

ASSESSING FINGERPRINT INDIVIDUALITY USING EPIC: A CASE STUDY IN THE ANALYSIS OF SPATIALLY DEPENDENT MARKED PROCESSES

BY CHAE YOUNG LIM AND SARAT C. DASS

Michigan State University

Fingerprint individuality refers to the extent of uniqueness of fingerprints and is governed by the distribution of fingerprint features, termed minutiae, in a population. This paper develops a flexible class of marked point processes for minutiae and associated methodology for assessing fingerprint individuality. Inference is carried out in a Bayesian MCMC framework. The flexibility of the model fit to different kinds of minutiae patterns is demonstrated using real fingerprints. Evidence of a Paired Impostor Correspondence (EPIC) is developed as a measure of fingerprint individuality and its value is obtained using a simulation procedure based on the fitted models.

1. Introduction. The fingertip pattern of an individual is unique to that person. This is the central premise of fingerprint based authentication systems used for identifying individuals. In practice, however, various sources of variability can confound this uniqueness information and cause erroneous decisions to be made. A central problem in fingerprint analysis, therefore, is to determine the amount of information in a fingerprint and assess the extent of uniqueness. These problems can be addressed by eliciting statistical models that adequately capture the different sources of variability. This paper develops statistical models for fingerprint features, called minutiae, that exhibit certain distributional characteristics, such as clustering tendencies and spatial dependence. Inference methodology is developed for these models and is utilized for quantifying the extent of uniqueness in a pair of fingerprints.

1.1. The Application. A salient characteristic of fingerprint images is the smooth flow-like patterns with alternating dark and light lines, termed as ridges and valleys, over the entire fingerprint domain. Figure 1 illustrates this salient pattern for two fingerprint images in a publicly available collection called the NIST database. Occasionally, the ridges abruptly end or bifurcate, and these anomalous terminations and bifurcations are termed as minutiae. Minutiae information consists of the location of the ridge anomaly and the orientation, which is the direction of the ridge flow measured as an angle with respect to the horizontal axis at the minutiae location. Minutiae locations take values in the fingerprint spatial domain of $S \equiv [0, 512] \times [0, 512]$. Opposite

Keywords and phrases: Fingerprint analysis, Marked point processes, Mixture distribution, Reversible jump Markov Chain Monte Carlo, Wrapped distributions



FIG 1. Two fingerprints exhibiting different ridge flow patterns and minutiae clustering characteristics. The white squares and lines denote minutiae locations and orientations, respectively. Panels (a) and (b), respectively, correspond to Data 1 and 2 in Section 7.

ridge flows are equivalent, and therefore, the domain of minutiae orientations is $\Omega \equiv (0, \pi]$. The information of a minutiae is thus $s = (x, w)$ where x and w are respectively, the location and orientation information.

The fingerprint matching process entails determining the degree of similarity between a pair of fingerprint images. One such similarity measure is to determine the number of minutiae matches between the image pair. A minutiae match is defined by a pair of minutiae, $s_j = (x_j, w_j)$, $j = 1, 2$, one from the first image and the other from the second, which satisfies

$$|x_1 - x_2|_d \leq r_0 \text{ and } |w_1 - w_2|_a \leq w_0 \quad (1)$$

for some fixed thresholds r_0 and w_0 , where $|\cdot|_d$ is the Euclidean distance in R^2 and $|\cdot|_a$ is the angular distance $|w|_a = \min\{|w|, \pi - |w|\}$.

The goal here is to determine whether the fingerprint images (or, in short, prints) are from the same *finger* or different fingers, called as a genuine and impostor match, respectively. The differences in a true match pair will be largely removed once rotation and translation corrections are made (which are caused by possibly different placements of the finger onto the sensor or any acquisition platform for obtaining the print). Fingerprint based authentication proceeds by first extracting all minutiae information to generate a minutiae set for each print. One minutiae set is then affinely transformed (i.e., by rotation and translation) to best match the minutiae set of the other, in the sense that the highest number of minutiae matches is produced. A high (respectively, low) number of minutiae matches provides support for a genuine (respectively, impostor) match. Additionally, if different fingers correspond to different individuals, the matching process decides between a genuine and an impostor identity. This is the case when we know that the two prints are prints of the right index finger, for example.

The two prints in Figure 1 are used to demonstrate this procedure. A matching between them

yields the number of minutiae matches to be $\tau_0 = 13$ when $r_0 = 5$ and $w_0 = 18$ degrees (see (1)). The central question now is “can we decide that the prints are from the same finger or are they an impostor pair?” Is $\tau_0 = 13$ matches large enough to convince us that the prints come from the same finger? For any such decision rule, what are the associated error rates? This is precisely the problem of fingerprint individuality.

There have been numerous court cases that have highlighted the need for an assessment of individuality in forensic testimony. In the case of *Daubert v. Merrell Dow Pharmaceuticals* (1993), the U.S. Supreme Court ruled that in order for expert forensic testimony to be allowed in courts, it had to be subject to principles of scientific validation. Following *Daubert*, forensic evidence based on fingerprints was first challenged in the 1999 case of *U.S. v. Byron C. Mitchell*, stating that the fundamental premise for asserting the uniqueness of fingerprints had not been objectively tested and its potential matching error rates were unknown. Subsequently, fingerprint based identification has been challenged in more than 20 court cases in the United States. To address these concerns, several research investigations have proposed measures that characterize the extent of uniqueness of fingerprints (i.e., fingerprint individuality); see Pankanti et al. (2002), Zhu et al. (2007) and the references therein. The primary aim of these measures is to capture the inherent variability and uncertainty when an individual is identified based on fingerprint evidence.

There are a couple of salient characteristics that are typical of minutiae in fingerprint images. First, the number of minutiae in a fingerprint image is random since the feature extraction process depends on the effectiveness of the algorithm as well as the quality of the input image. Second, minutiae locations tend to show clustering tendencies (Scolve 1979; Stoney and Thornton 1986; Zhu et al. 2007), and third, the orientations of spatially neighboring minutiae show high correlation. Previous approaches have either not modeled the clustering tendencies or have assumed that the minutiae information is independent and identically distributed. These models for fingerprint individuality, therefore, do not capture the observed spatial dependence between neighboring minutiae orientations. By viewing the minutiae information as a marked point process in R^2 (with orientations as the marks), our aim is to develop novel point process models in R^2 with spatially dependent distributions for the marks, and subsequently, utilize them for assessing individuality. The inference for the spatially varying marked processes is carried out in a Bayesian MCMC framework. One particular aspect of our MCMC methodology is the incorporation of a dimension-changing Reversible Jump step to update the number of clusters of the spatial point pattern. A key point of our models is that they are flexible; Section 7 demonstrates that the proposed models fit well to fingerprint images with different minutiae characteristics, which in turn, give reliable estimates of fingerprint individuality.

We denote the minutiae sets extracted from two prints (such as in Figure 1) during the fingerprint authentication process by M_1 and M_2 in that order, and the number of minutiae matches between M_1 and M_2 by the random variable \mathcal{S} . We model the minutiae set M_j as a realization from a spatially dependent marked point process, f_j for $j = 1, 2$. To address the question of finger-

print individuality, our approach is to develop a model for the distribution of $\mathcal{S} = \mathcal{S}(M_1, M_2)$ for all pairs of impostor prints (that is, prints from different fingers) that have minutiae distributions f_1 and f_2 . Further, M_1 and M_2 are assumed to be generated independently of each other.

Our model is designed to reflect the spatial dependence in a minutiae set. However, a truly genuine pair of prints will have much stronger dependence than is captured by the model. Two prints from the same finger will have $f_1 = f_2$, but will behave like dependent (and not independent) realizations from this model. Therefore, the number of matches from two independent realizations will behave like a result from an impostor pair, even if $f_1 = f_2$. We can use our model to investigate the distribution of \mathcal{S} when faced with an impostor pair. In general, we expect that it will be concentrated on small values, so sufficiently large values are an indication that we have a true match.

