Partial conservation laws and indexability: past, present, and future

José Niño-Mora
http://alum.mit.edu/www/jnimora

Department of Statistics

Universidad Carlos III de Madrid
www.uc3m.es

Keynote talk at Workshop on restless bandits, index policies and applications in reinforcement learning
Grenoble, 20–21 November, 2023

Work partially supported by grant PID2019-109196GB-I00 funded by MCIN/AEI/10.13039/501100011033
MDP model: optimal control of resource-fueled project

• Discrete-time **restless bandit**, modeling a resource-fueled project:
  - States (finite): \( X(t) \in \mathcal{N} \coloneqq \{1, 2, \ldots, n\} \)
  - Actions (binary): \( A(t) \in \mathcal{A} \coloneqq \{0, 1\} \)
  - One-period rewards: \( r_i^a \)
  - One-period quantity of resource consumed: \( q_i^a \) (\( 0 < q_i^1 > q_i^0 \geq 0 \))
  - Transition probabilities: \( p_{ij}^a \)
  - Discount factor: \( 0 < \beta < 1 \)
  - \( \Pi \): admissible policies (stationary is enough)

• For each **resource price** \( \lambda \in \mathbb{R} \), consider \( \lambda \)-price problem:

\[
\max_{\pi \in \Pi} \mathbb{E}_i^\pi \left[ \sum_{t=0}^{\infty} (r_i^{A(t)} X(t) - \lambda q_i^{A(t)} X(t)) \beta^t \right]
\]

• Optimal value function: \( V_i^*(\lambda) \)
\(\lambda\)-price problem & Bellman equations

- For each resource price \(\lambda \in \mathbb{R}\), consider \(\lambda\)-price problem:

\[
\text{maximize } \mathbb{E}_{\pi}^{\pi} \left[ \sum_{t=0}^{\infty} \left( r^{A(t)} X(t) - \lambda q^{A(t)} X(t) \right) \beta^t \right]
\]

- Opt. val. funct. \(V^*_i(\lambda)\) & opt. policies through Bellman equations:

\[
V^*_i(\lambda) = \max_{a \in \{0,1\}} \left[ r^a_i - \lambda q^a_i + \beta \sum_{j \in \mathcal{N}} p_{ij}^a V^*_j(\lambda), \quad i \in \mathcal{N} \right]
\]

- Write them as:

\[
V^*_i(\lambda) = \max_{a \in \{0,1\}} V^{(a,*)}_i(\lambda), \quad i \in \mathcal{N}
\]
Indexability and Whittle index

- $V_i^*(\lambda)$ and optimal policies determined by Bellman equations:

$$V_i^*(\lambda) = \max_{a \in \{0, 1\}} V_i^{(a,*)}(\lambda), \quad i \in \mathcal{N}$$

- Consider the marginal value function $v_i^*(\lambda) \equiv V_i^{(1,*)}(\lambda) - V_i^{(0,*)}(\lambda)$

Call the project indexable if, for each state $i$:

1. The eqn. $v_i^*(\lambda) = 0$ has a unique root $\lambda = \lambda_i^*$ (Whittle index)
2. $v_i^*(\lambda) > 0$ for $\lambda < \lambda_i^*$
3. $v_i^*(\lambda) < 0$ for $\lambda > \lambda_i^*$
The submodularity approach to indexability

• \( V^*_i(\lambda) \) and optimal policies determined by **Bellman equations**:

\[
V^*_i(\lambda) = \max_{a \in \{0,1\}} V^{(a,\ast)}_i(\lambda), \quad i \in \mathcal{N}
\]

• Write \( v^*_i(\lambda) \triangleq V^{(1,\ast)}_i(\lambda) - V^{(0,\ast)}_i(\lambda) \)

• Suppose can prove that \( V^{(a,\ast)}_i(\lambda) \) is strictly **supermodular** in \((i,a)\)

• i.e., \( v^*_i(\lambda) \) is **increasing** in \( i \) (for state ordering \( i = 1, \ldots, n \)):

\[
v^*_1(\lambda) < v^*_2(\lambda) < \cdots < v^*_n(\lambda)
\]

• This implies optimality of **threshold policies**
The submodularity approach to indexability

- Suppose one can prove that $\forall \lambda$, $v_i^*(\lambda)$ is increasing in $i = 1, \ldots, n$:
  \[
  v_1^*(\lambda) < v_2^*(\lambda) < \cdots < v_n^*(\lambda)
  \]

- This implies optimality of threshold policies

For indexability: prove $\Lambda_0 = (-\infty, \lambda_1^*]$, $\Lambda_i = (\lambda_i^*, \lambda_{i+1}^*]$, $\Lambda_n = (\lambda_n^*, \infty)$

