Generic effects of deviations from test design orthogonality on test power and regression modelling of Central-Composite Designs

Marco Arndt [a,*], Philipp Mell [a] and Martin Dazer [a]
[a] Institute of Machine Components (IMA), University of Stuttgart, Stuttgart, Germany
[*] marco.arndt@ima.uni-stuttgart.de

Abstract: In the context of design of experiments (DoE), for many cases the quantitative dependency of a nonlinear target parameter on a few factors is to be determined for the related parameter prediction. For these cases, from the group of response surface designs, test plans are used following the structure of Central-Composite Design (CCD). Their leverage value α predefines the relative directional distance beyond the center run for the star runs, which yield the required information for a quadratic model while still being highly efficient. The individual value determination of α as well as the specific arrangement of the test runs in the design matrix follow a generic mathematical approach to match required DoE properties. Here the most essential respective property is orthogonality. It is sufficiently required in order to consider uncorrelated and independent coefficients separately and to establish regression models, guaranteeing the narrowest possible confidence intervals for parameter prediction. It can be complied and determined analytically based on α and the relative amount of individual run types. However, in the current state of research it remains unclear to what extent renewed adjustments in the amount and arrangement of test runs or further deviations from orthogonality have a practicable effect on design efficiency, test power and precision in regression coefficient estimation. This paper presents a parameter study regarding generic orthogonality deviations in CCDs. For this purpose, various orthogonality deviations are mathematically identified, quantified and performed. Subsequently, potentials and deviations in the effect detection are calculated. Finally, tendencies and first recommendations for design adaptations are presented under consideration of parameter prediction and design efficiency. This includes the categorical exclusion of possible orthogonality deviations as well as the quantification of tolerance limits for minor orthogonality deviations.

1. INTRODUCTION

Design of Experiments (DoE) as a method for the empirical description of a target variable or a system behavior as an investigation objective is considered in many aspects as the most efficient method to determine the variable characteristics of an object or system holistically by descriptive influencing factors. *Fischer* was the first to define this approach fundamentally and to introduce the transfer of a variety of experimental observations into matrix notation [1]. In contrast to e.g., *One-Factor-at-a-Time* approach (OFAT), the information gain to be obtained is thus made more comprehensive and maximized with a putative minimum number of test runs. This avoids an unstructured and disordered implementation of experiments, whose integration into a sequential investigation procedure, potentially based on just previously gained experimental information in each run, is avoided. The indicated again has a strong positive effect on efficiency, both in monetary, temporal – in particular for lifetime testing – and scientific terms [2]. In contrast to a possible identification of a best investigation result with respect to a sequence-dependent target variable, a presumed global optimum can thus be identified. Within this process, a simultaneous and systematic change of several factors and the multiple utilization of experiments is used to generate additional benefit to system understanding in the evaluation of the results and interactions.
For the implementation of this systematic approach in experimental investigation, by this time various experimental designs are available, which meet different requirements in the respective application. Initially, factorial designs were introduced as a basis, and are further developed by *Box* and *Behnken*

(Box-Behnken experimental design), *Plackett-Burmann* (screening experimental designs) and *Box* and *Wilson* (*Central-Composite Designs, CCDs*) [3-5]. The latter in particular favor system response variable-performance tuning or specific application and extension of DoE plans in lifetime modeling over a multidimensional parameter space as established in reliability engineering [6,7]. As predestined use cases for this type of experimental design, this enables exemplarily the investigation of chemical reaction-behaviors with regard to their optima or likewise the consideration of reliability models with curved and low-order system model functions in combination with quantifiably narrow confidence intervals. *Response surface designs* (RSM), which include CCDs, namely allow optimization of the determined global optima beyond a linear relationship with respect to the k factors to be investigated [2].

However, in order to ensure this, certain boundary conditions must be considered. The mathematical description for test designs of this type delineates the demand to meet particular requirements according to their matrix notations: *balance* and *orthogonality*. Although both balance and orthogonality are characterized as equally considerable in mathematical terms, orthogonality may analytically determine the balance. Thus, within an orthogonal test design matrix, performed test runs are arithmetically independent, meaning derived model coefficients are uncorrelated and separately/multiply assignable to individually observed test results. Thereby each of these assignments regard the appropriate factor combinations [2]. In addition to uncertainties due to noise parameters in experimenting, which are usually not to be investigated factorially in the test designs, therefore a widening of the confidence intervals of derivable system models is minimized [8]. However, in order to ultimately use rating parameters evaluating the orthogonality of a test plan, in case there are deviations from the same, a number of criteria are provided by now [9]. These control procedures are capable of quantifying the type and scale of the deviations of orthogonality. Nevertheless, they do not provide practical information about the outcome model's power (*test power*) generated supposedly by non-orthogonal experimental designs. While one might encounter anomalies such as correlations via the construction of the fitted system function estimates and a residuals analysis, it does not give practical significance for actually identifying effects. A feasible evaluation of the practical impact of these deviations on detection performance of significant system effects remains thus and to date undetermined.

Using and applying commonly deployed CCDs, this paper thus provides a primal overview of generally potential and ordinarily occurring deviations from orthogonality. To this end, relevant concepts related to CCDs, regression models, orthogonality in experimental designs, and associated control criteria are initially highlighted. Eventually generic effects are recorded and exemplified for one study point in a second-order model. Other than just consulting control criteria, with the presentation of possible orthogonality deviations, the consideration of the design efficiency, the coefficient quality in regression analysis as well as the test power are evaluated. Finally, a first recommendation for action is formulated based on identified tendencies.

2. CENTRAL COMPOSITE DESIGNS (CCDs) AND BOUNDARY CONDITIONS

In this section, first an overview of properties and boundary conditions associated with CCDs, orthogonality and the use of regression analysis for statistical evaluation is provided. This overview also includes a selection of applicable control criteria available to evaluate CCD orthogonality. Additionally, well-documented standard literature such as [2,7,8] and a more detailed, elementary overview of relationships between factorial experimental designs, orthogonality, and control criteria according to [10] provide a quick-access information framework beyond the presented work, so they will not necessarily be discussed in detail here.

