SAND2022-4527C

Improving Digital Twins by Learning from a Fleet of Assets

Tom Seidl and John Jakeman
dtseidl@sandia.gov

SIAM Conference on Uncertainty Quantification
April 12-15 2022

Supported by the Laboratory Directed Research and Development program at Sandia National Laboratories.

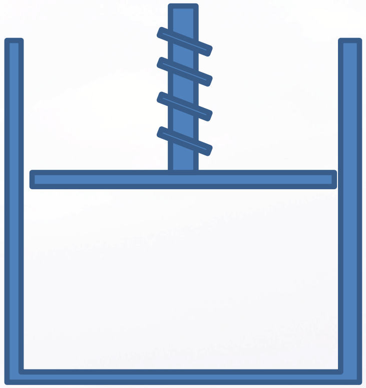
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Motivation and Main Idea

Digital Twins

- Single asset:



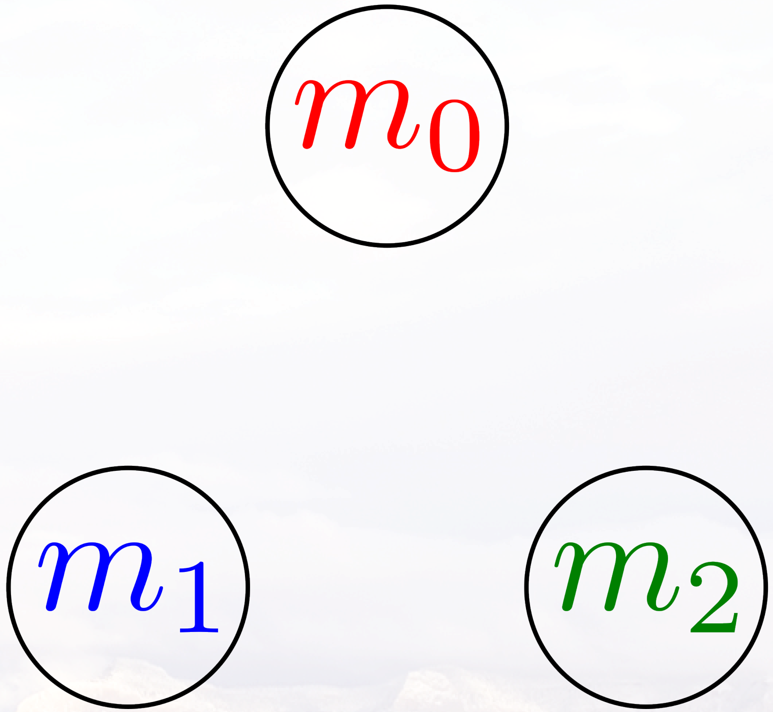
- Corresponding digital twin:

Fleet of Assets

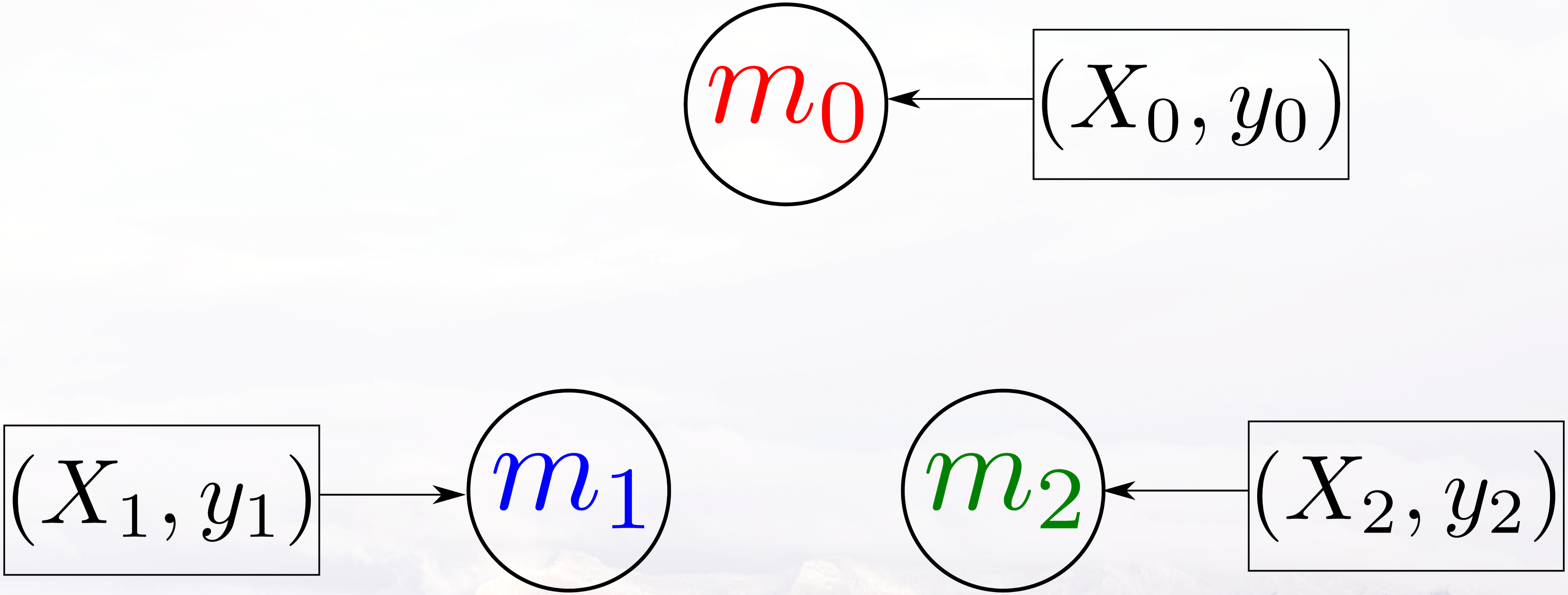
- Similar assets, different (unknown) operating conditions.

- Fuse information from the fleet to improve each individual digital twin.

