Finite Expression Method for Discovering Physical Laws from Data

Haizhao Yang
Department of Mathematics
University of Maryland College Park

Joint work with Zhongyi Jiang (U. Of Delaware) and Chunmei Wang (U. Of Florida)

Minisymposium: Reduced Order Modeling and Forecasting
in Geophysical Flows and Complex Dynamical Systems
SIAM Conference on Applications of Dynamical Systems
May 16th, 2023

Discovering Physical Laws from Data
Observational Data
• Extract information from massive data
• Make prediction by applying the learned laws
• Data-driven optimal control
Prediction and simulations etc.
SATURN
NEPTUNE
MARS
VENUS
JUPITER
MERCURY
SUN
EARTH
URANUS
Smart Grid

Problem Statement

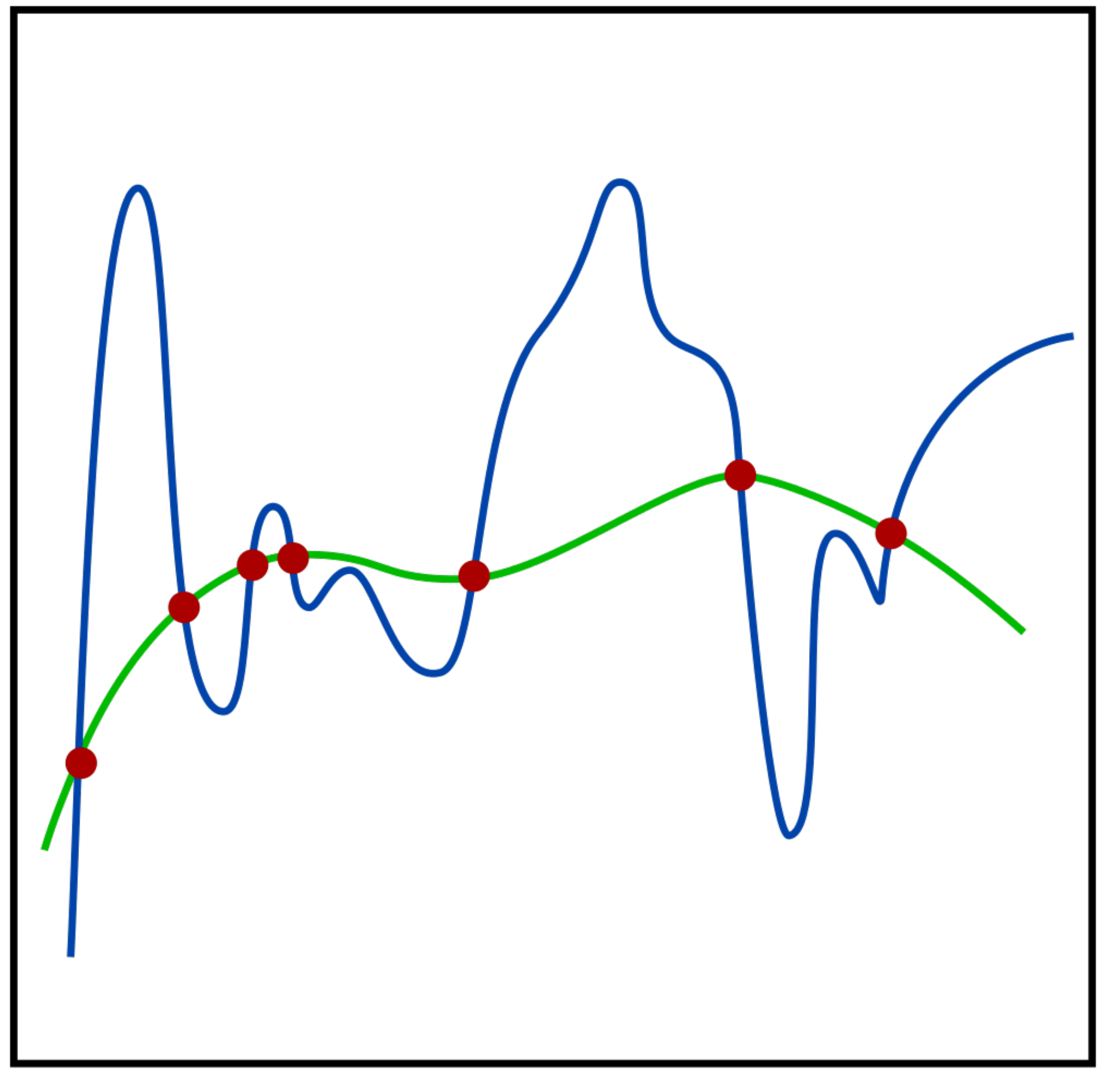
A Concrete Example: 1D Burgers Equation

- Given nonlinear operator $F : \mathcal{X} \to \mathcal{Y}$, e.g., $F(u) = \dfrac{\partial u}{\partial t}$
- Unknown nonlinear operator $G : \mathcal{X} \to \mathcal{Y}$, e.g., $G(u) = -u \cdot u_x + \nu u_{xx}$
- Assumption: a function $u(x, t) \in \mathcal{X}$ satisfies

$$F(u) = G(u) \qquad \Leftrightarrow \qquad \frac{\partial u}{\partial t} = -u \cdot u_x + \nu u_{xx}$$

- Given discrete data observations $u(x_i, t_j), i = 1,\ldots, m, j = 1,\ldots, n$
- Goal: identify G with $G \neq F$
- Challenges: 1) non-uniqueness of G (due to data fitting and discretization)

  2) noisy data



Problem Statement

An Abstract Framework

- Function spaces $\mathscr{X}$ and $\mathscr{Y}$
- Given nonlinear operator $F : \mathscr{X} \to \mathscr{Y}$
- Unknown nonlinear operator $G : \mathscr{X} \to \mathscr{Y}$
- Assumption: a function $u(x, t) \in \mathscr{X}$ satisfies

$$G(u) = F(u)$$

- Given data observations $u(x_i, t_j), i = 1, \ldots, m, j = 1, \ldots, n$
- Goal: identify G with $G \neq F$

