Environ Ecol Stat (2018) 25:363–385
https://doi.org/10.1007/s10651-018-0409-3

# A new multimodal and asymmetric bivariate circular distribution

**Fatemeh Hassanzadeh**[1] · **Zeynep Kalaylioglu**[2]

Received: 18 August 2017 / Revised: 21 May 2018 / Published online: 14 June 2018


**Abstract** Multimodal and asymmetric bivariate circular data arise in several different disciplines and fitting appropriate distribution plays an important role in the analysis of such data. In this paper, we propose a new bivariate circular distribution which can be used to model both asymmetric and multimodal bivariate circular data simultaneously. In fact the proposed density covers unimodality as well as multimodality, symmetry as well as asymmetry of circular bivariate data. A number of properties of the proposed density are presented. A Bayesian approach with MCMC scheme is employed for statistical inference. Three real datasets and a simulation study are provided to illustrate the performance of the proposed model in comparison with alternative models such as finite mixture Cosine model.

**Keywords** Asymmetric distribution · Circular bivariate data · Multimodal distribution · Protein structure · Wind direction

## 1 Introduction

Current interest in bivariate directional distributions has increased with their important applications in many different areas such as in modeling wind directions (e.g. Shieh and Johnson 2005; Kato 2009) and structural protein bioinformatics (e.g. see Green and Mardia 2006; Mardia et al. 2007; Dahl et al. 2008; Shieh et al. 2011; Mardia 2013; Fernández-Durán and Gregorio-Domínguez 2014). Here we give a brief historical sur-

Handling Editor: Pierre Dutilleul.

✉ Fatemeh Hassanzadeh
f.hassanzadeh@stat.ui.ac.ir

[1] Department of Statistics, University of Khansar, Khansar, Iran

[2] Department of Statistics, Middle East Technical University, 06800 Ankara, Turkey

vey on and the current status of bivariate circular distributions. A general von Mises bivariate circular model, which is referred as full bivariate von Mises distribution, was introduced by Mardia (1975a) in which correlation in circular context and dependence on torus were well laid out. In terms of wrapped approach, Thompson (1975) mentioned wrapping the bivariate normal distribution doubly around the torus to obtain a bivariate wrapped distribution. This approach was further elaborated by Wehrly and Johnson (1980) and they obtained a family of bivariate distributions with specified marginal distributions. Rivest (1988), Singh et al. (2002) and Mardia et al. (2007) considered submodels of the full bivariate von Mises distribution. More recently, Kato and Pewsey (2015) proposed a five-parameter bivariate wrapped Cauchy distribution as a unimodal model. Jones et al. (2015) developed a novel bivariate circular approach based on copulae on the circle. Mardia et al. (2007), Lennox et al. (2009) and Mardia (2013) are recent examples in the literature in which bivariate circular approaches are developed for bioinformatical studies of the dihedral angles that summarizes the structure of a protein. These distributions are symmetric and unimodal or bimodal.

Many bivariate circular datasets in the literature and ongoing researches include multimodal and/or asymmetric features. This type of data have been analyzed using a mixture of symmetric/asymmetric unimodal or bimodal circular distributions (e.g. Mardia et al. 2007; Dahl et al. 2008; Ferreira et al. 2008). Special asymmetric circular distributions are given in univariate circular context by e.g. Gatto and Jammalamadaka (2007), Umbach and Jammalamadaka (2009), Abe and Pewsey (2011) and Kim and SenGupta (2013). In terms of multivariate situations, recently, Kim et al. (2016) developed a new multivariate distribution with specified marginals for the analysis of bimodal asymmetric/ symmetric circular data. Their joint probability density function (pdf) is a function of marginal cumulative distribution functions which may be difficult to obtain as most circular distribution functions do not have closed forms. Our aim is to propose a general distribution for multimodal asymmetric bivariate circular data. The proposed distribution is a combination of the asymmetric generalized von Mises (AGvM) distribution introduced by Kim and SenGupta (2013) and the generalized von Mises distribution (GvM) (e.g. Cox 1975; Yfantis and Borgman 1982; Gatto and Jammalamadaka 2007). Some important inferential features for GvM distribution are presented in e.g. Yfantis and Borgman (1982). The proposed multimodal asymmetric bivariate distribution will be called hereafter MABvM. The pdf of MABvM is in a known form.

The remainder of the article is organized as follows. Section 2 introduces the MABvM distribution and gives its properties. A Bayesian analysis of the model is given in Sect. 3. In Sect. 4 we describe a small simulation study. Use of the distribution is illustrated in Sect. 5 on several examples including the protein dataset. Finally, the last section includes concluding remarks.

## 2 The distribution specification

Let $(\Theta_1, \Theta_2)^T$ and $(\theta_1, \theta_2)^T$ be bivariate circular random variables and their realized values respectively with joint pdf $f_{\Theta_1,\Theta_2}(\theta_1, \theta_2)$. The joint pdf can be factorized into a marginal and a conditional density functions such as

$$f_{\Theta_1,\Theta_2}(\theta_1,\theta_2) = f_{\Theta_1|\Theta_2}(\theta_1|\theta_2) f_{\Theta_2}(\theta_2).$$

This is also a simple method for generating bivariate distributions and has been used for obtaining dependent models. In this study, we consider AGvM (Kim and SenGupta 2013) and GvM (Gatto and Jammalamadaka 2007) distributions for $f_{\Theta_1|\Theta_2}(\theta_1|\theta_2)$ and $f_{\Theta_2}(\theta_2)$, respectively. The pdf of AGvM$(\mu, \kappa_1, \kappa_2)$ distribution is

$$f_\Theta(\theta) = \frac{1}{2\pi C\left(\frac{\pi}{4}, \kappa_1, \kappa_2\right)} \exp\{\kappa_1 \cos(\theta - \mu) + \kappa_2 \sin(2\theta - 2\mu)\};\ \theta \in [0, 2\pi), \tag{1}$$

where $\mu \in [0, 2\pi)$ is a location parameter, $\kappa_1 > 0$ and $\kappa_2 \in [-1, 1]$ are concentration and skewness parameters, respectively. The three parameter model belongs to an exponential family of distributions and can be used to model both asymmetric and bimodal univariate data. The pdf of GvM$(\mu_1, \mu_2, \kappa_1, \kappa_2)$ distribution is given below

$$f_\Theta(\theta) = \frac{1}{2\pi C(\delta, \kappa_1, \kappa_2)} \exp\{\kappa_1 \cos(\theta - \mu_1) + \kappa_2 \cos(2\theta - 2\mu_2)\};\ \theta \in [0, 2\pi), \tag{2}$$

where $\mu_1 \in [0, 2\pi)$, $\mu_2 \in [0, \pi)$ are location parameters, $\delta = \mu_1 - \mu_2$ and $\kappa_1, \kappa_2 > 0$ are the shape parameters. The GvM distribution covers unimodality as well as multimodality, symmetry as well as asymmetry in circular univariate data. The distributions (1) and (2) are the extensions of the von Mises (vM) distribution. Also, (1) is a subclass of (2), i.e. AGvM$(\mu, \kappa_1, \kappa_2)$=GvM$(\mu, \mu + \pi/4, \kappa_1, \kappa_2)$ with restriction on $\kappa_2$. Following definition introduces a new multimodal and asymmetric bivariate circular (MABvM) distribution.

**Definition 2.1** A bivariate circular random variable $(\Theta_1, \Theta_2)$ is said to follow the MABvM distribution, if its density is given by

$$\begin{aligned} f_{\Theta_1,\Theta_2}(\theta_1,\theta_2) = & \frac{1}{4\pi^2 C(\delta, \kappa_1, \kappa_2, \kappa_3, \kappa_4)} \\ & \times \exp\{\kappa_1 \cos(\theta_1 - \mu - \varphi\theta_2) + \kappa_2 \sin(2\theta_1 - 2\mu - 2\varphi\theta_2)\} \\ & \times \exp\{\kappa_3 \cos(\theta_2 - \mu_1) + \kappa_4 \cos(2\theta_2 - 2\mu_2)\}, \end{aligned} \tag{3}$$

where $0 \le \theta_1, \theta_2 < 2\pi$, $0 \le \mu, \mu_1 < 2\pi$, $0 \le \mu_2 < \pi$, $\delta = \mu_1 - \mu_2$, and $C(\delta, \kappa_1, \kappa_2, \kappa_3, \kappa_4)$ is the normalizing constant that is defined in Proposition 2.2. Also $\kappa_1, \kappa_3, \kappa_4 \ge 0$, $\varphi \in [-1, 1]$ and $\kappa_2 \in [-1, 1]$.

The MABvM is a periodic distribution i.e. $f_{\Theta_1,\Theta_2}(\theta_1, \theta_2) = f_{\Theta_1,\Theta_2}(\theta_1 + 2\pi m, \theta_2 + 2\pi m)$ for any integer $m$. The parameter $\varphi$ accounts for the statistical dependence between $\Theta_1$ and $\Theta_2$: if $\varphi = 0$ then $\Theta_1$ and $\Theta_2$ are independent. In this case, $\Theta_1$ follows an AGvM$(\mu,\kappa_1,\kappa_2)$ distribution with the location parameter $\mu$ and $\Theta_2$ follows a GvM$(\mu_1,\mu_2,\kappa_3,\kappa_4)$ distribution. The distribution has eight parameters where $\mu$, $\mu_1$ and $\mu_2$ are the location parameters. The parameters $\kappa_3$ and $\kappa_4$ are shape

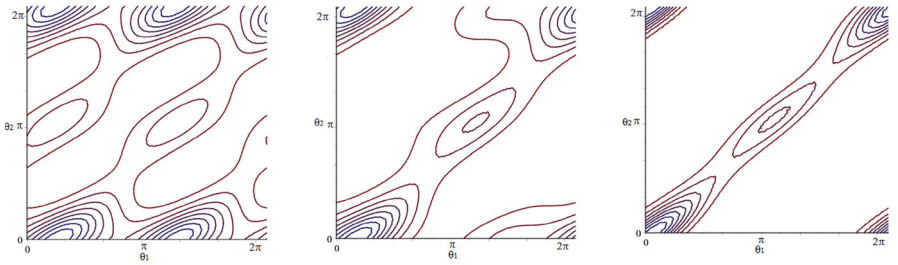


**Fig. 1** Contour of MABvM(0, 0, 0, $\kappa_1$, 0.5, 0.5, 0.5, 1) distribution with various values of $\kappa_1$. **a** $\kappa_1 = 0$, **b** $\kappa_1 = 1$, **c** $\kappa_1 = 5$

