

An adapted derivative-free optimization method for an optimal design application with mixed binary and continuous variables

Thi-Thoi Tran¹, Delphine Sinoquet¹, Sébastien Da Veiga², and Marcel
Mongeau³

¹ IFPEN, Rueil-Malmaison, France
{thi-thoi.tran,delphine.sinoquet}@ifpen.fr

² Safran, Magny-Les-Hameaux, France
sebastien.da-veiga@safran.fr

³ ENAC, Université de Toulouse, France
mongeau@recherche.enac.fr

Abstract. Numerous optimal design applications are black-box mixed integer nonlinear optimization problems: with objective function and constraints that are outputs of a black-box simulator involving mixed continuous and integer (discrete) variables. In this paper, we address an optimal design application for bladed disks of turbo-machines in aircraft. We discuss the formulation of an appropriate distance with respect to discrete variables which can deal with the cyclic symmetry property of the system under study. The necklace concept is introduced to characterize similar blade configurations and an adapted distance is proposed for discrete space exploration of a derivative-free optimization method. The results obtained with this method on a simplified industrial application are compared with results of state-of-the-art black-box optimization methods.

Keywords: Mixed integer non-linear programming · black-box simulation · derivative free trust region method · necklace distance.

1 Motivation

Air traffic is one of the most important means of transportation, especially in Europe. Besides, it is connected with very high costs of fuel [1,2], and also with high costs of maintenance and of manufacturing. Thus, reducing fuel consumption by increasing engine efficiency and maintenance savings by decreasing vibrations are two major concerns of the aviation industry.

There are several ways to optimize costs in aircraft: optimal trajectories, optimal seat arrangement designs, optimal cargo arrangements ... In our case, we want to optimize the design of turbo-machines, precisely, by maximizing the efficiency of compressor and minimizing the vibrations.

In the concrete application proposed by SAFRAN, the optimization variables are of 2 types: continuous shape parameters, x , e.g. thickness, length of blades

and binary variables y associated to each blade, with the value 0 for a reference shape and 1 for the other predefined shape (mistuning shape). Binary variables are used to locate these reference blade shapes on the disk. This parameterization provides the distribution of the two shapes around the turbine disk.

There is a strong symmetry in this problem that should be taken into account. Two bladed disks that differ only by a rotation of the blade pattern around the disk will lead to the same simulation outputs, such arrangements are considered as the same solution or called redundant solution : e.g. 001101001101 and 010011010011. Note that the number of identical solutions increases rapidly with the number of blades (see Table 1).

In practice, engineers try to overcome this difficulty in practical optimization by using Reduce Order Modeling Technique (ROMT), detailed in [15,6,8,7]. Briefly, grouping blades by sectors of 2 patterns - 00 or 11 and 10 or 01 - tends to reduce the number of discrete variables from n to $n/2$.

In general these mixed problems are NP-hard and difficult to solve. Especially, for real applications, the difficulty lies in evaluating computationally expensive cost function: it can take several hours or even days to compute the functions to be optimized. Moreover, the derivatives are often not available. Most of algorithms for solving black-box Mixed Integer Nonlinear Problem are based on genetic algorithms which require a lot of function evaluations which are particularly costly in our application. In the next section, we motivate the use of derivative-free trust-region method.

Total number of blades on the disk (N)	Number of distinct arrangements ($\sim 2^N/N$)	Total number of arrangements (2^N)
2	3	4
3	4	8
5	8	32
10	108	1024
12	352	4096
20	52488	1048576

Table 1: Number of distinct and total arrangements for a given number of blades

2 Derivative free trust-region method

Among Derivative-free optimization (DFO) methods, one distinguishes direct search methods (e.g., directional or Nelder Mead simplex) and trust region methods based on simple interpolation or regression models (linear or quadratic). The direct search methods require a large number of simulations, whereas the second type of methods is generally more efficient to converge to a local solution. Convergence results to local minima are proved for the later methods but it requires

adaptations to hope to converge to a global optimal solution (e.g., a multi-start approach with several initial points).

