



Optimal Design of Experiments for Powerful Equivalence Testing

Ying Chen

(Research funded by a grant from GlaxoSmithKline
Biologicals SA, Belgium)

Prof. dr. Peter Goos

(KU Leuven)

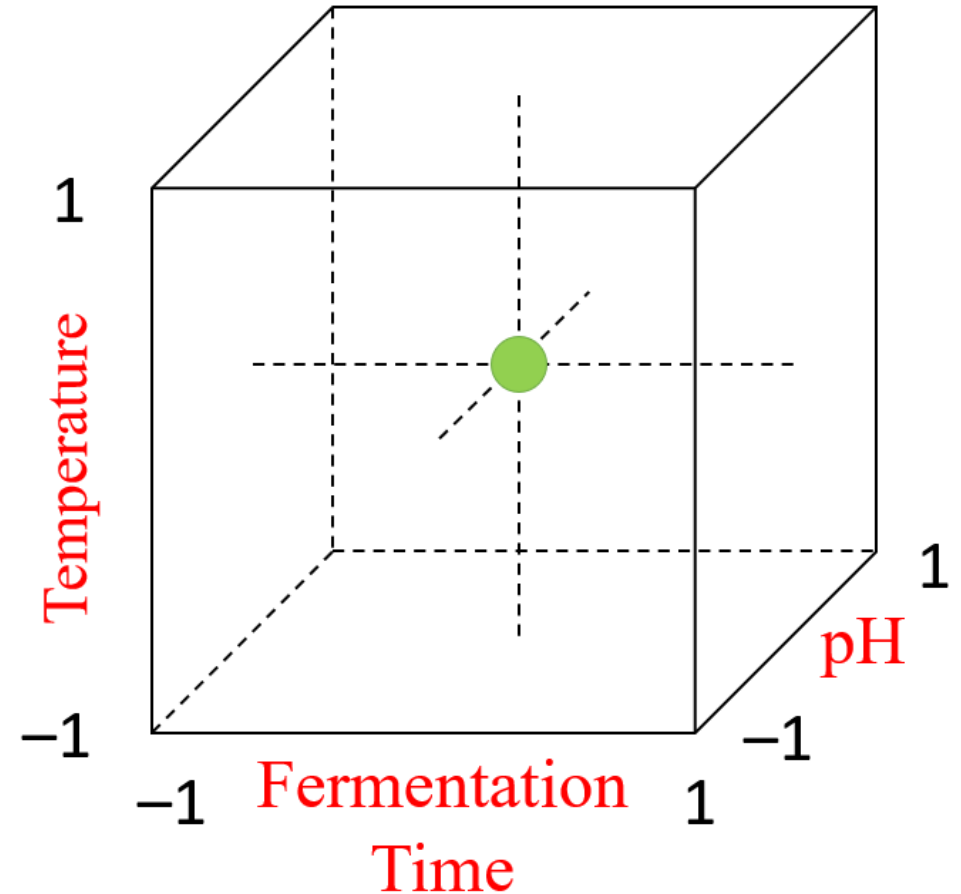
Dr. Bernard G Francq

(An employee of the GSK group of companies)



Context

- In manufacturing, an optimized process operates at **the target point t**
- Products manufactured under this condition have the intended quality (= the target response)
- Uncontrollable variations occur during routine production



Process robustness study

- Quantify to what extent the product quality is affected by variability in the process parameters
- We proposed the **GI_D-optimality criterion** to construct designs for process robustness studies (Technometrics, 2026)

$$\frac{\int_{\chi} \sigma_{\varepsilon}^2 [\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{t})]' (\mathbf{X}'\mathbf{X})^{-1} [\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{t})] d\mathbf{x}}{\int_{\chi} d\mathbf{x}}$$

Need for an alternative design strategy

- Analyzing the data from a process robustness study involves equivalence tests
- Need to demonstrate that responses under typical variations are equivalent to the target response
- The design needs to have enough power to demonstrate equivalence
 - correctly conclude equivalence when equivalence exists

