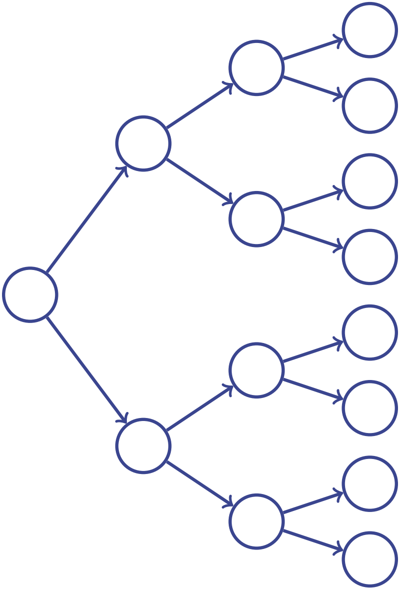
ISMP Bordeaux 2018

A deterministic algorithm for solving stochastic minimax dynamic programmes

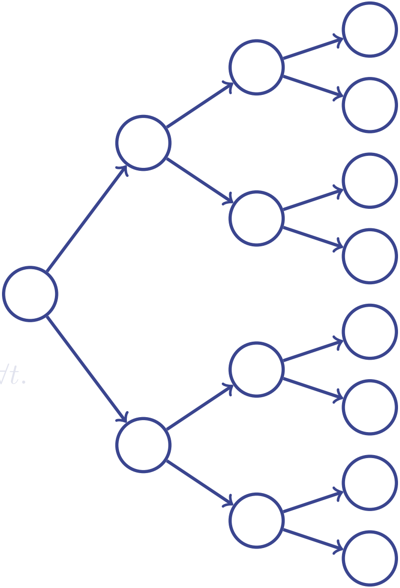
Regan Baucke,
Tony Downward, Golbon Zakeri

EPOC, The University of Auckland

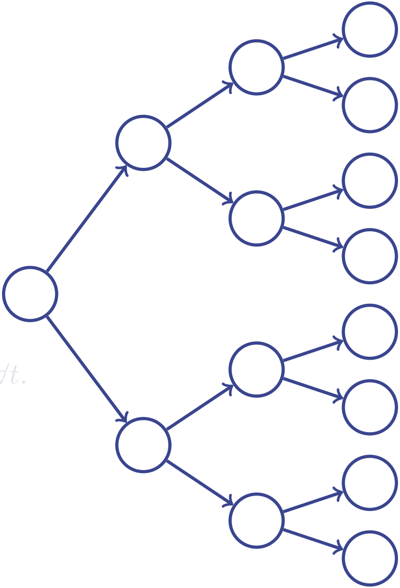
r.baucke@auckland.ac.nz



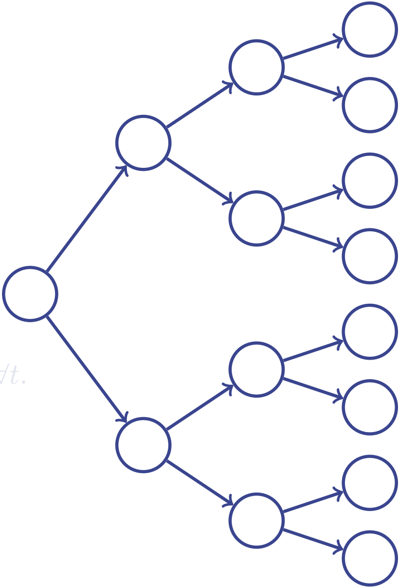
$$\begin{aligned}
\max_{v_t} \quad & \mathbb{E}_{\mathbb{P}}\Big(\sum_{t=0}^{T} g_{\omega(t)}(y_{\omega(t)}, v_{\omega(t)}) \\
& \qquad\qquad + G_{\omega(T)}(y_{\omega(T)}) \Big) \\
\text{s.t.} \quad & v_{\omega(t)} \in \mathcal{V}_{\omega(t)}(y_{\omega(t)}), \ \forall t, \\
& y_{\omega(t)} \in \mathcal{Y}_{\omega(t)}, \ \forall t, \\
& y_{\omega(t+1)} = f^{y}_{\omega(t+1)}(y_{\omega(t)}, v_{\omega(t)}), \ \forall t.
\end{aligned}$$



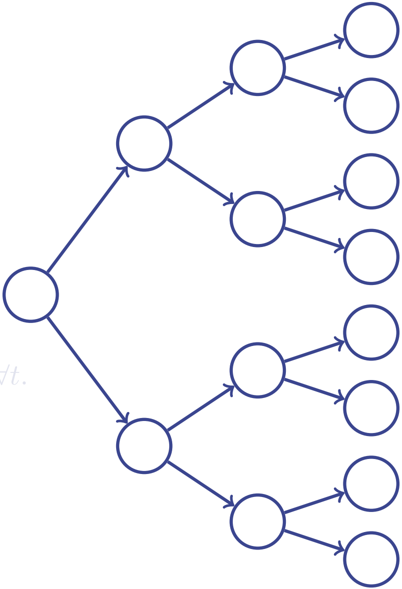
$$
\begin{aligned}
\max_{v_t} \quad & \mathbb{E}_{\mathbb{P}}\Big(\sum_{t=0}^{T} g_{\omega(t)}\big(y_{\omega(t)}, v_{\omega(t)}\big) \\
& \qquad\qquad + G_{\omega(T)}\big(y_{\omega(T)}\big)\Big) \\
\text{s.t.} \quad & v_{\omega(t)} \in \mathcal{V}_{\omega(t)}\big(y_{\omega(t)}\big), \ \forall t, \\
& y_{\omega(t)} \in \mathcal{Y}_{\omega(t)}, \ \forall t, \\
& y_{\omega(t+1)} = f^{y}_{\omega(t+1)}\big(y_{\omega(t)}, v_{\omega(t)}\big), \ \forall t.
\end{aligned}
$$



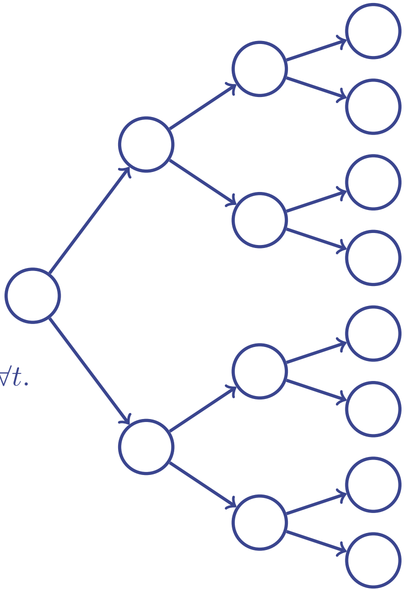
$$
\begin{aligned}
\max_{v_t} \quad & \mathbb{E}_{\mathbb{P}}\Big(\sum_{t=0}^{T} g_{\omega(t)}\big(y_{\omega(t)}, v_{\omega(t)}\big) \\
& \qquad\qquad + G_{\omega(T)}\big(y_{\omega(T)}\big)\Big) \\
\text{s.t.} \quad & v_{\omega(t)} \in \mathcal{V}_{\omega(t)}\big(y_{\omega(t)}\big), \ \forall t, \\
& y_{\omega(t)} \in \mathcal{Y}_{\omega(t)}, \ \forall t, \\
& y_{\omega(t+1)} = f^{y}_{\omega(t+1)}\big(y_{\omega(t)}, v_{\omega(t)}\big), \ \forall t.
\end{aligned}
$$



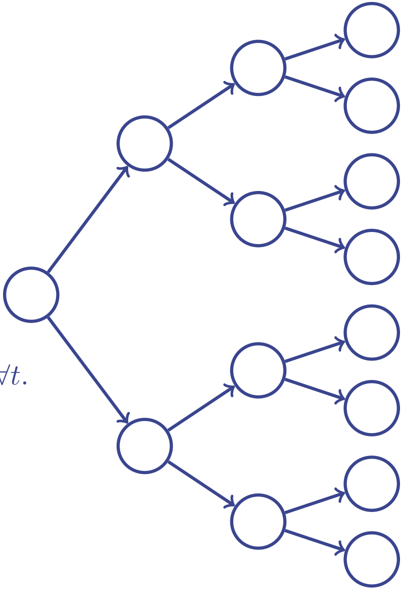
$$
\begin{aligned}
\max_{v_t} \quad & \mathbb{E}_{\mathbb{P}}\Big(\sum_{t=0}^{T} g_{\omega(t)}(y_{\omega(t)}, v_{\omega(t)}) \\
& \qquad\qquad + G_{\omega(T)}(y_{\omega(T)}) \Big) \\
\text{s.t.} \quad & v_{\omega(t)} \in \mathcal{V}_{\omega(t)}(y_{\omega(t)}), \ \forall t, \\
& y_{\omega(t)} \in \mathcal{Y}_{\omega(t)}, \ \forall t, \\
& y_{\omega(t+1)} = f^{y}_{\omega(t+1)}(y_{\omega(t)}, v_{\omega(t)}), \ \forall t.
\end{aligned}
$$



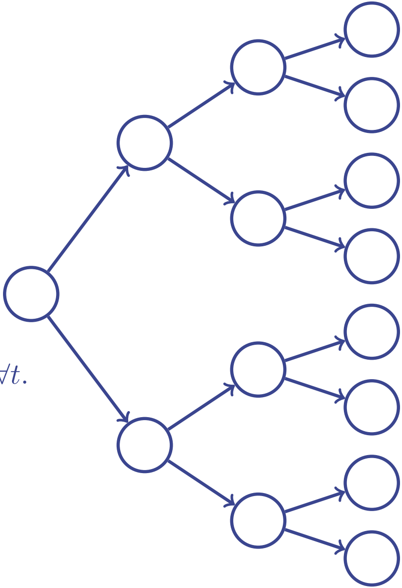
$$
\begin{aligned}
\max_{v_t} \quad & \mathbb{E}_{\mathbb{P}}\Big(\sum_{t=0}^{T} g_{\omega(t)}\big(y_{\omega(t)}, v_{\omega(t)}\big) \\
& \qquad\qquad + G_{\omega(T)}\big(y_{\omega(T)}\big)\Big) \\
\text{s.t.} \quad & v_{\omega(t)} \in \mathcal{V}_{\omega(t)}\big(y_{\omega(t)}\big), \ \forall t, \\
& y_{\omega(t)} \in \mathcal{Y}_{\omega(t)}, \ \forall t, \\
& y_{\omega(t+1)} = f^{y}_{\omega(t+1)}\big(y_{\omega(t)}, v_{\omega(t)}\big), \ \forall t.
\end{aligned}
$$



$$
\begin{aligned}
\max_{v_t} \quad & \rho\Big(\sum_{t=0}^{T} g_{\omega(t)}(y_{\omega(t)}, v_{\omega(t)}) \\
& \qquad\qquad + G_{\omega(T)}(y_{\omega(T)})\Big) \\
\text{s.t.} \quad & v_{\omega(t)} \in \mathcal{V}_{\omega(t)}(y_{\omega(t)}),\ \forall t, \\
& y_{\omega(t)} \in \mathcal{Y}_{\omega(t)},\ \forall t, \\
& y_{\omega(t+1)} = f^{y}_{\omega(t+1)}(y_{\omega(t)}, v_{\omega(t)}),\ \forall t.
\end{aligned}
$$



$$
\begin{aligned}
\max_{v_t} \quad & \min_{\mathbb{P} \in \mathcal{Q}^*} \mathbb{E}_{\mathbb{P}} \Big(\sum_{t=0}^{T} g_{\omega(t)}(y_{\omega(t)}, v_{\omega(t)}) \\
& \qquad\qquad\qquad + G_{\omega(T)}(y_{\omega(T)}) \Big) \\
\text{s.t.} \quad & v_{\omega(t)} \in \mathcal{V}_{\omega(t)}(y_{\omega(t)}), \ \forall t, \\
& y_{\omega(t)} \in \mathcal{Y}_{\omega(t)}, \ \forall t, \\
& y_{\omega(t+1)} = f^{y}_{\omega(t+1)}(y_{\omega(t)}, v_{\omega(t)}), \ \forall t.
\end{aligned}
$$



Decomposition

$$\min_{\mathbb{P}\in\mathcal{Q}^*} \mathbb{E}_{\mathbb{P}}\Big(\sum_{t=0}^{T} Z_t\Big)$$

$$=$$

$$\min_{\mathbb{P}_0\in\mathcal{Q}_0} \mathbb{E}_{\mathbb{P}_0}\left(Z_0 + \min_{\mathbb{P}_1\in\mathcal{Q}_1} \mathbb{E}_{\mathbb{P}_1}\left(Z_1 + \min_{\mathbb{P}_2\in\mathcal{Q}_2} \mathbb{E}_{\mathbb{P}_2}\left(Z_2\right)\right)\right)$$

Decomposition

$$\min_{\mathbb{P}\in\mathcal{Q}^*} \mathbb{E}_{\mathbb{P}}\Big(\sum_{t=0}^{T} Z_t\Big)$$

$$=$$

$$\min_{\mathbb{P}_0\in\mathcal{Q}_0} \mathbb{E}_{\mathbb{P}_0}\left(Z_0 + \min_{\mathbb{P}_1\in\mathcal{Q}_1} \mathbb{E}_{\mathbb{P}_1}\left(Z_1 + \min_{\mathbb{P}_2\in\mathcal{Q}_2} \mathbb{E}_{\mathbb{P}_2}\left(Z_2\right)\right)\right)$$

Decomposition

$$\min_{\mathbb{P}\in\mathcal{Q}^*} \mathbb{E}_{\mathbb{P}}\Big(\sum_{t=0}^{T} Z_t\Big)$$

$$=$$

$$\min_{\mathbb{P}_0\in\mathcal{Q}_0} \mathbb{E}_{\mathbb{P}_0}\left(Z_0 + \min_{\mathbb{P}_1\in\mathcal{Q}_1} \mathbb{E}_{\mathbb{P}_1}\left(Z_1 + \min_{\mathbb{P}_2\in\mathcal{Q}_2} \mathbb{E}_{\mathbb{P}_2}\left(Z_2\right)\right)\right)$$

Shapiro (2012)

Decomposition

$$\min_{\mathbb{P}\in\mathcal{Q}^*} \mathbb{E}_{\mathbb{P}}\Big(\sum_{t=0}^{T} Z_t\Big)$$

$$=$$

$$\min_{\mathbb{P}_0\in\mathcal{Q}_0} \mathbb{E}_{\mathbb{P}_0}\left(Z_0 + \min_{\mathbb{P}_1\in\mathcal{Q}_1} \mathbb{E}_{\mathbb{P}_1}\left(Z_1 + \min_{\mathbb{P}_2\in\mathcal{Q}_2} \mathbb{E}_{\mathbb{P}_2}\left(Z_2\right)\right)\right)$$

Shapiro (2012)

Decomposition

$$\min_{\mathbb{P}\in\mathcal{Q}^*} \mathbb{E}_{\mathbb{P}}\Big(\sum_{t=0}^{T} Z_t\Big)$$

$$=$$

$$\min_{\mathbb{P}_0\in\mathcal{Q}_0} \mathbb{E}_{\mathbb{P}_0}\left(Z_0 + \min_{\mathbb{P}_1\in\mathcal{Q}_1(\mathbb{P}_0)} \mathbb{E}_{\mathbb{P}_1}\left(Z_1 + \min_{\mathbb{P}_2\in\mathcal{Q}_2(\mathbb{P}_1)} \mathbb{E}_{\mathbb{P}_2}\left(Z_2\right)\right)\right)$$

Pflug and Pichler (2016), My Thesis

Decomposition

$$\min_{\mathbb{P}\in\mathcal{Q}^*} \mathbb{E}_{\mathbb{P}}\Big(\sum_{t=0}^{T} Z_t\Big)$$

$$=$$

$$\min_{\mathbb{P}_0\in\mathcal{Q}_0} \mathbb{E}_{\mathbb{P}_0}\left(Z_0 + \min_{\mathbb{P}_1\in\mathcal{Q}_1(\mathbb{P}_0)} \mathbb{E}_{\mathbb{P}_1}\left(Z_1 + \min_{\mathbb{P}_2\in\mathcal{Q}_2(\mathbb{P}_1)} \mathbb{E}_{\mathbb{P}_2}\left(Z_2\right)\right)\right)$$

Pflug and Pichler (2016), My Thesis

Formulations

Bounding functions

Algorithm

Applications

Extensive form

$$\begin{aligned} \max_{v_t} \min_{u_t} \quad & \mathbb{E}_{\mathbb{P}}\left[\sum_{t=0}^{T} g_{\omega(t)}(x_{\omega(t)}, y_{\omega(t)}, u_{\omega(t)}, v_{\omega(t)}) + G_{\omega(T)}(x_{\omega(T)}, y_{\omega(T)}) \right] \\ \text{s.t.} \quad & (u_{\omega(t)}, v_{\omega(t)}) \in \mathcal{U}_{\omega(t)}(x_{\omega(t)}) \times \mathcal{V}_{\omega(t)}(y_{\omega(t)}),\ \forall t, \\ & (x_{\omega(t)}, y_{\omega(t)}) \in \mathcal{X}_{\omega(t)} \times \mathcal{Y}_{\omega(t)},\ \forall t, \\ & x_{\omega(t+1)} = f^x_{\omega(t+1)}(x_{\omega(t)}, u_{\omega(t)}),\ \forall t, \\ & y_{\omega(t+1)} = f^y_{\omega(t+1)}(y_{\omega(t)}, v_{\omega(t)}),\ \forall t. \end{aligned}$$

Extensive form

$$
\begin{aligned}
\max_{v_t} \min_{u_t} \quad & \mathbb{E}_{\mathbb{P}}\Big[\sum_{t=0}^{T} g_{\omega(t)}(x_{\omega(t)}, y_{\omega(t)}, u_{\omega(t)}, v_{\omega(t)}) + G_{\omega(T)}(x_{\omega(T)}, y_{\omega(T)})\Big] \\
\text{s.t.} \quad & (u_{\omega(t)}, v_{\omega(t)}) \in \mathcal{U}_{\omega(t)}(x_{\omega(t)}) \times \mathcal{V}_{\omega(t)}(y_{\omega(t)}),\ \forall t, \\
& (x_{\omega(t)}, y_{\omega(t)}) \in \mathcal{X}_{\omega(t)} \times \mathcal{Y}_{\omega(t)},\ \forall t, \\
& x_{\omega(t+1)} = f^x_{\omega(t+1)}(x_{\omega(t)}, u_{\omega(t)}),\ \forall t, \\
& y_{\omega(t+1)} = f^y_{\omega(t+1)}(y_{\omega(t)}, v_{\omega(t)}),\ \forall t.
\end{aligned}
$$

Extensive form

$$\begin{aligned}
\max_{v_t} \min_{u_t} \quad & \mathbb{E}_{\mathbb{P}}\left[\sum_{t=0}^{T} g_{\omega(t)}(x_{\omega(t)}, y_{\omega(t)}, u_{\omega(t)}, v_{\omega(t)}) + G_{\omega(T)}(x_{\omega(T)}, y_{\omega(T)})\right] \\
\text{s.t.} \quad & (u_{\omega(t)}, v_{\omega(t)}) \in \mathcal{U}_{\omega(t)}(x_{\omega(t)}) \times \mathcal{V}_{\omega(t)}(y_{\omega(t)}),\ \forall t, \\
& (x_{\omega(t)}, y_{\omega(t)}) \in \mathcal{X}_{\omega(t)} \times \mathcal{Y}_{\omega(t)},\ \forall t, \\
& x_{\omega(t+1)} = f^{x}_{\omega(t+1)}(x_{\omega(t)}, u_{\omega(t)}),\ \forall t, \\
& y_{\omega(t+1)} = f^{y}_{\omega(t+1)}(y_{\omega(t)}, v_{\omega(t)}),\ \forall t.
\end{aligned}$$

Extensive form

$$
\begin{aligned}
\max_{v_t} \min_{u_t} \quad & \mathbb{E}_{\mathbb{P}}\left[\sum_{t=0}^{T} g_{\omega(t)}\left(x_{\omega(t)}, y_{\omega(t)}, u_{\omega(t)}, v_{\omega(t)}\right) + G_{\omega(T)}\left(x_{\omega(T)}, y_{\omega(T)}\right) \right] \\
\text{s.t.} \quad & \left(u_{\omega(t)}, v_{\omega(t)}\right) \in \mathcal{U}_{\omega(t)}\left(x_{\omega(t)}\right) \times \mathcal{V}_{\omega(t)}\left(y_{\omega(t)}\right),\ \forall t, \\
& \left(x_{\omega(t)}, y_{\omega(t)}\right) \in \mathcal{X}_{\omega(t)} \times \mathcal{Y}_{\omega(t)},\ \forall t, \\
& x_{\omega(t+1)} = f^{x}_{\omega(t+1)}\left(x_{\omega(t)}, u_{\omega(t)}\right),\ \forall t, \\
& y_{\omega(t+1)} = f^{y}_{\omega(t+1)}\left(y_{\omega(t)}, v_{\omega(t)}\right),\ \forall t.
\end{aligned}
$$

Extensive form

$$\begin{aligned} \max_{v_t} \min_{u_t} \quad & \mathbb{E}_{\mathbb{P}} \left[\sum_{t=0}^{T} g_{\omega(t)}(x_{\omega(t)}, y_{\omega(t)}, u_{\omega(t)}, v_{\omega(t)}) + G_{\omega(T)}(x_{\omega(T)}, y_{\omega(T)}) \right] \\ \text{s.t.} \quad & (u_{\omega(t)}, v_{\omega(t)}) \in \mathcal{U}_{\omega(t)}(x_{\omega(t)}) \times \mathcal{V}_{\omega(t)}(y_{\omega(t)}),\ \forall t, \\ & (x_{\omega(t)}, y_{\omega(t)}) \in \mathcal{X}_{\omega(t)} \times \mathcal{Y}_{\omega(t)},\ \forall t, \\ & x_{\omega(t+1)} = f^x_{\omega(t+1)}(x_{\omega(t)}, u_{\omega(t)}),\ \forall t, \\ & y_{\omega(t+1)} = f^y_{\omega(t+1)}(y_{\omega(t)}, v_{\omega(t)}),\ \forall t. \end{aligned}$$

Extensive form

$$
\begin{aligned}
\max_{v_t}\ \min_{u_t} \quad & \mathbb{E}_{\mathbb{P}}\left[\sum_{t=0}^{T} g_{\omega(t)}\big(x_{\omega(t)}, y_{\omega(t)}, u_{\omega(t)}, v_{\omega(t)}\big) + G_{\omega(T)}\big(x_{\omega(T)}, y_{\omega(T)}\big)\right] \\
\text{s.t.} \quad & \big(u_{\omega(t)}, v_{\omega(t)}\big) \in \mathcal{U}_{\omega(t)}\big(x_{\omega(t)}\big) \times \mathcal{V}_{\omega(t)}\big(y_{\omega(t)}\big),\ \forall t, \\
& \big(x_{\omega(t)}, y_{\omega(t)}\big) \in \mathcal{X}_{\omega(t)} \times \mathcal{Y}_{\omega(t)},\ \forall t, \\
& x_{\omega(t+1)} = f^{x}_{\omega(t+1)}\big(x_{\omega(t)}, u_{\omega(t)}\big),\ \forall t, \\
& y_{\omega(t+1)} = f^{y}_{\omega(t+1)}\big(y_{\omega(t)}, v_{\omega(t)}\big),\ \forall t.
\end{aligned}
$$

