Emulating ISM non equilibrium (photo-thermo)-chemistry with Deep Neural Operators

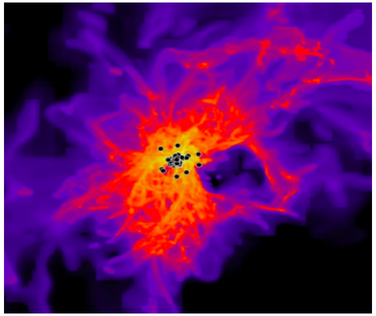
Lorenzo Branca

25/09/2024

Lorenzo Branca & Andrea Pallottini, A&A, 10.1051/0004-6361/202449193

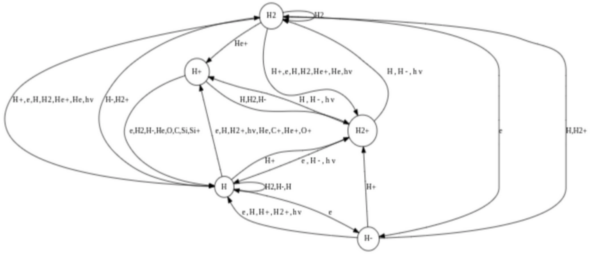
Introduction

- Context: realistic astrophysical simulations
  needs to compute several processes
  (gravity, fluid-dynamic, radiation and chemistry).



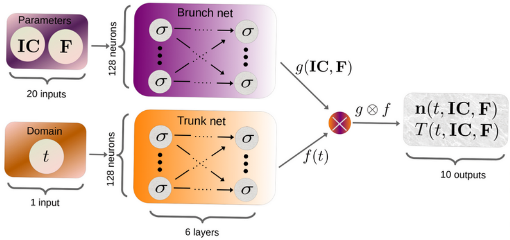
Decataldo et al;2020

- Problem: non-equilibrium chemistry is among the most difficult tasks to include in astrophysical simulations:
  - high (>40) number of reactions
  - short evolutionary timescales
  - non-linearity and stiffness of the associate ODEs
  - Load balancing for parallel computation



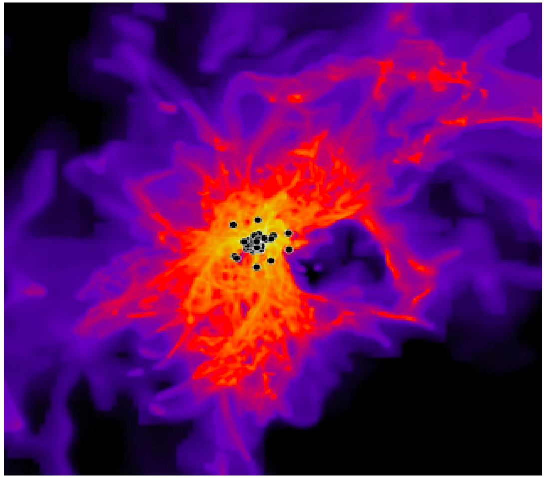
Bovino et al; 2015

- Aim: Replacing of classical ODEs solvers
  with fast deep learning based emulators



InterStellar Medium physics

Decataldo et al;2020

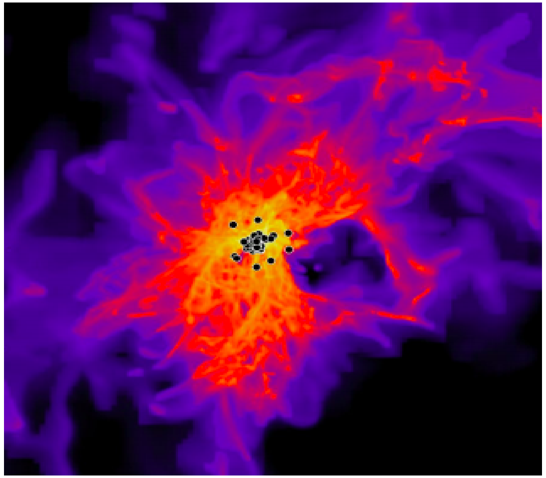


Molecular Cloud (MC)
simulation
about 10^8 finite elements

total of ~350 kCPUhr

InterStellar Medium physics

Decataldo et al;2020



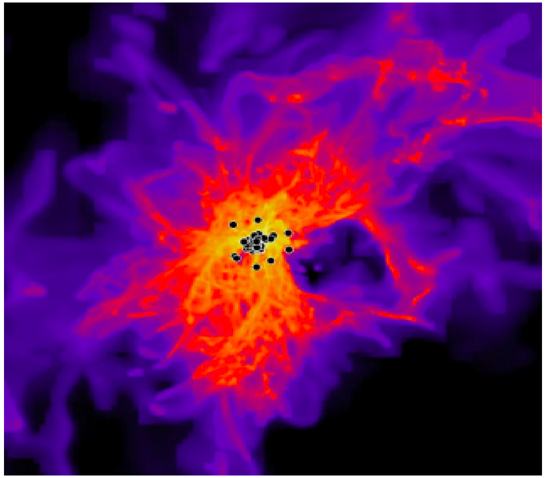
Molecular Cloud (MC)
simulation
about 10^8 finite elements

