Detection and localization of peaks in a smooth random field

Samuel Davenport (work with Fabian Telschow, Armin Schwartzman and Thomas E. Nichols)

University of California, San Diego

August 24, 2022

Peak Detection

Smoothing

- Suppose we have n subjects and that for each subject we observe

$$X_i = \mu_X + \epsilon_i$$

on a finite lattice $\mathcal{L} \subset S$.

- Smoothing is typically done in order to increase the signal to noise ratio. I.e. for each i, X_i is smoothed with a C^2 kernel K to give

$$Y_i(v) = \sum_{l \in \mathcal{L}} K(v-l)\mu_X(l) + \sum_{l \in \mathcal{L}} K(v-l)\epsilon_i(l)$$

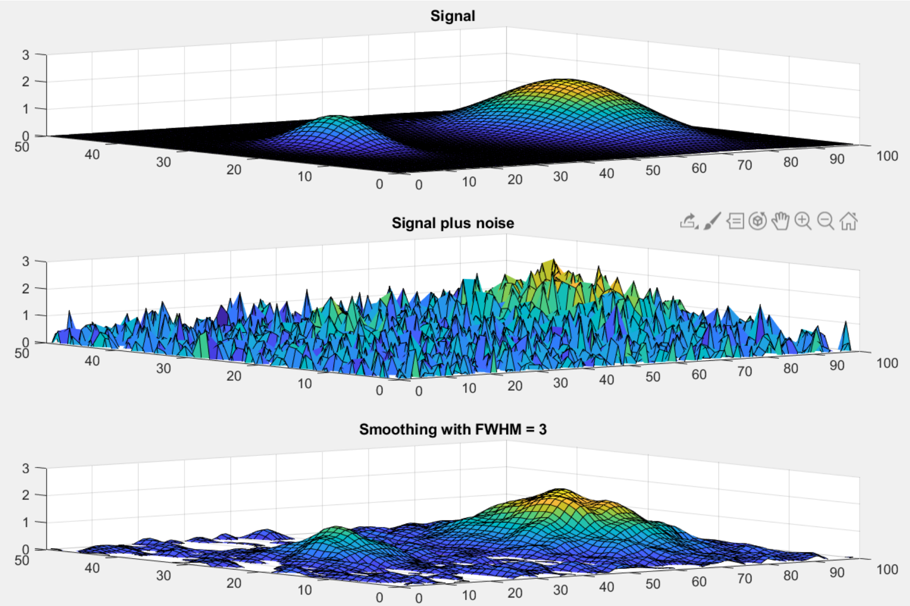
at each $v \in \mathcal{L}$.

- Let

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i = \sum_{l \in \mathcal{L}} K(v-l)\overline{X}_n(l)$$

where $\overline{X}_n = \sum_{i=1}^{n} X_i$.

Why you should smooth

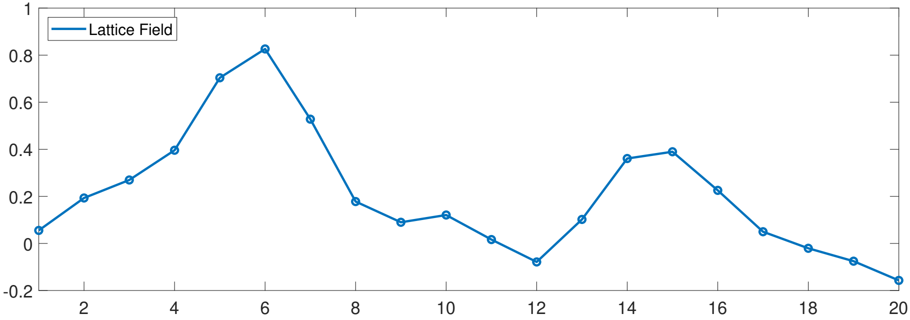


Lattice smoothing

In applications when smoothing is performed (e.g. in fMRI) you evaluate

$$\hat{\mu}_n(v) = \sum_{l \in \mathcal{L}} K(v-l)\bar{X}(l)$$

at every voxel $v \in \mathcal{L}$. An example evaluation $\hat{\mu}_n$ is shown below.



