Open Journal of Applied Sciences, 2025, 15(2), 463-479
https://www.scirp.org/journal/ojapps
ISSN Online: 2165-3925
ISSN Print: 2165-3917

Constructing Optimal Baseline Designs from Regular Designs

Shiyu Li

School of Mathematics and Statistics, Shandong Normal University, Jinan, China
Email: 19862141180@163.com

How to cite this paper: Li, S.Y. (2025) Constructing Optimal Baseline Designs from Regular Designs. *Open Journal of Applied Sciences*, **15**, 463-479.
https://doi.org/10.4236/ojapps.2025.152030

Received: January 19, 2025
Accepted: February 22, 2025
Published: February 25, 2025

Abstract

This paper delves into the baseline design under the baseline parameterization model in experimental design, focusing on the relationship between the K-aberration criterion and the word length pattern (WLP) of regular two-level designs. The paper provides a detailed analysis of the relationship between K_5 and the WLP for regular two-level designs with resolution $t=3$, and proposes corresponding theoretical results. These results not only theoretically reveal the connection between the orthogonal parameterization model and the baseline parameterization model but also provide theoretical support for finding the K-aberration optimal regular two-level baseline designs. It demonstrates how to apply these theories to evaluate and select the optimal experimental designs. In practical applications, experimental designers can utilize the theoretical results of this paper to quickly assess and select regular two-level baseline designs with minimal K-aberration by analyzing the WLP of the experimental design. This allows for the identification of key factors that significantly affect the experimental outcomes without frequently changing the factor levels, thereby maximizing the benefits of the experiment.

Keywords

Baseline Parameterization, K-Aberration Criterion, Regular Design, Word Length Pattern

1. Introduction

1.1. Research Background

Experimental design refers to the scientific and rational arrangement of experiments after clarifying the factors to be examined and the research objectives, in order to achieve the best experimental results. This process is also called experimental design. The purpose of experimental design is to explore the relationships between variables through systematic methods, so as to optimize processes and improve product quality.

The development of experimental design can be traced back to the early 20th century. British statistician R.A. Fisher published the first example of experimental design in collaboration with W.A. Mackenzie in 1923, and proposed the basic ideas of experimental design in 1926. In 1935, Fisher [1] published his famous book “The Design of Experiments”, in which he proposed three principles that experimental design should follow: randomization, local control, and replication. These principles aim to reduce the impact of accidental factors, so that experimental data has an appropriate mathematical model, which can be used for data analysis with the method of variance analysis.

In China, research on experimental design began in the 1950s, and there were new insights in the viewpoints, theories, and methods of orthogonal experimental design. The famous mathematician Professor Hua Luogeng actively advocated and popularized the “optimization method” in China, thus popularizing the concept of experimental design. In 1978, mathematicians Wang Yuan [2] and Fang Kaitai [3] proposed the uniform design, which considers how to scatter the design points evenly within the experimental range, so as to obtain the most information with fewer experimental points.

In the field of experimental design, with the continuous in-depth of scientific research and industrial applications, traditional orthogonal parameterization design methods have gradually shown limitations in some specific scenarios. Especially in experimental situations where it is necessary to screen key factors from many factors and hope to minimize changes to the existing process, baseline design has emerged, bringing new ideas and methods to the field of experimental design.

Baseline design is an experimental design in which each factor has a specified baseline level (default or preferred level), measuring the impact of changes in one or more factor levels on the response while other factors remain at the baseline level. This design parameterization method is different from orthogonal parameterization, which measures the impact of changes in one or more factor levels on the response, averaged over all possible level combinations of all other factors. The goal of baseline design is to identify designs with the minimum K-aberration.

1.2. Research Content

Based on the results of Miller and Tang [4], this paper further develops the relationship between K-aberration and word length pattern. Use 2^{m-p} to represent a regular two-level design with m factors and 2^{m-p} runs, and the two levels in the design are represented by 0, 1 respectively. For any regular 2^{m-p} design D, the set of all s column submatrices of D is denoted as $\Omega_s(D)$, where $s \le m$. Let $\alpha(W)$ represent the total number of rows in W where all elements are 1, where W is a subdesign composed of some columns of D. Under the baseline parameterization model, for a regular 2^{m-p} design D, K_s represents the total deviation caused by all s th-order factor interactions on the estimation of main effects in design D. Mukerjee and Hunter [5] proved that,

$$K_s = \left(4/N^2\right)\left(sT_1 + T_2\right),$$

where, $T_1 = \sum_{W \in \Omega_s(D)} \left(\alpha(W)\right)^2$, $T_2 = \sum_{W^* \in \Omega_{s+1}} \sum_{W^0 \in \Omega_s(W)} \left(2\alpha\left(W^*\right) - \alpha\left(W^0\right)\right)^2$. Given the number of rows and columns of the design, the two-level orthogonal design that minimizes the sequence

$$\left(K_2, K_3, \cdots, K_m\right)$$

in lexicographic order is called the K-aberration optimal design.

Based on the K_s expression as the theoretical foundation, this paper has thoroughly explored the research process of the K_s expression under various circumstances, analyzed the relationship between K_5 and the word length pattern for regular two-level designs with resolution $t = 3$, listed all possible defining word scenarios, and provided the expression results for K_5, offering corresponding theoretical results for the design of such experiments. That is, by sequentially minimizing the K_s values up to K_5, the design method can be obtained more quickly.

2. The Relationship between *K*-Criterion and Word Length Pattern

In the study of experimental design, we find that T_1 and T_2 are closely related to the subarrays W or W^* of the regular 2^{m-p} design. According to the definition of T_1 in the formula, when the subarray W composed of any s columns in design D does not contain defining words, when calculating T_1, the contribution of W to T_1 is $N/2^s$. When the subarray W composed of any s columns in design D contains defining words, that is, several columns in W form defining words, the situation becomes complicated. The existence of defining words means that some column combinations in W have special structures, which will affect the contribution of W to T_1. In addition, similar situations need to be analyzed when calculating T_2. Therefore, investigating whether the subarray W contains defining words is crucial for simplifying the calculation of K_s. For the convenience of the subsequent analysis of the K_s formula, the following lemmas are given to provide theoretical explanations for all possible cases of defining words.

