International Journal of Mechanical, Industrial and Aerospace Sciences
ISSN: 2517-9950
Vol:3, No:7, 2009

# Simulation and Optimization of Mechanisms made of Micro-molded Components

Albert Albers, Pablo Enrique Leslabay

*Abstract*— **The Institute of Product Development is dealing with the development, design and dimensioning of micro components and systems as a member of the Collaborative Research Centre 499 "Design, Production and Quality Assurance of Molded micro components made of Metallic and Ceramic Materials". Because of technological restrictions in the miniaturization of conventional manufacturing techniques, shape and material deviations cannot be scaled down in the same proportion as the micro parts, rendering components with relatively wide tolerance fields. Systems that include such components should be designed with this particularity in mind, often requiring large clearance. On the end, the output of such systems results variable and prone to dynamical instability. To save production time and resources, every study of these effects should happen early in the product development process and base on computer simulation to avoid costly prototypes. A suitable method is proposed here and exemplary applied to a micro technology demonstrator developed by the CRC499. It consists of a one stage planetary gear train in a sun-planet-ring configuration, with input through the sun gear and output through the carrier. The simulation procedure relies on ordinary Multi Body Simulation methods and subsequently adds other techniques to further investigate details of the system's behavior and to predict its response. The selection of the relevant parameters and output functions followed the engineering standards for regular sized gear trains. The first step is to quantify the variability and to reveal the most critical points of the system, performed through a whole-mechanism Sensitivity Analysis. Due to the lack of previous knowledge about the system's behavior, different DOE methods involving small and large amount of experiments were selected to perform the SA. In this particular case the parameter space can be divided into two well defined groups, one of them containing the gear's profile information and the other the components' spatial location. This has been exploited to explore the different DOE techniques more promptly. A reduced set of parameters is derived for further investigation and to feed the final optimization process, whether as optimization parameters or as external perturbation collective. The 10 most relevant perturbation factors and 4 to 6 prospective variable parameters are considered in a new, simplified model. All of the parameters are affected by the mentioned production variability. The objective functions of interest are based on scalar output's variability measures, so the problem becomes an optimization under robustness and reliability constrains. The study shows an initial step on the development path of a method to design and optimize complex micro mechanisms composed of wide tolerated elements accounting for the robustness and reliability of the systems' output.**

*Keywords*— **Micro molded components, Optimization, Robustness und Reliability, Simulation**

Manuscript received March 27, 2009. We are grateful for the support provided by German Research Foundation (DFG – Deutsche Forschungsgemeinschaft) within the scope of Collaborative Research Center 499 "Design, Production and Quality Assurance of Molded micro components made of Metallic and Ceramic Materials". The authors want to thank Dr. Sandeep Vijayakar, Advanced Numerical Solutions LLC for his valuable support in understanding and adapting the Calyx Solver.

A. Albers is with the Institute of Product Development, Universität Karlsruhe (TH), 76131 Karlsruhe, Germany. Associate Professor and author of correspondence, Phone: +49 (721) 608 2371, e-mail: albers@ipek.uni-karlsruhe.de

Pablo. E. Leslabay is with the Institute of Product Development, Universität Karlsruhe (TH), 76131 Karlsruhe, Germany. e-mail: leslabay@ipek.uni-karlsruhe.de

## I. Introduction

The development of smaller and smaller micro components and systems is an ongoing process. Within the scope of the Collaborative Research Center 499 (CRC) "Design, Production and Quality Assurance of Molded Micro Components made of Metallic and Ceramic Materials" fundamentals in a persistent process chain for micro components are acquired. As a participating member, the Institute of Product Development (IPEK) is dealing with the development, design and dimensioning of micro components and systems. High precision and very tight tolerances are properties commonly related to micro technologies, at least for a group of manufacturing processes and materials. But this is not the only reality in the micro world. The production of metallic and ceramic micro molded components is still in its research phase, and the components obtained with the actual technologies are suffering of a lack of repeatability, and therefore specified with wide tolerances. Thus, when creating functional systems that contain such components, the designer has to accommodate these wide tolerances, and allow significant clearance between the elements. A study of the effects of tolerance and clearance is therefore needed to understand the limits of such systems and to be able to forecast their output's performance. The following study will present and compare the ability of some common Design of Experiments techniques as a tool to qualify and quantify the relevance of varying parameters. With the reduced set of factors derived, investigations will be conducted to develop appropriate methods for determining the robustness or the reliability of the system as a whole. Finally, further reduction of the influence factor set and selection of prospective control ones' will enable to optimize single component geometrical properties in order to improve the system response as desired.

## II. Simulation Model

In order to exemplify the capabilities of the different techniques considered during the research, it is necessary to introduce a system test case. To make the comparisons between methods even more meaningful, it is desirable that system and model remain as constant as possible. In this case, a technology demonstrator developed by the CRC499 is taken as test case. It consists of a one stage planetary gear train in a sun-planet-ring configuration, with input through the sun gear and output through the carrier. Table I gives insights about the dimensions of the demonstrator, and Figure 1 shows an exploded view of the system. This demonstrator allows for a good balance between the quantity of parameters considered plus the output variability included and the modeling complexity of the components intervening. It also represents a system that has been deeply researched in the macroscopic scale so comparison points and typical output characteristics can be derived from the existing literature. Another advantage of choosing this demonstrator is the fact that it exist physically, and that it can be partially tested on a test bench developed by the IPEK for that purpose. Results of comparisons between simulation and test bench runs, based on a standard double flank test, are published in [1]. The gear train output load is set at 20 mNm, which is ambitious but could be confirmed in static tests, while the input speed remains constant. Rotation speed is kept low at 10 RPM as dynamical effects are of less interest here. The flank geometry represented, despite being parametric and highly modifiable, is still away from the profiles measured on real probes [2]. The intention of this setup is to assure that there is no contact interruption during the time each tooth is engaged.

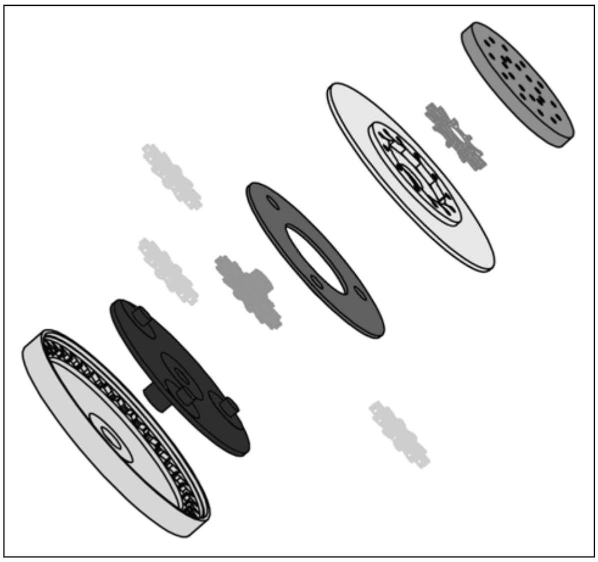

Fig. 1 CRC499 Demonstrator: compressed air driven turbine and one stage planetary gear train assembly.

