

Safe Learning in Tree-Form Sequential Decision Making

Handling Hard and Soft Constraints

Martino Bernasconi, Federico Cacciamani, Matteo Castiglioni, Alberto Marchesi, Nicola Gatti, Francesco Trovò
martino.bernasconideluca@polimi.it



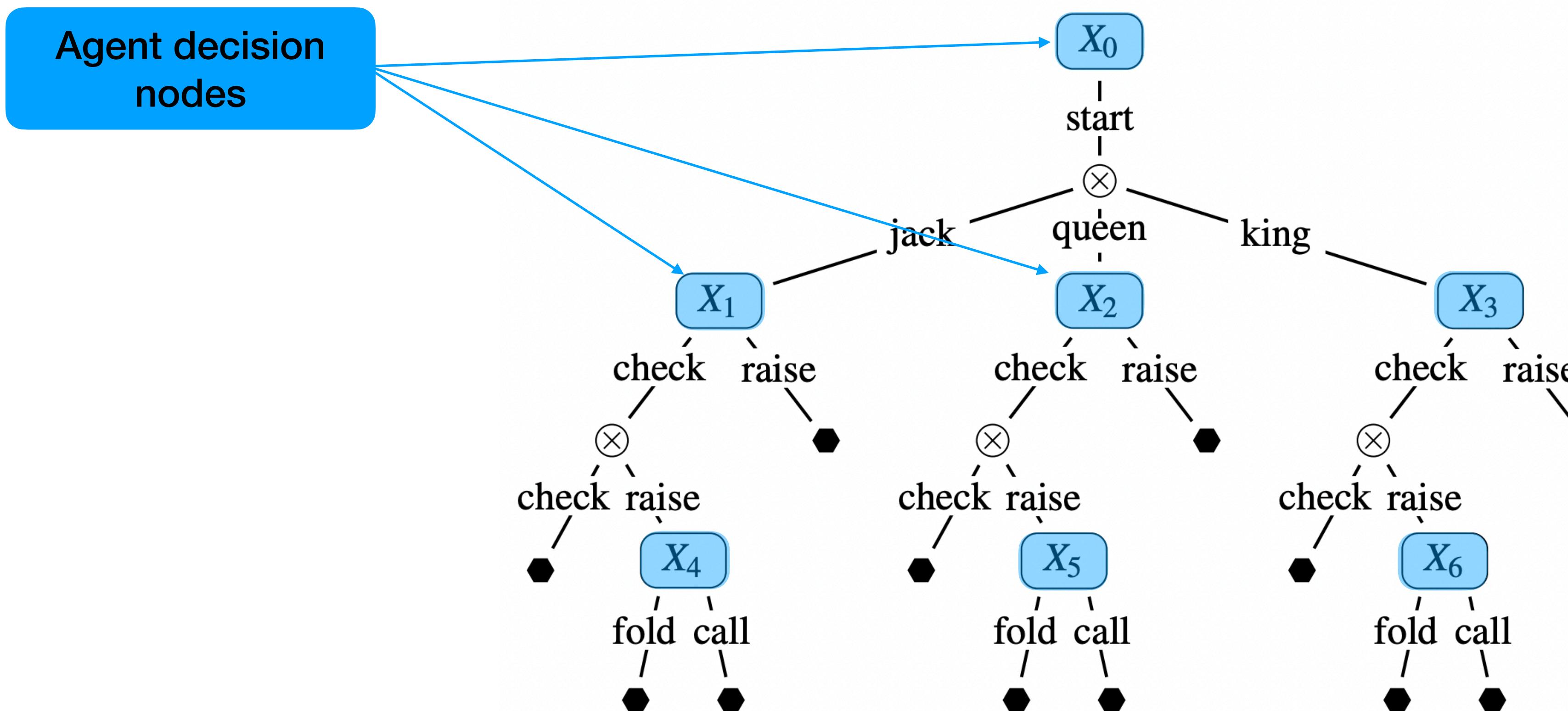
**POLITECNICO
MILANO 1863**



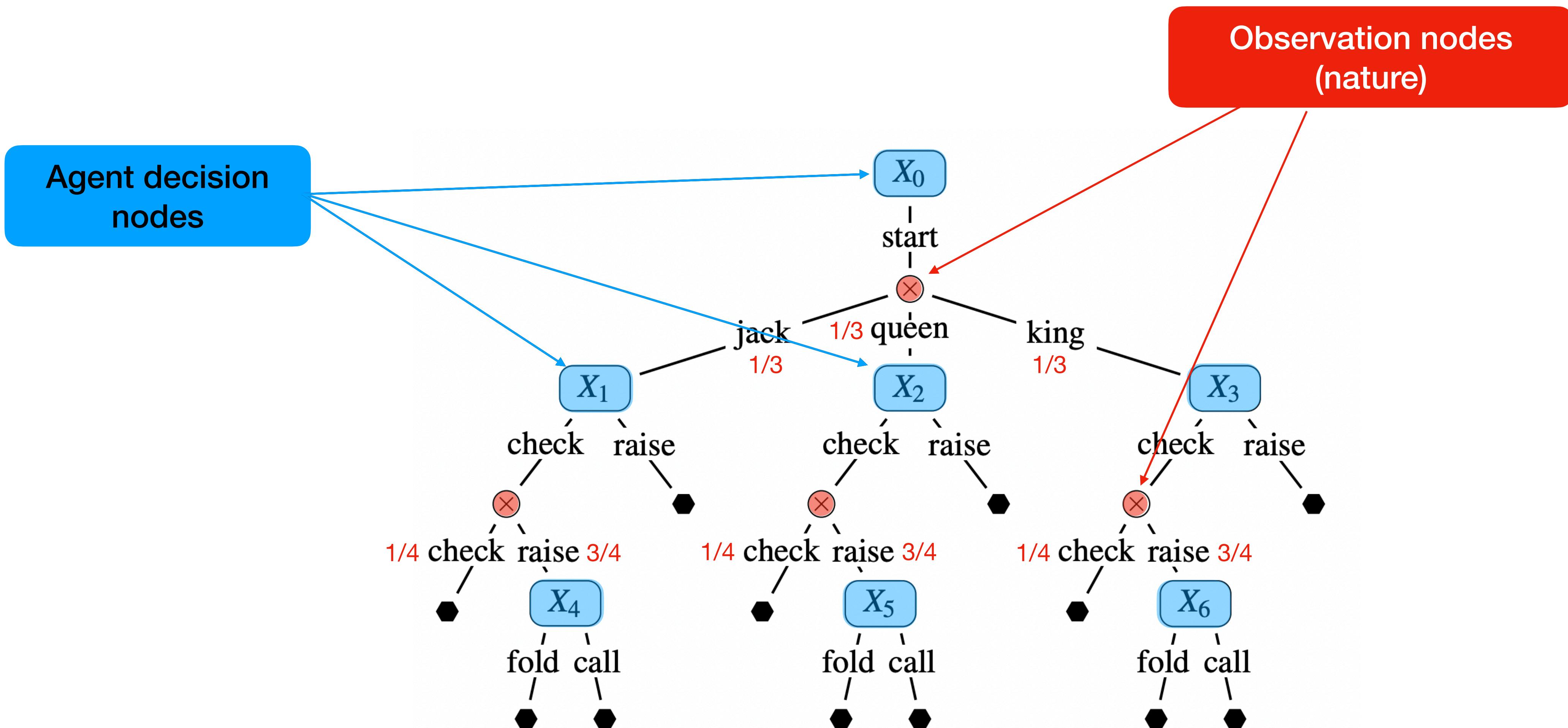
ICML
International Conference
On Machine Learning