total of ~350 kCPUhr

processes		approximate CPU cost
(self)gravity	$\nabla^2\Phi = 4\pi G\rho$	~10%

InterStellar Medium physics

Decataldo et al;2020



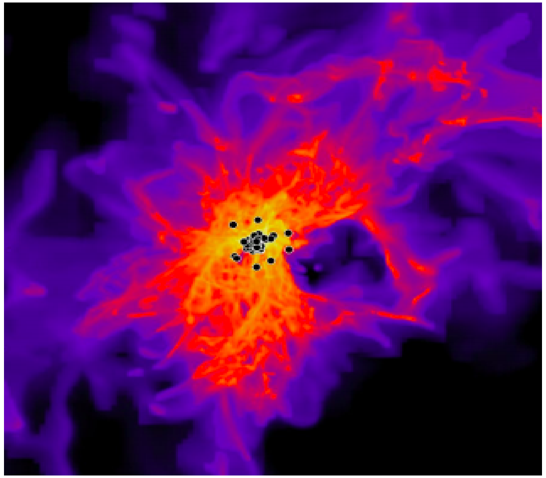
Molecular Cloud (MC)
simulation
about 10^8 finite elements

total of ~350 kCPUhr

processes		approximate CPU cost
(self)gravity	$\nabla^2 \Phi = 4\pi G \rho$	~10%
Fluid dynamics	$\partial_t \mathcal{U} + \nabla \cdot \mathcal{F} = \mathcal{S}$	~20%

InterStellar Medium physics

Decataldo et al;2020



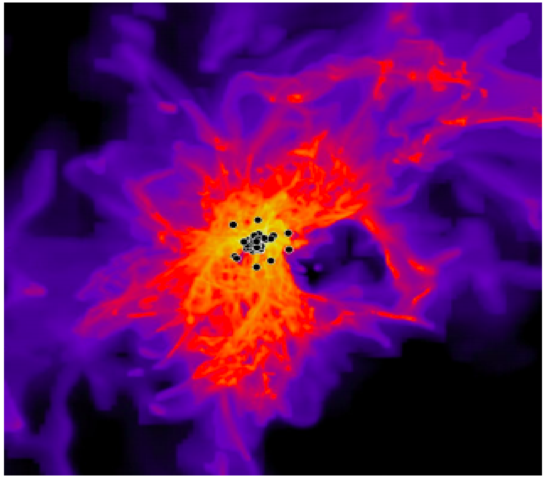
Molecular Cloud (MC) simulation
about 10^8 finite elements

total of ~350 kCPUhr

processes		approximate CPU cost
(self)gravity	$\nabla^2\Phi = 4\pi G\rho$	~10%
Fluid dynamics	$\partial_t\mathcal{U} + \nabla\cdot\mathcal{F} = \mathcal{S}$	~20%
Radiative transfer	$\partial_t I_\nu + \hat{n}\nabla\cdot I_\nu = j_\nu - k_\nu I_\nu$	~20%

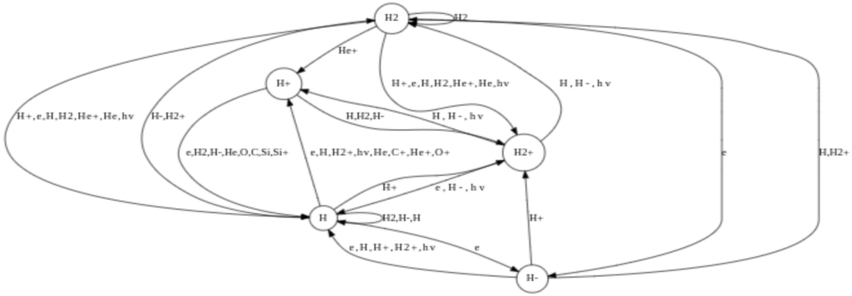
InterStellar Medium physics

Decataldo et al;2020



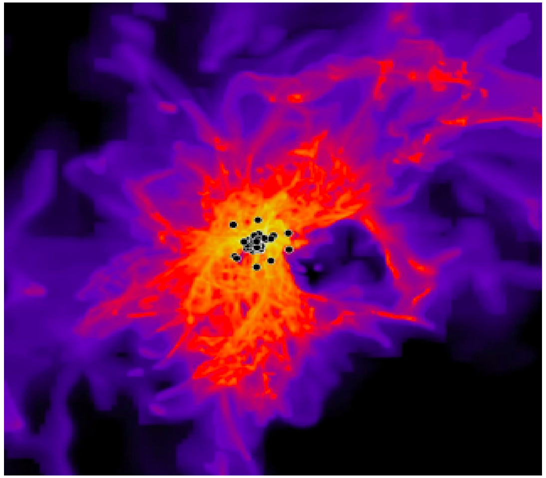
Molecular Cloud (MC) simulation
about 10^8 finite elements

