Hurst Estimation for Operator Scaling Random Fields

Jeonghwa Lee

Truman State University

NSF/CBMS Conference, August 2021

Operator Scaling Gaussian Random Fields

A scalar valued random field $\{X(x)\}_{x\in\mathbb{R}^d}$ is called operator-scaling if for some $d \times d$ matrix E with positive real parts of the eigenvalues and some $H > 0$ we have

$$\{X(c^E x)\}_{x\in\mathbb{R}^d} \stackrel{f.d.}{=} \{c^H X(x)\}_{x\in\mathbb{R}^d} \qquad \text{for all } c > 0, \tag{1}$$

where $\stackrel{f.d.}{=}$ denotes equality of all finite-dimensional marginal distributions, and $c^E = \exp(E \log c)$ where $\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}$ is the matrix exponential.

If E is diagonizable matrix, $Eu_i = \lambda_i u_i$, for $i = 1, \cdots, d$, then $\{X(ctu_i)\}_{t\in\mathbb{R}} \stackrel{f.d.}{=} \{c^{H/\lambda_i} X(tu_i)\}_{t\in\mathbb{R}}$ for any $c > 0$.

Operator Scaling Gaussian Random Fields

Explicit covariance functions of operator scaling Gaussian random field (OSGRF) were proposed by Bierme, H. and Lacaux, C. (2018). They define a function

$$\upsilon_{E,H}(x) = \Big(\sum_{i=1}^{d} |\langle x, u_i \rangle|^{2a_i} \Big)^H \qquad \text{for all } x \in \mathbb{R}^d, \tag{2}$$

where $H \in (0, 1]$, and $\{1/a_i = \lambda_i > 0, u_i, i = 1, 2, \cdots, d\}$ are eigenvalues and eigenvectors of a diagonizable matrix E.

(2) is a semi-variogram funtion for a centered Gaussian random field that has operator scaling property (1).

Recall semi-variogram of X is defined by

$$\upsilon_{E,H}(h) = \frac{1}{2} E(X(x+h) - X(x))^2.$$

Assumption

(A1) $\upsilon_{E,H}(x)$ in (2) with $H \leq 1$ and $a_1 < a_2, a_1, a_2 \in (0,1)$.
(A2) The matrix E in (1) is diagonizable with eigenvalues $\lambda_2 \neq \lambda_1$, and

$$E = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad \text{for } \theta \in (0,\pi).$$

By assumption (A2), X has self-similarity along the directions u_1, u_2 with the Hurst indices $h_1 = Ha_1 (= H/\lambda_1), h_2 = Ha_2 (= H/\lambda_2)$, respectively.

$$\{X(ctu_j)\}_{t\in\mathbb{R}} \overset{f.d}{=} \{c^{h_j} X(tu_j)\}_{t\in\mathbb{R}} \ \text{ for any } c > 0, j = 1,2, \tag{3}$$

where $u_1 = (\cos\theta, -\sin\theta)$ and $u_2 = (\sin\theta, \cos\theta)$.

Estimation Strategy

Below is a brief explanation on how we will proceed to estimate parameters, u_1, u_2, h_1, h_2.

First, using sample paths on the horizontal and vertical axis, $\{X(t,0), X(0,t), t \in [0,1]\}$, we obtain estimate of θ, $\hat{\theta}$, and $\hat{u}_1 = (\cos\hat{\theta}, -\sin\hat{\theta}),\ \hat{u}_2 = (\sin\hat{\theta}, \cos\hat{\theta})$.

Second, using the sample paths on the estimated eigenvector directions, $\{X(t\hat{u}_1), X(t\hat{u}_2), t \in [0,1]\}$, and self-similar property, we obtain Hurst estimates $\hat{h}_j, j = 1, 2$.

