FORMULATION OF POTENTIAL FOR DYNAMICAL PARTICLE SYSTEM APPLIED TO MONTE CARLO SAMPLING

Jan Mašek¹ and Miroslav Vořechovský²

¹ Brno University of Technology
Veveří 331/95, 602 00 Brno, Czech Republic
e-mail: jan.masek1@vut.cz

² Brno University of Technology
Veveří 331/95, 602 00 Brno, Czech Republic
e-mail: vorechovsky.m@vut.cz

Keywords: Particle dynamical system, statistical sampling optimization, periodic space, particle interaction laws, short- and long-range interactions, self-similar designs.

Abstract. The presented paper investigates the effect of formulation of energy potential of a dynamical particle system as used for optimization of statistical point sampling. The dynamical particle system, originally developed as a physical analogy of the Audze-Eglajs (AE) optimization criterion and its periodical modification (PAE), effectively demonstrated that the originally proposed potential performs well only in poorly applicable scenarios of design spaces of low dimension filled with rather high number of design points. A remedy lying in a refined formulation of energy potential as well as its derivation and reasoning are presented.

©2017 The Authors. Published by Ecimos Proceedia. Peer-review under responsibility of the organizing committee of UNCECOMP 2017. doi: 10.7712/120217.5384.17000
1 INTRODUCTION

Numerical integration of Monte-Carlo type requires sampling of integration points that are uniformly distributed the design domain. The layout of design points crucially affects the performance of such a numerical integration. The problem of using an ideally distributed set of integration points is also of interest of many other engineering and research fields. While sampling from a random vector or integrating an unknown function, using a uniform layout of integration points is the only possible way for minimization of the lower bound of the resulting error.

However, the “uniformity” itself is not a recognized property. Many criteria have been proposed during the recent years for evaluation of the uniformity of point layouts containing \( N \text{sim} \) points within a design space of the dimension of \( N \text{var} \). Typically, these criteria investigate point layouts with a tendency to prefer designs with points distributed equally distant from each other. Certain criteria are derived from analogies with physical problems.

Namely, the Audze-Eglājs (AE) criterion [1] may be considered as an elegant instance of these. The objective of the AE criterion is in minimization of potential energy of a system of mutually repelling particles. With their positions, these particles represent positions of sampling points within a unit hypercube design domain.

During the recent years, it has been proposed that the original Audze-Eglājs criterion does suffer from the existence of boundaries of the design space [2, 3]. A remedy of this behavior was also proposed [2, 3] assuming a periodically extended design hypercube in case of which the boundaries naturally disappear. Building on such a refined Periodic Audze-Eglājs criterion (PAE), it has been proved that usage of the PAE criterion leads to statistically uniform designs (from design to design) and to well distributed set of points for every single point layout.

2 AUDZE-EGLĀJS AND PHI CRITERIA

The original formulation of the AE criterion [1] (see also [4, 5, 6, 7]) considers the analogy between the sampling plan and a system of charged particles with repulsive forces. The potential energy of the system is a sum of energies \( 1/L_{ij}^2 \) accumulated by each pair of points \( i \) and \( j \).

Instead of the sum of energies, one can alternatively calculate the average potential energy, i.e. divide the total energy by the number of pairs of points \( N_p = \binom{N \text{sim}}{2} \). The AE criterion then reads:

\[
E_{AE}^n = \frac{1}{N_p} \sum_{i=1}^{N \text{sim}-1} \sum_{j=i+1}^{N \text{sim}} \frac{1}{L_{ij}^2}
\]

(1)

where \( L_{ij} \) is the inter-site Euclidean distance between points \( i \) and \( j \), depending on the Cartesian coordinates of said points, \( x_i = \{x_{i,v}\}, i = 1, \ldots, N \text{sim}, v = 1, \ldots, N \text{var} \):

\[
L_{ij} = \sqrt{\sum_{v=1}^{N \text{var}} (\Delta_{i,v})^2}
\]

(2)

where:

\[
\Delta_{i,v} = |x_{i,v} - x_{j,v}|
\]

(3)

is the projection of the distance \( L_{ij} \) onto the axis \( v \).

