

Stealthy Deception Attacks on Water SCADA Systems

Saurabh Amin¹ Xavier Litrico² Alexandre M. Bayen¹
S. Shankar Sastry¹

¹University of California, Berkeley

²Cemagref, Unité Mixte de Recherche G-EAU

TRUST Spring Conference
Washington DC, October 29, 2009

Outline

Recapitulation from last year

The Gignac Water SCADA System

Modeling of Cascade Canal Pools

Attacks on PI Control

Limits on Stability and Detectability

Outline

Recapitulation from last year

The Gignac Water SCADA System

Modeling of Cascade Canal Pools

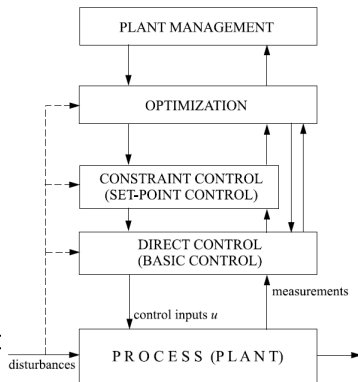
Attacks on PI Control

Limits on Stability and Detectability

Attacks to Control Systems.

Attacks can disrupt

- ▶ **Set points:** man-in-the-middle substitutions,
- ▶ **Control:** tuning parameter substitutions
- ▶ **Process value readings:** value substitutions,
- ▶ **Communication:** latency impact via DoS attack,
- ▶ **Process disruption:** disrupt connection to plant.



Multilayer Control Structure

(Tatjewski, '08)

Operational Goals and Security Attributes.

Operational Goals

- ▶ Maintain safe operational mode
 - Limit the probability of undesirable behavior,
- ▶ Meet production demands
 - Keep certain process values within prescribed limits,
- ▶ Maximize production profit.

Security attributes

- ▶ Mode of attack
 - Availability, integrity, confidentiality,
- ▶ Signature of attack
 - Targeted, resource constrained, random,
- ▶ Time of attack.

Control System Theory.

Theory

- ▶ linear, nonlinear, adaptive, robust, networked, fault-tolerant, distributed, ...

Practice

- ▶ PID (more than 80% of engineering practice)
- ▶ Model predictive control (widely accepted mediator between PID and PhD control)

Which tools to use

- ▶ To analyze of resiliency or defenses under malicious attacks?
- ▶ Our answer: Use any and all, but be grounded


Main goals of our work since last year

For SCADA systems

- ▶ From **attacker's viewpoint**: Provide intuition about *best* deception attack signatures that are most effective and least detectable. We define such attacks as **stealthy attacks**.
- ▶ From **defender's viewpoint**: Synthesize controllers that improve robustness margins with safety and cost as operational goals under a class of deception attack models.

Roadmap for secure water SCADA


Roadmap to Secure Control Systems in the Water Sector



March 2008

Developed by
Water Sector Coordinating Council Cyber Security Working Group

Sponsored by



American Water Works Association Homeland Security

Outline

Recapitulation from last year

The Gignac Water SCADA System

Modeling of Cascade Canal Pools

Attacks on PI Control

Limits on Stability and Detectability

The Gignac Water Distribution System.

The Gignac Project

- ▶ Built in 18XX, located near Montpellier, France
- ▶ Irrigates 2800 *Hc* of land by 50 *km* of primary, 270 *km* of secondary canals
- ▶ Used as experimental testbed for modeling and control methods
- ▶ Equipped with level and velocity sensors, and motorized gates with local slave controllers
- ▶ SCADA system architecture: centralized base station that communicates with field devices



The Gignac Water Distribution System



Communication system at GIGNAC SCADA



Hacking SCADA Wireless is Easy: PNNL



SCADA Communication System.



