

Modeling and design of experiments

Chapitre 7: Mixture designs

Dr Jean-Marie Fürbringer

École Polytechnique Fédérale de Lausanne

Fall 2025

7.1.1 Specificity of mixture spaces

- ▶ A mixture space is considered when the factors x_i are fractions of a whole and that the increase of one factor implies the decrease in proportion of one or more other factors.
- ▶ This situation then implies in addition to the model $f(\vec{x})$ to be determined, two other conditions :

1. a constant sum :
$$\sum_{i=1}^q x_i = 1$$

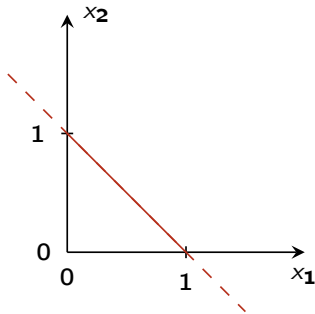
2. positive values $x_i \geq 0$

7.1.2 A mixture of two components

- ▶ Simple example to illustrate the principle
- ▶ Two variables x_1 and x_2 like the concentrations of two products :

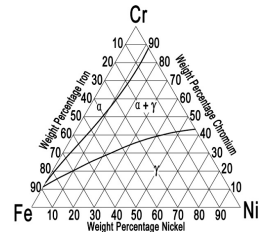
$$x_1 + x_2 = 1$$

- ▶ The experimental space has one degree of freedom only

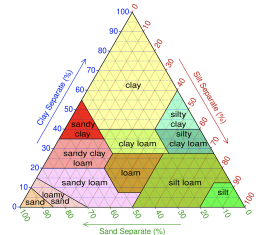


7.1.3 Ternary plot

- ▶ A ternary plot, ternary graph, triangle plot, simplex plot, or Gibbs triangle is a barycentric plot on three variables which sum to a constant.
- ▶ It graphically depicts the ratios of the three variables as positions in an equilateral triangle.
- ▶ Ternary plots are tools for analyzing compositional data in the three-dimensional case.



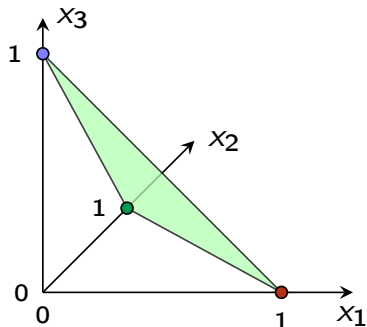
Stainless austenitic steel



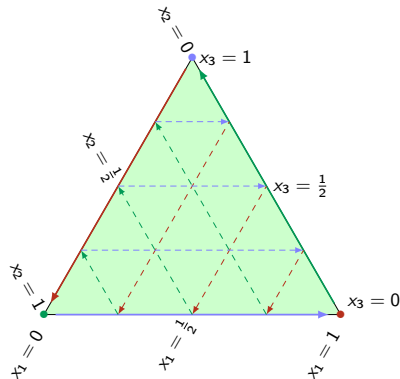
Soil types

7.1.4 Ternary plot : from 3D to 2D

For the case of 3 components, it is usual and useful to draw a ternary diagram.



$$x_1 + x_2 + x_3 = 1$$



7.1.5 Computing a ternary diagram

MATLAB

No native functions, but a few user defined routines on
<https://ch.mathworks.com/matlabcentral/fileexchange/2299-chemist-ternplot>

ternaxes(12) create the axis system
ternplot(A,B,C,'or') place points in the diagram
ternpcolor(A,B,Z) create a colorplot
ternsurf(A,B,Z) create a surfaceplot