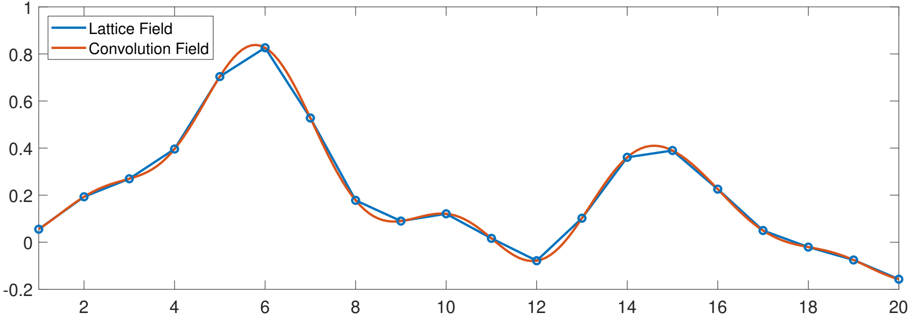
Convolution Random Fields

Definition

Instead we can evaluate the smoothing continuously, meaning that for all $s \in S$,

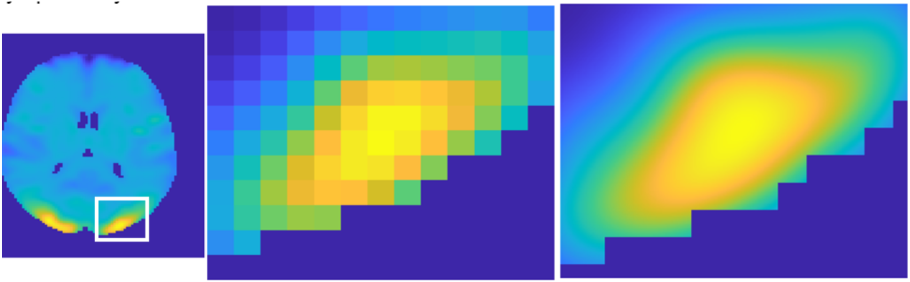
$$\hat{\mu}_n(s) := \sum_{l \in \mathcal{L}} K(s-l)\bar{X}(l).$$

We call this a convolution field.



Convolution Fields in Brain Imaging

Taking slices through a 3D convolution field generated from brain imaging data, you get the following images!



Voxelwise RFT

Let $M_u(\hat{\mu}_n)$ be the number of local maxima of $\hat{\mu}_{Y,n}$ above a threshold u then

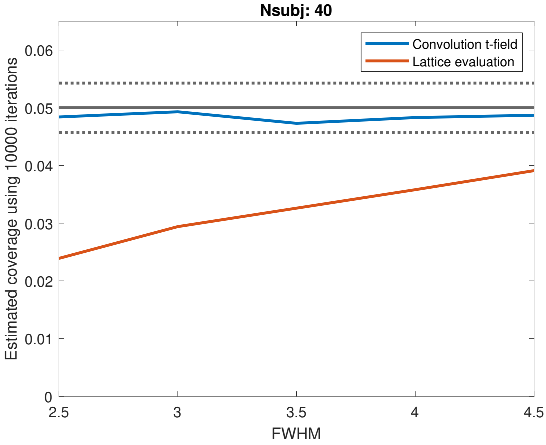
$$\mathbb{P}\left(\sup_{s\in S}\hat{\mu}_n(s) > u\right) = \mathbb{P}(M_u(\hat{\mu}_n) \geq 1) \leq \mathbb{E}[M_u(\hat{\mu}_n)]$$

because $\hat{\mu}$ exceeds u if and only if there is at least one local maxima above u.

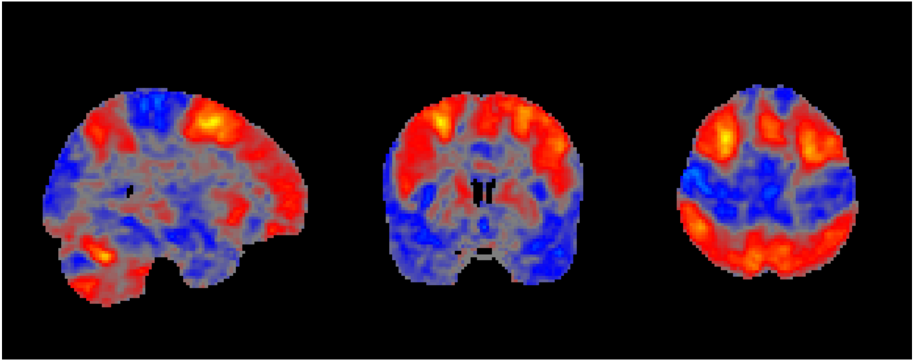
- Many good approximations exist to $\mathbb{E}[M_u(\hat{\mu}_n)]$ at high thresholds u so can control FWER.
- FWER over space is the same as FWER over peaks so we can use this to find significant peaks.
- Using convolution fields to do inference means that RFT works at any applied smoothness because the theory is valid for continuous random fields.

FWER for 2D Gaussian random fields

We generate mean-zero random fields by smoothing Gaussian white noise with an isotropic Gaussian kernel and take $n = 40$. (FWHM is proportional to the smoothing parameter of the kernel.)



Detection and Localization



Peak Confidence Regions

Asymptotic confidence regions

Assume that μ has J peaks at locations $\theta_1, \dots, \theta_J$ within disjoint balls $B_1, \dots, B_J$.

