

### PROBLEM 1

Let random variable  $X$  follow the Chebyshev probability density function

$$f_X(x) = \pi^{-1}(1 - x^2)^{-1/2}, \quad x \in (-1, +1). \quad (1)$$

Use the inversion method to generate 10,000 samples of  $X$ . To verify the quality of the generated samples, compare the empirical density and distribution functions of  $X$  (based on the generated samples) to the exact values. You may use MATLAB functions `ksdensity` and `ecdf` to compute the empirical density and distribution functions, respectively.

### METHODS AND DISCUSSION

In the inversion method, we integrate the pdf above and calculate the result's inverse,

$$F_X^{-1}(u) = \sin\left(\pi\left[u - \frac{1}{2}\right]\right), \quad (2)$$

from which random variables  $X_i$  can be generated using

$$X_i = F_X^{-1}(U_i), \quad U_i \sim U(0, 1), \quad i = 1, 2, \dots \quad (3)$$

As can be seen in Figure 1, the samples generated by the inversion method result in a pdf and CDF that are in excellent agreement with their analytical behavior.

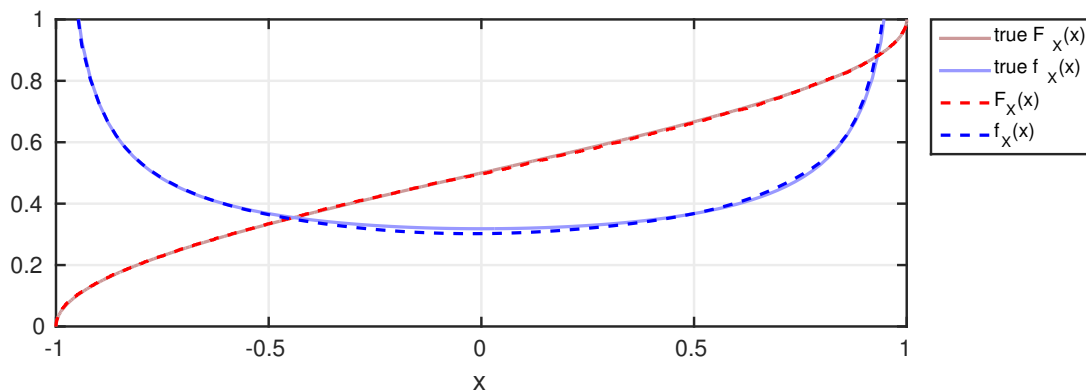
### PROBLEM 2

$G(x, \omega)$  is a Gaussian random process defined on  $(0, 1)$ . The mean and covariance functions of the Gaussian process  $G(x, \omega)$  are

$$\langle G(x) \rangle = 1.0, \quad x \in (0, 1), \quad (4)$$

and

$$C_{GG}(x_1, x_2) = \sigma^2 \exp\left(\frac{-|x_1 - x_2|}{\ell}\right), \quad (x_1, x_2) \in (0, 1) \times (0, 1), \quad (5)$$



**Figure 1:** Cumulative density function  $F_X(x)$  and probability density function  $f_X(x)$  based on analytical arguments and the inversion method using  $N = 10,000$  samples.

respectively. We would like to compute the eigen-pairs of  $C_{GG}(x_1, x_2)$  using the Galerkin approach with linear finite elements (or the Nyström approach) and compare them with the exact values provided in the notes.

For  $(\sigma, \ell)$  of  $(2, 2)$  and  $(2, 0.2)$ , we compute the number of terms required in the KL expansion of  $G(x, \omega)$  such that the mean-square error of truncation is roughly 10%. Then, we verify the increasing the finite element resolution reduces error of the Galerkin estimated eigenvalues. Finally, we draw 100,000 samples of the random process, plot seven of them, and compute the mean and variance of all samples, comparing them to the given values.

## METHODS

Eigenvalues and eigenfunctions of the system are found by solving the generalized eigenproblem

$$\mathbf{CZ} = \mathbf{MZA}, \quad (6)$$

where

$$\begin{aligned} \mathbf{C}_{ij} &= \int_0^1 \int_0^1 C_{GG}(x_1, x_2) \psi_i(x_2) \psi_j(x_1) dx_1 dx_2 \\ \mathbf{M}_{ij} &= \int_0^1 \psi_i(x) \psi_j(x) dx \\ \mathbf{Z}_{ij} &= z_j^{(k)} \\ \mathbf{\Lambda}_{ij} &= \delta_{jk} \lambda_k \end{aligned} \quad (7)$$

are the matrices that result from the Galerkin formulation of the eigenvalue problem. In our case, we use piecewise linear basis functions  $\psi_i(x)$ , with  $i = 1, 2, \dots, N$ .  $\mathbf{M}_{ij}$  is computed analytically, and is proportional to the spacing between nodes in one-dimension,  $\Delta x$ .  $\mathbf{G}_{ij}$  is computed on an element-wise basis in two-dimensions. The element-wise contributions to the integral are assembled in the standard manner.