We focus on surrogate DFO methods [10,4]. These methods aim to explore the optimization variable space by replacing the costly-to-evaluate functions by response surfaces, with a choice of the points to simulate based on a compromise between exploration (points far from points simulated at previous iterations) and exploitation of information captured by the response surface (points around potential optima). Most popular response surfaces for surrogate optimization are Radial Basis Functions (RBF) and Gaussian processes [17,14,11].

Besides, derivative-free trust-region algorithm is based on local quadratic models defined inside a trust region

$$m(z) = \alpha + g^T z + \frac{1}{2} z^T H z,$$

whose coefficients are determined as the solution of minimal Frobenius norm problem

$$\begin{cases} \min_{\alpha, g, H=H^T} \frac{1}{2} \|H\|_F^2 \\ m(z_i) = f_i, i = 0, \dots, p. \end{cases}$$

where $Z = (z_0, \dots, z_p)$ is the interpolation set for both continuous and binary variables $z = (x, y)$, f_i the associated objective function evaluations and $\|\cdot\|_F$ the Frobenius norm. The interpolation set is assumed to be poised, see details in [10,4].

The brief idea of the algorithm is to replace the initial problem, which involves the expensive simulations with quadratic optimization problem, simple to optimize. At the first step, we fix the binary variables to the best current solution and solve the quadratic sub-problem within the trust region with respect to continuous variables only. When a better point is obtained in this first step (a smaller objective function), the second step consists in minimizing the quadratic model with respect to both continuous and binary variables, thus solving a mixed binary continuous quadratic problem.

The model is built in order to efficiently approximate the function and fulfills the fully-linear or fully-quadratic model properties to ensure the local convergence of the algorithm. Thus, model improvement steps are performed in order to optimize the geometry of the interpolation set, see details in [10,9,4].

The introduction of binary (or integer) variables requires an adapted trust region definition. In [9], the authors introduce a l_1 -norm trust region for continuous variables and the Hamming distance trust region for binary variables defined as

$$\sum_{j:y_{c_j}=0} y_j + \sum_{j:y_{c_j}=1} (1 - y_j) \leq \Delta_y, \quad (1)$$

with Δ_y , the size of the trust region for binary variables.

When the algorithm reaches a "local solution" (no further improvement), an exploration phase is necessary in binary variable domain. It is performed by

adding a "no-good-cut" constraint to escape from this local solution y^* , defined as

$$\sum_{j:y_j^*=0} y_j + \sum_{j:y_j^*=1} (1 - y_j) > \Delta_y^*, \quad (2)$$

where (y^*, Δ_y^*) are respectively the local solution and the current radius of the "sufficiently explored" area around this solution. Note that this approach leads to a bunch of constraints, one for each explored local solution.

In the next section, we propose an adapted distance that will be used for the trust region associated with binary variables for our application.

3 Adapted distance for blade design application

To avoid the solution redundancy (illustrated in Table 1), engineers use ROMT which reduces the optimization problem size but has the disadvantage of removing a large number of feasible solutions. Safran's application with 12 blades has 352 distinct arrangements but only 28 distinct arrangements for the two sub-problems of ROMT, limiting a lot the explored configuration set with a high probability to remove "good" candidates.

Therefore, we attempt to define a new distance which can avoid the redundant solutions without arbitrary removal of configurations. The new distance should lead to simple constraints as the Hamming distance which leads to linear constraints (1).

In order to detect similar blade arrangements, we introduce the concept of "necklace" [12,13]. In combinatorics, a k -ary necklace of length n is an equivalence class of n -character strings over an alphabet $\sum^k = \{a_1, \dots, a_k\}$ of size k , taking all rotations as equivalent. It represents a structure with n circularly connected beads which have k available colors. Our blade design application can be seen as a 2-color necklace optimization with a fixed number of beads.

