COMPUTE THE P-VALUE OF DISCRETE LOCAL MAXIMA ON GAUSSIAN RANDOM FIELD

A PREPRINT

Tuo Lin*
Division of Biostatistics
Department of Family Medicine and Public Health
University of California, San Diego
San Diego, CA 92122
tulin@ucsd.edu

Armin Schwartzman
Division of Biostatistics and HalicioÄlu Data Science Institute
University of California, San Diego
San Diego, CA 92122
armins@ucsd.edu

September 9, 2020

ABSTRACT

***K*eywords** First keyword · Second keyword · More

1 Introduction

As functional MRI analysis becomes a popular topic these days, how to compute the exceedance probability or P-value of a local maxima in a finite region accurately turns into an ongoing research question. A standard procedure of spatial smoothing is employed by convolving the fMRI signal with a Gaussian function of a specific width. Researchers commonly use the Full Width at Half Maximum to measure the spatial size of the filter. In the literature, currently the best way of computing it is by approximating by results for the continuously sampled smooth random field, but when the FWHM is small, and random field is not smooth, so the adjacent lattice values are nearly independent, the density is not accurate. In this paper, we demonstrate that under such circumstance, we can improve the accuracy by using discrete local maxima method.

- What is the goal? Compute the exceedance probability or the P-value of a local maxima in a smooth random field sampled on the discrete grid.
- What has done before? Cheng and Schwartzman (2015), Schwartzman and Telschow (2019) develops the method to calculate the exact height distribution of local maxima in an isotropic Gaussian random field from discrete grid by random field theory, extends this method to a formal "smoothing and testing of maxima" procedure, and mentions that when the smoothing kernel has FWHM>7, this method is accurate, but typically the smoothing kernel has FWHM$=3$, then this method which works well for continuous field is too conservative to provide an accurate p-value. Worsley (2005) and Taylor et al. (2007) uses local maxima as a means of global maxima, which makes the calculation of height distribution for their paper possible. However, this distribution has a major limitation: it only considers immediate neighbor in its coordinate axis but not

*Use footnote for providing further information about author (webpage, alternative address)—*not* for acknowledging funding agencies.

diagonal ones, which makes the computation of the P-value inaccurate when dimension increases from 1 to 2 and 3.

- What contribution of this paper? We implement the idea of calculating the height distribution of the isotropic Gaussian random field by discrete local maxima method. First, we obtain the local maxima by a Monte Carlo simulation using joint gaussian distribution from a centered voxel and its immediate neighbors, including the diagonal neighbors. These local maxima consists of the empirical distribution of the local maxima in the specific smoothed random field. Then, we provide a smoothed look-up table for different spatial correlation size and height values from our simulation results in 1-D, 2-D and 3-D cases.
- Conclusion. For a random field with large FWHM, we recommend using the continuous method to calculate for the P-values of local maxima, but for a random field with small FWHM, we suggest using our look-up table to calculate for the P-values of local maxima.

2 Methods for calculating the height distribution of local maxima

2.1 Random field theory

This method aims for calculating the distribution for a random field in a continuous domain. Since the continuous domain is an idealized concept, so the method provides an approximation of the discrete case by continuous theory height of local maxima. The explicit density $h(u;\kappa)$ of the 1-D 2-D and 3-D cases are provided in Appendix. κ is a parameter depends on shape of covariance function, and we consider $\kappa = 1$ when correlation is squared exponential.

2.2 Discrete local maxima

DLM:

Disadvantage: not including diagonal.

Advantage: dealing with the boundary issue and FWHM is small.

Extending the idea from Tayler, Worsley and Grosselin (2007), we propose the DLM method to calculate the height distribution of local maxima. The density function has the form:

$$f_{\text{DLM}} = \frac{\prod_{d=1}^{D} Q(\rho_d, u)\phi(u)}{\int_{-\infty}^{\infty} \left(\prod_{d=1}^{D} Q(\rho_d, u) \right) \phi(u) du}, \tag{1}$$

where

$$h_d = \sqrt{\frac{1-\rho_d}{1+\rho_d}}, \quad \alpha_d = \sin^{-1}\left(\sqrt{(1-\rho^2)/2} \right),$$
$$Q(\rho_d, u) = 1 - 2\Phi(h_d u^+) + \frac{1}{\pi} \int_0^{\alpha_d} \exp(-\frac{1}{2} h_d^2 u^2 / \sin^2\theta) d\theta. \tag{2}$$

As we can see, the $Q(\rho_d, u)$ in above density function depends on u and ρ_d. u is a realization of the height variable, which is a pre-determined value that researchers are interested in. ρ_d is the spatial correlation between adjacent voxels along each lattice axis d. We need further assumptions to get the explicit formula for ρ_d, and we will discuss in the following subsection.

2.3 Spatial correlation ρ_d and FWHM

In this paper, we only focus on the Gaussian random field that is isotropic. Then, the spatial correlation between two voxels $x_1, x_2 \in \mathbb{R}^d$ in an isotropic field is

$$\exp(-(x_1 - x_2)' \Lambda (x_1 - x_2)/2), \tag{3}$$

where $\Lambda = diag(1/(2\nu^2), ..., 1/(2\nu^2))$, and ν is the standard deviation of the smoothing kernel.

Based on the assumption above, the correlation ρ_d between adjacent voxels along each lattice axis d is

$$\rho_d = \exp(-\frac{1}{2}(\frac{1}{2\nu^2})) = \exp(-\frac{1}{4\nu^2}), \tag{4}$$

so it does not depend on d under our assumption, and we simply denote it as ρ.

