Journal of Chromatography B, 910 (2012) 14–21

Contents lists available at SciVerse ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb

Review

Doehlert uniform shell designs and chromatography☆

Pedro Araujo[a,*], Steve Janagap[a,b,c]

[a] National Institute of Nutrition and Seafood Research (NIFES), PO Box 2029 Nordnes, N-5817 Bergen, Norway
[b] Department of Chemistry, University of Bergen, N-5009 Bergen, Norway
[c] Department of Chemistry, College of Arts and Sciences, University of the Philippines Visayas, 5023 Miagao, Iloilo, Philippines

ARTICLE INFO

Article history:
Received 28 January 2012
Accepted 10 May 2012
Available online 9 June 2012

Keywords:
Doehlert uniform shell designs
Uniform shell designs
Doehlert designs
Chromatography

ABSTRACT

The principles of the Doehlert uniform shell designs (aka Doehlert designs) and their importance in the context of chromatography are discussed. The confidence of different models generated by Doehlert designs is studied by means of the uncertainty of the experimental points. The article provides an overview of analytical applications in chromatography with focus on single and coupled techniques and also discusses some reported blunders regarding Doehlert designs.

Contents

1. Introduction

Chemometric experimental design can be defined as the rational process of planning experiments with sufficient statistical power, sample size and adequate type of data in order to provide maximum information from chemical data and address efficiently the challenges and goals imposed by an intended research. The literature on chemometric experimental design covers different branches of chemistry ranging from practical to theoretical chemistry and, most importantly, it contains a wide variety of experimental designs such as factorial designs [1–11], simplex designs [1–9], Plackett–Burman designs [1,3,5–10], Box–Behnken designs [2,5,7–10], star designs [1–3], central composite designs [1–3,5–11], Taguchi designs [1,8,10], and Doehlert designs [5–10,12]. It is possible to estimate the percentage of articles discussing specific types of designs published before and after the year 2000 (Fig. 1), by entering in Scopus (a well-known abstract and citation database) the keyword "experimental design",

☆ This paper belongs to the Special Issue Chemometrics in Chromatography, Edited by Pedro Araujo and Bjørn Grung.

* Corresponding author. Tel.: +47 95285039; fax: +47 55905299.
E-mail address: pedro.araujo@nifes.no (P. Araujo).

1570-0232/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jchromb.2012.05.019

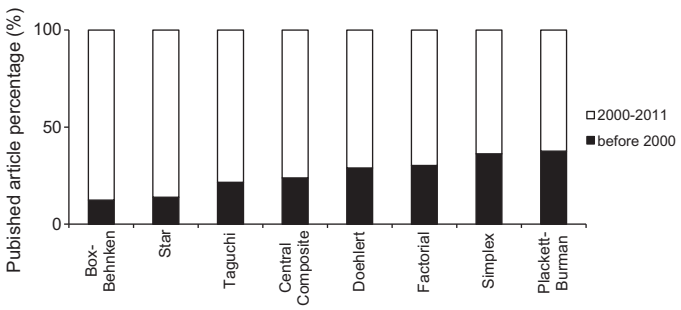


Fig. 1. Percentage of publication on specific types of designs published before and after the year 2000. The information was obtained from Scopus (10th December 2011).

searching within the results for the keyword "chemistry" and finally limiting the search to the name of every design mentioned above (e.g. "central composite design"). The Scopus results (Fig. 1) allow concluding that more than 50% of the total number of articles on chemometric experimental design has been published in the period 2000–2011. This upsurge of interest in experimental design application in the last decade may be seen as a product of the explosion of information technology on the web. Somewhat less known experimental designs are those proposed by Doehlert [12] four decades ago. Doehlert uniform shell designs were largely ignored in well-known reference books [1–4,11] and also in the most popular web based encyclopedias (e.g. Wikipedia) despite the increased percentage of publications on Doehlert design applications in the last decade (comparable with the more traditional factorial designs) and regardless of the chief merit of these types of designs when compared with classical designs (e.g. central composite design). It is surprising that Doehlert designs have only recently been acknowledged in some books [5–10].

This article discusses comprehensively the principles, importance and application of Doehlert uniform shell designs in different aspects of chromatography by using examples from the literature. The confidence of different models generated by Doehlert designs is visualized by estimating the associated uncertainty over the experimental domain and compared with widely used designs. The article also discusses some reported misconceptions about the Doehlert uniform shell designs with a view to prevent their erroneous application in chromatography.

2. Doehlert uniform shell designs

Doehlert designs are called "uniform shell designs" due to their regular distribution of the experimental points on the surface of spherical shells which in addition confer some important uniformity properties to be discussed in this article.

2.1. Generation of the design

Doehlert designs for k variables ($k > 3$) are always generated by allocating the experimental points on the surface of a hyper-sphere. For three and two variables the experimental points are circumscribed into a sphere and a circle of radius 1 respectively (Fig. 2). The combination of the different variables at every experimental point is obtained by projecting the cross section of the sphere in two dimensions (Fig. 2b) and the total number of experimental combinations (η) as a function of the number of variables (k) under study is given by the expression:

$$\eta = k^2 + k + 1 \quad (1)$$

For three and two variables the total numbers of experiments are 13 and 7 respectively and their spatial distribution are located at the vertices of a cuboctahedron and a hexagon with a point at their centers respectively (Fig. 2).