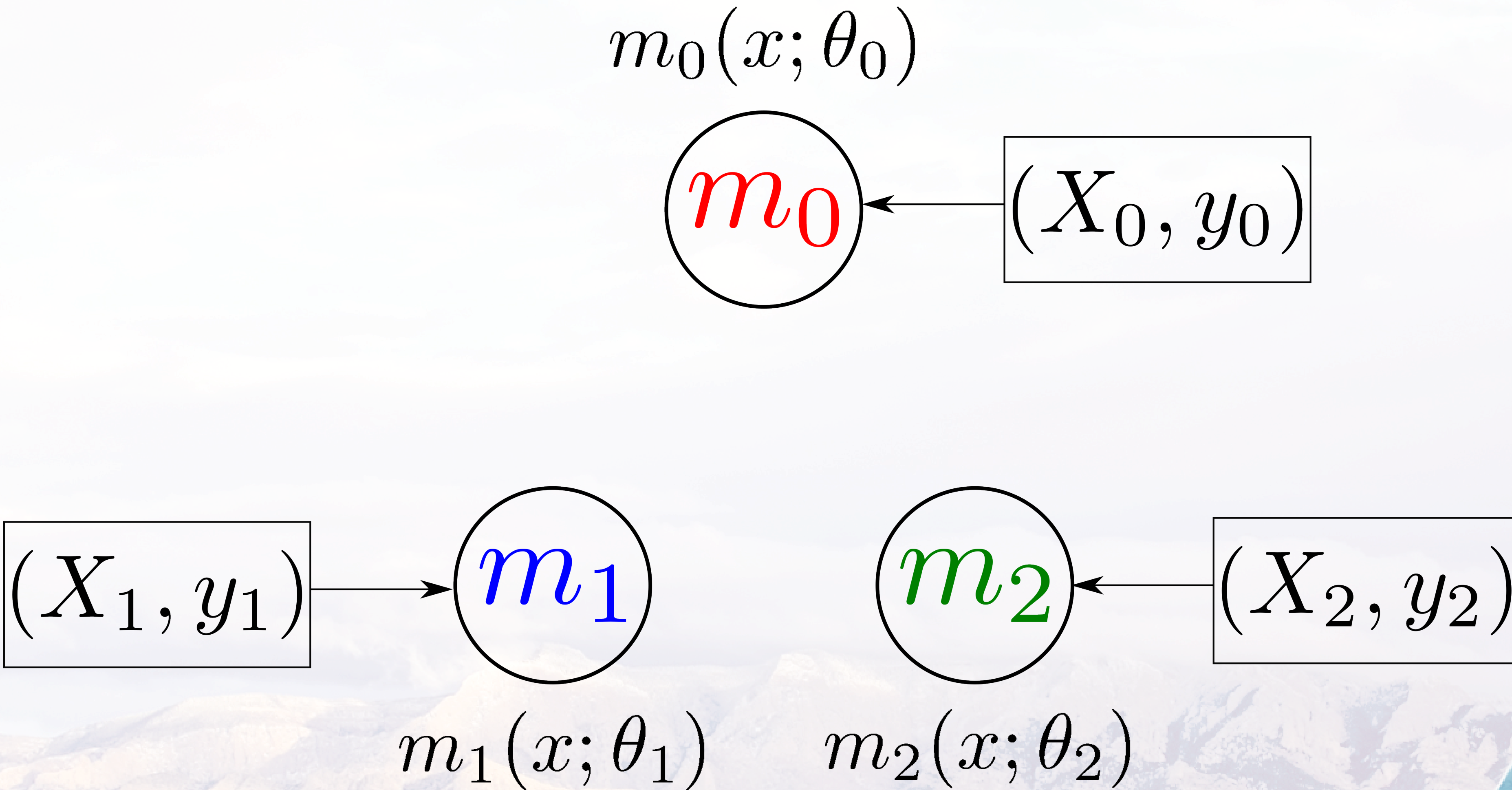
Individual Digital Twins



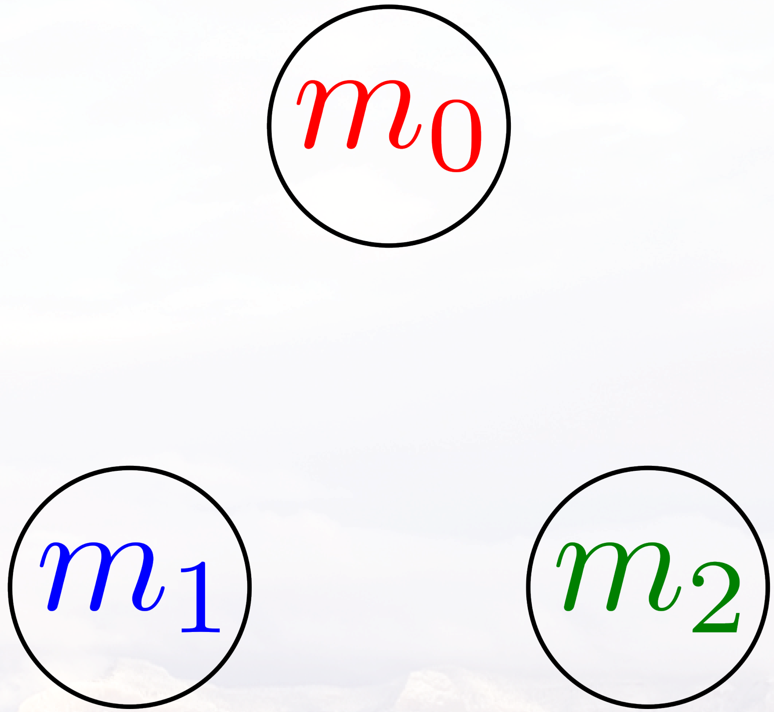
Individual Digital Twins



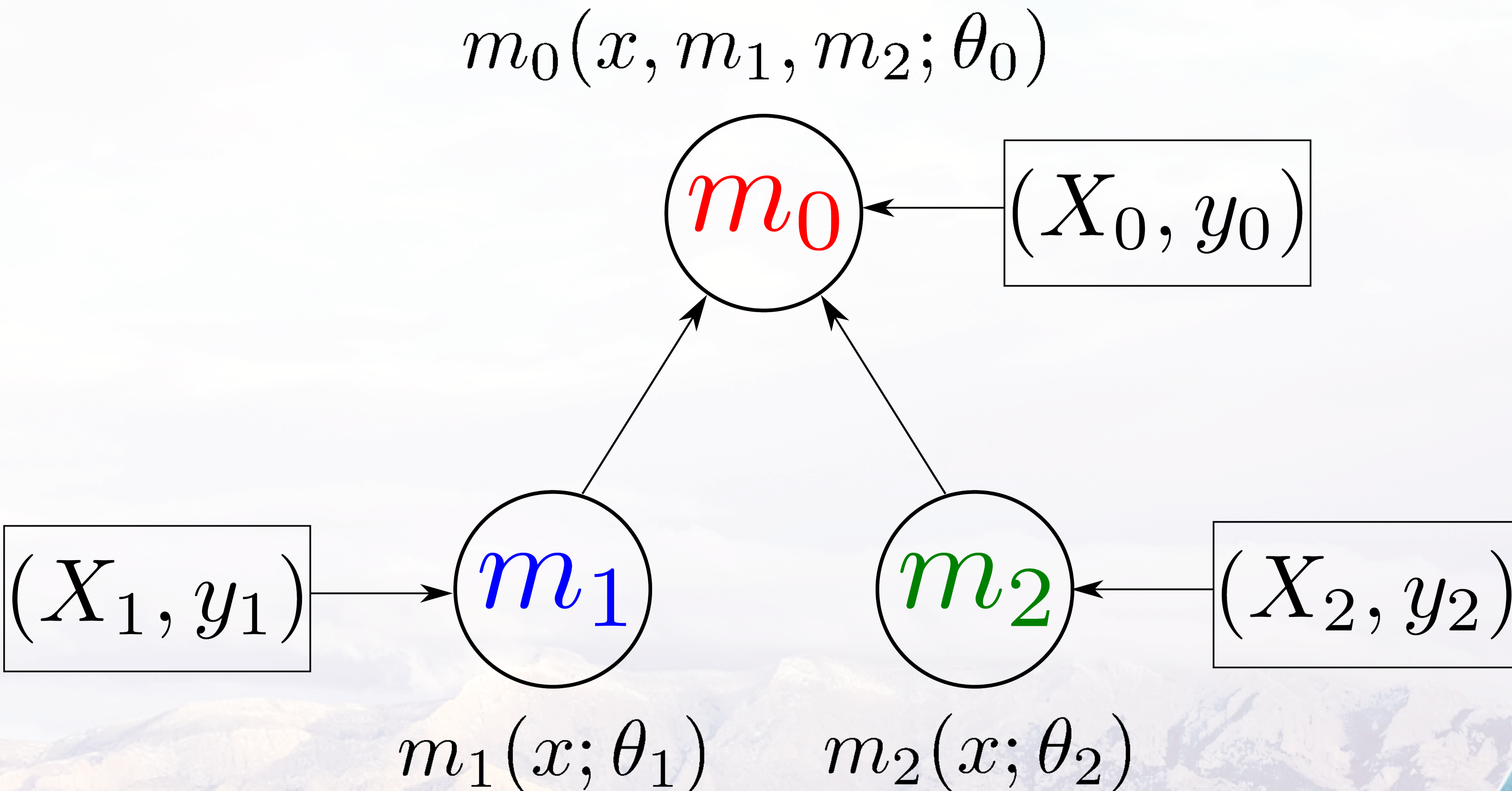
Individual Digital Twins



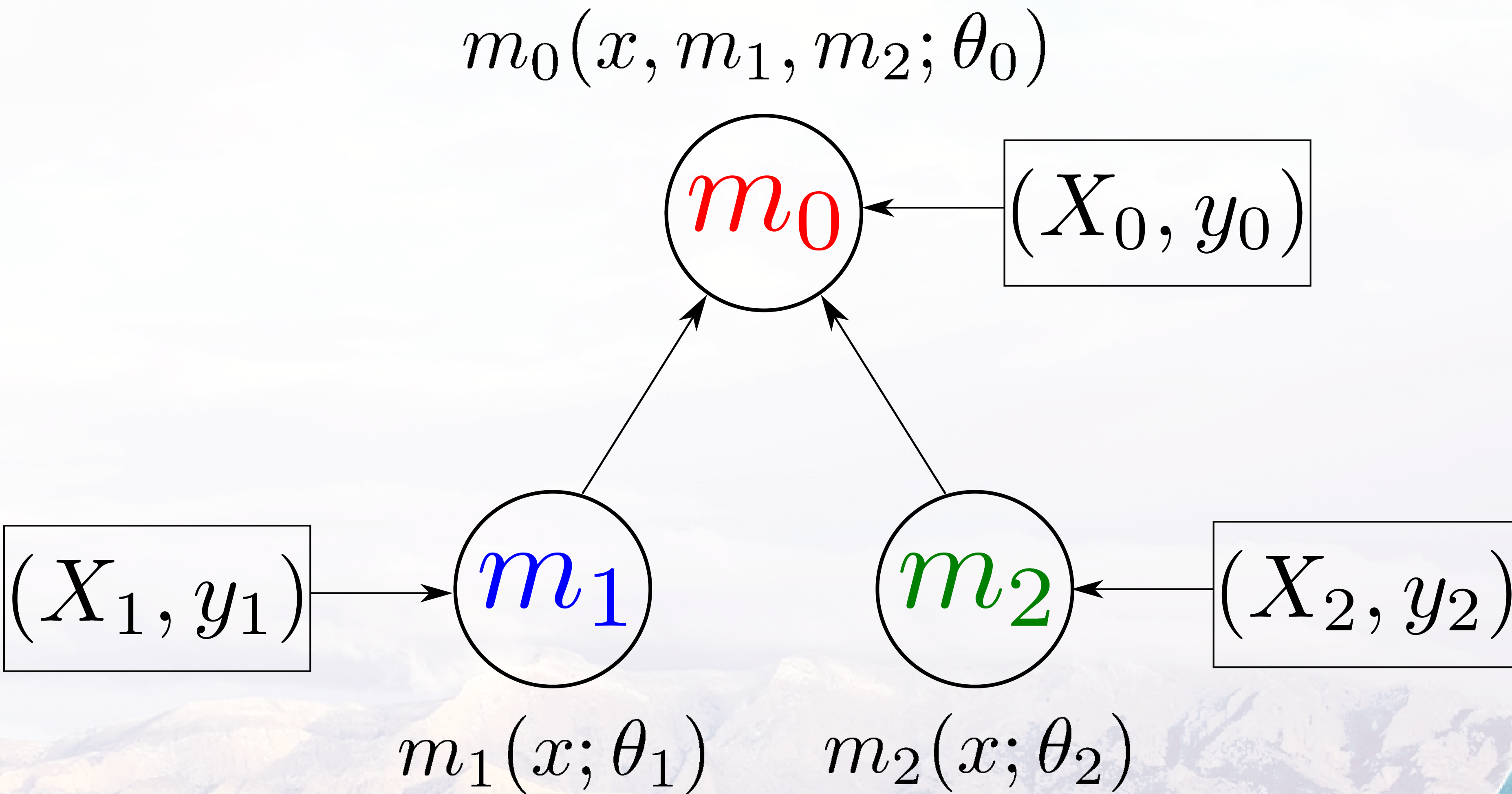
Network from the Fleet of Assets



Network from the Fleet of Assets



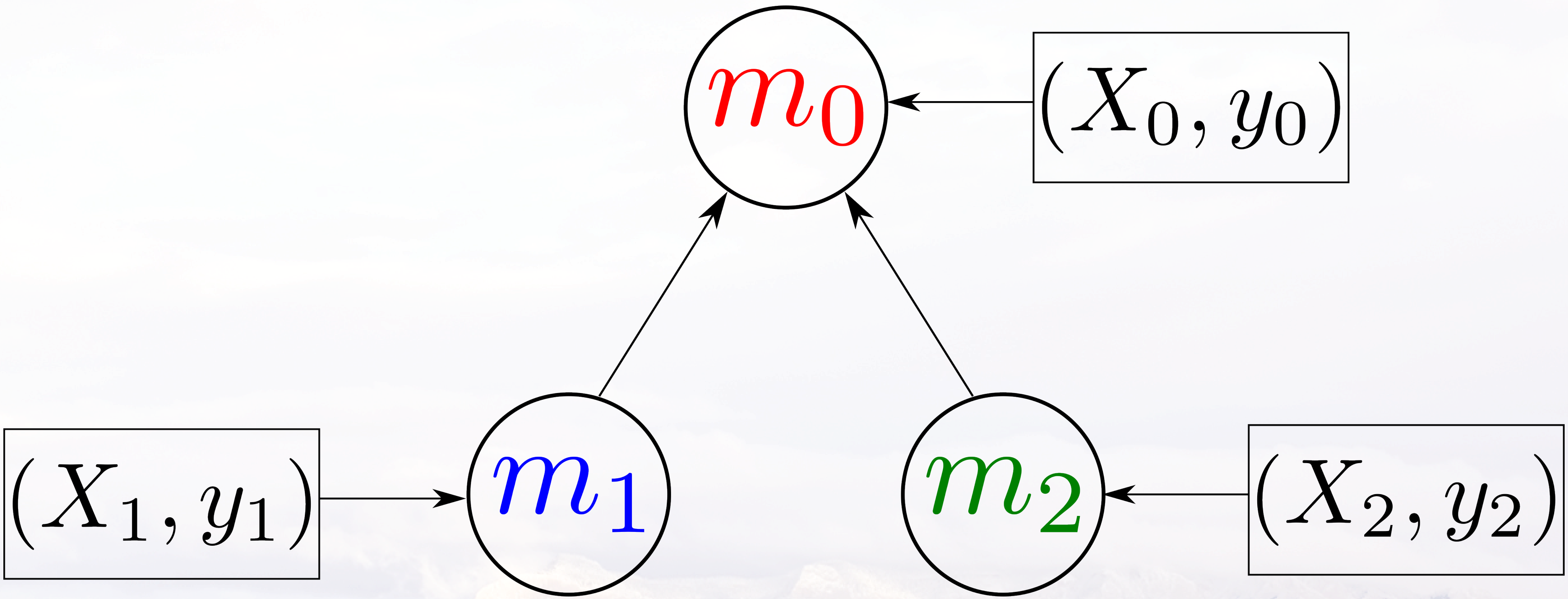
Network from the Fleet of Assets



- We can repeat this procedure for any asset of interest (AoI) in the fleet.

