Global Optimal Feedbacks for Stochastic Quantized Nonlinear Event Systems

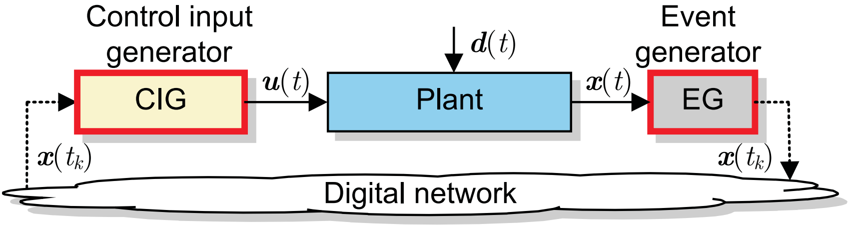
Oliver Junge

Fakultät für Mathematik
Technische Universität München

joint work with M. Post and S. Jerg

supported by the Priority Programme SPP 1305 of the DFG

Networked control systems



restricted communication:

- delays
- dropouts
-

The problem

- time-discrete control system

$$x_{k+1} = f(x_k, u_k), \quad k = 0, 1, \ldots,$$

- $+$ cost function

$$g(x, u) \geq 0, \quad g(x, u) = 0 \Leftrightarrow x = 0$$

- Goal: optimal feedback $u_k = u(x_k)$, s.t. for

$$x_{k+1} = f(x_k, u(x_k))$$

we get $x_k \to 0$ for all x_0.

The value function

- accumulated cost

$$J(x_0, (u_k)) = \sum_{k=0}^{\infty} g(x_k, u_k)$$

- value function

$$V(x) = \inf_{(u_k)} \{J(x, (u_k)) \mid x_k \to 0\}$$

The optimality principle

... and the Bellman operator

- optimality principle

$$V(x) = \inf_{u \in U} \{g(x,u) + V(f(x,u))\}$$

- Bellman-Operator

$$L[v](x) := \inf_{u \in U} \{g(x,u) + v(f(x,u))\}$$

- V is the unique fixed point of L with $V(0) = 0$.

Discretization

The discretized Bellman operator

- Partition $\mathcal{P}$ of phase space
- $\rho(x) = P \in \mathcal{P}$ with $x \in P$
- projection onto piecewise constant functions

$$\varphi[v](x) := \inf_{x' \in \rho(x)} v(x')$$

- discretized Bellman operator

$$L_{\mathcal{P}} := \varphi \circ L$$

Discretization

The discrete optimality principle

$$
\begin{aligned}
L_{\mathcal{P}}[v](x) &= \varphi\left(L[v](x)\right) \\
&= \inf_{x' \in \rho(x)} \left\{ \inf_{u \in U} \left\{ g(x', u) + v(f(x', u)) \right\} \right\} \\
&= \min_{P' \cap f(P,U) \neq \emptyset} \left\{ \underbrace{\inf_{x', u : f(x,u) \in P'} g(x', u)} + v(P') \right\} \\
&= \min_{P' \cap f(P,U) \neq \emptyset} \left\{ w(P, P') + v(P') \right\}
\end{aligned}
$$

- discrete optimality principle = shortest path problem

$$
V(P) = \min_{P' \cap f(P,U) \neq \emptyset} \{ w(P, P') + V(P') \}
$$

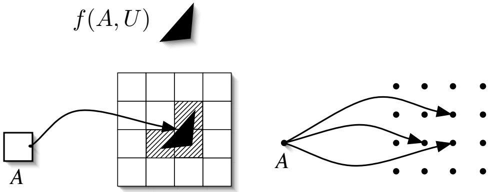
cf. [Tsitsiklis, 95], [Sethian, 96]

Discretization

The discrete graph

- graph $(\mathcal{P}, E)$ with

$$E = \{(A,B) \in \mathcal{P} \times \mathcal{P} \mid f(A,U) \cap B \neq \emptyset\}$$



