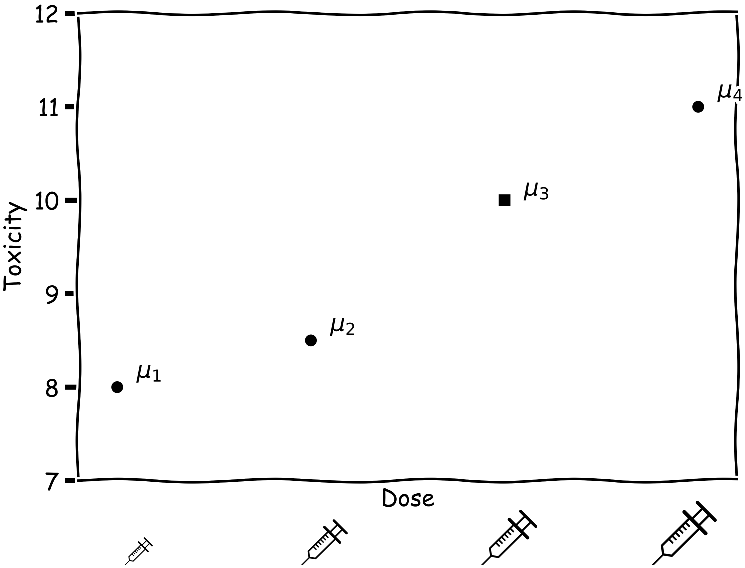
Threshold bandit for dose-ranging: The impact of monotonicity