Specialization to Gaussian Processes

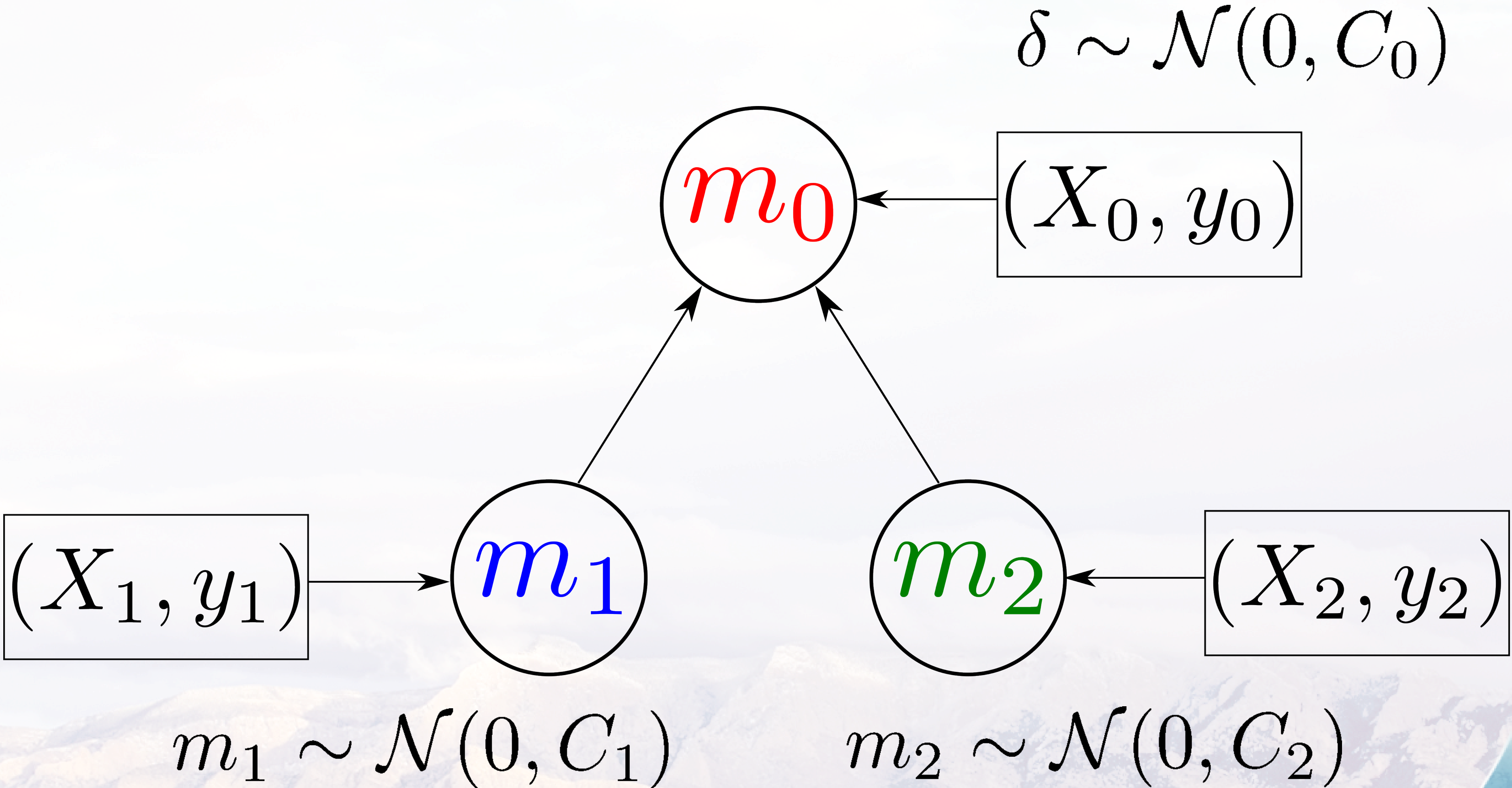
FA Network of Gaussian Processes



$$m_1 \sim \mathcal{N}(0, C_1) \qquad m_2 \sim \mathcal{N}(0, C_2)$$

- Introduce independent GP priors for the peers and a discrepancy GP.

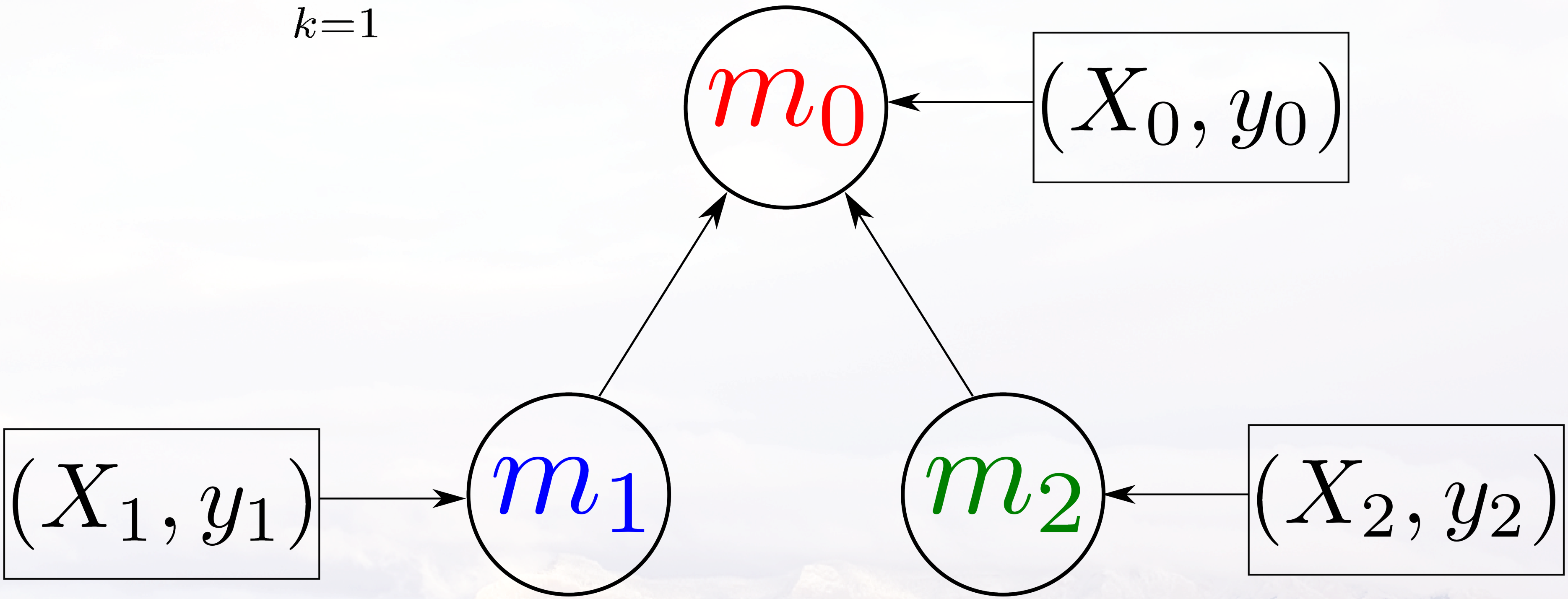
FA Network of Gaussian Processes



- Introduce independent GP priors for the peers and a discrepancy GP.

FA Network of Gaussian Processes

$$m_0(x) = \sum_{k=1}^{N_{\text{peer}}} \rho_k(x) m_k(x) + \delta(x) \qquad \delta \sim \mathcal{N}(0, C_0)$$



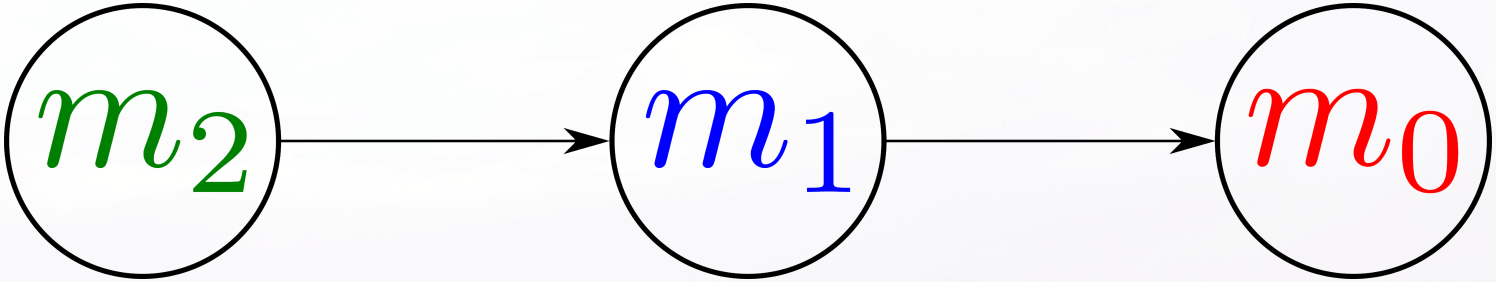
$$m_1 \sim \mathcal{N}(0, C_1) \qquad m_2 \sim \mathcal{N}(0, C_2)$$

- Treat ρ_k polynomial coefficients as hyperparameters and form a sparse prior for the covariance between asset GPs.

Sandia National Laboratories

Related Work

- Hierarchical / auto-regressive GPs (Kennedy and O'Hagan, Le Gratiet and Garnier):



  - Requires nested training data sets and noiseless observations for the parent models.
  - Hierarchical ordering not appropriate for a fleet of assets.

- MF Nets (Gorodetsky et al.): Our work is an instance of a peer directed acyclic graph of Gaussian processes.

FA Network of GPs via Co-Kriging

- The covariance matrix has a block structure that is determined by computing the covariance between the approximations $m_k(\boldsymbol{x}),\ k=1,\dots N_{\text{peer}}$ and $m_0(\boldsymbol{x}) = \sum_{k=1}^{N_{\text{peer}}} \rho_k(\boldsymbol{x}) m_k(\boldsymbol{x}) + \delta(\boldsymbol{x})$:

$$\mathbb{E}\left[m_k m_k\right] = C_k(X_k, X_k), \quad k = 1, \dots N_{\text{peer}},$$

$$\mathbb{E}\left[m_0 m_k\right] = \left[\mathbb{1}_{N_0} \otimes \boldsymbol{\rho}_k(X_k)\right] \odot C_k(X_0, X_k), \quad k = 1, \dots N_{\text{peer}},$$

$$\mathbb{E}\left[m_0 m_0\right] = C_0(X_0, X_0) + \sum_{k=1}^{N_{\text{peer}}} \left[\boldsymbol{\rho}_k(X_0) \otimes \boldsymbol{\rho}_k(X_0)\right] \odot C_k(X_0, X_0).$$

- The matrices $C_k(X_i, X_j),\ k = 1, \dots, N_{\text{peer}}$ are obtained by evaluating the kernels for $m_k(\boldsymbol{x})$ at all pairs of points in X_i, X_j.

- Similarly, the matrix $C_0(X_0, X_0)$ is computed by evaluating the kernel for the discrepancy GP $\delta(\boldsymbol{x})$.

Hyperparameter Estimation

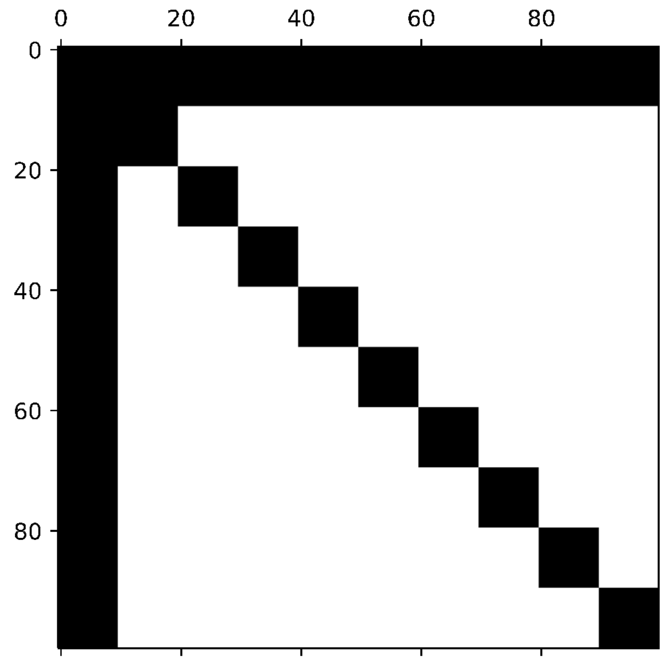
- Negative marginal log-likelihood:

$$\text{NLL} = \frac{1}{2}\log\left(\det\ \boldsymbol{C}\right) + \frac{1}{2}\boldsymbol{y}^T\boldsymbol{C}^{-1}\boldsymbol{y} + \frac{N}{2}\log\left(2\pi\right)$$

- Computational bottlenecks:
  - Cholesky factorization – $O\left(N^3\right)$.
  - NLL typically non-convex – multiple optimization runs from different starting points.
  - All-at-once and sequential approaches.

Linear Algebra for the FA Network of GPs

- The block covariance matrix is sparse.



- Efficient linear algebra techniques reduce the overall complexity of hyperparameter learning to $O(\max(N_k)^3)$.