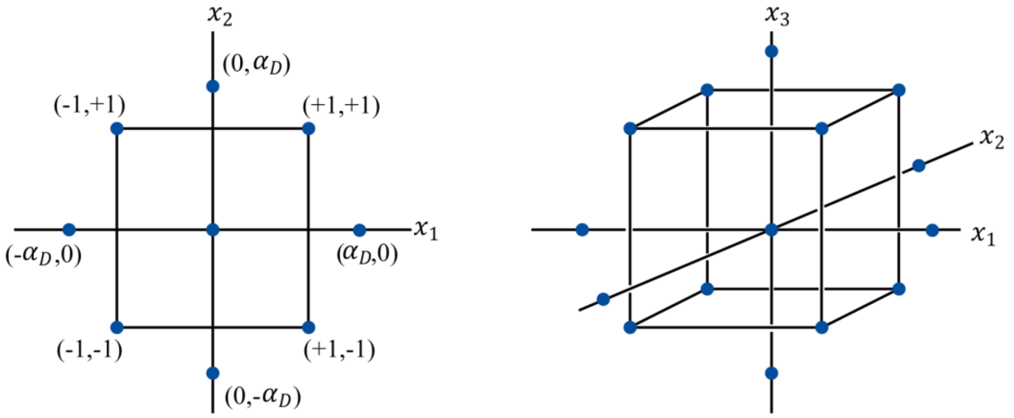
2.1. CCD Characteristics

The CCD is based on a 2^k full- or 2^{k-p} fractional-factorial experimental design with *r* replications per run, where $p \geq 0$ represents an extra quantity of parameters added to the model as influencing factors. The design for n_F *factorial* or *cube points* is expanded to include $2k$ *star points* in a symmetrical arrangement around the factor axes and additional n_C *central points*, cf. *Figure 1*. As a part of RSMs, this makes the CCD a viable, diversely utilized experimental design in two aspects, since it can be sequentially extended on a factorial experimental design, and nonlinearity can additionally be identified and mapped by 5 axial runs (star points) [2]. Accordingly, an original derivation is based on the central point to discover curvature of the model. The additional star points per factor provide information about the pure quadratic function of the detected curvature. The factorial points are. Moreover, the factorial runs are the only ones providing information about factor interactions [11]. Containing a random error term $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, for k factors the system's output can be determined in a second-order model by

$$y = \beta_o + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k-1} \sum_{i<j=1}^{k} \beta_{ij} x_i x_j + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \varepsilon, \quad (1)$$

which is based on the factors x_i, the output's mean β_o and β_i coefficients derived from the observed parameter influence.

Figure 1: CCDs with $k = 2$ (left) and $k = 3$ (right), cf. [2]



The axial distance of the star points from the central points is defined by the distance- or leverage-value[*] α_D, whereby this value generally varies from 1.0 to $\sqrt{k}$ depending on a face-centered or spherical positioning of star points on the factor axis [11]. However, both α_D and n_C needs be defined depending on various design characteristics: *orthogonality*, and as needed exemplarily combined with *rotatability* and *sphericity*. Here, the parameter space to be examined is defined decisively in its scope. To quantify prediction goodness of the model to be derived, this is done by assessing the consistency and robustness of the variance of all examination points x_i through k factors [2,11]. As *Box and Hunter* proposed in [12], the prediction variance of the model $\hat{y}$ to be fitted based on y_i observations will be stable in all study points on spheres of a second-order design if they have the same distance from the central point. Consequently, for this reason a test design with factorial points and star points becomes a *rotatable* central composite around the centre point by:

$$\alpha_D = (n_F)^{1/4}, \quad (2)$$

regardless of the value for n_C. The amount of center runs n_c, however, determines a reasonable stability of prediction variance within the entire design region [13], respectively the *orthogonality*, which will be discussed in the next Section by terms of model building. Since the sets of test runs have to be integers by nature, there are cases for some combinations of test run types where orthogonality and rotatability cannot be achieved exactly at the same time [13,14]. Exemplarily for *orthogonal* CCDs, the test points may depend on adjusted runs by individual numbers of replications r for

- factorial runs $n_f = r_f \cdot n_F$, forming the factorial portion with $x_i = -1, +1$ for $i = 1, \dots, k$;
- central runs $n_c = r_c \cdot n_C$ with $x_i = 0$ for $i = 1, \dots, k$; and
- axial runs (or star runs) $n_s = r_s \cdot 2k$ of the form $(0, \dots, x_i, \dots 0)$ with $x_i = -\alpha_D, \alpha_D$ for $i = 1, \dots, k$;

[*] In standard literature (e.g., [2,8,11]) usually called α – for reasons of clear differentiation in variable naming to *significance level* α: here adapted to "α_D".

Here the axial distance of star runs follows according to [2,13,15] as:

$$\alpha_D = \left[\frac{\sqrt{n_F\left(n_F + 2k \cdot \frac{n_S}{n_f} + \frac{n_C}{n_f}\right)} - n_F}{2 \cdot \frac{n_S}{n_f}} \right]^{1/2} . \tag{3}$$

Beside even more, also *spherical* CCDs ($\alpha_D = \sqrt{k}$), *face-centred* CCDs ($\alpha_D = 1$) and Box-Behnken designs (e.g., $\alpha_D = \sqrt{2}$ for $k = 3$) are commonly applied satisfying different investigation intents.
Since positioning of the star runs / axial points, the position of the (± 1)-coded factorial points is decisive, the influence of significance levels on effect detection is also examined in the following.