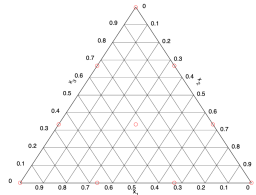
7.1.6 Example with *ternplot()*

MATLAB code

```

ternaxes; % ternary axes
ternlabel('x1','x2','x3'); % placing labels
F=(fullfact([5 5 5])-1)/3; % generating a FFD
index=sum(F,2)==1; % selecting coordinates
E=F(index,:); % essay matrix
ternplot(E(:,1),E(:,2),E(:,3),'r'); % plotting

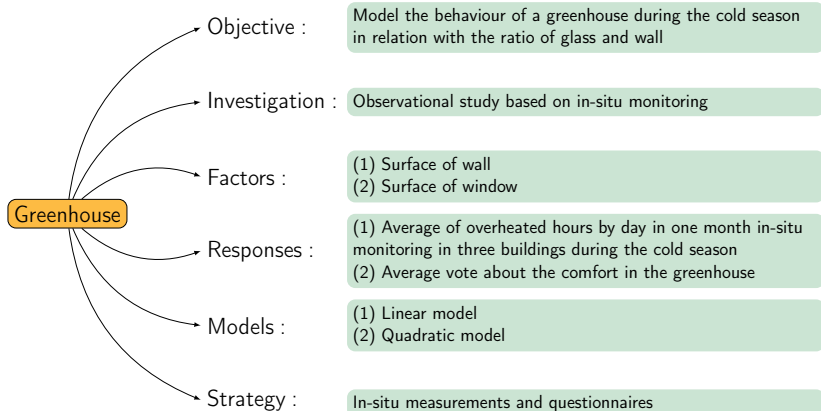
```



7.1.7 Example : Optimization of a greenhouse



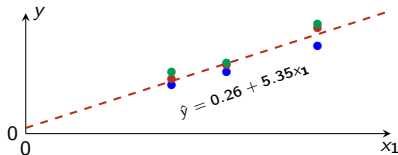
7.1.8 Optimization of a greenhouse : mindmap



7.1.9 Optimization of a greenhouse : data

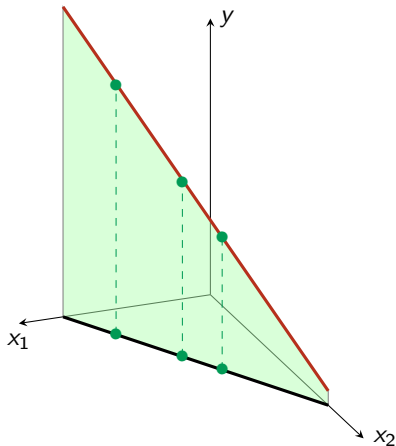
Table – Average overheating hours ($T_{in} \geq 25^{\circ}C$)

Building	A	B	C
Window	40%	55%	80%
D	2.2 hrs/D	2.8 hrs/D	4 hrs/D
J	2.5 hrs/D	3.1 hrs/D	4.8 hrs/D
F	2.8 hrs/D	3.2 hrs/D	5 hrs/D



7.1.10 Optimization of a greenhouse : model

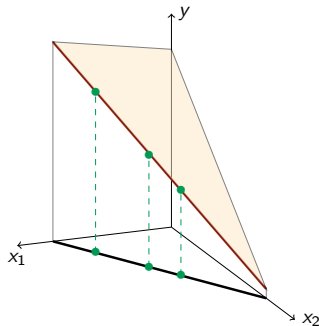
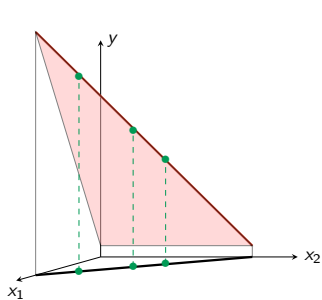
Here is the solution in the plane $x_1 + x_2 = 1$



7.1.11 There are an infinity of solutions

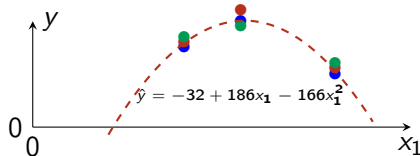
$$\hat{y} = 0.26 + 5.35x_1 = 0.26(x_1 + x_2) + 5.35x_1 = 5.61x_1 + 0.26x_2$$

