Probabilistic Methods for Model Validation, Verification and Refinement

Abhishek Halder

Department of Electrical and Computer Engineering
Texas A&M University
College Station, TX

July 8, 2014

Schlumberger-Doll Research Center
Cambridge, MA

Background

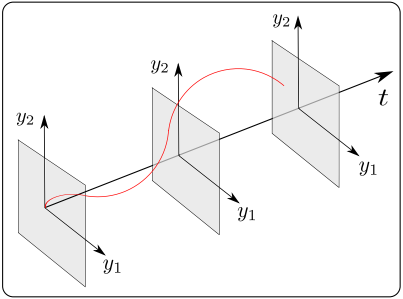
- **Postdoctoral Research Fellow (June 2014 - current)**
  - Department of Electrical and Computer Engineering
    Texas A&M University
- **Ph.D (Aug. 2008 - May 2014)**
  - Department of Aerospace Engineering
    Texas A&M University
  - Dissertation: *Probabilistic Methods for Model Validation*
- **Bachelors and Masters (June 2003 - June 2008)**
  - Department of Aerospace Engineering
    Indian Institute of Technology Kharagpur, India
  - Thesis: *Development of An Autonomous Reconfigurable UAV*

Model validation problem: introduction

- Given
  - Time varying measurements
    - vector (trajectory)
    - set
    - concentration / density
  - Candiate model
    - flow
    - map
  - Input
    - open loop command
    - stochastic disturbance
    - initial condition
- Question
  - How well does the model replicate the measurements?

Model validation problem: introduction

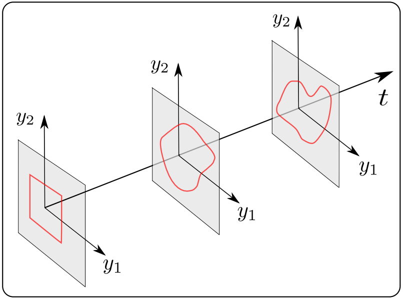
- Given
  - Time varying measurements
    - vector (trajectory)
    - set
    - concentration / density
  - Candiate model
    - flow
    - map
  - Input
    - open loop command
    - stochastic disturbance
    - initial condition



- Question
  - How well does the model replicate the measurements?

Model validation problem: introduction

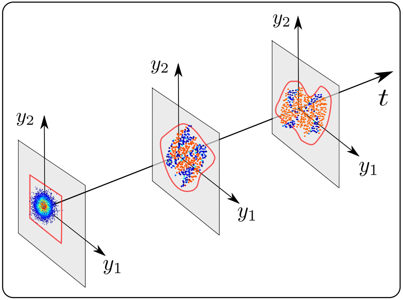
- Given
  - Time varying measurements
    - vector (trajectory)
    - set
    - concentration / density
  - Candiate model
    - flow
    - map
  - Input
    - open loop command
    - stochastic disturbance
    - initial condition



- Question
  - How well does the model replicate the measurements?

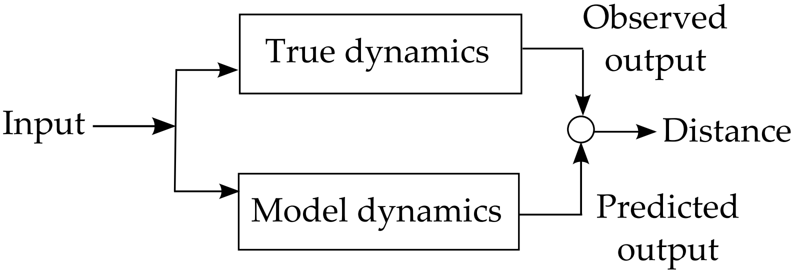
Model validation problem: introduction

- Given
  - Time varying measurements
    - vector (trajectory)
    - set
    - concentration / density
  - Candiate model
    - flow
    - map
  - Input
    - open loop command
    - stochastic disturbance
    - initial condition



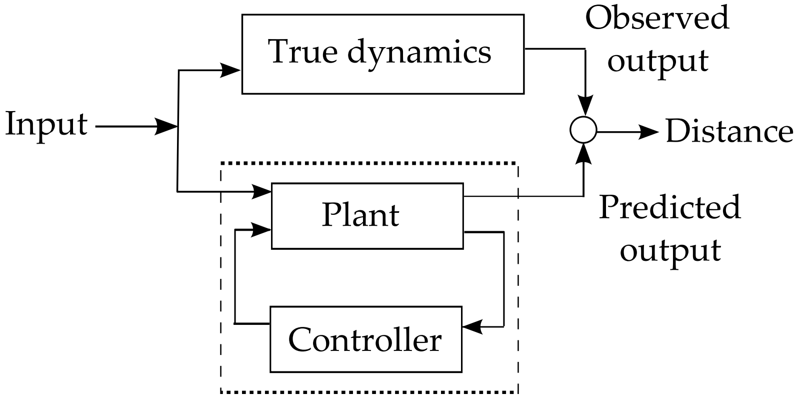
- Question
  - How well does the model replicate the measurements?

Generic validation problem: is the physics correct?



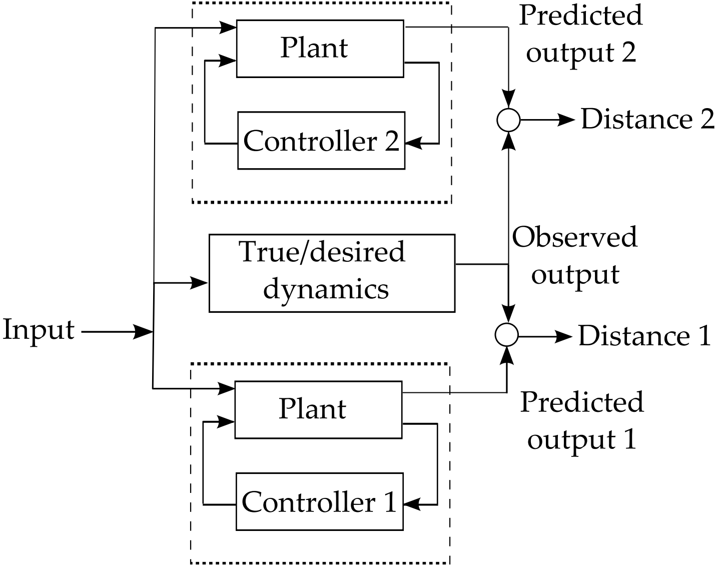
- Applications: predictive modeling
  - Systems biology
  - Atmospheric modeling in planetary entry-descent-landing (EDL)

Generic verification problem: is the implementation correct?

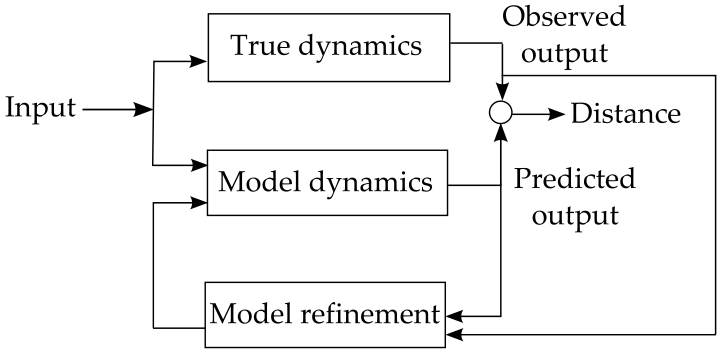


- Applications: performance assessment
  - Flight control software certification
  - Fault detection

Generic verification problem: which implementation is better?



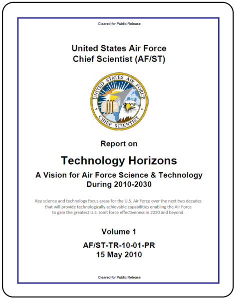
Generic refinement problem: how to improve the model?



- Applications
  - Data driven modeling
  - Density control
  - Fault reconfiguration

Model validation problem: motivation

- U.S. Air Force 2010 Report on Technology Horizons
  - "It is possible to develop systems having high levels of autonomy, but it is the lack of suitable V&V methods that prevents all but relatively low levels of autonomy from being certified for use."


Cleared for Public Release

United States Air Force
Chief Scientist (AF/ST)

Report on

Technology Horizons

A Vision for Air Force Science & Technology
During 2010-2030

Key science and technology focus areas for the U.S. Air Force over the next two decades that will provide technologically achievable capabilities enabling the Air Force to gain the greatest U.S. Joint force effectiveness in 2030 and beyond.

Volume 1

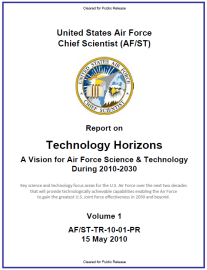
AF/ST-TR-10-01-PR
15 May 2010

Cleared for Public Release

Model validation problem: motivation

- U.S. Air Force 2010 Report on Technology Horizons
  - "It is possible to develop systems having high levels of autonomy, but it is the lack of suitable V&V methods that prevents all but relatively low levels of autonomy from being certified for use."

- F/A-18 Hornet Falling Leaf Mode
  - 47 out-of-control flights during 1983-2000
  - Controller revised in 2001
  - Linear analysis found insufficient


Cleared for Public Release

United States Air Force
Chief Scientist (AF/ST)

Report on

Technology Horizons

A Vision for Air Force Science & Technology
During 2010-2030

Key science and technology focus areas for the U.S. Air Force over the next two decades that will provide technologically achievable capabilities enabling the Air Force to gain the greatest U.S. Joint force effectiveness in 2030 and beyond.

Volume 1

AF/ST-TR-10-01-PR
15 May 2010

Cleared for Public Release

Model validation: state-of-the-art

- Linear model validation
  - Robust control framework
    - Smith & Doyle, 1992
    - Poolla *et. al.*, 1994
    - Smith & Dullerud, 1996
    - Chen & Wang, 1996
    - Steele & Vinnicombe, 2001
    - Gevers *et. al.*, 2003
  - Statistical setting
    - Lee & Poolla, 1996
    - Ljung & Guo, 1997

- Noninear model validation
  - Barrier certificate
    - Prajna, 2006
  - Polynomial chaos
    - Ghanem *et. al.*, 2008

"For the general case of nonparametric (uncertainty) models, the situation is significantly more complicated."

