# Planning Energy Investment under Uncertainty

**Claudia Sagastizábal**

joint with Felipe Atenas, University of Chile

**Cambridge, MESW02, March 19 2019**

**Electricity systems of the future:**
**incentives, regulation and analysis for efficient investment**

# Towards sustainable energy systems

- Sustainability concerns have increased the participation of renewable sources like wind and sun in electrical systems
- How to blend the different technologies to support the transition to a low carbon energy system?
- The success in this transition depends on the change of investment strategies

# Towards sustainable energy systems

- Sustainability concerns have increased the participation of renewable sources like wind and sun in electrical systems
- How to blend the different technologies to support the transition to a low carbon energy system?
- The success in this transition depends on the change of investment strategies

**Generation Expansion Planning**

# Generation Expansion Planning

Long term problem (10 years or more)

Minimize investment and operational costs s.t.

- satisfaction of demand (unknown)
- financial constraints: construction times, delays (unknown)
- environmental constraints: limits on carbon emissions (unknown)
- operational data: fuel costs (unknown)
- % wind power in the mix (unknown)
- inflows (unknown)

**Given:**

- **initial power system**
- **investment portfolio** { **more generation capacity**, **more transmission capacity (DC)** }

# Generation Expansion Planning

Long term problem (10 years or more)

Minimize investment and operational costs s.t.

- satisfaction of demand (unknown)
- financial constraints: construction times, delays (unknown)
- **environmental constraints**: limits on carbon emissions (unknown)
- operational data: fuel costs (unknown)
- % wind power in the mix (unknown)
- inflows (unknown)

**Given:**

- **initial power system**
- **investment portfolio** { **more generation capacity** / **more transmission capacity (DC)** }

# Why environmental constraints?

**Controlling emissions of coal based plants**

~~OK!~~

# Why environmental constraints?

**Controlling emissions of coal based plants**

~~OK!~~

**BUT:**

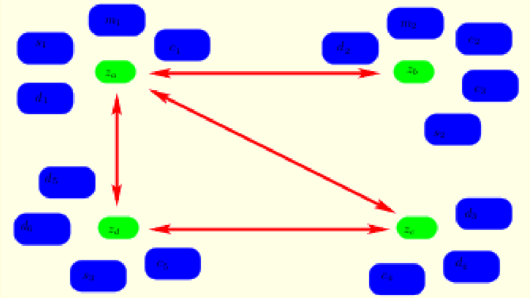
- They are already in the mix
- They can be built by private companies (“cheap and fast”)
- Peakers respond fast to sudden operational changes
- Gas fueled plants come with technology to reduce emissions

# Some notation and elements of the model

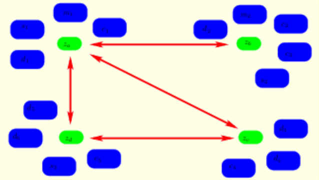
- Long time horizon $T$
- Investment decisions $x$, generation decisions $y$
- Multidimensional uncertainty $\xi$
- Generation assets on a continental scale

# Some notation and elements of the model

- Long time horizon $T$
- Investment decisions $x$, generation decisions $y$
- Multidimensional uncertainty $\xi$
- Generation assets on a continental scale
- Interconnected zones
- Zone balancing of supply and demand (inelastic)

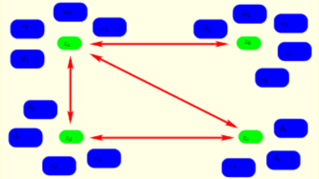


# Important model features



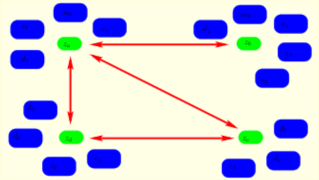
- Very large-scale stochastic problem

# Important model features



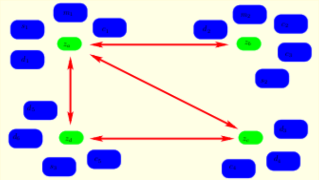
- Very large-scale stochastic problem
- Risk aversion (volume risk & price risk)

# Important model features



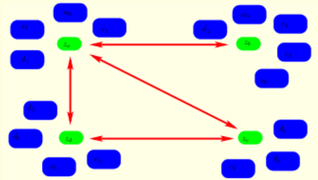
- Very large-scale stochastic problem
- Risk aversion (volume risk & price risk)
- Have good tools to define optimal generation for a **given** system configuration

# Important model features



- Very large-scale stochastic problem
- Risk aversion (volume risk & price risk)
- Have good tools to define optimal generation for a **given** system configuration

# Important model features



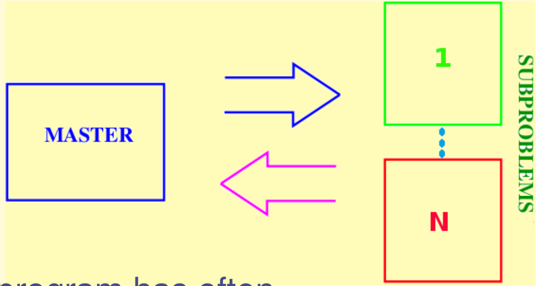
- Very large-scale stochastic problem
- Risk aversion (volume risk & price risk)
- Have good tools to define optimal generation for a **given** system configuration

**Decomposition techniques**

# Decomposition principle

if a problem is difficult to solve directly,

solve instead a sequence of easier subproblems.

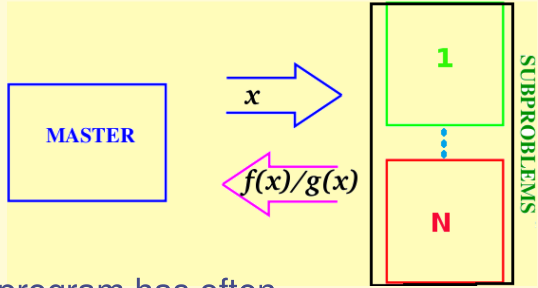



The **master** program has often
a nonsmooth objective function

## Decomposition principle

if a problem is difficult to solve directly,

solve instead a sequence of easier subproblems.