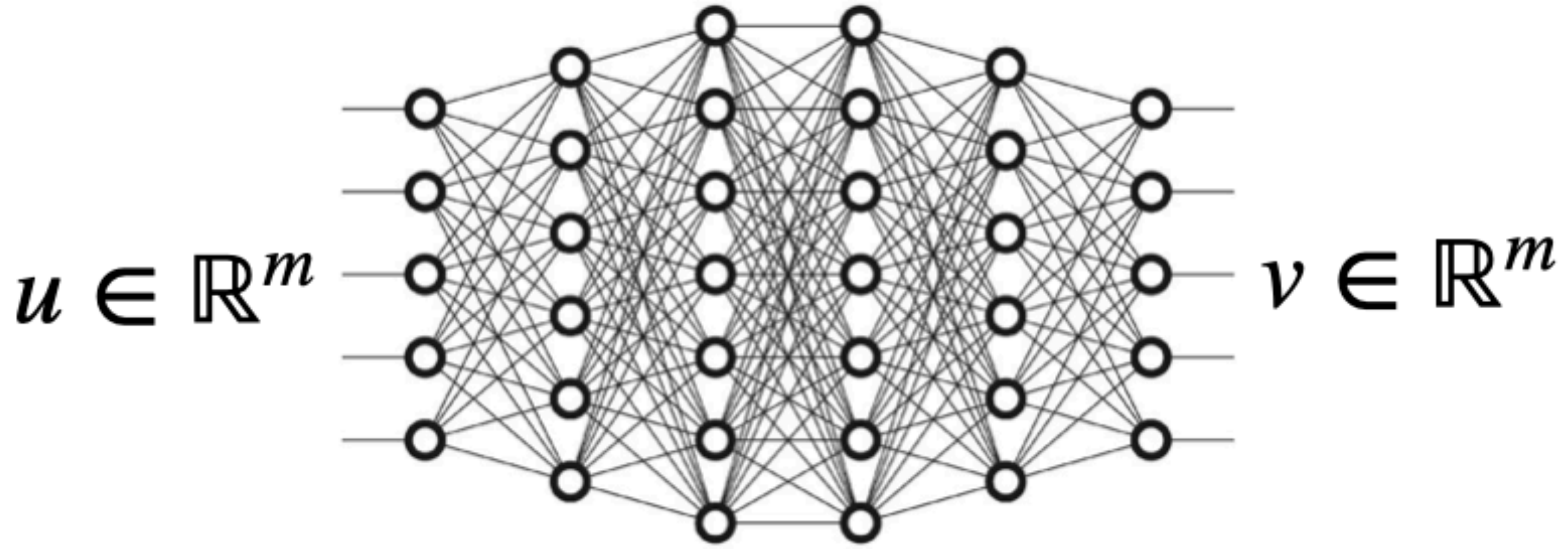
Problem Statement

A Concrete Example: 1D Burgers Equation

- Goal: identify $G(u) = -u \cdot u_x + \nu u_{xx}$
- What's operator G after discretization?

$$G : \mathbb{R}^m \rightarrow \mathbb{R}^m$$

  - A high-dimensional function
  - Traditional parametrization methods: **the curse of dimensionality**
  - Neural network parametrization: **no interpretability**



Finite Expression Method (FEX)

Jiang, Wang, Y. arXiv:2305.08342

Idea:

- Find an explicit expression that approximates the unknown operator G
- Function space with finite expressions
  - k-**finite expression**: a mathematical expression with at most k operators

    e.g., $\sin(2x) + 5e^x$ and $5\frac{\partial}{\partial t}(u)$
  - Function space: $\mathbb{S}_k$ as the set of s-finite expressions with $s \leq k$

Finite Expression Method (FEX)

Jiang, Wang, Y. arXiv:2305.08342

Advantages: No curse of dimensionality in approximation

- **Theorem** (Liang and Y. 2022) Suppose the function space is $\mathbb{S}_k$ generated with operators including ``+", ``-", ``×", ``/", ``max{0,x}", ``sin(x)", and ``2^x". Let $p \in [1, +\infty)$. For any f in the Holder function class $\mathscr{H}_{\mu}^{\alpha}([0,1]^d)$ and $\varepsilon > 0$, there exists a k-finite expression ϕ in $\mathbb{S}_k$ such that $\|f - \phi\|_{L^p} \leq \varepsilon$, if $k \geq \mathcal{O}(d^2(\log d + \log \frac{1}{\varepsilon})^2)$.
- e.g., $G(u) = -u \cdot u_x + \nu u_{xx}$ in the Burgers equation is a 3-dimensional polynomial $p(u, u_x, u_{xx})$

Finite Expression Method

Jiang, Wang, Y. arXiv:2305.08342

Least square based FEX

- e.g., $\frac{\partial u}{\partial t} = G(u) = -u \cdot u_x + \nu u_{xx}$
- A mathematical expression G^* to approximate the unknown operator via