Estimation of Eigenvectors

Define for $i = 2, \cdots, n$,

$$\nabla_1 X(i/n) = X\begin{pmatrix} i/n \\ 0 \end{pmatrix} - 2X\begin{pmatrix} (i-1)/n \\ 0 \end{pmatrix} + X\begin{pmatrix} (i-2)/n \\ 0 \end{pmatrix}, \tag{4}$$

$$\nabla_2 X(i/n) = X\begin{pmatrix} 0 \\ i/n \end{pmatrix} - 2X\begin{pmatrix} 0 \\ (i-1)/n \end{pmatrix} + X\begin{pmatrix} 0 \\ (i-2)/n \end{pmatrix}. \tag{5}$$

Note that $\{\nabla_1 X(i/n), \nabla_2 X(i/n), i = 2, \cdots, n\}$ are stationary processes for a fixed n.

Estimation of Eigenvectors

Define

$$P_n := \sum_{i=2}^{n} \frac{\nabla_1 X(i/n)^2}{n-1}, \qquad Q_n := \sum_{i=2}^{n} \frac{\nabla_2 X(i/n)^2}{n-1}, \tag{6}$$

and $\sigma_P^2 := \lim_n var(n^{.5+2h_1} P_n), \sigma_Q^2 := \lim_n var(n^{.5+2h_1} Q_n)$.

Lemma (2.2)

Under the assumptions $(A1, A2)$, P_n and Q_n are asymptotically independent in a sense that $\sigma_{PQ} := \lim_n n^{1+4h_1} cov(P_n, Q_n) = 0$, and

$$n^{.5+2h_1} \begin{pmatrix} P_n - EP_n \\ Q_n - EQ_n \end{pmatrix} \rightarrow_d N(0, \Sigma),$$

where $\Sigma_{11} = \sigma_P^2, \Sigma_{22} = \sigma_Q^2, \Sigma_{12} = \sigma_{PQ} = 0$.

Estimation of Eigenvectors

Define for $i = 4, \cdots, n$,

$$\nabla_1^* X(i/n) = X\begin{pmatrix} i/n \\ 0 \end{pmatrix} - 2X\begin{pmatrix} (i-2)/n \\ 0 \end{pmatrix} + X\begin{pmatrix} (i-4)/n \\ 0 \end{pmatrix}.$$

Let

$$P_n^* := \sum_{i=4}^{n} \frac{\nabla_1^* X(i/n)^2}{n-3}.$$

Lemma (2.3)

Under the assumptions $(A1, A2)$,

$$
\begin{aligned}
i) \quad & \frac{P_n}{Q_n} - \frac{EP_n}{EQ_n} = O_p(n^{-.5}), \\
ii) \quad & \frac{EP_n}{EQ_n} - \left(\frac{\cos\theta}{\sin\theta}\right)^{2h_1} = \\
& \qquad O\Big(n^{2(a_1-a_2)}(\sin^{2(a_2-h_1)}\theta \vee \cos^{2(a_2+h_1)}\theta/\sin^{4h_1}\theta)\Big), \\
iii) \quad & \frac{P_n^*}{P_n} - \frac{EP_n^*}{EP_n} = O_p(n^{-.5}), \\
iv) \quad & \frac{EP_n^*}{EP_n} - 2^{2h_1} = O(n^{2(a_1-a_2)}\sin^{2a_2}\theta/\cos^{2h_1}\theta).
\end{aligned}
$$

Estimation of Eigenvectors

The estimation method for θ is the following:
Step 1) Estimate $2h_1$ by the ratio of P_n^* and P_n.

$$2\tilde{h}_1 = \log\left(\frac{P_n^*}{P_n}\right) / \log 2. \tag{7}$$

Step 2) Estimate θ with $2\tilde{h}_1$ and $\frac{P_n}{Q_n}$.

$$\hat{\theta}_n := \cot^{-1}\left(\left(\frac{P_n}{Q_n}\right)^{1/2\tilde{h}_1}\right). \tag{8}$$