total of ~350 kCPUhr

processes		approximate CPU cost
(self)gravity	$$\nabla^2 \Phi = 4\pi G \rho$$	~10%
Fluid dynamics	$$\partial_t \mathcal{U} + \nabla \cdot \mathcal{F} = \mathcal{S}$$	~20%
Radiative transfer	$$\partial_t I_\nu + \hat{n} \nabla \cdot I_\nu = j_\nu - k_\nu I_\nu$$	~20%
chemistry		~50%

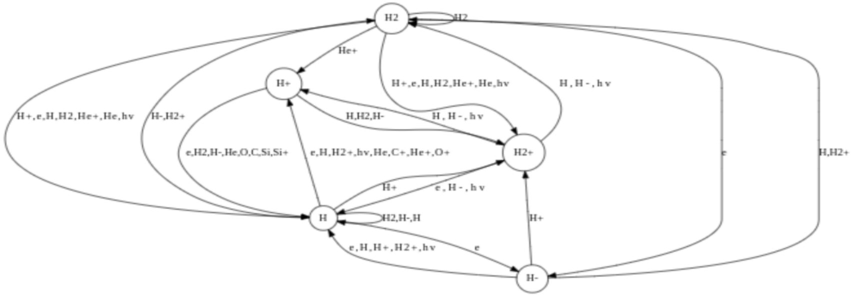
InterStellar Medium physics

Decataldo et al;2020



Molecular Cloud (MC) simulation
about 10^8 finite elements

total of ~350 kCPUhr

processes		approximate CPU cost
(self)gravity	$\nabla^2\Phi = 4\pi G\rho$	~10%
Fluid dynamics	$\partial_t\mathcal{U} + \nabla\cdot\mathcal{F} = \mathcal{S}$	~20%
Radiative transfer	$\partial_t I_\nu + \hat{n}\nabla\cdot I_\nu = j_\nu - k_\nu I_\nu$	~20%
chemistry		~50%

non-equilibrium chemistry evolution plays a crucial role in cosmological and astrophysical phenomena, especially in the study of InterStellar Medium (ISM)

ISM (photo-thermo)-chemistry

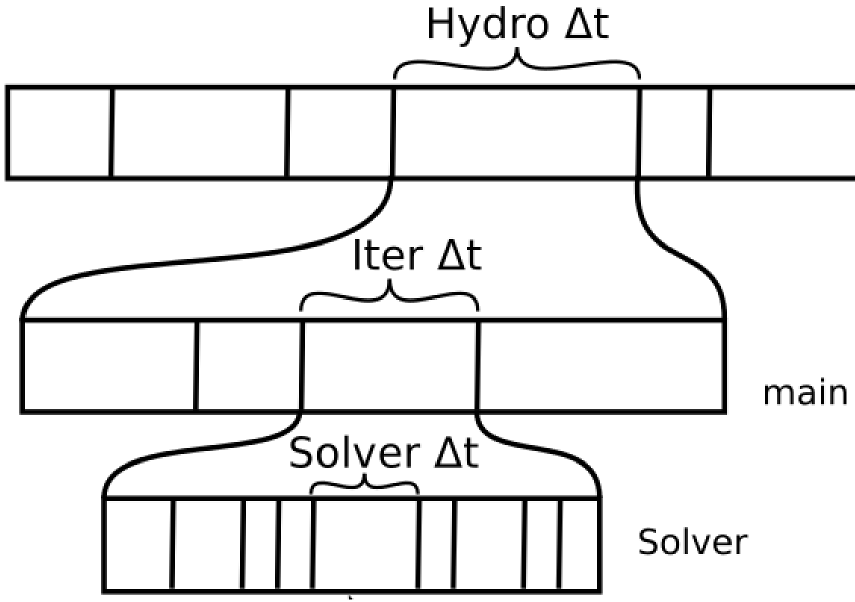
$$\dot{n}_k = A_k^{ij} n_i n_j + B_k^i n_i$$

