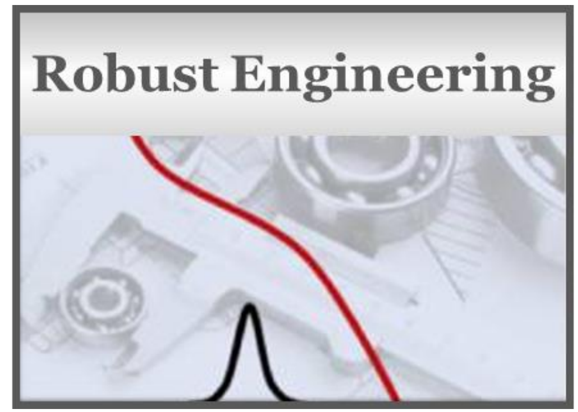
Module: Robust Engineering

Design of Experiments & Response Surface Modelling



Session 4: Planning a Two-Level Experiment

In This Session We Will…

- Describe the advantages of designed experiments, and some of the different types of experiment.
- Define a two-step strategy (screening + follow-up) that is useful in most engineering experimentation.
- Explain the concept of two-factor interaction and how it is represented in a twisted response surface.
- Discuss some designs that allow us to fit a ‘1^{st} order + twist’ equation.
- Define full factorial and fractional factorial designs.
- Show that regular fractions give us limited flexibility in choosing an efficient design and give an example of a useful non-regular fraction.

What Is The Advantage of A Designed Experiment?

- In Session 2 and 3 we showed that;
  - fitting a response surface to historical data can provide useful information.
  - the power of the analysis was compromised by the correlation between the x-variables.
- The purpose of 'DoE' is to structure our data collection carefully so that we get the maximum benefit from our data collection and analysis.

In the form of new knowledge about a process or product

We are aiming for maximum efficiency in our use of data

Different Types of Designed Experiment

- The first experiments with a mathematical structure were designed by R. A. Fisher at the Rothamsted agricultural research station in the 1920's.
  - these were comparative experiments, and the aim was usually to select the best variety of a crop (the one with the highest yield, or the most resistant to pests and disease).
- Fisher and his successor, F. Yates, introduced factorial (or multi-factor) experiments.
  - two or more factors are varied, using carefully selected combinations of levels.
  - the complete set of data is analysed to generate a statistical model which can be used (e.g.) to select the best combination of levels.

Different Types of Designed Experiment (cont.)

- The first factorial experiment in industry was probably run by L.H.C.Tippett in 1934.
- George Box, working at ICI around 1950, showed how multiple regression could be applied to the results of a factorial experiment to generate a response surface.
- Most experiments in engineering are factorial.

The experimenters do not always examine the response surface

… but it would usually add engineering value if they did!

Experimental Strategy

- Applications of DoE in engineering often use the sequential strategy, which is as follow:
  1. run a two-level experiment to measure trends and identify the most important factors (screening);
  2. run a multi-level follow-up experiment with a smaller number of factors, aiming to model the response surface accurately.

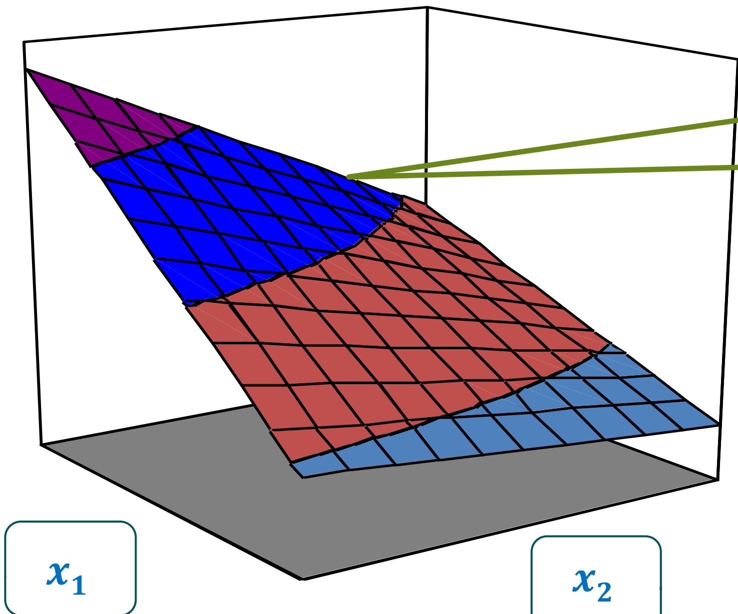
Experimental Strategy (cont.)

