omega$). Finally, if $\mathbf{S}$ follows the Haar measure, then $a_i(\mathbf{o}_i^T \mathbf{S}) = [3 - tr(\mathbf{S})]/2 = 1 - \cos R$ holds for a random variable R, having a Lebesgue density $f(r) = (1 - \cos r)/\pi$, $r \in [0, \pi]$ (Miles 1965); this implies

$$
\int_\Omega \frac{1}{a_i(\mathbf{o}_i^T \mathbf{S})} dH = \int_0^\pi \frac{f(r)}{1 - \cos r} dr = 1.
$$

Hence, the posterior kernel is finitely integrable over $\Omega \times (\kappa_0, \infty)$.

For $\kappa \in (0, \kappa_0]$ and $\mathbf{S} \in \Omega$, the portion $\kappa^{n/2} \exp[-\kappa b(\mathbf{S})] c(\kappa)$ of $g(\mathbf{S}, \kappa)$ is bounded, so that

$$
\int_{\Omega \times (0, \kappa_0]} g(\mathbf{S}, \kappa) d(H \times \lambda) \leq C \int_\Omega a(\mathbf{S}) dH < \infty
$$

follows by similar arguments. $\square$

References

Bingham, M. A., Nordman, D. J., and Vardeman, S. B. (2009a). "Finite-Sample Investigation of Likelihood and Bayes Inference for the Symmetric von Mises-Fisher Distribution." Preprint, Department of Statistics, Iowa State University. 609, 612, 625

— (2009b). "Modeling and Inference for Measured Crystal Orientations and a Tractable Class of Symmetric Distributions for Rotations in Three Dimensions." *Journal of the American Statistical Association*, to appear. 607, 608, 611, 612, 613, 614, 615, 616, 617, 622, 623, 624, 625

Chang, T. and Rivest, L.-P. (2001). "M-Estimation for Location and Regression Parameters in Group Models: A Case Study Using Stiefel Manifolds." *The Annals of Statistics*, 29: 784–814. 608

Chikuse, Y. (2003). *Statistics on Special Manifolds*. New York: Springer. 608

Demirel, M. C., El-Dasher, B. S., Adams, B. L., and Rollett, A. D. (2000). "Studies on the Accuracy of Electron Backscatter Diffraction Measurements." In Schwartz, A. J., Mukul, K., and Adams, B. L. (eds.), *Electron Backscatter Diffraction in Materials Science*, 407–418. New York: Kluwer Academic/Plenum Publishers. 608

Downs, T. D. (1972). "Orientation Statistics." *Biometrika*, 59: 665–676. 607, 608, 609, 610, 626

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis*. Boca Raton: Chapman & Hall. 613

Jupp, P. E. and Mardia, K. V. (1979). "Maximum Likelihood Estimators for the Matrix Von Mises-Fisher and Bingham Distributions." *The Annals of Statistics*, 7: 599–606. 608

Khatri, C. G. and Mardia, K. V. (1977). "The Von Mises-Fisher Matrix Distribution in Orientation Statistics." *Journal of the Royal Statistical Society, Ser. B*, 39: 95–106. 608, 609, 614

León, C. A., Massé, J.-C., and Rivest, L.-P. (2006). "A Statistical Model for Random Rotations." *Journal of Multivariate Analysis*, 97: 412–430. 609, 614

Mardia, K. V. and Jupp, P. E. (2000). *Directional Statistics*. Chichester: New York: John Wiley & Sons. 608, 609

Miles, R. E. (1965). "On Random Rotations in R^3." *Biometrika*, 52: 636–639. 610, 626, 627

Nordman, D. J., Vardeman, S. B., and Bingham, M. A. (2009). "Uniformly Hyper-Efficient Bayes Inference in a Class of Non-Regular Problems." To appear in *The American Statistician*. 617

Prentice, M. J. (1986). "Orientation Statistics without Parametric Assumptions." *Journal of the Royal Statistical Society, Ser. B*, 48: 214–222. 608

Rancourt, D., Rivest, L.-P., and Asselin, J. (2000). "Using Orientation Statistics to Investigate Variations in Human Kinematics." *Journal of the Royal Statistical Society, Ser. C*, 49: 81–94. 607, 608

Randle, V. (2003). *Microtexture Determination and its Applications*. London: Maney for The Institute of Materials, Minerals and Mining. 607, 608

Rivest, L.-P. and Chang, T. (2006). "Regression and Correlation for 3×3 Rotation Matrices." *The Canadian Journal of Statistics*, 34: 187–202. 608

Wilson, A. W. and Spanos, G. (2001). "Application of Orientation Imaging Microscopy to Study Phase Transformations in Steels." *Materials Characterization*, 46: 407–418. 608