L a b o r a t ó r i o
N a c i o n a l de
C o m p u t a ç ã o
C i e n t í f i c a

COMOD
Coordenação de Modelagem Computacional

Relatório de Atividades

Servidor: Marcio Borges

Petrópolis-RJ
October 6, 2023

1 KARHUNEN-LOÈVE EXPANSION

Due to incomplete knowledge about the rock properties that show variability at multiple length scales, input parameters such as the permeability field, $\boldsymbol{\kappa}(\mathbf{x}, \omega)$, are treated as random space functions with statistics inferred from geostatistical models (here $\mathbf{x} = (x_1, x_2, x_3)^\mathsf{T} \in \mathbb{R}^3$ and ω is a random element in the probability space). In line with **?** and **?** the permeability field is modeled as a log-normally distributed random space function

$$\boldsymbol{\kappa}(\mathbf{x}, \omega) = \beta \exp\Big[\rho \mathsf{Y}(\mathbf{x}, \omega)\Big], \tag{1}$$

where $\beta, \rho \in \mathbb{R}^+$ and $\mathsf{Y}(\mathbf{x}, \omega) \sim \mathbb{N}\left(\mu_\mathsf{Y}, \boldsymbol{\mathcal{C}}_\mathsf{Y}\right)$ is a Gaussian random field characterized by its mean $\mu_\mathsf{Y} = \langle \mathsf{Y} \rangle$ and two-point covariance function

$$\boldsymbol{\mathcal{C}}_\mathsf{Y}(\mathbf{x}, \mathbf{y}) = \mathsf{Cov}\left(\mathsf{Y}(\mathbf{x}), \mathsf{Y}(\mathbf{y})\right) = \mathsf{E}\Big[\big(\mathsf{Y}(\mathbf{x}) - \langle \mathsf{Y}(\mathbf{x}) \rangle\big)\big(\mathsf{Y}(\mathbf{y}) - \langle \mathsf{Y}(\mathbf{y}) \rangle\big)\Big]. \tag{2}$$

Moreover, in this work, Y is a second-order stationary process (**?**), that is:

$$\begin{aligned} \langle \mathsf{Y}(\mathbf{x}) \rangle &= \mu_\mathsf{Y}, \quad \text{(constant)} \\ \boldsymbol{\mathcal{C}}_\mathsf{Y}(\mathbf{x}, \mathbf{y}) &= \boldsymbol{\mathcal{C}}_\mathsf{Y}\left(\|\mathbf{x} - \mathbf{y}\|\right) = \boldsymbol{\mathcal{C}}_\mathsf{Y}\left(d\right). \end{aligned} \tag{3}$$

The Gaussian field Y can be represented as a series expansion involving a complete set of deterministic functions with correspondent random coefficients using the Karhunen-Loève (KL) expansion proposed independently by **?** and **?**. It is based on the eigen-decomposition of the covariance function. Depending on how fast the eigenvalues decay one may be able to retain only a small number of terms in a truncated expansion and, consequently, this procedure may reduce the search to a smaller parameter space. In uncertainty quantification methods for porous media flows, the KL expansion has been widely used to reduce the number of parameters used to represent the permeability field (**??????**). Another advantage of KL expansion lies on the fact that it provides orthogonal deterministic basis functions and uncorrelated random coefficients, allowing for the optimal encapsulation of the information contained in the random process into a set of discrete uncorrelated random variables (**?**). This remarkable feature can be used to simplify the Metropolis-Hastings McMC Algorithm in the sense of the search may be performed in the space of discrete uncorrelated random variables ($\boldsymbol{\theta}$), no longer in the space of permeabilities which have a more complex statistical structure.

