Supplementary Material

1 Simulation with randomly generated locations

We conduct extensive simulations to evaluate the FDR control as well as the power of our method with 1,764 randomly generated points over spatial domain $[0, 1] \times [0, 1]$.

1.1 Mean comparison

Consider two spatiotemporal random fields $X(\mathbf{s}, t)$ and $Y(\mathbf{s}, t)$ with spatial mean function $\mu_X(\mathbf{s})$ and $\mu_Y(\mathbf{s})$ and three different types of covariance structures as defined in Section 4.1 in the paper. Again, our hypothesis of interest is

$$H_{0,\mathbf{s}} : \mu_X(\mathbf{s}) = \mu_Y(\mathbf{s}) \quad vs. \quad H_{a,\mathbf{s}} : \mu_X(\mathbf{s}) \neq \mu_Y(\mathbf{s}). \tag{1.1}$$

All the notations and parameter settings are directly adopted from Section 4.1. We only present the nonparametric mirror procedure which is recommended in the paper and implement Benjamini and Hochberg (1995)(BS hereafter) and Storey (2002) (JS hereafter) as references.

Figure 1 and Figure 2 show the FDR and the power of the mean comparison, respectively. The results in these two figures are very similar as what we obtained in Section 4.1 of the paper.

1.2 Covariance comparison

Again, the spatial dependency structure of interest is represented by the covariance between observations at a fixed location $\mathbf{s}_0 \in D$ and every other location $\mathbf{s} \in D, \mathbf{s} \neq \mathbf{s}_0$.

