An Interventional Perspective on Identifiability in Gaussian LTI Systems with Independent Component Analysis

Goutham Rajendran

Carnegie Mellon University

Joint work with

Patrik Reizinger (MPI for Intelligent Systems),

Wieland Brendel (MPI for Intelligent Systems),

Pradeep Ravikumar (CMU)

Causal Learning and Reasoning 2024

April 3, 2024

Linear Time Invariant (LTI) Systems

- A simple model for temporal system dynamics
- Used in control theory, engineering and machine learning
- System identification: Identify system parameters from trajectories
  - Studied for over half a century, starting with [Kalman 1960]

- *Interventions* can control trajectories

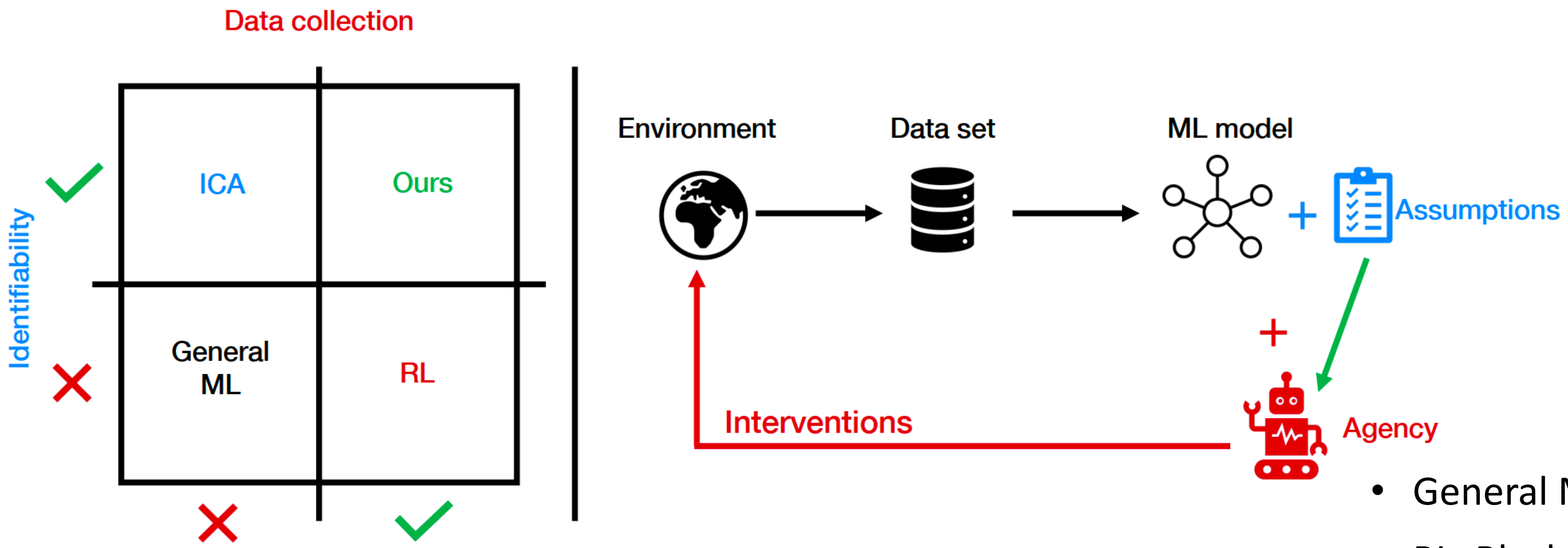
Independent Component Analysis

- A statistical model for representation learning
- Special case of latent variable modeling and causal representation learning (CRL)
- Main driver: Understand causation
- Identifiability: In what scenarios can we learn/recover causal model?