Theorem

For each $j = 1, \dots, J$ corresponding to a local maximum of μ, let $\hat{\theta}_{j,n} = \text{argmax}_{t \in B_j} \hat{\mu}_n(t)$ (and for the minima let $\hat{\theta}_{j,n} = \text{argmin}_{t \in B_j} \hat{\mu}_n(t)$) and let $\hat{\boldsymbol{\theta}}_{\boldsymbol{n}} := (\hat{\theta}_{1,n}^T, \dots, \hat{\theta}_{J,n}^T)^T$ and $\boldsymbol{\theta} := (\theta_1^T, \dots, \theta_J^T)^T$. Then, under regularity assumptions on μ and the noise,

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_{\boldsymbol{n}} - \boldsymbol{\theta}) \xrightarrow{d} N(0, \boldsymbol{A}\boldsymbol{\Lambda}\boldsymbol{A}^T)$$

as $n \longrightarrow \infty$. Here $\boldsymbol{A} \in \mathbb{R}^{DJ \times DJ}$ depends on $\nabla^2 \mu$ and $\boldsymbol{\Lambda} \in \mathbb{R}^{DJ \times DJ}$ depends on the covariance of ∇Y_1.

Accounting for the variance

Expanding we have,

$$\hat{\theta}_{j,n} - \theta_j = -\big(\nabla^2 \hat{\mu}_n(\theta^*_{j,n})\big)^{-1} \nabla^T \hat{\mu}_n(\theta_j)$$

As $n \to \infty$, $\left(\nabla^2 \hat{\mu}_n(\theta^*_{j,n})\right)^{-1} \to \big(\nabla^2 \hat{\mu}_n(\theta_j)\big)^{-1}$ so can approximated by $\left(\nabla^2 \hat{\mu}_n(\hat{\theta}_{j,n})\right)^{-1}$ when making asymptotic confidence intervals.

$$\begin{aligned}\hat{\theta}_{j,n} - \theta_j &= -\left(\nabla^2 \hat{\mu}_n(\theta_j) + \frac{1}{2}(\hat{\theta}_{j,n} - \theta_j)^T \nabla^3 \hat{\mu}_n(\tilde{\theta}_{j,n})\right)^{-1} \nabla^T \hat{\mu}_n(\theta_j) \\ &\approx -\big(\nabla^2 \hat{\mu}_n(\theta_j)\big)^{-1} \nabla^T \hat{\mu}_n(\theta_j)\end{aligned}$$

Monte Carlo simulation from the distribution

We have

$$\begin{pmatrix} \nabla^T \hat{\mu}_n(\theta_j) \\ \mathbf{vech}(\nabla^2 \hat{\mu}_n(\theta_j)) \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ \mathbf{vech}(\nabla^2 \mu_n(\theta_j)) \end{pmatrix}, \frac{1}{n} \begin{pmatrix} \Lambda & 0 \\ 0 & \Omega \end{pmatrix} \right)$$

and for $1 \leq k \leq K$ ($K \in \mathbb{N}$) we can approximate this by simulating from the following distribution

$$\begin{pmatrix} A_k \\ B_k \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ \mathbf{vech}(\nabla^2 \hat{\mu}_n(\hat{\theta}_{j,n})) \end{pmatrix}, \frac{1}{n} \begin{pmatrix} \hat{\Lambda} & 0 \\ 0 & \hat{\Omega} \end{pmatrix} \right).$$

and calculating $\delta_{k,n} = (\mathbf{vech}^{-1}(B_{k,n}))^{-1} A_{k,n}$.

Monte Carlo confidence regions

Let $\hat{\Sigma}'_j = (\nabla^2 \hat{\mu}_n(\hat{\theta}_j))^{-1} \hat{\Lambda} (\nabla^2 \hat{\mu}_n(\hat{\theta}_j))^{-1}$ and for $0 < \alpha < 1$, choose λ_α such that

$$\frac{1}{K} \sum_{k=1}^{K} 1\left[n(\hat{\delta}_{k,n}^T (\hat{\Sigma}'_j)^{-1} \delta_{k,n}) > \lambda_\alpha \right] = \frac{\lfloor \alpha K \rfloor}{K}.$$

Given this we define a $(1-\alpha)$ Monte Carlo confidence region to be

$$\left\{ \theta : n(\hat{\theta}_{j,n} - \theta)^T (\hat{\Sigma}'_j)^{-1} (\hat{\theta}_{j,n} - \theta) < \lambda_\alpha \right\}.$$

- These regions are asymptotically valid (for the same reason as the asymptotic cases)
- Under stationarity we can prove that these intervals are bigger than the asymptotic ones.