2.1. Three Basic Lemmas

Lemma 1. Suppose D is a regular 2^{m-p} design with resolution t, let $W \in \Omega_{t+3}(D)$, then

(i) When $t = 3$, W contains at most three independent defining words;

(ii) When $4 \le t \le 6$, W contains at most two independent defining words;

(iii) When $t \ge 7$, W contains at most one defining word,

where $m \ge t + 3$.

Proof:

(i) When $t=3$, $W=\{m_1,m_2,m_3,m_4,m_5,m_6\}$, $W\in\Omega_6(D)$, then the number of independent defining words contained in W is as follows:

(a1) W does not contain independent defining words;

(a2) W contains one independent defining word, the length of which may be 3 or 4 or 5 or 6;

(a3) W contains two independent defining words, their lengths may be 3 and 3 or 3 and 4;

(a4) W contains three independent defining words, all of whose lengths are 3.

Situations (a1) and (a2) are obvious. In situation (a3), when the lengths of the two independent defining words are 3 and 3, there are two cases for their composition, that is, they contain a common column (for example, $m_1m_2m_3=m_3m_4m_5$) and do not contain a common column (for example, $m_1m_2m_3=m_4m_5m_6$); when the lengths of the two independent defining words are 3 and 4, their composition is that they contain a common column (for example, $m_1m_2m_3=m_3m_4m_5m_6$). In situation (a4), if the length of one of the independent defining words exceeds 3, then it is impossible to have two other independent defining words with lengths ≤3. Therefore, when W contains three independent defining words, the length of the independent defining words can only be 3. Since 6 factors form 3 independent defining words, there must be a common column between every two independent defining words. Let m_1,m_2,m_3 be the common column factors between the three independent defining words, $*$ represents an unknown factor, then the relationship between the three independent defining words can be obtained as: $m_1*m_3=m_1*m_2=m_2*m_3$, such arrangement can supplement the $*$ place with the other three factors. Since these three independent defining words will generate a non-independent defining word with a length of 3, so when $t=3$, W cannot contain four independent defining words.

(ii) When $t=4$, $W=\{m_1,m_2,m_3,m_4,m_5,m_6,m_7\}$, $W\in\Omega_7(D)$, then the number of independent defining words contained in W is as follows:

(a1) W does not contain independent defining words;

(a2) W contains one independent defining word, the length of which may be 4 or 5 or 6 or 7;

(a3) W contains two independent defining words, their lengths may be 4 and 4 or 4 and 5.

Situations (a1) and (a2) are obvious. In situation (a3), the independent defining words with lengths 4 and 4 may have one common column or two common columns. The independent defining words with lengths 4 and 5 have two common columns. Consider whether W may contain three independent defining words. Suppose W contains three independent defining words, all of which have a length of 4, and two of the independent defining words have two common columns, denoted as m_3m_4, and there is the following defining relationship, that is, $m_1m_2m_3m_4=m_3m_4m_5m_6$, which can generate a non-independent defining word

$m_1m_2m_5m_6$, so that the other independent defining word has at most two common columns with these three defining words, thus two additional factors need to be added, exceeding the range of seven factors. Therefore, when $t=4$, W contains at most two independent defining words. The cases of $t=5$ and $t=6$ can be proved in the same way.

(iii) When $t=7$, W cannot have two independent defining words. Selecting two independent defining words with a length of 7 from 10 factors will inevitably obtain a non-independent defining word with a length ≤ 6, which does not conform to the condition of resolution 7, so when the resolution is 7, W contains at most one independent defining word. The cases of $t>7$ can be proved in the same way.

Lemma 2. Consider a regular matrix D, which is an orthogonal matrix of strength 2. If this design is used to create a baseline design, then the values of T_1 and T_2 in the expression satisfy the following conditions:

(i) When $v \leq t-2$, $T_1 = C_m^v \left(\frac{N}{2^v}\right)^2$, $T_2 = 0$

(ii) When $v = t-1$, $T_1 = C_m^{t-1}\left(\frac{N}{2^{t-1}}\right)^2$, $T_2 = t\left(\frac{N^2}{2^{2t-2}}\right)B_t$

(iii) When $v=t$ and t is odd

$$T_1 = \frac{N^2}{2^{2t}}\left(C_m^t - A_t^0 + 3A_t^1\right)$$

$$T_2 = \frac{N^2}{2^{2t}}\left((t+1)A_{t+1} + t(m-t)A_t^0 + t(m-t)A_t^1\right)$$

(iv) When $v=t$ and t is even

$$T_1 = \frac{N^2}{2^{2t}} + \left(C_m^t + 3A_t^0 - A_t^1\right)$$

$$T_2 = \frac{N^2}{2^{2t}}\left((t+1)A_{t+1} + t(m-t)A_t^0 + t(m-t)A_t^1\right)$$

Proof:

(i) For any regular two-level design, let W be the submatrix composed of any v columns of the design matrix. Let t be the resolution of the design, then when $v \leq t-1$, $\alpha(w) = N/2^v$. Substituting it into the expressions of T_1 and T_2, we can get $T_1 = C_m^v\left(\frac{N}{2^v}\right)^2$, $T_2 = \sum_{w^* \in \Omega^{v+1}} \sum_{w^0 \in \Omega^v(w^*)} \left(\frac{2N}{2^{v+1}} - \frac{N}{2^v}\right)^2 = 0$, proved.

(ii) When $v=t-1$, $\alpha(w)$ still satisfies the above condition, that is, $\alpha(w) = N/2^{t-1}$, so $T_1 = C_m^{t-1}\left(\frac{N}{2^{t-1}}\right)^2$.