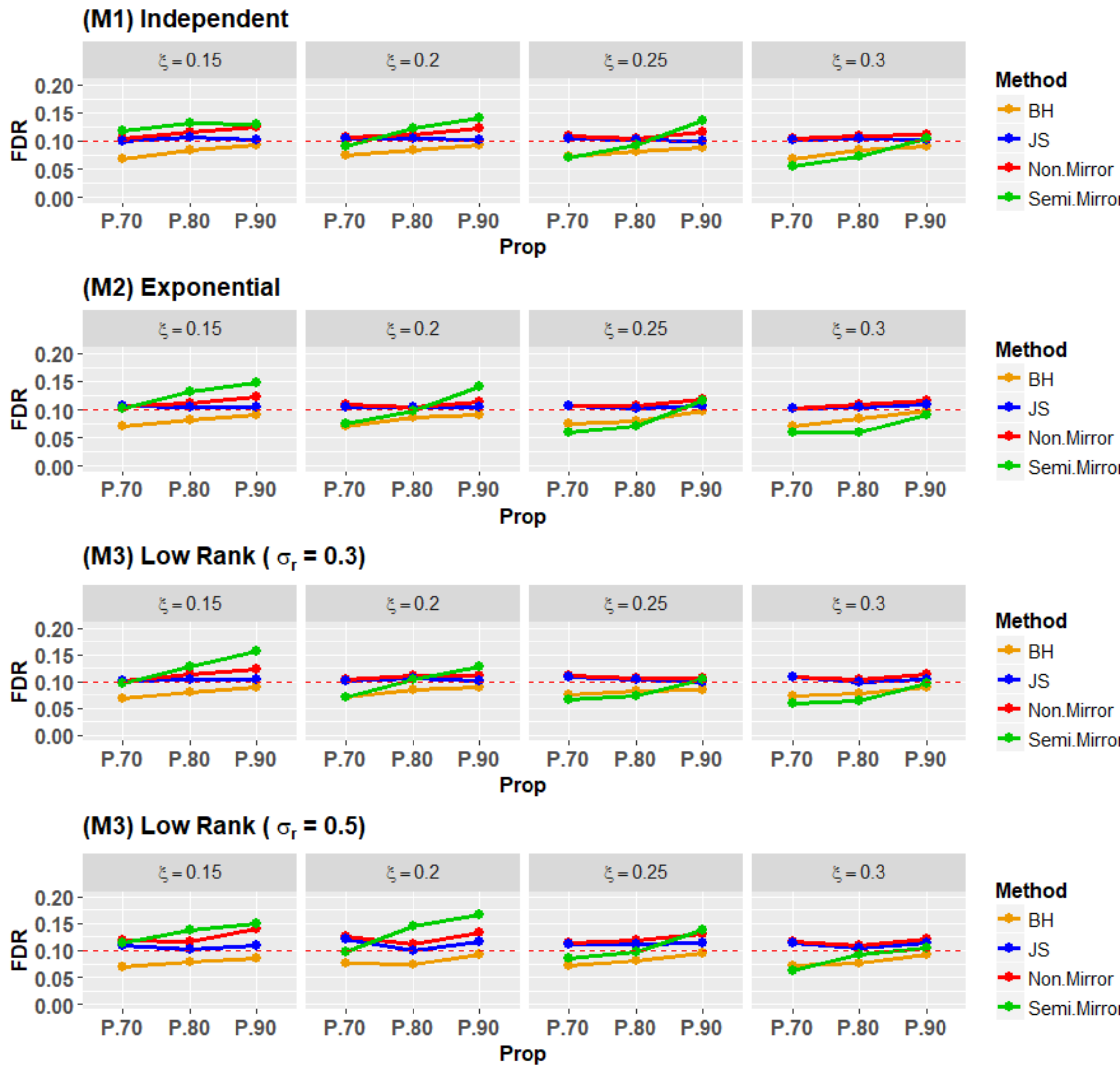

Figure 1: FDR for mean comparison between two randomly generated spatiotemporal fields. $P70, P80, P90$ indicate null proportions being 70%, 80%, and 90%, respectively.

Consider the test:

$$H_{0,\mathbf{s}} : C_X(\mathbf{s}, \mathbf{s}_0) = C_Y(\mathbf{s}, \mathbf{s}_0) \quad vs. \quad H_{a,\mathbf{s}} : C_X(\mathbf{s}, \mathbf{s}_0) \neq C_Y(\mathbf{s}, \mathbf{s}_0), \tag{1.2}$$

where $C_X(\mathbf{s}, \mathbf{s}_0)$ is the covariance between $X(\mathbf{s}, t)$ and $X(\mathbf{s}_0, t)$ and $C_Y(\mathbf{s}, \mathbf{s}_0)$ is defined similarly. We also consider the two covariance functions (C1) and (C2) and follow all notations and parameter settings in Section 4.2 of the paper.