These eigenvalues  $\lambda_k$  and eigenfunctions  $\phi_k(x)$  can be verified against the analytical solution presented in the course notes. With a sufficiently accurate eigenproblem solution, we can write the optimal orthogonal expansion of a random (not necessarily Gaussian) process as a Karhunen-Loève Expansion (KLE):

$$G_n(x, \omega) \equiv \sum_{i=1}^{\infty} \sqrt{\lambda_i} \phi_i(x) Y_i(\omega), \quad Y_i \stackrel{iid}{\sim} N(0, 1). \quad (8)$$

Since it is impractical to sum to  $\infty$ , only the first  $d$  are used. One common way of determining a sufficient and practical cut-off  $d$  is to reduce below some tolerance  $\alpha$  the mean-square norm of the approximation:

$$d = \min_j \left\{ \frac{\sum_1^j \lambda_i}{\sum_1^{\infty} \lambda_i} \leq \alpha \right\}. \quad (9)$$

Finally, realizations of the random process are generated with

$$X(x) = \langle X(x) \rangle + \sum_{i=1}^d \sqrt{\lambda_i} \phi_i(x) Y_i. \quad (10)$$

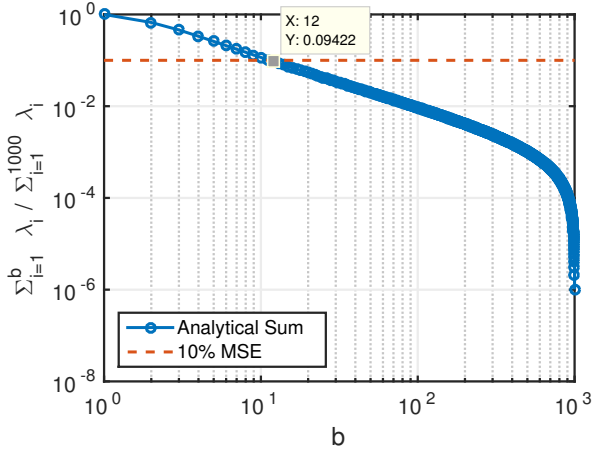
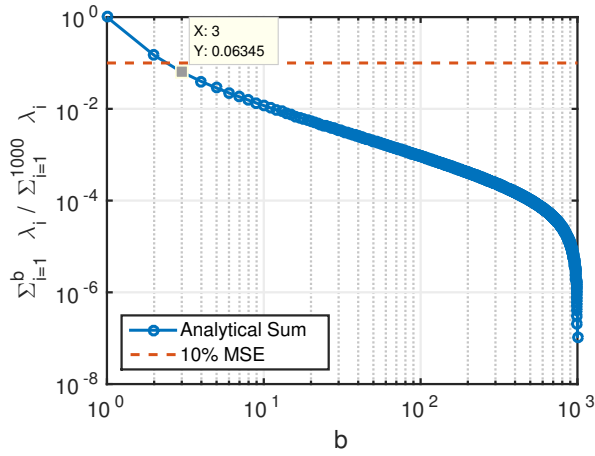
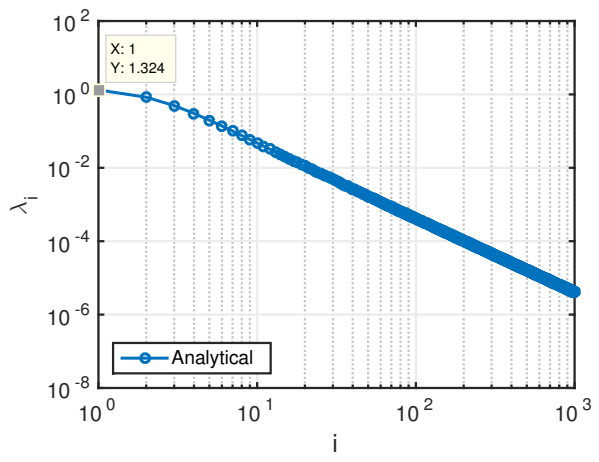
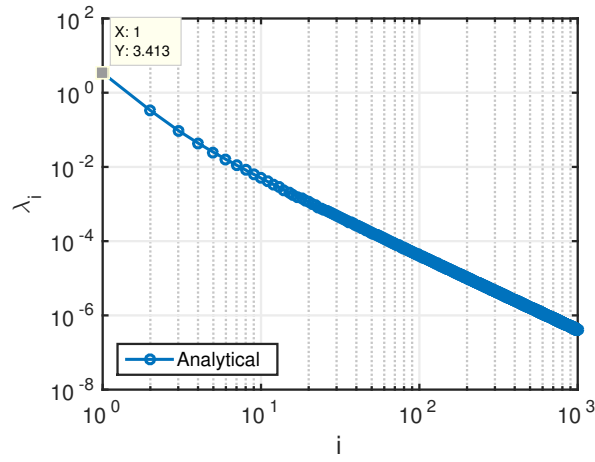
## DISCUSSION