Equivalently, need to further prove that $z^*(\lambda) \nearrow$, spanning $\{0, \ldots, n\}$
Reformulating thru project performance metrics

- **Reward metric:** \( F_{i}^{\pi} \triangleq \mathbb{E}_{i}^{\pi} \left[ \sum_{t=0}^{\infty} r_{X(t)}^A(t) \beta^t \right] \)

- **Resource (usage) metric:** \( G_{i}^{\pi} \triangleq \mathbb{E}_{i}^{\pi} \left[ \sum_{t=0}^{\infty} q_{X(t)}^A(t) \beta^t \right] \)

- **\( \lambda \)-price problem:** maximize \( \max_{\pi \in \Pi} F_{i}^{\pi} - \lambda G_{i}^{\pi} \)

- **Stationary policies are enough:** \( S \)-active policy, for \( S \in 2^N \)

- **\( \lambda \)-price problem:** maximize \( \max_{S \in 2^N} F_{i}^{S} - \lambda G_{i}^{S} \)

- **Optimal value function:** \( V_{i}^{*}(\lambda) = \max_{S \in 2^N} F_{i}^{S} - \lambda G_{i}^{S} \)
Marginal project performance metrics

- Marginal reward metric: \( f_i^S \triangleq F_i^{(1,S)} - F_i^{(0,S)} \)

- Marginal resource (usage) metric: \( g_i^S \triangleq G_i^{(1,S)} - G_i^{(0,S)} \)

- Marginal productivity metric: \( m_i^S \triangleq \frac{f_i^S}{g_i^S} \), provided that \( g_i^S \neq 0 \)
Properties of performance metrics

• Optimal value function: \( V^*_i(\lambda) = \max_{S \in 2^N} F^S_i - \lambda G^S_i \)

Properties of \( V^*_i(\lambda) \) (as a function of \( \lambda \)):
1. Convex (hence continuous)
2. Piecewise linear
3. Nonincreasing

• Recall: \( v^*_i(\lambda) \triangleq V^i_{(1,*)}(\lambda) - V^i_{(0,*)}(\lambda) \)

Properties of \( v^*_i(\lambda) \) (as a function of \( \lambda \)):
1. Difference of convex nonincreasing functions (hence continuous)
2. Piecewise linear
3. Need not be monotonic
Motivation of PCL-indexability conditions

• Suppose project is indexable w/ Whittle index satisfying

$$\lambda_1^* < \lambda_2^* < \cdots < \lambda_n^*$$

• Write $$S_z \triangleq \{ j \in \mathcal{N} : j > z \}$$ (active set of $$z$$-policy)

• Note: $$S_0 = \mathcal{N}$$, $$S_i = \{ i + 1, \ldots, n \}$$ for $$0 \leq i < n$$, $$S_n = \emptyset$$

• Then

$$V_i^*(\lambda) = \begin{cases} 
F_i^{S_0} - \lambda G_i^{S_0}, & \text{if } \lambda \leq \lambda_1^* \\
F_i^{S_1} - \lambda G_i^{S_1}, & \text{if } \lambda_1^* \leq \lambda \leq \lambda_2^* \\
\vdots & \vdots \\
F_i^{S_{n-1}} - \lambda G_i^{S_{n-1}}, & \text{if } \lambda_{n-1}^* \leq \lambda \leq \lambda_n^* \\
F_i^{S_n} - \lambda G_i^{S_n}, & \text{if } \lambda \geq \lambda_n^* 
\end{cases}$$
Motivation of PCL-indexability conditions

• Then

\[ V_i^*(\lambda) = \begin{cases} 
F_i^{S_0} - \lambda G_i^{S_0}, & \text{if } \lambda \leq \lambda_1^* \\
F_i^{S_1} - \lambda G_i^{S_1}, & \text{if } \lambda_1^* \leq \lambda \leq \lambda_2^* \\
\vdots & \vdots \\
F_i^{S_{n-1}} - \lambda G_i^{S_{n-1}}, & \text{if } \lambda_{n-1}^* \leq \lambda \leq \lambda_n^* \\
F_i^{S_n} - \lambda G_i^{S_n}, & \text{if } \lambda \geq \lambda_n^* 
\end{cases} \]

• By continuity of \( V_i^*(\lambda) \), it follows that the equation

\[ F_i^{S_j-1} - \lambda G_i^{S_j-1} = F_i^{S_j} - \lambda G_i^{S_j} \]

i.e.,

\[ F_i^{S_j-1} - F_i^{S_j} = \lambda (G_i^{S_j-1} - G_i^{S_j}) \]

has a unique root given by \( \lambda = \lambda_j^* \)
Some relations between performance metrics