– *[Lee and Poolla, 1996]*

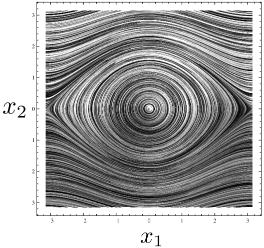
- Most existing methods focus on invalidation/falsification
- Overly conservative?
- Binary oracle vs. "degree" of (in)validation

Our approach: intuitive idea

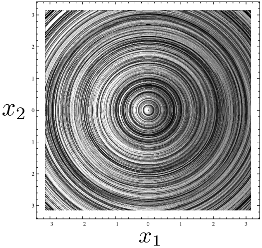
- Our proposal:
  - Compare shapes of the output PDFs at $\{t_j\}_{j=1}^{\tau}$
- Why PDFs instead of
  - trajectories?
  - sets?
  - moments?
- Why shapes?

Our approach: intuitive idea

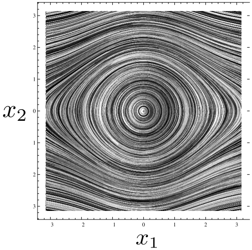
- Our proposal:
  - Compare shapes of the output PDFs at $\{t_j\}_{j=1}^{\tau}$
- Why PDFs instead of
  - trajectories?
  - sets?
  - moments?
- Why shapes?



$$\dot{x}_1 = -x_2, \quad \dot{x}_2 = \sin x_1,$$



$$\dot{x}_1 = -x_2, \quad \dot{x}_2 = x_1,$$

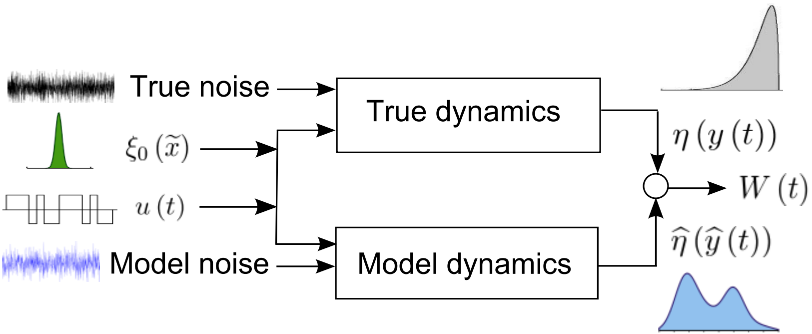


$$\dot{x}_1 = -x_2, \quad \dot{x}_2 = x_1 - \frac{x_1^3}{3!} + \frac{x_1^5}{5!},$$

Outline

- Introduction
- State-of-the-art
- Intuitive idea
- Problem formulation
- Uncertainty propagation
- Distributional comparison
- Examples
- Systems-theoretic results for probabilistic V&V
- Probabilistic model refinement
- Conclusions

Problem formulation



Proposed framework: Valid if $W(t_j) \leqslant \gamma_j, \forall j = 1, 2, \ldots, \tau$

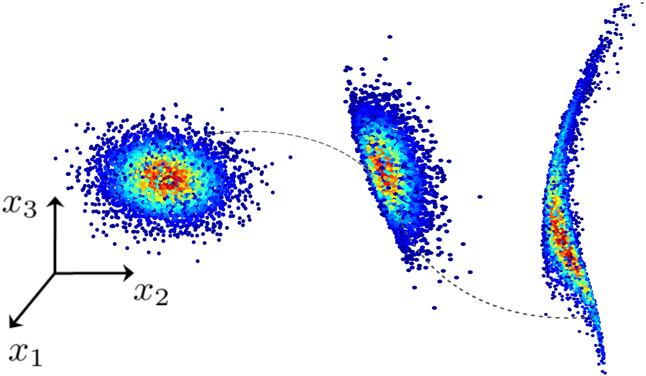
Step 1. Uncertainty propagation

Step 2. Distributional comparison

Uncertainty propagation: deterministic model

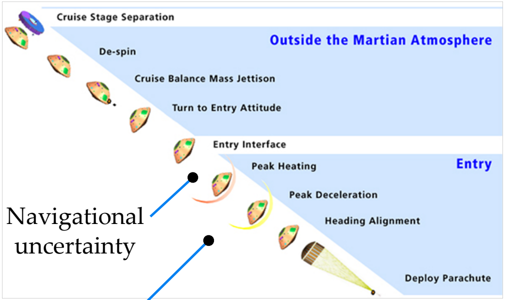
- Model
  - State equation: $\dot{\widehat{x}} = \widehat{f}\left(\widehat{x}, \widehat{p}\right), \widehat{x}\left(t\right) \in \widehat{\mathcal{X}} \subseteq \mathbb{R}^{\widehat{n}_s}, \widehat{p} \in \widehat{\mathcal{P}} \subseteq \mathbb{R}^{\widehat{n}_p}$
  - Extended state space form:
    $$\dot{\widehat{\widehat{x}}} = \widehat{\widehat{f}}\left(\widehat{\widehat{x}}\right), \widehat{\widehat{x}} \in \widehat{\mathcal{X}} \times \widehat{\mathcal{P}} \subseteq \mathbb{R}^{\widehat{n}_s + \widehat{n}_p}, \widehat{\widehat{f}} = \begin{Bmatrix} \widehat{f}_{\widehat{n}_s \times 1} \\ \mathbf{0}_{\widehat{n}_p \times 1} \end{Bmatrix}$$
  - Output equation: $\widehat{y} = \widehat{h}\left(\widehat{\widehat{x}}\right), \widehat{h} : \widehat{\mathcal{X}} \times \widehat{\mathcal{P}} \mapsto \widehat{\mathcal{Y}} \subseteq \mathbb{R}^{n_o}$
- PDF evolution
  - State PDF via Liouville equation: $\dfrac{\partial \widehat{\xi}}{\partial t} = -\displaystyle\sum_{i=1}^{\widehat{n}_s} \dfrac{\partial}{\partial \widehat{x}_i}\left(\widehat{\xi} \widehat{f}_i\right)$
  - MOC for Liouville equation: $\dfrac{d\widehat{\xi}}{dt} = -\widehat{\xi}\, \nabla . \widehat{f}$, initial PDF ξ_0
  - Output PDF: $\widehat{\eta}\left(\widehat{y}, t\right) = \displaystyle\sum_{j=1}^{\nu} \dfrac{\widehat{\xi}\left(\widehat{\widehat{x}}_j^{\star}, t\right)}{\left|\det\left(\mathcal{J}_{\widehat{h}}\left(\widehat{\widehat{x}}_j^{\star}, t\right)\right)\right|}$

Uncertainty propagation: deterministic model



MC simulation	Liouville MOC
Offline post-processing	Online
Histogram approximation	Exact arithmetic
Grid based	Meshless
$\widehat{n}_s$ ODEs per sample	$\widehat{n}_s + 1$ ODEs per sample

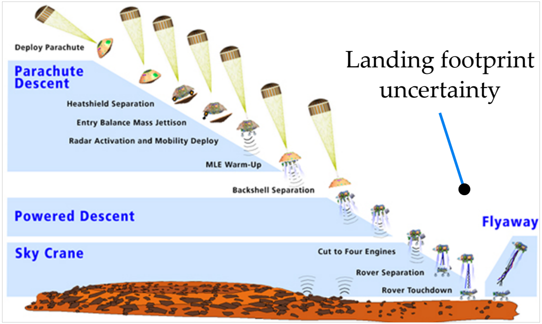
Case study: Mars EDL uncertainty analysis



Navigational uncertainty

Heating uncertainty

Chute deployment uncertainty



Landing footprint uncertainty

Image credit: NASA JPL

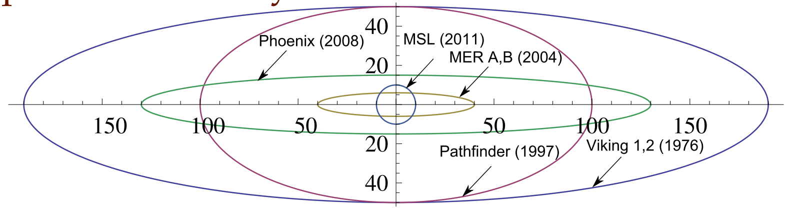
Case study: Mars EDL uncertainty analysis

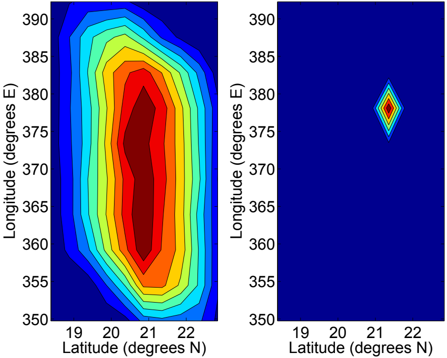
6-state Vinh's equation with 3 parameters: $\rho_0, B_c, \dfrac{C_L}{C_D}$

Atmospheric model: $\rho = \rho_0 \exp\left(\dfrac{h_2 - hR_0}{h_1}\right)$

$$
\begin{aligned}
\dot{h} &= V \sin\gamma \\
\dot{\zeta} &= \frac{V\cos\gamma\sin\chi}{(1+h)} \\
\dot{\lambda} &= \frac{V\cos\gamma\cos\chi}{(1+h)\cos\zeta} \\
\dot{V} &= -\frac{\rho R_0}{2B_c}V^2 - \frac{gR_0}{v_c^2}\sin\gamma + \frac{R_0^2\Omega^2}{v_c^2}(1+h)\cos\zeta\,(\sin\gamma\cos\zeta - \cos\gamma\sin\zeta\sin\chi) \\
\dot{\gamma} &= \frac{\rho R_0}{2B_c}\frac{C_L}{C_D}V\cos\sigma + \frac{gR_0}{v_c^2}\cos\gamma\left(\frac{V}{1+h} - \frac{1}{V}\right) \\
\dot{\chi} &= \frac{\rho R_0}{2B_c}\frac{C_L}{C_D}\frac{V\sin\sigma}{\cos\gamma} - \frac{V\cos\gamma}{(1+h)}\tan\zeta\cos\chi + \frac{2R_0\Omega}{v_c}(\tan\gamma\cos\zeta\sin\chi - \sin\zeta) - \\
&\quad \frac{R_0^2\Omega^2}{v_c^2}\frac{(1+h)}{V\cos\gamma}\sin\zeta\cos\zeta\cos\chi
\end{aligned}
$$