Start with one point $\mathbf{x}$

- **Notations**

- mean response at point $\mathbf{x}$: $\mu_{y|\mathbf{x}}$
- predicted response at point $\mathbf{x}$: $\hat{y}_{\mathbf{x}}$
- mean response at the target $\mathbf{t}$: $\mu_{y|\mathbf{t}}$
- predicted response at the target $\mathbf{t}$: $\hat{y}_{\mathbf{t}}$

- $\Delta = \mu_{y|\mathbf{x}} - \mu_{y|\mathbf{t}} = [\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{t})]' \boldsymbol{\beta}$
 - Estimator $\hat{y}_{\mathbf{x}} - \hat{y}_{\mathbf{t}}$ is normally distributed with mean Δ and variance $\sigma^2 = \sigma_{\varepsilon}^2 [\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{t})]' (\mathbf{X}'\mathbf{X})^{-1} [\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{t})]$

Equivalence test

- Hypotheses to be tested

$$H_0 : \Delta \notin [\theta_L, \theta_U] \quad \text{versus} \quad H_1 : \Delta \in [\theta_L, \theta_U].$$

- Rewrite as two sets of one-sided hypotheses:

$$H_{01} : \Delta \leq \theta_L \quad \text{versus} \quad H_{11} : \Delta > \theta_L,$$

$$H_{02} : \Delta \geq \theta_U \quad \text{versus} \quad H_{12} : \Delta < \theta_U.$$

- **Two One-Sided t -Tests (TOST)**

Equivalence test

- Two test statistics

$$T_L = \frac{\hat{y}_x - \hat{y}_t - \theta_L}{\hat{\sigma}} \quad T_U = \frac{\hat{y}_x - \hat{y}_t - \theta_U}{\hat{\sigma}}$$

- Decision rules

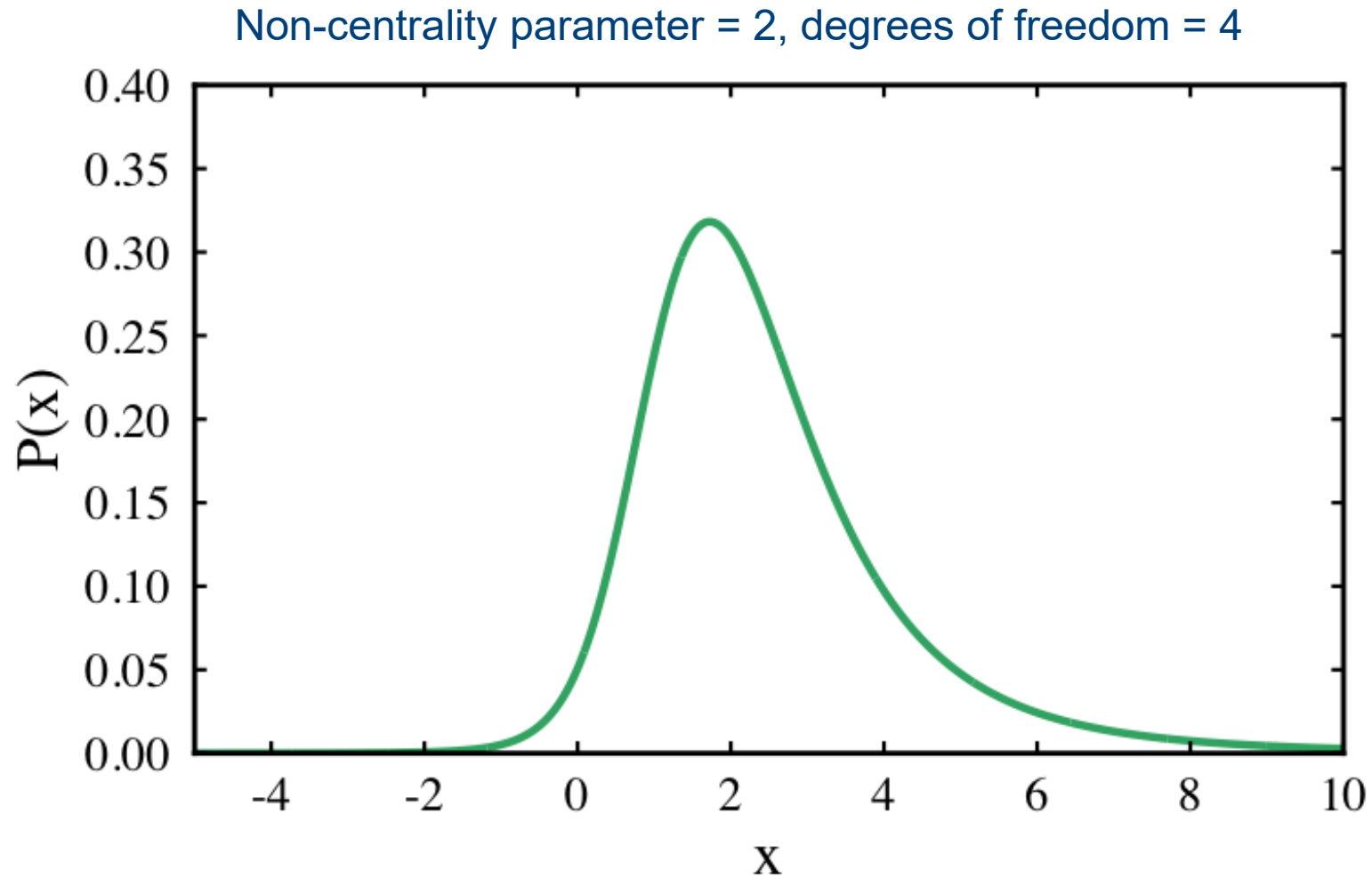
- reject H_{01} if $T_L > t_{1-\alpha, n-p}$
- reject H_{02} if $T_U < -t_{1-\alpha, n-p}$