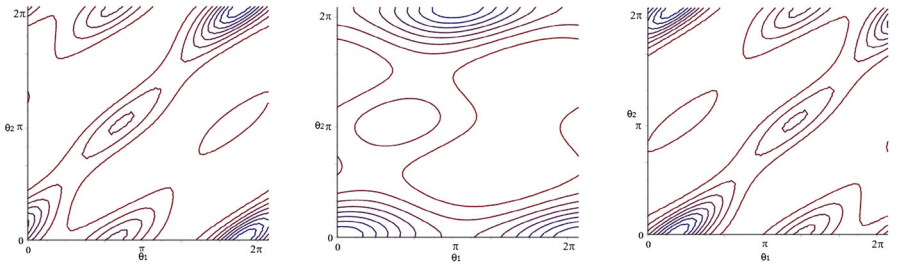


**Fig. 2** Contour of MABvM(0, 0, 0, 0.5, $\kappa_2$, 0.5, 0.5, 1) distribution with various values of $\kappa_2$. **a** $\kappa_2 = -1$, **b** $\kappa_2 = 0$, **c** $\kappa_2 = 1$

parameters. The parameters $\kappa_1$ and $\kappa_2$ model the circular association between $\Theta_1$ and $\Theta_2$. The role of $\kappa_1$ and $\kappa_2$ are illustrated by the two dimensional contours of the MABvM($\mu$, $\mu_1$, $\mu_2$, $\kappa_1$, $\kappa_2$, $\kappa_3$, $\kappa_4$, $\varphi$) distribution given in Figs. 1 and 2. As seen in these figures, as $\kappa_1$ increases from 0 to 5 two dimensional contours get centered around the $\theta_1 = \theta_2$ line. Similarly, as $|\kappa_2|$ increases, the mass centers around the $\theta_1 = \theta_2$ line. Thus the association between two variables increases when $\kappa_1$ or $|\kappa_2|$ increases.

For $\kappa_1 = \kappa_2 = \kappa_3 = \kappa_4 = 0$ the distribution reduces to a bivariate uniform distribution. For $\kappa_2 = 0$ the MABvM is a combination of an AGvM conditional and a vM marginal distribution. For $\kappa_4 = 0$ the MABvM is a combination of a GvM conditional and a vM marginal distribution and if $\kappa_2 = \kappa_4 = 0$ it is a combination of two vM distributions. Listed below are the properties of MABvM.

(i) **Normalizing Constant** Here we give an expression for $C(\delta, \kappa_1, \kappa_2, \kappa_3, \kappa_4)$ in terms of an infinite series, involving sequences of modified Bessel functions, which can be computed efficiently using the algorithm of Amos (1974).

**Proposition 2.2** *The normalizing constant in* (3) *is given by*

$$C(\delta,\kappa_1,\kappa_2,\kappa_3,\kappa_4) = \left\{ I_0(\kappa_1)I_0(\kappa_2) + \sum_{j=1}^{\infty} I_{2j}(\kappa_1)I_j(\kappa_2)\cos\left(\frac{j}{2\pi}\right)\right\} \times \left\{ I_0(\kappa_3)I_0(\kappa_4) + \sum_{l=1}^{\infty} I_{2l}(\kappa_3)I_l(\kappa_4)\cos(2l\delta)\right\}, \quad (4)$$

*where* $I_m(\cdot)$ *is the modified Bessel function of the first kind and order* $m$.

The proof is given in the Appendix.

(ii) **Conditional and Marginal Distributions** Conditional and marginal distributions of MABvM are given below which can be proved easily.

**Proposition 2.3** *Let* $(\Theta_1,\Theta_2)$ *follow the MABvM with pdf* (3). *Then*

(a) *conditional distribution of* $\Theta_1$ *given* $\Theta_2 = \theta_2$ *is AGvM*$(\mu+\varphi\theta_2, \kappa_1, \kappa_2)$ *as follows*

$$f_{\Theta_1|\Theta_2}(\theta_1|\theta_2) = \frac{1}{2\pi C\left(\frac{\pi}{4},\kappa_1,\kappa_2\right)} \exp\{\kappa_1\cos(\theta_1-\mu-\varphi\theta_2) + \kappa_2\sin 2(\theta_1-\mu-\varphi\theta_2)\}, \quad (5)$$

*where* $C\left(\frac{\pi}{4},\kappa_1,\kappa_2\right)$ *is the normalizing constant that is defined in the Appendix,* $-1 \le \varphi \le 1$ *shrinks the influence of* $\Theta_2 = \theta_2$, $\kappa_1 \ge 0$ *and* $\kappa_2 \in [-1, 1]$ *are concentration and skewness parameters, respectively.*

(b) *conditional distribution of* $\Theta_2$ *given* $\Theta_1 = \theta_1$ *is proportional to* $f_{\Theta_1,\Theta_2}(\theta_1,\theta_2)$ *in* (3). *As a special case when* $\varphi = 1$ *and* $\mu = 0$, *conditional distribution of* $\Theta_2$ *given* $\Theta_1 = \theta_1$ *is obtained as*

$$f_{\Theta_2|\Theta_1}(\theta_2|\theta_1) = C_2\exp\{A_1\cos(\theta_2) + B_1\sin(\theta_2) + C_1\cos(2\theta_2) + D_1\sin(2\theta_2)\},$$

*where* $A_1 = \kappa_1\cos(\theta_1) + \kappa_3\cos(\mu_1)$, $B_1 = \kappa_1\sin(\theta_1) + \kappa_3\sin(\mu_1)$, $C_1 = \kappa_2\sin(2\theta_1) + \kappa_4\cos(2\mu_2)$, $D_1 = \kappa_2\cos(2\theta_1) + \kappa_4\sin(2\mu_2)$ *and* $C_2$ *is the normalizing constant.*

(c) *marginal pdf of* $\Theta_2$ *is GvM*$(\mu_1,\mu_2,\kappa_3,\kappa_4)$ *distribution in* (2).

(d) *marginal pdf of* $\Theta_1$ *does not have a known form. Thus we approximate it by using Monte Carlo method as follows*

$$f(\theta_1) \approx \frac{1}{M}\sum_{m=1}^{M} f_{\Theta_1|\Theta_2}(\theta_1|\theta_{2,m})$$

*where* $\{\theta_{2,m}\}_{m=1}^{M}$ *are independent observations from the GvM*$(\mu_1,\mu_2,\kappa_3,\kappa_4)$ *distribution and* $f_{\Theta_1|\Theta_2}(\theta_1|\theta_{2,m})$ *denotes the pdf of the AGvM*$(\mu+\varphi\theta_{2,m},\kappa_1,\kappa_2)$ *distribution in* (5).

Proposition 2.3 (a) and (c) are straightforward since the normalizing constant in Definition 2.1 is written as the product of the normalizing constants for the conditional and marginal distributions. The marginal distribution of $\Theta_2$ and the conditional of $\Theta_1$ are quite flexible since AGvM and GvM can model either symmetric or asymmetric, unimodal or bimodal distributions, depending on the values of its four parameters. For $\kappa_4 = 0$ the marginal distribution of $\Theta_2$ is a vM distribution.

There is a relationship between the proposed model and the specified conditional models (Arnold and Strauss 1991) given in SenGupta (2004). The exponential part of (3) can be written as

$$p'(\theta_1)\begin{pmatrix} m & \kappa_3\cos(\mu_1) & \kappa_3\sin(\mu_1) & \kappa_4\cos(2\mu_2) & \kappa_4\sin(2\mu_2) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & c_1 & c_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -c_2 & c_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_3 & c_4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_4 & c_3 \end{pmatrix} q(\theta_2)$$

where $c_1 = \kappa_1\cos(\mu)$, $c_2 = \kappa_1\sin(\mu)$, $c_3 = \kappa_2\cos(2\mu)$,and $c_4 = -\kappa_2\sin(2\mu)$,

$$\begin{aligned} p'(\theta_1) &= (1, \cos(\theta_1), \sin(\theta_1), \cos(2\theta_1), \sin(2\theta_1))\,, \\ q'(\theta_2) &= (1, \cos(\theta_2), \sin(\theta_2), \cos(2\theta_2), \sin(2\theta_2), \\ &\qquad \cos(\varphi\theta_2), \sin(\varphi\theta_2), \cos(2\varphi\theta_2), \sin(2\varphi\theta_2)), \end{aligned}$$

and $m$ is a function of the other parameters. SenGupta (2004) has given an example of specified conditionals model (Arnold and Strauss 1991) on the torus with joint density

$$f(\theta_1, \theta_2) = p'(\theta_1) M q(\theta_2), \quad (\theta_1, \theta_2) \in [0, 2\pi)^2$$

Here $p'(\theta_1) = (1, \cos(\theta_1), \sin(\theta_1))$, $q'(\theta_2) = (1, \cos(\theta_2), \sin(\theta_2))$ and

$$M = \begin{pmatrix} m_{00} & m_{01} & m_{02} \\ m_{10} & m_{11} & m_{12} \\ m_{20} & m_{21} & m_{22} \end{pmatrix}$$

where $m_{00}$ is a function of the other eight $m_{jk}$s. The exponential part of the proposed distribution is similar to the joint pdf on the torus with vM conditionals in SenGupta (2004). In addition, model (3) with many zero-elements has advantage as it reduces the number of parameters.

(iii) **Multimodality** Distribution may be unimodal or multimodal, depending on the parameter values. Multimodality conditions are given by the following proposition.

**Proposition 2.4** *The joint pdf* (3) *has two modes if* $\kappa_1 < 2|\kappa_2|$ *or* $\kappa_3 < 4\kappa_4$ *and it has four modes if* $\kappa_1 < 2|\kappa_2|$ *and* $\kappa_3 < 4\kappa_4$*, otherwise it is a unimodal distribution for* $\kappa_1 \neq 0$ , $\kappa_2 \neq 0$, *and* $\kappa_4 \neq 0$.

The proof is given in the Appendix. Figure 3 provides a graphical illustration. Figure 3a–d represent different examples for varying ($\kappa_1$,$\kappa_2$,$\kappa_3$,$\kappa_4$).