$$\hat{y} = 5 + 5.35x_1 - 4.74(x_1 + x_2) = 5 + 0.61x_1 - 4.74x_2$$

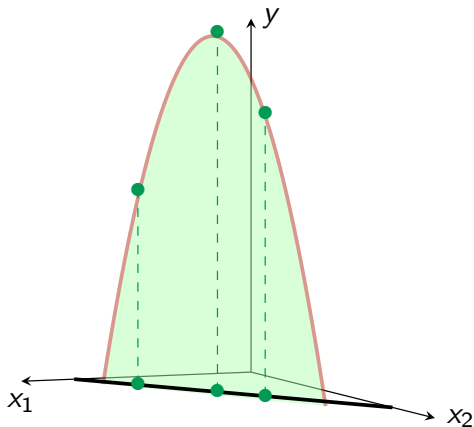


7.1.12 Average votes about the comfort

Building	A	B	C
Window	40%	55%	80%
P1	15	20	10
P2	16	22	11
P3	17	19	12

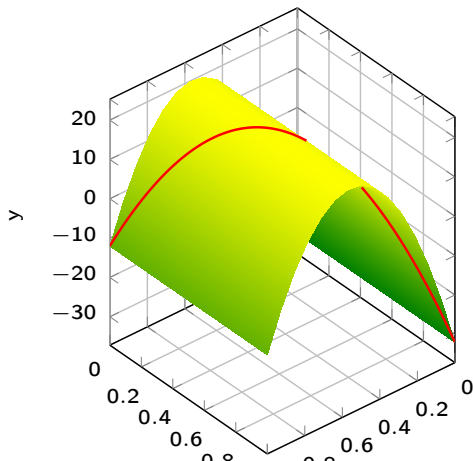


7.1.13 Solution in the plane $x_1 + x_2 = 1$



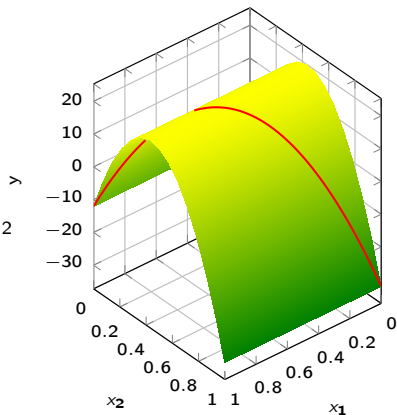
7.1.14 A possible model in 3 dimensions

- ▶ The function $y = -32 + 18x_1 - 166x_1^2$ can be interpreted as a paraboloid in the 3D space x_1x_2y
- ▶ But an infinity of functions have the same intersection with the vertical plane $x_1 + x_2 = 1$



7.1.15 Here is another possible model

$$\begin{aligned}y &= -32 + 18x_1 - 166x_1^2 \\ &= -32 + 186(1 - x_2) - 166(1 - x_2)^2 \\ &= -12 + 146x_2 - 166x_2^2\end{aligned}$$



7.2.1 Scheffé's and slack models

- ▶ The reduction to a $(N_{fact} - 1)$ dimension experiment space, due to the additional equation $\sum_{i=1}^q x_i = 1$, has the consequence that several models allows to model a given response.
- ▶ The imposed correlation between the factors reduces the rank of the essay matrix E of one unit.
- ▶ A *canonical* model is needed : a unique representation for the different models that represent the same response
- ▶ There are several possible choices
 - ▶ Scheffé's models (models without an intercept)
 - ▶ Slack models (models with an intercept)

7.2.2 Scheffé's linear mixture model

- ▶ Let's start with a standard linear model of rank $(q + 1)$:

$$y = a_o + \sum_{i=1}^q a_i x_i$$

- ▶ Let's introduce, at the level of the constant a_o , the proportionality constraint $\sum_{i=1}^q x_i = 1$
- ▶ The Scheffé's model (of rank q) is then :

$$y = \sum_{i=1}^q (a_o + a_i) x_i = \sum_{i=1}^q \beta_i x_i$$

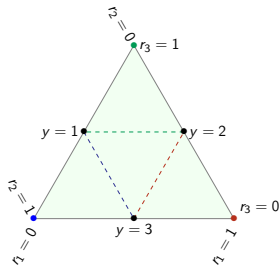
- ▶ With the Scheffé's linear coefficients $\beta_i = (a_o + a_i)$

7.2.3 Example of a Scheffé's linear model

1. A recipe is made with three ingredients x_1 , x_2 and x_3 so that $x_1 + x_2 + x_3 = 1$
2. Essays have been made at the middles of the vertices of the ternary scheme. The output y is a KPI of the process.

x_1	x_2	x_3	y
0.5	0.5	0	2
0.5	0	0.5	3
0	0.5	0.5	1

3. Determine the coefficients of the Scheffé's linear model



7.2.4 Identification of a Scheffé's linear model

$$\text{Model : } Y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$$

Linear system :

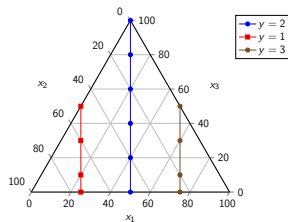
$$\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} \Leftrightarrow \begin{cases} \beta_1 + \beta_2 = 4 \\ \beta_1 + \beta_3 = 6 \\ \beta_2 + \beta_3 = 2 \end{cases}$$