Using the concept of necklace gives an exact formula of the number of distinct arrangements which is the number of necklace for given n beads : $1/n \sum_{d|n} \phi(d) 2^{n/d}$, where ϕ is Euler's totient function, the summation is taken over all divisors d of n .

There are numerous applications based on "necklace" concept and various distances, e.g. geometry distance for music, swap distance [22,23,24], Hamming distance with shift [16], and necklace alignment distance (NAD) [5].

Following the idea of NAD, we propose a new distance, that we call in the following the necklace distance,

$$d_{neck}(y, y') = \min_r d_H(y, Rot^r(y')), \quad (3)$$

where d_H denotes the Hamming distance, $Rot^r(y)$ is the rotation of r positions from y . It is clear that this distance satisfies

- the non-negativity property: $d_{neck}(y, y') \geq 0$,
- the reflexivity property : $d_{neck}(y, y) = 0$,

- the commutativity property: $d_{neck}(y, y') = d_{neck}(y', y)$,
- the triangle inequality property: $d_{neck}(y, y'') \leq d_{neck}(y, y') + d_{neck}(y', y'')$.

Besides these metric properties, d_{neck} has one important property

$$d_{neck}(y, y') = 0 \iff y \in Rot(y') \quad (4)$$

which ensures to detect all the necklaces in the set. The definition of the necklace distance is not linear because of "min" operator. We propose in the next section an adaptation of this distance for the trust region and no-good-cut constraints used in the trust region algorithm.

Reformulating the distance for trust region DFO algorithm

The mixed binary continuous sub-problem associated with the minimization of the quadratic model is written as

$$\begin{cases} \min_{x,y} m(x, y) \\ \min(g_1(y, y_c), \dots, g_n(y, y_c)) \leq \Delta_y \\ \|x - x_c\|_{l_1} \leq \Delta_x, \\ x \in \mathbb{R}^n, y \in \{0, 1\}^n, \end{cases} \quad (5)$$

with x_c and y_c the current centers of the two trust regions in continuous and binary variable space with Δ_x and Δ_y , the size of the trust region for respectively the continuous and binary variables and $g_i(y, y_c) = d_H(y, Rot^i(y_c)), i = 1, \dots, n$.

To deal with "min" operator in the constraints, we add slack variables t , integer, and $y_{n+1}, y_{n+2}, \dots, y_{n+n}$, binaries, and propose an exact reformulation,

$$\begin{cases} \min_{x,y,t} m_\mu(x, y, t) = \min_{x,y,t} m(x, y) + \mu t \\ t \geq g_1(y, y_c) - M y_{n+1} \\ \dots \\ t \geq g_n(y, y_c) - M y_{n+n} \\ \sum_{i=1}^n y_{n+i} = n - 1 \\ t \leq \Delta_y \\ \|x - x_c\|_{l_1} \leq \Delta_x, \\ x \in \mathbb{R}^n, y \in \{0, 1\}^{2n}, t \in \mathbb{Z}_+, \end{cases} \quad (6)$$

with a real parameter $\mu > 0$ and an integer parameter $M > n$.

As explained before, the exploration phase uses "no-good-cut" constraints (2) to enforce to explore new values for binary variables. The maximum number of constraints is $2^n - 1$. If we apply the same trick as in (6) for exploration phase, we highly increase the dimension of the problem due to the additional slack variables, $n + 1$ for each "no-good-cut" constraint. We use instead the equivalence

$$\min(g_1, g_2, \dots, g_n) \geq \Delta_y^* \iff g_i \geq \Delta_y^*, i = 1, \dots, n,$$

and thus replace no-good-cut constraints by at most n linear constraints

$$d_H(y, Rot(y^*)) \geq \Delta_y^*, \dots, d_H(y, Rot^n(y^*)) \geq \Delta_y^*.$$

Note that, in practice, we use $\Delta_y^* = 1$.

4 Toy problem

The problem of determining the best-case (or worst-case) mistuning pattern and maximizing the forced response vibration amplification by the addition of a small mistuning to a perfectly cyclical bladed disk is mentioned in literature [19,20,15,6,8,7,21]. We use a single degree-of-freedom (DOF) per blade disk model (see Figure 1).