The extent of fingerprint individuality is determined in terms of EPIC (Evidence of a Paired Impostor Correspondence) under the assumption that the prints are an impostor pair from the aforementioned population. Mathematically, EPIC is expressed as

$$EPIC(\tau_0) = P(\mathcal{S} \geq \tau_0) \quad (2)$$

with τ_0 denoting the observed number of matches. EPIC is a measure of fingerprint individuality wherein large values of EPIC indicate strong evidence for an impostor match and vice versa. For a truly impostor pair, τ_0 is typically small (but not exactly 0 due to some degree of random matching), usually falling within the variability bounds of \mathcal{S} . As a result, EPIC will be large, reflecting strong evidence for the impostor case. To illustrate this numerically, consider the fingerprints in Figure 1 which have a total of 76 and 59 minutiae for panels (a) and (b), respectively; these are prints f0001 and f0002 from the NIST database reported in Table 7. Figure 2 (a) gives the distribution of \mathcal{S} based on 5,000 realizations; the methodological details on how this is obtained are given in Section 5. Note that \mathcal{S} takes on small values ranging mostly around 10 matches. The value of $\tau_0 = 13$ is within the variability bounds of \mathcal{S} , thus giving high evidence for an impostor pair based on EPIC.

Next, we derive the EPIC value for a truly genuine pair. Figure 2 (b) gives the histogram based on 5,000 realizations of \mathcal{S} for a genuine pair of fingerprints, f0002 and s0002, from the NIST database in Table 7. Note that although f0002 and s0002 are a genuine pair, values of \mathcal{S} are typically small, ranging between 6 to 16 matches. The histogram in Figure 2 (b) reflects the variability of \mathcal{S} when a pair of impostor prints is randomly selected from the population with $f_1 = f_2 = f$. However, the observed number of matches $\tau_0 = 22$ in this case (see Table 7) is far from the variability bounds of \mathcal{S} . Thus, EPIC for a truly genuine pair will be small since most of the probability mass in (2) is concentrated on small \mathcal{S} values and τ_0 is usually a relatively larger number.

Compared to the probability of a random correspondence (PRC), an alternative fingerprint individuality measure (Zhu et al. 2007; Dass and Li 2009) which compares two arbitrary prints

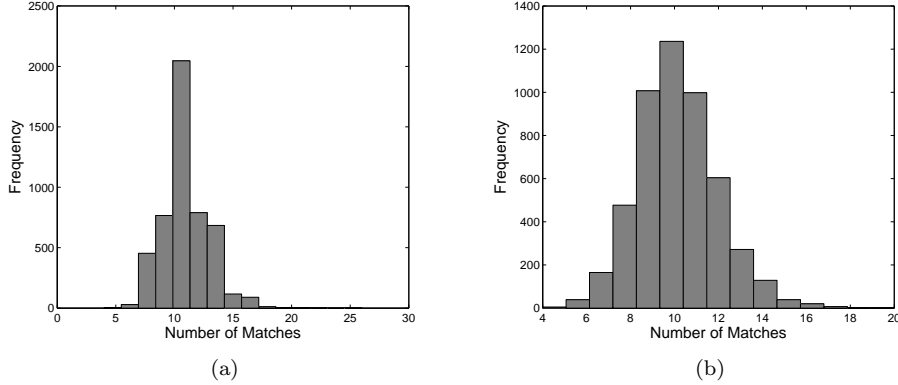


FIG 2. The distribution \mathcal{S} for impostor pairs when independent minutiae sets $M_j \sim f_j$ for $j = 1, 2$; panel (a) (respectively, (b)) corresponds to a truly impostor (respectively, genuine) pair with $f_1 \neq f_2$ (respectively, $f_1 = f_2$).

randomly selected from a super-population, EPIC performs a conditional analysis by restricting the population to only fingerprints having the same minutiae distributions as the two prints in question.

The rest of the paper is organized as follows: Section 2 discusses the general theory of marked processes and develops models that incorporate clustering and spatial dependence in these processes. Sections 3 and 4 describe specific marked spatial point process models for the analysis of fingerprints. Although our discussion is based on processes in R^2 , the discussion can be generalized to R^q , $q > 2$ in a straightforward manner. A fully Bayesian inferential framework is developed for fitting the models to the observed data in Section 5. Section 6 gives the simulation results for two special cases of our models. The two fingerprint images in Figure 1 are analyzed as an illustration of the proposed methodology in Section 7 together with other examples. Details of the computational procedure as well as other tables and plots are available online as supplementary material.

2. Spatial Point Processes. We consider only finite point processes subsequently. A finite spatial point process \mathbf{X} is a (finite) random subset of a bounded region $S \subseteq R^2$, with realizations given by the set of observations $\mathbf{x}_n \equiv \{x_1, x_2, \dots, x_n\}$ of $n \geq 0$ points contained in S . It follows that the realizations \mathbf{x}_n belong to the space $\cup_{n=0}^{\infty} S^n$ where S^n is the n -fold product space of S . The distribution of a finite spatial point process can be characterized by the following hierarchical

specification of distributions (Daley and Vere-Jones 1988):

- (i) A distribution $\{p_n\}$ for $n \geq 0$ determining the total number of points in \mathbf{X} with $\sum_{n=0}^{\infty} p_n = 1$, and
 - (ii) for each integer $n \geq 1$, a probability distribution $\Pi_n(\cdot)$ on S^n that determines the joint distribution of the points given their total number, n .
- (3)

The density (or, likelihood) of a spatial point process is given by the expression $j_n(\mathbf{x}_n)$ satisfying

$$j_n(\mathbf{x}_n) d\mathbf{x}_n = P\{\text{there are exactly } n \text{ points in the process, one in each of the distinct infinitesimal balls centered at } x_i, B_{x_i}(dx_i)\}. \quad (4)$$

The term $j_n(\mathbf{x}_n)$ corresponds to the density of the Janossy measure for spatial point processes (Daley and Vere-Jones 1988). If $f_n(x_1, \dots, x_n)$ is the symmetric density function corresponding to Π_n , the density of \mathbf{X} is given by

$$j_n(\mathbf{x}_n) = p_n f_n(x_1, \dots, x_n). \quad (5)$$

Note that the densities defined in (4) and (5) are valid only for *finite* point processes as is the case here.

A marked spatial point process is an extension of a spatial point process in which an additional random variable w_x (called a mark) is observed at each point x of \mathbf{x}_n . Typically, w_x represents some additional attribute(s) measured at the point x with a possibly different domain, $\Omega \subset R^p$ (Stoyan & Stoyan 1998; Møller and Waagepetersen 2003). Throughout this paper, a generic marked spatial point process will be denoted by $(\mathbf{x}_n, \mathbf{w}_{\mathbf{x}_n}) \equiv \{(x, w_x) : x \in \mathbf{x}_n\}$. The distribution of a marked spatial point process can be characterized by adding a third level to the hierarchical specification of (3), namely,

- (iii) Given n and \mathbf{x}_n , the marks $\mathbf{w}_{\mathbf{x}_n}$ have a distribution $\Pi_n^w(\cdot)$ with density g_n .
- (6)

It follows from the hierarchical specification of (i)-(iii) that the marked spatial point process $(\mathbf{x}_n, \mathbf{w}_{\mathbf{x}_n})$ has a density given by

$$j_n(\mathbf{x}_n, \mathbf{w}_{\mathbf{x}_n}) = p_n f_n(x_1, \dots, x_n) g_n(\mathbf{w}_{\mathbf{x}_n}), \quad (7)$$

where the density in (7) is a natural extension of the density of the Janossy measure in (4) to the marked spatial process $(\mathbf{x}_n, \mathbf{w}_{\mathbf{x}_n})$.

An important special case of the general formulation is the Poisson point process on R^2 defined in terms of a non-negative intensity measure $\lambda : S \rightarrow (0, \infty)$. In this case,

$$p_n = P(N = n) = \exp(-\lambda_0) \lambda_0^n / n! \quad (8)$$

where N follows a Poisson distribution with mean $\lambda_0 = \int_S \lambda(s) ds$, and the density x_1, x_2, \dots, x_n given n are iid with

$$f_n(x_1, x_2, \dots, x_n) = \prod_{i=1}^n (\lambda(x_i)/\lambda_0). \quad (9)$$

A marked Poisson point process additionally involves a third level elicitation involving the distribution of marks $\mathbf{w}_{\mathbf{x}_n}$ given n and \mathbf{x}_n , $g_n(\mathbf{w}_{\mathbf{x}_n})$ say, where g_n is an n -dimensional joint density on the n -fold space of the marks. Subsequently, a marked Poisson point process will be denoted by $\mathcal{P}(\lambda, g_n)$ with parameters λ and g_n as described above.

Our application to fingerprint images in Section 7 requires the development of marked spatial point processes with dependent marks; that is, marks corresponding to realized random points on S that are spatially close should exhibit a high degree of correlation. One way of incorporating such spatial dependence is to consider general marked spatial point processes whose densities are given by (7). Clustering tendencies by such processes can be ensured by requiring p_n , $f_n(\mathbf{x}_n)$ and $g_n(\mathbf{w}_{\mathbf{x}_n})$ to be functions with respect to some mixing distribution, Φ , say. The details are given in the subsequent sections.