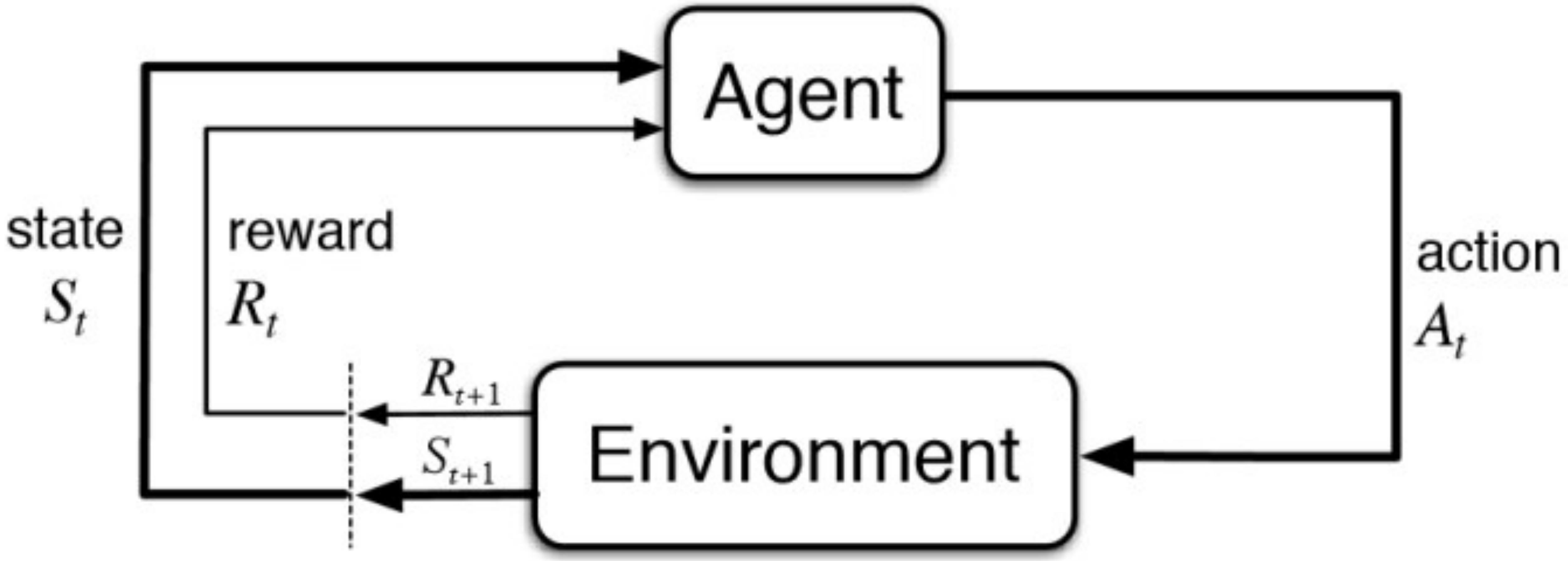
$$G^* = \arg\min_{G \in \mathbb{S}_k} \mathscr{L}(G) := \arg\min_{G \in \mathbb{S}_k} \|G(u) - u_t\|_2^2$$

- Or numerically

$$G^* = \arg\min_{G \in \mathbb{S}_k} \mathscr{L}(G) := \arg\min_{G \in \mathbb{S}_k} \frac{1}{mn} \sum_{j=1}^{n} \sum_{i=1}^{m} |G(u)(x_i, t_j) - u_t(x_i, t_j)|^2$$

○ Question: how to solve this combinatorial optimization problem?

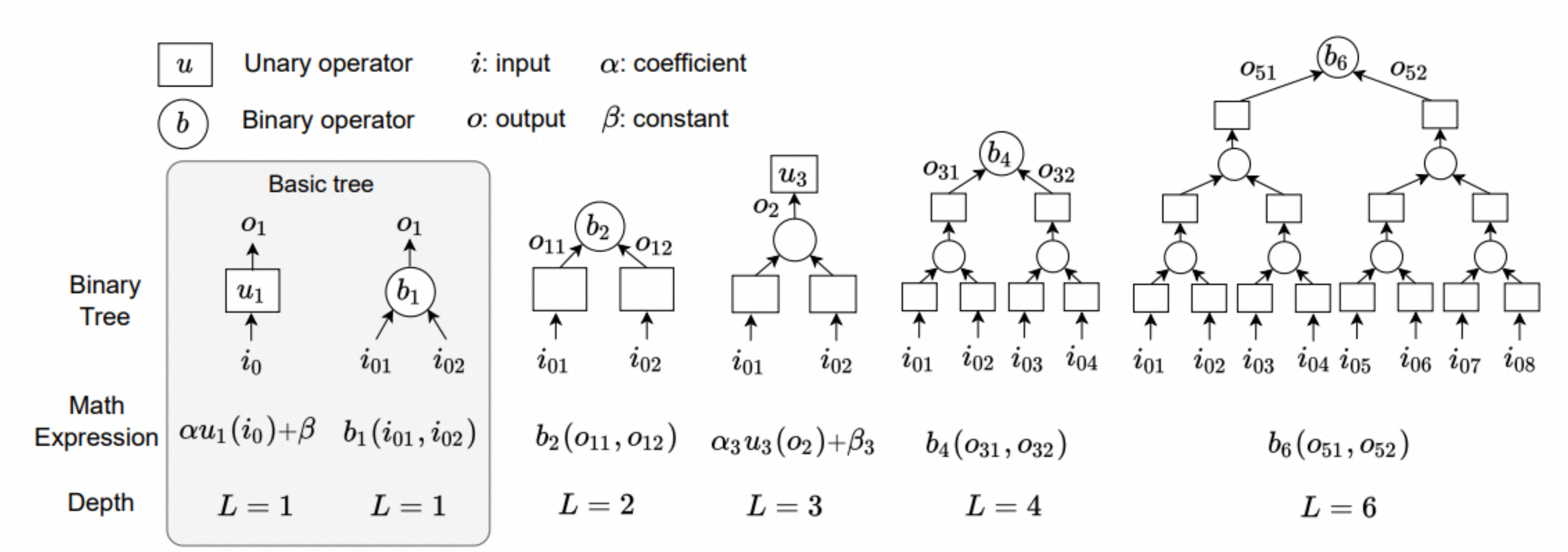
Reinforcement Learning for Combinatorial Optimization