It is also well-known that the relationship between FWHM and standard deviation:

$$\text{FWHM} = 2\sqrt{2ln2}\nu.$$

2.4 Discrete covariance function

The discrete covariance function is provided by:

$$E\left[Z(\mathbf{x})Z(\mathbf{x}+\mathbf{u})\right] = \sum_i \frac{1}{\nu^{2D}}\phi_D(\frac{\left\|\mathbf{x}-s(i)\right\|}{\nu})\phi_D(\frac{\left\|\mathbf{x}+\mathbf{u}-s(i)\right\|}{\nu})$$

2.5 Density plot

The derivation of this density function will be in Appendix. In this subsection, we investigate into the plot of the density function where D = 1, 2, 3.

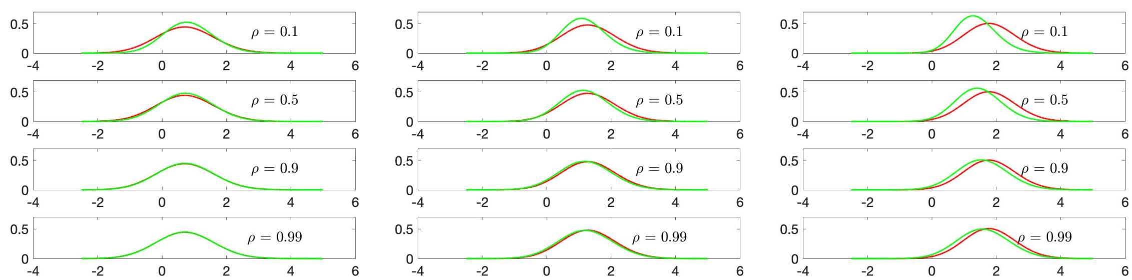

Figure 1: From left to right is density function of DLM in 1D, 2D and 3D cases. The green curve is from DLM method and red curve is from continuous method.

As we can see in Figure 1, in 3D case, even when ρ goes to 1, the DLM and continuous methods are not consistent in density functions. The reason is that the continuous methods include the diagonal points as the neighbors but in DLM the diagonal points are not considered as neighbors.

2.6 Monte Carlo simulation

The goal of this study is to calculate the distribution function of a local maxima in an fMRI image. For simplicity, we assume that our image is a 2D isotropic gaussian random field in this case. The local maxima is defined as the height of one voxel that is greater than its 8 neighbors, including 4 adjacent diagonal voxels. Since in 2D isotropic gaussian random field, the joint distribution of the adjacent diagonal voxels is hard to calculate theoretically, so we begin our calculation with generating an empirical distribution of the local maxima. The empirical distribution is obtained by simulating a $9 \times n$ matrix from the theoretical distribution of bivariate gaussian distribution n times. From this matrix, we select all the columns that have the local maxima and record all local maxima. These local maxima consists the empirical distribution of the local maxima and we can calculate the empirical p-value from it. In order to make this empirical p-value as accurate as possible, the sample size of this empirical local maxima distribution is increased to 100,000.

The bivariate gaussian distribution we mentioned above only depends on the spatial correlation of the field. We readily show that the covariance matrix is a kronecker product of a matrix A in appendix, where

$$A = \begin{pmatrix} 1 & \rho & \rho^4 \\ \rho & 1 & \rho \\ \rho^4 & \rho & 1, \end{pmatrix} \tag{5}$$

and ρ is the spatial correlation.

2.7 Look-up table method

In order to cut the time of running the simulation to get the p-value every time, we decide to use the look-up table to pre-record the simulation results for 100,000 possible local maxima values and ρ. The ρ is ranging from 0.01 to 0.99 with 0.01 in difference. For local maxima, since the empirical CDF is steeper in the middle of its support rather than the ends, so we consider not using uniform sampling but sampling from the empirical distribution to choose the local maxima values. To proceed, we first loop through all the ρ and get roughly 100,000 local maxima for each ρ, and then sample 100,000 from the union of these local maxima. Another issue should be noticed here is that different ρ will give different number of local maxima. The higher the ρ is, the less the number of local maxima, so more samples we need for simulating same numbers of local maxima. Luckily, we find there is some rough linear relationship between ρ and number of local maxima (see Figure 2 below), and we believe that the number of iteration samples in proportion to the number of local maxima. Thus, we fit a linear regression, predict for the number of iteration times that we need and finally get around 100,000 local maxima for all ρ. This is not working well at the extreme values of ρ in 3D case, in which the large ρ value predicts with some negative number of local maxima, so we use the exact fitting for $\rho = 0.96, 0.97, 0.98, 0.99$ instead. After sampling 100,000 values from the union of all the local maxima, we interpolate these 100,000 values through all ρ and obtain a matrix of p-values.

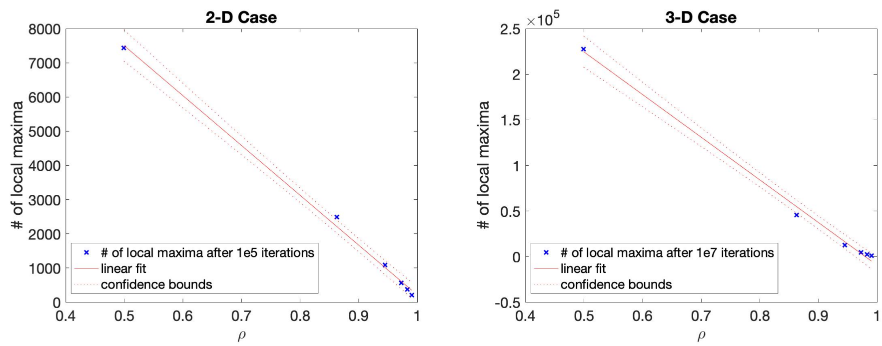

Figure 2: The figure above shows the relationship between ρ and number of local maxima we get after 1e5 iterations for 2D case and 1e7 iterations for 3D case. We use six ρ here: 0.5, 0.86, 0.945, 0.97, 0.98 and 0.99 . The left one is for 2D case and right one is for 3D case.

