



# MATEMATICKO-FYZIKÁLNÍ FAKULTA

Univerzita Karlova

---

**Vojtěch Votruba**

## **Rozpoznávání disipativních systémů pomocí strojového učení**

Obhajoba bakalářské práce

---

17. června 2025

- ▶ Zadána data čistě disipativního systém — chemické reakce, difundující částice...

- ▶ Zadána data čistě disipativního systém — chemické reakce, difundující částice...
- ▶ Stav  $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))$

- ▶ Zadána data čistě disipativního systém — chemické reakce, difundující částice...
- ▶ Stav  $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))$
- ▶ Úloha: Sestavení(=„rozpoznání“) modelu

- ▶ Zadána data čistě disipativního systém — chemické reakce, difundující částice...
- ▶ Stav  $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))$
- ▶ Úloha: Sestavení(=„rozpoznání“) modelu
- ▶ Termodynamická konzistence!

# Zobecněná gradientní dynamika

Jak popsat disipativní systém?

- ▶ Formulace: Grmela, Öttinger, 1997
- ▶ Entropie  $S = S(\mathbf{x}) \rightarrow$  Entropické sdružení  $\mathbf{x}^* = \frac{\delta S}{\delta \mathbf{x}}$  (funkcionální derivace)
- ▶ Disipační potenciál  $\Xi(\mathbf{x}, \mathbf{x}^*)$  tak, že
  1.  $\Xi(\mathbf{x}, \mathbf{x}^*) \geq 0$  a zároveň  $\Xi(\mathbf{x}, \mathbf{0}) = 0$
  2.  $\frac{\delta \Xi}{\delta \mathbf{x}^*} \cdot \mathbf{x}^* \geq 0, \quad \forall \mathbf{x}^*$
  3. poblíž  $\mathbf{x}^* = \mathbf{0}$  je  $\Xi$  konvexní
- ▶ Evoluční rovnice:

$$\dot{\mathbf{x}} = \left. \frac{\delta \Xi}{\delta \mathbf{x}^*} \right|_{\mathbf{x}^* = \frac{\delta S}{\delta \mathbf{x}}}$$

# Hluboké učení

Jak destilovat model z dat?

- ▶ Hluboké učení  $\in$  Strojové učení
- ▶ Konstrukce hlubokých neuronových sítí

$$\Lambda(\mathbf{x}; \theta) := \varphi_n(W_n \cdot (\varphi_{n-1}(\dots \varphi_1(W_1 \cdot \mathbf{x} + \mathbf{b}_1) + \mathbf{b}_2)) + \mathbf{b}_n)$$

- ▶  $\theta$  zahrnuje *váhy*  $\{W_i\}_{i=1}^n$  a *prahy*  $\{b_i\}_{i=1}^n$ .  $\varphi$  jsou *aktivační funkce*
- ▶ Neuronové sítě jsou *univerzální aproximátory funkcí* (Leshno et al., 1993)
- ▶ Trénování = optimalizace ztrátové funkce  $L : \mathbb{R}^M \rightarrow [0, +\infty)$

- ▶ Dvě neuronové sítě
  - ▶ Disipační neuronová síť — Konvexní v  $\mathbf{x}^*$
  - ▶ Entropická neuronová síť — Konkávní v  $\mathbf{x}$
  - ▶ (Variational Onsager Neural Networks, Reina et al., 2021)
- ▶ Reparametrizace pro splnění vazeb

$$L \propto \|\dot{\mathbf{x}}^{(\text{nn.})} - \dot{\mathbf{x}}^{(\text{true})}\|^2 = \left\| \frac{\partial \Xi^{(\text{nn.})}}{\partial \mathbf{x}^*} \bigg|_{\mathbf{x}^* = \frac{\partial S^{(\text{nn.})}}{\partial \mathbf{x}}} - \dot{\mathbf{x}}^{(\text{true})} \right\|^2$$

