KU LEUVEN

Bayesian D- and I-optimal designs for choice experiments involving mixtures and process variables

Mario Becerra
with Peter Goos

September 14th, 2023
Scientific Research Network on Choice modelling
Second International Workshop
Leuven, Belgium

Outline

1. Choice modeling and choice experiments
2. Mixture experiments and models
3. Combining choice models and mixture models
4. Optimality criteria for choice experiments
5. Examples

Choice modeling and choice experiments

Discrete choice experiments

- Quantify preferences
- Preference data is collected
- Respondents choose between sets of alternatives (choice sets)
  - Example: choosing to buy product A, B or C
- Latent utility function → probability of making each decision

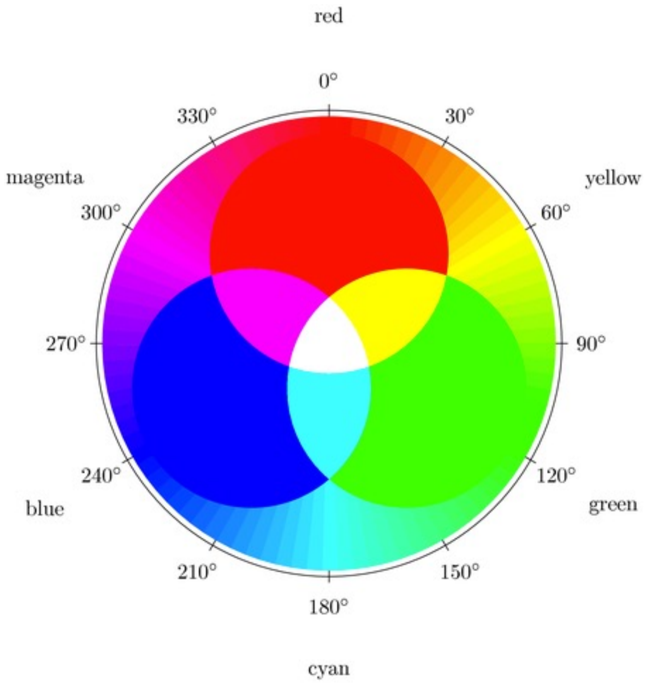
KU LEUVEN

Mixture experiments

Mixtures

- Many products and services can be described as mixtures of ingredients
- Examples:
  - ingredients of bread
  - ingredients used to make a cocktail
  - sand, water and cement to make concrete
  - primary colors to make new colors



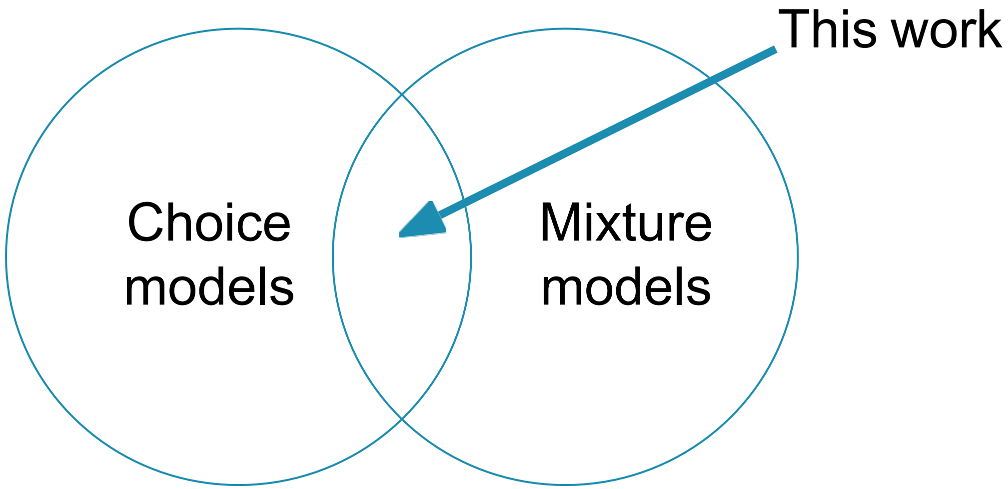
KU LEUVEN

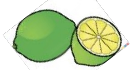
Mixtures

- In mixture experiments, products are expressed as combinations of **proportions** of ingredients
- The researchers' interest is generally in one or more characteristics of the mixture
- In this work, the characteristic of interest is the **preference** of respondents
- Choice experiments are ideal

Combining choice models and mixture models

Choice experiments with mixtures

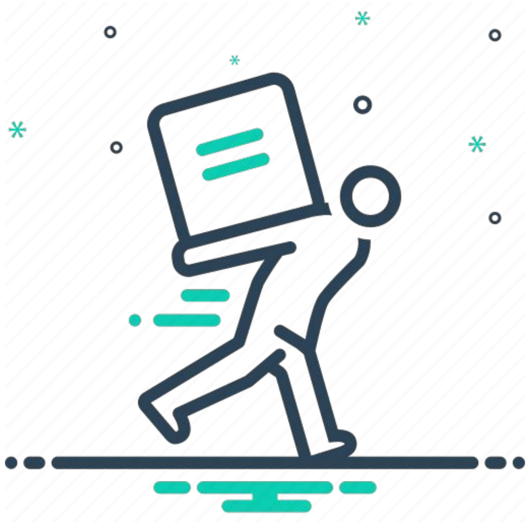


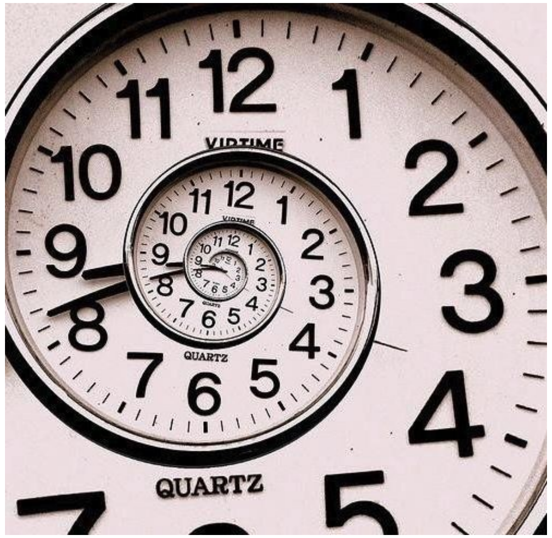
- First example by Courcoux and Séménou (1997), preferences for cocktails
  - mango juice 
  - lemon juice 
  - blackcurrant syrup 
- 60 people, each making 8 pairwise comparisons

KU LEUVEN

Designing choice experiments with mixtures

- Experiments are expensive, cumbersome and time-consuming
- Efficient experimental designs → reliable information
- Optimal design of experiments: the branch of statistics that deals with the construction of **efficient experimental designs**





Optimality criteria for choice experiments

Optimal choice experiments with mixtures

- **D-optimal** experimental designs → low-variance estimators
- We want to find a mixture that maximizes consumer preference
- Precise predictions are crucial
- **I-optimal** experimental designs → low-variance prediction

Models for data from mixture experiments

- Mixture models assume two or more ingredients and a response variable that depends only on the **relative** proportions of the ingredients in the mixture
- Each mixture is described as a combination of *q* ingredient proportions (0 to 1)
- Constraint: proportions sum up to one → perfect collinearity
- Special-cubic Scheffé model:

$$Y = \sum_{i=1}^{q} \beta_i x_i + \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} \beta_{ij} x_i x_j + \sum_{i=1}^{q-2} \sum_{j=i+1}^{q-1} \sum_{k=j+1}^{q} \beta_{ijk} x_i x_j x_k + \varepsilon$$

Including process variables

- The result of a mixture may depend on other characteristics
- Additional variables → **process variables**
- Second-order Scheffé model

$$Y = \sum_{k=1}^{q} \gamma_k^0 x_k + \sum_{k=1}^{q-1} \sum_{l=k+1}^{q} \gamma_{kl}^0 x_k x_l + \sum_{i=1}^{r} \sum_{k=1}^{q} \gamma_k^i x_k z_i + \sum_{i=1}^{r-1} \sum_{j=i+1}^{r} \alpha_{ij} z_i z_j + \sum_{i=1}^{r} \alpha_i z_i^2 + \varepsilon$$

Multinomial logit model for choice data

- A respondent faces *S* choice sets involving *J* alternatives each
- Respondent chooses the alternative that has the highest perceived utility
- The probability that a respondent chooses alternative *j* in choice set *s* is

$$p_{js} = \frac{\exp\left[\boldsymbol{f}^T(\boldsymbol{x}_{js})\boldsymbol{\beta}\right]}{\sum_{t=1}^{J} \exp\left[\boldsymbol{f}^T(\boldsymbol{x}_{ts})\boldsymbol{\beta}\right]}$$