Model

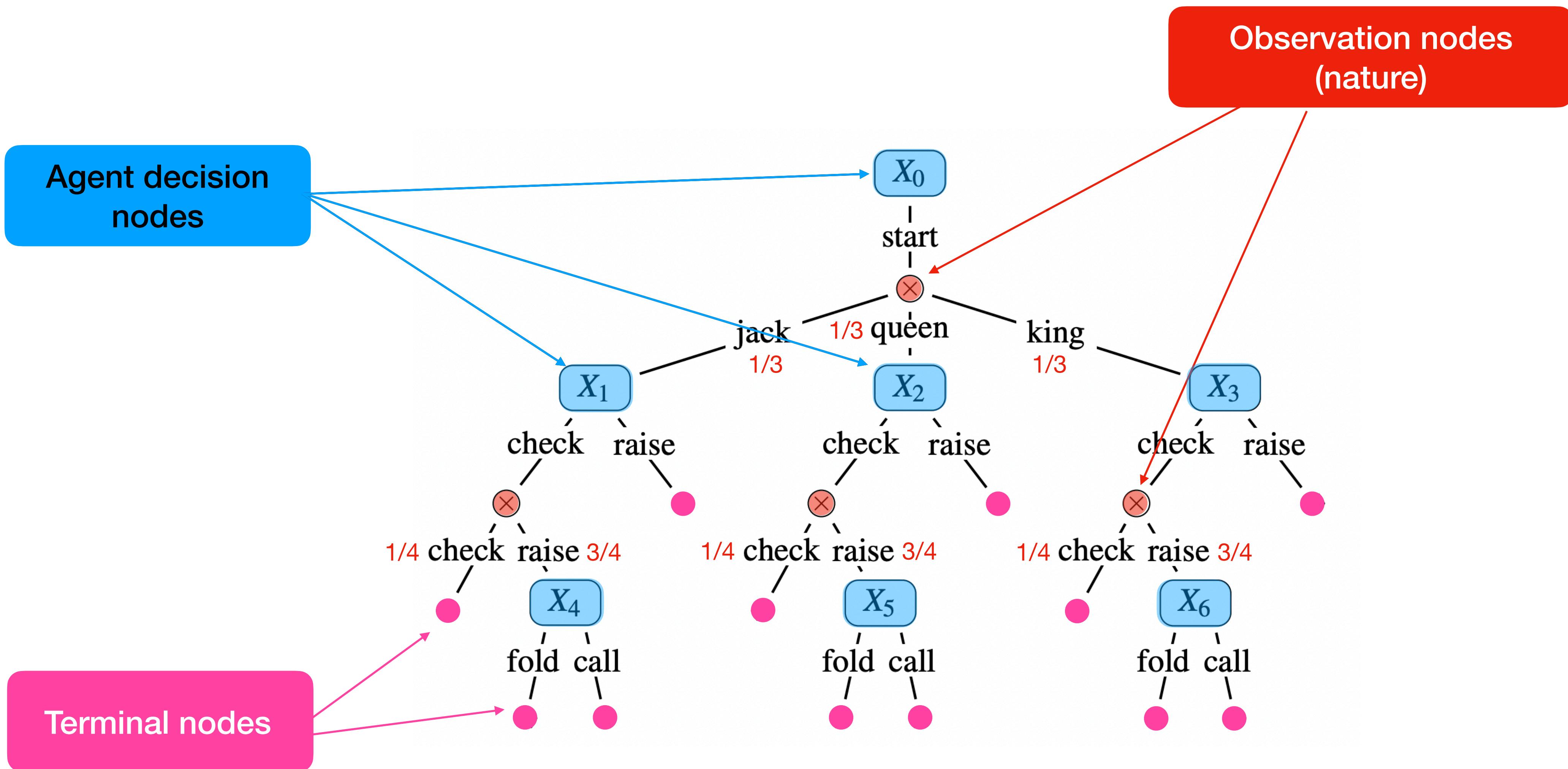
Tree-form sequential decision making



Tree-form sequential decision making



Tree-form sequential decision making



Objective

Utility maximization, subject to cost constraints satisfaction

x : strategy

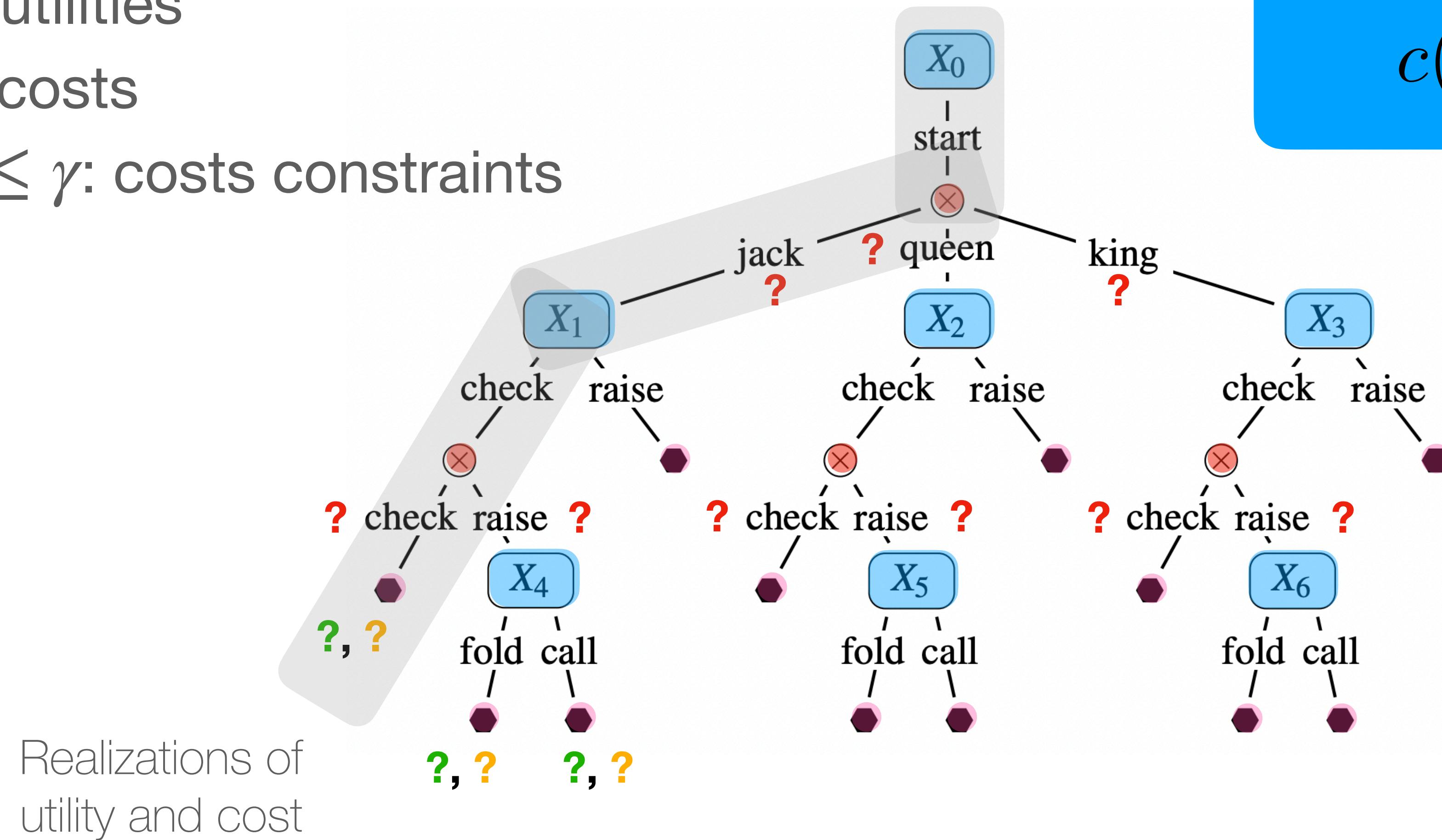
$u(x)$: utilities

$c(x)$: costs

$c(x) \leq \gamma$: costs constraints

Assumption:

$$c(x^\diamond) = \gamma - \lambda$$



Online learning

- Agent chooses $\boldsymbol{x}_t \in \mathcal{X}$
- Utility collected $u(\boldsymbol{x}_t)$
- Violation incurred $v_t := \max(0, c(\boldsymbol{x}) - \gamma)$