Extensive form

$$\max_{v_t} \min_{u_t} \quad \mathbb{E}_{\mathbb{P}}\left[\sum_{t=0}^{T} g_{\omega(t)}(x_{\omega(t)}, y_{\omega(t)}, u_{\omega(t)}, v_{\omega(t)}) + G_{\omega(T)}(x_{\omega(T)}, y_{\omega(T)})\right]$$

$$\text{s.t.} \quad (u_{\omega(t)}, v_{\omega(t)}) \in \mathcal{U}_{\omega(t)}(x_{\omega(t)}) \times \mathcal{V}_{\omega(t)}(y_{\omega(t)}),\ \forall t,$$

$$(x_{\omega(t)}, y_{\omega(t)}) \in \mathcal{X}_{\omega(t)} \times \mathcal{Y}_{\omega(t)},\ \forall t,$$

$$x_{\omega(t+1)} = f^{x}_{\omega(t+1)}(x_{\omega(t)}, u_{\omega(t)}),\ \forall t,$$

$$y_{\omega(t+1)} = f^{y}_{\omega(t+1)}(y_{\omega(t)}, v_{\omega(t)}),\ \forall t.$$

Extensive form

$$\begin{aligned} \max_{v_t} \min_{u_t} \quad & \mathbb{E}_{\mathbb{P}}\left[\sum_{t=0}^{T} g_{\omega(t)}(x_{\omega(t)}, y_{\omega(t)}, u_{\omega(t)}, v_{\omega(t)}) + G_{\omega(T)}(x_{\omega(T)}, y_{\omega(T)})\right] \\ \text{s.t.} \quad & (u_{\omega(t)}, v_{\omega(t)}) \in \mathcal{U}_{\omega(t)}(x_{\omega(t)}) \times \mathcal{V}_{\omega(t)}(y_{\omega(t)}),\ \forall t, \\ & (x_{\omega(t)}, y_{\omega(t)}) \in \mathcal{X}_{\omega(t)} \times \mathcal{Y}_{\omega(t)},\ \forall t, \\ & x_{\omega(t+1)} = f^x_{\omega(t+1)}(x_{\omega(t)}, u_{\omega(t)}),\ \forall t, \\ & y_{\omega(t+1)} = f^y_{\omega(t+1)}(y_{\omega(t)}, v_{\omega(t)}),\ \forall t. \end{aligned}$$

Dynamic form

$$
\begin{aligned}
G_n(x_n, y_n) = \max_{v_n} \min_{u_n} \quad & g_n(x_n, y_n, u_n, v_n) + \mathbb{E}[G_m(x_m, y_m)] \\
\text{s.t.} \quad & (u_n, v_n) \in \mathcal{U}_n(x_n) \times \mathcal{V}_n(y_n), \\
& (x_n, y_n) \in \mathcal{X}_n \times \mathcal{Y}_t, \\
& x_m = f_m^x(x_n, u_n), \ \forall m \in C(n), \\
& y_m = f_m^y(y_n, v_n), \ \forall m \in C(n).
\end{aligned}
$$

Dynamic form

$$
\begin{aligned}
G_n(x_n, y_n) = \max_{v_n} \min_{u_n} \quad & g_n(x_n, y_n, u_n, v_n) + \mathbb{E}\big[G_m(x_m, y_m)\big] \\
\text{s.t.} \quad & (u_n, v_n) \in \mathcal{U}_n(x_n) \times \mathcal{V}_n(y_n), \\
& (x_n, y_n) \in \mathcal{X}_n \times \mathcal{Y}_t, \\
& x_m = f^x_m(x_n, u_n), \ \forall m \in C(n), \\
& y_m = f^y_m(y_n, v_n), \ \forall m \in C(n).
\end{aligned}
$$

Dynamic form

$$
\begin{aligned}
G_n(x_n, y_n) = \max_{v_n} \min_{u_n} \quad & g_n(x_n, y_n, u_n, v_n) + \mathbb{E}\big[G_m(x_m, y_m)\big] \\
\text{s.t.} \quad & (u_n, v_n) \in \mathcal{U}_n(x_n) \times \mathcal{V}_n(y_n), \\
& (x_n, y_n) \in \mathcal{X}_n \times \mathcal{Y}_t, \\
& x_m = f^x_m(x_n, u_n),\ \forall m \in C(n), \\
& y_m = f^y_m(y_n, v_n),\ \forall m \in C(n).
\end{aligned}
$$

Dynamic form

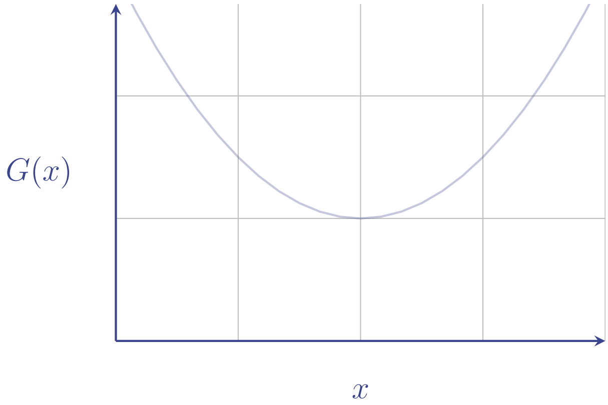
$$
\begin{aligned}
G_n(x_n, y_n) = \max_{v_n} \min_{u_n} \quad & g_n(x_n, y_n, u_n, v_n) + \mathbb{E}\big[G_m(x_m, y_m)\big] \\
\text{s.t.} \quad & (u_n, v_n) \in \mathcal{U}_n(x_n) \times \mathcal{V}_n(y_n), \\
& (x_n, y_n) \in \mathcal{X}_n \times \mathcal{Y}_t, \\
& x_m = f_m^x(x_n, u_n),\ \forall m \in C(n), \\
& y_m = f_m^y(y_n, v_n),\ \forall m \in C(n).
\end{aligned}
$$

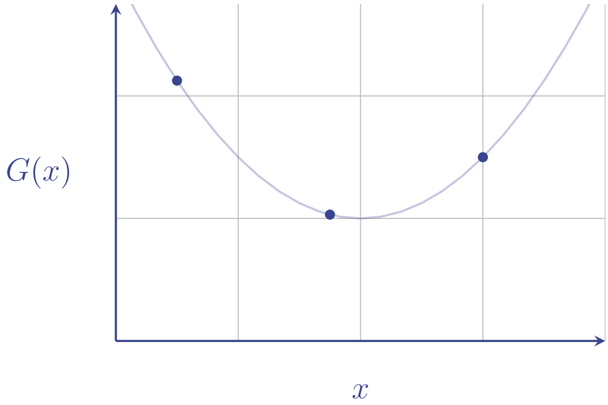
Dynamic form

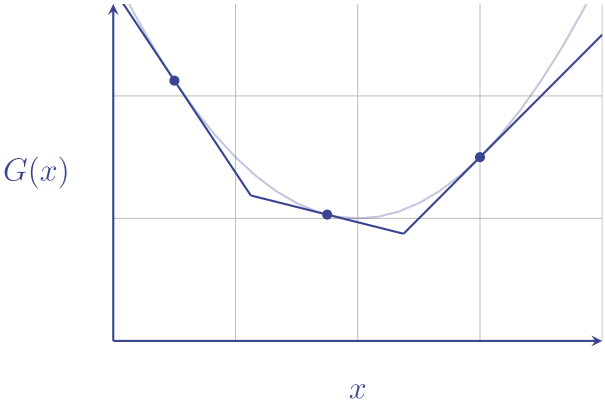
$$
\begin{aligned}
G_n(x_n, y_n) = \max_{v_n} \min_{u_n} \quad & g_n(x_n, y_n, u_n, v_n) + \mathbb{E}\big[G_m(x_m, y_m)\big] \\
\text{s.t.} \quad & (u_n, v_n) \in \mathcal{U}_n(x_n) \times \mathcal{V}_n(y_n), \\
& (x_n, y_n) \in \mathcal{X}_n \times \mathcal{Y}_t, \\
& x_m = f^x_m(x_n, u_n),\ \forall m \in C(n), \\
& y_m = f^y_m(y_n, v_n),\ \forall m \in C(n).
\end{aligned}
$$

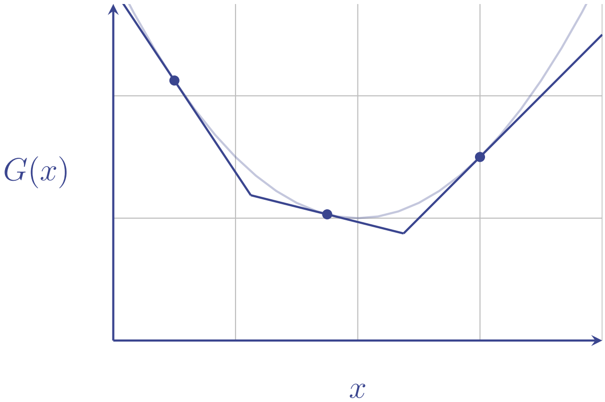
Dynamic form

$$
\begin{aligned}
G_n(x_n, y_n) = \max_{v_n} \min_{u_n} \quad & g_n(x_n, y_n, u_n, v_n) + \mathbb{E}\big[G_m(x_m, y_m)\big] \\
\text{s.t.} \quad & (u_n, v_n) \in \mathcal{U}_n(x_n) \times \mathcal{V}_n(y_n), \\
& (x_n, y_n) \in \mathcal{X}_n \times \mathcal{Y}_t, \\
& x_m = f^x_m(x_n, u_n),\ \forall m \in C(n), \\
& y_m = f^y_m(y_n, v_n),\ \forall m \in C(n).
\end{aligned}
$$

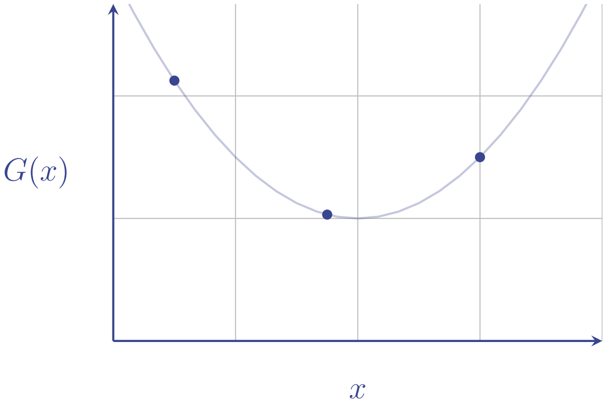

$G(x)$
x

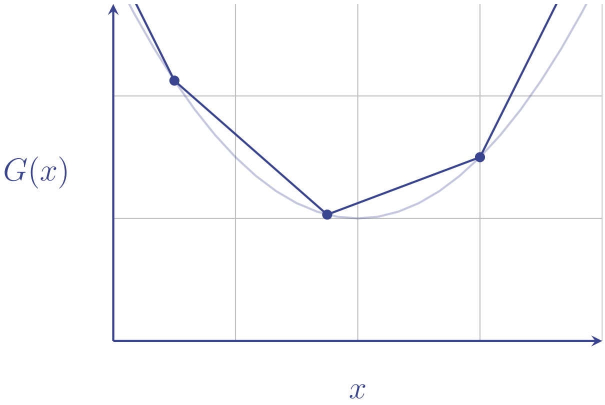

$G(x)$
x

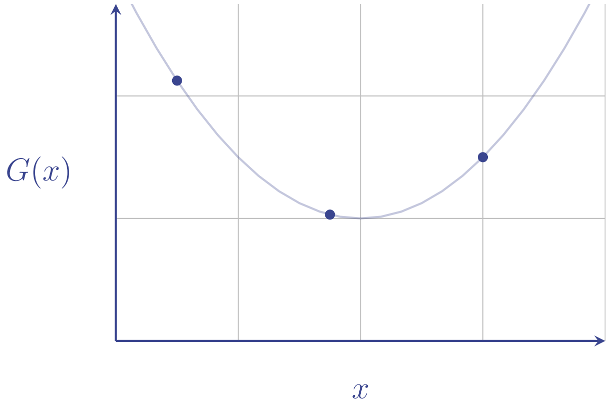

G(x)
x



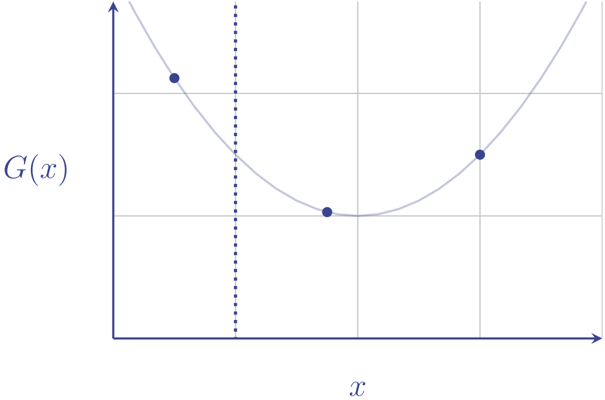
$$\underline{G}(x) = \min_{\mu \in \mathbb{R}} \quad \mu$$
$$\text{s.t.} \quad \mu \geq G(\hat{x}) + \langle d_{\hat{x}}, x - \hat{x} \rangle, \ \forall \hat{x},$$


$G(x)$
x

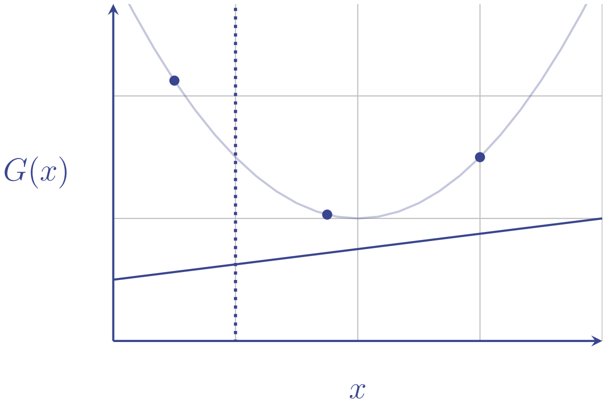

$G(x)$
x



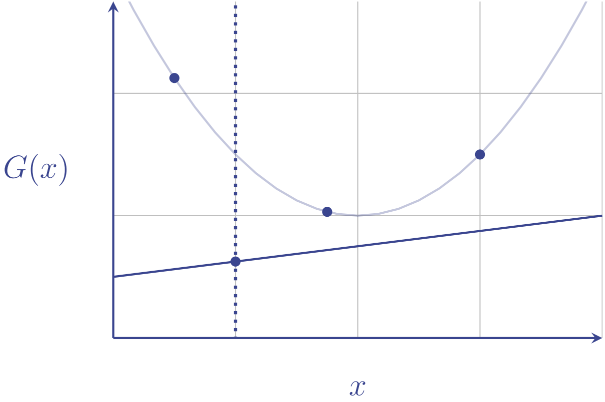
$G(x)$
x

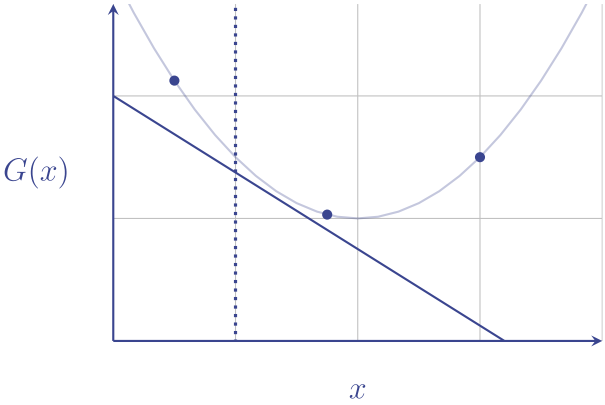


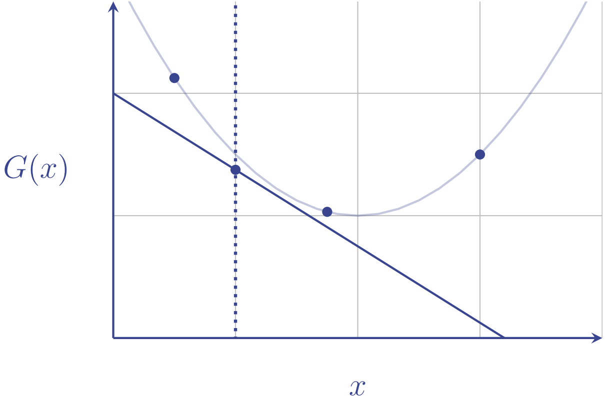
$G(x)$
x



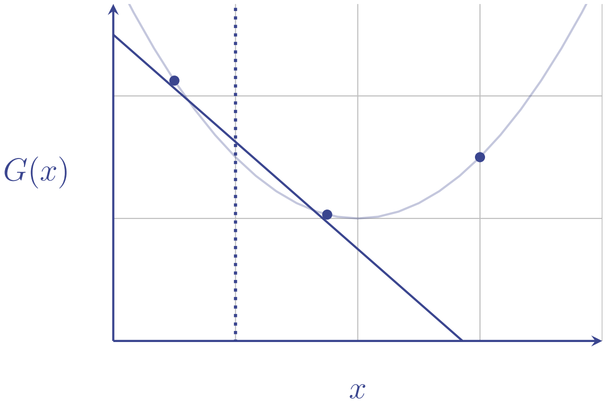
$G(x)$
x

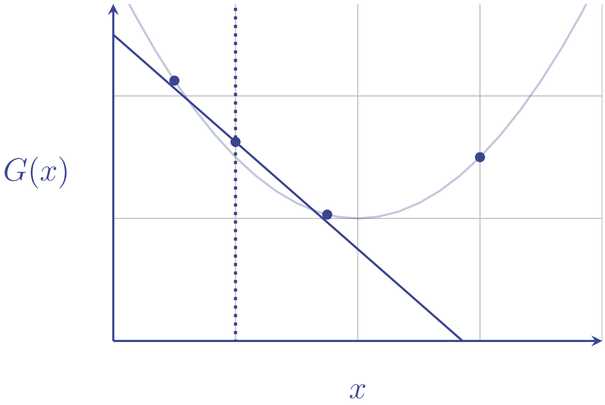

$G(x)$
x

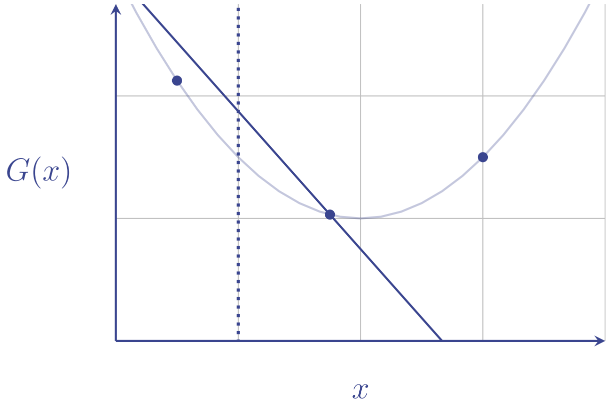

$G(x)$
x

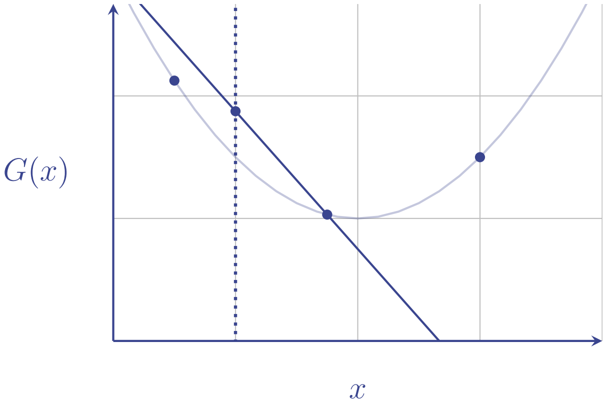


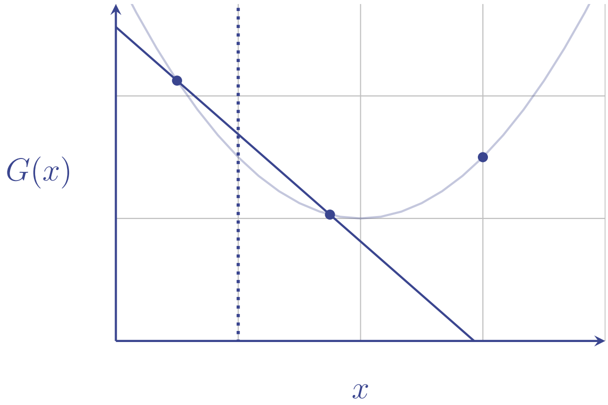
$G(x)$
x

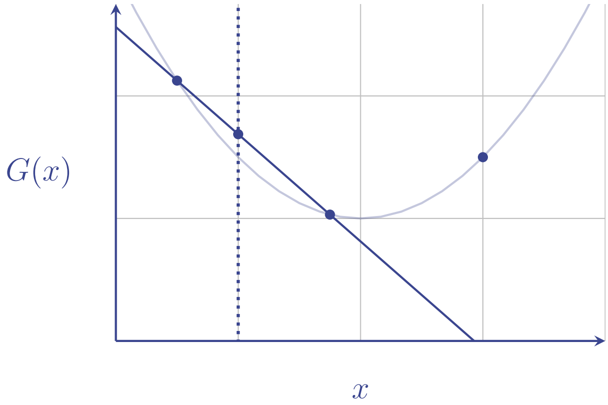

$G(x)$
x


$G(x)$
x

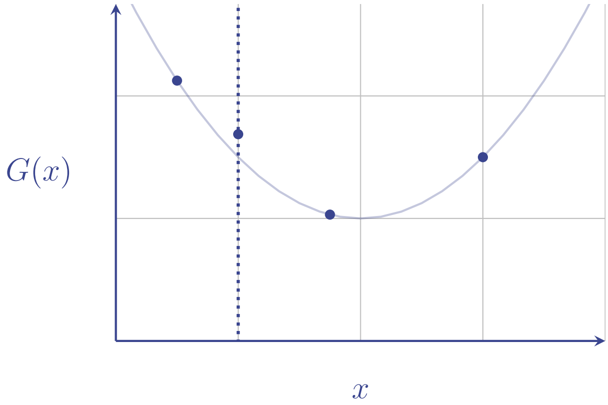

$G(x)$
x


$G(x)$
x


$G(x)$
x



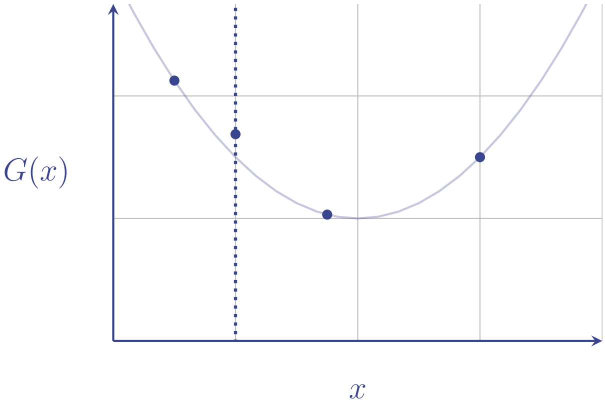
$G(x)$
x



$$\bar{G}(x) = \max_{\mu \in \mathbb{R}, \lambda \in \mathbb{R}^n} \quad \mu + \langle \lambda, x \rangle$$

$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x} \rangle \leq G(\hat{x}), \ \forall \hat{x},$$