- and weights

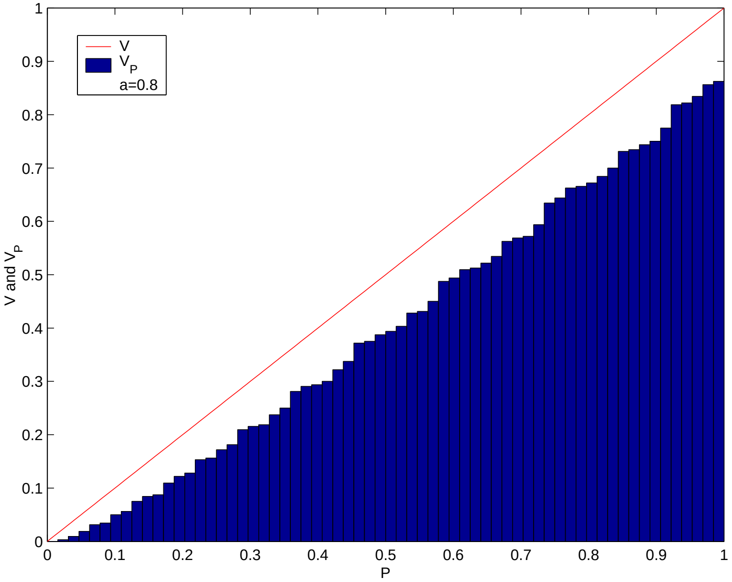
$$w(A,B) = \min_{x \in A, u \in U} \{g(x,u) \mid f(x,u) \in B\}$$

A simple example

- $$x_{k+1} = (1 + a u_k) x_k,$$
  $x_k \in X = [0, 1]$,
  $u_k \in U = [-1, 1]$ and
  $a \in (0, 1)$ fixed.
- cost function
  $$g(x, u) = ax.$$
- optimal control sequence: $\mathbf{u} = (-1, -1, \ldots)$.
- value function: $V(x) = x$.

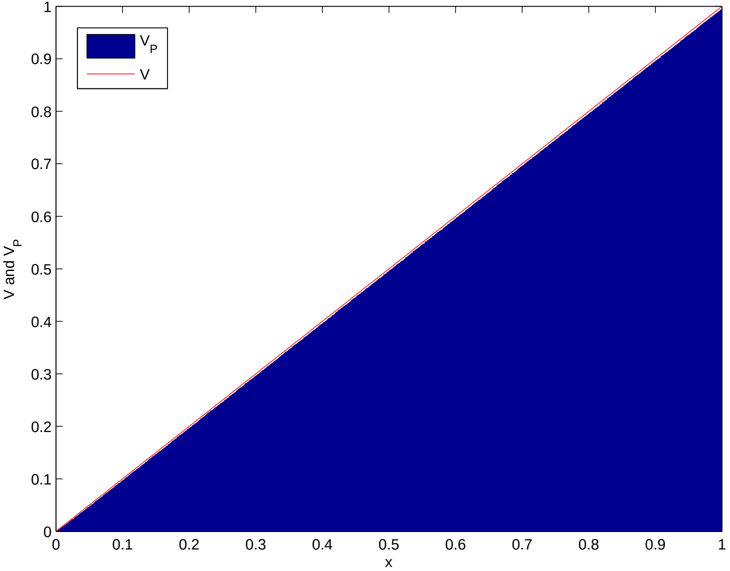
A simple example

Approximate value function: 64 cells



A simple example

Approximate value function: 4096 cells



Convergence

- $V_{\mathcal{P}}$: approximation to V on partition $\mathcal{P}$
- $V_{\mathcal{P}}(x) \leq V(x)$ for all x
- $\mathcal{P}^\ell$: sequence of *nested* partitions with *diam* $\mathcal{P}^\ell \to 0$

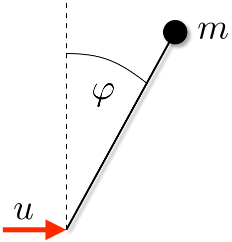
Thm [J., Osinga, 04]

$$V_{\mathcal{P}^\ell}(x) \to V(x) \qquad \text{as } \ell \to \infty$$

on the stabilizable set.