3. The Multivariate Normal and Wrapped Normal Distributions. Before defining the marked processes on minutiae, we first need the following two specific classes of multivariate distributions.

The d -variate normal density with mean $\boldsymbol{\mu} \in R^d$ and covariance matrix $\Sigma \in R^d \times R^d$ is given by

$$\phi_d(\mathbf{x} | \boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{d/2} \det(\Sigma)^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \quad (10)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_d)'$, and $\det(\Sigma)$ is the determinant of the covariance matrix Σ . In the special case when $d = 2$ and Σ is diagonal, say $\Sigma = \text{diag}(\sigma_1^2, \sigma_2^2)$, we denote the corresponding density in (10) by

$$\phi_d(\mathbf{x} | \boldsymbol{\mu}, \Sigma) \equiv \phi_2(\mathbf{x} | \boldsymbol{\mu}, \sigma_1^2, \sigma_2^2). \quad (11)$$

The d -variate wrapped normal distribution on $(0, 2\pi/L]^d$ is derived from a regular multivariate normal distribution by wrapping the latter onto the set $(0, 2\pi/L]^d$. If $\mathbf{X} = (X_1, X_2, \dots, X_d)'$ follows (10), the vector $\mathbf{W} = (W_1, W_2, \dots, W_d)'$ defined by

$$W_j = X_j \bmod \left(\frac{2\pi}{L} \right) \quad \text{for } j = 1, 2, \dots, d, \quad (12)$$

is said to have a multivariate wrapped normal distribution with parameters $\boldsymbol{\mu}$ and Σ . The density corresponding to a d -variate wrapped normal density is given by

$$V_{d,L}(\mathbf{w} | \boldsymbol{\mu}, \Sigma) = \sum_{t_1=-\infty}^{\infty} \sum_{t_2=-\infty}^{\infty} \cdots \sum_{t_d=-\infty}^{\infty} \phi_d \left(\mathbf{w} + \left(\frac{2\pi}{L} \right) \mathbf{t} \mid \boldsymbol{\mu}, \Sigma \right) \quad (13)$$

with mean angle $\boldsymbol{\mu} \in (0, 2\pi/L]^d$, $\boldsymbol{w} = (w_1, w_2, \dots, w_d)' \in (0, 2\pi/L]^d$ and $\boldsymbol{t} = (t_1, t_2, \dots, t_d)' \in \mathbb{Z}^d$, \mathbb{Z} is the set of integers.

Univariate wrapped distributions have been extensively used as models for angular random variables (Fisher 1993; Fisher and Lee 1994; Mardia and Jupp 1999). Developing bivariate and multivariate distributions for angular random variables is still an active field of research; see, for example, Singh et al. (2002), Shieh and Johnson (2005) and Mardia et al. (2007). The multivariate extension of wrapped distributions we propose here can easily incorporate dependence between the W_{j_s} via the covariance matrix $\boldsymbol{\Sigma}$ and be analyzed using a Bayesian computational scheme.

Parametric inference for the wrapped distributions is generally carried out via a data augmentation technique (Breckling 1989; Coles 1998; Ravindran 2002). Essentially, the “missing” information in (13) is the wrapping numbers \boldsymbol{t} . By augmenting \boldsymbol{t} to \boldsymbol{w} so that the complete data becomes $(\boldsymbol{w}, \boldsymbol{t})$, the joint density of the complete data becomes

$$V(\boldsymbol{w}, \boldsymbol{t} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) \equiv \phi_d\left(\boldsymbol{w} + \left(\frac{2\pi}{L}\right)\boldsymbol{t} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}\right). \quad (14)$$

Note that the wrapped density $V_{d,L}$ in (13) is the marginal of V in (14) after summing out \boldsymbol{t} .

4. Spatial Point Processes with Dependent Marks for Fingerprint Minutiae. Let $\boldsymbol{x}_n \equiv \{x_i, i = 1, 2, \dots, n\}$ denote the collection of n minutiae locations, and for each $x \in \boldsymbol{x}_n$, the minutiae orientation w_x denotes the corresponding mark, which takes values in $(0, \pi]$ (i.e., $L = 2$). The distribution of minutiae in a fingerprint image is best described in terms of a hierarchical model involving all random entities. Recall that $\mathcal{P}(\lambda, h)$ denotes a marked Poisson process with λ and h , respectively, denoting the intensity measure and joint density function of marks. The hierarchical model for $(\boldsymbol{x}_n, \boldsymbol{w}_{\boldsymbol{x}_n})$ is given by

$$(\boldsymbol{\theta}, \boldsymbol{m}_{\boldsymbol{\theta}}) \equiv \Phi \sim \mathcal{P}(\lambda_1, h_1), \quad (15)$$

$$(\boldsymbol{x}^{(k)}, \boldsymbol{w}_{\boldsymbol{x}^{(k)}}) | \Phi \stackrel{ind}{\sim} \mathcal{P}(\lambda_{2k}, g_k), \text{ for } k = 1, 2, \dots, K, \text{ and} \quad (16)$$

$$(\boldsymbol{x}_n, \boldsymbol{w}_{\boldsymbol{x}_n}) = \bigcup_{k=1}^K (\boldsymbol{x}^{(k)}, \boldsymbol{w}_{\boldsymbol{x}^{(k)}}), \quad (17)$$

where K is the number of elements in $\boldsymbol{\theta}$; in the above formulation, the intensity measure for $\boldsymbol{\theta}$ is the constant function

$$\lambda_1(s) = \begin{cases} K_0/\text{area}(S_0) & \text{if } s \in S_0, \\ 0, & \text{otherwise,} \end{cases} \quad (18)$$

with S_0 and K_0 , respectively, denoting a bounded rectangular region of S and a fixed (and known) positive real number. The model (15-17) entails that the observations $(\boldsymbol{x}_n, \boldsymbol{w}_{\boldsymbol{x}_n})$ are the total collection (i.e., union) of realizations from K component marked Poisson processes, and it is not known which of the K components gave rise to each observation in the collection. An equivalent formulation of the model, which is perhaps more straightforward to understand, is

given in (25-30) and used to carry out the computational steps for Bayesian inference. In our application, the fingerprint sensing domains are rectangular and this is the reason for our choice of bounded S_0 . The mark corresponding to θ is $m_\theta \equiv (\gamma, \sigma_1^2, \sigma_2^2, \eta, \rho, \delta^2)$ with density h_1 defined by its component densities

$$\gamma \sim G(\alpha_\gamma, \beta_\gamma), \quad \sigma_1^2 \sim \text{IG}(\alpha_1, \beta_1), \quad \sigma_2^2 \sim \text{IG}(\alpha_2, \beta_2), \quad (19)$$

$$\eta \sim U(0, 2\pi/L), \quad \rho \sim U(\rho_{\min}, \rho_{\max}) \quad \text{and} \quad \delta^2 \sim \text{IG}(\alpha_\delta, \beta_\delta), \quad (20)$$

independently of each other; in (19) and (20), $G(\alpha, \beta)$ and $\text{IG}(\alpha, \beta)$ are, respectively, the Gamma and inverse Gamma distributions with shape and scale parameters given by α and β , and $U(a, b)$ is the uniform distribution from a to b . It is clear from the above specification that the γ and σ_j^2 , $j = 1, 2$, components of m_θ should be positive: $\gamma > 0$, $\sigma_j^2 > 0$ for $j = 1, 2$. The rest of the components are required to satisfy $\eta \in (0, 2\pi/L]$, $\rho \in (\rho_{\min}, \rho_{\max})$, and $\delta^2 > 0$. The intensity measure λ_{2k} for $\mathbf{x}^{(k)}$ in (16) is

$$\lambda_{2k}(s) = \begin{cases} \gamma_k \phi_2(s | \theta_k, \sigma_{1k}^2, \sigma_{2k}^2) & \text{if } s \in S_0, \\ 0, & \text{otherwise,} \end{cases} \quad (21)$$

with g_k (conditional on $\mathbf{x}^{(k)}$) given by

$$g_k(\mathbf{w} | \eta_k, \rho_k, \delta_k^2) = V_{n_k, L} \left(\mathbf{w} | \eta_k, \sum_k (\rho_k, \delta_k^2) \right) \quad (22)$$