Numerical Examples

Analytical Assets Example

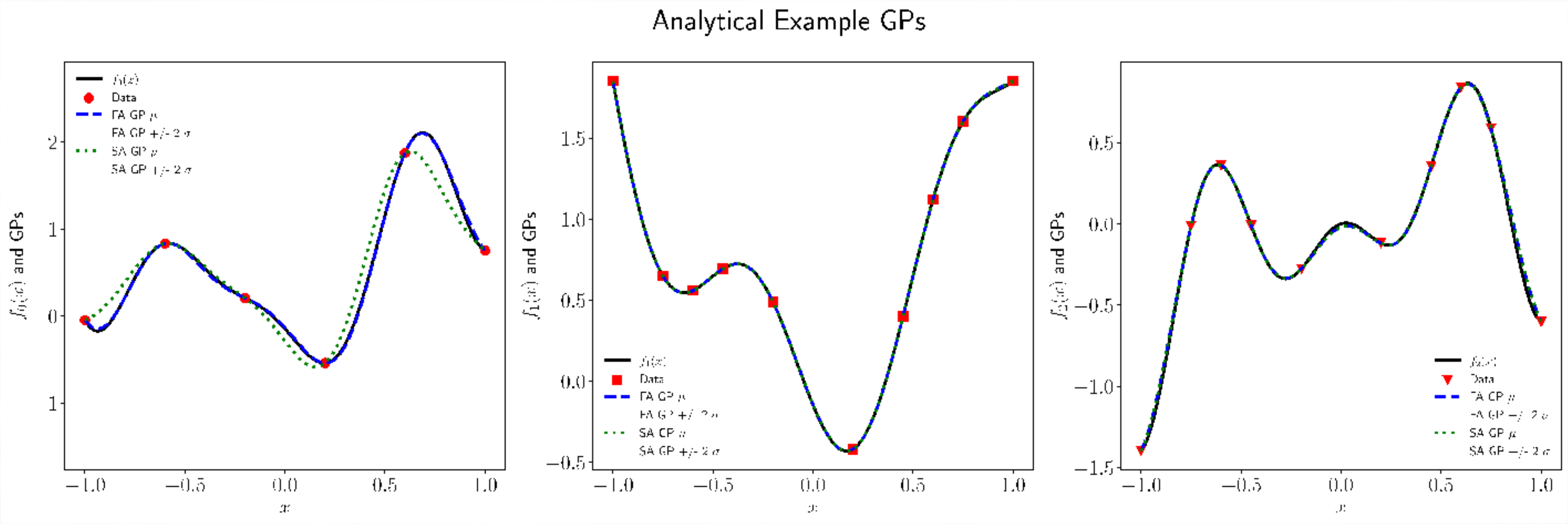
$$f_2(x) = 0.4x - x \sin\left(\frac{5\pi x}{2}\right)$$

$$f_1(x) = 2x^2 - \frac{1}{2}\sin\left(2\pi x + 0.3\right)$$

$$f_0(x) = -\frac{1}{2}x^3 \sin\left(\frac{\pi x}{2}\right) + f_1(x) + f_2(x)$$

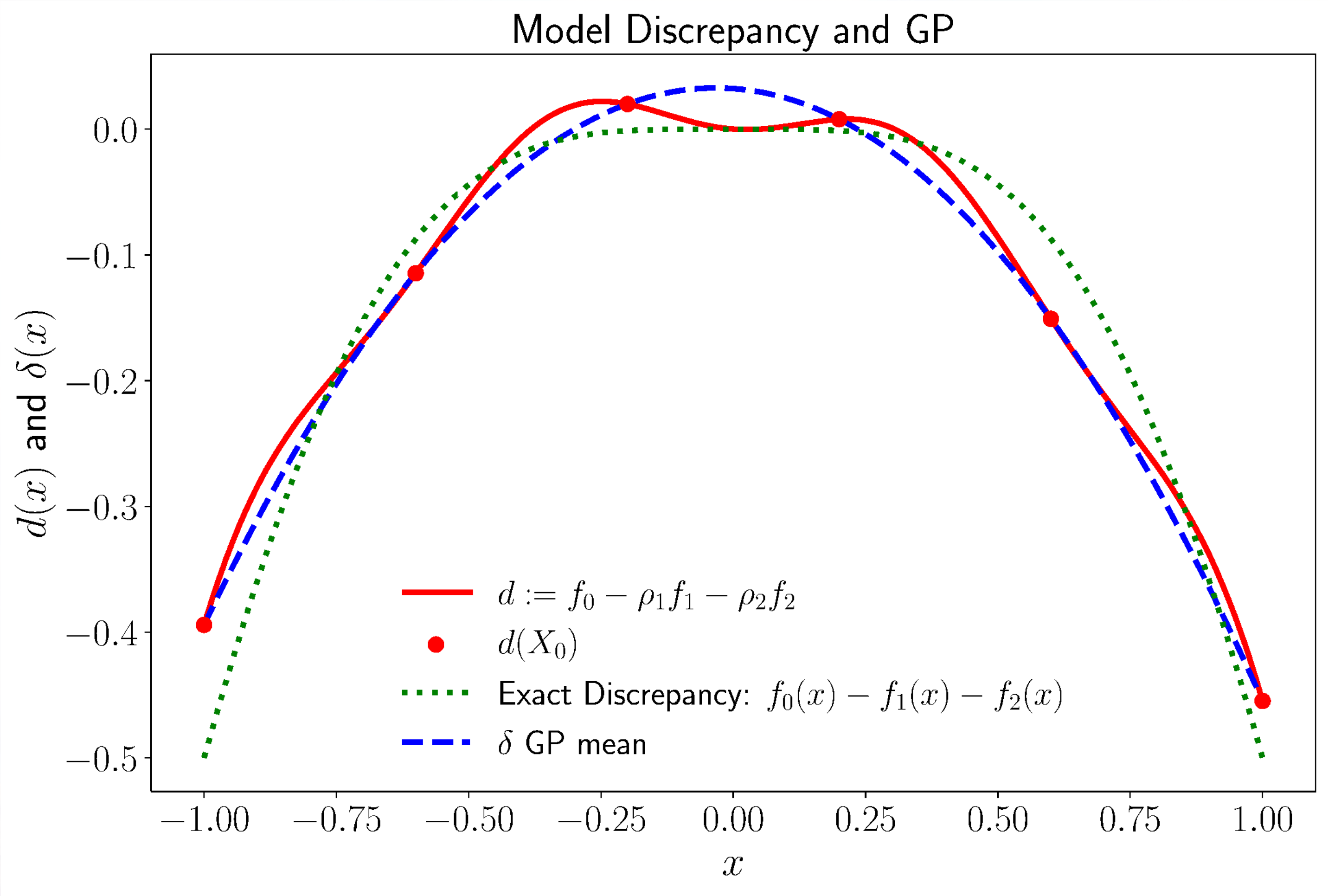
- Analytical functions for the assets that conform to the additive / multiplicative discrepancy structure.

Nested Samples Analytical DT GP Results



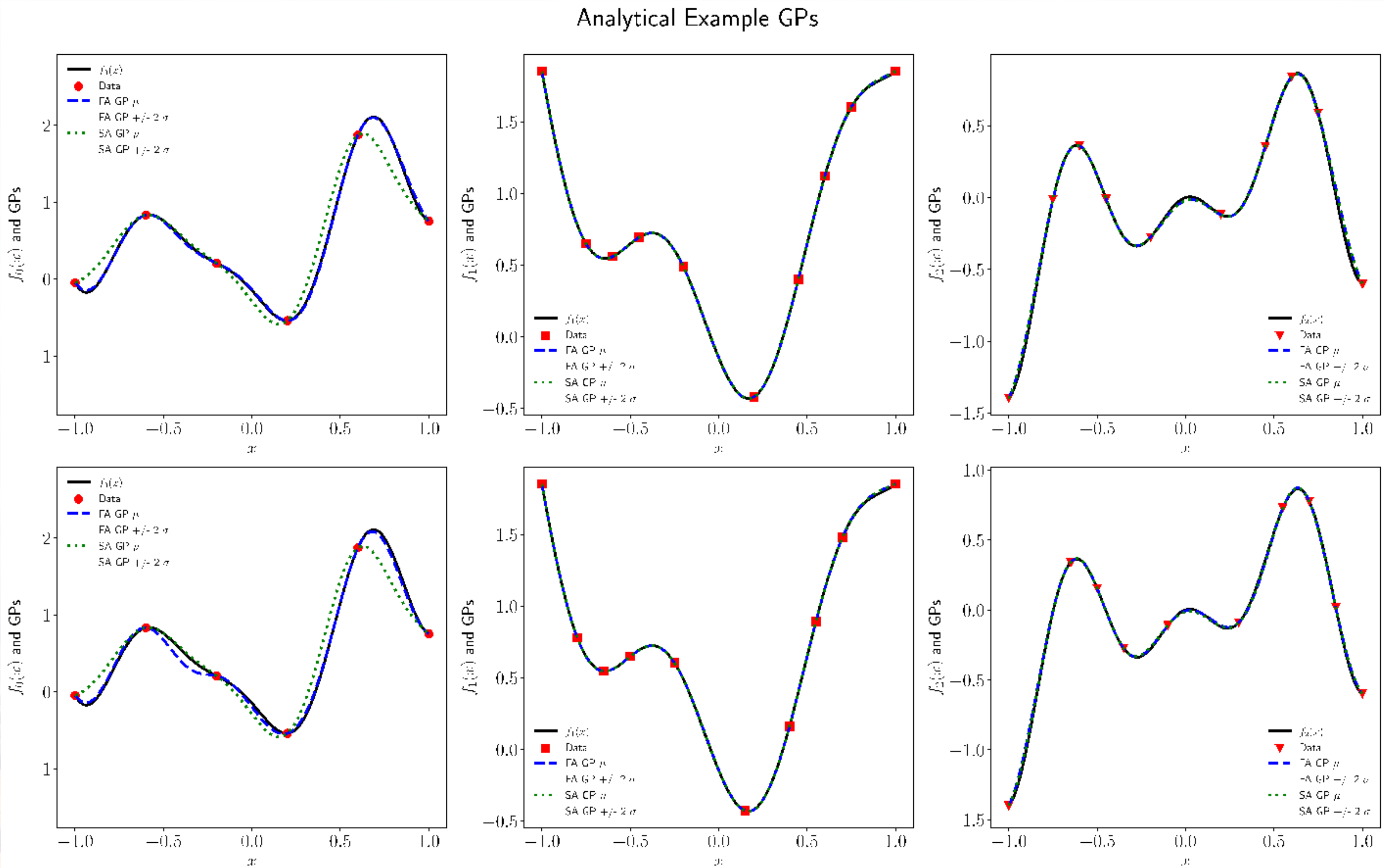
- Model 0 is the asset of interest (AoI) in these results.
- Fleet of assets (FA) vs single asset (SA).
- MLE results in $\rho_1 = 1.0, \rho_2 = 1.08.$