Another important characteristic of this type of design is the unequal number of experimental levels at the different axes. For instance, the three axes in Fig. 2 suggest that the factors x_1, x_2 and x_3 should be analyzed at 5, 3 and 7 different experimental levels respectively. The unequal number of levels is an important feature in cases where the factors under study are subjected to different and unavoidable constraints. For example, when applying coupling techniques such as liquid chromatography–mass spectrometry (LC–MS) or gas chromatography–MS (GC–MS) it is expected that the chromatography and MS associated factors will exhibit different ranges of variation due to the different principles of the instrumental techniques. For instance, a Doehlert design has been used to determine the influence of instrumental parameters on the quantitative determination of tri-α-linolenoylglycerol by LC-ion-trap-MS [13]. The drying gas flow rate and nebulizer gas pressure were investigated at 3 and 5 levels respectively due to their intrinsic shorter span imposed by the MS instrument, while the variable chromatographic flow rate was investigated at 7 levels due to its wider range. This imbalance in the number of levels represents an advantage over more traditional designs exhibiting the same number of levels in all directions (e.g. star designs, central composite designs).

In cases where the variables have been screened and the instrumental system does not impose any restriction on the significant variables, the levels dictated by the Doehlert design are assigned according to the magnitude of the effect of the variable. The highest the effect of a variable on the experimental response the highest the number of levels assigned to it. For example, if six variables are

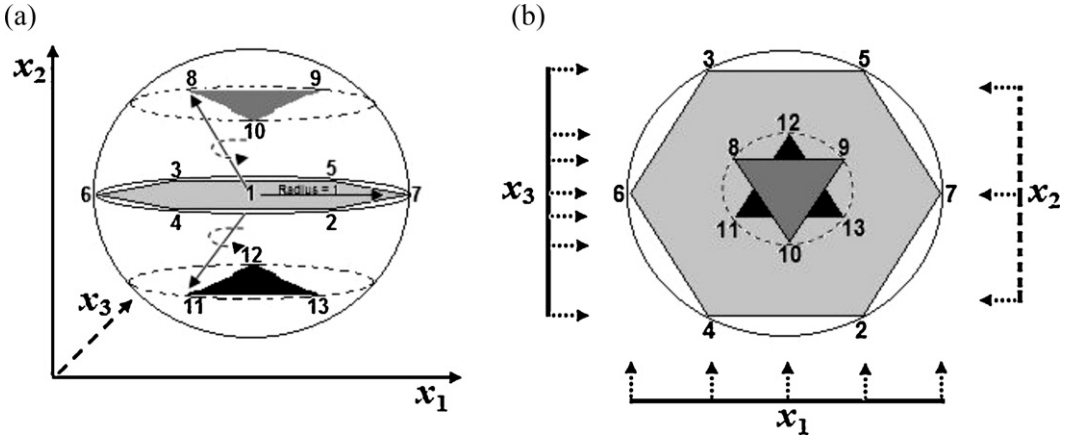


Fig. 2. Spatial distribution of the experimental points in a Doehlert design (a) for 2 (points 1–7) and 3 (points 1–13) factors; (b) cross section projection of the initial sphere.

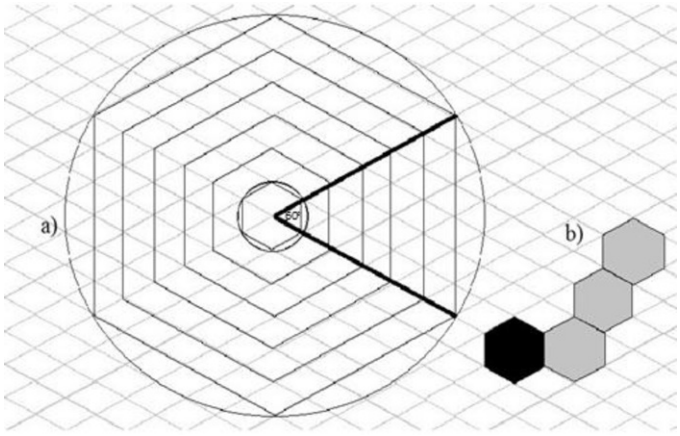


Fig. 3. (a) Translational rhombic tiling with a constant vertex angle of 60° generated by increasing gradually the radius of a Doehlert design in concentric circles. (b) Extension of an initial Doehlert design (black hexagon) to neighboring experimental domains by reusing previous experimental points.

evaluated by a factorial design and a Pareto chart of standardized effects reveals that only three variables are statistically significant, then the variable with the highest effect is evaluated at 7 levels, the variable with the second highest effect is evaluated at 5 levels and the variable with third highest effect is evaluated at 3 levels. This selection criterion has been applied in the screening and optimization of experimental factors by GC [14,15]. For instance, the optimization of the variables that affect the purity and yield of beef tallow biodiesel production (x_1: temperature, x_2: reaction time, x_3: catalyst concentration and x_4: alcohol:tallow molar ratio) was firstly screened by using a two-level factorial design [14]. The results indicated that x_4 and x_3 with Pareto scores of 62.0 and 39.2 were the most significant variables and were subsequently optimized by means of a two-factor Doehlert design at 5 and 3 levels respectively. One important property of a Doehlert design is that the enlargement of the original arrangement of experimental points which is carried out by increasing gradually the radius in concentric spheres or circles will always create a translational rhombic tiling with vertex angles of 60° (Fig. 3) regardless of which point is chosen as the center. The generated rhombic network can accommodate neighboring regions or new variables in any direction by extending the initial arrangement to another experimental domain by reusing previous experimental points adjacent to the new experimental region. This particular property has been used in the optimization of the factors that influence the methylation yield of polyunsaturated fatty acids in cod liver oil by gas chromatography [16] by implementing initially a Doehlert design for 2 factors (x_1: reaction time and x_2: alkali concentration) and performing 7 chromatographic runs (points 1–7 in Fig. 2). After analyzing the results, a second Doehlert design was considered with 2 additional factors (x_3: alkali volume and x_4: bath temperature) and 14 extra chromatographic runs performed (points 8–13 in Fig. 2 plus 8 additional points in a fourth dimension). The final model in this study was constructed by using the 21 experimental responses from both Doehlert designs.