A generalization of the AE criterion is the \( \phi \) criterion [8]:

\[
\phi(d,\lambda) = \frac{1}{N_p} \left( \sum_{i=1}^{N \text{sim}-1} \sum_{j=i+1}^{N \text{sim}} \frac{1}{d^\lambda (x_i, x_j)} \right)^{\frac{1}{\lambda}}
\]

(4)
This criterion considers a general power, $\lambda$, of the distance (metric) $d$. The combination of $\lambda = 2$ and the Euclidean intersite distance $d$ leads makes $\Phi(d, \lambda)$ identical to the AE criterion. As $\lambda \to \infty$, the criterion increasingly prioritizes designs where the minimal distances are maximized which is, in limit, the Maximin distance criterion [9]. Using the analogy with the system of charged particles, one can say that with increasing power $\lambda$ a greater portion of energy is stored in the short-range interactions.

The power $1/\lambda$ upon the entire sum is, in fact, a monotonous transformation (the difference between designs is not distorted) and can be dropped as well. Therefore, a simplified version of the criterion with parameters of a used metric $d(x_i, x_j)$ and exponent value $\lambda$ can be considered:

$$\Phi(d, \lambda) = \frac{1}{N_p} \sum_{i=1}^{N_{sim}-1} \sum_{j=i+1}^{N_{sim}} \frac{1}{d^\lambda(x_i, x_j)}$$

(5)

The standard way of measuring the inter-point distances is the Euclidean length defined above in Eq. 2: $d(x_i, x_j) = L_{ij}$. It has been shown that due to the presence of boundaries of the design domain, this metric used in the criteria Eqs. 1, 4, 5 leads to non-uniform point distribution [2, 10, 11]. Moreover, the authors of [2, 10, 11] have shown that if the metric $d(x_i, x_j)$ is modified such that it considers distances measured in the periodically repeated design domain (a periodic metric), the criterion becomes invariant with respect to arbitrary shifts along individual dimensions. A simplified version of the periodic space considers only the shortest distances, i.e. the distance is taken as $d(x_i, x_j) = L_{ij}$, where $L_{ij}$ is the Euclidean distance between the $i$-th point and the closest image of the $j$-th point within the periodic space:

$$L_{ij} = \sqrt{\sum_{v=1}^{N_{var}} (\Delta_{ij,v})^2}$$

(6)

where:

$$\Delta_{ij,v} = \min(\Delta_{ij,v}, 1 - \Delta_{ij,v})$$

(7)

is the shortest projection of the distance between point $i$ and the nearest image of point $j$ onto the axis $v$.

Using this definition of distance, Eq. 5 reads:

$$\Phi(L, \lambda) = \frac{1}{N_p} \sum_{i=1}^{N_{sim}-1} \sum_{j=i+1}^{N_{sim}} \frac{1}{L_{ij}^\lambda}$$

(8)

Utilizing the above-described nomenclature, the AE criterion as written in can be denoted as $\Phi_{(L,2)}$. The periodic version of the AE criterion (the PAE criterion [2]) uses a combination of exponent $\lambda = 2$ and the shortest (periodic) metric $L_{ij}$ and therefore it can be denoted as $\Phi_{(L,2)}$.

The authors of [2] argue that already for exponent $\lambda = 2$, the consideration of the shortest distance $L_{ij}$ suffices to deliver invariance with respect to random shifts along individual dimension and thus prioritizes designs leading to statistical uniformity of coverage. Moreover, it is argued that the shortest distance is the one associated with the highest contribution to the criterion and therefore the PAE criterion captures the important features of the full periodic repetition of the design space. The next section generalizes the criterion in Eq. 8 by considering a higher number of copies of the design domain.
3 PERIODIC EXTENSION OF THE DESIGN DOMAIN