$$\dot{T} = \Gamma - \Lambda$$

ISM (photo-thermo)-chemistry

$$\dot{n}_k = A_k^{ij} n_i n_j + B_k^i n_i$$

$$\dot{T} = \Gamma - \Lambda$$

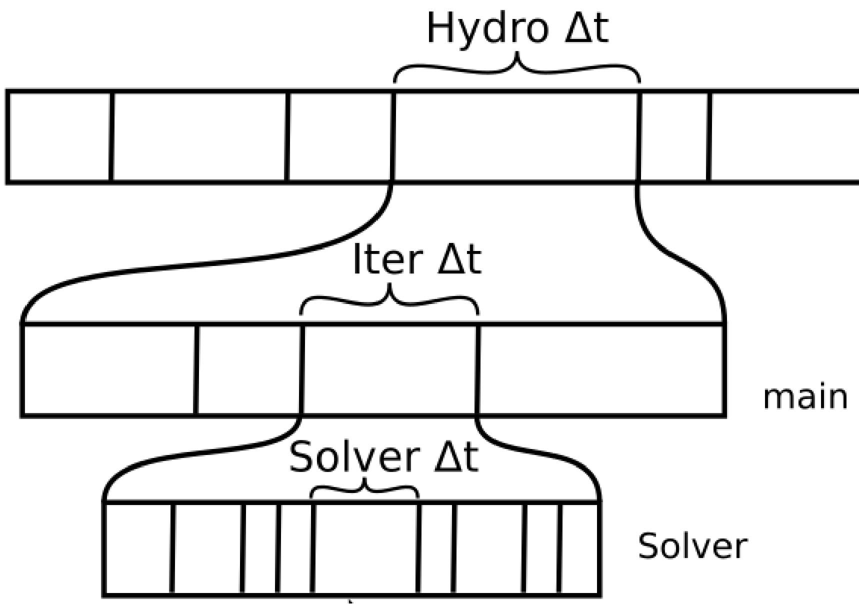


ISM (photo-thermo)-chemistry

$$\dot{n}_k = A_k^{ij} n_i n_j + B_k^i n_i$$

$$\tau_{chem}/\tau_{HD} < 10^{-4}$$

$$\dot{T} = \Gamma - \Lambda$$



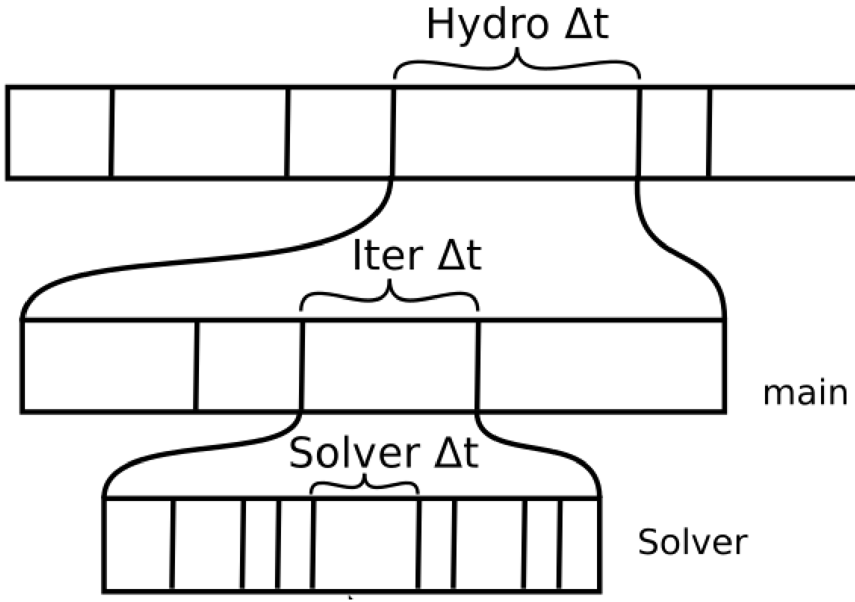
ISM (photo-thermo)-chemistry

$$\dot{n}_k = A_k^{ij} n_i n_j + B_k^i n_i$$

$$\dot{T} = \Gamma - \Lambda$$