Estimation of Eigenvectors

Theorem (2.4)

Under the assumptions $(A1, A2)$,

$$i) \qquad 2\tilde{h}_1 - 2h_1 = O_p\left(n^{2(a_1-a_2)}\frac{\sin^{2a_2}\theta}{\cos^{2h_1}\theta} \vee n^{-.5}\right), \tag{9}$$

$$ii) \qquad \hat{\theta}_n - \theta = O_p(n^{2(a_1-a_2)} \vee n^{-.5}\sin^2\theta). \tag{10}$$

Estimation of Hurst Indices

Define vector-valued functions F_{u_1}, F_{u_2}, for any $\theta_c \in (0, \pi/2)$,

$$F_{u_1}(\theta_c) := (\cos\theta_c, -\sin\theta_c), \qquad F_{u_2}(\theta_c) := (\sin\theta_c, \cos\theta_c).$$

Our estimators for $u_j, j = 1, 2$, are

$$\hat{u}_1 := F_{u_1}(\hat{\theta}_n) = (\cos\hat{\theta}_n, -\sin\hat{\theta}_n), \qquad \hat{u}_2 := F_{u_2}(\hat{\theta}_n) = (\sin\hat{\theta}_n, \cos\hat{\theta}_n).$$

For an integer $2^m << n$, and a vector $u \in [0, 1]^2, |u| = 1$, define

$$\nabla^m(u)X(i/n) := X\left(\frac{i - 2^{m+1}}{n}u\right) - 2X\left(\frac{i - 2^m}{n}u\right) + X\left(\frac{i}{n}u\right) \tag{11}$$

for $i = 2^{m+1}, \cdots, n$.

Estimation of Hurst Indices

Generalizing P_n, Q_n in (6), define for any constant $\theta_c \in (0, \pi/2)$,

$$P_n^m(\theta_c) := \sum_{i=2^{m+1}}^{n} \frac{\nabla^m(F_{u_1}(\theta_c))X(i/n)^2}{n - 2^{m+1} + 1}, \tag{12}$$

$$Q_n^m(\theta_c) := \sum_{i=2^{m+1}}^{n} \frac{\nabla^m(F_{u_2}(\theta_c))X(i/n)^2}{n - 2^{m+1} + 1}. \tag{13}$$

Since,

$$P_n^m(\hat{\theta}_n) \approx P_n^m(\theta), \qquad Q_n^m(\hat{\theta}_n) \approx Q_n^m(\theta),$$

and

$$P_n^m(\theta) \overset{d}{=} 2^{2mh_1} P_n^1(\theta), \qquad Q_n^m(\theta) \overset{d}{=} 2^{2mh_2} P_n^1(\theta),$$

(by self-similarity of X along the directions u_1, u_2)

Estimation of Hurst Indices

we estimate h_1 and h_2 by log-regression of

$$\{P_n^m(\hat{\theta}_n); m = 1, 2, \cdots, \ell_n\}, \{Q_n^m(\hat{\theta}_n); m = 1, 2, \cdots, \ell_n\}$$

on $\{2m \log 2, m = 1, 2, \cdots, \ell_n\}$, respectively, for a fixed integer $\ell_n << n$, i.e.,

$$\hat{h}_1 = \frac{1}{2} \sum_{m=1}^{\ell_n} w_m \log_2 P_n^m(\hat{\theta}_n), \qquad \hat{h}_2 = \frac{1}{2} \sum_{m=1}^{\ell_n} w_m \log_2 Q_n^m(\hat{\theta}_n), \tag{14}$$

where $\sum_{m=1}^{\ell_n} w_m = 0, \sum_{m=1}^{\ell_n} m w_m = 1$.