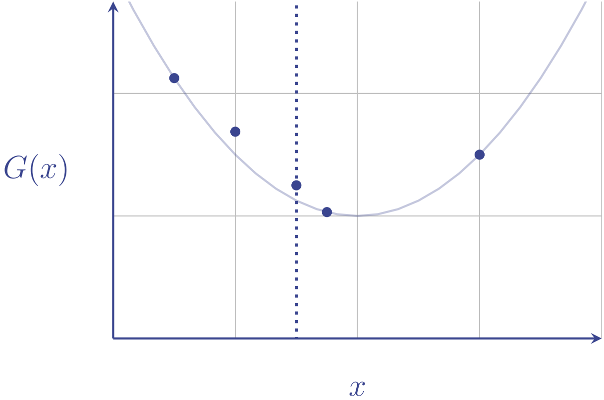
$$||\lambda||_* \leq \alpha.$$



$$\bar{G}(x) = \max_{\mu\in\mathbb{R},\lambda\in\mathbb{R}^n} \quad \mu + \langle \lambda, x \rangle$$

$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x} \rangle \leq G(\hat{x}),\ \forall \hat{x},$$

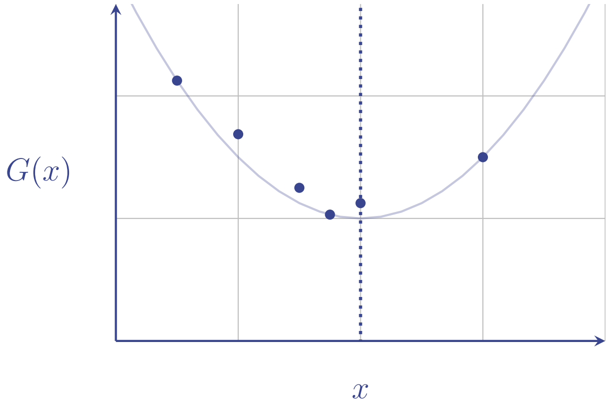
$$||\lambda||_* \leq \alpha.$$



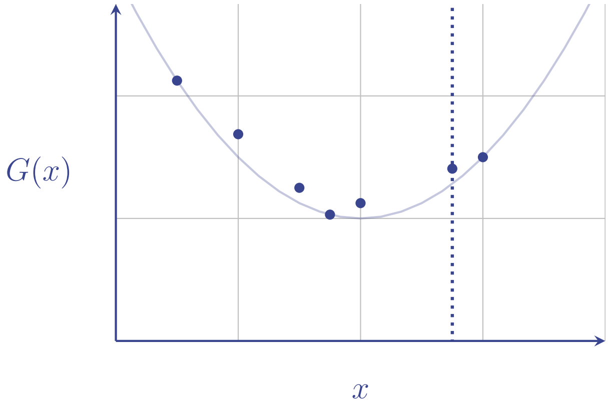
$$\bar{G}(x) = \max_{\mu \in \mathbb{R}, \lambda \in \mathbb{R}^n} \quad \mu + \langle \lambda, x \rangle$$

$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x} \rangle \leq G(\hat{x}), \ \forall \hat{x},$$

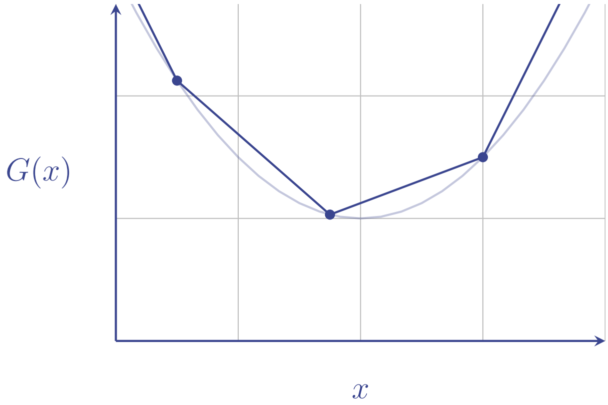
$$||\lambda||_* \leq \alpha.$$



$$\begin{aligned} \bar{G}(x) = \max_{\mu \in \mathbb{R}, \lambda \in \mathbb{R}^n} \quad & \mu + \langle \lambda, x \rangle \\ \text{s.t.} \quad & \mu + \langle \lambda, \hat{x} \rangle \leq G(\hat{x}),\ \forall \hat{x}, \\ & ||\lambda||_* \leq \alpha. \end{aligned}$$



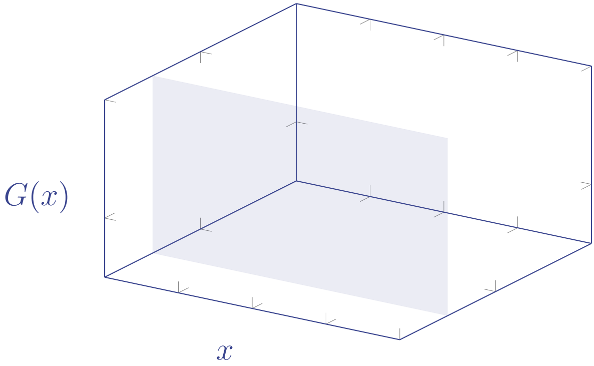
$$\bar{G}(x) = \max_{\mu\in\mathbb{R},\lambda\in\mathbb{R}^n} \quad \mu + \langle \lambda, x\rangle$$
$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x}\rangle \leq G(\hat{x}),\ \forall \hat{x},$$
$$||\lambda||_* \leq \alpha.$$

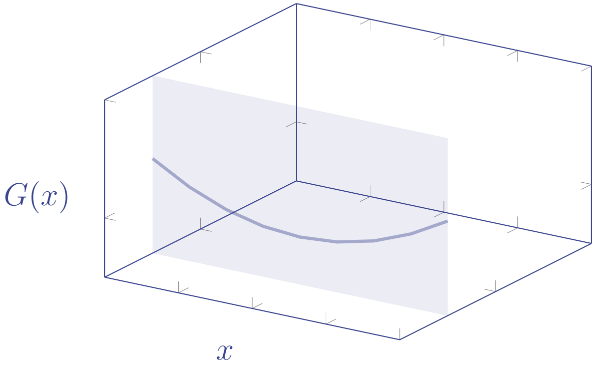


$$\bar{G}(x) = \max_{\mu\in\mathbb{R},\lambda\in\mathbb{R}^n} \quad \mu + \langle \lambda, x\rangle$$
$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x}\rangle \leq G(\hat{x}),\ \forall \hat{x},$$
$$||\lambda||_* \leq \alpha.$$

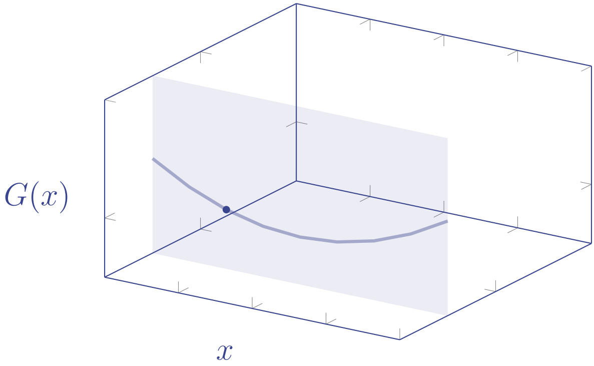
Theorem

These convex bounding functions are “nice”.

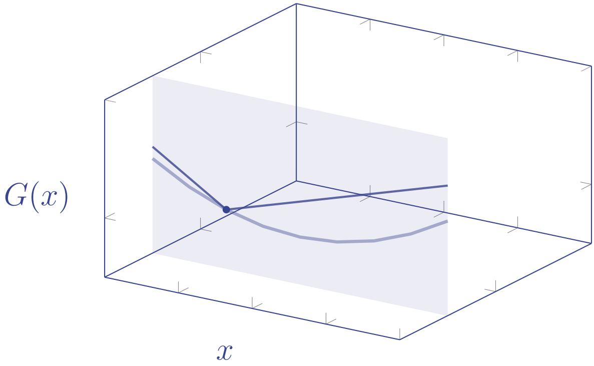

$G(x)$
x

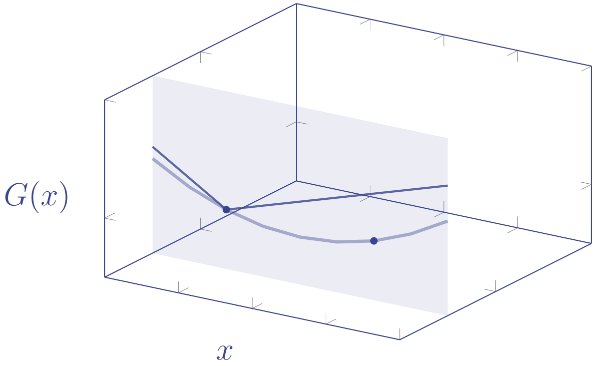


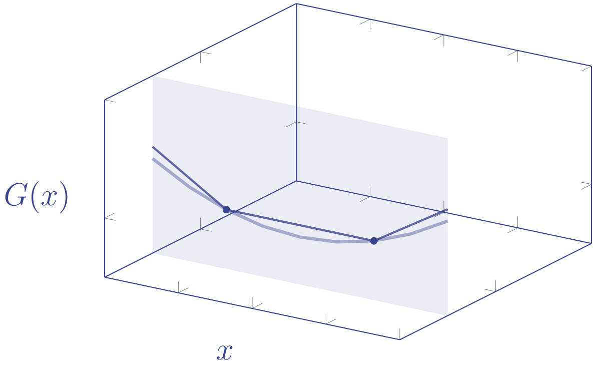
$G(x)$
x



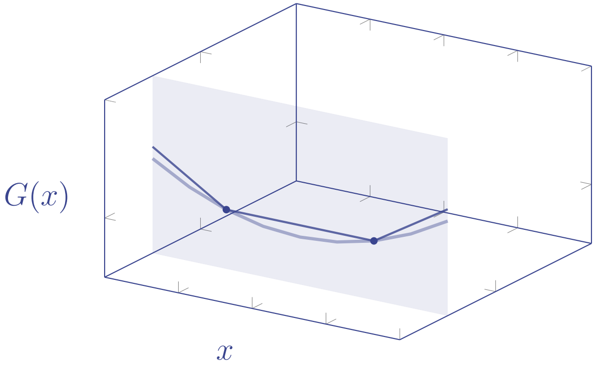
$G(x)$
x


G(x)
x


G(x)
x



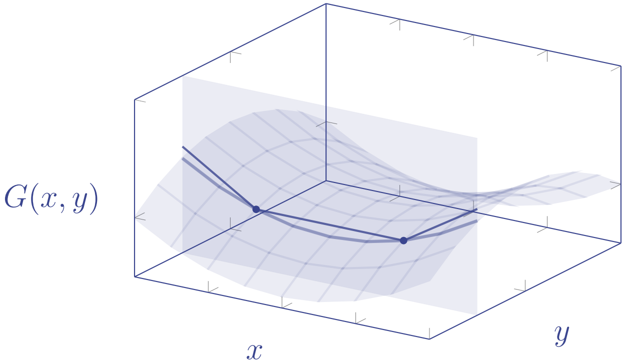
$G(x)$
x



$$\bar{G}(x) = \max_{\mu \in \mathbb{R}, \lambda \in \mathbb{R}^n} \quad \mu + \langle \lambda, x \rangle$$

$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x} \rangle \leq G(\hat{x}), \ \forall \hat{x},$$

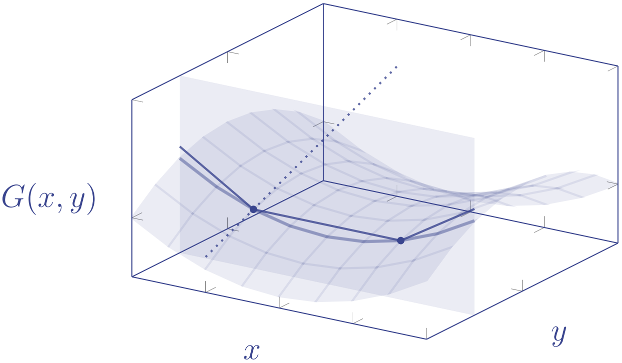
$$||\lambda||_* \leq \alpha.$$



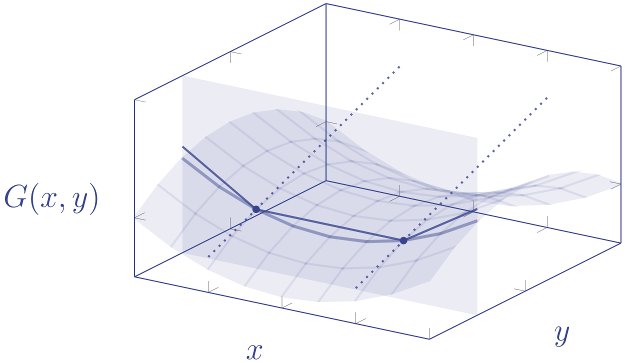
$$\bar{G}(x,\hat{y}) = \max_{\mu\in\mathbb{R},\lambda\in\mathbb{R}^n} \quad \mu + \langle \lambda, x\rangle$$

$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x}\rangle \leq G(\hat{x},\hat{y}) \qquad , \ \forall(\hat{x},\hat{y}),$$

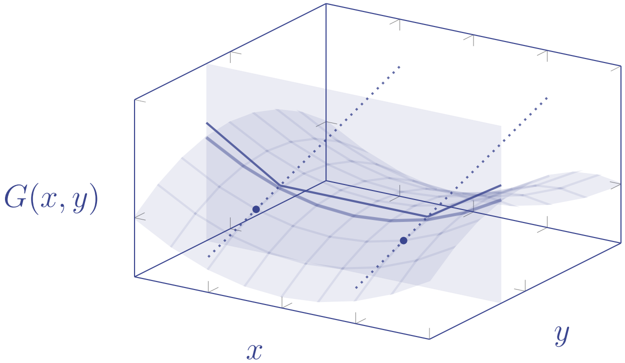
$$||\lambda||_* \leq \alpha.$$



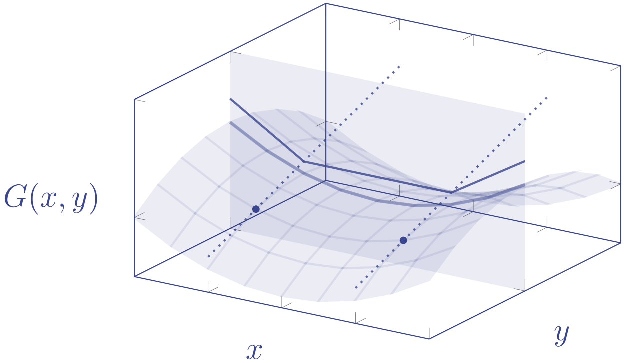
$$
\begin{aligned}
\bar{G}(x, \hat{y}) = \max_{\mu \in \mathbb{R}, \lambda \in \mathbb{R}^n} \quad & \mu + \langle \lambda, x \rangle \\
\text{s.t.} \quad & \mu + \langle \lambda, \hat{x} \rangle \leq G(\hat{x}, \hat{y}) \qquad , \ \forall (\hat{x}, \hat{y}), \\
& ||\lambda||_* \leq \alpha.
\end{aligned}
$$



$$\bar{G}(x,\hat{y}) = \max_{\mu\in\mathbb{R},\lambda\in\mathbb{R}^n} \quad \mu + \langle \lambda, x\rangle$$
$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x}\rangle \leq G(\hat{x},\hat{y}) \qquad , \ \forall(\hat{x},\hat{y}),$$
$$||\lambda||_* \leq \alpha.$$



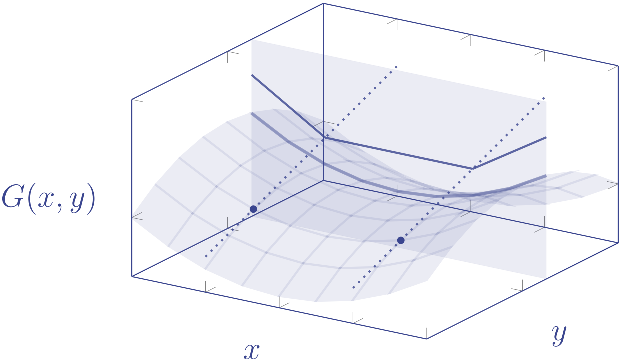
$$\bar{G}(x, y) = \max_{\mu \in \mathbb{R}, \lambda \in \mathbb{R}^n} \quad \mu + \langle \lambda, x \rangle$$
$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x} \rangle \leq G(\hat{x}, \hat{y}) + \langle d_{\hat{y}}, y - \hat{y} \rangle, \ \forall (\hat{x}, \hat{y}),$$
$$||\lambda||_* \leq \alpha.$$



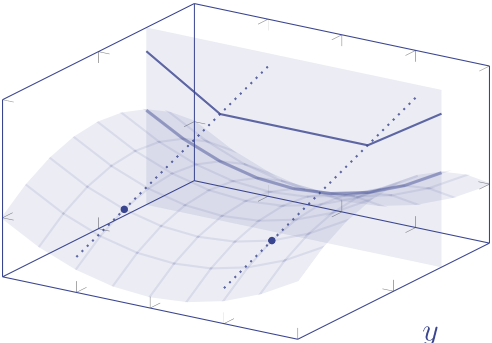
$$\bar{G}(x,y) = \max_{\mu\in\mathbb{R},\lambda\in\mathbb{R}^n} \quad \mu + \langle \lambda, x\rangle$$

$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x}\rangle \leq G(\hat{x},\hat{y}) + \langle d_{\hat{y}}, y - \hat{y}\rangle, \ \forall(\hat{x},\hat{y}),$$

$$||\lambda||_* \leq \alpha.$$



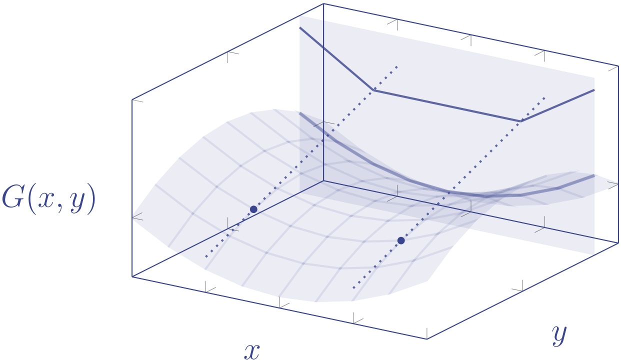
$$\bar{G}(x,y) = \max_{\mu\in\mathbb{R},\lambda\in\mathbb{R}^n} \quad \mu + \langle \lambda, x\rangle$$
$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x}\rangle \leq G(\hat{x},\hat{y}) + \langle d_{\hat{y}}, y - \hat{y}\rangle,\ \forall(\hat{x},\hat{y}),$$
$$||\lambda||_* \leq \alpha.$$



$$\bar{G}(x, y) = \max_{\mu\in\mathbb{R},\lambda\in\mathbb{R}^n} \quad \mu + \langle \lambda, x\rangle$$

$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x}\rangle \leq G(\hat{x}, \hat{y}) + \langle d_{\hat{y}}, y - \hat{y}\rangle, \ \forall(\hat{x}, \hat{y}),$$

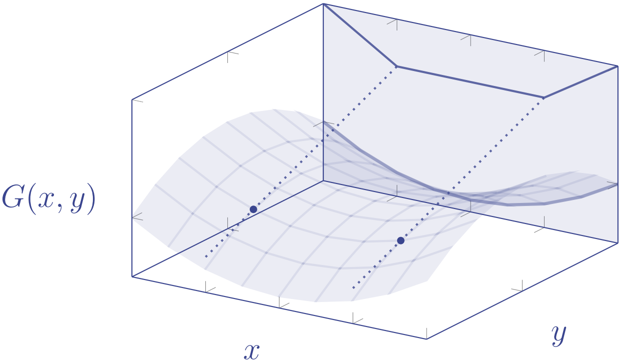
$$||\lambda||_* \leq \alpha.$$



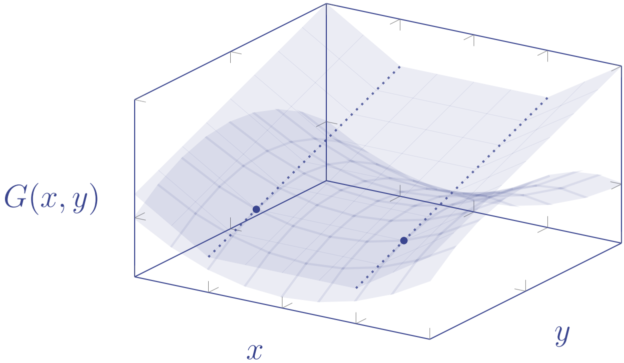
$$\bar{G}(x, y) = \max_{\mu \in \mathbb{R}, \lambda \in \mathbb{R}^n} \quad \mu + \langle \lambda, x \rangle$$

$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x} \rangle \leq G(\hat{x}, \hat{y}) + \langle d_{\hat{y}}, y - \hat{y} \rangle, \ \forall (\hat{x}, \hat{y}),$$

$$||\lambda||_* \leq \alpha.$$



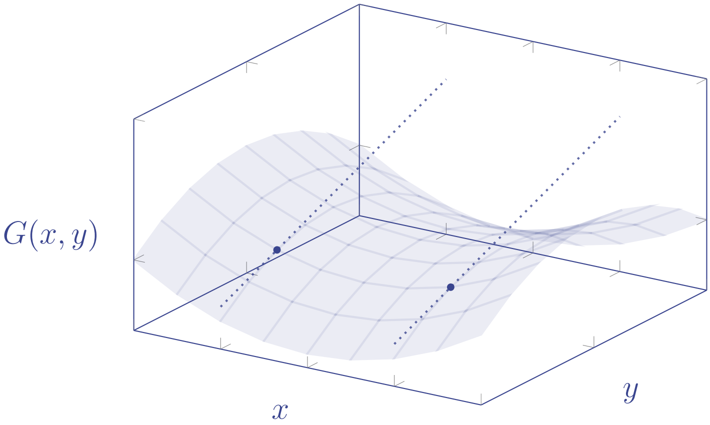
$$\bar{G}(x, {\color{red}y}) = \max_{\mu\in\mathbb{R},\lambda\in\mathbb{R}^n} \quad \mu + \langle \lambda, x\rangle$$
$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x}\rangle \le G(\hat{x}, \hat{y}) {\color{red}+ \langle d_{\hat{y}}, y - \hat{y}\rangle}, \ \forall(\hat{x}, \hat{y}),$$
$$||\lambda||_* \le \alpha.$$



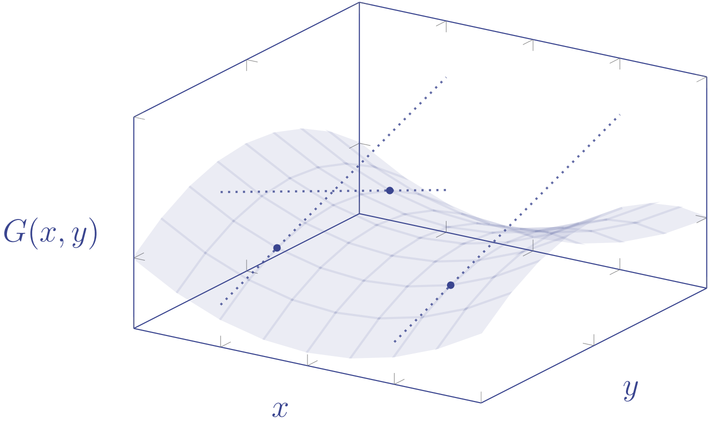
$$\bar{G}(x, y) = \max_{\mu \in \mathbb{R}, \lambda \in \mathbb{R}^n} \quad \mu + \langle \lambda, x \rangle$$