Smoothing is the last issue after we create a look-up table from all the steps we discussed above since the look-up table we create so far is very noisy across both ρ and u, the local maxima value. Cubic spline smoothing is the strategy that we consider in this case. Here is the procedure of our smoothing:

Step1: Use the Cubic spline smoothing to smooth the matrix across ρ;

Step2: Use the Cubic spline smoothing to smooth the matrix we smoothed in step1 across u.

In doing so, we aim to reduce the violation of monotonicity across x but also retain the smoothness when we smooth the matrix across u and ρ simultaneously. Indeed, we reduce the minimum difference between any two adjacent elements across x from -6.36e-06 to -1.92e-07 through this procedure. Furthermore, the smoothing parameters in Cubic spline smoothing is selected by 5-fold cross validation in each scenario separately.

2.8 Why smoothing good

Below is 50 column samples of the pre-smoothed table and after-smoothed table in 3D case:

We can see in this plot that smoothing across u has the fit that is not visible.

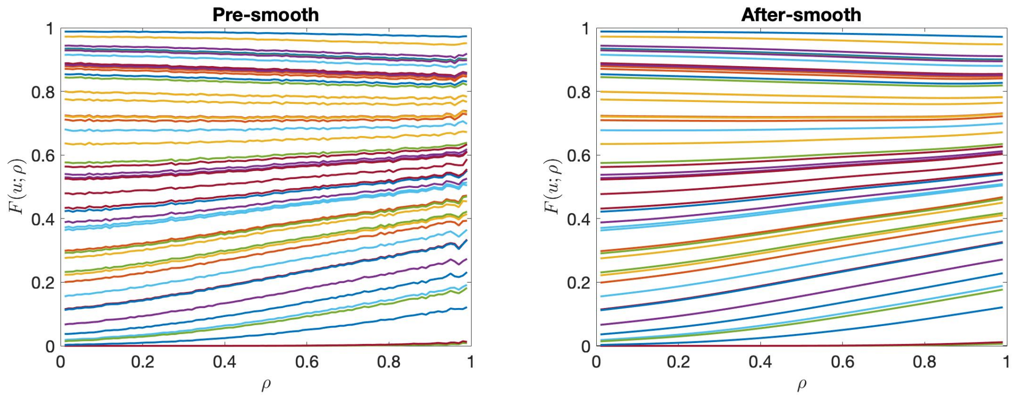

Figure 3: This figure shows selected 50 samples' $F(u; \rho)$ across ρ from Pre-smooth table (left) and after-smooth table (right).

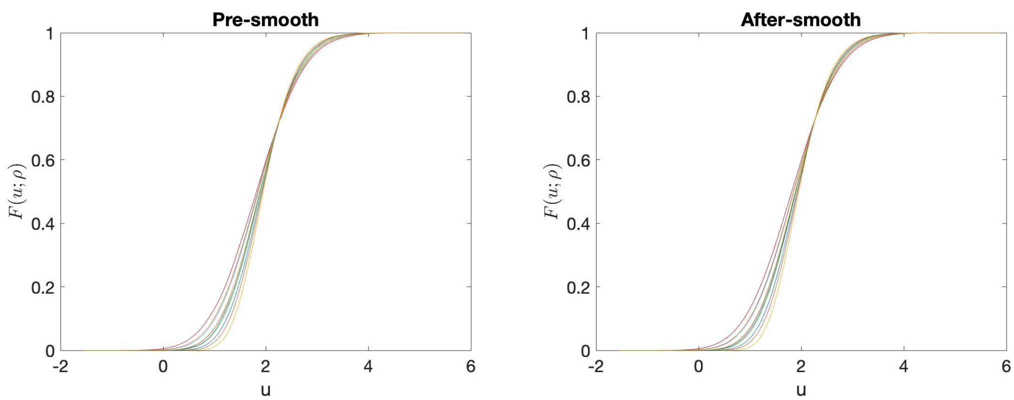

Figure 4: This figure shows selected 10 samples' $F(u; \rho)$ across u from Pre-smooth table (left) and after-smooth table (right). Same color is used for the same sample before and after-smoothing.

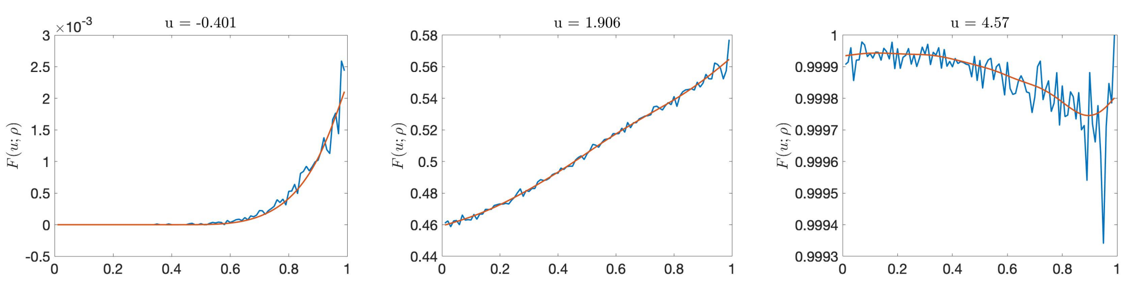

Figure 5: From left to right are the 10, 50000 and 99990 slice of the pre-smoothed CDF table and after smoothed CDF table. Same color is used for the same sample before and after-smoothing.