For \( j \in S^c \),

\[
F_i^{S\cup\{j\}} - F_i^S = f_{ij} x_{ij}^{1,S\cup\{j\}} = f_{ij} x_{ij}^{0,S}
\]

\[
G_i^{S\cup\{j\}} - G_i^S = g_{ij} x_{ij}^{1,S\cup\{j\}} = g_{ij} x_{ij}^{0,S}
\]

where

\[
x_{ij}^{a,\pi} \triangleq \mathbb{E}_i^\pi \left[ \sum_{t=0}^{\infty} 1\{A(t)=a\} \beta^t \right]
\]

Equivalently: for \( j \in S \),

\[
F_i^S - F_i^{S\setminus\{j\}} = f_{ij}^{S\setminus\{j\}} x_{ij}^{1,S} = f_{ij} x_{ij}^{0,S\setminus\{j\}}
\]

\[
G_i^S - G_i^{S\setminus\{j\}} = g_{ij}^{S\setminus\{j\}} x_{ij}^{1,S} = g_{ij} x_{ij}^{0,S\setminus\{j\}}
\]
Some relations between performance metrics

For $j \in S^c$ (since $x_{jj}^{1, S \cup \{j\}}, x_{jj}^{0, S} > 0$),

$$\text{sgn} \left( G_j^{S \cup \{j\}} - G_j^S \right) = \text{sgn} g_j^S = \text{sgn} g_j^{S \cup \{j\}}$$

Equivalently: for $j \in S$,

$$\text{sgn} \left( G_j^S - G_j^{S \setminus \{j\}} \right) = \text{sgn} g_j^{S \setminus \{j\}} = \text{sgn} g_j^S$$
Some relations between performance metrics

For $j \in S^c$, if $g_j^S \neq 0$ (recall $m_j^S \triangleq f_j^S / g_j^S$),

$$F_{i}^{S \cup \{j\}} - F_{i}^{S} = m_j^S (G_{i}^{S \cup \{j\}} - G_{i}^{S}) = m_j^{S \cup \{j\}} (G_{i}^{S \cup \{j\}} - G_{i}^{S})$$

Hence,

$$m_j^{S} = m_j^{S \cup \{j\}}$$

Equivalently: for $j \in S$, if $g_j^S \neq 0$,

$$F_{i}^{S} - F_{i}^{S \setminus \{j\}} = m_j^{S \setminus \{j\}} (G_{i}^{S} - G_{i}^{S \setminus \{j\}}) = m_j^{S} (G_{i}^{S} - G_{i}^{S \setminus \{j\}})$$

Hence,

$$m_j^{S \setminus \{j\}} = m_j^{S}$$
Some relations between performance metrics

For $j \in S^c$, if $g_j^S \neq 0$ (recall $m_j^S \triangleq f_j^S / g_j^S$),

$$F_i^{S \cup \{j\}} - F_i^S = m_j^S (G_i^{S \cup \{j\}} - G_i^S) = m_j^{S \cup \{j\}} (G_i^{S \cup \{j\}} - G_i^S)$$

Hence,

$$m_j^S = m_j^{S \cup \{j\}}$$

- Recall: By continuity of $V_i^*(\lambda)$, it follows that

$$F_i^{S_j^{-1}} - F_i^{S_j} = \lambda_j^* (G_i^{S_j^{-1}} - G_i^{S_j})$$

Hence, if $g_j^S \neq 0$, we have

$$\lambda_j^* = m_j^{S_j^{-1}} = m_j^{S_j}$$
Some relations between performance metrics

- Suppose indexable with $\lambda_1^* < \cdots < \lambda_n^*$

$$V_i^*(\lambda) = \begin{cases} F_i^{S_0} - \lambda G_i^{S_0}, & \text{if } \lambda \leq \lambda_1^* \\ F_i^{S_1} - \lambda G_i^{S_1}, & \text{if } \lambda_1^* \leq \lambda \leq \lambda_2^* \\ \vdots & \vdots \\ F_i^{S_{n-1}} - \lambda G_i^{S_{n-1}}, & \text{if } \lambda_{n-1}^* \leq \lambda \leq \lambda_n^* \\ F_i^{S_n} - \lambda G_i^{S_n}, & \text{if } \lambda \geq \lambda_n^* \end{cases}$$

- For $\lambda < \lambda_1^*$, $F_1^{S_0} - F_1^{S_1} > \lambda(G_1^{S_0} - G_1^{S_1})$

- For $\lambda_1^* < \lambda < \lambda_2^*$, $F_1^{S_0} - F_1^{S_1} < \lambda(G_1^{S_0} - G_1^{S_1})$