2.2. Measures of uniformity of the models

Two measures of uniformity are proposed by Doehlert in his original article, namely: the distance from the center to the external points and the distance from a selected external point to the other points (excluding the central point).

For a better understanding of the impact of these two measures of uniformity on the modeling process, it is important to discuss in advance the parameter uncertainty.

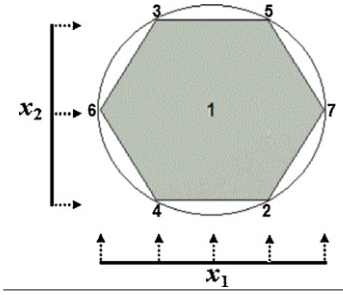
2.2.1. Uncertainty and confidence in the design

Uncertainty can be defined as a parameter that gives a quantitative indication of the confidence associated with a determined experimental design prior to performing any experiment. The uncertainty of an experimental design can be determined by means of the Working-Hotelling confidence limits [17] expressed as:

$$y_{\pm} = s_0 \sqrt{m \times F_{m \cdot m-n} \times h} \quad (2)$$

The terms as they appear in Eq. (2) represent the measurement, the root mean square overall error, the number of parameters in the model, the Fisher variance ratio and the uncertainty. The subscript n represents the total number of experiments. The most important feature of Eq. (2) is the presence of both non-design dependent (e.g. y, s_0, m) and design dependent (e.g. h) terms. On closer inspection, the design dependent term h in Eq. (2) measures the potential influence of an observation on a selected response (y) and it has an inverse relationship with $y_{\pm}$. An experimental design exhibiting a high value of h at a particular point is synonymous with widely separated confidence bands around y and lack of confidence in the design to predict the value of this point. On the contrary, a low value of h is associated with narrow confidence bands around y and a high degree of confidence in the design. The term h in Eq.

Table 1
Generation of a design matrix X for a 2-factor Doehlert design and calculation of the uncertainty (h) at every experimental level.

Doehlert design for $k=2$	Experimental point	Coded variables		Design matrix **X**						$h = \mathbf{x}_n(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_n^T$
		x_1	x_2	x_o	x_1	x_2	x_1x_2	x_1^2	x_2^2	
	1→	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	→0.333
	1→	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	→0.333
	1→	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	→0.333
	2→	0.500	−0.866	1.000	0.500	−0.866	−0.433	0.250	0.750	→0.833
	3→	−0.500	0.866	1.000	−0.500	0.866	−0.433	0.250	0.750	→0.833
	4→	−0.500	−0.866	1.000	−0.500	−0.866	0.433	0.250	0.750	→0.833
	5→	0.500	0.866	1.000	0.500	0.866	0.433	0.250	0.750	→0.833
	6→	−1.000	0.000	1.000	−1.000	0.000	0.000	1.000	0.000	→0.833
	7→	1.000	0.000	1.000	1.000	0.000	0.000	1.000	0.000	→0.833
	$n=9$ experiments									$\sum_{i=1}^{n} h_i = 6$

$\sum_{i=1}^{n} h_i$ = number of coefficients in the model.

(2) emerges as a powerful indicator of the confidence that could be attributed to the design and can be expressed mathematically by:

$$h = \mathbf{x}_n(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_n^T \quad (3)$$

where $\mathbf{x}_n$ represents the n-row of the design matrix $\mathbf{X}$. A comprehensive example for generating the design matrix $\mathbf{X}$ is given in Table 1 for a 2-factor Doehlert design with three replicates at the center and intended at studying a model of the form:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_1x_2 + b_4x_1^2 + b_5x_2^2 \quad (4)$$

The $\mathbf{X}$ matrix should have a number of rows equal to the total number of experiments and a number of columns in accordance with the number of coefficients in Eq. (4), in that way the dimension of the $\mathbf{X}$ matrix in Table 1 will be 9×6. The most important feature of Eq. (3) is that it allows assessing how confident the design predicts the data at every experimental point without performing any experiment. For instance, the experimental point number 3 (−0.500, 0.866) of the Doehlert design described in Table 1, which corresponds to the row number 5 ($\mathbf{x}_5 = 1$, −0.500, 0.866, −0.433, 0.250, 0.750) of the $\mathbf{X}$ matrix has an associated uncertainty value of 0.833 calculated by the expression $h = \mathbf{x}_5(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_5^T$. One important property of Eq. (3) is that the sum of h over all experimental point equals the number of coefficients in the model ($\sum h_i = 6$ coefficients). Although it is not the intention of the authors to give a full account of the calculations and graphical representation, Eq. (3) can be used to visualize uncertainty contour plots over an entire experimental domain. The reader interested in the mathematical calculations behind the contour plots and their generation is referred to the various comprehensive articles published by the authors elsewhere [18–21].

2.2.2. Distance from the center to the external points

This particular distance is commonly referred and described in related applications and is kept constant, as explained above, by circumscribing all the experimental points on the surface of a hyper-sphere (if $k > 3$), a sphere (if $k = 3$) or a circle (if $k = 2$) of radius 1 (Fig. 2), so that it is possible to produce a constant uncertainty at every experimental point around the center. For example, the experimental points around the center of the 2-factor Doehlert design described in Table 1 display a constant uncertainty value of 0.833 (points 2–7) while the minimum uncertainty is found at the center of the design. In addition, the allocation of the experimental points at a constant distance from the center, enables the incorporation of additional factors without influencing the constancy of the uncertainty. This property is of importance when additional factors are included into a particular mathematical model. For instance, the effect of the mass spectrometer parameters (x_1: nebulizer gas pressure and x_2: drying gas flow rate) on the analytical response (y: peak area) could be modeled as:

$$y = b_0 + b_1x_1 + b_2x_2 \quad (5)$$

By using Eq. (3) an uncertainty value of 0.476 is recorded for all the points around the center (Table 2, $k = 2$). However, if an additional variable is considered (x_3: chromatographic flow rate) the original model could be expanded to:

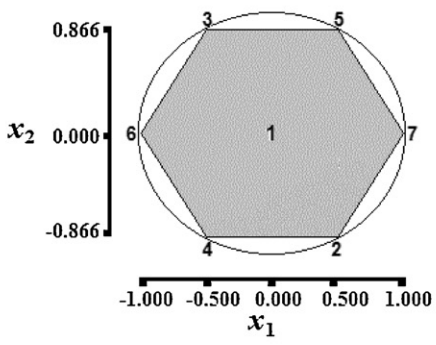
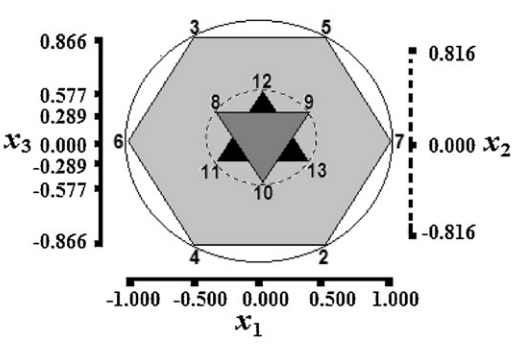
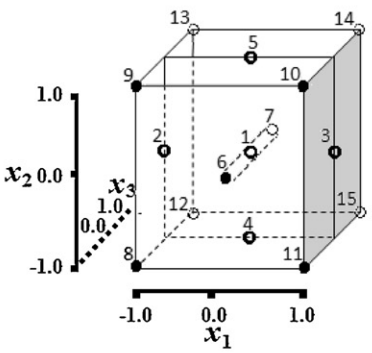
$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 \quad (6)$$

Six extra experiments should be added and distributed spatially as is shown in Table 2 for $k = 3$. The associated uncertainty of the 13 total experimental conditions is calculated by using Eq. (3) and it revealed that by combining both, the seven initial (points 1–7) and the extra six (points 8–13) experiments, a constant uncertainty value of 0.327 is computed around the central point (Table 2, $k = 3$) regardless of the incorporation of the chromatographic flow rate as an additional factor. The only observed change after increasing the number of variables from $k = 2$ to 3 is the reduction in the magnitude of the uncertainty as a result of increasing the number of experimental runs from $\eta = 7$ to 13.

The previous results for the 3-factor Doehlert design are compared to a 3-factor face-centered cubic design (the most popular type of central composite design according to a review of the literature) which required a minimum of 15 experiments ($\eta = 2^k + 2k + 1 = 2^3 + 2 \times 3 + 1 = 15$). The lack of uniformity in the distance from the center to the external points results in unequal values of uncertainty, 0.167 and 0.367, for the experimental points spatially distributed in the star part (points 2–7 in Table 2) and factorial part (points 8–15) of the face-centered cubic design respectively. It is evident that the observed uniformity in the designs proposed by Doehlert is advantageous to generate models that are sufficiently smooth to permit interpolation. In addition, the statistical reliability of the constant uncertainty can be visualized as a regular scattering of the experimental responses at every

Table 2
Calculation of the uncertainty values for Doehlert designs with 2 and 3 factors and a face-centered cubic designs with 3 factors.

Designs

Doehlert		Face-Centred Cubic
$k = 2$	$k = 3$	$k = 3$
		
Uncertainty at every experimental point estimated by $h = \mathbf{x}_n(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_n^T$		
Point 1 = 0.143 Points 2-7 = 0.476	Point 1 = 0.077 Points 2-13 = 0.327	Point 1 = 0.067 Points 2-7 = 0.167 (star points) Points 8-15 = 0.367 (factorial points)

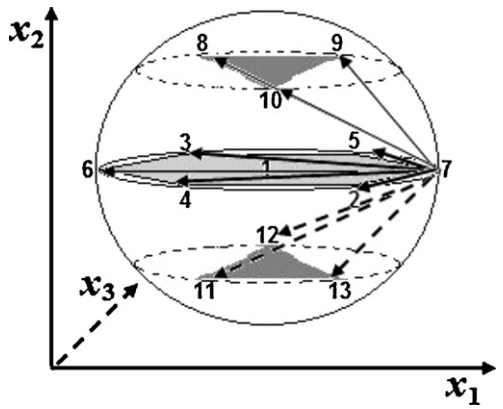


Fig. 4. Distance vectors from a selected reference point (number 7) to the other points and by excluding the central point (1).

experimental point and consequently as a measure of how closely the models can fit the data.

It must be mentioned that models different from those described by Eqs. (5) and (6) could be also considered, for instance, the quadratic model described by Eq. (4). However, in order to judge the adequacy of more complex models, the inclusion of replicate measures is required to ensure a good balance between the number of parameters and degrees of freedom.

2.2.3. The distance from a selected external point to the other points (excluding the central point)

Regardless of its importance, this particular distance is rarely acknowledged in publications. It can be demonstrated by using graphical representations that it is possible to compute 5 and 11 vectors for $k=2$ and 3 respectively by taking as reference the distance from point number 7 to the rest of the experimental points lying on the surface of the sphere of radius 1 and excluding the central point (Fig. 4). Simple algebra allows estimating, for $k=2$ and reference point 7, two vectors with a distance of 1 (vectors 7-5 and 7-2 in Fig. 4), two vectors with a distance of $\sqrt{3}$ (vectors 7-3 and 7-4 in Fig. 4) and one vector with a distance of 2. These 5 vectors for $k=2$ can be represented by a distance pattern vector of the form:

$$[2\times\sqrt{1},\quad 2\times\sqrt{3},\quad 1\times\sqrt{4}]$$

Similarly, for $k=3$ the 11 distance vectors illustrated in Fig. 4 can be represented as:

$$[4\times\sqrt{1},\quad 2\times\sqrt{2},\quad 4\times\sqrt{3},\quad 1\times\sqrt{4}]$$