Model for choice data concerning mixtures

- We assume vector $\boldsymbol{x}_{js}$ contains the *q* ingredient proportions and *r* process variables
- Perceived utility modeled as

$$
\begin{aligned}
u_{js} &= \boldsymbol{f}(\boldsymbol{x}_{js})^T \boldsymbol{\beta} \\
&= \sum_{i=1}^{q-1} \gamma_i^{0*} x_{ijs} + \sum_{i=1}^{q-1} \sum_{k=i+1}^{q} \gamma_{ik}^{0} x_{ijs} x_{kjs} + \sum_{i=1}^{r} \sum_{k=1}^{q} \gamma_k^{i} x_{kjs} z_{ijs} + \\
&\quad \sum_{i=1}^{r-1} \sum_{k=i+1}^{r} \alpha_{ik} z_{ijs} z_{kjs} + \sum_{i=1}^{r} \alpha_i z_{ijs}^2
\end{aligned}
$$

D-optimal designs

- D-optimality criterion

$$\mathcal{D} = \det\left(\boldsymbol{I}^{-1}(\boldsymbol{X}, \boldsymbol{\beta})\right) \longrightarrow \text{prior distribution } \pi(\boldsymbol{\beta})$$

D-optimal designs

- D-optimality criterion

$$\mathcal{D} = \det\left(\boldsymbol{I}^{-1}(\boldsymbol{X}, \boldsymbol{\beta})\right)$$

- Bayesian D-optimality criterion

$$\mathcal{D}_B = \int_{\mathbb{R}^m} \det\left(\boldsymbol{I}^{-1}(\boldsymbol{X}, \boldsymbol{\beta})\right) \pi(\boldsymbol{\beta}) d\boldsymbol{\beta}$$

- Numerical approximation to Bayesian D-optimality criterion

$$\mathcal{D}_B \approx \frac{1}{R} \sum_{i=1}^{R} \det\left(\boldsymbol{I}^{-1}(\boldsymbol{X}, \boldsymbol{\beta}^{(i)})\right)$$

I-optimal designs

- I-optimality criterion

$$\mathcal{I} = \int_{\chi} \boldsymbol{f}^T(\boldsymbol{x}_{js})\boldsymbol{I}^{-1}(\boldsymbol{X},\boldsymbol{\beta})\boldsymbol{f}(\boldsymbol{x}_{js})d\boldsymbol{x}_{js}$$

$$= \mathrm{tr}\left[\boldsymbol{I}^{-1}(\boldsymbol{X},\boldsymbol{\beta})\boldsymbol{W}_u\right]$$

- Bayesian I-optimality criterion

$$\mathcal{I}_B = \int_{\mathbb{R}^m} \mathrm{tr}\left[\boldsymbol{I}^{-1}(\boldsymbol{X},\boldsymbol{\beta})\boldsymbol{W}_u\right]\pi(\boldsymbol{\beta})d\boldsymbol{\beta}$$

- Numerical approximation to Bayesian I-optimality criterion

$$\mathcal{I}_B \approx \frac{1}{R}\sum_{i=1}^{R} \mathrm{tr}\left[\boldsymbol{I}^{-1}(\boldsymbol{X},\boldsymbol{\beta}^{(i)})\boldsymbol{W}_u\right]$$

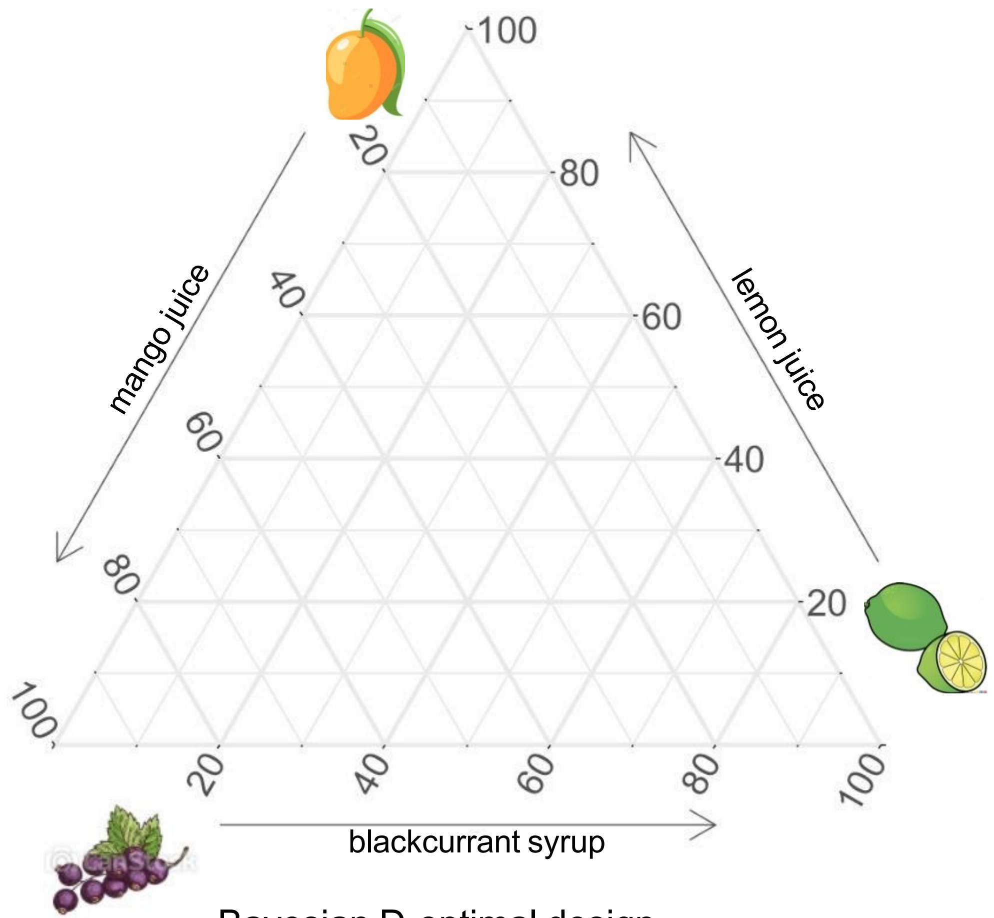
$$\boldsymbol{W}_u = \int_{\chi} \boldsymbol{f}(\boldsymbol{x}_{js})\boldsymbol{f}^T(\boldsymbol{x}_{js})d\boldsymbol{x}_{js}$$

Example

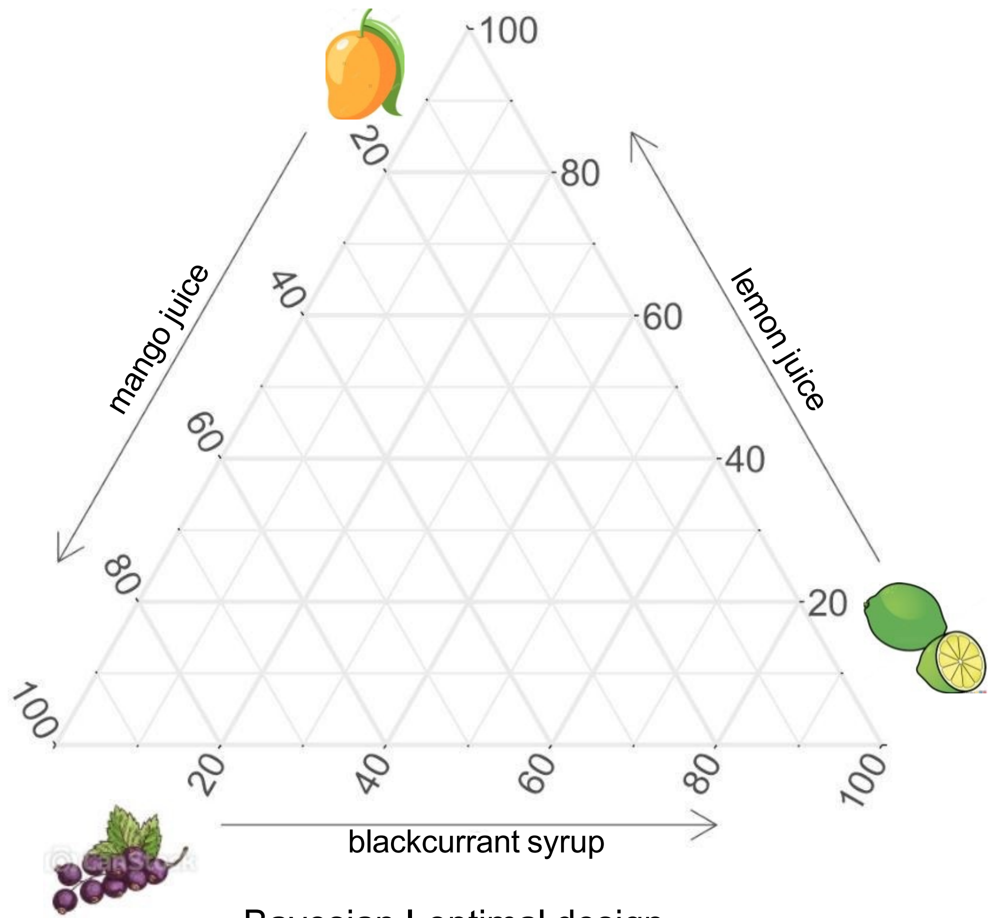
Cocktail preferences

- Original experiment by Courcoux and Semenou
- September 2019: students from KU Leuven replicated the experiment with 35 respondents
- Each respondent tasted 4 choice sets of size 2
- Simulated responses for temperature (process variable) → $\boldsymbol{\beta}$ parameter vector
- $\boldsymbol{\beta}$ used as prior distribution in a second-order Scheffé model and MNL model for **Bayesian D- and I-optimal** designs