3 Compare of Methods

3.1 Including the diagonal case

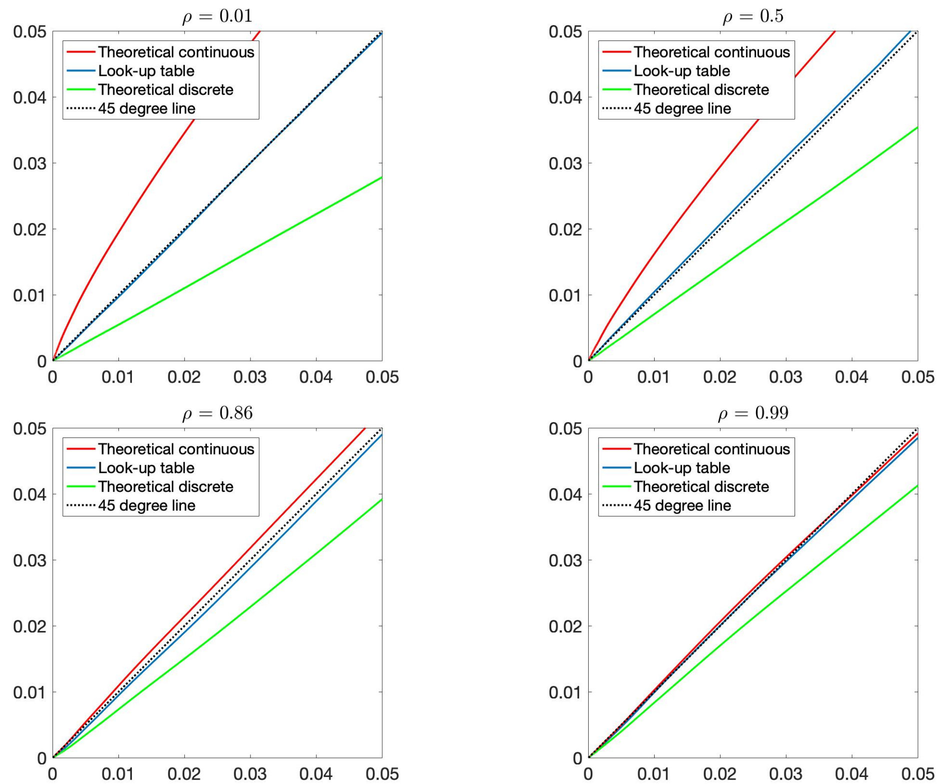

Figure 6: From top left to bottom right: $\rho = 0.01, 0.5, 0.86, 0.99$. The figure is generated based on the comparison among look-up table, simulation directly from the field, DLM and continuous method in 2D case. The x axis is the true p-value, which we generate from the simulated field.

From figure 6 and figure 7, when ρ is small, our look-up table method works better than theoretical p from both GRF theory and DLM method; when ρ is big, the theoretical p from GRF theory overlaps with the 45-degree line, and our look-up table method is a little bit off but acceptable. In addition, we observe that our look-up table works better when ρ increases, the reason is that the multivariate gaussian distribution we used to simulate the data is based on continuous covariance function. We will discuss about using discrete covariance matrix in the simulation in later section.

3.2 Not including the diagonal case

3.3 Applying the discrete covariance function

3.4 comparison based on not using the look-up table

3.5 Discussion of computational issues in using DLM

Since the integration takes a lot of time, so the Riemann integral is used here for integration, and interval length is set to be 0.001 to shorten the time. Moreover, since the density function includes the calculation of double integration and we estimate the computing time will be several days, so we try to use the look-up table here to shorten the computing time. In fact, it only takes around 10 minutes to create a look-up table with different z and rho, where z is from -2 to 6, with interval 0.001 and rho from 0 to 0.999, with interval 0.001. The interpolation only takes no more than one seconds for 500,000 local maximum.