By Richard S. Sutton and Andrew G. Barto.

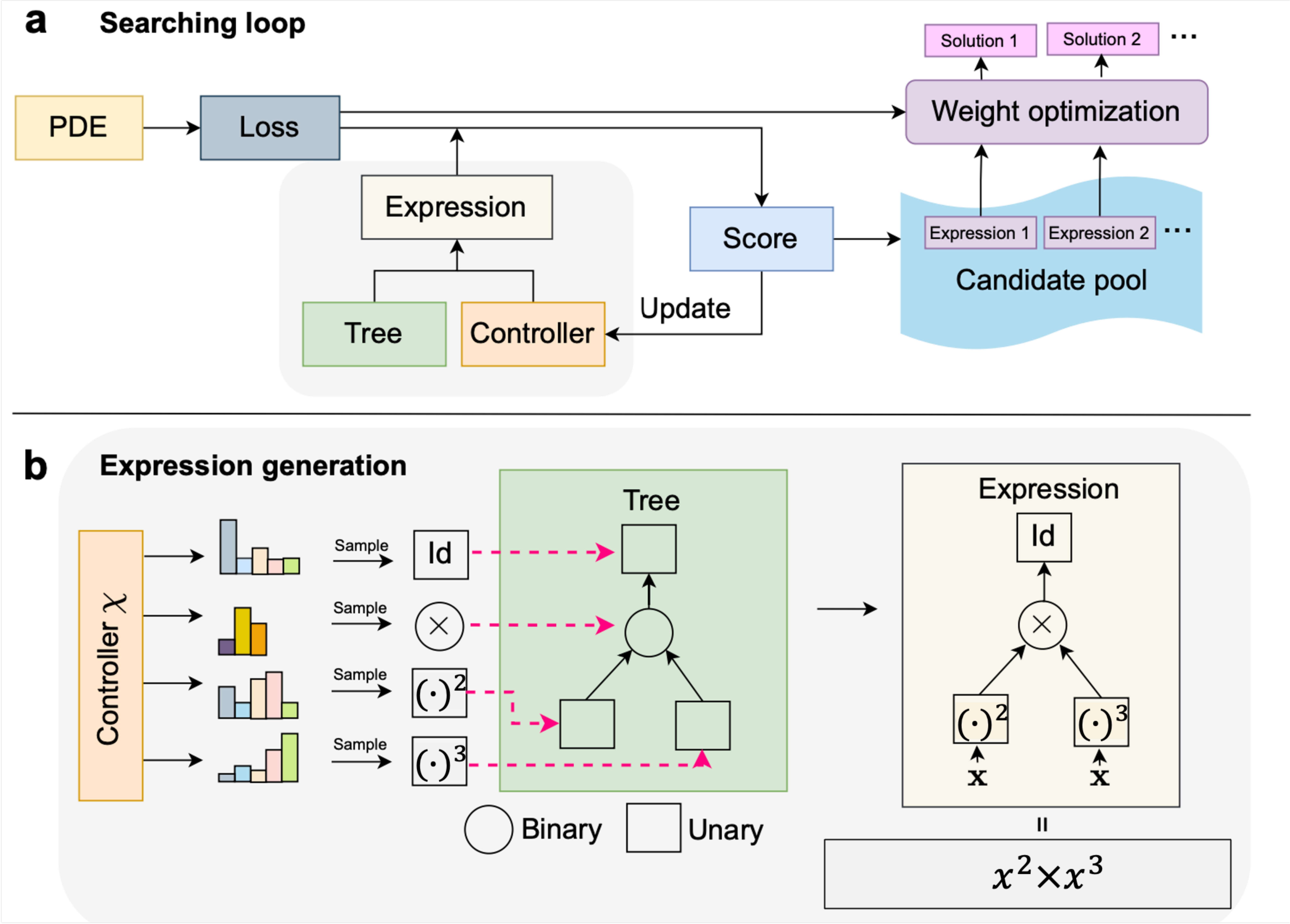
- **Goal:** Apply RL to select mathematical expressions to identify a governing equation
- **Ideas:**
    1. Reformulate the sequential (selection, realization, evaluation) procedure as a sequence of (action, state, reward)
    2. Reformulate the decision strategy for selection as the policy to take actions
    3. The data fitting quality as the reward