It is difficult to visualize the number of vectors for $k\gg 3$ or their distance from a reference point by using graphical display or simple algebra respectively. However, an increase in the number of factors k will result in a general distance pattern vector of the form:

$$\left[\alpha\times\sqrt{1},\quad \beta\times\sqrt{2},\quad \gamma\times\sqrt{3},\quad \delta\times\sqrt{4}\right] \tag{7}$$

where α, β, γ and δ represent the number of vectors with a distance of 1, $\sqrt{2}$, $\sqrt{3}$ and 2 respectively, constructed by excluding the central point and defined by:

$$\alpha=\gamma=2\times(k-1) \tag{8}$$

$$\beta=(k-1)\times(k-2) \tag{9}$$

$$\delta=1 \tag{10}$$

The total number of vectors $(\alpha+\beta+\gamma+\delta)$ as a function of the number of factors is given by the expression:

$$\omega=k^2+k-1 \tag{11}$$

The application of Eqs. (8)–(11) for computing the number of distance vectors generated by increasing the number of factors from 2 to 10 is presented in Table 3. The uniform distance patterns described in Table 3 is found for all dimensions and its configuration (α, β, γ and δ) and magnitude (1, $\sqrt{2}$, $\sqrt{3}$ and 2) is the same no matter which point is selected as reference. In addition, the uniform distance patterns exhibited by the uniform shell designs confer some distinctive rugged properties to the models such as a constant experimental domain volume and consequently a high confidence in the predictions over the experimental domain.

2.3. Doehlert designs, efficiency and degrees of freedom

The efficiency (ε) is a function that relates the number of parameters p (b_0, b_1, b_2, b_3, etc.) in a model and the total number of experiments η required to generate the model in question and it is expressed as:

$$\varepsilon=\frac{p}{\eta}\times 100 \tag{12}$$

It is clear from Eq. (12) that the worst values of ε are those obtained when the number of experiments equals the number of parameters ($\lim_{\eta\to p}\varepsilon=100\%$). For instance, a 2-factor Doehlert design ($\eta=7$) used to estimate a seven coefficients model of the form:

$$y=b_0+b_1x_1+b_2x_2+b_3x_1x_2+b_4x_1^2+b_5x_2^2+b_6x_1^2x_2^2 \tag{13}$$

will yield a $\varepsilon=100\%$ with no degrees of freedom ($\eta-p=0$) to judge the adequacy of this model. Although the ideal ε values are those close to zero ($\lim_{\eta\to\infty}\varepsilon=0$ is ideal but impractical due to the high number of experiments involved), values under 60% can be rated as optimal. It should be noted that $\varepsilon<60\%$ implies in practical terms the existence of an appropriate balance between the number of experiments and the number of parameters. As a rule of thumb, the minimum acceptable difference between η and p should be 3 ($\eta-p=3$).

Table 3
Number of distance vectors generated by increasing the number of factors from 2 to 10 and by using Eqs. (8)–(10).

Number of factors (k)	Number of vectors at a distance				Total number of distance vectors (ω)
	$\alpha=1$	$\beta=1.414$	$\gamma=1.732$	$\delta=2$	$\omega=k^2+k-1$
2	2	0	2	1	5
3	4	2	4	1	11
4	6	6	6	1	19
5	8	12	8	1	29
6	10	20	10	1	41
7	12	30	12	1	55
8	14	42	14	1	71
9	16	56	16	1	89
10	18	72	18	1	109

3. Misconceptions about Doehlert uniform shell designs

It is not the intention of the present section to discredit any particular research article or to discourage anyone from reading them. For these reasons the reference numbers of the criticized articles are not provided. However, specific terminology, definitions or statements textually taken from these articles are given in italics. It must be emphasized that the main objective of this section is to help readers avoid using incorrect definitions in connection with Doehlert uniform shell designs.

The general definition of Doehlert uniform shell designs has sometimes been misclassified as "*when a simplex optimization with two variables comes to the point where it encircles the optimum*". Such an oversimplified and erroneous definition confines the optimum to lie invariably at the center of the design and it is in direct contradiction with well-known examples from the literature. For instance, the first application of a 2-factor Doehlert design in synthesis optimization described the use of a GC instrument equipped with a flame ionization detector (FID) to monitor the yield of the reaction [22] and reported an optimum yield away from the central experimental point. Perhaps, the above mentioned misconception is due to the fact that in his article Doehlert used a regular simplex to explain the generation of the experimental points. However, it must be emphasized that in no instance Doehlert's original article [12] mentioned that a simplex design optimization should be carried out to obtain an encircling optimum. The characteristics of both designs, simplex and uniform shell, are quite different and have been described in the literature [23].

One serious issue regarding the application of uniform shell designs in chromatography is the belief that *a Doehlert design is more efficient because it estimates more coefficients with fewer experiments. For instance, the six coefficients of a quadratic model with two factors are estimated with seven experiments by the Doehlert design (R-efficiency 85.71%) and with nine experiments by the central composite design (R-efficiency 66.67%)*. The reader must be aware that efficiency (ε) defined in Eq. (12) measures the relation between the number of coefficients (p) in a model and the total number of experiments (η) and should be analyzed not only by considering its numerical value but also by taking into account the associated number of degrees of freedom. For instance an experimental design with $\varepsilon = 100\%$ yields an overfitted model ($p = \eta$) that cannot be validated statistically to check its predictive performance due to the lack of degrees of freedom ($DF = \eta - p$). Experimental designs with $\varepsilon < 60\%$ are better choice for modeling purposes [24]. In this context, if the italicized statement implies using a chromatographic system that has demonstrated to yield highly precise measurements, making the pure error too small as to be considered of practical importance, then a 6-coefficient model generated from the central composite design with lower efficiency ($\varepsilon = 66.67\%$) and 3 degrees of freedom ($DF = 9 - 6 = 3$) will offer more reliable predictions than a model derived from a Doehlert design with higher efficiency ($\varepsilon = 85.71\%$) and minimum number of degrees of freedom ($DF = 7 - 6 = 1$). Doehlert uniform shell designs are valuable tools for modeling a wide variety of analytical responses as functions of sample preparation or instrumental related factors and offer several advantages over the widely applied central composite designs. However their correct implementation requires a good knowledge of mathematics and statistics.