Cumulative regret

$$R_T := \max_{\boldsymbol{x}: u(\boldsymbol{x}) \leq \gamma} \sum_{t=1}^T u(\boldsymbol{x}) - \sum_{t=1}^T u(\boldsymbol{x}_t)$$

Cumulative violation

$$V_T := \sum_{t=1}^T v_t(\boldsymbol{x}_t)$$

***Hard* and *soft* constraints**

Cumulative violations

$$V_T := \sum_{t=1}^T v_t(\mathbf{x}_t)$$

Hard threshold (w.h.p.)

$$V_T \leq 0$$

Soft threshold (w.h.p.)

$$V_T = o(T)$$

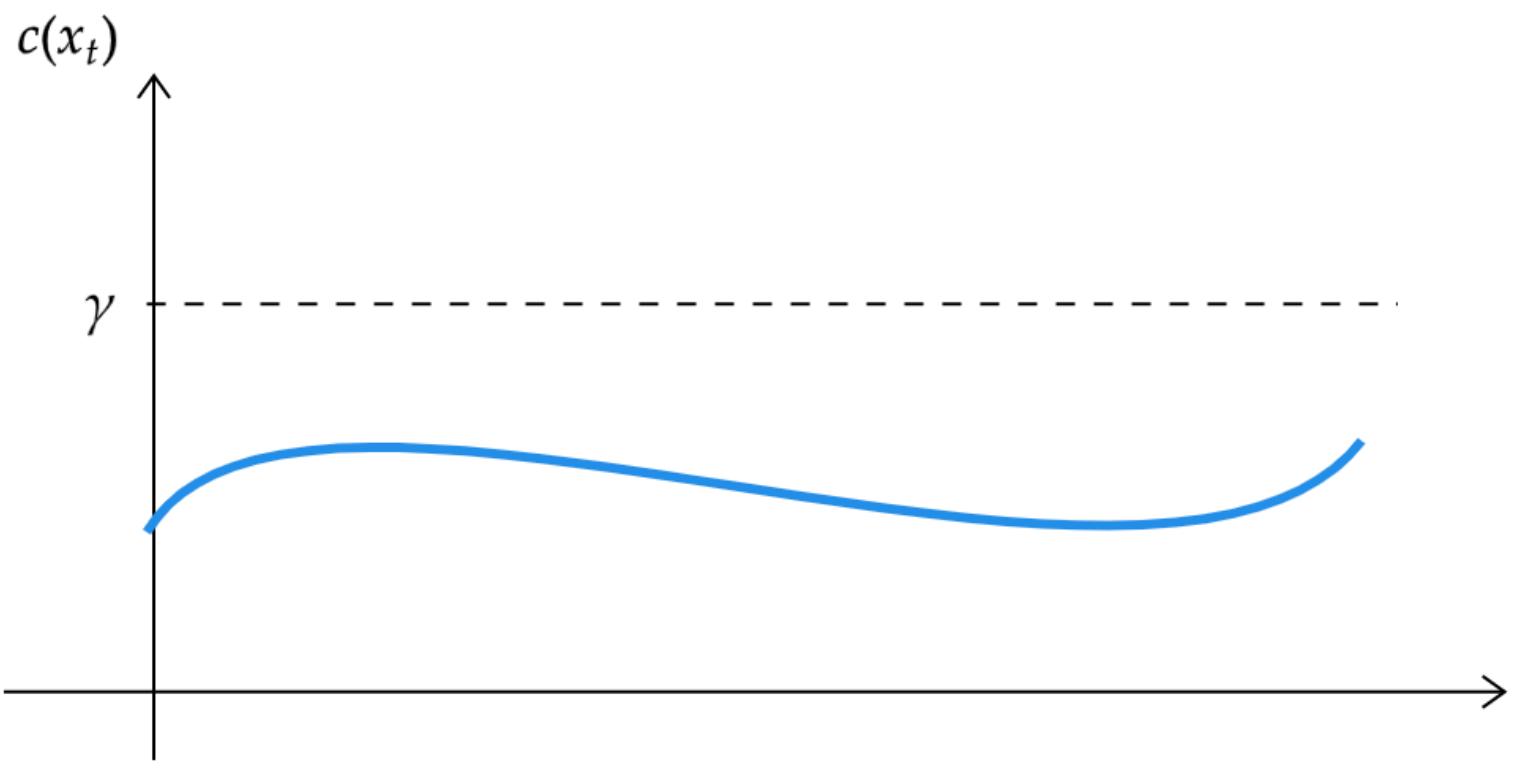
Hard and soft constraints

Cumulative violations

$$V_T := \sum_{t=1}^T v_t(\mathbf{x}_t)$$

Hard threshold (w.h.p.)