- *Interventions* enable identifiability

Our contributions: Interventional learning

- Intervene smartly, collect diverse data, gain identifiability



- General ML: Black components
- RL: Black + Red
- ICA: Black + Blue
- Ours: Black + Blue + Green + Red

Linear Time Invariant (LTI) Systems

- For each time step t, we have
  - Hidden state: $\boldsymbol{x}_t \in \mathbb{R}^{d_{\boldsymbol{x}}}$
  - Observed state: $\boldsymbol{y}_t \in \mathbb{R}^{d_{\boldsymbol{y}}}$
  - Control signal: $\boldsymbol{u}_t \in \mathbb{R}^{d_{\boldsymbol{u}}}$

- System matrices: $\mathbf{A} \in \mathbb{R}^{d_{\boldsymbol{x}} \times d_{\boldsymbol{x}}}, \mathbf{B} \in \mathbb{R}^{d_{\boldsymbol{x}} \times d_{\boldsymbol{u}}}, \mathbf{C} \in \mathbb{R}^{d_{\boldsymbol{y}} \times d_{\boldsymbol{x}}}$
- LTI system:

$$\boldsymbol{x}_{t+1} = \mathbf{A}\boldsymbol{x}_t + \mathbf{B}\boldsymbol{u}_t + \boldsymbol{\varepsilon}_t^{\boldsymbol{x}}$$
$$\boldsymbol{y}_t = \mathbf{C}\boldsymbol{x}_t + \boldsymbol{\varepsilon}_t^{\boldsymbol{y}}$$

- Noise variables are independent

LTI System Identification

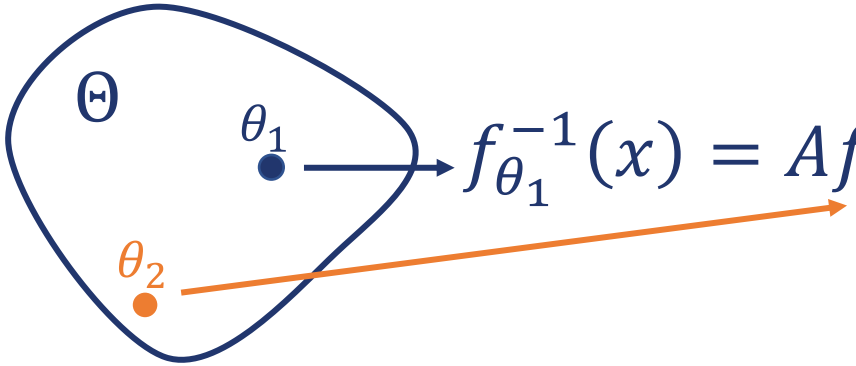
- Task: Recover system matrices A, B, C from observed states y.
- Necessary assumptions [Kalman]:
  - Controllability: $\mathbf{M}_c = [\mathbf{B}; \mathbf{AB}; \ldots; \mathbf{A}^{d_{\mathbf{x}}-1}\mathbf{B}]$ is full rank
  - Observability: $\mathbf{M}_o = \begin{pmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{d_{\mathbf{x}}-1} \end{pmatrix}$ is full rank
  - Stability: Eigenvalues of A are less than 1
- Enough to recover Markov Parameter matrix

$$\mathbf{G} = [\mathbf{I}, \mathbf{CB}, \mathbf{CAB}, \ldots, \mathbf{CA}^{T-1}\mathbf{B}]$$

  - This leads to system identification (Ho-Kalman algorithm)

Why identifiability?

- If we know ground truth, can use that to reliability control systems
- Identifiability is widely studied in ICA/CRL literature
  - Identifiability:



$$f_{\theta_1}^{-1}(x) = A f_{\theta_2}^{-1}(x) + c$$

  - Step towards better ML models that are reliable, stable, etc.
- Our contribution (and hope): Recent advances in ICA/CRL can/will apply to engineering/robotics

Related works

- LTI systems identification: Started with [Kalman 1960]
  - Recent works have studied polynomial time complexity [Bakshi et al. 2023]
  - Mixtures of LTI Systems [Chen-Poor, 2023]
- Nonlinear ICA: [Comon 1994, Hyvarinen-Oja 2000]
- Causality: [Spirtes et al. 2000, Pearl 2009]

- Each of these are vast fields!

Related works – Causal Representation Learning

- CRL is the field of identifying ground truth representations from raw data [Schölkopf et al. 2021]
- General principle: Multiple environments lead to identifiability [Khemakhem et al. 2020, Buchholz et al. 2023]
  - Some works have traded off environments for inductive bias [Kivva et al. 2020]
- Special case: Interventional Environments
  - [Lippe et al. 2022, Squires et al. 2023, Buchholz et al. 2023]
  - Various assumptions: Paired data, known intervention targets, etc.