**TABLE I** GEOMETRICAL CONFIGURATION OF THE PLANETARY GEAR TRAIN

| | Sun | Planet | Ring | Carrier |
|---|---|---|---|---|
| Pitch [mm] | 5.917 | 5.917 | 5.917 | 0.177 |
| N | 14 | 12 | 37 | 1 |
| Depth [mm] | 0.156 | 0.156 | 0.200 | 0.140 |
| Mass [mg] | 5.54 | 2.79 | - | 24.73 |

### A. *Parameter selection*

The selection of relevant variation parameters for the demonstrator described above began by consulting german normative [3, 4, 5] and classical textbooks about macroscopic gear trains [6, 7]. With this information, the specifics of micro gears and gear trains were investigated, again citing existing literature [8, 9, 10] and the internal experience collected by the CRC499. To avoid adding unnecessary complexity to the model, those parameters accounting for small tolerance windows were not included. As a result of this selection work, following parameters were considered as relevant variation agents: radial and tangential position of two of the three planet pins on the carrier; clearance in each of the planet bearings; position of the planet bearing related to its pitch diameter; concentricity of the carrier output shaft to the pins' locator diameter; clearance in the output bearing; concentricity of the input shaft to the sun's pitch diameter and of the input bearing to its theoretical central position; clearance in the input bearing; 4 geometrical parameters for the profile description of each of the five present gears. This gives a grand total of 43 degrees of freedom. Figure 2 describes the meaning of these parameters graphically.

Following the same procedure, outputs of interest could be selected: maximal stress in the tooth's root zone, for durability concerns; highest contact pressure registered, for wear estimation; output waviness (shaft angular deviation from its ideal mathematical position) as an estimative of system's transfer quality. In order to differentiate the effects of all of these parameters, it is necessary to let several teeth to engage and disengage. Actually, it is sought that every component rotates for about 90°, so that non-concentricity effects can be covered with two orthogonal parameters.

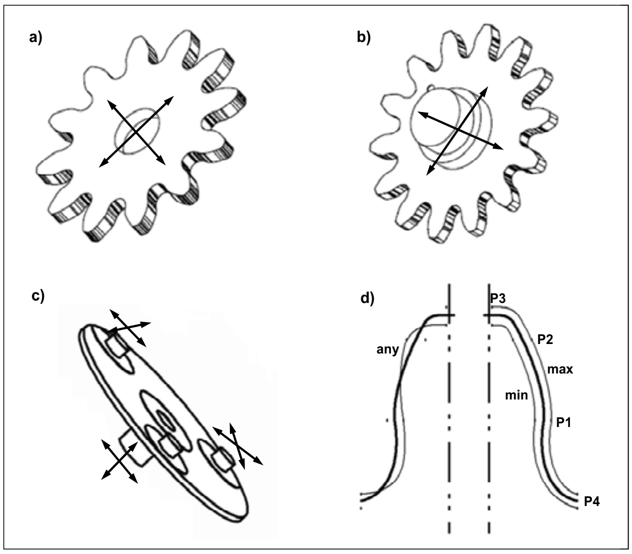


Fig. 2 Geometrical parameters considered in the model: a) bearing position in planet gear; b) axis position on sun gear; c) pin and axis position on carrier; d) profile description parameter with profile variation examples.

### B. *Simulation process*

The investigation is based on a hybrid FEM-MBS procedure. The FEM mesh with higher order elements on the contour assures an adequate approximation of the gear flank in all conditions, while the MBS solver handles the contact force reactions and the component's positions. The 200 steps simulation with step time $\Delta t$ = 7.5 ms runs a four point discretization scheme after the Houbolt decomposition [11] which is a special case of the four-point one to assure unconditional stability. It becomes a requisite to run a fully scriptable simulation tool to perform a complete sensitivity analysis and optimization task. Specialized gear train simulation software from Advanced Numerical Solutions LLC seems adequate for the task. The gear geometry generation module Multyx and the dynamic and contact solver Calyx were adapted to reduce the computation time as much as possible. Step time was increased just until the relevant details began to vanish. Contrary as in other FEM based approaches [12], the simulation step width Δt is not a trouble source for the solver due to the manner in which the static and dynamic effects are computed, see eq. 1.

$$M\ddot{x} + C\dot{x} + Kx = f_{ext} + f_{inertial} \quad (1)$$

where x is the vector of d.o.f. of the structure, $f_{ext}$ is the vector of externally applied loads, $f_{inertial}$ is the vector of inertial loads, M is the structure mass matrix, C the damping matrix and K the stiffness matrix. By the end, the simulation time was reduced to 23% of its original, from 30 to 7 minutes. Currently the process has a weakness, because the geometry files corresponding to the 4 profile parameters have to be created previously. For a level study this is not a problem, but when it comes to the optimization, the geometry parameters will vary in a continuous form, thus requiring new geometry files for each iteration. The creation of these files requires with the actual state-of-the-art about the same time as the simulation itself. A direct mesh morpher is very hard to implement because of the higher order FE mesh used by Calyx. The direct variation of the internal parameters in the geometry generator Multyx does not assure a 2-way correlation between the changed values and the profile generated, rendering the approach useless. To overcome the problem, the 81 geometry variations for each gear were pre-computed drawing on an intermediate MatLab step that includes an optimization cycle using the simplex search method [13] to near the Multyx geometry to the desired one. Figure 3 shows the complete simulation process graphically. Up to 4 parallel Calyx solver runs could be managed.

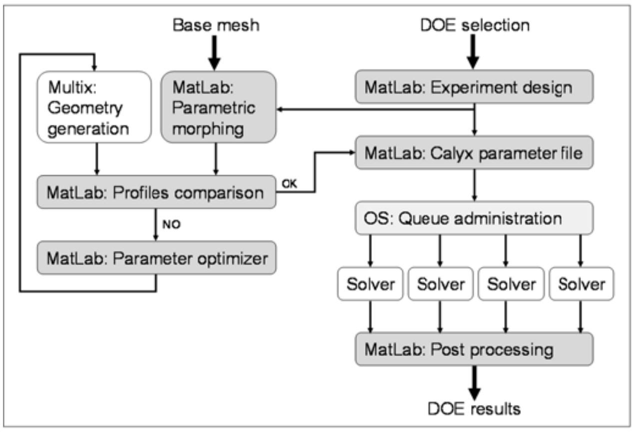


Fig. 3 Simulation and profile variation processes.