Another important aspect introduced in the literature is the use of some designs exhibiting a non-uniform distance from the center to the external points and labeled as deformed Doehlert designs [25,26]. Unfortunately, these deformed designs have been mistakenly reported in some research articles as genuine Doehlert designs in recent years. The readers should be aware that uniformity is a property overemphasized and demonstrated by Doehlert in his original article from different perspectives. Any attempt to change the spatial disposition of the experimental points proposed by a Doehlert design can have an impact on the uniformity around the central point, the uncertainty over the entire experimental domain and consequently the confidence in the generated models.

4. Applications in chromatography

4.1. Gas chromatography

The first two publications on the application of Doehlert designs involved the study of two and four factors using GC for monitoring the yield of the synthetic product 4-(*N*,*N*-dimethyl-amino)-acetophenone and the saponification of cod liver oil respectively [16,22]. After this first application, uniform shell designs have been primarily used in the development of GC sample treatment procedures by optimizing the recovery [14,15,27–29] and the intensity of the analytical signals [24,30–33]. The optimization of the GC instrumental parameters by means of Doehlert designs has been rarely considered in the literature. To the best of our knowledge, the only reported article focused on the optimization of the three column related factors that can affect the chromatographic performance (x_1: temperature, x_2: pressure and x_3: length) by considering 10 specific response functions namely the flow rate (y_1), the analysis time (y_2), the height equivalent to a theoretical plate under carbon monoxide and xenon (y_3 and y_4 respectively) to characterize the column efficiency, the capacity factor under argon, nitrogen, carbon monoxide and xenon (y_5, y_6, y_7 and y_8 respectively) for measuring the retention and the valley height (y_9 and y_{10}) to characterize the chromatographic resolution [34].

4.2. Liquid chromatography

Contrary to GC, Doehlert designs have been equally applied in LC for developing sample preparation protocols and optimizing the instrumental conditions. The main optimized parameters considered in the development of sample protocols have been the reaction yield [35], the recovery [36–38] and the analytical signal intensity [39,40]. The majority of the studies on the improvement of the instrumental conditions focused generally on the optimization of the mobile phase [41–47]. Other instrumental parameters considered in a lesser extent are the flow rate [41,46,48], the analysis time [48], and column temperature [46].

4.3. Gas chromatography–mass spectrometry

An increased implementation of Doehlert uniform shell designs in GC–MS has been observed in the last decade for developing efficient sample treatment methods by monitoring different analytical functions such as recovery [49–55], analytical signal [56–61] and reaction yield [51,62]. Table 4 also shows that regardless of the potential benefits of uniform shell designs, they have not gained widespread popularity among researchers for optimizing systematically and simultaneously GC–MS instrumental variables. The few reported works in this area have been focused on optimizing the temperature gradient program and gas velocity [63], the splitless time and split flow [52] and the final temperatures of the first and second ramp of a chromatographic temperature program [64].

4.4. Liquid chromatography–mass spectrometry

Table 4 shows that the number of publications on the use of uniform shell designs and LC–MS is less compared to the above mentioned chromatographic techniques. Most of the applications have focused on modeling the internal standard response factor as a function of the analyte and internal standard concentrations

Table 4
Overview of the applications of Doehlert uniform shell designs in chromatography.

Methodology	Parameter optimized	Number of factors (k)	Area of implementation		Reference
			Sample protocol	Instrument	
GC	Reaction yield	2	×		[22]
		4	×		[16]
	Recovery	2		×	[14]
		2	×		[15]
		3	×		[27,29]
		4	×		[28]
	Analytical signal	2	×		[24,32]
		3	×		[30,33]
		4	×		[31]
	Instrumental[a]: CT, CP, CL	3		×	[34]
LC	Reaction yield	2	×		[35]
	Recovery	3	×		[36–38]
	Analytical signal	3	×		[40]
		4	×		[39]
	Instrumental[a,b]: CT, MP, FR, AT, GL,	2		×	[43–45,47,48]
		3		×	[41,42,46]
GC–MS	Reaction yield	2	×		[51]
		4	×		[62]
	Recovery	2	×		[49,50,52–55]
		3	×		[51]
	Analytical signal	2	×		[56,59–61]
		3	×		[57,58]
	Instrumental[c]: ST, SF, TP, GP, GV	2		×	[52,64]
		3		×	[63]
LC–MS	Recovery	2	×		[67]
		3	×		[68,69]
	Analytical signal	2	×		[65,66]
	Instrumental[b,d]: MP, GF, V, GNP	3		×	[70,13]

[a] CT, column temperature; CP, column pressure; CL, column length.
[b] MP, mobile phase; FR, flow rate; AT, analysis time; GL, gradient length.
[c] ST, splitless time; SF, split flow; TP, temperature programming; GP, gradient program; GV, gas velocity.
[d] GF, MS gas flow; V, ESI needle voltage; GNP, MS nebulizer gas pressure.

[65,66] and on determining the optimal recovery conditions of developed analytical methods [67–69]. Articles regarding instrumental optimization are mainly concerned with the 3 factors such as x_1: sheath gas flow, x_2: ESI needle voltage and x_3: pH of the mobile phase [70] or x_1: MS gas flow rate, x_2: MS nebulizer gas pressure and x_3: LC flow rate [13].