- The responses are equivalent iff both H_{01} and H_{02} are rejected

$$P(T_L > t_{1-\alpha, n-p} \quad \text{and} \quad T_U < -t_{1-\alpha, n-p} \mid \theta_L < \Delta < \theta_U)$$

T_L and T_U follow a bivariate non-central t -distribution

Univariate non-central t -distribution



T_L and T_U follow a bivariate non-central t -distribution

- **Special case:** T_L and T_U are identical except for their non-centrality parameters
- Owen (Biometrika, 1965) provided exact formulas for calculating probabilities of joint events with this specific t -distribution
- The power for the equivalence test is
$$P(T_L > t_{1-\alpha, n-p} \quad \text{and} \quad T_U < -t_{1-\alpha, n-p} \mid \theta_L < \Delta < \theta_U)$$
$$= Q_{n-p}(-t_{1-\alpha; n-p}, \delta_U; 0, R) - Q_{n-p}(t_{1-\alpha; n-p}, \delta_L; 0, R)$$
- **Difficulty:** power calculation depends on β
 - Construct Bayesian optimal designs by assuming a prior distribution for β

PE-Optimality Criterion (Powerful Equivalence Testing)

The expected proportion of locations in $\mathcal{X}$ where the power of the TOST exceeds the threshold, averaged over the prior distribution of β