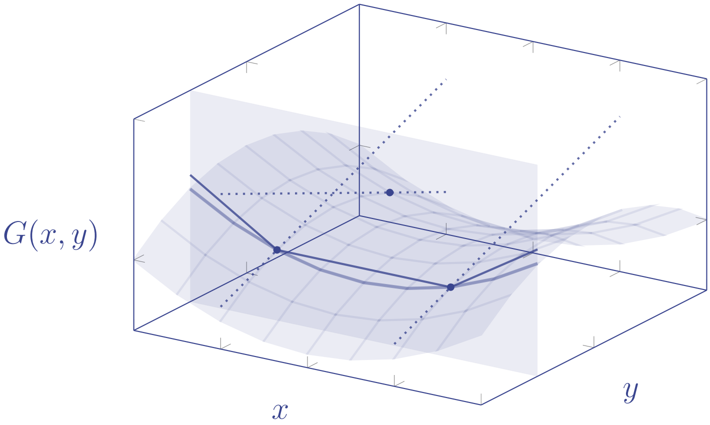
$$\text{s.t.} \quad \mu + \langle \lambda, \hat{x} \rangle \leq G(\hat{x}, \hat{y}) + \langle d_{\hat{y}}, y - \hat{y} \rangle, \ \forall (\hat{x}, \hat{y}),$$

$$||\lambda||_* \leq \alpha.$$

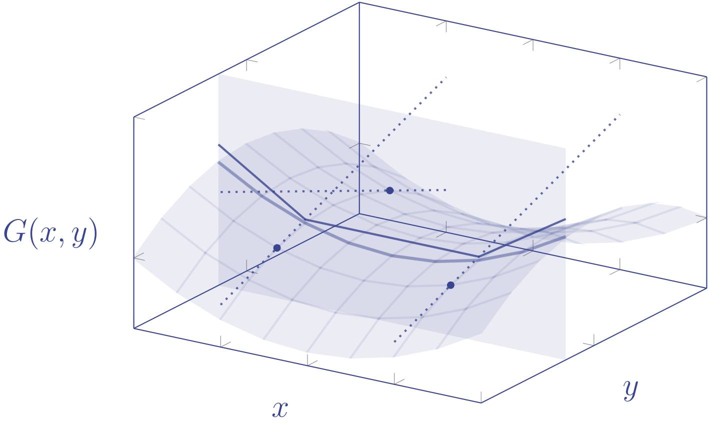


Bounding functions | Saddle bounding functions

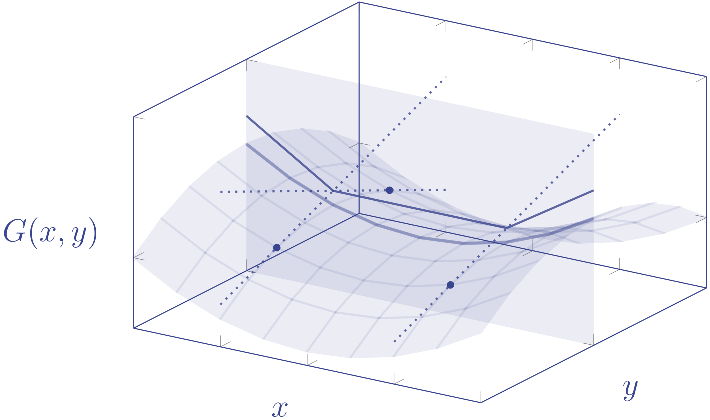

$G(x, y)$
x
y



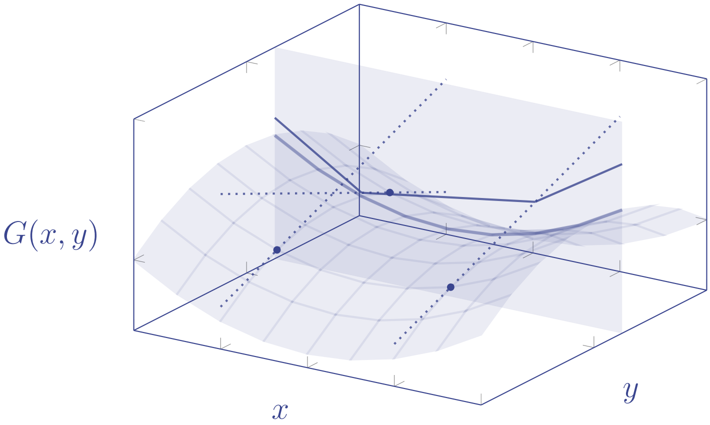
$G(x, y)$
x
y



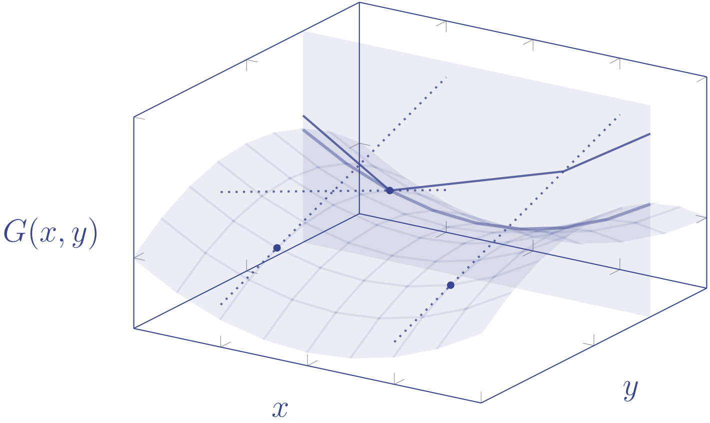
$G(x, y)$
x
y



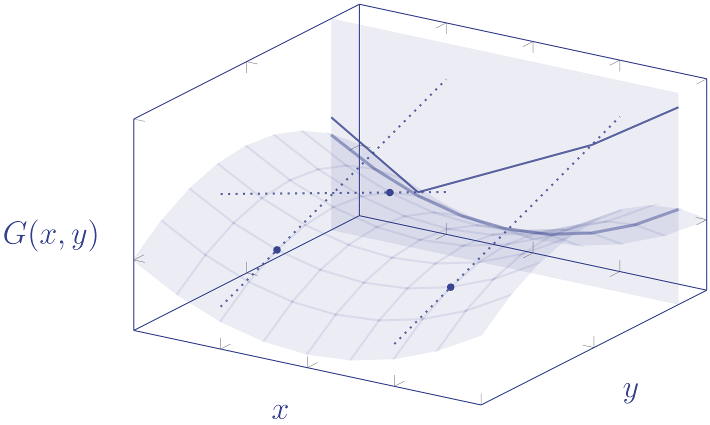
$G(x,y)$
x
y

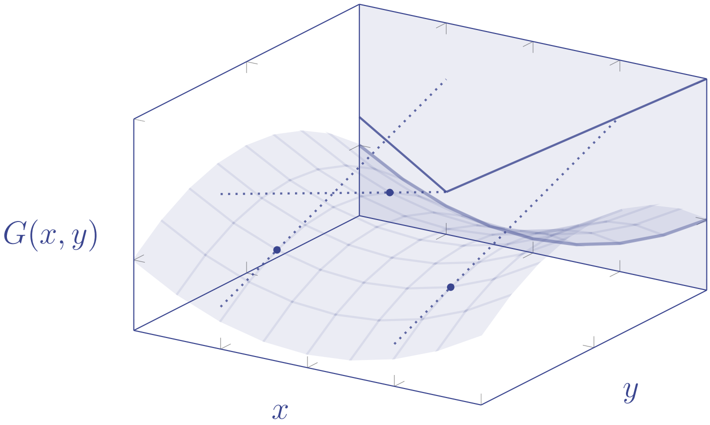


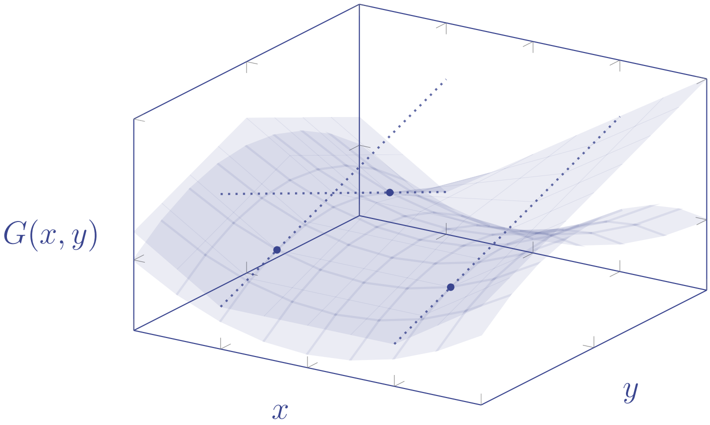
Bounding functions | Saddle bounding functions



$G(x,y)$
x
y


$G(x, y)$
x
y

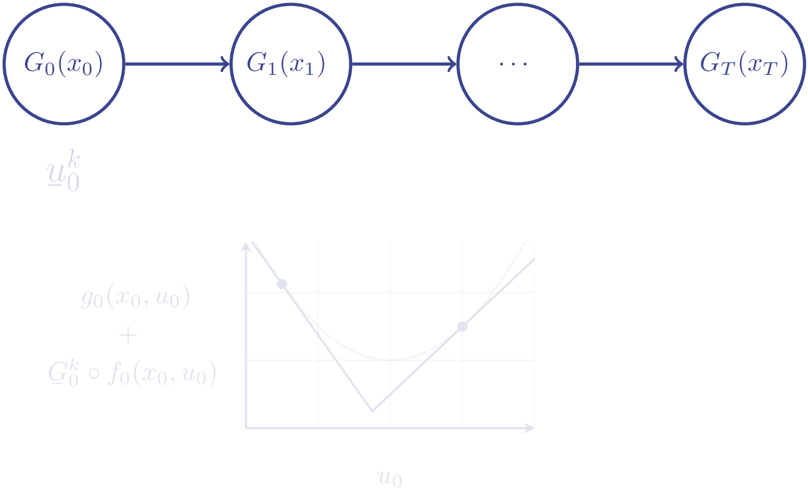

$G(x, y)$
x
y

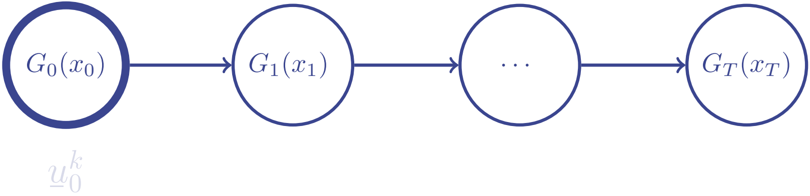


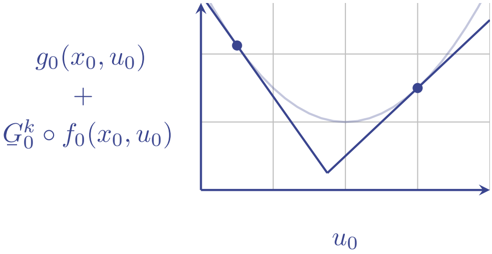
Bounding functions | Saddle bounding functions

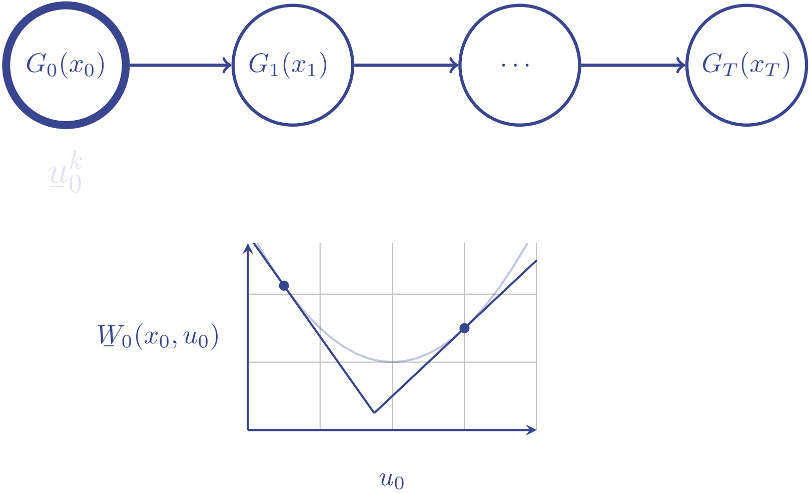
Theorem

These saddle bounding functions
are also “nice”.

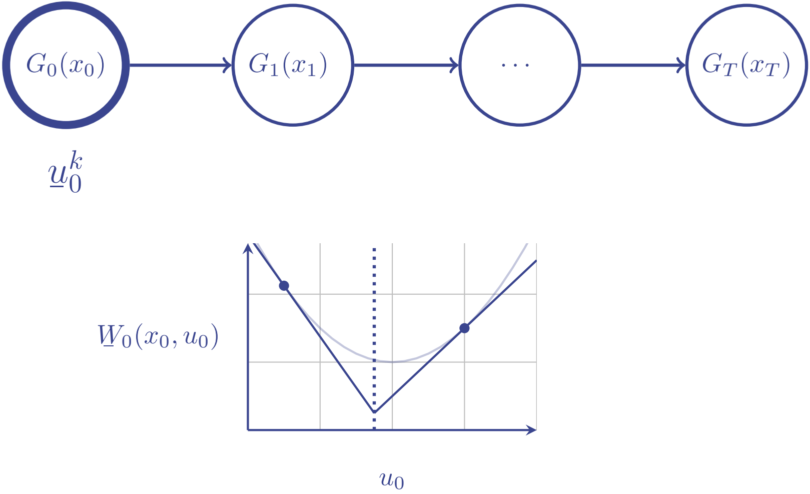

$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$
$\underline{u}_0^k$
$g_0(x_0, u_0)$
$+$
$\underline{G}_0^k \circ f_0(x_0, u_0)$
u_0


$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$
$\underline{u}_0^k$

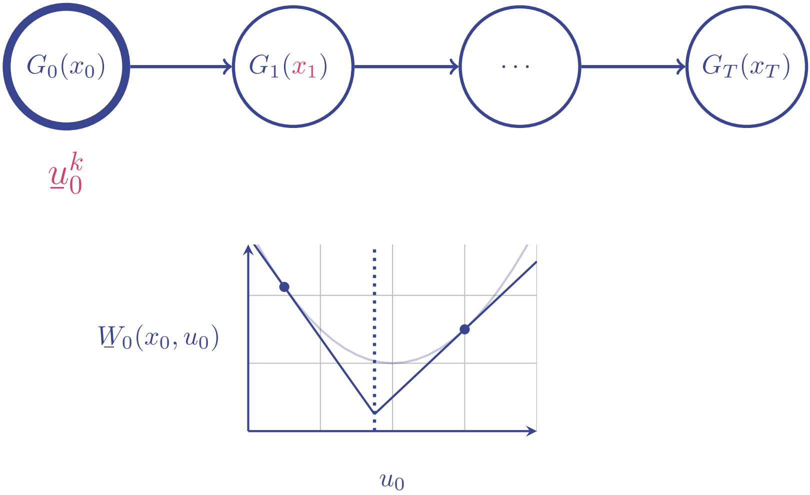

$g_0(x_0, u_0)$
$+$
$\underline{G}_0^k \circ f_0(x_0, u_0)$
u_0

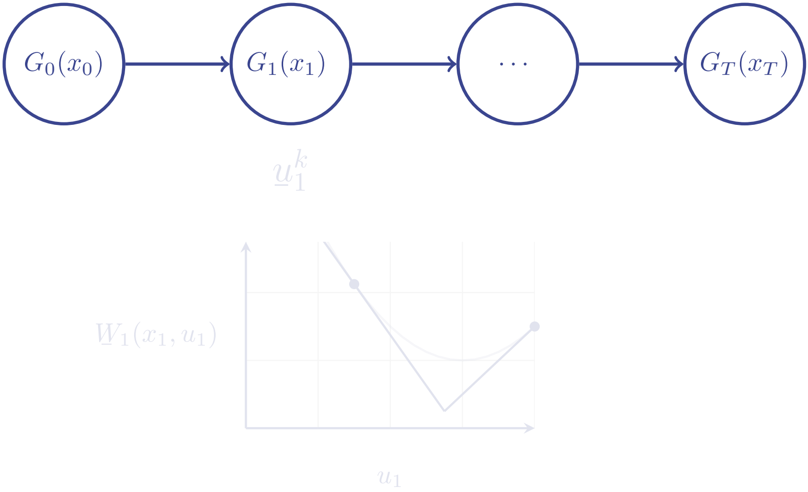


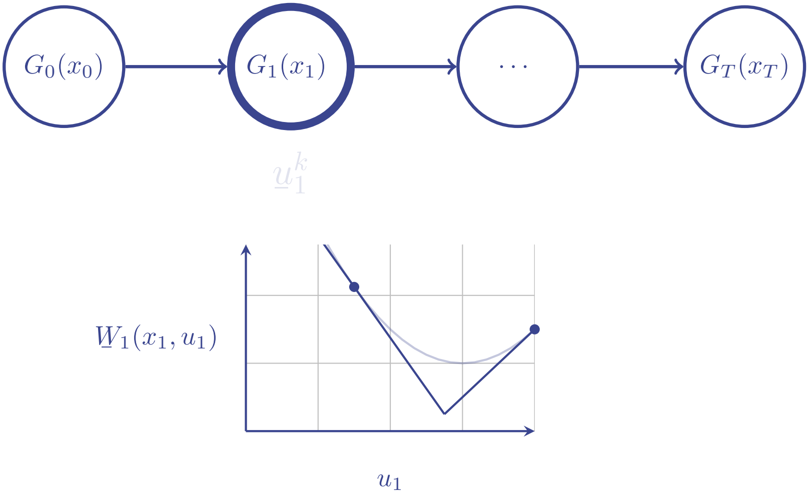
$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$
$\underline{u}_0^k$
$\underline{W}_0(x_0, u_0)$
u_0

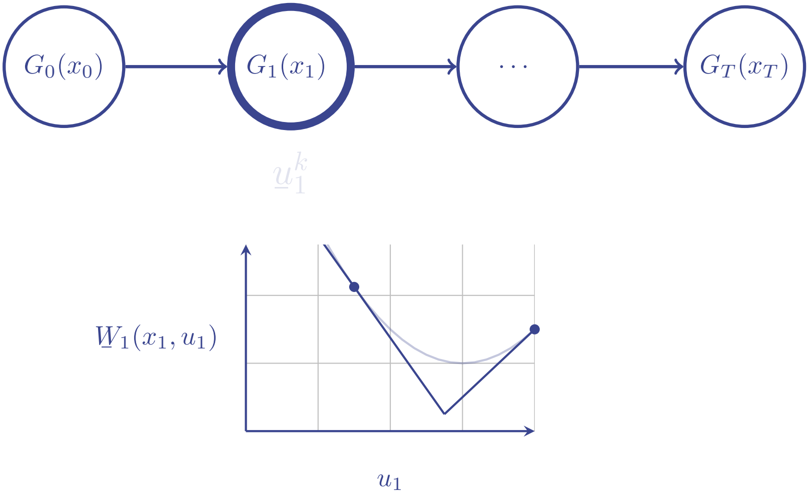


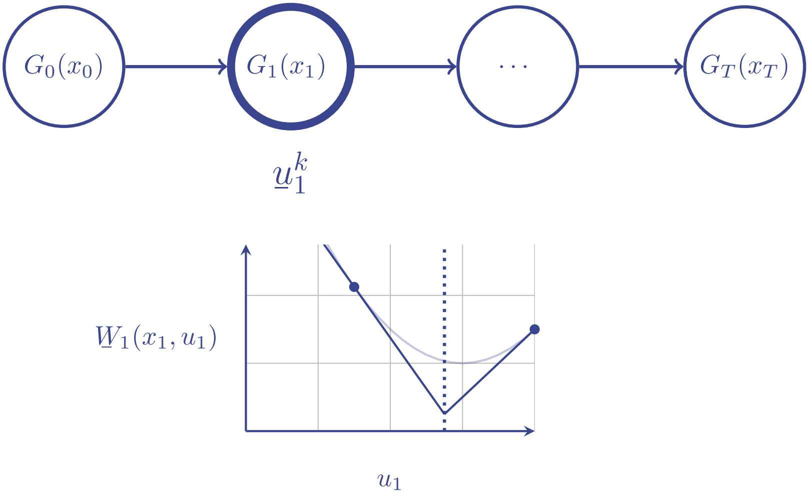
$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$
$\underline{u}_0^k$
$\underline{W}_0(x_0, u_0)$
u_0

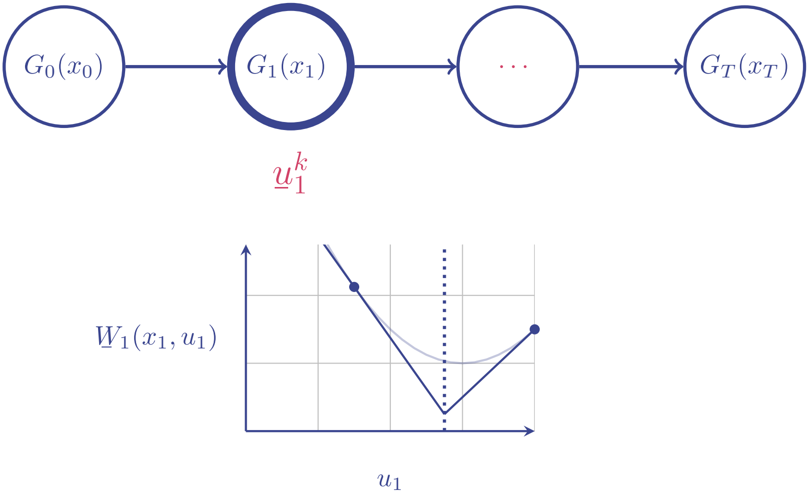

$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$
$\underline{u}_0^k$
$\underline{W}_0(x_0, u_0)$
u_0

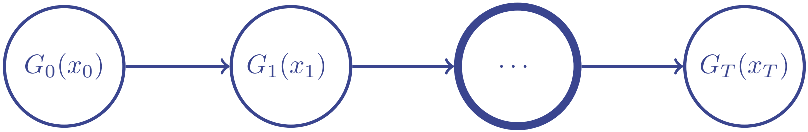

$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$
$\underline{u}_1^k$
$\underline{W}_1(x_1, u_1)$
u_1


$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$
$\underline{u}_1^k$
$\underline{W}_1(x_1, u_1)$
u_1

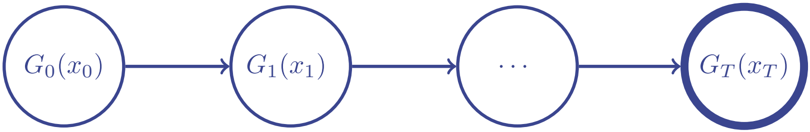

$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$
$\underline{u}_1^k$
$\underline{W}_1(x_1, u_1)$
u_1


$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$
$\underline{u}_1^k$
$\underline{W}_1(x_1, u_1)$
u_1

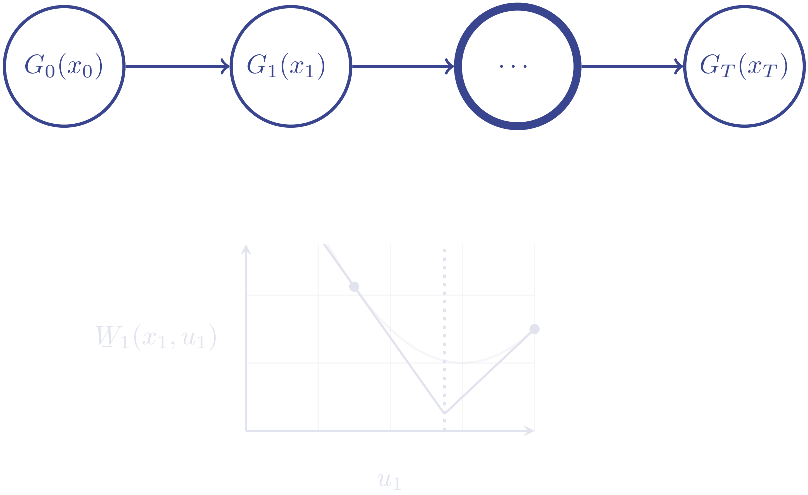

$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$
$\underline{u}_1^k$
$\underline{W}_1(x_1, u_1)$
u_1