where n_k is the number of elements in $\mathbf{x}^{(k)}$, $V_{n_k, L}$ is the n_k -variate wrapped normal distribution given in (13), with mean $\eta_k = (\eta, \eta, \dots, \eta)' \in R^{n_k}$ for some fixed but unknown constant $\eta \in (0, 2\pi/L]$, and covariance matrix $\sum_k (\rho_k, \delta_k^2)$. Given $\mathbf{x}^{(k)}$, the marks $\mathbf{w}_{\mathbf{x}^{(k)}}^{(k)}$ can be thought to be observations from a geo-spatial model where spatial dependencies are modeled via a covariance function, q ; if W_x and W_y are the observable marks at locations x and y , respectively, we have

$$\text{Cov}(W_x, W_y) = q(x, y; \rho, \delta^2) \quad (23)$$

where $x, y \in R^2$, and ρ and δ^2 are fixed but unknown parameters that govern the spatial correlation and variance of the marks, respectively. The exponential covariance function, for example, yields the covariance matrix $\sum_k \equiv \sum_k (\rho, \delta) = (\sigma_{rs}^*)$, $r, s = 1, 2, \dots, n_k$ with entries given by

$$\sigma_{rs}^* = \delta^2 \exp(-\rho \|x_r - x_s\|), \quad (24)$$

where x_r and x_s are a pair of points from $\mathbf{x}^{(k)}$ and $\|\cdot\|$ is the R^2 -Euclidean norm. Low values of ρ (close to 0) yield high spatial dependence between neighboring marks w_{x_r} and w_{x_s} , which is the scenario we observe with fingerprint minutiae.

An alternative but equivalent formulation of the hierarchical model (15)-(17) can be described with some additional notation. Denote the class label set $\mathbf{c}_n \equiv \{c_i, i = 1, 2, \dots, n\}$ with c_i ,

corresponding to $x_i \in \mathbf{x}_n$, taking values in the set $\{1, 2, \dots, K\}$. Also, let $B_k \equiv \{x_i : c_i = k\}$, for $k = 1, 2, \dots, K$ denote a partition of \mathbf{x}_n which induces a partition of $(\mathbf{x}_n, \mathbf{w}_{\mathbf{x}_n})$ into K sets, given by $(\mathbf{x}_{B_k}, \mathbf{w}_{B_k})$ for $k = 1, 2, \dots, K$. The equivalent formulation of (15-17) is

$$(\boldsymbol{\theta}, \mathbf{m}_{\boldsymbol{\theta}}) \equiv \Phi \sim \mathcal{P}(\lambda_1, h_1), \quad (25)$$

$$n | \Phi \sim p_n = \exp(-T) \frac{T^n}{n!}, \quad (26)$$

$$\mathbf{c}_n | n, \Phi \sim \prod_{i=1}^n \left[\frac{\gamma_{c_i} D(\theta_{c_i}, \sigma_{1c_i}^2, \sigma_{2c_i}^2)}{T} \right], \quad (27)$$

$$\mathbf{x}_n | \mathbf{c}_n, n, \Phi \sim \prod_{i=1}^n \left[\frac{\phi_2(x_i | \theta_{c_i}, \sigma_{1c_i}^2, \sigma_{2c_i}^2)}{D(\theta_{c_i}, \sigma_{1c_i}^2, \sigma_{2c_i}^2)} \right], \text{ and} \quad (28)$$

$$(\mathbf{w}_{\mathbf{x}_n} | \mathbf{x}_n, \mathbf{c}_n, n, \Phi) \sim \prod_{k=1}^K g_k(\mathbf{w}_{B_k} | \eta_k, \rho_k, \delta_k^2), \quad (29)$$

where T is defined as

$$T \equiv \sum_{k=1}^K \gamma_k D(\theta_k, \sigma_{1k}^2, \sigma_{2k}^2), \quad (30)$$

with

$$D(\theta_k, \sigma_{1k}^2, \sigma_{2k}^2) \equiv \int_{S_0} \phi_2(s | \theta_k, \sigma_{1k}^2, \sigma_{2k}^2) ds,$$

and the density g_k is as given in (22). The data augmentation technique in (14) gives rise to the augmented density

$$(\mathbf{w}_{\mathbf{x}_n}, \mathbf{t}_{\mathbf{x}_n} | \mathbf{x}_n, \mathbf{c}_n, n, \Phi) \sim \prod_{k=1}^K \phi_{n_k} \left(\mathbf{w}_{B_k} + \pi \mathbf{t}_{B_k} \middle| \eta_k, \sum_k (\rho_k, \delta_k^2) \right) \quad (31)$$

where $\mathbf{t}_{B_k} = \{t_x : x \in B_k\}$.

This hierarchical model entails that each c_i is distributed independently according to the common distribution in (27), and conditional on c_i , each x_i in (28) is distributed independently as a bivariate normal density $\phi_2(\cdot | \theta_{c_i}, \sigma_{1c_i}^2, \sigma_{2c_i}^2)$ restricted to the domain S_0 . Independent normal kernels allow for two separate one dimensional integrations when evaluating $D(\theta_k, \sigma_{1k}^2, \sigma_{2k}^2)$. However, one can still perform the integration on a two dimensional space with a general covariance structure with marginal increase in computational complexity. The reason we choose diagonal covariance matrices is because earlier studies conducted found strong evidence for this model in the case of fingerprint images (Dass and Li 2009).

The two-Poisson process specification above seems an unnecessary and complicated way to develop a single Poisson process for *locations*, with non-homogeneous mixture intensity and correlation between marks governed by mixture component membership. Indeed, the model can be recast in this way but the aim here is to make the mechanism for inducing correlation between

marks more apparent. Note also that the resulting marginal process for the spatial locations is a Cox process. Similar modeling approaches for the marginal point process have been reported in the literature. Wolpert and Ickstadt (1998) and Kottas and Sanso (2007) present a flexible class of non-homogeneous Poisson process models that make use of mixture specifications with an unknown number of components. In another paper, Ickstadt and Wolpert (1999) develop models with directional marks (daily commutes to work). We note that the works above utilize nonparametric stochastic process priors as opposed to the fully parametric approach developed here.

5. A Bayesian Framework for Statistical Inference. We develop a fully Bayesian inferential framework for the model described in Section 4 based on the observed data $(n, \mathbf{x}_n, \mathbf{w}_{\mathbf{x}_n})$. The hierarchical model specification of (26)-(28) and (31) yields the complete (or, augmented) likelihood

$$\begin{aligned} \ell(n, \mathbf{c}_n, \mathbf{x}_n, \mathbf{w}_{\mathbf{x}_n}, \mathbf{t}_{\mathbf{x}_n} | \Phi) &= p_n \times \ell_0(\mathbf{c}_n | n, \Phi) \times \ell_0(\mathbf{x}_n | \mathbf{c}_n, n, \Phi) \\ &\times \ell_0(\mathbf{w}_{\mathbf{x}_n}, \mathbf{t}_{\mathbf{x}_n} | \mathbf{x}_n, \mathbf{c}_n, n, \Phi) \end{aligned} \quad (32)$$

where $\ell_0(A, B, \dots | C, D, \dots)$ denotes the density of random variables A, B, \dots conditional on C, D, \dots given in (26)-(28) and (31). In (32), $\Phi = \{(\theta_k, \gamma_k, \sigma_{1k}, \sigma_{2k}, \eta_k, \rho_k, \delta_k), k = 1, 2, \dots, K\}$ denotes the collection of all unknown parameters: K denotes the number of clusters with cluster k having (i) spatial mean θ_k , (ii) spatial variances σ_{jk}^2 for $j = 1, 2$, (iii) mean of marks η_k , (iv) covariances between marks governed by the correlation and variance parameters ρ_k and δ_k^2 , respectively, and (v) γ_k denoting the intensity of the k -th cluster which determine the expected total number of points in cluster k . For implementing a Bayesian framework of inference, the prior we adopt on Φ is \mathcal{P} as given in (15). The hyper-parameters in (19) and (20) will be assumed to be fixed and known.

5.1. Posterior Inference. Posterior inference for the likelihood of the hierarchical model in (32) is carried out based on a Markov Chain Monte Carlo (MCMC) algorithm. The MCMC updating steps are (1) update K , and for fixed K ,

- (2) update $(\theta_k, \sigma_{1k}^2, \sigma_{2k}^2, \eta_k, \rho_k, \delta_k), k = 1, 2, \dots, K$,
- (3) update \mathbf{c}_n , and
- (4) update $\mathbf{t}_{\mathbf{x}_n}$.