$$V_T \leq 0$$



Soft threshold (w.h.p.)

$$V_T = o(T)$$

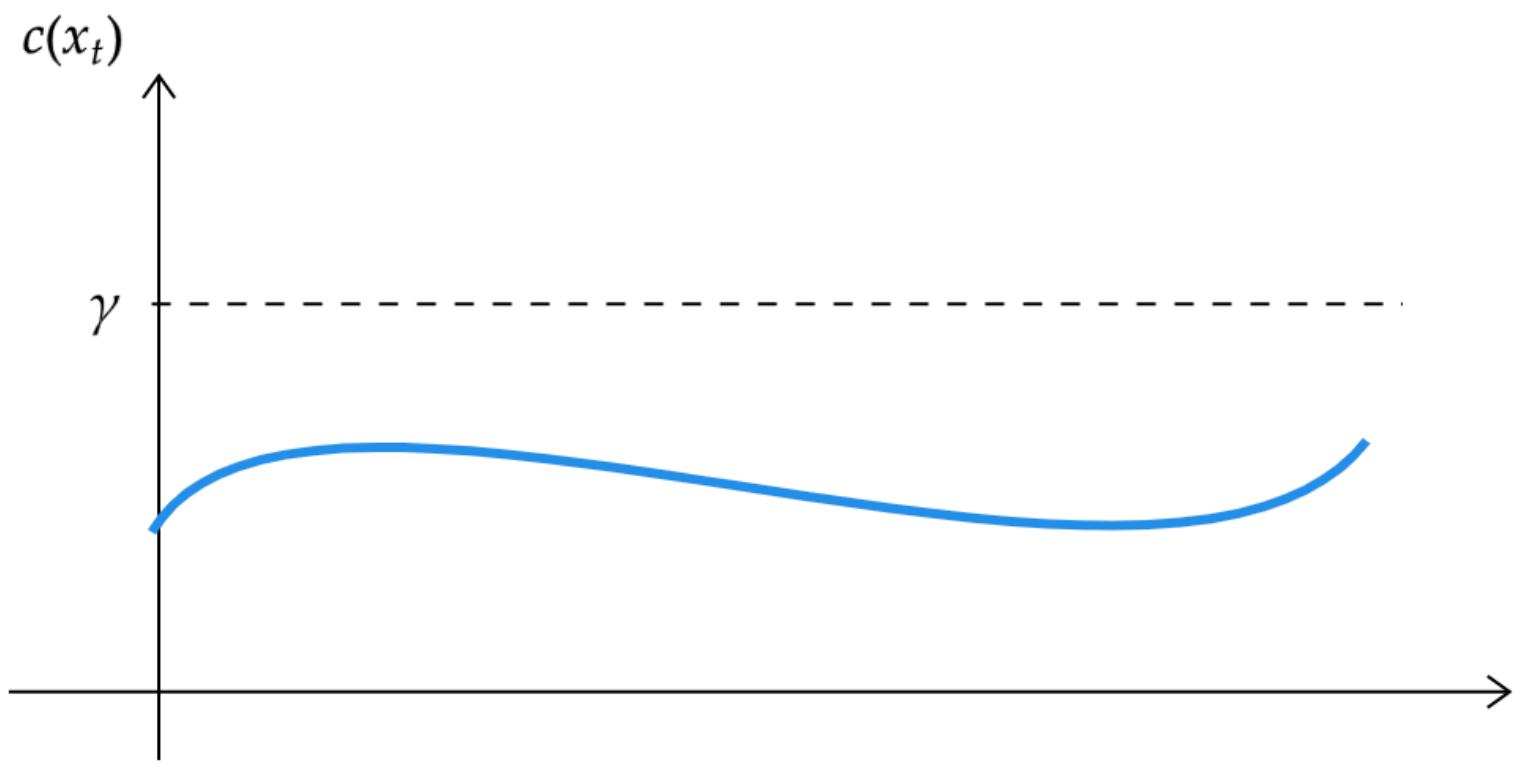
Hard and soft constraints

Cumulative violations