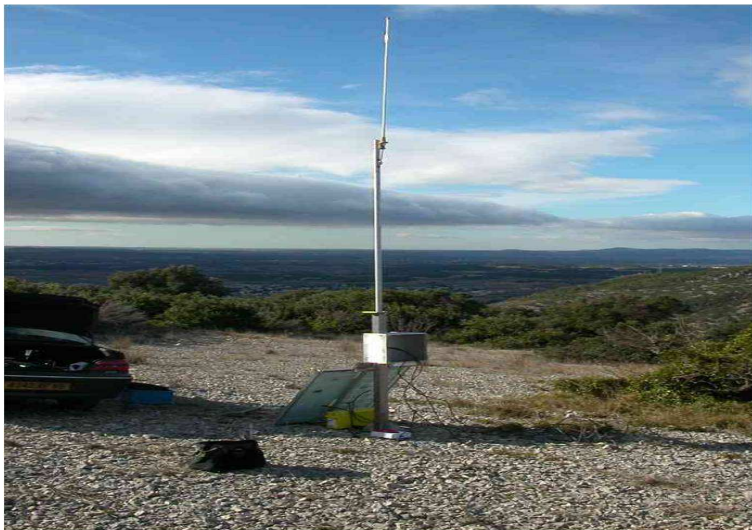
Attack on Local Slave Controller.



Stealing Water by Compromising Level Sensors.



Stolen Solar Panels.



SCADA Supervisory Interface Under Physical Attack.

The screenshot displays the 'asa frontal1' SCADA interface for the 'Canal de Gignac' system. The interface is organized into several panels, each representing a different station. Each panel includes a station name, a 'Horodate API' (date and time), and various monitoring data points such as 'Niveau amont' (upstream level), 'Niveau aval' (downstream level), 'Décharge' (discharge), 'Déversoir' (weir), and 'Physiometrie' (hydraulic characteristics). The interface also features status indicators (alarms, trends) and control buttons for each station.

Key stations visible include:

- Station Prise de la Combe du cor:** Niveau amont: 0.0 cm, Niveau aval: 0.0 cm, Niveau aval canal: 0.0 m.
- Station BELBEZET:** Niveau amont: 113.2 cm, Niveau trop haut: 1331.4 m. Décharge: 64 L/s, 529 L/s, 430 L/s, 291 L/s. Déversoir: 11 L/s.
- Station REPETEUR:** Horodate API: 00:00:00 / 00 / 0. Repeteur.
- Station AVENCQ:** Niveau amont: 91.2 cm, Décharge: 0.0 cm, Déversoir: 652 L/s. Niveau aval: 58.0 cm, Obj: 551 L/s, Obj: 601 L/s.
- Station LAGAREL:** Niveau amont: 95.4 cm, Niveau amont van.: 94.2 cm. Décharge: 3.2 cm, 48.8 cm, 497 L/s, 0 L/s. Déversoir: 0 L/s, 0 L/s. Déversoir mince: 0 L/s. Niveau aval: 65.8 cm, Déversoir total: 0 L/s.
- Station MAS DE ROUVIERE:** Niveau amont: 81.5 cm, Décharge: 60 L/s, Déversoir: 115 L/s.
- Station CEYRIAS:** Niveau amont: 56.57 cm, Décharge: 0.0 cm, Déversoir: 11 L/s, Déversoir rejet: 0 L/s, Rejet total: 10 L/s.
- Station GIGNAC:** Niveau amont: 77.0 cm, Déversoir: 555 L/s.
- Station AURELLE:** Niveau amont: 0.0 cm, Déversoir canal: 0 L/s, Déversoir rejet: 0 L/s.
- Station PONT LUSSAC:** Niveau amont: 66.21 cm, Point: 88.6 cm, Conduite: 0 L/s, Déversoir rejet: 84.3 L/s, Rejet total: 84.3 L/s.
- Station SAINT JEAN:** Niveau amont: 0.0 cm, Décharge: 0.0 cm, Déversoir rejet: 0 L/s. Physiometrie: 0.0 mm, 0.0 mm, 0.0 mm. Rejet total: 0 L/s.

The interface also shows a Windows taskbar at the bottom with the system clock at 11:41 and various application icons.

Other Cyber Attacks on Water SCADA Systems.

- ▶ Tehama colusa canal incident
- ▶ Maroochy water breach incident
- ▶ Harrisburg water filtering plant incident

Outline

Recapitulation from last year

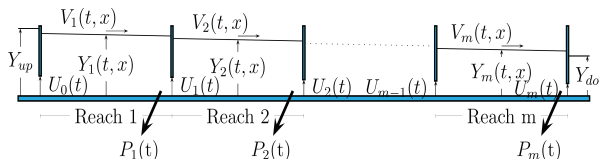
The Gignac Water SCADA System

Modeling of Cascade Canal Pools

Attacks on PI Control

Limits on Stability and Detectability

Cascade of Canal Pools.