Estimation of Hurst Indices

For any integers m, m', $\lim_n cov(n^{.5+2h_1} P_n^{m'}(\theta_1), n^{.5+2h_1} P_n^{m}(\theta_2)) = 0$.
Therefore, for any constant $\theta_c \in (0, \pi/2)$,

$$n^{.5+2h_1} \begin{pmatrix} P_n^m(\theta_c) - EP_n^m(\theta_c) \\ Q_n^m(\theta_c) - EQ_n^m(\theta_c) \\ P_n^0(0) - EP_n^0(0) \\ Q_n^0(0) - EQ_n^0(0) \end{pmatrix} \to_d N(0, \Sigma^*),$$

where $\Sigma^* = \begin{pmatrix} \Sigma_{\theta_c} & 0 \\ 0 & \Sigma \end{pmatrix}$, $\Sigma, \Sigma_{\theta_c}$ are diagonal matrices.

$$\Rightarrow n^{.5+2h_1} \begin{pmatrix} P_n^m(\theta_c) - EP_n^m(\theta_c) \\ Q_n^m(\theta_c) - EQ_n^m(\theta_c) \end{pmatrix} \Bigg| \begin{pmatrix} P_n^0(0) \\ Q_n^0(0) \end{pmatrix} \to_d N(0, \Sigma_{\theta_c}).$$

Estimation of Hurst Indices

Since $\hat{\theta}_n$ is obtained from $P_n = P_n^0(0), Q_n = Q_n^0(0)$, given $\hat{\theta}_n = \theta_c$,

$$n^{.5+2h_1} \begin{pmatrix} P_n^m(\theta_c) - EP_n^m(\theta_c) \\ Q_n^m(\theta_c) - EQ_n^m(\theta_c) \end{pmatrix} \rightarrow_d N(0, \Sigma_{\theta_c}). \tag{15}$$

Estimation of Hurst Indices

For any constant $\theta_c \in (0, \pi/2)$, define

$$h_1^{\theta_c} := \frac{1}{2} \sum_{m=1}^{\ell_n} w_m \log_2 E(P_n^m(\theta_c)), \qquad h_2^{\theta_c} := \frac{1}{2} \sum_{m=1}^{\ell_n} w_m \log_2 E(Q_n^m(\theta_c)).$$

Given that $\hat{\theta}_n = \theta_c$, $\hat{h}_j$ converges to $h_j^{\theta_c}$, for $j = 1, 2$. Also, as $\hat{\theta}_n$ converges to θ, $h_j^{\hat{\theta}_n}$ converges to h_j, for $j = 1, 2$.

Estimation of Hurst Indices

Theorem (2.5)

In a fixed domain with the assumptions $(A1, A2)$,
i)

$$|\hat{h}_1 - h_1| = \begin{cases} \mathcal{O}_P\big(n^{-.5-2h_1} + n^{-2(a_2-a_1)-a_2}\big) & \text{if } a_2 - a_1 > .25, \\ \mathcal{O}_P\big(n^{-.5-2h_1} + n^{-(2+4a_2)(a_2-a_1)}\big) & \text{if } a_2 - a_1 \leq .25. \end{cases}$$

ii)

$$|\hat{h}_2 - h_2| = \mathcal{O}_P\big(n^{-.5-2h_1} + n^{-(4a_1-2)(a_2-a_1)}\big) \quad \text{if } a_2 - a_1 \leq .25, a_1 > .5.$$

Estimation of Hurst Indices

proof. i) By (15) and the delta method, given $\hat{\theta}_n = \theta_c$, the estimator $\hat{h}_1$ behaves as follows:

$$n^{.5+2h_1}(\hat{h}_1 - h_1^{\theta_c}) \to_d N(0, \sigma_{\hat{u}_1}). \tag{16}$$

