# From Bad Models to Good Policies: an Intertwined Story about Energy and Reinforcement Learning

**2014 NIPS Workshop « From Bad Models to Good Policies Workshop (Sequential Decision Making under Uncertainty) »**
**Montreal, December 12th, 2014**

Raphael Fonteneau, University of Liège, Belgium
@R_Fonteneau

Joint work with Damien Ernst, Susan A. Murphy, Louis Wehenkel, Quentin Louveaux, Bernard Boigelot - thanks to many other people and to F.R.S.-FNRS

# Outline

- Stories of energy
- Learning from trajectories: batch mode RL
- Bad model bonus
- Trajectories of societies
- Transitioning

# Intertwined Stories

Fiesco via Wikipedia

Mosaique du Grand Plalais, Constantinople via Wikipedia

# The Roman Empire in 117 AD


Andrei Nacu via Wikipedia

# Deterministic RL

- Dynamics $x_{t+1} = f(x_t, u_t) \quad t = 0, \ldots, T-1 \quad T \in \mathbb{N} \setminus \{0\}$

$$\mathcal{X} \subset \mathbb{R}^d \qquad \mathcal{U} = \left\{u^{(1)}, \ldots, u^{(m)}\right\}$$

- Reward function $r_t = \rho(x_t, u_t) \in \mathbb{R}$

- Return

$$\forall (u_0, \ldots, u_{T-1}) \in \mathcal{U}^T, \qquad J(u_0, \ldots, u_{T-1}) \triangleq \sum_{t=0}^{T-1} \rho(x_t, u_t)$$

- Optimality $J_T^* \triangleq \max_{(u_0, \ldots, u_{T-1}) \in \mathcal{U}^T} J(u_0, \ldots, u_{T-1})$

# Batch Mode RL

- Dynamics and reward function are **unknown**
- Instead, we have access to trajectories (« bad model »):

$$\mathcal{F}^{(u)} = \left\{ \left( x^{(u),k}, r^{(u),k}, y^{(u),k} \right) \right\}_{k=1}^{n^{(u)}}$$

$$y^{(u),k} = f\left(x^{(u),k}, u\right) \qquad r^{(u),k} = \rho\left(x^{(u),k}, u\right)$$

$$\forall u \in \mathcal{U}, n^{(u)} > 0 \qquad \mathcal{F} = \mathcal{F}^{(1)} \cup \ldots \cup \mathcal{F}^{(m)}$$

# Lipschitz Continuity

$$\forall (x, x') \in \mathcal{X}^2, \forall u \in \mathcal{U}, \qquad \begin{aligned} \|f(x,u) - f(x',u)\| &\leq L_f \|x - x'\| \\ |\rho(x,u) - \rho(x',u)| &\leq L_\rho \|x - x'\| \end{aligned}$$

$$L_f, L_\rho \in \mathbb{R}$$

# Lipschitz Compatibility

$$\mathcal{L}_{\mathcal{F}}^{f} = \left\{ f' : \mathcal{X} \times \mathcal{U} \rightarrow \mathcal{X} \middle| \begin{cases} \forall x', x'' \in \mathcal{X}, \forall u \in \mathcal{U} \ , \\ \|f'(x', u) - f'(x'', u)\| \leq L_f \|x' - x''\| \ , \\ \forall k \in \{1, \ldots, n^{(u)}\}, f'(x^{(u),k}, u) = f(x^{(u),k}, u) = y^{(u),k} \end{cases} \right\}$$

$$\mathcal{L}_{\mathcal{F}}^{\rho} = \left\{ \rho' : \mathcal{X} \times \mathcal{U} \rightarrow \mathbb{R} \middle| \begin{cases} \forall x', x'' \in \mathcal{X}, \forall u \in \mathcal{U} \ , \\ |\rho'(x', u) - \rho'(x'', u)| \leq L_\rho \|x' - x''\| \ , \\ \forall k \in \{1, \ldots, n^{(u)}\}, \rho'(x^{(u),k}, u) = \rho(x^{(u),k}, u) = r^{(u),k} \end{cases} \right\}$$

$$\forall (f', \rho') \in \mathcal{L}_{\mathcal{F}}^{f} \times \mathcal{L}_{\mathcal{F}}^{\rho}, J_{(f', \rho')}(u_0, \ldots, u_{T-1}) = \sum_{t=0}^{T-1} \rho'(x'_t, u_t)$$

$$x'_{t+1} = f'(x'_t, u_t)$$

# Minmax Generalization

- Define:

$$B^* \left(\mathcal{F}, u0, \ldots, u_{T-1}\right) = \min_{(f',\rho') \in \mathcal{L}_{\mathcal{F}}^{f} \times \mathcal{L}_{\mathcal{F}}^{\rho}} J_{(f',\rho')(u_0,\ldots,u_{T-1})}$$

- The minmax generalization solution is defined as:

$$\left(u_0, \ldots, u_{T-1}\right) \in \underset{(u_0,\ldots,u_{T-1}) \in \mathcal{U}^T}{\arg\max} B^* \left(\mathcal{F}, u0, \ldots, u_{T-1}\right)$$

- Here, we focus on the min part

# Minmax Generalization

$(\mathcal{P}(\mathcal{F}, L_f, L_\rho, x_0, u_0, \ldots, u_{T-1})):$

$$\min_{\substack{\hat{\mathbf{r}}_0 \quad \ldots \quad \hat{\mathbf{r}}_{T-1} \in \mathbb{R} \\ \hat{\mathbf{x}}_0 \quad \ldots \quad \hat{\mathbf{x}}_{T-1} \in \mathcal{X}}} \sum_{t=0}^{T-1} \hat{\mathbf{r}}_t,$$

subject to

$$\left|\hat{\mathbf{r}}_t - r^{(u_t),k_t}\right|^2 \leq L_\rho^2 \left\|\hat{\mathbf{x}}_t - x^{(u_t),k_t}\right\|^2, \forall (t, k_t) \in \{0, \ldots, T-1\} \times \left\{1, \ldots, n^{(u_t)}\right\}, \tag{3.1}$$

$$\left\|\hat{\mathbf{x}}_{t+1} - y^{(u_t),k_t}\right\|^2 \leq L_f^2 \left\|\hat{\mathbf{x}}_t - x^{(u_t),k_t}\right\|^2, \forall (t, k_t) \in \{0, \ldots, T-1\} \times \left\{1, \ldots, n^{(u_t)}\right\}, \tag{3.2}$$

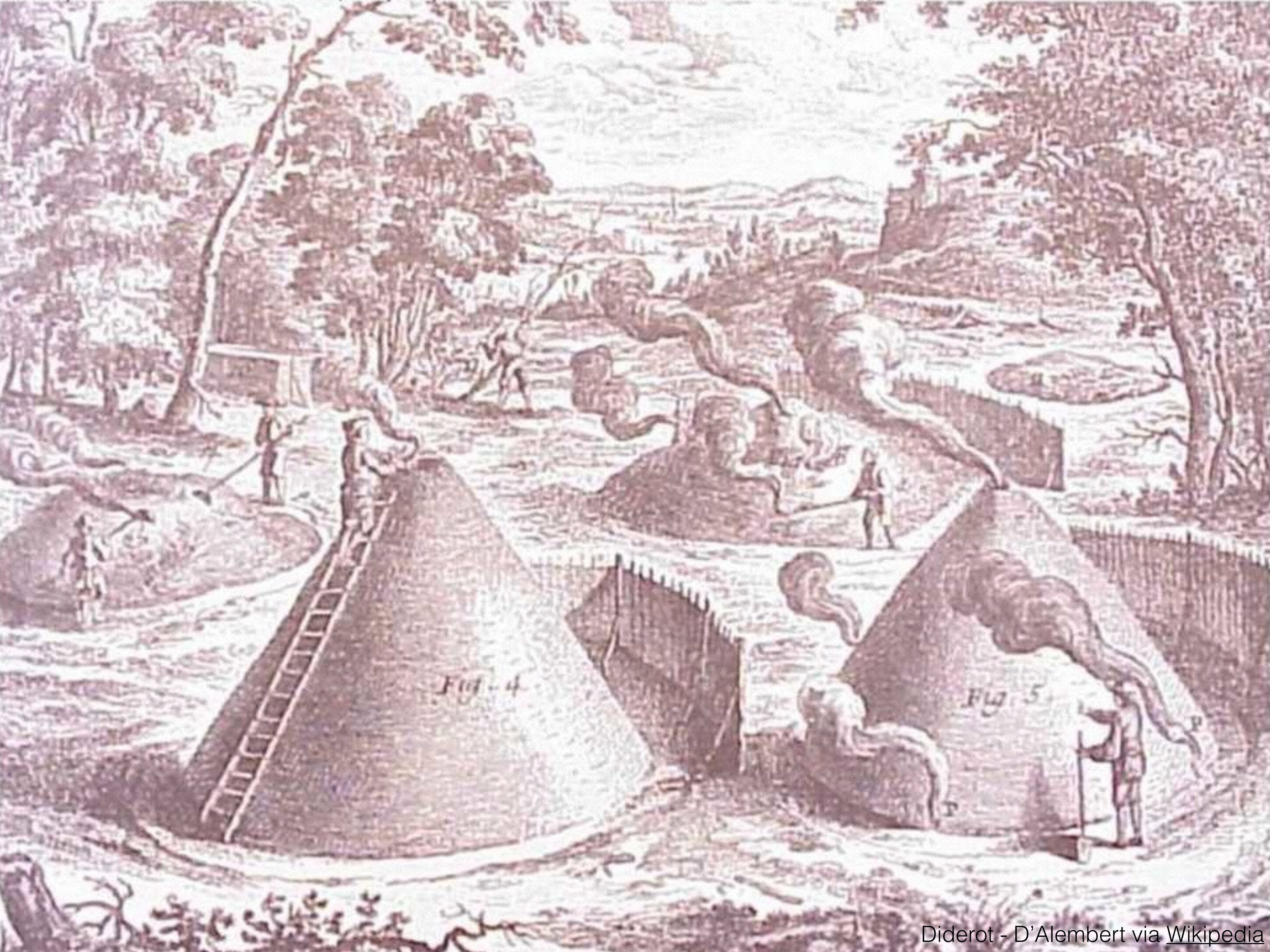
$$\left|\hat{\mathbf{r}}_t - \hat{\mathbf{r}}_{t'}\right|^2 \leq L_\rho^2 \left\|\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_{t'}\right\|^2, \forall t, t' \in \{0, \ldots, T-1 | u_t = u_{t'}\}, \tag{3.3}$$

$$\left\|\hat{\mathbf{x}}_{t+1} - \hat{\mathbf{x}}_{t'+1}\right\|^2 \leq L_f^2 \left\|\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_{t'}\right\|^2, \forall t, t' \in \{0, \ldots, T-2 | u_t = u_{t'}\}, \tag{3.4}$$