## III. Design of Experiments

Design of Experiment (DOE) is a structured, organized method used to determine the relationship between the different factors (X's) affecting a process and the output of that process (Y). This method was first developed in the 1920s and 1930s, by Sir Ronald A. Fisher, a renowned mathematician and geneticist. Design of Experiment involves designing a set of experiments in which all relevant factors are varied systematically. The results of these experiments help to identify optimal conditions, the factors that most influence the results, and those that do not, as well as details such as the existence of interactions and synergies between factors. Their ability to determine which factors are relevant to the desired output and which not is the reason for their introduction into this study. Reducing the initial set of 41 parameters present in the model can be very advantageous for future nominal-geometry optimization steps, which is the ultimate goal of the study. This goal constitutes a priori a typical screening exercise and the corresponding techniques are the first to be explored. No special care needs to be taken about randomizing or blocking the simulation runs, because there are no uncontrolled factors affecting the simulation model. Actually, there is no reason to repeat experiments neither (replication), because no divergent points have been detected during a run; runs with the same initial parameter set generate exactly the same results [14]. Although it is possible and computationally affordable to study the influence of all factors simultaneously, more methods can be tested and compared when operating with smaller subsets of similar parameters. In the case of the micro planetary gear train, it is possible to divide the set into two sections, one containing the parameters purely related with the teeth profile and the other merging the geometrical ones. This particularity has been exploited to derive a subset of 12 contour parameters, 4 for each of the gear types involved (sun, planet, ring), affecting the root, tip relief and real profile angle as depicted in Figure 2d. The following part of the study involving factorial and related designs have been conducted for this subset.

### A. *Analysis of profile parameters*

To start with the simplest design case, two possibilities for screening procedures have been considered. Firstly, a typical Resolution III fractional factorial design with 16 runs for the 12 parameters is created. As the tables summarized in [15, 16] do not contain this case explicitly, the design is created from a mathematical algorithm within MatLab. The design produced maintains equivalent properties. Secondly, a 'saturated main effect' or Plackett-Burman design is generated. Due to the parameters number this accomplishes not from the tables published in [17] but from a Hadamard [18] matrix. Starting with a 16 by 16 matrix, any 12 columns except for the first one can be selected to create the desired design. After screening, the study focuses on gaining more insight in the model's outputs, being the next simplest design to consider the 2 level fractional and full factorial designs. The intention is to compare the prediction quality of several fractional designs with the objective of avoiding unnecessary computations in the future. A full factorial study (4096 runs) will be conducted for this case only as a comparison base, and then, fractional designs of Resolution IV (32 runs), V (256 runs) and VI (512 runs) will be extracted and their predictions compared. Keeping in mind the ultimate target of the investigation, the optimization of the system's outputs, considerations about the design requirements for the generation of response surfaces should be taken into account when analyzing different experiment designs. Box summarized these requirements in [19]: rotatability, residuals minimization, proper 'lack of fit' detection, internal error estimation, constant variance check, sequential construction for higher order designs, blocking capability, etc. Two level designs are clearly not adequate for the task; three to five levels are commonly used for this purpose. If two level full factorials where difficult to conduct, three or more levels are completely impracticable; fractional designs become very tricky to define and are not rotatable. Two multilevel design families derived from other methods are widely adopted: the Central Composite Designs (CCD) and the Box-Behnken designs. From the three CCD variations existing, circumscribed, inscribed and face centered, the latter (CCF) has been selected due to its minimal level requirements. The CCF design starts from a two level full factorial and adds 24 runs for the 12 star points and 236 centre point runs. Its major withdraw is its lack of rotatability. Box-Behnken (BB) designs are independent quadratic designs that do not contain an embedded factorial or fractional factorial design, their mayor withdraw is its poor prediction capability at the corners of the parameter space due to the missing probes. The design used for this study contains 192 edge runs and 12 centre point runs. There is another family of experiment designs, more adequate for a response surface fitting, which offers a better distribution of the prediction error. Called optimality criteria designs, are based either on the prediction variance as defined in Eq. 2 or on information about the variance of the regression parameters described by the covariance matrix. To minimize Cov(β), a characteristic scalar value for the matrix evaluation has to be chosen.

$$Cov(\beta) = \sigma^2 \left(X^T X\right)^{-1} \qquad (2)$$

Among the many possibilities (a short description of each is included in [20]), the D-Optimality criteria is adopted. It maximizes the determinant of Fisher's information matrix $X^TX$. This matrix is proportional to the inverse of the covariance matrix of the parameters. So maximizing its determinant is equivalent to minimizing the determinant of the covariance of the parameters [21] and also minimizes the volume of the regression estimates' confidence ellipsoid for the linear model parameters β. Several approaches exist to find this optimal designs; this study uses the coordinate exchange algorithm [22] as implemented in MatLab. Two different designs have been created, one with 256 and another with 512 experiments, both of them aimed to accommodate a quadratic response surface. After analyzing the results obtained, a progressive augmentation up to 4096 experiments of the later design has been conducted for further investigation. Complete results can be found in §IV and §V.

### *B. Analysis of geometrical parameters*

Besides of benchmarking different DOE methods for the specific problem of wide tolerances, the analysis of §III.A identified the essential parameters in representing the gear's profile variations. The next step is to repeat the study for the geometry parameters. This is accomplished with the best performing method, the 3 level D-Optimality criteria, for a set of 29 parameters including the 4 relevant profile parameters. One important geometrical parameter, the planet's centre radius, is excluded from this part of the study as it is the principal target candidate for the coming optimization step. The first attempt involved a design with 1500 probes, but difficulties with the solver arose. Up to 40% of the parameter combinations could not be solved for, due to instable or even impossible starting conditions. In other cases, although being able to start, the solver diverged at some point, mostly when one of the planets could not establish contact with either sun or ring. Both situations were also observed in some physical tests and contributed to trigger this research work. The set of failed runs and its variance were investigated to pinpoint the parameters responsible for the abnormal condition. Unfortunately, no conclusive responsible could be drawn. To resolve the problem in the simulations, every side of parameter space metacube was gradually reduced to its half, thus losing some significance in the space corners. The modification reduced the amount of failed simulations to about 25%, and the total number of probes was increased to 2400 in an attempt to compensate for the lost information. Refer to §IV and §V for the final results of the procedure.