Example

An inverted pendulum

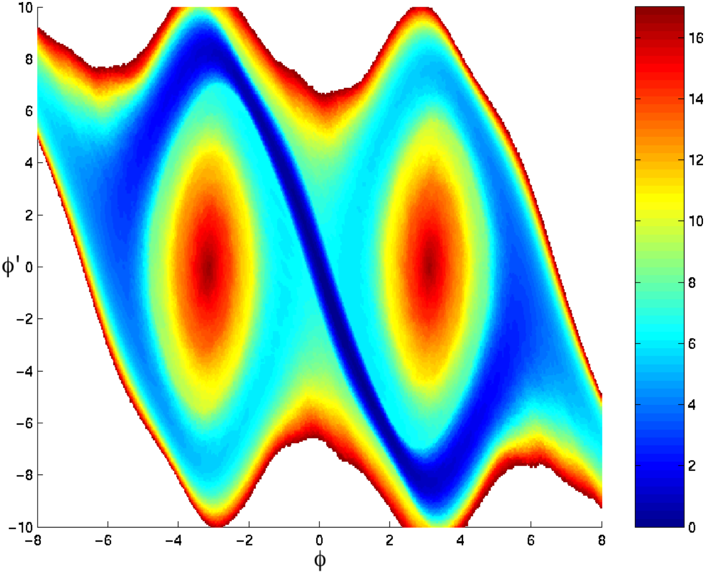


- state: $x = (\varphi, \dot{\varphi})$
- control system: $f(x, u) = \Phi^T(x, u)$
- cost function

$$g(x, u) = \int_0^T q_1 \varphi^2(t) + q_2 \dot{\varphi}^2(t)\, dt + T q_3 u^2$$

Example

The approximate value function



Optimal feedback

- optimality principle

$$V(x) = \inf_{u \in U} \{g(x,u) + V(f(x,u))\}$$

- optimal feedback

$$u(x) = \operatorname*{argmin}_{u \in U} \{g(x,u) + V(f(x,u))\}$$

Optimal feedback

- optimality principle

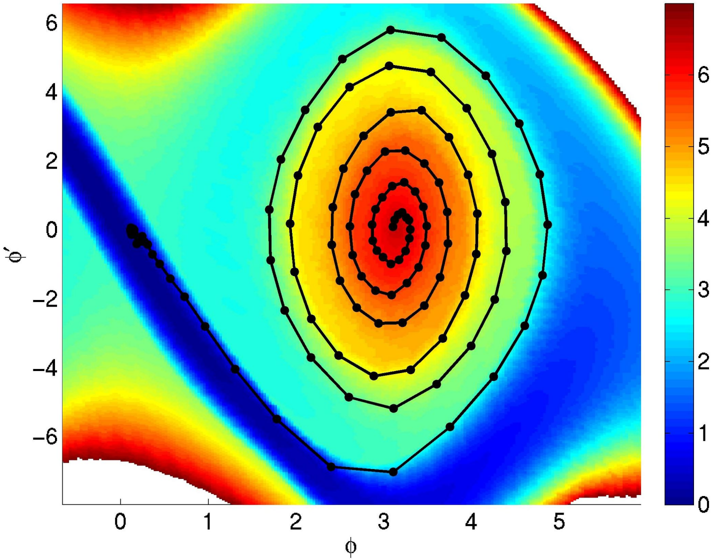
$$V(x) = \inf_{u \in U} \{g(x, u) + V(f(x, u))\}$$

- approximate optimal feedback

$$u_{\mathcal{P}}(x) = \operatorname*{argmin}_{u \in U} \{g(x, u) + V_{\mathcal{P}}(f(x, u))\}$$