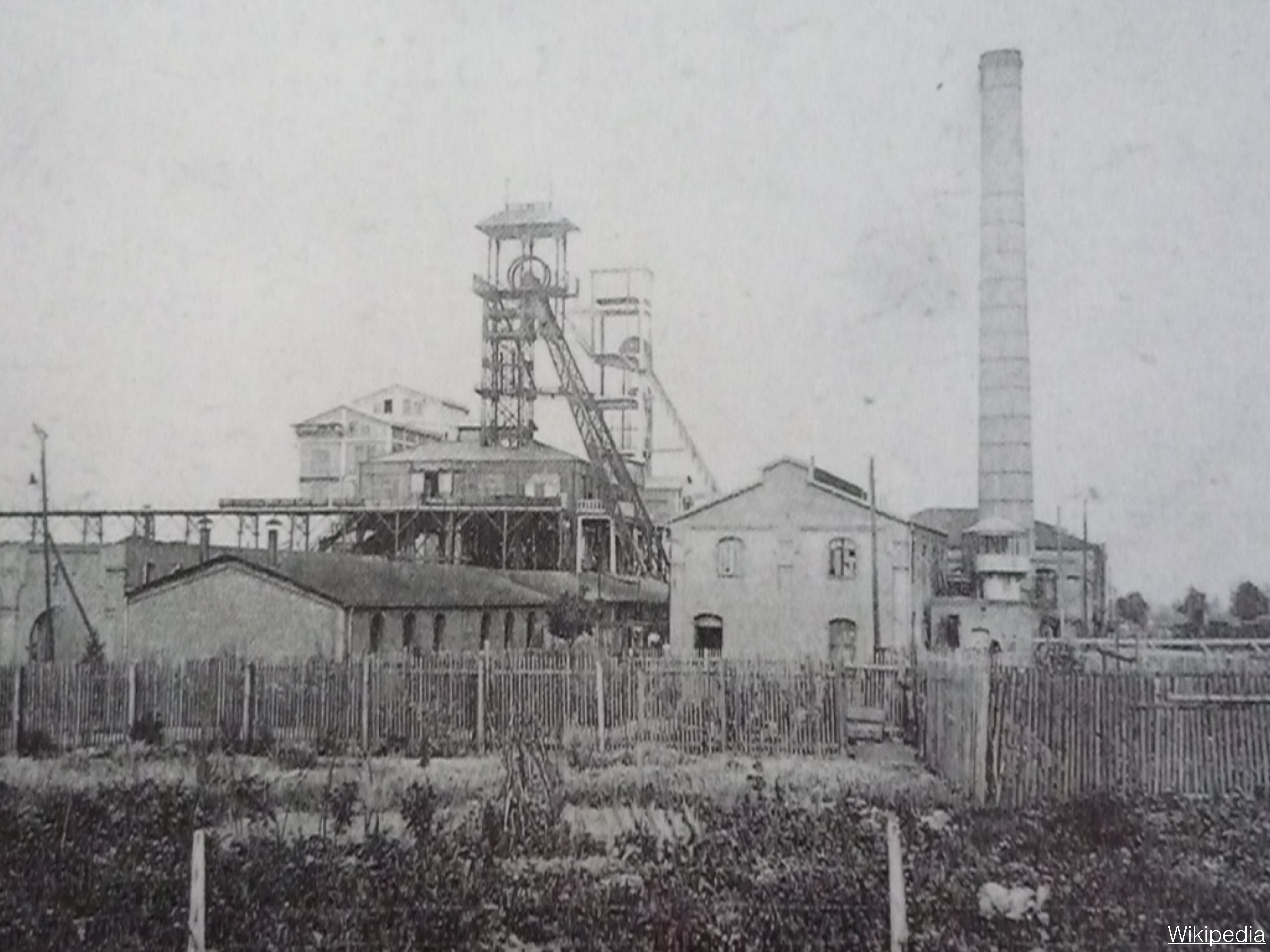
$$\hat{\mathbf{x}}_0 = x_0. \tag{3.5}$$

Jacob van Ruisdael via Wikipedia


Hendrick Cornelis Vroom via Wikipedia



Diderot - D'Alembert via Wikipedia



Wikipedia

Hartmut Reiche via Wikipedia

Eric Kounce via Wikipedia

# Minmax Generalization

$(\mathcal{P}(\mathcal{F}, L_f, L_\rho, x_0, u_0, \ldots, u_{T-1}))$ :

$$\min_{\substack{\hat{\mathbf{r}}_0 \quad \ldots \quad \hat{\mathbf{r}}_{T-1} \in \mathbb{R} \\ \hat{\mathbf{x}}_0 \quad \ldots \quad \hat{\mathbf{x}}_{T-1} \in \mathcal{X}}} \quad \sum_{t=0}^{T-1} \hat{\mathbf{r}}_t,$$

subject to

$$\left|\hat{\mathbf{r}}_t - r^{(u_t),k_t}\right|^2 \leq L_\rho^2 \left\|\hat{\mathbf{x}}_t - x^{(u_t),k_t}\right\|^2, \forall (t, k_t) \in \{0, \ldots, T-1\} \times \left\{1, \ldots, n^{(u_t)}\right\}, \tag{3.1}$$

$$\left\|\hat{\mathbf{x}}_{t+1} - y^{(u_t),k_t}\right\|^2 \leq L_f^2 \left\|\hat{\mathbf{x}}_t - x^{(u_t),k_t}\right\|^2, \forall (t, k_t) \in \{0, \ldots, T-1\} \times \left\{1, \ldots, n^{(u_t)}\right\}, \tag{3.2}$$

$$\left|\hat{\mathbf{r}}_t - \hat{\mathbf{r}}_{t'}\right|^2 \leq L_\rho^2 \left\|\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_{t'}\right\|^2, \forall t, t' \in \{0, \ldots, T-1 | u_t = u_{t'}\}, \tag{3.3}$$

$$\left\|\hat{\mathbf{x}}_{t+1} - \hat{\mathbf{x}}_{t'+1}\right\|^2 \leq L_f^2 \left\|\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_{t'}\right\|^2, \forall t, t' \in \{0, \ldots, T-2 | u_t = u_{t'}\}, \tag{3.4}$$

$$\hat{\mathbf{x}}_0 = x_0. \tag{3.5}$$

# Minmax Generalization

- One can show that constraint (3.3) are redundant

LEMMA 4.1. *Consider* $(\hat{\mathbf{r}}^*, \hat{\mathbf{x}}^*) \in \mathbb{R}^T \times \mathcal{X}^T$ *an optimal solution to* $\bar{\mathcal{P}}\left(\mathcal{F}, u_0, \ldots, u_{T-1}\right)$. *Then, for all* $t, t'$ *such that* $u_t = u_{t'}$,

$$\left|\hat{\mathbf{r}}_t^* - \hat{\mathbf{r}}_{t'}^*\right|^2 \leq L_\rho^2 \left\|\hat{\mathbf{x}}_t^* - \hat{\mathbf{x}}_{t'}^*\right\|^2 .$$

- In particular, this implies that optimal reward for the first stage (t=0) can also be computed

LEMMA 4.2. *The solution of the problem* $(\mathcal{P}'(\mathcal{F}, u_0))$ *is*

$$\hat{\mathbf{r}}_0^* = \max_{k_0 \in \left\{1, \ldots, n^{(u_0)}\right\}} r^{(u_0), k_0} - L_\rho \left\| x_0 - x^{(u_0), k_0} \right\| .$$