But for T_2, $w^* \in \Omega^{v+1}$ that is, $w^* \in \Omega^t$, then it is necessary to consider whether this t columns contain defining words. Let $J_k(w) = \left|2\Phi(w) - N\right|$, where $\Phi(w)$ represents the value of the sum of the rows of the W matrix modulo 2 and then added together. When W is a non-defining word matrix, $\Phi(w) = N/2$, $J_k(w) = 0$. When W is a defining word matrix, $\Phi(w) = 0$ or

N, $J_k(w)=N$. Therefore, when this t columns do not contain defining words, $J_k(w)=0$, $\alpha(w^*)=N/2^t$, $\alpha(w^0)=N/2^{t-1}$, it can be concluded that $T_2=0$ when this t columns contain defining words, $J_k(w)=N$. Among them, this submatrix contains $(N-J_t(w))/2^t$ 2^t designs and $(J_t(w))/2^t$ half of the 2^t designs, that is, 0 2^t designs and $N/2^{t-1}$ half of the 2^t designs. Whether the $N/2^{t-1}$ half of the 2^t designs contain all 1 rows needs to consider the parity of t and the ϕ value of the defining words. The following four situations are divided:

(a1) When t is odd, $\phi(w)=0_N$, that is, there is no all 1 row, $\alpha(w^*)=0$, $\alpha(w^0)=N/2^{t-1}$, $2\alpha(w^*)-\alpha(w^0)=-N/2^{t-1}$;

(a2) When t is odd, $\phi(w)=1_N$, that is, there is an all 1 row, $\alpha(w^*)=N/2^{t-1}$, $\alpha(w^0)=N/2^{t-1}$, $2\alpha(w^*)-\alpha(w^0)=N/2^{t-1}$;

(a3) When t is even, $\phi(w)=0_N$, that is, there is an all 1 row, $\alpha(w^*)=N/2^{t-1}$, $\alpha(w^0)=N/2^{t-1}$, $2\alpha(w^*)-\alpha(w^0)=N/2^{t-1}$;

(a4) When t is even, $\phi(w)=1_N$, that is, there is no all 1 row, $\alpha(w^*)=0$, $\alpha(w^0)=N/2^{t-1}$, $2\alpha(w^*)-\alpha(w^0)=-N/2^{t-1}$.

The $\left|2\alpha(w^*)-\alpha(w^0)\right|$ obtained in the four situations are all $N/2^{t-1}$, so we can get:

$$\begin{aligned}
T_2 &= \sum_{w^*\in\Omega^t}\sum_{w^0\in\Omega^{t-1}(w^*)}\frac{N^2}{2^{2t-2}} \\
&= \sum_{w^*\in\Omega^t} t\frac{N^2}{2^{2t-2}} \\
&= t\frac{N^2}{2^{2t-2}}\sum_{w^*\in\Omega^t}\left(\frac{J_t(w^*)}{N}\right)^2 \\
&= t\frac{N^2}{2^{2t-2}}B_t
\end{aligned}$$

(iii) The calculation of T_1 needs to consider three situations:

Table 1. Calculation situations of T_1.

Defining word situation	$\alpha(W)$	n_W
This t columns do not contain any defining words	$\frac{N}{2^t}$	$\binom{m}{t}-A_t^0-A_t^1$
This t columns form defining words and $\phi=0_N$	0	A_t^0
This t columns form defining words and $\phi=1_N$	$\frac{N}{2^{t-1}}$	A_t^1

n_W: The number of W in each situation.

According to **Table 1**, the calculation result of T_1 can be obtained.

$$T_1 = \left[\binom{m}{t} - A_t^0 - A_t^1\right] \cdot \left(\frac{N}{2^t}\right)^2 + 0 \cdot A_t^0 + \left(\frac{N}{2^{t-1}}\right)^2 \cdot A_t^1$$
$$= \frac{N^2}{2^{2t}}\left(C_m^t - A_t^0 + 3A_t^1\right)$$

The calculation of T_2 needs to consider four situations:

(a1) W^* is a defining word with a length of $t+1$ and $\phi = 0_N$;

(a2) W^* is a defining word with a length of $t+1$ and $\phi = 1_N$;

(a3) W^* contains a defining word with a length of t and $\phi = 0_N$;

(a4) W^* contains a defining word with a length of t and $\phi = 1_N$;

Table 2. Calculation situations of T_2.

Situation	$\alpha(W^*)$	$\alpha(W^0)$: n_{W^0}	n_W
(a1)	$\frac{N}{2^t}$	$\frac{N}{2^t}:t+1$	A_{t+1}^0
(a2)	0	$\frac{N}{2^t}:t+1$	A_{t+1}^1
(a3)	0	$0:1$ and $\frac{N}{2^t}:t$	$(m-t)A_t^0$
(a4)	$\frac{N}{2^t}$	$\frac{N}{2^{t-1}}:1$ and $\frac{N}{2^t}:t$	$(m-t)A_t^1$

n_{W^*}: The number of each $\alpha(W^0)$ value corresponding to W^*. n_W: The number of W^* in (a1)-(a4).

According to Table 2, the calculation result of T_2 can be obtained.

$$T_2 = \frac{N^2}{2^{2t}}\left((t+1)A_{t+1} + t(m-t)A_t^0 + t(m-t)A_t^1\right)$$

(iii) is proved, and the result of (iv) can be obtained in the same way.

Lemma 3. The ϕ value of the defining word with a length of 5 obtained by the combination of independent defining words with lengths 3 and 4 containing a common column is determined. Let the ϕ values of the above defining words with lengths 3 and 4 be ϕ_1 and ϕ_2 respectively, and the ϕ value of the defining word with a length of 5 obtained by the combination be ϕ_3, the following results can be obtained:

(i) If ϕ_1 and ϕ_2 are both 0_N, then ϕ_3 is 0_N;

(ii) If ϕ_1 and ϕ_2 are both 1_N, then ϕ_3 is 0_N;

(iii) If one of ϕ_1 and ϕ_2 is 0_N and the other is 1_N, then ϕ_3 is 1_N.