- Step 1 (screening) can itself be implemented in several different ways:
  - fit a 1st order response surface, ignoring interactions,
  - fit selected interactions,
  - fit a complete '1st order + twist' response surface.
- From a teaching/learning viewpoint, it is easiest to begin with the last of these.

In the rest of this session we focus on the last option

Interaction

- Here is a '1st order + twist' response surface in two factors



The effect of changing x_1 is quite different at different levels of x_2 (and vice versa)

The effect may even change sign if the twist is big enough

The technical term for twist is **two-factor interaction** (**2fi**)

Interaction (cont.)

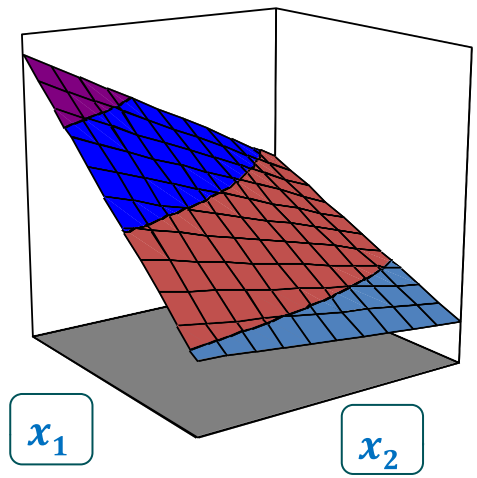
- The equation for this surface can be expressed like this:

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2$$

Think of this coefficient as ‘b 1,2’ not ‘b twelve’

This product term puts the twist into the surface

A larger value for b_{12} implies more twist



Measuring All 2fi's

- In this experimental strategy we look for a design which will allow us to fit an equation with the following terms:
  - an intercept,
  - a linear term in each factor,
  - a product (2fi) term in each pair of factors.
- For example, the equation for three factors is:

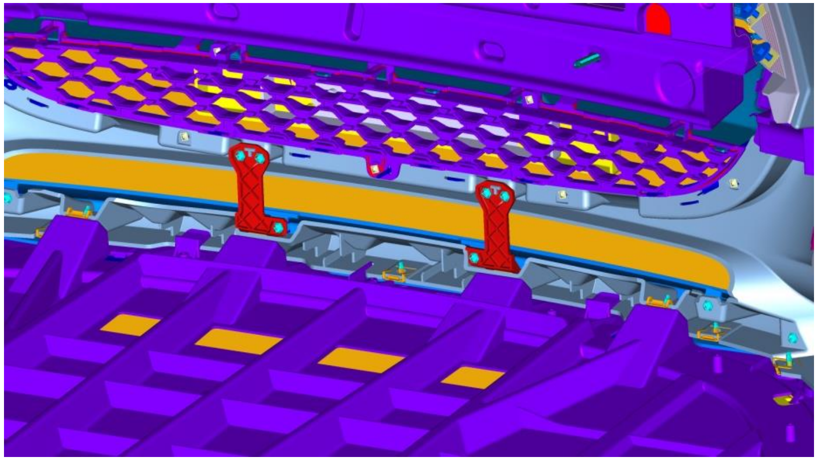
$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3$$