Expression Generation



An expression tree as a sequence of node values by using its pre-order traversal, e.g., $2\frac{\partial(\cdot)}{\partial t}+3$ and $\frac{\partial}{\partial x}(\cdot)+\frac{\partial}{\partial y}(\cdot)$

Computation Flow of FEX



Key Features of FEX

- **No curse of dimensionality** in approximation theory v.s. traditional methods
- **Interpretable** learning outcomes v.s. blackbox neural networks
- **Higher accuracy** v.s. existing symbolic regression tools
- A nonlinear approach to generate **a large set** of expressions from **a small collection** of operators
  - SINDy[1]: require a large manually designed dictionary
  - PDE-Net[2]: only capable of polynomials of operators
  - GP: Genetic programming with poor accuracy
  - SPL[3]: Monte Carlo tree search with poor accuracy

1. Brunton, Proctor, Nathan, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, PNAS, 2016
2. Long, Lu, Dong, PDE-Net 2.0: Learning PDEs from data with a numeric-symbolic hybrid deep network, Journal of Computational Physics 2019
3. Sun et al. Symbolic Physics Learner: Discovering governing equations via Monte Carlo tree search. ICLR 2023

Numerical Example 1:

2D Burgers equation with periodic boundary conditions on $(x, y, t) \in [0,2\pi]^2 \times [0,10]$:

$$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x} - v\frac{\partial u}{\partial y} + \nu(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2})$$

$$\frac{\partial v}{\partial t} = -u\frac{\partial v}{\partial x} - v\frac{\partial v}{\partial y} + \nu(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2})$$

$$u(x, y,0) = u_0(x, y)$$

$$v(x, y,0) = v_0(x, y)$$

$$\nu = 0.1$$

Numerical Example 1:

Correct PDE	PDE-Net 2.0	SINDy	GP	SPL	FEX
$-uu_x$	$-1.00uu_x$	$0uu_x$	$0uu_x$	$0uu_x$	$1.00uu_x$
$-vu_y$	$-1.00vu_y$	$0vu_y$	$0vu_y$	$1.00vu_y$	$1.00vu_y$
$0.05u_{xx}$	$0.0503u_{xx}$	$0.0832u_{xx}$	$0u_{xx}$	$0u_{xx}$	$0.0498u_{xx}$
$0.05u_{yy}$	$0.0503u_{yy}$	$0u_{yy}$	$0u_{yy}$	$0u_{yy}$	$0.0502u_{yy}$
	$5.98 \times 10^{-3}uv$	$5.55 \times 10^{-1}u^2u_x$	$1.695v_x$	u_x	$4.21 \times 10^{-4}u^2$
	$1.56 \times 10^{-3}u$	$-4.62 \times 10^{-1}u_yv^3$	$-vv_x$	0	$4 \times 10^{-4}u^2$
	$-8.02 \times 10^{-4}uu_y$	$-4.76e^{-1}u^2u_x$	$4.6 \times 10^{-2}v$	0	$2 \times 10^{-4}u_x$
	$-6.78e^{-4}uu_xu_y$	$4.36 \times 10^{-2}uu_xu_{yy}$	$4.4 \times 10^{-2}v_{xx}$	0	$2 \times 10^{-4}u_{xy}$
	$-6.62 \times 10^{-4}v_x$	$-4.26 \times 10^-2v^2v_y$	$4.4 \times 10^{-2}v_xu_{yy}$	0	$2 \times 10^{-4}v_x$
	-5.82×10^{-4}	$-3.5 \times 10^{-2}u_yvv_{yy}$	$4.4 \times 10^{-2}v_xu_{xx}$	0	$2 \times 10^{-4}v_y$