In this section we consider a generalized model in which a certain number of periodic repetitions of the original design domain are considered. Using the nearest image of point \( j \) with respect to point \( i \), as considered in Eq. 8, does not cover a true periodic repetition of the design domain. We argue that the above presented approach is a simplification that can be shown to yield identical results to the fully repeated system in case of sufficient point count \( N_{\text{sim}} \). If the number of points in the original domain is too small to carry enough information about the pattern of a periodically repeated system, making a periodic extension to a sufficient level is desirable. In a true periodic domain, infinite number of images of point \( j \) would interact with point \( i \). When a finite number of copies of the design domain is considered, not only the real particle \( j \), but also all periodically repeated images of the particle \( j \) will contribute to the potential:

\[
\phi(L, \lambda, k_{\text{max}}) = \frac{N_{\text{sim}}}{i=1} \sum_{j=i+1}^{N_{\text{sim}}} \left( \frac{1/N_p}{L_{ij}} + \sum_{k=1}^{k_{\text{max}}} \sum_{c=1}^{c_{\text{max}}} \frac{1/N_p}{L_{ij}^c(x_i, x_j + s_c)} \right)
\]

where \( k_{\text{max}} \) introduced as an additional parameter is the number of added periodical extensions (envelopes) of the design space. In the fully repeated system \( k_{\text{max}} = \infty \) and analogically, for a non-extended system \( k_{\text{max}} = 0 \). Therefore \( \phi(L, \lambda) = \phi(L, \lambda, 0) \), compare Eqs. 8 and 9.

When a certain number of envelopes \( k_{\text{max}} \) is considered, the number of copies of the design domain is denoted as \( c_{\text{max}} = 0 \). The vector \( s_c \) is the vector needed for shifting the original point \( j \) to the particular periodically repeated version indexed by \( c \). Let us denote that the distances to the periodically repeated images of the point \( j \) must be measured as the standard Euclidean distances.

A single level of periodic extension adds another envelope of periodically repeated images of all other particles around each point, see Fig. 1 and 2e. Such an extension does provide additional information about the point layout within the domain.

The level of the periodic extension is quantified by a positive integer \( k_{\text{max}} \). Within an extended periodic domain of finite value of \( k_{\text{max}} \), the particle \( i \) interacts not only with the actual particle \( j \), but with all of \( c_{\text{max}} = [(2k_{\text{max}} + 1)^{N_{\text{var}}} - 1] \) images of the particle \( j \) as well, see Fig. 1 for \( N_{\text{var}} = 1 \). The envelopes are considered to be centered around the shortest distance with \( L_{ij} \).

![Figure 1: 1D example of the periodic extension of level \( k_{\text{max}} = 2 \)](image)

4 THE EXponent \( \lambda \)

This section focuses on the exponent \( \lambda \) in the periodic \( \phi \) criterion [10], see Eq. 8. In the original AE criterion and also in the periodic version (PAE), the potential energy between each
pair of points is not dependent on the dimension, $N_{\text{var}}$. It has been found \[12\] that the character of the criterion is different for various $N_{\text{var}}$ and also for various numbers of points, $N_{\text{sim}}$. In 1D situation, the energy tends to infinity linearly with increasing $N_{\text{sim}}$. In a 2D, the energy tends to infinity as $\ln(N_{\text{sim}})$ which is not a power law. For dimensions $N_{\text{var}} \geq 3$, the energy tends to a constant for increasing $N_{\text{sim}}$.

This behavior can be explained by the fact that for a given $N_{\text{var}} > 1$, various numbers of points yield to different proportions between energy due to the long-range and short-range interactions. The higher is the number of points, the higher the proportion of energy stored in long interactions is. This may not be desirable behavior as the criterion in high dimensions and also for high number of points becomes insensitive to local clusters of points: it becomes dominated by long-range interactions.

The power is suggested to be at least $\lambda \geq N_{\text{var}} + 1$. With this power, the interaction is dominated by short-range interactions. With such a sufficient exponent $\lambda$, the convergence of the potential energy $\phi(L, \lambda)$ or better $\phi(L, \lambda)$ towards infinity for a uniform distribution of points is a power law. Such a convergence signalizes self-similarity of the problem or absence of a length scale. In other words, a zoom into sufficiently dense uniform design with a window greater than a certain size (see below) carries all features of the full design and the energy value can be easily scaled from the value corresponding to the smaller zoom.