The **master** program has often
a nonsmooth objective function

Separate **subproblems** allow for fast oracle calculations

(the sum of *N* terms)

# Features preventing separability in the GEP

- **Spatial coupling:**
- **Temporal coupling:**
- **Stagewise coupling:**
- **Scenario coupling:**

# Features preventing separability in the GEP

- **Spatial coupling:** **demand constraint**

  all generation assets at each time step

# Features preventing separability in the GEP

- **Spatial coupling:** **demand constraint**

  all generation assets at each time step

- **Temporal coupling:** **ramps, water balance**

  one/few generation assets

# Features preventing separability in the GEP

- **Spatial coupling:** **demand constraint**

  all generation assets at each time step

- **Temporal coupling:** **ramps, water balance**

  one/few generation assets

- **Stagewise coupling:** **capacity constraint**

# Features preventing separability in the GEP

- **Spatial coupling:** **demand constraint**
  all generation assets at each time step
- **Temporal coupling:** **ramps, water balance**
  one/few generation assets
- **Stagewise coupling:** **capacity constraint**
  investment and generation decisions
- **Scenario coupling:** **induced by risk aversion**
  value-at-risk

# Features preventing separability in the GEP

- **Spatial coupling:** **demand constraint**

  all generation assets at each time step: $\sum_p y_p^t(\xi) \geq d^t(\xi)$, for each $t$

- **Temporal coupling:** **ramps, water balance**

  one/few generation assets: $\left(y_p^1(\xi), \ldots, y_p^T(\xi)\right) \in \mathcal{D}_p(\xi)$, for each $p$

- **Stagewise coupling:** **capacity constraint**

  investment and generation decisions: $y_p^t(\xi) \leq cap_p + x$, for each $t, p$

- **Scenario coupling:** **induced by risk aversion**

  value-at-risk: **VaR(gen.cost) couples ALL** $y(\xi)$

# Mathematical Model

(2-stage for simplicity)

$$\left\{ \begin{array}{l} \min \quad c^\top x + CVaR_\alpha(q(\xi)^\top y(\xi)) \\ \\ \\ \\ \end{array} \right.$$

# Mathematical Model

(2-stage for simplicity)

$$\begin{cases} \min \quad c^\top x + CVaR_\alpha(q(\xi)^\top y(\xi)) \\ \text{s.t.} \quad x^t \in \mathcal{S}_{x^t}, \quad y^t(\xi) \in \mathcal{S}_{y^t}(\xi) \quad \text{for all } t\text{, a.s. } \xi \end{cases}$$

# Mathematical Model

(2-stage for simplicity)

$$\begin{cases} \min & c^\top x + CVaR_\alpha(q(\xi)^\top y(\xi)) \\ \text{s.t.} & x^t \in \mathcal{S}_{x^t}, \quad y^t(\xi) \in \mathcal{S}_{y^t}(\xi) \quad \text{for all } t \text{, a.s. } \xi \\ & x_p \in \mathcal{D}_{x_p}, y_p(\xi) \in \mathcal{D}_{y_p}(\xi) \quad \text{for all } p \text{, a.s. } \xi \end{cases}$$

# Mathematical Model

(2-stage for simplicity)

$$\begin{cases} \min & c^\top x + CVaR_\alpha(q(\xi)^\top y(\xi)) & \\ \text{s.t.} & x^t \in \mathcal{S}_{x^t}, \quad y^t(\xi) \in \mathcal{S}_{y^t}(\xi) & \text{for all } t\text{, a.s. } \xi \\ & x_p \in \mathcal{D}_{x_p}, y_p(\xi) \in \mathcal{D}_{y_p}(\xi) & \text{for all } p\text{, a.s. } \xi \\ & \left(x_p^t, y_p^t(\xi)\right) \in \mathcal{C}_p^t(\xi) & \text{for all } t, p\text{, a.s. } \xi \end{cases}$$

# Mathematical Model

(2-stage for simplicity)

$$
\begin{cases}
\min & c^\top x + CVaR_\alpha(q(\xi)^\top y(\xi)) \\
\text{s.t.} & x^t \in \mathcal{S}_{x^t}, \quad y^t(\xi) \in \mathcal{S}_{y^t}(\xi) \quad \text{for all } t, \text{ a.s. } \xi \\
 & x_p \in \mathcal{D}_{x_p}, y_p(\xi) \in \mathcal{D}_{y_p}(\xi) \quad \text{for all } p, \text{ a.s. } \xi \\
 & \left(x_p^t, y_p^t(\xi)\right) \in \mathcal{C}_p^t(\xi) \quad \text{for all } t, p, \text{ a.s. } \xi
\end{cases}
$$

$$
\begin{aligned}
\text{with } CVaR_\alpha(X) &= \min\left\{u + \tfrac{1}{1-\alpha}\mathbb{E}[\max(0, X-u)] : u \in \mathbb{R}\right\} \\
&= \min\left\{u + \tfrac{1}{1-\alpha}\mathbb{E}[v(X))]] : u \in \mathbb{R}, v(X) = \max(0, X-u)\right\}
\end{aligned}
$$

# Mathematical Model

(2-stage for simplicity)

$$
\begin{cases}
\min & c^\top x + CVaR_\alpha(q(\xi)^\top y(\xi)) & \\
\text{s.t.} & x^t \in \mathcal{S}_{x^t}, \quad y^t(\xi) \in \mathcal{S}_{y^t}(\xi) & \text{for all } t, \text{ a.s. } \xi \\
& x_p \in \mathcal{D}_{x_p}, y_p(\xi) \in \mathcal{D}_{y_p}(\xi) & \text{for all } p, \text{ a.s. } \xi \\
& \left(x_p^t, y_p^t(\xi)\right) \in \mathcal{C}_p^t(\xi) & \text{for all } t, p, \text{ a.s. } \xi
\end{cases}
$$

$$
\begin{aligned}
\text{with } CVaR_\alpha(X) &= \min\left\{u + \tfrac{1}{1-\alpha}\mathbb{E}[\max(0, X-u)] : u \in \mathbb{R}\right\} \\
&= \min\left\{u + \tfrac{1}{1-\alpha}\mathbb{E}[v(X))]] : u \in \mathbb{R}, v(X) = \max(0, X-u)\right\}
\end{aligned}
$$

