ARE MADE AND MISSED DIFFERENT? AN ANALYSIS OF FIELD GOAL ATTEMPTS OF PROFESSIONAL BASKETBALL PLAYERS VIA DEPTH BASED TESTING PROCEDURE

BY KAI QI[1,a], GUANYU HU[2,b] AND WEI WU[3,c]

[1]*Microsoft,* [a]*kq15b@my.fsu.edu*

[2]*Center for Spatial Temporal Modeling for Applications in Population Sciences, Department of Biostatistics and Data Science, University of Texas Health Science Center at Houston,* [b]*guanyu.hu@uth.tmc.edu*

[3]*Department of Statistics, Florida State University,* [c]*wwu@fsu.edu*

In this paper we develop a novel depth-based testing procedure on spatial point processes to examine the difference in made and missed field goal attempts for NBA players. Specifically, our testing procedure can statistically detect the differences between made and missed field goal attempts for NBA players. We first obtain the depths of two processes under the polar coordinate system. A two-dimensional Kolmogorov–Smirnov test is then performed to test the difference between the depths of the two processes. Throughout extensive simulation studies, we show our testing procedure with good frequentist properties under both null hypothesis and alternative hypothesis. A comparison against the competing methods shows that our proposed procedure has better testing reliability and testing power. Application to the shot chart data of 191 NBA players in the 2017–2018 regular season offers interesting insights about these players' made and missed shot patterns.

1. Introduction. Do made and missed field goal attempts follow different spatial processes for Stephen Curry? Does LeBron James have more field goal attempts in the region where he has a higher field goal percentage? In the National Basketball Association (NBA), analyzing field goal attempts for different players is a crucial problem for each team. From the players' point of view, it helps them optimize their field goal attempts and design their training plans to improve their shooting skills. The optimal selection of field goal attempts for a player is that he or she can make field goal attempts on all locations without differences. From the teams' perspective, it helps them improve their training plans for different players or different positions. In order to answer the questions above, we need to know if there is any difference between the made and missed locations of field goal attempts. This paper aims to provide a statistical testing procedure to discriminate made and missed processes.

Quantitative analytics has been a key driving force in advancing modern professional sports, and there is no exception for professional basketball (Kubatko et al. (2007)). In the NBA, shot chart data that contain both the location and the result of each field goal attempt will offer important insights about players' attacking styles and shed light on the evolution of defensive tactics. There has been much literature discussing shot chart data (see, e.g., Erčulj and Štrumbelj (2015), Gómez Ruano, Alarcón López and Ortega Toro (2015), Hu, Yang and Xue (2021), Hu et al. (2023), Jiao, Hu and Yan (2021a), Miller et al. (2014), Reich et al. (2006), Shortridge, Goldsberry and Adams (2014)). Most existing studies rely on developing statistical models to analyze the spatial patterns of different players' field goal attempts or find the latent subgroups among different players. Reich et al. (2006) proposed a hierarchical regression model under point-reference data framework for both field goal attempt frequency and accuracy over grids on the basketball court. However, they failed to capture the randomness of field goal attempts since the shot locations were fixed in their study. Considering

randomness of field goal attempts in a basketball game is very important for quantitative analysis of shot charts data. Compared to traditional point reference analysis in Reich et al. (2006), spatial point processes (Diggle (2014)) are natural to model randomness of locations. There is a large body of literature in the field of spatial point pattern data (see, e.g., Baddeley (2017), Diggle (2014), Geng, Shi and Hu (2021), Guan (2006), Guan and Shen (2010), Illian et al. (2008), Jiao, Hu and Yan (2021b)) under both parametric and nonparametric frameworks, which thoroughly analyze the first-order(intensity) and second-order (pair correlation) properties of spatial point processe. Miller et al. (2014), Yin et al. (2022) analyzed the intensity surface of field goal attempts of professional players, which offered interesting insights about the shooting patterns. Based on intensity estimates, Hu, Yang and Xue (2021), Yin, Hu and Shen (2023) discovered latent subgroups among NBA players to gain important insight about shooting similarities between different players. Jiao, Hu and Yan (2021a) proposed a Bayesian marked point process framework for jointly modeling shot frequency and accuracy to discover the relationship between them. Although quantitative analysis (regression analysis and clustering analysis) of shot charts based on spatial point process has been well-explored, discussions of the testing procedure for two or more processes are minimal.

The initial problem for testing is to select a proper statistical measure for the spatial point process. The statistical depth provides an effective framework to describe the central tendency (Liu and Wu (2017)) and the ordered property (Qi, Chen and Wu (2021)) for the spatial point process. The notion of statistical depth was first introduced (Tukey (1975)) as a tool for visualizing bivariate data sets and has been extended to multivariate data over the last few decades. The depth is a measure of the centrality of a point with respect to a certain data cloud, which helps to set up center-outward ordering rules of ranks. Based on different centrality criteria, a large class of depths has been proposed, including the halfspace depth (Tukey (1975)), convex hull peeling depth (Barnett (1976)), simplicial depth (Liu (1990)), L_1-depth (Vardi and Zhang (2000)), and projection depth (Zuo (2003)). The concept of statistical depth has been widely applied in outlier detection (Donoho and Gasko (1992)), multivariate density estimation (Fraiman, Liu and Meloche (1997)), nonparametric description of multivariate distributions (Liu, Parelius and Singh (1999)), and depth-based classification and clustering (Christmann (2002)).

On the other hand, the depth-based statistical tests are still underexplored in both theory and applications. The earliest related work in this area, though different in the objective, was the bivariant rank test based on Jia's concept of ordering (Oja (1983), Hannu and Jukka (1989), Brown et al. (1992)). Then for multivariate data, a Wilcoxon's type of test is proposed by Liu and Singh (1993) (henceforth, Liu–Singh rank sum test) since the data depth is naturally related to ranking. As a theoretical foundation, Zuo and He (2006) elaborately studied the limiting distribution and asymptotic of Liu–Singh's rank sum test. Multivariate depth functions can transfer data into a one-dimensional space, and many one-dimensional testing techniques can then be applied to the depth values. In this respect a well-composed generalization can be found in Zhang, Xiang and Shen (2012). Another graphic form of depth-based test, the DD (depth vs. depth) plot, was introduced by Liu, Parelius and Singh (1999) as an example of the potential application of depth function. Then Li and Liu (2004) systematically studied the DD plot by examining its performance w.r.t. location shift and scale expansion or contraction. In this paper we focus on depth-based testing procedures for the spatial point process to explore the difference between made and missed field goal attempts among different players in the NBA.

The contribution of this paper is two-fold. First, we introduce the two-dimensional depth calculation for the spatial point process under the polar coordinate. Compared with the Cartesian coordinate, the polar coordinate represents the field goal attempts more appropriately, since the distance and angle are the two most important factors for shooting selections by

professional players (Reich et al. (2006)). Then we propose a new depth-based Kolmogorov–Smirnov test for two spatial point processes based on a two-dimensional depth.

The rest of this article is organized as follows. In Section 2 we select one representative player for each position from the 2017–2018 NBA regular season and present their shot charts. In Section 3 we first briefly discuss the spatial point process model and depth calculation for spatial point pattern data. Then a depth-based testing procedure is proposed to discriminate the difference between made and missed processes. Extensive simulation studies are conducted in Section 4 to investigate the empirical performance of the proposed testing procedure. Applications of the proposed testing procedure to 191 NBA players in the 2017–2018 regular season are reported in Section 5. We conclude this paper with a brief discussion in Section 6.

2. Motivating data. The NBA shot charts are freely available on the official NBA site nba.com/stats, and our study focuses on the regular season in 2017–2018. The data for each player contain information about all field goal attempts in the regular season. From the data the half-court is positioned on a Cartesian coordinate system centered at the center of the rim, with x ranging from -250 to 250 and y ranging from -50 to 420, both with the unit of 0.1 foot, as the size of an actual NBA basketball half-court is 50 feet $\times$ 47 feet. We pick up one representative player for each position (Point Guard (PG), Shooting Guard (SG), Small Forward (SF), Power Forward (PF), or Center (C)), such as Stephen Curry, DeMar DeRozan, LeBron James, Giannis Antetokounmpo, and DeAndre Jordan. Figure 1 shows their field goal attempts' locations, colored to indicate made or missed for each attempt.

As we can see from the plots, DeAndre Jordan does not have any field goal attempts from 3-point line. Most of the shots of Stephen Curry are made either close to the rim or right outside the arc (i.e., 3-point line). DeMar DeRozan tends to have a uniformly distributed shot pattern throughout the court. LeBron James has relatively more shots in the right-hand side
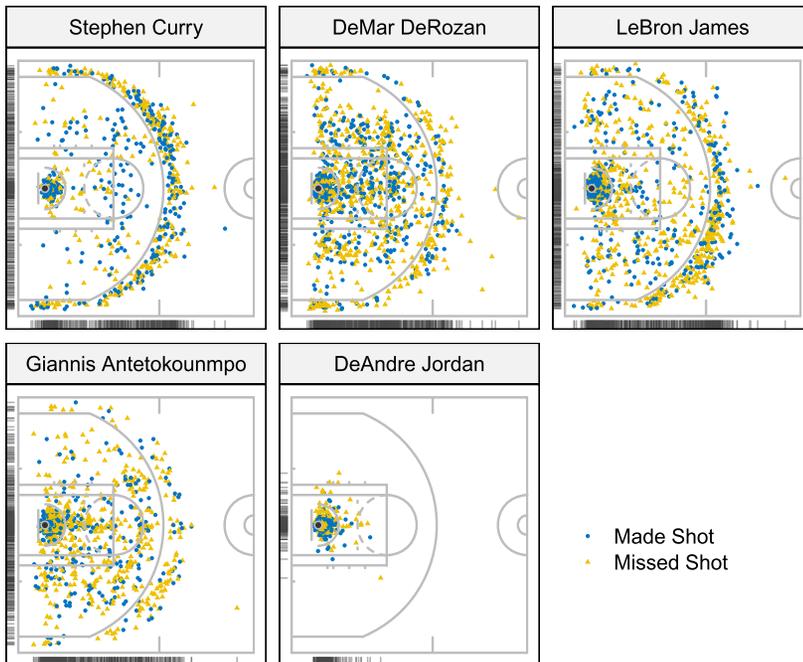
FIG. 1. *Field Goal Attempts Data Display. On half court image, each point represents one shot. The marginal distributions of the shot locations are shown on the left and down sides of each plot, respectively.* (*Players*: *Stephen Curry, DeMar DeRozan, LeBron James, Giannis Antetokounmpo, DeAndre Jordan*).

area toward the basket. Giannis Antetokounmpo apparently has more missed attempts from 3-point line.