Case study: Mars EDL uncertainty analysis – landing footprint uncertainty

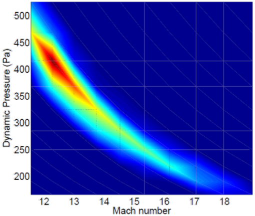


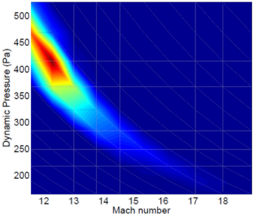


Case study: Mars EDL uncertainty analysis

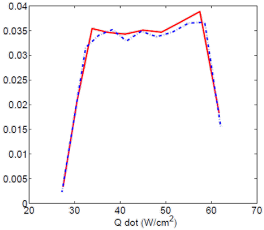
MC | Liouville MOC

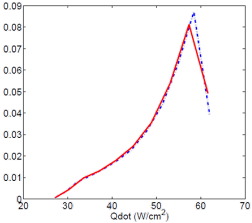
Chute deployment uncertainty





Heating rate uncertainty





Uncertainty propagation: stochastic model

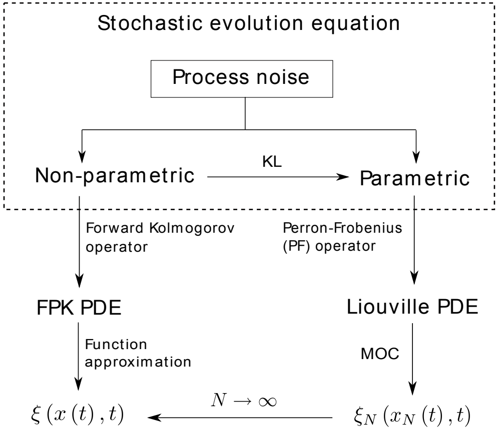
- Model
  - State equation: $d\widehat{\widehat{x}}(t) = \widehat{\widehat{\widetilde{f}}}\left(\widehat{\widehat{x}}(t)\right)\, dt + \widehat{g}\left(\widehat{\widehat{x}}(t)\right)\, d\mathcal{W}$,
  - Output equation: $\widehat{y}(t) = \widehat{h}\left(\widehat{\widehat{x}}(t)\right)$,
- PDF evolution
  - State PDF via Fokker-Planck equation:
    $$\frac{\partial\widehat{\xi}}{\partial t} = -\sum_{i=1}^{\widehat{n}_s}\frac{\partial}{\partial\widehat{x}_i}\left(\widehat{\xi}\widehat{\widetilde{f}}_i\right) + \sum_{i=1}^{\widehat{n}_s}\sum_{j=1}^{\widehat{n}_s}\frac{\partial^2}{\partial\widehat{x}_i\partial\widehat{x}_j}\left(\left(\widehat{g}Q\widehat{g}^\top\right)_{ij}\widehat{\xi}\right),$$
  - Main idea: design a dynamics whose Liouville MOC approximates the Fokker-Planck soltuion
  - Proposed KL + MOC formulation:
    $$\dot{\widehat{x}}_N^{(j)} = \widehat{f}^{(j)}\left(\widehat{x}_N, t\right) + \sum_{k=1}^{n_{\text{noise}}} g^{(j,k)}\left(\widehat{x}_N, t\right)\text{KL}_N^{(k)},$$

Uncertainty propagation: stochastic model

- Theorem
  - Asymptotic consistency: $x_N(t) \xrightarrow[\text{m.s.}]{N\to\infty} x(t)$, $\forall\, t > 0$.
  - Rate-of-convergence: $\lesssim \exp(-N)$ for OU, GBM.

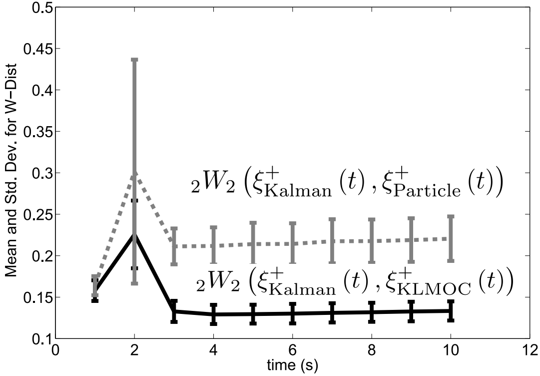
Uncertainty propagation: stochastic model

- Theorem
  - Asymptotic consistency: $x_N(t) \xrightarrow[\text{m.s.}]{N \to \infty} x(t), \forall\, t > 0.$
  - Rate-of-convergence: $\lesssim \exp(-N)$ for OU, GBM.



Numerical results: Kalman filter

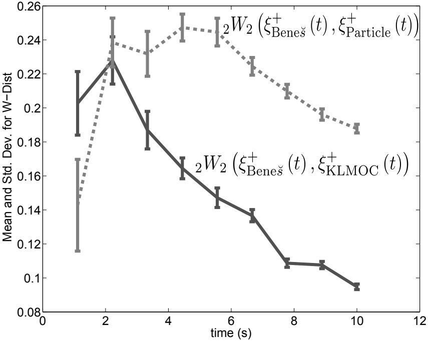
- Process model: $\dot{x}(t) = -0.05 I_2\, x(t) + \begin{bmatrix} 1 & 1 \end{bmatrix}^\top \eta(t)$
- Measurement model: $y_k = \begin{bmatrix} 1 & 1 \end{bmatrix} x_k + v_k, \qquad k \in \mathbb{N}$
- $\eta(t), v_k$ are independent GWN with $Q = 1/8$, $R = 1/4$
- Initial PDF $\varphi_0 = \mathcal{N}\left(\begin{bmatrix} 1 & 1 \end{bmatrix}^\top, \mathrm{diag}(1,1) \right)$
- From ξ_0, we draw 100 sample sets, each with 500 samples



Numerical results: Beneš filter

- Process model: $dx(t) = \dfrac{\kappa e^{x} - e^{-x}}{\kappa e^{x} + e^{-x}}\, dt + d\mathcal{W}(t)$
- Measurement model: $dy(t) = x(t)\, dt + d\mathcal{V}(t)$
- Noise variances: $Q = 1$, $R = 10$; deterministic x_0
- Normalized posterior:

$$\xi^{+}\left(x(t), t|\mathcal{Y}_t\right) = \sqrt{\frac{\coth(t)}{2\pi}} \left(\frac{\kappa e^{x} + e^{-x}}{\kappa e^{I_t(y)} + e^{-I_t(y)}}\right) \exp\left(-\tfrac{1}{2}\Gamma(t)\right)$$



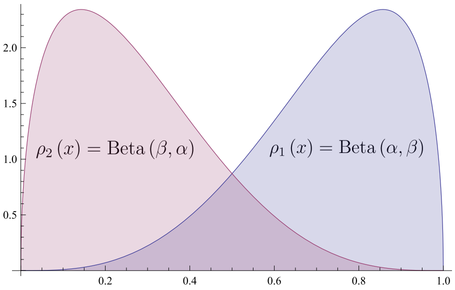
Distributional comparison: axiomatic approach

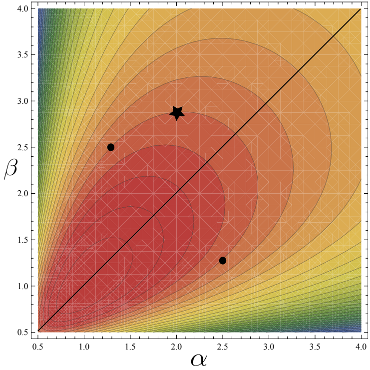
- Candidates for validation distance
  - Kullback-Leibler divergence $D_{KL}(\rho_1||\rho_2) := \int_{\mathbb{R}^d} \rho_1(x) \log\left(\frac{\rho_1(x)}{\rho_2(x)}\right) dx$
  - Symmetric KL divergence
    $D_{KL}^{\text{symm}}(\rho_1||\rho_2) := \frac{1}{2}\left(D_{KL}(\rho_1||\rho_2) + D_{KL}(\rho_2||\rho_1)\right)$
  - Wasserstein distance
    ${}_pW_q(\mu_1, \mu_2) := \left[\inf_{\mu \in \mathcal{M}_2(\mu_1,\mu_2)} \int_\Omega \| \underline{x} - \underline{y} \|_p^q \, d\mu\left(\underline{x}, \underline{y}\right)\right]^{1/q}$

What we want	D_{KL}	D_{KL}^{symm}	W
$\geqslant 0$	✓	✓	✓
Symmetry	×	✓	✓
Triangle inequality	×	×	✓
$\text{supp}(\eta) \neq \text{supp}(\widehat{\eta})$	×	×	✓
$\dim(\text{supp}(\eta)) \neq \dim(\text{supp}(\widehat{\eta}))$	×	×	✓
$\#\text{sample}(\eta) \neq \#\text{sample}(\widehat{\eta})$	×	×	✓
Convexity	✓	✓	✓
Finite range	$[0, \infty)$	$[0, \infty)$	$[0, \text{diam}(\Omega)]$

Distributional comparison: axiomatic approach

- Counterexample 1: randomness $\neq$ shape
  $W(\rho_1, \rho_2) \neq 0$, for $\alpha \neq \beta$ (e.g. $\alpha = 4$, $\beta = \frac{3}{2}$ below)