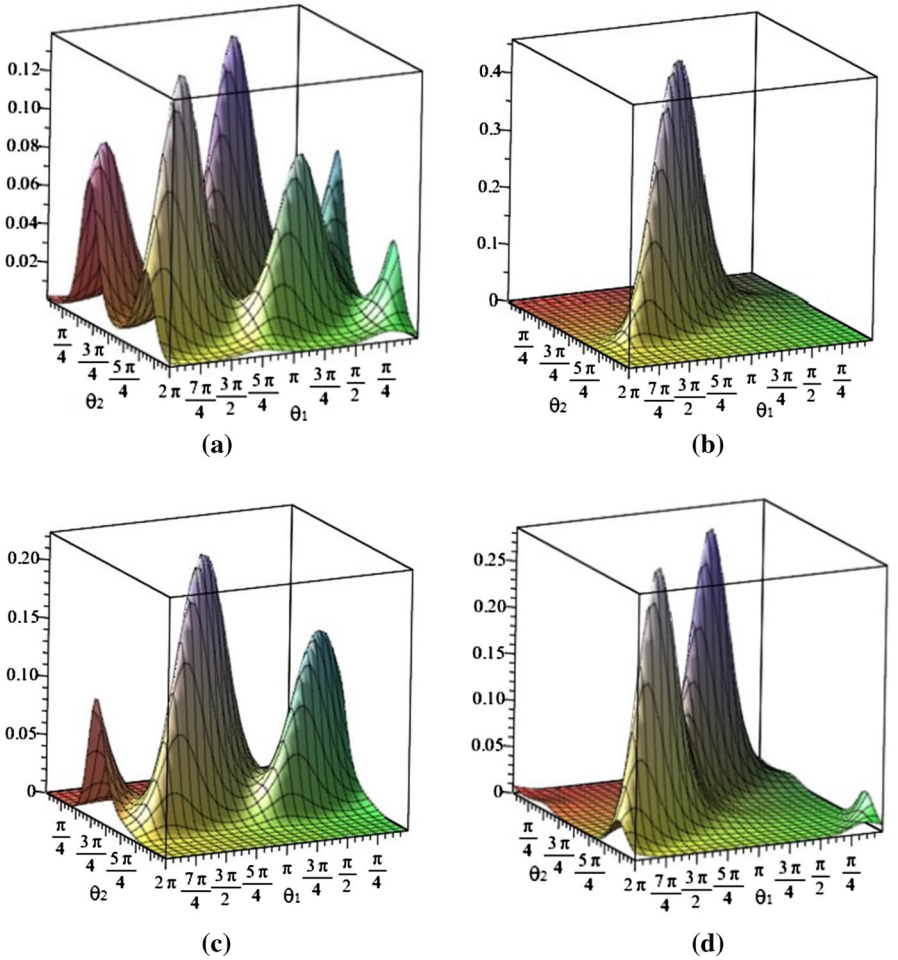


**Fig. 3** Joint pdf of MABvM $(0, \pi, \frac{\pi}{2}, \kappa_1, \kappa_2, \kappa_3, \kappa_4, 1)$. **a** $(\kappa_1, \kappa_2, \kappa_3, \kappa_4) = (0.3, 1, 1, 1)$, **b** $(\kappa_1, \kappa_2, \kappa_3, \kappa_4) = (3, 1, 5, 1)$, **c** $(\kappa_1, \kappa_2, \kappa_3, \kappa_4) = (0.3, 1, 5, 1)$, **d** $(\kappa_1, \kappa_2, \kappa_3, \kappa_4) = (3, 1, 1, 1)$

(iv) **Exchangeability** Two random variables $\Theta_1$ and $\Theta_2$ are exchangeable if $(\Theta_2, \Theta_1)$ and $(\Theta_1, \Theta_2)$ are identically distributed i.e. $f_{\Theta_2,\Theta_1}(\theta_2, \theta_1) = f_{\Theta_1,\Theta_2}(\theta_1, \theta_2)$. Based on Definition 2.1, it is clear that $f_{\Theta_2,\Theta_1}(\theta_2, \theta_1) \neq f_{\Theta_1,\Theta_2}(\theta_1, \theta_2)$. That is, exchangeability does not hold. It means that the properties of the proposed distribution is not invariant, similar to Shieh and Johnson (2005), under permutations of its arguments. General definition of exchangeability can be found in De Finetti (1972).

(v) **Asymmetry** If a multivariate distribution is symmetric then all marginals are symmetric (Mardia et al. 2008). Conclusively, a multivariate distribution is asymmetric if the mass is distributed asymmetrically with respect to a reference point in at least one of its dimensions. Note that marginal pdf of $\Theta_2$ can be asymmetric as a result of Proposition 2.3(d) which in turn leads to an asymmetric MABvM. Then MABvM can be symmetric or asymmetric.

(vi) **Maximum Entropy** Mardia (1975b) introduced maximum entropy distributions based on specifying the first marginal of the first variable and joint harmonics. Accordingly, Sect. 2.3. in Mardia (1975b), suppose that distributions defined over a space S are to be represented by densities relative to some familiar measure such as Lebesgue and Haar. Let $t_l, \ldots, t_q$ represent $q$ given real valued functions over S such that no linear combination of $t_l, \ldots, t_q$ is constant. If, for a density function $f(x)$,

i. $S_1$ is the support of $f(x)$ where $x \in S_1$, $S_1 \subset S$,
ii. $E\{t_i(x)\} = a_i$ (fixed) , $t = 1, \ldots, q$,
iii. The entropy is maximized,

Then $f(x)$ should be of the form

$$f(x) = \exp\left\{b_0 + \sum_{i=1}^{q} b_i t_i(x)\right\}, \; x \in S_1 \tag{6}$$

provided there exist $b_0, b_1, \ldots, b_q$ such that (6) satisfies (i) and (ii).

The proposed distribution (3) is a maximum entropy distribution subject to $a_1 = E(\cos(\theta_1)\cos(\varphi\theta_2))$, $a_2 = E(\sin(\theta_1)\sin(\varphi\theta_2))$, $a_3 = E(\cos(\theta_1)\sin(\varphi\theta_2))$, $a_4 = E(\sin(\theta_1)\cos(\varphi\theta_2))$, $a_5 = E(\sin(2\theta_1)\cos(2\varphi\theta_2))$, $a_6 = E(\cos(2\theta_1)\sin(2\varphi\theta_2))$, $a_7 = E(\cos(2\theta_1)\cos(2\varphi\theta_2))$, $a_8 = E(\sin(2\theta_1)\sin(2\varphi\theta_2))$, $a_9 = E(\cos(\theta_2))$, $a_{10} = E(\sin(\theta_2))$, $a_{11} = E(\cos(2\theta_2))$, and $a_{12} = E(\sin(2\theta_2))$ taking specified values consistent with the expectations with respect to the proposed distribution. Since Cosine and Sine functions are bounded, the $a_k$, for $k = 1, 2, \ldots, 12$ , exists. Also we obtain $b_1 = b_4 = \kappa_1\cos(\mu)$, $b_3 = -b_2 = \kappa_1\sin(\mu)$, $b_5 = b_7 = \kappa_2\cos(2\mu)$, $b_6 = -b_8 = \kappa_2\sin(2\mu)$, $b_9 = \kappa_3\cos(\mu_1)$, $b_{10} = \kappa_3\sin(\mu_1)$, $b_{11} = \kappa_4\cos(2\mu_2)$, and $b_{12} = \kappa_4\sin(2\mu_2)$.

## 3 Bayesian estimation

We adopt Bayesian methods based on Markov chain Monte Carlo (MCMC) algorithms to obtain statistical inference of a MABvM distribution. We use the hierarchical representation of the MABvM for an efficient Gibbs sampling. Gibbs sampler provides a way for generating samples from the posteriors that are specified up to a proportionality. The scheme proceeds iteratively by generating variates in a cyclic manner from the various full conditionals, specified up to proportionality, using suitably chosen initial values for starting the process. The iterations are carried out until a systematic pattern of convergence is achieved through the generating scheme. It has been observed that after a large number of iterations the generated samples converge in distribution to a random sample from the true posterior distribution under Ergodicity conditions [see, for example, Smith and Roberts (1993) and Robert and Casella (2004) for a detailed discussion].

Let data $(\theta_{11}, \theta_{21}), (\theta_{12}, \theta_{22}), \ldots, (\theta_{1n}, \theta_{2n})$ be a sample of size $n$ from a bivariate MABvM distribution. Hierarchical specification of MABvM for $i = 1, 2, \ldots, n$, is given by

$$\theta_{1i}|\theta_{2i} \sim AGvM(\mu + \varphi\theta_{2i}, \kappa_1, \kappa_2),$$
$$\theta_{2i} \sim GvM(\mu_1, \mu_2, \kappa_3, \kappa_4).$$

Then the corresponding likelihood function is

$$L(\vartheta|\theta_1, \theta_2) = \prod_{i=1}^{n} f_{AGvM}(\theta_{1i}|\theta_{2i}) f_{GvM}(\theta_{2i}),$$

where $\vartheta$ is a vector of parameters i.e. $\vartheta = (\mu, \mu_1, \mu_2, \kappa_1, \kappa_2, \kappa_3, \kappa_4, \varphi)$. We assume a conjugate prior for $(\mu_1, \mu_2)$ and $\mu$ as follows

$$f(\mu_1, \mu_2) = \{C(\delta, \kappa_3 R_{01}, \kappa_4 R_{02})\}^{-r} \exp\{\kappa_3 R_{01} \cos(\mu_1 - \mu_{01}) + \kappa_4 R_{02} \cos(2\mu_2 - 2\mu_{02})\},$$

and

$$f(\mu) = \left\{C\left(\frac{\pi}{4}, \kappa_1 L_{01}, \kappa_2 L_{02}\right)\right\}^{-l} \exp\{\kappa_1 L_{01} \cos(\mu - \mu_0) + \kappa_2 L_{02} \sin(2\mu - 2\mu_0)\},$$

where $0 \le \mu, \mu_1 < 2\pi$, $0 \le \mu_2 < \pi$, $r$ and $l$ are integers which show the number of realizations from the prior distributions, $C(\delta, \kappa_3 R_{01}, \kappa_4 R_{02})$ and $C\left(\frac{\pi}{4}, \kappa_1 L_{01}, \kappa_2 L_{02}\right)$ are the normalizing constant of the GvM and AGvM distributions, respectively, and $(\mu_0, \mu_{01}, \mu_{02}, R_{01}, R_{02}, L_{01}, L_{02})$ are the fixed hyper-parameters. All the other parameters are assumed to be apriori independent. We consider $U(-1, 1)$ for prior distributions of $\kappa_2$ and $\varphi$ , and $Ga\left(\alpha_j, \beta_j\right)$ for $\kappa_j$, $j = 1, 3, 4$ . Then the joint posterior distribution is given by