$$\tau_{chem}/\tau_{HD} < 10^{-4}$$

$$N_{rea} \propto N_s^{\alpha}, \alpha > 1$$



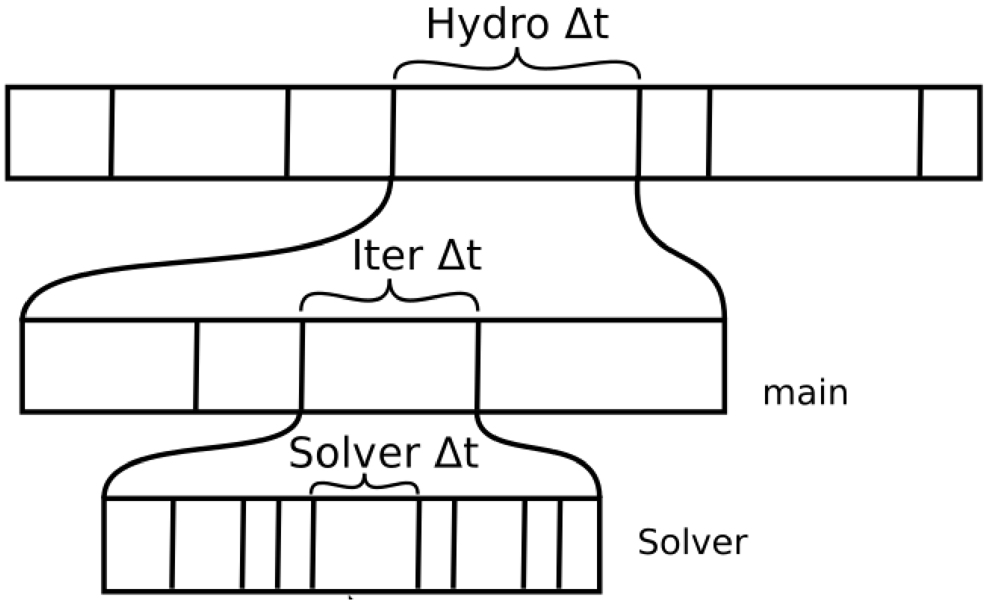
ISM (photo-thermo)-chemistry

$$\dot{n}_k = A_k^{ij} n_i n_j + B_k^i n_i$$

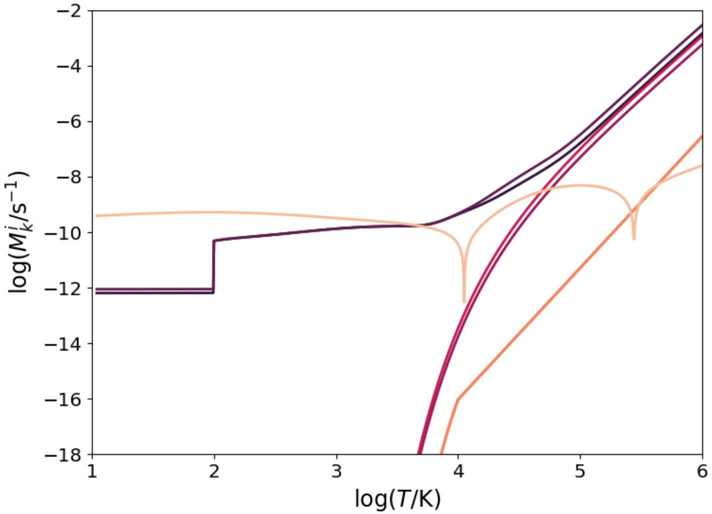
$$\dot{T} = \Gamma - \Lambda$$

$$\tau_{chem}/\tau_{HD} < 10^{-4}$$

$$N_{rea} \propto N_s^{\alpha}, \alpha > 1$$



Stifness



ISM (photo-thermo)-chemistry

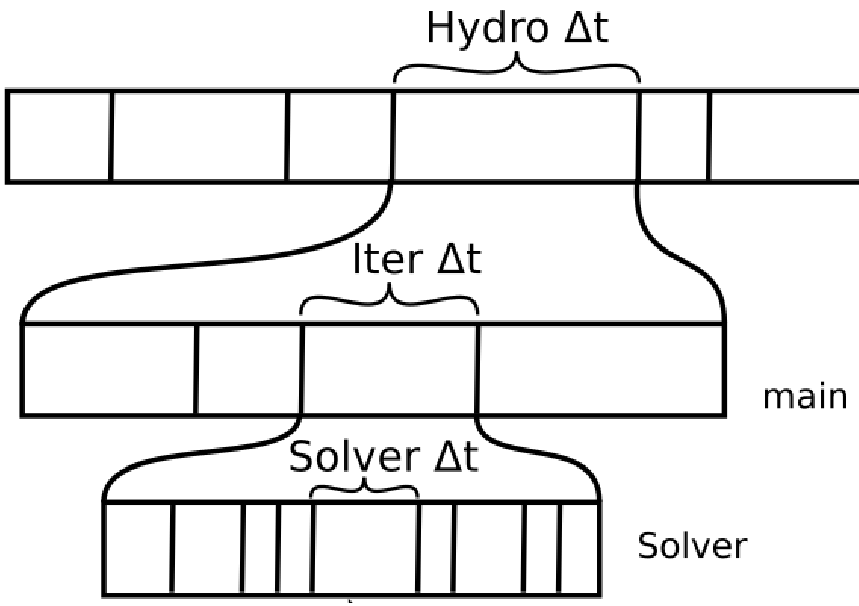
$$\dot{n}_k = A_k^{ij} n_i n_j + B_k^i n_i$$