$$V_T := \sum_{t=1}^T v_t(\mathbf{x}_t)$$

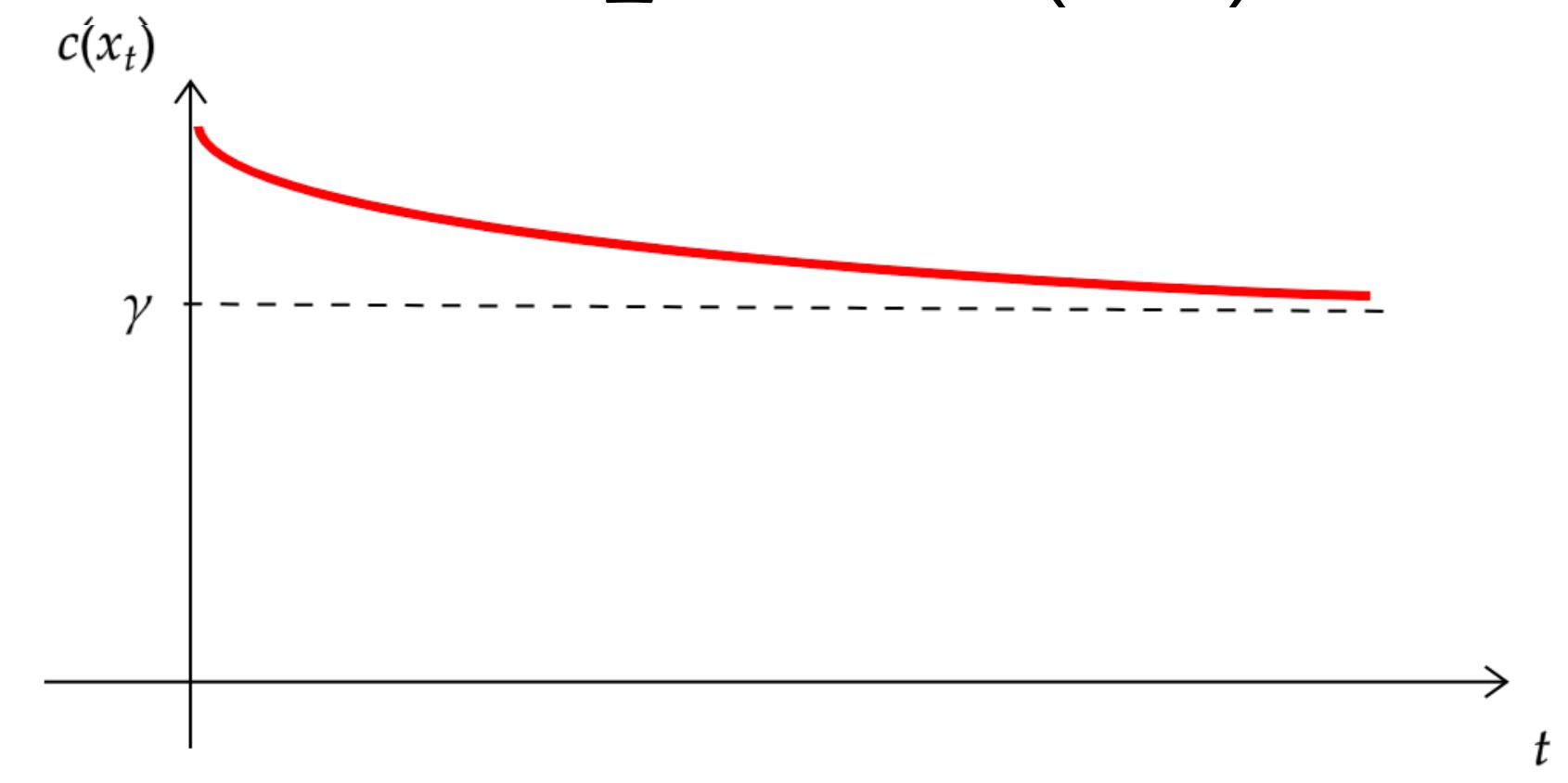
Hard threshold (w.h.p.)

$$V_T \leq 0$$



Soft threshold (w.h.p.)