Here we recall the basic facts about the Karhunen-Loève expansion. Consider a random field $\mathsf{Y}(\mathbf{x}, \omega)$ defined on a probability space $(\Omega, \mathcal{A}, \mathcal{P})$ composed by the sample space, the ensemble of events and a probability measure, respectively, and indexed on a bounded domain $\mathcal{D} \in \mathbb{R}^3$. The process Y can be expressed as

$$\mathsf{Y}(\mathbf{x}, \omega) = \langle \mathsf{Y}(\mathbf{x}) \rangle + \sum_{i=1}^{\infty} \sqrt{\lambda_i} \phi_i(\mathbf{x}) \theta_i(\omega), \tag{4}$$

where λ_i and ϕ_i are the eigenvalues and eigenfunctions of the covariance function $\mathcal{C}_{\mathsf{Y}}(\mathbf{x},\mathbf{y})$, respectively. By definition, $\mathcal{C}_{\mathsf{Y}}(\mathbf{x},\mathbf{y})$ is bounded, symmetric and positive definite and has the following eigen-decomposition:

$$\mathcal{C}_{\mathsf{Y}}(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{\infty} \lambda_i \phi_i(\mathbf{x})\phi_i(\mathbf{y}).$$

The eigenvalues and eigenfunctions of Eq. (4) are the solution of the homogeneous Fredholm integral equation of second kind given by

$$\int_{\mathcal{D}} \mathcal{C}_{\mathsf{Y}}(\mathbf{x},\mathbf{y})\phi(\mathbf{x})d\mathbf{x} = \lambda\phi(\mathbf{y}). \tag{5}$$

The solution of Eq. (5) forms a complete set of a square-integrable orthogonal eigenfunctions that satisfy the equation

$$\int_{\mathcal{D}} \phi_i(\mathbf{x})\phi_j(\mathbf{x}) = \delta_{ij},$$

in which δ_{ij} is the Kronecker-delta function. $\theta_i(\omega)$ is a set of independent random variables which can be expressed as

$$\theta_i(\omega) = \frac{1}{\sqrt{\lambda_i}} \int_{\mathcal{D}} \tilde{\mathsf{Y}}\phi_i(\mathbf{x})d\mathbf{x},$$

where $\tilde{\mathsf{Y}} = \mathsf{Y} - \langle \mathsf{Y} \rangle$ is the fluctuation. For practical implementations of the KL expansion the eigenvalues are arranged from the largest to smallest and the series is approximated by a finite number of terms, say the first m, giving

$$\mathsf{Y}(\mathbf{x},\omega) \approx \langle \mathsf{Y}(\mathbf{x}) \rangle + \sum_{i=1}^{\mathsf{m}} \sqrt{\lambda_i}\phi_i(\mathbf{x})\theta_i(\omega). \tag{6}$$

The corresponding covariance function is given by

$$\mathcal{C}_{\hat{\mathsf{Y}}}(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{\mathsf{m}} \lambda_i \phi_i(\mathbf{x})\phi_i(\mathbf{y}).$$

The factors affecting the convergence of the Karhunen-Loève series are the ratio of the length of the process over correlation parameter, the form of the covariance function, and the solution method for the eigensolutions of the covariance function (see **?**). Next we discuss the field conditioning using the KL expansion.

The number of terms used in the series can be chosen based on the energy represented by the sum of the eigenvalues. Then we define the relative energy for n terms as