# Mathematical Model

(2-stage for simplicity)

$$
\begin{cases}
\min & c^\top x + u + \frac{1}{1-\alpha}\mathbb{E}[v(\xi)] & \\
\text{s.t.} & x^t \in \mathcal{S}_{x^t}, \quad y^t(\xi) \in \mathcal{S}_{y^t}(\xi) & \text{for all } t \text{, a.s. } \xi \\
& x_p \in \mathcal{D}_{x_p}, y_p(\xi) \in \mathcal{D}_{y_p}(\xi) & \text{for all } p \text{, a.s. } \xi \\
& \left(x_p^t, y_p^t(\xi)\right) \in \mathcal{C}_p^t(\xi) & \text{for all } t, p \text{, a.s. } \xi \\
& u \in \mathbb{R}, v(\xi) \geq 0 & \\
& v(\xi) \geq q(\xi)^\top y(\xi) - u & \text{for all } \xi
\end{cases}
$$

# Mathematical Model

$$
\left\{
\begin{array}{ll}
\min & c^\top x + u + \frac{1}{1-\alpha}\mathbb{E}[v(\xi)] \\
\text{s.t.} & x^t \in \mathcal{S}_{x^t}, \quad y^t(\xi) \in \mathcal{S}_{y^t}( \\
& x_p \in \mathcal{D}_{x_p}, y_p(\xi) \in \mathcal{D}_{y_p}( \\
& \left(x_p^t, y_p^t(\xi)\right) \in \mathcal{C}_p^t(\xi) \\
& u \in \mathbb{R}, v(\xi) \geq 0 \\
& v(\xi) \geq q(\xi)^\top y(\xi) - u
\end{array}
\right.
$$

# Uncoupling stages 

## Progressive Hedging Algorithm

## Uncoupling stages 

### Progressive Hedging Algorithm

- Reformulates an expected-value objective function,

$$\mathbb{E}[c^\top x + q(\xi)^\top y(\xi)] = \sum_{\xi} p(\xi)\Big(c^\top x + q(\xi)^\top y(\xi)\Big)$$

# Uncoupling stages 

## Progressive Hedging Algorithm

- Reformulates an expected-value objective function,

$$\mathbb{E}[c^\top x + q(\xi)^\top y(\xi)] = \sum_{\xi} p(\xi)\Big(c^\top x + q(\xi)^\top y(\xi)\Big)$$

- by creating **copies** of the 1st-stage variable $x$

$$\mathbb{E}[c^\top x + q(\xi)^\top y(\xi)] = \sum_{\xi} p(\xi)\Big(c^\top \hat{x}(\xi) + q(\xi)^\top y(\xi)\Big).$$

# Uncoupling stages 

## Progressive Hedging Algorithm

- Reformulates an expected-value objective function,

$$\mathbb{E}[c^\top x + q(\xi)^\top y(\xi)] = \sum_\xi p(\xi)\Big(c^\top x + q(\xi)^\top y(\xi)\Big)$$

- by creating **copies** of the 1st-stage variable $x$

$$\mathbb{E}[c^\top x + q(\xi)^\top y(\xi)] = \sum_\xi p(\xi)\Big(c^\top \hat{x}(\xi) + q(\xi)^\top y(\xi)\Big).$$

- The **nonanticipativity** constraint

$$\hat{x}(\xi) = x \quad \text{a.e.}\,\xi,$$

is relaxed à la Augmented Lagrangian

# Uncoupling stages 

## Risk-Neutral Progressive Hedging Algorithm (R&W 1987)

- Reformulates an expected-value objective function,

$$\mathbb{E}[c^\top x + q(\xi)^\top y(\xi)] = \sum_{\xi} p(\xi)\Big(c^\top x + q(\xi)^\top y(\xi)\Big)$$

- by creating copies of the 1st-stage variable $x$

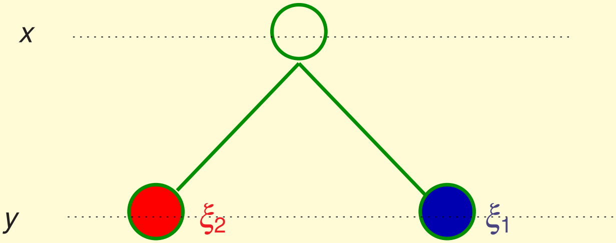
$$\mathbb{E}[c^\top x + q(\xi)^\top y(\xi)] = \sum_{\xi} p(\xi)\Big(c^\top \hat{x}(\xi) + q(\xi)^\top y(\xi)\Big).$$

- The nonanticipativity constraint,

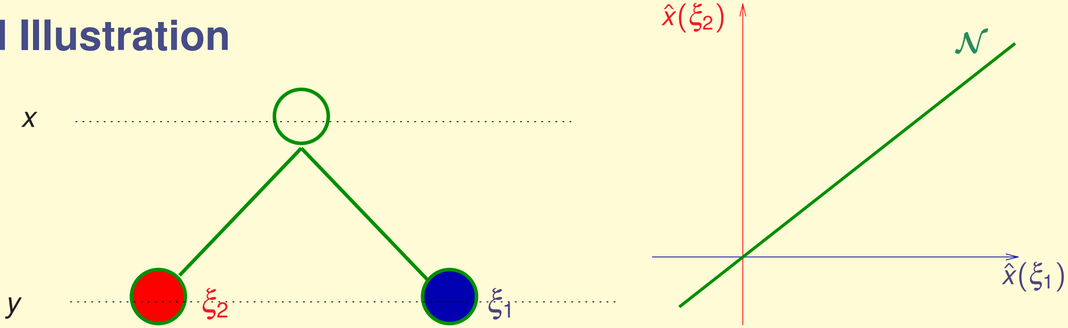
$$\hat{x}(\xi) = x \quad \text{a.e.}\,\xi,$$