This implies $G_1^{S_0} > G_1^{S_1}$, i.e., $g_1^{S_0}, g_1^{S_1} > 0$, and $\lambda_1^* = m_1^{S_0}$
Some relations between performance metrics

- Suppose indexable w/ $\lambda_1^* < \cdots < \lambda_n^*$

$$V_i^*(\lambda) = \begin{cases} 
F_i^{S_0} - \lambda G_i^{S_0}, & \text{if } \lambda \leq \lambda_1^* \\
F_i^{S_1} - \lambda G_i^{S_1}, & \text{if } \lambda_1^* \leq \lambda \leq \lambda_2^* \\
\vdots & \vdots \\
F_i^{S_{n-1}} - \lambda G_i^{S_{n-1}}, & \text{if } \lambda_{n-1}^* \leq \lambda \leq \lambda_n^* \\
F_i^{S_n} - \lambda G_i^{S_n}, & \text{if } \lambda \geq \lambda_n^* 
\end{cases}$$

- For $\lambda_1^* < \lambda < \lambda_2^*$, $F_2^{S_1} - F_2^{S_2} < \lambda (G_2^{S_1} - G_2^{S_2})$
- For $\lambda_2^* < \lambda < \lambda_3^*$, $F_2^{S_1} - F_2^{S_2} < \lambda (G_2^{S_1} - G_2^{S_2})$

This implies $G_2^{S_1} > G_2^{S_2}$, i.e., $g_2^{S_1}, g_2^{S_2} > 0$, and $\lambda_2^* = m_2^{S_1} = m_2^{S_2}$
Some relations between performance metrics

• Suppose indexable w/ $\lambda_1^* < \cdots < \lambda_n^*$

$$V_i^*(\lambda) = \begin{cases} F_i^{S_0} - \lambda G_i^{S_0}, & \text{if } \lambda \leq \lambda_1^* \\ F_i^{S_1} - \lambda G_i^{S_1}, & \text{if } \lambda_1^* \leq \lambda \leq \lambda_2^* \\ \vdots & \vdots \\ F_i^{S_{n-1}} - \lambda G_i^{S_{n-1}}, & \text{if } \lambda_{n-1}^* \leq \lambda \leq \lambda_n^* \\ F_i^{S_n} - \lambda G_i^{S_n}, & \text{if } \lambda \geq \lambda_n^* \end{cases}$$

This implies $G_i^{S_{i-1}} > G_i^{S_i}$, i.e., $g_i^{S_{i-1}}, g_i^{S_i} > 0$, and $\lambda_i^* = m_i^{S_{i-1}}$.
Reformulating marginal value function \( v_i^*(\lambda) \)

- Suppose indexable w/ \( \lambda_1^* < \cdots < \lambda_n^* \)

\[
v_i^*(\lambda) = \begin{cases} 
F_i^{(1,S_0)} - F_i^{(0,S_0)} - \lambda(G_i^{(1,S_0)} - G_i^{(0,S_0)}), & \lambda \leq \lambda_1^* \\
F_i^{(1,S_1)} - F_i^{(0,S_1)} - \lambda(G_i^{(1,S_1)} - G_i^{(0,S_1)}), & \lambda_1^* \leq \lambda \leq \lambda_2^* \\
\vdots & \\
F_i^{(1,S_{n-1})} - F_i^{(0,S_{n-1})} - \lambda(G_i^{(1,S_{n-1})} - G_i^{(0,S_{n-1})}), & \lambda_{n-1}^* \leq \lambda \leq \lambda_n^* \\
F_i^{(1,S_n)} - F_i^{(0,S_n)} - \lambda(G_i^{(1,S_n)} - G_i^{(0,S_n)}), & \lambda \geq \lambda_n^*
\end{cases}
\]

i.e.,

\[
v_i^*(\lambda) = \begin{cases} 
f_i^{S_0} - \lambda g_i^{S_0}, & \lambda \leq \lambda_1^* \\
f_i^{S_1} - \lambda g_i^{S_1}, & \lambda_1^* \leq \lambda \leq \lambda_2^* \\
\vdots & \\
f_i^{S_{n-1}} - \lambda g_i^{S_{n-1}}, & \lambda_{n-1}^* \leq \lambda \leq \lambda_n^* \\
f_i^{S_n} - \lambda g_i^{S_n}, & \lambda \geq \lambda_n^*
\end{cases}
\]
Reformulation of of marginal value function