Out of the four updating steps above, only the update of (1) involves parameter spaces of varying dimensions (the remaining updating steps (2-4) are regular Metropolis-Hastings (MH) steps). To obtain posterior inference for such a space of models, Green (1995) and Green and Richardson (1997) developed the Reversible Jump Markov Chain Monte Carlo (RJMCMC) approach for the Bayesian inferential framework. Since its introduction, RJMCMC has been successfully used

to obtain posterior inference in many situations. We develop a Reversible Jump Markov Chain Monte Carlo approach to explore the posterior distribution in updating step (1). The updating steps are explained in detail in the Appendix of the supplementary material.

5.2. Convergence Diagnostics. The assessment of convergence of the RJMCMC is carried out based on the methodology of Brooks and Guidici (1999, 2000). A total of $I \geq 2$ chains are run from different starting values representing overdispersed states, and quantities that maintain the same interpretation across different models are monitored. We recommend running at least $I = 3$ chains to ensure correctness of the convergence diagnostics. The diagnostics for assessing convergence utilize the following six quantities: the overall variance, \hat{V} , the within chain variance, W_c , within model variance W_m , within model within chain variance $W_m W_c$, the between model variance, B_m and the between model within chain variance, $B_m W_c$. For each monitoring parameter three figures are obtained: (i) \hat{V} and W_c , (ii) W_m and $W_m W_c$, and (iii) B_m and $B_m W_c$ versus the number of iterations. The two plotted lines in each figure should be close to each other to indicate sufficient mixing. Our choice of the monitoring parameter is the logarithm of the complete likelihood (32), namely, $\log \ell(n, \mathbf{c}_n, \mathbf{x}_n, \mathbf{w}_{\mathbf{x}_n}, \mathbf{t}_{\mathbf{x}_n} | \Phi)$, based on the hierarchical model specified by (26)-(28) and (31). Note that our choice of the monitoring parameter was motivated by the difficulty of evaluating the likelihood corresponding to only the observed data $(n, \mathbf{x}_n, \mathbf{w}_{\mathbf{x}_n})$ due to high dimensional summations involved to integrate out \mathbf{c}_n and $\mathbf{t}_{\mathbf{x}_n}$ from the complete likelihood. Furthermore, the complete likelihood is a function of the parameters that maintains a consistent interpretation across different models for varying K , as required by Brooks and Guidici (1999, 2000).

5.3. Inference based on EPIC. Since it is not known whether the pair of prints is genuine or impostor, two marked processes are fitted separately to each print in the pair. After convergence is established, we generate N_0 independent samples from each of the fitted marked processes (and independently of each other). Each sample, M_j^l , constitutes a minutiae set from the j -th marked process, for $l = 1, 2, \dots, N_0$ and $j = 1, 2$. The M_j^l , $j = 1, 2$, samples are then paired with each other to yield the N_0 pairs (M_1^l, M_2^l) , $l = 1, 2, \dots, N_0$. The number of minutiae matches $S_l = \mathcal{S}(M_1^l, M_2^l)$ is obtained by applying the minutiae matcher described in the Introduction to each pair for $l = 1, 2, \dots, N_0$. The empirical estimate of EPIC is given by the proportion of pairs that have τ_0 or more matches out of N_0 , mathematically expressed as $EPIC(\tau_0) = (1/N_0) \sum_{l=1}^{N_0} I\{S_l \geq \tau_0\}$, where $I\{A\}$ is 1 if A is true, and 0 otherwise. A high value of EPIC implies that there is weak evidence to reject the hypothesis that the prints form an impostor pair. This form of analysis is, thus, similar to obtaining estimates of posterior predictive p-values based on a sampling procedure.

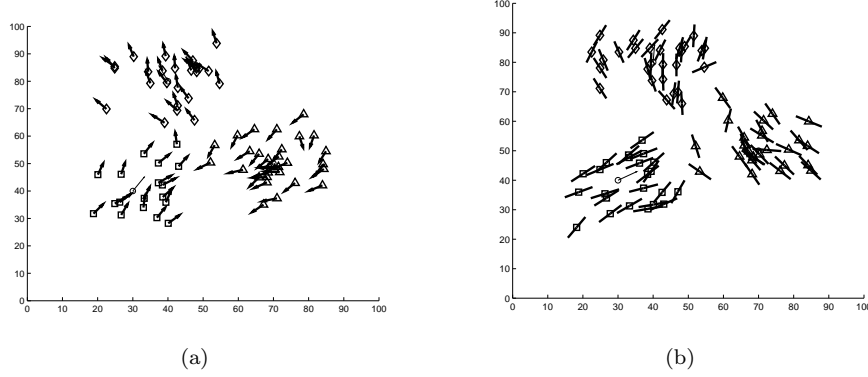


FIG 3. Simulated point patterns with directional marks attached as lines. Panel (a) is for $L = 1$. Panel (b) is for $L = 2$. The labels \square , \diamond and \triangle denote the different clusters.

6. Simulation. Two simulation experiments were carried out for the cases $L = 1$ and $L = 2$, corresponding to the mark spaces of $(0, 2\pi]$ and $(0, \pi]$, respectively. We consider the spatial domain $S_0 = [0, 100] \times [0, 100]$. Hyper-parameters in (19) and (20) are set as follows: α_γ and β_γ are derived from $E(\gamma) = 15$ and $Var(\gamma) = 100$ by solving two equations in the two unknowns. Similarly we set α_j and β_j so that $E(\sigma_j^2) = 75$ and $Var(\sigma_j^2) = 10,000$ for $j = 1, 2$. We set α_δ and β_δ so that $E(\delta^2) = 0.1$ and $Var(\delta^2) = 1$. Also, we set $K_0 = 3$, $(\rho_{\min}, \rho_{\max}) = (0.01, 5)$, $K_{\min} = 2$ and $K_{\max} = 5$. We took the probabilities of selecting move types to be $r_m = r_{m'} = 0.5$ corresponding to the moves $(m, m') = (K\text{-split}, K\text{-merge})$ for $K = K_{\min} + 1, \dots, K_{\max} - 1$. Also, when $K = K_{\min}$, $r_m = 1 = 1 - r_{m'}$ and $r_m = 0 = 1 - r_{m'}$ for $K = K_{\max}$. Given $K = 3$, two sets of point patterns with angular marks were simulated with $L = 1$ and $L = 2$, respectively. Plots of simulated data are given in Figure 3. We monitor convergence of $I = 5$ chains with starting values that represent *over-dispersion* in the chains. For both $L = 1$ and $L = 2$, the RJMCMC converged after 40,000 iterations. For example, see Figure 4. Tables 1 gives the associated statistical inference for the unknown parameters for $L = 1$; the true values are compared with the posterior means as well as the 99% credible intervals based on the last 1,000 values from each of the 5 chains. The estimated probability for $K = 3$ (the true value for the unknown K) is 0.918 and 0.610 for $L = 1$ and $L = 2$, respectively. For $L = 2$, the second highest value of K is 4 with posterior probability 0.262. The high posterior probability of $K = 4$ can be attributed to the splitting of the second cluster with true spatial mean $(\theta_{12}, \theta_{22}) = (40, 80)$ (i.e., topmost cluster in Figure 3 (b)). The RJMCMC favors splitting this cluster into two groups (left and right) with different values for the average orientations given by 1.23 and 1.57 radians, respectively. The computational time taken for every 10,000 iterations is around 1.15 hours on a Dell Latitude E6500 laptop with Intel Core 2 Duo processor.

From panel (a) of Figure 4, it seems possible that the predictive distribution based on param-

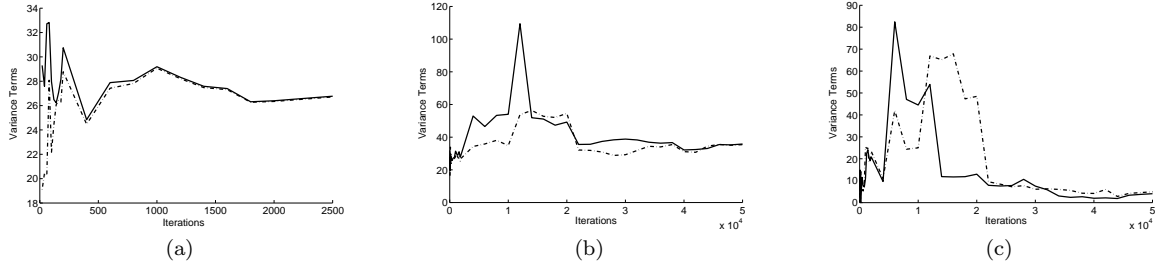


FIG 4. Convergence diagnostics for $L = 1$. Panels (a), (b) and (c) show the plots of (\hat{V}, W_c) , $(W_c, W_m W_c)$ and $(B_m, B_m W_c)$, respectively. A solid line represents the first entry and a dotted line represents the second entry. The x -axis in Panel (a) is taken from 0 to 2,500 to show the initial over-dispersion of the chains. The two lines in Panel (a) coincide all the way up to 50,000 iterations.