Our setting: Multi-environment LTI Systems

- We have multiple environments (set E). In environment e,

$$\mathbf{x}^e_{t+1} = \mathbf{A}\mathbf{x}^e_t + \mathbf{B}\mathbf{u}^e_t$$
$$\mathbf{y}^e_t = \mathbf{C}\mathbf{x}^e_t + \boldsymbol{\varepsilon}^e_t.$$

- Assumption 1: Control signals are Gaussian

$$\mathbf{u}^e_t \sim \prod_{i=1}^{d_{\mathbf{u}}} \mathcal{N}\left(0; (\boldsymbol{\sigma}^e_i)^2\right)$$

- Main assumption [Sufficient variability]: The environment variability matrix $\boldsymbol{\Delta} \in \mathbb{R}^{|E| \times d_{\mathbf{u}}}$ has full rank

$$\boldsymbol{\Delta}_{e,i} = \frac{1}{(\boldsymbol{\sigma}^e_i)^2} - \frac{1}{(\boldsymbol{\sigma}^0_i)^2}.$$

Main result

- Markov parameter matrix $\mathbf{G} = [\mathbf{I}, \mathbf{CB}, \mathbf{CAB}, \dots, \mathbf{CA}^{T-1}\mathbf{B}]$
- Under sufficient environment variability
  - **Theorem:** G is identifiable
  - **Corollary:** A, B, C are identifiable (up to similarity transformations)

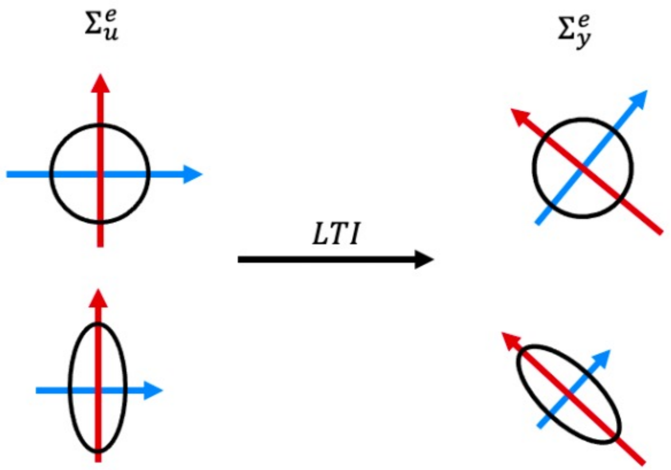
Main result

- Markov parameter matrix $\mathbf{G} = [\mathbf{I}, \mathbf{CB}, \mathbf{CAB}, \ldots, \mathbf{CA}^{T-1}\mathbf{B}]$
- Under sufficient environment variability
  - **Theorem:** G is identifiable
  - **Corollary:** A, B, C are identifiable (up to similarity transformations)

- Takeaways:
  - Data collector should ensure data diversity
  - How? Estimate rank of environment variability matrix, use as proxy for system identifiability
  - Exact rank may not be useful, perhaps condition number or stable rank?
  - Also relates to identifiability of *transfer function*.

Main result - Intuition

- Assume for simplicity: Dim = 2 and system dynamics are rotations
  - Covariance of Gaussian is rotated
- If we start with isotropic Gaussian, 1 environment is not enough due to rotational symmetry.
- However, with another environment, we can identify the rotation.



Main Identifiability result – Proof Sketch

- **Theorem:** Under sufficient environment variability, the Markov parameter matrix G of the LTI system is identifiable.
- Proof sketch:
  - Each environment controls a *rank-1 facet* of the state space
  - So, enough diverse environments => probe and learn the entire state space
  - Formally, estimate log-odds as quadratic functions of signals
  - Differentiate twice to get linear system
  - Use environment variability to estimate parameters

Connection to Causal de Finetti

- Causal de Finetti theorem [Guo et al. 2022]: A statistical formalization of the Independent Causal Mechanisms (ICM) principle
  - Consider an *exchangeable* sequence of random variables (relaxes iid assumption)
  - Assume they satisfy various conditional independencies
  - Then we can factorize joint probability distribution (conditioned on independent parameters)
- Assumes categorical variables, hypothesized to also hold for continuous variables
- Our theorem: Can be interpreted as a generalization of the CdF theorem to continuous variables (for the special case of LTI systems)