Frequency domain model of canal cascade

- ▶ Control input variables: Upstream and downstream discharge,
Controlled variable: the downstream water level,
- ▶ Frequency domain input-output relationship for pool i :

$$y_i(s) = G_i(s)\mu_i(s) + \tilde{G}_i(s)[\mu_{i+1}(s) + p_i(s)]$$

- ▶ $G_i(s)$ and $\tilde{G}_i(s)$ are infinite dimensional transfer functions belonging to the **Callier-Desoer algebra**

Reduced Order Model for Adversary.

Low frequency approximation

- ▶ Transfer functions $G_i(s)$ and $\tilde{G}_i(s)$ can be approximated by an Integrator Delay (ID) model

$$G_i(s) = \frac{\exp(-\tau_i s)}{A_i s}, \quad \tilde{G}_i(s) = -\frac{1}{A_i s}$$

where τ_i is the propagation delay of the i -th canal pool (in s) and A_i is the backwater area (in m^2)

- ▶ Classically used to design PI controllers
- ▶ Multi-pool representation of canal is obtained as

$$y = G\mu + \tilde{G}p.$$

Outline

Recapitulation from last year

The Gignac Water SCADA System

Modeling of Cascade Canal Pools

Attacks on PI Control

Limits on Stability and Detectability

Local Upstream Control of Avencq Cross Regulator.



PI Controller for Local Upstream Control.

- ▶ Integrator-delay model of single pool

$$y_1(s) = G_1(s)\mu_1(s) + \tilde{G}_1(s)[\mu_2(s) + p_1(s)]$$

- ▶ $G_1(s) = \exp(-\tau_1 s)/A_1 s$ and $\tilde{G}_1(s) = -1/A_1 s$.
- ▶ Measured variable: water level upstream of the gate, Control action variable: Gate opening
- ▶ $K_1(s) = 0$ and $K_2(s) = k_p \left(1 + \frac{1}{T_i s}\right)$: k_p proportional gain, T_i integral time
- ▶ Tracking error $\epsilon_1 = r_1 - y_1$

$$\epsilon_1 = (1 + \tilde{G}_1(s)K_2(s))^{-1}[r_1 - \tilde{G}_1(s)p_1(s)] \quad \text{for local u.s.}$$

- ▶ Disturbance rejection is characterized by modulus of transfer function.

Robust PI Tuning Rules.

- ▶ Robustness margins: Gain margin ΔG dB and Phase margin $\Delta\Theta^\circ$ – directly related to the time-domain performance of the closed-loop system
- ▶ Tuning Rules

$$k_p = k_u \frac{\pi^2}{8} 10^{-\Delta G/20} \sin \left(\frac{\pi}{180} \Delta\Theta + \frac{\pi}{2} 10^{-\Delta G/20} \right)$$

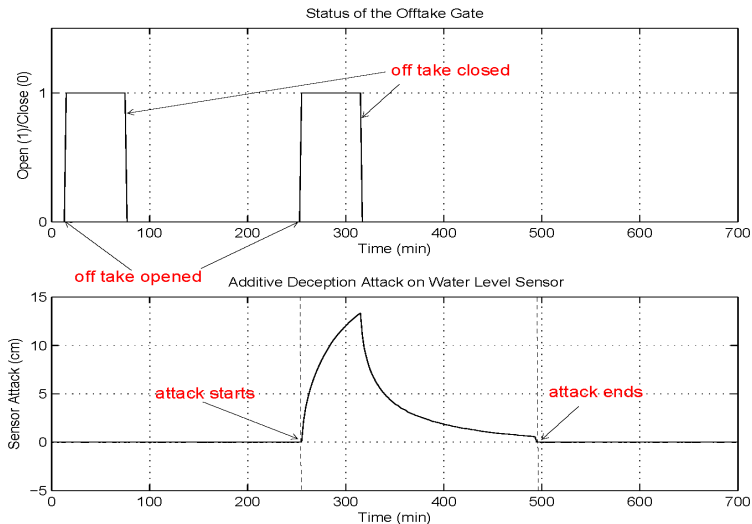
$$T_i = \frac{T_u}{2\pi} 10^{\Delta G/20} \tan \left(\frac{\pi}{180} \Delta\Theta + \frac{\pi}{2} 10^{-\Delta G/20} \right)$$