	PDE-Net 2.0	SINDy	GP	SPL	FEX
Mean Absolute Error	1.086×10^{-3}	3.239×10^{-1}	4.973×10^{-1}	2.1×10^{-1}	2.021×10^{-4}

Numerical Example 1:

Noise Robustness

Numerical Results of the Burger's equation by PDE-Net with different levels of noise

Correct PDE	$C = 0.001$	$C = 0.005$	$C = 0.01$
$-uu_x$	$-1.00uu_x$	$-1.01uu_x$	$-0.88uu_x$
$-vu_y$	$-1.00vu_y$	$-0.92vu_y$	$-0.80vu_y$
$0.05u_{xx}$	$0.0503u_{xx}$	$0.01u_{xx}$	$0.01u_{xx}$
$0.05u_{yy}$	$0.0503u_{yy}$	$0.02u_{yy}$	$0.01u_{yy}$

Numerical Results of the Burger's equation by FEX with different levels of noise

Correct PDE	$C = 0.001$	$C = 0.005$	$C = 0.01$
$-uu_x$	$-1.00uu_x$	$-1.006uu_x$	$-1.025uu_x$
$-vu_y$	$-1.00vu_y$	$-1.002vu_y$	$-0.926vu_y$
$0.05u_{xx}$	$0.0498u_{xx}$	$0.0534u_{xx}$	$0.0617u_{xx}$
$0.05u_{yy}$	$0.0502u_{yy}$	$0.0543u_{yy}$	$0.0612u_{yy}$

Numerical Example 2:

PDE with varying coefficients and periodic boundary conditions:

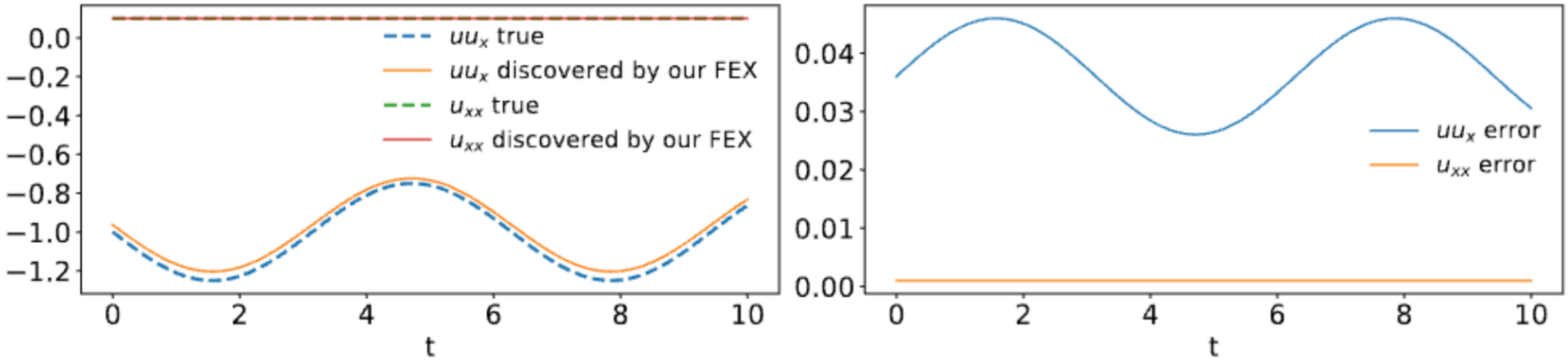
$$u_t(x,t) = a(t)uu_x + \nu u_{xx}, \qquad \forall (x,t) \in [-8,8] \times [0,10],$$

$$u(x,0) = \exp(-(x+1)^2)$$

$$a(t) = 1 + \frac{1}{4}\sin t$$

$$\nu = 0.1$$

Numerical Example 2:



Visualization of the recovery error of varying coefficients

Numerical Example 3:

2D Hopf normal form:

Correct equations	$\dot{x} = \mu x - y - x\left(x^2 + y^2\right)$ $\dot{y} = x + \mu y - y\left(x^2 + y^2\right)$
PDE-Net 2.0	$\dot{x} = 0.912\mu x - 0.998y - 0.930x^3 - 0.935xy^2$ $\dot{y} = 0.999x + 0.960\mu y - 0.979yx^2 - 0.930y^3$
SINDy	$\dot{x} = -0.993y$ $\dot{y} = 0.995x + 1.006\mu y - 1.006yx^2 - 1.008y^3$
GP	$\dot{x} = -y$ $\dot{y} = x$
SPL	$\dot{x} = -0.9391y$ $\dot{y} = x$
FEX	$\dot{x} = 0.984\mu x - 0.997y - 1.003x^3 - 1.004xy^2$ $\dot{y} = 0.996x + 1.014\mu y - 0.995yx^2 - 1.000y^3$

Numerical Example 3:

Correct PDE for $\dot{x}$	PDE-Net 2.0	FEX	Correct PDE for $\dot{y}$	PDE-Net 2.0	FEX
μx	$0.912\mu x$	$0.984\mu x$	x	$0.999x$	$0.996x$
$-y$	$-0.998y$	$-0.997y$	μy	$0.960\mu y$	$1.014\mu y$
$-x^3$	$-0.930x^3$	$-1.003x^3$	$-yx^2$	$-0.979yx^2$	$-0.995yx^2$
$-xy^2$	$-0.935xy^2$	$-1.004xy^2$	$-y^3$	$-0.930y^3$	$-1.000y^3$
	$-3.33 \times 10^{-1}xy^3$	$7.66 \times 10^{-2}y^3$		$-2.45 \times 10^{-1}\mu y^3$	$1.75 \times 10^{-2}\mu^3$
	$3.22 \times 10^{-1}\mu xy$	$6.35 \times 10^{-2}\mu y$		$1.25 \times 10^{-1}x^2y^2$	$-4.05 \times 10^{-3}\mu^2$
	$-3.20 \times 10^{-1}x^3y$	$2.16 \times 10^{-2}x\mu^2$		$1.20 \times 10^{-1}y^4$	$3.99 \times 10^{-3}\mu x^2$
	$3.09 \times 10^{-1}y^4$	$-5.41 \times 10^{-3}\mu^3$		$1.13 \times 10^{-1}\mu xy$	$2.80 \times 10^{-3}xy$
	$-3.04 \times 10^{-1}\mu y^2$	$2.70 \times 10^{-3}x$		$-1.09 \times 10^{-1}\mu xy^2$	$2.20 \times 10^{-3}y$
	$2.72 \times 10^{-1}x^2y^2$	$2.16 \times 10^{-3}x\mu^2$		$-1.03 \times 10^{-1}xy^4$	$1.64 \times 10^{-3}x^3$

	PDE-Net for $\dot{x}$	FEX for $\dot{x}$	PDE-Net for $\dot{y}$	FEX for $\dot{x}$
Mean Absolute Error	2.085×10^{-1}	1.977×10^{-2}	9.470×10^{-2}	5.518×10^{-3}

Numerical Example 4:

Johnson-Mehl-Avrami-Kolmogorov nonlinear equation:

$$y = 1 - \exp\left(-kt^n\right)$$

Correct function	$y = 1 - \exp\left(-0.6t^2\right)$
PDE-Net 2.0	$y = 0.0538t + 0.5013t^2 - 0.0888t^3 + 0.0048t^4 - 0.0024$
SINDy	$y = 0.6015t^2 - 0.2042t^4 + 0.0537t^5$
GP	$y = 0.994 - \exp\left(-0.58t^2\right)$
SPL	$y = 0.5165t^2$
FEX	$y = 1.000 - \exp\left(-0.6001t^2\right)$

Acknowledgement