Experimental Design for Left Ventricular Biomechanics

Diogo Amaro

October 8, 2025

1 Experimental Design

To ensure thorough coverage of the input parameter space, it is important to design a framework capable of exploring every possible combination as effectively as possible. Experimental design refers to the systematic planning of physical experiments to efficiently explore a parameter space and extract meaningful insights while minimizing both number of simulations and time. Basically, a good experimental design should aim to minimize the number of runs needed to acquire as much information as possible [1]. To achieve this, various techniques have been developed and applied, four of which were initially considered and are briefly described below.

1.1 Rectangular Grid

Rectangular grid sampling involves discretizing each dimension of the input space into evenly spaced intervals and evaluating all possible combinations of these values across dimensions [2]. For a d-dimensional parameter space with n discretization points per dimension, the total number of sample points is given by n^d. Let each parameter $q_i \in [a_i, b_i]$, for $i = 1, \ldots, d$. Then, the values for q_i are computed as

$$q_i^{(j)} = a_i + \frac{j-1}{n-1}(b_i - a_i), \quad j = 1, \ldots, n. \tag{1}$$

The Cartesian product of these values across all dimensions forms the complete grid of input vectors. While straightforward, this method presents significant limitations for high-dimensional parameter spaces due to the curse of dimensionality. For a 4-dimensional parameter space, achieving reasonable resolution requires n^4 sample points, where n represents the number of discretization points per dimension. Even modest discretization (e.g., $n = 20$) yields 160,000 required samples, making the approach computationally prohibitive for ABAQUS simulations. Because we defined our number of simulations at 10,000, the exponential scaling of computational requirements make rectangular grid sampling impractical for this study's objectives.

1.2 Uniform Distribution

Uniform random sampling, also known as Monte Carlo sampling, consists of independently drawing each parameter from a uniform distribution over its domain. For each parameter $q_i \in [a_i, b_i]$, a sample is drawn as

$$q_i^{(j)} \sim \mathcal{U}(a_i, b_i), \quad j = 1, \ldots, N. \tag{2}$$

This results in N independent samples

$$\mathbf{q}^{(j)} = \left(q_1^{(j)}, q_2^{(j)}, \ldots, q_d^{(j)} \right) \tag{3}$$

Although sampling from a uniform distribution is conceptually simple, it lacks the stratification mechanisms of more advanced sampling methods that are capable of better capturing the nonlinear relations between parameters.

1.3 Sobol Sequence

Sobol sequences are quasi-random low-discrepancy sequences designed to generate points that uniformly fill a multidimensional space [3]. They are commonly used for numerical integration and design of experiments due to their improved space-filling properties over purely random sampling. Mathematically, a Sobol sequence generates a sequence of points:

$$\mathbf{x}_i = (x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(d)}), \quad i = 1, 2, \ldots, N, \tag{4}$$

where each component $x_i^{(j)} \in [0, 1]$ is constructed using direction numbers and bitwise operations involving the binary representation of the index i [4]. The key idea is to minimize the discrepancy D_N of the point set, which quantifies how uniformly the samples cover the domain:

$$D_N = \sup_{B \subset [0,1]^d} \left| \frac{A(B; N)}{N} - \lambda(B) \right|, \tag{5}$$

where $A(B; N)$ is the number of sample points falling inside the axis-aligned box B, and $\lambda(B)$ is its Lebesgue measure.

Despite the theoretical plausibility, Sobol sampling is not particularly advantageous in our case, as it assumes independent and uniformly distributed dimensions, whereas the physiological parameter space considered here represents complex interactions. Moreover, if the response surface is locally nonlinear or exhibits abrupt changes, as is often the case in biomechanical systems, Sobol points may not allocate sufficient resolution to those region [5] . However, other similar studies have achieved satisfactory simulation accuracy by sampling from a Sobol sequence [6].

1.4 Latin Hypercube Sampling

The Latin hypercube sampling is a stratified sampling technique, first introduced by McKay et al. [7], to generate diverse and well-distributed parameter combinations across the input space of reduced material parameters $\mathbf{q} = (q_1, q_2, q_3, q_4) \in \mathbb{R}^4$.