• Suppose indexable w/ $\lambda_1^* < \cdots < \lambda_n^*$

$$v_i^*(\lambda) = \begin{cases} 
 f_i^{S_0} - \lambda g_i^{S_0}, & \lambda \leq \lambda_1^* \\
 f_i^{S_1} - \lambda g_i^{S_1}, & \lambda_1^* \leq \lambda \leq \lambda_2^* \\
 \vdots & \vdots \\
 f_i^{S_{n-1}} - \lambda g_i^{S_{n-1}}, & \lambda_{n-1}^* \leq \lambda \leq \lambda_n^* \\
 f_i^{S_n} - \lambda g_i^{S_n}, & \lambda \geq \lambda_n^*
\end{cases}$$

• Note that, e.g.:

$$v_1^*(\lambda) = f_1^{S_0} - \lambda g_1^{S_0} > 0, \quad \lambda < \lambda_1^*$$

$$v_1^*(\lambda) = f_1^{S_1} - \lambda g_1^{S_1} < 0, \quad \lambda_1^* < \lambda < \lambda_2^*$$
Some implications of submodularity-based conditions

• Suppose indexable w/ $\lambda_1^* < \cdots < \lambda_n^*$

$$v_i^*(\lambda) = \begin{cases} 
  f_i^{S_0} - \lambda g_i^{S_0}, & \lambda \leq \lambda_1^* \\
  f_i^{S_1} - \lambda g_i^{S_1}, & \lambda_1^* \leq \lambda \leq \lambda_2^* \\
  \vdots & \vdots \\
  f_i^{S_{n-1}} - \lambda g_i^{S_{n-1}}, & \lambda_{n-1}^* \leq \lambda \leq \lambda_n^* \\
  f_i^{S_n} - \lambda g_i^{S_n}, & \lambda \geq \lambda_n^* 
\end{cases}$$

$v_1^*(\lambda) < v_2^*(\lambda) < \cdots < v_n^*(\lambda), \quad \lambda \leq \lambda_1^*$

i.e.,

$$f_1^{S_0} - \lambda g_1^{S_0} < f_2^{S_0} - \lambda g_2^{S_0} < \cdots < f_n^{S_0} - \lambda g_n^{S_0}, \quad \lambda \leq \lambda_1^*$$

• This implies the following (not required by PCLI conditions!):

$$(0 <) \quad g_1^{S_0} < g_2^{S_0} < \cdots < g_n^{S_0}$$
Further implications of submodularity-based cond.

• Suppose indexable w/ $\lambda_1^* < \cdots < \lambda_n^*$

$$v_i^*(\lambda) = \begin{cases} 
  f_i^{S_0} - \lambda g_i^{S_0}, & \lambda \leq \lambda_1^* \\
  f_i^{S_1} - \lambda g_i^{S_1}, & \lambda_1^* \leq \lambda \leq \lambda_2^* \\
  \vdots & \vdots \\
  f_i^{S_{n-1}} - \lambda g_i^{S_{n-1}}, & \lambda_{n-1}^* \leq \lambda \leq \lambda_n^* \\
  f_i^{S_n} - \lambda g_i^{S_n}, & \lambda \geq \lambda_n^* 
\end{cases}$$

$$v_1^*(\lambda) < v_2^*(\lambda) < \cdots < v_n^*(\lambda), \quad \lambda \geq \lambda_n^*$$

i.e.,

$$f_1^{S_n} - \lambda g_1^{S_n} < f_2^{S_n} - \lambda g_2^{S_n} < \cdots < f_n^{S_n} - \lambda g_n^{S_n}, \quad \lambda \geq \lambda_n^*$$

• This implies the following (not required by PCLI conditions!):

$$g_1^{S_n} > g_2^{S_n} > \cdots > g_n^{S_n} \ (> 0)$$
PCL-indexability conditions wrt state ordering $1, \ldots, n$

- Wanted: indexability consistently w/ optim. of threshold policies
  $S_0 = \{1, 2, \ldots\}$, $S_1 = \{2, 3, \ldots\}$, \ldots $S_n = \emptyset$
- Active set family: $\mathcal{F} \triangleq \{S_0, S_1, \ldots, S_n\}$

PCL($\mathcal{F}$)-indexability conditions:
- (PCLI1) $g_i^S > 0$ for every $i \in \mathcal{N}$, $S \in \mathcal{F}$
- (PCLI2) For MP index $m_i^* \triangleq m_i^{S_i-1} = m_i^{S_i} : m_1^* \leq \cdots \leq m_n^*$

Verification theorem (part (b) more recent)
(a) (PCLI1)+(PCLI2) $\implies$ indexable w/ $\lambda_i^* = m_i^*$
(b) Under (PCLI1), indexable w/ $\lambda_i^* = m_i^* \iff$ (PCLI2)
Example: optimal admission control to a queue

- Holding cost $h_i$, rejection cost $\nu$, active action = reject

\[ \lambda_i \rightarrow \text{Entry gate} \]

\[ \mu_i \]
Ex: optimal admission control to a queue (NM ‘02)

• Write $d_i \triangleq \mu_i - \lambda_i$

• Then, $d_i$ concave nondecreasing $\implies$ (PCLI1)

• If, further, $h_i$ is convex nondecreasing $\implies$ (PCLI2)

• Such conditions ensure indexability wrt threshold policies

• Sharpest conditions (as far as I know)
PCL-indexability conditions wrt state ordering 1, 2, . . .