There is one well-known type of design which will allow us to fit this equation

The two-level full factorial

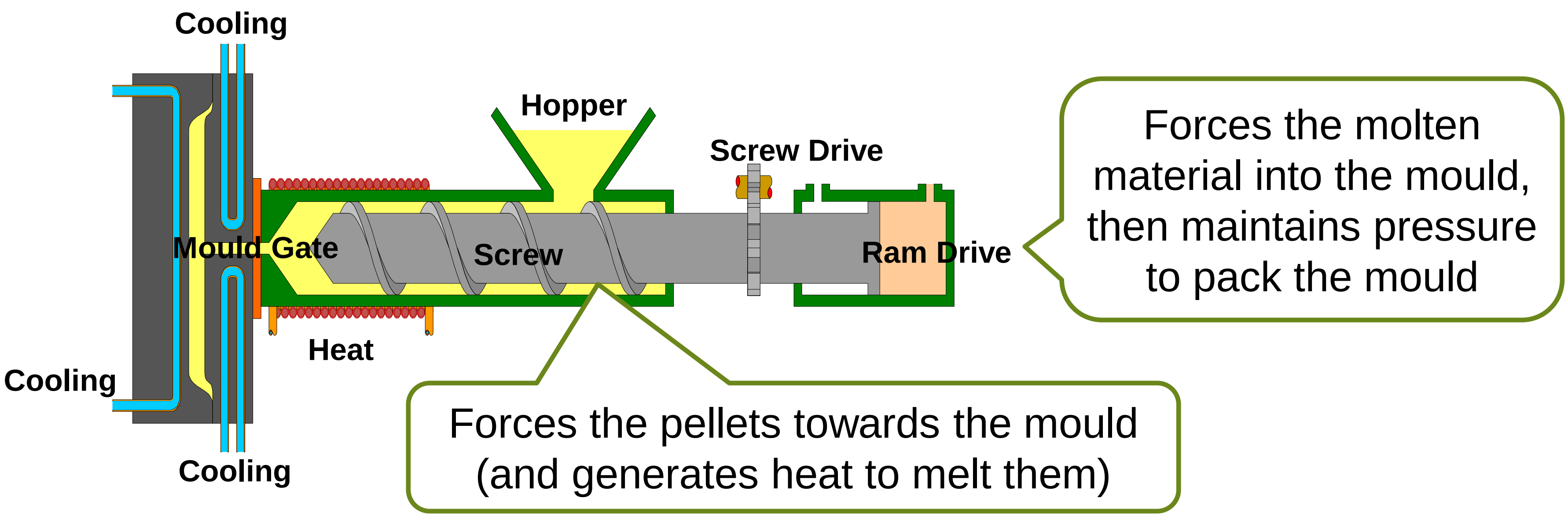
Example

- This example, and all the examples in the next 7 sessions, will relate to a study of potential alternative materials for bumper bracket mouldings.



Example

- The brackets are injection moulded (IM).
- In our examples, the factors are settings on the IM machine.
- For example, the factors in a three-factor full factorial might be the melt temperature, screw speed and holding pressure .



Factor Settings For A Full Factorial

We use ‘Row’ rather than ‘Run’ because the design need not be run in this order

We will discuss the response variable later

Row	Melt temperature (°C)	Screw speed (rpm)	Hold pressure (MPa)	Response
1	230	50	50	
2	270	50	50	
3	230	300	50	
4	270	300	50	
5	230	50	250	
6	270	50	250	
7	230	300	250	
8	270	300	250	

In a full factorial we run all possible combinations

Coding A Full Factorial

- As in Session 2, we can code the factor levels, using –1 for the lowest level, +1 for the highest:

Row	Melt temp (℃)	Screw speed (rpm)	Hold pressure (MPa)	x_1	x_2	x_3
1	230	50	50	–1	–1	–1
2	270	50	50	+1	–1	–1
3	230	300	50	–1	+1	–1
4	270	300	50	+1	+1	–1
5	230	50	250	–1	–1	+1
6	270	50	250	+1	–1	+1
7	230	300	250	–1	+1	+1
8	270	300	250	+1	+1	+1

We now use generic labels for the factors

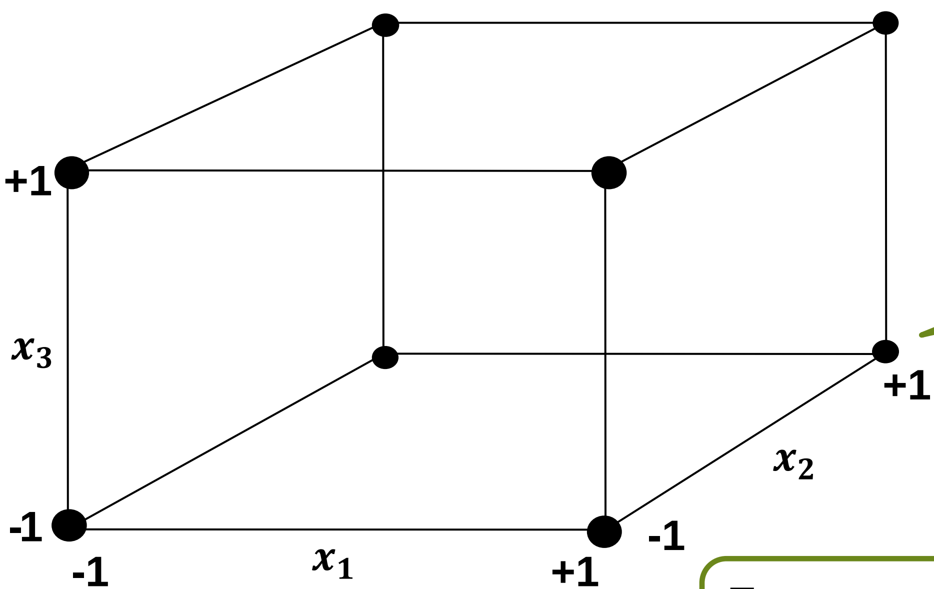
This is called the 2^3 design

i.e. $(\#\ levels)^{(\#\ factors)}$

This neat notation also tells us the no. of runs

The 2^3 Design In Factor Space

- Factor space is the multi-dimensional space in which the x -variables 'live'
  - it does not have a y-dimension.