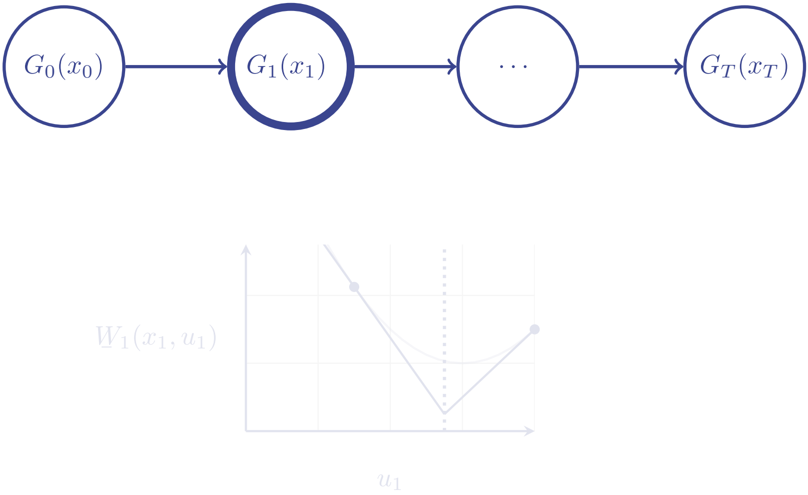
Algorithm | Dual Dynamic Programming

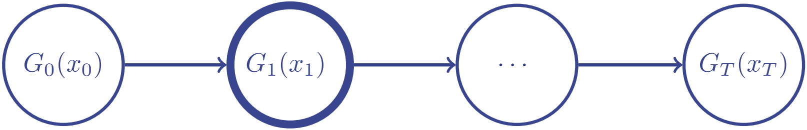


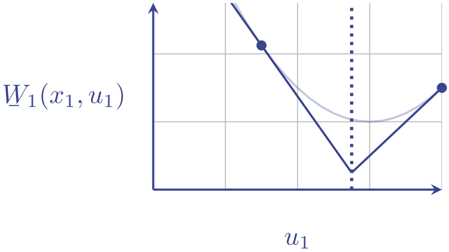
Algorithm | Dual Dynamic Programming

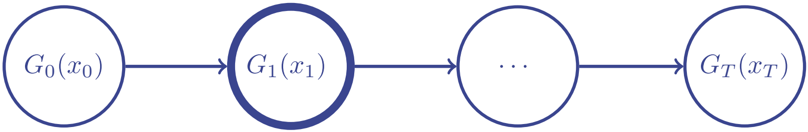


Algorithm | Dual Dynamic Programming

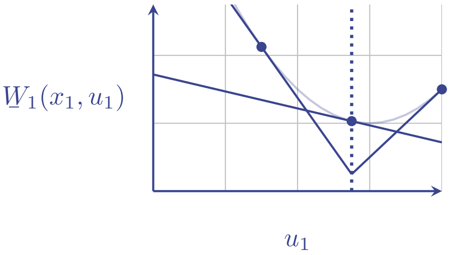

$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$
$\underline{W}_1(x_1, u_1)$
u_1


$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$

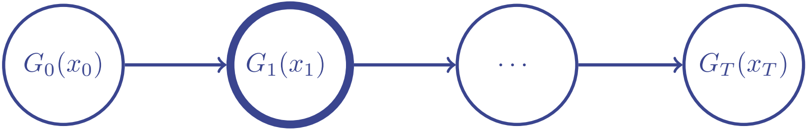

$\underline{W}_1(x_1, u_1)$
u_1

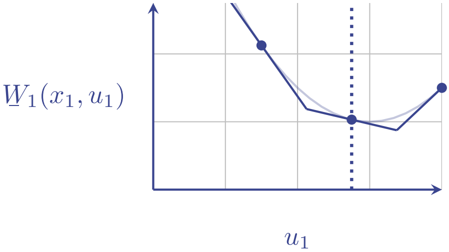


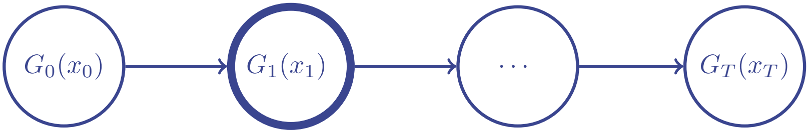
$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$



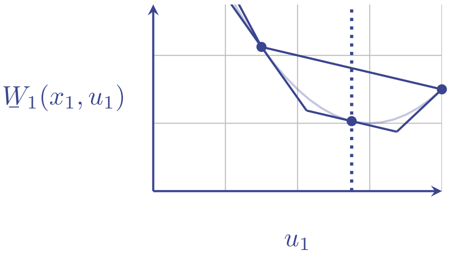
$\underline{W}_1(x_1, u_1)$
u_1


$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$

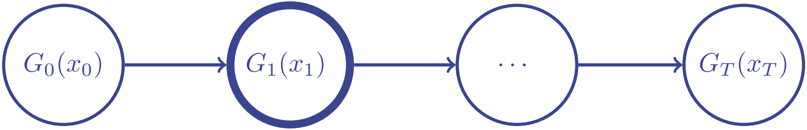

$\underline{W}_1(x_1, u_1)$
u_1

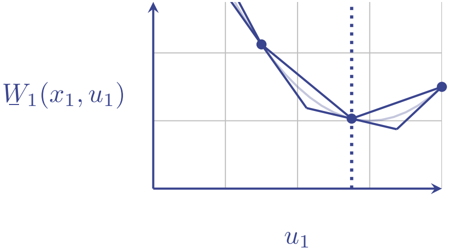


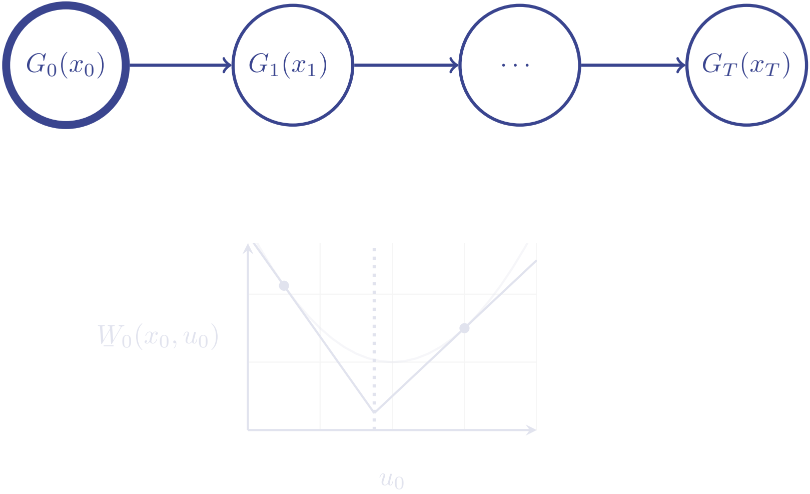
$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$

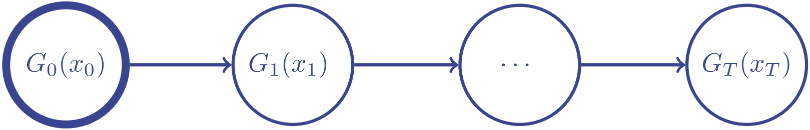


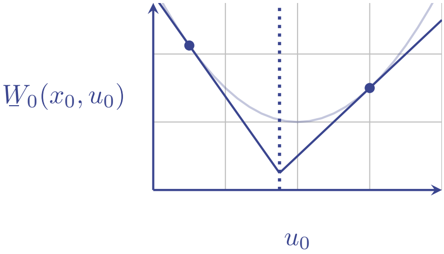
$\underline{W}_1(x_1, u_1)$
u_1

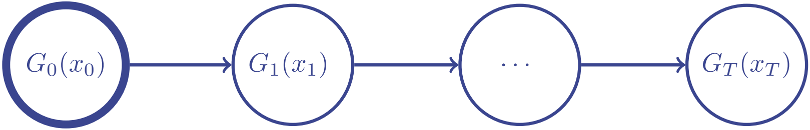

$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$

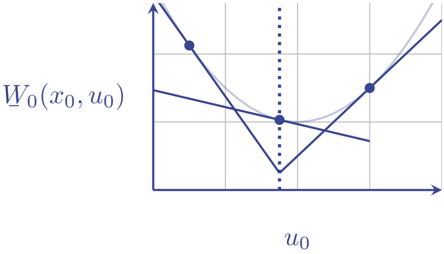

$\underline{W}_1(x_1, u_1)$
u_1


$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$
$\underline{W}_0(x_0, u_0)$
u_0

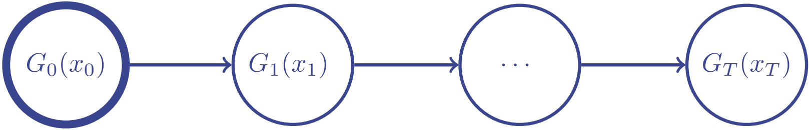

$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$

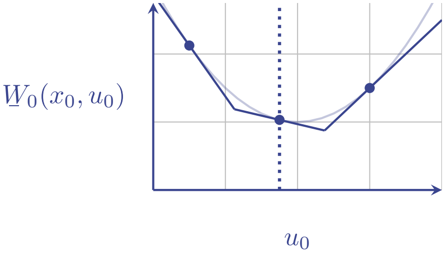

$\underline{W}_0(x_0, u_0)$
u_0

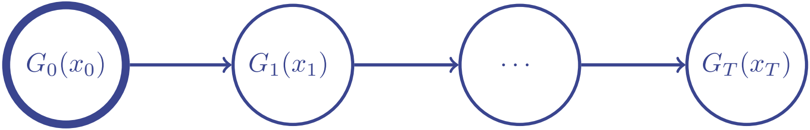




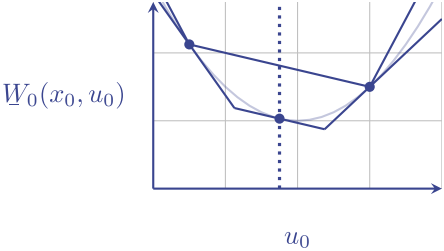
Algorithm | Dual Dynamic Programming


$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$

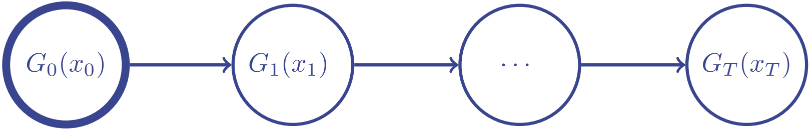

$\underline{W}_0(x_0, u_0)$
u_0

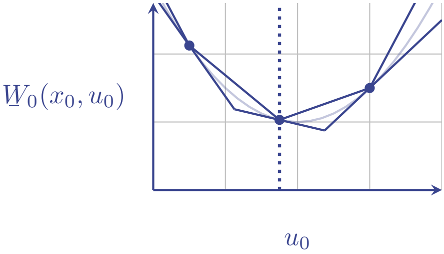


$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$



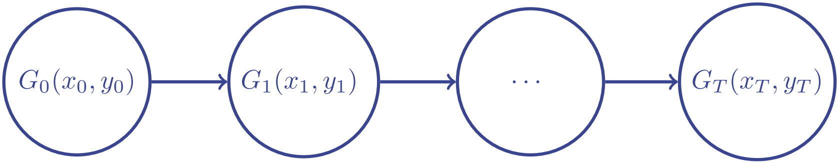
$\underline{W}_0(x_0, u_0)$
u_0


$G_0(x_0)$
$G_1(x_1)$
$\cdots$
$G_T(x_T)$

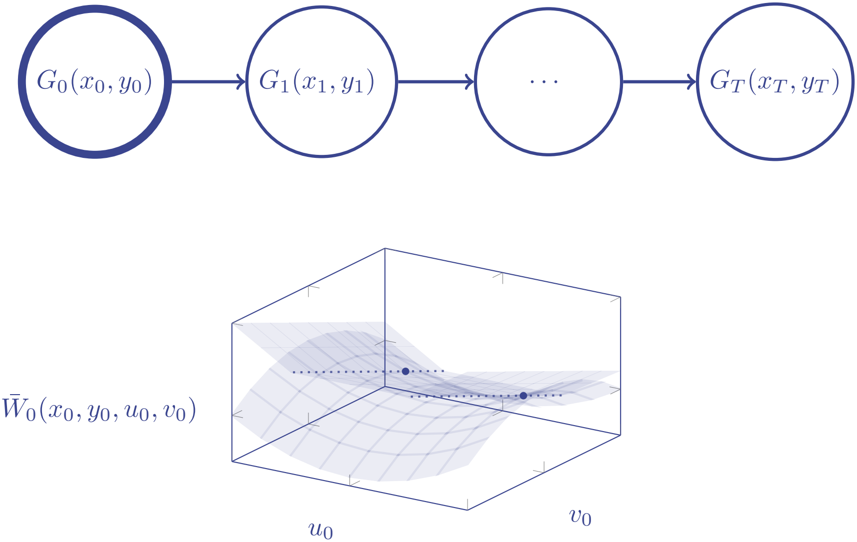

$\underline{W}_0(x_0, u_0)$
u_0

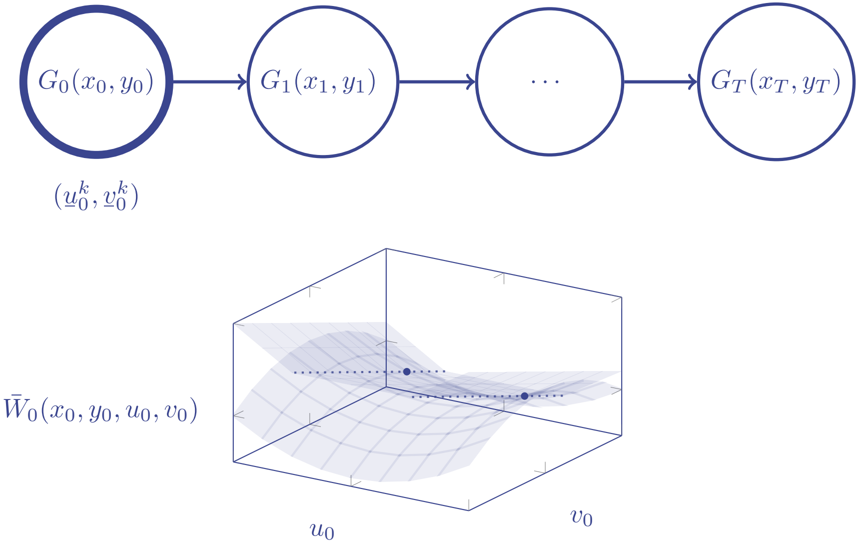
Theorem

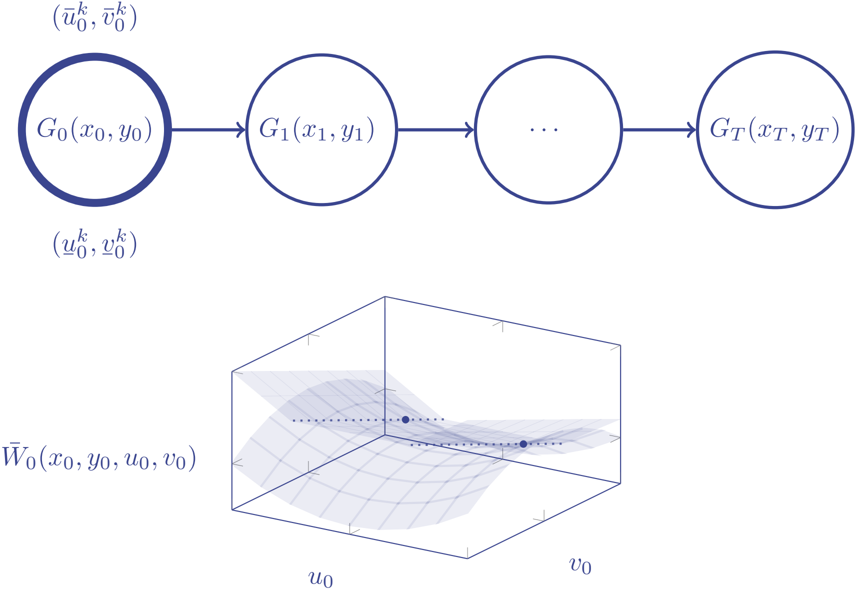
The minimums of $\bar{W}_t(x_t^k, \cdot)$ and $\underline{W}_t(x_t^k, \cdot)$ converge as $k \to \infty$.

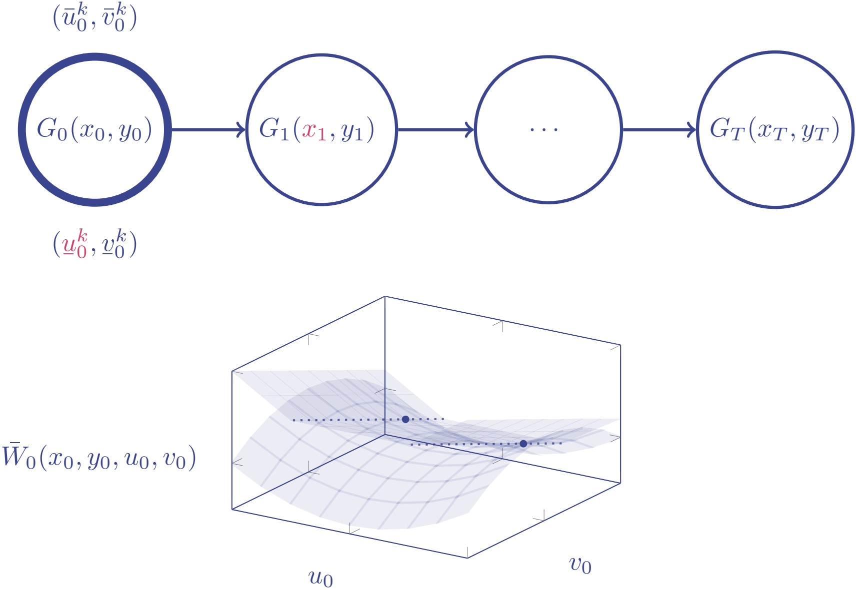

$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$

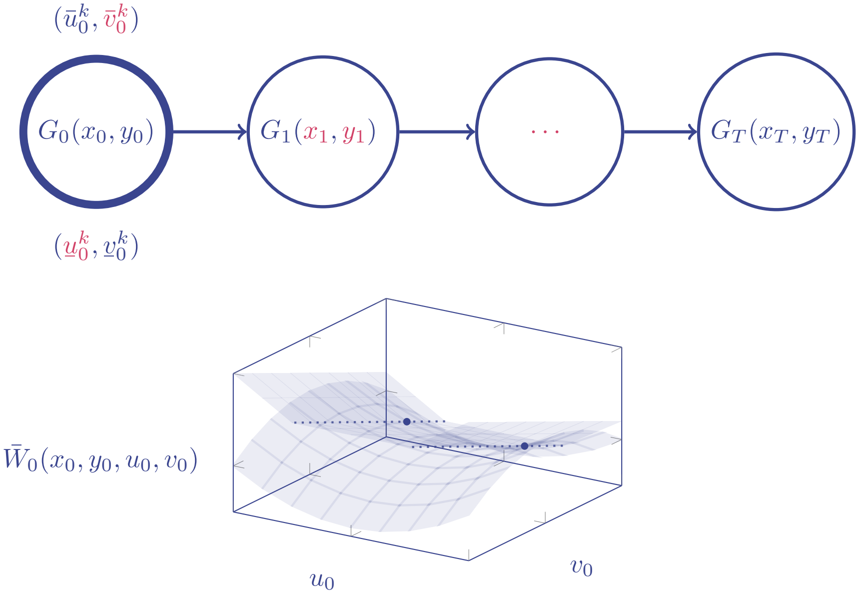
$\bar{W}_0(x_0, y_0, u_0, v_0)$
u_0
v_0

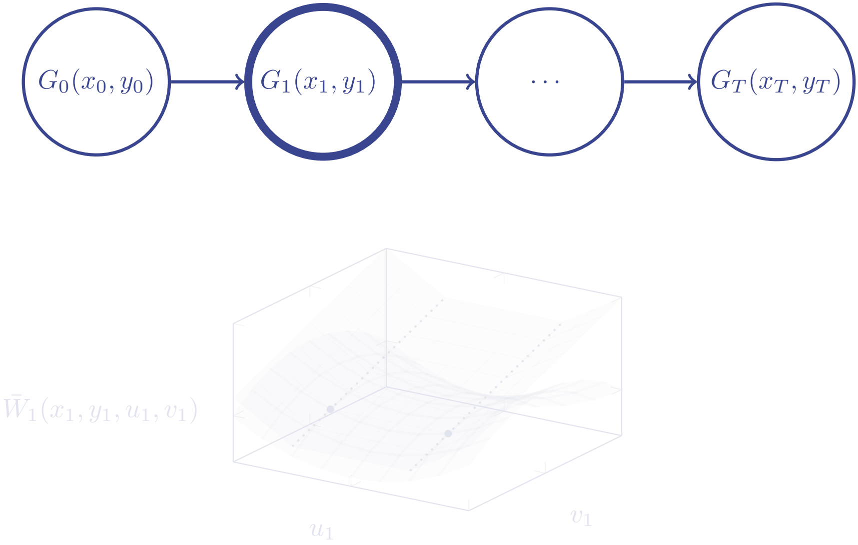

$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$\bar{W}_0(x_0, y_0, u_0, v_0)$
u_0
v_0


$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$(\underline{u}_0^k, \underline{v}_0^k)$
$\bar{W}_0(x_0, y_0, u_0, v_0)$
u_0
v_0

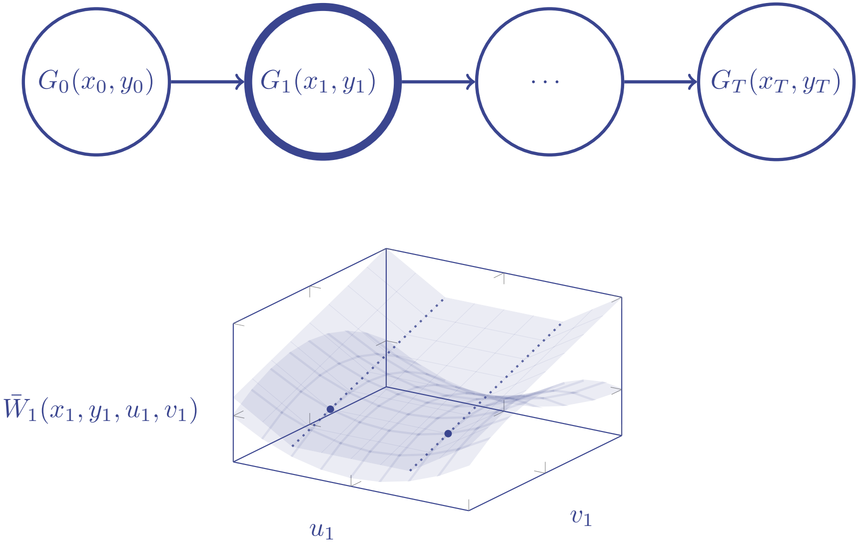

$(\bar{u}_0^k, \bar{v}_0^k)$
$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$(\underline{u}_0^k, \underline{v}_0^k)$
$\bar{W}_0(x_0, y_0, u_0, v_0)$
u_0
v_0