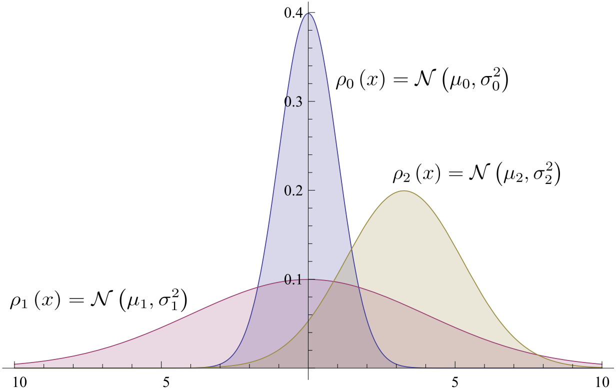


$$H(\rho_1) = H(\rho_2) = \log B(\alpha, \beta) - (\alpha - 1)(\Psi(\alpha) - \Psi(\alpha + \beta)) - (\beta - 1)(\Psi(\beta) - \Psi(\alpha + \beta))$$

Distributional comparison: axiomatic approach

- Counterexample 2: $D_{KL} \neq$ shape

$(\mu_0, \sigma_0) = (0, 1)\,;\ (\mu_1, \sigma_1) = (0, 4)\,;\ (\mu_2, \sigma_2) = \left(\sqrt{12 - 2\log 2}, 2\right)$



$$D_{KL}(\rho_1, \rho_0) = D_{KL}(\rho_2, \rho_0), \text{ but } W(\rho_1, \rho_0) \neq W(\rho_2, \rho_0)$$

Distributional comparison: W for model validation

Wasserstein distance in validation context

- $_pW_q(\eta,\widehat{\eta}) = \left(\inf_{\varrho\in\mathcal{M}_2(\eta,\widehat{\eta})} \int_{\mathcal{Y}\times\widehat{\mathcal{Y}}} \| y-\widehat{y} \|_p^q \; \varrho(y,\widehat{y})\, dy d\widehat{y}\right)^{1/q}$
- Minimum effort required to convert one shape to another
- We choose $p=q=2$, and denote $_2W_2$ as W

When can we write W in closed-form:

- Single output case: $W^2(\eta,\widehat{\eta}) = \int_0^1 \left(F^{-1}(u) - G^{-1}(u)\right)^2 du$
- Multivariate Normal case (comparing Linear Gaussian systems): $W\left((A,C);\left(\widehat{A},\widehat{C}\right)\right) = W(\eta,\widehat{\eta}) = W\left(\mathcal{N}(\mu_1,\Sigma_1), \mathcal{N}(\mu_2,\Sigma_2)\right) = \sqrt{\| \mu_1-\mu_2 \|_2^2 + \operatorname{tr}(\Sigma_1) + \operatorname{tr}(\Sigma_2) - 2\operatorname{tr}\left(\left(\sqrt{\Sigma_1}\Sigma_2\sqrt{\Sigma_1}\right)^{1/2}\right)}$

Distributional comparison: computing W

- At each time $\{t_j\}_{j=1}^{\tau}$, we have two sets of colored scattered data
- Construct complete, weighted, directed bipartite graph $K_{m,n}(U \cup V, E)$ with $\#(U) = m$ and $\#(V) = n$
- Assign edge weight $c_{ij} := \| u_i - v_j \|_{\ell_2}^2$, $u_i \in U$, $v_j \in V$
- minimize $\displaystyle\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}\,\varphi_{ij}$ subject to

$$\sum_{j=1}^{n} \varphi_{ij} = \alpha_i, \qquad \forall\, u_i \in U, \qquad \text{(C1)}$$

$$\sum_{i=1}^{m} \varphi_{ij} = \beta_j, \qquad \forall\, v_j \in V, \qquad \text{(C2)}$$

$$\varphi_{ij} \geqslant 0, \qquad \forall\, (u_i, v_j) \in U \times V. \qquad \text{(C3)}$$

- Necessary feasibility condition: $\displaystyle\sum_{i=1}^{m} \alpha_i = \sum_{j=1}^{n} \beta_j$

Distributional comparison: computing W

Sample complexity

- Rate-of-convergence of empirical Wasserstein estimate

$$\mathbb{P}\left(\left|W\left(\eta_m, \widehat{\eta}_n\right) - W\left(\eta, \widehat{\eta}\right)\right| > \epsilon\right) \leqslant K_1 \exp\left(-\frac{m\epsilon^2}{32C_1}\right) + K_2 \exp\left(-\frac{n\epsilon^2}{32C_2}\right)$$

Runtime complexity

- An LP with mn unknowns and $(m+n+mn)$ constraints
- For $m = n$, runtime is $\mathcal{O}\left(n_o n^{2.5} \log n\right)$

Storage complexity

- For $m = n$, constraint is a binary matrix of size $2n \times n^2$
- Each row has n ones. Total # of ones = $2n^2$
- At a given snapshot, sparse storage complexity is $2n\left(3n + n_o + 1\right) = \mathcal{O}\left(n^2\right)$
- Non-sparse storage complexity is $2n\left(n^2 + n_o + 1\right) = \mathcal{O}\left(n^3\right)$

Distributional comparison: computing W

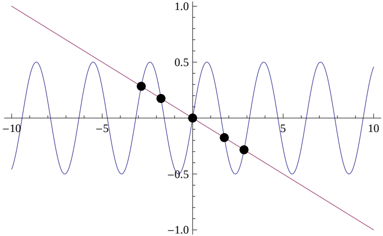
In standard LP form

$$\underset{x \geqslant 0}{\text{minimize}} \quad \tilde{c}^\top x,$$
$$\text{subject to} \quad Ax = b,$$

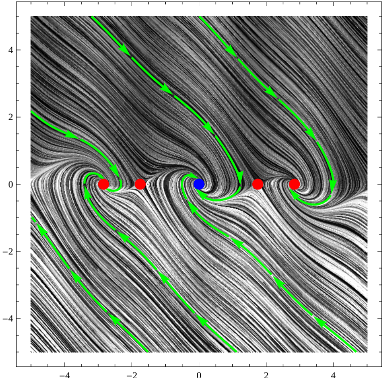
with $\tilde{c}_{mn \times 1} = \text{vec}(c)$, $x_{mn \times 1} = \text{vec}(\varphi)$, $b_{(m+n) \times 1} = [\alpha_{m \times 1}; \beta_{n \times 1}]^\top$, Let $e_n = \Big[\underbrace{1, 1, \ldots, 1}_{n \text{ times}}\Big]^\top$. Then fast construction of $A_{(m+n) \times mn} = \begin{bmatrix} e_n^\top \otimes I_m \\ I_n \otimes e_m^\top \end{bmatrix}$.

Solver used: Large scale sparse LP solver MOSEK®

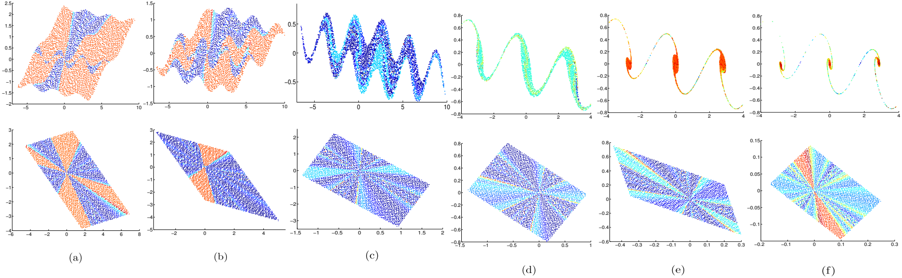
Example 1: model validation

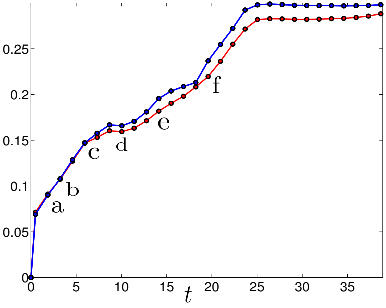


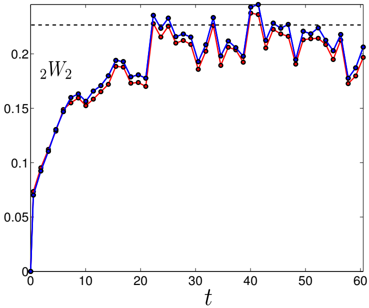
- Truth: $\ddot{x} = -ax - b\sin 2x - c\dot{x}$, $a = 0.1, b = 0.5, c = 1$.
- Five equilibria
- Model: Linearization about origin
- $\xi_0 = \mathcal{U}\left([-4,6] \times [-4,6]\right)$
- Let $y_1 = x$, $y_2 = \dot{x}$
- We plot time history of $W\left(\eta_k, \widehat{\eta}_k\right)$



Example 1 (contd.): W vs. t



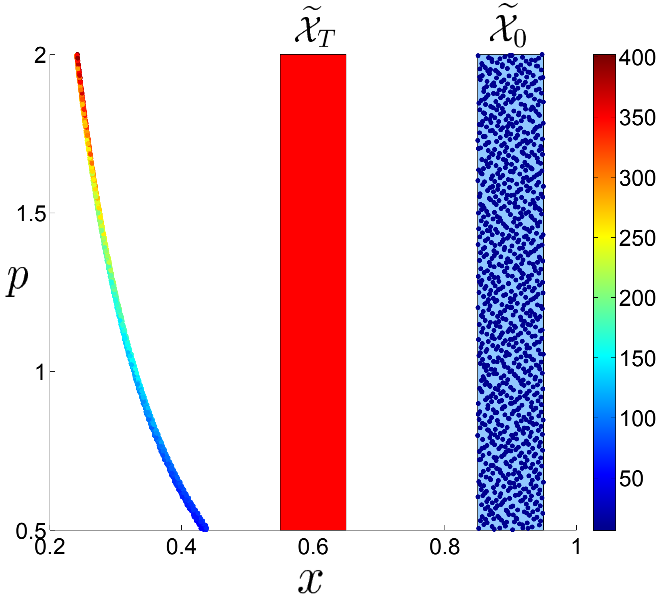




Example 2: model falsification

- Model: $\dot{x} = -px^3$,
- Parameter: $p \in \mathcal{P} = [0.5, 2]$,
- Measurement data: $\mathcal{X}_0 = [0.85, 0.95]$ at $t = 0$, and $\mathcal{X}_T = [0.55, 0.65]$ at $t = T = 4$,
- Prajna's Barrier certificate (from SOS optimization):
  $B(x,t) = B_1(x) + tB_2(x)$,
  $B_1(x) = 8.35x + 10.40x^2 - 21.50x^3 + 9.86x^4$,
  $B_2(x) = -1.78 + 6.58x - 4.12x^2 - 1.19x^3 + 1.54x^4$.
- Our approach: Show that the final measure $\xi_T(x_T, p, T) \sim \mathcal{U}(x_T, p) = \dfrac{1}{\text{vol}\left(\widetilde{\mathcal{X}}_T\right)}$ is not reachable from the initial measure $\xi_0(x_0, p) \sim \mathcal{U}(x_0, p) = \dfrac{1}{\text{vol}\left(\widetilde{\mathcal{X}}_0\right)}$ in $T = 4$.