2.2. Test Power and Statistical Interference

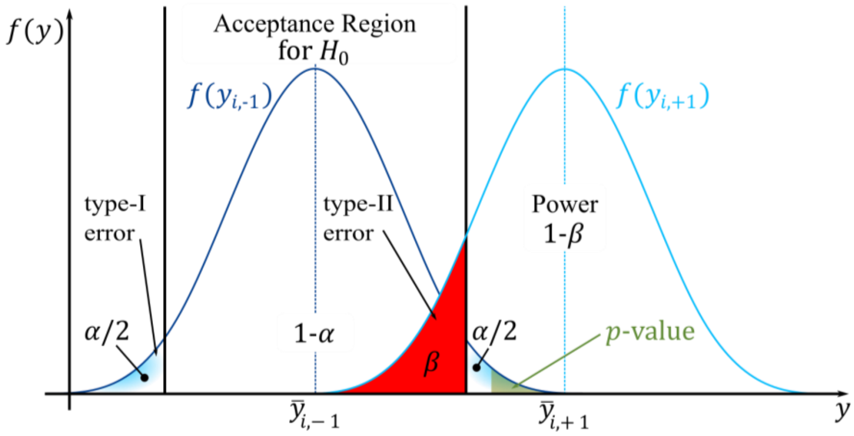
With the conversion of experimentally observed effects by factor level change into an empirical model, it needs to be guaranteed that these are not subject to an effect, which is exclusively due to a (normally distributed) error of randomness. For this and the intention of this paper, a brief overview of the consideration of significance levels is given here. Therefore, hypothesis testing is used to detect the significance of an effect, which is derived by the putatively observed difference of mean values $\bar{y}$ through the stochastic system response variable $y \sim \mathcal{N}(\bar{y}, \sigma^2)$, supposedly existing in simultaneous coexistence with a random variance σ^2 for a respective combination of factor levels changes [2]. To this end, two complementary hypotheses (null hypothesis H_0 and alternative hypothesis H_1) are formed to evaluate an effect as existing or not existing by difference of the means after factor level ($x_i = -1, +1$) adjustments:

$$H_0: \bar{y}_{i,-1} = \bar{y}_{i,+1}; \tag{4}$$

$$H_1: \bar{y}_{i,-1} \neq \bar{y}_{i,+1}. \tag{5}$$

With this approach, two wrong decisions may be likely due to chance: H_0 *is* rejected although it is *true* (*type-I-error*); or H_0 is *not* rejected although it is *false* (*type-II-error*) [2,8], see *Figure 2*.

Figure 2: Statistical Interference and Test Power, cf. [2,8]



The probabilities of these errors result in α for the type-I and in β for the type-II error, where the power of test is defined as the probability to identify an existing influence on the effect correctly by

$$\text{Power} = 1 - \beta. \tag{6}$$

The level of significance complementary to the error probability can thus be specified via α for the effect evaluation, usually with $\alpha = 0.05$ or less [2,15]. However, in order to obtain an information on the significance of the result before the test is performed, the p-value and t-statistic is generally used. The p-value is the probability that the test statistic will adopt the value that is at least as extreme as the observed value of the statistic in case that H_0 according to *Equation (4)* is true [2]. Accordingly, the p-value is the probability of being wrong when H_0 is rejected [9].
In addition, if any number of parameters are relevant to the application, Analysis of Variance (ANOVA) and the $F_{k,n-k-1}$-distribution can be used, where $p = 1 - P(F > F_{k,n-k-1})$ is derived from [2,8,11].

A detailed description of this statistical analysis is outside the focus of this paper and well documented in [2,8,11,17]. Instead, for background on orthogonality deviations, an overview of the model building through regression analysis and aspects of orthogonality therein are described below.

2.3. Model Fitting of the Second-Order Response Surface

In order to evaluate effects on the system response y from the experimental observations as significant and to finally transfer them into a fit for the RSM, the regression coefficients β_i need to be estimated to $\hat{\beta}_i$ from the effect observations. Among others, the Maximum-Likelihood Estimation (MLE) and the Method of Least Squares are suitable for this purpose [2]. Since, regardless of the shape of the surface fitted by the regression model, all regression models are linear as long as the regression parameters are linear, and therefore second-order models with interactions can also be described through this [17]. Conclusively, this also includes RSM with curvatures.

Illustrating the above, exemplarily the quadratic function of an RSM with $k = 2$ factors and interactions

$$y = \beta_o + \beta_1 \mathrm{x}_1 + \beta_2 \mathrm{x}_2 + \beta_{11} \mathrm{x}_1^2 + \beta_{22} \mathrm{x}_2^2 + \beta_{12} \mathrm{x}_1 \mathrm{x}_2 + \varepsilon \tag{7}$$

may hold substitutes as stated in [11,17] for

$$\mathrm{x}_3 = \mathrm{x}_1^2, \mathrm{x}_4 = \mathrm{x}_2^2, \mathrm{x}_5 = \mathrm{x}_1 \mathrm{x}_2 \text{ and } \beta_3 = \beta_{11}, \beta_4 = \beta_{22}, \beta_5 = \beta_{12}.$$

Dealing with $n > k$ observations, then exemplarily the least-squares estimators $\hat{\beta}_i$ can be derived from *multiple linear regression analysis* via matrix notation of the system response through

$$\mathrm{y} = \mathrm{X}\upbeta + \upvarepsilon, \tag{8}$$

where

$$\mathrm{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathrm{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \upbeta = \begin{bmatrix} \beta_o \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \text{ and } \upvarepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

Concluding, the vector of least-squares estimators $\hat{\upbeta}$ is tried to be determined, so that the least-squares function

$$S(\upbeta) = \sum_{i=1}^{n} \varepsilon_i^2 = \upvarepsilon'\upvarepsilon = (\mathrm{y} - \mathrm{X}\upbeta)'(\mathrm{y} - \mathrm{X}\upbeta) \tag{9}$$

with *Equation (8)* rearranged to ε is minimized, corresponding to

$$\left.\frac{\partial S}{\partial \upbeta}\right|_{\hat{\upbeta}} = -2\mathrm{X}'\mathrm{y} + 2\mathrm{X}'\mathrm{X}\hat{\upbeta} = 0. \tag{10}$$