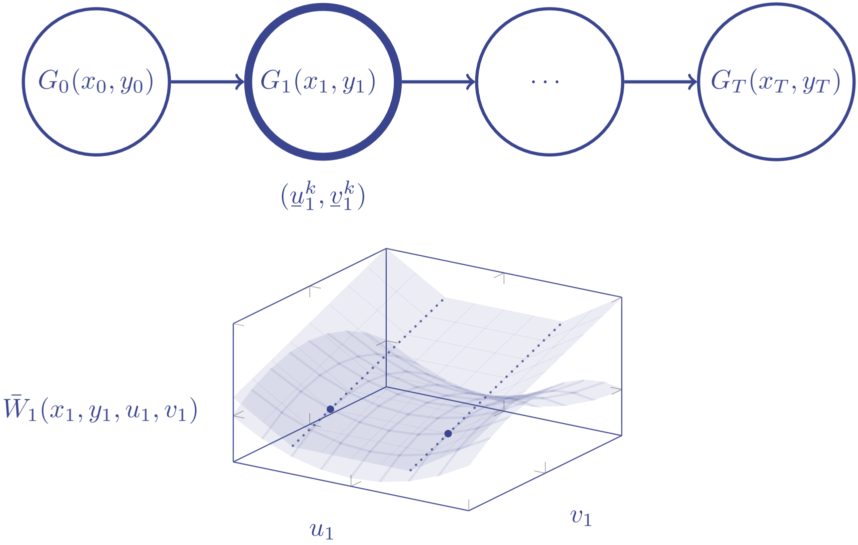

$(\bar{u}_0^k, \bar{v}_0^k)$
$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$(\underline{u}_0^k, \underline{v}_0^k)$
$\bar{W}_0(x_0, y_0, u_0, v_0)$
u_0
v_0


$(\bar{u}_0^k, \bar{v}_0^k)$
$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$(\underline{u}_0^k, \underline{v}_0^k)$
$\bar{W}_0(x_0, y_0, u_0, v_0)$
u_0
v_0

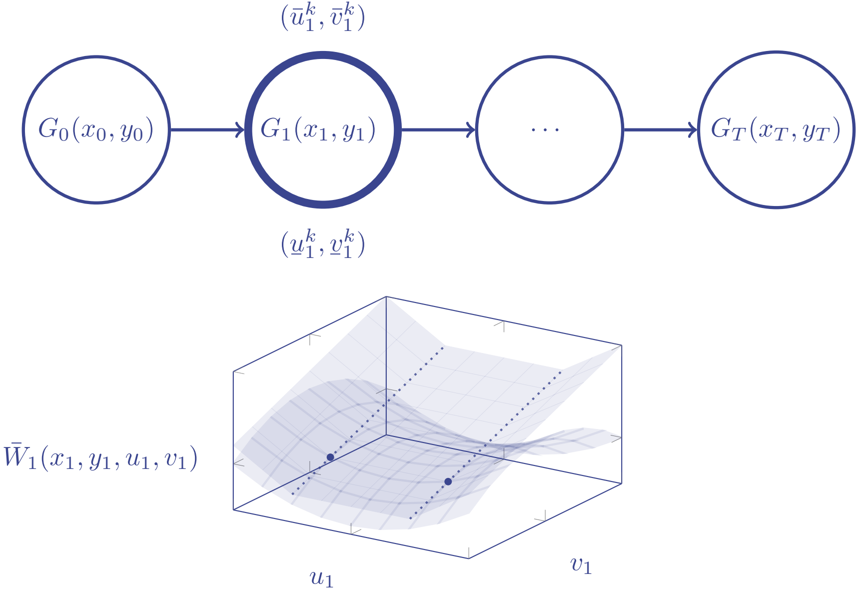


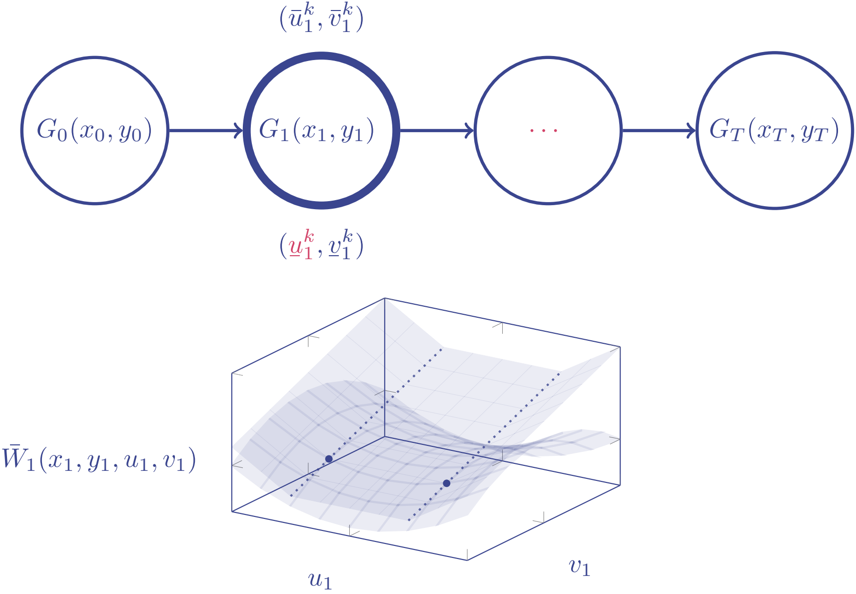
$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$\bar{W}_1(x_1, y_1, u_1, v_1)$
u_1
v_1

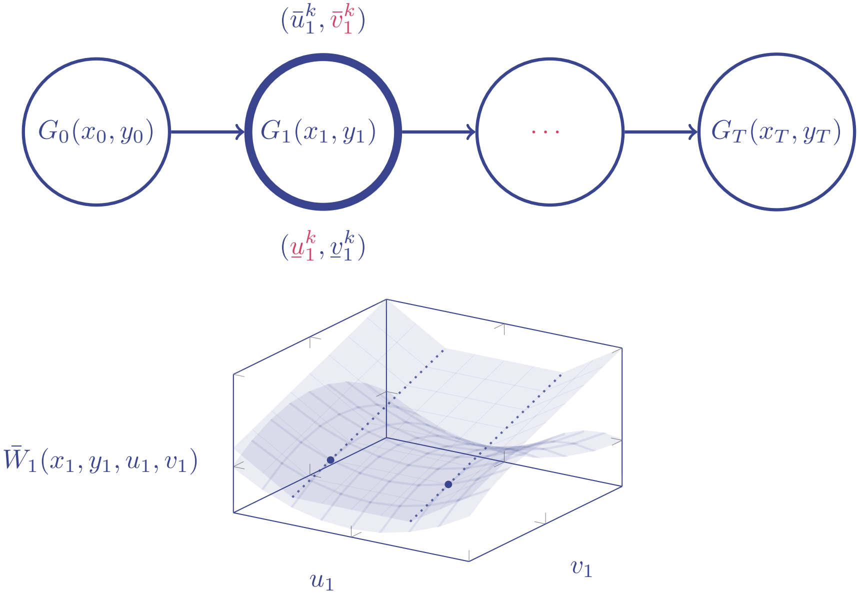

$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$\bar{W}_1(x_1, y_1, u_1, v_1)$
u_1
v_1



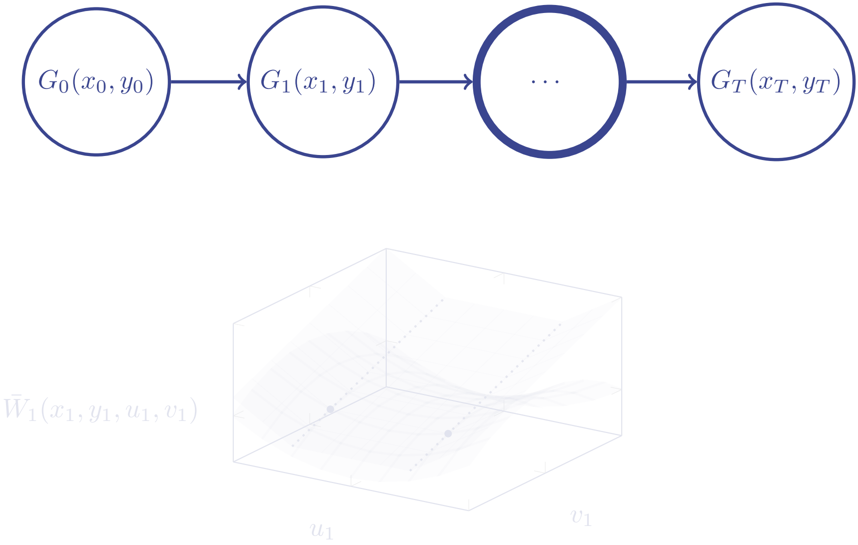
Algorithm | Saddle Algorithm


$(\bar{u}_1^k, \bar{v}_1^k)$
$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$(\underline{u}_1^k, \underline{v}_1^k)$
$\bar{W}_1(x_1, y_1, u_1, v_1)$
u_1
v_1

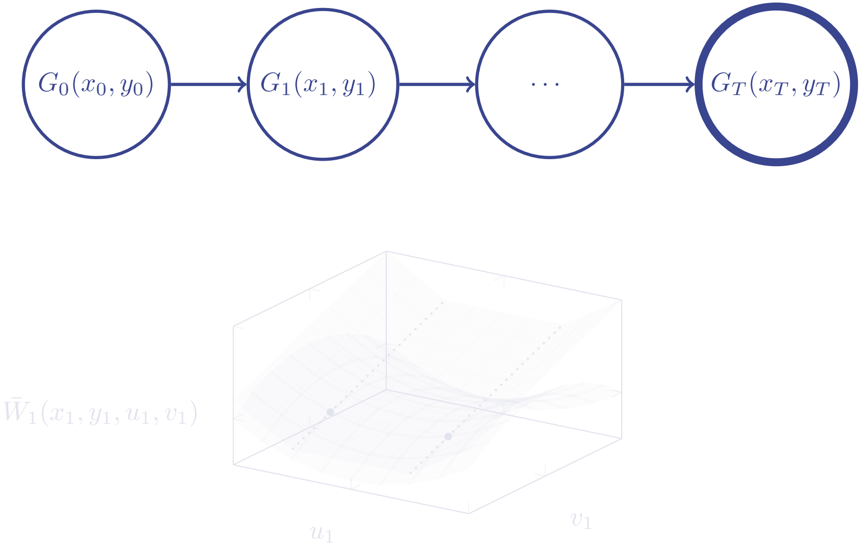

$(\bar{u}_1^k, \bar{v}_1^k)$
$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$(\underline{u}_1^k, \underline{v}_1^k)$
$\bar{W}_1(x_1, y_1, u_1, v_1)$
u_1
v_1

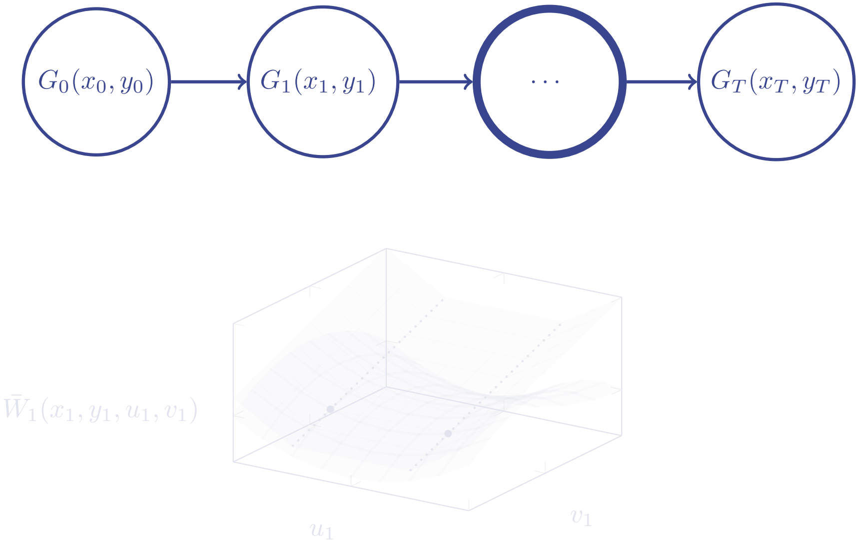


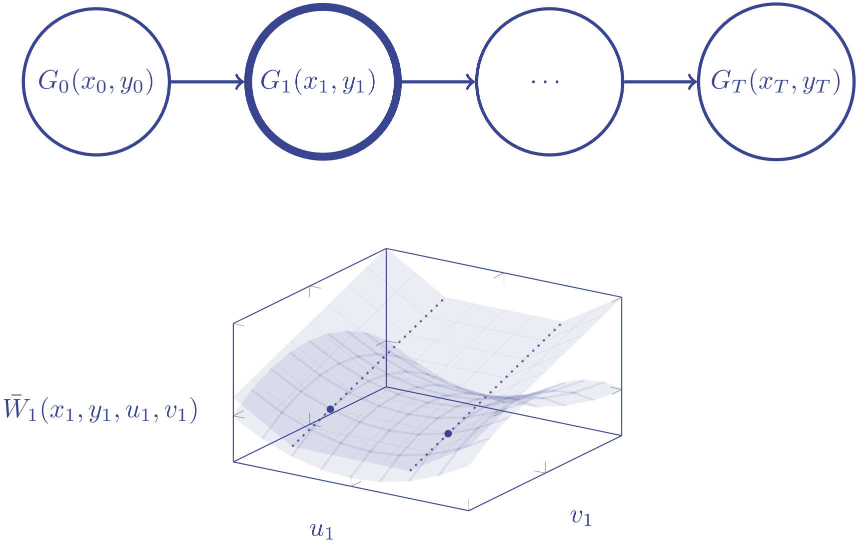
Algorithm | Saddle Algorithm

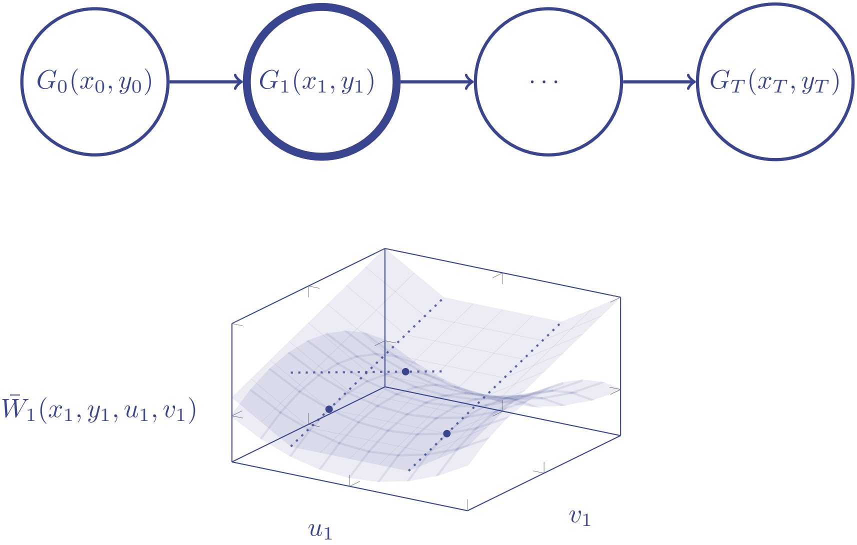


$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$\bar{W}_1(x_1, y_1, u_1, v_1)$
u_1
v_1

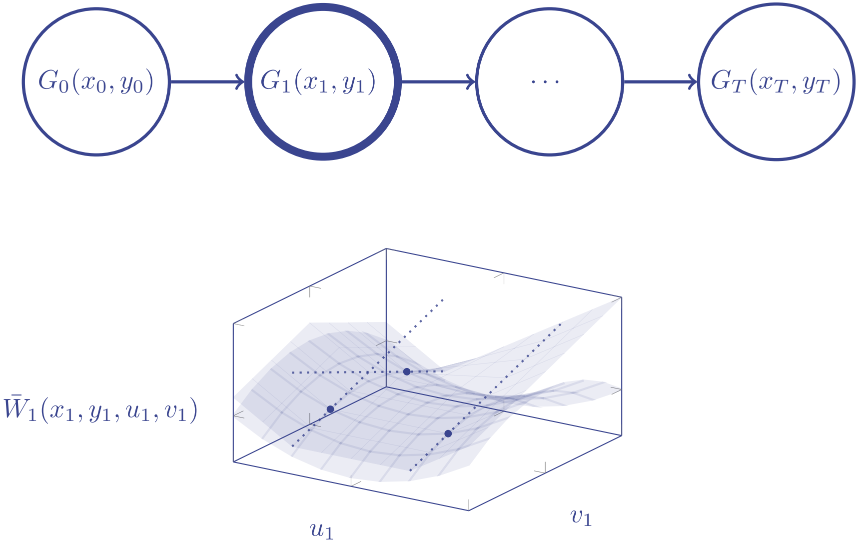

$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$\bar{W}_1(x_1, y_1, u_1, v_1)$
u_1
v_1

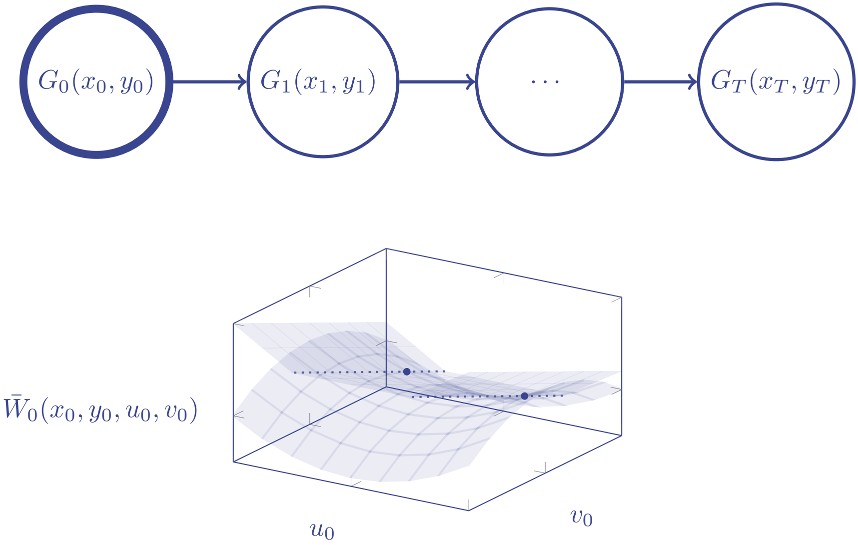

$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$\bar{W}_1(x_1, y_1, u_1, v_1)$
u_1
v_1


$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$\bar{W}_1(x_1, y_1, u_1, v_1)$
u_1
v_1

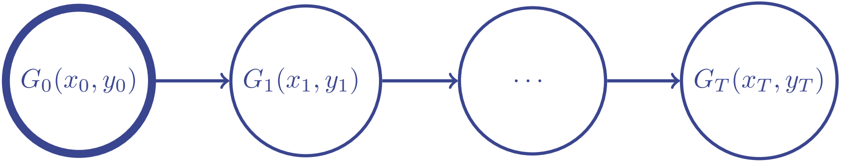


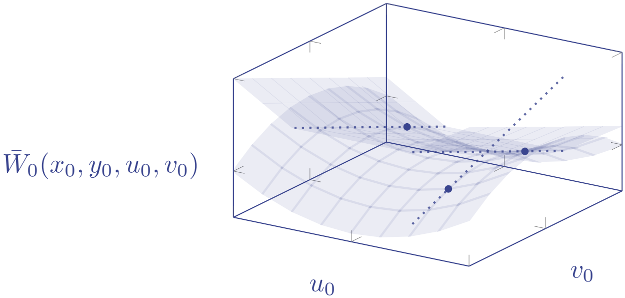
$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$\bar{W}_1(x_1, y_1, u_1, v_1)$
u_1
v_1

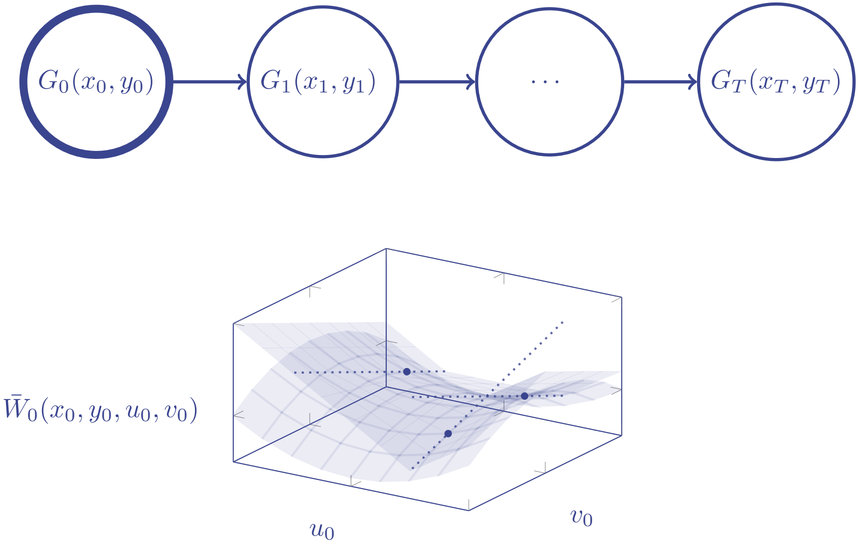

$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$\bar{W}_1(x_1, y_1, u_1, v_1)$
u_1
v_1