Aurélien Garivier *Pierre Ménard* Laurent Rossi

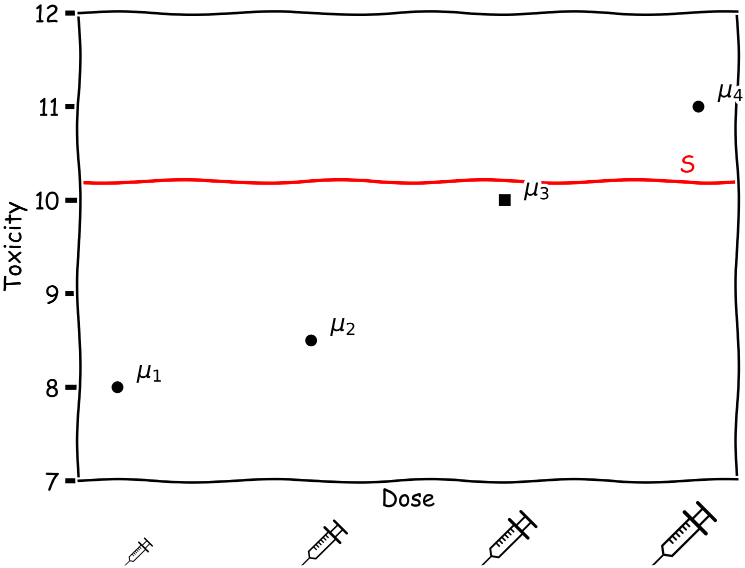
Institut de Mathématiques de Toulouse

May 24, 2018

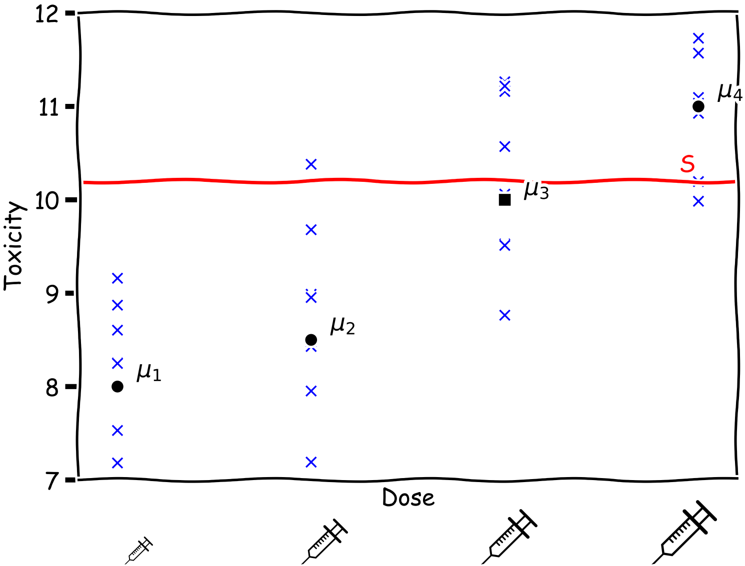
Dose-ranging



Dose-ranging



Dose-ranging



Best arm identification

$$\boldsymbol{\mu} = \quad (\mathcal{N}(\mu_1, 1) \quad \cdots \quad \mathcal{N}(\mu_a, 1) \quad \cdots \quad \mathcal{N}(\mu_K, 1))$$

 $\dots$  $\dots$ 

Best arm identification

$$\begin{array}{ccccccc} \boldsymbol{\mu} = & (\mathcal{N}(\mu_1, 1) & \cdots & \mathcal{N}(\mu_a, 1) & \cdots & \mathcal{N}(\mu_K, 1)) \\ \sim & [\mu_1 & \cdots & \mu_a & \cdots & \mu_K] \end{array}$$

Goal: Find the optimal arm (dose): $a^*_{\boldsymbol{\mu}} \in \arg\min_{1 \leqslant a \leqslant K} |\mu_a - S|$

Best arm identification

$$\begin{array}{cccccc} \boldsymbol{\mu} = & (\mathcal{N}(\mu_1, 1) & \cdots & \mathcal{N}(\mu_a, 1) & \cdots & \mathcal{N}(\mu_K, 1)) \\ \sim & [\mu_1 & \cdots & \mu_a & \cdots & \mu_K] \end{array}$$

Goal: Find the optimal arm (dose): $a^*_{\boldsymbol{\mu}} \in \arg\min_{1 \leqslant a \leqslant K} |\mu_a - S|$
Game: while $t < \tau$:

1. Player pulls arm (dose) $A_t \in \{1, .., K\}$.
2. He gets an observation (toxicity) $Y_t \sim \mathcal{N}(\mu_{A_t}, 1)$.

Predict best arm $\widehat{a}_\tau$.

Best arm identification

$$\begin{array}{cccccc} \boldsymbol{\mu} = & (\mathcal{N}(\mu_1, 1) & \cdots & \mathcal{N}(\mu_a, 1) & \cdots & \mathcal{N}(\mu_K, 1)) \\ \sim & [\mu_1 & \cdots & \mu_a & \cdots & \mu_K] \end{array}$$

Goal: Find the optimal arm (dose): $a^*_{\boldsymbol{\mu}} \in \arg\min_{1 \leqslant a \leqslant K} |\mu_a - S|$
Game: while $t < \tau$:

1. Player pulls arm (dose) $A_t \in \{1, .., K\}$.
2. He gets an observation (toxicity) $Y_t \sim \mathcal{N}(\mu_{A_t}, 1)$.

Predict best arm $\widehat{a}_\tau$.

δ-**correct algorithm.** ($\mathcal{F}_t = \sigma(Y_1, \ldots, Y_t)$ information available at step t)

- a **sampling rule** $(A_t)_{t \geqslant 1}$, where A_t is $\mathcal{F}_{t-1}$-measurable;
- a **stopping rule** τ, stopping time with respect to the filtration $(\mathcal{F}_t)_{t \geqslant 1}$;
- a $\mathcal{F}_\tau$-measurable **decision rule** $\widehat{a}_\tau$;

An algorithm is δ-correct if $\mathbb{P}_{\boldsymbol{\mu}}(\widehat{a}_\tau \neq a^*_{\boldsymbol{\mu}}) \leqslant \delta$ and $\mathbb{P}_{\boldsymbol{\mu}}(\tau < +\infty) = 1$.

Goal: find a δ-correct algorithm that minimize $\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]$.

$\rightarrow$ lower bound on $\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]$?