Here the least-squares estimator of β is derived by

$$\hat{\upbeta} = (\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'\mathrm{y} \tag{11}$$

and supplemented in the fitted regression model with $\mathrm{x}' = [1, x_1, x_2, \ldots, x_k]$

$$\hat{y} = \mathrm{x}'\hat{\upbeta} = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \text{ ,for } i \neq j \text{ and } j = 1, \ldots, k. \tag{12}$$

Determining $\mathrm{X}'\mathrm{X}$ as a $(k + 1 \times k + 1)$ matrix, note that here diagonal elements are sums of squares of column elements of X, as off-diagonal elements are respective cross products. Corresponding the vector of fitted response values $\hat{y}_i$ to the observed system responses y_i the $(n \times n)$ hat-matrix $\mathrm{H} = \mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'$ mapping the vector of observed values into the vector of fitted values is obtained for

$$\hat{\mathrm{y}} = \mathrm{X}\hat{\upbeta} = \mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'\mathrm{y} = \mathrm{H}\mathrm{y} \tag{13}$$

Eventually, for the residual $e_i = y_i - \hat{y}_i$, the n differences between the observations and the fit follows (cf. [17])

$$\mathrm{e} = \mathrm{y} - \hat{\mathrm{y}} = \mathrm{y} - \mathrm{X}\hat{\upbeta}. \tag{14}$$

Analogously, the fit to a second-order model according to *Equation (1)* is given by [2,11] as follows

$$\hat{y} = \hat{\beta}_0 + \mathrm{x}'\hat{\upbeta} + \mathrm{x}'\mathrm{B}\mathrm{x}, \tag{15}$$

where with consideration of re-substitution

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix},\ \hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} \text{ and } \mathrm{B} = \begin{bmatrix} \hat{\beta}_{11} & \hat{\beta}_{12}/2 & \cdots & \hat{\beta}_{1k}/2 \\ & \hat{\beta}_{22} & \cdots & \hat{\beta}_{2k}/2 \\ & & \ddots & \vdots \\ \text{sym.} & & & \hat{\beta}_{kk} \end{bmatrix}$$

In this shape $\hat{\beta}$ forms a $(k \times 1)$ first order regression coefficients vector, B a symmetric $(k \times k)$ matrix, containing quadratic coefficients $(\hat{\beta}_{ii})$ on the main diagonal and mixed coefficients $\hat{\beta}_{ij}, i \neq j$ as else.
Following the idea of hypothesis testing and ANOVA from *Section 2.2.* considering p parameters, the residual mean square with *n-p* degrees of freedom results through *Equation (14)* in an unbiased estimator of the unknown model-dependent variance σ^2 as follows [17]:

$$\sigma^2 \approx \hat{\sigma}^2 = \frac{\mathrm{y}'\mathrm{y} - \hat{\beta}'\mathrm{X}'y}{n-p} = \frac{\mathrm{e}'\mathrm{e}}{n-p}. \tag{16}$$

Considering the matrix $\mathrm{M} = (\mathrm{X}'\mathrm{X})$, the property of the variance of $\hat{\beta}$ can be determined [11,17] by the covariance matrix

$$Var(\hat{\beta}) = \sigma^2 \mathrm{M}^{-1}, \tag{17}$$

resulting in the variance of $\hat{\beta}_j$ through $\sigma^2 M_{jj}^{-1}$ and the covariance between $\hat{\beta}_i$ and $\hat{\beta}_j$ in $\sigma^2 M_{ij}^{-1}$.
If last based on M also the confidence interval around an estimated value of the system response $\hat{y}(\mathrm{x})$ in the overall analysis process to RSM is to be described as stated in [11,17,18], this is called prediction variance and describes how well ones predicts with the model:

$$PV(\mathrm{x}) = Var(\hat{y}(\mathrm{x})) = \sigma^2 \mathrm{x}^{(m)\prime}\, \mathrm{M}^{-1} \mathrm{x}^{(m)\prime}\,. \tag{18}$$

Here, (m) in $\mathrm{x}^{(m)}$ reflects the obtained model, where exemplarily a first-order model defines $\mathrm{x}^{(1)\prime} = (1, x_1, \ldots, x_k)$. Scaling the $PV(\mathrm{x})$ with a *per observation basis* by N/σ^2, the scaled prediction variance is derived by

$$SPV(\mathrm{x}) = \frac{N\, Var(\hat{y}(\mathrm{x}))}{\sigma^2} = N\mathrm{x}^{(m)\prime}\, \mathrm{M}^{-1} \mathrm{x}^{(m)\prime} = 1 + \sum_{i=1}^{k} x_i^2. \tag{19}$$

The scaled prediction variance is constant on spheres, therefore designs with same $N\, Var(\hat{y}(\mathrm{x}))/\sigma^2$ are rotatable.
From this understanding of the model construction by (multiple) linear regression shown within this *Section*, the following relation should be clear: if off-diagonal entries of $\mathrm{X}'\mathrm{X}$ go towards zero and the entries on the main diagonal are as large as possible, the main diagonal of M^{-1} and thus an explained part of the model variance $Var(\hat{\beta})/\sigma^2$ as stated in *Equation (17)* is minimized [17].
Thus, the test plan property orthogonality can be outlined in the following.

2.4. CCD Orthogonality

As delineated in the previous *Section 2.3.* and from a model fitting point of view, an orthogonal test design minimizes $Var(\hat{\beta})$ and therefore enhances model prediction accuracy [11,17]. This is based on linear independency of input parameters with $x_i = \pm 1$ for all levels in the factorial test design and $i = 1, \ldots, k$. Therefore, an *orthogonal* test design features a matrix $\mathrm{M} = (\mathrm{X}'\mathrm{X})$ which consists of a diagonal matrix, where the columns of X are apparently also mutually orthogonal.
The CCD becomes *orthogonal* as in addition the ratio between the amount of design points and leverage value $(x_i = \pm\alpha_D)$ is satisfied according to *Equation (3)*, cf. ***Figure 1*** [17]. For this, in any case, the transformations by *Equation (7)* must be considered – at least according to this procedure, since with the quadratic terms $\mathrm{X}'\mathrm{X}$ is otherwise not composable. Because of this reason and with just this particularity about leverage values α_D of CCDs, mainly the deviations of orthogonality and thus sphericity by changing these values will be discussed in the context of this paper.