Nested Samples Analytical DT GP Results



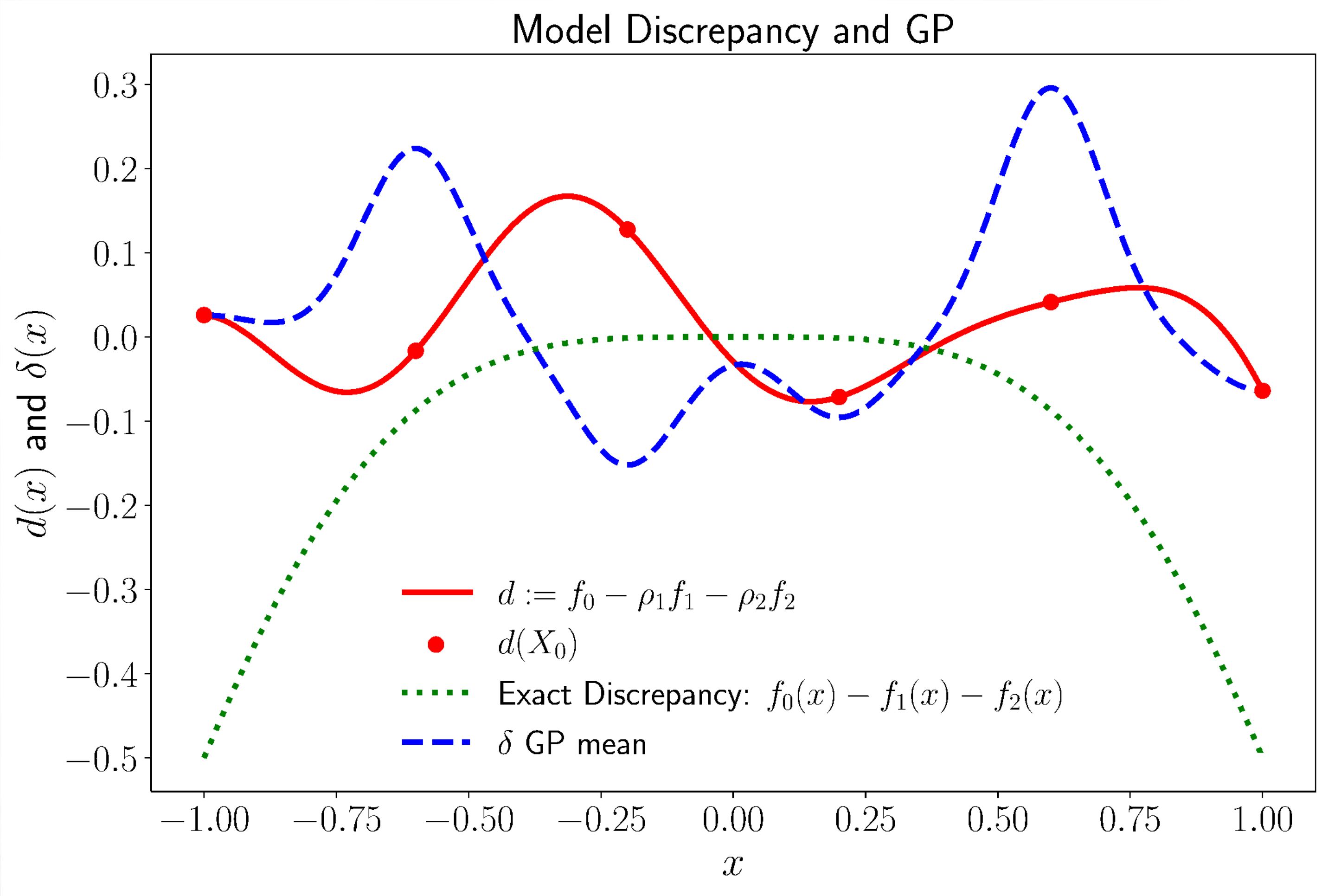
- The discrepancy GP approximates the discrepancy when the training data is nested and peer datasets are noiseless.

Non-nested Samples Analytical DT GP Results



- FA approach still performs better than single asset, but …

Non-nested Samples Analytical DT GP Results



- ... the discrepancy GP is doing something different.

Piston Model

- Benchmark problem for surrogate modeling: https://www.sfu.ca/~ssurjano/piston.html .

$$C(\boldsymbol{x}) = 2\pi \sqrt{\frac{M}{k + S^2 \frac{P_0 V_0}{T_0} \frac{T_a}{V^2}}}$$

$$V = \frac{S}{2k} \left(\sqrt{A^2 + 4k \frac{P_0 V_0}{T_0} T_a} - A \right)$$

$$A = P_0 S + 19.62 M - \frac{k V_0}{S}$$

$M \in [30, 60]$	piston weight (kg)
$S \in [0.005, 0.020]$	piston surface area (m^2)
$V_0 \in [0.002, 0.010]$	initial gas volume (m^3)
$k \in [1000, 5000]$	spring coefficient (N/m)
$P_0 \in [90000, 110000]$	atmospheric pressure ($\mathrm{N/m}^2$)
$T_a \in [290, 296]$	ambient temperature (K)
$T_0 \in [340, 360]$	filling gas temperature (K)

Piston Model: 1D Problem

- Benchmark problem for surrogate modeling: https://www.sfu.ca/~ssurjano/piston.html .

$$C(\boldsymbol{x}) = 2\pi\sqrt{\frac{M}{k + S^2 \frac{P_0 V_0}{T_0} \frac{T_a}{V^2}}}$$

$$V = \frac{S}{2k}\left(\sqrt{A^2 + 4k\frac{P_0 V_0}{T_0} T_a} - A\right)$$

Input variable

$$A = P_0 S + 19.62M - \frac{kV_0}{S}$$

$M \in [30, 60]$	piston weight (kg)
$S \in [0.005, 0.020]$	piston surface area (m^2)
$V_0 \in [0.002, 0.010]$	initial gas volume (m^3)
$k \in [1000, 5000]$	spring coefficient (N/m)
$P_0 \in [90000, 110000]$	atmospheric pressure (N/m^2)
$T_a \in [290, 296]$	ambient temperature (K)
$T_0 \in [340, 360]$	filling gas temperature (K)

Sandia National Laboratories

Piston Model: 1D Problem

- Benchmark problem for surrogate modeling: https://www.sfu.ca/~ssurjano/piston.html .

$$C(\boldsymbol{x}) = 2\pi\sqrt{\frac{M}{k + S^2 \frac{P_0 V_0}{T_0} \frac{T_a}{V^2}}}$$

$$V = \frac{S}{2k}\left(\sqrt{A^2 + 4k\frac{P_0 V_0}{T_0} T_a} - A\right)$$

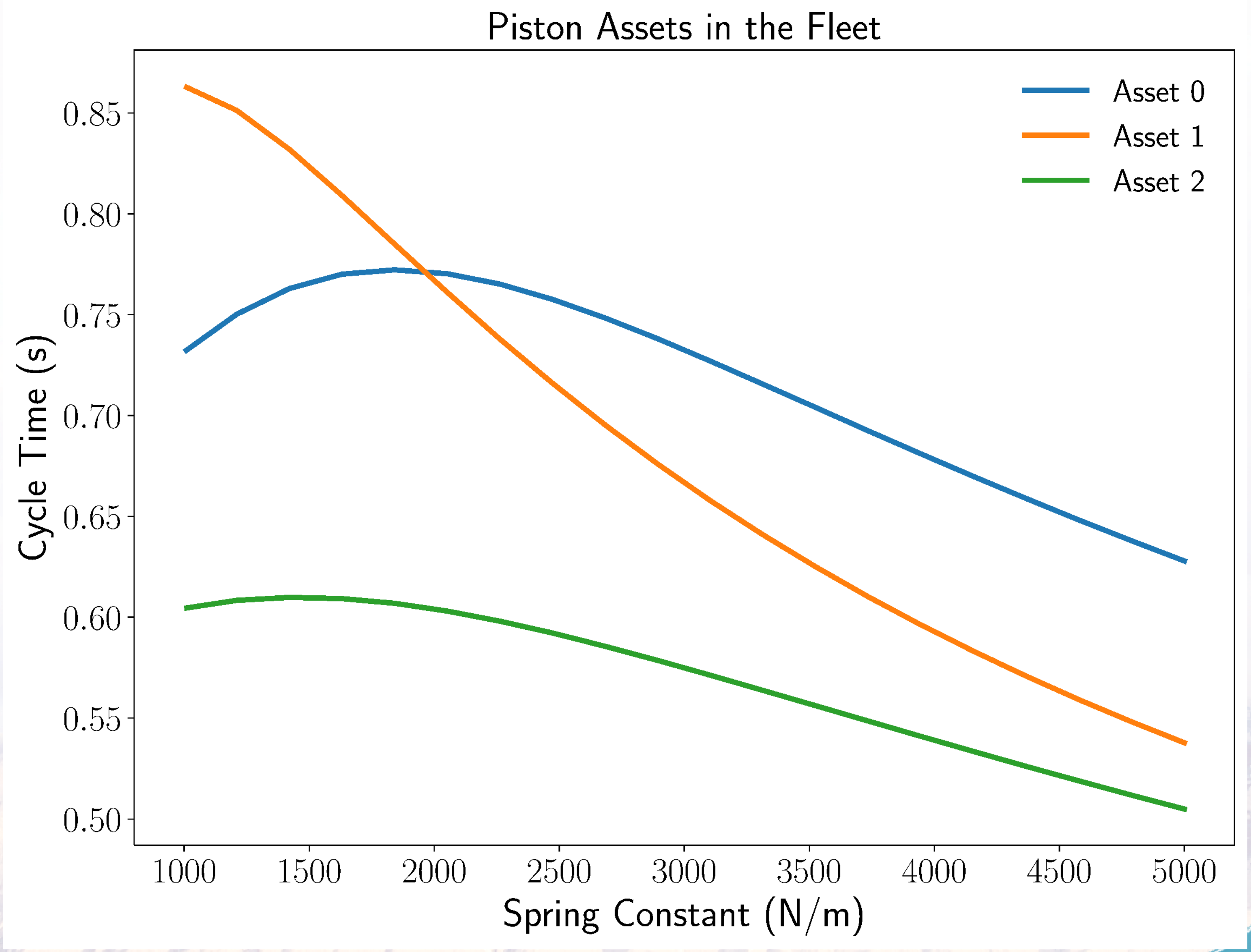
Input variable

$$A = P_0 S + 19.62M - \frac{kV_0}{S}$$

Latent variables

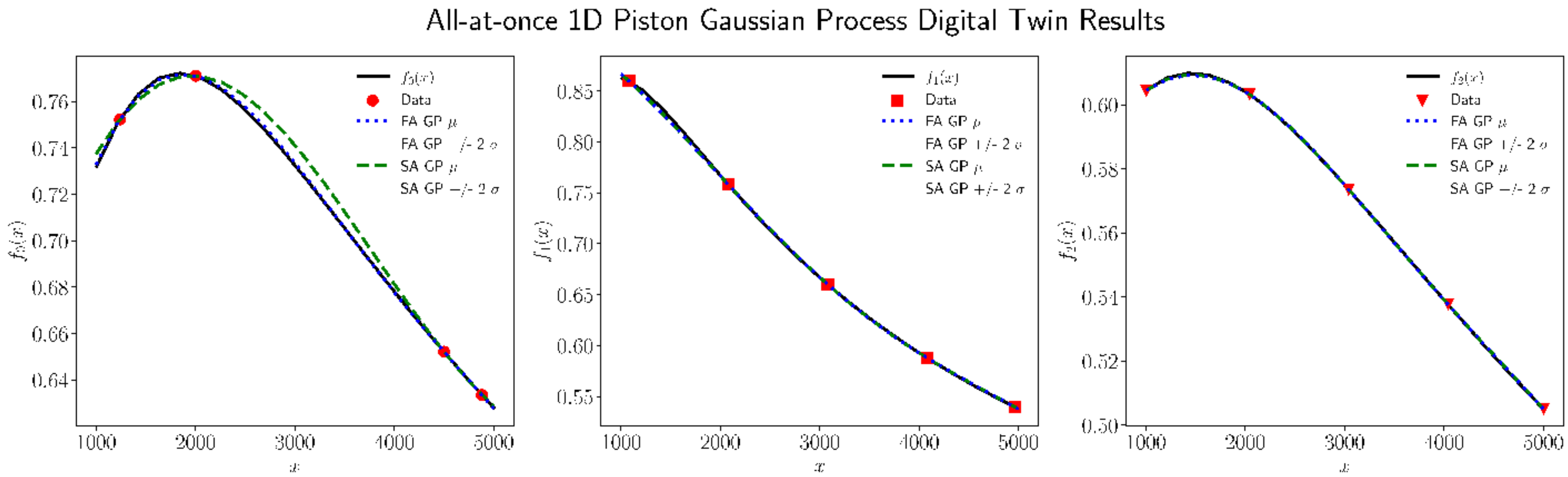
$M \in [30, 60]$	piston weight (kg)
$S \in [0.005, 0.020]$	piston surface area (m^2)
$V_0 \in [0.002, 0.010]$	initial gas volume (m^3)
$k \in [1000, 5000]$	spring coefficient (N/m)
$P_0 \in [90000, 110000]$	atmospheric pressure (N/m^2)
$T_a \in [290, 296]$	ambient temperature (K)
$T_0 \in [340, 360]$	filling gas temperature (K)