In Figure 2, we see that a longer correlation length  $\ell$  is associated with a more rapid decay in eigenvalues. This is expected, since because more eigenfunctions are necessary to capture the higher-frequency dynamics of systems with small correlation length, their eigenvalues decay less rapidly.

In Figure 3, as expected, increasing the number of finite elements decreases the relative error of our approximated eigenvalues. The larger the magnitude of  $\lambda_i$ , the lower its relative error. Smaller correlation length  $\ell$  corresponds to reduced eigenvalue accuracy for a given number of finite elements. Matlab's `eigs` solver only solves for  $N-2$  eigenpairs without producing warnings.

Finally, in Figure 4, we see the high-frequency behavior implied by smaller  $\ell$  manifest in the seven individual realizations. For both  $\ell = 2$  and  $\ell = 0.2$ , the sample mean matches the true mean very well; this is a direct consequence of reducing the mean square error to  $< 10\%$ . The variance is less accurate, however. Larger  $\ell$  results in a better approximation of variance, but this could be due to the mean-square error being reduced to 6.3% for  $\ell = 2$ , compared to 9.4% for  $\ell = 0.2$ . These are the highest values below 10% for each case.

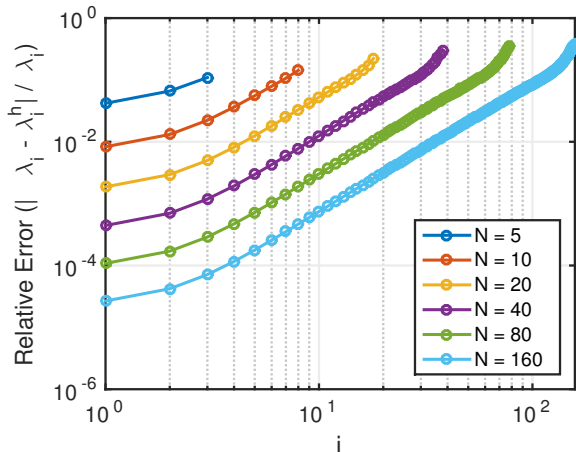
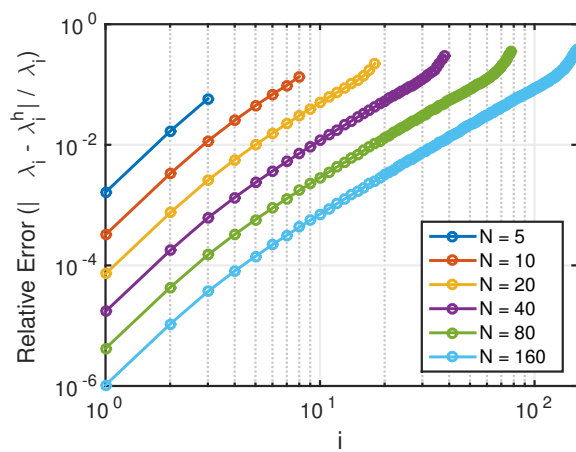
To investigate further, Figure 5 increases the KLE cutoff to include  $b = 17$  terms for the  $\ell = 0.2$  case, matching the  $\ell = 2$  case's MSE within 1% using (9). Despite nearly identical MSE's, the  $\ell = 0.2$  variance is still  $\sim 2.5\%$  further away from the true value of  $\sigma^2 = 4$  than the  $\ell = 2$  case's variance. I believe this results from variance being a higher-order statistic than the mean, which makes it inherently more difficult to capture with a small number of KLE terms. This effect is pronounced when the process has a shorter correlation length.