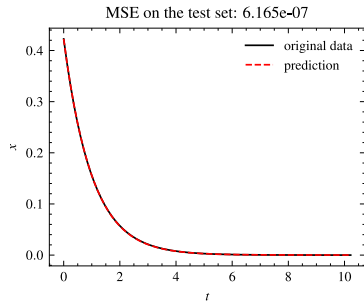
- ▶ Data z konečných diferencí



- ▶ Evoluční rovnice  $\dot{\mathbf{x}} = -\gamma \mathbf{x}$ , kde  $\gamma > 0$
- ▶ Velmi jednoduchá dynamika

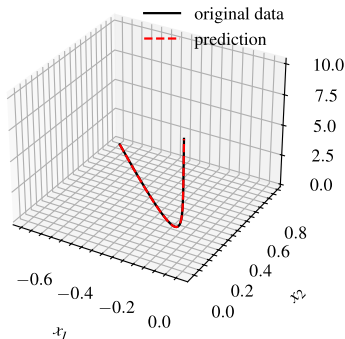
$$S(\mathbf{x}) = -\frac{1}{2} \|\mathbf{x}\|^2, \quad \Xi(\mathbf{x}, \mathbf{x}^*) = \frac{1}{2} \gamma \|\mathbf{x}\|^2$$

### ► Výsledné trajektorie



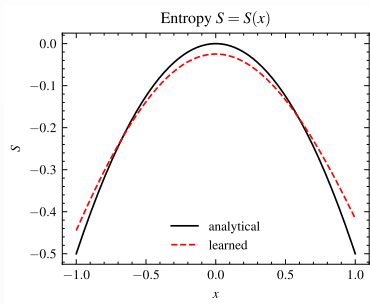
(a) 1D

MSE on the test set:  $1.894\text{e-}06$

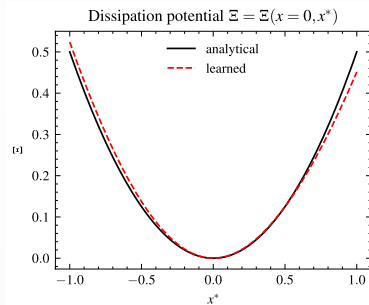


(b) 2D

### ► Naučené funkce 1D

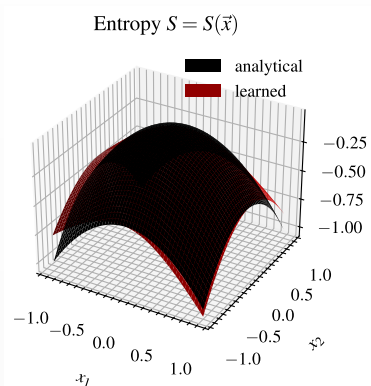


(c) Entropie

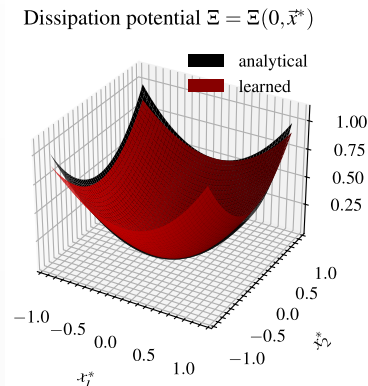


(d) Disipační potenciál

### ► Naučené funkce 2D



(e) Entropie

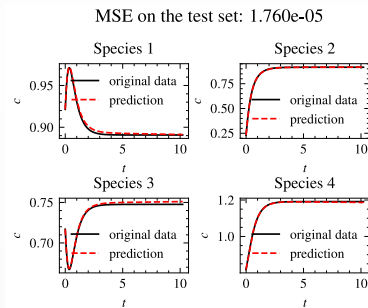


(f) Disipační potenciál

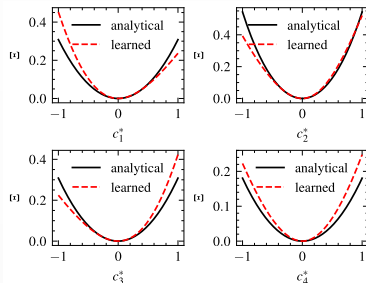
- ▶  $N$  chemických reakcí s  $M$  látkami  $\mathbf{c} = (c^1, \dots, c^M)$
- ▶ Zákon působících hmot (Guldberg, Wagee 1864)
- ▶ Entropie  $S(\mathbf{c}) = -\sum_{m=1}^M c_m (\ln c_m + Q_m)$
- ▶ Disipační potenciál  $\Xi(\mathbf{c}, \mathbf{c}^*) = \sum_{n=1}^N W_n(\mathbf{c}) (\cosh(X_n/2) - 1)$
- ▶ Kde  $X_n = -\sum_{m=1}^M c_m^* \nu_{mn}$ ,  $W_n = W_0 \sqrt{\prod_{m=1}^M c_m^{|\nu_{mn}|}}$
- ▶ Netriviální nejednoznačnost  $\implies$  fixujeme jednu funkci