Cocktail preferences



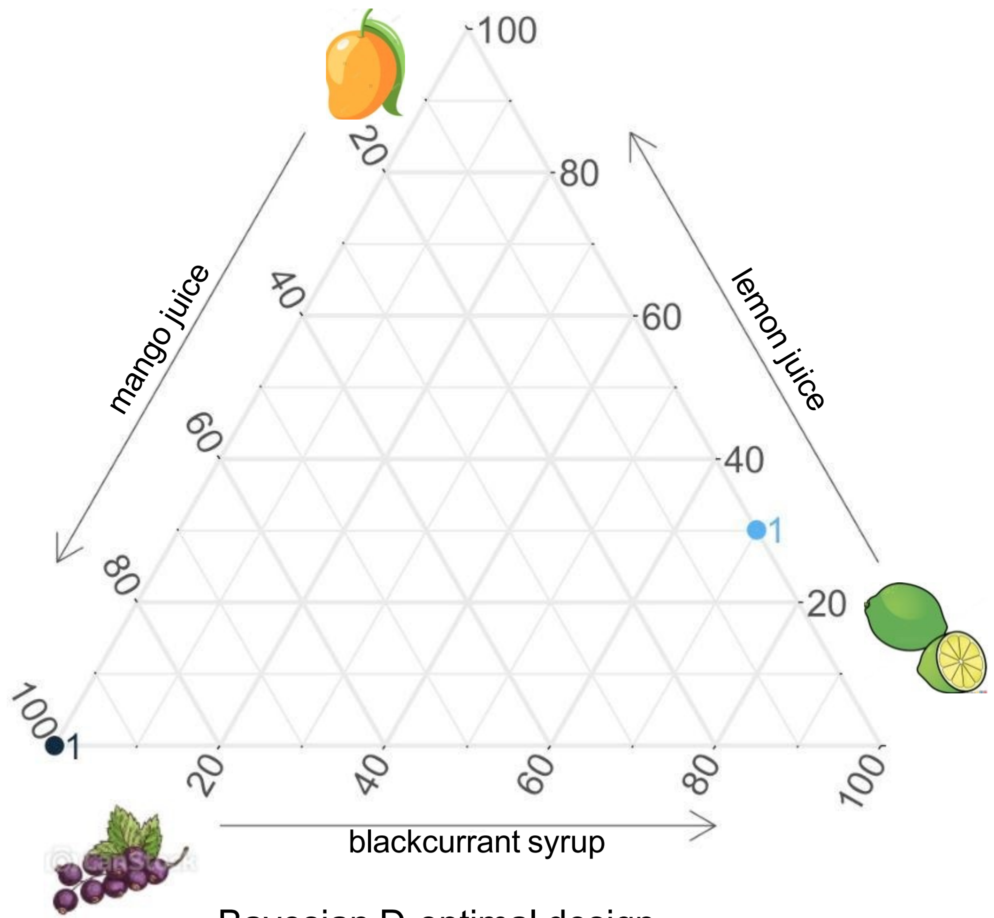
Bayesian D-optimal design



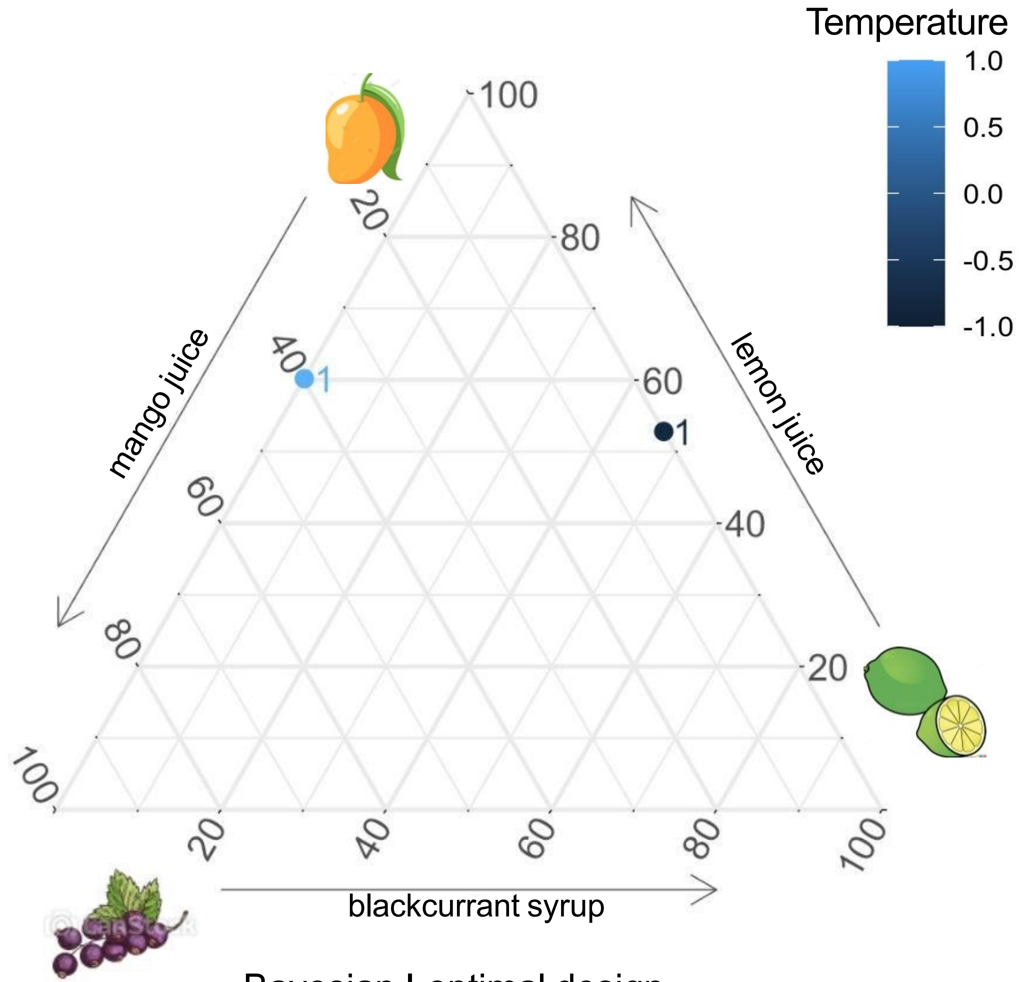
Bayesian I-optimal design

KU LEUVEN

Cocktail preferences

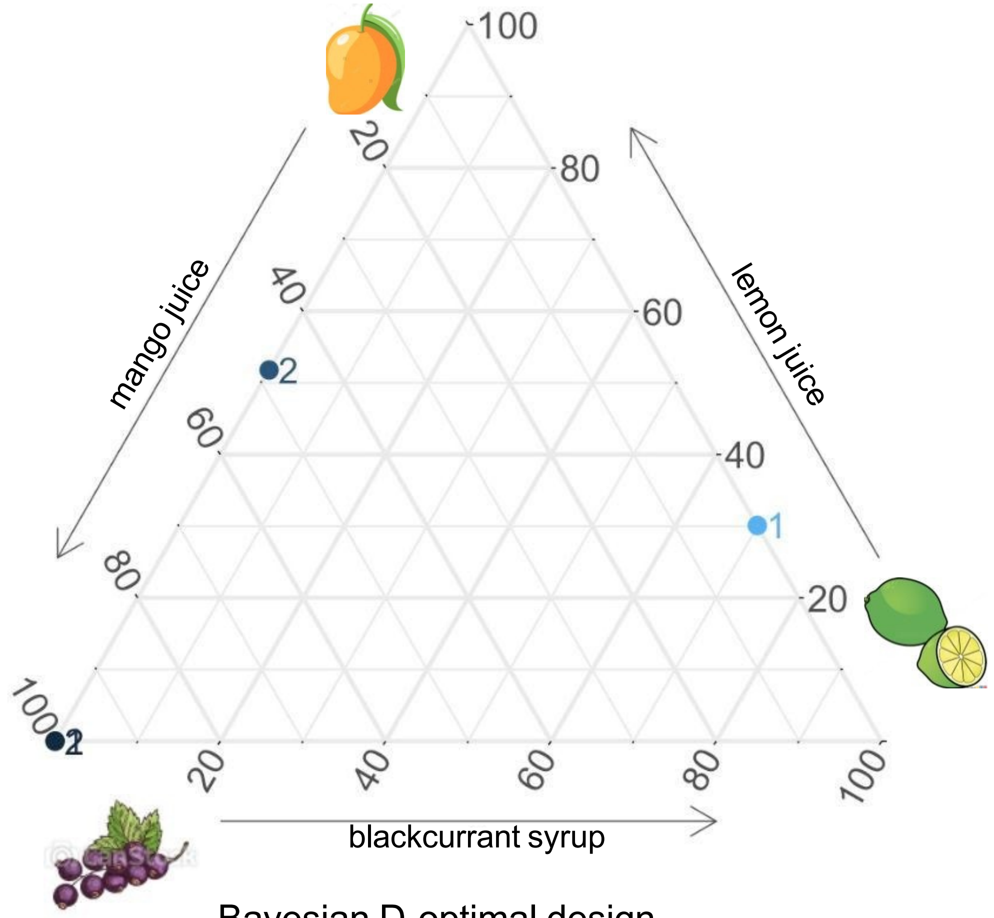


Bayesian D-optimal design

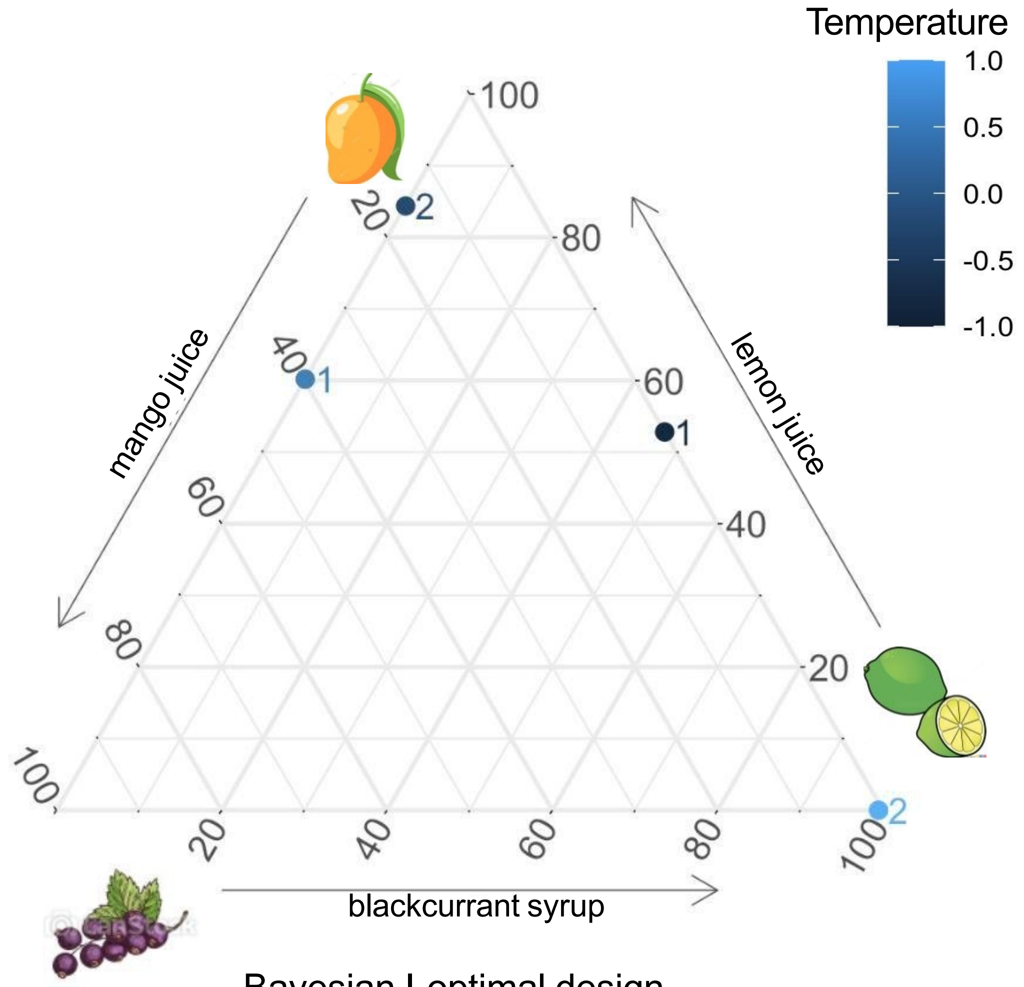


Bayesian I-optimal design

Cocktail preferences