3. Model setup.

3.1. *Spatial point process.* For a particular player, the observed shot chart can be represented by two spatial point processes $\mathbf{S}_{\text{made}} = (\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_{n_1})$ and $\mathbf{S}_{\text{missed}} = (\mathbf{s}'_1, \mathbf{s}'_2, \ldots, \mathbf{s}'_{n_2})$, where $\mathbf{S}_{\text{made}}$ and $\mathbf{S}_{\text{missed}}$ are the collections of the locations of made and missed field goal attempts and n_1 and n_2 are the numbers of shots for made and missed processes, respectively.

Spatial point process model provides a natural framework for capturing the random behavior of event location data. For example, in the made process, let $\mathbf{s}_i = (x_i, y_i)$, $i = 1, \ldots, n_1$ be the set of x- and y-coordinate locations for points that are observed in a predefined, bounded region $\mathcal{B} \subseteq \mathbb{R}^2$, which is the half basketball court. We denote the underlying stochastic mechanism that gives rise to the observed point pattern $\mathbf{S}_{\text{made}}$ as the spatial point process $\mathbf{Y}$. The process $N_{\mathbf{Y}}(A) = \sum_{i=1}^{n_1} \mathbb{1}(\mathbf{s}_i \in A)$ is a counting process associated with the spatial point process $\mathbf{Y}$, which counts the number of points that fall into the area $A \subseteq \mathcal{B}$.

A large body of literature on probability distributions exists for spatial point processes (see, e.g., Diggle (2014) and references therein), including Poisson processes, Gibbs processes, and Cox processes, etc. In that literature the most widely adopted class of models is the non-homogeneous Poisson processes (NHPP), which assumes conditionally independent event locations with a deterministic intensity $\lambda(\mathbf{s})$. Under the NHPP framework, the number of events in area A, $N_{\mathbf{Y}}(A)$, follows a Poisson distribution with rate parameter $\Lambda(A) = \int_A \lambda(\mathbf{s}) \, d\mathbf{s}$. In addition, $N_{\mathbf{Y}}(A_1)$ and $N_{\mathbf{Y}}(A_2)$ are independent if two areas $A_1 \subseteq \mathcal{B}$ and $A_2 \subseteq \mathcal{B}$ are disjoint. Given the observed point pattern $\mathbf{S}_{\text{made}}$ on fixed region $\mathcal{B}$, the likelihood of NHPP is

$$L = \frac{\prod_{i=1}^{n_1} \lambda(\mathbf{s}_i)}{\exp(\int_{\mathcal{B}} \lambda(\mathbf{s}) \, d\mathbf{s})},$$

where $\lambda(\mathbf{s}_i)$ is the intensity function evaluated at location $\mathbf{s}_i$. As mentioned in Section 1, most existing literature focuses on intensity estimation or pair correlation estimation and joint model for marked point process. In this paper our goal is to propose a testing procedure for the null hypothesis

$$H_0 : \mathbf{S}_{\text{made}} \text{ and } \mathbf{S}_{\text{missed}} \text{ are the same process.}$$

Our testing procedure is based on the depth of spatial point process.

3.2. *Depth calculation.* Let $\mathbf{S}$ denotes a collection of locations in a bounded set $\mathcal{B} \subseteq \mathbb{R}^2$; a depth function essentially is a map from $\mathcal{B}$ to $\mathbb{R}^+$ which generates an appropriate center-outward rank for locations of points.

Depth functions for multivariate data can be broadly categorized as nonparametric depth (e.g., half-space/Tukey depth) and parametric forms (e.g., Mahalanobis depth). While different depth functions have various properties (such as shapes and contours), one should choose an appropriate depth based on the nature of data. Here we briefly describe two classical depth functions as follows.

Let $\mathbf{P}$ be the probability measure on $\mathbb{R}^d$, $d \geq 1$, then for a point $z \in \mathbb{R}^d$, the Mahalanobis depth and Tukey's depths are defined in the following forms:

Mahalanobis Depth (MD) (Liu and Singh (1993)): Let μ and Σ denote the mean vector and variance-covariance matrix of $\mathbf{P}$, the Mahalanobis depth (MD) for z w.r.t. $\mathbf{P}$ is

$$(3.1) \qquad MD(z; \mathbf{P}) = \left[1 + (z - \mu)^T \Sigma^{-1} (z - \mu)\right]^{-1}.$$

Half-space/Tukey's Depth (HD) (Tukey (1975)): The HD of z w.r.t. **P** is defined as

$$(3.2) \qquad HD(z; \mathbf{P}) = \inf_H \{ P(H) : \text{ H is a closed halfspace in } \mathbb{R}^d \text{ containing z} \}.$$

By replacing **P** in equations (3.1) and (3.2) with empirical distribution $\mathbf{P}_n$, we have the sample version of Mahalanobis depth and Tukey's depth. Both Mahalanobis depth and Tukey's depth for multivariate data are well-established. In contrast to the stereotyped elliptic depth contours of Mahalanobis depth, Tukey's depth provides more flexible and appropriate contours. As a nonparametric depth function, Tukey's depth does not rely on the estimation of mean and variance. In the following sections, we will adopt Tukey's depth except otherwise specified.

For $d > 1$, the computation relies on effectively finding a hyperplane containing z_i that satisfies the definition of Tukey's depth in equation (3.2). Such computation can be extremely inefficient or infeasible, especially for high-dimensional data. In this paper the computation of depth has never been extended to more than two dimensions. For $d = 2$, a straightforward algorithm for Tukey's depth needs $\mathcal{O}(n^2)$ steps. We will adopt a faster algorithm proposed by Rousseeuw and Ruts (1996), which only requires $\mathcal{O}(n \log n)$ steps, where n denotes sample size.

3.3. *Depth based testing.*

The performance of parametric statistical tests often relies on the normality assumption of the underlying distribution. In contrast, some nonparametric tests can perform better when dealing with symmetrically distributed data (e.g., KS test (Mohd Razali and Yap (2011)), Wilcoxon's test). To address this, a data transformation function is needed to improve the reliability and consistency of tests. On the other hand, it is generally challenging to extend the one-dimensional tests to multivariate space directly. For example, since the maximum difference of two joint cumulative distribution functions is not clearly defined, the Kolmogorov–Smirnov test (KS test) cannot be simply generalized for high-dimensional data. One common approach to generalizing those tests to high-dimensional space is through dimensional reduction techniques.

In general, a multivariate depth function is a continuous function mapping data from $\mathbb{R}^d$ to $\mathbb{R}^+$ that holds the following four properties (Zuo and Serfling (2000)): affine invariance, maximality at the center, monotonicity relative to the deepest point, and vanishing at infinity. All these properties make multivariate depth function an ideal predefined transformation when performing statistical tests on high-dimensional data.

Here, we briefly review the depth-based goodness-of-fit tests that treat depth functions as transformations to reduce data dimensionality, include Liu–Singh rank-sum test, depth-based Pearson's chi-square test, and depth-based KS test and Cramer–Von Mises test. We will then propose our new depth-based test.

3.3.1. *Liu–Singh rank sum test.*

The well-known Wilcoxon rank sum can naturally incorporate depth functions. In this respect, Liu and Singh (1993), Zuo and He (2006) systematically generalized the Wilcoxon rank sum test to multivariate data through depth functions (refer to as Liu–Singh depth-based rank sum test). Briefly, the Liu–Singh depth-based rank sum test can be described as follows.

Let $X \sim H$ and $Y \sim G$ be two independent random variables in $\mathbb{R}^d$, and $D(\cdot; H)$ be a depth function of a given distribution H that maps $\mathbb{R}^d$ to $[0, 1]$. Then the outlyingness of y w.r.t. H can be measured by $R(y; H) = P_H(X : D(X; H) \leq D(y; H))$, and we define

$$Q(H, G) := \int R(y; H) \, dG(y) = P\{D(X; H) \leq D(Y; H) | X \sim H, Y \sim G\}.$$

Following Proposition 3.1 in Liu and Singh (1993), for continuously distributed $D(\cdot; H)$, $R(Y; H)$ follows a uniform distribution in $[0, 1]$. Under the null hypothesis $H_0 : H = G$, the

test statistic $Q(H, G) = \frac{1}{2}$ and $Q(H, G) < \frac{1}{2}$ indicates the possible location shift or scale increase from H to G.

In practice, one can compare two empirical distributions H_m and G_n through the two-sample version of Liu–Singh depth-based rank sum test such that

$$(3.3) \qquad Q(H_m, G_n) := \int R(y; H_m) \, dG_n(y) = \frac{1}{n} \sum_{j=1}^{n} R(Y_j; H_m).$$

While the asymptotic results rely on the dimensionality of data and properties of depth function, for dimension $d = 1$, under null hypothesis, Liu and Singh (1993) have shown that

$$((1/m + 1/n)/12)^{-1/2} (Q(H_m, G_n) - 1/2) \to N(0, 1).$$

By assuming more conditions on the depth function, Zuo and He (2006) extended the asymptotic studies of Liu–Singh depth-based rank sum test to higher dimensions under both null and alternative hypothesis.

3.3.2. *Depth-based Pearson's chi-square test.*

Another important depth-based test is derived from Pearson's chi-square test, which involves discretizing and separating the continuous depth values into multiple intervals. By counting the number of observations that fall into those predefined depth value intervals, one can form a classical Pearson's chi-squared test against expected frequencies. Here we briefly describe the test procedure as follows.