Let the parameter space be defined as a d-dimensional hypercube, with each parameter $q_i \in [a_i, b_i]$, for $i = 1, \ldots, d$. Suppose we wish to generate N samples. In Latin hypercube sampling, each interval $[a_i, b_i]$ is divided into N equally probable, non-overlapping intervals. Then, one value is randomly sampled from each interval without replacement. The sampled values along each dimension are then randomly paired across dimensions to form N d-dimensional input vectors. Formally, for each variable i, define the partition:

$$P_{ij} = \left[a_i + \frac{j-1}{N}(b_i - a_i), \ a_i + \frac{j}{N}(b_i - a_i) \right], \quad j = 1, \ldots, N. \tag{6}$$

Then, for each i, randomly permute the indices $\{1, 2, \ldots, N\}$ to obtain a vector π_i. Sample a point $x_i^{(j)} \in P_{i, \pi_i(j)}$ uniformly, and construct the full sample:

$$\mathbf{q}^{(j)} = \left(x_1^{(j)}, x_2^{(j)}, \ldots, x_d^{(j)} \right), \quad j = 1, \ldots, N. \tag{7}$$

This process ensures that each marginal distribution is uniformly sampled, while achieving better coverage of the joint space compared to pure Monte Carlo sampling. However, for designs that depend strongly on the correlation of two variables, Latin hypercube sampling has little advantage over traditional Monte Carlo sampling [8] as much as it is guaranteed not to perform much worse. Specifically, Owen has proven that the variance of a n_t-point Latin hypercube sample, σ_{LHS}^2, is related to the variance of a traditional n_t-point Monte Carlo sample, σ_{MC}^2, by

$$\sigma_{LHS}^2 \leq \frac{n_t}{n_t - 1} \sigma_{MC}^2, \qquad \text{where } n_t > 1, \tag{8}$$

meaning that a Latin hypercube sample with n_t points never leads to a variance greater than a Monte Carlo sample with $n_t - 1$ points [9].
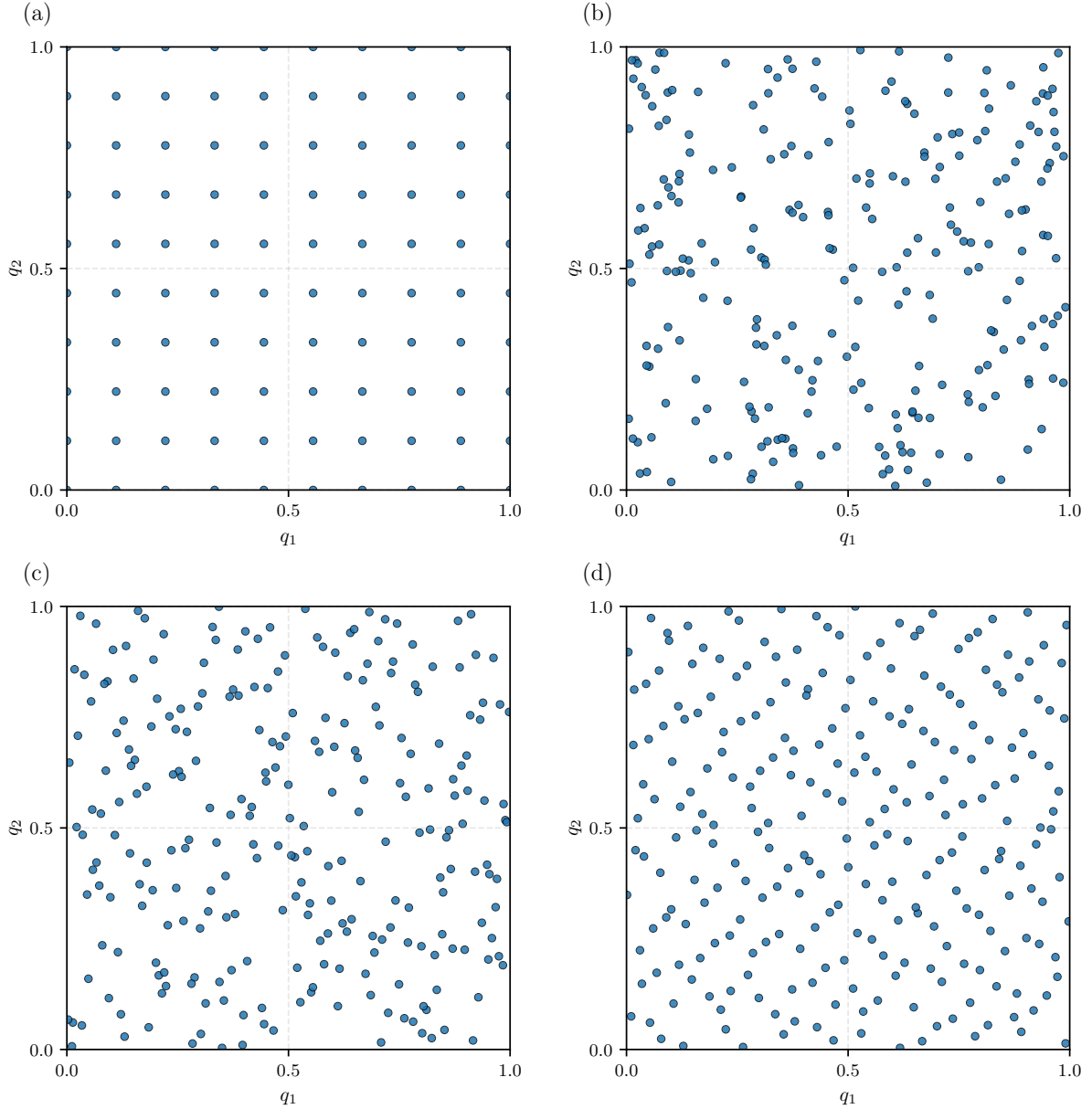