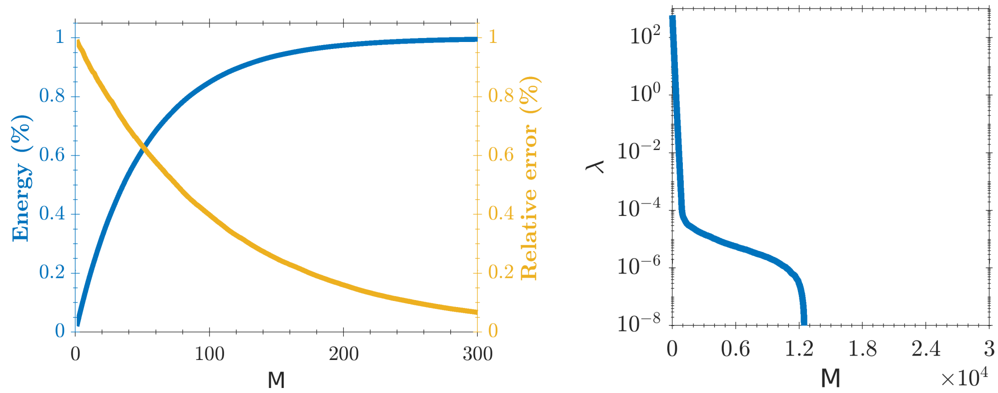
$$\mathsf{ER_n} = \frac{\sum_{i=1}^{n} \lambda_i}{\sum_{j=1}^{m\to\infty} \lambda_j}. \tag{7}$$

The eigenvalues associated with the eigenfunctions provide a measure of the energy contained in the respective mode. To quantify the energy involved in a truncated KL expansion (with M modes) we define the total relative energy as

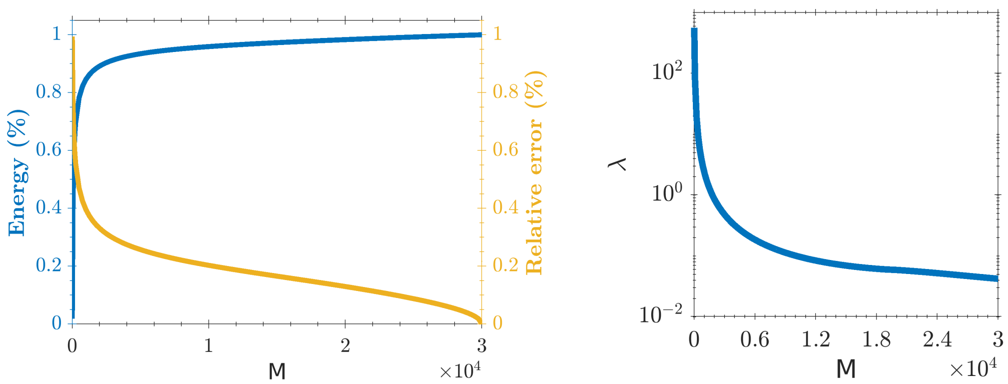
$$\mathcal{E}_{\mathsf{M}} = \frac{\sum_{j=1}^{\mathsf{M}} \lambda_j}{\sum_{i=1}^{m \to \infty} \lambda_i}, \tag{8}$$

in which the denominator will be approximated by $m = 30,000$ eigenvalues.

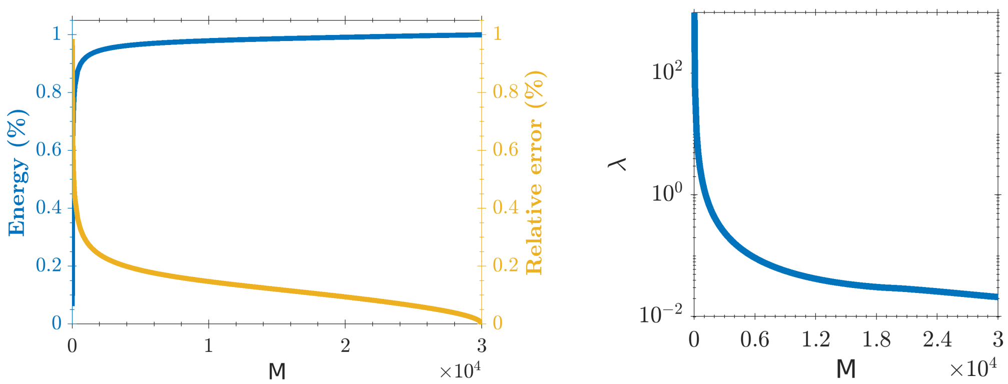
The Fig. 1 shows the behavior of the eigenvalues, the energy contained, and the average error as a function of the number of modes (M) for the three covariance functions considered in this study. The decay of the eigenvalues of the square exponential case is much faster than the exponential case. The latter requires a more significant number of modes to capture the same energy when compared to the first one.