Lower bound

$\mathcal{M} = \{\boldsymbol{\mu} \in \mathbb{R}^K : a^*_{\boldsymbol{\mu}} \text{ is unique}\}$ $\quad$ $\mathcal{I} = \{\boldsymbol{\mu} \in \mathcal{M} : \mu_1 < \ldots < \mu_K\}$

Alternative set for $\mathcal{S} \in \{\mathcal{M}, \mathcal{I}\}$: $\mathcal{A}lt(\boldsymbol{\mu}, \mathcal{S}) := \{\boldsymbol{\lambda} \in \mathcal{S} : a^*_{\boldsymbol{\lambda}} \neq a^*_{\boldsymbol{\mu}}\}$.

Lower bound

$$\mathcal{M} = \{\boldsymbol{\mu} \in \mathbb{R}^K : a^*_{\boldsymbol{\mu}} \text{ is unique}\} \qquad \mathcal{I} = \{\boldsymbol{\mu} \in \mathcal{M} : \mu_1 < \dots < \mu_K\}$$

Alternative set for $\mathcal{S} \in \{\mathcal{M}, \mathcal{I}\}$: $\mathcal{A}lt(\boldsymbol{\mu}, \mathcal{S}) := \{\boldsymbol{\lambda} \in \mathcal{S} : a^*_{\boldsymbol{\lambda}} \neq a^*_{\boldsymbol{\mu}}\}$.

Theorem

Let $\mathcal{S} \in \{\mathcal{M}, \mathcal{I}\}$. For all δ-correct algorithm,

$$\liminf_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} \geqslant T^*_{\mathcal{S}}(\boldsymbol{\mu}),$$

where the characteristic time is

$$T^*_{\mathcal{S}}(\boldsymbol{\mu})^{-1} = \sup_{\omega \in \Sigma_K} \inf_{\boldsymbol{\lambda} \in \mathcal{A}lt(\boldsymbol{\mu}, \mathcal{S})} \sum_{a=1}^{K} \omega_a \frac{(\mu_a - \lambda_a)^2}{2},$$

where Σ_K is the simplex of dimension $K - 1$.

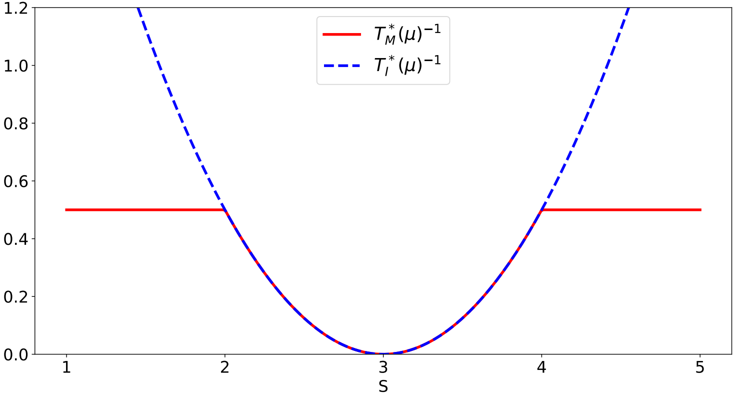
Two arms

If $K = 2$,

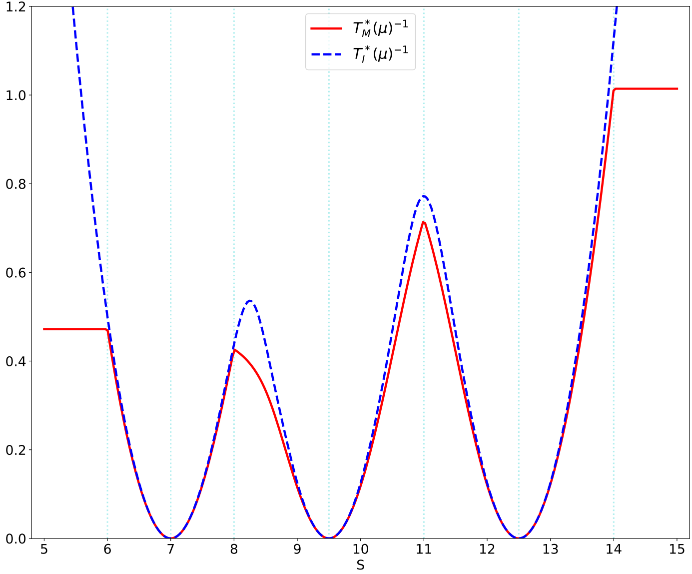
$$T^*_{\mathcal{I}}(\boldsymbol{\mu})^{-1} = (2S - \mu_1 - \mu_2)^2/8$$