Let $X \sim F$ be a random variable in $\mathbb{R}^d$ and $D(\cdot; F)$ be a depth function that maps X to $[0, 1]$. Given a random sample of size n, under $H_0 : F = F_0$, denote $z_{j, F_0} = D(x_j; F_0)$ as the depth of x_j with respect to F_0, $j = 1, 2, \ldots, n$. Assuming the depth values can be separated by $a_i, i = 1, \ldots, r$ such that

$$0 = a_0 < a_1 < \cdots < a_{r-1} < a_r = 1.$$

Define

$$A_i = \{X : a_{i-1} < D(X; F_0) \leq a_i, X \in R^d\}, \quad i = 1, \ldots, r,$$

$$p_i = P_{F_0}(A_i), n_i = \#\{j : a_{i-1} < z_{j, F_0} \leq a_i, j = 1, \ldots, n\}, \quad i = 1, \ldots, r.$$

Now, we have the expected frequency np_i and observed frequency n_i for r regions: $A_1, A_2, \ldots, A_r \in \mathbb{R}^d$, and the Pearson's chi-square test statistic, shown in Section 2.1 in Zhang, Xiang and Shen (2012), is computed by

$$(3.4) \qquad \chi^2 = \sum_{i=1}^{r} \frac{(n_i - np_i)^2}{np_i} \sim \chi^2(r - 1).$$

To make the depth-based chi-square test powerful, one does not only require the sample size $n \geq 50$ but also need to ensure the theoretical frequencies $np_i \geq 5$ for each segmented region $A_i, i = 1, \ldots, r$.

3.3.3. *Depth-based Cramer–Von Mises and Kolmogorov–Smirnov test.*

A multivariate depth function transforms data from high-dimensional space to one dimension. Therefore, some general one-dimensional tests, such as the classical Kolmogorov–Smirnov (KS) test and Cramer–Von Mises (CM) test, can be performed on the transformed data.

Denote $F_0(x)$ as a continuous distribution function and $F_n(x)$ as the empirical distribution function of sample $X_1, X_2, \ldots, X_n \in \mathbb{R}^d$. Then the KS and CM statistics can be denoted as

$$\text{KS} = \sup_{x \in R^d} |F_n(x) - F_0(x)|, \text{CM} = \int_{R^d} [F_n(x) - F_0(x)]^2 \, dF_0(x).$$

The KS test or CM test depends largely on the distribution F_0, and technically, they are not nonparametric tests. But one can apply a depth-based transformation $R(\cdot)$ (Zhang, Xiang and Shen (2012)) of Liu–Singh's test as

$$T_i = P_{F_0}(X : D(X; F_0) \geq D(X_i; F_0)), \quad i = 1, 2, \ldots, n.$$

Under Proposal 3.1 in Liu and Singh (1993), T_i follows a uniform distribution in $[0, 1]$ under null hypothesis. The test of equal distribution is equivalent to: $H_0 : T_1, \ldots T_n$ is a random sample from $U[0, 1]$. Denote $G_n(t)$ as the empirical distribution function of $\{T_i, i = 1, \ldots, n\}$, then the one-dimensional KS test statistic and CM test statistic, based on $G_n(t)$, are

$$\text{KS}^* = \sup_{t \in [0,1]} |G_n(t) - t|, \text{CM}^* = \int_0^1 [G_n(t) - t]^2 \, dt.$$

Notice that both test statistics KS^* and CM^* do not depend on F_0, and they are nonparametric tests.

3.4. *Two-dimensional KS test.* The depth-based statistical tests mentioned above include two essential steps: (1) apply a multivariate depth function to reduce the dimensionality of each data point to one and (2) perform a one-dimensional test on the scalar depth values. Inevitably, part of dimensional information (e.g., shape) is lost during the dimension reduction. In this section we seek an alternative approach of incorporating depth into hypothesis testing.

Among many popular multivariate tests of equal distribution, it is well-known that the KS test is sensitive to location shift and shape change. In practice, the generalization of the KS statistic to multivariate settings is challenging since the maximum difference between two joint cumulative distribution functions is not well-defined. One solution proposed by Justel, Peña and Zamar (1997) is Rosenblatt's transformation, but the computational process of Rosenblatt's transformation is rather complicated to use. Another natural approach is to compare the CDFs of two samples from all possible orders. In this aspect, Peacock (1983) proposed a two-dimensional KS test, which is hard to extend to more than two dimensions due to the computational cost. Fasano and Franceschini (1987) further improved Peacock's work by proposing a faster version, which also considers the correlation of data and sample size and then generalizes it to a three-dimensional case. Although the theoretical study of Peacock's test for dimensionality beyond three is lacking, we mainly focus on two-dimensional data in this paper.

For dimension $d = 2$, Peacock's method compares the integrated probability in each of four natural quadrants around a given point (x_i, y_i), namely, the total probabilities in $(x > x_i, y > y_i)$, $(x < x_i, y > y_i)$, $(x < x_i, y < y_i)$, $(x > x_i, y < y_i)$. Then the two-dimensional KS statistic T_{KS} is the maximum difference of the corresponding integrated probabilities over all data points and all quadrants.

Note that the distribution of statistic T_{KS}, in general, may not be independent of the shape of underlying two-dimensional distribution under the null hypothesis. Fasano and Franceschini (1987) approximated the distribution of T_{KS} as a function of sample size and correlation through extensive Monte Carlo simulations. Furthermore, by studying the Monte Carlo simulation results, the significant level of the two-dimensional KS test can be approximated by

$$(3.5) \qquad \Pr(T_{\text{KS}} > observed) = Q_{\text{KS}}\left(\frac{\sqrt{N} T_{\text{KS}}}{1 + \sqrt{1 - r^2}(0.25 - 0.75/\sqrt{N})}\right),$$

where N is the sample size,

$$Q_{\text{KS}}(x) = 2 \sum_{j=1}^{\infty} (-1)^{j-1} e^{-2j^2 x^2},$$

and r is the correlation coefficient. For comparing two samples with sizes n_1 and n_2, approximation equation (3.5) still holds with $N = \frac{n_1 n_2}{n_1 + n_2}$. Notice that the above approximation formula is accurate enough when $N > 20$ and when the indicated probability (significance level) is less than 0.20.

3.5. *A new two-dimensional depth-based KS test.* We now turn to propose our new statistical test based on statistical depths for basketball players' shot charts. The statistical scheme we presented in this section is built upon the polar representation of shot locations. Polar coordinates have been widely applied in shooting data analysis (Gudmundsson and Horton (2017), Reich et al. (2006)). During a basketball game (among many other games), players move around and focus on a fixed point (the basket); therefore, it is more natural to interpret a shot location as polar coordinates than the conventional Cartesian coordinate.

Then we are interested in testing if two shooting patterns are the same. Suppose one shooting pattern follows a Poisson process with intensity function $\lambda_1(\cdot)$ and another follows a spatial point process with intensity function $\lambda_2(\cdot)$. The goal of the proposed procedure is to test

$$H_0 : \lambda_1(\cdot) = \lambda_2(\cdot).$$

Notice that their intensity function completely defines the two spatial point processes; therefore, the null hypothesis is equivalent to

$$H_0 : F_1(\cdot) = F_2(\cdot),$$

where F_1 and F_2 are the corresponding distribution functions of two spatial point processes.

Suppose two independent spatial point process samples $\mathbf{S}_1 = (s_1, \ldots, s_{n_1})$ and $\mathbf{S}_2 = (s'_1, \ldots, s'_{n_2})$ have underlying intensity function λ_1 and λ_2, respectively, where $s_i = (x_i, y_i)$, $i = 1, \ldots, n_1$ and $s'_j = (x'_j, y'_j)$, $j = 1, \ldots, n_2$ are the observed locations in the pre-defined, bounded region $\mathcal{B} \in \mathbb{R}^2$. Then the proposed test procedures can be list as the following steps:

1. Transfer Cartesian coordinate (x, y) to polar coordinate (r, θ) by

$$r = \sqrt{(x^2 + y^2)}, \theta = \arctan\left(\frac{y}{x}\right).$$

Denote the transformed $\mathbf{S}_1$, $\mathbf{S}_2$ as $\mathbf{S}_1^* = (s_1^*, \ldots, s_{n_1}^*)$ and $\mathbf{S}_2^* = (s_1^{*'}, \ldots, s_{n_2}^{*'})$, where $s_i^* = (r_i, \theta_i)$ and $s_j^{*'} = (r'_j, \theta'_j)$. For basketball shot locations in this paper, we set the basket as the origin.