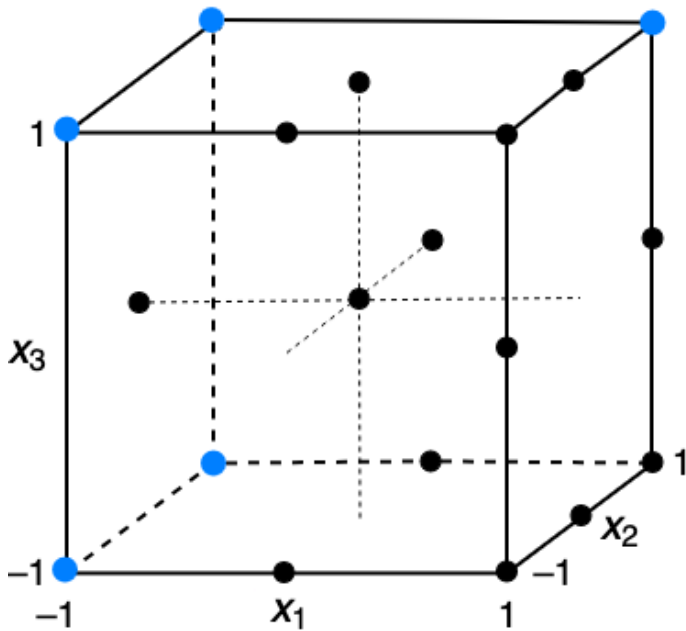
- Sample a large number of points from the region
- For each draw β^i ($i = 1, 2, \dots, N$)
 - Compute the TOST power at every sampled location
 - Calculate the percentage of locations where power exceeds the threshold, π_i
 - Weight the result π_i with w_i
- The quality of a design: $S_{\text{PE}} = \sum_{i=1}^N w_i \pi_i$. Higher value is better

Proof-of-concept PE-optimal designs

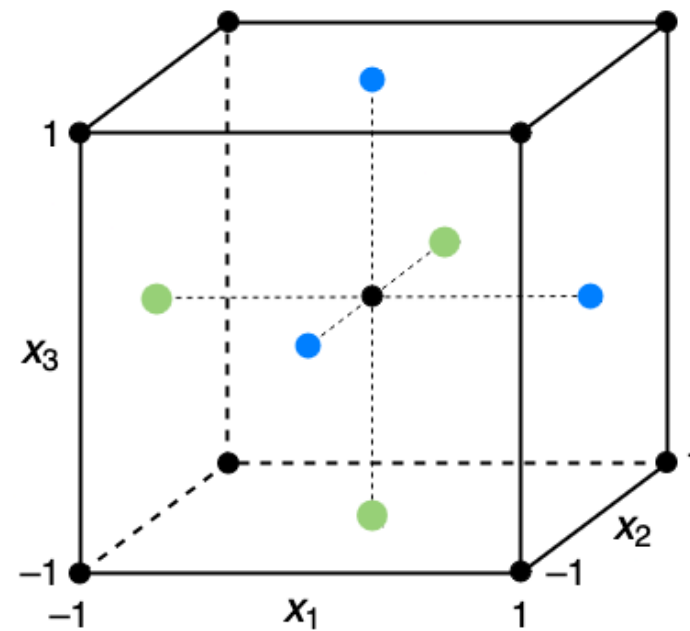
- For the TOST procedure
 - $\theta_L = -2, \theta_U = 2$
 - $\alpha = 0.05$
 - Power threshold: 0.8
- Prior distribution of β : multivariate standard normal