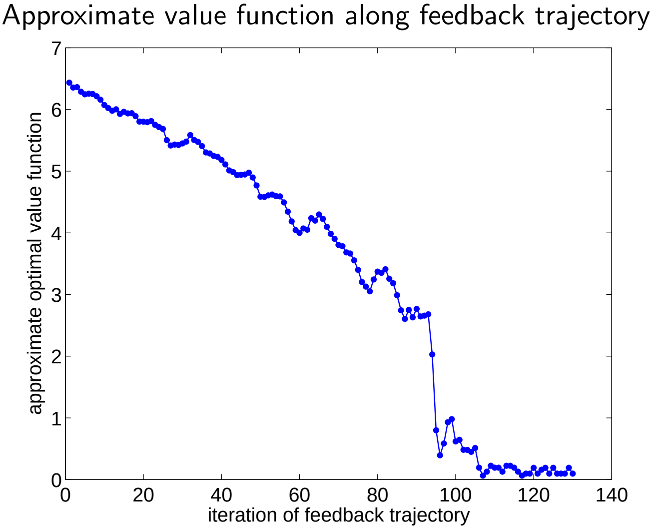
Optimal feedback

... for the inverted pendulum



Optimal feedback

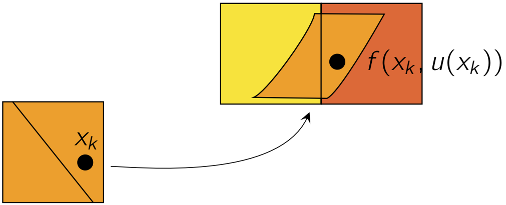
... for the inverted pendulum



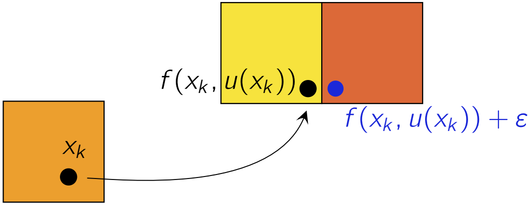
Optimal feedback

Artefacts of the discretization

- problem 1: boxes too large



- Problem 2: feedback is not robust



Incorporating perturbations

dynamic game

$$x_{k+1} = f(x_k, u_k, w_k), \quad k = 0, 1, \ldots.$$

+ cost function

Goal

construct feedback $u_k = u(x_k)$, s.t.

$$x_k \to B_\varepsilon(0) \quad \text{for all sequences } (w_k).$$

Game theoretic viewpoint

- controller u_k wants to minimize J,
- perturber w_k wants to maximize J.
- upper value function

$$V(x) = \inf_{u(x)} \sup_{(w_k)} J(x, u(x), (w_k))$$

- optimality principle

$$V(x) = \inf_{u \in U} \left\{ g(x, u) + \sup_{w \in W} V(f(x, u, w)) \right\}$$

- feedback

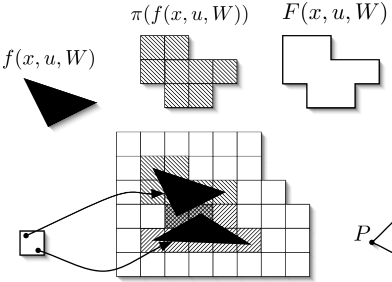
$$u(x) = \operatorname*{argmin}_{u \in U} \left\{ g(x, u) + \sup_{w \in W} V(f(x, u, w)) \right\}$$

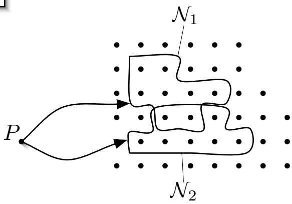
Discretization

- discrete optimality principle

$$V_{\mathcal{P}}(P) = \inf_{\mathcal{N} \in \mathcal{F}(P)} \left\{ \mathcal{G}(P, \mathcal{N}) + \sup_{N \in \mathcal{N}} V_{\mathcal{P}}(N) \right\}$$