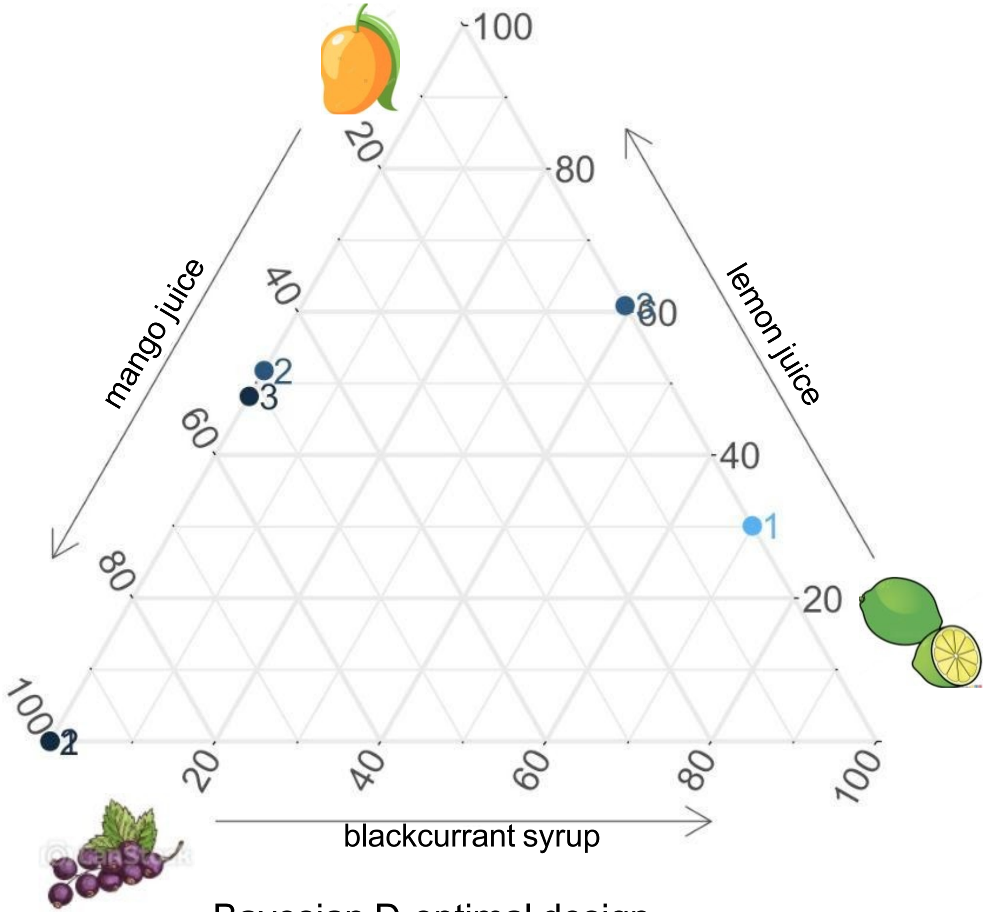


Bayesian D-optimal design

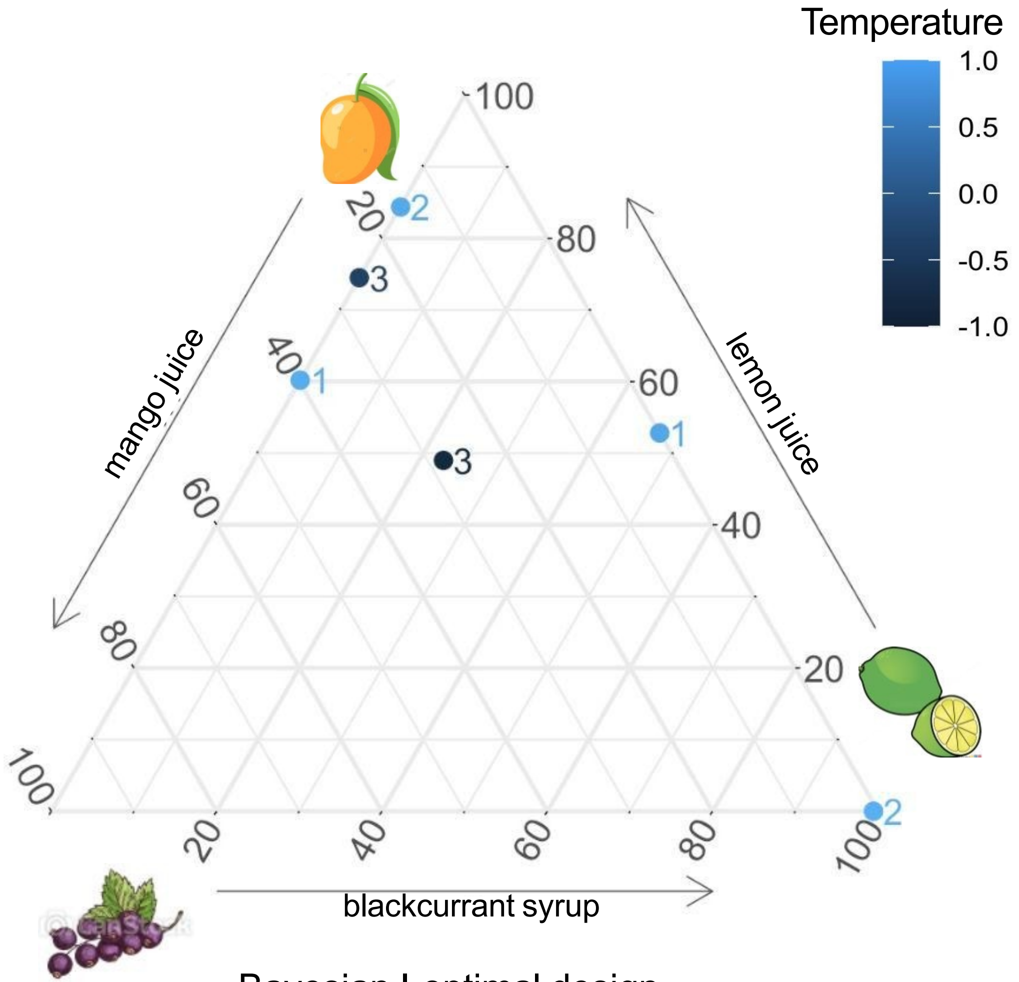


Bayesian I-optimal design

Cocktail preferences



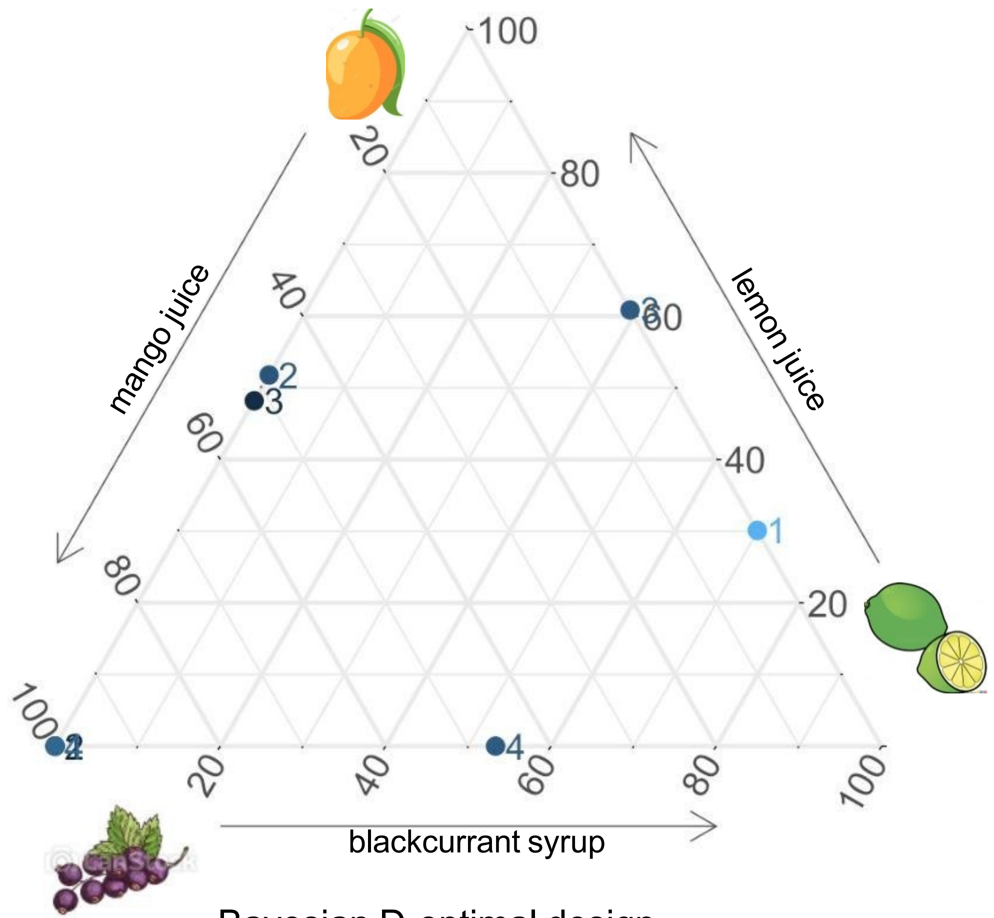
Bayesian D-optimal design



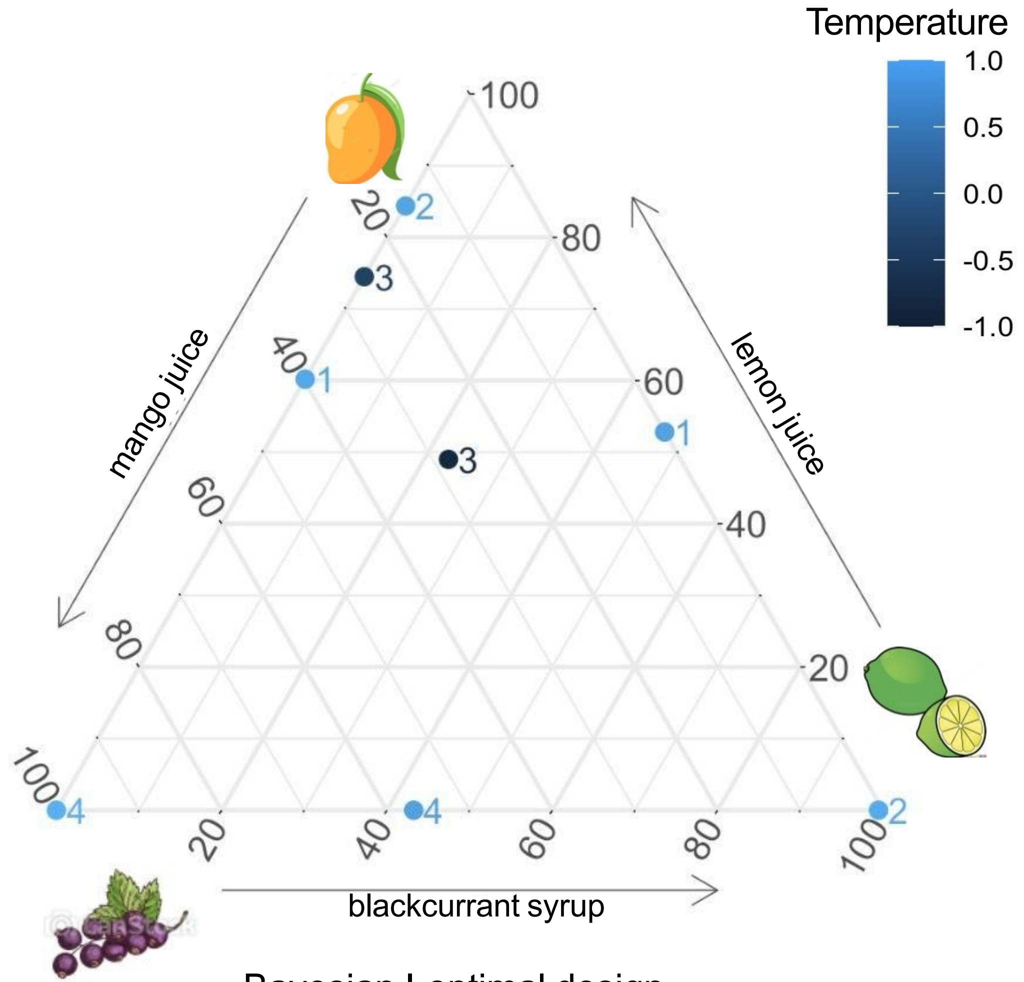
Bayesian I-optimal design

KU LEUVEN

Cocktail preferences

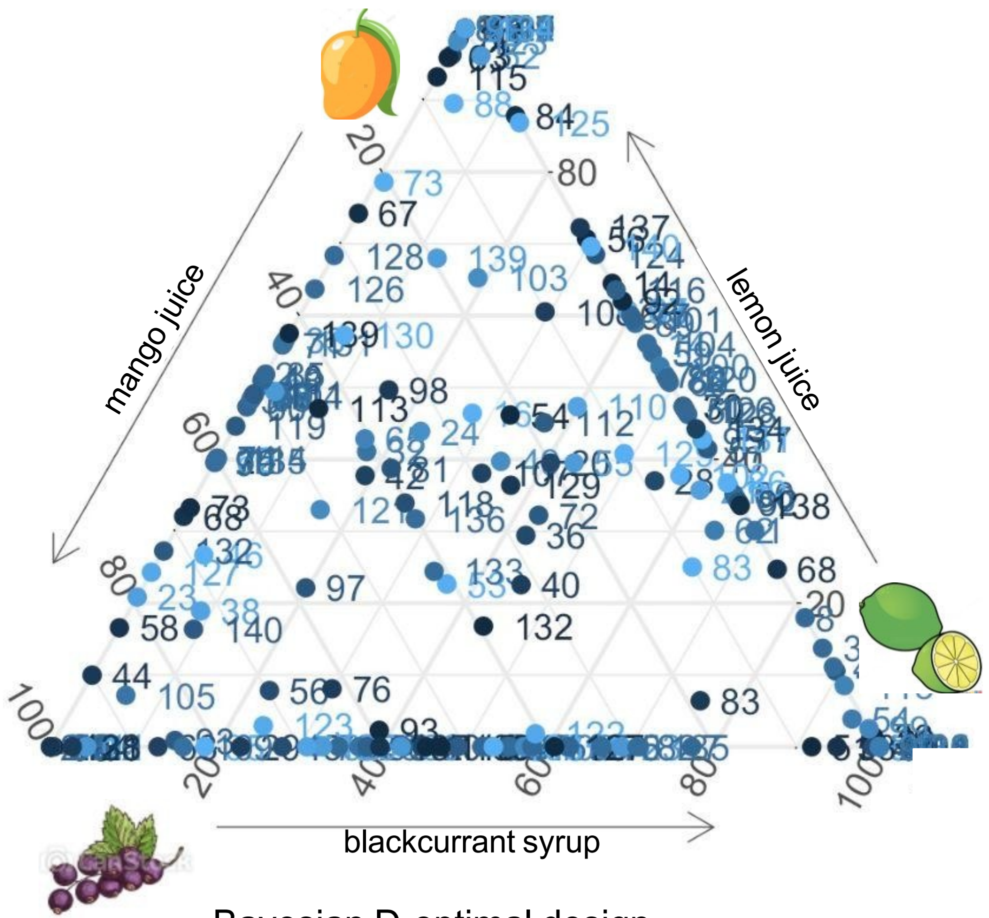


Bayesian D-optimal design