$L = 1$									
Parameter	True	Mean	Sd	CI	Parameter	True	Mean	Sd	CI
θ_{11}	30	32.63	1.91	[27.35, 37.88]	θ_{21}	40	39.48	1.93	[34.49, 44.53]
θ_{12}	40	41.28	1.82	[36.29, 45.92]	θ_{22}	80	80.13	1.97	[75.27, 85.47]
θ_{13}	70	70.78	1.60	[66.60, 75.20]	θ_{23}	50	51.00	1.47	[47.14, 54.84]
σ_{11}^2	70	58.28	21.02	[26.96, 142.75]	σ_{21}^2	70	63.09	22.12	[29.28, 153.60]
σ_{12}^2	70	82.34	23.72	[39.14, 171.06]	σ_{22}^2	70	82.94	28.27	[36.59, 207.46]
σ_{13}^2	70	73.48	19.23	[38.27, 146.75]	σ_{23}^2	70	61.68	16.03	[33.07, 123.39]
η_1	1	1.00	0.10	[0.69, 1.32]	δ_1	0.16	0.08	0.03	[0.03, 0.21]
η_2	2	2.01	0.08	[1.77, 2.24]	δ_2	0.16	0.09	0.03	[0.04, 0.23]
η_3	4	3.93	0.15	[3.44, 4.38]	δ_3	0.16	0.17	0.05	[0.08, 0.43]
ρ_1	0.15	0.60	0.76	[0.07, 4.58]	γ_1	21	16.98	3.88	[8.39, 28.30]
ρ_2	0.15	0.91	0.93	[0.09, 4.72]	γ_2	21	23.84	4.49	[13.78, 36.82]
ρ_3	0.15	0.23	0.13	[0.06, 0.72]	γ_3	28	27.14	4.86	[15.90, 41.28]

TABLE 1

The results of posterior inference for $L = 1$ based on simulated data.

eter realizations from the 5 chains has converged much faster than the methodology proposed by Brooks and Guidici (1999, 2000). Two measures that determine the predictive characteristics are considered here. For each iteration of the chains, one realization of the marked spatial point process, $\mathcal{X}^* \equiv (n, \mathbf{x}_n, \mathbf{w}_{\mathbf{x}_n})$ is generated according to (27) - (29). The bounded domain S_0 is subdivided into a grid of 4 rectangular regions, and the following two quantities (1) the proportion of points in \mathcal{X}^* falling in each subregion and (2) the average of the marks within each subregion are obtained. Figure 5 shows 3 sample plots of \hat{V} versus W_c for $L = 1$ based on the 5 chains. It is clear that convergence takes place more rapidly, that is, by 2,500 iterations for $L = 1$. Table 2 gives the 99% credible intervals for (1) and (2) corresponding to the 4 subregions. The values obtained from observed data fall inside this credible interval for all the cases considered in Table 2, indicating that the predictive capabilities of the model have already converged. The corresponding plots and table for $L = 2$ are available in the supplementary material.

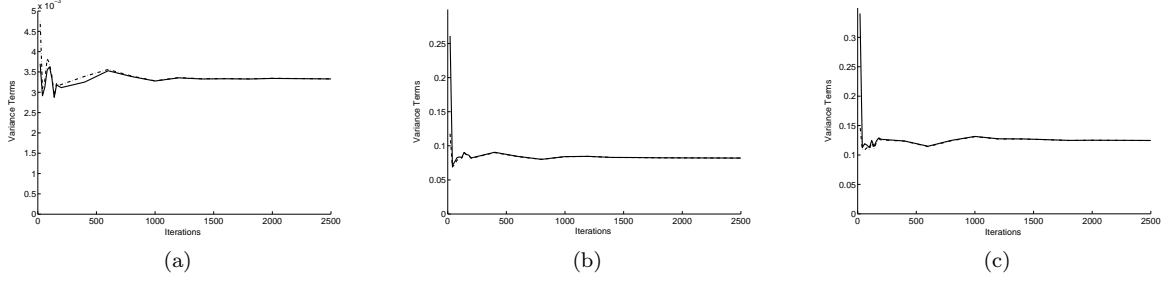


FIG 5. Convergence diagnostics based on predictive characteristics for $L = 1$. Panels (a), (b) and (c) show the plot of (\hat{V}, W_c) , for the proportion of points in region 3 and average marks for regions 3 and 4, respectively. A solid line represents \hat{V} and a dotted line represents W_c . The x-axis is taken from 0 to 2,500 to show the initial over-dispersed state. The two lines in each figure coincide all the way up to 50,000 iterations.

$L = 1$			$L = 2$		
Prop. of points	CI	Obs. value	Prop. of points	CI	Obs. Value
Region 1	[0.0814, 0.3750]	0.2113	Region 1	[0.1224, 0.4180]	0.2857
Region 2	[0.1542, 0.4928]	0.3380	Region 2	[0.1752, 0.4938]	0.3247
Region 3	[0.0506, 0.3529]	0.1972	Region 3	[0.0360, 0.3000]	0.1558
Region 4	[0.1222, 0.4667]	0.2535	Region 4	[0.1006, 0.4387]	0.2338
Average mark	CI	Obs. value	Average mark	CI	Obs. Value
Region 1	[0.4446, 1.5443]	0.9769	Region 1	[0.2356, 1.2810]	0.6237
Region 2	[1.5295, 2.4014]	1.9496	Region 2	[0.3507, 2.4771]	1.4646
Region 3	[2.9807, 4.7490]	3.5653	Region 3	[1.4979, 2.8632]	2.2827
Region 4	[2.6273, 4.5219]	3.8965	Region 4	[1.1694, 2.8522]	2.0807

TABLE 2

Measure of predictive capabilities of the RJMCMC by comparing the 99% equal tail credible interval with the value obtained from observed data.

7. Fingerprint Analysis. Two fingerprint images in the NIST database are chosen to illustrate our methodology (see Figure 1 panels (a) and (b)). The NIST fingerprint database is publicly available and consists of 2,000 8-bit gray scale fingerprint image pairs of size 512-by-512 pixels. The first image of each of the 2,000 fingers in the database is labeled by ‘f’ while the second image is labeled by ‘s’. Thus, the label ‘f0280’, for example, is the first impression of finger 280 in this database. Data 1 (‘f0001’) and Data 2 (‘f0002’) contain minutiae locations and orientations of the respective fingerprints in panel (a) and panel (b) of Figure 1, respectively. The domains for minutiae locations and orientations are $S_0 = [0, 512] \times [0, 512]$ and $(0, \pi]$ (corresponding to $L = 2$). Hyper-parameter specifications are chosen as in the simulation experiments. A total of $I = 5$ chains was run using the RJMCMC algorithm with different starting values representing overdispersion between the chains. The five initial estimates are obtained using an agglomerative clustering procedure with different settings for the number of clusters. The convergence of model parameters is achieved at 80,000 for Data 1 and nearly achieved for Data 2. Panels (a) and

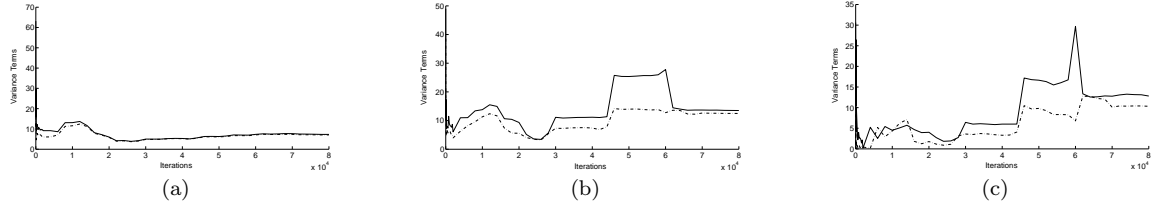


FIG 6. Convergence diagnostics for Data 2. Panels (a), (b) and (c) show the plots of (\hat{V}, W_c) , $(W_c, W_m W_c)$ and $(B_m, B_m W_c)$, respectively. A solid line represents the first entry and a dotted line represents the second entry. The x-axis unit is 10,000 iterations.