(a)  $\ell = 2$

(b)  $\ell = 0.2$

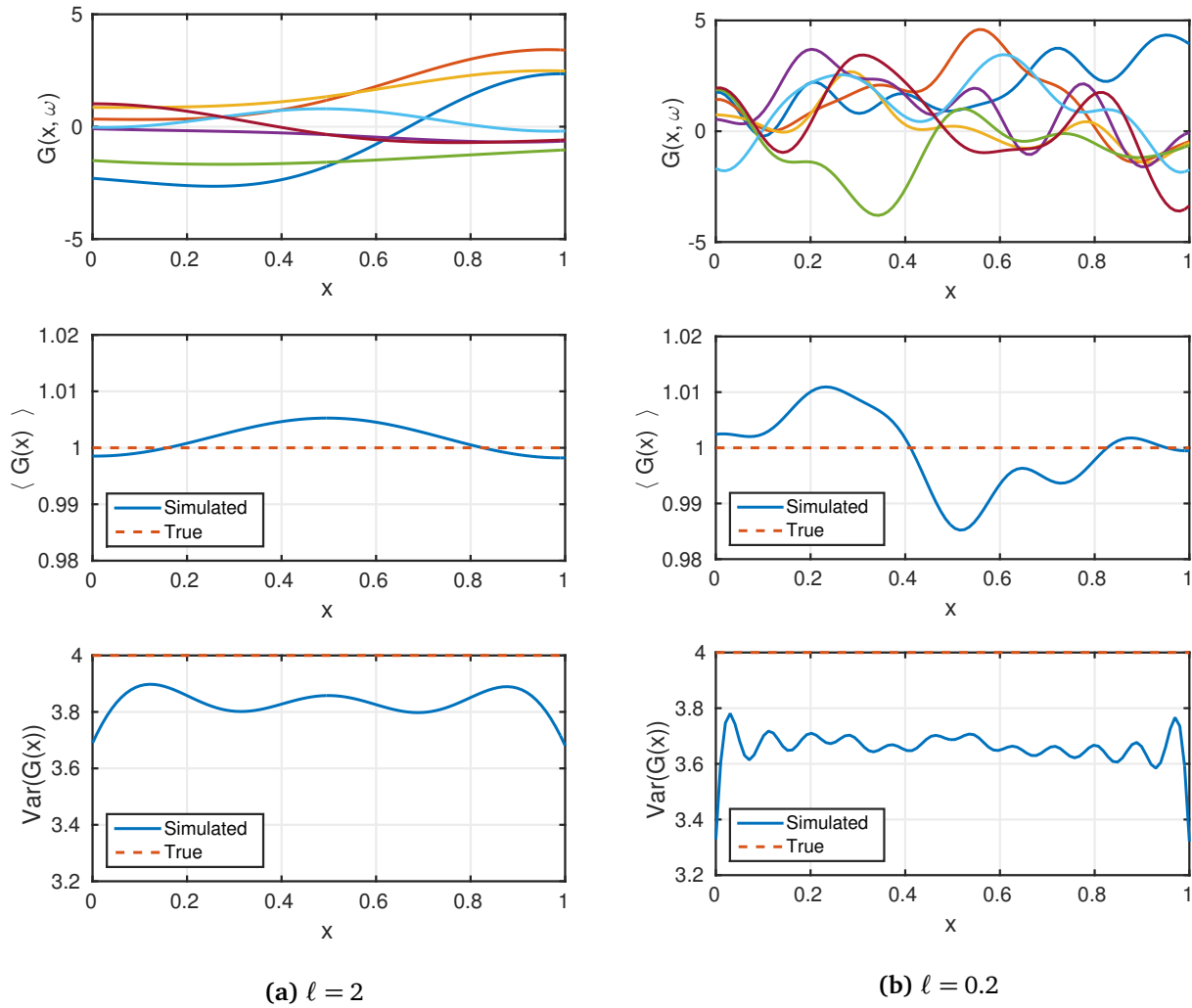
**Figure 2:** 1,000 analytical eigenvalues and normalized sum of the largest  $b$  eigenvalues using  $\sigma = 2$ . Mean-square error of truncation  $< 10\%$  is first achieved with (a)  $b = 3$  and (b)  $b = 12$  terms in the KL expansion.