$$\tau_{chem}/\tau_{HD} < 10^{-4}$$

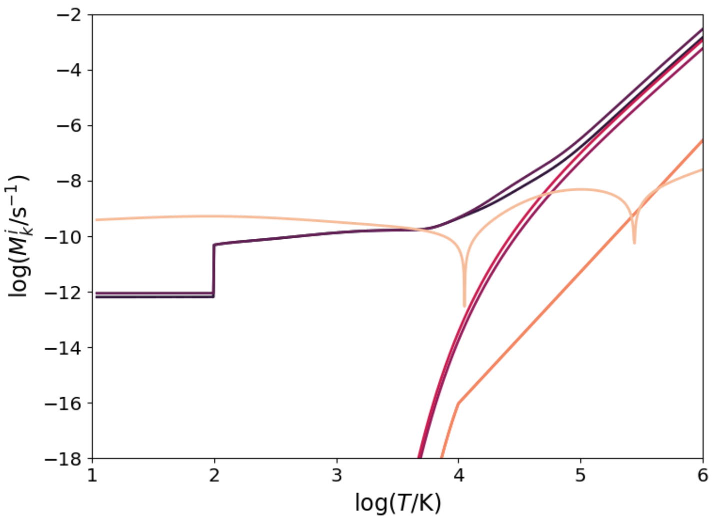
$$\dot{T} = \Gamma - \Lambda$$

Load balancing

$$N_{rea} \propto N_s^{\alpha}, \alpha > 1$$

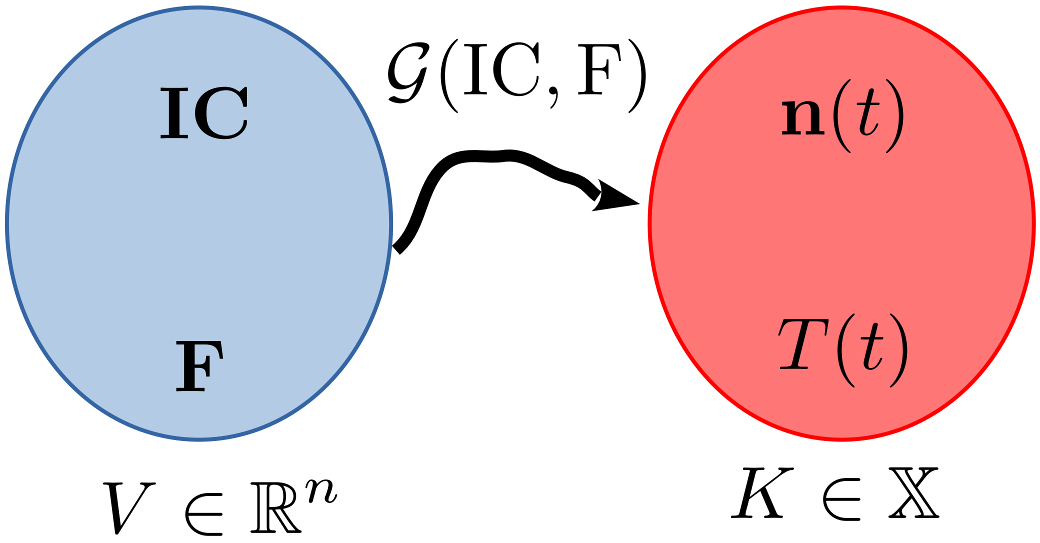


Stifness

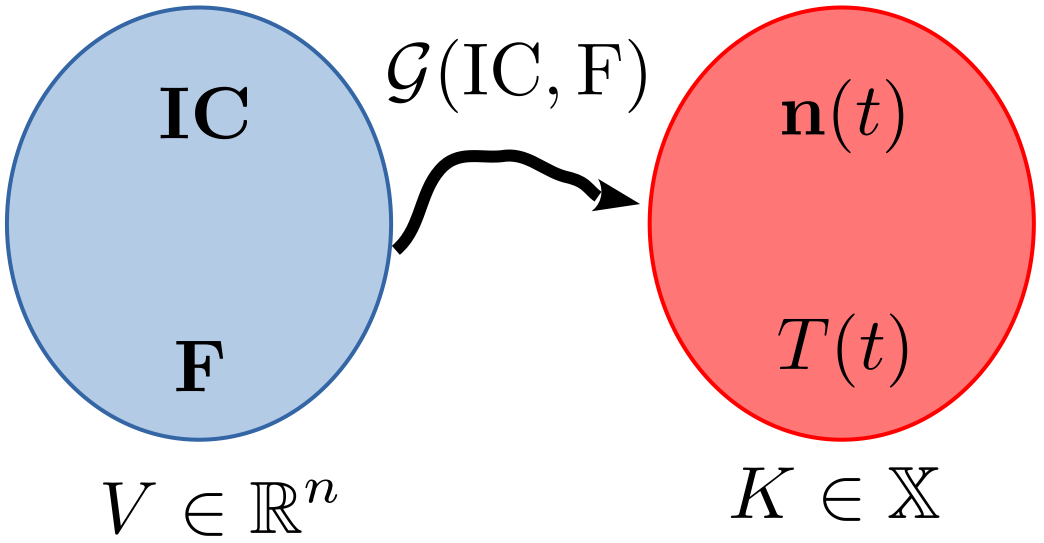


ODEs emulator

ODEs emulator

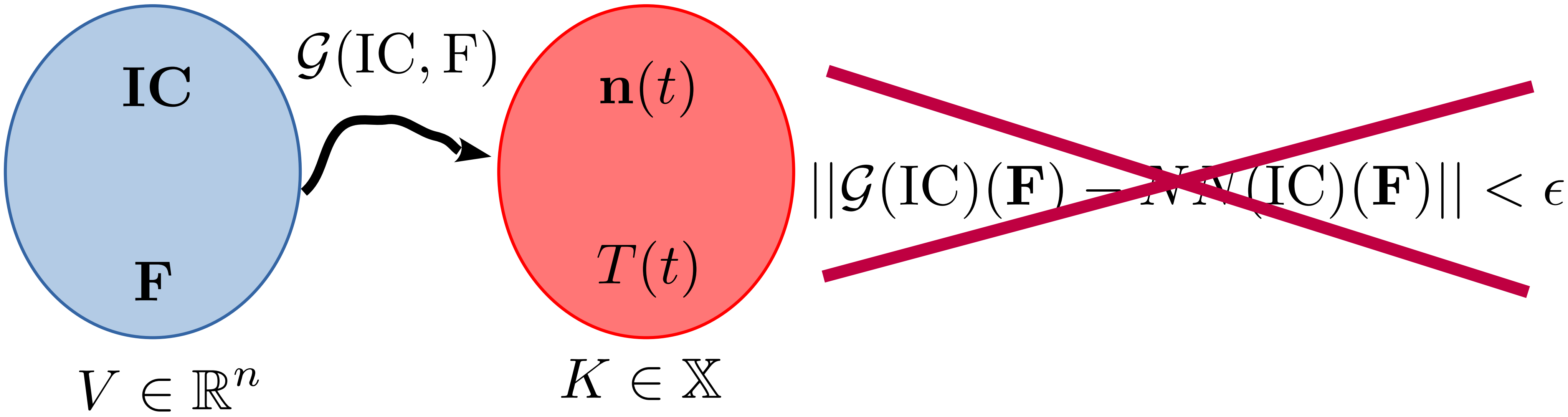


ODEs emulator



$$||\mathcal{G}(\mathrm{IC})(\mathbf{F}) - NN(\mathrm{IC})(\mathbf{F})|| < \epsilon$$

ODEs emulator



Universal approximation theorem — Let $C(X, \mathbb{R}^m)$ denote the set of continuous functions from a subset X of a Euclidean $\mathbb{R}^n$ space to a Euclidean space $\mathbb{R}^m$. Let $\sigma \in C(\mathbb{R}, \mathbb{R})$. Note that $(\sigma \circ x)_i = \sigma(x_i)$, so $\sigma \circ x$ denotes σ applied to each component of x.

Then σ is not polynomial if and only if for every $n \in \mathbb{N}$, $m \in \mathbb{N}$, compact $K \subseteq \mathbb{R}^n$, $f \in C(K, \mathbb{R}^m), \varepsilon > 0$ there exist $k \in \mathbb{N}$, $A \in \mathbb{R}^{k \times n}$, $b \in \mathbb{R}^k$, $C \in \mathbb{R}^{m \times k}$ such that