with $\Delta\Theta < 90(1 - 10^{-\Delta G/20})$

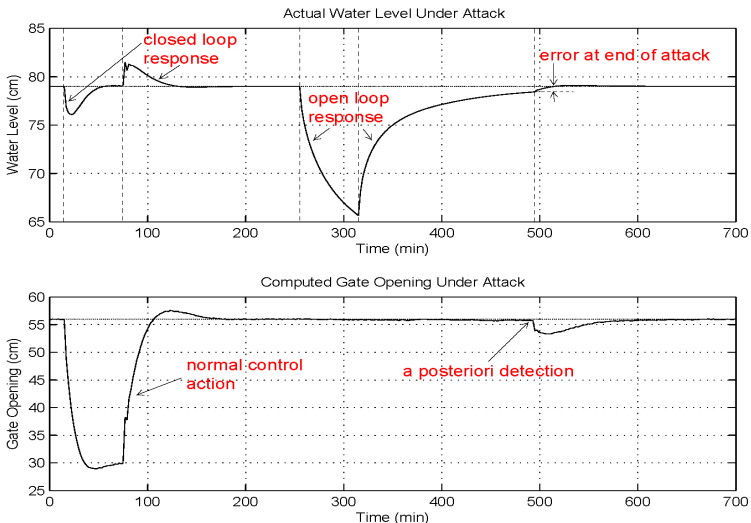
- ▶ The ultimate cycle parameters by ATV method (Astrom and Hagglung)

$$k_u = \frac{4A_1}{\pi\tau_1}, \quad T_u = 4\tau_1$$

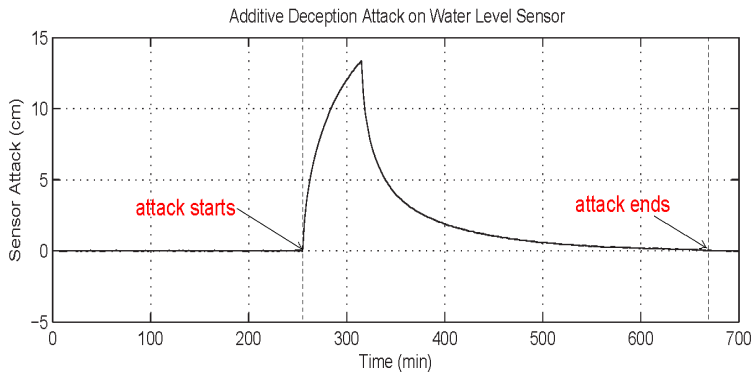
Hacking Level Sensor at Avencq: Simulation



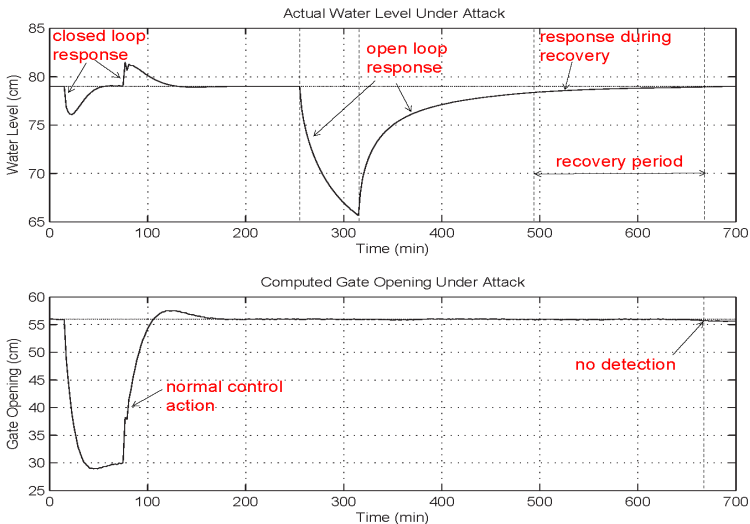
Performance of Local PI Controller: Simulation