# Minmax Generalization

$(\mathcal{P}''(\mathcal{F}, u_0, \ldots, u_{T-1}))$ :

$$\min_{\substack{\hat{\mathbf{r}}_1 \quad \ldots \quad \hat{\mathbf{r}}_{T-1} \in \mathbb{R} \\ \hat{\mathbf{x}}_0 \quad \ldots \quad \hat{\mathbf{x}}_{T-1} \in \mathcal{X}}} \quad \sum_{t=1}^{T-1} \hat{\mathbf{r}}_t,$$

subject to

$$\left|\hat{\mathbf{r}}_t - r^{(u_t),k_t}\right|^2 \leq L_\rho^2 \left\|\hat{\mathbf{x}}_t - x^{(u_t),k_t}\right\|^2 , \forall (t, k_t) \in \{1, \ldots, T-1\} \times \left\{1, \ldots, n^{(u_t)}\right\}, \tag{5.1}$$

$$\left\|\hat{\mathbf{x}}_{t+1} - y^{(u_t),k_t}\right\|^2 \leq L_f^2 \left\|\hat{\mathbf{x}}_t - x^{(u_t),k_t}\right\|^2, \forall (t, k_t) \in \{0, \ldots, T-1\} \times \left\{1, \ldots, n^{(u_t)}\right\}, \tag{5.2}$$

$$\|\hat{\mathbf{x}}_{t+1} - \hat{\mathbf{x}}_{t'+1}\|^2 \leq L_f^2 \|\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_{t'}\|^2 , \forall t, t' \in \{0, \ldots, T-2 | u_t = u_{t'}\}, \tag{5.3}$$

$$\hat{\mathbf{x}}_0 = x_0 . \tag{5.4}$$

# Minmax Generalization

- We show that this problem is **NP-hard**
    - Reduction from {0,1}-programming feasibility problem
- We then decide to look for relaxation schemes of **polynomial complexity**
- We want these relaxation schemes to preserve the philosophy of the original problem
    - **Lower bounds**

# Relaxation Schemes

- First approach: remove constraints until the problem becomes polynomial

$(\mathcal{P}''(\mathcal{F}, u_0, \ldots, u_{T-1}))$ :

$$\min_{\substack{\hat{\mathbf{r}}_1 \quad \ldots \quad \hat{\mathbf{r}}_{T-1} \in \mathbb{R} \\ \hat{\mathbf{x}}_0 \quad \ldots \quad \hat{\mathbf{x}}_{T-1} \in \mathcal{X}}} \sum_{t=1}^{T-1} \hat{\mathbf{r}}_t,$$

Only one

subject to

$$\left| \hat{\mathbf{r}}_t - r^{(u_t),k_t} \right|^2 \leq L_\rho^2 \left\| \hat{\mathbf{x}}_t - x^{(u_t),k_t} \right\|^2, \forall (t, k_t) \in \{1, \ldots, T-1\} \times \left\{1, \ldots, n^{(u_t)}\right\}, \quad (5.1)$$

$$\left\| \hat{\mathbf{x}}_{t+1} - y^{(u_t),k_t} \right\|^2 \leq L_f^2 \left\| \hat{\mathbf{x}}_t - x^{(u_t),k_t} \right\|^2, \forall (t, k_t) \in \{0, \ldots, T-1\} \times \left\{1, \ldots, n^{(u_t)}\right\}, \quad (5.2)$$

$$\|\hat{\mathbf{x}}_{t+1} - \hat{\mathbf{x}}_{t'+1}\|^2 \leq L_f^2 \|\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_{t'}\|^2, \forall t, t' \in \{0, \ldots, T-2 | u_t = u_{t'}\}, \quad (5.3)$$

$$\hat{\mathbf{x}}_0 = x_0 . \quad (5.4)$$

# Relaxation Schemes

- We get the « Intertwined Trust-Region » scheme:

$$\left(\mathcal{P}''_{ITR}(\mathcal{F}, u_0, \ldots, u_{T-1}, \bar{k}_0, \ldots, \bar{k}_{T-1})\right):$$

$$\min_{\substack{\hat{\mathbf{r}}_1, \ldots, \hat{\mathbf{r}}_{T-1} \in \mathbb{R} \\ \hat{\mathbf{x}}_0, \ldots, \hat{\mathbf{x}}_{T-1} \in \mathcal{X}}} \sum_{t=1}^{T-1} \hat{\mathbf{r}}_t$$

subject to

$$\left|\hat{\mathbf{r}}_t - r^{(u_t), \bar{k}_t}\right|^2 \leq L_\rho^2 \left\|\hat{\mathbf{x}}_t - x^{(u_t), \bar{k}_t}\right\|^2 \qquad t \in \{1, \ldots, T-1\} \tag{5.5}$$

$$\left\|\hat{\mathbf{x}}_t - y^{(u_{t-1}), \bar{k}_{t-1}}\right\|^2 \leq L_f^2 \left\|\hat{\mathbf{x}}_{t-1} - x^{(u_{t-1}), \bar{k}_{t-1}}\right\|^2 \qquad t \in \{1, \ldots, T-1\} \tag{5.6}$$

$$\hat{\mathbf{x}}_0 = x_0 \tag{5.7}$$

# Relaxation Schemes

- This problem can be solved by induction. Define:

$$\left(\mathcal{Q}''_{ITR}(\mathcal{F}, u_0, \ldots, u_j, \bar{k}_0, \ldots, \bar{k}_j)\right):$$

$$\max_{\substack{\hat{\mathbf{r}}_1, \ldots, \hat{\mathbf{r}}_j \in \mathbb{R} \\ \hat{\mathbf{x}}_0, \ldots, \hat{\mathbf{x}}_j \in \mathcal{X}}} \left\| \hat{\mathbf{x}}_j - x^{(u_j), \bar{k}_j} \right\|$$

*subject to*

$$\left| \hat{\mathbf{r}}_t - r^{(u_t), \bar{k}_t} \right|^2 \leq L_\rho^2 \left\| \hat{\mathbf{x}}_t - x^{(u_i), \bar{k}_t} \right\|^2 \qquad t \in \{1, \ldots, j\} \tag{5.8}$$

$$\left\| \hat{\mathbf{x}}_t - y^{(u_{t-1}), \bar{k}_{t-1}} \right\|^2 \leq L_f^2 \left\| \hat{\mathbf{x}}_{t-1} - x^{(u_{t-1}), \bar{k}_{t-1}} \right\|^2 \qquad t \in \{1, \ldots, j\} \tag{5.9}$$

$$\hat{\mathbf{x}}_0 = x_0 \tag{5.10}$$

# Relaxation Schemes

LEMMA 5.2. *The optimal solution* $D''_{ITR}(u_0, u_1, \bar{k}_0, \bar{k}_1)$ *to* $\left(\mathcal{Q}''_{ITR}(\mathcal{F}, u_0, u_1, \bar{k}_0, \bar{k}_1)\right)$ *is given by*

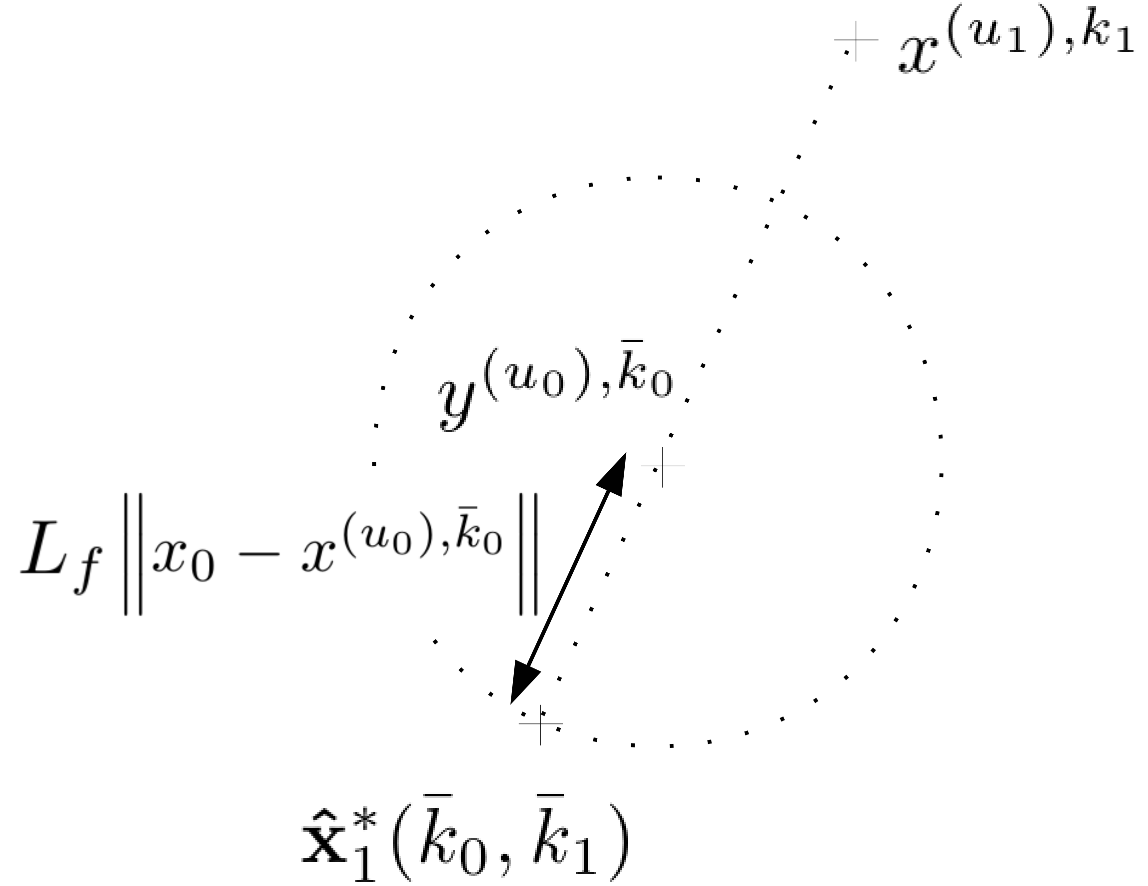
$$D''_{ITR}(u_0, u_1, \bar{k}_0, \bar{k}_1) = \left\| \hat{\mathbf{x}}_1^*(\bar{k}_0, \bar{k}_1) - x^{(u_1),\bar{k}_1} \right\|,$$

*where*

$$\hat{\mathbf{x}}_1^*(\bar{k}_0, \bar{k}_1) \doteq y^{(u_0),\bar{k}_0} + L_f \frac{\left\| x_0 - x^{(u_0),\bar{k}_0} \right\|}{\left\| y^{(u_0),\bar{k}_0} - x^{(u_1),\bar{k}_1} \right\|} \left( y^{(u_0),\bar{k}_0} - x^{(u_1),\bar{k}_1} \right) \textit{ if } y^{(u_0),\bar{k}_0} \neq x^{(u_1),\bar{k}_1}$$

*and, if* $y^{(u_0),\bar{k}_0} = x^{(u_1),\bar{k}_1}$, $\hat{\mathbf{x}}_1^*(\bar{k}_0, \bar{k}_1)$ *can be any point of the sphere centered in* $y^{(u_0),\bar{k}_0} = x^{(u_1),\bar{k}_1}$ *with radius* $L_f \|x_0 - x^{(u_0),\bar{k}_0}\|$.