$$\sup_{x \in K} \|f(x) - g(x)\| < \varepsilon$$

where $g(x) = C \cdot (\sigma \circ (A \cdot x + b))$

Operators approximation

Operators approximation

Theorem 1 (Universal Approximation Theorem for Operator). *Suppose that σ is a continuous non-polynomial function, X is a Banach space, $K_1 \subset X$, $K_2 \subset \mathbb{R}^d$ are two compact sets in X and $\mathbb{R}^d$, respectively, V is a compact set in $C(K_1)$, G is a nonlinear continuous operator, which maps V into $C(K_2)$. Then for any $\epsilon > 0$, there are positive integers n, p and m, constants c_i^k, ξ_{ij}^k, θ_i^k, $\zeta_k \in \mathbb{R}$, $w_k \in \mathbb{R}^d$, $x_j \in K_1$, $i = 1, \ldots, n$, $k = 1, \ldots, p$ and $j = 1, \ldots, m$, such that*

$$\left| G(u)(y) - \sum_{k=1}^{p} \underbrace{\sum_{i=1}^{n} c_i^k \sigma \left(\sum_{j=1}^{m} \xi_{ij}^k u(x_j) + \theta_i^k \right)}_{\text{branch}} \underbrace{\sigma(w_k \cdot y + \zeta_k)}_{\text{trunk}} \right| < \epsilon \tag{1}$$

holds for all $u \in V$ and $y \in K_2$. Here, $C(K)$ is the Banach space of all continuous functions defined on K with norm $\|f\|_{C(K)} = \max_{x \in K} |f(x)|$.