## IV. Benchmark of Metamodeling Methods

The 3 outputs of interest mentioned in §II.A give place to 8 different output data families which in turn allow 54 post processing rules for the 200 measurement points per run. Different regression models can be fitted to each of these outputs: linear, linear interactions, extended interactions (3, 4 and 5 way aliasing free models) and quadratic models with and without extended interactions. They will be called here: Lin, Int, Int3 /-4/-5, Quad, QI3 /-4/-5. The whole procedure is then repeated for different experiment design creation methods. All the possible combinations would be worth a DOE study themselves! Due to its simple evaluation and easy comparison, the criteria to benchmark the metamodels will be the R^2 and Radj^2 statistics. We concentrate here in showing differences in the prediction ability of the methods and select therefore just 3 prospective data fields, from the 54 available, to compare them. The outputs are the carrier axe angular position error variation during the run (Out1), the teeth contact force (Out2), and the tooth bending moment (Out3). In all three cases, the 10 maximal values registered in the run are averaged to avoid peaks arising from numerical instability or other solver induced noise. Starting with the screening exercise, the only possible prediction is about the main factors so a linear regression model is selected. As seen in Table II, both models fit the data quite well, being the Plackett-Burman slightly more robust.

**Table II** Model fit statistics for screening DOE

| | Out1 | | Out2 | | Out3 | |
|---|---|---|---|---|---|---|
| | **P-B** | **Res III** | **P-B** | **Res III** | **P-B** | **Res III** |
| R^2 | 0.978 | 0.9185 | 0.986 | 0.8801 | 0.926 | 0.8783 |
| Radj^2 | 0.890 | 0.5926 | 0.928 | 0.4004 | 0.630 | 0.3915 |

Then the attention is focused on the level 2 factorial designs. As the number of probes increases, more comprehensive models can be fitted and the importance of the Radj^2 statistical becomes clear, as an indicator of unnecessary regressor terms. Table III summarizes the model fit statistics for increasingly complex models and four different experiment designs. The fractional factorial design definition is given in §III.A, "Full" means a full factorial design. To save space, only the worst performing output from the screening fit is displayed here, to show the effects of increasing the design resolution. Augmenting the design resolution above the theoretical needs is not beneficial for the linear model or for the linear interactions model. It is not clear why the Res VI design was not able to fit an Int3 model by yielding a bad conditioned matrix. This prevented another interesting comparison point with the full factorial design. The latter design shows no trend of getting 'saturated' as the model terms augment and the Radj^2 value does not start to sink, indicating that higher order interactions are relevant for the model.

**TABLE III** MODEL FIT STATISTICS FOR FACTORIAL DOE, OUTPUT 3

| | | Out3 | | | |
|---|---|---|---|---|---|
| **Model** | **Stat.** | **Res IV** | **Res V** | **Res VI** | **Full** |
| Lin | R^2 | 0.3946 | 0.3394 | 0.2583 | 0.2764 |
| | Radj^2 | 0.0122 | 0.3068 | 0.2405 | 0.2743 |
| Int | R^2 | - | 0.7512 | 0.6033 | 0.5895 |
| | Radj^2 | - | 0.6415 | 0.5319 | 0.5815 |
| Int3 | R^2 | - | - | * | 0.6753 |
| | Radj^2 | - | - | * | 0.6644 |
| Int4 | R^2 | - | - | - | 0.6847 |
| | Radj^2 | - | - | - | 0.6704 |
| Int5 | R^2 | - | - | - | 0.6942 |
| | Radj^2 | - | - | - | 0.6773 |

Linear models are very efficient for sorting out relevant parameters, but are not adequate for more complicated tasks like robustness or reliability assessment or optimization. A good method to check if a linear model is able to deliver a good system prediction is to compare the prediction for the parameter space's centre point with a simulation run. The result in this case is that the linear model is a poor metamodel for the system's behavior. Therefore, the study is augmented by including designs that can accommodate quadratic models. Table IV summarizes the model fit statistics for increasingly complex models and five different experiment designs, see §III.A for a terms explanation. Increasing the design complexity enhances the fit's quality, as indicated by the rising Radj^2 statistic. The decreasing R^2 is due to the increase in the probes to interpolate, but its meaning should not be misunderstood. As for the model complexity, only the larger experiment designs provide enough information for the higher order model to fit. In most of the cases a QI3 model is enough as indicated by the staging or falling Radj^2 statistic. Adding more model terms is redundant and only increases R^2 falsely. The full factorial roots of the CCF design are reflected in its ability to accommodate higher models. The Box-Behnken design is only suited for the targeted quadratic model, and the D-optimal designs become rapidly saturated for complex models.

Finally, from the knowledge won so far, the model fitting procedure for the second part of the DOE as explained in §III.B went straightforward. Table V shows the 3 considered outputs fit statistics, both for the first failed attempt as for the more successful second one. As the parameter space is now considerably larger, higher order interaction models would require too many experiments, and thus the modeling effort is limited to QI3 models. From the table it is clear that output 1 is not being well fitted, and this tendency sustained for other related outputs not considered for this work. Conclusion should be extracted carefully from them. For the remaining outputs, the selected model fits the data very satisfactorily.

**TABLE IV** MODEL FIT STATISTICS FOR 3 LEVEL DOE, OUTPUT 3

| | | | | Out3 | | |
|---|---|---|---|---|---|---|
| **Model** | **Stat.** | **CCF** | **BB** | **D-256** | **D-512** | **D-2048** |
| Int4 | R^2 | 0.6566 | - | 0.8134 | 0.7153 | 0.6156 |
| | Radj^2 | 0.6419 | - | 0.3821 | 0.5632 | 0.5790 |
| Quad | R^2 | 0.6048 | 0.6871 | 0.6504 | 0.5898 | 0.5412 |
| | Radj^2 | 0.5964 | 0.4379 | 0.4598 | 0.5021 | 0.5201 |
| QI3 | R^2 | 0.6870 | - | 0.7955 | 0.6938 | 0.6206 |
| | Radj^2 | 0.6763 | - | 0.5260 | 0.5725 | 0.5917 |
| QI4 | R^2 | 0.6961 | - | 0.8742 | 0.7315 | 0.6324 |
| | Radj^2 | 0.6822 | - | 0.5063 | 0.5726 | 0.5948 |
| QI5 | R^2 | 0.7052 | - | 0.9412 | 0.7729 | 0.6474 |
| | Radj^2 | 0.6890 | - | 0.4833 | 0.5928 | 0.6036 |