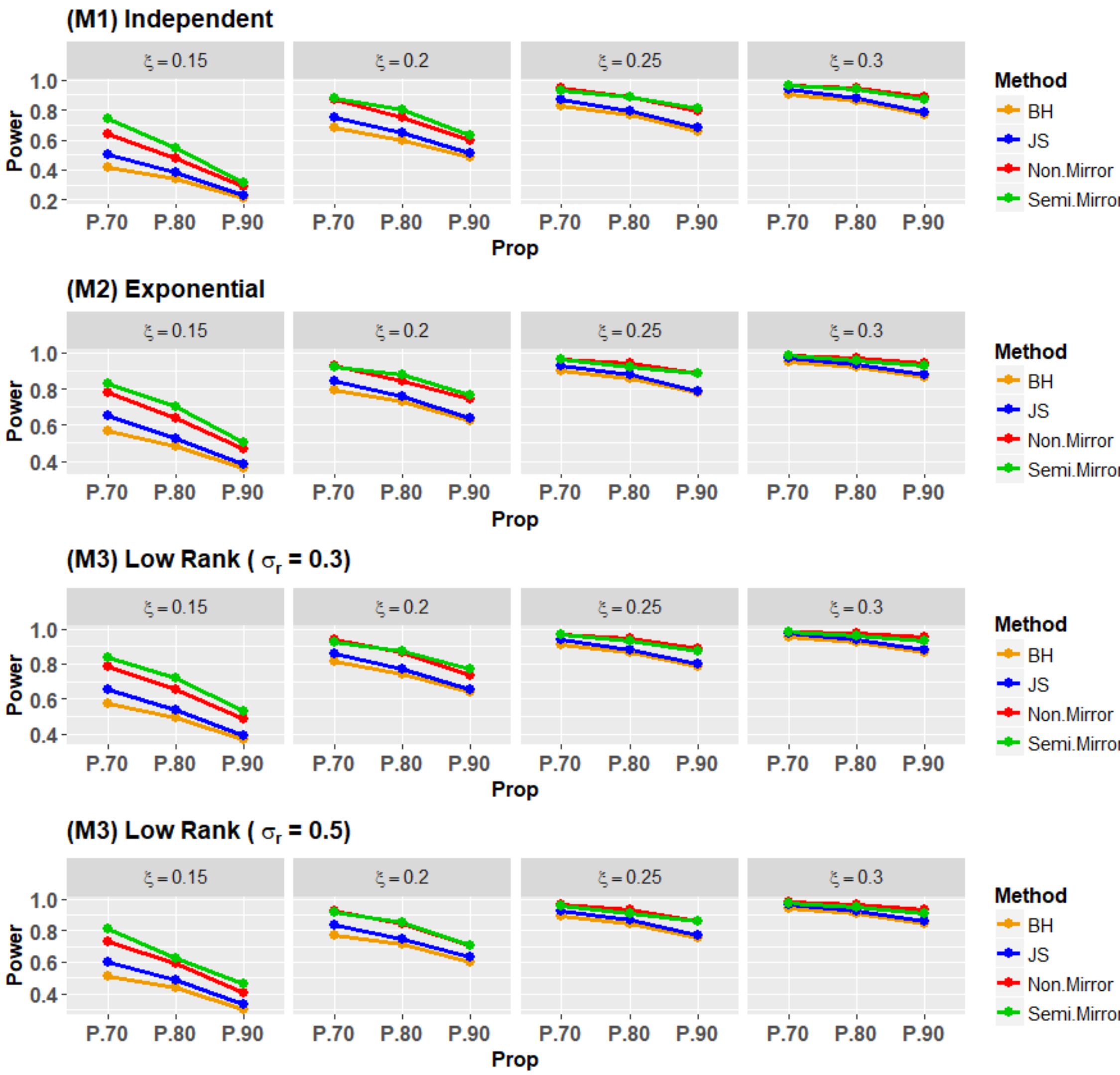

Figure 2: Power for mean comparison between two randomly generated spatiotemporal fields. $P70, P80, P90$ indicate null proportions being $70\%, 80\%$, and 90%, respectively.

Figure 3 and Figure 4 show the FDR and the power of the covariance comparison, respectively. The simulation results with randomly generated locations lead to the same conclusion as with the gridded data.

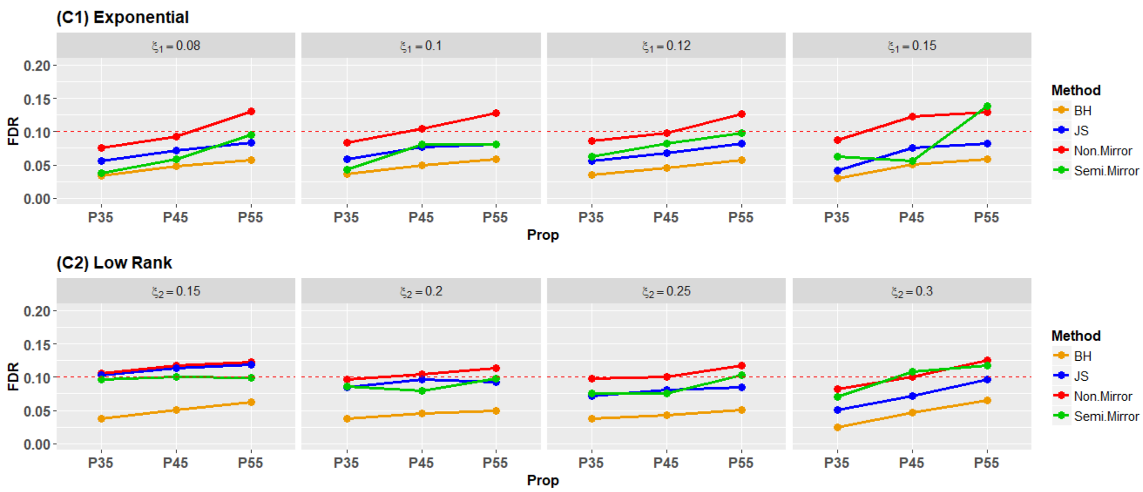

Figure 3: FDR of covariance comparison between two spatiotemporal random fields. $P.40$, $P.49$ and $P.57$ indicate null proportions being 40%, 49%, and 57% respectively for (C1) model; and $P.35$, $P.45$, $P.55$ indicate null proportions being 35%, 45%, and 55% respectively for (C2) model.