  is relaxed à la Augmented Lagrangian

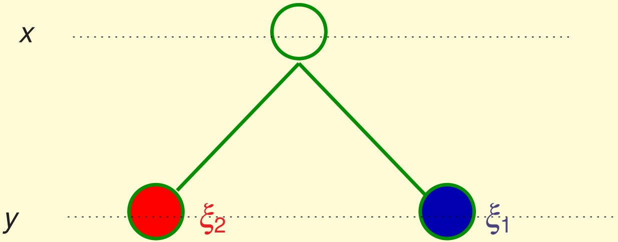
# PH Illustration

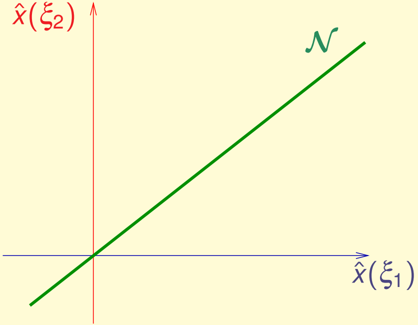


# PH Illustration



# PH Illustration

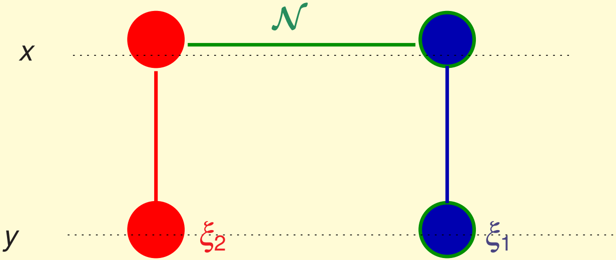


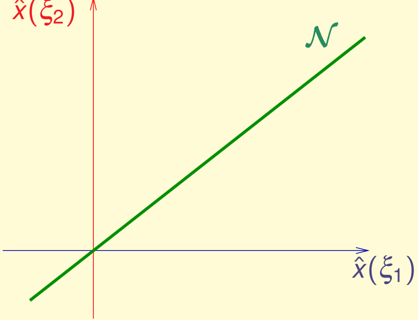




- Nonanticipativity constraint is a projection $x(\xi) = Proj_{\mathcal{N}}(\hat{x}(\xi))$
- On the linear subspace $\mathcal{N} := \{(x_1, x_2) : x_1 = x_2\}$ projection is **explicit**

$$Proj_{\mathcal{N}}(\hat{x}) = \frac{\hat{x}(\xi_1) + \hat{x}(\xi_2)}{2}$$

# PH Illustration

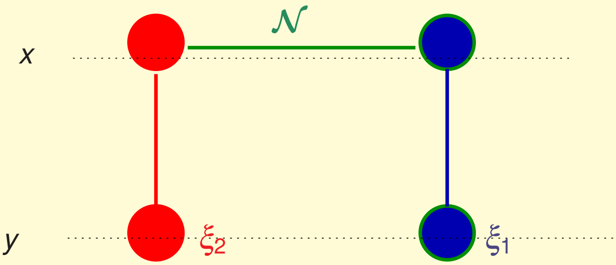


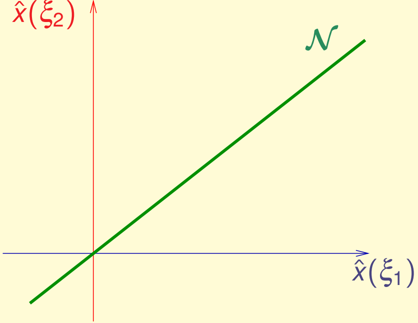




- Nonanticipativity constraint is a projection $x(\xi) = Proj_{\mathcal{N}}(\hat{x}(\xi))$
- On the linear subspace $\mathcal{N} := \{(x_1, x_2) : x_1 = x_2\}$ projection is **explicit**

$$Proj_{\mathcal{N}}(\hat{x}) = \frac{\hat{x}(\xi_1)+\hat{x}(\xi_2)}{2}$$

# PH Illustration



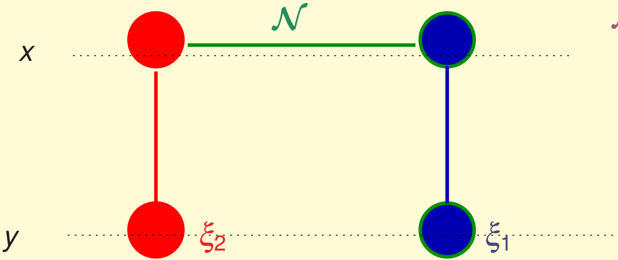



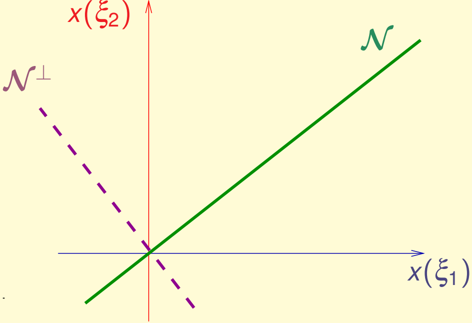

- Nonanticipativity constraint is a projection $x(\xi) = Proj_{\mathcal{N}}(\hat{x}(\xi))$
- On the linear subspace $\mathcal{N} := \{(x_1, x_2) : x_1 = x_2\}$ projection is **explicit**

$$Proj_{\mathcal{N}}(\hat{x}) = \frac{\hat{x}(\xi_1) + \hat{x}(\xi_2)}{2}$$

- Similarly on $\mathcal{N}^{\perp}$ $Proj_{\mathcal{N}^{\perp}}(\hat{x})$

# PH Illustration



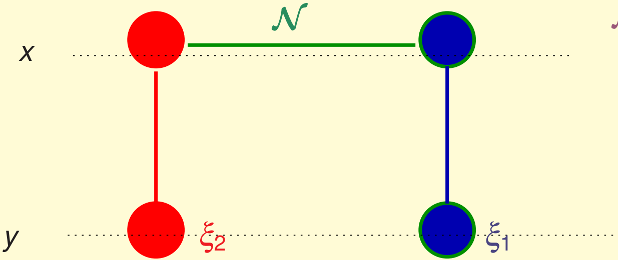



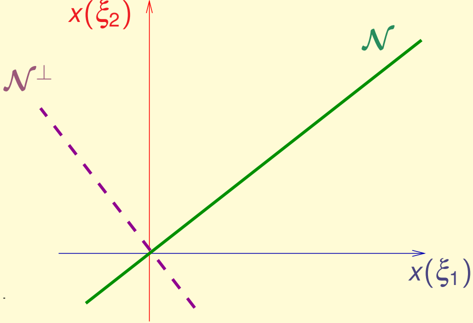

- Nonanticipativity constraint is a projection $x(\xi) = Proj_{\mathcal{N}}(\hat{x})$
- On the linear subspace $\mathcal{N} := \{(x_1, x_2) : x_1 = x_2\}$ projection is **explicit**

$$Proj_{\mathcal{N}}(\hat{x}) = \frac{\hat{x}(\xi_1)+\hat{x}(\xi_2)}{2}$$

- Similarly on $\mathcal{N}^\perp$ $Proj_{\mathcal{N}^\perp}(\hat{x}) = \frac{\hat{x}(\xi_1)-\hat{x}(\xi_2)}{2}$