**TABLE V** MODEL FIT STATISTICS FOR GEOMETRICAL DOE

| | Out1 | | Out2 | | Out3 | |
|---|---|---|---|---|---|---|
| | **Att. 1** | **Att. 2** | **Att. 1** | **Att. 2** | **Att. 1** | **Att. 2** |
| R^2 | 0.7724 | 0.7226 | 0.8972 | 0.8878 | 0.9347 | 0.9126 |
| Radj^2 | 0.4945 | 0.4607 | 0.7716 | 0.7819 | 0.8549 | 0.8302 |

## V. SENSITIVITY ANALYSIS RESULTS

After analyzing the results seeking for the most adequate metamodels to predict the system's behavior, the factors are sorted after their for the outputs' variance relevance. This more challenging task will yield now some previously favored models as inadequate. To better understand the following tables, the parameter ordering remains the same through the gear profile study, with factor X1 corresponding to P1 for the ring gear, factor X2 to P2, …, factor X5 to P1 for the sun, …, factor X9 to P1 for the planet, and so on. Please refer to Figure 2 for details about the geometry parameters P. Looking at the screening results in Table VI, the methods do not agree in the importance assigned to the factors and no clear tendency across the three outputs could be found. The reliability of screening methods is then for our test case low and no conclusions should be extracted from such a study.

**TABLE VI** RESULTS FOR SCREENING DOE

| Out1 | | Out2 | | Out3 | |
|---|---|---|---|---|---|
| **P-B** | **Res III** | **P-B** | **Res III** | **P-B** | **Res III** |
| X3 | X1 | X1 | X1 | X11 | X11 |
| X1 | X9 | X3 | X8 | X6 | X10 |
| X11 | X8 | X2 | X9 | X10 | X6 |
| X12 | X11 | X11 | X11 | X1 | X5 |
| X5 | X2 | X10 | X4 | X9 | X8 |
| X8 | X12 | X9 | X12 | X5 | X9 |
| X7 | X10 | X12 | X3 | X4 | X7 |
| X4 | X3 | X8 | X7 | X7 | X2 |

As can be seen in Table VII, higher resolution fractional factorial designs perform as well as the full factorial, saving up to 7/8 of the experiments. The ranks for the resolution VI and the full factorial designs are quite concordant, despite the different models used. The Res VI design can only fit an interactions model, while the full factorial became an I5 model. The importance of the interaction terms for this system can be confirmed when comparing the new ranked list against the one obtained from the screening experiment. The following tables will constrain the output to the 10 most relevant factors.

**TABLE VII** RESULTS FOR FACTORIAL DOE, OUTPUT 3

| Out3 | | | | |
|---|---|---|---|---|
| **Res III** | **Res IV** | **Res V** | **Res VI** | **Full** |
| X11 | X11 | X10 | X10 | X10 |
| X10 | X1 | X9 | X9 | X9 |
| X6 | X5 | X6 | X11X12 | X11X12 |
| X5 | X10 | X1X2 | X2X9 | X1X2 |
| X8 | X9 | X11X12 | X2X4 | X2X10 |
| X9 | X12 | X2X9 | X1X4 | X2X9 |
| X7 | X8 | X1X10 | X9X12 | X2X4 |
| X2 | X6 | X2X10 | X2X10 | X1X4 |
| X4 | X2 | X1X4 | X1 | X1X10 |
| X3 | X4 | X11 | X5 | X1X2X10 |

Table VIII shows the performance of the 3 level designs fitted with the best meaningful models possible. These are QI4, Quad, QI3, QI3 and QI4 respectively. The CCF model over estimates the importance of the quadratic factors, maybe because of the small amount of information added by the 24 star-point runs. The Box-Behnken design overcomes this problem but the ranking derived shows no convergent tendency with the D-optimal designs. These in fact tend to repeat the ranking as the design size and model fit quality increases. This behavior gives some confidence about the decisions derived from those designs.

**TABLE VIII** RESULTS FOR FACTORIAL DOE, OUTPUT 3

| Out3 | | | | |
|---|---|---|---|---|
| **CCF** | **BB** | **D-256** | **D-512** | **D-2048** |
| X2^2 | X2^2 | X1^2 | X2^2 | X2^2 |
| X11^2 | X2X3 | X2^2 | X10 | X10 |
| X3^2 | X6^2 | X7^2 | X9 | X9 |
| X1^2 | X10X12 | X4^2 | X1^2 | X1^2 |
| X10 | X4X12 | X3^2 | X1X2 | X1X2 |
| X6^2 | X5^2 | X10 | X1X10 | X11X12 |
| X5^2 | X9X11 | X9 | X4^2 | X2X4 |
| X9 | X10X11 | X1X2X10 | X11X12 | X1X10 |
| X8^2 | X1X12 | X1X10 | X7^2 | X2X10 |
| X11X12 | X11X12 | X11X12 | X2X10 | X1X4 |

Finally, from the analysis performed for tables VII and VIII, two designs and their correspondent metamodel were selected to complete the sensitivity analysis and extract the four or five most important parameters affecting the gear's profile. A resolution V design fitted with a linear interactions model represents the 2 level designs. Until the Res VI becomes able to fit the I3 design, there are no advantages in using it. From the quadratic designs, the D-Optimal method consisting of 2048 probes and fitted with a QI4 model is selected. The ranking from both designs for the 3 outputs considered are shown in Table IX. The same analysis was repeated for other output values as well, before the final decision about the parameters to continue proceeding with was taken. These are parameters X1 and X2, corresponding to the ring's main profile, and X9 and X10, the homonymous for the planets. It is, as usual in engineering, a compromise solution, as the sun in example is no longer represented. The sun's parameters are only relevant for a few outputs, but the selected ones are present in every list and represent a large proportion of the model's variance.

**TABLE IX** FINAL RESULTS FOR THE PROFILES DOE

| Out1 | | Out2 | | Out3 | |
|---|---|---|---|---|---|
| **Res V** | **D-2048** | **Res V** | **D-2048** | **Res V** | **D-2048** |
| X1 | X1 | X1 | X1 | X10 | X2^2 |
| X1X10 | X1X2 | X1X2 | X1X2 | X9 | X10 |
| X1X2 | X1^2 | X1X3 | X2 | X6 | X9 |
| X1X3 | X1X10 | X2 | X1^2 | X1X2 | X1^2 |
| X12 | X2^2 | X1X10 | X3^2 | X11X12 | X1X2 |
| X3 | X7^2 | X2X9 | X2X10 | X2X9 | X11X12 |
| X1X9 | X1X9 | X2X10 | X1X10 | X1X10 | X2X4 |
| X11X12 | X12 | X12 | X9X12 | X2X10 | X1X10 |
| X4X10 | X9X12 | X3 | X2X9 | X1X4 | X2X10 |
| X4X9 | X2 | X3X9 | X2X3 | X11 | X1X4 |