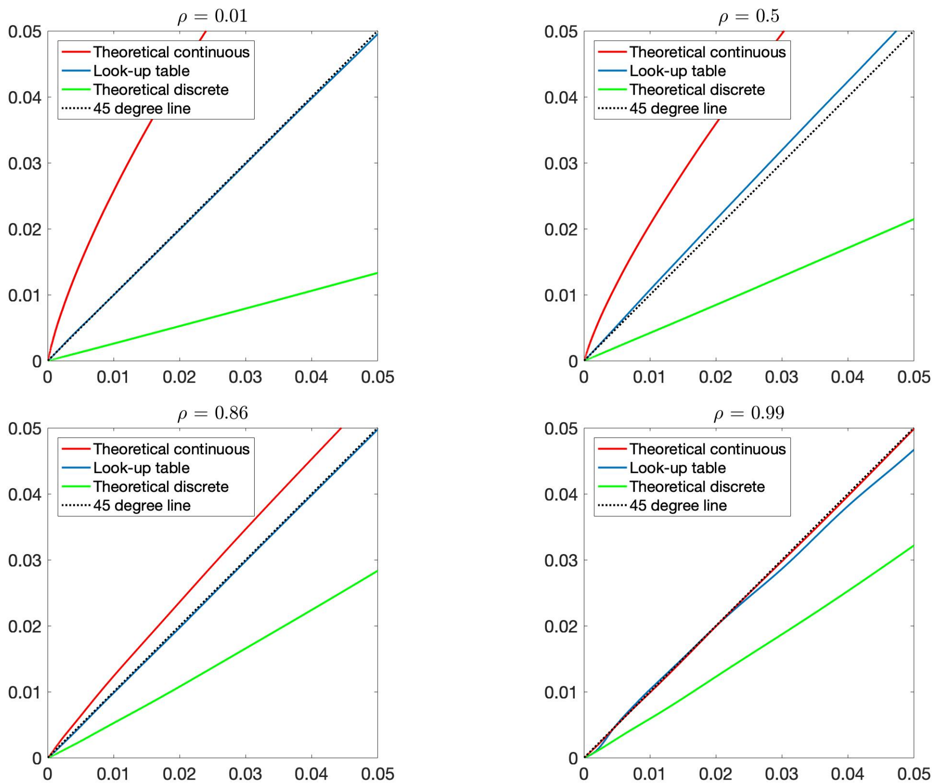

Figure 7: From top left to bottom right: $\rho = 0.01, 0.5, 0.86, 0.99$. The figure is generated based on the comparison among look-up table, simulation directly from the field, DLM and continuous method in 3D case. The x axis is the true p-value, which we generate from the simulated field.

4 Methods extension to the T-field

5 Estimate the covariance locally

6 More General: Stationary and non-stationary case

A Theoretical derivation of $Q(\rho_d, z)$ when ρ_d changes

A.1 Deriving $Q(\rho_d, z)$

Using the polar coordinate strategy, we can derive $\Phi(h_d z) = \frac{1}{2\pi}\int_0^\pi \exp(-\frac{1}{2}h_d^2 z^2/\sin^2\theta)d\theta$, where $h_d = \sqrt{(1-\rho_d)/(1+\rho_d)}$.
Moreover, for $z < 0$, $Q_{xd}(z) = \frac{1}{\pi}\int_0^{\alpha_d} \exp(-\frac{1}{2}h_d^2 z^2/\sin^2\theta)d\theta$, where $\alpha_d = \sin^{-1}\left(\sqrt{(1-\rho_d^2)/2}\right)$.
Combining with the formula of $Q(\rho_d, z)$ when $z > 0$, we get

$$Q(\rho_d, z) = 1 - 2\Phi(h_d z^+) + \frac{1}{\pi}\int_0^{\alpha_d} \exp(-\frac{1}{2}h_d^2 z^2/\sin^2\theta)d\theta. \tag{6}$$

When $z > 0$,

$$\begin{aligned} Q(\rho_d, z) &= 1 - 2 * \frac{1}{2\pi}\int_0^\pi \exp(-\frac{1}{2}h_d^2 z^2/\sin^2\theta)d\theta + \frac{1}{\pi}\int_0^{\alpha_d} \exp(-\frac{1}{2}h_d^2 z^2/\sin^2\theta)d\theta \\ &= 1 - \frac{1}{\pi}\int_{\alpha_d}^\pi \exp(-\frac{1}{2}h_d^2 z^2/\sin^2\theta)d\theta \end{aligned}$$

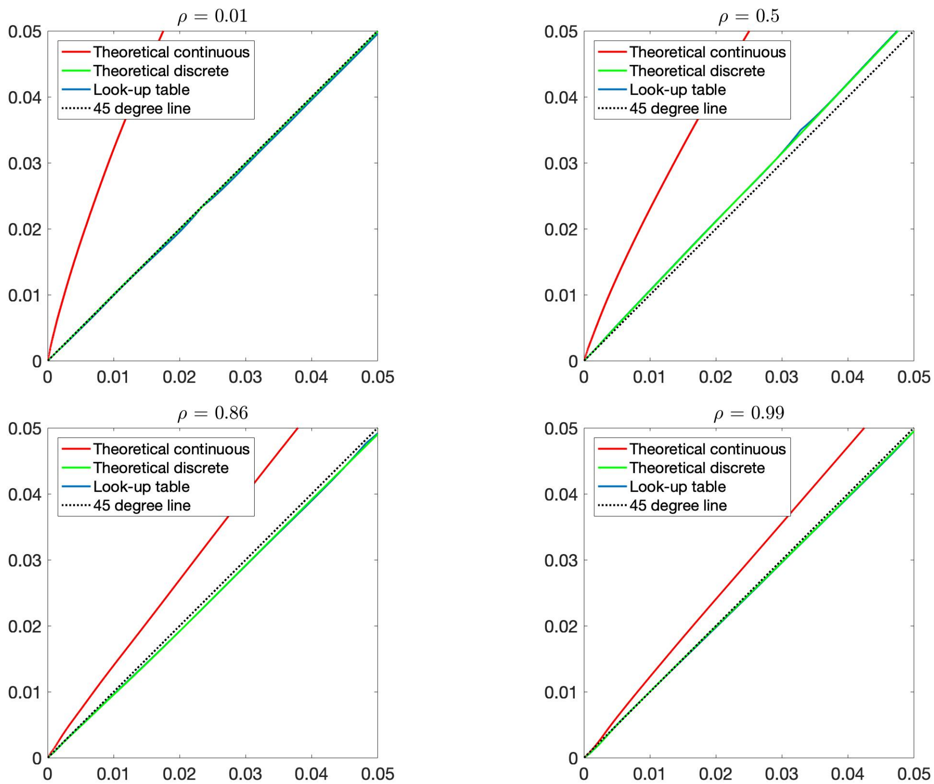

Figure 8: From top left to bottom right: $\rho = 0.01, 0.5, 0.86, 0.99$. The figure is generated based on the comparison among look-up table, simulation directly from the field, DLM and continuous method in 2D non diagonal case.