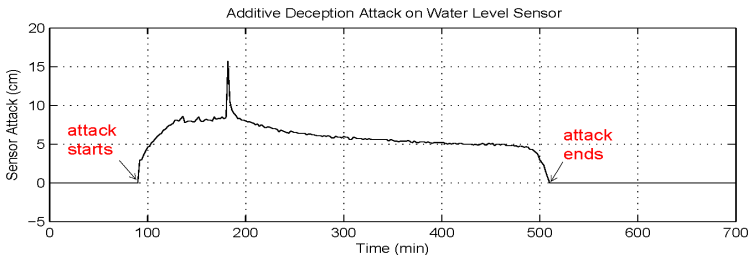
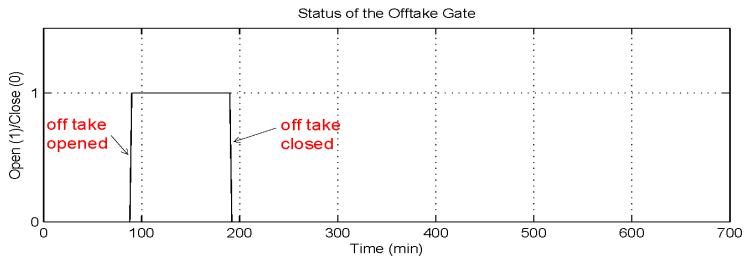
Stealthily Hacking Level Sensor: Simulation



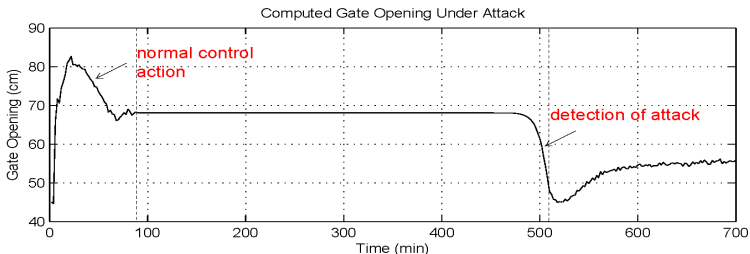
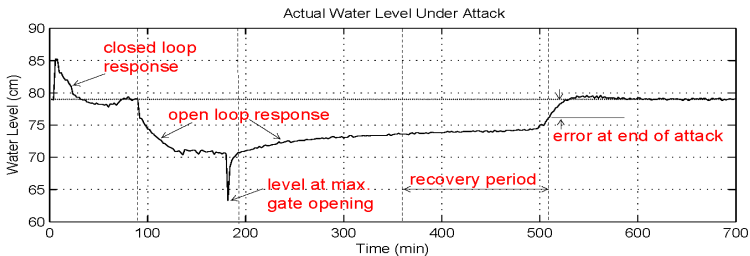
Performance of Local PI Controller: Simulation



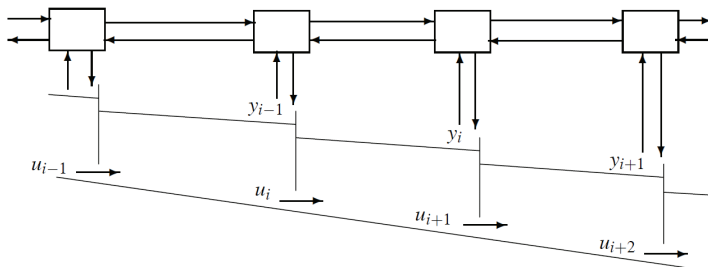
Hacking Level Sensor at Avencq: Real Experiment



Hacking Level Sensor at Avencq: Real Experiment



Extension to decentralized, multivariable PI controllers



- ▶ Compensating effect of water withdrawal at the boundaries by manipulating sensor readings,
- ▶ Such that the multivariable controller does not react to actual perturbation.

Outline

Recapitulation from last year

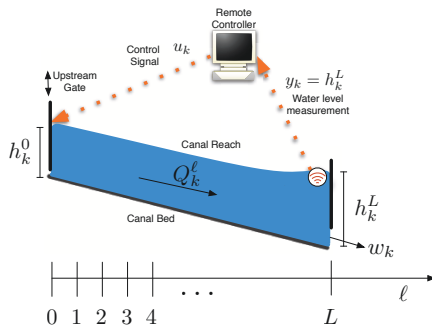
The Gignac Water SCADA System

Modeling of Cascade Canal Pools

Attacks on PI Control

Limits on Stability and Detectability

Modeling deception attacks as switched hyperbolic systems



For regulatory control of an open channel, switching from

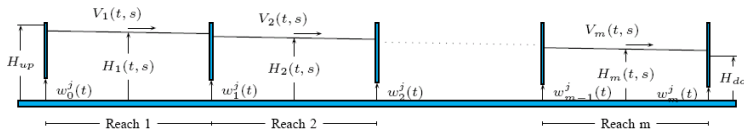
1. Intermittent offtake withdrawals.
2. Certain deception/DoS attacks on sensor and control data.