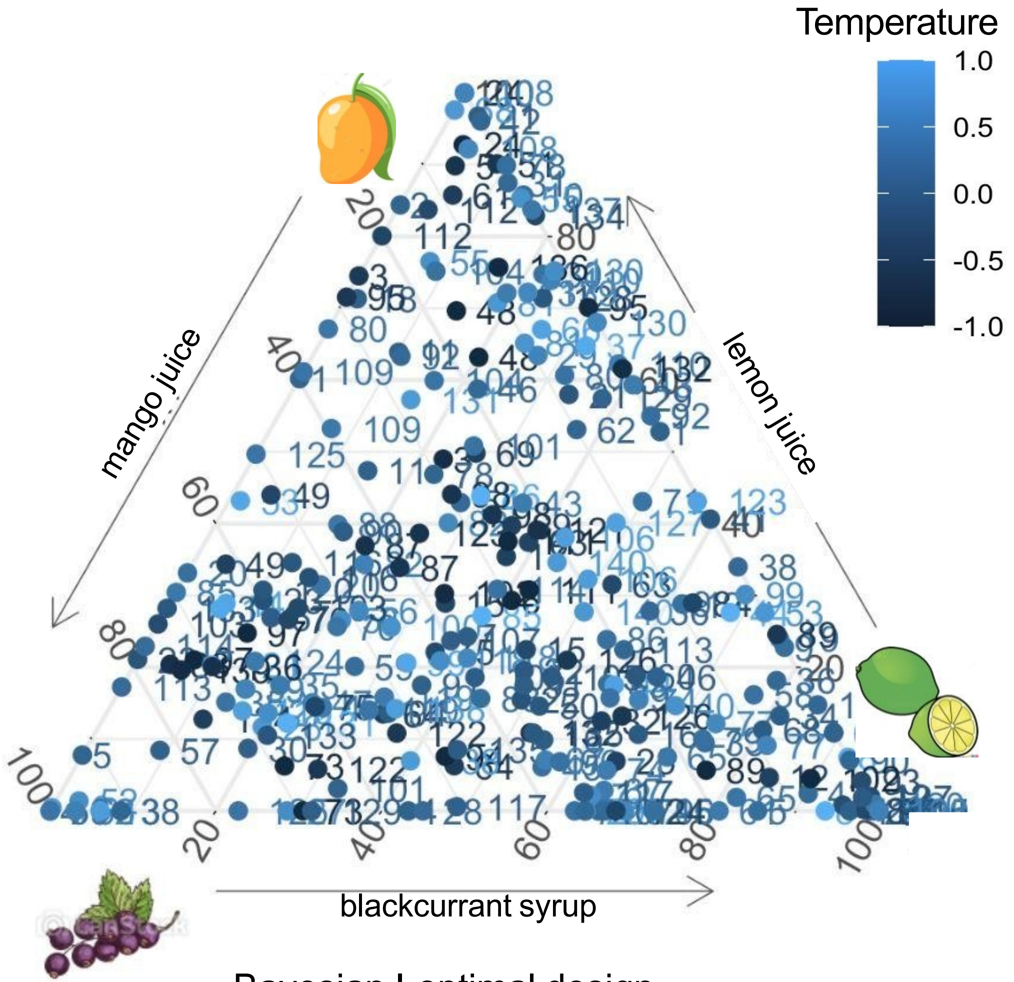


Bayesian I-optimal design

Cocktail preferences



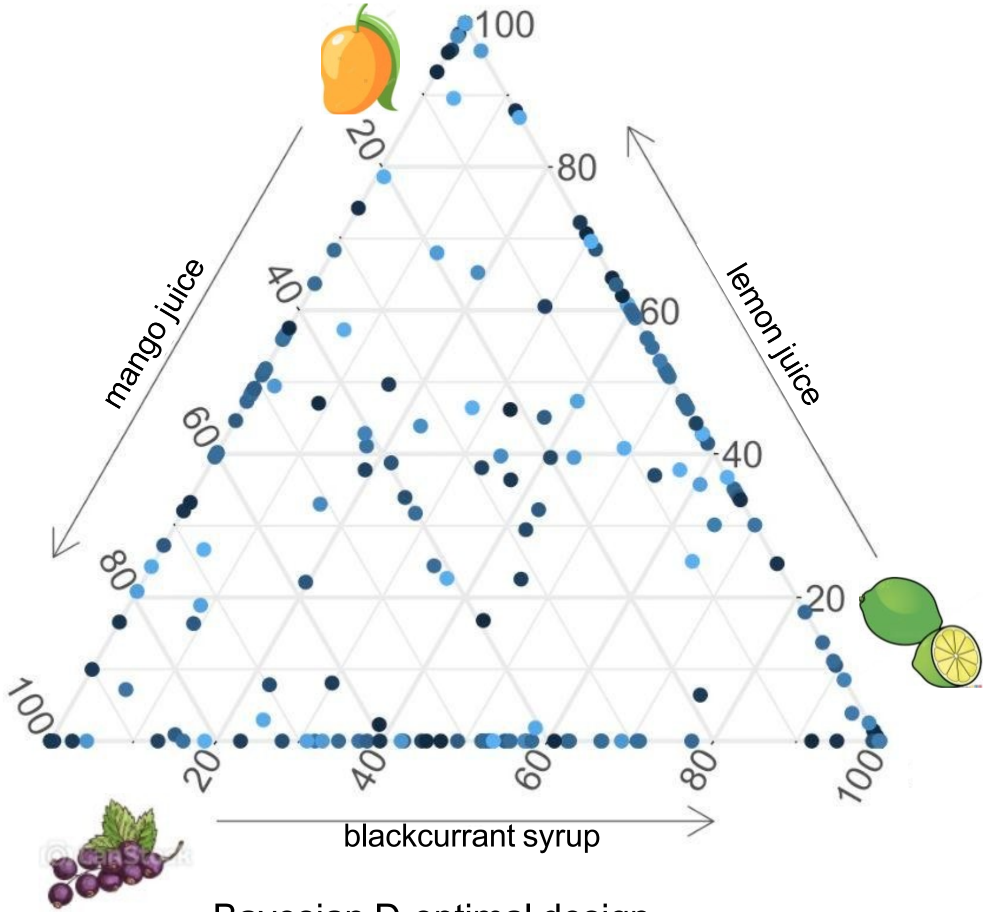
Bayesian D-optimal design



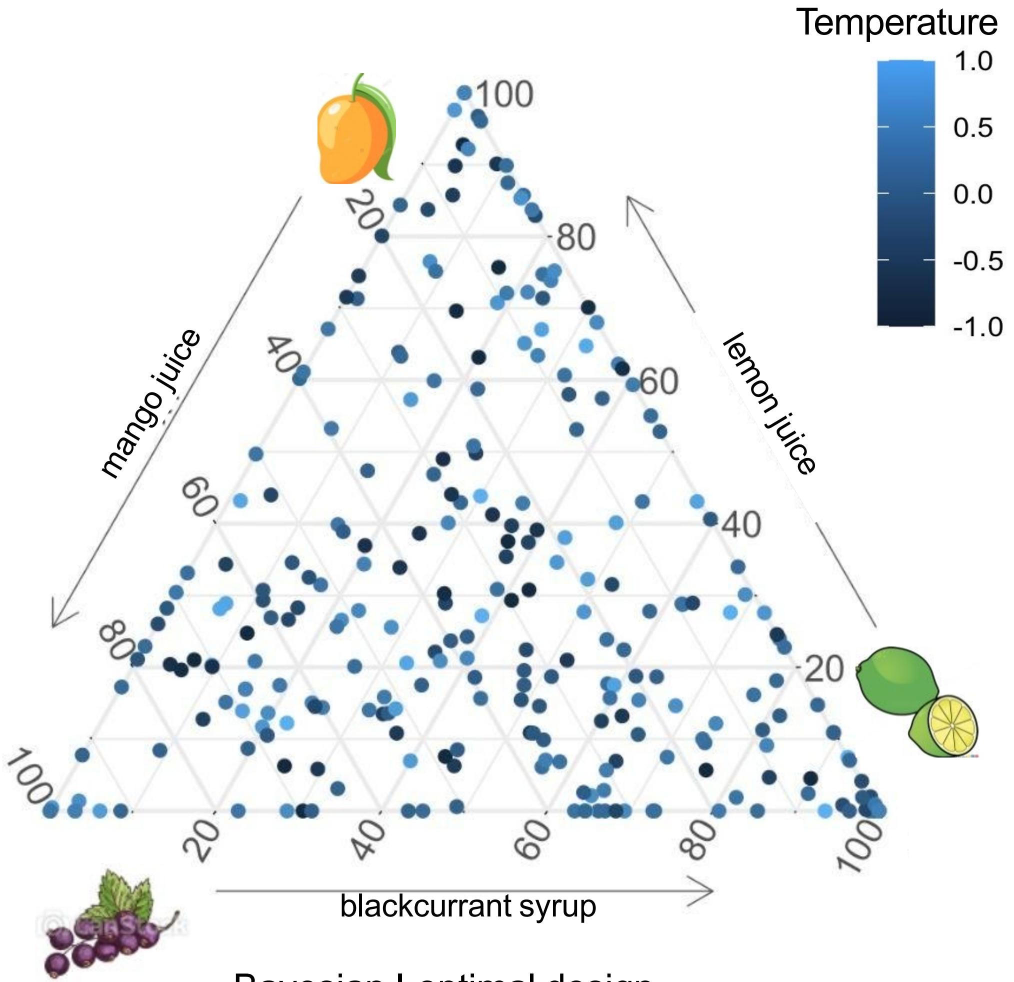
Bayesian I-optimal design

KU LEUVEN

Cocktail preferences

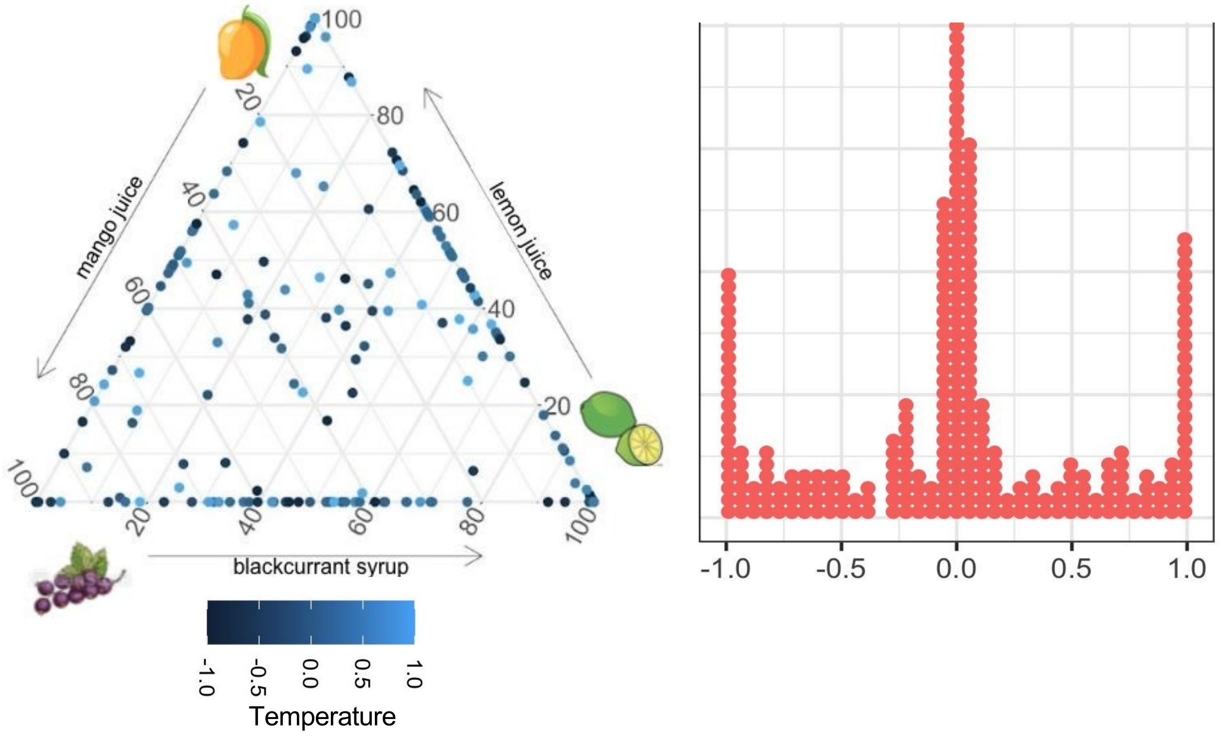


Bayesian D-optimal design

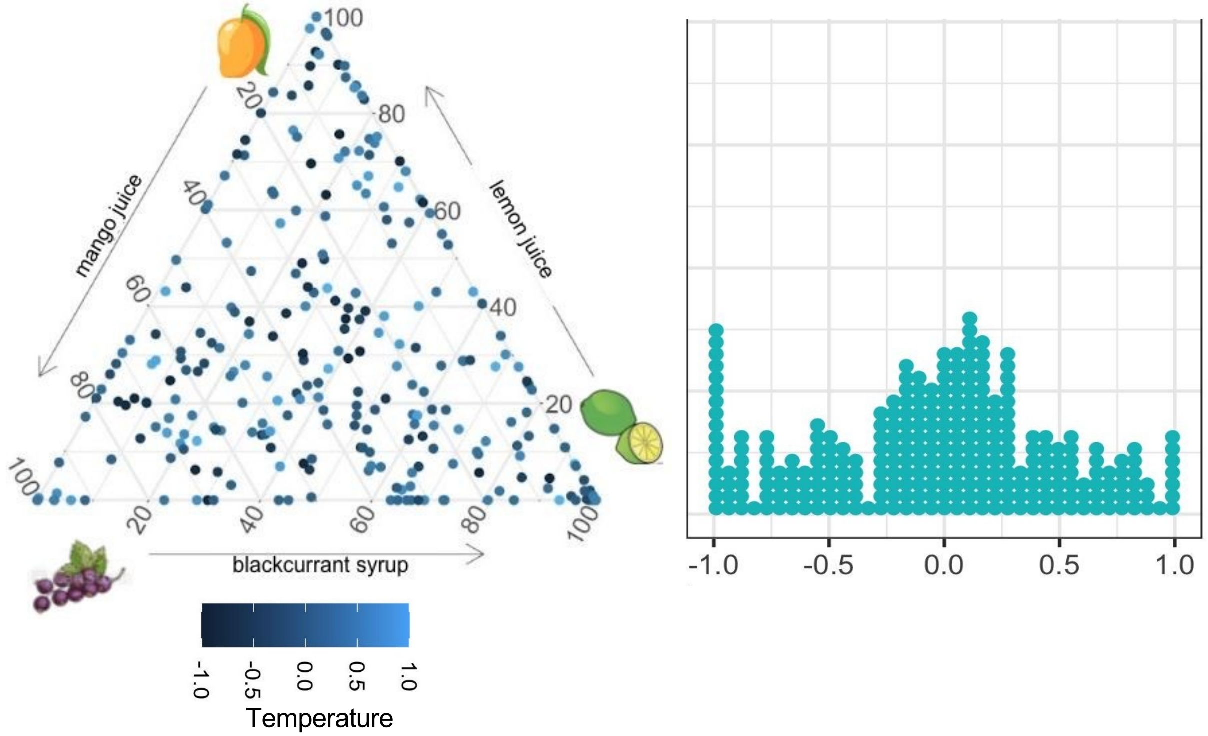