- data structure: directed weighted hypergraph





The discretization is the perturber

- we want $V_{\mathcal{P}}$ to decrease monotonically
- $\rightsquigarrow$ need to jump within current box
- idea: view jump as perturbation
- $\rightsquigarrow$ put

  $$\hat{f}(x, u, w) = f(w(x), u),$$

  w choice function, chooses arbitrary point from current box

  $$\hat{g}(x, u) = \sup_w g(w(x), u).$$

The discretization is the perturber

- Thm [Grüne, J., 07]:

$$V(x) \leq V_{\mathcal{P}}(x)$$

  and

$$V_{\mathcal{P}}(x) \geq \min_{u \in U} \{\hat{g}(x,u) + V_{\mathcal{P}}(\hat{f}(x,u))\},$$

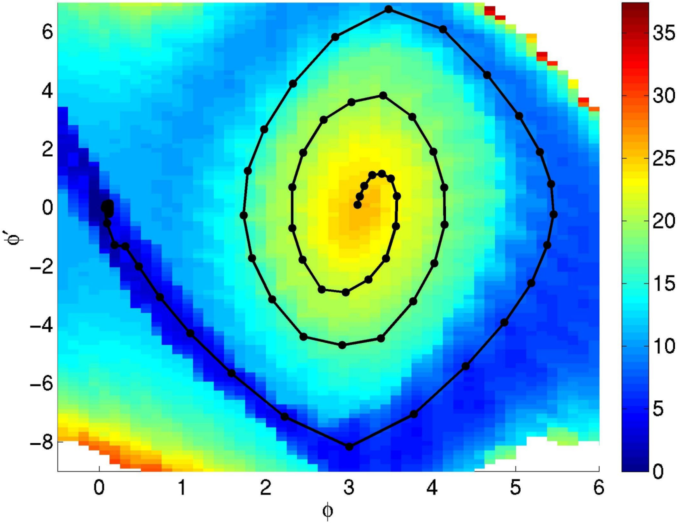
  i.e. $V_{\mathcal{P}}$ is a *Lyapunov function* for the closed loop system.
- feedback

$$u_{\mathcal{P}}(x) = \operatorname*{argmin}_{u \in U} \left\{ \hat{g}(x,u) + \sup_{x' \in \hat{f}(\rho(x),u)} V_{\mathcal{P}}(x') \right\}.$$

- constant on each box!
- $\rightsquigarrow$ constructed offline

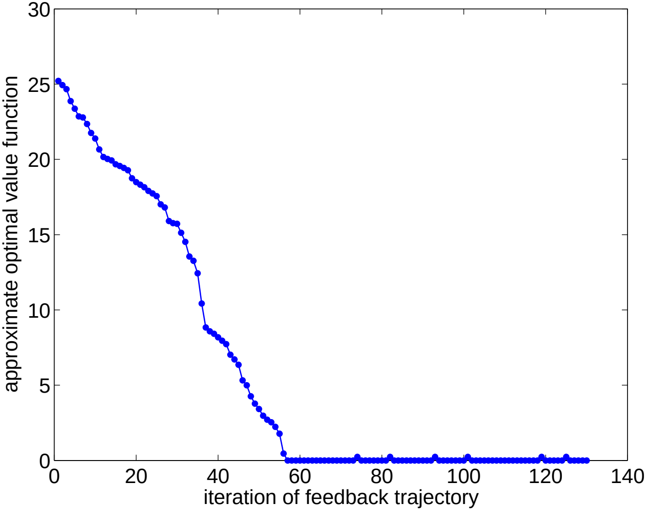
Example: inverted pendulum

Robust feedback trajectory



Example: inverted pendulum

Value function along feedback trajectory



Event systems

event function

$$r : X \times U \to \mathbb{N}$$

event system

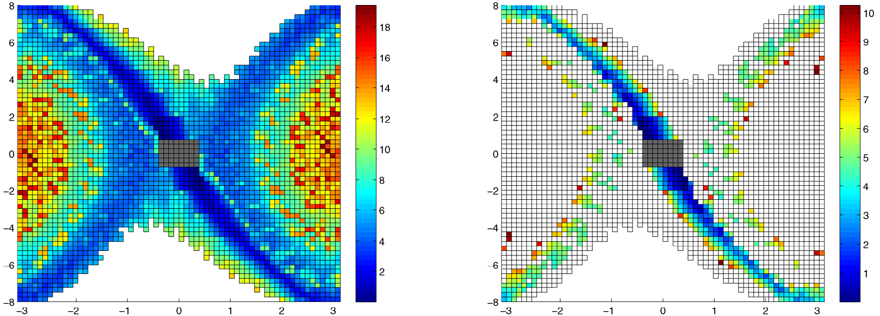
$$x_{\ell+1} = f^r(x_\ell, u_\ell), \quad r = r(x_\ell, u_\ell),$$

