Let $\mathcal{R}_k : \mathbb{R}^n \longrightarrow \mathbb{R}^{km}$ be the linear map that sends the pixelwise absorption values of a pixalated domain onto k X-ray projection images of m degrees of freedom. The map $\mathcal{R}_k$ can be identified with a matrix

$$\mathcal{R}_k = \begin{bmatrix} R_1 \\ \vdots \\ R_k \end{bmatrix} \in \mathbb{R}^{km \times n},$$

where $R_1, \dots, R_k \in \mathbb{R}^{m \times n}$ are the projection matrices of individual X-ray illuminations.

Unlike in [Burger et al, 2021], let us assume that the projection geometries for $R_1, \dots, R_k$ are fixed. Let us also assume that the unknown pixelwise absorption values are modeled as a random vector $X \in \mathbb{R}^n$ with a Gaussian (prior) distribution, $X \sim \mathcal{N}(x_{pr}, \Gamma_{pr})$, and that the measurement model is

belongs to $\mathbb{R}^{n \times n}$ and is symmetric pos. def.

$$Y = \mathcal{R}_k X + N,$$

where the additive noise N is indpendent of X and follows a zero-mean Gaussian distribution, $N \sim \mathcal{N}(0, \Gamma_{noise})$.

Under these assumptions, the posterior distribution of X, given a realized measurement $y = Y$, is $X|y \sim \mathcal{N}(x_{post}, \Gamma_{post})$, where

$$(*) \qquad \Gamma_{post} = \Gamma_{pr} - \Gamma_{pr}\mathcal{R}_k^T(\mathcal{R}_k\Gamma_{pr}\mathcal{R}_k^T + \Gamma_{noise})^{-1}\mathcal{R}\Gamma_{pr},$$

$$x_{post} = x_{pr} - \Gamma_{pr}\mathcal{R}_k^T(\mathcal{R}_k\Gamma_{pr}\mathcal{R}^T + \Gamma_{noise})^{-1}(y - \mathcal{R}_k x_{pr}).$$

The idea of Bayesian experimental design is to make Γ_{post} "as small as possible" in a certain sense. In your thesis the preferred measure of "smallness" is the A-optimality criterion measured by the A-optimality target function

$$\Phi_A(d) = \mathrm{tr}(A\Gamma_{post}(d)A^T), \quad A \in \mathbb{R}^{n\times n},$$

where $d \in \mathbb{R}^q$ is the design parameter and A is a predetermined weight matrix that can, e.g., define a region of interest inside the examined domain. If A is the identity matrix, then Φ_A measures the expected squared reconstruction error, i.e., $\mathbb{E}_X(|X - x_{post}|^2)$, where $|\cdot|$ denotes the Euclidean norm.

$q = km$

In order to define the design parameter $d \in \mathbb{R}^{km}$, let us consider first the noise model. Assume that the noise covariance Γ_{noise} is diagonal and that the corresponding diagonal elements are almost the squares of d, i.e., $\Gamma_{noise} = \mathrm{diag}(d_1^2, \ldots, d_{km}^2) + \varepsilon^2 I$, where ε^2 accounts for, e.g., digital noise that cannot be removed by increasing the radiation dose. (The mathematical reason for including ε in the model is guaranteeing that the inverse matrix in (*) exists.)
In other words, $d \in \mathbb{R}^{km}$ ~~almost~~ is almost the vector of standard deviations for N; note that it is also implicitly assumed that the components of noise are independent.

From (*) it is rather obvious that the minimizer of Φ_A without any extra constraints is $d = 0$. This would mean minimizing noise without any restrictions on the radiation dose.

Let us assume that the radiation dose is inversely propotional to the variance of noise, i.e.,

$$\mathrm{dose}_j = \frac{c}{d_j^2},$$

where $c>0$ is a positive constant. (I am not sure how physically accurate this model is, but at least it has a suitable behavior for our purposes.)

The assumption is that the overall dose is not allowed to exceed some predefined limit, say, $D>0$.

The complete minimization task thus is

[probably ~~concave~~ convex in d]

$$\begin{cases} d^* = \underset{d\in\mathbb{R}^{mk}}{\mathrm{argmin}}\Big(\underbrace{\mathrm{tr}(A\Gamma_{post}(d)A^T)}_{\Phi_A(d)}, \\ \displaystyle\sum_{j=1}^{km} \frac{1}{d_j^2} \le \frac{D}{c}, \end{cases} \qquad (\dagger)$$

[almost surely the minimizer satisfies equality here]

where

$$\Gamma_{post} = \Gamma_{pr} - \Gamma_{pr}\mathcal{R}_k^T(\mathcal{R}_k\Gamma_{pr}\mathcal{R}_k^T + \Gamma_{noise}(d))^{-1}\mathcal{R}_k\Gamma_{prior}.$$

The main steps are (approximately) as follows:

[June] 1. Try to understand these notes and the relevant parts of [Burger et al, 2021] and the program codes I sent you.

[June] 2. Deduce the gradient (and possibly) also the Hessian of the target function $\Phi_A(d)$ with respect to d. Help can be found, e.g., in the master's thesis of Jarno Maaninen.

[June] 3. Consider techniques of constraint optimization and choose suitable ones for tackling (†). (Material will be provided a bit later.)

[June & July] 4. Implement the algorithm and test it with some ~~sets of sim~~ settings in the spirit of the provided codes.

[July & August] 5. Write the thesis.