In this case we have a 3D factor space

The 8 runs are at the corners of the cube

The symmetry of this pattern gives it good statistical properties

For example, the x-variables are uncorrelated ….

…. as we show on the next slide

Zero Correlations In The 2^3 Design

- Take, for example, x_1 and x_2;
- The squared Pearson correlation is:

$$r^2 = \frac{\{SPx_1x_2\}^2}{SSx_1 \cdot SSx_2}$$

We need to show that this sum-product is 0

$$\text{where } \mathrm{SPx_1x_2} = \sum_{i=1}^{n} (x_{1_i} - \bar{x}_1)(x_{2\mathrm{i}} - \bar{x}_2)$$

Row	x_1	x_2	x_3
1	–1	–1	–1
2	+1	–1	–1
3	–1	+1	–1
4	+1	+1	–1
5	–1	–1	+1
6	+1	–1	+1
7	–1	+1	+1
8	+1	+1	+1

- For the coded full factorial, the variables have zero means, so

$$\mathrm{SPx_1x_2} = \sum_{i=1}^{n} \mathrm{x_{1i}x_{2i}}$$

Sum-product of the x_1 and x_2 columns

…. and the sum-product is 0 for every pair of columns

Would The Full Factorial Meet Our Requirement?

- How do we know that the full factorial design will allow us to measure all the 2fi's, i.e. to fit this equation:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3\ ?$$

- To check this, we have to construct the model matrix M, as in Session 2.
- In this case we need 7 columns
  - the intercept column is all 1's, as usual
  - to get the columns for the linear terms we just copy the levels in the design columns
  - to get the columns for the 2fi terms we multiply the linear columns row by row

M has a row for every run and a column for every term in the equation we want to fit

The Model Matrix M

Intercept	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3
1	-1	-1	-1	1	1	1
1	1	-1	-1	-1	-1	1
1	-1	1	-1	-1	1	-1
1	1	1	-1	1	-1	-1
1	-1	-1	1	1	-1	-1
1	1	-1	1	-1	1	-1
1	-1	1	1	-1	-1	1
1	1	1	1	1	1	1

Column of 1's

These 3 columns are a copy of the coded design

To get the x_1x_2 column, multiply the x_1 and x_2 columns row by row

We usually omit the '+' when writing M out

e.g. $-1 \times -1 = +1$

Checking The Model Matrix

Intercept	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3
1	-1	-1	-1	1	1	1
1	1	-1	-1	-1	-1	1
1	-1	1	-1	-1	1	-1
1	1	1	-1	1	-1	-1
1	-1	-1	1	1	-1	-1
1	1	-1	1	-1	1	-1
1	-1	1	1	-1	-1	1
1	1	1	1	1	1	1

- Every pair of columns in M is orthogonal (sum-product 0).
- This implies that the full factorial design will allow us to measure all the 2fi's.

Mathematical note: orthogonality is a sufficient but not necessary condition

The necessary condition is: M'M must be a non-singular matrix

Fractional Factorials

- The number of runs in a two-level full factorial doubles with each extra factor
  - but do we really need all these runs?
  - could we fit the '1st order + twist' model with fewer runs?
  - or perhaps study more factors with the same number of runs?