cost function

$$\tilde{c}(x_\ell, u_\ell) = \sum_{k=0}^{r(x,u)-1} c(x_{\ell+k}, u_{\ell+k})$$

Effect of communication delays

inverted pendulum: stabilizable set for quantized feedback

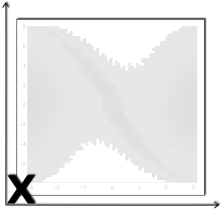


without delays

with delays

Extended state space

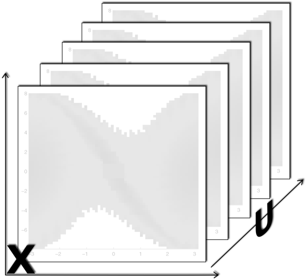
- Standard state space: X



Extended state space

- Extended state space: $Z = X \times U$
- New state: $z = (x, w)$

  x: current state
  w: last control

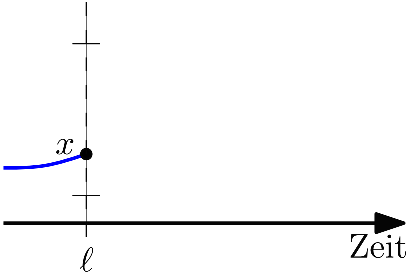


Modelling delays

system

$$z_{\ell+1} = g(z_\ell, u_\ell, \delta_\ell), \quad \ell = 0, 1, 2, \ldots$$

- random delays: $\delta_\ell \in \mathbb{N}_\infty \sim \pi$, i.i.d.
- $g(z, u, \delta) = \begin{bmatrix} f^s(f^t(z), u) \\ w' \end{bmatrix}$
- $t = \min\{\delta, r(z)\}$, $\quad r$: event time

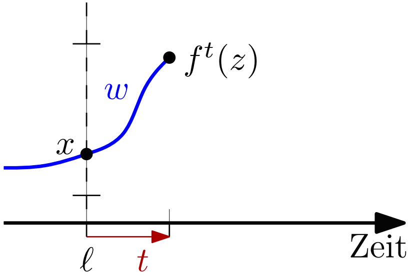


Modelling delays

system

$$z_{\ell+1} = g(z_\ell, u_\ell, \delta_\ell), \quad \ell = 0, 1, 2, \ldots$$

- random delays: $\delta_\ell \in \mathbb{N}_\infty \sim \pi$, i.i.d.
- $g(z, u, \delta) = \begin{bmatrix} f^s(f^t(z), u) \\ w' \end{bmatrix}$
- $t = \min\{\delta, r(z)\}$, $\quad r$: event time

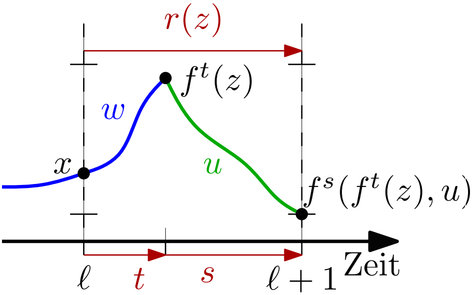


Modelling delays

system

$$z_{\ell+1} = g(z_\ell, u_\ell, \delta_\ell), \quad \ell = 0, 1, 2, \dots$$

- random delays: $\delta_\ell \in \mathbb{N}_\infty \sim \pi$, i.i.d.
- $g(z, u, \delta) = \begin{bmatrix} f^s(f^t(z), u) \\ w' \end{bmatrix}$
- $t = \min\{\delta, r(z)\}, \quad r:$ event time