$$T^*_{\mathcal{M}}(\boldsymbol{\mu})^{-1} = \min\Big((2S - \mu_1 - \mu_2)^2, (\mu_1 - \mu_2)^2\Big)/8\,.$$

$\boldsymbol{\mu} = [2,\ 4]$



K arms: $\boldsymbol{\mu} = [6,\ 8,\ 11,\ 14]$



K arms: optimal weights

Optimal proportions of arm draws: (Σ_K simplex of dimension $K-1$.)

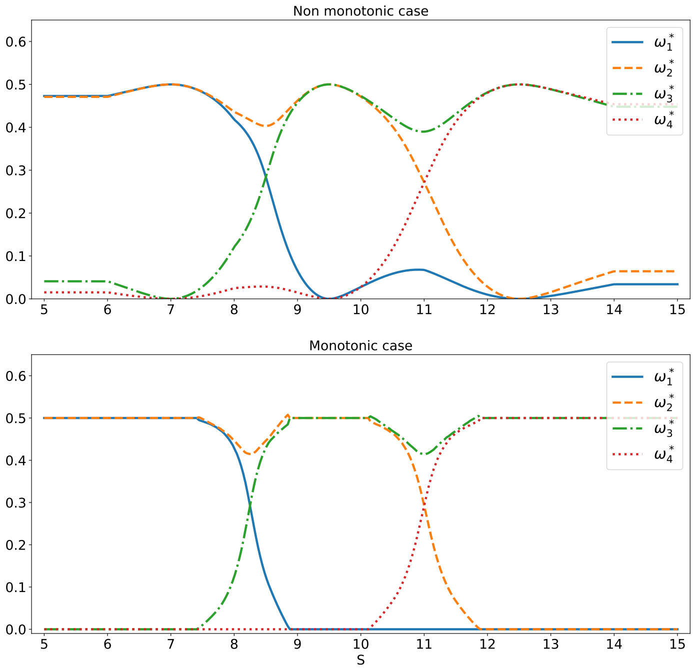
$$\omega^*(\boldsymbol{\mu}) := \arg\max_{\omega\in\Sigma_K} \inf_{\boldsymbol{\lambda}\in\mathcal{A}lt(\boldsymbol{\mu},\mathcal{S})} \sum_{a=1}^{K} \omega_a \frac{(\mu_a-\lambda_a)^2}{2} .$$

Case $S=\mathcal{M}$: Similar properties as in the BAI problem, in particular

- $\omega_a^*(\boldsymbol{\mu}) > 0$ for all a, for all $\boldsymbol{\mu}\in S$

Case $S=\mathcal{I}$:

- $\omega_a^*(\boldsymbol{\mu}) > 0$ for $a\in\{a^*_{\boldsymbol{\mu}}-1, a^*_{\boldsymbol{\mu}}, a^*_{\boldsymbol{\mu}}+1\}$, otherwise $\omega_a^*(\boldsymbol{\mu}) = 0$


Non monotonic case
ω_1^*
ω_2^*
ω_3^*
ω_4^*
0.0
0.1
0.2
0.3
0.4
0.5
0.6
5
6
7
8
9
10
11
12
13
14
15
Monotonic case
ω_1^*
ω_2^*
ω_3^*
ω_4^*
S