- ▶ Studovaný systému reakcí:



## ► Předepsaná entropie

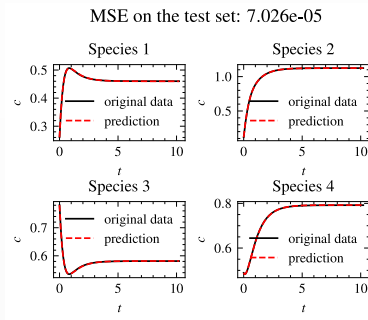


(g) Vývoj pro každou látku

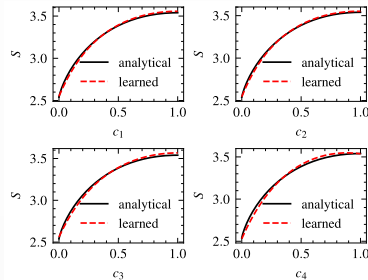


(h) Disipační potenciál

## ► Předepsaný dis. potenciál



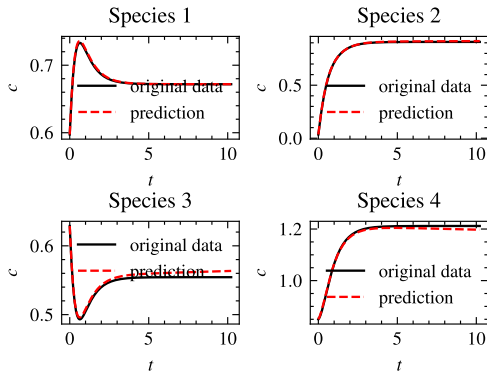
(i) Vývoj pro každou látku



(j) Entropie

► Bez předpisu (pouze dynamika)

MSE on the test set: 2.051e-05

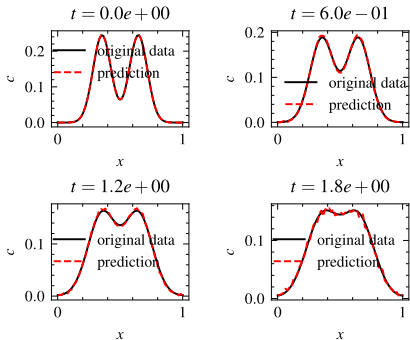




- ▶ Rovnice (lineární) difuze:  $\dot{c}(\mathbf{r}) = D\Delta c(\mathbf{r})$
- ▶ Skrze gradientní dynamiku:
  - ▶ Entropie  $S[c(\mathbf{r})] = - \int c \log(c) d^3r$
  - ▶ Disipační potenciál  $\Xi[c(\mathbf{r}), \nabla c^*(\mathbf{r})] = \frac{1}{2} \int Dc \|\nabla c^*\|^2 d^3r$
- ▶ Trénování na fourierovských módech s period, okraj, podmínkami

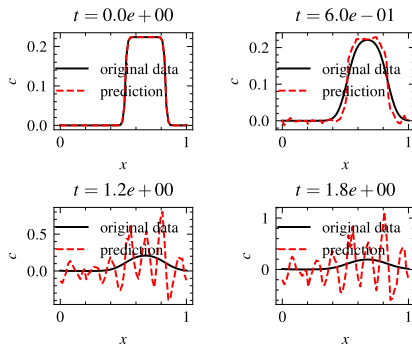
### ► Dva gaussové píky

MSE on the test set:  $2.617e-06$



### ► Nespojitá poč. podmínka

MSE on the test set:  $2.341e-03$



- ▶ Integrace s reverzibilními sítěmi pomocí teorie GENERIC
  - ▶  $\dot{\mathbf{x}} = \{\mathbf{x}, H\} + \frac{\delta \Xi}{\delta \mathbf{x}^*} \Big|_{\mathbf{x}^* = \frac{\delta S}{\delta \mathbf{x}}}$
- ▶ Rozšíření na prostory funkcí pomocí neuronových operátorů (Kovachki, Li et al, 2023)