2. Estimated the pooled depth values for every location in $\mathbf{S}_1^*$ and $\mathbf{S}_2^*$.
Under $H_0 : \lambda_1 = \lambda_2$, we merge two sets as

$$\mathbf{S}_{\text{pool}}^* = (s_1^*, \ldots, s_{n_1}^*, s_1^{*'}, \ldots, s_{n_2}^{*'}).$$

As discussed previously, instead of using a two-dimensional depth, we apply an one-dimensional halfspace depth function $D(\cdot)$ on r and θ separately. For one-dimensional halfspace depth, $D(\cdot)$ is uniformly distributed in $[0, \frac{1}{2}]$ if the underlying distribution is continuous. Since the one-dimensional halfspace depth can be written as $D(x) = \min(F(x), 1 - F(x))$, where $F(x)$ is the CDF of a continuous variable X. It is well-known that $F(x)$ is uniformly distributed between $[0, 1]$; hence, $0 \leq D(x) \leq 1/2$, and

$$P\{D(x) \leq d\} = P\{F(x) \leq d, F(x) \leq 1/2\} + P\{1 - F(x) \leq d, F(x) \geq 1/2\}$$
$$= 2 \cdot P\{F(x) \leq d, F(x) \leq 1/2)\} = 2d.$$

The pooled depth of location (r, θ), based on the pooled set, therefore, is a two-dimensional value $(D(r; \mathbf{S}_{\text{pool}}^*), D(\theta; \mathbf{S}_{\text{pool}}^*))$, where $D(r; \mathbf{S}_{\text{pool}}^*)$ and $D(\theta; \mathbf{S}_{\text{pool}}^*)$ measure the

center-outward ranks of (r, θ) based on $\mathbf{S}^*_{\text{pool}}$ w.r.t. distance r and angular θ, respectively. $D(r; \mathbf{S}^*_{\text{pool}})$ and $D(\theta; \mathbf{S}^*_{\text{pool}})$ can be calculated as

$$D(r; \mathbf{S}^*_{\text{pool}}) = \min\left(\frac{\#\{r_k \in \mathbf{S}^*_{\text{pool}} : r_k \leq r, k = 1, \ldots, n_1 + n_2\}}{n_1 + n_2}, \right.$$
$$\left. \frac{\#\{r_k \in \mathbf{S}^*_{\text{pool}} : r_k \geq r, k = 1, \ldots, n_1 + n_2\}}{n_1 + n_2}\right),$$

and

$$D(\theta; \mathbf{S}^*_{\text{pool}}) = \min\left(\frac{\#\{\theta_k \in \mathbf{S}^*_{\text{pool}} : \theta_k \leq \theta, k = 1, \ldots, n_1 + n_2\}}{n_1 + n_2}, \right.$$
$$\left. \frac{\#\{\theta_k \in \mathbf{S}^*_{\text{pool}} : \theta_k \geq \theta, k = 1, \ldots, n_1 + n_2\}}{n_1 + n_2}\right).$$

By applying above functions to $\mathbf{S}^*_1$ and $\mathbf{S}^*_2$, we get the two-dimensional depth representation for points in $\mathbf{S}^*_1$ and $\mathbf{S}^*_2$, denote as $\mathcal{D}_1$ and $\mathcal{D}_2$.

3. Calculate the two-dimensional KS statistic T_{KS} and approximate the P-value.

The two-dimensional KS statistic T_{KS} is the maximum difference of the corresponding integrated probabilities over all data points and four natural quadrants. But when comparing two samples, T_{KS} depends on which dataset ranged over. As proposed by Press et al. (1986), one solution is to use the average of T_{KS} values obtained from ranging two samples.

For notation purpose, we denote $D(r; \mathbf{S}^*_{\text{pool}})$ as d_r, $D(\theta; \mathbf{S}^*_{\text{pool}})$ as d_θ. Let $C^1_{\mathcal{D}_1}(d_r, d_\theta)$, $C^2_{\mathcal{D}_1}(d_r, d_\theta), C^3_{\mathcal{D}_1}(d_r, d_\theta), C^4_{\mathcal{D}_1}(d_r, d_\theta)$ denote the counts of points of set $\mathcal{D}_1 = \{(d_{r_i}, d_{\theta_i}), i = 1, \ldots, n_1\}$ in quadrants $(d_r > d_{r_i}, d_\theta > d_{\theta_i})$, $(d_r \leq d_{r_i}, d_\theta > d_{\theta_i})$, $(d_r \leq d_{r_i}, d_\theta \leq d_{\theta_i})$, $(d_r > d_{r_i}, d_\theta \leq d_{\theta_i})$, respectively. Then

$$T_{\text{KS1}} = \max_{(d_{r_i}, d_{\theta_i}) \in \mathcal{D}_1}\left[\left|\frac{C^1_{\mathcal{D}_1}(d_{r_i}, d_{\theta_i})}{n_1} - \frac{C^1_{\mathcal{D}_2}(d_{r_i}, d_{\theta_i})}{n_2}\right|, \left|\frac{C^2_{\mathcal{D}_1}(d_{r_i}, d_{\theta_i})}{n_1} - \frac{C^2_{\mathcal{D}_2}(d_{r_i}, d_{\theta_i})}{n_2}\right|, \right.$$
$$\left. \left|\frac{C^3_{\mathcal{D}_1}(d_{r_i}, d_{\theta_i})}{n_1} - \frac{C^3_{\mathcal{D}_2}(d_{r_i}, d_{\theta_i})}{n_2}\right|, \left|\frac{C^4_{\mathcal{D}_1}(d_{r_i}, d_{\theta_i})}{n_1} - \frac{C^4_{\mathcal{D}_2}(d_{r_i}, d_{\theta_i})}{n_2}\right|\right],$$

$$T_{\text{KS2}} = \max_{(d_{r'_i}, d_{\theta'_i}) \in \mathcal{D}_2}\left[\left|\frac{C^1_{\mathcal{D}_1}(d_{r'_i}, d_{\theta'_i})}{n_1} - \frac{C^1_{\mathcal{D}_2}(d_{r'_i}, d_{\theta'_i})}{n_2}\right|, \left|\frac{C^2_{\mathcal{D}_1}(d_{r'_i}, d_{\theta'_i})}{n_1} - \frac{C^2_{\mathcal{D}_2}(d_{r'_i}, d_{\theta'_i})}{n_2}\right|, \right.$$
$$\left. \left|\frac{C^3_{\mathcal{D}_1}(d_{r'_i}, d_{\theta'_i})}{n_1} - \frac{C^3_{\mathcal{D}_2}(d_{r'_i}, d_{\theta'_i})}{n_2}\right|, \left|\frac{C^4_{\mathcal{D}_1}(d_{r'_i}, d_{\theta'_i})}{n_1} - \frac{C^4_{\mathcal{D}_2}(d_{r'_i}, d_{\theta'_i})}{n_2}\right|\right],$$

and

$$T_{\text{KS}} = \frac{T_{\text{KS1}} + T_{\text{KS2}}}{2}.$$

The p-value can be estimated using equation (3.5). The detailed algorithm can be found in Section 14 of Press et al. (1986). Once the p-value is calculated, one can compare it with the critical value (e.g., 5%) to decide whether to reject the null hypothesis.

Both proposed depth-based KS test and KS test we described in Section 3.3.3. Assume that $\mathbf{S}_1$ and $\mathbf{S}_2$ are independent. Comparing with depth-based tests we discussed in Section 3.3, the proposed framework is computationally efficient. This is because one only needs to loop through four quadrants rather than all possible directions. The computation time can be further reduced by the approximation methods we described. Besides, we expect the proposed

test to achieve a better performance since it utilizes more information to form test statistics. In the following section, we will compare our proposed test with other tests on simulations and real data.

4. Simulation study. In this section we mainly focus on testing the performance of our proposed test method through simulations. For comparing two spatial processes, an effective statistical testing scheme is the one that is sensitive to minor variations in distributions/intensities (i.e., power of the test), yet insensitive to noises when the underlying distributions are the same (i.e., type I error). Therefore, the simulations in this section are designed to evaluate the performance of our proposed test w.r.t. controlling type I error and enhancing the power of the test.

Our proposed test scheme consists of three components: adoption of the polar system, transformation through one-dimensional depth function, and a two-dimensional KS test. We want to test the role and performance of each component. Also, we would like to compare our proposed test scheme with other classical depth-based tests or traditional "goodness-of-fit" tests. For comparison, we will perform Liu–Singh's rank sum test, the two-dimensional KS test, and our proposed test on the same simulated datasets, on both the original Cartesian system and polar system.

The detailed test schemes are listed as follows:

- *Method 1*: Liu–Singh's rank sum test, as introduced in Section 3.3. To be consistent, we adopt a two-dimensional Tukey's depth (equation (3.2)) as the transformation depth function. Then by Corollary 2 and Theorem 1 in Zuo and He (2006), assume H is continuous; the following asymptotic result holds for Tukey's depth:

$$(\sigma_{GH}^2/m + \sigma_{HG}^2/n)^{-1/2}(Q(H_m, G_n) - Q(H, G)) \xrightarrow{d} N(0, 1) \quad \text{as } \min(m, n) \to \infty,$$

where under the null hypothesis $H = G$, $Q(H, G) = 1/2$ and $\sigma_{GH}^2 = \sigma_{HG}^2 = 1/12$. This implies that the p-value can be estimated by $N(0, 1)$ according to the asymptotic result above.

- *Method 2*: A variant of Method 1 by adopting the polar coordinate. We introduce an additional transformation layer that maps data from Cartesian coordinates to polar coordinates before employing the two-dimensional Tukey's depth. The purpose of this method is to isolate the impact of the polar system and evaluate it separately.

- *Method 3*: The popular two-dimensional KS test. Since we have elaborately described the two-dimensional KS test in Section 3.5, we will omit the test details here.

- *Method 4*: The two-dimensional KS test on the polar coordinate system. Similar to Method 2, we perform an two-dimensional KS test on the polar coordinate.

- *Method 5*: Same as our proposed test scheme in Section 3.5, except that we omit the first step and directly work on the original Cartesian coordinates. That is, this method contains two steps: first, estimate the pooled depth of location (x, y), and then perform a two-dimensional KS test on pooled depth.

- *Method 6*: Our proposed method as described in Section 3.5.

4.1. *Performance on controlling type I error.* In this section we randomly generate two groups of spatial point process realizations, denote as $\mathbb{G}_1 = \{\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_{100}\}$ and $\mathbb{G}_2 = \{\mathbf{S}_1', \mathbf{S}_2', \ldots, \mathbf{S}_{100}'\}$. Each group contains 100 realizations; realizations in $\mathbb{G}_1$ follow a Poisson process with intensity map $\lambda_1(s)$, and realizations in $\mathbb{G}_2$ follow a Poisson process with intensity map $\lambda_2(s)$. The intensity maps are the discretized versions of intensity function, which can be denoted as a 112×112 grid. For each grid the associated number stands for the total intensity of the grid. The log intensity maps of λ_1 (Design 1) and λ_2 (Design 2) are shown as Figure 2.