Example simulations

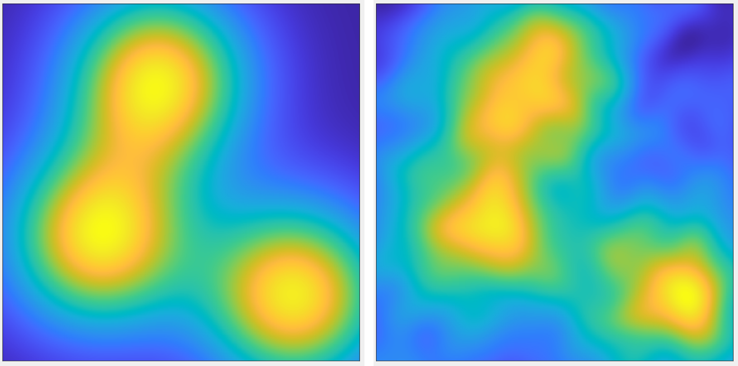
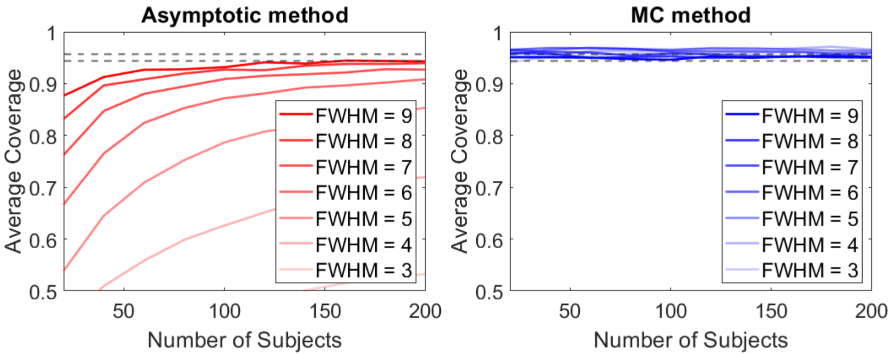


Figure 1: Left: True signal. Right: one realisation.

Comparing coverage rates

Simulate n



Application to fMRI

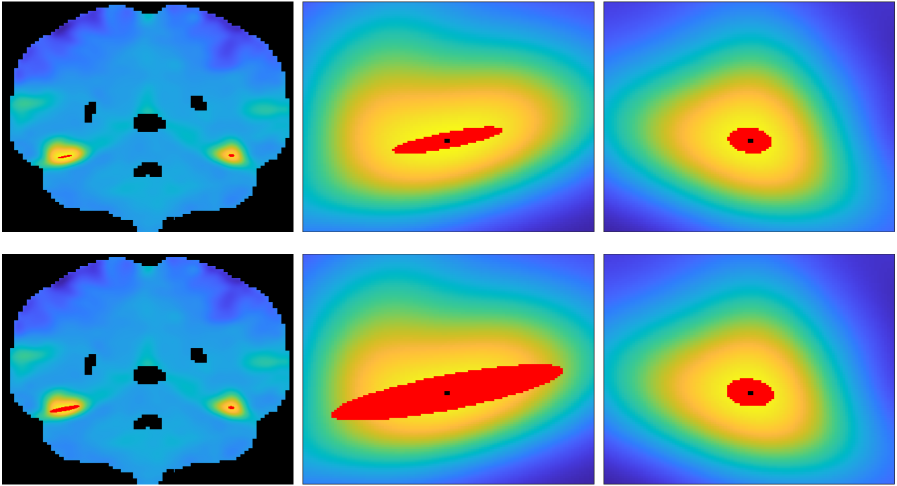


Figure 3: Peaks of the mean of 125 subjects

Conclusion

- Using convolution fields accurately and quickly controls the FWER at the right level. (This is valid under non-stationarity.)
- Here we assumed Gaussianity but the framework works for t-fields and F-fields as well.
- You can derive asymptotic confidence regions and improve upon these if you additionally assume stationarity.
- Software (in MATLAB) to perform inference on random fields is available at the RFTtoolbox (github.com/sjdavenport/RFTtoolbox). (E.g. for LKC estimation, Peak Inference, Peak Height distribution, confidence regions)
- Slides available at `sjdavenport.github.io/talks`.

Pre-prints

- Pre-print on convolution fields soon to be out, in the mean time you can read a lot of the details in my thesis here: https://sjdavenport.github.io/research/papers/thesis.pdf.
- Pre-print on peak confidence regions is available on arxiv (Davenport, Nichols, & Schwarzman, 2022). (see https://sjdavenport.github.io/research/)

Bibliography

Davenport, S., Nichols, T. E., & Schwarzman, A. (2022). Confidence regions for the location of peaks of a smooth random field. *arXiv preprint arXiv:2208.00251*.