Also, for $h_1^{\theta_c} - h_1$, note that

$$\begin{aligned} E\Big(\nabla^m(F_{u_1}(\theta_c))X(i/n)^2\Big) &= 8\bigg(b_{a_1}\left(\frac{2^{2ma_1}}{n^{2a_1}}\right) + \epsilon_1^{2a_2}\left(\frac{2^{2ma_2}}{n^{2a_2}}\right)\bigg)^H \\ &\quad - 2\bigg(b_{a_1}\left(\frac{2^{2(m+1)a_1}}{n^{2a_1}}\right) + \epsilon_1^{2a_2}\left(\frac{2^{2(m+1)a_2}}{n^{2a_2}}\right)\bigg)^H \\ = c_{h_1}^*\left(\frac{2^{2mh_1}}{n^{2h_1}}\right)&\left(1 + \frac{\tilde{c}_{a_2}}{c_{h_1}^*}H\epsilon_1^{2a_2}2^{2m(a_2-a_1)}n^{2a_1-2a_2} + o(\epsilon_1^{2a_2}n^{2a_1-2a_2})\right), \end{aligned} \tag{17}$$

where $b_{a_1} = (u_1' F_{u_1}(\theta_c))^{2a_1}, \epsilon_1 = u_2' F_{u_1}(\theta_c)$, and $c_{h_1}^* = (8 - 2^{2h_1+1})(u_1' F_{u_1}(\theta_c))^{2h_1}, \tilde{c}_{a_2} = (8 - 2^{2(a_2-a_1+h_1)+1})b_{a_1}^{H-1}$. The last equality follows by Taylor expansion. (17) implies that $h_1^{\theta_c} - h_1 = O(\epsilon_1^{2a_2}n^{2a_1-2a_2})$.

Estimation of Hurst Indices

Therefore, combining (16-17),

$$|\hat{h}_1 - h_1| \leq |\hat{h}_1 - h_1^{\theta_c}| + |h_1^{\theta_c} - h_1| = \mathcal{O}_P(n^{-.5-2h_1}) + \mathcal{O}(\epsilon_1^{2a_2} n^{2a_1-2a_2}). \quad (18)$$

(18) was derived for fixed $\hat{\theta}_n = \theta_c$.
Since $\epsilon_1 = u_2' F_{u_1}(\theta_c) = u_2'(\cos\theta_c, -\sin\theta_c)$, ϵ_1 varies with $\hat{\theta}_n$, and it has the order of (10). Therefore, $\epsilon_1^{2a_2} n^{2a_1-2a_2}$ is of order $n^{2(a_1-a_2)-a_2}$ or $n^{-(2+4a_2)(a_2-a_1)}$ depending on whether $a_2 - a_1 > .25$ or $a_2 - a_1 \leq .25$, and the result follows.

Estimation of Hurst Indices

ii) By (15) and the delta method, given $\hat{\theta}_n = \theta_c$,

$$n^{.5+2h_1}(\hat{h}_2 - h_2^{\theta_c}) \to_d N(0, \sigma_{\hat{u}_2}). \tag{19}$$

For $h_2^{\theta_c} - h_2$,

$$E\Big(\nabla^m(F_{u_2}(\theta_c))X(i/n)^2\Big) = 8\left(\epsilon_2^{2a_1}\left(\frac{2^{2ma_1}}{n^{2a_1}}\right) + b_{a_2}\left(\frac{2^{2ma_2}}{n^{2a_2}}\right)\right)^H$$

$$- 2\left(\epsilon_2^{2a_1}\left(\frac{2^{2(m+1)a_1}}{n^{2a_1}}\right) + b_{a_2}\left(\frac{2^{2(m+1)a_2}}{n^{2a_2}}\right)\right)^H$$

$$= c_{h_2}^*\left(\frac{2^{2mh_2}}{n^{2h_2}}\right)\left(1 + \frac{\tilde{c}_{a_1}}{c_{h_2}^*}H\epsilon_2^{2a_1}2^{2m(a_1-a_2)}n^{2a_2-2a_1} + o(\epsilon_2^{2a_1}n^{2a_2-2a_1})\right), \tag{20}$$

where
$b_{a_2} = (u_2' F_{u_2}(\theta_c))^{2a_2}, \epsilon_2 = u_1' F_{u_2}(\theta_c), c_{h_2}^* = (8 - 2^{2h_2+1})(u_2' F_{u_2}(\theta_c))^{2h_2}$,
and $\tilde{c}_{a_1} = (8 - 2^{2(a_1-a_2+h_2)+1})b_{a_2}^{H-1}$.