When $z < 0$,

$$\begin{aligned} Q(\rho_d, z) &= 1 - 2\Phi(0) + \frac{1}{\pi}\int_0^{\alpha_d} \exp(-\frac{1}{2}h_d^2 z^2/\sin^2\theta)d\theta \\ &= \frac{1}{\pi}\int_0^{\alpha_d} \exp(-\frac{1}{2}h_d^2 z^2/\sin^2\theta)d\theta \end{aligned}$$

A.2 Check for the extreme cases of ρ

First,

$$q(\rho, z) = \frac{1}{\pi}\int_0^{\alpha} \exp(-\frac{1}{2}h^2 z^2/\sin^2\theta)d\theta. \tag{7}$$

When $\rho \to 0$, $h \to 1$ and $\alpha \to \frac{\pi}{4}$. So, at the extreme case,

$$\begin{aligned} q(\rho, z) &= \frac{1}{\pi}\int_0^{\frac{\pi}{4}} \exp(-\frac{1}{2}z^2/\sin^2\theta)d\theta, \\ Q(\rho, z) &= 1 - 2\Phi(z^+) + q(\rho, z). \end{aligned}$$

When $\rho \to 1$, $h \to 0$ and $\alpha \to 0$. For lower limits, $-\frac{1}{2}h^2z^2/\sin^2 0 = -\infty$, so $\exp(-\frac{1}{2}h^2z^2/\sin^2 0) = \exp(-\infty) = 0$. For upper limits, $\sin^2\alpha = \frac{1-\rho^2}{2}$, and $h^2 = \frac{1-\rho}{1+\rho}$, so $-\frac{1}{2}h^2z^2/\sin^2\alpha = -\frac{z^2}{(1+\rho)^2}$. When z near zero, $\exp(-\frac{1}{2}h^2z^2/\sin^2\alpha)$ is bounded. So, at the extreme case and z near zero,

$$\begin{aligned} q(\rho, z) &= 0, \\ Q(\rho, z) &= 1 - 2\Phi(0) + q(\rho, z) = 1 - 1 + 0 = 0. \end{aligned}$$

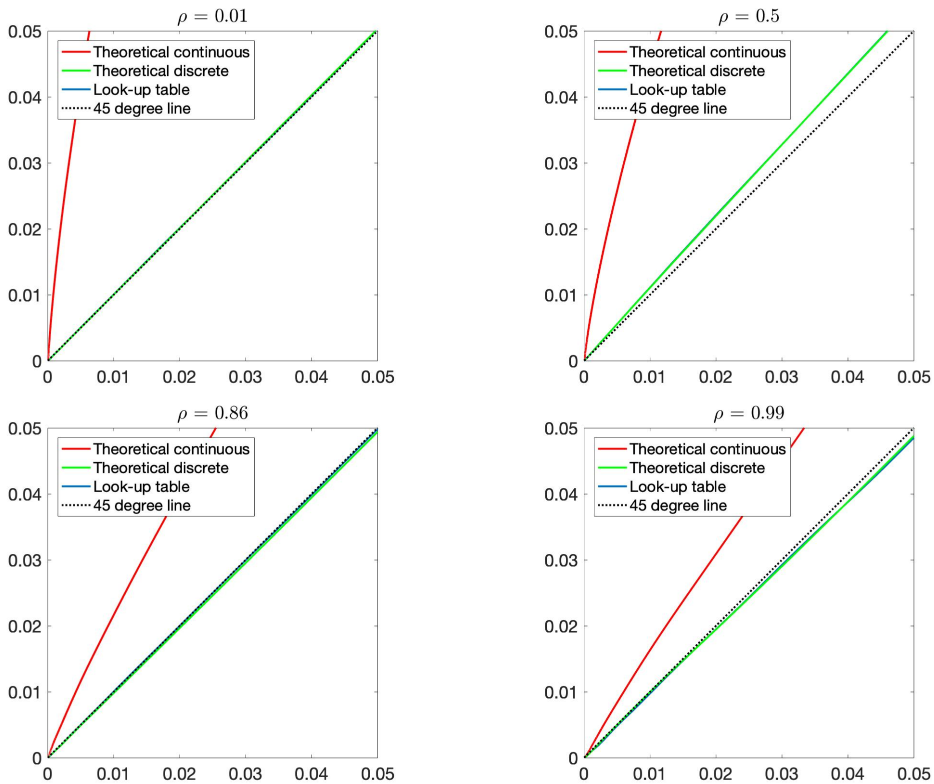

Figure 9: From top left to bottom right: $\rho = 0.01, 0.5, 0.86, 0.99$. The figure is generated based on the comparison among look-up table, simulation directly from the field, DLM and continuous method in 3D non diagonal case.

