Constructing Cascading Latin Hypercubes
C. Devon Lin, Derek Bingham, Randy Sitter, and Boxin Tang
Simon Fraser University, Burnaby, BC, V5A 1S6, Canada

Abstract
Computer experiments commonly use space-filling designs. As the number of factors increases, the sparsity of the design points increases. Space-filling designs place all the points about the same distance (quite far) apart. If the spatial correlation length is also small relative to the space, there are no points close enough together to give reliable estimates of the correlation parameters. Handcock (1991) introduced cascading Latin hypercube designs (CLHD) to alleviate this issue. We develop systematic methods for constructing a rich class of CLHDs.

Notation
• \( n \) and \( m \) represent the run size and the number of factors, respectively.
• The levels are chosen to be centered, equally-spaced and in integer-valued.
• \(-n+1/2\), \(-n+3/2\), \(-n+1\) and \(-n-1/2\) if \( n \) is odd.

Definitions
Definition (Handcock, 1991): A CLHD of size \( n \times n \) is a random design with levels \( (n_1,\ldots,n_p) \) as a \( n_p \)-point LHD about each point in \( (n_1,\ldots,n_p) \) centered CLHD.

New Definition: Define a matrix \( U \) of \( LHD(n,m) = (L_{ij}) \) to have \((i,j)\)-th element
\[
U_{ij} = \begin{cases} \lceil \frac{L_{ij} + (n + 1)/2 \mod n} \rceil, & \text{if \( n \) is even;} \\ \lceil \frac{L_{ij} + (n + 1)/2 \mod n} \rceil, & \text{if \( n \) is odd.} \\ \end{cases}
\]
where \( [q] \) be the nearest integers greater than or equal to \( q \). A LHD is then termed a two-level CLHD of \( n \) points in \( m \)-dimensional with level \((n_1,n_2)\) if matrix \( U \) has \( n_1 \) distinct rows and each distinct row has \( n_2 \) replicates. \( p \)-level CLHD can be defined in a similar manner.

Construction Methods
Define
• \( A \) be an \( n_1 \times m \) design with \( n_{\alpha} \geq 1 \)
• \( B \) be an \( n_2 \times n_1 \) Latin hypercube design
• \( C \) be an \( n_3 \times n_2 \) Latin hypercube design
• \( D \) be an \( n_4 \times n_3 \) design with \( d_j \geq 1 \)
• \( \alpha \) and \( \beta \) be any positive real number
• \( \otimes \) represents Kronecker product

Basic Method:
\[
L = \alpha A \otimes B + \beta C \otimes D. \tag{1}
\]

Generalization Method: For each \( j = 1,\ldots,n_4 \), let \( B_j \) be an \( n_2 \times n_3 \) LHD.
\[
L = (a_{ij}B_j + \beta c_{ij}D) = \begin{cases} a_{ij}B_j + \beta c_{ij}D, & \text{if \( n \) is even;} \\ a_{ij}B_j + \beta c_{ij}D, & \text{if \( n \) is odd.} \\ \end{cases}
\tag{2}
\]

Remarks:
• Method (2) is proposed for better projection property.
• If design \( C \) are constructed via (1) or (2), the resulting design \( L \) in (1) or (2) will be a three-level CLHD.

Proposition: Let \( D \) be an \( n_1 \times m_2 \) matrix of unit elements.
• \( A \otimes (B \otimes B) \) and \( C \) control the design points locally and globally, respectively.
• The proposition below tells us the value of \( \alpha \), \( \beta \) and design \( D \) in (1) and (2) in order to obtain a two-level CLHD.

Example
Let \( n_1 = 9, n_2 = 3, n_3 = 4, m_2 = 3 \). Design \( D \) is a \( 3 \times 3 \) matrix of unit elements. Designs \( A \) and \( C \) are defined as
\[
A = \begin{pmatrix} -1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix},
\quad
C = \begin{pmatrix} 0 & 1 & 4 & 4 \\ 3 & 2 & 1 & 4 \\ 1 & 2 & 1 & 3 \\ 2 & 3 & 1 & 4 \\ 4 & 2 & 1 & -3 \end{pmatrix}
\]

Let \( B \) be a row permutation of \( B \). Set \( \alpha = 1 \) and \( \beta = 3 \) based on the aforementioned proposition. The resulting design \( L \) via (2) is a two-level CLHD with 27 points in levels \((9, 3)\). The first and fifth columns of design \( L \) correspond to the 27 circles in Figure 2.

Results and Conclusions

Results:
• Consider large input dimensions \( d = 20 \).
• CLHDs with 192 points in level (48, 4), (24, 8) and (16, 12) are generated.
• Maximum LHD with 192 points is generated.
• Power exponential correlation function with \( \theta = 1 \) is used.
• The average correlation between each design point and its \( k \)-nearest neighbors is computed, \( k = 4, 8, 12 \).
• Euclidean distance is used to find \( k \)-nearest neighbors.

We observe that when the dimension of inputs is relatively large, maximum Latin hypercube designs fail to provide close design points.
• Cascading Latin hypercube designs have close design points to detect the relationship between the inputs.

Figure 3: Comparisons of average correlation between each design point and its \( k \)-nearest points. \( k = 4, 8, 12 \). CLHD, MLHD, and CLHD_{A, k} are similar for \( k = 4 \); CLHD_{A, k} and MLHD_{A, k} are similar for \( k = 12 \).

Conclusion:
• When the input dimension is large, design points provided by space-filling designs are too sparse for Gaussian process to be effective.
• We provide methods for systematically constructing a rich class of designs with cascading structure.
• Cascading Latin hypercube designs provide local clustered points to enhance estimation of correlation parameters.
• The reliable estimation of correlation parameters allows us to achieve the goal of screening factors.

References