A fractional factorial is a portion of a full factorial design

A regular fraction is ½, ¼, ⅛, … of a full factorial

Example: Five Factors, 16 Runs

Row	Melt temp (°C)	Screw speed (rpm)	Hold press (MPa)	Hold time (sec)	Injection rate (mm/sec)
1	230	50	50	15	300
2	270	50	50	15	50
3	230	300	50	15	50
4	270	300	50	15	300
5	230	50	250	15	50
6	270	50	250	15	300
7	230	300	250	15	300
8	270	300	250	15	50
9	230	50	50	35	50
10	270	50	50	35	300
11	230	300	50	35	300
12	270	300	50	35	50
13	230	50	250	35	300
14	270	50	250	35	50
15	230	300	250	35	50
16	270	300	250	35	300

Speed of the ram drive during injection of molten material

Half of the 2^5 full factorial

The Coded Five-factor Design

Row	x_1	x_2	x_3	x_4	x_5
1	–1	–1	–1	–1	+1
2	+1	–1	–1	–1	–1
3	–1	+1	–1	–1	–1
4	+1	+1	–1	–1	+1
5	–1	–1	+1	–1	–1
6	+1	–1	+1	–1	+1
7	–1	+1	+1	–1	+1
8	+1	+1	+1	–1	–1
9	–1	–1	–1	+1	–1
10	+1	–1	–1	+1	+1
11	–1	+1	–1	+1	+1
12	+1	+1	–1	+1	–1
13	–1	–1	+1	+1	+1
14	+1	–1	+1	+1	–1
15	–1	+1	+1	+1	–1
16	+1	+1	+1	+1	+1

It is called the 2^{5-1} design because

$$^1/_2 \cdot 2^5 = 2^{-1} \cdot 2^5 = 2^{5-1}$$

The 2^{5-1} Design

- This half-fraction was not chosen randomly from the full factorial!
- These 16 combinations were carefully selected to allow us to fit the '1st order + twist' response surface equation.
  - for the theory of choosing regular fractions, see the book by Box, Hunter and Hunter.
- How do we know that this design will do the job?
- As before, if we construct the model matrix M, we find that every pair of columns has sum-product 0.

The design is said to be completely orthogonal for fitting a '1st order + twist' equation

Other Regular Fractions

- Apart from the 2^{5-1}, the following regular fractions can be used to fit a '1st order + twist' response surface equation.

Designation	No. of runs
2^{6-1}	32
2^{7-1}	64
2^{8-2}	64
2^{9-1}	128
2^{10-2}	128

These are generally available in software packages

A Problem With Regular Fractions

- The disadvantage of only using regular fractional designs is the inflexibility of the number of runs.
  - E.g. with 7 factors we need 64 runs.
  - but with 7 factors, the '1st order + twist' equation has 29 terms, so it can be fitted to a much smaller design.

Designation	No. of runs	No. of terms	Residual DF
2^{6-1}	32	22	10
2^{7-1}	64	29	35
2^{8-2}	64	37	27
2^{9-1}	128	46	82
2^{10-2}	128	56	72

Non-regular Fractional Designs

- To give us more flexibility, we can use a non-regular design
  - e.g. the following slide lists a ¾ fraction of the 2^5 full factorial
  - or we can use a Custom Design algorithm to generate a design with any specified number of runs (To be discussed in later sessions).
- These designs are not completely orthogonal.

In the model matrix M, some pairs of columns will have small non-zero correlations

Example: Non-regular Design For 5 Factors In 24 Runs

x_1	x_2	x_3	x_4	x_5
−1	−1	−1	−1	−1
−1	+1	−1	−1	−1
+1	+1	−1	−1	−1
−1	−1	+1	−1	−1
+1	−1	+1	−1	−1
−1	+1	+1	−1	−1
−1	−1	−1	+1	−1
+1	−1	−1	+1	−1
+1	+1	−1	+1	−1
+1	−1	+1	+1	−1
−1	+1	+1	+1	−1
+1	+1	+1	+1	−1
−1	−1	−1	−1	+1
+1	−1	−1	−1	+1
+1	+1	−1	−1	+1
+1	−1	+1	−1	+1
−1	+1	+1	−1	+1
+1	+1	+1	−1	+1
−1	−1	−1	+1	+1
−1	+1	−1	+1	+1
+1	+1	−1	+1	+1
−1	−1	+1	+1	+1
+1	−1	+1	+1	+1
−1	+1	+1	+1	+1

If we construct the matrix M, the highest r^2 between any pair of columns is 0.111

In This Session We Have…

- Described the advantages of designed experiments, and some of the different types of experiment.
- Defined a two-step strategy (screening + follow-up) that is useful in most engineering experimentation.
- Explained the concept of two-factor interaction (2fi) and how it is represented in a twisted response surface.
- Discussed some designs that allow us to fit a ‘1st order + twist’ equation.
- Defined full factorial and fractional factorial designs.
- Listed some useful regular fractions, but explained that regular fractions give us limited flexibility in choosing an efficient design.
- Given an example of a non-regular fractional design.