Solution :

$$\begin{cases} \beta_3 = 2 \\ \beta_2 = 0 \\ \beta_1 = 4 \end{cases} \Rightarrow \hat{Y} = 4x_1 + 2x_3$$

Isolines :

$$\begin{cases} x_1 = x_1 \\ x_2 = x_1 - \frac{y}{2} + 1 \\ x_3 = \frac{y}{2} - 2x_1 \end{cases}$$



7.2.5 Quadratic Scheffé's model

- ▶ The standard quadratic model of rank $(q + 1)(q + 2)/2$ is :

$$y = a_o + \sum_{i=1}^q a_i x_i + \sum_{i \leq j}^q a_{ij} x_i x_j$$

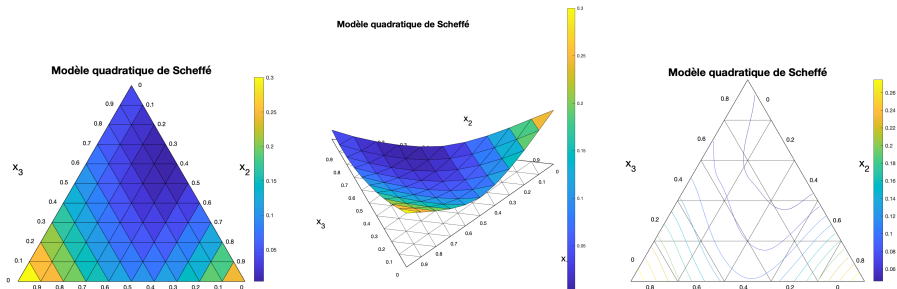
- ▶ The constraint of proportionality allows to write :

$$\begin{cases} a_o = a_o \cdot 1 = a_o \sum_{i=1}^q x_i = \sum_{i=1}^q a_o x_i \\ a_{ij} x_i^2 = a_{ij} x_i \left(1 - \sum_{j \neq i}^q x_j \right) \end{cases}$$

7.2.6 Quadratic Scheffé's model (2)

- When recombining the coefficients the quadratic Scheffé's model of rank $q(q + 1)/2$ is then

$$y = \sum_{i=1}^q \beta_i x_i + \sum_{i < j} \beta_{ij} x_i x_j$$



7.2.7 Cubic Scheffé's model

- ▶ The standard cubic model is :

$$y = a_o + \sum_{i=1}^q a_i x_i + \sum_{i \leq j} a_{ij} x_i x_j + \sum_{i \leq j \leq k} a_{ijk} x_i x_j x_k$$

- ▶ Integrating the constraints coming from the ratios of the mixture, the cubic Scheffé's model becomes

$$y = \sum_{i=1}^q \beta_i x_i + \sum_{i < j} \beta_{ij} x_i x_j + \sum_{i < j} \gamma_{ij} x_i x_j (x_i - x_j) + \sum_{i < j < k} \beta_{ijk} x_i x_j x_k$$

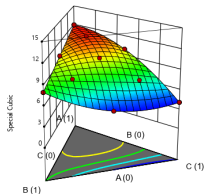
- ▶ Its rank is $q(q+1)(q+2)/3!$
- ▶ Truncated form if $\gamma_{ij} = 0$ (allows to diminish the number of runs)

7.2.7 Example of cubic Scheffé's model

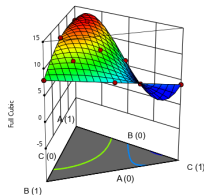
$$y = 12x_1 + 8x_2 + 4x_3 + 8x_1x_3 - 8x_2x_3 + 54x_1x_2x_3$$

$$y = 2x_1 + 8x_2 + 4x_3 + 8x_1x_2 - 8x_1x_3 + 54x_1x_2x_3 + 48x_1x_3(x_1 - x_3)$$

Scheffé Special cubic



Scheffé full cubic



7.2.8 Slack model

- ▶ It is in neglecting one factor in the model. It represents a risk if the neglected factor is active. So it is better to reserve this for factors having only linear effect.
- ▶ SV means Slack Variable
- ▶ Full linear SV model

$$y = a_0 + \sum_{i=1}^{q-1} a_i x_i$$

- ▶ Full quadratic SV model

$$y = a_0 + \sum_{i=1}^{q-1} a_i x_i + \sum_{i=1}^{q-1} a_{ii} x_i^2 + \sum_{i=1}^{q-2} \sum_{j=i+1}^{q-1} a_{ij} x_i x_j$$