# PH Illustration




- Nonanticipativity constraint is a projection $x(\xi) = Proj_{\mathcal{N}}(\hat{x})$
- On the linear subspace $\mathcal{N} := \{(x_1, x_2) : x_1 = x_2\}$ projection is **explicit**

$$Proj_{\mathcal{N}}(\hat{x}) = \frac{\hat{x}(\xi_1)+\hat{x}(\xi_2)}{2}$$

- Similarly on $\mathcal{N}^{\perp}$ $Proj_{\mathcal{N}^{\perp}}(\hat{x}) = \frac{\hat{x}(\xi_1)-\hat{x}(\xi_2)}{2}$

# PH algorithm with 2 stages

Finds saddle point $(x^*, w^*)$ of Lagrangian

$$L(x(\cdot), w(\cdot)) = \mathbb{E}\left[f(x(\cdot), \cdot)\right] + \langle x(\cdot), w(\cdot)\rangle$$

that is **separable** along scenarios $\xi_i$

# PH algorithm with 2 stages

Finds saddle point $(x^*, w^*)$ of Lagrangian

$$L(x(\cdot), w(\cdot)) = \mathbb{E}\left[f(x(\cdot), \cdot)\right] + \langle x(\cdot), w(\cdot)\rangle$$

that is **separable** along scenarios $\xi_i$

At iteration $k$, given $\left(x^k(\xi_i), w^k(\xi_i)\right)$ for all $\xi_i$,

- For each $\xi_i$,

  $\hat{x}^k(\xi_i)$ minimizes the augmented Lagrangian with fixed $w^k(\xi_i)$
- $x^{k+1}(\cdot) = Proj_{\mathcal{N}}(\hat{x}^k(\cdot))$
- $w^{k+1}(\cdot) = w^k(\cdot) + rProj_{\mathcal{N}^\perp}(\hat{x}^k(\cdot))$

## PH algorithm with 2 stages

Finds **saddle** point $(x^*, w^*)$ of Lagrangian

$$L(x(\cdot), w(\cdot)) = \mathbb{E}\left[f(x(\cdot), \cdot)\right] + \langle x(\cdot), w(\cdot)\rangle$$

that is separable along scenarios $\xi_i$.

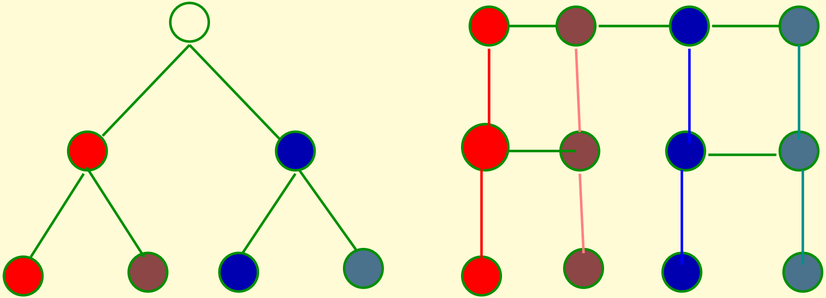
At iteration $k$, given $\left(x^k(\xi_i), w^k(\xi_i)\right)$ for all $\xi_i$,

- For each $\xi_i$,

  $\hat{x}^k(\xi_i)$ minimizes the augmented Lagrangian with fixed $w^k(\xi_i)$
- $x^{k+1}(\cdot) = Proj_{\mathcal{N}}(\hat{x}^k(\cdot))$ $\Leftarrow$ min
- $w^{k+1}(\cdot) = w^k(\cdot) + rProj_{\mathcal{N}^\perp}(\hat{x}^k(\cdot))$ $\Leftarrow$ max

# PH for more than 2 stages

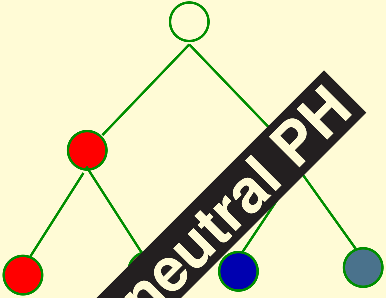
Same reasoning

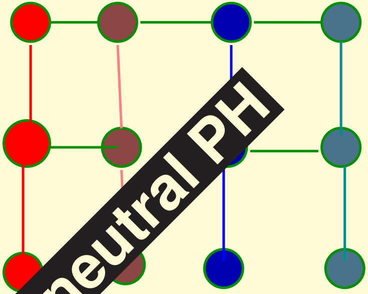


- $x_t = Proj_{\mathcal{N}}(x_t) \iff x_t = \mathbb{E}_{\xi|\xi_{[t]}}[\hat{x}_t]$
- $x_t = Proj_{\mathcal{N}^{\perp}}(x_t) \iff \mathbb{E}[x_t] = 0$

# PH for more than 2 stages

Same reasoning



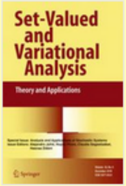


Risk neutral PH

Risk neutral PH

- $\ldots Proj_{\mathcal{N}}(x_t) \iff x_t \ldots {}_{\xi|\xi_{[t]}}[\hat{x}_t]$
- $\ldots x_t = Proj_{\mathcal{N}^{\perp}}(x_t) \iff \ldots [x_t] = 0$

# Uncoupling stages  : +30 years later for CVaR!




Set-Valued and Variational Analysis

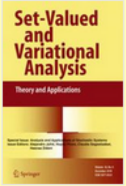
December 2018, Volume 26, Issue 4, pp 759–768 | Cite as

## Solving Stochastic Programming Problems with Risk Measures by Progressive Hedging

Authors | Authors and affiliations

R. Tyrrell Rockafellar

# Uncoupling stages  : +30 years later for CVaR!




Set-Valued and Variational Analysis
December 2018, Volume 26, Issue 4, pp 759–768 | Cite as

## Solving Stochastic Programming Problems with Risk Measures by Progressive Hedging

Authors Authors and affiliations

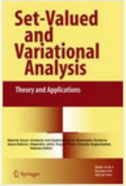
R. Tyrrell Rockafellar

- Creates copies $u(\xi)$ for the VaR variable, $u$, sets

$$u(\xi) = u \quad \text{a.e.}\xi,$$

- and relaxes the nonanticipativity constraint à la Augmented Lagrangian

# Uncoupling stages  : +30 years later for CVaR!




Set-Valued and Variational Analysis
December 2018, Volume 26, Issue 4, pp 759–768 | Cite as