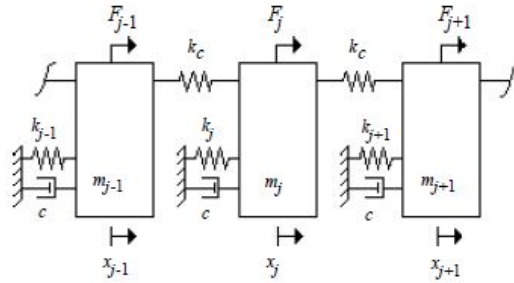


Fig. 1: Single DOF per blade disk model ([7])

The DOF problem is formulated as

$$\begin{aligned} & \underset{\omega, y}{\text{minimize}} && \|A\|_{\infty} \\ & \text{subject to} && \omega \in [\omega_{min}, \omega_{max}], \\ & && y_i \in \{0, 1\}, \end{aligned} \quad (7)$$

$$\text{where } A = T^{-1}\bar{\mathcal{F}}, \bar{F}_i = F_0 e^{j\phi_i}, T = \begin{pmatrix} T_0 & -k_c & 0 & \dots & 0 & -k_c \\ -k_c & T_1 & -k_c & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -k_c & 0 & 0 & \dots & -k_c & T_{N-1} \end{pmatrix},$$

$T_i = -m_i\omega^2 + j\omega c + 2k_c + (1 + \delta_i)k_b$. The nomenclature is detailed in [7].

5 Preliminary numerical results

We present some preliminary results of our trust-region method, called DFOb in the following, applied to the toy problem and to a simplified application provided by SAFRAN.

Results for the Toy problem

In general, optimization researchers ideally seek for an algorithm which provides the smallest objective function value comparing to alternative algorithms.

But, for derivative free optimization methods dealing with expensive-to-evaluate simulations, the efficiency of an optimization algorithm is generally measured by comparing the objective function value for a given budget of simulations (fixed number of simulations).

On this toy example, we run 10 times our method DFOb with different initial random points with two different distances for binary variables : Hamming distance and the necklace distance proposed in previous section (equation 3). The new implementation allows to reach a better point than the Hamming distance implementation as shown in Figure 2. The number of simulations to reach the same objective value, $9.68272e - 04$, is larger with Hamming distance (90 simulations) than with the necklace distance (60 simulations).

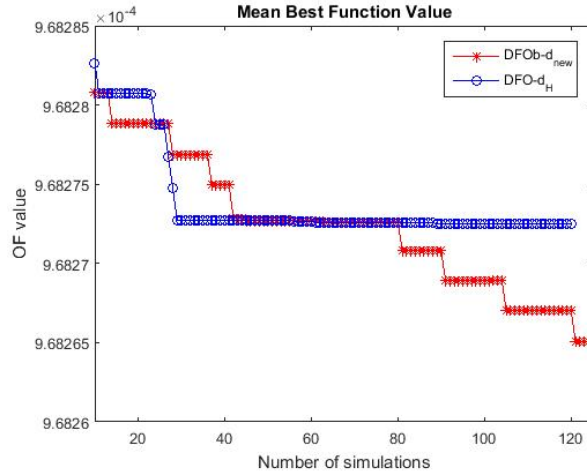


Fig. 2: Mean best average objective function obtained with 10 random initial datasets with Hamming distance and with the proposed distance.