Example 2: model falsification (contd.)



Example 3: F-16 controller robustness verification

Constant altitude longitudinal flight: $x = (\theta, V, \alpha, q)^{\top}$, $u = (T, \delta_e)^{\top}$

$$\begin{aligned}
\dot{\theta} &= q, \\
\dot{V} &= \frac{1}{m}\cos\alpha\left[T - mg\sin\theta + \bar{q}S\left(C_X + \frac{\bar{c}}{2V}C_{X_q}q\right)\right] + \frac{1}{m}\sin\alpha\left[mg\cos\theta + \bar{q}S\left(C_Z + \frac{\bar{c}}{2V}C_{Z_q}q\right)\right], \\
\dot{\alpha} &= q - \frac{\sin\alpha}{mV}\left[T - mg\sin\theta + \bar{q}S\left(C_X + \frac{\bar{c}}{2V}C_{X_q}q\right)\right] + \frac{\cos\alpha}{mV}\left[mg\cos\theta + \bar{q}S\left(C_Z + \frac{\bar{c}}{2V}C_{Z_q}q\right)\right], \\
\dot{q} &= \frac{\bar{q}S\bar{c}}{J_{yy}}\left[C_m + \frac{\bar{c}}{2V}C_{m_q}q + \frac{\left(x_{\text{cg}}^{\text{ref}} - x_{\text{cg}}\right)}{\bar{c}}\left(C_Z + \frac{\bar{c}}{2V}C_{Z_q}q\right)\right].
\end{aligned}$$

Stochastic initial condition: $x_0 = \underbrace{x_{\text{trim}}}_{\text{from SNOPT}} + x_{\text{pert}}$.

Admissible perturbation:

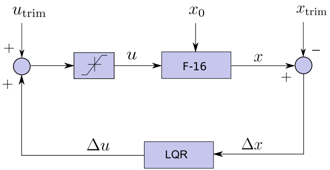
$x_{\text{pert}} \sim \mathcal{U}\left([-35^\circ, 35^\circ] \times [-50\text{ ft/s}, 50\text{ ft/s}] \times [-10^\circ, 45^\circ] \times [-60^\circ/\text{s}, 60^\circ/\text{s}]\right)$

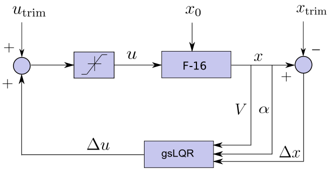
Control objective: $\min_u \mathcal{J} = \int_0^\infty \left(x^\top Q x + u^\top R u\right) dt$,

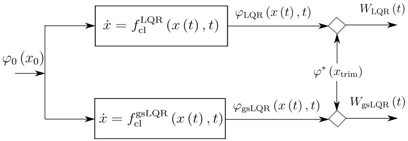
$Q = \text{diag}\left(100, 0.25, 100, 10^{-4}\right), R = \text{diag}\left(10^{-6}, 625\right)$.

Control saturation: $1000\text{ lb} \leqslant T \leqslant 28{,}000\text{ lb}, \quad -25^\circ \leqslant \delta_e \leqslant +25^\circ$.

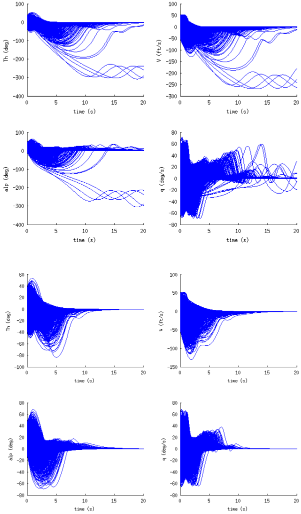
Case Study: F-16 controller robustness verification



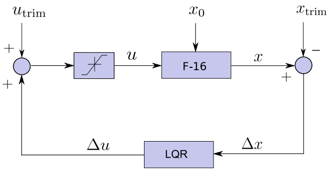


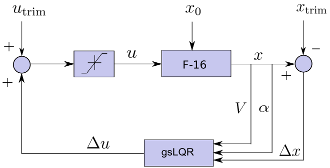


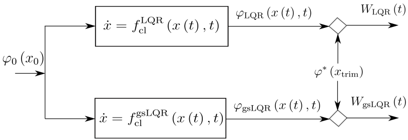
LQR vs. gsLQR Results: (MC)



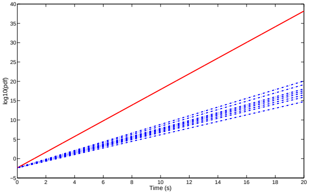
Case Study: F-16 controller robustness verification

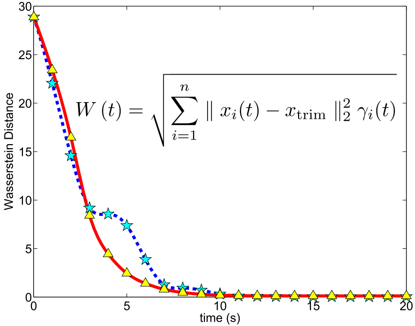






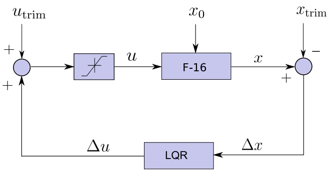
LQR vs. gsLQR Results: (MOC)

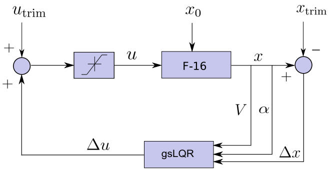


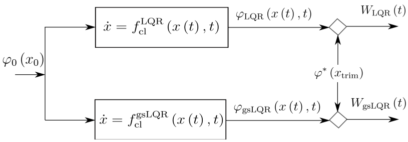


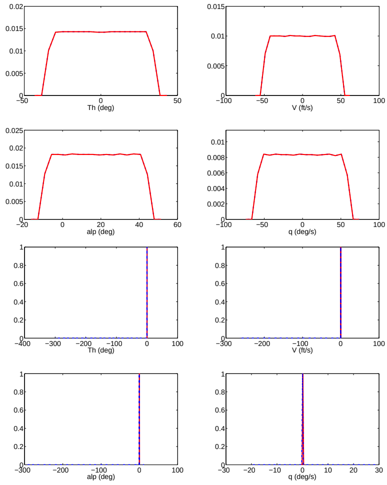
Case Study: F-16 controller robustness verification

Error marginals at $t = 0.01$ and 20 s

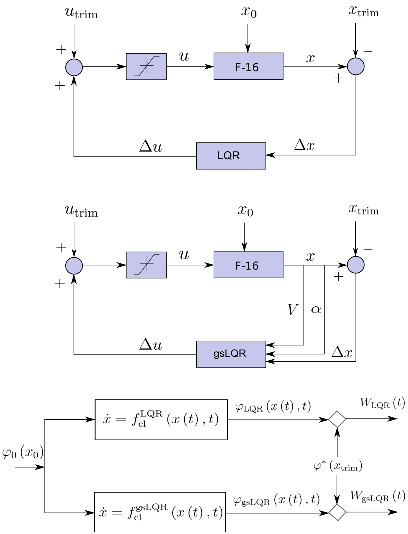




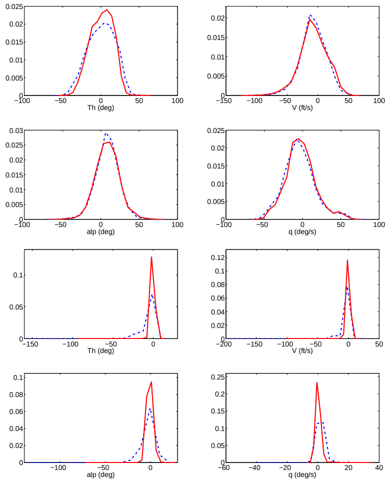




Case Study: F-16 controller robustness verification



Error marginals at $t = 1$ and 5 s



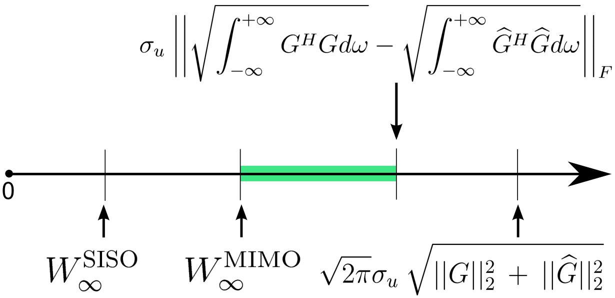
Input-Output Model Validation for LTI Systems