Experiments

- Methodology: MLE for multi-environment data with shared parameters
- Log likelihood:

$$\mathcal{L} = \sum_e \sum_t \ln p_{\mathcal{N}} \left(\mathbf{u}_t^e | \mathbf{y}_{t+1}^e, \mathbf{y}_t^e, \mathbf{A}, \mathbf{B}, \mathbf{C} \right)$$

- Because of Gaussianity, essentially least-squares

$$\mathcal{L} \propto \sum_e \sum_t (\mathbf{u}_t^e)^\top \mathbf{\Sigma}^e \mathbf{u}_t^e$$

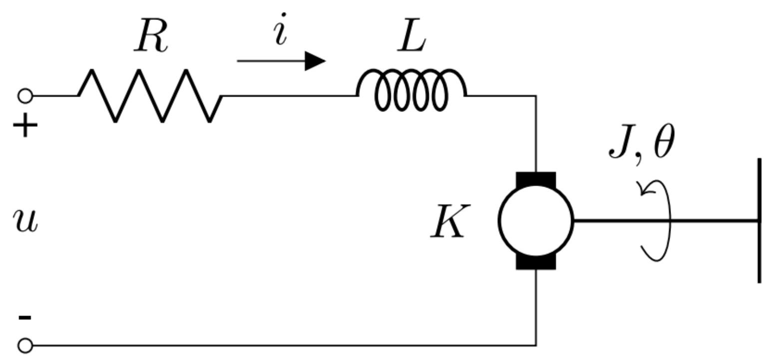
- Optimize via SGD

Experiments – Evaluation

- We run real-world and synthetic experiments
- To quantify identifiability, we report MCC between learned and grouth truth control signals

- Mean Correlation Coefficient (MCC) is a proxy for identifiability
  - Measures linear correlation up to permutation of components
  - Computes best permutation using linear sum assignment
  - Values in [0, 1], higher is better.

Experiments – DC Motor



- Continuous LTI system
  - Control: u – voltage
  - States: i - Current, θ - Rotor angle
  - R – resistance, L – conductance, K – Electromotive force constant, J – Inertia, D – Damping coefficient

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} i \\ \theta \end{bmatrix} = \begin{bmatrix} -R/L & K/L \\ -K/J & -D/J \end{bmatrix}\begin{bmatrix} i \\ \theta \end{bmatrix} + \begin{bmatrix} 1/L \\ 0 \end{bmatrix} u; \qquad y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} i \\ \theta \end{bmatrix}$$

- We discretize the dynamics, generate 3 environments and run MLE.
- MCC of recovered parameters with ground truth close to 0.99

Experiments – Synthetic

- We generate synthetic data, allowing us to experiment with
  - Both identity and non-identity observation matrix
  - Nonzero and zero means
  - Robustness to noise, large dimension, etc.

- Verifying our theory (minimal # environments):