The geometrical DOE was constructed around the knowledge obtained so far, so no linear designs were included, just D-optimal. The factors X have now a new physical correlation: X1 to X3 represent the carrier axis position and bearing clearance, X4 to X7 the pin position as shown in Figure 2.c; X8 to X12 the sun axis and bearing position and clearance (Fig 2.b); X13 to X21 the bearing position and gap for each planet as shown in Fig 2.a; and finally X22 to X29 the gear profile variations for the ring gear and each of the planets. Table X is the homonymous to Table IX and related to Table V for the latest results. Again, only the information about the outputs described in this work is included, but the study was quite extensive before choosing the final parameters to transfer to the reliability and robustness assessment and optimization procedures. These parameters are, sorted after importance, X4 and X6, X1 and X2, X22, X16 and X8. Later, X5, X7, X9 and X17 were included to completely define the corresponding spatial positions. Besides the initially designated planets centre radius, other 5 parameters were identified as possible optimization control factors. These are the 5 bearing clearances X3, X12, X15, X18 and X21.

**TABLE X** RESULTS FOR GEOMETRICAL DOE

| Out1 | | Out2 | | Out3 | |
|---|---|---|---|---|---|
| **Att. 1** | **Att. 2** | **Att. 1** | **Att. 2** | **Att. 1** | **Att. 2** |
| X6^2 | X22^2 | X17^2 | X4 | X4 | X4 |
| X5^2 | X8^2 | X5^2 | X6 | X6 | X6 |
| X22^2 | X7^2 | X4 | X22^2 | X1 | X6^2 |
| X15^2 | X4^2 | X6 | X4^2 | X16^2 | X4^2 |
| X1 | X2^2 | X16^2 | X7^2 | X7^2 | X22^2 |
| X7^2 | X5^2 | X1 | X6^2 | X5^2 | X1 |
| X18^2 | X9^2 | X18^2 | X8^2 | X10^2 | X4X6 |
| X3^2 | X6^2 | X10^2 | X1 | X3^2 | X5^2 |
| X28^2 | X18^2 | X7^2 | X12^2 | X8^2 | X26^2 |
| X25^2 | X14^2 | X8^2 | X10^2 | X17^2 | X3^2 |

## VI. FIRST OPTIMIZATION ATTEMPTS

The optimization of systems resting only upon direct experimental data can become very time consuming if the number of control inputs is high. If the system under study includes parameters (whether control or noise) with stochastic characteristics, the rapidly increasing amount of observations will turn this absolutely unviable. The methods described so far are intended to be applied to such systems and to help reducing the amount of relevant parameters to consider. Until now, the study relied on deterministic designs, where the factor's values were logically assigned. But these factors represent the components physical tolerance field arising from manufacturing uncertainty. So any general assertion about the system's outputs will have to deal with the random variation of these parameters. Different strategies exist to accomplish this task [23, 24] that will not be discussed here. All of them end up providing some scalar values that represent the output's variability due to (random) parameter variation. These values are then feed to the optimization algorithms. Spall explores in [25] methods to exploit knowledge about the stochastic variation's nature including it into the search and optimization procedure.

An important amend needs to be implemented to the simulation model as for now on the parameters will not vary between fixed levels anymore, but continuously along the complete tolerance field. This sets a high hurdle as already explained in §II.B the possibility to vary the gears' profiles stepless is not given yet. To overcome the obstacle the profile variation's range was laboriously discretized into 10 levels for the planet and ring gears' main profile, generating 100 contour variations for each gear. In §V, a set of plausible control factors has been derived from the analysis. The first step of the optimization procedure will be then to confirm the potential of these factors under service conditions, which is with stochastically variating noise parameters, not the design-assigned anymore. Five candidates were considered: the planets distance to the carrier axis; the clearance or gap in the bearings from carrier, sun and planets; and finally the overall size of the planets. This last should not be confounded with the parameter study of §III.A. These parameters are not deterministic themselves and have also a variation range. To account for all of the variation sources, noise and control, the nominal point around which the affected control tolerance field is defined, is displaced alternatively to the extremes of the original field (called xx+ and xx- later) thus extending it by one half its width in each direction, see Figure 4. During the computation of the designs for §III.B, it was noticed that difficulties arise with extreme geometrical conditions. This would be aggravated for this part of the investigation as the search field becomes extended. To minimize possible problems, the variations were conducted one factor at a time, accepting the lost of the interaction information but achieving very less failed simulations, fewer than 5% of the total. Table 11 summarizes the variation that was registered in the mean value and the standard deviation, typical robustness indicators, of the three outputs under investigation. The calculations were also used to test methods to spare probes in the assessment of the output's variability as mentioned before. As the planets' size results to be not more important than the other candidates and the difficulty of modelling it is quite large, it has been removed from the control parameters for the initial attempts.

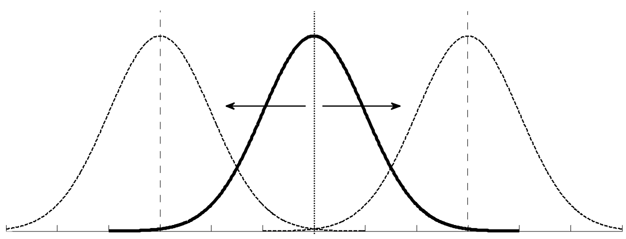

Fig. 4 Geometry nominal point displaced to field ends.

TABLE II STOCHASTIC VARIATION OF PROPOSED CONTROL PARAMETERS

| | Out1 | | Out2 | | Out3 | |
|---|---|---|---|---|---|---|
| | **Mean** | **Std.** | **Mean** | **Std.** | **Mean** | **Std.** |
| Nominal | 2.536 | 1.027 | 3.340 | 1.225 | 5.070 | 1.807 |
| CenterD- | 2.676 | 0.972 | 3.378 | 1.238 | 4.992 | 1.773 |
| CenterD+ | 2.650 | 1.077 | 3.410 | 1.249 | 5.254 | 1.872 |
| GapCar- | 2.373 | 1.269 | 3.362 | 1.210 | 5.122 | 1.855 |
| GapCar+ | 2.697 | 1.085 | 3.281 | 1.096 | 5.006 | 1.654 |
| GapPla- | 2.546 | 1.092 | 3.327 | 1.252 | 5.048 | 1.842 |
| GapPla+ | 2.521 | 1.040 | 3.382 | 1.194 | 5.176 | 1.770 |
| GapSun- | 2.160 | 0.941 | 3.388 | 1.217 | 5.116 | 1.752 |
| GapSun+ | 3.101 | 1.224 | 3.389 | 1.216 | 5.168 | 1.781 |
| PlaSize- | 2.625 | 1.113 | 3.285 | 1.198 | 4.958 | 1.735 |
| PlaSize+ | 2.326 | 1.027 | 3.333 | 1.207 | 5.709 | 1.798 |