(a) Squared exponential covariance ($\ell = 0.1$)

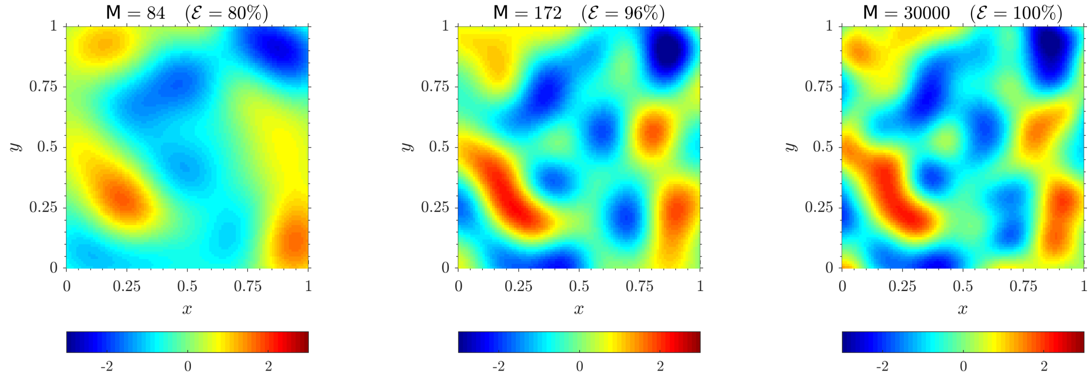


(b) Exponential covariance ($\ell = 0.1$)

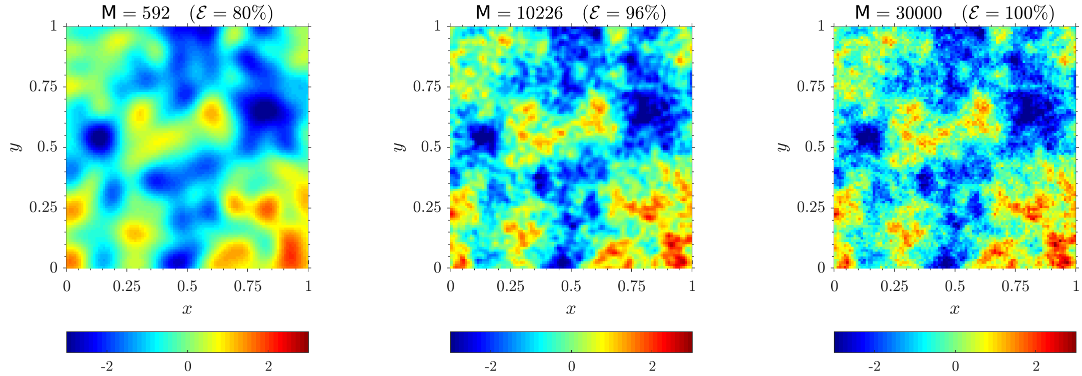


(c) Exponential covariance ($\ell = 0.2$)