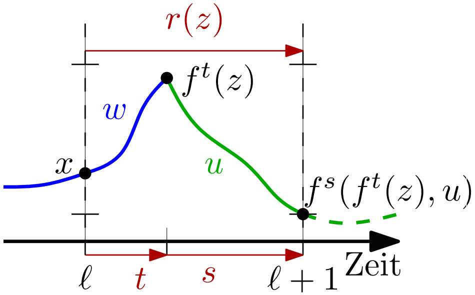


Modelling delays

system

$$z_{\ell+1} = g(z_\ell, u_\ell, \delta_\ell), \quad \ell = 0, 1, 2, \ldots$$

- random delays: $\delta_\ell \in \mathbb{N}_\infty \sim \pi$, i.i.d.
- $g(z, u, \delta) = \begin{bmatrix} f^s(f^t(z), u) \\ w' \end{bmatrix}$
- $t = \min\{\delta, r(z)\}, \quad r:$ event time



Modelling delays

system

$$z_{\ell+1} = g(z_\ell, u_\ell, \delta_\ell), \quad \ell = 0, 1, 2, \dots$$

- random delays: $\delta_\ell \in \mathbb{N}_\infty \sim \pi$, i.i.d.
- $g(z, u, \delta) = \begin{bmatrix} f^s(f^t(z), u) \\ w' \end{bmatrix}$
- $t = \min\{\delta, r(z)\}, \quad r:$ event time

If $\delta < r(z)$:

- remaining time: $s = r(f^t(z), u)$
- new control history: $w' = u$
- current control $u(z, \delta) = u(z)$

Otherwise:

- event gets lost $\rightarrow$ keep old control
- $s = 0,\ w' = w,\ u(z, \delta) = w$

Extension of the method to stochastic systems

Stochastic optimality principle:

$$V(P) = inf_{u \in U} \left\{ \tilde{c}(P,u) + \sup_w \mathbb{E} \left\{ V(F(P,u,w,\delta)) \right\} \right\}$$

optimal feedback for P *(constant w.r.t. δ)* :

$$K(P) = \underset{u}{\operatorname{argmin}} \left\{ \tilde{c}(P,u) + \sup_w \mathbb{E} \left\{ V(F(P,u,w,\delta)) \right\} \right\}$$

Numerical solution:

- discretization: extended hypergraph
- stochastic shortest path problem: value iteration

Extension of the method to stochastic systems

Stochastic optimality principle:

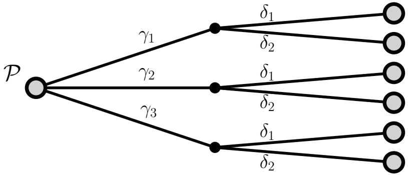
$$V(P) = inf_{u \in U} \left\{ \tilde{c}(P,u) + \sup_{w} \mathbb{E}\left\{ V(F(P,u,w,\delta)) \right\} \right\}$$

optimal feedback for P *(constant w.r.t. δ)* :

$$K(P) = \operatorname*{argmin}_{u} \left\{ \tilde{c}(P,u) + \sup_{w} \mathbb{E}\left\{ V(F(P,u,w,\delta)) \right\} \right\}$$

Numerical solution:

- discretization: extended hypergraph
- stochastic shortest path problem: value iteration



Extension of the method to stochastic systems

Stochastic optimality principle:

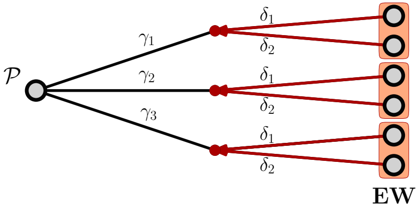
$$V(P) = inf_{u \in U} \Big\{ \tilde{c}(P,u) + \sup_{w} \mathbb{E}\left\{ V(F(P,u,w,\delta)) \right\} \Big\}$$

optimal feedback for P *(constant w.r.t. δ)* :

$$K(P) = \operatorname*{argmin}_{u} \Big\{ \tilde{c}(P,u) + \sup_{w} \mathbb{E}\left\{ V(F(P,u,w,\delta)) \right\} \Big\}$$