Asymptotically optimal algorithm

Algorithm 1: Direct-tracking

Sampling rule

$$A_{t+1} \in \begin{cases} \underset{a \in U_t}{\text{argmin}}\ N_a(t) \text{ if one } N_a(t) \text{ "too small"} & (\textit{forced exploration}) \\ \underset{1 \leqslant a \leqslant K}{\text{argmax}}\ \omega_a^*(\widehat{\boldsymbol{\mu}}(t)) - N_a(t)/t & (\textit{direct tracking}) \end{cases}$$

Stopping rule

$$\tau_\delta = \inf \left\{ t \in \mathbb{N}^* : \widehat{\boldsymbol{\mu}}(t) \in \mathcal{M} \text{ and } \inf_{\boldsymbol{\lambda} \in \mathcal{A}lt\left(\widehat{\boldsymbol{\mu}}(t), \mathcal{S}\right)} \sum_{a=1}^{K} N_a(t) \frac{(\widehat{\mu}_a(t) - \lambda_a)^2}{2} > \beta(t, \delta) \right\}$$

Decision rule

$$\widehat{a}_\tau \in \underset{1 \leqslant a \leqslant K}{\arg\min} \left| \widehat{\mu}_a(\tau) - S \right| .$$

Theorem (Asymptotic optimality)

*For $\mathcal{S} \in \{\mathcal{I}, \mathcal{M}\}$, for $\beta(t,\delta) = \log(tC/\delta) + (3K+2)\,\mathrm{loglog}(tC/\delta)$ Direct-tracking is δ-**correct on** $\mathcal{S}$ and **asymptotically optimal**, i.e.*

$$\limsup_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} \leqslant T^*_{\mathcal{S}}(\boldsymbol{\mu})\,.$$

where

$$C := e^{K+1}\left(\frac{2}{K}\right)^K (2(3K+2))^{3K}\frac{4}{\log(3)}\,.$$

Practical implementation

How to compute the optimal weights ?

$$\omega^*(\boldsymbol{\mu}) = \operatorname*{arg\,max}_{\omega \in \Sigma_K} \inf_{\boldsymbol{\lambda} \in \mathcal{A}lt(\boldsymbol{\mu}, \mathcal{S})} \sum_{a=1}^{K} \omega_a \frac{(\mu_a - \lambda_a)^2}{2} .$$

Non-monotonous case: $\mathcal{S} = \mathcal{M} \rightarrow$ boils down to solve one scalar equation
$\rightarrow$ Fast !

Increasing case: $\mathcal{S} = \mathcal{I} \rightarrow$gradient ascent with isotonic regression
$\rightarrow$ Relatively fast !

Monotonous case: $\mathcal{S} = \mathcal{I}$

$\mathcal{I}_b := \{\boldsymbol{\lambda} \in \mathcal{I}, a^*_\lambda = b\}$

$$F : \omega \mapsto \inf_{\boldsymbol{\lambda} \in \mathcal{A}lt(\boldsymbol{\mu}, \mathcal{I})} \sum_{a=1}^{K} \omega_a \frac{(\mu_a - \lambda_a)^2}{2} = \min_{b \neq a^*_{\boldsymbol{\mu}}} \inf_{\boldsymbol{\lambda} \in \mathcal{I}_b} \sum_{a=1}^{K} \omega_a \frac{(\mu_a - \lambda_a)^2}{2}, \quad (1)$$

F is concave $\rightarrow$ sub-gradient ascent on the simplex.

Monotonous case: $\mathcal{S} = \mathcal{I}$

$\mathcal{I}_b := \{\boldsymbol{\lambda} \in \mathcal{I}, a^*_\lambda = b\}$

$$F : \omega \mapsto \inf_{\boldsymbol{\lambda} \in \mathcal{A}lt(\boldsymbol{\mu}, \mathcal{I})} \sum_{a=1}^{K} \omega_a \frac{(\mu_a - \lambda_a)^2}{2} = \min_{b \neq a^*_{\boldsymbol{\mu}}} \inf_{\boldsymbol{\lambda} \in \mathcal{I}_b} \sum_{a=1}^{K} \omega_a \frac{(\mu_a - \lambda_a)^2}{2}, \qquad (1)$$