This can be shown by studying the behavior of the radial part of the integral of the potential over the volume $V$ of $N_{\text{var}}$-dimensional domain. The potential energy for a uniform design reads:

$$ I = \int \frac{1}{L^\lambda} d^{N_{\text{var}}} V $$

where $L$ is used to denote one-dimensional distance between points (the symbol $d$ is not used to avoid confusion with the symbol $d$ for the differential). Transforming this into polar coordinated gives:

$$ I = \int \varphi d^{N_{\text{var}} - 1} V |J| \frac{1}{L^\lambda} dL $$

where $|J|$ is the Jacobian. The volume element is thereby given as:

$$ d^{N_{\text{var}}} V = L^{N_{\text{var}} - 1} dL \cdot d\varphi \prod_{i=1}^{N_{\text{var}} - 2} \sin^{N_{\text{var}} - 1 - i}(\varphi_i) $$

Therefore, the integral is performed over the product $L^{N_{\text{var}} - 1 - i}$. Performing just the radial integration leads to:

$$ I_r = \int \frac{L^{N_{\text{var}} - 1}}{L^\lambda} dL = \int L^{N_{\text{var}} - 1 - \lambda} dL $$

For $\lambda = 2$ as used in the AE criterion, we get the behavior described above. Using $\lambda = N_{\text{var}}$ leads to:

$$ I_r = \int L^{-1} dL = \ln(L) $$

which diverges logarithmically and the interaction is still long-ranged. Using $\lambda = N_{\text{var}} + 1$ yields

$$ I_r = \int L^{-2} dL = \frac{1}{L} $$
which has the desired asymptotic behavior dominated by short-ranged interactions. Using powers \( \lambda > N_{\text{var}} + 1 \) only increases the (asymptotically constant) ratio between short-range and long-range interactions.

Fig. 2 shows the convergence of the normalized potential energy \( \phi(L, \lambda) \) with rise of the number of particles, \( N_{\text{sim}} \). Instead of presenting the results for the point count, \( N_{\text{sim}} \), we introduce a variable \( l_{\text{char}} \), the characteristic length that involves also the dimension of the space. The characteristic length is defined as:

It can be seen that with the original exponent value \( \lambda = 2 \) in the dimension \( N_{\text{var}} = 2 \) (or generally \( \lambda = N_{\text{var}} \)), the potential energy of the system does not converge to a power law but diverges logarithmically, roughly

\[
\phi(L, \lambda) \approx \pi \ln(N_{\text{sim}}) + \frac{1}{\sqrt{N_{\text{sim}}}} - \frac{1}{N_{\text{sim}}}
\]

In higher dimensions \( N_{\text{var}} \geq 3 \), the exponent \( \lambda = 2 \) further leads to convergence of the potential energy to a constant [12].

Using the above proposed exponent \( \lambda = N_{\text{var}} + 1 \), the potential energy value tends to a power law as \( N_{\text{sim}} \to \infty \):

\[
\phi(L, N_{\text{var}} + 1) \approx \frac{1}{l_{\text{char}}}
\]

Such a behavior is desired as the designs for a given dimension \( N_{\text{var}} \) tend to have a universal self-similar pattern and the dependence on sample size disappears (no length scale is present). Thus the character of the criterion is kept independent of \( N_{\text{sim}} \) and the proportion between short-range interactions and long-range interactions is constant. This stabilization is obtained for a sufficient number of points within the design domain (a kind of tile). The self-similarity manifested is by the power law dependence (a straight line in Fig. 2). When the exponent is taken even higher (\( \lambda > N_{\text{var}} + 1 \)), the self-similar regime is achieved for even smaller number of points (greater \( l_{\text{char}} \)).

Graphs in Fig. 2 suggest that there must be link between (a) the exponent (responsible for the proportion between long- and short-range interactions) and, (b) the number of “dummy” copies of the design domain that also modify the proportions. This aspect is discussed in the next section.