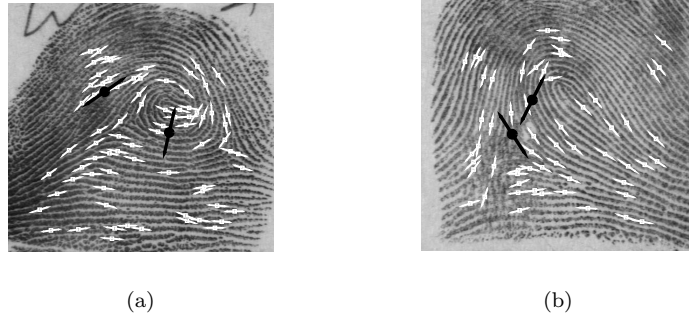


FIG 7. Posterior location and orientation means overlaid on the fingerprint images.

(b) of Figure 6 indicate convergence of the predictive characteristics and model parameters for $L = 2$ but panel (c) indicates that only an approximate inference can be made for the number of mixture components. Note that the predictive characteristics of both chains converged much faster, by 30,000 and 40,000 iterations for Data 1 and Data 2, respectively. The corresponding plots and tables are available in the supplementary material.

For Data 1 and 2, 1,000 posterior samples (after convergence) are obtained from each of the 5 chains. The posterior probabilities for the total number of clusters, K , are given in Table 3. There is strong evidence of the existence of multiple clusters in the point patterns of both fingerprint images; for both Data 1 and Data 2, $K = 2$ is strongly supported by the posterior (with probabilities 0.9774 and 0.9452, respectively). Although panel (c) of Figure 6 indicates convergence has not taken place, the high posterior probability of $K = 2$ in the case of Data 2 is evidence enough that this is the K suggested by the data. Tables 4 and 5 give the estimates

Fingerprint/ K	2	3	4	5
Data 1	0.9774	0.0224	0.0002	0.0000
Data 2	0.9452	0.0510	0.0038	0.0000

TABLE 3
Posterior probabilities of K for fingerprint data.

Data 1							
Parameter	Mean	Sd	CI	Parameter	Mean	Sd	CI
θ_{11}	186.70	40.16	[110.02, 328.16]	θ_{21}	171.01	66.51	[100.68, 324.46]
θ_{12}	311.67	50.01	[174.56, 398.24]	θ_{22}	249.64	64.10	[107.85, 326.37]
σ_{11}^2	11,344	19,881	[25, 56, 960]	σ_{21}^2	4,922	6,6260	[25, 18, 105]
σ_{12}^2	15,673	16,059	[29, 48, 117]	σ_{22}^2	9,232	5,540	[28, 16, 815]
η_1	2.53	0.56	[0.05, 3.11]	δ_1	1.94	7.86	[0.02, 55.16]
η_2	1.79	0.98	[0.02, 3.12]	δ_2	6.50	13.08	[0.02, 66.45]
ρ_1	0.54	0.95	[0.01, 4.69]	γ_1	30.86	33.90	[2.09, 114.98]
ρ_2	0.42	0.88	[0.01, 4.62]	γ_2	50.51	28.59	[1.20, 104.93]

TABLE 4

The results of posterior inference for $K = 2$ based on Data 1.

Data 2							
Parameter	Mean	Sd	CI	Parameter	Mean	Sd	CI
θ_{11}	170.22	35.16	[101.28, 258.92]	θ_{21}	263.25	29.78	[108.06, 322.97]
θ_{12}	207.49	31.09	[127.72, 337.22]	θ_{22}	199.43	97.71	[100.89, 379.83]
σ_{11}^2	45,904	18,012	[31, 61, 400]	σ_{21}^2	15,968	6,217	[50, 25, 857]
σ_{12}^2	1,701	9,128	[13, 57, 500]	σ_{22}^2	727.80	3,107	[16, 18, 957]
η_1	0.99	0.31	[0.23, 2.56]	δ_1	0.39	0.33	[0.07, 1.29]
η_2	2.03	0.74	[0.02, 3.12]	δ_2	0.73	2.61	[0.02, 15.59]
ρ_1	0.04	0.18	[0.01, 1.24]	γ_1	67.67	15.99	[2.40, 102.07]
ρ_2	0.77	1.09	[0.01, 4.75]	γ_2	6.64	10.76	[0.62, 79.51]

TABLE 5

The results of posterior inference for $K = 2$ based on Data 2.

of parameters based on the K with the highest posterior probabilities. For comparison purposes, the parameter estimates of θ and η are plotted on the original fingerprint images together with observed data in Figure 7. Both panels indicate that the proposed models provide a reasonable fit to the observed clusters of minutiae locations and orientations.

The small mean values of ρ in Tables 4 and 5 indicate the presence of strong spatial dependence between neighboring orientation values. As an assessment of goodness of fit in terms of capturing the extent of spatial dependence, the DIC (Deviance Information Criteria; see Celeux et al. 2006) is computed for the proposed model as well as for a model with no spatial dependence (i.e., $\rho_k = \infty$ in (24)). We call these models Model 1 and 2, respectively. The results are given in Table 6 which clearly demonstrates the superiority of Model 1 (lower DIC) for fingerprint data.

The EPIC measure of fingerprint individuality (see Section 5.3) based on $N_0 = 5,000$ samples corresponding to $\tau_0 = 13$ matches is $836/5000 = 0.1672$. The high value indicates weak evidence that this is not an impostor pair, therefore, suggesting (and correctly) that the two fingerprints in Figure 1 are possibly an impostor pair.

	Model 1	Model 2
Data 1	52.2	56.4
Data 2	30.8	44.9

TABLE 6

Goodness of fit assessment based on DIC.

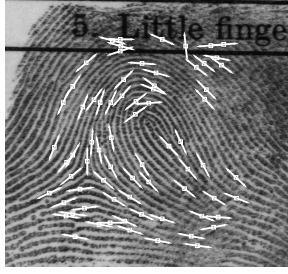
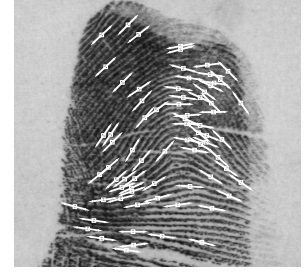
(a) $(m, q) = (72, 0.365)$ (b) $(m, q) = (64, 0.718)$ (c) $(m, q) = (74, 0.286)$

FIG 8. Additional fingerprints from the NIST database: m and q are, respectively, the total number of minutiae and image quality measure.

Fingerprint pairs	m_1, m_2	q_1, q_2	τ_0	EPIC
Impostor pairs				
f0001,f0002	76, 59	0.447, 0.419	13	0.1627
f0009,f0031	55, 53	0.722, 0.709	13	0.0982
f0091,f0092	57, 74	0.335, 0.286	17	0.2756
f0009,s0015	55, 55	0.722, 0.894	13	0.1770
s0091,s0031	48, 53	0.467, 0.560	14	0.0880
s0001,f0092	86, 74	0.331, 0.286	19	0.1516
Genuine pairs				
f0002,s0002	59, 72	0.419, 0.365	22	0.0000
f0009,s0009	55, 64	0.722, 0.718	28	0.0000
f0092,s0092	74, 68	0.286, 0.303	18	0.3128
f0031,s0031	53, 53	0.709, 0.560	32	0.0000
f0091,s0091	57, 48	0.335, 0.467	20	0.0056
f0015,s0015	40, 55	0.810, 0.894	16	0.0002

TABLE 7

Reporting EPIC for additional fingerprint pairs: For $j = 1, 2$, m_j and q_j are the total number of minutiae and quality measure for the j -th fingerprint.

It is reasonable to expect EPIC to possess certain properties with respect to image attributes. For example, EPIC values should be much larger in lower, compared to higher, quality images for the same number of observed minutiae matches. Also, when the total number of minutiae increases, one would expect EPIC to increase since it is more likely to obtain a random match. To investigate these conjectures, we performed matching experiments with eleven additional pairs of fingerprints from the NIST database. Image quality is obtained using the quality extractor reported in Dass (2010), which is based on a wavelet transform that measures the clarity of the ridge and valley structures. The quality of a fingerprint image is a number between 0 and 1 with higher values indicating better quality. Figure 8 gives some examples of the images used with associated quality and total number of minutiae. The RJMCMC algorithm was run until convergence for each fingerprint image and EPIC values were obtained for all pairs listed in Table 7. The first six pairs in Table 7 are impostors whereas the last six are genuine. It is clear from the entries of Table 7 that the reported EPIC values, indeed, behave in the expected way. For example, impostor fingerprints in the second row are of better quality compared to the first, they have the same observed matching number, τ_0 , and the EPIC values in the second row are lower than the first; we assume that the total number of minutiae are comparable although 76 is a bit higher but not too much. The same argument can be made when comparing the first and third rows, as well as the second and third rows, by the monotonicity of EPIC. For the truly genuine pairs, the reported EPIC values correctly reflect strong evidence against the hypothesis that the pairs are impostors for pairs with different quality measures; see, for example, the first and second rows for genuine pairs. In the third row, note that the EPIC value is large, indicating high uncertainty when the underlying quality of the images is poor.