# Relaxation Schemes




*A simple geometric algorithm to solve* $(\mathcal{Q}''_{ITR}(\mathcal{F}, u_0, u_1, \bar{k}_0, \bar{k}_1))$

# Relaxation Schemes

LEMMA 5.3. *The optimal solution to* $\left(\mathcal{Q}''_{ITR}(\mathcal{F}, u_0, \ldots, u_j, \bar{k}_0, \ldots, \bar{k}_j)\right)$ *is given by:*

$$\forall t \in \{1, \ldots, j\}, \quad \hat{\mathbf{x}}^*_t(\bar{k}_0, \ldots, \bar{k}_t) \doteq y^{(u_{t-1}), \bar{k}_{t-1}} + L_f \frac{\left\| \hat{\mathbf{x}}^*_{t-1}(\bar{k}_0, \ldots, \bar{k}_{t-1}) - x^{(u_{t-1}), \bar{k}_{t-1}} \right\|}{\left\| y^{(u_{t-1}), \bar{k}_{t-1}} - x^{(u_t), \bar{k}_t} \right\|} \left( y^{(u_{t-1}), \bar{k}_{t-1}} - x^{(u_t), \bar{k}_t} \right) \quad \textit{if } y^{(u_{t-1}), \bar{k}_{t-1}} \neq x^{(u_t), \bar{k}_t}$$

*and, if* $y^{(u_{t-1}), \bar{k}_{t-1}} = x^{(u_t), \bar{k}_t}$, $\hat{\mathbf{x}}^*_t(\bar{k}_0, \ldots, \bar{k}_t)$ *can be any point of the sphere centered in* $y^{(u_{t-1}), \bar{k}_{t-1}} = x^{(u_t), \bar{k}_t}$ *with radius* $L_f \|\hat{\mathbf{x}}^*_{t-1}(\bar{k}_0, \ldots, \bar{k}_{t-1}) - x^{(u_{t-1}), \bar{k}_{t-1}}\|$.

# Relaxation Schemes

THEOREM 5.4. *The solution to* $\left(\mathcal{P}''_{ITR}(\mathcal{F}, u_0, \ldots, u_{T-1}, \bar{k}_0, \ldots, \bar{k}_{T-1})\right)$ *is given by:*

$$B''_{ITR}(\mathcal{F}, u_0, \ldots, u_{T-1}, \bar{k}_0, \ldots, \bar{k}_{T-1}) = \sum_{t=1}^{T-1} \hat{\mathbf{r}}^*_t$$

*where*

$$\hat{\mathbf{r}}^*_t = r^{(u_t),\bar{k}_t} - L_\rho \left\| \hat{\mathbf{x}}^*_t(\bar{k}_0, \ldots, \bar{k}_t) - x^{(u_t),\bar{k}_t} \right\|,$$

$$\hat{\mathbf{x}}^*_t(\bar{k}_0, \ldots, \bar{k}_t) \doteq y^{(u_{t-1}),\bar{k}_{t-1}} + L_f \frac{\left\| \hat{\mathbf{x}}^*_{t-1}(\bar{k}_0, \ldots, \bar{k}_{t-1}) - x^{(u_{t-1}),\bar{k}_{t-1}} \right\|}{\left\| y^{(u_{t-1}),\bar{k}_{t-1}} - x^{(u_t),\bar{k}_t} \right\|} \left( y^{(u_{t-1}),\bar{k}_{t-1}} - x^{(u_t),\bar{k}_t} \right)$$

*if* $y^{(u_{t-1}),\bar{k}_{t-1}} \neq x^{(u_t),\bar{k}_t}$

*and, if* $y^{(u_{t-1}),\bar{k}_{t-1}} = x^{(u_t),\bar{k}_t}$, $\hat{\mathbf{x}}^*_t(\bar{k}_0, \ldots, \bar{k}_t)$ *can be any point of the sphere centered in* $y^{(u_{t-1}),\bar{k}_{t-1}} = x^{(u_t),\bar{k}_t}$ *with radius* $L_f \| \hat{\mathbf{x}}^*_{t-1}(\bar{k}_0, \ldots, \bar{k}_{t-1}) - x^{(u_{t-1}),\bar{k}_{t-1}} \|$.

# Relaxation Schemes

- One can look for the best bound among all possible configurations

DEFINITION 5.5 (Intertwined Trust-region Bound $B_{ITR}(\mathcal{F}, u_0, \ldots, u_{T-1})$).

$$
\begin{aligned}
&B_{ITR}(\mathcal{F}, u_0, \ldots, u_{T-1}) \triangleq \hat{\mathbf{r}}_0^* \\
&+ \max_{\substack{\bar{k}_{T-1} \in \{1, \ldots, n^{(u_{T-1})}\} \\ \cdots \\ \bar{k}_0 \in \{1, \ldots, n^{(u_0)}\}}} B''_{ITR}(\mathcal{F}, u_0, \ldots, u_{T-1}, \bar{k}_0, \ldots, \bar{k}_{T-1}).
\end{aligned}
$$

# Relaxation Schemes

$(\mathcal{P}''_{LD}(\mathcal{F}, u_0, \ldots, u_{T-1})):$

$$\max_{\substack{\nu_{t,t'} \in \mathbb{R} \\ \lambda_{t,k_t} \in \mathbb{R} \\ \mu_{t,k_t} \in \mathbb{R}}} \quad \min_{\substack{\hat{\mathbf{r}}_1, \ldots, \hat{\mathbf{r}}_{T-1} \in \mathbb{R} \\ \hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_{T-1} \in \mathcal{X}}}$$

$$\hat{\mathbf{r}}_1 + \cdots + \hat{\mathbf{r}}_{T-1} +$$

$$+ \sum_{(t,k_t) \in \{1,\ldots,T-1\} \times \left\{1,\ldots,n^{(u_t)}\right\}} \mu_{t,k_t} \left( \left| \hat{\mathbf{r}}_t - r^{(u_t),k_t} \right|^2 - L_\rho^2 \left\| \hat{\mathbf{x}}_t - x^{(u_t),k_t} \right\|^2 \right)$$

$$+ \sum_{(t,k_t) \in \{1,\ldots,T-1\} \times \left\{1,\ldots,n^{(u_t)}\right\}} \lambda_{t,k_t} \left( \left\| \hat{\mathbf{x}}_{t+1} - y^{(u_t),k_t} \right\|^2 - L_f^2 \left\| \hat{\mathbf{x}}_t - x^{(u_t),k_t} \right\|^2 \right)$$

$$+ \sum_{t,t' \in \{0,\ldots,T-2 \mid u_t = u_{t'}\}} \nu_{t,t'} \left( \left\| \hat{\mathbf{x}}_{t+1} - \hat{\mathbf{x}}_{t'+1} \right\|^2 - L_f^2 \left\| \hat{\mathbf{x}}_t - \hat{\mathbf{x}}_{t'} \right\|^2 \right)$$

# Relaxation Schemes

- The Lagrangian Relaxation provides a **lower bound** on the optimal bound, in a **polynomial time**

DEFINITION 5.7 (Lagrandian Bound $B_{LD}\left(\mathcal{F}, u_0, \ldots, u_{T-1}\right)$). *Let* $B''_{LD}(\mathcal{F}, u_0, \ldots, u_{T-1})$ *be the optimal Lagrangian dual of* $(\mathcal{P}''_{LD}(\mathcal{F}, u_0, \ldots, u_{T-1}))$. *Then,*

$$B_{LD}(\mathcal{F}, u_0, \ldots, u_{T-1}) = \mathbf{r}_0^* + B''_{LD}(\mathcal{F}, u_0, \ldots, u_{T-1}) \; .$$