Thereby it is to be understood, that in second-order models both orthogonality, despite still elemental, and the coefficient estimation take subordinate roles in relation to the fitted model, as mainly dominated by the importance of the scaled prediction variance $N\,Var(\hat{y}(\mathrm{x}))/\sigma^2$, cf. *Equation (19)*, and derived design characteristics (α_D, cf. *Section 2.1.*). According to the analysis of the response surface design procedure [11], the goodness of fit based on the estimate to the model response $\hat{y}(x)$ here is more important than the exact determination of the parameter ratios in the model. However, since both as consequences rely on orthogonality described here, control procedures for evaluation of the same are important as shown below. This motivates the first step in assessing generic effects through orthogonality deviations, as presented in this paper.

2.5. Controlling Orthogonality

In order to use orthogonality as a criterion derived from the best test design, it is useful to measure and control this size quantitatively and qualitatively. Several criteria can be found in literature [2,8,9]. Besides a simple evaluation of the design matrix X and X′X as described in *Section 2.4.*, respectively the parameter matrix and matrix column orthogonality determined as existing or not existing, the following methods are to be mentioned: *correlation matrix* $Corr(\mathrm{X})$ [17], A*-Optimality* [18], D*-Optimality* [12,19], G*-Optimality* [12]. More specifically, the correlation matrix and A-optimality evaluate the model estimates over correlations, D-optimality evaluates the estimates over variances and covariances, and G-optimality evaluates the scaled prediction variance of the estimated model, each with associated criteria. For $R = Corr(\mathrm{X}) = \mathrm{X}'\mathrm{X} = [\rho_{ij}]$,

$$A_2 = \sum_{i<j} \rho_{ij}^2 \tag{20}$$

measures as a criterium for A*-Optimality* the non-orthogonality while $A_2 = 0$ when orthogonality is present. Knowing well that beyond that other criteria are available, not all of them are applicable on the basis of the star point values in CCDs. Nevertheless, a measure of how much a test design deviates from orthogonality can be determined with the aid of these criteria.

3. ORTHOGONALITY DEVIATIONS OF CCDS

With the fundamentals shown in *Sections 2.3. - 2.5.* for the mathematical construct of orthogonality in test designs, as well as the outlined advantages and control criteria for it, the determining principles are thus explained. Now, with the help of an evaluation for the statistical significance of recognized coefficients and effects in the estimated model according to *Section 2.2.*, a far more palpable evaluation variable for deviations in orthogonality shall be introduced.

Building on the work of [10], which already presents first basic results for full factorial experimental designs, this initial foundation is extended step by step to RSMs using the example of CCDs in the context of this paper. For this purpose, an approach with relevant and realistic orthogonality deviations in star runs as well as the implementation of a second-order design model will be described. Therefore, the model described below is used to consider the probability of actually detecting an existing effect in coefficient determination correctly, in other words, the power as a performance measure.

3.1. Default System Model Setup

To illustrate the effects of orthogonality deviations in the use case of CCDs within RSMs, a two-dimensional model is implemented in the presented study: a CCD with a quadratic response surface via two factors x_1 and x_2. Consequently, this results in a second-order model as stated in *Equation (7)*.

In order to create a comparable database and for simplicity, a given *default system model* is attributed with an equivalent pre-specified value for the regression coefficients for each factor, cf. ***Figure 4***. For both the prediction (or model) error ε and the estimation of the coefficient values (effects) $\beta_{1,2}$ a normally distributed error is assumed, which can be defined by a given standard deviation $\sigma = 0.1$, cf. *Equation (9)* and *Equation (16)*. In the context of this paper, this definition is initially considered as a first initial setting point, which is subject to variation in further investigations. Consequently, this results in the model composition according to Table 1. Eventually, this model is investigated in the simulation described in Section 3.3. (also see ***Figure 4***).

Table 1: Coefficients and Standard Deviation for Model Setup

Model order type	$\boldsymbol{\beta_o}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_{11}}$	$\boldsymbol{\beta_{22}}$	$\boldsymbol{\beta_{12}}$	$\boldsymbol{\sigma}$
quadratic	30	0.1	0.1	0.1	0.1	0.1	0.1

3.2. Orthogonality Deviations

For a deliberate implementation of deviations from orthogonality for an experimental design for the investigation of generic effects, variants for the same are considered, which can be derived from reasonable and practicable contexts, cf. also [10]. Thus, on the one hand, these assumptions are well realizable by simulation, but they also have a direct relation to practical problems.
With respect to CCDs studied here, these contexts are first outlined in the course of this work and understood to be generally plausible in real-world test scenarios based on their foundation, as they may generally arise for reasons of precision and/or time constraints, as well as monetary and/or physical capacity limitations in the execution of the test design. These may include scenarios where a single test run, particularly a critical factor level combination at the star point,

- is not physically or mechanically feasible, due to e.g., time and effort or ambient testing conditions;
- is just realisable at a decreased level of the original the axial run definition, therefore $\pm\alpha_D$ shrinks;
- or a factor stage combination is encountered where a pure axial run can no longer be performed due to displacements, scatter or dependencies with another factor (e.g., factor correlation).