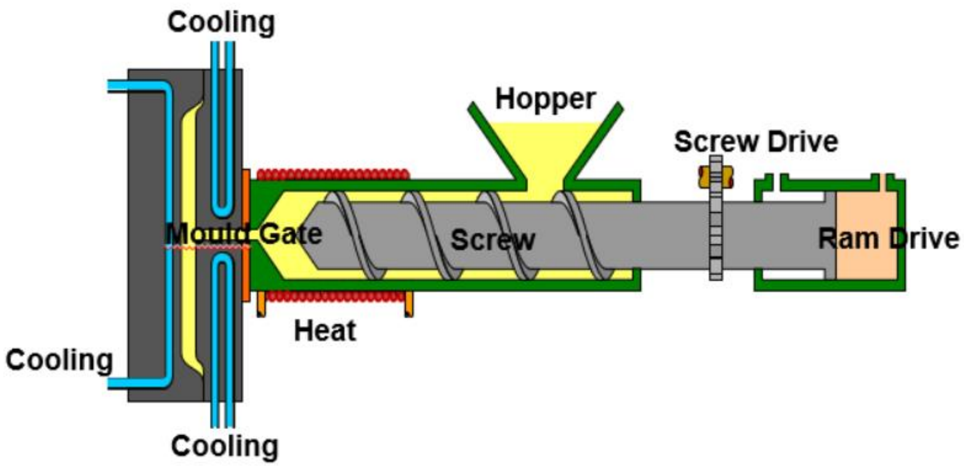
Session 4: Planning a Two Level Experiment

Tutorial and Exercise

Tutorial

- **Session TS04: Planning a Two Level Experiment**
- **Objectives**

  Create a two level full factorial design of experiments.
- **Engineering Scenario**
- A study of potential alternative materials for bumper bracket mouldings.
- The brackets are injection moulded (IM).
- In our example, the factors are settings on the IM machine.



Tutorial

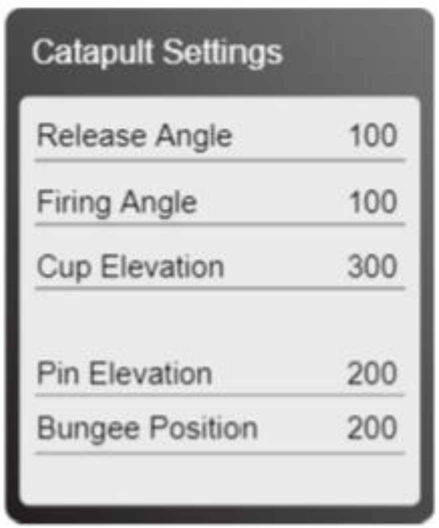
- **Session TS04: Planning a Two Level Experiment**
- **Python Environment**

  A self-guided tutorial has been created as a Colab notebook with pre-designed Python code and notes. For this tutorial, follow the instructions in the notes, upload data files and run the code. No modification of code is required. Interpret the results in accordance with the Technical session.
- **Tutorial Task**
  1. Generate a coded 2^3 full factorial design and export to be manipulated externally.
  2. Convert coded design into engineering units.
  3. Generate the model matrix for fitting a ' 1^{st} order + twist' response surface in three factors.
  4. Check if the generated design is orthogonal.
  5. Generate a coded 2^{5-1} fractional factorial design.

Exercise

- **Session TS04+05: Planning a Two Level Experiment**
- **Objective**
- To plan a screening experiment for a Virtual Catapult, run the experiment and collect the responses.

https://sigmazone.com/catapult/



Exercise

- **Session TS04+05: Planning a Two Level Experiment**
- **Python Environment**

  The exercise has been created as a Colab notebook with notes. For this exercise, follow the instructions in the notes, and create your own code using the accompanying tutorial as a guide. Interpret the results in accordance with the Technical session. **Stop after Exercise 04, Task 5.**

- **Guidelines**
  - The response is the horizontal distance travelled by the ball.
  - For screening experiment you should include all the 5 factors- show in the "Catapult Settings" dialogue box.
  - In Exercise 06 you will select 3 of them and run a three-level experiment to investigate their effects in more detail.
  - You should first “play” with the simulator to identify some reasonable limits for the DoE space your experiment (i.e. the minimum and maximum setting for each factor).
  - In choosing a design, you need to decide what kind of response surface you want to fit, i.e. 1st order only (ignoring interactions) or 1st order + twist.
  - Then select a design that allows you to fit this equation.