F is concave $\rightarrow$ sub-gradient ascent on the simplex.
Optimal alternative in $\overline{\mathcal{I}}_b$:

$$\boldsymbol{\lambda}^b := \arg\min_{\lambda \in \overline{\mathcal{I}}_b} \sum_{a=1}^{K} \omega_a \frac{(\mu_a - \lambda_a)^2}{2},$$

Sub-gradient of F at ω

$$\partial F(\omega) = \operatorname*{Conv}_{b \in B_{Opt}} \left[\frac{(\mu_a - \lambda^b_a)^2}{2} \right]_{a \in \{1, \cdots, K\}},$$

where Conv denotes the convex hull and B_{Opt} the set of points that attain the minimum in (1).

How to compute $\boldsymbol{\lambda}^b$?

$$\boldsymbol{\lambda}^b := \operatorname*{arg\,min}_{\lambda \in \overline{\mathcal{I}}_b} \sum_{a=1}^{K} \omega_a \frac{(\mu_a - \lambda_a)^2}{2} ,$$

$\sim$ to compute the Unimodal regression $\widehat{\boldsymbol{\lambda}}^b$ of $\boldsymbol{\mu}'$ with mode at b:

$$\widehat{\boldsymbol{\lambda}}^b = \operatorname*{arg\,min}_{\substack{\lambda'_1 \leqslant \ldots \leqslant \lambda'_b \\ \lambda'_K \leqslant \ldots \leqslant \lambda'_b}} \sum_{a=1}^{K} \omega_a \frac{(\mu'_a - \lambda'_a)^2}{2} .$$

Unimodal regression can be efficiently computed via isotonic regressions (Pool Adjacent Violators Algorithm) in $O(K)$.

$\rightarrow$ sub-gradient in $O(K^2)$.

Open questions

- Extend to any one-exponential family ? For example

$$\boldsymbol{\mu} = (\mathcal{B}(\mu_1), \dots, \ \mathcal{B}(\mu_a), \dots, \ \mathcal{B}(\mu_K)),$$

- Compute the sub-gradient in $O(K)$.
- Find algorithm with better theoretical/practical properties ("finite δ" bound).

How to compute $\boldsymbol{\lambda}^b$?

$$\mathcal{I}_b = \{\boldsymbol{\lambda} \in \mathcal{I} : a^*_{\boldsymbol{\lambda}} = b\} = \{\boldsymbol{\lambda} \in \mathcal{M} :$$
$$\lambda_1 < \ldots < \lambda_{b-1} < \min(\lambda_b, 2S - \lambda_b) \leqslant \max(\lambda_b, 2S - \lambda_b) < \lambda_{b+1} < \ldots < \lambda_K\}\,.$$

How to compute $\boldsymbol{\lambda}^b$?

$$\mathcal{I}_b = \{\boldsymbol{\lambda} \in \mathcal{I} : a^*_{\boldsymbol{\lambda}} = b\} = \{\boldsymbol{\lambda} \in \mathcal{M} : \\ \lambda_1 < \ldots < \lambda_{b-1} < \min(\lambda_b, 2S - \lambda_b) \leqslant \max(\lambda_b, 2S - \lambda_b) < \lambda_{b+1} < \ldots < \lambda_K\}\,.$$

Suppose $\mu_b \leqslant S$ then $\lambda_b^b < S \rightarrow$ replace $\mathcal{I}_b$ with

$$\begin{aligned}\{\boldsymbol{\lambda} \in \mathcal{M} : &\lambda_1 < \ldots < \lambda_{b-1} < \lambda_b, \\ &2S - \lambda_K < \ldots < 2S - \lambda_{b+1} < \lambda_b, \\ &\lambda_b \leqslant S\}\,.\end{aligned}$$

How to compute $\boldsymbol{\lambda}^b$?