Accordingly, the following deviations from orthogonality for generic effects are considered in this paper, while they are intended as first steps for further studies:

- One star run (originally defined by $0, \pm\alpha_D$) is not set to the predefined axial distance: $|\alpha_D|$ might accidentally or deliberately not be set correctly (cf. *Figure 3: I*);
- A star point is no longer performed as a pure axial run and obtains a second value unequal to 0 with respect to the second factor (cf. *Figure 3: II*);
- The star runs through α_D undergo a given scattering in the setting values, e.g., based on systematic variance of the operating hardware capacities in the control, regulation and sensor technology (cf. *Figure 3: III*);
- One star run through α_D is entirely omitted in the test design (cf. *Figure 3: IV*).

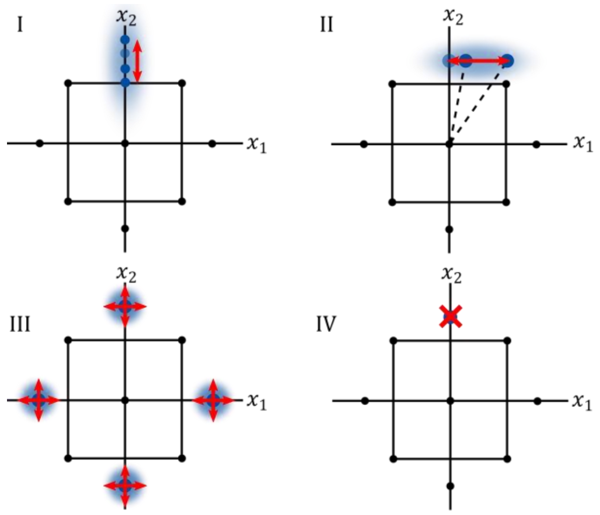
3.3. Study Approach

For the investigation of generic effects of orthogonality deviations summarized in *Section 3.2.*, a numerical simulation model is chosen for this work with the implementation of a Monte Carlo (MC) performing 100.000 repetitions, cf. ***Figure 4***. According to Section *3.1.*, the second order regression model is used to capture generic effects and impacts. Effects and coefficients considered in this model are taken over with normalized factor change around the average of the system response as defined in *Table 1* with an originally given system model. Finally, deviations of the presented cases *I – IV* as stated in ***Figure 3*** are implemented in such a way that for the design matrix X contains

- a star run taking $\alpha_D = 1$ (case *I-1*), $\alpha_D = 0.8 \cdot \alpha_D$ (case *I-2*) or $\alpha_D = 1.2 \cdot \alpha_D$ (case *I-3);*
- the combined factor at the star run α_D, adopting $x_1 = 0.1$ (case *II-1*) or $x_1 = 1$ (case *II-2*) causing correlation;
- star runs α_D equally scattering each over their original value within a 10 % scatter region (corresponding to case *III*); and
- an omitted star run (corresponding to case *IV*).

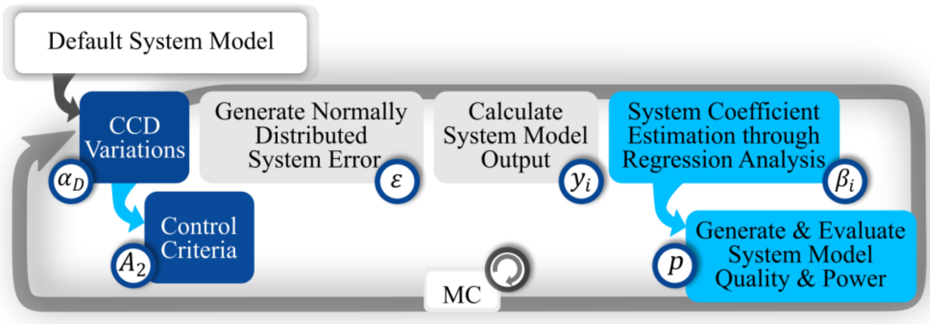
To map a statistical basis of the regression error, each experimental point is replicated three times (if present). For each deviation, the model approximation is then formed using the regression analysis presented above in Section 2.3.

Figure 3: Implementation of orthogonality deviations regarding axial runs (star points): *I)* varying over axial distance in one case; *II)* varying through factor correlation in one case; *III)* randomly scattered over all cases; *IV)* omitted in one case.



The resulting means and coefficients are finally evaluated with respect to their power at a significance level of $\alpha = 0.05$. The subsequent comparison ultimately provides information about supposed tendencies and optimization possibilities. The simulation process is shown in ***Figure 4***.

Figure 4: Simulation Study Approach: Flowchart



3.4. Results

In order to evaluate generic effects of the orthogonality deviations in CCDs as described in Section 3.2 and along the simulation approach from Section 3.3., in a first attempt the following measures are determined and compared for the estimated models: First, the A_2-criterion is utilized to measure non-orthogonality in a dimensionless way and stated in *Table 2*. Second, also the power of the estimated model coefficients is recorded in *Table 2* measuring the chance of correctly capturing them. Last, the percentage deviations of the coefficient estimates are listed in *Table 3* to finally show the extent of deviation of the estimated model from the default model depending on the overlaid errors and randomness.
According to the results in *Table 2*, the estimated coefficient β_o as the mean value of the system model is always found with a probability equivalent to certainty in all cases for the second-order model.
With regard to the coefficients β_1 and β_2 as linear fractions of the system model, an axial shift of the star point relative to x_2 has a slightly negative effect on the power of the factor-associated coefficient $\beta_2 = 67.2\ \%$ compared with β_1 (68.9 %), if the star run is face-centred (*I-1*). Compared with this, a

deviation in not fully reached distance above the face (*I-2*) results in an improved power for β_2 with similar power for β_1, an oversized α_D (*I-3*) reduces this power again. Also, a transversal shift with direction x_1 to $(0.1 \text{ and } 1.0, \alpha_D)$ reduces the power of the factor-associated coefficient β_2 compared to the orthogonal case (*II-1*), but shows that β_1 has slightly improved power with larger shift (*II-2*). If the settings of all star points are given a 10 % uniformly distributed scatter (*III*), the power of the linear coefficients is reduced by up to 2.4 percentage points. Last, according to the simulation results gathered here, omitting the star point has a comparable effect on the power of the linear regression coefficient β_1 as within the cases described before but causes a power drop for β_2 to 61.0 %.
With regard to the quadratic coefficients it follows, that the coefficient β_{11} is recorded with a power loss of 3 to 4 percentage points, cf. *Table 2* for $\beta_{11} = 72.1$ % in the orthogonal case against all other cases. Note that the factor star point of this quadratic effect is not manipulated by simulation, except in case (*III*). Orthogonality deviations through the single star point of x_2 obviously improve the power of β_{22} for a star point value between face-centered and alpha (73.2 %) and holds within random scatter (72.0 %).