5. Concluding remarks

Doehlert uniform shell designs are generally used for determining the optimal combination of the factors that have the strongest influence on selected single or multiple experimental responses.

The implementation of Doehlert uniform shell designs in chromatography has been focused on the development and improvement of sample preparation procedures by exploring up to four factors and on the optimization of specific instrumental parameters by exploring up to three factors. It has been recommended not to examine more than three factors due to the inherent high number of experiments [71]. However, considering that the chromatographic system alone could be influenced by more than 50 factors [72], then it would be interesting to investigate the performance of Doehlert uniform shell designs in the optimization of a high number of chromatographic instrumental parameters ($k \gg 3$).

References

[1] R. Brereton, Chemometrics Data Analysis for the Laboratory and Chemical Plant, John Wiley & Sons Ltd., West Sussex, England, 2003.
[2] M. Otto, Chemometrics: Statistics and Computer Application in Analytical Chemistry, Wiley-Vch, Weinheim, Germany, 1999.
[3] D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte, L. Kaufman, Chemometrics: A Textbook, Elsevier Science Publishers B.V, Amsterdam, The Netherlands, 1988.
[4] J.N. Miller, J.C. Miller, Statistics and Chemometrics for Analytical Chemistry, 4th ed., Pearson Educational Limited, Essex, England, 2000.
[5] S.D. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive Chemometrics: Chemical and Biochemical Data Analysis, 1st ed., Elsevier B.V, Amsterdam, The Netherlands, 2009.
[6] K. Valkó (Ed.), Handbook of Analytical Separations: Separation Methods in Drug Synthesis and Purification, Elsevier Science B.V, Amsterdam, The Netherlands, 2000.
[7] G. Hanrahan, F.A. Gomez (Eds.), Chemometric Methods in Capillary Electrophoresis, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2010.
[8] R.E. Bruns, I.S. Scarminio, B.D.B. Neto, Statistical Design—Chemometrics, Elsevier B.V, Amsterdam, The Netherlands, 2006.
[9] D.B. Hibbert, Quality Assurance for the Analytical Chemistry Laboratory, Oxford University Press, Inc, New York, USA, 2007.
[10] T.P. Ryan, Modern Experimental Design, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2007.
[11] S.N. Deming, S.L. Morgan, Experimental Design: A Chemometric Approach, 2nd revised and expanded ed., Elsevier, New York, 1993.
[12] D.H. Doehlert, Appl. Stat. 19 (1970) 231.
[13] P. Araujo, T.T. Nguyen, L. Frøyland, J. Am. Soc. Mass Spectrom. 16 (2005) 388.
[14] D.R. Mendoça, H.M.C. Andrade, P.R.B. Guimarães, R.F. Vianna, S.M.P. Meneghetti, L.A.M. Pontes, L.S.G. Teixeira, Fuel Process. Technol. 92 (2011) 342.
[15] M.S. dos Santos, D. Budziak, E. Martendal, E. Carasek, Anal. Sci. 25 (2009) 865.
[16] F. Lamothe, D. Peyronel, M. Sergent, M.C. Iatrides, J. Artaud, R. Phan-Tan-Luu, J. Am. Oil Chem. Soc. 65 (1988) 652.
[17] H. Working, H. Hotelling, J. Am. Stat. Assoc. 24 (1929) 73.
[18] P.W. Araujo, R.G. Brereton, Analyst 122 (1997) 621.
[19] P.W. Araujo, R.G. Brereton, TRAC – Trends Anal. Chem. 15 (1996) 156.
[20] P. Araujo, L. Frøyland, Accredit. Qual. Assur. 10 (2005) 185.
[21] P. Araujo, L. Frøyland, Anal. Chim. Acta 555 (2006) 348.
[22] T. Lundstedt, P. Thorén, R. Carlson, Acta Chem. Scand. B 38 (1984) 717.
[23] P.W. Araujo, R.G. Brereton, TRAC – Trends Anal. Chem. 15 (1996) 63.
[24] P. Araujo, F. Couillard, E. Leirnes, K. Ask, A. Bøkevoll, L. Frøyland, J. Chromatogr. A 1121 (2006) 99.
[25] M. Nechar, M.F.M. Molina, L.C. Rodriguez, J.M.B. Sendra, Anal. Chim. Acta 316 (1995) 185.
[26] A.M.G. Campaña, L.C. Rodríguez, A.L. Gonzalez, F.A. Barrero, M.R. Ceba, Anal. Chim. Acta 348 (1997) 237.
[27] D. Budziak, E. Martendal, E. Carasek, Microchim. Acta 164 (2009) 197.
[28] C.H. Kowalski, G.A. da Silva, R.J. Poppi, H.T. Godoy, F. Augusto, Anal. Chim. Acta 585 (2007) 66.