# Bounds Tightness & Convergence

- Tightness

THEOREM 5.18. $\forall\,(u_0,\dots,u_{T-1}) \in \mathcal{U}^T$,

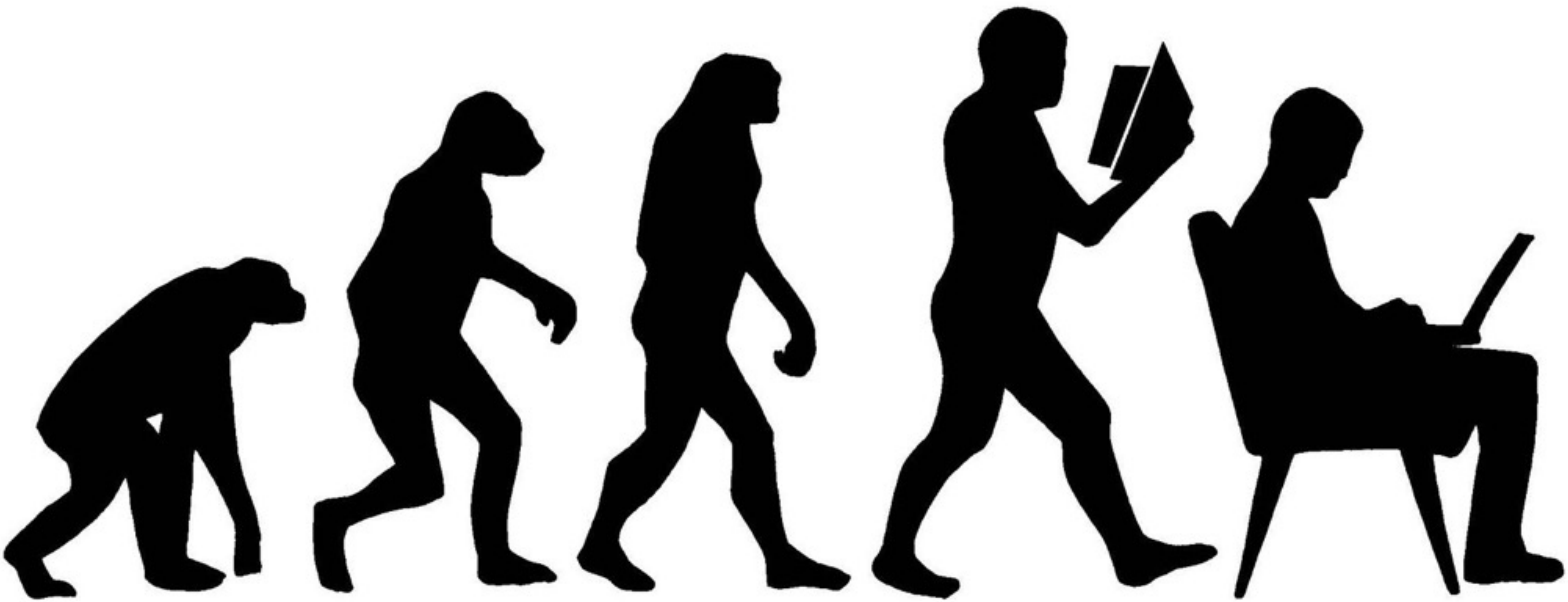
$$\begin{aligned} B_{CGRL}(\mathcal{F},u_0,\dots,u_{T-1}) &\leq B_{ITR}(\mathcal{F},u_0,\dots,u_{T-1}) \\ &\leq B_{LD}(\mathcal{F},u_0,\dots,u_{T-1}) \\ &\leq B^*(\mathcal{F},u_0,\dots,u_{T-1}) \\ &\leq J(u_0,\dots,u_{T-1}). \end{aligned}$$

- Convergence as the sample dispersion of the sample of trajectories goes to 0

THEOREM 5.21. $\forall(u_0,\dots,u_{T-1}) \in \mathcal{U}^T$,
$\forall\beta \in \{B_{CGRL}(\mathcal{F},u_0,\dots,u_{T-1}), B_{ITR}(\mathcal{F},u_0,\dots,u_{T-1}), B_{LD}(\mathcal{F},u_0,\dots,u_{T-1})\}$,

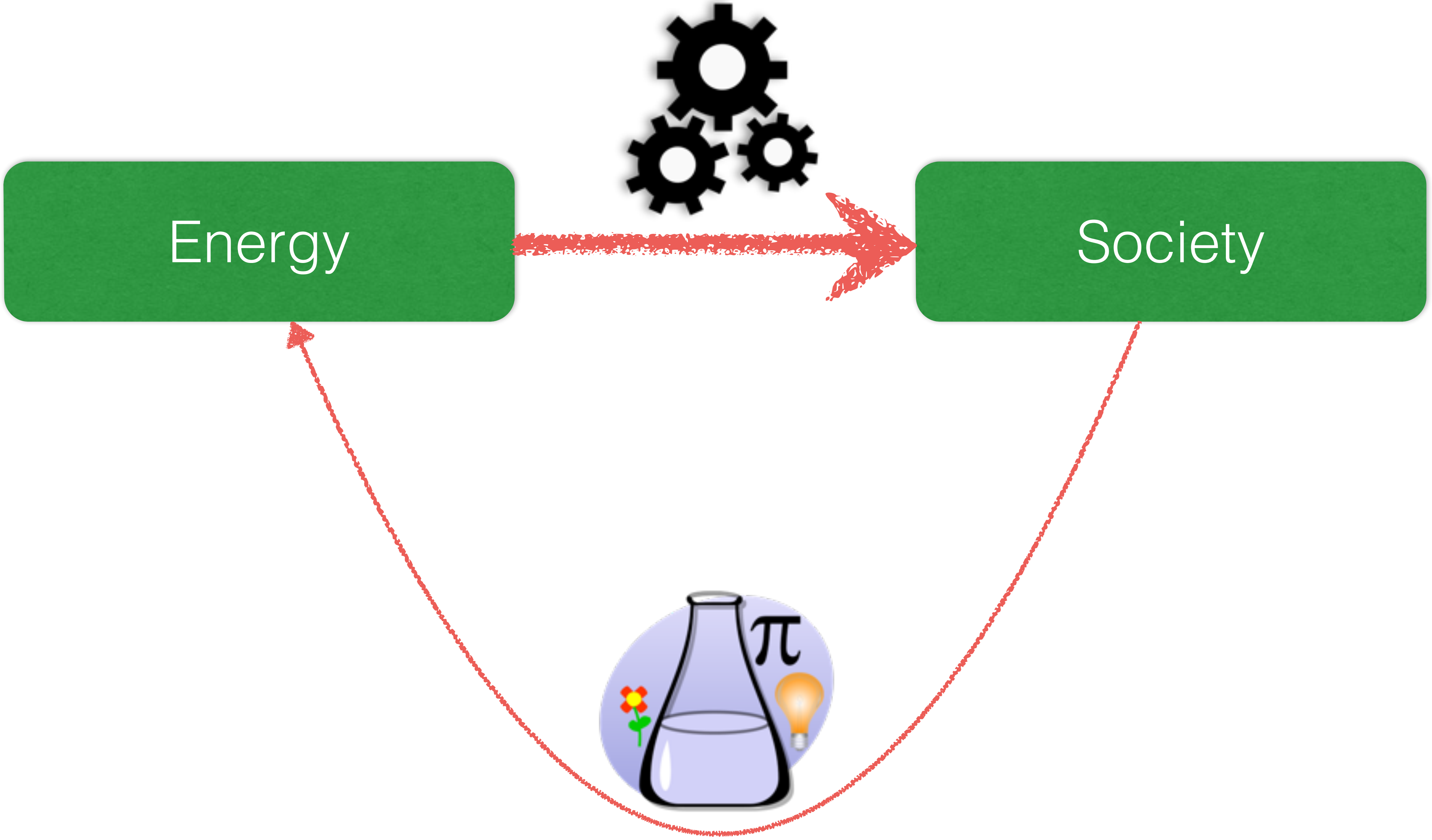
$$\lim_{\alpha^*(\mathcal{F})\to 0} \quad J(u_0,\dots,u_{T-1}) - \beta = 0\,.$$