(−6.908, −2.303]	(−2.303, −1.609]	(−1.609, −1.204]	(−1.204, −0.916]	(−0.916, −0.693]	(−0.693, 0.000]	(0.000, 0.693]	(0.693, 1.386]	(1.386, 1.792]
(1.792, 2.079]	(2.079, 2.303]	(2.303, 2.485]	(2.485, 2.639]	(2.639, 2.773]	(2.773, 2.890]	(2.890, 2.996]	(2.996, 3.219]	
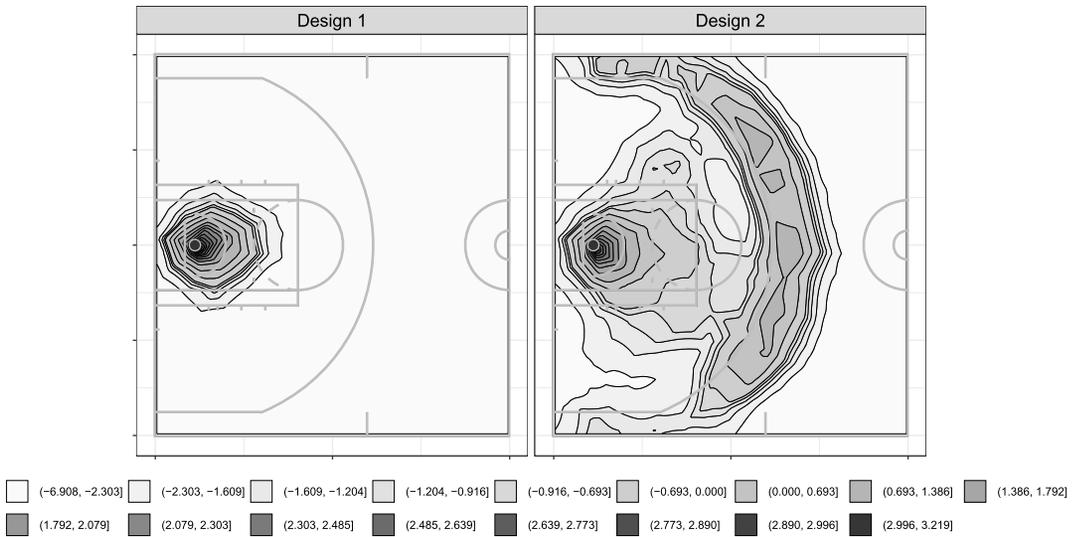
FIG. 2. *Contour Plots for Intensity Maps of Two Simulation Designs.*

Those two intensity maps, which are calculated by R-package *lgcp* (Taylor et al. (2013)), are picked from two representative NBA players (Steve Adams and James Harden). The first pattern corresponds to players most of whose shots are in the painted area, and another pattern represents shooting locations at the three-point line and inside the painted area. Under our simulation settings, the locations of field goal attempts are generated via the R-package *spatstat* (Baddeley and Turner (2005)) based on intensity maps shown in Figure 2. As we can see from the above intensity maps, the intensity surfaces of Group 1 and Group 2 are different visually.

As the first step, we test to see if all test schemes could differentiate realizations in $\mathbb{G}_1$ from $\mathbb{G}_2$. We randomly select one realization from Group 1 and Group 2 separately (denoted as $\mathbf{S}$ and $\mathbf{S}'$) without replacement for 100 times and form 100 hypothesis tests with the null hypothesis

$$H_0 : \lambda(s; \mathbf{S}) = \lambda(s; \mathbf{S}'),$$

or equivalently, $H_0 : F(s; \mathbf{S}) = F(s; \mathbf{S}')$, where $F(s; \mathbf{S})$ and $F(s; \mathbf{S}')$ are the corresponding distribution function of $\mathbf{S}$ and $\mathbf{S}'$. Here we use the conventional significant level 5% as criterion, and all 100 tests are rejected for all test schemes (Method 1–Method 6). This result is expected since the two intensity functions are very distinct. Then we focus on testing performances regarding controlling type I error.

To estimate the type I error of different test schemes, we randomly select *two* realizations from each group for 485 times, then test against each other. The null hypothesis is $H_0 : F(s; \mathbf{S}) = F(s; \mathbf{S}')$, where $\mathbf{S}$ and $\mathbf{S}'$ sampled from the same group. We still use 5% as the significant level and count the rejection number of different test schemes. The results are shown in Table 1.

From results shown in Table 1, we see that our proposed test (Method 6) is closest to the 5% rejection boundary for all 970 tests, which achieved the lowest type I error rate. We believe this is due to the combined effect of the two-dimensional KS test, polar system and adoption of depth functions. If we compare the depth-based Liu–Singh's tests (Methods 1 and 2) with the two-dimensional KS tests (Methods 3 and 4), the two-dimensional KS tests perform better than Liu–Singh's tests, especially for Design 2. The partial reason is that Liu–Singh's test focuses more on data ranks and does not utilize the complete information of data.

TABLE 1
Number of rejected tests for each group

Method	Design 1	Design 2	Sum
Liu–Singh's test on Cartesian coordinate (Method 1)	51	86	137
Liu–Singh's test on Polar coordinate (Method 2)	43	70	113
Two-dimensional KS test on Cartesian coordinate (Method 3)	52	55	107
Two-dimensional KS test on Polar coordinate (Method 4)	36	52	88
Our proposed test on Cartesian coordinate (Method 5)	**23**	45	68
Our proposed test on Polar coordinate (Method 6)	**25**	**25**	**50**

Note: There are 485 tests performed for each design based on methods we mentioned above. In total, for each method we performed 970 tests

Similarly, comparing the tests on the Cartesian system (Methods 1, 3, and 5) with tests on the polar coordinate (Methods 2, 4, and 6) in Table 1, we notice that the usage of the polar system improves the overall test performance on controlling type I error.

In addition, we compare the two-dimensional KS tests (Methods 3 and 4) with our proposed tests (Methods 5 and 6). Although all these tests are based on the two-dimensional KS test, one can see that the type I errors are lower after adopting depth transformation (especially for Group 2). It further suggests that our proposed test outperformed other tests we considered in controlling type I error.

4.2. *Performance on power of test.* In this section we emphasize on examining the sensitivity of our proposed test schemes with respect to small perturbations, that is, the sensitivity of power with respect to the magnitude of shifts in distribution. This section aims to examine if the adoption of the polar system and depth function will improve the power of the test. Hence, we will mainly consider three different KS test schemes for comparison:

- The two-dimensional KS test (Method 3).
- The two-dimensional KS test on the polar coordinate system (Method 4).
- Our proposed depth-based test scheme (Method 6).

First, for each realization $\mathbf{S}_i = (s_{i,1}, \ldots, s_{i,n_i})$ and $\mathbf{S}'_i = (s'_{i,1}, \ldots, s'_{i,m_i})$, $i = 1, \ldots, 100$ in $\mathbb{G}_1$ and $\mathbb{G}_2$ simulated previously, we add a random location-shift noise $\xi \sim$ Gaussian$([0, 0], 0.25rI)$ on each location. That is,

$$\mathbf{SS}_i = (s_{i,j} + \xi, j = 1, \ldots, n_i),$$
$$\mathbf{SS}'_i = (s'_{i,j} + \xi, j = 1, \ldots, m_i),$$

where n_i and m_i are the cardinalities of realization $\mathbf{S}_i$ and $\mathbf{S}'_i$ and I is the two-dimensional identity matrix. Here we introduce a scale parameter r to adjust the magnitude of the noise; as r gets larger, the shifted distribution is more separated from the original distribution.

Note that the noise ξ will only change the conditional distribution of location $F(\mathbf{S}_i || \mathbf{S}_i | = n_i)$, and the number of points (cardinality) n_i remains untouched. For each group we can perform 100 hypothesis tests with null hypothesis

$$H_0 : F(\mathbf{S}_i) = F(\mathbf{SS}_i) \quad \text{for } i \in [1, \ldots, 100],$$

or

$$H_0 : F(\mathbf{S}'_i) = F(\mathbf{SS}'_i) \quad \text{for } i \in [1, \ldots, 100].$$

By setting scale parameter r to different values in $[0, 1, 2, 3, 4, 5, 6, 7]$ and calculating the average p-value of 100 tests, we have the result in Figure 3. From Figure 3 we see that our

(a) Number of rejected tests

(b) Average P-value
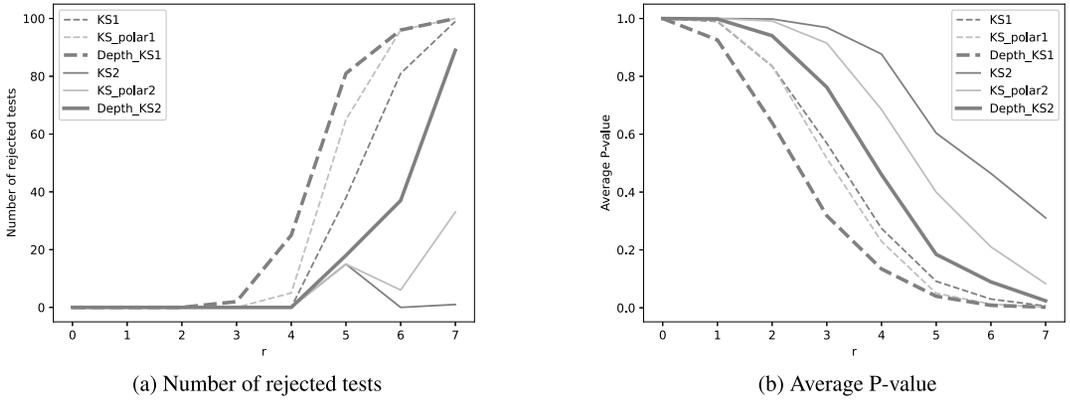
FIG. 3. *"KS1" is the two-dimensional KS test on Group 1 (Method 3), "KS-polar1" is the two-dimensional KS test based on polar coordinate for Group 1 (Method 4), and "Depth-KS1" is the proposed test scheme for Group 1. Similar notation applies to Group 2: (a). Plot of number of rejected tests ($\alpha = 0.05$) for Group 1 (dash line) and Group 2 (solid line) based on the proposed test scheme, KS test, and KS test on polar coordinate. (b). Corresponding average P-value plot.*

proposed test is more sensitive to small Gaussian noises than the traditional KS test or KS test on the polar system. The average P-values of our proposed KS test scheme decrease more rapidly than the other two tests, suggesting that the adoption of polar coordinate and depth function could improve the test power.