Figure 1: A comparison of different design choices for the training inputs. The plots show 100 points q_i in the two-dimensional space $[0, 1]^2$ using different design choices: (a) regular grid, (b) sampled from a uniform distribution, (c) Latin hypercube design and (d) from a Sobol sequence.

Given the above considerations, the input parameter vector $\mathbf{q} = (q_1, q_2, q_3, q_4)$ was sampled from the uniform domain $[0.1, 5]^4$ using Latin hypercube sampling with $N = 10,000$ samples. The parameter bounds were adopted from previous literature [6, 10], where they were shown to capture physiologically realistic variations in myocardial material properties. Latin hypercube sampling ensures that our surrogate models are trained on a dataset that broadly covers the entire biomechanical response space of the left ventricle. A representation of the mentioned sampling techniques is provided in Figure 1.

References

[1] K.-T. Fang, R. Li, and A. Sudjianto, *Design and modeling for computer experiments*. Chapman and Hall/CRC, 2005.

[2] D. P. Young, R. G. Melvin, M. B. Bieterman, F. T. Johnson, S. S. Samant, and J. E. Bussoletti, "A locally refined rectangular grid finite element method: application to computational fluid dynamics and computational physics," *Journal of Computational Physics*, vol. 92, no. 1, pp. 1–66, 1991.

[3] C. Lemieux, *Monte Carlo and Quasi-Monte Carlo Sampling*. 01 2009.

[4] S. Joe and F. Y. Kuo, "Constructing sobol sequences with better two-dimensional projections," *SIAM Journal on Scientific Computing*, vol. 30, no. 5, pp. 2635–2654, 2008.

[5] S. Burhenne, D. Jacob, and G. P. Henze, "Sampling based on sobol'sequences for monte carlo techniques applied to building simulations," in *Building Simulation 2011*, vol. 12, pp. 1816–1823, IBPSA, 2011.

[6] U. Noè, A. Lazarus, H. Gao, V. Davies, B. Macdonald, K. Mangion, C. Berry, X. Luo, and D. Husmeier, "Gaussian process emulation to accelerate parameter estimation in a mechanical model of the left ventricle: a critical step towards clinical end-user relevance," *Journal of the Royal Society Interface*, vol. 16, no. 156, p. 20190114, 2019.

[7] M. D. McKay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, no. 2, pp. 239–245, 1979.

[8] G. B. Rybicki and W. H. Press, "Interpolation, realization, and reconstruction of noisy, irregularly sampled data," *Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 398, no. 1, Oct. 10, 1992, p. 169-176.*, vol. 398, pp. 169–176, 1992.

[9] A. B. Owen, "Monte carlo variance of scrambled net quadrature," *SIAM Journal on Numerical Analysis*, vol. 34, no. 5, pp. 1884–1910, 1997.

[10] V. Davies, U. Noè, A. Lazarus, H. Gao, B. Macdonald, C. Berry, X. Luo, and D. Husmeier, "Fast parameter inference in a biomechanical model of the left ventricle by using statistical emulation," *Journal of the Royal Statistical Society Series C: Applied Statistics*, vol. 68, no. 5, pp. 1555–1576, 2019.