Proof:

Let $W = \{m_1, m_2, m_3, m_4, m_5, m_6\}$, the defining relationship is $I = m_1m_2m_3 = m_3m_4m_5m_6 = m_1m_2m_4m_5m_6$, let $W_1 = \{m_1, m_2, m_3\}$, $W_2 = \{m_3, m_4, m_5, m_6\}$, where the three factors in W_1 form a defining word with a length of 3, $m_1m_2m_3$, the four factors in W_2 form a defining word with a length of 4, $m_3m_4m_5m_6$, their common column is m_3.

$\phi_1 = 0_N$ indicates that there are an even number of 1 s in each row of the three columns m_1, m_2, m_3, $\phi_2 = 0_N$ indicates that there are an even number of 1 s in each row of the three columns m_3, m_4, m_5, m_6. When the column m_3 is 1, the corresponding two columns m_1, m_2 have an odd number of 1 s, and m_4, m_5, m_6 have an even number of 1 s. When the column m_3 is 0, the corresponding two columns m_1, m_2 have an even number of 1 s, and m_4, m_5, m_6 have an even number of 1 s, that is, m_1, m_2, m_4, m_5, m_6 have an even number of 1 s, so when ϕ_1 and ϕ_2 are both 0_N, then ϕ_3 is 0_N. Thus, (i) is proved, and (ii) and (iii) can be proved in the same way.

2.2. The Relationship between K_5 and WLP When the Resolution $t = 3$

This section mainly studies the relationship between K_5 and WLP in regular 2^{m-p} designs with resolution $t = 3$. By discussing the defining words contained in $W \in \Omega_5(D)$ and $W \in \Omega_6(D)$, the analysis process of T_1 and T_2 is given, and finally the relationship between K_5 and WLP in regular 2^{m-p} designs with resolution $t = 3$ is obtained.

Theorem 1. Suppose D is a regular 2^{m-p} design with resolution $t = 3$, then

$$K_5 = \frac{4}{N^2}(5T_1 + T_2)$$

where

$$\begin{aligned} T_1 &= \sum_{w \in \Omega^5} (\alpha(w))^2 \\ &= \left(\frac{N}{8}\right)^2 \cdot A_3^{1,1} + \left(\frac{N}{16}\right)^2 \cdot \left[\binom{m-3}{2} A_3^1 + (m-4) A_4^0 - A_3^{0,0} - 3A_3^{1,1} - A_3^{0,1} + A_5^1\right] \\ &\quad + \left(\frac{N}{32}\right)^2 \cdot I_1^5 \end{aligned}$$

$$\begin{aligned} T_2 = \left(\frac{N}{2^5}\right)^2 & \left[6\left(A_6^0 + A_6^1\right) + 5(m-5)\left(A_5^0 + A_5^1\right)\right] \\ & + 4\binom{m-4}{2}\left(A_4^0 + A_4^1\right) + 3\binom{m-3}{3}\left(A_3^0 + A_3^1\right) \\ & - 10A_3^{0,0} - 2A_3^{0,1} + 10A_3^{1,1} + 6\left(\Delta A_3^{(0,0)} - \Delta A_3^{(0,1)} + \Delta A_3^{(1,1)}\right) \\ & - 21A_3^{(0,0,0)} - 21A_3^{(1,0,0)} - 5A_3^{(1,1,0)} + 75A_3^{(1,1,1)} - 8A_{(3,4)}^{(0,1)} + 12A_{(3,4)}^{(1,0)} \end{aligned}$$

Proof: First, calculate the expression of T_1, let $W = \{m_1, m_2, m_3, m_4, m_5\} \in \Omega_5(D)$, then there are ten possible situations for the columns in W:

(a1) W contains two independent defining words with a length of 3 and $\phi = 0_N$;

(a2) W contains two independent defining words with a length of 3 and $\phi = 1_N$;

(a3) W contains two independent defining words with a length of 3, one of which has $\phi = 0_N$ and the other has $\phi = 1_N$;

(a4) W only contains one defining word with a length of 3 and $\phi = 0_N$;

(a5) W only contains one defining word with a length of 3 and $\phi = 1_N$;

(a6) W only contains one defining word with a length of 4 and $\phi = 0_N$;

(a7) W only contains one defining word with a length of 4 and $\phi = 1_N$;

(a8) W only contains one defining word with a length of 5 and $\phi = 0_N$;

(a9) W only contains one defining word with a length of 5 and $\phi = 1_N$;

(a10) W contains 5 independent columns.

The reason why (a1), (a2), (a3) appear independently from other situations is that two defining words can combine to form other situations, and to avoid repeated calculations, they are listed separately. The defining words in (a1) and (a2) can combine to form an independent defining word with a length of 4 and $\phi = 0_N$. The defining words in (a3) can combine to form an independent defining word with a length of 4 and $\phi = 1_N$.

Below are the proofs of the values of $\alpha(w)$ and the number of each W in situations (a1)-(a10).

For (a1), since there exists a defining word with a length of 3 and $\phi = 0_N$, there cannot be an all-one row in W, $\alpha(w)$ is 0, the number is $A_3^{0,0}$.

For (a2), two independent defining words with a length of 3 and $\phi = 1_N$ can combine to form an independent defining word with a length of 4 and $\phi = 0_N$. The cross column of the two independent defining words is affected by the other two columns of the defining words. When the other two columns of the defining words are all 1, the cross column must be 1, so there is an all-one row in this W, $\alpha(w)$ is $N/8$, the number is $A_3^{1,1}$.