Table 2: Results for Orthogonality Deviations through Star Runs in CCDs for Second-Order Models: A_2 and Coefficient Power

Case		Star Run	A_2	Power [%]					
				β_o	β_1	β_2	β_{11}	β_{22}	β_{12}
	Orth.	$(0, \alpha_D)$	0	100.0	72.1	71.9	72.1	72.3	62.1
	I-1	$(0,1)$	0.04	100.0	68.9	67.2	69.1	65.1	58.3
	I-2	$(0, 0.8 \cdot \alpha_D)$	0.04	100.0	69.1	70.0	69.1	73.2	57.9
	I-3	$(0,1.2 \cdot \alpha_D)$	0.07	100.0	69.0	66.2	69.1	63.5	58.0
	II-1	$(0.1, \alpha_D)$	0.01	100.0	68.9	69.1	68.8	68.7	58.3
	II-2	$(1, \alpha_D)$	0.35	100.0	69.5	66.7	68.1	66.7	60.8
	III	$(0, \alpha_D) \pm 10$ %	0.01	100.0	68.7	69.7	68.8	72.0	58.1
	IV	(NaN,NaN)	NaN	100.0	69.1	61.0	68.1	60.3	57.8

Table 3: Results for Orthogonality Deviations through Star Runs in CCDs for Second-Order Models: Coefficient Estimation Error, grey values indicate relative errors greater than 5 %

Case		Star Run	Coefficient Estimation Amount Error [%]					
			β_o	β_1	β_2	β_{11}	β_{22}	β_{12}
	Orth.	$(0, \alpha_D)$	0.00	0.74	-0.47	0.77	0.34	0.33
	I-1	$(0,1)$	-0.03	-2.88	1.55	-3.24	1.24	-8.60
	I-2	$(0, 0.8 \cdot \alpha_D)$	0.00	0.23	4.03	2.20	2.19	6.19
	I-3	$(0,1.2 \cdot \alpha_D)$	0.01	-2.97	0.05	-3.19	4.63	4.98
	II-1	$(0.1, \alpha_D)$	0.00	-0.09	5.50	5.05	1.93	1.32
	II-2	$(1, \alpha_D)$	0.00	-1.01	0.82	-0.46	7.71	3.98
	III	$(0, \alpha_D) \pm 10$ %	0.00	-1.53	2.77	-2.95	1.61	-1.65
	IV	(NaN,NaN)	-0.01	3.30	-0.89	-8.69	-0.20	-5.75

The interaction effect (β_{12}), assuming only 62.1 % power in the orthogonal case while considering the global scattering system error in the simulation, deteriorates in its power in every case, but least in the case of the strongest opportunity of factor correlation (*II-2*), which is also captured most strongly in relative terms by the implemented A_2-criterion (= 0.352). A relatively lower power for β_{12} compared to the previously mentioned coefficients can be explained by the axially oriented experimental setup in the CCD. As expected, the A_2-criterion, which uses the trace of the correlation matrix according to

Equation (20), becomes noticeably larger exclusively by increasing correlation of the star point with a second factor (*II-2*) in comparison to all other cases.

Eventually, for a general assessment of the orthogonality deviations, the relative deviations of the estimated values for the model coefficients are compared with the default model coefficients. *Table 3* shows that the estimated value of the coefficient β_2, which is axially affected by the star point manipulation, deviates strongly (5.50 %) when the star point is slightly correlated with the second factor (*II-1*). The same holds for the squared coefficient β_{11} (5.05 %) as well as when the x_2 star point above (*IV*) is not performed or observed (-8.69 %).

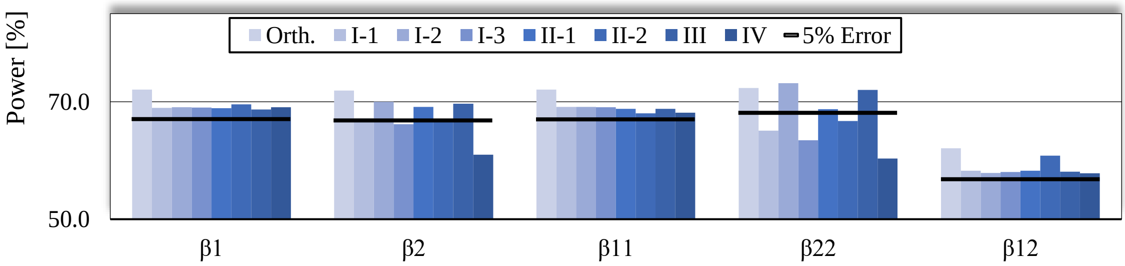
The transversal shift of the considered x_2 star point in case *II-2* also results in a strongly deviating estimated value (7.71 %) for the associated second-order coefficient. The interaction value varies over all deviation manipulations by up to -8.60 % – least, on the other hand, if the star points are subject to scatter (-1.65 %) or if one correlates only slightly with the second factor (1.32 %).