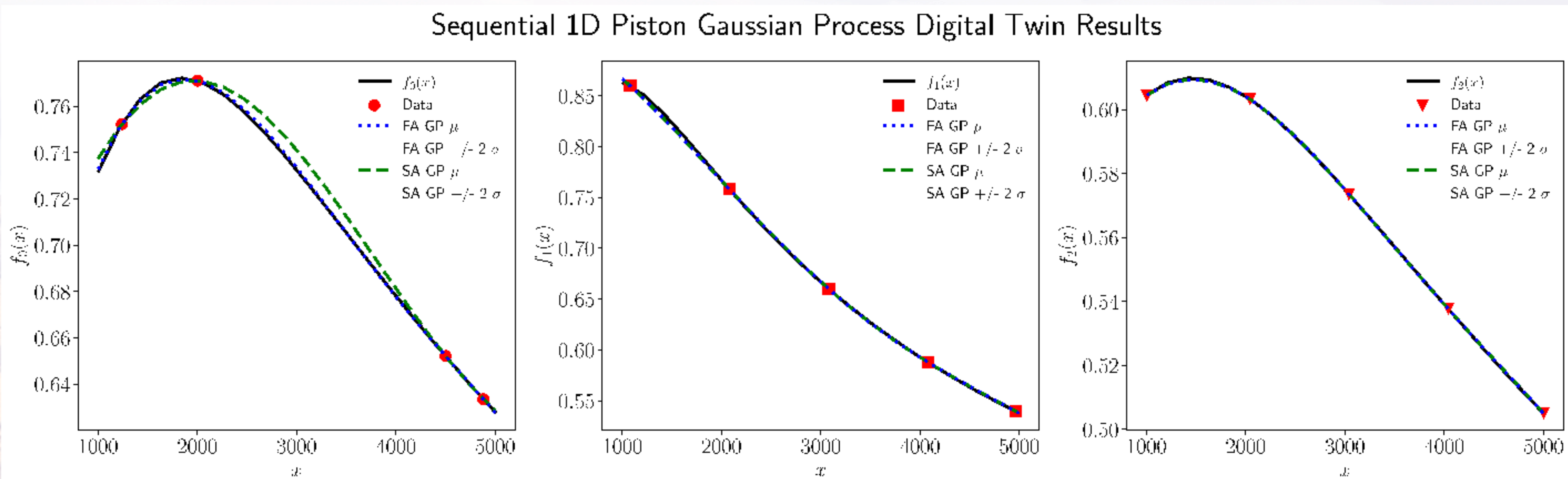
Piston Functions



Noiseless 1D Piston Results

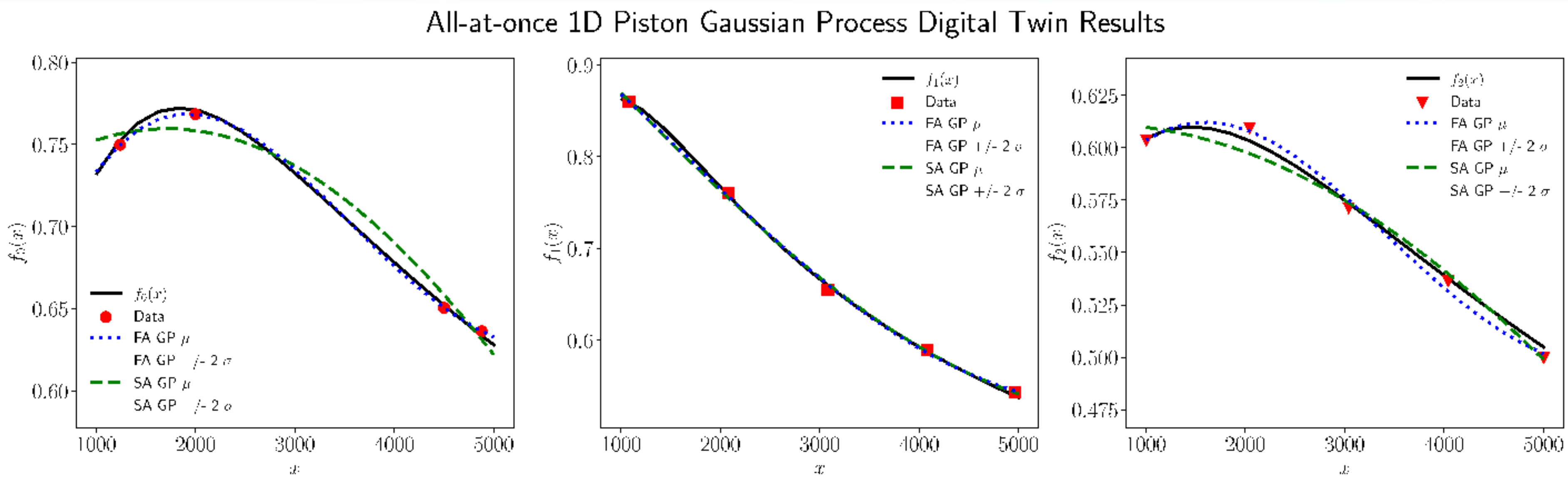
- All-at-once vs. sequential training of hyperparameters:

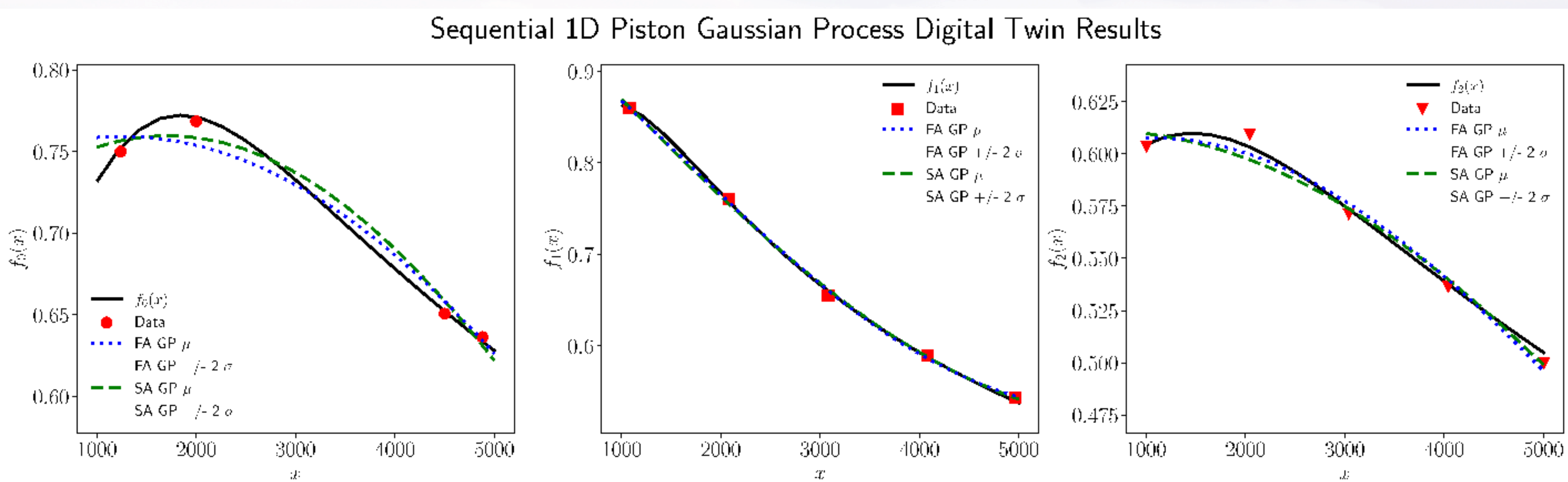




Noise-corrupted 1D Piston Results

- All-at-once vs. sequential training of hyperparameters:





Piston Model: 2D Problem

- Benchmark problem for surrogate modeling: https://www.sfu.ca/~ssurjano/piston.html .

$$C(\boldsymbol{x}) = 2\pi\sqrt{\frac{M}{k + S^2 \frac{P_0 V_0}{T_0} \frac{T_a}{V^2}}}$$