B Theoretical derivation of probability density function of DLM:

B.1 Goal:

$$\tilde{P}_{DLM} := \int_t^\infty \left(\prod_{d=1}^{D} Q_{xd}(z) \right) \phi(z) dz$$

If the spatial correlation is separable then

$$P_{DLM} = \tilde{P}_{DLM}$$

From above 1.1,

$$P_{DLM} = Pr[\{Z(x) > t\} \cap_{y \in \mathcal{N}} \{Z(y) < Z(x)\}]$$

by law of iterated expectations,

$$\begin{aligned}
&= \int_{-\infty}^{\infty} Pr\big[\{z > t\} \cap (Z(y) < z, \forall y \in \mathcal{N} | Z(x) = z)\big] \phi(z) dz \\
&= \int_{-\infty}^{\infty} I_{\{z>t\}} \cdot Pr\big[(Z(y) < z, \forall y \in \mathcal{N} | Z(x) = z)\big] \phi(z) dz \\
&= \int_{t}^{\infty} Pr\big[(Z(y) < z, \forall y \in \mathcal{N}(x) | Z(x) = z)\big] \phi(z) dz.
\end{aligned}$$

Since we assume that the spatial correlation function is locally separable and dimension 3, also, $Q_{xd}(z)$ is defined as:

$$Q_{xd}(z) = Pr\{Z(y) < z, \ \forall y = x \pm e_d \in S | Z(x) = z\},$$

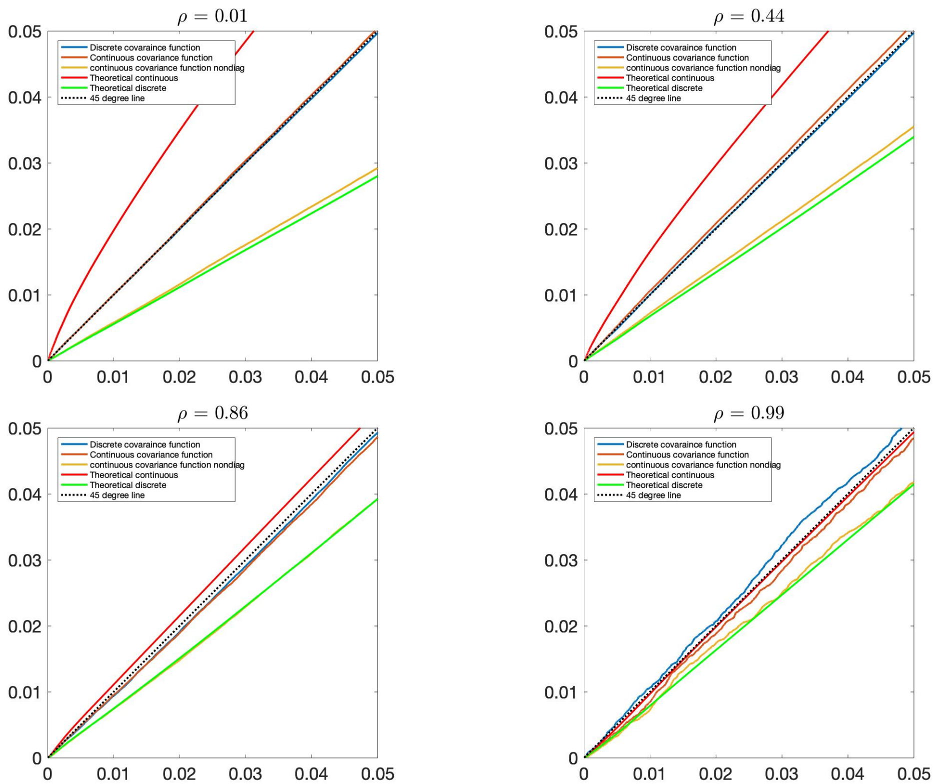

Figure 10: From top left to bottom right: $\rho = 0.01, 0.5, 0.86, 0.99$. The figure is generated based on the 2D comparison among using continuous covariance function method, using discrete covariance function method, simulation directly from the field, DLM and continuous method in non diagonal case.

so,

$$
\begin{aligned}
Pr\big[(Z(y) < z, \forall y \in \mathcal{N}(x) | Z(x) = z)\big] &= Pr\Big[\bigcap_{d=1}^{D} (Z(y) < z,\ \forall y = x \pm e_d \in S | Z(x) = z)\Big] \\
&= \prod_{d=1}^{D} Q_{xd}(z).
\end{aligned}
$$