3.5. General Findings

If the results briefly summarized in section 3.4. are to be subsumed, the following observations can first be drawn in general terms from the simulation results that have been carried out and some of which are presented in this paper:

- Generic effects of orthogonality deviations are detected on a factor-specific basis in the corresponding coefficient powers and estimates;
- The factor-dependent power deviation for the coefficients shows a relation to the corresponding orthogonality deviations (cf. *Table 2*): the axial value of the star directly determines the power and estimation quality of its coefficient; increasing correlation with x_1 increases power and estimation quality of the partner factor and vice versa;
- Star run scattering (± 10 %) keeps the power and quality of model estimation within manageable proportions (max power loss: -4.0 %, max estimate error: -2.95 %);
- The A_2-criterion has only limited potential as a suitable generic tool for the dimensionless measurement of correlation effects due to deliberately placed orthogonality deviations within the second-order models, as long as factor correlations are not directly implemented within;
- The simulation setup (default model calculation validated with commercial software Minitab) is appropriate in its structure;
- The results are suitable as a basis for calculating a tradeoff of exemplary 5 % estimation error and 5 % power loss versus deviation in orthogonal experimental design settings, which can be used deliberately or unconsciously for the improvement of test design efficiency (cf. ***Figure 5***).

Figure 5: Power of Model Coefficients Compared to a Tolerance Value of Error: for the Demonstrated Cases, *I-1, I-3, II-2* and *IV* cause a power loss > |5| % for linear and quadratic coefficients of the Manipulated Factor x_2.



4. CONCLUSION

In the context of this paper, the tools for hypothesis testing are introduced in addition to the explanation of the statistical background of an experimental design within DoE, the subsequent model building via regression analysis and the use of control criteria with respect to orthogonality. Besides the meaning of orthogonality described at the beginning for CCDs, possible deviations from the same are presented here, with the intention of identifying effects and deriving generic findings within second-order models.

Eventually, the statistical significance was taken from the fundamentals in order to use the power of model coefficients, i.e. the probability of actually detecting significant effects by means of test design orthogonality deviations, as a measure besides a calculated estimation error for model coefficients. To implement this investigation numerically, an MC is used.

Even though with this first approach presented within this paper a simple model setup with primitive orthogonality deviations through star points was investigated first, results and insights gained from this already serve as a great foundation for findings. This allows insight into opportunities to tune further test designs in terms of effort and cost to cogently compare this against model performance. Accordingly, further attempts are intended to consider more comprehensive orthogonality deviation variations and combinations, to combine them with further power levels, and to investigate model variations with further effect sizes and error terms in more detail.

References

[1] R. A. Fisher, "*The Design of Experiments*", Oliver & Boyd, 1935, Edinburgh.

[2] D. C. Montgomery, "*Design and Analysis of Experiments*", (10th ed.), John Wiley & Sons Inc, 2020, Hoboken NJ.

[3] G. E. P. Box and D. W. Behnken: "*Some New Three Level Designs for the Study of quantitative Variables*", Technometrics, vol. 2, pp. 455-475 (1960).

[4] R. L. Plackett and J. P. Burman: "*The design of optimum multifactorial experiments*", Biometrica, vol. 33, pp. 305-325 (1946).

[5] G. E. P. Box and K. B. Wilson: "*On the experimental attainment of optimum conditions*", Journal of the Royal Statistical Society, vol. 13, pp. 1-45 (1951).

[6] B. Bertsche, "*Reliability in Automotive and Mechanical Engineering: Determination of Component and System Reliability*", (1st ed.), Springer, 2008, Berlin.

[7] C. J. Wu and M. Hamada, "*Experiments: Planning, Analysis, and Parameter Design Optimization*", (1st ed.), John Wiley & Sons Inc, 2000, Hoboken NJ.

[8] W. Kleppmann, "*Versuchsplanung: Produkte und Prozesse optimieren*", (10th ed.), Hanser, 2020, München.

[9] K. Siebertz, D. van Bebber and T, Hochkirchen, "*Statistische Versuchsplanung: Design of Experiments (DoE)*", (2nd ed.), Springer, 2017, Berlin.

[10] P. Mell, M. Arndt and M. Dazer, "*Non-orthogonality in test design: practical relevance of the theoretical concept in terms of regression quality and test plan efficiency*" [Manuscript submitted for publication], PSAM 16, 2022, Honolulu, Hawaii.

[11] R. M. Myers, D. C. Montgomery and C. M. Anderson-Cook, "*Response Surface Methodology: Process and Product Optimization Using Designed Experiments*", (4th ed.), John Wiley & Sons Inc, 2016, Hoboken NJ.

[12] G. E. P. Box and J. S. Hunter: "*Multi-Factor Experimental Designs for Exploring Response Surfaces*", The Annals of Mathematical Statistics, vol. 28, No. 1, pp. 195-241 (1957).

[13] G. E. P. Box and N. R. Draper: "*Response Surfaces, Mixtures, and Ridge Analyses*", (2nd ed.), John Wiley & Sons Inc, 2007, Hoboken NJ.

[14] C. F. J. Wu and M. S. Hamada, "*Experiments: Planning, Analysis, and Optimization*", (2nd ed.), John Wiley & Sons Inc, 2009, Hoboken NJ.

[15] H. Ahn: "*Central Composite Design for the Experiments with Replicate Runs at Factorial and Axial Points*", Industrial Engineering, vol 349, pp. 969–979 (2015).

[16] D. Rasch and D. Schott, "*Mathematical Statistics*", (1st ed.), John Wiley & Sons Inc, 2018, Hoboken NJ.

[17] D. C. Montgomery, E. A. Peck and G. G. Vining, "*Introduction to Linear Regression Analysis*", (5th ed.), John Wiley & Sons Inc, 2012, Hoboken NJ.

[18] A. C. Rencher and G. B. Schaalje, "*Linear Models in Statistics*", (2nd ed.), John Wiley & Sons Inc, 2008, Hoboken NJ.

[19] A. Wald, "*On the Efficient Design of Statistical Investigations*", The Annals of Mathematical Statistics, vol. 14, No. 2, pp.134-140 (1943).

[20] H. Xu, "*An Algorithm for Constructing Orthogonal and Nearly-Orthogonal Arrays with Mixed Levels and Small Runs*", Technometrics, vol. 44, No. 4, pp.356-368 (2002).