Bayesian I-optimal design

Cocktail preferences



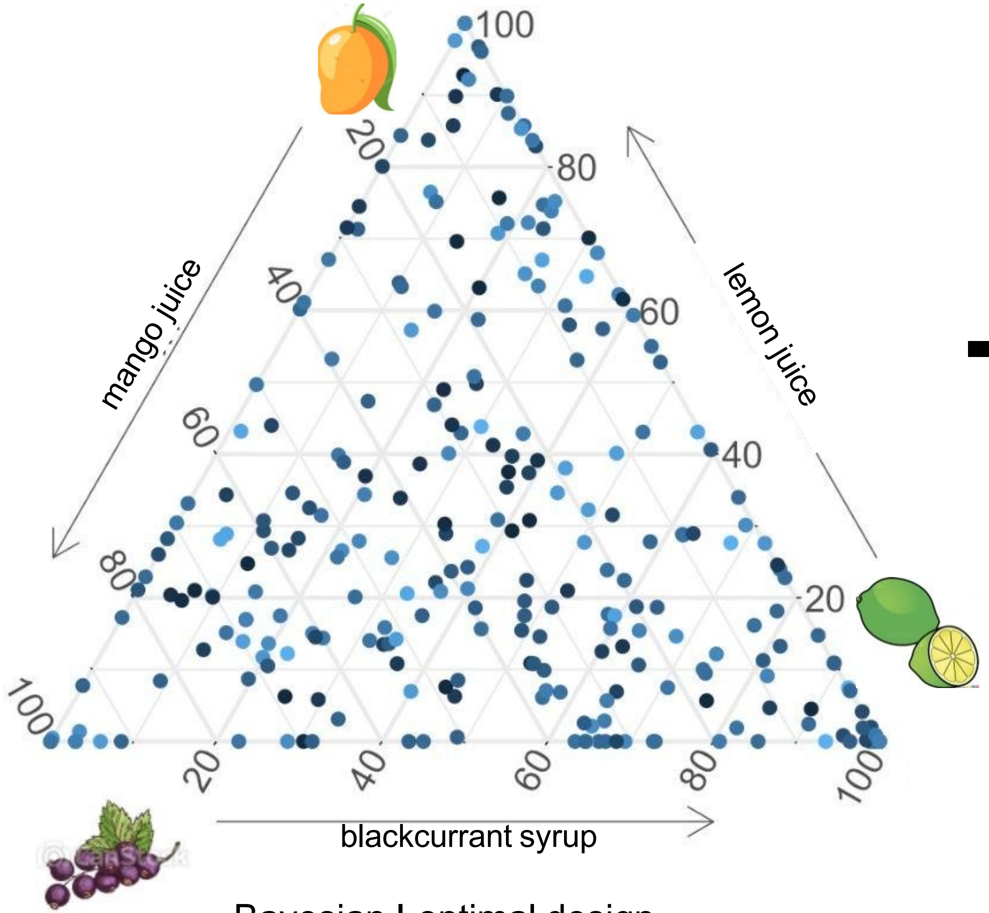
Bayesian D-optimal design



Bayesian I-optimal design

Cocktail preferences

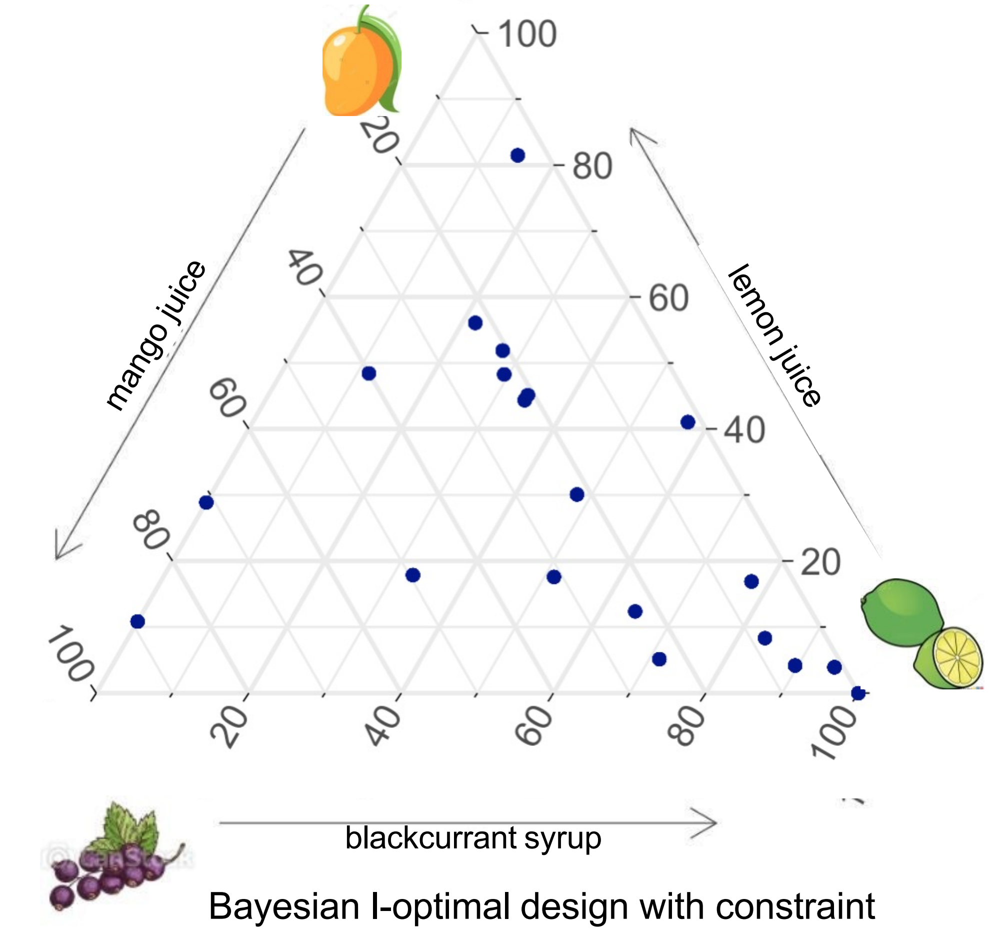
Upper bound on the number of distinct mixtures



Bayesian I-optimal design

261 different mixtures

I-opt = 2.72

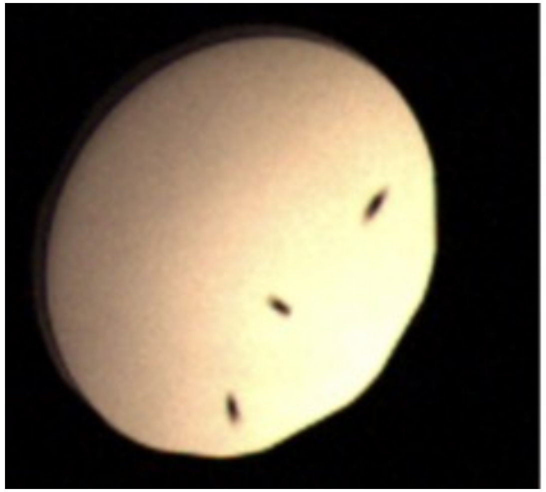