# Trajectories of Societies



Johanna Pung via Wikipedia

# Trajectories of Societies



# From Bad Models to Good Policies

The Energy Transition Case

# 1. World primary energy consumption

Non renewable

> 80% - < 20%

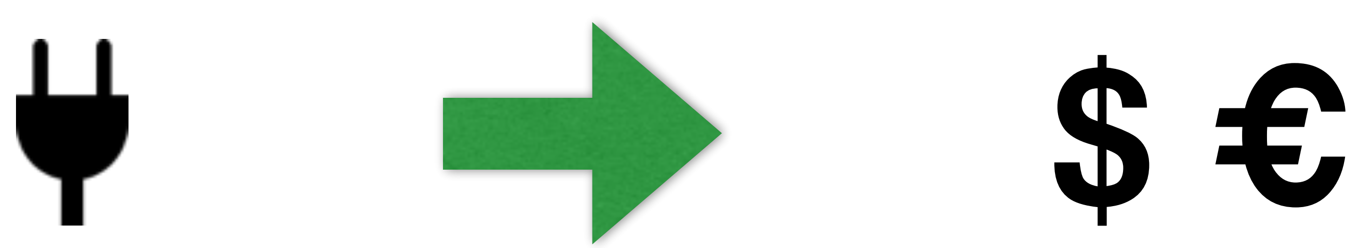
Renewable

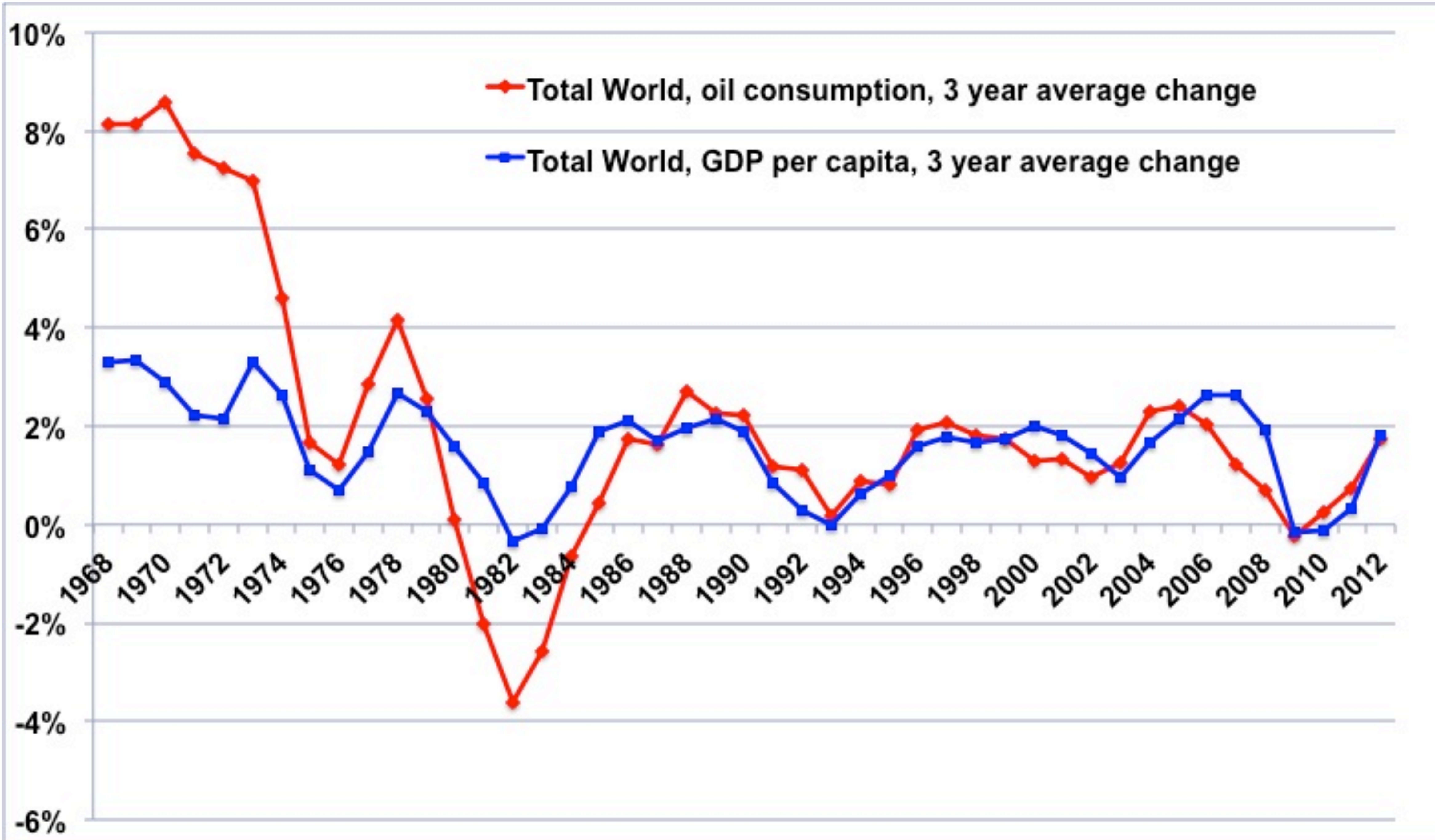
# 2. Energy <-> Economy

- Recent research in Economics has shown that:
    - The empirical elasticity (measured from time series among OECD countries over the last 50 years) of the consumption of primary energy into the GDP is about 60%, which is 10 times higher that what is predicted by the Cost Share Theorem

*Elasticity can be quantified as the ratio of the percentage change in one variable to the percentage change in another variable*

- There is a causality link between the consumption of primary energy and the GDP in the direction Energy -> GDP






Variation of the world oil consumption (red) and GDP per inhabitant (blue) - Data from the the World Bank for GDP and BP stat for energy

Source (in French): Jean-Marc Jancovici, « L'économie aurait-elle un vague rapport avec l'énergie? », LH Forum, 27 septembre 2013

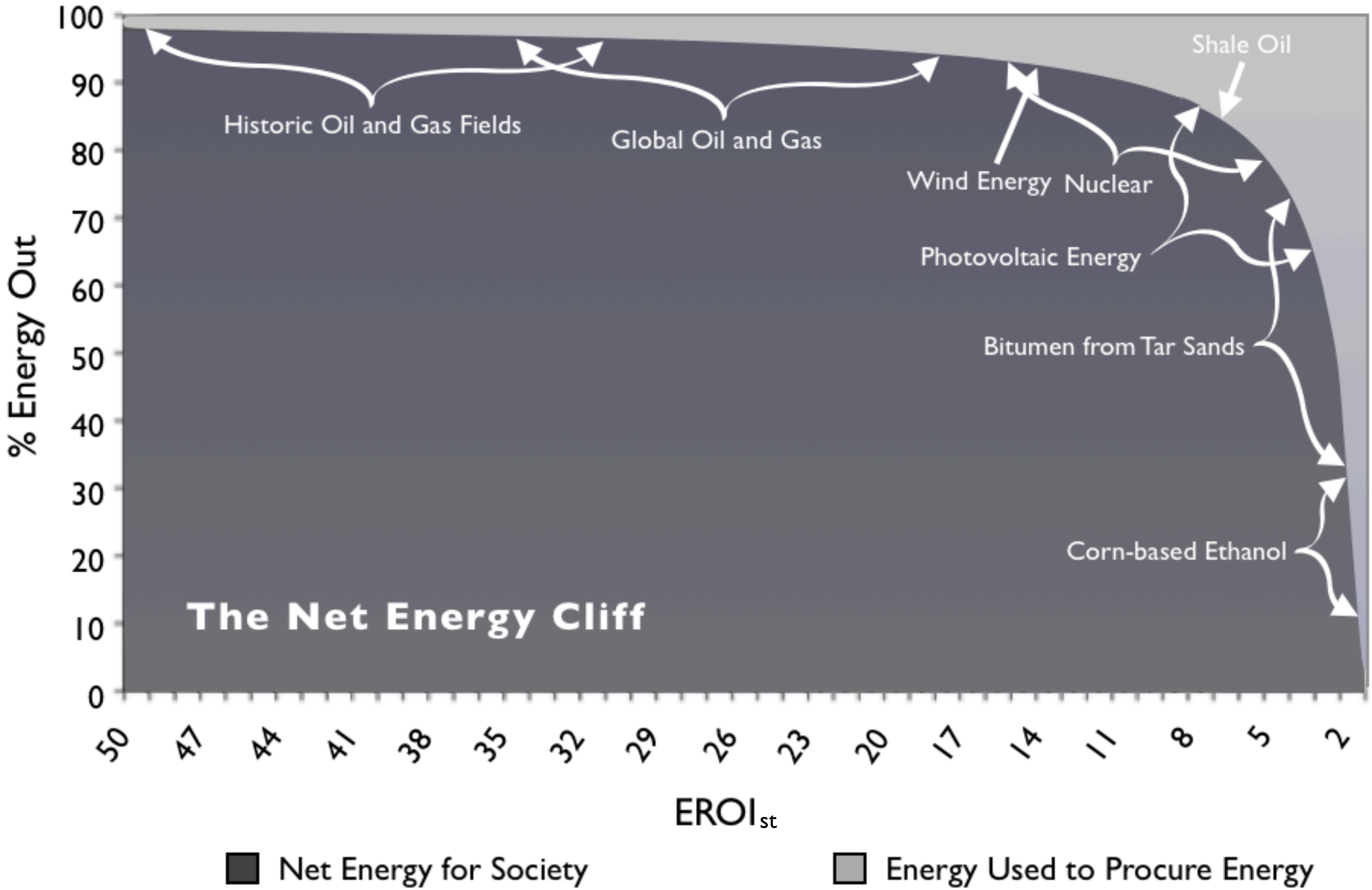
# 3. ERoEI

- **ERoEI for « Energy Return over Energy Investment »** (also called EROI) is the ratio of the amount of usable energy acquired from a particular energy resource to the amount of energy expended to obtain that energy resource:

$$EROI = \frac{Usable\ Acquired\ Energy}{Energy\ Expended}$$

- The highest this ratio, the more energy a technology brings back to society

- Notation : 1:X




*Source: EROI of Global Energy Resources - Preliminary Status and Trends - Jessica Lambert, Charles Hall, Steve Balogh, Alex Poisson, and Ajay Gupta State University of New York, College of Environmental Science and Forestry Report 1 - Revised Submitted - 2 November 2012 DFID - 59717*

# Another Bad Model

- A discrete-time model of the deployment of « renewable energy » production capacities

- Budget of non-renewable energy

$$\forall t \in \{0, \ldots, T-1\}, B_t \geq 0$$

$$\exists r > 0, \exists \tau > 0, \exists t_0 \in \mathbb{R} : \forall t \in \{0, \ldots, T-1\},$$

$$B_t = \frac{1}{r} \frac{e^{\frac{-(t-t_0)}{\tau}}}{\left(1 + e^{\frac{-(t-t_0)}{\tau}}\right)^2}$$

# Another Bad Model

- Set of renewable energy production technologies:

$$\forall n \in \{1, \ldots, N\}, \forall t \in \{0, \ldots, T-1\}, R_{n,t} \geq 0$$

- Characteristics $\Delta_{n,t} \geq 0$

$$EROEI_{n,t} \geq 0$$

- Deployment strategy

$$R_{n,t+1} = (1 + \alpha_{n,t}) R_{n,t} \qquad \alpha_{n,t} \in [-1, \infty[$$

# Another Bad Model

- Energy costs for growth and long-term replacement

$$\forall n \in \{1, \ldots, N\}, \forall t \in \{0, \ldots, T-1\},$$

$$C_{n,t}\left(R_{n,t}, \alpha_{n,t}\right) \geq 0 \qquad\qquad M_{n,t} \geq 0$$

- Total energy and net energy to society

$$\forall t \in \{0, \ldots, T-1\}, E_t = B_t + \sum_{n=1}^{N} R_{n,t}$$

$$S_t = E_t - \left(\sum_{n=1}^{N} C_{n,t}(R_{n,t}, \alpha_{n,t}) + M_{n,t}\right)$$