In addition, we add tiny variations directly on the intensity map, regenerate 100 realizations, and then test against the original realizations. As we mentioned earlier, the intensity map can be viewed as a 112×112 grid, and the number in each grid stands for the total intensity of the corresponding area. For each grid we add a noise term $\epsilon \sim F(r)$, where $F(r)$ is a distribution with a parameter r to adjust the magnitude of noise.

The distributions of the noise term we consider are:

- Half-Gaussian distribution: $\epsilon \sim |N(0, r)|$,
- Lognormal distribution $\log(\epsilon) \sim N(0, r)$.

By adjusting r from $[0.025, 0.05, 0.075, 0.125, 0.15]$ for the half-Gaussian noise and $[0.1, 0.5, 0.75, 1, 1.25]$ for the lognormal noise, we get the test results shown in Table 2.

TABLE 2
Number of rejected tests out of 100 tests

Type of noise	r	KS (method 3)	KS polar(method 4)	The proposed test (method 6)
$\epsilon \sim |N|$	0.025	59|9	73|12	57|15
	0.05	100|13	100|27	100|29
	0.075	100|23	100|37	100|42
	0.125	100|72	100|94	100|91
	0.15	100|93	100|100	100|99
$\log(\epsilon) \sim N$	0.1	6|5	2|3	12|5
	0.5	8|8	8|4	13|8
	0.75	23|11	23|15	25|20
	1	46|34	42|34	57|41
	1.25	96|94	99|99	99|100

Note: The test results are based on criteria $\alpha = 0.05$. The number in the front of separate '|' is the number of rejected tests for Group 1, the number after '|' is the number of rejection for Group 2, and each group contain 100 tests

TABLE 3
Number of rejected tests out of 100 *location-shift tests*

Shift by c pixel	KS (method 3)	KS polar (method 4)	The proposed test (method 6)
1	24	30	36
2	76	82	85
3	100	100	100
4	100	100	100
5	100	100	100

Note: The test results are based on criteria $\alpha = 0.05$

Finally, we shift the intensity map by c pixel horizontally, regenerate 100 realizations according to the shifted intensity map, and test against the original realizations. Since the intensity map of Group 1 (λ_1) is too concentrated and sensitive to small location shifts, all tests are rejected after one pixel shift. Here we discard tests results for λ_1 and only test λ_2 vs. λ_2^r (the shifted intensity map by r pixel). The results are shown in Table 3.

From Table 2, one can see that for half-Gaussian noises, our proposed test scheme is more sensitive to small noises in intensity maps, especially in Group 2 (λ_2). This indicates our proposed test has the ability to identifying small differences among underlying intensity maps. Although the test results for Method 4 and Method 6 are similar, the slight improvement suggests the adoption of depth transformation is suitable when dealing with half-Gaussian noises. The test results for lognormal noises further support our supposition for which the improvement after adopting depth transformation is more evident by comparing Method 6 with Method 3 and 4.

In contrast in Table 3, the test results suggest both polar and depth transformations will improve the sensitivity of the test. By comparing Method 3 and Method 4, the only difference is whether the KS test on polar coordinates are used, and the enhancement is solely based on the implementation of the polar system. On the other hand, the improvement in sensitivity from Method 4 to Method 6 attributes to the usage of depth transformation, which is the only difference between these two methods.

5. Professional basketball data analysis.

5.1. *Data overview.*   Our data consists of both made and missed field goal attempt locations from the offensive half court of games in the 2017–2018 National Basketball Association (NBA) regular season. The visualization of selected players is shown in Section 2. In this section we focus on players who have made more than 400 field goal attempts. Also, the rookie players in that season, such as Lonzo Ball and Jayson Tatum, are not considered. In total, we have 191 players who meet the two criteria above in our analysis.

5.2. *Hypothesis testing for NBA players.*   Our goal is to discriminate the difference between made process and missed process for NBA players. This section also includes the KS test and KS test with polar coordinates in contrast to our proposed depth-based test for all 191 players. Our testing procedure rejects the null hypothesis for 147 players under $\alpha = 0.05$ and fails to reject the null hypothesis for 44 players. However, the other two methods fail to reject the null hypothesis for just 13 players. The details of players' names for the three testing procedures are shown in the Supplementary Material (Qi, Hu and Wu (2024)). In order to have a closer look at our testing results, we pick several representative players to visualize their made and missed locations in Figure 4 and Figure 5.
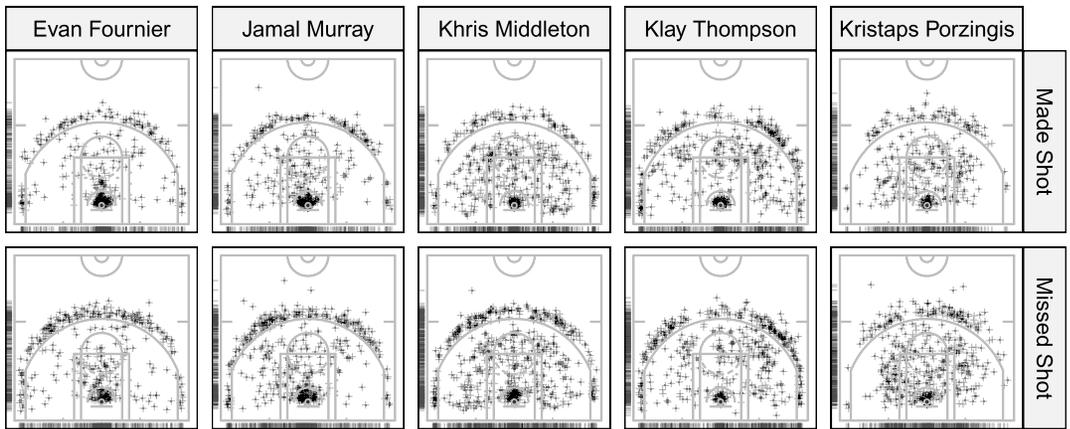
FIG. 4. *Made (Top) and missed (Bottom) shot locations of five example players (Evan Fournier, Jamal Murray, Khris Middleton, Klay Thompson, Kristaps Porzingis) whose missed and made shot patterns are not significantly different (with α = 0.05) by our proposed test but are different by the other two methods (KS test and KS test after Polar transition) in our comparison.*

Figure 4 contains five players. Our testing procedure does not reject the null hypothesis, but the other two testing procedures reject the null hypothesis for these five players. From this figure we see that the missed and made shot patterns for these players are very similar.

Figure 5 contains three players. Our testing procedure rejects the null hypothesis, but the other two testing procedures do not for these three players. Clearly, the missed and made shot patterns of these three players are quite different: Hood has more missed shots on top
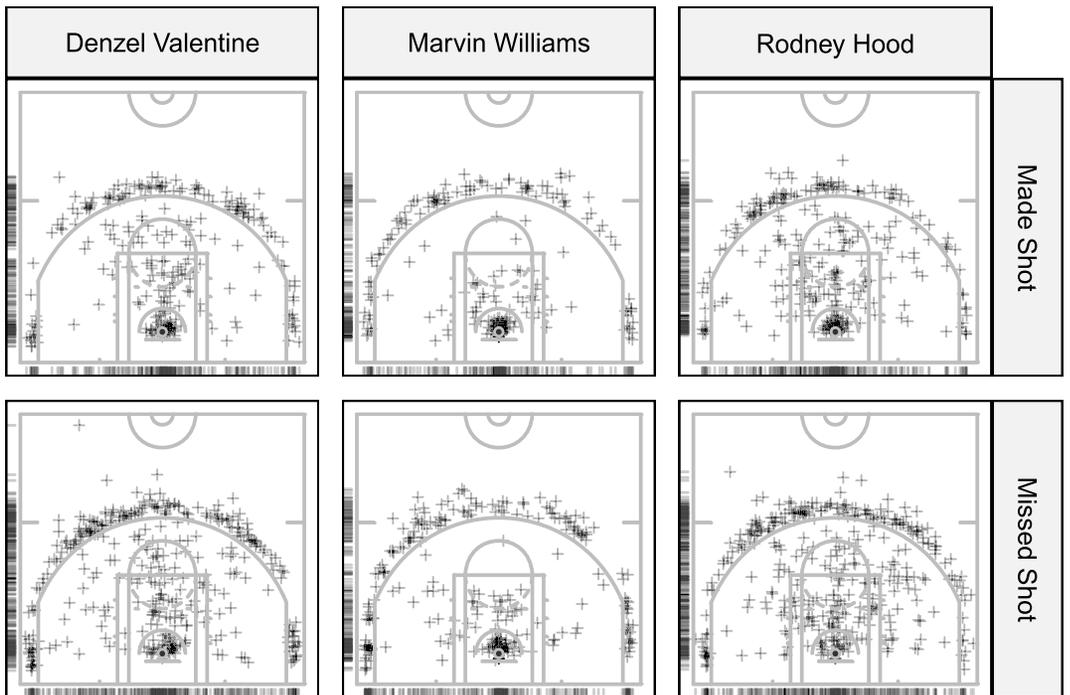
FIG. 5. *Made (Top) and missed (Bottom) shot locations of three example players (Denzel Valentine, Marvin Williams, Rodney Hood) whose missed and made shot patterns are significantly different (with α = 0.05) by our proposed test but are not different by the other two methods (KS test and KS test after Polar transition) in our comparison.*
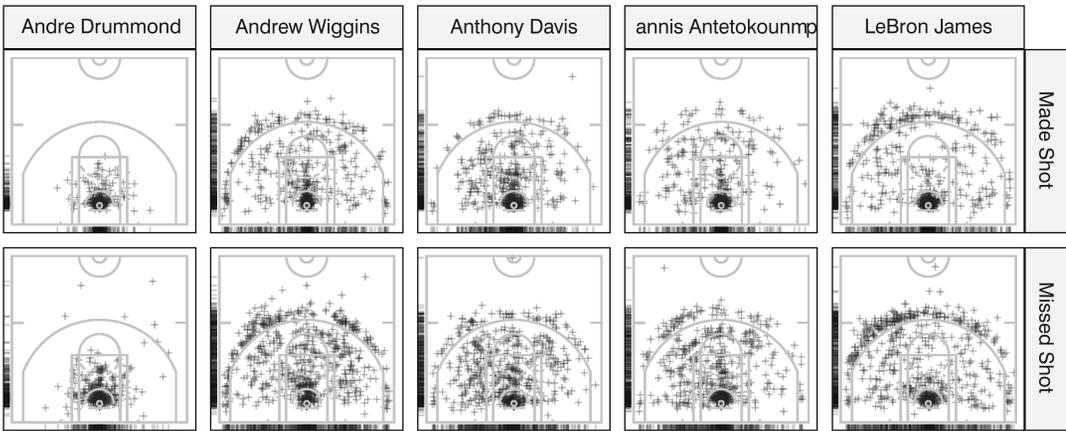
FIG. 6. *Top five Players with smallest P-values based on our proposed test framework. Each column indicates the Made pattern (Top) and Missed pattern (Bottom) of the respective player.*

3-pointers; Valentine has different patterns on long 2-pointers; Williams has more missed on the right and left top 3-pointers.