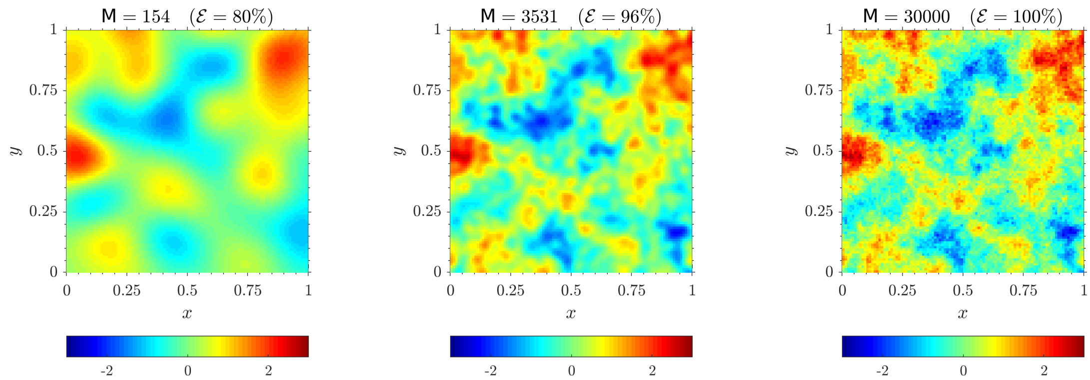
Figure 1: Decay of eigenvalues and contained energy as a function of the number of terms in the expansion.



(a) Squared exponential, $\ell = 0.1$



(b) Exponential, $\ell = 0.1$

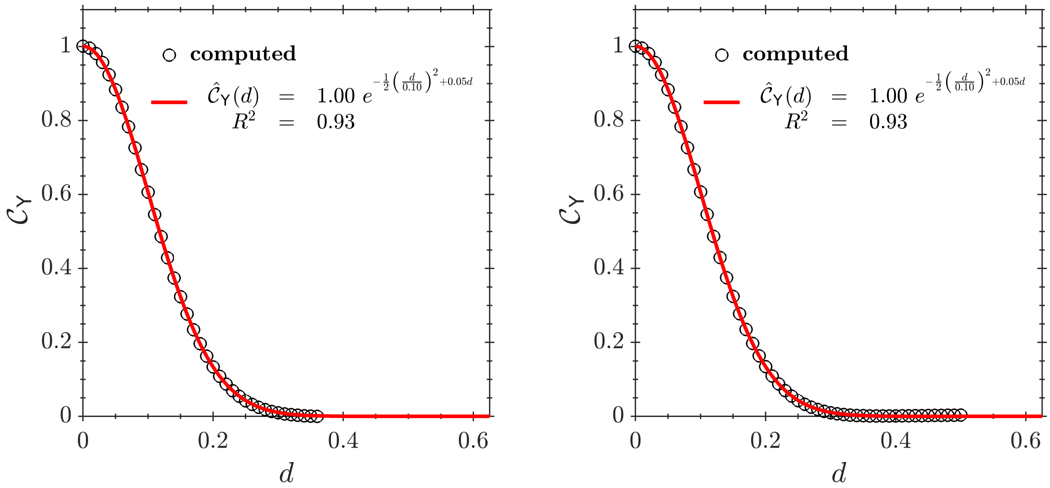


(c) Exponential, $\ell = 0.2$

Figure 2: Random fields generated with different numbers of modes M (or, equivalently, amount of energy).

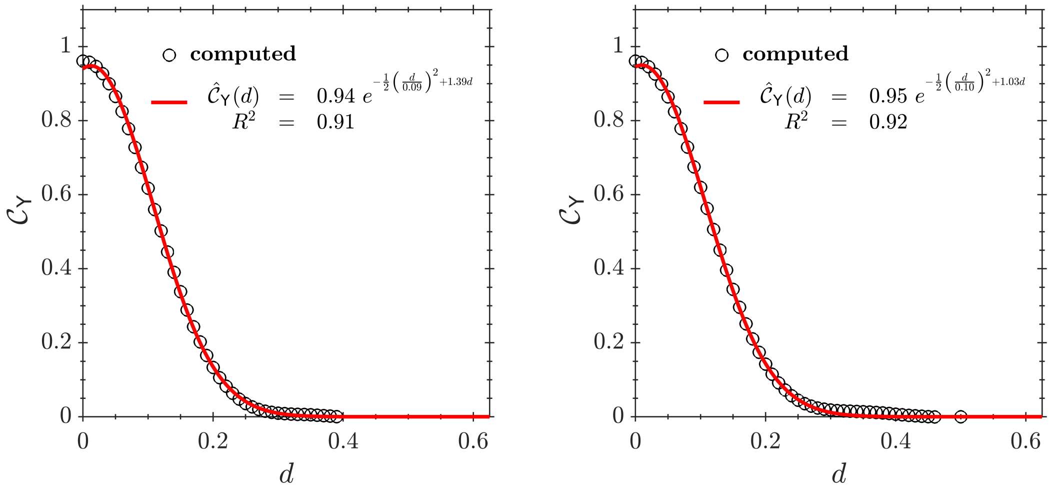
Table 1: Number of KL expansion terms needed to obtain an given energy level