Stability of linearized flow dynamics in open channels

Regulatory switching of multi-mode underflow sluice gates controlling the flow of water in a cascade of canals

- For horizontal m -cascaded canals with frictionless walls and rectangular cross-section

$$\partial_t \begin{pmatrix} H_i \\ V_i \end{pmatrix} + \begin{pmatrix} V_i & H_i \\ g & V_i \end{pmatrix} \partial_s \begin{pmatrix} H_i \\ V_i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$



Stability of linearized flow dynamics in open channels

Regulatory switching of multi-mode underflow sluice gates controlling the flow of water in a cascade of canals

- ▶ For horizontal m -cascaded canals with frictionless walls and rectangular cross-section

$$\partial_t \begin{pmatrix} H_i \\ V_i \end{pmatrix} + \begin{pmatrix} V_i & H_i \\ g & V_i \end{pmatrix} \partial_s \begin{pmatrix} H_i \\ V_i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

- ▶ Under constant boundary conditions, each canal attains a uniform steady state (\bar{H}_i, \bar{V}_i) .
- ▶ Using $v_i(x, t) = V_i(x, t) - \bar{V}_i$ and $h_i(x, t) = H_i(x, t) - \bar{H}_i$, the linearized model can be written as

$$\partial_t \begin{pmatrix} h_i \\ v_i \end{pmatrix} + \begin{pmatrix} \bar{V}_i & \bar{H}_i \\ g & \bar{V}_i \end{pmatrix} \partial_s \begin{pmatrix} h_i \\ v_i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Stability of linearized flow dynamics in open channels

- ▶ By coordinate change $\xi_i(t, s) = h_i(t, s) + v_i \sqrt{\bar{H}_i/g}$,
 $\xi_{m+i}(t, s) = h_i(t, s) - v_i \sqrt{\bar{H}_i/g}$

$$\frac{\partial}{\partial t} \begin{pmatrix} \xi_i \\ \xi_{m+i} \end{pmatrix} + \begin{pmatrix} \lambda_i & 0 \\ 0 & \lambda_{m+i} \end{pmatrix} \frac{\partial}{\partial s} \begin{pmatrix} \xi_i \\ \xi_{m+i} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

with $\lambda_i = (\sqrt{g\bar{H}_i} - \bar{V}_i)$ and $\lambda_{m+i} = (\sqrt{g\bar{H}_i} + \bar{V}_i)$.

- ▶ Under sub-critical flow, the eigenvalues satisfy $\lambda_i < 0 < \lambda_{m+i}$.

$$\partial_t \xi + \Lambda \partial_s \xi = 0,$$

- ▶ The boundary conditions in linearized form for each j

$$\xi_{II}(t, 0) = G_L^j \xi_I(t, 0) \quad \xi_I(t, 1) = G_R^j \xi_{II}(t, 1) \quad (1)$$

with appropriately defined G_L^j, G_R^j .

- ▶ Our results provide sufficient conditions to decay for any admissible regulatory control action.

Switched linear hyperbolic systems

Subsystems for the unknown $u(t, s) \in \mathbb{R}^n$ with dynamics

$$\partial_t u(t, s) + A^j(s) \partial_s u(t, s) + B^j(s) u(t, s) = 0, \quad s \in (a, b), \quad t > 0$$

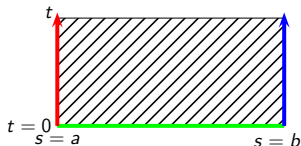
with **initial data**

$$u(0, s) = \bar{u}(s), \quad s \in (a, b)$$

and **left** and **right** boundary conditions

$$D_L^j u(t, a) = 0, \quad D_R^j u(t, b) = 0, \quad t \in [0, \infty)$$

where j belongs to discrete set $\mathcal{Q} = \{1, \dots, N\}$.