In addition, we have the top five players (Giannis Antetokounmpo, LeBron James, Anthony Davis, Andrew Wiggins, Andre Drummond) with the smallest *p*-value in Figure 6. We see that the listed five players' made processes are much different from their missed processes. For example, Andre Drummond does not have successful FTAs out of paint region, and Antetokounmpo has many more missed shots on right-wing 3-pointers and long 2-pointers.

Furthermore, we pull the top five players with the highest field goal attempts of each position (C, PF, SF, SG, and PG). The three testing procedures have consistent results, rejecting the null hypothesis at $\alpha = 0.05$ for all five positions. The results are shown in the Supplementary Materials.

In a brief summary, for most players (nearly 77% players in our analysis), the made process and the missed process are different based on our testing procedures. This means most players have their hot zones and cold zones on the court. The best defense strategy is to force those players to shoot more on their cold zones than hot zones. For the offense side, the coach needs to design offensive tactics, such as screening, to make their players have more comfortable shots during the game.

5.3. *Players classification.* Naturally, a significant test and *P*-value can be applied to classify point processes, as it measures how extremely different one point process is from another under the null hypothesis that the two processes are the same. In this section we will apply our proposed test framework to classify different players into different groups. The test steps can be generalized as follows:

1. Randomly choose one player, and use his missed and made pattern as the benchmark.
2. Calculate the pairedwise *P*-value of made and missed pattern of the benchmark player with all other players, denote as $P_{\text{missed}}(player1, player2)$ and $P_{\text{made}}(player1, player2)$.
3. Compare the calculated *P*-values with a predefined significant level(threshold), for example, 0.025. If both $P_{\text{missed}}(player1, player2)$ and $P_{\text{made}}(player1, player2)$ are above a certain threshold, then classify it to the benchmark player's group. Since we are doing multiple testing, that is, comparing made and missed simultaneously, we applied the Bonferroni Correction to adjust the significance level $\alpha_{\text{adj}} = 0.05/2 = 0.025$.
4. Randomly choose another player as threshold, and repeat steps 1–3.

TABLE 4
Classification results of five selected players

Benchmark player	Players in same group
DeAndre Jordan	N/A
DeMar DeRozan	Rondae HollisJefferson, Cory Joseph
Giannis Antetokounmpo	N/A
LeBron James	Eric Bledsoe, Aaron Gordon, Jeff Green, James Johnson, Mario Hezonja, DAngelo Russell, Andrew Harrison
Stephen Curry	Jaylen Brown, Terry Rozier, Wilson Chandler, Trey Lyles, Reggie Bullock, DJ Augustin, Stanley Johnson, Manu Ginobili, Justise Winslow, Lou Williams

Based on the procedures above, the classification results of five selected players are given in Table 4. Based on results shown in Table 4, we see that Giannis Antetokounmpo and DeAndre Jordan are two special players we can not find any players with similar missed and made patterns. In addition, LeBron James's shooting pattern is more like Power Forward.

5.4. *LeBron James vs. Stephen Curry.* Now, we will address the initial question we posed: Do made and missed field goal attempts follow different spatial processes for Stephen Curry? Additionally, we will explore whether LeBron James has a higher number of field goal attempts in the region where he has a higher field goal percentage.

The answer to the first question is straightforward. Based on the KS test, KS test after polar transformation, and our proposed test framework, all indicate that the patterns of made and missed shots for Stephen Curry differ significantly at the 5% significance level. This suggests that he has specific shooting areas where he achieves a higher success rate compared to misses. However, at the 1% significance level, our proposed test (as well as the KS test) would reject the null hypothesis, implying that his preferred shooting spots may not be as strong or as numerous. Considering the injury he sustained this season, this outcome is somewhat expected.

Regarding the second question, all the tests in our comparison indicate that LeBron James' made patterns and missed patterns are statistically different. This suggests that, in certain areas, he has a higher success rate in made shots compared to misses or vice versa. When we examine the made and missed patterns in Figure 6, it becomes evident that he attempts more field goals in the upper left 3-point line and has a higher success rate in that area as well. However, around the upper right 3-point line and the free-throw area, his success rate is less satisfactory, resulting in numerous missed attempts.

Furthermore, based on the classification results in Table 4, we can offer some suggestions to coaches who aim to train specific types of players or compete against them. For example, we observed that LeBron James has similar shooting and missed shot patterns to Eric Bledsoe, Aaron Gordon, Jeff Green, James Johnson, Mario Hezonja, D'Angelo Russell, and Andrew Harrison. From a coaching perspective, this suggests that those players could benefit from a similar training plan as LeBron James, who was recognized as one of the most successful players in the 2017–2018 NBA regular season and has had the highest points per game (PPG) in the past 10 years. Additionally, if there is an effective defensive plan against players like LeBron James, it could apply to other players as well.

6. Discussion. In this paper we propose a depth-based testing procedure for discriminating two spatial point processes. Building upon the polar coordinate system, we develop a two-dimensional KS test on one-dimensional depth. Unlike several benchmark test procedures, our proposed methods apply a multivariate KS test with a one-dimensional depth

function. The depth function then acts like a Box–Cox transformation, transforming data into a different shape rather than reducing dimensions. By comparing with several benchmark testing procedures, the numerical results showed that the proposed method can provide a robust testing procedure under both null and alternative hypotheses. In the analysis of the NBA shot charts data, our proposed method reveals several interesting findings for different NBA players and hence provides more *objective* and principled analysis for shooting patterns of the players in NBA. Compared with traditional point process methods for shot charts of NBA players, our proposed testing procedures successfully discriminate the differences between made and missed shots of NBA players. For some top players, such as Anthony Davis, LeBron James, and Giannis Antetokounmpo, they have smallest p-values. The comfortable zones for those players are very obvious. In addition, we can see that Giannis Antetokounmpo is a very distinct player in NBA. We cannot find any players who have similar missed and made pattern with him.

Unlike previous studies that have focused on goal attempt frequency and accuracy of specific players (e.g., Reich et al. (2006)), our test framework aims to determine whether two shot patterns differ from each other while also capturing the spatial randomness of field goal attempts. Instead of using complex models to analyze the relationships between defenders and offenders or estimating shoot intensity maps (e.g., Miller et al. (2014)), we have developed a simple yet powerful statistical test framework to differentiate spatial patterns based on the center-outward tendency of players' field-goal attempts.

Our framework significantly differs from the methodologies discussed in the Introduction. It can effectively summarize the overall differences in shot patterns, making it particularly valuable from coaching and recruiting perspectives. By using this approach, teams can identify and hire players with desired shot patterns, and its potential in classification allows recruiters to quickly identify groups of players with similar shot patterns.

However, it also comes with certain limitations. First, our proposed test framework primarily aims to determine if one spatial pattern significantly differs from another. But without additional information, such as performance-related summary statistics of players, our model lacks the ability to identify which shooting pattern is superior. Furthermore, in this research we made the assumption that the shot locations adhere to a Poisson distribution. This assumption stems from the belief that shot locations are independent of previous locations, which unfortunately prevents us from capturing the temporal relationships between shot positions of teammates or defenders, as explored in other studies on spatiotemporal processes (e.g., Cervone et al. (2016) proposed a framework based on Gaussian processes to model the evolution of a basketball possession). Additionally, our analysis using shot chart data is incapable of accounting for complicated time dependence, in contrast to studies such as Franks et al. (2015) and Cervone et al. (2016) that utilize real-time tracking data. Therefore, our model is unable to comprehend the influence of defenders on the shot effectiveness and correlations between successful and missed shot attempts.

Some topics beyond the scope of this paper are worth further investigation. First, an extension of the proposed testing framework to more than two-dimensional spatial point processes will provide broader applicability. Second, developing a depth-based testing procedure for pair correlation functions of point processes is another exciting topic. Moreover, false discovery rate control in multiple testing problems represents an interesting direction for future work. From the application point of view, having a closer look at court locations with the most significant effects will provide more data-driven information for NBA teams.

SUPPLEMENTARY MATERIAL

Testing results (DOI: [10.1214/24-AOAS1899SUPP](https://doi.org/10.1214/24-AOAS1899SUPP); .pdf). Details of players names and results.

REFERENCES

BADDELEY, A. (2017). Local composite likelihood for spatial point processes. *Spat. Stat.* **22** 261–295. MR3732850 https://doi.org/10.1016/j.spasta.2017.03.001

BADDELEY, A. and TURNER, R. (2005). spatstat: An R package for analyzing spatial point patterns. *J. Stat. Softw.* **12** 1–42.