[29] C. Aguilar, A. Peñalver, E. Pocurull, J. Ferré, F. Borrull, R.M. Marcé, J. Chromatogr. A 844 (1999) 425.
[30] B. Zhou, Y. Li, J. Gillespie, G.Q. He, R. Horsley, P. Schwarz, J. Agric. Food Chem. 55 (2007) 10141.
[31] G.A. da Silva, F. Augusto, R.J. Poppi, J. Chromatogr. A 1138 (2007) 251.
[32] D. Budziak, E. Martendal, E. Carasek, J. Sep. Sci. 31 (2008) 2875.
[33] A. Martínez-Uruñuela, J.M. González-Sáiz, C. Pizarro, J. Chromatogr. A 1056 (2004) 49.
[34] E. de Vanssay, S. Zubrzycki, R. Sternberg, F. Raulin, M. Sergent, R. Phan-Tan-Luu, J. Chromatogr. A 688 (1994) 161.
[35] J. Quesada, Eur. J. Lipid Sci. Technol. 105 (2003) 281.
[36] L. Bartolomé, J. Lezamiz, N. Etxebarria, O. Zuloaga, J.A. Jönsson, J. Sep. Sci. 30 (2007) 2144.
[37] D.M. Brum, R.J. Cassella, A.D.P. Netto, Talanta 74 (2008) 1392.
[38] P. Fernández, C. Vázquez, R.A. Lorenzo, A.M. Carro, I. Álvarez, P. Cabarcos, Anal. Bioanal. Chem. 397 (2010) 677.
[39] R. Romero, M.G. Bagur, M. Sánchez-Viñas, D. Gázquez, Chromatographia 51 (2000) 404.
[40] F. Masson, A. Lebert, R. Talon, M.C. Montel, J. Appl. Microbiol. 83 (1997) 36.
[41] T.R.I. Cataldi, D. Orlando, D. Nardiello, A. Rubino, G. Bianco, S. Abate, R. Ciriello, A. Guerrieri, Anal. Chim. Acta 597 (2007) 129.
[42] J.A. Hernández-Arteseros, I. Boronat, R. Compañó, M.D. Prat, Chromatographia 52 (2000) 295.
[43] R. Bonfilio, C.R.T. Tarley, G.R. Pereira, H.R.N. Salgado, M.B. de Araújo, Talanta 80 (2009) 236.
[44] F. Goutelard, C. Caussignac, R. Brennetot, G. Stadelmann, C. Gautier, J. Radioanal. Nucl. Chem. 282 (2009) 669.
[45] B. Bourguignon, F. Marcenac, H.R. Keller, P.F. de Aguilar, D.L. Massart, J. Chromatogr. 628 (1993) 171.
[46] P.A. Babay, R.T. Gettar, J.F. Magallanes, E.T. Becquart, B. Thiele, D.A. Batistoni, J. Chromatogr. A 1157 (2007) 227.
[47] E.J. Klein, S.L. Rivera, J.E. Porter, Biotechnol. Prog. 16 (2000) 506.
[48] P. Araujo, TRAC – Trends Anal. Chem. 19 (2000) 524.
[49] P. González, I. Racamonde, A.M. Carro, R.A. Lorenzo, J. Sep. Sci. 34 (2011) 2697.
[50] I. Racamonde, P. González, R.A. Lorenzo, A.M. Carro, J. Chromatogr. A 1218 (2011) 6878.
[51] P.C.F. de Lima Gomes, J.Y. Barletta, C.E.D. Nazario, A.J. Santos-Neto, M.A. Von Wolff, C.M.R. Coneglian, G.A. Umbuzeiro, F.M. Lancas, J. Sep. Sci. 34 (2011) 436.
[52] D. García-Rodríguez, A.M. Carro-Díaz, R.A. Lorenzo-Ferreira, R. Cela-Torrijos, J. Chromatogr. A 1217 (2010) 2940.
[53] E.A. Suchara, D. Budziak, E. Martendal, L.L.F. Costa, E. Carasek, Anal. Chim. Acta 613 (2008) 169.
[54] R. Rodil, A.M. Carro, R.A. Lorenzo, R. Cela, Chemosphere 67 (2007) 1453.
[55] S. Risticevic, E. Carasek, J. Pawliszyn, Anal. Chim. Acta 617 (2008) 72.
[56] A.G. Souza, L.M. Costa, R. Augusti, Z.L. Cardeal, Water Air Soil Pollut. 211 (2010) 427.
[57] D. Arroyo, M.C. Ortiz, L.A. Sarabia, J. Chromatogr. A 1157 (2007) 358.
[58] A. Navalón, A. Prieto, L. Araujo, J.L. Vílchez, J. Chromatogr. A 946 (2002) 239.
[59] E. Carasek, J. Pawliszyn, J. Agric. Food Chem. 54 (2006) 8688.
[60] Y. Zhou, Q. Jiang, Q. Peng, D. Xuan, W. Qu, Chemosphere 70 (2007) 256.
[61] E. Carasek, E. Cudjoe, J. Pawliszyn, J. Chromatogr. A 1138 (2007) 10.
[62] T. Dagnac, J.M. Guillot, P. Le Cloirec, J. Anal. Appl. Pyrol. 37 (1996) 33.
[63] L.K. Skartland, S.A. Mjøs, B. Grung, J. Chromatogr. A 1218 (2011) 6823.
[64] I. García, L.A. Sarabia, M.C. Ortiz, J.M. Aldama, Anal. Chim. Acta 544 (2005) 26.
[65] P. Araujo, T. Bjørkkjaer, A. Berstad, L. Frøyland, Prostaglandins Leukot. Essent. 77 (2007) 9.
[66] K. Gregersen, L. Frøyland, A. Berstad, P. Araujo, Talanta 75 (2008) 466.
[67] M. García-Chao, M.J. Agruña, G.F. Calvete, V. Sakkas, M. Llompart, T. Dagnac, Anal. Chim. Acta 672 (2010) 107.
[68] V. Tak, D. Pardasani, P.K. Kanaujia, D.K. Dubey, J. Chromatogr. A 1216 (2009) 4319.
[69] T. Dagnac, A. Padró, R. Rubio, G. Rauret, Anal. Chim. Acta 364 (1998) 19.
[70] V. Tak, P.K. Kanaujia, D. Pardasani, R. Kumar, R.K. Srivastava, A.K. Gupta, D.K. Dubey, J. Chromatogr. A 1161 (2007) 198.
[71] B. Dejaegher, Y.V. Heyden, LCGC Eur. 22 (2009) 256.
[72] J.A. van Leeuwen, L.M.C. Buydens, B.G.M. Vandeginste, G. Kateman, P.J. Schoenmakers, M. Mulholland, Chemometr. Intell. Lab. Syst. 10 (1991) 337.