• Wanted: indexability consistently w/ optim. of threshold policies
  \( S_0 = \{1, 2, \ldots\} \), \( S_1 = \{2, 3, \ldots\} \), \ldots (countable)

• Active set family: \( \mathcal{F} \triangleq \{S_0, S_1, \ldots\} \)

PCL(\( \mathcal{F} \))-indexability conditions:
- (PCLI1) \( g^S_i > 0 \) for every \( i \in \mathcal{N}, S \in \mathcal{F} \)
- (PCLI2) For MP index \( m^*_i \triangleq m^{S_{i-1}}_i = m^{S_i}_i: m^*_1 \leq m^*_2 \leq \ldots \)

Verification theorem: (part (b) more recent)
(a) (PCLI1)+(PCLI2) \( \implies \) indexable w/ \( \lambda^*_i = m^*_i \)
(b) Under (PCLI1), indexable w/ \( \lambda^*_i = m^*_i \) \( \iff \) (PCLI2)
Optimal control of a MTO/MTS M/G/1 queue, NM ’06

- Net backorder cost $h_i$, service cost $\nu$, active action = serve
- (PCLI1) holds
- if $h_i$ convex, (PCLI2) holds. Hence, indexable wrt threshold policies
PCL-indexability conditions w/ unknown state ordering

- Typical situation in multidimensional state models
- Need to postulate a structured family of policies, w/ active sets $\mathcal{F}$, which one thinks might be optimal (based on insight)
- Wanted: indexability consistently w/ optimality of $\mathcal{F}$-policies
- Note: need $\emptyset, \mathcal{N} \in \mathcal{F}$, and natural connected properties of $\mathcal{F}$

$(\text{PCLI}1)$ $g_i^S > 0$ for every $i \in \mathcal{N}, S \in \mathcal{F}$

- How to define $(\text{PCLI}2)$? Don’t know a priori the “right” state ordering
- Will construct it adaptively
Adaptive-greedy($\mathcal{F}$) algorithm and (PCLI2)

- Start w/ $S_0 \triangleq \mathcal{N}$ (which must be in $\mathcal{F}$)
- Pick $i_1 \in \arg \max_{i \in S_0: \{i\} \in \mathcal{F}} m_{i_1}^{S_0}; \quad m_{i_1}^* := m_{i_1}^{S_0}; \quad S_1 := S_0 \setminus \{i_1\}$
- Pick $i_2 \in \arg \max_{i \in S_1: \{i\} \in \mathcal{F}} m_{i_2}^{S_1}; \quad m_{i_2}^* := m_{i_2}^{S_1}; \quad S_2 := S_1 \setminus \{i_2\}$
- And so on
- $i_n \in \arg \max_{i: S_{n-1} \setminus \{i\} \in \mathcal{F}} m_{i_n}^{S_{n-1}}; \quad m_{i_n}^* := m_{i_n}^{S_{n-1}}; \quad S_n := S_{n-1} \setminus \{i_n\} = \emptyset$

(PCLI2): $m_{i_1}^* \leq m_{i_2}^* \leq \cdots \leq m_{i_n}^*$

Verification theorem (NM 2001, 2002): (b) more recent)

(a) (PCLI1)+(PCLI2) $\implies$ indexable w/ $\lambda_i^* = m_i^*$

(b) Under (PCLI1), indexable w/ $\lambda_i^* = m_i^* \iff$ (PCLI2)
Extension to countably infinite state

• Start w/ \( S_0 \triangleq \mathcal{N} \) (which must be in \( \mathcal{F} \))
• Pick \( i_1 \in \arg \max_{i \in S_0 : S_0 \setminus \{i\} \in \mathcal{F}} m_{i_1}^{S_0} \); \( m_{i_1}^{*} := m_{i_1}^{S_0} \); \( S_1 := S_0 \setminus \{i_1\} \)
• Pick \( i_2 \in \arg \max_{i \in S_1 : S_1 \setminus \{i\} \in \mathcal{F}} m_{i_2}^{S_1} \); \( m_{i_2}^{*} := m_{i_2}^{S_1} \); \( S_2 := S_1 \setminus \{i_2\} \)
• And so on