# Another Bad Model

- Constraint on the quantity of energy invested for energy production

$$\forall t \in \{0, \ldots, T-1\},$$

$$\exists \sigma_t : C_{n,t}(R_{n,t}, \alpha_{n,t}) + M_{n,t} \leq \frac{1}{\sigma_t} E_t$$

# Another Bad Model

- Further assumptions
  - Energy cost for growth is proportional to growth, and done initially:

$$C_{n,t}\left(R_{n,t},\alpha_{n,t}\right)=\frac{\Delta_{n,t}}{EROEI_{n,t}}\alpha_{n,t}R_{n,t}\text{ if }\alpha_{n,t}\geq 0$$

  - Long-term replacement cost is (i) proportional and (ii) annualized

$$M_{n,t}\left(R_{n,t}\right)=\frac{1}{EROEI_{n,t}}R_{n,t}$$

$$E_0 = 1$$

$$B_0 = 0.85E_0$$

$$R_{1,0} = 0.01E_0$$

$$\sum_{n=2}^{N} R_{n,0} = 0.14E_0$$

$$ERoEI_{1,t} = 9$$

$$\Delta_{1,t} = 20$$

$$\sigma_t = 14$$

Constant growth if possible, else max admissible

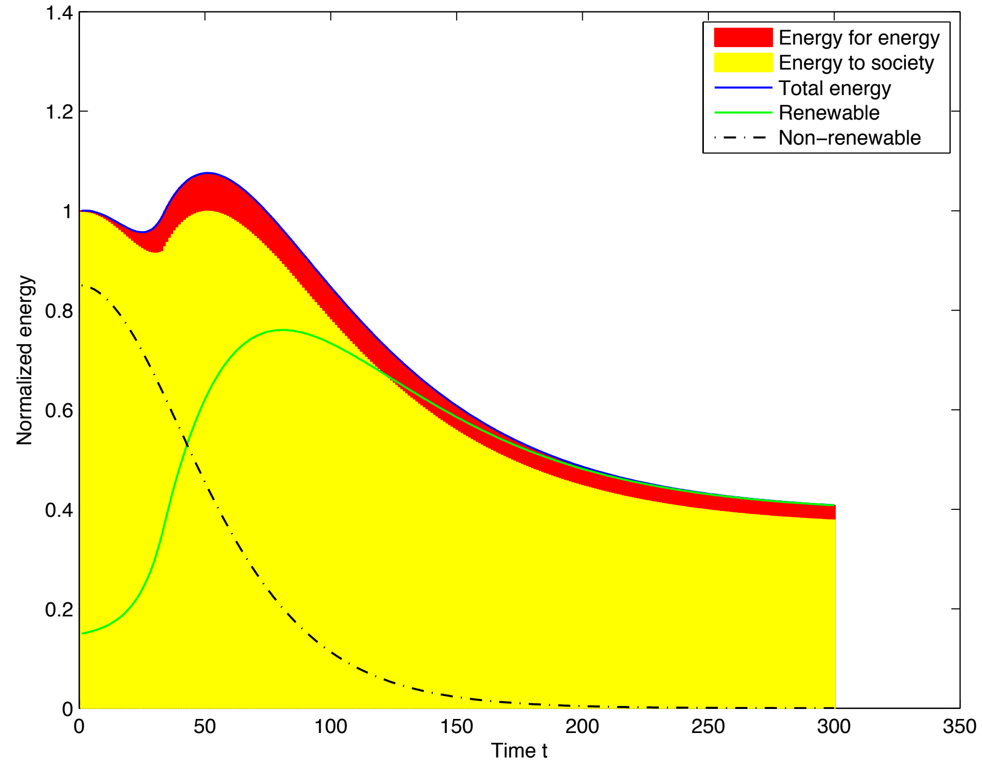



**Fig. 2.** Scenario "peak at time t=0"

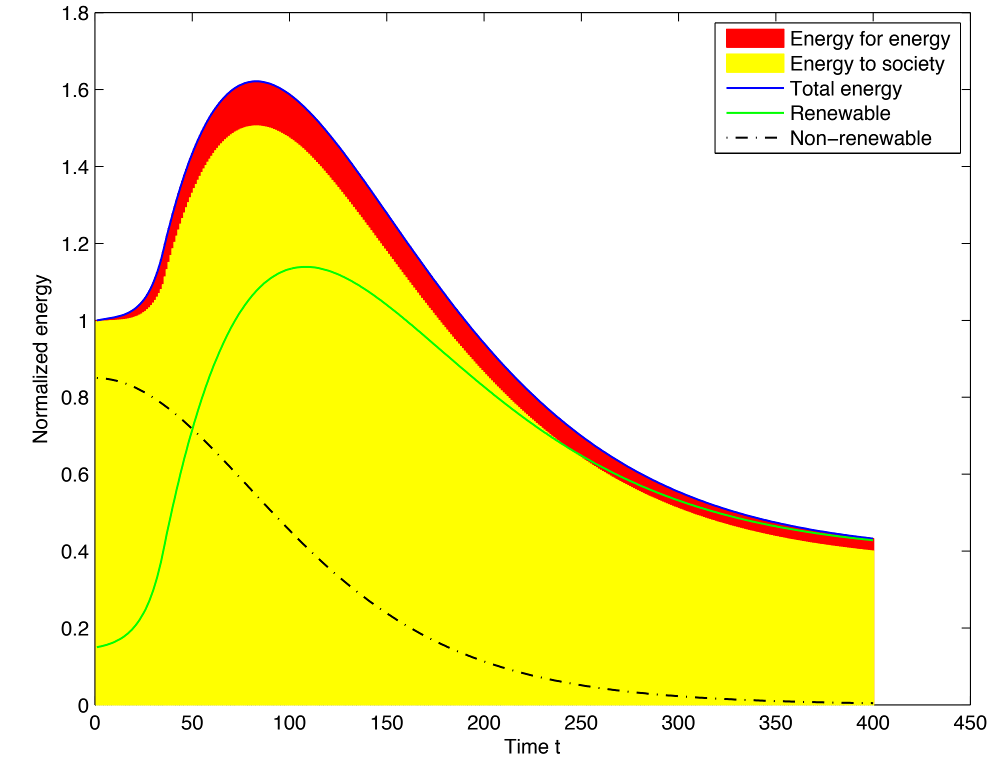



**Fig. 3.** Scenario "plateau at time t=0"

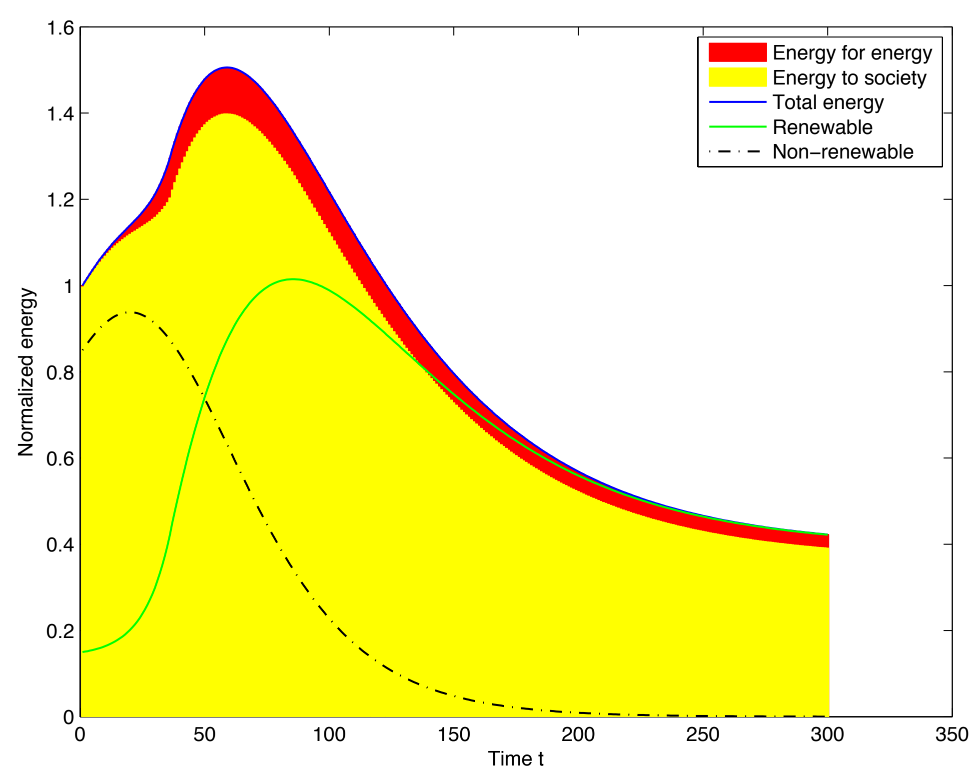



**Fig. 4.** Scenario "peak at time t=20"

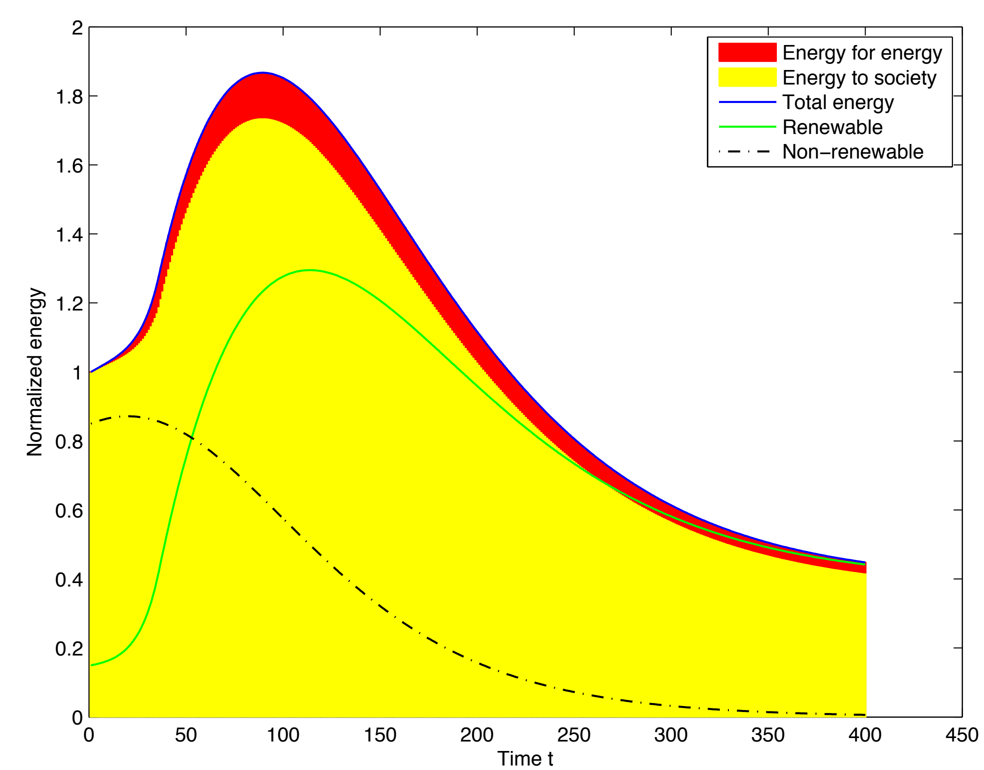
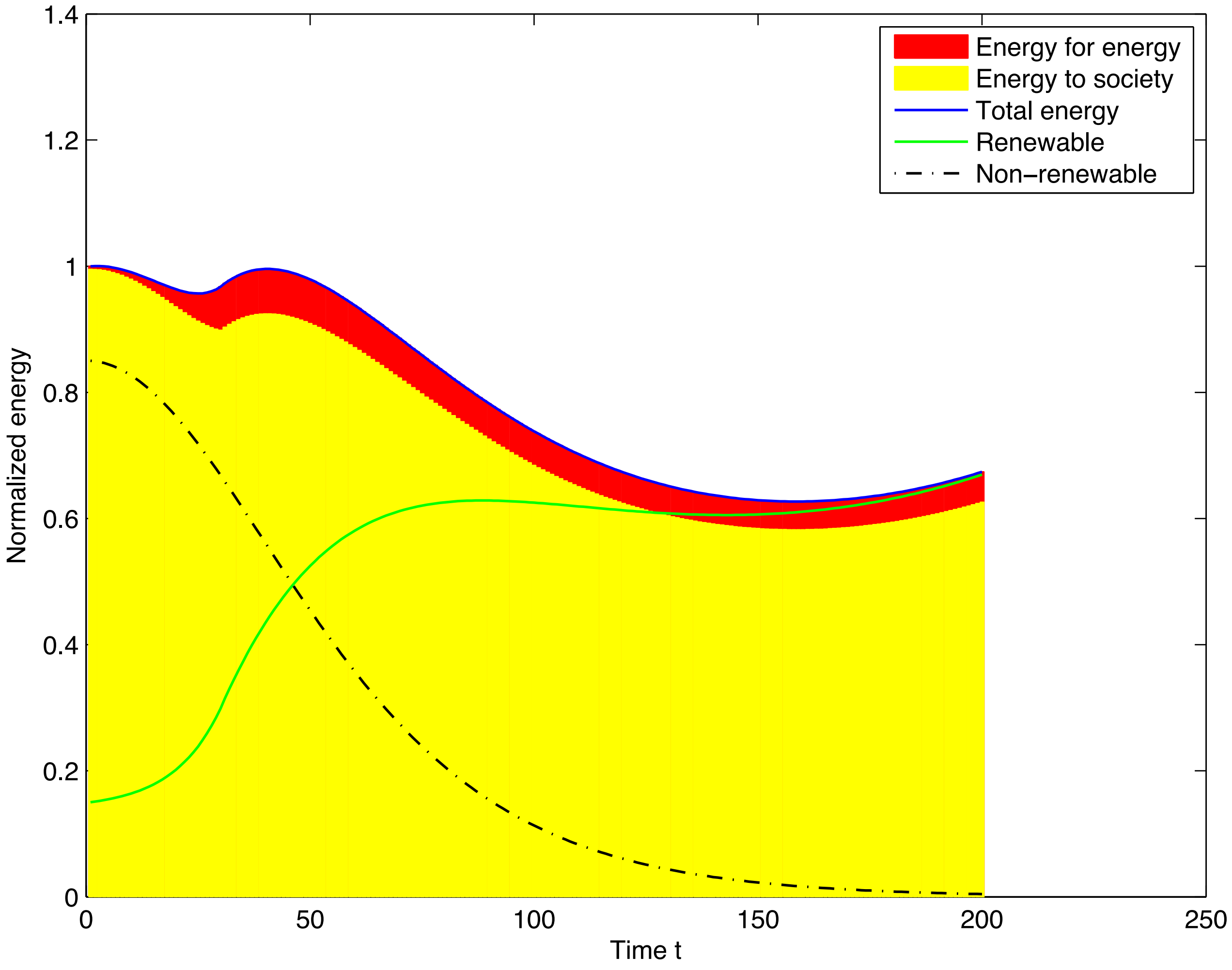



**Fig. 5.** Scenario "plateau at time t=20"

# Another Bad Model

- Increasing the ERoEI parameter




$$\forall t \in \{0, \ldots, T-1\}, ERoEI_{1,t} = 9 + \frac{t}{T}(12-9)$$

# Good Policies?

- What kind of « good policy » can be suggested by such a « bad model »?

    - **Energy efficiency**: « do better with less »

    -> Lots of decision making under uncertainty problems to solve here

- For people interested in Smart Grids: below is link toward a simulator for Active Network Management (ANM) developed by my colleagues at the University of Liège:

    http://www.montefiore.ulg.ac.be/~anm/

# Epilogue

PhR61via [Wikipedia]

# References


[1] Wikipedia, Feu, Domestication par l'Homme
[2] Auzanneau, M. (2011). L'empire romain et la société d'opulence énergétique : un parallèle via lemonde.fr
[3] Tainter, J. (1990). The Collapse of Complex Societies.
[4] Gimel, J. - The Medieval Machine : the industrial Revolution of the Middle Ages, Penguin Books, 1976 (ISBN 978-0-7088-1546-5)
[5] Maddison, A. « When and Why did the West get Richer than the Rest ? »
[6] Wikipedia, Dutch Golden Age, Causes of the Golden Age
[7] Wikipedia, Histoire de la production de l'acier