Estimation of Hurst Indices

By (19-20), for fixed $\hat{\theta}_c = \theta_c$,

$$|\hat{h}_2 - h_2| \leq |\hat{h}_2 - h_2^{\theta_c}| + |h_2^{\theta_c} - h_2| = \mathcal{O}_P(n^{-.5-2h_2}) + \mathcal{O}(\epsilon_2^{2a_1} n^{2a_2-2a_1}). \quad (21)$$

$\epsilon_2 = u_1' F_{u_2}(\theta_c)$ varies with $\hat{\theta}_n = \theta_c$, and it is of order $n^{-(4a_1-2)(a_2-a_1)}$ when $a_2 - a_1 \leq .25$ and $a_1 > .5$, or divergent otherwise, therefore, the results follow. $\square$

Simulation

Simulation method for OSGRF on a grid was developed by Bierme, H. and Lacaux, C. (2018) when semi-variogram is (2) with diagonal matrix E.

However, since we have diagonizable matrix E, the algorithm cannot be used.

Moreover, the samples we need for the whole estimation procedure do not fit a grid in a fixed domain, since we need not only $\{X(i/n, 0), X(0, i/n), i = 0, 1, \cdots, n\}$, but also $\{X((i/n)\hat{u}_j), i = 0, 1, \cdots, n, j = 1, 2\}$ which do not lie in $\{X(i/n, j/n), i, j = 0, 1, \cdots, n\}$.

Simulation

Instead, we take the following approach:
(1) From OSGRF with fixed parameters (θ, h_1, h_2, H), we obtained two sample paths,

$$\{X(i/n, 0), i = 0, 1, \cdots, n\}, \{X(0, i/n), i = 0, 1, \cdots, n\},$$

that are independent of each other.
(2) From sample paths in (1), $\hat{\theta}$ is obtained.
(3) Along the estimated directions $\hat{u}_1, \hat{u}_2$, sample paths are obtained,

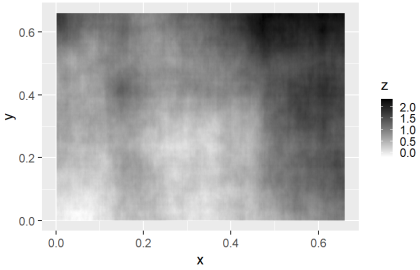
$$\{X\big((i/n)\hat{u}_1\big), i = 0, 1, \cdots, n\}, \{X\big((i/n)\hat{u}_2\big), i = 0, 1, \cdots, n\},$$

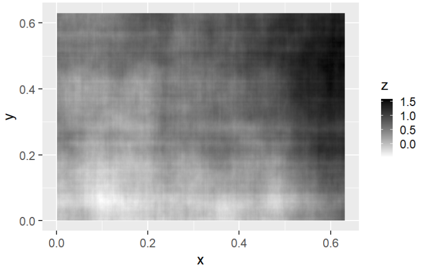
that are independent of each other and also independent of sample paths in (1).
(4) Using the sample paths in (3), $\hat{h}_1, \hat{h}_2$, are computed.
We repeat (1)-(4) with various parameters (θ, h_1, h_2, H).