Results on SAFRAN application

We apply our method DFOb on the two sub-problems of ROMT and compare the results with NOMAD, [3,18] and RBFOpt [11] optimization methods with a fixed budget of 100 simulations:

1. ROMT sub-problem 1 with patterns 00 – 11,
2. ROMT sub-problem 2 with patterns 10 – 01.

DFOb and RBFOpt share the same initial points (random choice from RBFOpt method). NOMAD's initial point is chosen as the best point of the initial set with regard to objective function value. Some constraints are added in order to avoid trivial solutions with only one type of patterns : only zero or one values in vector of binaries y . Figure 3 illustrates the efficiency of DFOb coupled with the necklace distance compared to Hamming distance for binary variables for

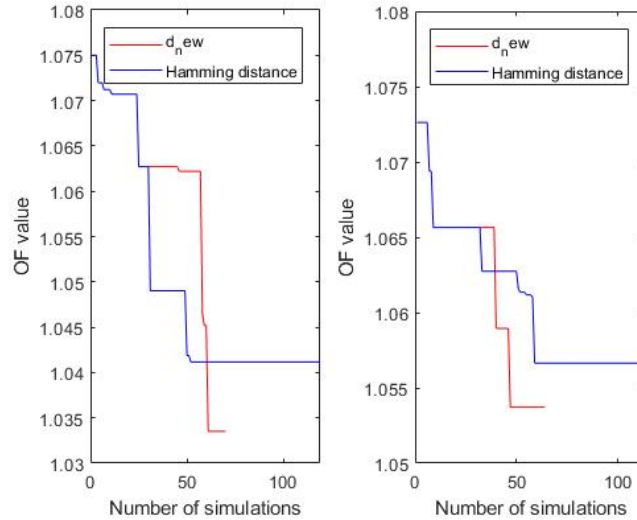


Fig. 3: Comparing results obtained with Hamming and necklace distances for the two ROMT sub-problems with patterns 00 – 11 (left) and 10 – 01 (right) of Safran application

the two sub-problems of ROMT. For ROMT sub-problem with 00 – 11 patterns, DFOb with the necklace distance provides an objective function of 1.033509 after 57 simulations while the same algorithm with the Hamming distance does not reach this value after more than 100 simulations. For the first ROMT sub-problem (see Figure 4), DFOb method coupled with necklace distance has very good performance compared to RBFopt and NOMAD. NOMAD and RBFopt provide infeasible points during the exploration phasis (with small objective function values), the constraints being handled as soft constraints. In DFOb, the constraints are taken into account explicitly.

For the second ROMT sub-problem (see Figure 5), DFOb finds good points rapidly after less than 10 simulations, nevertheless, it does not reach the best objective function value obtained by NOMAD and RBFopt within the fixed budget of simulations (100). With a different initial set, DFOb is able to find the global solution within 100 simulations. A future study will focus on the sensitivity of the results of our method to the initial set and of its size.

6 Conclusions

In this study we address black-box optimization problems with costly-to-evaluate objective functions. A trust region derivative free optimization method adapted to mixed binary and continuous variables is presented. In order to improve the

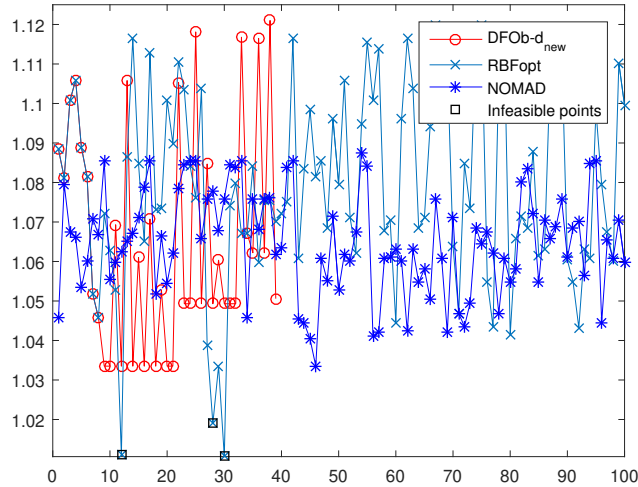


Fig. 4: Compare DFOb, Nomad and RBFopt results obtained for sub-problem with patterns 00 – 11 of Safran application. Square black symbols indicate infeasible points.

exploration phase of this algorithm for a blade design application, we introduce a distance in binary variable domain, that takes into account the symmetry of the problem. Preliminary results illustrate the performances of this new distance compared to classical Hamming distance. The comparison of the proposed method with two state-of-the-art methods NOMAD and RBFopt shows some encouraging results.