$$V = \frac{S}{2k}\left(\sqrt{A^2 + 4k\frac{P_0 V_0}{T_0} T_a} - A\right)$$

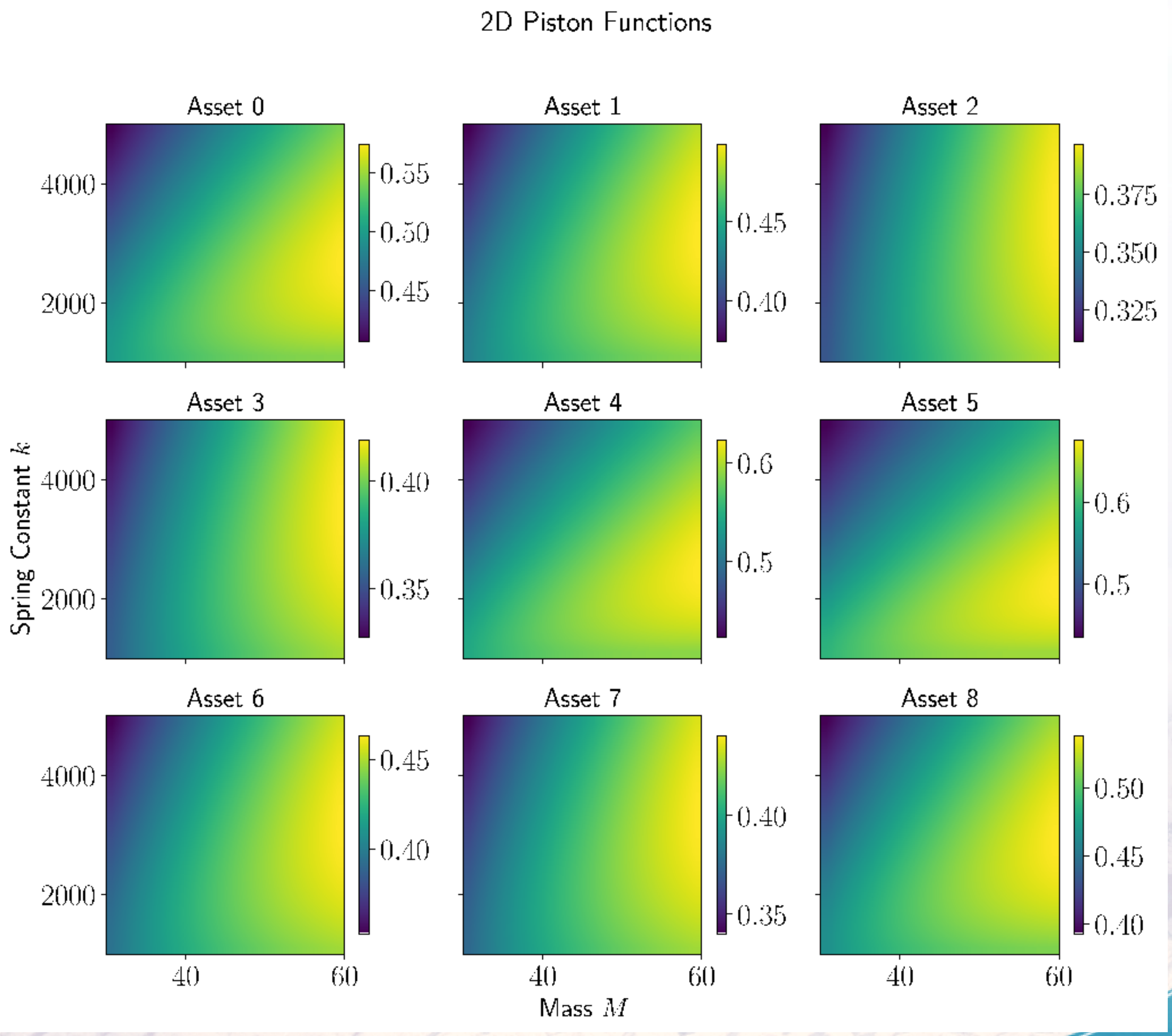
Input variable

$$A = P_0 S + 19.62M - \frac{kV_0}{S}$$

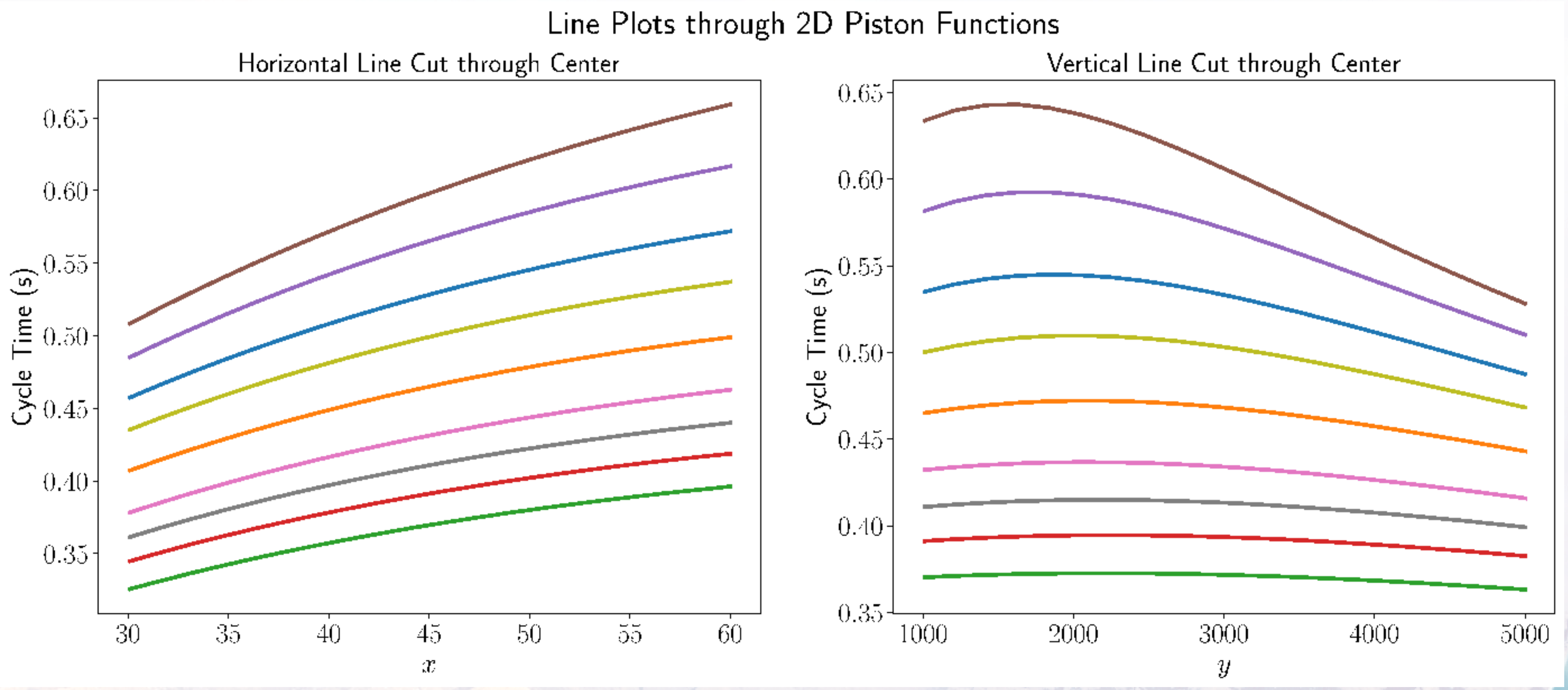
Latent variables

$M \in [30, 60]$	piston weight (kg)
$S \in [0.005, 0.020]$	piston surface area (m^2)
$V_0 \in [0.002, 0.010]$	initial gas volume (m^3)
$k \in [1000, 5000]$	spring coefficient (N/m)
$P_0 \in [90000, 110000]$	atmospheric pressure (N/m^2)
$T_a \in [290, 296]$	ambient temperature (K)
$T_0 \in [340, 360]$	filling gas temperature (K)