## Děkuji za pozornost

Za vedení práce a možnost prezentovat její výsledky na konferenci IWNET vřele  
děkuji svému vedoucímu doc. Michalu Pavelkovi.

- „Jak souvisí nejpravděpodobnější dráha s derivací v rovnici pod (1.5) dle  $\eta$ ?”
- Rovnice (1.5) s derivací dle  $\eta$  byla zkratkovitá. Plné pochopení problematiky by vyžadovalo hlubší znalosti konvexní analýzy, nicméně z podmínky  $\mathcal{L} = 0$ , jež platí pro nejpravděpodobnější dráhu, má platit

$$\Psi^*(\mathbf{x}, \dot{\mathbf{x}}) = -\eta \cdot \dot{\mathbf{x}} - \Psi(\mathbf{x}, -\eta).$$

A zároveň dle definice

$$\begin{aligned}\Psi^*(\mathbf{x}, \dot{\mathbf{x}}) &= \sup_{\eta} \{-\eta \cdot \dot{\mathbf{x}} - \Psi(\mathbf{x}, -\eta)\} \\ \implies \frac{\delta}{\delta \eta} (-\eta \cdot \dot{\mathbf{x}} - \Psi(\mathbf{x}, -\eta)) &= 0.\end{aligned}$$

- ▶ „Existuje nějaká klíčová vlastnost, kterou  $\varphi$  v neuronových sítích, rovnice (2.1), mají mít?“
- ▶ Klíčovou vlastností je pouze nelinearita. Obzvlášť pro naše aplikace se dále hodí diferencovatelnost a při nevhodné volbě  $\varphi$  mohou dělat problém vybuchující/mizející gradienty (např. kdybychom použili  $\exp(x)$ ).

- ▶ „V předpisu pro  $\theta_{i+1}$  v *second-order methods* na str. 16 užíváte  $(\nabla^2 L)^{-1}$  inverzi Hessovy matice. Není toto však problém pro funkce typu ReLU (viz též předchozí bod)?“
- ▶ Použití ReLU s L-BFGS nemusí být ideální, neb  $\frac{\partial^2}{\partial y^2} \text{ReLU}(y) = 0$ , nicméně pro složenou funkci derivace vychází jako  $\frac{\partial^2}{\partial y^2} \text{ReLU}(f(y)) = f''(y) \frac{\partial}{\partial f} \text{ReLU}(f(y))$ . Pokud dotaz naráží na nespojitost derivace v počátku, tak to standardně není problém.



- ▶ „Při učení dynamiky předpokládáte od začátku znalost  $\mathbf{x}$ . Lze rozšířit metodu tak, aby se učila i stavové proměnné?“
- ▶ Nejspíše ne. Neumím si příliš představit, jak by pak měl vypadat vstup do neuronové sítě ve smyslu naměřených dat. Teoreticky je ovšem možné model upravit tak, aby byl agnostický vůči dimenzi dat, tomu bych se rád věnoval budoucnu (viz neuronové operátory).

- ▶ „Nejednoznačnost úlohy učení v  $S, \Xi$  je zřejmá. Neznamená to však, že numerické aproximativní metody musí selhávat při hledání řešení, dokud se tato nejednoznačnost neodstraní? Např. v Figure 4.6 vidíme, že shoda není ideální, ale přijde mi, že jde o nevhodnou volbu  $C, K$ . Nešlo by přidat tyto dvě konstanty do učení modelu neuronovými sítěmi? Podobně je tomu asi v dalších příkladech. Např. výsledky v Figure 4.10, si myslím, lze vylepšit vhodnější volbou shift a scalingu.“
- ▶ Neznamená. Máme-li například několik stejně hlubokých minim, Newtonova metoda může jednoduše zkonvergovat do jednoho z nich a neznamená to její selhání. Vzhledem k tomu, že neuronová síť nemá mít a priori přístup ke správnému potenciálu, zahrnutí  $C, K$  není možné. Vzorce pro výpočet těchto škálovacích konstant jsou uvedeny v práci a byly zvoleny podle metody nejmenších čtverců.