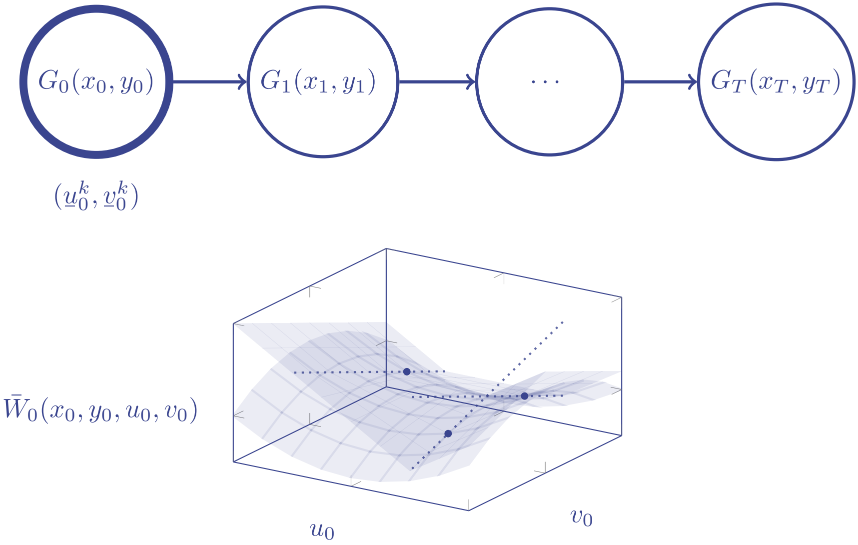


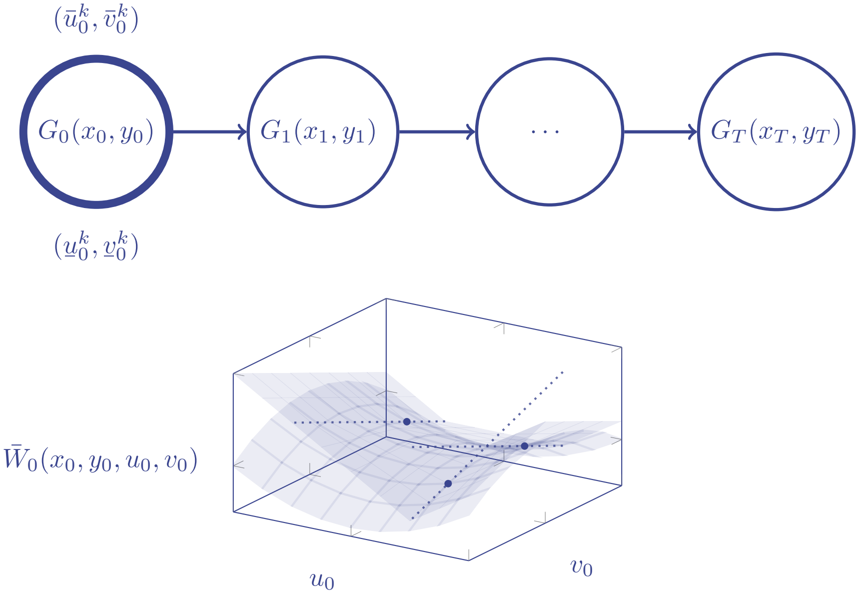
Algorithm | Saddle Algorithm

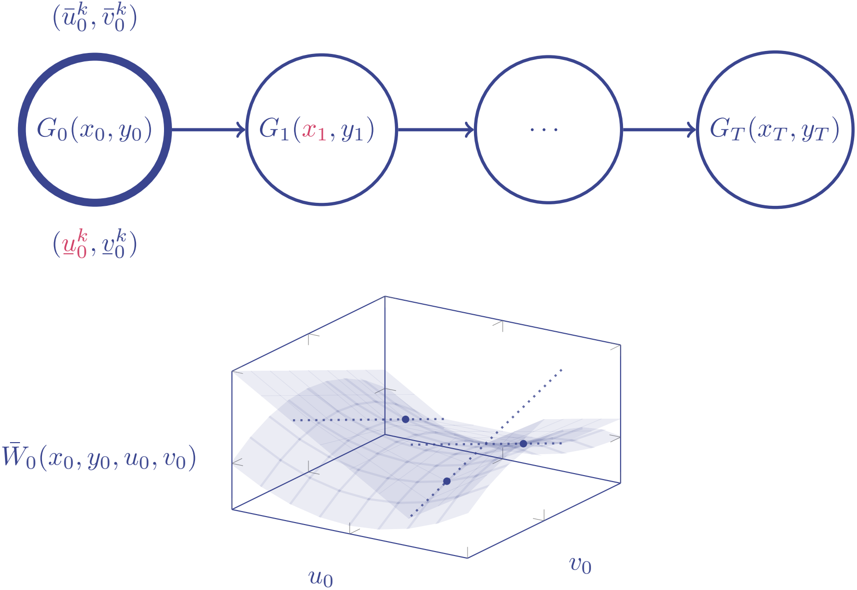

$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$

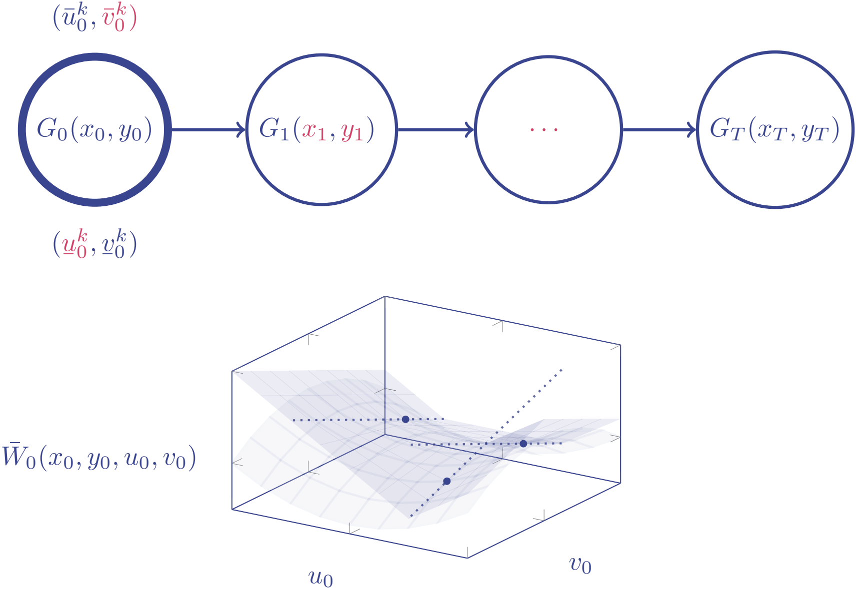

$\bar{W}_0(x_0, y_0, u_0, v_0)$
u_0
v_0


$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$\bar{W}_0(x_0, y_0, u_0, v_0)$
u_0
v_0


$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$(\underline{u}_0^k, \underline{v}_0^k)$
$\bar{W}_0(x_0, y_0, u_0, v_0)$
u_0
v_0


$(\bar{u}_0^k, \bar{v}_0^k)$
$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$(\underline{u}_0^k, \underline{v}_0^k)$
$\bar{W}_0(x_0, y_0, u_0, v_0)$
u_0
v_0


$(\bar{u}_0^k, \bar{v}_0^k)$
$G_0(x_0, y_0)$
$G_1(x_1, y_1)$
$\cdots$
$G_T(x_T, y_T)$
$(\underline{u}_0^k, \underline{v}_0^k)$
$\bar{W}_0(x_0, y_0, u_0, v_0)$
u_0
v_0



Algorithm | Saddle Algorithm

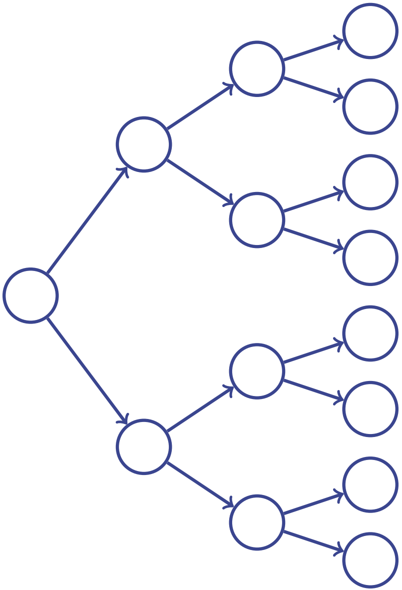
Theorem

The saddlepoints of $\bar{W}_t^k(x_t^k, y_t^k, \cdot, \cdot)$ and $\underline{W}_t^k(x_t^k, y_t^k, \cdot, \cdot)$ converge as $k \to \infty$.