Switched linear hyperbolic systems

Consider switching in time among these
with a piecewise constant **switching signal**

$$\sigma(t) = j_k, \text{ with } j_k \in \mathcal{Q}, \text{ for } t \in [\tau_k, \tau_{k+1}).$$

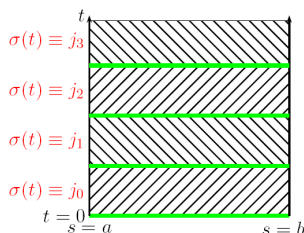
For initial data $\bar{u}(\cdot)$, define $\mathbf{u}(t) = u(t, \cdot)$ as

$$u(t, \cdot) = u^{j_k}(t, \cdot), \text{ for } t \in (\tau_k, \tau_{k+1})$$

where $u^{j_k}(t, \cdot)$ is a solution of subsystem for mode j_k .

At switching times τ_k

$$u^{j_k}(\tau_k, \cdot) = \begin{cases} \lim_{t \rightarrow \tau_k, t < \tau_k} u^{j_{k-1}}(t, \cdot) & \text{if } k > 0, \\ \bar{u}(\cdot) & \text{if } k = 0. \end{cases}$$



Exponential stability for boundary control actions

Exponential stability

- ▶ The system is **exponentially stable** with respect to a norm $\|\cdot\|$ if there exists $c > 0$ and $\beta > 0$ such that for every initial condition $u(0, \cdot)$, the solution $u(t, \cdot)$ satisfies

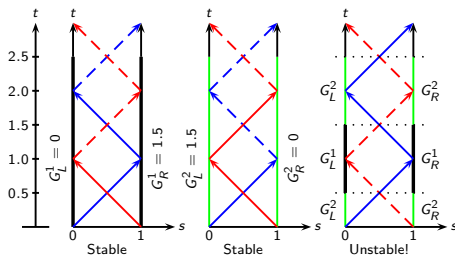
$$\|u(t, \cdot)\| \leq c \exp(-\beta t) \|\bar{u}(\cdot)\|.$$

- ▶ We then say that the switched system is **absolutely exponentially stable** if it is exponentially stable for all switching sequences $\sigma(\cdot)$.

Instability by switching

PDE counterpart to the classical ODE result ($n=2, j \in \{1, 2\}$)

Let $A^j = \text{diag}(-1, 1)$, $G_L^j = 1.5(j-1)$, $G_R^j = 1.5(2-j)$, $\bar{u} \equiv 1$.



We have for $[a, b] = [0, 1]$

- ▶ unswitched: $\|u(t)\| = 0$ for $j \in \{1, 2\}$ and $t > 2$.
- ▶ switch at $t = 0.5, 1.5, 2.5, \dots$: $\|u(t)\|$ unbounded as $t \rightarrow \infty$.

Stability criterion for switched hyperbolic systems

Theorem (Absolute spectral radius condition)

Assume that all matrices A^j are diagonal. Further, assume that boundary data satisfy the N^2 absolute spectral radius conditions

$$(A2) \quad \max_{j,j' \in \{1, \dots, N\}} \rho \left(\begin{pmatrix} 0 & |G_R^{j'}| \\ |G_L^j| & 0 \end{pmatrix} \right) < 1.$$

Then there exists $\epsilon > 0$ such that if $\|B^j(s)\|_\infty < \epsilon$, the switched hyperbolic system is **absolutely exponentially L^∞ -stable**. \square

Proof.

Integrating along characteristic paths and forming desired upper bounds. \square

Stability criterion for switched hyperbolic systems

Theorem (Absolute spectral radius condition)

Assume that all matrices A^j are diagonal. Further, assume that boundary data satisfy the N^2 absolute spectral radius conditions

$$(A2) \quad \max_{j,j' \in \{1, \dots, N\}} \rho \left(\begin{pmatrix} 0 & |G_R^{j'}| \\ |G_L^j| & 0 \end{pmatrix} \right) < 1.$$

Then there exists $\epsilon > 0$ such that if $\|B^j(s)\|_\infty < \epsilon$, the switched hyperbolic system is **absolutely exponentially L^∞ -stable**. \square

Special case ($n = 2$)

$$(A2) \iff \max_{j,j' \in \{1, \dots, N\}} \sqrt{|G_L^j| |G_R^{j'}|} < 1.$$

Stability criterion for switched hyperbolic systems

Theorem (Absolute spectral radius condition)