- ▶ „*Nebylo by lepší nagenеровaná numerická data trochu zašumit před aplikací metody?*“
- ▶ Nejspíše ano. Zašumění je forma regularizace a zároveň odpovídá fyzikální realitě. Tato práce sloužila spíše jako *proof of concept*. Při aplikaci na komplikovanější systémy by zašumění jistě bylo správným krokem.

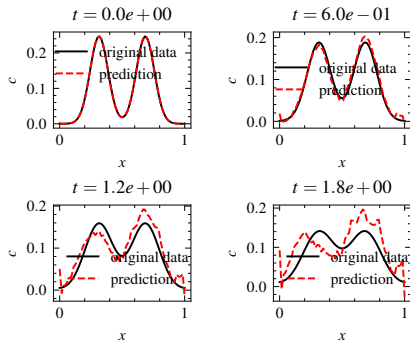
- ▶ „V chemických reakcích typu (4.2) bývá několik zákonů zachování. Není toto pro předkládanou metodu problém? A lze se tyto zákony zachování naučit (odhalit je) neuronovou sítí?“
- ▶ Nevidím žádný důvod, kvůli kterému by zákony zachování měly být problém. Naopak, pokud bychom je znali předem, můžeme je přičíst k loss funkci a zefektivnit tak trénink.  
Hledání zachovávajících se veličin je obecně možné a jde o dobrý podnět do budoucího výzkumu. Jistým způsobem to bylo provedeno v článku:  
[Discovering symmetry invariants and conserved quantities by interpreting siamese neural networks](#), S. J. Wetzal et al., 2020.

- ▶ „V obecné situaci v chemických reakcích, str. 36-37, se podařilo naučit dynamiku, ale strukturu (entropii) již méně úspěšně. Prosím však o kontrolu správnosti Figure 4.15 a 4.16 — první říká, že entropie roste ve všech  $c_j$ , což se zdá být v rozporu s druhým obrázkem. Navíc, pokud se nepletu, ani variace (v abs. hodnotě) si neodpovídají.“
- ▶ Na str. 35 je detailně popsáno, že (bez újmy na dynamiku) entropie může být afinně posunutá o hodnotu  $\mathbf{q} \cdot \mathbf{c}$ , pakliže  $\mathbf{q} \cdot \boldsymbol{\nu} = 0$ . Tento posun pak samozřejmě mění znaménko i hodnotu derivací (dokonce každou posouvá jinak). Figure 4.15 obsahuje grafy již posunuté entropie, tak aby se dala co nejlépe porovnat s analytickou hodnotou. Naproti tomu Figure 4.16 bere "surovou" neposunutou entropii a počítá derivace pouze proto, abychom ukázali, že nemá lineární průběh. Proto se obrázky neshodují.

- ▶ „Uvažovaná počáteční podmínka ve fickovské difuzi může být příliš speciální (Fourierovské módy velmi úzce souvisí s difuzí). Nemůže toto souviset s charakterem výsledku (oscilace) v Figure 4.20? Jak by vypadal výsledek učení pro náhodnou počáteční podmínku?“
- ▶ Nejsem si jistý, co je přesně myšleno náhodnou počáteční podmínkou. Jistá míra náhody je zařazená již ve fourierovských módech formou koeficientů a jejich počtu. Pokusil jsem se tedy trénovat model na poč. podmínce, kdy je každému bodu v mříži přiřazena hodnota z  $(-1, 1)$  dle uniformního rozdělení.

- Dva gaussovské píky
- MSE se zvýšila

MSE on the test set:  $1.755e-04$



- Nespojité poč. podmínky
- MSE se snížila