$$\pi(\vartheta|\theta_1, \theta_2) \propto L(\vartheta|\theta_1, \theta_2) f(\mu_1, \mu_2|\mu_{01}, \mu_{02}, R_{01}, R_{02}) f(\mu|\mu_0, L_{01}, L_{02}) f(\kappa_1|\alpha_1, \beta_1) \times f(\kappa_3|\alpha_3, \beta_3) f(\kappa_4|\alpha_4, \beta_4) I_{\varphi}[-1, 1] I_{\kappa_2}[-1, 1].$$

where $I$ denotes an indicator function. The MCMC methods, including Gibbs sampling, are based upon the full conditional posterior distributions. Joint full conditional posterior for $(\mu_1, \mu_2)$ is given by

$$\pi(\mu_1, \mu_2|\cdot) = \{C(\delta, \kappa_3, \kappa_4)\}^{-m} \exp\left\{\kappa_3 R_{n_1} \cos\left(\mu_1 - \mu_{n_1}\right) + \kappa_4 R_{n_2} \cos 2\left(\mu_2 - \mu_{n_2}\right)\right\}, \quad (7)$$

where $m = r + n$. Also $\mu_{n_1}, \mu_{n_2}, R_{n_1}$ and $R_{n_2}$ are satisfied in the following equations $R_{n_1} \cos\left(\mu_{n_1}\right) = R_{01} \cos(\mu_{01}) + \sum_i \cos(\theta_{2i})$, $R_{n_2} \cos\left(2\mu_{n_2}\right) = R_{02} \cos(2\mu_{02}) + \sum_i \cos(2\theta_{2i})$, $R_{n_1} \sin\left(\mu_{n_1}\right) = R_{01} \sin(\mu_{01}) + \sum_i \sin(\theta_{2i})$ and $R_{n_2} \sin\left(2\mu_{n_2}\right) = R_{02} \sin(2\mu_{02}) + \sum_i \sin(2\theta_{2i})$. Posterior inferences for $\mu_1$ and $\mu_2$ are obtained using the following marginal full conditionals derived from (7)

$$\pi(\mu_1|\cdot) \propto \exp\left\{\kappa_3 R_{n_1} \cos\left(\mu_1 - \mu_{n_1}\right)\right\}, \tag{8}$$

$$\pi(\mu_2|\cdot) \propto \exp\left\{\kappa_4 R_{n_2} \cos 2\left(\mu_2 - \mu_{n_2}\right)\right\}. \tag{9}$$

Full conditionals given in (8) and (9) are kernels of $vM\left(\mu_{n_1}, \kappa_3 R_{n_1}\right)$ and $vM\left(2\mu_{n_2}, \kappa_4 R_{n_2}\right)$ respectively. Several straightforward methods to sample data from the von Mises distribution are available (Best and Fisher 1979). Thus, the Gibbs sampler algorithm can be easily applied to (8) and (9) to simulate random draws from the corresponding posterior. The full conditional distribution for $\mu$ is obtained as follows

$$\pi(\mu|\cdot) = \left\{C\left(\frac{\pi}{4}, \kappa_3, \kappa_4\right)\right\}^{-M} \exp\left\{\kappa_1 L_{n_1} \cos(\mu_1 - \mu_n) + \kappa_2 L_{n_2} \sin 2(\mu_2 - \mu_n)\right\},$$

where $M = l + n$. Also $\mu_n$, $L_{n_1}$ and $L_{n_2}$ are satisfied in the following equations $L_{n_1}\cos(\mu_n) = L_{10}\cos(\mu_0) + \sum_i \cos(\theta_{1i} - \varphi\theta_{2i})$, $L_{n_2}\cos(2\mu_n) = L_{20}\cos(2\mu_0) - \sum_i \cos(2\theta_{1i} - 2\varphi\theta_{2i})$, $L_{n_1}\sin\left(\mu_{n_1}\right) = L_{10}\sin(\mu_0) + \sum_i \sin(\theta_{1i} - \varphi\theta_{2i})$ and $L_{n_2}\sin(2\mu_n) = L_{20}\sin(2\mu_0) - \sum_i \sin(2\theta_{1i} - 2\varphi\theta_{2i})$. The other full conditional distributions are as follows

$$\pi(\kappa_1|\cdot) \propto \left\{C\left(\frac{\pi}{4}, \kappa_1, \kappa_2\right)\right\}^{-n} \kappa_1^{\alpha_1 - 1} \exp\left\{\kappa_1\left(\sum_{i=1}^{n} \cos(\theta_{1i} - \mu - \varphi\theta_{2i}) - \beta_1\right)\right\},$$

$$\pi(\kappa_2|\cdot) \propto \left\{C\left(\frac{\pi}{4}, \kappa_1, \kappa_2\right)\right\}^{-n} \exp\left\{\kappa_2 \sum_{i=1}^{n} \sin 2(\theta_{1i} - \mu - \varphi\theta_{2i})\right\} I_{\kappa_2}[-1, 1],$$

$$\pi(\kappa_3|\cdot) \propto \{C(\delta, \kappa_3, \kappa_4)\}^{-m} \kappa_3^{\alpha_3 - 1} \exp\left\{\kappa_3\left(\sum_{i=1}^{n} \cos(\theta_{2i} - \mu_2) - \beta_3\right)\right\},$$

$$\pi(\kappa_4|\cdot) \propto \{C(\delta, \kappa_3, \kappa_4)\}^{-m} \kappa_4^{\alpha_4 - 1} \exp\left\{\kappa_4\left(\sum_{i=1}^{n} \cos 2(\theta_{2i} - \mu_2) - \beta_4\right)\right\},$$

$$\pi(\varphi|\cdot) \propto \exp\left\{\kappa_1 \sum_{i=1}^{n} \cos(\theta_{1i} - \varphi\theta_{2i}) + \kappa_2 \sum_{i=1}^{n} \sin 2(\theta_{1i} - \varphi\theta_{2i})\right\} I_{\varphi}[-1, 1],$$

where $C(.)$ are normalizing constants of a GvM distribution. These are not analytically tractable. Thus Bayesian inference will be performed using MCMC methods such as Metropolis-within-Gibbs algorithm (e.g. Geweke and Tanizaki 2001). Convergence is monitored via MCMC chain histories, Gelman-Rubin diagnostic, autocorrelation and density plots.

## 4 A simulation study

In this section we conduct a simulation experiment to assess the performance of the Bayesian estimation for MABvM distribution. The experiment is controlled for the sample size and the strength of the association between the two circular random variables. The data are generated from a MABvM distribution. We chose different values

**Table 1** Posterior mean and SD of the parameters, $\varphi = -0.8$

| | $n = 30$ | | $n = 200$ | | $n = 500$ | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| $\mu_1(\pi)$ | 181.1 | 7.646 | 179.4 | 2.947 | 179.4 | 1.674 |
| $\mu(0)$ | 0.093 | 0.164 | 0.059 | 0.075 | 0.059 | 0.065 |
| $\kappa_1(0.5)$ | 0.476 | 0.183 | 0.495 | 0.081 | 0.500 | 0.055 |
| $\kappa_2(0.8)$ | 0.375 | 0.293 | 0.907 | 0.073 | 0.797 | 0.076 |
| $\kappa_3(3)$ | 3.021 | 0.583 | 2.990 | 0.325 | 2.990 | 0.256 |
| $\kappa_4(1)$ | 1.007 | 0.227 | 0.997 | 0.114 | 0.997 | 0.104 |
| $\varphi(0.8)$ | −0.788 | 0.362 | −0.796 | 0.015 | −0.796 | 0.015 |

**Table 2** Posterior mean and SD of the parameters, $\varphi = 0$

| | $n = 30$ | | $n = 200$ | | $n = 500$ | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| $\mu_1(\pi)$ | 176.6 | 8.053 | 176.6 | 2.985 | 180.2 | 2.958 |
| $\mu(0)$ | 0.724 | 0.599 | 0.163 | 0.071 | 0.083 | 0.071 |
| $\kappa_1(0.5)$ | 0.480 | 0.128 | 0.487 | 0.077 | 0.503 | 0.071 |
| $\kappa_2(0.8)$ | 0.795 | 0.135 | 0.799 | 0.125 | 0.799 | 0.106 |
| $\kappa_3(3)$ | 2.913 | 0.699 | 2.922 | 0.286 | 3.005 | 0.248 |
| $\kappa_4(1)$ | 0.765 | 0.372 | 0.881 | 0.162 | 1.003 | 0.135 |
| $\varphi(0)$ | −0.028 | 0.058 | 0.014 | 0.018 | −0.009 | 0.013 |

for $\varphi$, representing different degrees of associations, namely $\varphi = -0.8, 0, 0.8$. When $\varphi = 0$ the variables $\Theta_1$ and $\Theta_2$ are independent. The parameters are fixed at certain values seen in Tables 1, 2 and 3. To reduce the number of parameters we set $\mu_2 = 0.5\mu_1$. Also, we consider different sample sizes, namely $n = 30, 200, 500$. Using the OpenBUGs software, we run 11,000 samples after removing 5000 as burn-in period. There is no evidence of lack of convergence based on examinations of histories, Gelman-Rubin diagnostic, kernel density, and autocorrelation plots. For example, posterior plots of $\varphi$ for $n = 30$ and true value $\varphi = 0$ are given in the Appendix.

Results of the experiment are given in Tables 1, 2 and 3. Labels *Mean* and *SD* are posterior means and standard deviations respectively. True values are included in the parentheses. We observe that with $n = 30$, the bias and SD of each estimator is moderate and, decrease as $n$ increases. In small samples, when the two circular variables are negatively associated, $\kappa_2$ estimation has considerable bias. Also by using various values of hyperparameters we obtained similar results implying that posterior estimates are not sensitive to the prior in this Bayesian analysis.