\((\text{PCLI2}): m_{i_1}^{*} \leq m_{i_2}^{*} \leq \cdots, \text{w/ } \{i_k : k = 1, 2, \ldots\} = \mathcal{N}\)

Verification theorem (NM 2006): ((b) more recent)
(a) \((\text{PCLI1}) + (\text{PCLI2}) \iff \text{indexable w/ } \lambda_i^{*} = m_i^{*}\)
(b) Under (PCLI1), indexable w/ \( \lambda_i^{*} = m_i^{*} \iff (\text{PCLI2}) \)
Further examples of PCL-indexable models

- Finite-buffer delay-/loss-sensitive $M/M/1$ queue (NM ’06)
- Bandits w/ switching costs (NM ’08)
- Finite-horizon bandits (NM ’11)
- Web crawling model (NM ’14)
- Age-of-Information scheduling model (NM ’23)
What if the model is not PCL-indexable?

- Example: bandits w/ switching delays (NM ’07, ’21)

- In NM ’07: extension of adaptive-greedy index algorithm which can handle that (relaxes some PCL-indexability requirements, but still ensures indexability)

- It works for bandits w/ switching delays
Fast index computation

- Fast block-implementations of index algorithms (both w/ and w/out PCLs): NM ’07, NM ’20) for given $\mathcal{F}$

- $O(n^3)$ time if $\mathcal{F} = 2^\mathcal{N}$ w/ $\mathcal{N} = \{1, \ldots, n\}$

- But faster if $\mathcal{F} \subset 2^\mathcal{N}$

- Recent faster implementation for $\mathcal{F} = 2^\mathcal{N}$: Gast, Gaujal and Khun ’23

- See also $O(n^3)$ algorithm of Akbarzadeh & Mahajan ’22 (for $\mathcal{F} = 2^\mathcal{N}$)
A (too) brief history of conservation laws (CLs)

• CLs: fundamental invariance relations on performance metrics for stochastic scheduling models, explain optimality of index policies

• Kleinrock’s ’65 work CL, multiclass M/G/1 queue: \( \sum_{j \in \mathcal{N}} \rho_j \bar{W}_j^\pi \equiv b^\mathcal{N} \)

• Coffman & Mitrani ’80, Gelenbe & Mitrani ’80: polyhedral characterization of waiting time performance in multiclass M/G/1 queue

\[
\sum_{j \in S} \rho_j \bar{W}_j^\pi \geq b^S, \quad S \subset \mathcal{N}
\]

• Shanthikumar & Yao ’92: Framework of strong CLs

• Bertsimas & NM ’96: generalized CLs, (nonrestless) MABP:

\[
\sum_{j \in S} g_j^S x_j^\pi \geq b^S, \quad S \subset \mathcal{N}; \quad \sum_{j \in \mathcal{N}} g_j^\mathcal{N} x_j^\pi \equiv b^\mathcal{N}
\]

...
Partial CLs (PCLs)

- Bertsimas & NM ’96: **generalized CLs**, (nonrestless) MABP:
  \[
  \sum_{j \in S} g_j^S x_j^\pi \geq b^S, \quad S \subseteq \mathcal{N}; \quad \sum_{j \in \mathcal{N}} g_j^\mathcal{N} x_j^\pi \equiv b^\mathcal{N}
  \]

- NM ’01, ’02: **partial CLs**, single restless project:
  \[
  G_i^\pi + \sum_{j \in S} g_j^S x_{ij}^{0,\pi} \geq G_i^S, \quad \mathcal{N} \neq S \in \mathcal{F}; \quad G_i^\pi + \sum_{j \in \mathcal{N}} g_j^\mathcal{N} x_{ij}^{0,\pi} \equiv G_i^\mathcal{N}, \ldots
  \]

- \( \mathcal{F} \subseteq 2^\mathcal{N} \), but typically one takes \( \mathcal{F} \subset 2^\mathcal{N} \), i.e., \( \mathcal{F} \) is a partial collection of subsets of \( \mathcal{N} \)
PCLs & indices for multi-gear restless bandits (NM ’22)

- Weber ’07 sketched extension of Whittle index to multi-action bandits
- NM ’08: outlined extension of PCLs for multi-act; NM ’22: full analysis
- A multi-gear project can be operated in multiple gears $a = 0, 1, \ldots, A$
- Higher gears entail larger resource consumption:

  \[
  0 \leq q_i^0 < q_i^1 < \cdots < q_i^A, \quad i \in \mathcal{N} = \{1, \ldots, N\}
  \]