Energy	M			**Mean relative error**		
(%)	Squared Exp. ($\ell = 0.1$)	Exponential ($\ell = 0.1$)	Exponential ($\ell = 0.2$)	Squared Exp. ($\ell = 0.1$)	Exponential ($\ell = 0.1$)	Exponential ($\ell = 0.2$)
80	84	592	154	0.46	0.45	0.47
90	122	2,323	612	0.32	0.32	0.33
94	150	5,760	1,655	0.25	0.25	0.25
96	172	10,226	3,531	0.21	0.20	0.21
98	211	18,327	10,239	0.14	0.14	0.15
100	~30,000	~30,000	~30,000	0.00	0.00	0.00



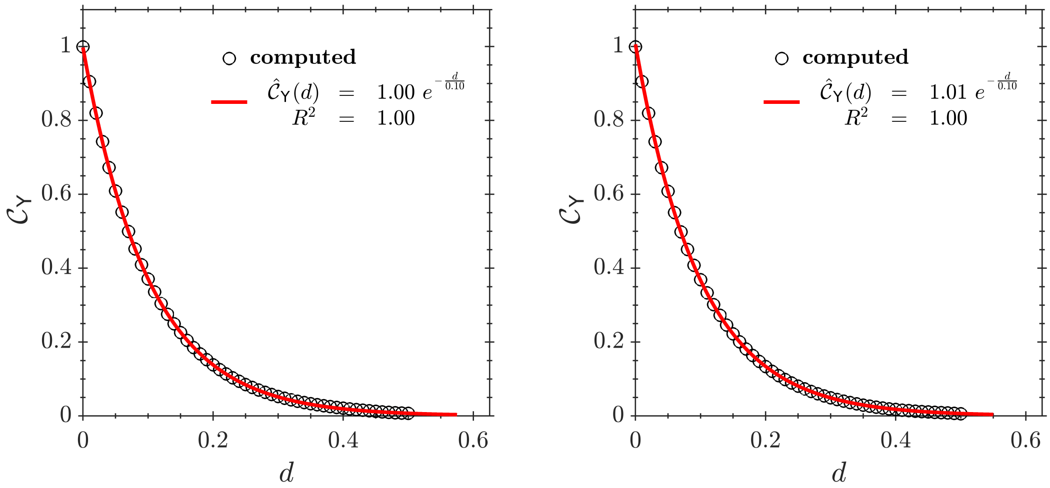
(a) x direction (b) y direction

Figure 3: Estimated covariance in a sample of $5,000$ fields generate by `KLE` with $30,000$ terms. Squared exponential case with $\ell = 0.1$.