## Solving Stochastic Programming Problems with Risk Measures by Progressive Hedging

Authors Authors and affiliations

R. Tyrrell Rockafellar

Convergence properties: the primal-dual error sequence

$$\left\{ \left\| \begin{pmatrix} x^k(\cdot) \\ u^k(\cdot) \end{pmatrix} - \begin{pmatrix} x^*(\cdot) \\ u^*(\cdot) \end{pmatrix} \right\|^2 + \frac{1}{r^2} \left\| \begin{pmatrix} w^k(\cdot) \\ w_u^k(\cdot) \end{pmatrix} - \begin{pmatrix} w^*(\cdot) \\ w_u^*(\cdot) \end{pmatrix} \right\|^2 \right\}$$

is decreasing

# Mathematical Model

(2-stage for simplicity)

$$\left\{\begin{array}{lll} \min & c^\top x + u + \frac{1}{1-\alpha}\mathbb{E}\left[v(\xi)\right] & \\ \text{s.t.} & x^t \in \mathcal{S}_{x^t}, \quad y^t(\xi) \in \mathcal{S}_{y^t}(\xi) & \text{for all } t, \text{ a.s. } \xi \\ & x_p \in \mathcal{D}_{x_p}, y_p(\xi) \in \mathcal{D}_{y_p}(\xi) & \text{for all } p, \text{ a.s. } \xi \\ & \left(x_p^t, y_p^t(\xi)\right) \in \mathcal{C}_p^t(\xi) & \text{for all } t, p, \text{ a.s. } \xi \\ & u \in \mathbb{R}, v(\xi) \geq 0 & \\ & v(\xi) \geq q(\xi)^\top y(\xi) - u & \text{for all } \xi \end{array}\right.$$

# Mathematical Model

$$
\begin{cases}
\min & c^\top x + u + \frac{1}{1-\alpha}\mathbb{E}\left[v(\xi)\right] \\
\text{s.t.} & x^t \in \mathcal{S}_{x^t}\,, \quad y^t(\xi) \in \mathcal{S}_{y^t}( \\
& x_p \in \mathcal{D}_{x_p}\,, y_p(\xi) \in \mathcal{D}_{y_p}( \\
& \left(x_p^t, y_p^t(\xi)\right) \in \mathcal{C}_p^t(\xi) \\
& u \in \mathbb{R}\,, v(\xi) \geq 0 \\
& v(\xi) \geq q(\xi)^\top y(\xi) - u
\end{cases}
$$

# Mathematical Model

$$
\begin{cases}
\min & c^\top x + u + \frac{1}{1-\alpha}\mathbb{E}[v(\xi)] \\
\text{s.t.} & x^t \in \mathcal{S}_{x^t}, \quad y^t(\xi) \in \mathcal{S}_{y^t}( \\
& x_p \in \mathcal{D}_{x_p}, y_p(\xi) \in \mathcal{D}_{y_p}( \\
& \left(x_p^t, y_p^t(\xi)\right) \in \mathcal{C}_p^t(\xi) \\
& u \in \mathbb{R}, v(\xi) \geq 0 \\
& v(\xi) \geq q(\xi)^\top y(\xi) - u
\end{cases}
$$

CVaR coupling resolved by new PH variant

# Features preventing separability in the GEP

- **Spatial coupling:** **demand constraint**

  all generation assets at each time step: $\sum_p y_p^t(\xi) \geq d^t(\xi)$, for each $t$

- **Temporal coupling:** **ramps, water balance**

  one/few generation assets: $\left(y_p^1(\xi), \ldots, y_p^T(\xi)\right) \in \mathcal{D}_p(\xi)$, for each $p$

- **Stagewise coupling:** **capacity constraint**

  investment and generation decisions: $y_p^t(\xi) \leq cap_p + x$, for each $t, p$

- **Scenario coupling:** **induced by risk aversion**

  value-at-risk: VaR(gen.cost) couples ALL $y(\xi)$

# Features preventing separability in the GEP

- **Spatial coupling:** **demand constraint**

  all generation assets at each time step: $\sum_p y_p^t(\xi) \geq d^t(\xi)$, for each $t$

- **Temporal coupling:** **ramps, water balance**

  one/few generation assets: $\left(y_p^1(\xi), \ldots, y_p^T(\xi)\right) \in \mathcal{D}_p(\xi)$, for each $p$

- **Stagewise coupling:** **capacity constraint**

  investment and generation decisions: $y_p^t(\xi) \leq cap_p + x$, for each $t, p$

- **Scenario coupling:** **induced by risk aversion**

  value-at-risk: **VaR(gen.cost) couples ALL** $y(\xi)$

# Features preventing separability in the GEP

- **Spatial coupling?** **demand constraint**

  all generation assets at each time step: $\sum_p y_p^t(\xi) \geq d^t(\xi)$, for each $t$

- **Temporal coupling:** **ramps, water balance**

  one/few generation assets: $\left(y_p^1(\xi), \ldots, y_p^T(\xi)\right) \in \mathcal{D}_p(\xi)$, for each $p$

- **Stagewise coupling:** **capacity constraint**

  investment and generation decisions: $y_p^t(\xi) \leq cap_p + x$, for each $t, p$

- **Scenario coupling:** **induced by risk aversion**

  value-at-risk: VaR(gen.cost) couples ALL $y(\xi)$

## Spatial coupling: variable splitting

To induce separability along power plants,

- some variables are duplicated
- the corresponding constraints are relaxed
  (subproblems are separable for power plant)
- a proximal bundle method updates the multipliers

## Spatial coupling: variable splitting

To induce separability along power plants,

- some variables are duplicated
- the corresponding constraints are relaxed (subproblems are separable for power plant)
- a proximal bundle method updates the multipliers

The bundle method requires minimizing subproblems below certain level, not until optimality (faster!)

**inexact proximal bundle method with noise attenuation**

# Handling inexact problem solution in PH

PH finds certain **saddle point**.

- $x^{k+1}(\cdot) = Proj_{\mathcal{N}}(\hat{x}^k(\cdot))$ $\Leftarrow$ min
- $w^{k+1}(\cdot) = w^k(\cdot) + rProj_{\mathcal{N}^\perp}(\hat{x}^k(\cdot))$ $\Leftarrow$ max

## Handling inexact problem solution in PH

PH finds certain **saddle point**.