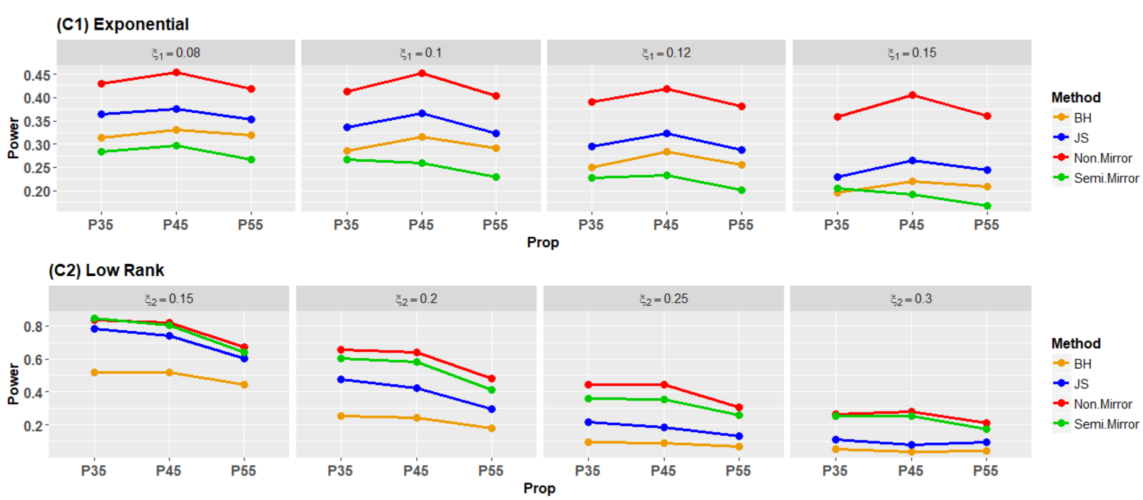

Figure 4: Power of covariance comparison between two spatiotemporal random fields. $P.40$, $P.49$ and $P.57$ indicate null proportions being 40%, 49%, and 57% respectively for (C1) model; and $P.35$, $P.45$, $P.55$ indicate null proportions being 35%, 45%, and 55% respectively for (C2) model.

2 Two sample self-normalization test

In this subsection, we describe the two sample self-normalization test for testing

$$H_{0,\mathbf{s}} : \theta_X(\mathbf{s}) = \theta_Y(\mathbf{s}) \quad \text{versus} \quad H_{a,\mathbf{s}} : \theta_X(\mathbf{s}) \neq \theta_Y(\mathbf{s})$$

at each location $\mathbf{s} \in D$, where $\theta_X(\mathbf{s})$ and $\theta_Y(\mathbf{s})$ are q-dimensional parameters. Set $n = n_1 + n_2$. Let $\widehat{\theta}_{X,k}(\mathbf{s})$ and $\widehat{\theta}_{Y,k'}(\mathbf{s})$ be the recursive estimates of $\theta_X(\mathbf{s})$ and $\theta_Y(\mathbf{s})$ calculated based on the subsamples $\{X(\mathbf{s},t)\}_{t=1}^{k}$ and $\{Y(\mathbf{s},t)\}_{t=1}^{k'}$ respectively. Define $\widehat{\theta}_k(\mathbf{s}) = \widehat{\theta}_{X,\lfloor kn_1/n \rfloor}(\mathbf{s}) - \widehat{\theta}_{Y,\lfloor kn_2/n \rfloor}(\mathbf{s})$ for $1 \leq k \leq n$, where $\lfloor a \rfloor$ denotes the integer part of a. The self-normalization test is defined as

$$G_n = n^3 \widehat{\theta}_n(\mathbf{s})^\top \left\{ \sum_{k=1}^{n} k^2 (\widehat{\theta}_k(\mathbf{s}) - \widehat{\theta}_n(\mathbf{s}))(\widehat{\theta}_k(\mathbf{s}) - \widehat{\theta}_n(\mathbf{s}))^\top \right\}^{-1} \widehat{\theta}_n(\mathbf{s}).$$

Under suitable weak dependence assumption, it can be shown that

$$G_n \to^d G_\infty = B_q(1)^\top \left\{ \int_0^1 (B_q(r) - rB_q(1))(B_q(r) - rB_q(1))^\top \right\}^{-1} B_q(1)$$

where $B_q(r)$ is a q-dimensional vector of independent Brownian motions. Thus the p-value can be calculated as $p_{\mathbf{s}} = P(G_\infty \geq G_n | G_n)$. In practice, the distribution of G_∞ can be obtained via simulation.