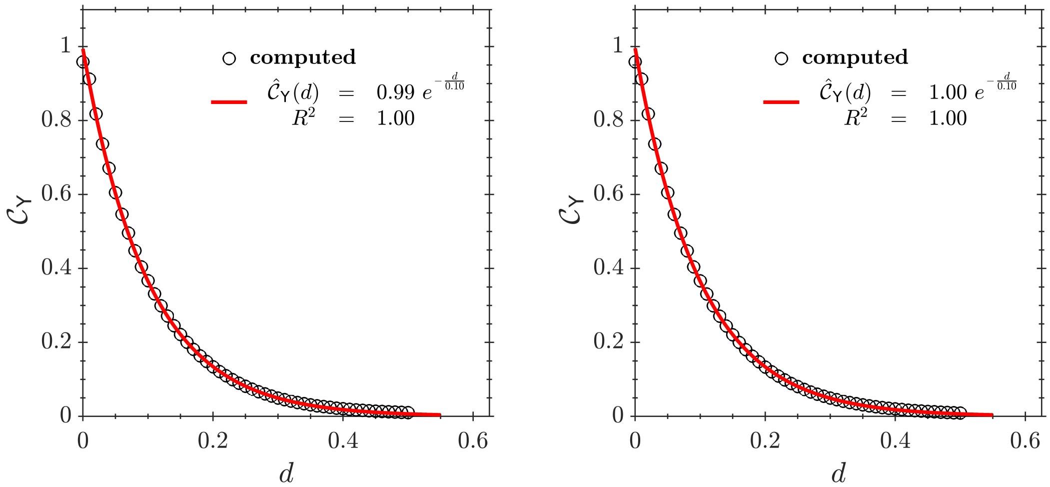
(a) x direction (b) y direction

Figure 4: Estimated covariance in a sample of $5,000$ fields generate by KLE with 172 terms. Squared exponential case with $\ell = 0.1$.



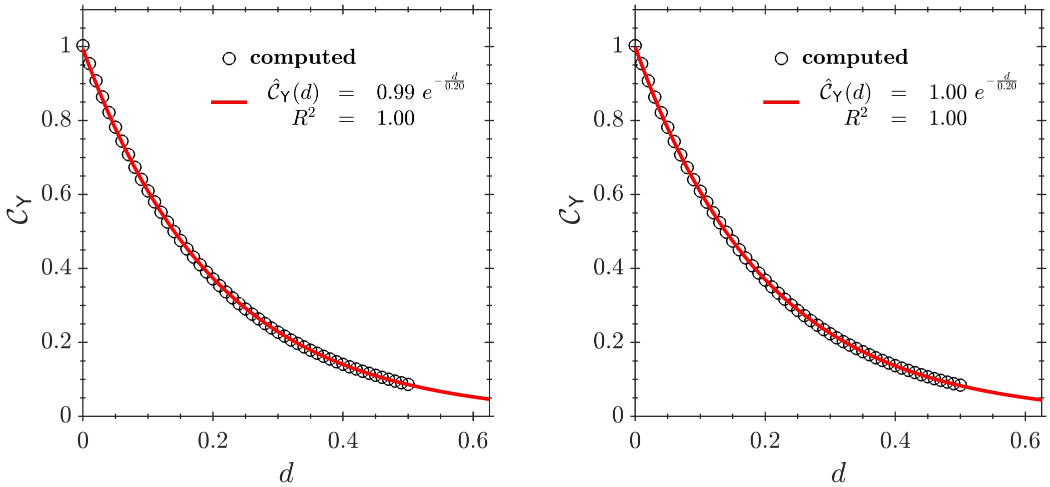
(a) x direction (b) y direction

Figure 5: Estimated covariance in a sample of $5,000$ fields generate by KLE with $30,000$ terms. Exponential case with $\ell = 0.1$.



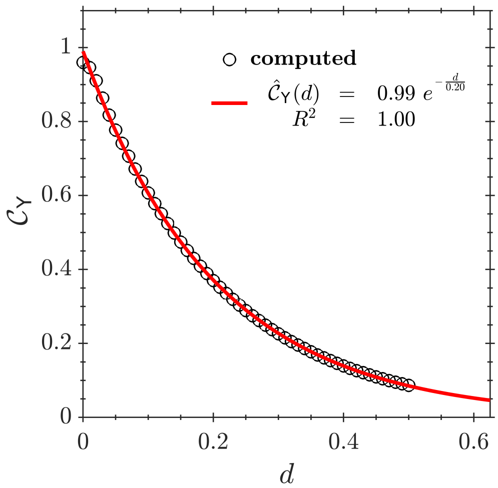
(a) x direction (b) y direction

Figure 6: Estimated covariance in a sample of $5,000$ fields generate by KLE with $10,226$ terms. Exponential case with $\ell = 0.1$.