Theorem

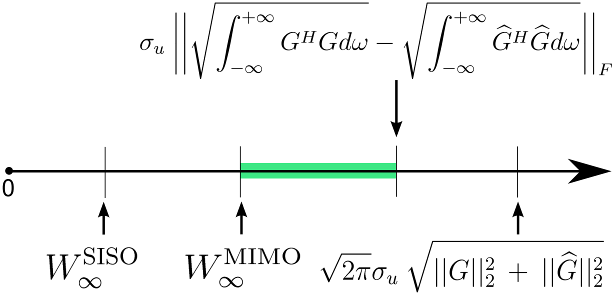
1. **SISO and MISO:** $W_\infty\left(G,\widehat{G}\right) = \sqrt{2\pi}\sigma_u \left| \, ||G(j\omega)||_2 - ||\widehat{G}(j\omega)||_2 \, \right|,$

2. **MIMO:** $$W_\infty\left(G,\widehat{G}\right) = \sqrt{2\pi}\sigma_u \left(||G(j\omega)||_2^2 + ||\widehat{G}(j\omega)||_2^2 - 2\,\mathrm{tr}\left[\left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} G^H(j\omega)\,G(j\omega)\,d\omega\right)^{1/2}\left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} \widehat{G}^H(j\omega)\,\widehat{G}(j\omega)\,d\omega\right)\left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} G^H(j\omega)\,G(j\omega)\,d\omega\right)^{1/2}\right]^{1/2}\right)^{1/2}$$

Bounds for MIMO W_∞

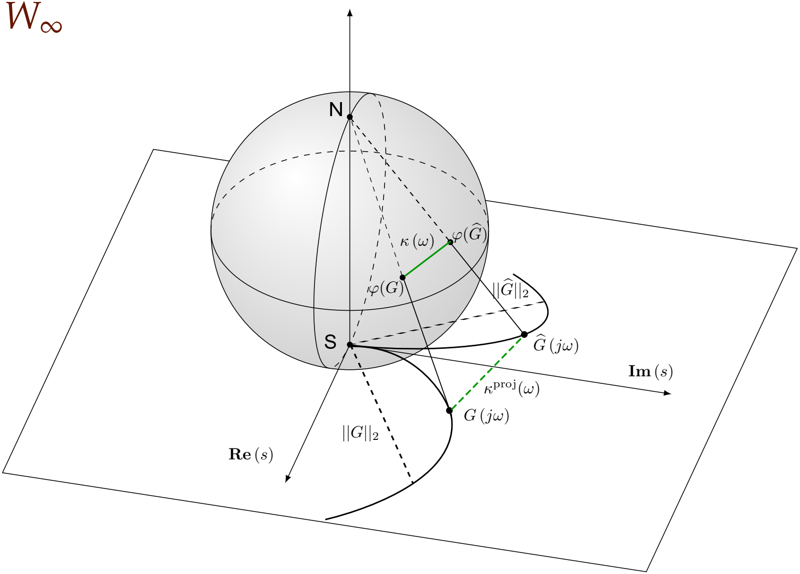


Bounds for MIMO W_∞

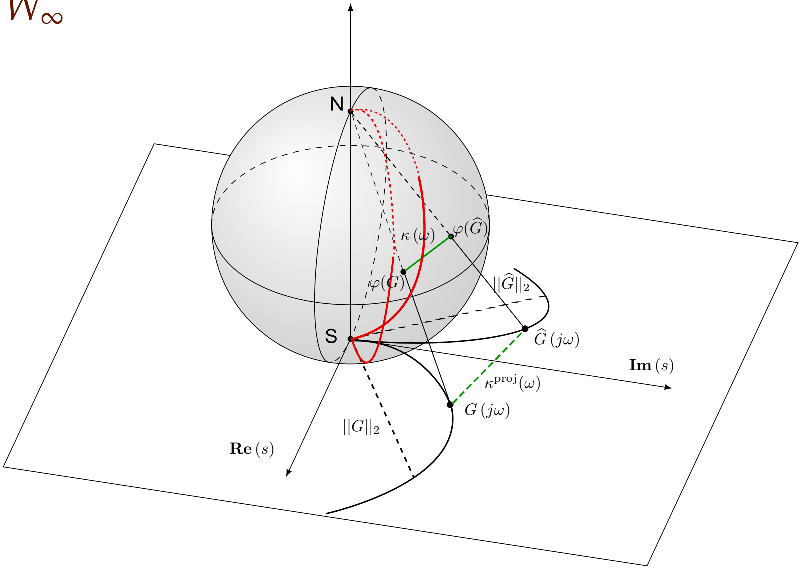


Observation: the **"green gap"** $\to 0$, if $\left[\Sigma_\infty, \widehat{\Sigma}_\infty\right] \to 0$.

Geometric Meaning & Intrinsic Normalization of SISO W_∞



Geometric Meaning & Intrinsic Normalization of SISO W_∞

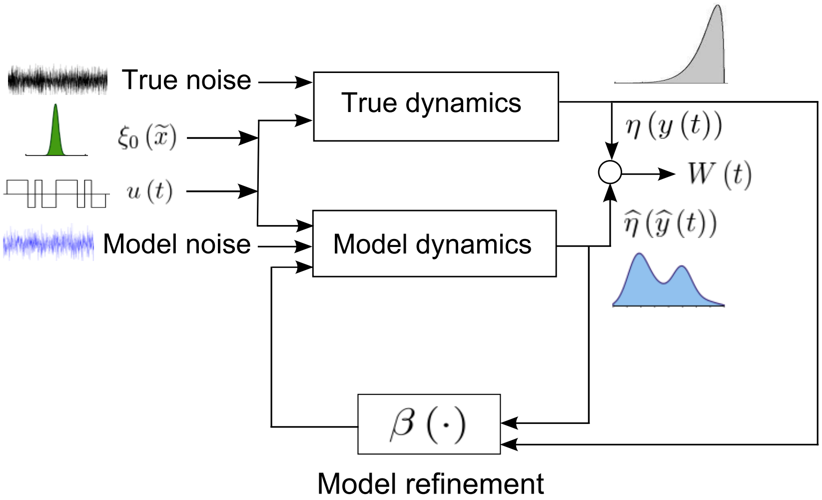


Comparing W_∞ and $\delta_\nu := \sup_\omega \kappa(\omega)$

- **Un-normalized comparison on Complex plane:**
  $\sup_\omega \kappa^{\text{proj}}(\omega) \geqslant W_\infty$

- **Normalized comparison on Riemann sphere:**
  $\overline{W}_{\mathcal{S}}\left(G, \widehat{G}\right) = \frac{2}{\pi}\left|\ \arctan||G||_2 - \arctan||\widehat{G}||_2\right|$, we find $\delta_\nu \geqslant \overline{W}_{\mathcal{S}}$ under some technical conditions.

Probabilistic model refinement



Probabilistic model refinement: formulation

- Startegy: Only refine the output model (why?)
- For example, consider proposed model $\dot{\widehat{x}} = \widehat{f}\left(\widehat{x}\right), \widehat{y} = \widehat{h}\left(\widehat{x}\right)$
- Call $\widehat{y}_j^- \triangleq \widehat{y}\left(t_j\right)$. We know η_j and $\widehat{\eta}_j$.
- We seek $\beta_j : \mathbb{R}^{n_o} \mapsto \mathbb{R}^{n_o}$, so that $\widehat{y}_j^+ = \beta_j\left(\widehat{y}_j^-\right)$ satisying $\widehat{y}_j^+ \sim \eta_j$ and $\widehat{y}_j^- \sim \widehat{\eta}_j$
- Then the refined model is: $\dot{\widehat{x}} = \widehat{f}\left(\widehat{x}\right), \widehat{y}_j^+ = \beta_j \circ \widehat{h}\left(\widehat{x}\right)$
- Seek optimal push-forward:
  $\inf\limits_{\beta(\cdot)} \underbrace{\int_{\widehat{\mathcal{Y}}} \| \beta_j\left(\widehat{y}_j^-\right) - \widehat{y}_j^- \|^2_{\ell_2(\mathbb{R}^{n_o})} \; \widehat{\eta}_j d\widehat{y}_j^-}_{J_2(\beta)}$, subject to $\eta_j = \beta_j \sharp \widehat{\eta}_j$.

Probabilistic model refinement: some background

- (Brenier, 1991): optimal $\beta^{\star}(\cdot)$ exists and is unique. Further, $\beta^{\star}(\cdot) = \nabla\psi$. Here $\psi : \mathbb{R}^{n_o} \mapsto \mathbb{R}$, and is convex.

- (Benamou & Brenier, 2001): Consider the space-time variational formulation

  $$T \inf_{(\varphi,v)} \underbrace{\int_{\mathbb{R}^{n_o}} \int_0^T \varphi(\widehat{y},s) \parallel v(\widehat{y},s) \parallel^2_{\ell_2(\mathbb{R}^{n_o})} \, d\widehat{y} ds}_{J_3(\varphi,v)}$$ subject to

  $$\frac{\partial\varphi}{\partial s} + \nabla\cdot(\varphi v) = 0,\ \varphi(\cdot,0) = \widehat{\eta},\ \varphi(\cdot,T) = \eta.$$ Then $J_3^{\star} = W^2$ and $v^{\star}$ is gradient flow.

- $$W^2 = \underbrace{\inf_{\varrho \in \mathcal{P}_2(\rho,\widehat{\rho})} J_1(\varrho)}_{\text{infinite dimensional LP}} = \underbrace{\inf_{\beta : c(\beta)=0} J_2(\beta)}_{\text{Nonlinear nonconvex optimization}} = \underbrace{T \inf_{(\varphi,v)} J_3(\varphi,v)}_{\text{Nonsmooth convex optimization}}$$

Linear Gaussian model refinement

- Theorem: Consider discrete-time deterministic LTI pairs: (A, C), $\left(\widehat{A}, \widehat{C}\right)$, sarting with $\xi_0 = \mathcal{N}\left(\mu_0, \Sigma_0\right)$. Then refined model is: $\widehat{x}_{j+1} = \widehat{A}\widehat{x}_j$, $\widehat{y}_j^+ = \Theta_j \widehat{C}\widehat{x}_j + \theta_j$.