7.3.0 Fixed Component Constraint

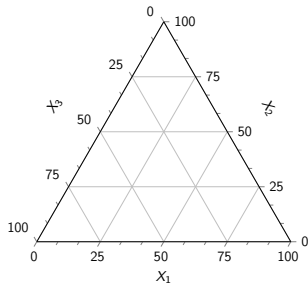
$$x_1 = a, \quad 0 \leq a \leq 1$$

Typical use cases :

- ▶ Sensitivity analysis on one component
- ▶ Conditional binary mixture analysis
- ▶ Fixing an additive or catalyst concentration

Key implication :

- ▶ Dimensionality reduction of the mixture space



7.3.0 Individual Lower and Upper Bounds

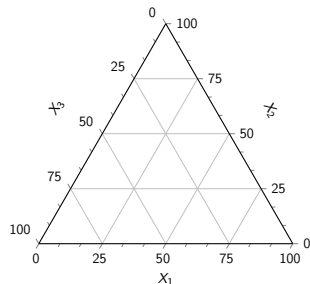
$$L_i \leq x_i \leq U_i$$

Typical use cases :

- ▶ Safety or regulatory limits
- ▶ Solubility or process constraints
- ▶ Minimum functional concentration

Geometric effect :

- ▶ Truncation of the simplex
- ▶ Reduced feasible mixture region



7.3.0 Linear Combination Constraints

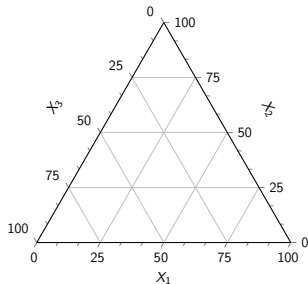
$$ax_1 + bx_2 + cx_3 \leq d$$

Typical use cases :

- ▶ Cost limitations
- ▶ Energy content or emissions limits
- ▶ Environmental or life-cycle impact constraints

Interpretation :

- ▶ Intersection of the simplex with a half-space
- ▶ Linear feasibility region



7.3.1 Ratio constraints

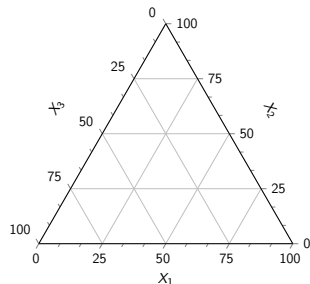
$$\frac{x_i}{x_j} \leq r \iff x_i - rx_j \leq 0$$

Typical use cases :

- ▶ Additive-to-base material ratios
- ▶ Catalyst or inhibitor formulations
- ▶ Process stability requirements

Key message :

- ▶ Apparently nonlinear constraints may be linear in mixture space

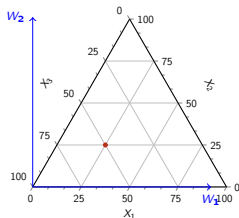


7.3.2 Transformation ternary to Cartesian

- ▶ Ternary variables x_1 , x_2 et x_3
- ▶ Cartesian variables W_1 et W_2
- ▶ The matrices for the change of base are

$$\begin{pmatrix} W_1 \\ W_2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 0 & \frac{\sqrt{3}}{2} & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{\sqrt{3}}{3} & 0 \\ 0 & \frac{2\sqrt{3}}{3} & 0 \\ -1 & -\frac{\sqrt{3}}{3} & 1 \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \\ 1 \end{pmatrix}$$

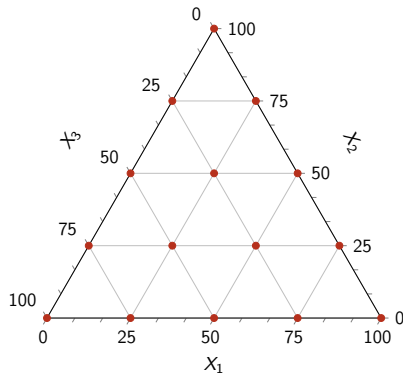


7.4.1 Simplex lattice design $\{q, m\}$

Factorial design for q components with levels $(\frac{0}{m}, \frac{1}{m}, \dots, \frac{m}{m})$