The payoff of having eliminated irrelevant parameters from the model becomes now evident. Determination of the scalar variability indicators require in the best case (FOSM based on gradients) one solver run per control or noise stochastic variable. Every parameter, including the optimization control ones, is affected by stochastical variation and included in the FOSM determination. If gradients are not adequate, the Latin Hypercube method will use about 2 probes per variable, and a pure random Monte Carlo needs around 5. In our test case it implies 16, 32 or 80 solver passes for each determination! Thus, the possible use of metamodel based approaches like Response Surfaces should be kept in eye. Unfortunately, a FOSM approach is not applicable in this case. The reason can be best visualized in Figure 5. The method is originally conceived for tolerance fields narrow enough for a linear approximation (gradient) to be valid; the gradient calculation is then repeated as the X value changes, see the light probability distribution curves in the figure. But for tolerance fields that are comparable in width to the optimization space itself, the method's linear behaviour ground hypothesis is not sustainable, as the example with the darker pdf curve shows, where the gradient obtained approximates the right-hand one precisely, but the left-hand one not at all. Accepting it will imply that a linear approximation is a valid response model for Y, which has been already proven false in §IV, and that the optimization could be also performed very quickly and inexpensively over a linear metamodel.

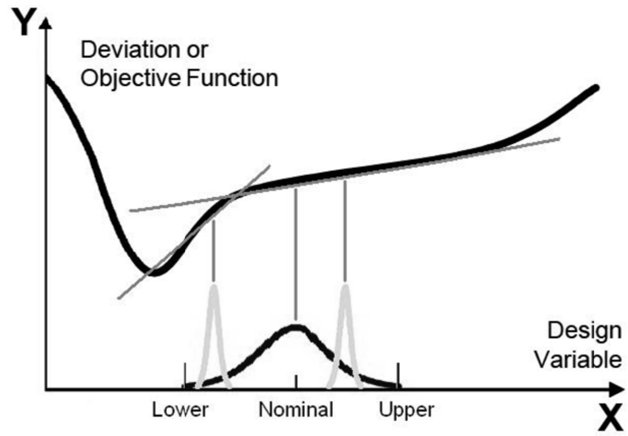


Fig. 5 First order second moment (FOSM) method limitations due to excessively wide approximation space.

The same argumentation undermines the sense of introducing gradient based optimization methods. However, some of them were tested (with scepticism) against random based

ISSN: 2517-9950
Vol:3, No:7, 2009

evolutionary algorithms to confirm that none is converging to a global solution. To simplify the task, the stochastic variation of the parameters of interested was ignored for this example. Slightly changing the starting point yields different results, and the methods do not obtain the same solution from identical starting conditions. Tables XII and XIII illustrate these sayings. From a Mixed Integer Programming (to give the gear profile parameters a chance), a Nonlinear Programming Quadratic Line Search, a Simulated Annealing, a Differential Evolution strategy and a Self Adaptive Evolution, the NPQLS performed the best regarding amount of required probes and progress stability, but with a very limited search space. A global search process relying solely on evolutionary algorithms or other stochastically driven methods cannot be sustained due the enormous computational expense, as every element requires the mentioned solver passes to determine the robustness indicators. As seen in Table XII, an extremely optimistic minimum of 200 probes are required for such a search, which multiplied by the 30 solver passes needed to determine the robustness via Latin Hypercube, rockets the total calculations to at least 6000!

**TABLE XII** OPTIMIZATION STRATEGIES COMPARISON

| Algorithm | Probes | Start | End | Improvement |
|---|---|---|---|---|
| Mixed Integer | 180 | 0.8086 | 0.5072 | 37% |
| NLPQL | 210 | 0.8086 | 0.4366 | 46% |
| Simulated annealing | 200 | 0.8086 | 0.4305 | 47% |
| Diff. Evolution | 1200 | 0.8086 | 0.2340 | 71% |
| Self. Adap evo | 400 | 0.8086 | 0.2714 | 66% |

The introduction of metamodels in the optimization loop becomes mandatory. Preliminary results obtained from a broadly accepted algorithm known as Efficient Global Optimization (EGO) [26], which bases on a kriging hypersurface and the expected improvement criterion, are very promising. The output's variability could be reduced for every output individually, up over 60% from the nominal design's mark, but the parameter configurations obtained diverge considerably, see Table XIII. The parameter values in italics have reached the field limit. The difference between the procedures labelled EGO and EGO 2 resides in the methodology followed to determine the position of the additional interpolation points as the optimization advances, also called iterations. For the first one, the new promising point is the result of a global optimization on a surface which is the Kriging model minus a selectable number of times the estimated standard deviation. In the second case the point is determined via a global optimization on the expected improvement function, which is a probabilistic unction using the Kriging surface and the estimated standard deviation. As the initial search of the algorithm remains unchanged, both cases start the actual optimization from the same point, but arrive to different results. Under Probes, the (small) amount of interpolation points and the correspondent charge of solver passes is indicated. Increasing it is not always a guarantee of better results, a desirable behavior.

Performing a multi-criteria optimization to overcome this limitation is unfortunately not realistic in the near future with the current solution times as it would imply performing several nested optimization with similar consequences as the global search.

Figure 6 shows a typical EGO progress chart. The parameter space limits are represented by the x-axis and the dotted lines, the dashes indicate the output values from the current iteration, the solid lines the position of the best combination so far. It is interesting to note that, after EGO completes the initial search and optimization, most of the configurations explored (iterations) render worse solutions than the best mark so far, which should not be confounded with the initial nominal configuration!, showing again the complexity of the objective functions under study. It can be also read from the last Tables, that the starting point from the EGO is usually worse than the condition at the nominal point, so the improvement achieved by the method should be read carefully. It becomes also important to allocate enough resources to the determination of this starting point.

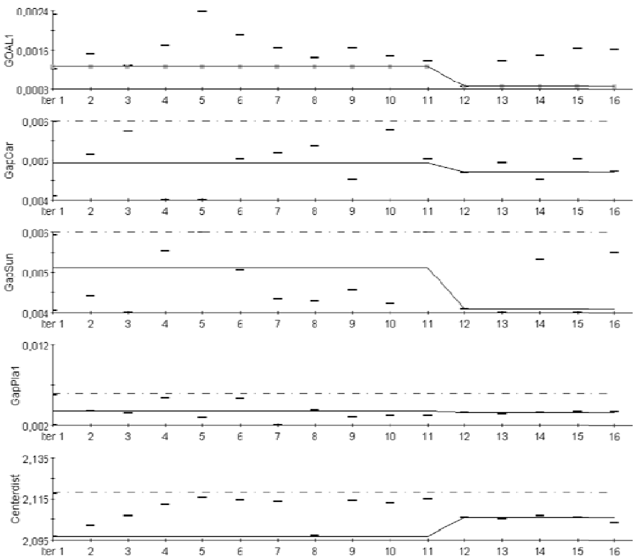


Fig. 6 Typical EGO progress chart.