For (a3) and (a4), the situation is the same as (a1), $\alpha(w)$, the numbers are $A_3^{0,1}$ and $\binom{m-3}{2}A_3^0 - 2A_3^{0,0} - A_3^{0,1}$ respectively. In situation (a4), each W has and only has one defining word with a length of 3 and $\phi = 0_N$. The number of such defining words is A_3^0, $W \in \Omega_5(D)$, that is, selecting 5 factors from m factors that meet the conditions of (a4), which is $\binom{m-3}{2}A_3^0$. To avoid the appearance of repeated situations, it is necessary to remove the situations where $\phi = 0_N$ in (a1) and (a2). Since (a1) contains two $\phi = 0_N$ defining words in W, the final number in situation (a4) is $\binom{m-3}{2}A_3^0 - 2A_3^{0,0} - A_3^{0,1}$.

For (a5), each W^* contains a defining word with a length of 3 and $\phi = 1_N$, so the number of all-one rows in the defining word, combined with the other two independent columns, is $\frac{N}{16}$. In this situation, each W^* contains 6 W^0, two of which are combinations of the defining word and one independent column, their $\alpha(W^0)$ is $\frac{N}{16}$, and the other four W^0 are independent columns, their

$\alpha(W^0)$ is $\frac{N}{32}$. The number is based on $\binom{m-3}{2}A_3^1$ and needs to remove the non-independent defining words with a length of 5 and $\phi=0_N$ generated in (a9), (a10), (a15), (a16), (a18). Thus, the number is $\binom{m-3}{2}A_3^1-2A_3^{1,1}-A_3^{0,1}$.

For (a6), each W contains a defining word with a length of 4 and $\phi=0_N$, so the number of all-one rows in the defining word is $N/8$, and the number of all-one rows combined with another independent column is $N/16$, that is, $\alpha(w)$ is $N/16$, the number is $(m-4)A_4^0-2A_3^{0,0}-A_3^{1,1}$.

For (a7), since there exists a defining word with a length of 4 and $\phi=1_N$, there cannot be an all-one row in W, $\alpha(w)$ is 0, the number is $(m-4)A_4^1-A_3^{0,1}$.

For (a8), since there exists a defining word with a length of 5 and $\phi=0_N$, there cannot be an all-one row in W, $\alpha(w)$ is 0, the number is A_5^0.

For (a9), W contains an all-one row, five factors form a defining word, so $\alpha(w)$ is $N/16$, the number is A_5^1.

For (a10), since W contains 5 independent columns, $\alpha(w)$ is $N/32$. Selecting 5 factors from m factors has C_m^5 possibilities, removing all previous situations, the number is

$\binom{m}{5}+2A_3^{0,0}+2A_3^{1,1}+2A_3^{0,1}-\binom{m-3}{2}(A_3^0+A_3^1)-(m-4)A_4^1-A_5^0-A_5^1$, denoted as I_1^5.

Summarizing the above situations, we obtain Table 3:

Table 3. Calculation of T_1 in Theorem 1.

Situation	$\alpha(W)$	n_W
(a1)	0	$A_3^{0,0}$
(a2)	$N/8$	$A_3^{1,1}$
(a3)	0	$A_3^{0,1}$
(a4)	0	$\binom{m-3}{2}A_3^0-2A_3^{0,0}-A_3^{0,1}$
(a5)	$N/16$	$\binom{m-3}{2}A_3^1-2A_3^{1,1}-A_3^{0,1}$
(a6)	$N/16$	$(m-4)A_4^0-A_3^{0,0}-A_3^{1,1}$
(a7)	0	$(m-4)A_4^1-A_3^{0,1}$
(a8)	0	A_5^0
(a9)	$N/16$	A_5^1
(a10)	$N/32$	I_1^5

n_W: The number of W in each situation.

Thus, the expression of T_1 is obtained.

Next, calculate the expression of T_2. When $t=3$, $W\in\Omega_6(D)$ has the following twenty situations:

(a1) W contains only one defining word with a length of 6 and $\phi = 0_N$;

(a2) W contains only one defining word with a length of 6 and $\phi = 1_N$;

(a3) W contains only one defining word with a length of 5 and $\phi = 0_N$;

(a4) W contains only one defining word with a length of 5 and $\phi = 1_N$;

(a5) W contains only one defining word with a length of 4 and $\phi = 0_N$;

(a6) W contains only one defining word with a length of 4 and $\phi = 1_N$;

(a7) W contains only one defining word with a length of 3 and $\phi = 0_N$;

(a8) W contains only one defining word with a length of 3 and $\phi = 1_N$;

(a9) W contains two independent defining words with a length of 3 and $\phi = 0_N$, sharing one common column;

(a10) W contains two independent defining words with a length of 3 and $\phi = 1_N$, sharing one common column;

(a11) W contains two independent defining words with a length of 3, one with $\phi = 0_N$ and the other with $\phi = 1_N$, sharing one common column;

(a12) W contains two independent defining words with a length of 3 and $\phi = 0_N$, without sharing any common column;

(a13) W contains two independent defining words with a length of 3 and $\phi = 1_N$, without sharing any common column;

(a14) W contains two independent defining words with a length of 3, one with $\phi = 0_N$ and the other with $\phi = 1_N$, without sharing any common column;

(a15) W contains three independent defining words with a length of 3, each pair sharing one common column, divided into the following four cases:

(1) W does not contain any defining word with $\phi = 1_N$;

(2) W contains only one defining word with $\phi = 1_N$;

(3) W contains only two defining words with $\phi = 1_N$;

(4) W contains only three defining words with $\phi = 1_N$.

(a16) W contains one independent defining word with a length of 3 and $\phi = 0_N$ and one independent defining word with a length of 4 and $\phi = 0_N$, sharing one common column;

(a17) W contains one independent defining word with a length of 3 and $\phi = 0_N$ and one independent defining word with a length of 4 and $\phi = 1_N$, sharing one common column;

(a18) W contains one independent defining word with a length of 3 and $\phi = 1_N$ and one independent defining word with a length of 4 and $\phi = 0_N$, sharing one common column;

(a19) W contains one independent defining word with a length of 3 and $\phi = 1_N$ and one independent defining word with a length of 4 and $\phi = 1_N$, sharing one common column;

(a20) W contains 6 independent columns.