- ▶ Example $\{q=3, m=4\}$
- ▶ The levels are $(0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1)$
- ▶ The matrix of experiments

$$\begin{pmatrix} 1 & 0 & 0 \\ 0.75 & 0.25 & 0 \\ 0.75 & 0 & 0.25 \\ 0.5 & 0.5 & 0 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.25 & 0.25 \\ 0.25 & 0.75 & 0 \\ 0.25 & 0 & 0.75 \\ 0.25 & 0.5 & 0.25 \\ 0.25 & 0.25 & 0.5 \\ 0 & 1 & 0 \\ 0 & 0.75 & 0.25 \\ 0 & 0.5 & 0.5 \\ 0 & 0.25 & 0.75 \\ 0 & 0 & 1 \end{pmatrix}$$



7.4.2 Simplex lattice design

$$N_{exp} = \binom{q+m-1}{m} = \frac{(q+m-1)!}{m!(q-1)!}$$

Number of points per design

q	m		
	2	3	4
3	6	10	15
4	10	20	35
5	15	35	70
6	21	56	126
7	28	84	210

Number of coefficients per model

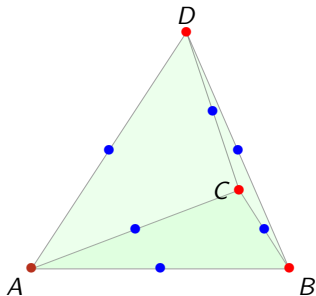
q	linear	quad- ratic	special cubic	full cubic
3	3	6	7	10
4	4	10	14	20
5	5	15	25	35
6	6	21	41	56
7	7	28	63	84

7.4.3 Simplex lattice design {4,2}

- ▶ $N_{exp} = 10$
- ▶ Sufficient for a quadratic model

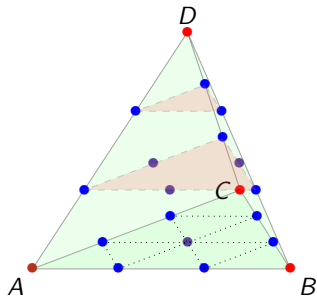
$$y = \sum_{i=1}^q \beta_i x_i + \sum_{i<j}^q \beta_{ij} x_i x_j$$

- ▶ But No point within the domain : no full mixture



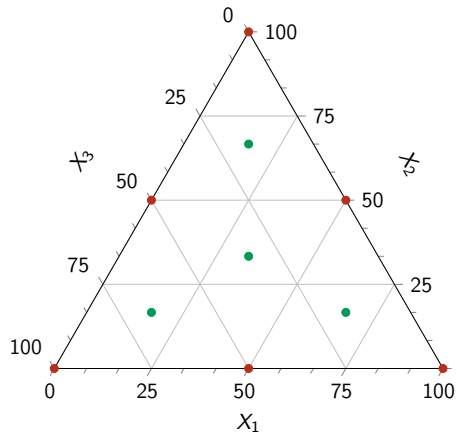
7.4.4 Simplex lattice design {4,3}

- ▶ $N_{exp} = \binom{4+3-1}{3} = \binom{6}{3} = 20$
- ▶ Sufficient for a full cubic model
- ▶ But No point within the domain : no full mixture
- ▶ We need another type of design to be sure to enter in the domain



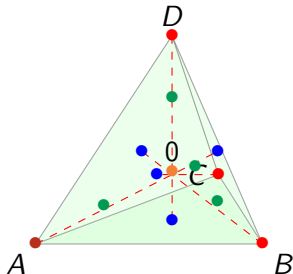
7.4.5 Simplex response surface

- ▶ If $m < q$ there is no full mixture
- ▶ Then designs $\{q, 2\}$ have to be completed with $q + 1$ points
- ▶ One point at the center $(1/q, 1/q, \dots)$
- ▶ q points at the middle of the distance between the center and the vertices (axial check blends)
- ▶ This has the advantage of offering additional degrees of freedom



7.4.6 Simplex screening design

- ▶ As its name suggests, it is a question of determining the variables that can be eliminated from the problem.
- ▶ The design is composed by
 - ▶ the vertices - q points
 - ▶ the central point - 1 point
 - ▶ the points in the middle of the distance from the center to the vertices (axial check blends) - q points
 - ▶ the end points - q points
- ▶ With 3 factors, it is the same design as the Simplex response surface



7.4.7 Conclusions

- ▶ A **comprehensive overview** of important approaches in mixture design.
- ▶ These approaches offer a versatile toolkit for tackling diverse mixture design problems allowing **flexibility and adaptability**.
- ▶ They reduce the need for extensive experimentation, saving time and resources, providing robust solutions less susceptible to variability and uncertainty : **efficiency and robustness**
- ▶ **Improved Decision-Making** : enhance our understanding of the relationships between components, enabling process adjustments and improving product performance.