**Table 3** Posterior mean and SD of the parameters, $\varphi = 0.8$

| | $n = 30$ | | $n = 200$ | | $n = 500$ | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| $\mu_1(\pi)$ | 182.3 | 5.695 | 179.4 | 0.975 | 179.4 | 0.974 |
| $\mu(0)$ | 0.504 | 0.482 | 0.059 | 0.076 | 0.059 | 0.075 |
| $\kappa_1(0.5)$ | 0.575 | 0.085 | 0.502 | 0.081 | 0.500 | 0.051 |
| $\kappa_2(0.8)$ | 0.874 | 0.095 | 0.797 | 0.091 | 0.799 | 0.063 |
| $\kappa_3(3)$ | 2.854 | 0.458 | 2.99 | 0.315 | 2.99 | 0.201 |
| $\kappa_4(1)$ | 0.981 | 0.157 | 0.997 | 0.131 | 0.998 | 0.104 |
| $\varphi(0.8)$ | 0.798 | 0.049 | 0.787 | 0.013 | 0.779 | 0.007 |

## 5 Numerical examples

This section provides three real data applications to assess the performance of the suggested Bayesian algorithm for MABvM and the performance of the proposed model in relation to the existent models. The datasets have previously been analyzed in several other circular papers for various different purposes. First two of the competitive models considered here are the Sine and Cosine models by Mardia (1975a) that are produced from the full bivariate vM pdf that is proportional to

$$\begin{aligned}&\exp[\kappa_1\cos(\theta_1-\mu_1)+\kappa_2\cos(\theta_2-\mu_2)+\alpha\cos(\theta_1-\mu_1)\cos(\theta_2-\mu_2)\\&\quad+\beta\sin(\theta_1-\mu_1)\sin(\theta_2-\mu_2)],\end{aligned}\tag{10}$$

where $0 \le \theta_1, \theta_2 < 2\pi$, $0 \le \mu_1, \mu_2 < 2\pi$, $\kappa_1, \kappa_2 \ge 0$, and $\alpha, \beta \in \mathbb{R}$. This pdf boils down to the Sine and Cosine models for $(\alpha, \beta) = (0, \lambda)$ and $(\alpha, \beta) = (\kappa_3, -\kappa_3)$ respectively. In addition, we consider mixtures of Cosine and Sine models as competitive models here. Mixtures of Cosine and Sine models are discussed more broadly in Mardia et al. (2007). The other competitive bivariate circular models are presented by Shieh and Johnson (2005) and Kato (2009). The pdf of Shieh and Johnson model is proportional to

$$\exp[\kappa_1\cos(\theta_1-\mu_1)+\kappa_2\cos(\theta_2-\mu_2)+\kappa_3\cos[2\pi\{F_1(\theta_1)-F_2(\theta_2)\}-\mu_3]],\tag{11}$$

and the pdf of Kato model is proportional to

$$\begin{aligned}&\exp[\kappa_1\cos(\theta_1-\mu_1)+\kappa_2\cos(\theta_2-\mu_2)]\\&\times\left[1+|\psi|^2-2|\psi|\cos\left[2\pi\{F_1(\theta_1)-F_2(\theta_2)\}-\arg(\psi)\right]\right]^{-1},\end{aligned}\tag{12}$$

where $0 \le \theta_1, \theta_2 < 2\pi$, $\mu_j \in [0, 2\pi)$, $\kappa_j \ge 0$, for $j = 1, 2, 3$, and $F_\kappa(\cdot)$ are the distribution functions of $\text{vM}(\mu_k, \kappa_k)$, $k = 1, 2$. Details about $\psi$ are given in Kato (2009). These models have the vM marginal distributions. We also fit the exchanged MABvM (EMABvM) as a MABvM model where the angles are denoted by $\Theta_2$ and

$\Theta_1$, respectively. We employ Deviance Information Criterion (DIC) to compare the models with each other.

We also wish to evaluate if the considered distribution well captures the multimodality, if present, in the dataset. Let $C_k$, $k = 2, 4$ be a modality condition, e.g. $C_2 = (\widetilde{\kappa}_1 < 2|\widetilde{\kappa}_2|\text{or}\ \widetilde{\kappa}_3 < 4|\widetilde{\kappa}_4|)$ for bimodality and $C_4 = (\widetilde{\kappa}_1 < 2|\widetilde{\kappa}_2|\text{and}\ \widetilde{\kappa}_3 < 4|\widetilde{\kappa}_4|)$ for four-modes. We used the posterior probability $P_k = P(C_k|\theta_1, \theta_2)$ to make inference about the modality. Small $P_k$ indicates unimodality.

MCMC technique is implemented using OpenBugs software version 3.2.3. In the Bayesian analysis, the first 5000 iterations are used as burn-in. The next 10,000 iterations are used for posterior inference. Convergence was monitored via MCMC chain histories, autocorrelation and density plots.

### 5.1 Protein data

Protein structures are closely related with molecular ecology. One of the important characteristics of these structures is protein conformational angles that is a natural pairing of angles denoted by $(\theta_1, \theta_2)$ with each angle in $[-\pi, \pi)$. Protein angles data can be accessed in the Protein Structure Databank (PDB) (Berman et al. 2000) that consists of large number of angle pairs. Different samples of the protein conformational angles have been employed previously by Mardia et al. (2007), fitting mixture of bivariate von Mises distributions. We use a subsample consisting of 255 paired observations to investigate the performances of MABvM and the other models considered herein.

Scatter plots and contours of fitted bivariate distributions are given in Fig. 4. In each panel, scatter plot of the angles, the contour of the fitted models, and the DICs are given for visual and numerical assessment of the fit respectively. Accordingly, circular scatterplots of $\theta_1$ and $\theta_2$ display multimodality represented by different clusters seen in the figures. The similar plot in Mardia et al. (2007) confirms this general property about the protein conformational angles as well. The fitted density of model MABvM seems to show a satisfactory fit. Also $P_2 = 1$ demonstrates that MABvM is appropriate to analyze this dataset.

Based on the DIC values, the MABvM model seems to be the most efficient among all in terms of balancing the goodness of the fit with model complexity. Between the competitive models, mixture of bivariate Cosine model is performing slightly better than EMABvM. Also mixture models perform better than the Sine and Cosine models. The difference between the models, MABvM and bivariate mixture models in particular, for this data are further investigated in Fig. 8 given in the Appendix. The parameter estimates of MABvM are $\widetilde{\mu} = 0.001\,(0.006)$, $\widetilde{\mu_1} = 0.026\,(0.045)$, $\widetilde{\mu_2} = 0.016\,(0.015)$, $\widetilde{\kappa}_1 = 0.225\,(0.104)$, $\widetilde{\kappa}_2 = 0.080\,(0.0.067)$, $\widetilde{\kappa}_3 = 0.244\,(0.143)$, $\widetilde{\kappa}_4 = 1.445\,(0.219)$ and $\widetilde{\varphi} = -0.589\,(0.148)$, with posterior standard deviations given in parenthesis. The test for independence is equivalent to testing $\varphi = 0$. Since 95% credible intervals of $\widetilde{\varphi}$ is $(-0.819, -0.149)$, circular dependency between $\Theta_1$ and $\Theta_2$ is significant.

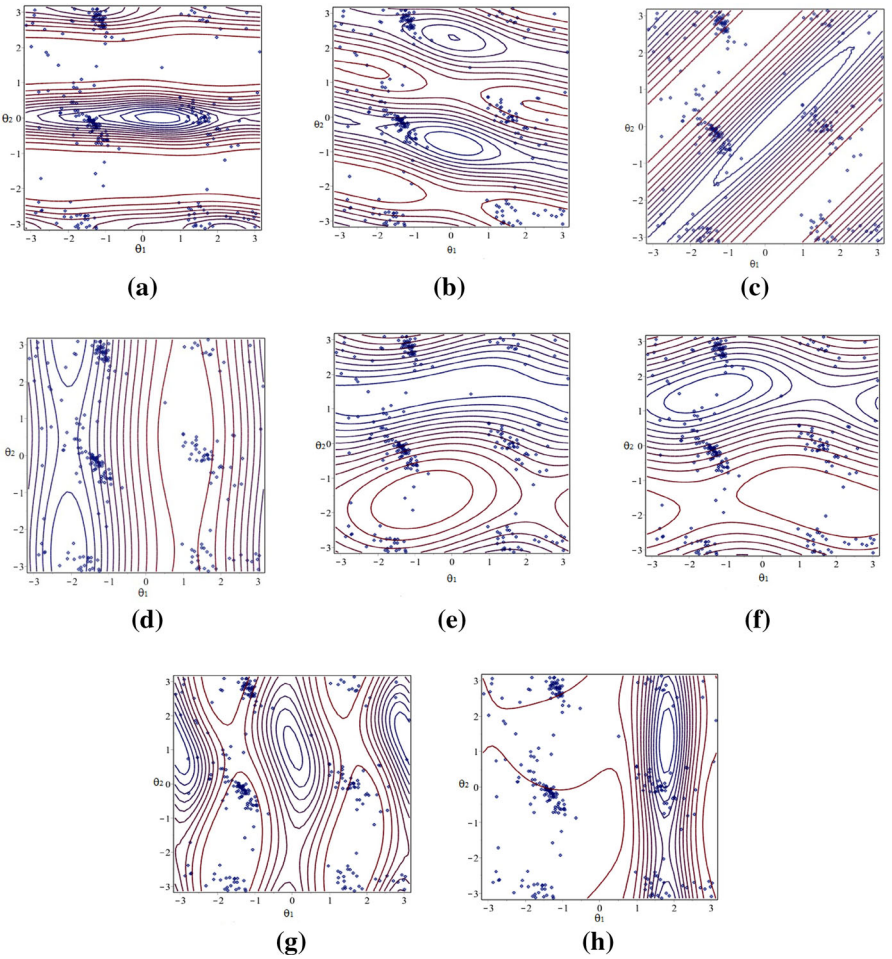


**Fig. 4** The scatter plots, the contours, and DICs of the fitted models for protein data. **a** MABvM model, DIC = 6350.2, **b** EMABvM model, DIC = 6359.1, **c** Shieh and Johnson model, DIC = 6496, **d** Kato model, DIC = 6540, **e** bivariate Cosine model, DIC = 6541, **f** bivariate Sine model, DIC = 6543, **g** mixture of bivariate Cosine model, DIC = 6355, **h** mixture of bivariate Sine model, DIC = 6407

## 5.2 Wind data

Wind direction is important as it impacts a wide variety of environmental elements, wind power and wildfires to name only a few. Previously, various wind directions have been analyzed for various different applications and used for different theoretical statistical researches (e.g. see Shieh and Johnson 2005 and Kato 2009). The two particular data sets considered here have previously been used to illustrate the models developed in the seminal circular articles. In these examples each angles representing wind directions belong to $[0, 2\pi)$. Our aim here is to compare the performance of

the proposed model with those other bivariate circular models applied on these well-known data sets.