BARNETT, V. (1976). The ordering of multivariate data. *J. Roy. Statist. Soc. Ser. A* **139** 318–355. MR0445726 https://doi.org/10.2307/2344839

BROWN, B. M., HETTMANSPERGER, T. P., NYBLOM, J. and OJA, H. (1992). On certain bivariate sign tests and medians. *J. Amer. Statist. Assoc.* **87** 127–135. MR1158630

CERVONE, D., D'AMOUR, A., BORNN, L. and GOLDSBERRY, K. (2016). A multiresolution stochastic process model for predicting basketball possession outcomes. *J. Amer. Statist. Assoc.* **111** 585–599. MR3538688 https://doi.org/10.1080/01621459.2016.1141685

CHRISTMANN, A. (2002). Classification based on the support vector machine and on regression depth. In *Statistical Data Analysis Based on the L_1-Norm and Related Methods* (Neuchâtel, 2002). *Stat. Ind. Technol.* 341–352. Birkhäuser, Basel. MR2001327

DIGGLE, P. J. (2014). *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*, 3rd ed. *Monographs on Statistics and Applied Probability* **128**. CRC Press, Boca Raton, FL. MR3113855

DONOHO, D. L. and GASKO, M. (1992). Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *Ann. Statist.* **20** 1803–1827. MR1193313 https://doi.org/10.1214/aos/1176348890

ERČULJ, F. and ŠTRUMBELJ, E. (2015). Basketball shot types and shot success in different levels of competitive basketball. *PLoS ONE* **10** e0128885.

FASANO, G. and FRANCESCHINI, A. (1987). A multidimensional version of the Kolmogorov–Smirnov test. *Mon. Not. R. Astron. Soc.* **225** 155–170.

FRAIMAN, R., LIU, R. Y. and MELOCHE, J. (1997). Multivariate density estimation by probing depth. In *L_1-Statistical Procedures and Related Topics* (Neuchâtel, 1997). *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **31** 415–430. IMS, Hayward, CA. MR1833602 https://doi.org/10.1214/lnms/1215454155

FRANKS, A., MILLER, A., BORNN, L. and GOLDSBERRY, K. (2015). Characterizing the spatial structure of defensive skill in professional basketball. *Ann. Appl. Stat.* **9** 94–121. MR3341109 https://doi.org/10.1214/14-AOAS799

GENG, J., SHI, W. and HU, G. (2021). Bayesian nonparametric nonhomogeneous Poisson process with applications to USGS earthquake data. *Spat. Stat.* **41** Paper No. 100495, 26. MR4214392 https://doi.org/10.1016/j.spasta.2021.100495

GÓMEZ RUANO, M. Á., ALARCÓN LÓPEZ, F. and ORTEGA TORO, E. (2015). Analysis of shooting effectiveness in elite basketball according to match status. *Rev. Psicol. Deporte* **24** 0037–41.

GUAN, Y. (2006). A composite likelihood approach in fitting spatial point process models. *J. Amer. Statist. Assoc.* **101** 1502–1512. MR2279475 https://doi.org/10.1198/016214506000000500

GUAN, Y. and SHEN, Y. (2010). A weighted estimating equation approach for inhomogeneous spatial point processes. *Biometrika* **97** 867–880. MR2746157 https://doi.org/10.1093/biomet/asq043

GUDMUNDSSON, J. and HORTON, M. (2017). Spatio-temporal analysis of team sports. *ACM Comput. Surv.* **50** 1–34.

HANNU, O. and JUKKA, N. (1989). Bivariate sign tests. *J. Amer. Statist. Assoc.* **84** 249–259. MR0999686

HU, G., YANG, H.-C. and XUE, Y. (2021). Bayesian group learning for shot selection of professional basketball players. *Stat* **10** Paper No. e324, 12. MR4276018 https://doi.org/10.1002/sta4.324

HU, G., YANG, H.-C., XUE, Y. and DEY, D. K. (2023). Zero-inflated Poisson model with clustered regression coefficients: Application to heterogeneity learning of field goal attempts of professional basketball players. *Canad. J. Statist.* **51** 157–172. MR4551819 https://doi.org/10.1002/cjs.11684

ILLIAN, J., PENTTINEN, A., STOYAN, H. and STOYAN, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*. *Statistics in Practice*. Wiley, Chichester. MR2384630

JIAO, J., HU, G. and YAN, J. (2021a). A Bayesian marked spatial point processes model for basketball shot chart. *J. Quant. Anal. Sports* **17** 77–90. https://doi.org/10.1515/jqas-2019-0106

JIAO, J., HU, G. and YAN, J. (2021b). Heterogeneity pursuit for spatial point pattern with application to tree locations: A Bayesian semiparametric recourse. *Environmetrics* **32** Paper No. e2694, 15. MR4323206 https://doi.org/10.1002/env.2694

JUSTEL, A., PEÑA, D. and ZAMAR, R. (1997). A multivariate Kolmogorov–Smirnov test of goodness of fit. *Statist. Probab. Lett.* **35** 251–259. MR1484961 https://doi.org/10.1016/S0167-7152(97)00020-5

KUBATKO, J., OLIVER, D., PELTON, K. and ROSENBAUM, D. T. (2007). A starting point for analyzing basketball statistics. *J. Quant. Anal. Sports* **3** Art. 1, 24. MR2326663 https://doi.org/10.2202/1559-0410.1070

LI, J. and LIU, R. Y. (2004). New nonparametric tests of multivariate locations and scales using data depth. *Statist. Sci.* **19** 686–696. MR2185590 https://doi.org/10.1214/088342304000000594

LIU, R. Y. (1990). On a notion of data depth based on random simplices. *Ann. Statist.* **18** 405–414. MR1041400 https://doi.org/10.1214/aos/1176347507

LIU, R. Y., PARELIUS, J. M. and SINGH, K. (1999). Multivariate analysis by data depth: Descriptive statistics, graphics and inference. *Ann. Statist.* **27** 783–858. With discussion and a rejoinder by Liu and Singh. MR1724033 https://doi.org/10.1214/aos/1018031260

LIU, R. Y. and SINGH, K. (1993). A quality index based on data depth and multivariate rank tests. *J. Amer. Statist. Assoc.* **88** 252–260. MR1212489

LIU, S. and WU, W. (2017). Generalized Mahalanobis depth in point process and its application in neural coding. *Ann. Appl. Stat.* **11** 992–1010. MR3693555 https://doi.org/10.1214/17-AOAS1030

MILLER, A., BORNN, L., ADAMS, R. and GOLDSBERRY, K. (2014). Factorized point process intensities: A spatial analysis of professional basketball. In *International Conference on Machine Learning* 235–243.

MOHD RAZALI, N. and YAP, B. (2011). Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson–Darling tests. *J. Stat. Model. Analytics* **2**.

OJA, H. (1983). Descriptive statistics for multivariate distributions. *Statist. Probab. Lett.* **1** 327–332. MR0721446 https://doi.org/10.1016/0167-7152(83)90054-8

PEACOCK, J. A. (1983). Two-dimensional goodness-of-fit testing in astronomy. *Mon. Not. R. Astron. Soc.* **202** 615–627.

PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A. and VETTERLING, W. T. (1986). *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge Univ. Press, Cambridge. MR0833288

QI, K., CHEN, Y. and WU, W. (2021). Dirichlet depths for point process. *Electron. J. Stat.* **15** 3574–3610. MR4298977 https://doi.org/10.1214/21-ejs1867

QI, K., HU, G. and WU, W. (2024). Supplement to "Are made and missed different? An analysis of field goal attempts of professional basketball players via depth based testing procedure." https://doi.org/10.1214/24-AOAS1899SUPP

REICH, B. J., HODGES, J. S., CARLIN, B. P. and REICH, A. M. (2006). A spatial analysis of basketball shot chart data. *Amer. Statist.* **60** 3–12. MR2224131 https://doi.org/10.1198/000313006X90305

ROUSSEEUW, P. J. and RUTS, I. (1996). Bivariate location depth. *J. R. Stat. Soc., Ser. C* **45** 516–526.

SHORTRIDGE, A., GOLDSBERRY, K. and ADAMS, M. (2014). Creating space to shoot: Quantifying spatial relative field goal efficiency in basketball. *J. Quant. Anal. Sports* **10** 303–313.

TAYLOR, B. M., DAVIES, T. M., ROWLINGSON, B. S. and DIGGLE, P. J. (2013). lgcp: An R package for inference with spatial and spatio-temporal log-Gaussian Cox processes. *J. Stat. Softw.* **52** 1–40.

TUKEY, J. W. (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians (Vancouver, B.C., 1974), Vol. 2* 523–531. Canad. Math. Congr., Montreal, QC. MR0426989

VARDI, Y. and ZHANG, C.-H. (2000). The multivariate L_1-median and associated data depth. *Proc. Natl. Acad. Sci. USA* **97** 1423–1426. MR1740461 https://doi.org/10.1073/pnas.97.4.1423

YIN, F., HU, G. and SHEN, W. (2023). Analysis of professional basketball field goal attempts via a Bayesian matrix clustering approach. *J. Comput. Graph. Statist.* **32** 49–60. MR4552936 https://doi.org/10.1080/10618600.2022.2085727

YIN, F., JIAO, J., YAN, J. and HU, G. (2022). Bayesian nonparametric learning for point processes with spatial homogeneity: A spatial analysis of NBA shot locations. In *International Conference on Machine Learning* 25523–25551. PMLR, Baltimore.

ZHANG, C., XIANG, Y. and SHEN, X. (2012). Some multivariate goodness-of-fit tests based on data depth. *J. Appl. Stat.* **39** 385–397. MR2879827 https://doi.org/10.1080/02664763.2011.594033

ZUO, Y. (2003). Projection-based depth functions and associated medians. *Ann. Statist.* **31** 1460–1490. MR2012822 https://doi.org/10.1214/aos/1065705115

ZUO, Y. and HE, X. (2006). On the limiting distributions of multivariate depth-based rank sum statistics and related tests. *Ann. Statist.* **34** 2879–2896. MR2329471 https://doi.org/10.1214/009053606000000876

ZUO, Y. and SERFLING, R. (2000). General notions of statistical depth function. *Ann. Statist.* **28** 461–482. MR1790005 https://doi.org/10.1214/aos/1016218226