Therefore,

$$
P_{DLM} = \int_t^{\infty} \Big(\prod_{d=1}^{D} Q_{xd}(z) \Big) \phi(z) dz.
$$

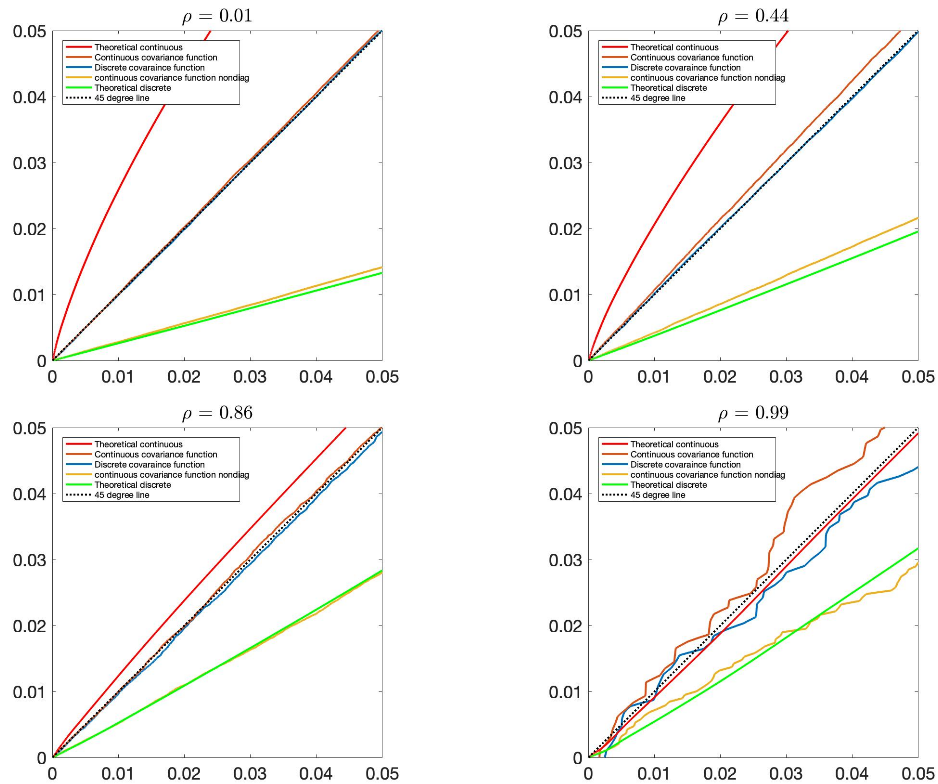

Figure 11: From top left to bottom right: $\rho = 0.01, 0.5, 0.86, 0.99$. The figure is generated based on the 3D comparison among using continuous covariance function method, using discrete covariance function mothod, simulation directly from the field, DLM and continuous method in non diagonal case.

B.2 Goal:

Get the density function of $Pr[Z(x) > t | Z(y) < Z(x), \forall y \in N]$.
By Bayes Rule,

$$
\begin{aligned}
Pr[Z(x) > t | Z(y) < Z(x), \forall y \in N] &= \frac{Pr[\{Z(x) > t\} \cap \{Z(y) < Z(x), \forall y \in N\}]}{Pr[Z(y) < Z(x), \forall y \in N]} \\
&= \frac{\int_t^\infty Pr\big[(Z(y) < z, \forall y \in \mathcal{N}(x) | Z(x) = z)\big]\phi(z)dz}{\int_{-\infty}^\infty Pr\big[(Z(y) < z, \forall y \in \mathcal{N}(x) | Z(x) = z)\big]\phi(z)dz} \\
&= \frac{\int_t^\infty \left(\prod_{d=1}^D Q_{xd}(z)\right)\phi(z)dz}{\int_{-\infty}^\infty \left(\prod_{d=1}^D Q_{xd}(z)\right)\phi(z)dz}
\end{aligned}
$$

C Kronecker proof

Claim: In a 3D Gaussian random field, denote 2 points in the neighbor of $3^3 = 27$ points to be $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$, where x_i and y_i vary from $0, 1, 2$ for $i = 1, 2, 3$. Define $Z(\mathbf{x})$ and $Z(\mathbf{y})$ as the height value of points $\mathbf{x}$ and $\mathbf{y}$. Then

$$
\mathrm{Cov}(Z(\mathbf{x}), Z(\mathbf{y})) = [A \otimes A \otimes A]_{m,n} \,,
$$

where

$$
A = \begin{pmatrix} 1 & \rho & \rho^4 \\ \rho & 1 & \rho \\ \rho^4 & \rho & 1, \end{pmatrix}
$$

and ρ is the correlation between adjacent voxels, i.e., $||\mathbf{x}-\mathbf{y}||=1$. Moreover, m and n denotes the m-th row and n-th column of the matrix, and $m=x_1+3x_2+9x_3+1$, $n=y_1+3y_2+9y_3+1$, which is the index of vectoring the $3\times 3\times 3$ neighbor matrix.

Proof: Since $\mathrm{Cov}(Z(\mathbf{x}),Z(\mathbf{y}))=\exp(-u'\Lambda u/2)$, where u is the euclidean distance between $\mathbf{x}$ and $\mathbf{y}$ and $\Lambda=(1/(2\nu^2),...,1/(2\nu^2))$, then $\mathrm{Cov}(Z(\mathbf{x}),Z(\mathbf{y}))=\rho^{||\mathbf{x}-\mathbf{y}||^2}$, where ρ is the correlation between adjacent voxels. Now,

$$||\mathbf{x}-\mathbf{y}||^2=||x_1-y_1||^2+||x_2-y_2||^2+||x_3-y_3||^2,$$

Thus, for the 27×27 covariance matrix, the (m,n)-th entry, which is defined in the claim, with respect to the covariance between $\mathbf{x}$ and $\mathbf{y}$ has value

$$\mathrm{Cov}(Z(\mathbf{x}),Z(\mathbf{y}))=\rho^{||\mathbf{x}-\mathbf{y}||^2}$$

For the matrix of $A\otimes A\otimes A$, the entry with respect to covariance between $\mathbf{x}$ and $\mathbf{y}$ has value $A_{x_1y_1}\cdot A_{x_2y_2}\cdot A_{x_3y_3}$. If $||x_i-y_i||=0$, $A_{x_iy_i}=1=\rho^0$; if $||x_i-y_i||=1$, $A_{x_iy_i}=1=\rho$; if $||x_i-y_i||=2$, $A_{x_iy_i}=\rho^4$. Thus,

$$\begin{aligned}
&A_{x_iy_i}=\rho^{||x_i-y_i||^2}\\
&A_{x_1y_1}\cdot A_{x_2y_2}\cdot A_{x_3y_3}=\rho^{||x_1-y_1||^2+||x_2-y_2||^2+||x_3-y_3||^2}=\rho^{||x-y||^2}.
\end{aligned}$$