*Example 1* In this example, we analyze a dataset consisting pairs of wind directions measured at a weather station in Texas. The dataset is a sample of a larger dataset which is taken from a website http://data.eol.ucar.edu/codiac/dss/id=85.034. The original dataset contains hourly resolution surface meteorological data from the Texas Natural Resources Conservation Commission Air Quality Monitoring Network. In this dataset, we consider 30 pairs of wind directions recorded at 6 a.m. and 7 a.m. each day at a weather station from June 1 to June 30, 2003. The same subset has previously been used by Kato (2009) to illustrate their model. Figure 5 shows scatter plot of the real data, the contour of the fitted models and the corresponding DICs. The scatterplot displays unimodality and suggests that there is an association between the wind directions at 6 a.m. and 7 a.m.

MABvM model provides the smallest DIC among all the competitive models. The estimated parameters of MABvM model are given by $\widetilde{\mu} = 0.056\,(0.021)$, $\widetilde{\mu_1} = 2.411\,(0.219)$, $\widetilde{\mu_2} = 1.110\,(0.177)$, $\widetilde{\kappa_1} = 2.300\,(0.609)$, $\widetilde{\kappa_2} = 0.372\,(0.435)$, $\widetilde{\kappa_3} = 1.432\,(0.361)$, $\widetilde{\kappa_4} = 0.178\,(0.094)$ and $\widetilde{\varphi} = 0.886\,(0.068)$. Also $P_2 = 0$ and $P_4 = 0$ show that MABvM model is unimodal. Parameter $\varphi$ controls the dependency between the two circular variables and $\widetilde{\varphi}$ with the 95% credible interval $(0.777, 0.927)$ confirms that the wind directions at 6 a.m. and 7 a.m. are associated. This result is in line with the result of Kato (2009) about the association of these two variables.

*Example 2* We illustrate the utility of model (3) also by employing it for the analysis of a set of 21 paired observations of $\Theta_1$ and $\Theta_2$, wind directions recorded for 21 consecutive days at 6 a.m. and noon respectively at a Milwaukee weather station. The same set has previously been analyzed by Johnson and Wehrly (1977) in their seminal article as well as more recently by Shieh and Johnson (2005). Figure 6 gives the scatter plots, the contours, and DICs of the fitted models. The dependency structure between the two variables is not as clear as in the previous example. Based on the plots given in the figure, it appears that the MABvM, EMABvM and mixture of bivariate cosine distribution with two components may be particularly appropriate for modelling these directions.

DIC values suggest that EMABvM is the best fitting model while the model MABvM and the mixture of cosine models are also competitive. The DICs of MABvM and the mixture of cosine model are quite close. One should note that the number of parameters MABvM is less than that in the mixture of cosine model. In EMABvM model, the estimates are $\widetilde{\mu} = 0.110\,(0.009)$, $\widetilde{\mu_1} = 0.382\,(0.545)$, $\widetilde{\mu_2} = 0.203\,(0.199)$, $\widetilde{\kappa_1} = 0.338\,(0.148)$, $\widetilde{\kappa_2} = 0.632\,(0.256)$, $\widetilde{\kappa_3} = 0.311\,(0.167)$, $\widetilde{\kappa_4} = 0.233\,(0.128)$ and $\widetilde{\varphi} = 0.119\,(0.053)$. In this case, $P_4 = 1$ thus the fitted EMABvM is a multimodal model. Also the test for independence is equivalent to testing $\varphi = 0$. Since 95% credible interval of $\widetilde{\varphi}$ is $(0.017, 0.229)$, the circular dependency between the directions is significant. This is in line with the likelihood ratio test results in Shieh and Johnson (2005).

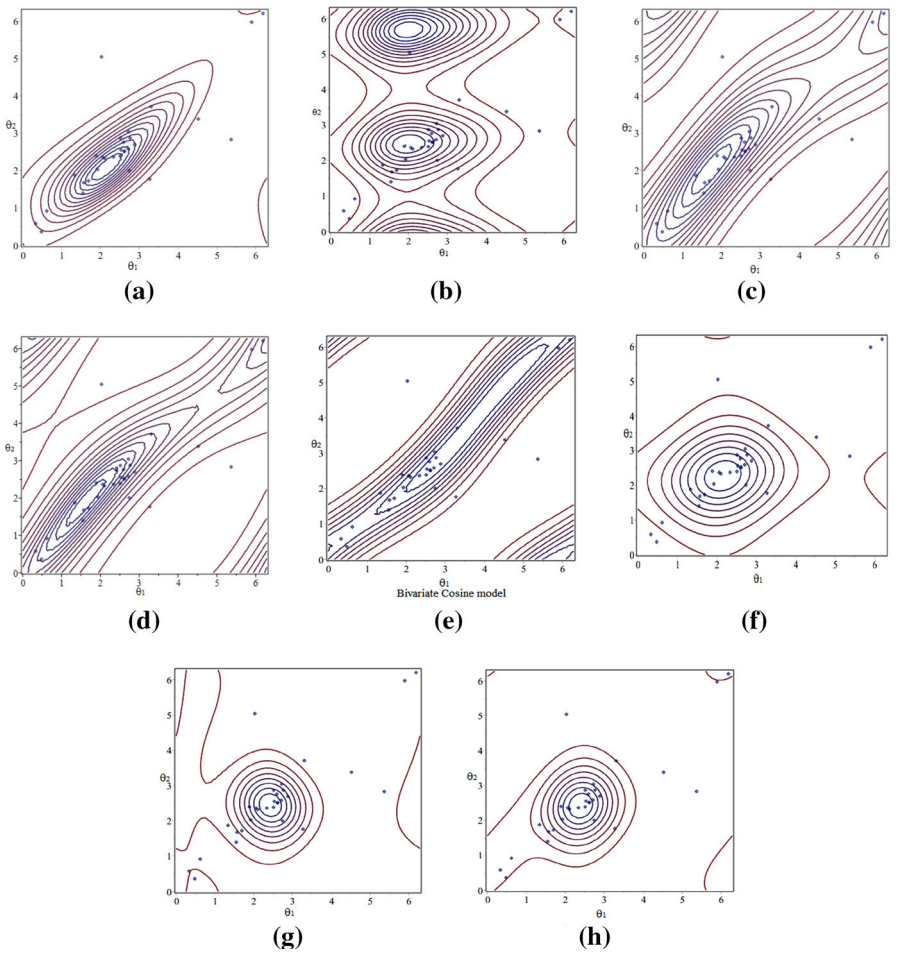


**Fig. 5** The scatter plots, the contours, and DICs of the fitted models for the 30 samples of wind data (Example 1). **a** MABvM model, DIC = 707.2, **b** EMABvM model, DIC = 747.7, **c** Shieh and Johnson model, DIC = 717.2, **d** Kato model, DIC = 722.2, **e** bivariate Cosine model, DIC = 716.9, **f** bivariate Sine model, DIC = 736.2, **g** mixture of bivariate Cosine model, DIC = 714.2, **h** mixture of bivariate Sine model, DIC = 720

## 6 Concluding remarks

In the present paper, the asymmetry and multimodality of bivariate circular random variables are addressed by utilizing a novel distribution that is denoted by MABvM distribution. The proposed distribution that is based on both the GvM and AGvM distributions achieves a more flexible distributions for the analysis of bivariate circular data. It is notably easy to implement due to its hierarchical form based on GvM and AGvM distributions. We addressed several important features of this distribution such as obtaining the normalizing constant as a function of modified Bessel function and periodicity.

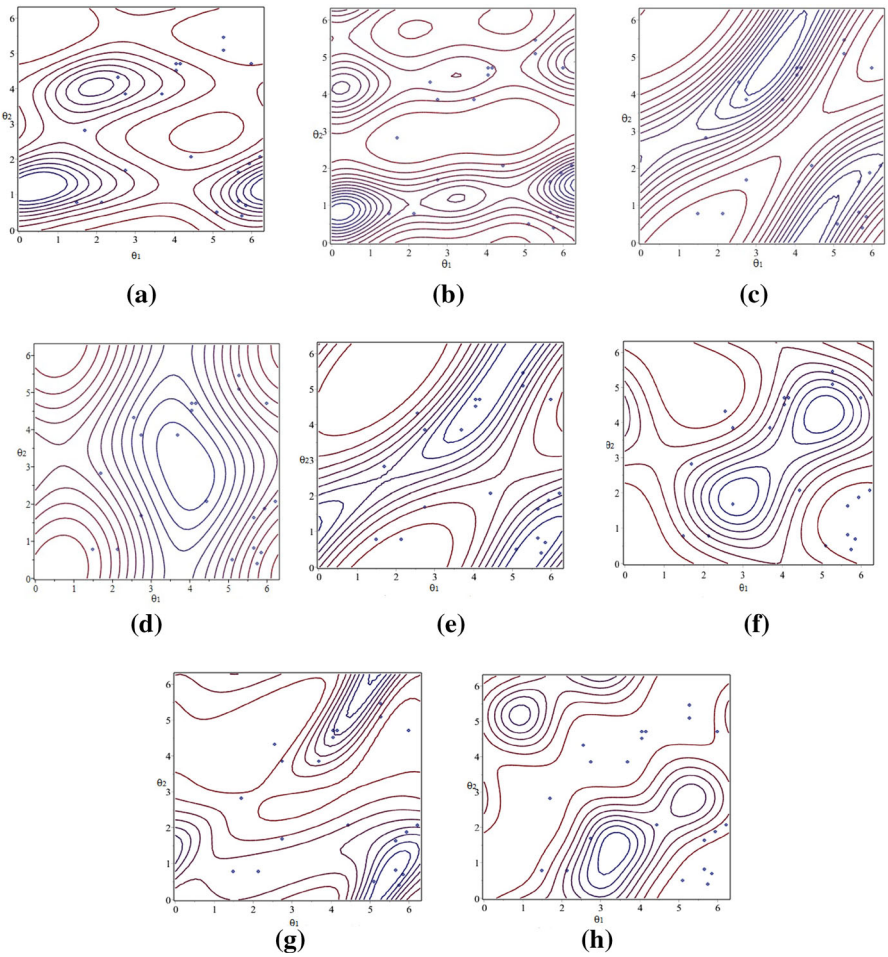


**Fig. 6** The scatter plots, the contours, and DICs of the fitted models for the 21 samples of wind data (Example 2). **a** MABvM model, DIC = 534.3, **b** EMABvM model, DIC = 532.1, **c** Shieh and Johnson model, DIC = 537.7, **d** Kato model, DIC = 542, **e** bivariate Cosine model, DIC = 538.0, **f** bivariate Sine model, DIC = 540.0, **g** mixture of bivariate Cosine model, DIC = 534.6, **h** mixture of bivariate Sine model, DIC = 538.8

The proposed MABvM is applied to a small simulation study and also to three datasets. We used Bayesian estimation and the Gibbs sampling algorithms for the estimation. Other estimation techniques, e.g. frequentist, can also be applied in this framework. Based on the DIC values, the MABvM model or the exchanged version of it (EMABvM) seems to be most efficient among all the competing models considered herein, in terms of balancing the goodness of the fit with model complexity. For the future work we suggest applying the MABvM regression model to investigate the effect of covariates on the location parameters.