References

1. Air transport action group (atag), <http://www.atag.org/facts-and-figures.html>
2. International air transport association, iata price analysis, <http://www.iata.org/publications/economics/fuel-monitor/Pages/price-analysis.aspx>
3. Abramson, M., Audet, C., Couture, G., Dennis, Jr., J., Le Digabel, S., Tribes, C.: The NOMAD project. Software available at <https://www.gerad.ca/nomad/>, <https://www.gerad.ca/nomad/>
4. Audet, C., Hare, W.: Derivative-Free and Blackbox Optimization. Springer Series in Operations Research and Financial Engineering, Springer International Publishing, Cham, Switzerland (2017). <https://doi.org/10.1007/978-3-319-68913-5>
5. Bremner, D., Chan, T.M., Demaine, E.D., Erickson, J., Hurtado, F., Iacono, J., Langerman, S., Patrascu, M., Taslakian, P.: Necklaces, convolutions, and X+Y. CoRR **abs/1212.4771** (2012), <http://arxiv.org/abs/1212.4771>
6. Choi, B., Lentz, J., Rivas-Guerra, A., Mignolet, M.: Optimization of intentional mistuning patterns for the reduction of the forced response effects of unintentional

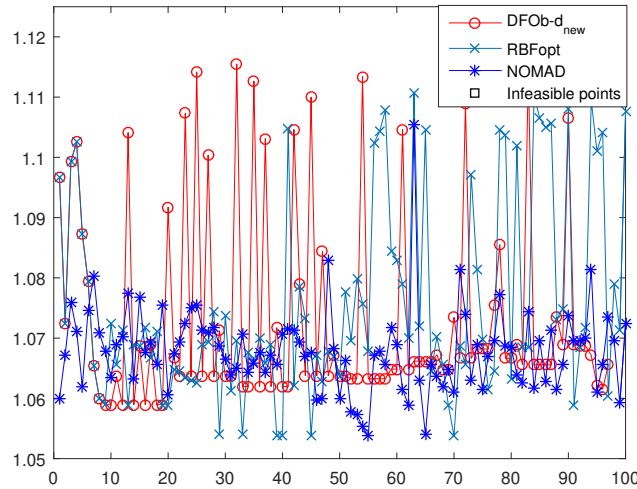


Fig. 5: Compare DFOb, Nomad and RBFopt results obtained for sub-problem with patterns 10 – 01 of Safran application.

- mistuning: Formulation and assessment. *Journal of Engineering for Gas Turbines and Power* **125**(1), 131–140 (1 2003). <https://doi.org/10.1115/1.1498270>
7. Choi, B.: Pattern optimization of intentional blade mistuning for the reduction of the forced response using genetic algorithm. *KSME International Journal* **17**(7), 966–977 (Jul 2003). <https://doi.org/10.1007/BF02982981>, <https://doi.org/10.1007/BF02982981>
 8. Choi, B., Eun, K.H., Jung, K.H., Haneol, J., DongSik, G., Kwan, K.M.: Optimization of intentional mistuning for bladed disk : Intentional mistuning intensity effect. In: Mathew, J., Kennedy, J., Ma, L., Tan, A., Anderson, D. (eds.) *Engineering Asset Management*. pp. 1024–1029. Springer London, London (2006)
 9. Conn, R., D'Ambrosio, C., Liberti, L., Sinoquet, D.: A trust region method for solving grey-box mixed integer nonlinear problems with industrial applications, <https://mode2016.sciencesconf.org/file/223761>
 10. Conn, R., Scheinberg, K., Vicente, L.: *Introduction to Derivative-Free Optimization*. Society for Industrial and Applied Mathematics (2009). <https://doi.org/10.1137/1.9780898718768>, <https://epubs.siam.org/doi/abs/10.1137/1.9780898718768>
 11. Costa, A., Nannicini, G.: Rbfopt: an open-source library for black-box optimization with costly function evaluations. *Mathematical Programming Computation* **10**(4), 597–629 (Dec 2018). <https://doi.org/10.1007/s12532-018-0144-7>, <https://doi.org/10.1007/s12532-018-0144-7>
 12. Fredricksen, H., Kessler, I.J.: An algorithm for generating necklaces of beads in two colors. *Discrete Mathematics* **61**(2), 181 – 188 (1986). [https://doi.org/https://doi.org/10.1016/0012-365X\(86\)90089-0](https://doi.org/https://doi.org/10.1016/0012-365X(86)90089-0), <http://www.sciencedirect.com/science/article/pii/0012365X86900890>