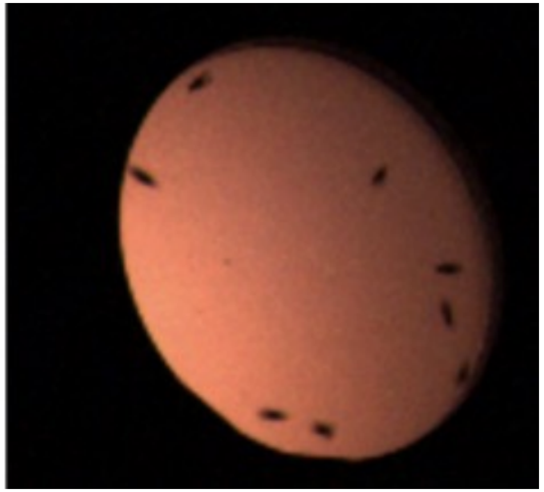
Bayesian I-optimal design with constraint

20 different mixtures

I-opt = 2.85

Fruit flies’ color preferences





More information

- *Bayesian I-optimal designs for choice experiments with mixtures* by Mario Becerra and Peter Goos. Chemometrics and Intelligent Laboratory Systems 217 (2021): 104395. DOI: 10.1016/j.chemolab.2021.104395

- *Bayesian D- and I-optimal designs for choice experiments involving mixtures and process variables* by Mario Becerra and Peter Goos. Food Quality and Preference. DOI: 10.1016/j.foodqual.2023.104928

- R package with our algorithms (https://github.com/mariobecerra/opdesmixr)

- Mario Becerra's website (with links to papers, R package, and code): mariobecerra.github.io/