24-run designs for a response surface model in three factors

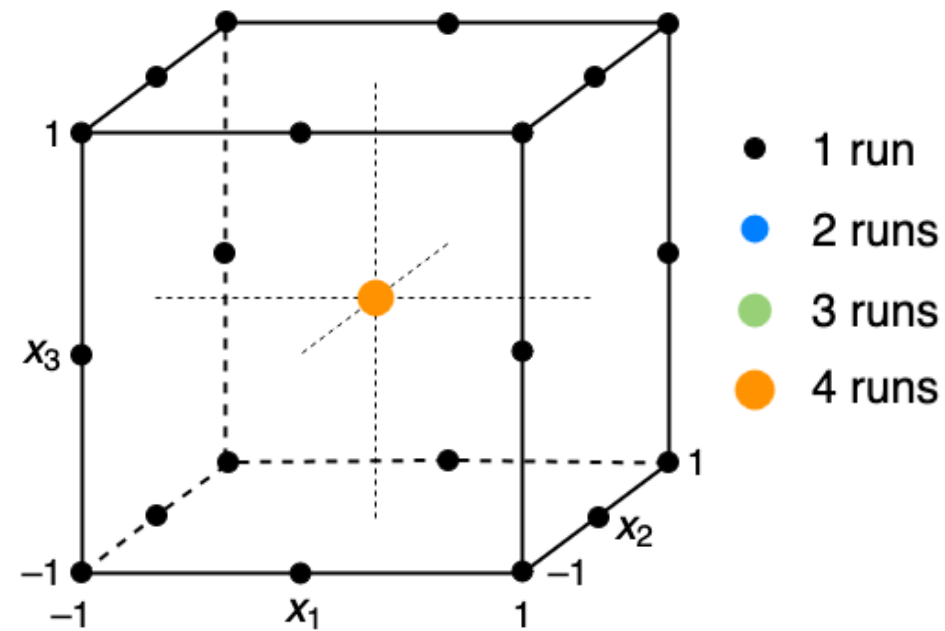
D-optimal design



I-optimal design



A-optimal design

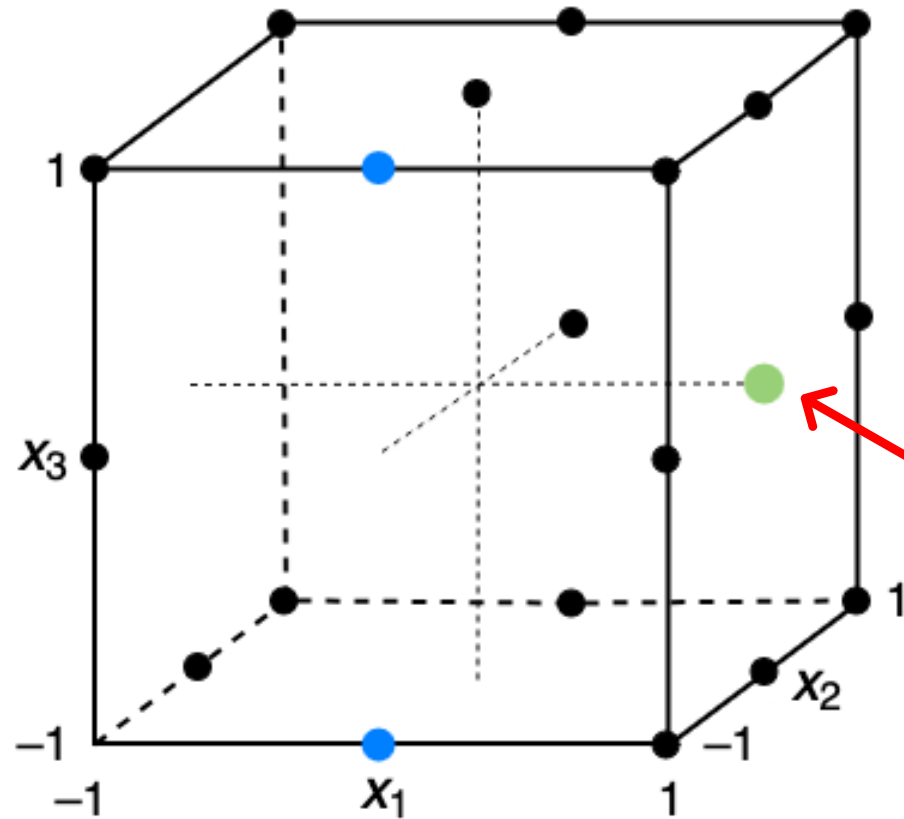


GI_D - and PE-optimal design

Target = $[0, 0, 0]$

Target = [1, 0, 0]