[8] Wikipedia, Houille
[9] Giraud, G. & Kahraman, Z. (2014). On the Output Elasticity of Primary Energy in OECD countries (1970-2012). Center for European Studies, Working Paper.
[10] Stern, D.I. (2011). From correlation to Granger causality. Crawford School Research Papers. Crawford School Research Paper No 13.
[11] Stern, D.I. & Enflo, K. (2013). Causality Between Energy and Output in the Long-Run. Energy Economics, 2013 - Elsevier.
[12] Auzanneau, M. (2014). Gaël Giraud, du CNRS : « Le vrai rôle de l'énergie va obliger les économistes à changer de dogme » via lemonde.fr
[13] Jancovici, J.M. (2013). Transition énergétique pour tous ! ce que les politiques n'osent pas vous dire, Éditions Odile Jacob, avril 2013. See also J.M. Jancovici's website.
[14] Meilhan, N. (2014). Comprendre ce qui cloche avec l'énergie (et la croissance économique) en 7 slides et 3 minutes.
[15] Wikipedia, Decline of the Roman Empire
[16] Lambert, J., Hall, C., Balogh, S., Poisson, A. and Gupta, A. (2012). EROI of Global Energy Resources - Preliminary Status and Trends - J State University of New York, College of Environmental Science and Forestry Report 1 - Revised Submitted - 2 November 2012 DFID - 59717
[17] Jancovici, J.M. « L'économie aurait-elle un vague rapport avec l'énergie? »(2013), LH Forum, 27 septembre 2013
[18] Fonteneau, R., Murphy, S.A., Wehenkel, L. and Ernst, D. Towards min max generalization in reinforcement learning, in Agents and Artificial Intelligence: International Conference, ICAART 2010, Valencia, Spain, January 2010, Revised Selected Papers. Series: Communications in Computer and Information Science (CCIS), vol. 129, Springer, Heidelberg, 2011, pp. 61–77
[19] Fonteneau, R., Ernst, D., Boigelot, B. and Louveaux, Q. (2013). Min Max Generalization for Deterministic Batch Mode Reinforcement Learning: Relaxation Schemes. SIAM Journal on Control and Optimization
[20] Fonteneau, R. and Ernst, D. On the Dynamics of the Deployment of Renewable Energy Production Capacities. Submitted
[21] Kümmel, R., Ayres, R.U. and Linderberger, D. (2010).Thermodynamic Laws, Economic Methods and the Productive Power of Energy. Journal of Non-Equilibrium Thermodynamics, in press
[22] Gemine, Q., Ernst, D. and Cornelusse, B. (2015). Active network management for electrical distribution systems: problem formulation and benchmark. In press