Simulation

- We proceed (1)-(4) with $n = 2^{13}$ for each set of (θ, h_1, h_2, H) where $\theta \in \{i/20 * \pi/2, i = 1, \cdots, 19\}$ with various h_1, h_2, H, and obtain $\hat{\theta}$ and $\hat{h}_1, \hat{h}_2$ with $m = 1, 2, 3, 4 (= \ell_n)$.
- Sample paths were simulated by circulant embedding method developed by Dietrich, C.R. and Newsam, G.N. (1997). This method was independently applied to each sample path in (1) and (3), which results in zero covariance between any two samples from different sample paths. This is not assumed covariance in our model.
- However, $P_n^m(\theta_1), P_n^{m'}(\theta_2)$ are asymptotically independent when $\theta_1 \neq \theta_2$. Since we only use $P_n^m(\theta)$ from each sample path with different θ for different sample path, our approach is justified to investigate how estimators perform.

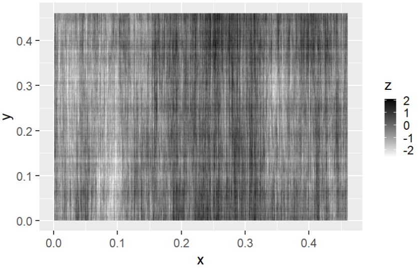
Simulation Results

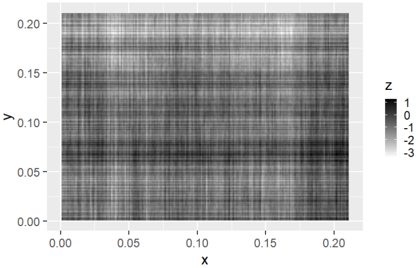




(A): $h_1 = .6, h_2 = .7, H = 75$ (B): $h_1 = .6, h_2 = .7, H = .85,$

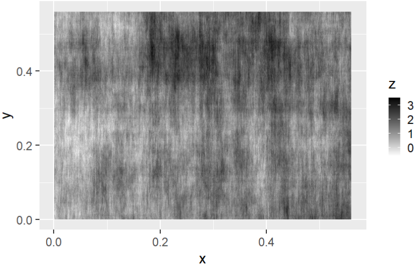
Simulation Results





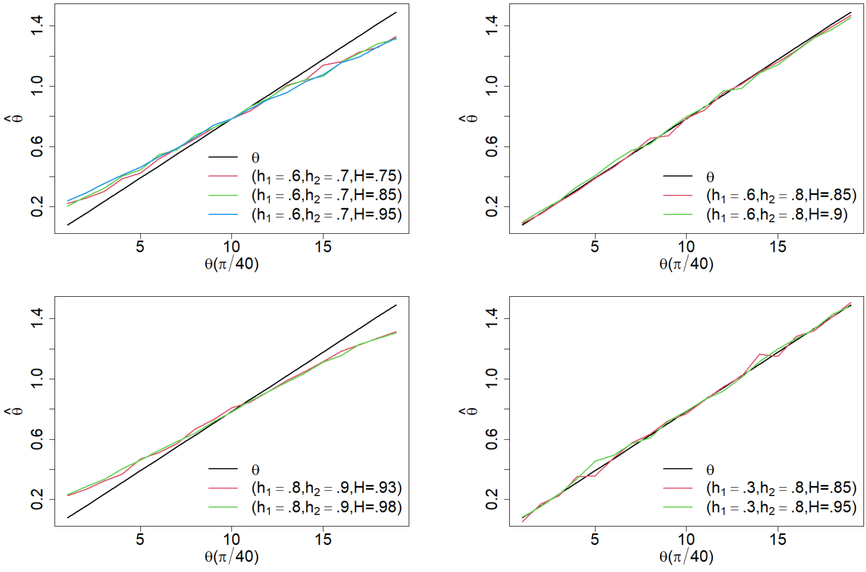