Next, using Lemma 2's calculation method, the values of $\alpha(w_*)$ and $\alpha(w_0)$ and the number of each W in situations (a1)-(a20) are proved.

For (a1), each W^* contains a defining word with a length of 6 and $\phi = 0_N$, so the number of all-one rows in the defining word is $\frac{N}{2^5}$. In this situation, each

W^* contains 6 W^0 which are all independent columns, that is, $\alpha\left(W^0\right)$ is $\frac{N}{2^5}$. The number is based on A_6^0 and needs to remove the non-independent defining words with a length of 6 and $\phi=0_N$ generated by two independent defining words with a length of 3 in (a12) and (a14). Thus, the number is $A_6^0-\Delta A_3^{(0,0)}-\Delta A_3^{(1,1)}$;

For (a2), since there exists a defining word with a length of 6 and $\phi=0_N$, there cannot be an all-one row in W^*, $\alpha\left(W_*\right)$ is 0. In this situation, each W^* contains 6 W^0 which are all independent columns, that is, $\alpha\left(W^0\right)$ is $\frac{N}{2^5}$. The number is based on A_6^1 and needs to remove the non-independent defining words with a length of 6 and $\phi=1_N$ generated by two independent defining words with a length of 3 in (a13). Thus, the number is $A_6^0-\Delta A_3^{(0,1)}$;

For (a3), since there exists a defining word with a length of 5 and $\phi=0_N$, there cannot be an all-one row in W^*, $\alpha\left(W^*\right)$ is 0. In this situation, each W_* contains 6 w_0, one of which is a defining word, its $\alpha\left(W^0\right)$ is 0, and the other five w^0 are independent columns, their $\alpha\left(W^0\right)$ is $\frac{N}{2^5}$. The number is based on $\left(m-5\right)A_5^0$ and needs to remove the non-independent defining words with a length of 5 and $\phi=0_N$ generated by two independent defining words with lengths of 3 and 4 in (a16) and (a19). Thus, the number is $\left(m-5\right)A_5^0-A_{(3,4)}^{(0,0)}-A_{(3,4)}^{(1,1)}$;

For (a4), each W^* contains a defining word with a length of 5 and $\phi=1_N$, so the number of all-one rows in the defining word, combined with another independent column, is $\frac{N}{2^5}$. In this situation, each W^* contains 6 W^0, one of which is a defining word, its $\alpha\left(W^0\right)$ is $\frac{N}{2^4}$, and the other five W^0 are independent columns, their $\alpha\left(W^0\right)$ is $\frac{N}{2^5}$. The number is based on $\left(m-5\right)A_5^1$ and needs to remove the non-independent defining words with a length of 5 and $\phi=1_N$ generated by two independent defining words with lengths of 3 and 4 in (a16) and (a19). Thus, the number is $\left(m-5\right)A_5^1-A_{(3,4)}^{(0,1)}-A_{(3,4)}^{(1,0)}$;

For (a5), each W^* contains a defining word with a length of 4 and $\phi=0_N$, so the number of all-one rows in the defining word, combined with the other two independent columns, is $\frac{N}{2^5}$. In this situation, each W^* contains 6 W^0, two of which are combinations of defining words with a length of 4 and one independent column, their $\alpha\left(W^0\right)$ is $\frac{N}{2^4}$, and the other four W^0 are independent columns, their $\alpha\left(W^0\right)$ is $\frac{N}{2^5}$. The number is based on $\left(m-5\right)A_5^1$ and needs to remove the non-independent defining words with a length of 4 and $\phi=0_N$ generated in (a9), (a10), (a15), (a16), (a18). Thus, the number is $\binom{m-4}{2}A_4^0-A_{(3,4)}^{(0,0)}-A_{(3,4)}^{(1,0)}-3A_3^{(0,0,0)}-A_3^{(1,0,0)}-A_3^{(1,1,0)}-3A_3^{(1,1,1)}-A_3^{0,0}-A_3^{1,1}$, denoted as N_{a5};

For (a6), since there exists a defining word with a length of 4 and $\phi = 1_N$, there cannot be an all-one row in W^*, $\alpha\left(W^*\right)$ is 0. In this situation, each W^* contains 6 W^0, two of which are combinations of defining words with a length of 4 and one independent column, their $\alpha\left(W^0\right)$ is 0, and the other four W^0 are independent columns, their $\alpha\left(W^0\right)$ is $\frac{N}{2^5}$. The number is similar to the analysis in (a5), which is $\binom{m-4}{2}A_4^1 - A_{(3,4)}^{(0,1)} - A_{(3,4)}^{(1,1)} - 2A_3^{(1,0,0)} - 2A_3^{(1,1,0)} - A_3^{0,1}$, denoted as N_{a6};

For (a7), since there exists a defining word with a length of 3 and $\phi = 0_N$, there cannot be an all-one row in W^*, $\alpha\left(W^*\right)$ is 0. In this situation, each W^* contains 6 W^0, three of which are combinations of defining words with a length of 3 and two independent columns, their $\alpha\left(W^0\right)$ is 0, and the other three W^0 are independent columns, their $\alpha\left(W^0\right)$ is $\frac{N}{2^5}$. The number needs to remove (a9), (a11), (a12), (a14), (a15), (a16), (a17), which is

$$\binom{m-3}{3}A_3^0 - 2A_3^{0,0} - A_3^{0,1} - 2\Delta A_3^{(0,0)} - \Delta A_3^{(0,1)} - 3A_3^{(0,0,0)} - 2A_3^{(1,0,0)} - A_3^{(1,1,0)} - A_{(3,4)}^{(0,0)} - A_{(3,4)}^{(0,1)},$$

denoted as N_{a7};