$d_{\mathbf{u}}$	$\|E\|$	MCC ↑								
		$\boldsymbol{\mu}_{\mathbf{u}}^e \neq \mathbf{0}$				$\boldsymbol{\mu}_{\mathbf{u}}^e = \mathbf{0}$				$\boldsymbol{\mu}_{\mathbf{u}}^e \neq \mathbf{0}$
		$\mathbf{B} = \mathbf{I}$		$\mathbf{B} \neq \mathbf{I}$		$\mathbf{B} = \mathbf{I}$		$\mathbf{B} \neq \mathbf{I}$		$\mathbf{B} \neq \mathbf{I}$
		$\mathbf{C} = \mathbf{I}$	$\mathbf{C} \neq \mathbf{I}$	$\mathbf{C} = \mathbf{I}$	$\mathbf{C} \neq \mathbf{I}$	$\mathbf{C} = \mathbf{I}$	$\mathbf{C} \neq \mathbf{I}$	$\mathbf{C} = \mathbf{I}$	$\mathbf{C} \neq \mathbf{I}$	$\mathbf{C} \neq \mathbf{I}$
2	3	$0.866_{\pm 0.033}$	$0.767_{\pm 0.158}$	$0.730_{\pm 0.191}$	$0.697_{\pm 0.090}$	$0.633_{\pm 0.104}$	$0.675_{\pm 0.132}$	$0.734_{\pm 0.158}$	$0.725_{\pm 0.115}$	$0.968_{\pm 0.055}$
3	4	$0.901_{\pm 0.054}$	$0.910_{\pm 0.061}$	$0.916_{\pm 0.044}$	$0.861_{\pm 0.062}$	$0.659_{\pm 0.201}$	$0.618_{\pm 0.241}$	$0.633_{\pm 0.110}$	$0.667_{\pm 0.109}$	$1.000_{\pm 0.000}$
5	6	$0.892_{\pm 0.057}$	$0.929_{\pm 0.025}$	$0.928_{\pm 0.026}$	$0.911_{\pm 0.045}$	$0.657_{\pm 0.116}$	$0.618_{\pm 0.079}$	$0.620_{\pm 0.025}$	$0.539_{\pm 0.078}$	$0.995_{\pm 0.009}$
8	9	$0.943_{\pm 0.006}$	$0.940_{\pm 0.008}$	$0.941_{\pm 0.009}$	$0.867_{\pm 0.039}$	$0.585_{\pm 0.144}$	$0.479_{\pm 0.129}$	$0.523_{\pm 0.016}$	$0.414_{\pm 0.031}$	$0.977_{\pm 0.011}$
10	11	$0.939_{\pm 0.011}$	$0.915_{\pm 0.023}$	$0.925_{\pm 0.017}$	$0.924_{\pm 0.042}$	$0.708_{\pm 0.042}$	$0.611_{\pm 0.043}$	$0.604_{\pm 0.063}$	$0.525_{\pm 0.097}$	$0.996_{\pm 0.006}$

Experiments – Synthetic

- Some observations:
  - MCCs comparable to other CRL results
  - MLE works better when we have non-zero means
  - Better conditioning of environment variability matrix helps identifiability

- Robustness to observation noise:

$d_{\mathbf{u}}$	$\|E\|$	MCC ↑				
		$\boldsymbol{\sigma}_{\boldsymbol{\varepsilon}}^2 = 0$	$\boldsymbol{\sigma}_{\boldsymbol{\varepsilon}}^2 = 1\text{e}{-4}$	$\boldsymbol{\sigma}_{\boldsymbol{\varepsilon}}^2 = 1\text{e}{-2}$	$\boldsymbol{\sigma}_{\boldsymbol{\varepsilon}}^2 = 1\text{e}{-1}$	$\boldsymbol{\sigma}_{\boldsymbol{\varepsilon}}^2 = 1$
2	3	$0.627_{\pm 0.090}$	$0.679_{\pm 0.162}$	$0.643_{\pm 0.068}$	$0.590_{\pm 0.034}$	$0.625_{\pm 0.038}$
3	4	$0.886_{\pm 0.057}$	$0.823_{\pm 0.052}$	$0.718_{\pm 0.168}$	$0.637_{\pm 0.194}$	$0.796_{\pm 0.022}$

Summary

- Theory of ICA/CRL suggest new perspectives for optimal intervention design in engineering/physics/robotics.
- ICA/CRL take a passive perspective, while in robotics, we have control over the system and data collection
- Our work: Use identifiability theory insights to design data collection policies/interventions

Summary

- Theory of ICA/CRL suggest new perspectives for optimal intervention design in engineering/physics/robotics.
- ICA/CRL take a passive perspective, while in robotics, we have control over the system and data collection
- Our work: Use identifiability theory insights to design data collection policies/interventions

Future work

- What if LTI system is not controllable/observable?
- Non-linear transition dynamics?

Summary

- Theory of ICA/CRL suggest new perspectives for optimal intervention design in engineering/physics/robotics.
- ICA/CRL take a passive perspective, while in robotics, we have control over the system and data collection
- Our work: Use identifiability theory insights to design data collection policies/interventions

Future work

- What if LTI system is not controllable/observable?
- Non-linear transition dynamics?

Thank you!