Assume that all matrices A^j are diagonal. Further, assume that boundary data satisfy the N^2 absolute spectral radius conditions

$$(A2) \quad \max_{j,j' \in \{1, \dots, N\}} \rho \left(\begin{pmatrix} 0 & |G_R^{j'}| \\ |G_L^j| & 0 \end{pmatrix} \right) < 1.$$

Then there exists $\epsilon > 0$ such that if $\|B^j(s)\|_\infty < \epsilon$, the switched hyperbolic system is **absolutely exponentially L^∞ -stable**. \square

For our example

$$\rho \begin{pmatrix} 0 & 1.5 \\ 0 & 0 \end{pmatrix} = \rho \begin{pmatrix} 0 & 0 \\ 0 & 1.5 \end{pmatrix} = \rho \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0, \quad \rho \begin{pmatrix} 0 & 1.5 \\ 0 & 1.5 \end{pmatrix} = 1.5.$$

Dwell-time result for switched hyperbolic systems

Theorem (Dwell-time)

Assume that all matrices A^j are diagonal. Further, assume that the boundary conditions satisfy the **only N** spectral radius conditions

$$\max_{j \in \{1, \dots, N\}} \rho \left(\begin{pmatrix} 0 & |G_R^j| \\ |G_L^j| & 0 \end{pmatrix} \right) < 1.$$

and define the following value

$$\bar{\tau} := \max_{j \in \{1, \dots, N\}} (b - a) \left\{ \left(\min_{\substack{s \in [a, b] \\ i=1, \dots, m_j}} |\lambda_i^j(s)| \right)^{-1} + \left(\min_{\substack{s \in [a, b] \\ i=m_j+1, \dots, n}} |\lambda_i^j(s)| \right)^{-1} \right\}.$$

Then, for any $\tau \geq \bar{\tau}$ assumed as **dwell-time** for the switching signal $\sigma(\cdot)$, i. e. if the intervals between consecutive switches are no shorter than τ , the switched hyperbolic system is **exponentially L^∞ -stable**.



Static Data Reconciliation

- ▶ Similar setting to power system state estimation (*false data injection attacks, CCS 2009.*)
- ▶ Collected n sensor measurements

$$Y_m(k) = [y_{m_1}(k), \dots, y_{m_n}(k)]^\top.$$
- ▶ Physical model as a static relationship of m equations

$$MY(k) = R \quad (2)$$

- ▶ True, unknown measurements $Y(k) = [y_1(k), \dots, y_n(k)]^\top$
- ▶ Additive sensor model: $\epsilon(k)$ Gaussian mean 0, covariance V

$$Y_m(k) = Y(k) + \epsilon(k)$$

- ▶ Solution by solving weighted least squares

$$\hat{Y} = (I_n - SM)Y_m + SR$$

with $S = VM^\top(MVM^\top)^{-1}$.

Detection of Bad Data

- ▶ m -dimensional residual $r = M(Y_m - \hat{Y})$
- ▶ When no bad data r has mean 0, covariance $V_r = MVM^T$,
- ▶ Statistical hypothesis test

$$H_0 : \text{No bad data} \quad \text{and} \quad H_1 : \text{Bad data}$$

- ▶ Hypothesis test can now be expressed as

$$\text{If } \begin{cases} \hat{\phi} > \tau, & \text{reject } H_0 \\ \hat{\phi} \leq \tau, & \text{accept } H_0. \end{cases}$$

where the statistic $\phi := r^T V_r^{-1} r$ and τ depends on allowed false alarm probability.

Stealthy Attack on Data Reconciliation

Theorem (Stealthy Attacks)

Assume that the original measurements Y_m can pass the bad data detection test. Then malicious measurements $Y_a = Y_m + A$ can also pass the bad measurement detection if the adversary with knowledge of the original measurements Y_m , the model M, R , the error variance V and the threshold τ , chooses A such that

$$\| MY_m - R + MA \|_{(MVM)^{-1}}^2 \leq \tau$$



Acknowledgements

- ▶ This research is sponsored by the Team for Research in Ubiquitous Secure Technology (TRUST) at UC Berkeley
- ▶ The support of the Cemagref is gratefully acknowledged
- ▶ We thank the project leaders of SIC simulation software: J.P. Baume & P.O. Malaterre