- $\lambda$-price problem:

  \[
  \min_{\pi \in \Pi} \mathbb{E}_{\pi} \left[ \sum_{t=0}^{\infty} \left( h_{s(t)}^{a(t)} + \lambda q_{s(t)}^{a(t)} \right) \beta^t \right]
  \]
Indexability of multi-gear restless bandits

- **λ-price problem:**

\[
\min_{\pi \in \Pi} \mathbb{E}_i \left[ \sum_{t=0}^{\infty} \left( h_{i}(t) + \lambda q_{i}(t) \right) \beta^t \right]
\]

**Definition** We call the above multi-gear bandit model *indexable* if there exist critical resource prices \( \lambda_{i}^{*,a} \) for every state \( i \) and active action (gear) \( a \geq 1 \) satisfying \( \lambda_{i}^{*,A} \leq \cdots \leq \lambda_{i}^{*,1} \), such that, for any such state and resource price \( \lambda \in \mathbb{R} \): (i) action 0 is \( \lambda \)-optimal in state \( i \) iff \( \lambda \geq \lambda_{i}^{*,1} \); (ii) action \( 1 \leq a \leq A - 1 \) is \( \lambda \)-optimal in state \( i \) iff \( \lambda_{i}^{*,a+1} \leq \lambda \leq \lambda_{i}^{*,a} \); and (iii) action \( A \) is \( \lambda \)-optimal in state \( i \) iff \( \lambda \leq \lambda_{i}^{*,A} \). We call \( \lambda_{i}^{*,a} \) the model’s *dynamic allocation index (DAI)*, viewed as a function of \( (i, a) \).
**Definition** We call a multi-gear bandit model PCL-indexable with respect to $\mathcal{F}$-policies, or $\text{PCL}(\mathcal{F})$-indexable, if:

- (PCLI1) $g_{j,a-1}^a(S) > 0$ for every policy $S \in \mathcal{F}$, active action $a \geq 1$, and state $j \in \mathcal{N}$;

- (PCLI2) Downshift adaptive-greedy algorithm $\text{DS}(\mathcal{F})$ computes the MP index $m_{j,k}^{*,a}$ in order:
  
  $$m_{j_1}^{*,a_1} \leq m_{j_2}^{*,a_2} \leq \cdots \leq m_{j_K}^{*,a_K}.$$  

**Theorem**

If a multi-gear bandit model is $\text{PCL}(\mathcal{F})$-indexable, then it is $\mathcal{F}$-indexable with its DAI being given by its MPI, i.e., $\lambda_{j,a}^{*,a} = m_{j,a}^{*,a}$. 


PCLs for real-state restless bandits

- Real-state restless bandits: sensor scheduling & target tracking POMDP models
- Real-state: probability of channel on (in cognitive radio), posterior variance (target tracking)
- Early results on indexability of real-state restless bandits: Liu & Zhao ’08, ’10, Le Ny et al. (2008), cognitive radio
- For target tracking, La Scala & Moran ’06, Kalman filter model, yet no tools for analysis
- NM ’08: outline of PCLs for real-state restless bandits, experiments
PCLs for real-state restless bandits

• Notation: \( g(x, z) = g_x^{(z, \infty)} \), etc.

\[
\begin{align*}
\text{(PCLI1)} & \quad g(x, z) > 0 \text{ for every state } x \text{ and threshold } z \\
\text{(PCLI2)} & \quad \text{MP index } m(x) \triangleq f(x, x)/g(x, x) \text{ continuous } & \nearrow \\
\text{(PCLI3)} & \quad F(x, z_2) - F(x, z_1) = \int_{(z_1, z_2]} m(z) G(x, dz) \text{ (Lebesgue–Stieltjes)}, \text{ i.e., } m(\cdot): \text{Radon–Nikodym deriv. of } F(x, \cdot) \text{ wrt } G(x, \cdot)
\end{align*}
\]

**Verification theorem (NM ’15, ’20):** (PCLI1)+(PCLI2)+(PCLI3) \( \implies \) indexable w/ Whittle index \( \lambda^*(x) = m(x) \)

• Best application to date: Dance & Silander ’19 (Kalman filter RBs)
• Empirical application on model extension (submitted ’23): (joint work w/ Hao et al.)
Future challenges

- Proving indexability of multi-dimensional discrete-state RB models
- PCL-indexability conditions for multi-dimensional continuous-state RB models
- Implementing & testing adaptive-greedy algorithm for multi-gear bandits
- PCL-indexability conditions for multi-gear continuous-state RB models
- . . .
- Anybody wants to join in?
- Note: some references in final slide
Some references

- NM 2007. Dynamic priority allocation via restless bandit marginal productivity indices (with discussion). *TOP*
- NM 2023. Markovian restless bandits and index policies: A review. *Mathematics*