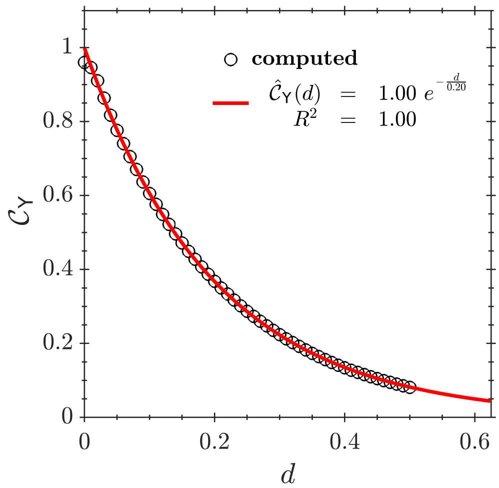


(a) x direction (b) y direction

Figure 7: Estimated covariance in a sample of $5,000$ fields generate by KLE with $30,000$ terms. Exponential case with $\ell = 0.2$.



(a) x direction



(b) y direction

Figure 8: Estimated covariance in a sample of $5,000$ fields generate by KLE with $3,531$ terms. Exponential case with $\ell = 0.2$.

References

G. Dagan. *Flow and transport in porous formations*. Springer-Verlag, 1989.

N. N. Das, B. P. Mohanty, and Y. Efendiev. Characterization of effective saturated hydraulic conductivity in an agricultural field using Karhunen-Loève expansion with the Markov chain Monte Carlo technique. *Water Resources Research*, 46, W06521 2010. doi: 10.1029/2008WR007100.

Y. Efendiev, A. Datta-gupta, V. Ginting, X. Ma, and B. Mallick. An efficient two-stage Markov chain Monte Carlo method for dynamic data integration. *Water Resources Research*, 41: 12423, 2005.

Y. Efendiev, T. Hou, and W. Luo. Preconditioning Markov Chain Monte Carlo Simulations Using Coarse-Scale Models. *SIAM J. Sci. Comput.*, 28:776–803, 2006.

L. W. Gelhar. *Stochastic subsurface hydrology*. Englewood Cliffs. Prentice-Hall, 1993.

R. Ghanem and P.D. Spanos. *Stochastic Finite Element: A Spectral Approach*. Springer, New York, 1991.

V. Ginting, F. Pereira, M. Presho, and S. Wo. Application of the two-stage Markov chain Monte Carlo method for characterization of fractured reservoirs using a surrogate flow model. *Computational Geosciences*, 15:691–707, 2011.

V. Ginting, F. Pereira, and A. Rahunanthan. Multiple Markov chains Monte Carlo approach for flow forecasting in porous media. *Procedia Computer Science*, 5:707–716, 2012.

S. P. Huang, S. T. Quek, and K. K. Phoon. Convergence study of the truncated Karhunen-Loève expansion for simulation of stochastic processes. *International Journal for Numerical Methods in Engineering*, 52(9):1029–1043, 2001.

K. Karhunen. Zur spektraltheorie stochastischer prozesse. *Ann. Acad. Sci. Fennicae*, 1946.

M. M. Loève. *Probability Theory*. Princeton, N.J., 1955.

A. Mondal, Y. Efendiev, B. Mallick, and A. Datta-Gupta. Bayesian uncertainty quantification for flows in heterogeneous porous media using reversible jump Markov chain Monte Carlo methods. *Advances in Water Resources*, 33(3):241 – 256, 2010.

INDEX