8. Summary and Conclusion. A family of marked spatial point processes with dependent marks is developed in this paper for fingerprint minutiae. The proposed models are flexible in the sense that they can capture salient minutiae characteristics in different fingerprint images. A Bayesian framework of inference is developed to estimate the unknown parameters in this family based on an RJMCMC algorithm. Our analysis establishes the presence of clustering and strong spatial dependence between minutiae orientations which are typical of all fingerprints. The inferential methodology based on the RJMCMC is somewhat slow to converge. Predictive characteristics of the chain, however, converge fairly quickly and can be utilized for the assessment of fingerprint individuality. This is possible since the calculation of the EPIC is based on the number of matches between minutiae configurations in both the fingerprint images, which is a function of the predictive distribution of the proposed model.

One of the referees has pointed out that it may be worthwhile to formulate the hierarchical model by eliciting the distribution in (15) in terms of Dirichlet processes (DP). Further, such a scheme would not require the use of RJMCMC trans-dimensional sampling. Our future goal will be to investigate this issue further. Another important area for research will be to adapt the

distributions in this paper so that they reflect uncertainty due to other image attributes, such as image quality, *in the model elicitation*. This investigation has been carried out for a simpler model than the one proposed here; see, for example, Dass (2010). While EPIC provides a quantitative measure for assessing the extent of individuality, determining sensitivity and specificity (in terms of thresholds based on EPIC) has yet to be investigated and remains an area where more work is needed.

Our model for \mathcal{S} is not able to capture any strong dependence between minutiae sets, for example, minutiae sets that are extracted for a truly genuine pair of prints in the authentication process. To capture this strong dependence, such a model should be able to represent the second minutiae set as a small perturbation of the first. Although this is the general idea, developing a model for \mathcal{S} in the genuine case that adequately describes the number of matches in real databases is far more complicated; we have to incorporate other extraneous noise variability for real fingerprint images such as the rates of missed and false minutiae correspondences. Our future research will address this problem. Such a model will be useful in the development of Bayes factors and posterior probabilities (corresponding to an impostor as well as a genuine match) as measures of fingerprint individuality.

Acknowledgment. The authors wish to acknowledge the support of NSF DMS grant no. 0706385 while conducting this research.

Supplementary Materials. Additional experimental results and the Appendix describing details of the RJMCMC algorithm are available online at the Supplementary Materials link.

REFERENCES

- Bognar, M. (2008), “Bayesian modeling of continuously marked spatial point patterns,” *Computational Statistics*, 23, 361-379.
- Breking, J. (1989), *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, New York: Springer Verlag.
- Brix, A. (1998), “Generalized gamma measures and shot-noise Cox processes,” *Advances in Applied Probability*, 31, 929-953.
- Brooks, S. P., and Giudici, P. (1999), “Convergence assessment for Reversible Jump MCMC Simulations,” in *Bayesian Statistics 6*, eds. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, Oxford: Oxford University Press, 733-742.
- (2000), “Markov Chain Monte Carlo convergence assessment via Two-Way analysis of variance,” *Journal of Computational and Graphical Statistics*, 9, 266-285.
- Celeux, G., Forbes, F., Robert, C. P., and Titterton, D. M. (2006), “Deviation information criteria for missing data models,” *Bayesian Analysis*, 1, 651-706.
- Coles, S. (1998), “Inference for circular distributions and processes,” *Statistics and Computing*, 8, 105-113.
- Daley, D. J., and Vere-Jones, D. (1988), *An Introduction to the Theory of Point Processes*, New York: Springer Verlag.
- Dass, S., and Li, M. (2009), “Hierarchical Mixture Models for Assessing Fingerprint Individuality,” *Annals of Applied Statistics*, 3, 1448-1466.
- Dass, S. (2010), “Assessing Fingerprint Individuality In Presence Of Noisy Minutiae,” *IEEE Transactions on Information Forensics and Security*, 5, 62-70.

- Daubert v. Merrel Dow Pharmaceuticals Inc. (1993), 509 U.S. 579, 113 S. Ct. 2786, 125 L.Ed.2d 469.
- Diggle, P. J. (2003), *Statistical Analysis of Spatial Point Patterns*, 2nd edition, London: Edward Arnold.
- Fisher, N. I. (1993), *Statistical Analysis of Circular Data*, New York: Cambridge University Press.
- Fisher, N. I., and Lee, A. J. (1994), "Time Series Analysis of Circular Data," *Journal of the Royal Statistical Society*, Ser. B, 56, 327-339.
- Green, P. (1995), "Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination," *Biometrika*, 82, 711-732.
- Green, P., and Richardson, S. (1997), "On the Bayesian analysis of mixtures with an unknown number of components," *Journal of the Royal Statistical Society*, Ser. B, 59, 731-792.
- Ickstadt, K., and Wolpert, R. (1999), "Spatial regression for marked point processes" (with discussion), in *Bayesian Statistics 6*, eds. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, Oxford: Oxford University Press, pp. 323-341.
- Kottas, A., and Sanso, B. (2007), "Bayesian Mixture Modeling for Spatial Poisson Process Intensities, with Applications to Extreme Value Analysis," *Journal of Statistical Planning and Inference* (Special Issue on Bayesian Inference in Stochastic Processes), 137, 3151-3163.
- Lawson, A. B., and Denison G. T. (eds) (2002), *Spatial Cluster Modeling*, Boca Raton: Chapman & Hall/CRC.
- Mardia, K. V., and Jupp, P. E. (1999), *Directional statistics*, Chichester: Wiley.
- Mardia, K. V., Taylor, C. C., and Subramaniam, G. K. (2007), "Protein Bioinformatics and Mixtures of Bivariate von Mises Distributions for Angular Data," *Biometrics*, 63, 505-512.
- Møller, J., and Waagepetersen, R. P. (2003), *Statistical Inference and Simulation for Spatial Point Processes*, Boca Raton: Chapman & Hall/CRC.
- Pankanti, S., Prabhakar, S., and Jain, A. K. (2002), "On the Individuality of Fingerprints," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24, 1010-1025.
- Ravindran, P. (2002), "Bayesian Analysis of Circular Data Using Wrapped Distributions," unpublished Ph.D. dissertation, North Carolina State University, Dept. of Statistics.
- Scolve, S. C. (1979), "The occurrence of fingerprint characteristics as a two dimensional process," *Journal of the American Statistical Association*, 74, 588-595.
- Shieh, G., and Johnson, R. (2005), "Inferences based on a bivariate distribution with von Mises marginals," *Annals of the Institute of Statistical Mathematics*, 57, 789-802.
- Singh, H., Hnizdo, V., and Demchuk, E. (2002), "Probabilistic model for two dependent circular variables," *Biometrika*, 89, 719-723.
- Stoney, D. A., and Thornton, J. I. (1986), "A critical analysis of quantitative fingerprint individuality models," *Journal of Forensic Sciences*, 31, 1187-1216.
- Stoyan, D., and Stoyan, H. (1998), "Non-homogeneous Gibbs process models for forestry-a case study," *Biometrical Journal*, 40, 521-531.
- U. S. v. Byron Mitchell (1999), Criminal Action No. 96-407, U. S. District Court for the Eastern District of Pennsylvania.
- U. S. v. Crisp (2003), 324 F 3d 261 (4th Cir).
- Wolpert, R. L., and Ickstadt, K. (1998), "Poisson/gamma random field models for spatial statistics," *Biometrika*, 85, 251-267.
- Zhu, Y., Dass, S. C., and Jain, A. K. (2007), "Statistical Models for Assessing the Individuality of Fingerprints," *IEEE Transactions on Information Forensics and Security*, 2, 391-401.

CHAE Y. LIM,
ASSISTANT PROFESSOR,
DEPARTMENT OF STATISTICS & PROBABILITY
MICHIGAN STATE UNIVERSITY,
EAST LANSING, MI 48824
E-MAIL: lim@stt.msu.edu
URL: <http://www.stt.msu.edu/~sdass>

SARAT C. DASS,
ASSOCIATE PROFESSOR,
DEPARTMENT OF STATISTICS & PROBABILITY,
MICHIGAN STATE UNIVERSITY,
EAST LANSING, MI 48824
E-MAIL: sdass@stt.msu.edu