For (a8), each W^* contains a defining word with a length of 3 and $\phi = 1_N$, so the number of all-one rows in the defining word, combined with the other three independent columns, is $\frac{N}{2^5}$. In this situation, each W^* contains 6 W^0, three of which are combinations of defining words with a length of 3 and two independent columns, their $\alpha\left(W^0\right)$ is $\frac{N}{2^4}$, and the other three W^0 are independent columns, their $\alpha\left(W^0\right)$ is $\frac{N}{2^5}$. The number is similar to the situation of (a8), which is

$$\binom{m-3}{3}A_3^1 - A_3^{0,1} - 2A_3^{1,1} - 2\Delta A_3^{(1,1)} - \Delta A_3^{(0,1)} - A_3^{(1,0,0)} - 2A_3^{(1,1,0)} - 3A_3^{(0,0,0)} - A_{(3,4)}^{(1,0)} - A_{(3,4)}^{(1,1)},$$

denoted as N_{a8};

For (a9), since there exists a defining word with a length of 3 and $\phi = 0_N$, there cannot be an all-one row in W^*, $\alpha\left(W^*\right)$ is 0. In this situation, each W^* contains 6 W^0, their $\alpha\left(W^0\right)$ are all 0, and there is no all-one column in the case without common columns. The number is $A_3^{0,0}$;

For (a10), W^* contains an all-one row, $\alpha\left(W^*\right)$ is $\frac{N}{2^4}$. In this situation, each W^* contains 6 W^0, the $\alpha\left(W^0\right)$ in the case without common columns is $\frac{N}{8}$, and the $\alpha\left(W^0\right)$ in the other five cases is $\frac{N}{2^4}$. The number is $A_3^{1,1}$;

For (a11), since there exists a defining word with a length of 3 and $\phi = 0_N$, there cannot be an all-one row in W^*, $\alpha\left(W^*\right)$ is 0. In this situation, each W^* contains 6 W^0, two of which have $\alpha\left(W^0\right)$ as $\frac{N}{2^4}$, and four have $\alpha\left(W^0\right)$ as

0. The number is $A_3^{0,1}$;

For (a12), since there exists a defining word with a length of 3 and $\phi = 0_N$, there cannot be an all-one row in W^*, $\alpha(W^*)$ is 0. In this situation, each W^* contains 6 W^0 which are all 0. The number is $\Delta A_3^{(0,0)}$;

For (a13), W^* contains an all-one row, $\alpha(W^*)$ is $\frac{N}{2^5}$. In this situation, each W^* contains 6 W^0 which are all $\frac{N}{2^4}$. The number is $\Delta A_3^{(1,1)}$;

For (a14), since there exists a defining word with a length of 3 and $\phi = 0_N$, there cannot be an all-one row in W^*, $\alpha(W^*)$ is 0. In this situation, each W^* contains 6 W^0, three of which are 0, and the other three are $\frac{N}{2^4}$. The number is $\Delta A_3^{(0,1)}$;

For (a15), three independent defining words with a length of 3 will generate a non-independent defining word with a length of 3, the ϕ value of which is determined by the ϕ values of the three independent defining words. As long as there exists a defining word with a length of 3 and $\phi = 0_N$, there cannot be an all-one row in W^*, so the $\alpha(W^*)$ values of situations (1), (2), and (3) are all 0, and the $\alpha(W^*)$ value of situation (4) is $\frac{N}{8}$. When there are two or more defining words with $\phi = 0_N$, each W^* contains 6 W^0 which are all 0, so the $\alpha(W^0)$ of situations (1) and (2) are all 0. In situation (3), each W^* contains 6 W^0, one of which is $\frac{N}{8}$, and the other five are 0. In situation (4), each W^* contains 6 W^0 whose $\alpha(W^0)$ are all $\frac{N}{8}$. Their numbers are denoted as $A_3^{(0,0,0)}$, $A_3^{(1,0,0)}$, $A_3^{(1,1,0)}$, $A_3^{(1,1,1)}$ respectively;

For (a16), since there exists a defining word with a length of 3 and $\phi = 0_N$, there cannot be an all-one row in W^*, $\alpha(W^*)$ is 0. In this situation, each W^* contains 6 W^0, three of which are 0, and the other three are $\frac{N}{2^4}$. The number is $A_{(3,4)}^{(0,0)}$;

For (a17), since there exists a defining word with a length of 3 and $\phi = 0_N$, there cannot be an all-one row in W^*, $\alpha(W^*)$ is 0. In this situation, each W^* contains 6 W^0 whose $\alpha(W^0)$ are all $\frac{N}{2^4}$. The number is $A_{(3,4)}^{(0,1)}$;

For (a18), according to Lemma 3, a defining word with a length of 3 and $\phi = 1_N$ and a defining word with a length of 4 and $\phi = 0_N$ will generate a defining word with a length of 5 and $\phi = 1_N$, $\alpha(W^*)$ is $\frac{N}{2^4}$. In this situation, each W^* contains 6 W^0, five of which are 0, and the other one is $\frac{N}{2^4}$. The number is $A_{(3,4)}^{(1,0)}$;

For (a19), since there exists a defining word with a length of 4 and $\phi = 1_N$, there cannot be an all-one row in W^*, $\alpha(W^*)$ is 0. In this situation, each W^* contains 6 W^0, three of which are 0, and the other three are $\frac{N}{2^4}$. The number is $A_{(3,4)}^{(1,1)}$;

Table 4. Calculation of T_2 in Theorem 1.