$$\mathcal{I}_b = \{\boldsymbol{\lambda} \in \mathcal{I} : a^*_{\boldsymbol{\lambda}} = b\} = \{\boldsymbol{\lambda} \in \mathcal{M} : \\ \lambda_1 < \ldots < \lambda_{b-1} < \min(\lambda_b, 2S - \lambda_b) \leqslant \max(\lambda_b, 2S - \lambda_b) < \lambda_{b+1} < \ldots < \lambda_K\}\,.$$

Suppose $\mu_b \leqslant S$ then $\lambda^b_b < S \to$ replace $\mathcal{I}_b$ with

$$\begin{aligned}\{\boldsymbol{\lambda} \in \mathcal{M} :& \lambda_1 < \ldots < \lambda_{b-1} < \lambda_b, \\ & 2S - \lambda_K < \ldots < 2S - \lambda_{b+1} < \lambda_b, \\ & \lambda_b \leqslant S\}\,.\end{aligned}$$

$\boldsymbol{\lambda}'$ change of variables

$$\lambda'_a = \begin{cases} \lambda_a & \text{if } 1 \leqslant a \leqslant b \\ 2S - \lambda_a & \text{else}\,, \end{cases}$$

How to compute $\boldsymbol{\lambda}^b$?

$$\mathcal{I}_b = \{\boldsymbol{\lambda} \in \mathcal{I} : a^*_{\boldsymbol{\lambda}} = b\} = \{\boldsymbol{\lambda} \in \mathcal{M} : \lambda_1 < \ldots < \lambda_{b-1} < \min(\lambda_b, 2S - \lambda_b) \leqslant \max(\lambda_b, 2S - \lambda_b) < \lambda_{b+1} < \ldots < \lambda_K\}\,.$$

Suppose $\mu_b \leqslant S$ then $\lambda^b_b < S \to$ replace $\mathcal{I}_b$ with

$$\begin{aligned}\{\lambda' \in \mathcal{M} :\, & \lambda'_1 < \ldots < \lambda'_{b-1} < \lambda'_b, \\ & \lambda'_K < \ldots < \lambda'_{b+1} < \lambda'_b, \\ & \lambda'_b \leqslant S\}\,.\end{aligned}$$

$\boldsymbol{\lambda}'$ change of variables

$$\lambda'_a = \begin{cases} \lambda_a & \text{if } 1 \leqslant a \leqslant b \\ 2S - \lambda_a & \text{else}\,, \end{cases}$$

Unimodal regression

$$\boldsymbol{\lambda}^{b\prime} = \mathop{\arg\min}_{\substack{\lambda'_1 \leqslant \ldots \leqslant \lambda'_b \\ \lambda'_K \leqslant \ldots \leqslant \lambda'_b \\ \lambda'_b \leqslant S}} \sum_{a=1}^{K} \omega_a \frac{(\mu'_a - \lambda'_a)^2}{2} \,.$$

Unimodal regression

$$\boldsymbol{\lambda}^{b\prime} = \mathop{\arg\min}_{\substack{\lambda'_1 \leqslant \ldots \leqslant \lambda'_b \\ \lambda'_K \leqslant \ldots \leqslant \lambda'_b \\ \lambda'_b \leqslant S}} \sum_{a=1}^{K} \omega_a \frac{(\mu'_a - \lambda'_a)^2}{2} .$$

Then $\boldsymbol{\lambda}^{b\prime} = \min(\widehat{\boldsymbol{\lambda}}^b, S)$ where $\widehat{\boldsymbol{\lambda}}^b$ is the unimodal regression of $\boldsymbol{\mu}$ with mode at b:

$$\widehat{\boldsymbol{\lambda}}^b := \mathop{\arg\min}_{\substack{\lambda'_1 \leqslant \ldots \leqslant \lambda'_b \\ \lambda'_K \leqslant \ldots \leqslant \lambda'_b}} \sum_{a=1}^{K} \omega_a \frac{(\mu'_a - \lambda'_a)^2}{2} ,$$

Unimodal regression can be efficiently computed via isotonic regressions (Pool Adjacent Violators Algorithm) in $O(K)$.

$$\rightarrow \text{sub-gradient in } O(K^2).$$