**TABLE XIII** OPTIMIZATION RESULTS FROM TWO EGO METHODS

| Algorithm | Probes | Goal | Nominal | Start | End | Improvement | Par1 | Par2 | Par3 | Par4 |
|---|---|---|---|---|---|---|---|---|---|---|
| EGO | 35 / 1050 | mp5 | 0.240 | 0.294 | 0.230 | 4% | *6.000* | 4.118 | 5.726 | 2.110 |
| EGO | 35 / 1050 | dCm10 | 1.034 | 1.250 | 1.004 | 3% | 5.244 | 4.468 | 3.398 | 2.106 |
| EGO | 50 / 1500 | fp5 | 1.585 | 1.570 | 1.496 | 6% | 4.118 | 4.015 | 2.440 | 2.107 |
| EGO | 40 / 1200 | dSm5 | 0.538 | 0.322 | 0.200 | 63% | 4.318 | 5.283 | *6.000* | *2.096* |
| EGO 2 | 35 / 1050 | dCm10 | 1.034 | 1.250 | 0.847 | 18% | 4.700 | 4.088 | 3.548 | 2.106 |
| EGO 2 | 50 / 1500 | fp5 | 1.585 | 1.570 | 1.530 | 3% | 5.418 | 4.042 | 5.915 | *2.118* |

## VII. Conclusion

The computer simulation model of a micro planetary gear train introduced in this work could be successfully deployed as a technical example of a mechanical system that is affected by imprecise components manufacture. Such a system should be deeply investigated before starting with the mass production to assure that its correct function will be assured up to a desired level. A possible methodology to perform this task has been proposed, successfully demonstrated and its results analyzed.

A key step in the procedure is to eliminate variation elements that increase the complexity but not contribute to the solution from the model. A Design of Experiments and a subsequent sensitivity analysis are very adequate methods for this task. The application of both has been extensively discussed in chapters III to VI. As of DOE concerns, optimality criteria designs have been proved better performing than classical screening, (fractional) factorial and 3 level tabulated designs.

It has not been extensively discussed in this study, but the selection of the functions representing the system's output(s) of interest indicates the importance of counting with the complete dataset to easily play variations with.

The work denotes a possible way for first optimization trials on dynamically unstable system, relying on the experience won with the model through the sensitivity analysis step of the process. Thus the optimization process will be very dependent on the output focused in, and further research is required when trying to optimize for multiple outputs.

## References


[1] Albers, A.; et all, An integrated Approach for Validating Micro Mechanical Systems based on Simulation and Test, *MicroSystem Technologies*, Springer, 2008.

[2] Albers, A.; et all, Dealing with Uncertainty of Micro Gears - Integration of Dimensional Measurement, Virtual and Physical Testing, *ASME International Mechanical Engineering Congress & Exposition*, 2008.

[3] DIN 867:1986-02, *Basic rack tooth profiles for involute teeth of cylindrical gears for general engineering and heavy engineering*, Beuth Verlag, 1986.

[4] DIN 3960:1987-03, *Definitions, parameters and equations for involute cylindrical gears and gear pairs*, Beuth Verlag, 1987.

[5] DIN 3990 series, *Calculation of load capacity of cylindrical gears*, Beuth Verlag, 1987.

[6] Bausch, T. et all, *Moderne Zahnradfertigung*, 2. Auflage, Expert Verlag, 1994.

[7] Roth, K., *Zahnradtechnik Band II*, Springer Verlag, 1989.

[8] DIN 58400:1984-06, *Basic rack for involute teeth of cylindrical gears for fine mechanics*, Beuth Verlag, 1984.

[9] DIN 58425 series, *Gears with round flanks for fine mechanics*, Beuth Verlag, 1980.

[10] Menz, W., et all, *Mikrosystemtechnik für Ingenieure*, 3. Auflage, Wiley-VCH, 2005.

[11] Houbolt, J.C., A recurrence matrix solution for the dynamic response of elastic aircraft, *J. Aeron. Sci. Vol. 17*, pp. 540-550, 1950.

[12] Albers, A.; Enkler, H.-G.; Leslabay, P., On the Simulation of Molded Micro Components and Systems, *MicroSystem Technologies*, Springer, 2008.

[13] Lagarias, J.C., et all, Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions, *SIAM Journal of Optimization*, Vol. 9 Number 1, pp. 112-147, 1998.

[14] Sacks, J., et all, Design and Analysis of Computer Experiments, *Statistical Science*, Vol. 4 Number 4, 1989.

[15] Box, G. A. F., W. G. Hunter, and J. S. Hunter, *Statistics for Experimenters*, Wiley, 1978.

[16] Montgomery, D. C., *Design and Analysis of Experiments*, Wiley, 2001.

[17] Plackett, R.L., Burman, J.P., The Design of Optimal Multifactorial Experiments, *Biometrika Vol. 33*, 1946.

[18] Ryser, H. J., *Combinatorial Mathematics*, John Wiley & Sons, 1963.

[19] Box, G. E. P., Draper, N.R., *Empirical Model Building and Response Surfaces*, John Wiley & Sons, New York, 1987.

[20] Jurecka, F., Robust design optimization based on metamodeling techniques, *Dissertation,* TU München, 2007.

[21] Box, M.J., Draper, N.R., Factorial designs, the |xtx| criterion, and some related matters, *Technometrics*, Vol. 13, pp. 731-742, 1971.

[22] Atkinson, A. C., and A. N. Donev, *Optimum Experimental Designs*, Oxford University Press, 1992.

[23] Doltsinis, I., Kang, Z., Robust design of structures using optimization methods, *Comput. Methods Appl. Mech. Engrg. Vol 193*, pp. 2221–2237, 2004.

[24] Unger, J., Roos, D., Investigation and benchmark of algorithms for reliability analysis, *Weimarer Optimierungs- und Stochastiktage 1.0*, 2004.

[25] Spall, J. C., *Introduction to Stochastic search and Optimization*, Wiley, 2003.

[26] Jones, D.R., Schonlau, M., Welch, W.J., Efficient global optimization of expensive black-box functions, *Journal of Global Optimization*, Vol. 13-4, pp455-492, 1998.