5 ON THE SEEMINGLY SIMILAR EFFECT OF RISING THE EXPONENT \( \lambda \) AND INCREASING THE LEVEL OF PERIODIC EXTENSION \( k_{\text{max}} \)

Let us consider a few-body \( (N_{\text{sim}} = 3) \) particle system in design space of \( N_{\text{var}} = 2 \) while using a potential with the exponent \( \lambda = 2 \). When considering the point layout via the \( \phi(L, \lambda) \) interaction, for each point, there exist only two mutual distances to other points, i.e. two forces acting upon each particle.

With an exponent of such insufficient magnitude, these forces do not differ significantly enough to represent correctly which particle shall be considered to be close (short-range interaction) and which to be far (long-range interaction). A rise of the exponent above certain threshold (discussed above) does lead to the needed qualitative change of the ratio between the acting forces: the closer particles start to act as short-range and the farther particles as long-range. In fact, the higher the exponent, the larger portion of potential energy will be stored in the short-range interactions.
Figure 2: Convergence of the normalized potential energy \( \phi(L, l) \) depending on the exponent \( l \). Coloring that designate \( N_{\text{var}} \) is identical for the two bundles of curves. Solid circles are accompanied by the sample count corresponding to the \( l_{\text{char}} \) and \( N_{\text{var}} \).

The analogy between the effect of increasing the number of envelopes in and rising the exponent in is evident. While using the original (low) value of the exponent \( l \), majority of the potential energy is stored in the long-range interactions. However, there is not a sufficient number of particles for the criterion to distinguish between short and long range. All particles seem to be at similar distance as the design is not filled enough.

The \( \phi(L, l, k_{\text{max}}) \) interaction adds one or multiple additional envelopes of neighboring images of actual particles. These images, naturally, will act as long-range. Even longer-range than the real particles previously acting as long-range. This leads to a qualitatively more accurate distribution of forces acting upon the real particles. Hence the identical behavior of the \( \phi(L, l, k_{\text{max}}) \) and the \( \phi(L, l) \) interaction with a correct exponent \( l \):

- the \( \phi(L, l, k_{\text{max}}) \) interaction does add long-range points for the actual particles to seem closer,
- the \( \phi(L, l) \) interaction with corrected exponent changes the ratio between forces for the close particles to seem closer and the distant particles to seem farther.

It can be therefore shown that while simulating a few-body problem with the \( \phi(L, l) \) interaction, it is advised to rise the exponent \( l \) even above the lower bound of \( N_{\text{var}} + 1 \) to force the desired self-similarity for various \( N_{\text{var}} \).

When using the \( \phi(L, l, k_{\text{max}}) \) interaction, especially for few-body problems, greater context of the pattern is carried within the interaction for there is considered \( 3N_{\text{var}} \) images of each particle. Effectively, a system mimicking \( N_{\text{sim}} \cdot 3N_{\text{var}} \) particles is being simulated and the identical pattern should be obtained, see Fig. 3b and 3f.

On a side note, the simulation of a greater (extended) system might be the slower option compared with the correction of the exponent in the energy potential as the number of additional
6 COMPUTER IMPLEMENTATION

A simulation of a particle system is typically a computationally demanding task. However, it is possible and beneficiary to conduct the implementation of solution as parallel as possible. The degree up to which the parallelism can be reached depends dominantly on the nature of the problem at hand and also on the possibilities of the used hardware. The implementation of solution of the proposed dynamical particle system has been conducted using the nVidia CUDA platform.

The equations of motion of the dynamical particle system contain independent accelerations on the left-hand side. This means that accelerations of particles can be solved separately, without solving a system of equations. Furthermore, computation of the mutual distances as well as the numerical integration of equations of motion using the semi-implicit Euler method can be conducted in parallel.

The derivation of equations of motion, the nature of implementation and speed-up are not trivial and were already covered in detail in concurrent publications [3, 13].

7 RESULTS OF NUMERICAL SIMULATIONS

The following section presents examples of numerical simulations of the dynamical particle system, mainly focusing on the effect of the value of the exponent $\lambda$ in the energy potential as well as on the effect of periodical extension.