- $x^{k+1}(\cdot) = Proj_{\mathcal{N}}(\hat{x}^k(\cdot))$ $\Leftarrow$ min
- $w^{k+1}(\cdot) = w^k(\cdot) + r Proj_{\mathcal{N}^\perp}(\hat{x}^k(\cdot))$ $\Leftarrow$ max

After decomposing, the PH Lagrangian decomposes in three terms

1. separable by plants, with dynamic constraints
2. with static constraints
3. with the nonanticipativity terms induced by risk-averse PH

## Handling inexact problem solution in PH

PH finds certain **saddle point**.

- $x^{k+1}(\cdot) = Proj_{\mathcal{N}}(\hat{x}^k(\cdot))$ $\Leftarrow$ min
- $w^{k+1}(\cdot) = w^k(\cdot) + rProj_{\mathcal{N}^\perp}(\hat{x}^k(\cdot))$ $\Leftarrow$ max

After decomposing, the PH Lagrangian decomposes in three terms

1. separable by plants, with dynamic constraints
2. with static constraints
3. with the nonanticipativity terms induced by risk-averse PH

Where is the impact of inexact subproblem solution?

## Handling inexact problem solution in PH

PH finds certain **saddle point**.

- $x^{k+1}(\cdot) = Proj_{\mathcal{N}}(\hat{\mathbf{x}}^{\mathbf{k}}(\cdot))$ $\Leftarrow$ min
- $w^{k+1}(\cdot) = w^k(\cdot) + rProj_{\mathcal{N}^\perp}(\hat{x}^k(\cdot))$ $\Leftarrow$ max

After decomposing, the PH Lagrangian decomposes in three terms

1. separable by plants, with dynamic constraints
2. with static constraints
3. with the nonanticipativity terms induced by risk-averse PH

Where is the impact of inexact subproblem solution?

# Handling inexact problem solution in PH

PH finds certain **saddle point**.

- $x^{k+1}(\cdot) = Proj_{\mathcal{N}}( \hat{\mathbf{x}}^{\mathbf{k}}(\cdot) )$ $\Leftarrow$ min
- $w^{k+1}(\cdot) = w^k(\cdot) + rProj_{\mathcal{N}^\perp}(\hat{x}^k(\cdot))$ $\Leftarrow$ max

After decomposing, the PH Lagrangian decomposes in three terms

1. separable by plants, with dynamic constraints
2. with static constraints
3. with the nonanticipativity terms induced by risk-averse PH

Where is the impact of inexact subproblem solution?

The **dual** update must adapt (fundamental!)

# Handling inexact problem solution in PH

PH finds certain **saddle point**.

- $x^{k+1}(\cdot) = Proj_{\mathcal{N}}(\, \hat{\mathbf{x}}^{\mathbf{k}}(\cdot)\,)$ $\Leftarrow$ min
- $\mathbf{w}^{\mathbf{k+1}}(\cdot) = w^k(\cdot) + rProj_{\mathcal{N}^\perp}(\hat{x}^k(\cdot))$ $\Leftarrow$ max

After decomposing, the PH Lagrangian decomposes in three terms

1. separable by plants, with dynamic constraints
2. with static constraints
3. with the nonanticipativity terms induced by risk-averse PH

Where is the impact of inexact subproblem solution?

The **dual** update must adapt (fundamental!)

# Handling inexact problem solution

PH finds certain **saddle point**.

- $x^{k+1}(\cdot) = Proj_{\mathcal{N}}( \hat{\mathbf{x}}^{\mathbf{k}}(\cdot) )$ $\Leftarrow$ min
- $w^{k+1}(\cdot) = w^k(\cdot) + rProj_{\mathcal{N}^\perp}(\hat{x}^k(\cdot))$ $\Leftarrow$ max

After decomposing, the PH Lagrangian decomposes in three terms

1. separable by plants, with dynamic constraints
2. with static constraints
3. with the nonanticipativity terms induced by risk-averse PH

Where is the impact of inexact subproblem solution?

The **dual** update must adapt (fundamental!) $\Longrightarrow$ **NSO**

# Dealing with inexact problem solution

- The dual update is just a gradient-like step to maximize the augmented dual function

# Dealing with inexact problem solution

- The dual update is just a gradient-like step to maximize the augmented dual function
- Augmented dual function is smooth, but only inexact information is available

# Dealing with inexact problem solution

- The dual update is just a gradient-like step to maximize the augmented dual function
- Augmented dual function is smooth, but only inexact information is available
- NSO guides the handling of such inexactness, by keeping memory of the iterative process (“bundle”)

# Dealing with inexact problem solution

- The dual update is just a gradient-like step to maximize the augmented dual function
- Augmented dual function is smooth, but only inexact information is available
- NSO guides the handling of such inexactness, by keeping memory of the iterative process (“bundle”)
- The approach is versatile and flexible: *inexactness* can be on-demand, solving more accurately some subproblems
- Theory gives optimality certificates

# Optimality certificates

In both primal and dual variables, up to the accuracy in the subproblem solution

# Optimality certificates

In both primal and dual variables, up to the accuracy in the subproblem solution

- Convergence theory combines
  - R. T. Rockafellar. Solving stochastic programming problems with risk measures by progressive hedging. Set-Valued Var. Anal, 4:759-768, 2018
  - W. de Oliveira W., C. Sagastizábal, C. Lemaréchal. Convex proximal bundle methods in depth: a unified analysis for inexact oracles. Math. Program., 148:241-277, 2014

# Optimality certificates

In both primal and dual variables, up to the accuracy in the subproblem solution

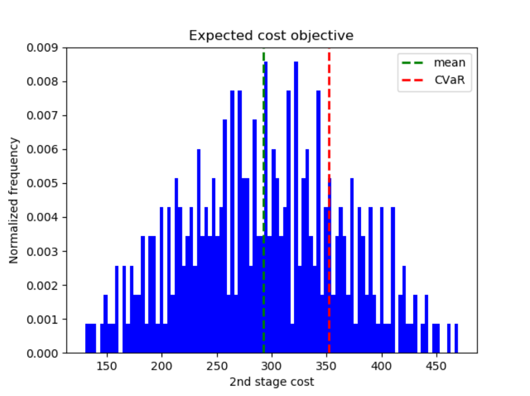
- Convergence theory combines
  - R. T. Rockafellar. Solving stochastic programming problems with risk measures by progressive hedging. Set-Valued Var. Anal, 4:759-768, 2018
  - W. de Oliveira W., C. Sagastizábal, C. Lemaréchal. Convex proximal bundle methods in depth: a unified analysis for inexact oracles. Math. Program., 148:241-277, 2014
- Algorithm solves quadratic and linear programming problems at each iteration