$$V_T = o(T)$$



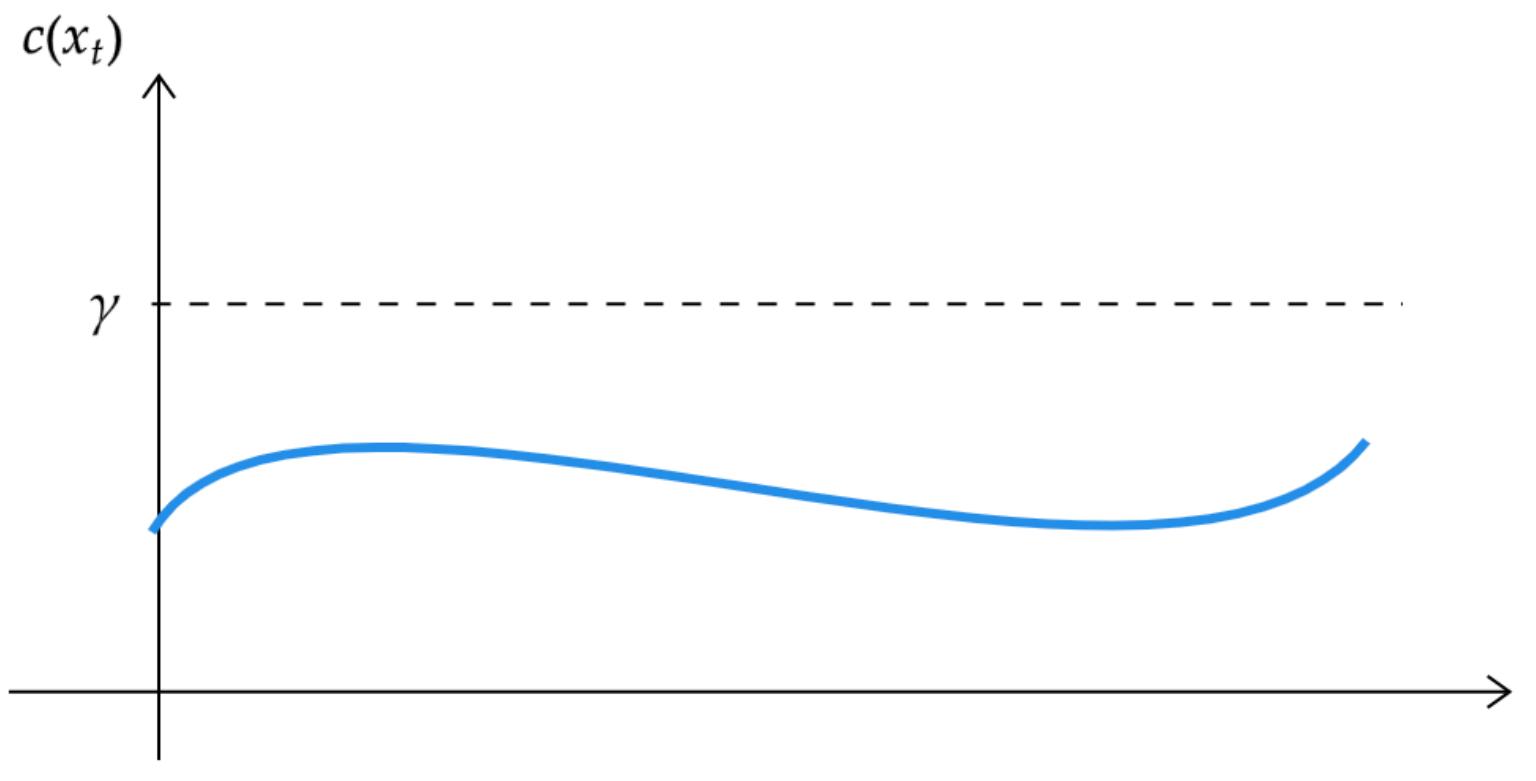
Hard and soft constraints

Cumulative violations

$$V_T := \sum_{t=1}^T v_t(\mathbf{x}_t)$$

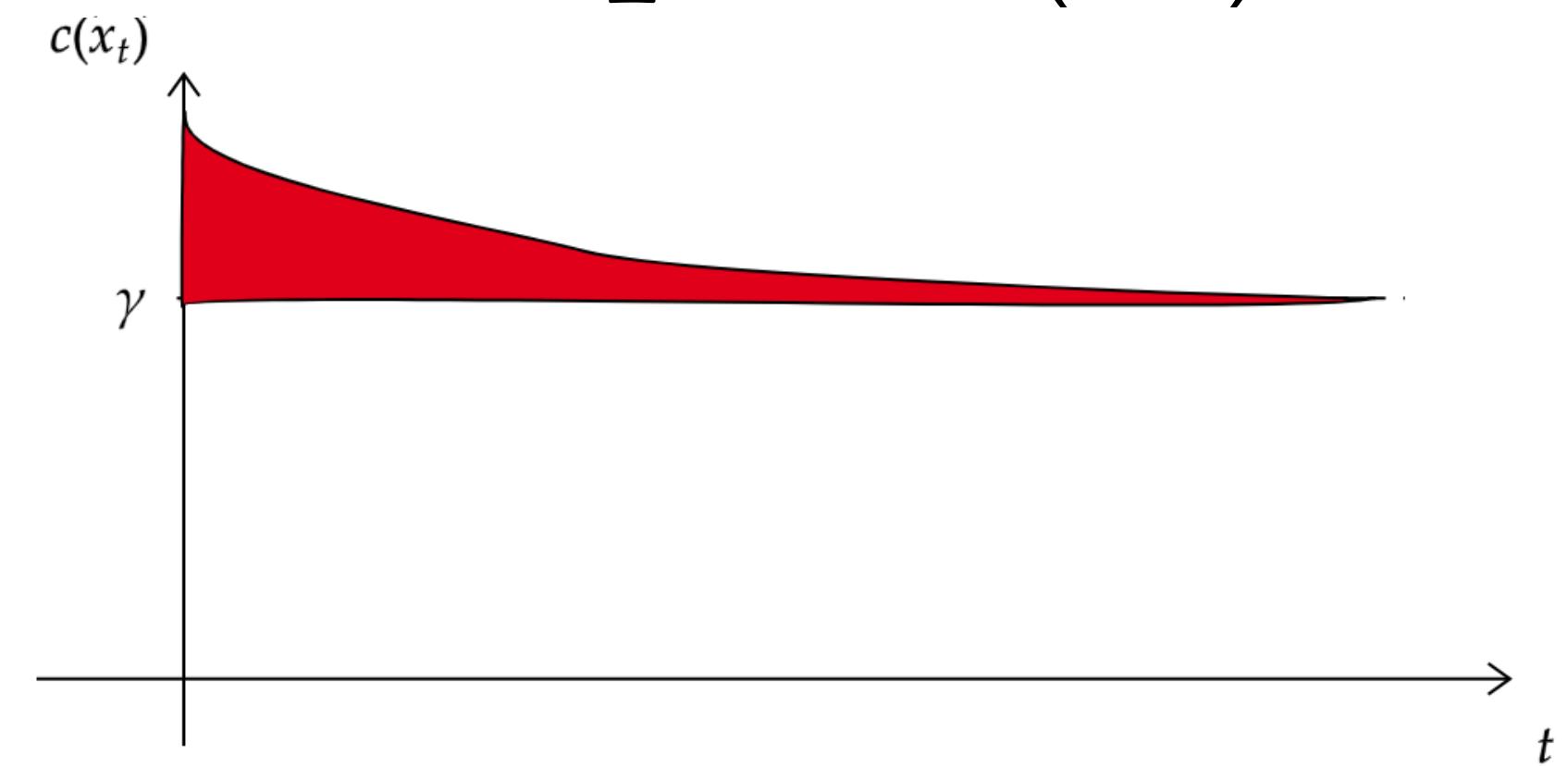
Hard threshold (w.h.p.)