**Acknowledgements** Authors gratefully acknowledge Editor-in-Chief and Reviewers for their valuable comments. The first author is grateful to The Scientific and Technological Research Council of Turkey (TUBITAK) for the support.

## Appendix

*Proof of Proposition 2.2* The proof is easily obtained since

$$C(\delta, \kappa_1, \kappa_2, \kappa_3, \kappa_4) = C\left(\frac{\pi}{4}, \kappa_1, \kappa_2\right) C(\delta, \kappa_3, \kappa_4)$$

where $C\left(\frac{\pi}{4}, \kappa_1, \kappa_2\right)$ and $C(\delta, \kappa_3, \kappa_4)$ are normalizing constants of AGvM and GvM, respectively. Gatto and Jammalamadaka (2007) maintained

$$C(\delta, \kappa_3, \kappa_4) = I_0(\kappa_3) I_0(\kappa_4) + \sum_{l=1}^{\infty} I_{2l}(\kappa_3) I_l(\kappa_4) \cos(2l\delta).$$

Also $C\left(\frac{\pi}{4}, \kappa_1, \kappa_2\right)$ can be obtained by choosing $\delta = \frac{\pi}{4}$ in Eq. (4).

*Proof of Proposition 2.4* We can derive the conditions for number of modes of the MABvM density when $\mu = \mu_1 = \mu_2 = 0$ and $\varphi \neq 0$, without loss of generality. To locate the critical values, we take partial derivatives of $\log(f(\theta_1, \theta_2))$ in Eq. (3) w.r.t

**Table 4** Critical points of MABvM when $\theta_2 = 0$

| $\theta_1$ | $\kappa_2 > 0$ | $\kappa_2 < 0$ | Conditions |
|---|---|---|---|
| $\arctan\left(\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_1 - 1\right), \sqrt{R_1}\right)$ | Mode | Saddle point | – |
| $\arctan\left(\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_1 - 1\right), -\sqrt{R_1}\right)$ | Saddle point | Mode | – |
| $\arctan\left(\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_2 - 1\right), \sqrt{R_2}\right)$ | Saddle point | Mode | $\kappa_1 < 2\lvert\kappa_2\rvert$ |
| $\arctan\left(\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_2 - 1\right), -\sqrt{R_2}\right)$ | Mode | Saddle point | $\kappa_1 < 2\lvert\kappa_2\rvert$ |

**Table 5** Critical points of MABvM when $\theta_2 = \pi$

| $\theta_1$ | $\kappa_2 > 0$ | $\kappa_2 < 0$ | Conditions |
|---|---|---|---|
| $\arctan\left(-\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_1 - 1\right), \sqrt{R_1}\right)$ | Saddle point | Mode | $\kappa_3 < 4\kappa_4$ |
| $\arctan\left(-\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_1 - 1\right), -\sqrt{R_1}\right)$ | Mode | Saddle point | $\kappa_3 < 4\kappa_4$ |
| $\arctan\left(-\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_2 - 1\right), \sqrt{R_2}\right)$ | Mode | Saddle point | $\kappa_1 < 2\lvert\kappa_2\rvert, \kappa_3 < 4\kappa_4$ |
| $\arctan\left(-\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_2 - 1\right), -\sqrt{R_2}\right)$ | Saddle point | Mode | $\kappa_1 < 2\lvert\kappa_2\rvert, \kappa_3 < 4\kappa_4$ |

**Table 6** Critical points of MABvM when $\theta_2 = \arctan(R_3, R_4)$ with condition $\kappa_1 < 2|\kappa_2|$ and $\kappa_3 < 4\kappa_4$

| $\theta_1$ | $\kappa_2 > 0$ | $\kappa_2 < 0$ |
|---|---|---|
| $\arctan\left(\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_1 - 1\right), \sqrt{R_1}\right) + \arctan(R_3, R_4)$ | Saddle point | Anti-mode |
| $\arctan\left(\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_1 - 1\right), -\sqrt{R_1}\right) + \arctan(R_3, R_4)$ | Anti-mode | Saddle point |
| $\arctan\left(\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_2 - 1\right), \sqrt{R_2}\right) + \arctan(R_3, R_4)$ | Anti-mode | Saddle point |
| $\arctan\left(\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_2 - 1\right), -\sqrt{R_2}\right) + \arctan(R_3, R_4)$ | Saddle point | Anti-mode |

**Table 7** Critical points of MABvM when $\theta_2 = \arctan(-R_3, R_4)$ with condition $\kappa_1 < 2|\kappa_2|$ and $\kappa_3 < 4\kappa_4$

| $\theta_1$ | $\kappa_2 > 0$ | $\kappa_2 < 0$ |
|---|---|---|
| $\arctan\left(\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_1 - 1\right), \sqrt{R_1}\right) + \arctan(-R_3, R_4)$ | Anti-mode | Saddle point |
| $\arctan\left(\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_1 - 1\right), -\sqrt{R_1}\right) + \arctan(-R_3, R_4)$ | Saddle point | Anti-mode |
| $\arctan\left(\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_2 - 1\right), \sqrt{R_2}\right) + \arctan(-R_3, R_4)$ | Saddle point | Anti-mode |
| $\arctan\left(\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_2 - 1\right), -\sqrt{R_2}\right) + \arctan(-R_3, R_4)$ | Anti-mode | Saddle point |

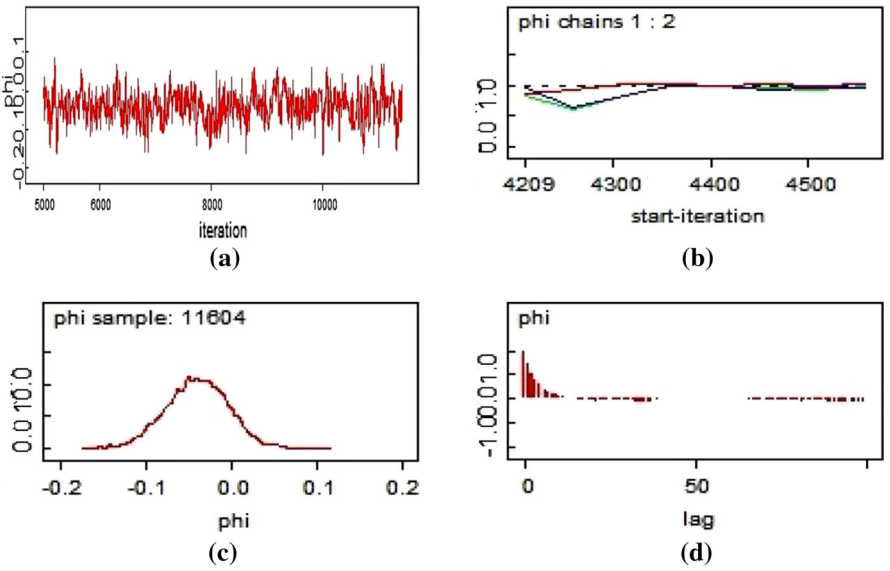


**Fig. 7** The posterior plots of $\varphi$ in simulation study with real value $\varphi = 0$ when $n = 30$. **a** History plot, **b** Gelman–Rubin diagnostic, **c** Kernel density plot, **d** autoregressive plot

$\theta_1$ and $\theta_2$. Let $g = \log(f(\theta_1, \theta_2)) = \log C + \kappa_1 \cos(\theta_1 - \varphi\theta_2) + \kappa_2 \sin(2\theta_1 - 2\varphi\theta_2) + \kappa_3 \cos(\theta_2) + \kappa_4 \cos(2\theta_2)$. Then

$$G_{\theta_1} = \frac{\partial g}{\partial \theta_1} = \kappa_1 \sin(\varphi\theta_2 - \theta_1) + 2\kappa_2 \cos(2\varphi\theta_2 - 2\theta_1)$$

and

$$G_{\theta_2} = \frac{\partial g}{\partial \theta_2} = -G_{\theta_1} + \kappa_3 \sin(-\theta_2) + 2\kappa_4 \sin(-2\theta_2)$$

There are sixteen critical points for $G_{\theta_1} = G_{\theta_2} = 0$. Let $D_{\theta_1,\theta_2} = G_{\theta_1,\theta_1} G_{\theta_2,\theta_2} - G^2_{\theta_1,\theta_2}$,

(i) if $D_{\theta_1,\theta_2} > 0$ and $G_{\theta_1,\theta_1} < 0$ then $(\theta_1, \theta_2)$ is a mode.
(ii) if $D_{\theta_1,\theta_2} > 0$ and $G_{\theta_1,\theta_1} > 0$ then $(\theta_1, \theta_2)$ is an anti-mode.
(iii) if $D_{\theta_1,\theta_2} < 0$ then $(\theta_1, \theta_2)$ is a saddle point.

We can find the number of modes after simplification of underlying algebraic equations. The critical points are given in Tables 4, 5, 6 and 7 where $R_1 = (4\kappa_2)^{-2}\left(-\kappa_1^2 + 16\kappa_2^2 + \sqrt{\kappa_1^4 + 32\kappa_1^2\kappa_2^2}\right)$, $R_2 = (4\kappa_2)^{-2}\left(-2\kappa_1^2 + 32\kappa_2^2 - 2\sqrt{\kappa_1^4 + 32\kappa_1^2\kappa_2^2}\right)$, $R_3 = \frac{\sqrt{-\kappa_3^2 + 16\kappa_4^2}}{4\kappa_4}$, and $R_4 = -\frac{\kappa_3}{4\kappa_4}$.
In these tables $\arctan(x, y)$ gives the arc tangent of $\frac{y}{x}$. Table 4 illustrates that there is no restrictions on unimodality when $\theta_1 = \arctan\left(\frac{2\kappa_2}{\kappa_1}\left(\frac{1}{2}R_1 - 1\right), \sqrt{R_1}\right)$ and $\theta_2 = 0$. Checking all of the critical points shows that if $\kappa_1 < 2|\kappa_2|$ or $\kappa_3 < 4\kappa_4$ then MABvM has two modes. Also if $\kappa_1 < 2|\kappa_2|$ and $\kappa_3 < 4\kappa_4$ then MABvM has four modes.