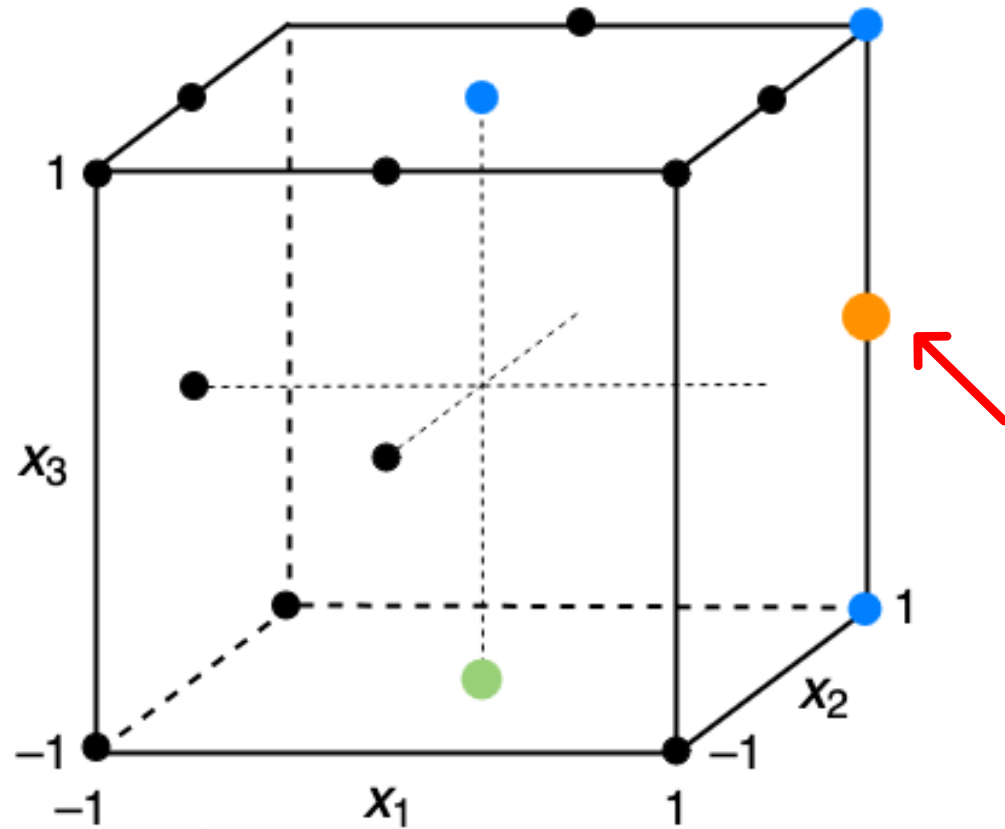
GI_D - and PE-optimal design



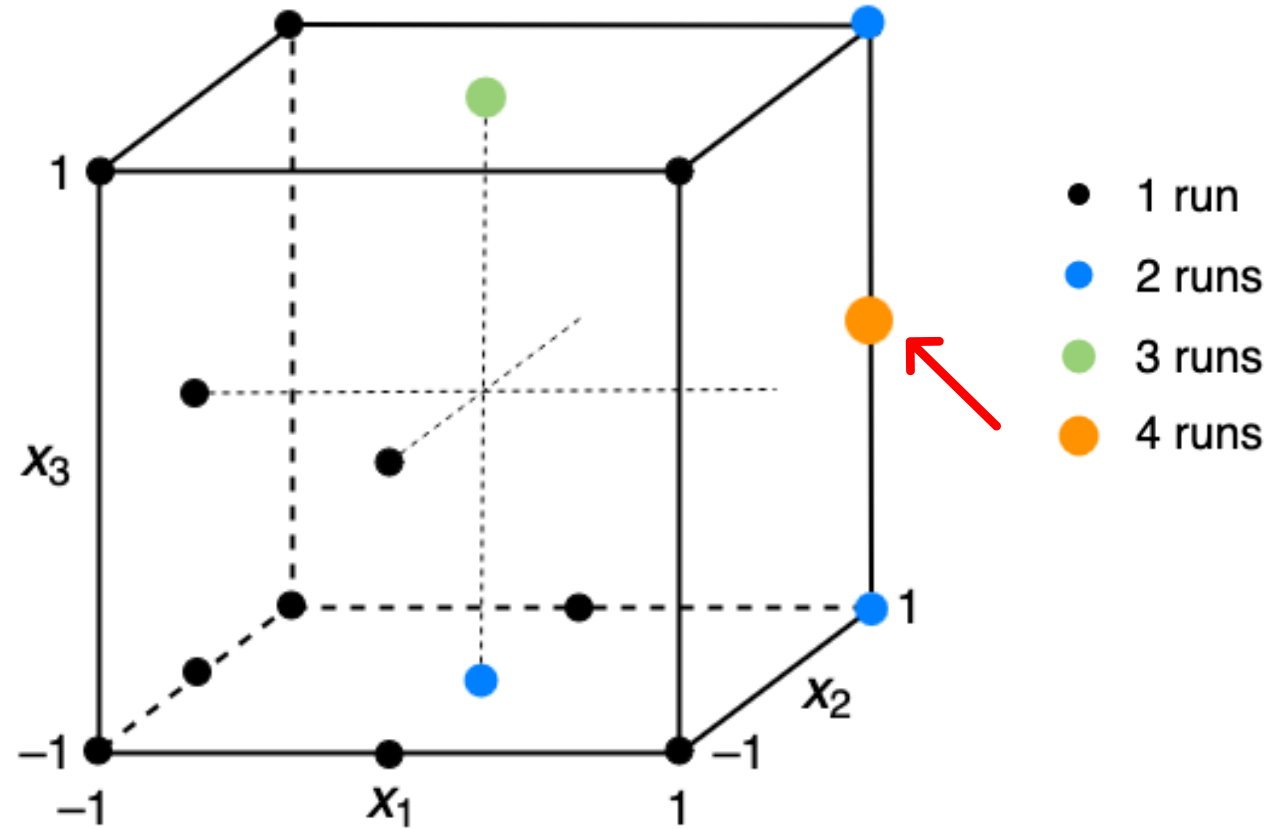
- 1 run
- 2 runs
- 3 runs

Target = [1, 1, 0]

PE-optimal design



GI_D -optimal design



PE-optimality criterion values

Optimal Design	$\mathbf{t}' = [0, 0, 0]$	$\mathbf{t}' = [1, 0, 0]$	$\mathbf{t}' = [1, 1, 0]$
D	53.81%	23.48%	5.73%
I	53.58%	22.04%	5.77%
A	56.07%	23.87%	6.33%
GI _D	56.07%	32.93%	18.25%
PE	56.07%	32.93%	18.86%

Robustness design for a vaccine manufacturing process

- 28-run design, 4 quantitative factors for a response surface model,
 $\mathbf{t}' = [0, 0, 0, 0]$
- PE-optimal design
 - $\boldsymbol{\beta} \sim N(\hat{\boldsymbol{\beta}}_{\text{OLS}}, \widehat{\text{Var}}(\hat{\boldsymbol{\beta}}_{\text{OLS}}))$
 - $\theta_{\text{L}} = -8$ and $\theta_{\text{U}} = 8$
 - $\hat{\sigma}_{\varepsilon} = 2.4381$

Optimal Design	PE Value
D	50.87%
A	53.81%
I	54.09%
GI _D	54.42%
PE	55.12%

Conclusion

- New PE-optimality criterion, for selecting a design for process robustness studies focusing on equivalence testing
 - Explicitly incorporates information relevant to equivalence testing
 - Maximizes the power of the TOST procedure across the experimental region
- PE-optimal designs outperform D-, I- and A-optimal designs, especially when the target point is not at the center
- PE-optimal designs outperform GI_D -optimal designs when the designs differ
- When reliable prior information is available, PE-optimal designs that incorporate this information perform best



Optimal Design of Experiments for Powerful Equivalence Testing

Ying Chen

(Research funded by a grant from GlaxoSmithKline
Biologicals SA, Belgium)

Prof. dr. Peter Goos

(KU Leuven)

Dr. Bernard G Francq

(An employee of the GSK group of companies)