For start, let us show how the ability of a self-similar design (power-law quality) disappears while the number of particles $N_{\text{sim}}$ decreases. In a design space of dimension $N_{\text{var}} = 2$, the original exponent value yields ideal distribution only if the design space is filled enough. For $N_{\text{var}} > 2$, the original exponent leads to primitive patterns of even incomplete ortho-grids.

Studying the design space of dimension $N_{\text{var}} = 2$, Fig. 3a, 3b and 3c show that an ideal pattern of a triangular grid (if possible) is reached as long as the information needed to creating of such a self-similar pattern is provided (the number of points is sufficient). When the number of particles becomes insufficient, the interaction is not approximated well enough (in a pursuit of a perfect triangular grid) and primitive (even incomplete) ortho-grids are produced, see Fig. 3d.

A remedy, as described above, can be conducted by two seemingly dissimilar approaches. First, see Fig. 3e for an example of a layout obtained with a sufficiently high value of the exponent $\lambda$. Second, a result of a simulation using a periodic extension of a single envelope ($k_{\text{max}} = 1$) is provided. Such a simulation effectively considers $4 \cdot 3^2 = 36$ points (4 of which are the actual particles), see Fig. 3f. Hence, the entire extended space is filled equivalently to the non-extended ($k_{\text{max}} = 0$) simulation of 36 particles, see Fig. 1b. An identical point layout of 36 points is therefore achieved and, in a way, the resulting layout for the system of $N_{\text{sim}} = 4$ is cut out of a greater scenario.

It is worth noting that, when simulating such an extended system (36 particles instead of 4 particles), the power-law quality is easier to achieve as the higher theoretical $N_{\text{sim}}$ (lower $l_{\text{char}}$ value) occurs closer to the power-law asymptote, see Fig 4.

Another portion of simulation examples concerns the results of simulations within a three-dimensional design space ($N_{\text{var}} = 3$). For this and higher dimensions, the original formulation of the (P)AE potential is already malfunctioning and yields poor designs based dominantly on ortho-grids, see Fig. 4a. The result of the remedy of rising the value of the exponent $\lambda$ is
Figure 3: Disappearing of the quality of a self-similar design as yielded by the original PAE formulation and the effect of remedies proposed.

provided, see Fig. 4b.

Such a refined formulation of the energy potential, we believe, will lead to perfect designs not only within the design space of the complete dimension, but also in all sub-spaces of lower dimension, see Fig. 4.

8 RESULTS OF NUMERICAL SIMULATIONS

The paper investigates the formulation of the energetic potential of the Audze-Eglājs optimization criterion and its periodical modification PAE. Remedies of an incorrect behavior of systems of low numbers of particles and high dimensions are proposed.

An indirect solution leading for appropriate layouts of few-body systems lies in simulating a greater scenario: considering additional envelopes of the periodically repeated system. That way, a richer information about the pattern is provided and the optimal point layout can be obtained. However, such a scenario mimics simulation of a system of higher number of particles and does not lead to a correction of the malfunctioning energy potential.

An overarching remedy was therefore pursued, leading to a potential which takes into account the dimension of the problem, \( N_{\text{var}} \). First, a generalization of the potential based on the \( f \) criterion is provided so the crucial parameters of the potential, the metric \( d(\bullet, \bullet) \) and the exponent \( \lambda \), can become subject of a refinement.

Further derivation was based on the desire for an ability of creating self-similar patterns of point layouts for various point counts. With the proposed value of the exponent \( \lambda \), convergence of the potential energy of the criterion towards infinity for a uniform distribution of points is a power law. Such a convergence signalizes self-similarity of the problem or absence of a length scale. Using such a refined interaction, optimal (self-similar) designs are produced even for scenarios of arbitrary dimension, \( N_{\text{var}} \), or few-body systems, as was shown.

In this way, the role of the exponent featured in the \( f \) criterion is explained using the analogy with a system of mutually repelling particles.
ACKNOWLEDGEMENT

This work has been supported by the Grant agency of the Czech Republic under projects Nos. GA16-22230S and GA15-07730S. This support is gratefully acknowledged.

The authors would like to thank Dr. Godehard Sutmann for valuable discussions about physics of interacting particles.
REFERENCES