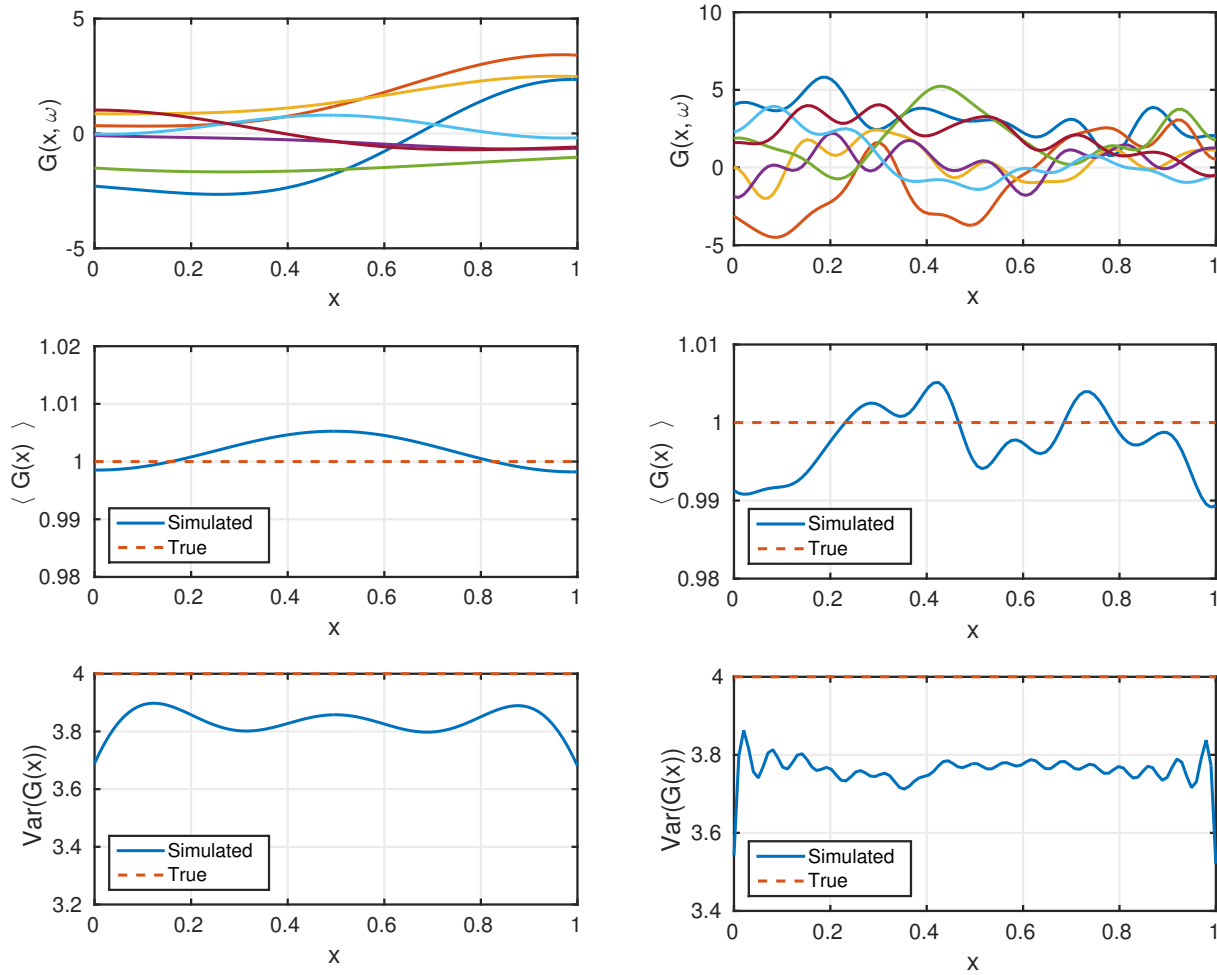
(a)  $\ell = 2$

(b)  $\ell = 0.2$

**Figure 3:** Increasing the number of finite elements ( $N_{\text{el}} = [N - 1]^2$ ) increases accuracy of eigenvalues. For both plots,  $\sigma = 2$ .



**Figure 4:** From top to bottom: seven realizations of the random process  $G(x, \omega)$ , sample mean  $\langle G(x) \rangle$ , and sample variance  $\text{Var}(G(x))$ , using 100,000 samples. For both plots,  $\sigma = 2$  and the KL expansion is used with  $b(\ell)$  terms.



(a)  $\ell = 2$ , MSE 6.34%

(b)  $\ell = 0.2$ , MSE 6.40%

**Figure 5:** Same as Figure 4, but incorporating  $b = 17$  terms in the KL expansion of  $\ell = 0.2$  to match MSE with the  $\ell = 2$  case. For both plots,  $\sigma = 2$ . Variance is still approximated better for case with longer correlation length.