  $\Theta_j = \Sigma_j^{1/2}\left(\Sigma_j^{1/2}\widehat{\Sigma}_j\Sigma_j^{1/2}\right)^{-1/2}\Sigma_j^{1/2}$, and $\theta_j = \mu_j - \widehat{\mu}_j$.

  The s^{th} synthetic time PDF at j^{th} physical time is: $\mathcal{N}\left(\mu_{\widehat{y}\to y}\left(s\right), \Sigma_{\widehat{y}\to y}\left(s\right)\right)$, where

$$\mu_{\widehat{y}\to y}\left(s\right) = \left[\left(1-s\right)\,\widehat{C}\widehat{A}^j + s\,CA^j\right]\mu_0,$$

$$\Sigma_{\widehat{y}\to y}\left(s\right) = \left[\left(1-s\right)\,I + s\,\Theta\left(j\right)\right]\left(\left(\widehat{C}\widehat{A}^j\right)\Sigma_0\left(\widehat{C}\widehat{A}^j\right)^\top\right)\left[\left(1-s\right)\,I + s\,\Theta\left(j\right)\right].$$

Linear Gaussian model refinement

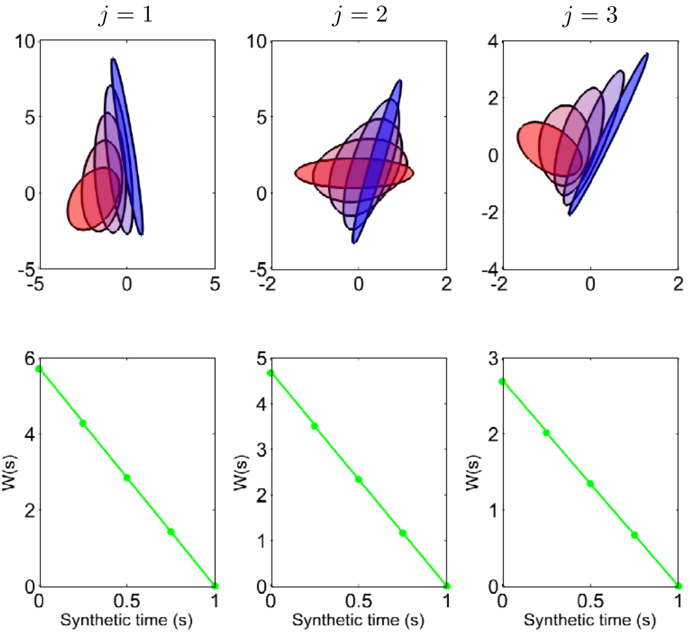
- Theorem: Consider discrete-time deterministic LTI pairs: (A, C), $\left(\widehat{A}, \widehat{C}\right)$, sarting with $\xi_0 = \mathcal{N}\left(\mu_0, \Sigma_0\right)$. Then refined model is: $\widehat{x}_{j+1} = \widehat{A}\widehat{x}_j$, $\widehat{y}_j^+ = \Theta_j \widehat{C}\widehat{x}_j + \theta_j$.
  $\Theta_j = \Sigma_j^{1/2}\left(\Sigma_j^{1/2}\widehat{\Sigma}_j\Sigma_j^{1/2}\right)^{-1/2}\Sigma_j^{1/2}$, and $\theta_j = \mu_j - \widehat{\mu}_j$.
  The s^{th} synthetic time PDF at j^{th} physical time is: $\mathcal{N}\left(\mu_{\widehat{y}\to y}(s), \Sigma_{\widehat{y}\to y}(s)\right)$, where

$$\mu_{\widehat{y}\to y}(s) = \left[(1-s)\ \widehat{C}\widehat{A}^j + s\, CA^j\right]\mu_0,$$

$$\Sigma_{\widehat{y}\to y}(s) = \left[(1-s)\ I + s\,\Theta(j)\right]\left(\left(\widehat{C}\widehat{A}^j\right)\Sigma_0\left(\widehat{C}\widehat{A}^j\right)^\top\right)\left[(1-s)\ I + s\,\Theta(j)\right].$$

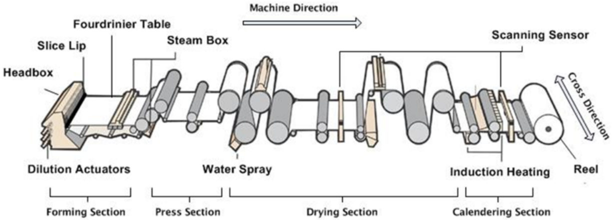
- Example: $A = \begin{bmatrix} 0.4 & -0.1 \\ 2 & 0.6 \end{bmatrix}$, $\widehat{A} = \begin{bmatrix} 0.2 & -0.7 \\ -0.7 & 0.1 \end{bmatrix}$, $C = \begin{bmatrix} -1 & 0.03 \\ -0.2 & 0.8 \end{bmatrix}$, $\widehat{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. $\mu_0 = \{1, 3\}^\top, \Sigma_0 = \begin{bmatrix} 10 & 6 \\ 6 & 7 \end{bmatrix}$

Linear Gaussian model refinement: example



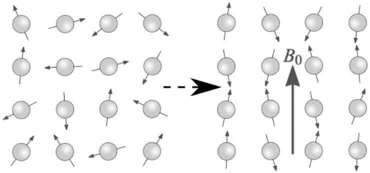
Application: finite horizon density tracking

- Process industry applications

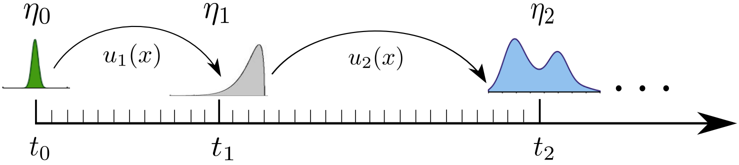


Source: Chu *et.al.* (2011), "Model Predictive Control and Optimization for Papermaking Process", doi: 10.5772/18535.

- NMR spectroscopy and MRI applications



Application: linear Gaussian tracking

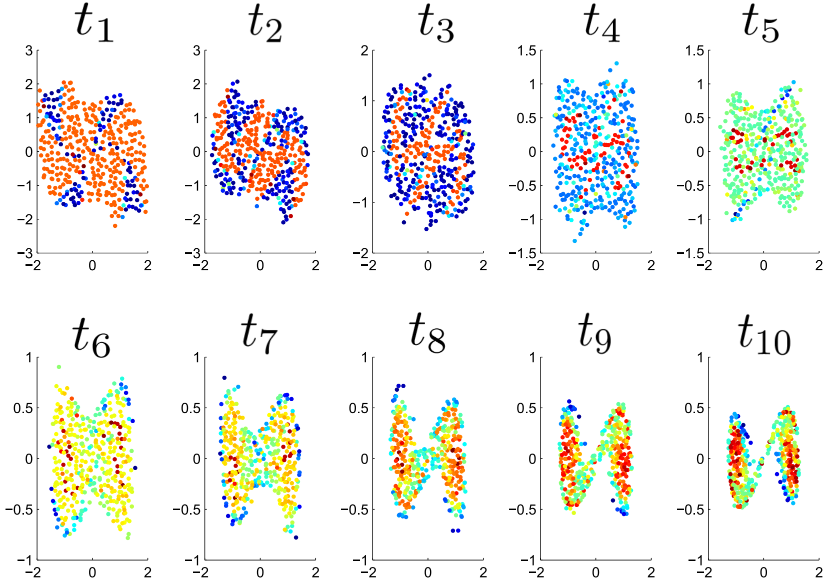


- Theorem: Consider tracking Gaussians $\eta_j = \mathcal{N}\left(\mu_j, \Sigma_j\right)$, under LTI structure $x_{j+1} = Ax_j + Bu_j$. The state feedback $u_j^{\star} \triangleq u^{\star}\left(x_j\right)$ guaranteeing optimal transport
  1. exists iff $\left(\Theta_j - A\right), \theta_j \in \ker\left(I - BB^{\dagger}\right)$
  2. if exists, then must be affine form $u_j^{\star} = K_j x_j + \kappa_j$, where $K_j = B^{\dagger}\left(\Theta_j - A\right) - \left(I - BB^{\dagger}\right) R$, and $\kappa_j = B^{\dagger}\theta_j - \left(I - BB^{\dagger}\right) r$
  3. is unique, if B is full rank.

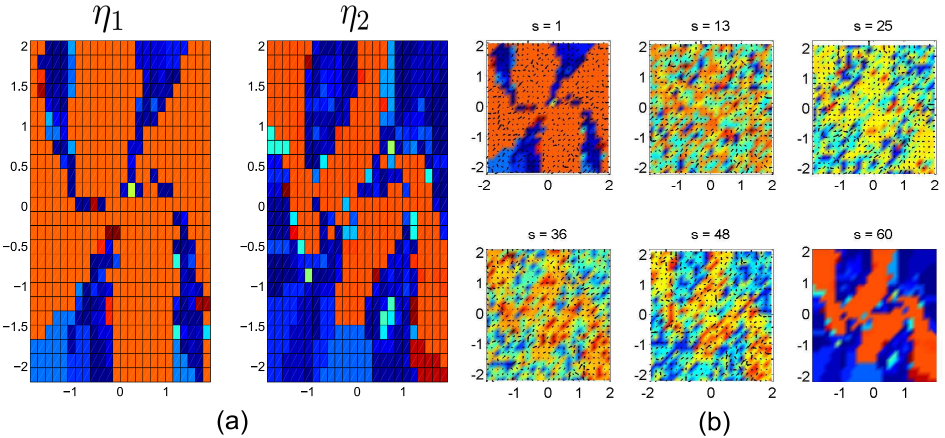
Application: data-driven modeling

- Duffing vector field (unknown to modeler) to generate data: $\dot{x}_1 = x_2, \quad \dot{x}_2 = -\alpha x_1^3 - \beta x_1 - \delta x_2$, $y = \{x_1, x_2\}^\top$, $\alpha = 1$, $\beta = -1$, $\delta = 0.5$
- Lioville MOC with 500 samples from $\xi_0 = \mathcal{U}\left([-2, 2]^2\right)$
- 10 snapshot data $\{t_j, \eta_j\}_{j=1}^{10}$
- Subdivided each of the 10 intervals $[t_j, t_{j+1})$, $j = 0, \ldots, 9$ into 60 sub-intervals.
- Want to compute optimal transport vector field $v_{j \to j+1}$ for each of those intervals