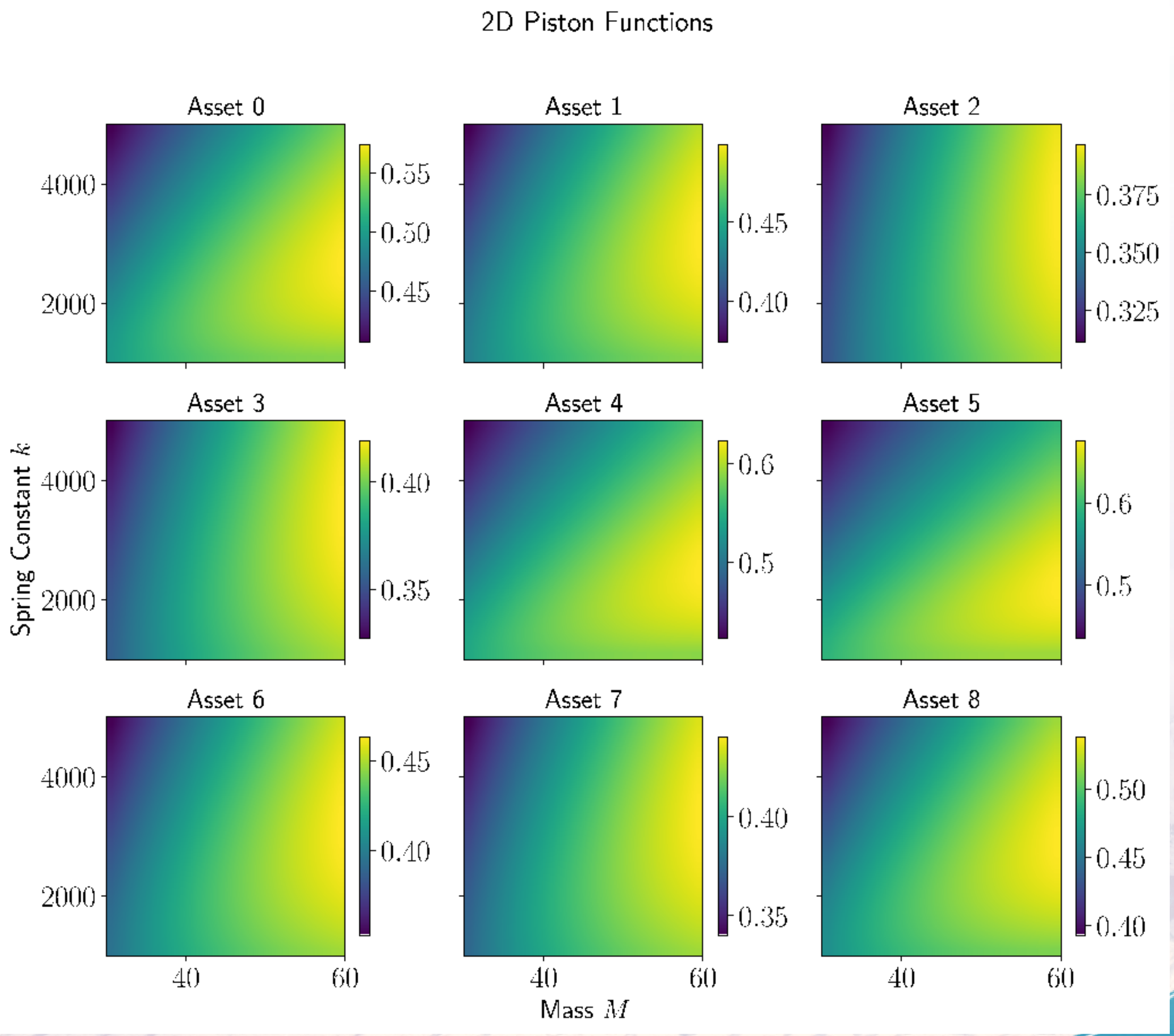
Piston Model: 2D Problem



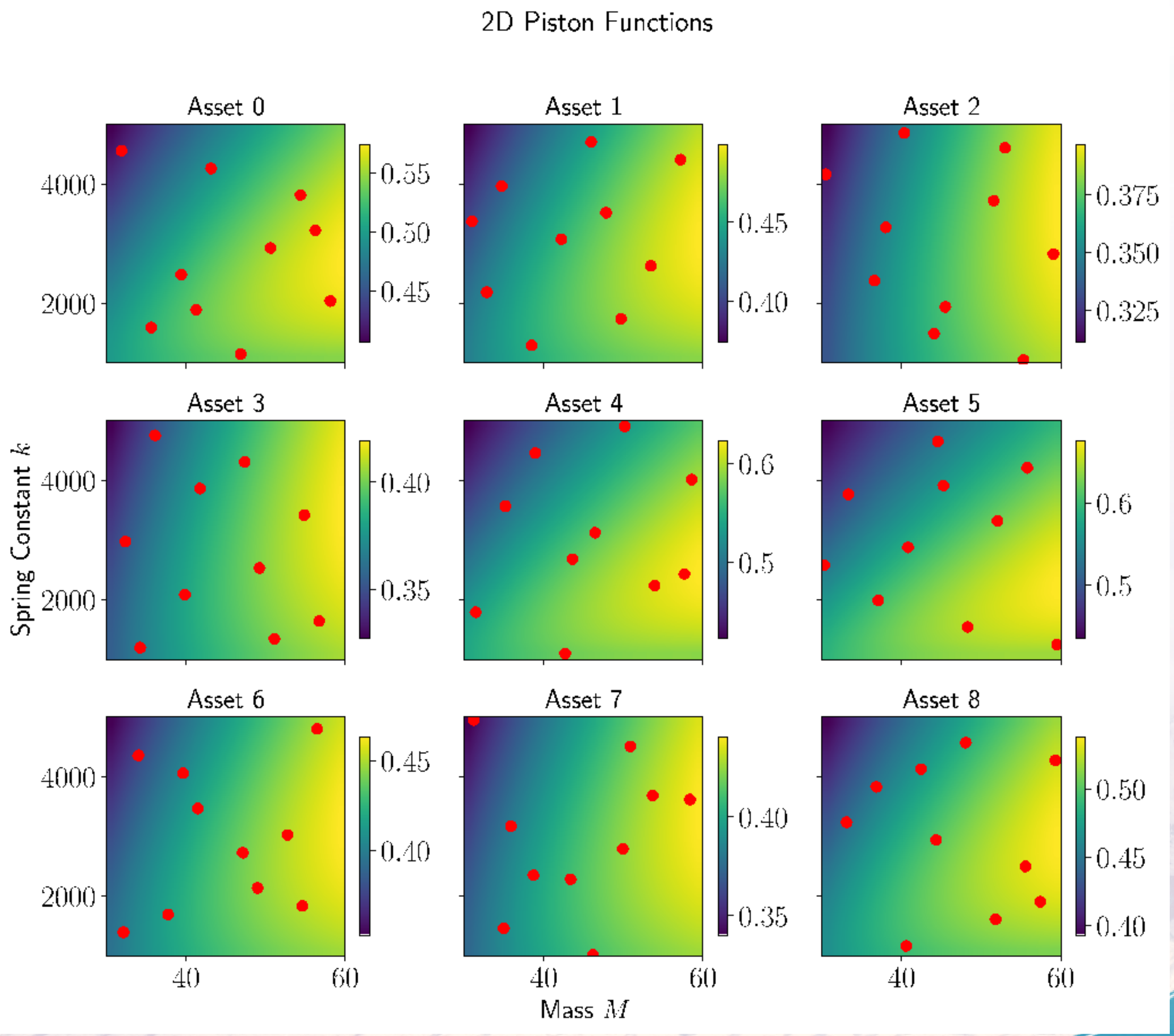
Piston Model: 2D Problem



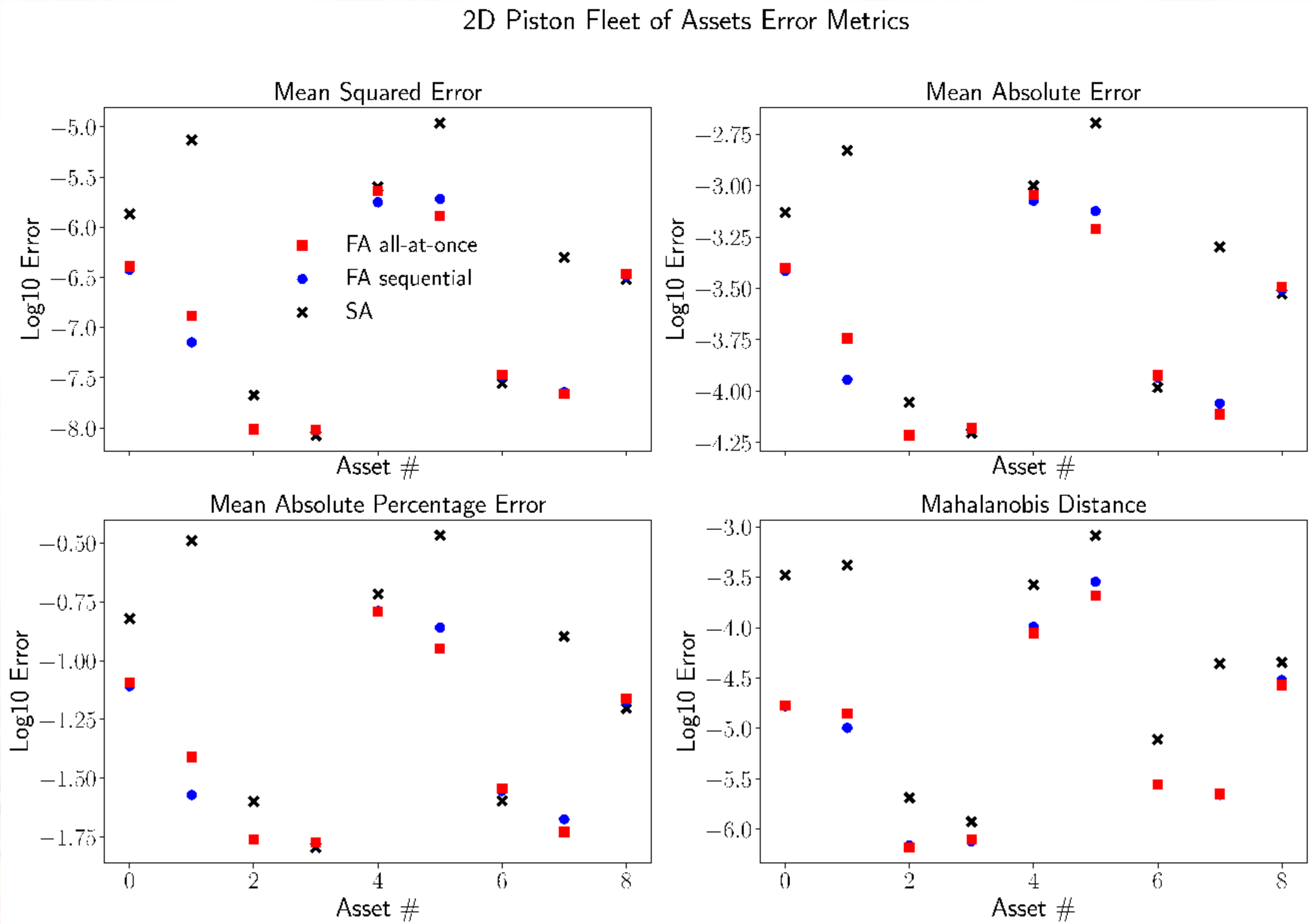
Piston Model: 2D Problem



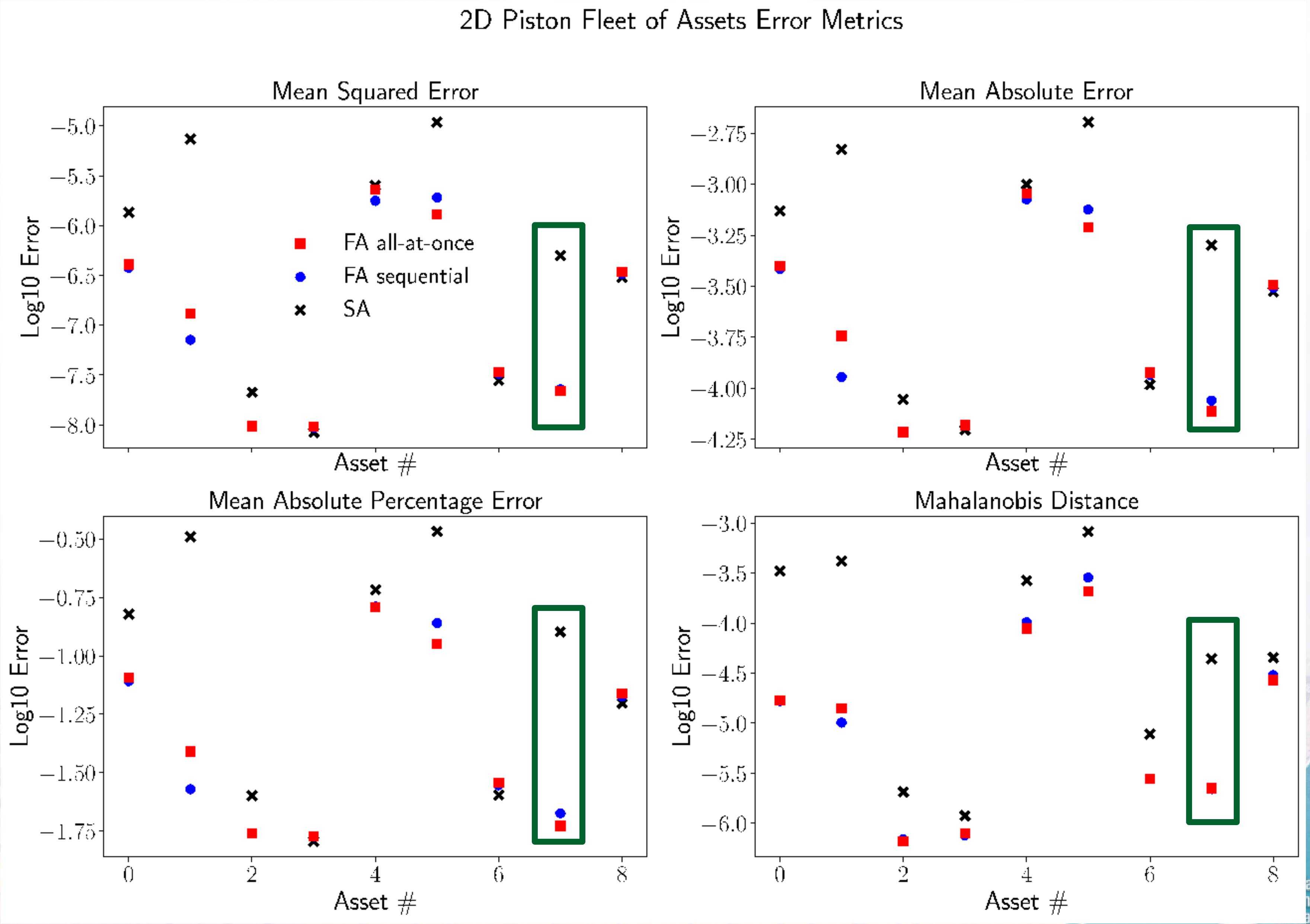
Piston Model: 2D Problem



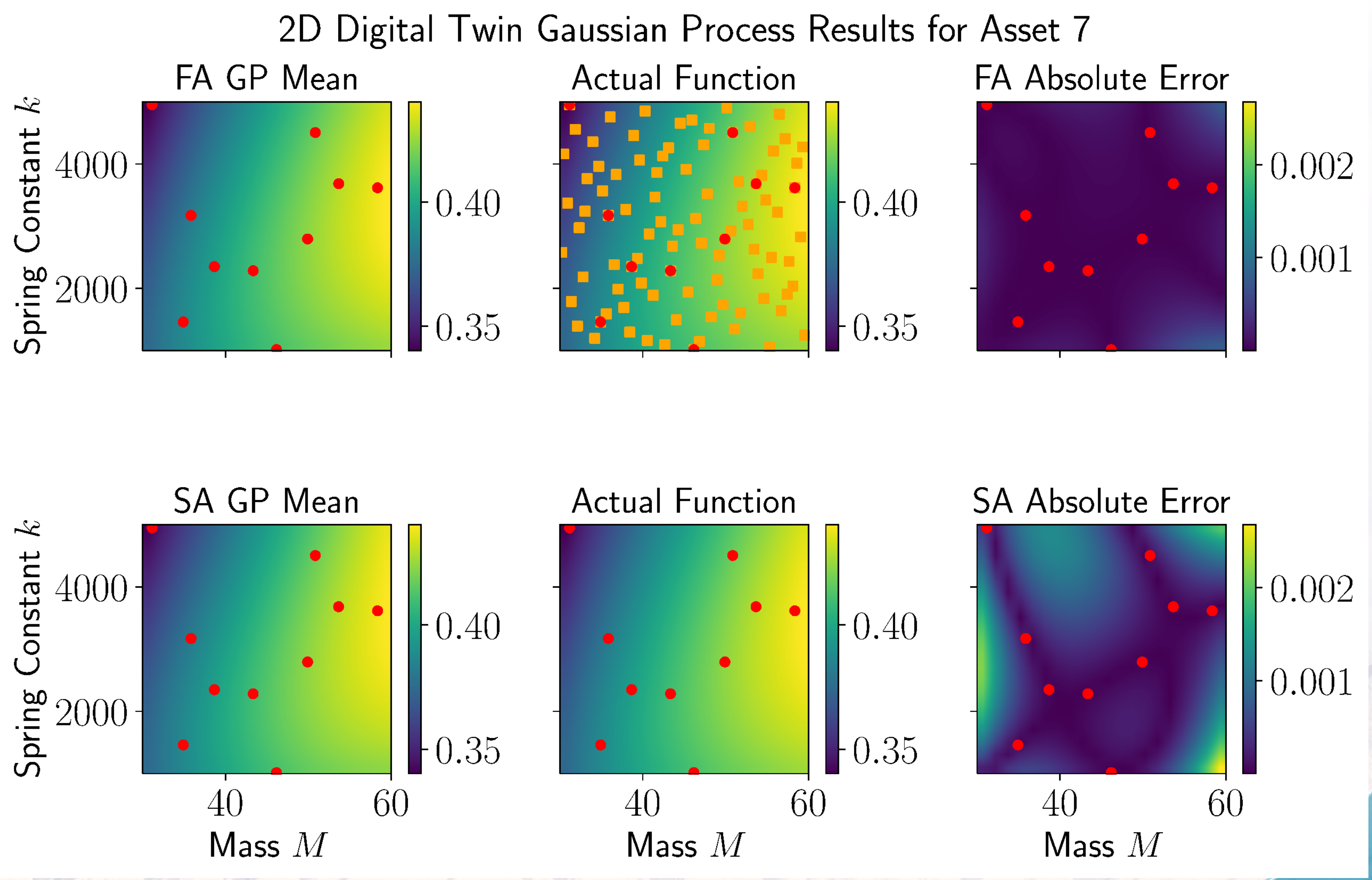
Piston Model: 2D Problem



Piston Model: 2D Problem



Piston Model: 2D Problem



Conclusions

- Limited data (possibly non-nested, noisy) for each asset …
- … but the dataset for the entire fleet is rich and diverse.
- Demonstrated that fleet data can be used to increase prediction accuracy for any single asset.
- Efficient linear algebra is critical for reducing the cost of hyperparameter estimation.

Conclusions

- Limited data (possibly non-nested, noisy) for each asset …
- … but the dataset for the entire fleet is rich and diverse.
- Demonstrated that fleet data can be used to increase prediction accuracy for any single asset.
- Efficient linear algebra is critical for reducing the cost of hyperparameter estimation.

Thank you!