*Posterior plots in simulation study when* $n = 30$ *and* $\varphi = 0$ Figure 7 shows the history, Gelman–Rubin diagnostic, kernel density, and autocorrelation plots from the simulation study. We define two sequences from different starting points. The Gelman–Rubin diagnostic shows that the behavior of the sequence of chains is the same. Therefore, the variance within the chains is the same as the variance across the chains. Also, the autocorrelation plot reveals there is low correlation between successive samples and the history plot moves up and down around the mode of the distribution. Thus the samples will reach a stationary distribution.

*3-D plots of protein data and the fitted models* The difference between the proposed models and the alternative models is an advantage of our study, since we have defined a different model to analysis bivariate circular data. An additional figure and comment on the difference are now provided as follows. Accordingly, the MABvM model seems to provide a better fit for the ruggedness of the histogram (i.e. the frequencies that should not be ignored). The 3-D histogram of protein data and the kernel density plots of the competitive models are given in Fig. 8.

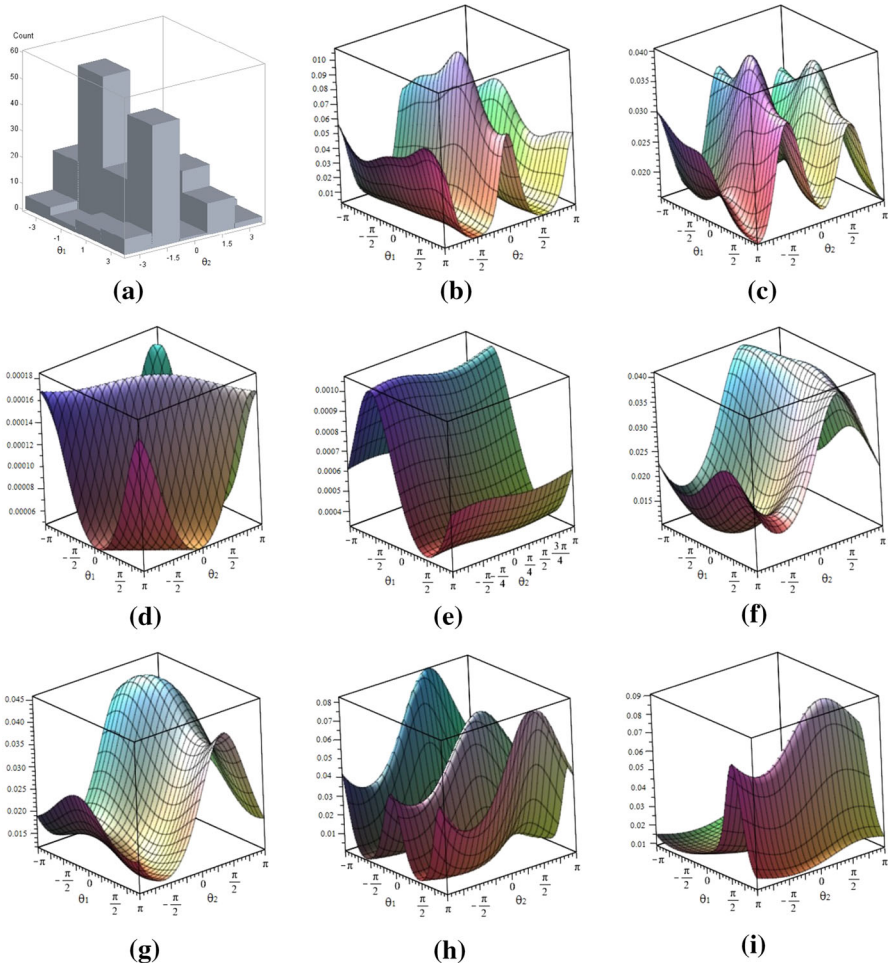


**Fig. 8** The 3-D histogram of protein data and kernel density plots of the fitted models. **a** Histogram plot, **b** MABvM model, **c** EMABvM model, **d** Shieh and Johnson model, **e** Kato model, **f** bivariate Cosine model, **g** bivariate Sine model, **h** mixture of bivariate Cosine model, **i** mixture of bivariate Sine model

## References

Abe T, Pewsey A (2011) Sine-skewed circular distributions. Stat Pap 52(3):683–707

Amos DE (1974) Computation of modified Bessel functions and their ratios. Math Comput 28(125):239–251

Arnold BC, Strauss DJ (1991) Bivariate distributions with conditionals in prescribed exponential families. J R Stat Soc Ser B Methodol 53:365–375

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

Best D, Fisher NI (1979) Efficient simulation of the von Mises distribution. Appl Stat 28:152–157

Cox DR (1975) Contribution to discussion of Mardia (1975a). J R Stat Soc Ser B Methodol 37:380–381

Dahl DB, Bohannan Z, Mo Q, Vannucci M, Tsai J (2008) Assessing side-chain perturbations of the protein backbone: a knowledge-based classification of residue Ramachandran space. J Mol Biol 378(3):749–758

De Finetti B (1972) Probability, induction, and statistics. Wiley, New York
Fernández-Durán JJ, Gregorio-Domínguez MM (2014) Modeling angles in proteins and circular genomes using multivariate angular distributions based on multiple nonnegative trigonometric sums. Stat Appl Genet Mol Biol 13(1):1–18
Ferreira JT, Jurez MA, Steel MF (2008) Directional log-spline distributions. Bayesian Anal 3(2):297–316
Gatto R, Jammalamadaka SR (2007) The generalized von Mises distribution. Stat Methodol 4(3):341–353
Geweke J, Tanizaki H (2001) Bayesian estimation of state-space models using the Metropolis Hastings algorithm within Gibbs sampling. Comput Stat Data Anal 37(2):151–170
Green PJ, Mardia KV (2006) Bayesian alignment using hierarchical models, with applications in protein bioinformatics. Biometrika 93(2):235–254
Johnson RA, Wehrly T (1977) Measures and models for angular correlation and angular-linear correlation. J R Stat Soc Ser B Methodol 39(2):222–229
Jones MC, Pewsey A, Kato S (2015) On a class of circulas: copulas for circular distributions. Ann Inst Stat Math 67(5):843–862
Kato S (2009) A distribution for a pair of unit vectors generated by Brownian motion. Bernoulli 15(3):898–921
Kato S, Pewsey A (2015) A Möbius transformation-induced distribution on the torus. Biometrika 102(2):359–370
Kim S, SenGupta A (2013) A three-parameter generalized von Mises distribution. Stat Pap 54(3):685–693
Kim S, SenGupta A, Arnold BC (2016) A multivariate circular distribution with applications to the protein structure prediction problem. J Multivar Anal 143:374–382
Lennox KP, Dahl DB, Vannucci M, Tsai JW (2009) Density estimation for protein conformation angles using a bivariate von Mises distribution and Bayesian nonparametrics. J Am Stat Assoc 104(486):586–596
Mardia KV (1975a) Statistics of directional data (with discussion). J R Stat Soc Ser B Methodol 37:349–393
Mardia KV (1975b) Characterizations of directional distributions. In: Patil GP, Kotz S, Ord JK (eds) Statistical distributions in scientific work, vol 3. Reidel, Dordrecht, pp 365–386
Mardia KV (2013) Statistical approaches to three key challenges in protein structural bioinformatics. J R Stat Soc Ser C Appl Stat 62(3):487–514
Mardia KV, Taylor CC, Subramaniam GK (2007) Protein bioinformatics and mixtures of bivariate von Mises distributions for angular data. Biometrics 63(2):505–512
Mardia KV, Hughes G, Taylor CC, Singh H (2008) A multivariate von Mises distribution with applications to bioinformatics. Can J Stat 36(1):99–109
Rivest LP (1988) A distribution for dependent unit vectors. Commun Stat Theory Methods 17(2):461–483
Robert CP, Casella G (2004) Monte Carlo statistical methods. Springer texts in statistics. Springer, Berlin
SenGupta A (2004) On the constructions of probability distributions for directional data. Bull Calcutta Math Soc 96:139–154
Shieh GS, Johnson RA (2005) Inferences based on a bivariate distribution with von Mises marginal. Ann Inst Stat Math 57(4):789–802
Shieh GS, Zheng S, Johnson RA, Chang Y-F, Shimizu K, Wang C-C, Tang S-L (2011) Modeling and comparing the organization of circular genomes. Bioinformatics 27:912–918
Singh H, Hnizdo V, Demchuk E (2002) Probabilistic model for two dependent circular variables. Biometrika 89(3):719–723
Smith AFM, Roberts GO (1993) Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. J R Stat Soc Ser B Methodol 55:3–25
Thompson JW (1975) Contribution to discussion of paper by K. V. Mardia. J R Stat Soc Ser B Methodol 37:379
Umbach D, Jammalamadaka SR (2009) Building asymmetry into circular distributions. Stat Probab Lett 79(5):659–663
Wehrly TE, Johnson RA (1980) Bivariate models for dependence of angular observations and a related Markov process. Biometrika 67(1):255–256
Yfantis EA, Borgman LE (1982) An extension of the von Mises distribution. Commun Stat Theory Methods 11:1695–1706

**Fatemeh Hassanzadeh** received her Ph.D. in Statistics from University of Isfahan. Her thesis title was 'Analysis of Count Data with Flexible Mixed Poisson Models'. She was a post doctoral researcher at the same university on the topic 'Analysis of Binomial Data with Extra Variation'. She also was a postdoctoral

researcher at the Middle East Technical University, Department of Statistics and her research focused on multimodal bivariate circular data analysis. Then she accepted a position at the Department of Statistics, University of Khansar.

**Zeynep Kalaylioglu** is currently Associate Professor of Statistics at the Department of Statistics, Middle East Technical University (METU), Ankara, Turkey. She earned her Ph.D. in Statistics at North Carolina State University, USA, in 2002, working with Prof. Sastry Pantula and Prof. Sujit K. Ghosh. In her thesis, she developed a Bayesian unit root test for stochastic volatility models. After completing her Ph.D., she accepted a position at Information Management Services where she worked for 5 years working with NIH/DCEG researchers in developing methodologies for the analysis of genetic associations, gene-environment interactions, etiological heterogeneity by cancer subtypes and risk prediction for important cancer types. In 2007, she accepted a position at the Department of Statistics, METU, where she has been since then. Her research focuses on developing Bayesian methodologies for health and biological data, in particular model selection and nonrandom missingness.