Situation	$\alpha(W)$	$\alpha(W^*)$: n_{W^*}	n_W
(a1)	$\frac{N}{2^5}$	$\frac{N}{2^5}:6$	$A_6^0-\Delta A_3^{(0,0)}-\Delta A_3^{(1,1)}$
(a2)	0	$\frac{N}{2^5}:6$	$A_6^0-\Delta A_3^{(0,1)}$
(a3)	0	$\frac{N}{2^5}:5$ and $0:1$	$(m-5)A_5^0-A_{(3,4)}^{(0,0)}-A_{(3,4)}^{(1,1)}$
(a4)	$\frac{N}{2^5}$	$\frac{N}{2^5}:5$ and $\frac{N}{2^4}:1$	$(m-5)A_5^1-A_{(3,4)}^{(0,1)}-A_{(3,4)}^{(1,0)}$
(a5)	$\frac{N}{2^5}$	$\frac{N}{2^5}:4$ and $\frac{N}{2^4}:2$	N_{a5}
(a6)	0	$\frac{N}{2^5}:4$ and $0:2$	N_{a6}
(a7)	0	$\frac{N}{2^5}:3$ and $0:3$	N_{a7}
(a8)	$\frac{N}{2^5}$	$\frac{N}{2^5}:3$ and $\frac{N}{2^4}:3$	N_{a8}
(a9)	0	$0:6$	$A_3^{0,0}$
(a10)	$\frac{N}{2^4}$	$\frac{N}{8}:1$ and $\frac{N}{2^4}:5$	$A_3^{1,1}$
(a11)	0	$\frac{N}{2^4}:2$ and $0:4$	$A_3^{0,1}$
(a12)	0	$0:6$	$\Delta A_3^{(0,0)}$
(a13)	$\frac{N}{2^5}$	$\frac{N}{2^4}:6$	$\Delta A_3^{(1,1)}$
(a14)	0	$\frac{N}{2^4}:3$ and $0:3$	$\Delta A_3^{(0,1)}$
(a15-1)	0	$0:6$	$A_3^{(0,0,0)}$
(a15-2)	0	$0:6$	$A_3^{(1,0,0)}$
(a15-3)	0	$\frac{N}{8}:1$ and $0:5$	$A_3^{(1,1,0)}$
(a15-4)	$\frac{N}{8}$	$\frac{N}{8}:6$	$A_3^{(1,1,1)}$
(a16)	0	$\frac{N}{2^4}:3$ and $0:3$	$A_{(3,4)}^{(0,0)}$
(a17)	0	$\frac{N}{2^4}:1$ and $0:5$	$A_{(3,4)}^{(0,1)}$
(a18)	$\frac{N}{2^4}$	$\frac{N}{2^4}:6$	$A_{(3,4)}^{(1,0)}$
(a19)	0	$\frac{N}{2^4}:3$ and $0:3$	$A_{(3,4)}^{(1,1)}$
(a20)	$\frac{N}{2^6}$	$\frac{N}{2^5}:6$	I_2^5

n_W: The number of W in each situation.

For (a20), since W^* contains 6 independent columns, $\alpha(W^*)$ is $\frac{N}{2^6}$. In this situation, each W^* contains 6 W^0 whose $\alpha\left(W^0\right)$ are all $\frac{N}{2^5}$. Selecting 6 factors from m factors has C_m^6 possibilities, removing all previous situations, the number is denoted as I_2^5.

Table 4 summarizes the above discussions.

According to Table 4, the calculation result of T_2 can be obtained.

The above completes the proof of Theorem 1.

3. Applications

To verify the effectiveness of the K_5 expression in orthogonal two-level fractional factorial designs, a resolution 3 experimental scheme can be designed to assess its computational advantages and practical application value through simulation studies.

Suppose we are studying a chemical reaction process and need to consider 9 factors (A,B,C,D,E,F,G,H,I) that affect the reaction efficiency. Due to the limited number of experimental runs, we choose a fractional factorial design, specifically a 2^{9-5} fractional factorial design, which means 9 factors and 2^{9-5} experimental runs. According to the design catalog proposed by Xu [6], we select one type of experiment with a word length pattern of (4, 14, 8, 0, 4, 1, 0), and its defining relations are as follows:

$$
\begin{aligned}
I &= ABE = ACF = ADG = AGI \\
&= BCDH = BCE = BDEG = CDEI = CDFG = BFGH = CFHI \\
&= BDFI = BCGI = DEFH = CEGH = BEHI = DGHI = EFGI \\
&= ABCDI = ACDEH = ABDFH = ABCGH \\
&= ADEFI = ACEGI = ABFGI = AEFGH \\
&= ABCDEFGH = ABDEGHI = ABCEFHI = ACDFGHI \\
&= BCDEFGH
\end{aligned}
$$

Based on the defining relations, the K expression is calculated, and the K value is minimized in sequence by using level permutation to seek the optimal design. According to Lemma 2, we can obtain $K_2 = 21$, where K_2 is a constant. $K_3 = \frac{1}{16}\left(3C_9^3 + 15A_3^0 + 27A_3^1\right)$, by which the ϕ values of the defining words with length 3 can be set to 0_N to minimize K_3. Assuming that K_4 has reached the minimum value, we next consider K_5. Using C to represent the constant, the expression for K_5 is obtained as $K_5 = \frac{25}{64}A_4^0 - \frac{25}{64}A_3^{0,0} - \frac{1}{4}A_{(3,4)}^{(0,1)} + C$, by which the ϕ values of the defining words with length 4 can be set to 1_N to minimize K_5. After verification, it can be concluded that this design is the optimal design under the K -aberration criterion.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Fisher, R.A. (1935) The Design of Experiments. Hafner Press.

[2] Wang, Y. and Fang, K.T. (1981) A Note on Uniform Distribution and Experimental Design. *A Monthly Journal of Science*, **26**, 485-489.

[3] Fang, K.T. (1980) The Uniform Design: Application of Number-Theoretic Methods in Experimental Design. *Acta Mathematicae Applicatae Sinica*, **3**, 363-372.

[4] Miller, A. and Tang, B. (2016) Using Regular Fractions of Two-Level Designs to Find Baseline Designs. *Statistica Sinica*, **26**, 745-759.

[5] Mukerjee, R. and Tang, B. (2011) Optimal Fractions of Two-Level Factorials under a Baseline Parameterization. *Biometrika*, **99**, 71-84. https://doi.org/10.1093/biomet/asr071

[6] Xu, H. (2009) Algorithmic Construction of Efficient Fractional Factorial Designs with Large Run Sizes. *Technometrics*, **51**, 262-277. https://doi.org/10.1198/tech.2009.07120