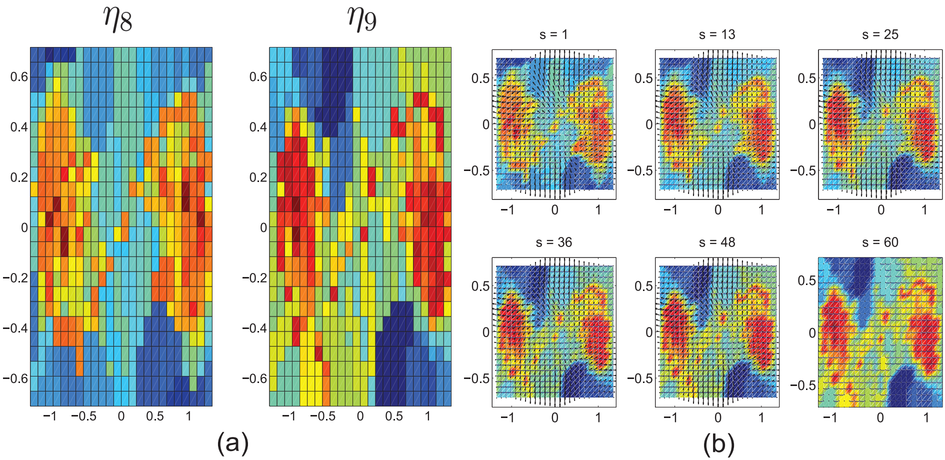
Application: data-driven modeling



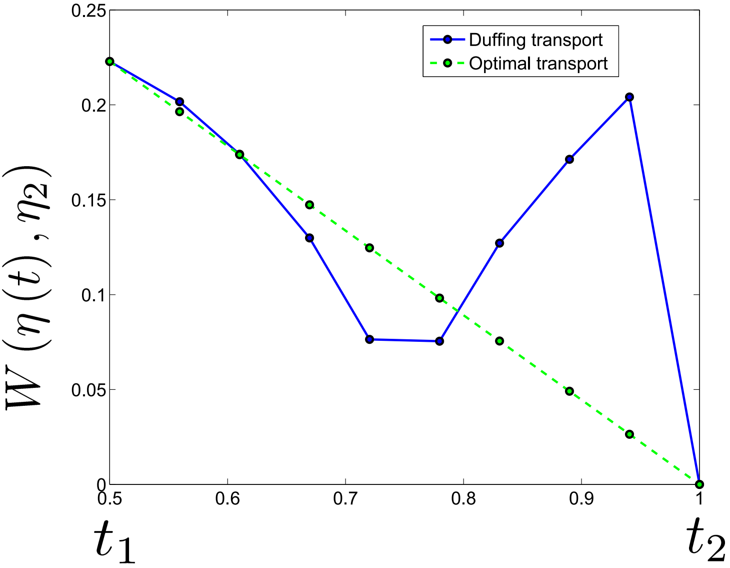
Application: data-driven modeling of $v_{1\to2}$



Application: data-driven modeling of $v_{8\to 9}$



Application: Duffing transport vs. optimal transport for $[t_1, t_2)$



Conclusions

- Unifying framework for probabilistic V&V
- Transport-theoretic Wasserstein distance as (in)validation measure
- Probabilistic framework for model refinement
- Possible extensions:
  (i) compositionality in probabilistic V&V
  (ii) optimal transport based model reduction
  (iii) application to ensemble tracking, e.g. MRI and NMR

Conclusions

- Unifying framework for probabilistic V&V
- Transport-theoretic Wasserstein distance as (in)validation measure
- Probabilistic framework for model refinement
- Possible extensions:
  (i) compositionality in probabilistic V&V
  (ii) optimal transport based model reduction
  (iii) application to ensemble tracking, e.g. MRI and NMR

Thank You

Backup Slides

Some details on noise KL expansion

- KL expansion of $\eta(\omega, t) \in L_2(\Omega, \mathscr{F}, \mathbb{P})$, is:
$\eta(\omega, t) \overset{\text{m.s.}}{=} \sum_{i=1}^{\infty} \sqrt{\Lambda_i} \zeta_i(\omega) e_i(t)$
- Covariance function $C(t_1, t_2) \triangleq$
$\text{cov}(\eta(\omega, t_1) - \mathbb{E}[\eta(\omega, t_1)], \eta(\omega, t_2) - \mathbb{E}[\eta(\omega, t_2)])$,
$t_1, t_2 \in [0, T]$
- Fredholm integral eqn. of second kind:
$\int_0^T C(t_1, t_2) e_i(t_2) dt_2 = \Lambda_i e_i(t_1)$

Noise $\mathcal{W}(\omega, t)$ in SDE	$C(t_1, t_2)$ **for** $\mathcal{W}(\omega, t)$	$\eta(\omega, t)$	**KL expansion of** $\eta(\omega, t)$**,** $0 < t \leqslant T$
Wiener process	$\sigma^2 (t_1 \wedge t_2)$	GWN	$\sqrt{\frac{2}{T}} \sum_{i=1}^{\infty} \zeta_i(\omega) \cos\left(\left(i - \frac{1}{2}\right) \frac{\pi t}{T}\right)$
Compound Poisson process	$\lambda\sigma^2 (t_1 \wedge t_2) + (\lambda\mu)^2 t_1 t_2$	PWN	$\sum_{i=1}^{\infty} \bar{\zeta}_i(\omega) \frac{\frac{2}{\beta_i} \sqrt{\Lambda_i}}{\sqrt{2T - \beta_i \sin \frac{2T}{\beta_i}}} \cos\left(\frac{t}{\beta_i}\right)$

On the KL expansion of compound Poisson process

$\overline{\zeta}_i(\omega)$ are i.i.d random variables from $\mathcal{N}(0,1)$, $\beta_i \triangleq \sqrt{\frac{\Lambda_i}{\lambda\sigma^2}}$, $\forall\, i \in \mathbb{N}$, and $\Lambda_i > 0$ solves

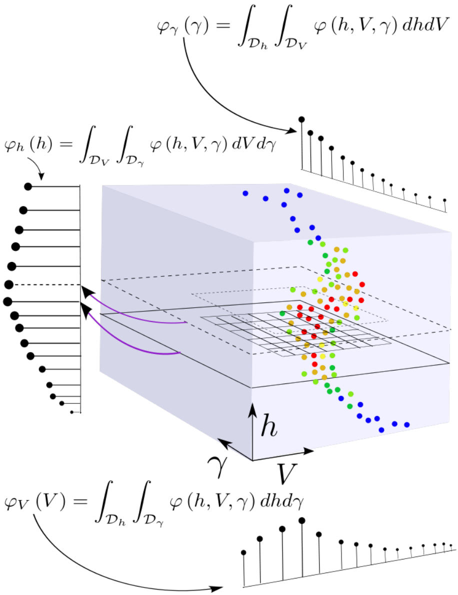
$$\tan\left(\sigma T\sqrt{\frac{\lambda}{\Lambda_i}}\right) = \left[1 + \frac{1}{\lambda T}\left(\frac{\sigma}{\mu}\right)^2\right]\left(\sigma T\sqrt{\frac{\lambda}{\Lambda_i}}\right),$$

where the parameters $\lambda, \sigma, \mu, T > 0$.

Application of KL+ MOC algorithm to nonlinear estimation

- Compute particle filter posterior: $\xi^{+}_{\text{Particle}}\left(x(t), t\right)$
- Compute posterior from our proposed method: $\xi^{+}_{\text{KLMOC}}\left(x(t), t\right)$
- Compute the "distances" of $\xi^{+}_{\text{Particle}}\left(x(t), t\right)$ and $\xi^{+}_{\text{KLMOC}}\left(x(t), t\right)$ from the true posterior $\xi^{+}_{\text{True}}\left(x(t), t\right)$
- Distance metric on the space of PDFs: Wasserstein distance W
- $W \triangleq \left(\inf\limits_{\gamma \in \mathcal{M}(\varphi, \widehat{\varphi})} \mathbb{E}\left[\| x - \widehat{x} \|_2^2\right]\right)^{\frac{1}{2}}$, $\mathcal{M}\left(\varphi, \widehat{\varphi}\right) = \{\text{All joint PDFs}\, \gamma\left(x, \widehat{x}\right) \,: x \sim \varphi, \widehat{x} \sim \widehat{\varphi}\}$.
- W = minimum amount of work needed to morph one PDF to other

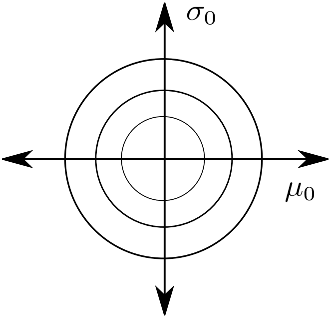
Marginal computation



Worst case initial PDF: model discrimination problem

Scalar linear systems

- Deterministic (continuous and discrete time): gap $\propto \sqrt{m_{20}}$
- Construction:



- Stochastic: depends on m_{20}, m_{10} and
  $s\left(F_0\right) := \sqrt{2}\,\mathbb{E}\left[x_0 \operatorname{erf}^{-1}\left(2F_0\left(x_0\right) - 1\right)\right]$.

Vector linear systems: Conjecture

- Deterministic: gap $\propto \sqrt{\parallel \mu_0 \parallel_2^2 + \left(\operatorname{tr}\left(P_0\right)\right)^2}$
- Can prove for Gaussian family

Worst case initial PDF: example

Uniform PDF $\not\Leftrightarrow \sup_{\rho_0} {}_2W_2(t)$