$$V_T \leq 0$$



Soft threshold (w.h.p.)

$$V_T = o(T)$$



Results

Summary of the technical results

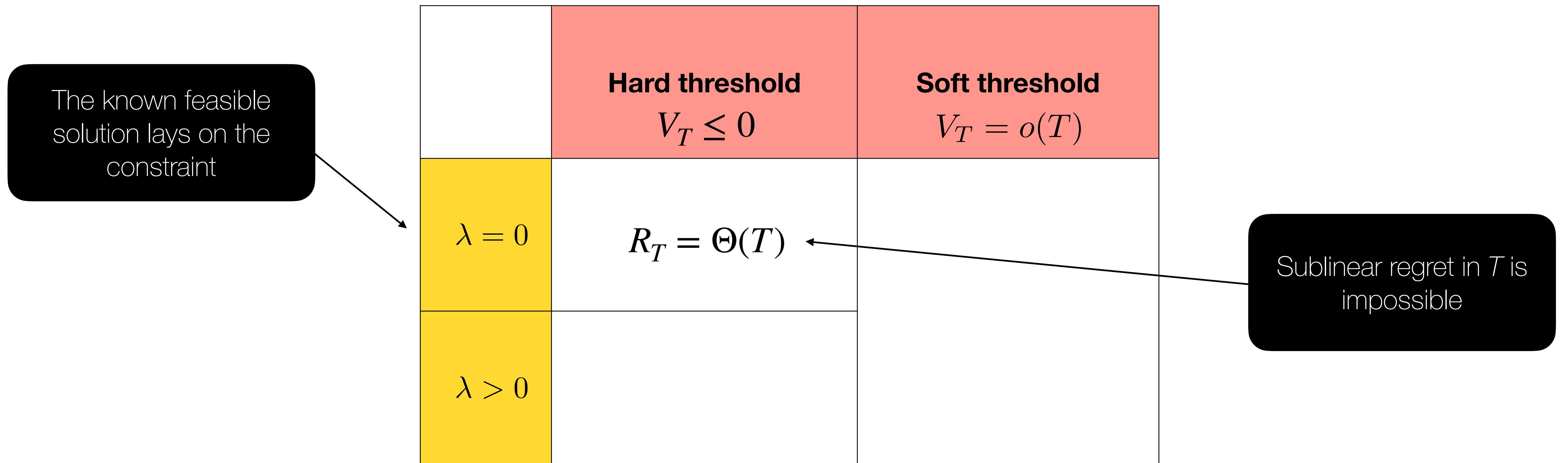
	Hard threshold $V_T \leq 0$	Soft threshold $V_T = o(T)$
$\lambda = 0$		
$\lambda > 0$		

Summary of the technical results

	Hard threshold $V_T \leq 0$	Soft threshold $V_T = o(T)$
$\lambda = 0$	$R_T = \Theta(T)$	
$\lambda > 0$		

The known feasible solution lays on the constraint

Sublinear regret in T is impossible



Summary of the technical results

	Hard threshold $V_T \leq 0$	Soft threshold $V_T = o(T)$
$\lambda = 0$	$R_T = \Theta(T)$	
$\lambda > 0$	$R_T = \Theta\left(\sqrt{\frac{T}{\lambda}}\right)$	

The known feasible solution lays in the inner of the feasible space

Sublinear regret in T is possible

The larger the distance from the constraint the lower the regret

Summary of the technical results

	Hard threshold $V_T \leq 0$	Soft threshold $V_T = o(T)$
$\lambda = 0$	$R_T = \Theta(T)$	$R_T, V_T = \Theta\left(\sqrt{T}\right)$
$\lambda > 0$	$R_T = \Theta\left(\sqrt{\frac{T}{\lambda}}\right)$	

Sublinear regret in T is possible

The regret does not depend on the known feasible solution slack

Thanks for the attention!

Martino Bernasconi, Federico Cacciamani, Matteo Castiglioni, Alberto Marchesi, Nicola Gatti, Francesco Trovò

martino.bernasconideluca@polimi.it



**POLITECNICO
MILANO 1863**



ICML
International Conference
On Machine Learning