13. Gabric, D., Sawada, J.: Constructing de bruijn sequences by concatenating smaller universal cycles. *Theoretical Computer Science* **743**, 12 – 22 (2018). <https://doi.org/https://doi.org/10.1016/j.tcs.2018.06.039>, <http://www.sciencedirect.com/science/article/pii/S0304397518304559>
14. Gutmann, H.M.: A radial basis function method for global optimization. *Journal of Global Optimization* **19**(3), 201–227 (Mar 2001). <https://doi.org/10.1023/A:1011255519438>, <https://doi.org/10.1023/A:1011255519438>
15. Han, Y., Murthy, R., Mignolet, M.P., Lentz, J.: Optimization of Intentional Mistuning Patterns for the Mitigation of the Effects of Random Mistuning. *Journal of Engineering for Gas Turbines and Power* **136**(6) (02 2014). <https://doi.org/10.1115/1.4026141>, <https://doi.org/10.1115/1.4026141>, 062505
16. Jiang, M.: On the sum of distances along a circle. *Discrete Mathematics* **308**(10), 2038 – 2045 (2008). <https://doi.org/https://doi.org/10.1016/j.disc.2007.04.025>, <http://www.sciencedirect.com/science/article/pii/S0012365X07002555>
17. Jones, D.R., Schonlau, M., Welch, W.J.: Efficient global optimization of expensive black-box functions. *Journal of Global Optimization* **13**(4), 455–492 (Dec 1998). <https://doi.org/10.1023/A:1008306431147>, <https://doi.org/10.1023/A:1008306431147>
18. Le Digabel, S.: Algorithm 909: NOMAD: Nonlinear optimization with the MADS algorithm. *ACM Transactions on Mathematical Software* **37**(4), 1–15 (2011)
19. Liao, H., Wang, J., Yao, J., Li, Q.: Mistuning Forced Response Characteristics Analysis of Mistuned Bladed Disks. *Journal of Engineering for Gas Turbines and Power* **132**(12) (08 2010). <https://doi.org/10.1115/1.4001054>, <https://doi.org/10.1115/1.4001054>, 122501
20. Óttarsson, G.: Dynamic modeling and vibration analysis of mistuned bladed disks. Theses, University of Michigan (May 1994), <https://tel.archives-ouvertes.fr/tel-00598068>
21. Schonlau, M.: Conception robuste en vibration et aéroélasticité des roues aubagées de turbomachines. Ph.D. thesis, Université Paris-Est Marne la vallée (2009), <https://tel.archives-ouvertes.fr/tel-00529002v2/document>
22. Toussaint, G.: A mathematical analysis of african, brazilian, and cuban clave rhythms. In: Townson University. pp. 157–168 (2002)
23. Toussaint, G.: The geometry of musical rhythm. In: Akiyama, J., Kano, M., Tan, X. (eds.) *Discrete and Computational Geometry*. pp. 198–212. Springer Berlin Heidelberg, Berlin, Heidelberg (2005)
24. Toussaint, G.: Computational geometric aspects of rhythm, melody, and voice-leading. *Computational Geometry* **43**(1), 2 – 22 (2010). <https://doi.org/https://doi.org/10.1016/j.comgeo.2007.01.003>, <http://www.sciencedirect.com/science/article/pii/S092577210900042X>, special Issue on the 14th Annual Fall Workshop