Numerical solution:

- discretization: extended hypergraph
- stochastic shortest path problem: value iteration



Extension of the method to stochastic systems

Stochastic optimality principle:

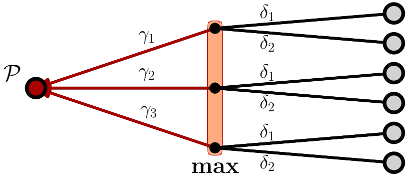
$$V(P) = inf_{u \in U} \Big\{ \tilde{c}(P,u) + \sup_w \mathbb{E} \{ V(F(P,u,w,\delta)) \} \Big\}$$

optimal feedback for P *(constant w.r.t. δ)* :

$$K(P) = \operatorname*{argmin}_u \Big\{ \tilde{c}(P,u) + \sup_w \mathbb{E} \{ V(F(P,u,w,\delta)) \} \Big\}$$

Numerical solution:

- discretization: extended hypergraph
- stochastic shortest path problem: value iteration



Extension of the method to stochastic systems

Stochastic optimality principle:

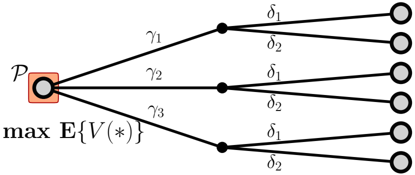
$$V(P) = inf_{u\in U}\left\{\tilde{c}(P,u) + \sup_w \mathbb{E}\left\{V(F(P,u,w,\delta))\right\}\right\}$$

optimal feedback for P *(constant w.r.t. δ)* :

$$K(P) = \operatorname*{argmin}_u \left\{\tilde{c}(P,u) + \sup_w \mathbb{E}\left\{V(F(P,u,w,\delta))\right\}\right\}$$

Numerical solution:

- discretization: extended hypergraph
- stochastic shortest path problem: value iteration



Analysis

closed system with new feedback

$$z_{\ell+1} = g(z_\ell, K(z_\ell), \delta_\ell), \quad \ell = 0, 1, 2, \ldots,$$

λ level set: $\mathcal{S}_\lambda = \{z \in \mathcal{Z} : V([z]) < \lambda\}$, $\quad \lambda > 0$

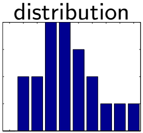
Theorem (Jerg, Junge, Post, 09)

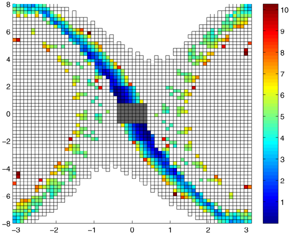
- *Every random trajectory with initial value $z(0) \in \mathcal{S}_\lambda$ remains in $\mathcal{S}_\lambda$ with probability $1 - V(z(0))/\lambda$.*
- *Almost all trajectories $z_\ell = (x_\ell, w_\ell)$, $\ell = 0, 1, 2, \ldots,$ which remain in $\mathcal{S}_\lambda$ satisfy $x_\ell \to \Omega$.*

Numerical example

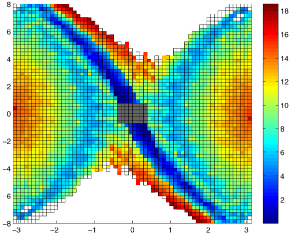
Inverted pendulum: stabilizable set

- 100 random delay sequences
  $\delta \in \{10, 20, \ldots, 90\}$*ms*
- 25 initial values per box





standard feedback
$<$ 25% *stabilizable*



new feedback
$>$ 95% *stabilizable*

Conclusion/Outlook

- piecewise constant value function
- efficient shortest path algorithm
- robustified by game theoretic approach
- unified treatment of hybrid systems
- adaptive partitioning possible

Current/future directions (SPP 1305)

- incorporating past information
- minimizing the number of partition cells
- minimizing the transmitted information
- incorporating delays
- extension to networked systems