Average-cut

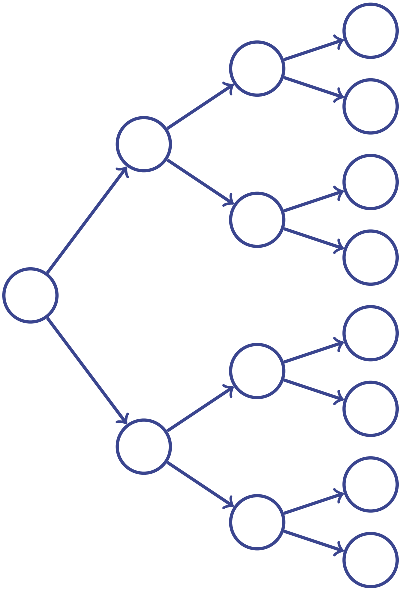
Multi-cut

Multi-cut + extra work →
Deterministic convergence



Average-cut
Multi-cut

Multi-cut + extra work →
Deterministic convergence

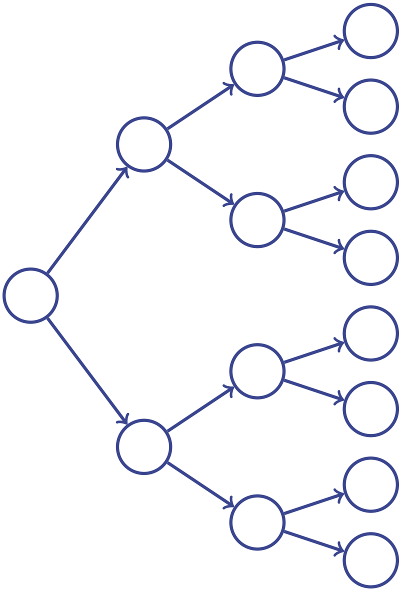


Average-cut

Multi-cut

Multi-cut + extra work →

Deterministic convergence

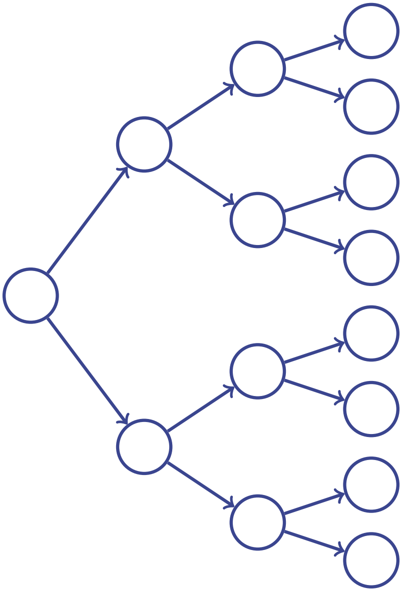


Average-cut

Multi-cut

Multi-cut + extra work →

Deterministic convergence

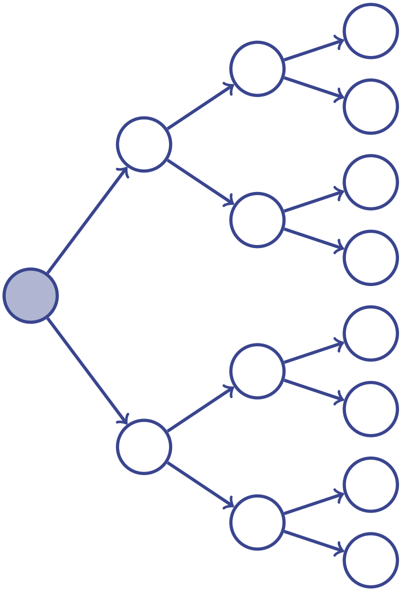


Average-cut

Multi-cut

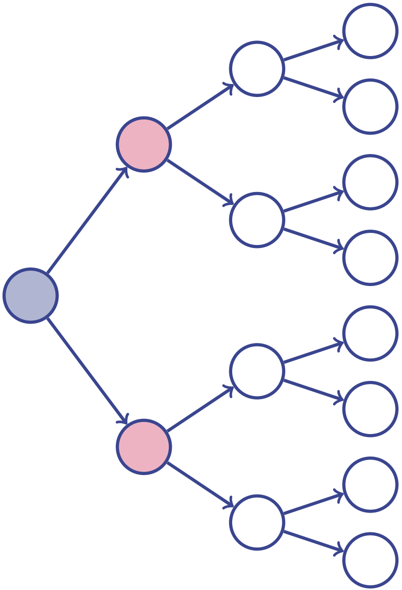
Multi-cut $+$ extra work $\rightarrow$

Deterministic convergence



Average-cut
Multi-cut

Multi-cut $+$ extra work $\rightarrow$
Deterministic convergence

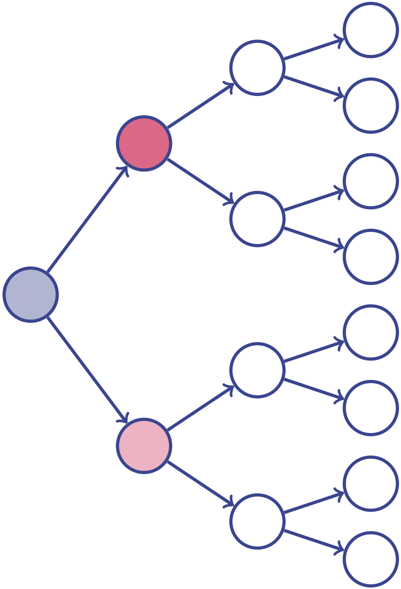


Average-cut

Multi-cut

Multi-cut $+$ extra work $\rightarrow$

Deterministic convergence

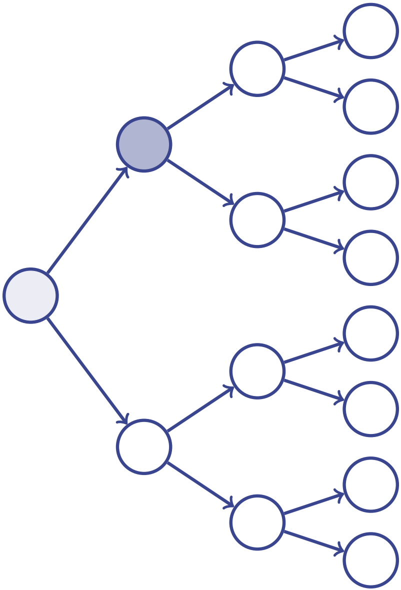


Average-cut

Multi-cut

Multi-cut + extra work →

Deterministic convergence

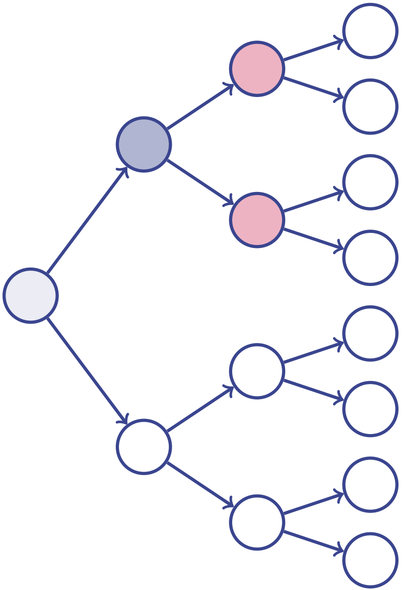


Average-cut

Multi-cut

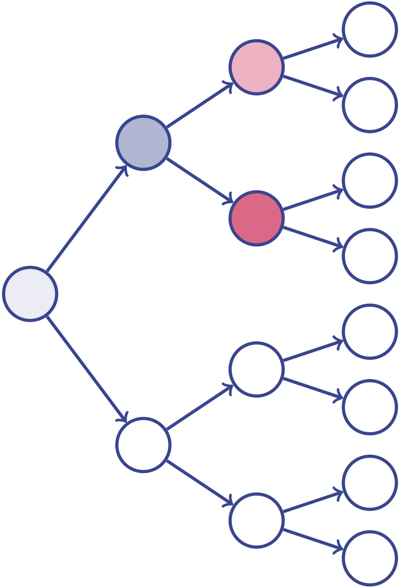
Multi-cut + extra work →

Deterministic convergence



Average-cut
Multi-cut

Multi-cut + extra work →
Deterministic convergence

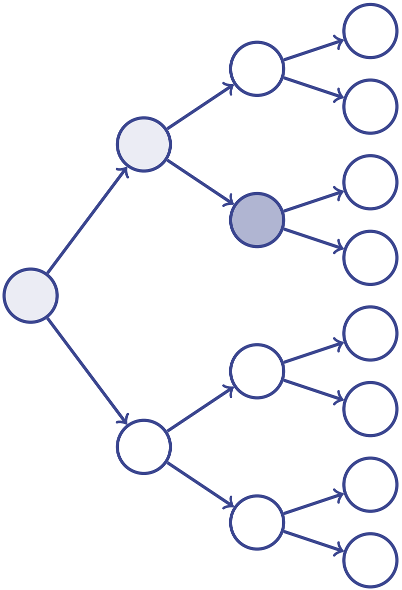


Average-cut

Multi-cut

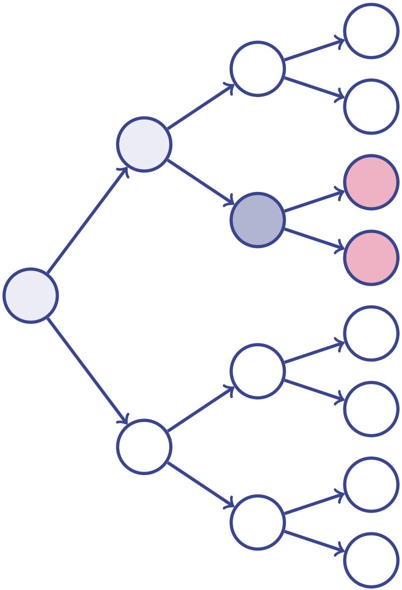
Multi-cut + extra work →

Deterministic convergence



Average-cut
Multi-cut

Multi-cut + extra work →
Deterministic convergence

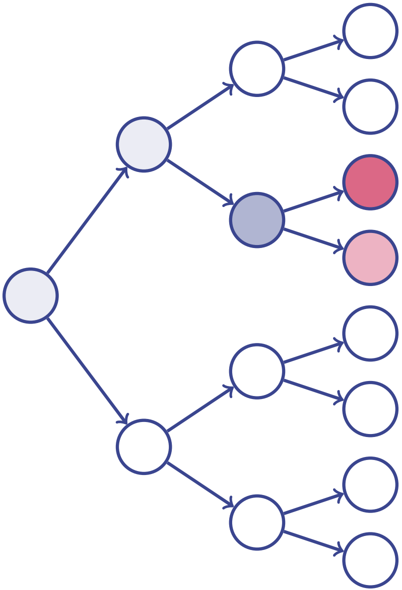


Average-cut

Multi-cut

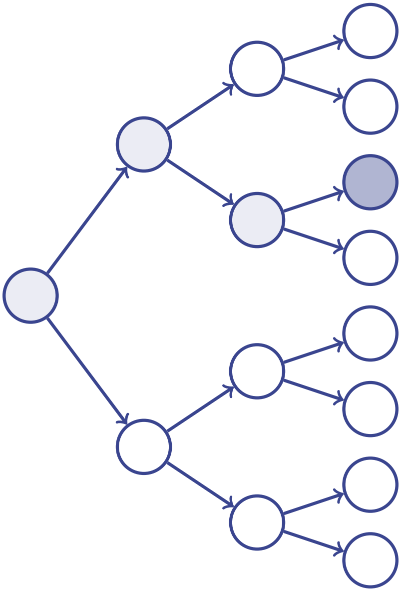
Multi-cut $+$ extra work $\rightarrow$

Deterministic convergence



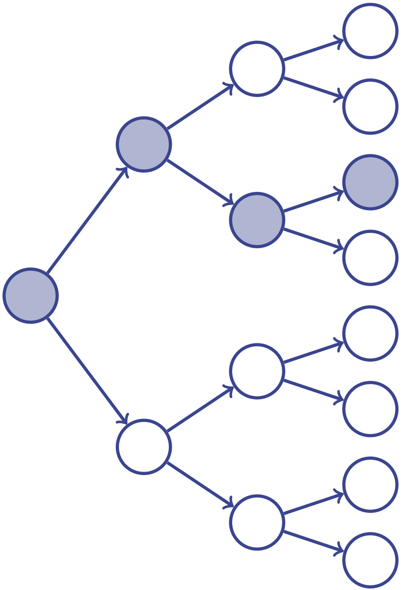
Average-cut
Multi-cut

Multi-cut + extra work →
Deterministic convergence



Average-cut
Multi-cut

Multi-cut + extra work →
Deterministic convergence

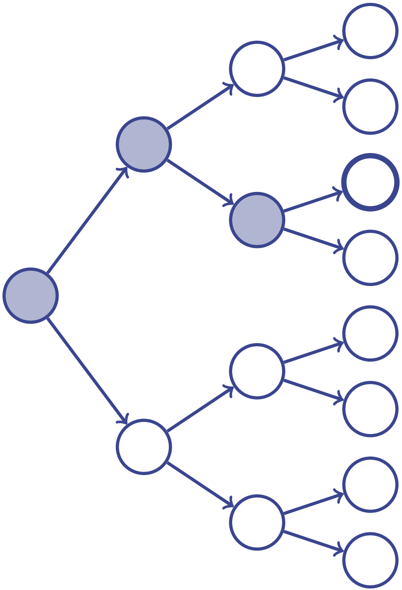


Average-cut

Multi-cut

Multi-cut + extra work →

Deterministic convergence

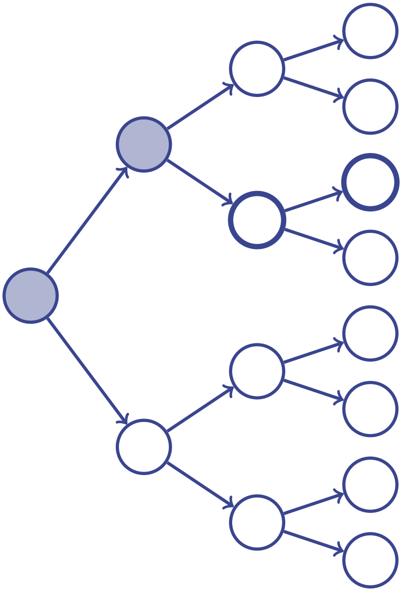


Average-cut

Multi-cut

Multi-cut + extra work →

Deterministic convergence

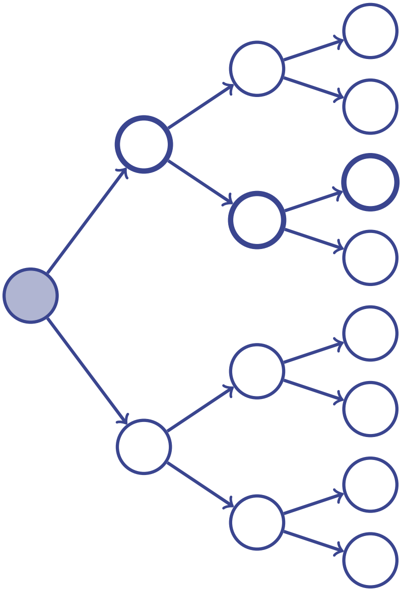


Average-cut

Multi-cut

Multi-cut + extra work →

Deterministic convergence

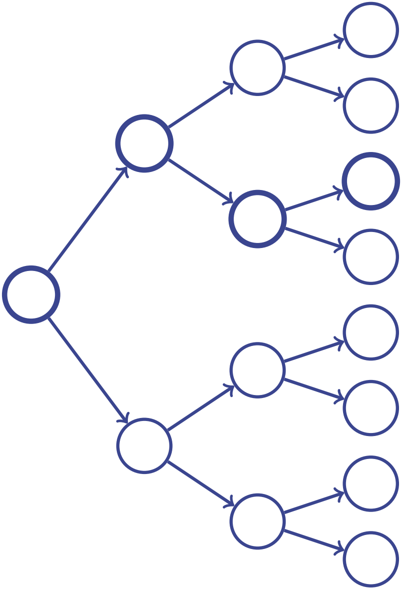


Average-cut

Multi-cut

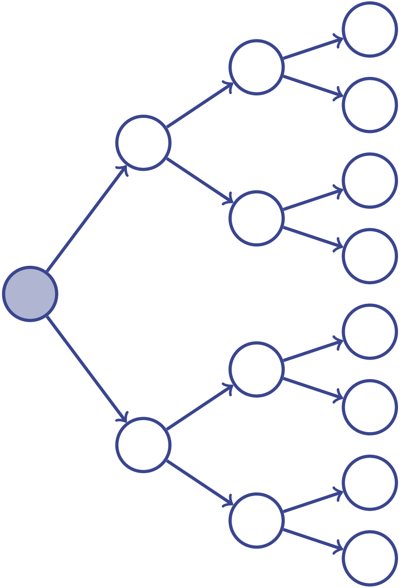
Multi-cut $+$ extra work $\rightarrow$

Deterministic convergence



Average-cut
Multi-cut

Multi-cut + extra work →
Deterministic convergence

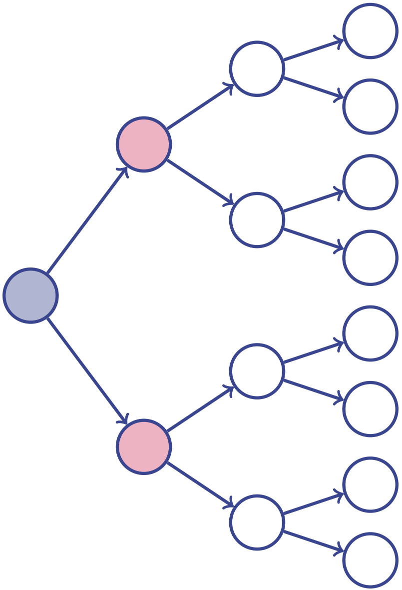


Average-cut

Multi-cut

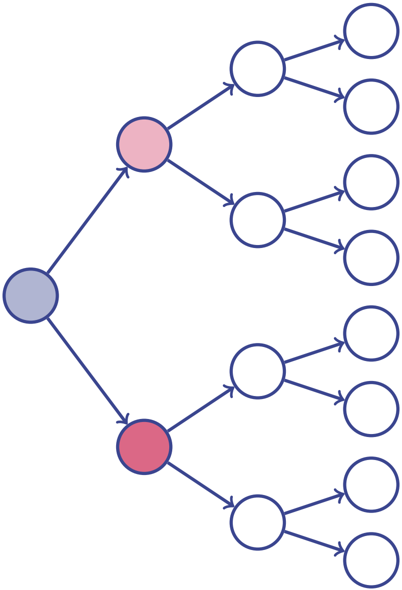
Multi-cut + extra work →

Deterministic convergence



Average-cut
Multi-cut

Multi-cut + extra work →
Deterministic convergence

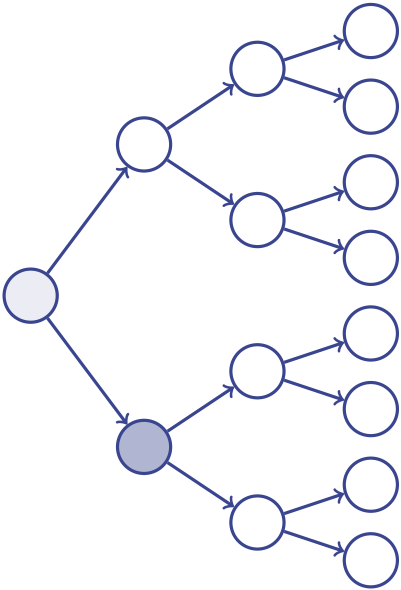


Average-cut

Multi-cut

Multi-cut + extra work →

Deterministic convergence

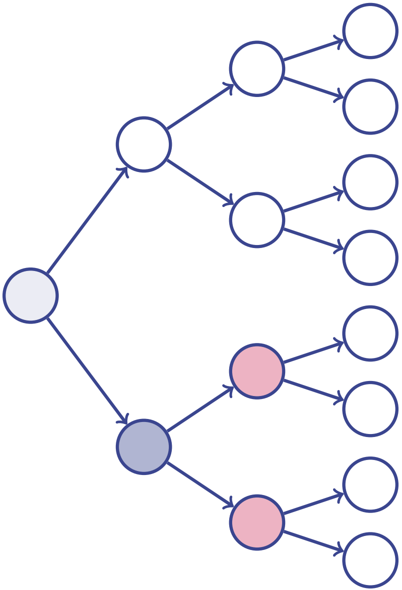


Average-cut

Multi-cut

Multi-cut + extra work →

Deterministic convergence

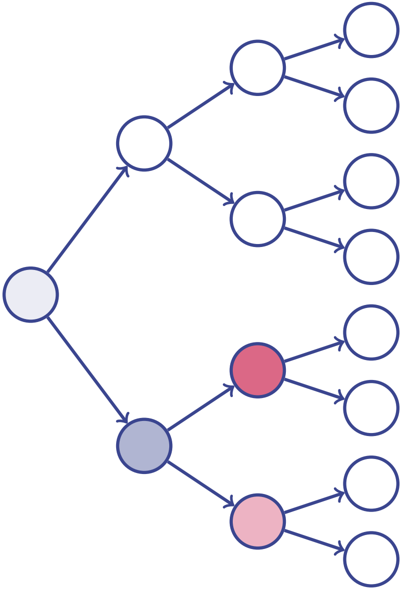


Average-cut

Multi-cut

Multi-cut + extra work →

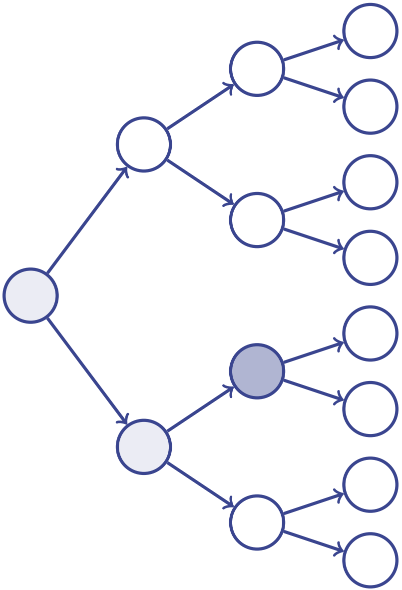
Deterministic convergence



Average-cut

Multi-cut

Multi-cut + extra work →
Deterministic convergence

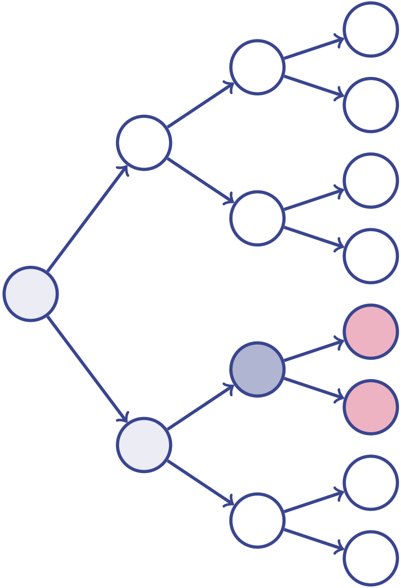


Average-cut

Multi-cut

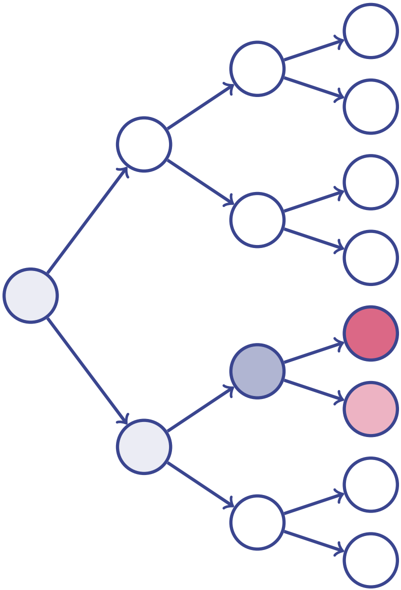
Multi-cut $+$ extra work $\rightarrow$

Deterministic convergence



Average-cut
Multi-cut

Multi-cut + extra work →
Deterministic convergence

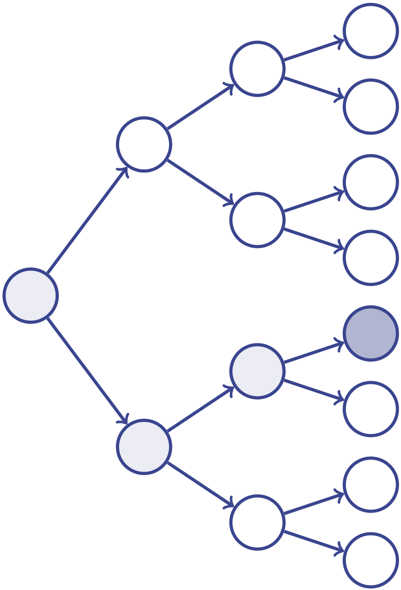


Average-cut

Multi-cut

Multi-cut + extra work →

Deterministic convergence

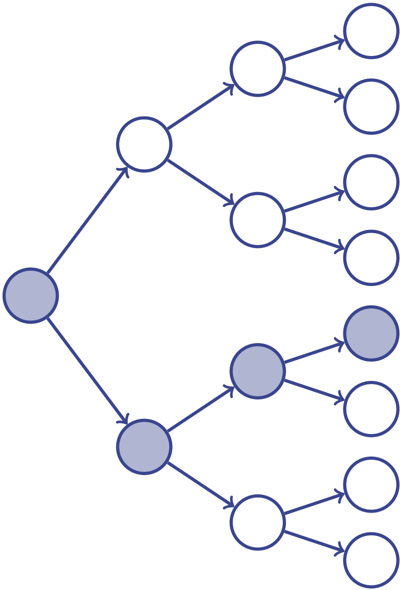


Average-cut

Multi-cut

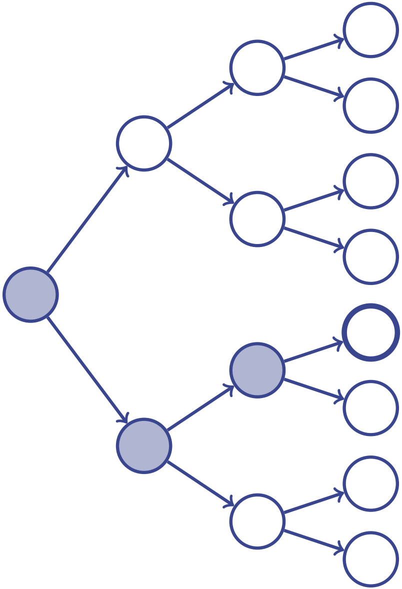
Multi-cut + extra work →

Deterministic convergence



Average-cut
Multi-cut

Multi-cut + extra work →
Deterministic convergence

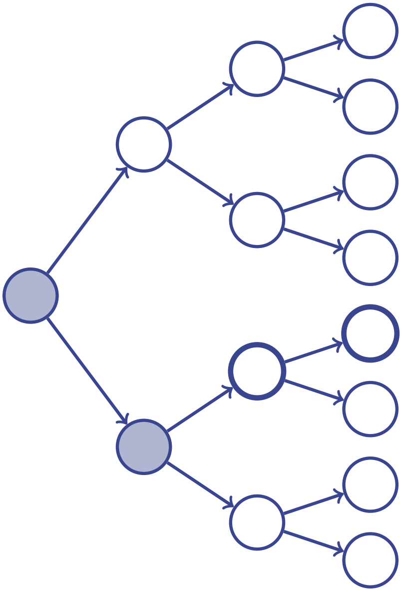


Average-cut

Multi-cut

Multi-cut + extra work →

Deterministic convergence

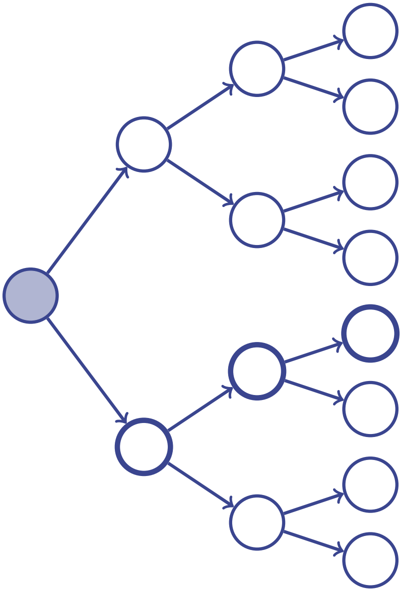


Average-cut

Multi-cut

Multi-cut + extra work →

Deterministic convergence

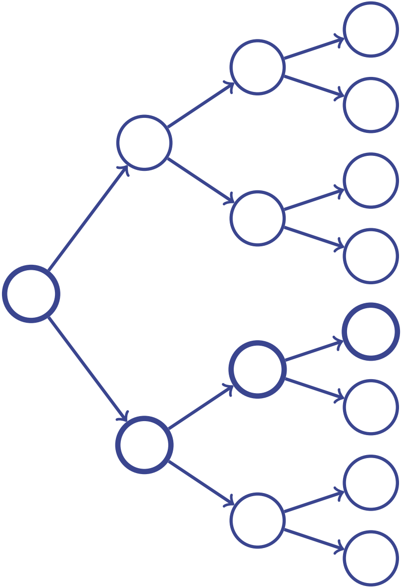


Average-cut

Multi-cut

Multi-cut $+$ extra work $\rightarrow$

Deterministic convergence



Portfolio optimisation problem:

3 securities

12 stages

6 children nodes

tree of 2.6 billion nodes

Portfolio optimisation problem:

3 securities

12 stages

6 children nodes

tree of 2.6 billion nodes

independence

Portfolio optimisation problem:

3 securities

12 stages

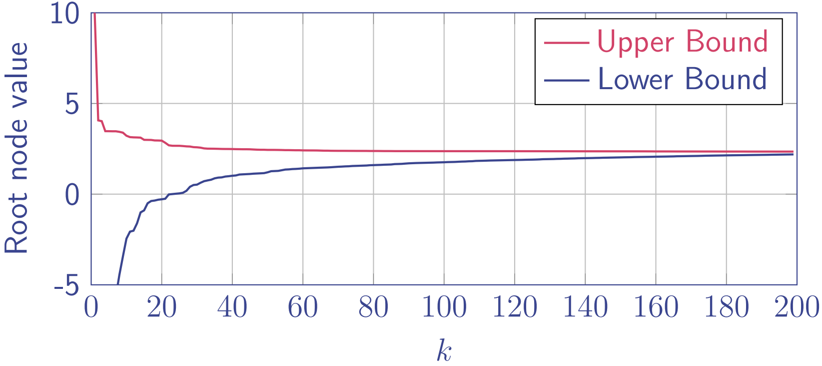
6 children nodes

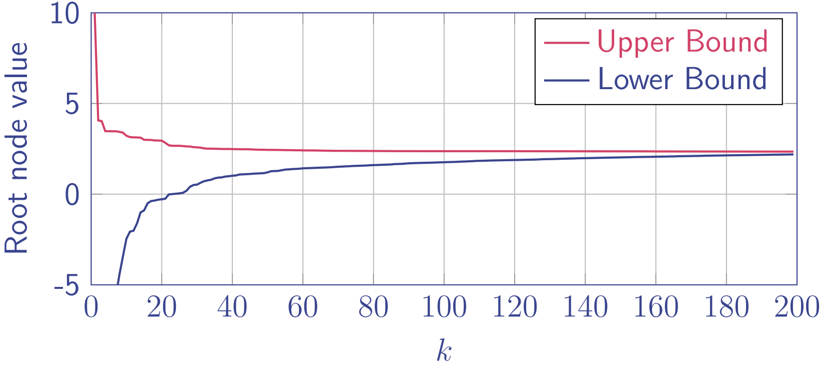
tree of 2.6 billion nodes

independence

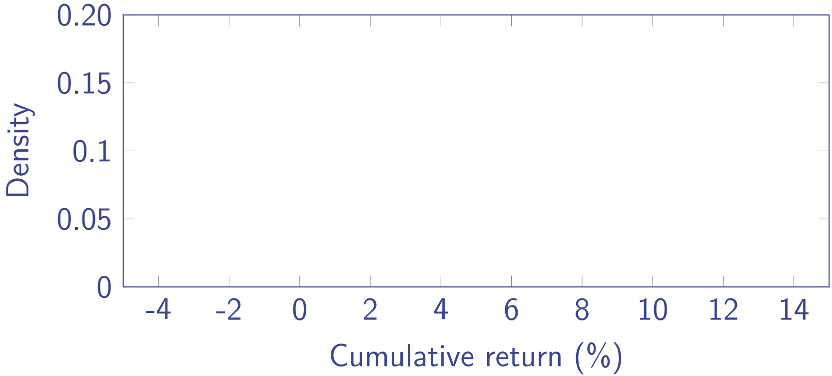
Maximising the risk adjusted cumulative return

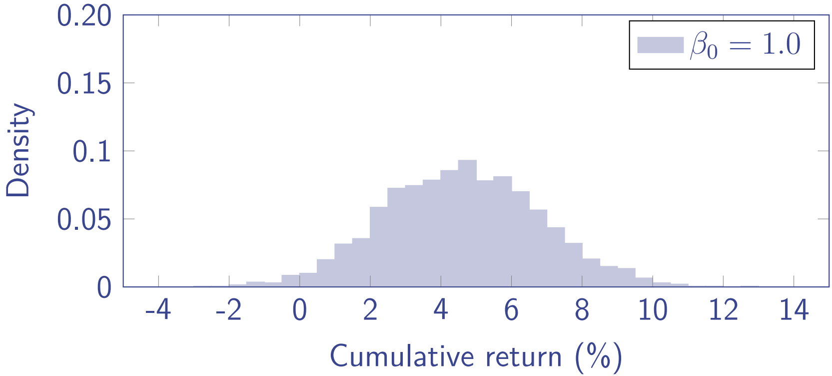
using CVaR_β


Upper Bound
Lower Bound
Root node value
10
5
0
-5
0
20
40
60
80
100
120
140
160
180
200
k

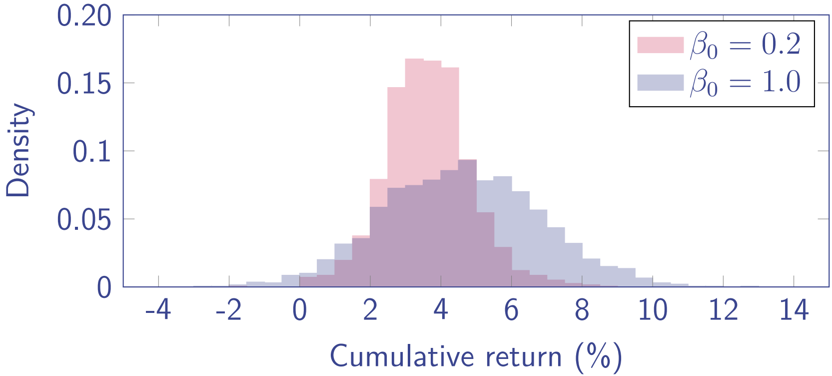


3 minutes 47 seconds


0.20
0.15
0.1
0.05
0
Density
-4
-2
0
2
4
6
8
10
12
14
Cumulative return (%)



0.20
0.15
0.1
0.05
0
Density
-4
-2
0
2
4
6
8
10
12
14
Cumulative return (%)
$\beta_0 = 1.0$


0.20
0.15
0.1
0.05
0
Density
-4
-2
0
2
4
6
8
10
12
14
Cumulative return (%)
$\beta_0 = 0.2$
$\beta_0 = 1.0$

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	1.52%	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	2.68%	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	4.09%

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	1.52%	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	2.68%	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	4.09%

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	1.52%	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	2.68%	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	4.09%

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	1.52%	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	2.68%	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	4.09%

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	1.52%	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	2.68%	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	4.09%

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	1.52%	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	2.68%	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	4.09%

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	1.52%	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	2.68%	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	4.09%

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	1.52%	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	2.68%	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	4.09%

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	1.52%	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	2.68%	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	4.09%

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	**1.52%**	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	2.68%	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	4.09%

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	**1.52%**	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	2.68%	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	4.09%

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	**1.52%**	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	**2.68%**	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	4.09%

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	**1.52%**	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	**2.68%**	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	4.09%

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	**1.52%**	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	**2.68%**	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	**4.09%**

Stochastic

$$
\begin{aligned}
G_n(x_n, y_n) = \min_{u_n} \max_{v_n} \quad & g_n(x_n, y_n, u_n, v_n) + \mathbb{E}[G_m(x_m, y_m)] \\
\text{s.t.} \quad & (x_n, y_n) \in \mathcal{X}_n \times \mathcal{Y}_n, \\
& (u_n, v_n) \in \mathcal{U}_n(x_n) \times \mathcal{V}_n(y_n), \\
& x_m = f^x_m(x_n, u_n),\ \forall m \in C(n), \\
& y_m = f^y_m(y_n, v_n),\ \forall m \in C(n).
\end{aligned}
$$

Stochastic

$$
\begin{aligned}
G_n(x_n, y_n) = \min_{u_n} \max_{v_n} \quad & g_n(x_n, y_n, u_n, v_n) + \mathbb{E}[G_m(x_m, y_m)] \\
\text{s.t.} \quad & (x_n, y_n) \in \mathcal{X}_n \times \mathcal{Y}_n, \\
& (u_n, v_n) \in \mathcal{U}_n(x_n) \times \mathcal{V}_n(y_n), \\
& x_m = f_m^x(x_n, u_n), \ \forall m \in C(n), \\
& y_m = f_m^y(y_n, v_n), \ \forall m \in C(n).
\end{aligned}
$$

Stochastic

$$
\begin{aligned}
G_n(x_n, y_n) = \min_{u_n} \max_{v_n} \quad & g_n(x_n, y_n, u_n, v_n) + \mathbb{E}[G_m(x_m, y_m)] \\
\text{s.t.} \quad & (x_n, y_n) \in \mathcal{X}_n \times \mathcal{Y}_n, \\
& (u_n, v_n) \in \mathcal{U}_n(x_n) \times \mathcal{V}_n(y_n), \\
& x_m = f_m^x(x_n, u_n),\ \forall m \in C(n), \\
& y_m = f_m^y(y_n, v_n),\ \forall m \in C(n).
\end{aligned}
$$

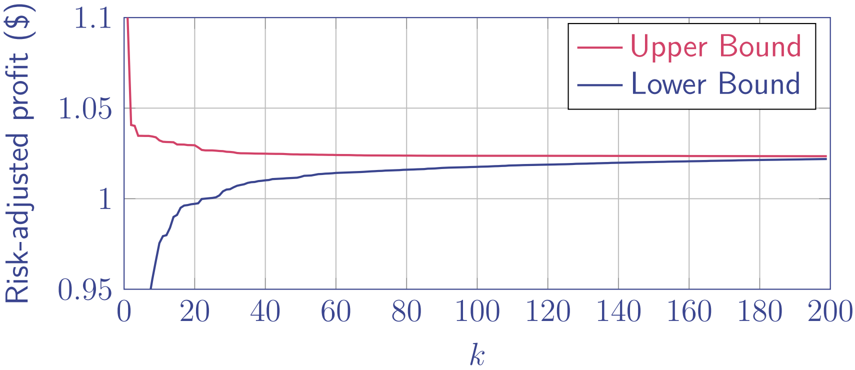
Downward, A., Dowson, O., and Baucke, R. (2018). Stochastic dual dynamic programming with stagewise dependent objective uncertainty. *www.optimization-online.org*

Conclusions

- Non-rectangularity,
- bounding saddle functions,
- iterative (deterministic) algorithm,
- applications in game-theory, budget constrained robust optimisation, and many more!

[1] A. Downward, O. Dowson, and R. Baucke. Stochastic dual dynamic programming with stagewise dependent objective uncertainty. *www.optimization-online.org*, 2018.

[2] R. Baucke, A. Downward, and G. Zakeri. A deterministic algorithm for solving stochastic minimax dynamic programmes. *www.optimization-online.org*, 2018.



3 minutes 47 seconds

CVaR_β Estimation	Optimal Policy		
	$\beta_0 = 0.2$	$\beta_0 = 0.6$	$\beta_0 = 1.0$
$\beta_0 = 0.2$	**1.52%**	1.30%	1.05%
$\beta_0 = 0.6$	2.38%	**2.68%**	2.66%
$\beta_0 = 1.0$	3.14%	3.73%	**4.09%**