(C): $h_1 = .1, h_2 = .3, H = .4$

(D): $h_1 = .1, h_2 = .3, H = .8$

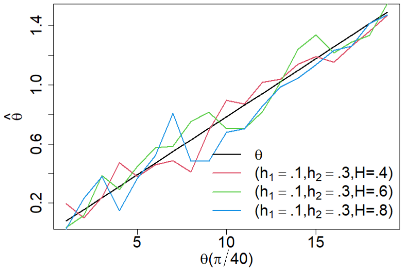


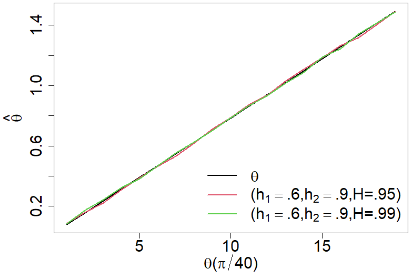
(E): $h_1 = .2, h_2 = .3, H = .4$

Simulation Results

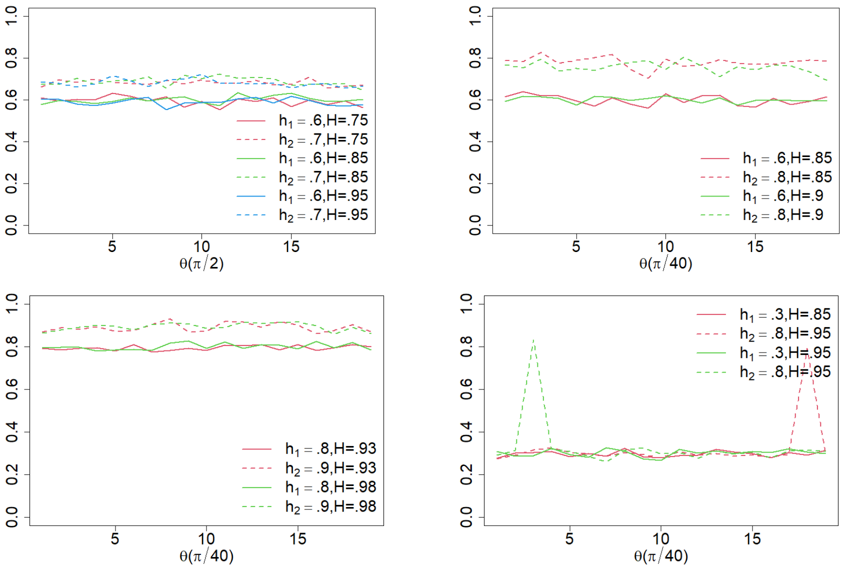


Simulation Results

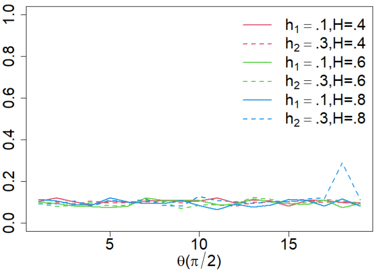


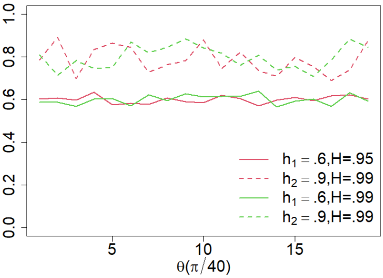


Simulation Results



Simulation Results





References

- Biermé, H., Meerschaert, M.M., and Scheffler, H.-P. (2007) *Operator scaling stable random fields*. Stochastic Processes and their Applications, 117, 312-332.
- Bierme, H. and Lacaux, C. (2018) *Fast and exact synthesis of some operator scaling Gaussian random fields*. Applied and Computational Harmonic Analysis.
- Didier, G. and Pipiras, V. (2011) *Integral representations and properties of operator fractional Brownian motions*. Bernoulli, 17, 1-33.
- Dietrich, C.R. and Newsam, G.N. (1997) *Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix*. SIAM Journal on Scientific Computing, 18, 1088-1107.
- Lee, J. (2020) *Online supplement: Hurst estimation for operator scaling random fields*.

The End