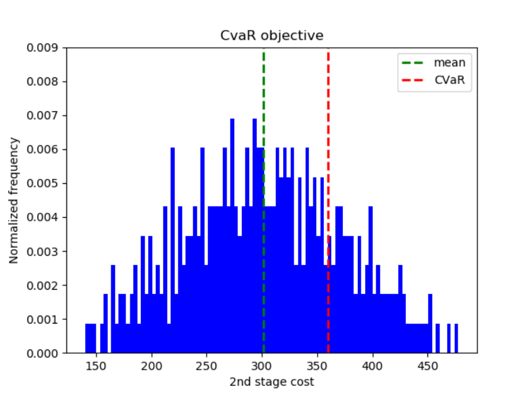
# Numerical assessment on a 2-stage toy problem

- Consistent output, comparged to solving without decomposing
- Risk-neutral (left) and risk-averse (right), no decomposition

# Numerical assessment on a 2-stage toy problem

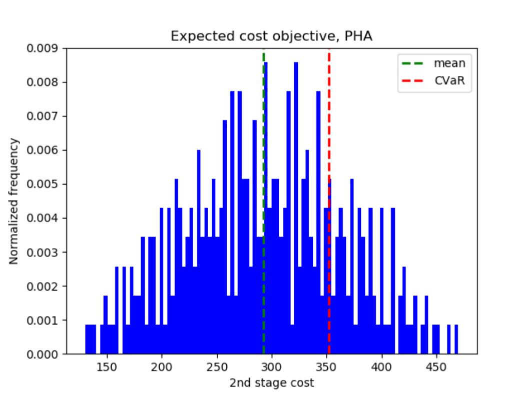
- Consistent output, comparged to solving without decomposing
- Risk-neutral (left) and risk-averse (right), no decomposition

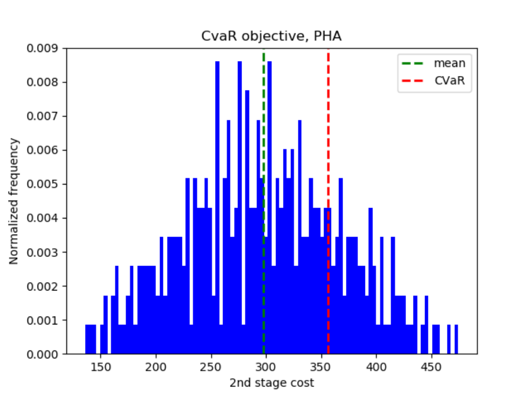


# Numerical assessment on a 2-stage toy problem

- Consistent output, comparged to solving without decomposing
- Risk-neutral (left) and risk-averse (right), our proposal, PHA



# Final remarks

**This is on-going work!** Points we need to explore

- Can we device a stopping test for PHA, à la bundle method

$$\left\{ \left\| \begin{pmatrix} x^k(\cdot) \\ u^k(\cdot) \end{pmatrix} - \begin{pmatrix} \mathbf{x}^*(\cdot) \\ \mathbf{u}^*(\cdot) \end{pmatrix} \right\|^2 + \frac{1}{r^2} \left\| \begin{pmatrix} w^k(\cdot) \\ w_u^k(\cdot) \end{pmatrix} - \begin{pmatrix} \mathbf{w}^*(\cdot) \\ \mathbf{w}_{\mathbf{u}}^*(\cdot) \end{pmatrix} \right\|^2 \right\}$$

# Final remarks

**This is on-going work!** Points we need to explore

- Can we device a stopping test for PHA, à la bundle method

$$\left\{ \left\| \begin{pmatrix} x^k(\cdot) \\ u^k(\cdot) \end{pmatrix} - \begin{pmatrix} \mathbf{x}^*(\cdot) \\ \mathbf{u}^*(\cdot) \end{pmatrix} \right\|^2 + \frac{1}{r^2} \left\| \begin{pmatrix} w^k(\cdot) \\ w_u^k(\cdot) \end{pmatrix} - \begin{pmatrix} \mathbf{w}^*(\cdot) \\ \mathbf{w}_\mathbf{u}^*(\cdot) \end{pmatrix} \right\|^2 \right\}$$

- Can the PH parameter $r$ be updated without impairing convergence?

# Final remarks

**This is on-going work!** Points we need to explore

- Can we device a stopping test for PHA, à la bundle method

$$\left\{ \left\| \begin{pmatrix} x^k(\cdot) \\ u^k(\cdot) \end{pmatrix} - \begin{pmatrix} \mathbf{x}^*(\cdot) \\ \mathbf{u}^*(\cdot) \end{pmatrix} \right\|^2 + \frac{1}{r^2} \left\| \begin{pmatrix} w^k(\cdot) \\ w_u^k(\cdot) \end{pmatrix} - \begin{pmatrix} \mathbf{w}^*(\cdot) \\ \mathbf{w}_\mathbf{u}^*(\cdot) \end{pmatrix} \right\|^2 \right\}$$

- Can the PH parameter $r$ be updated without impairing convergence?
- Run on EdF's instances: PGMO project joint with Wim van Ackooij (EdF labs) and Welington de Oliveira (Mines Paris Tech)

# Final remarks

**This is on-going work!** Points we need to explore

- Can we device a stopping test for PHA, à la bundle method

$$\left\{ \left\| \begin{pmatrix} x^k(\cdot) \\ u^k(\cdot) \end{pmatrix} - \begin{pmatrix} \mathbf{x}^*(\cdot) \\ \mathbf{u}^*(\cdot) \end{pmatrix} \right\|^2 + \frac{1}{r^2} \left\| \begin{pmatrix} w^k(\cdot) \\ w_u^k(\cdot) \end{pmatrix} - \begin{pmatrix} \mathbf{w}^*(\cdot) \\ \mathbf{w}_\mathbf{u}^*(\cdot) \end{pmatrix} \right\|^2 \right\}$$

- Can the PH parameter $r$ be updated without impairing convergence?
- Run on EdF's instances: PGMO project joint with Wim van Ackooij (EdF labs) and Welington de Oliveira (Mines Paris Tech)
- Final comment: PH with inexact subproblem solution applies in multistage setting, Benders not so much