Let consider a general mixture model. We suppose that each component is known. We will focus on the estimation on the weights $\boldsymbol{\pi} \in \mathbb{R}^K$ vector. We observe N random variables $X_1, X_2, \dots, X_N$ drawn from the density $f_{\boldsymbol{\pi}}$:

$$f_{\boldsymbol{\pi}}(x) = \sum_{j=1}^{K} \boldsymbol{\pi}_j f_j(x) \tag{1}$$

Each component density f_i are known, not necessarily Gaussian. We assume that the weight vector $\boldsymbol{\pi}$ is sparse. We will focus on the Maximum Likelihhod Estimator's performance and it's asymptotical behaviour regarding the risk. We define $\Phi_N(\boldsymbol{\pi})$ as:

$$\Phi_N(\boldsymbol{\pi}) = -\frac{1}{N} \sum_{i=1}^{N} \log f_{\boldsymbol{\pi}}(x_i) \tag{2}$$

We can rewrite the minimization problem:

$$\widehat{\boldsymbol{\pi}} \in \arg\min_{\boldsymbol{\pi} \in \Pi} \left\{ \Phi_N(\boldsymbol{\pi}) \right\}, \quad \Pi = \left\{ \boldsymbol{\pi} \in [0,1]^K : \sum_{j=1}^{K} \pi_j = 1 \right\} \tag{3}$$

For theoretical objectives, we will make the following asumption:

Hypothesis 1. *All realizations are not probably unlikely to be observed. Therefore* $f_{\boldsymbol{\pi}}(x) \geq m > 0$ *for all* $x \in \{x_1, \dots, x_N\}$.

Let denote $M = max_{x \in x_1, \dots, x_N, j \in [K]} \{f_j(x)\}$, since $\boldsymbol{\pi}^T \mathbf{1} = 1$ then $f_{\boldsymbol{\pi}} \leq M$. Therefore, $\forall \boldsymbol{\pi} \in \Pi, f_{\boldsymbol{\pi}}(x_1, \dots, x_N) \in [m, M]$. We have the following lemma:

Lemma 1. *Under hypothesis 1,* Φ_N *is Lipschitz-smooth and strongly convex.*

Proof. $\forall i \in [K], g_{x_i}(\boldsymbol{\pi}) = f_{\boldsymbol{\pi}}(x_i)$ is a linear function defined on Π and it's image is the interval $[m, M]$ where $m > 0$. We will prove that $-log$ is strongly convex on $[m, M]$.

$$\forall x \in [m, M], \frac{1}{M^2} \leq \frac{d^2(-log)}{dx^2}(x) = \frac{1}{x^2} \leq \frac{1}{m^2} \tag{4}$$

The first inequality proves the $1/M^2$ strong convexity of $-log$, the second proves that it is $1/m^2$ Lipschitz smooth. The sum of strongly convex functions is strongly convex. Therefore, Φ_N is strongly convex. □

The minimization problem can be rewritten as following:

$$\widehat{\boldsymbol{\pi}} \in \arg\min_{\boldsymbol{\pi} \in \Pi} \left\{ \Phi_N(\boldsymbol{\pi}) \right\}, \quad \Pi = \left\{ \boldsymbol{\pi} \in [0,1]^K : \boldsymbol{\pi}^T \mathbf{1} = 1, \forall i \in [N], \sum_{j=1}^{K} \pi_j f_j(x_i) \geq m \right\} \tag{5}$$

We will study different loss function for the risk: $||\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi}^*||_1$, $||\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi}^*||_2$ and $dist(f_{\widehat{\boldsymbol{\pi}}}, f_{\boldsymbol{\pi}^*})$

This problem is close to the regression with random design since we can consider Φ_N as a function fo two random variable X_i and $\boldsymbol{\pi}$:

$$\Phi_N(\boldsymbol{\pi}) = \frac{1}{N}\sum_{i=1}^{N}\varphi(x_i, \boldsymbol{\pi}) \tag{6}$$

In this setting $\varphi(.,.)$ is strongly convex and Lipschitz smooth regarding the second variable