Operators approximation

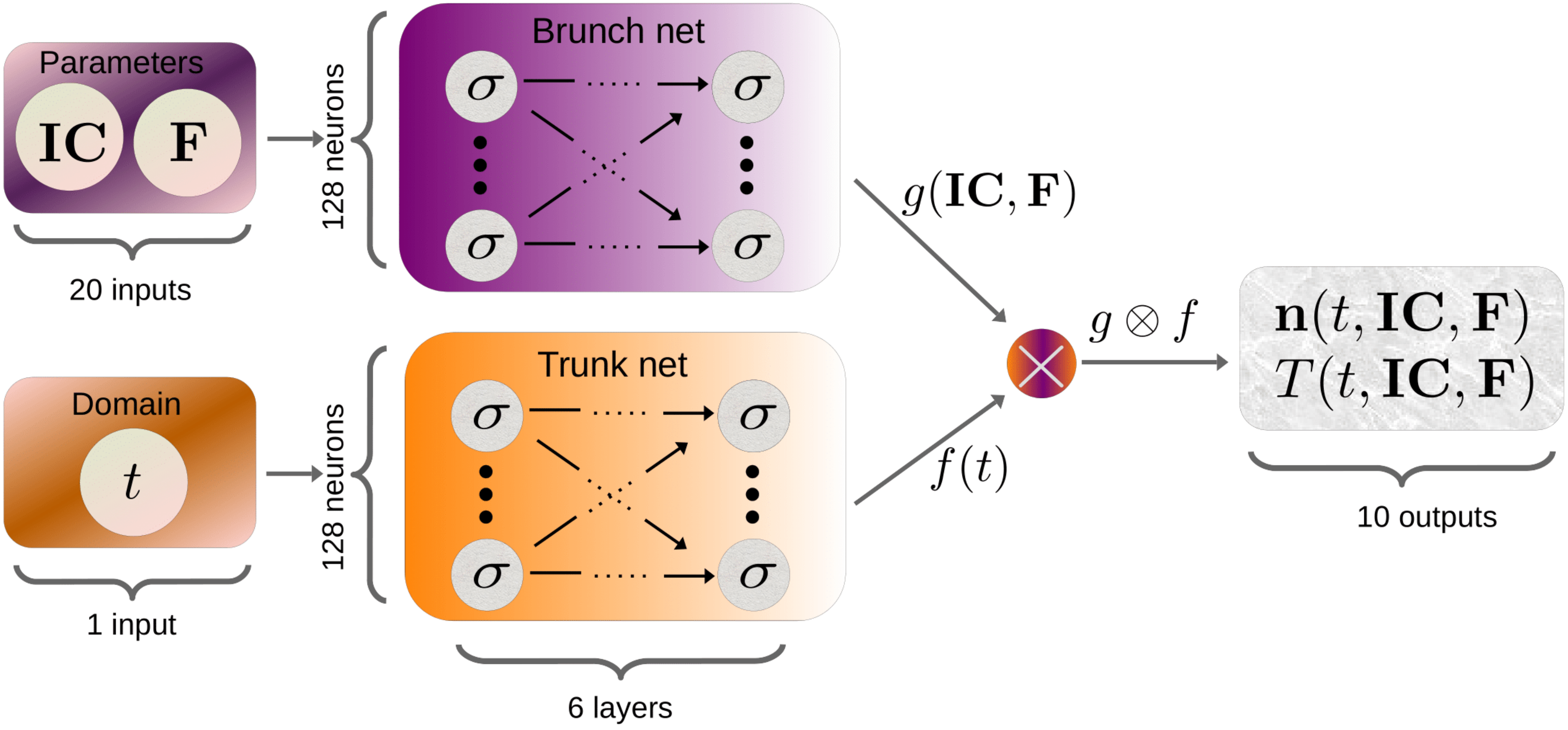
Theorem 1 (Universal Approximation Theorem for Operator). *Suppose that σ is a continuous non-polynomial function, X is a Banach space, $K_1 \subset X$, $K_2 \subset \mathbb{R}^d$ are two compact sets in X and $\mathbb{R}^d$, respectively, V is a compact set in $C(K_1)$, G is a nonlinear continuous operator, which maps V into $C(K_2)$. Then for any $\epsilon > 0$, there are positive integers n, p and m, constants $c_i^k, \xi_{ij}^k, \theta_i^k, \zeta_k \in \mathbb{R}$, $w_k \in \mathbb{R}^d$, $x_j \in K_1$, $i = 1, \ldots, n$, $k = 1, \ldots, p$ and $j = 1, \ldots, m$, such that*

$$\left| G(u)(y) - \underbrace{\sum_{k=1}^{p} \sum_{i=1}^{n} c_i^k \sigma\left(\sum_{j=1}^{m} \xi_{ij}^k u(x_j) + \theta_i^k\right)}_{\text{branch}} \underbrace{\sigma(w_k \cdot y + \zeta_k)}_{\text{trunk}} \right| < \epsilon \quad (1)$$

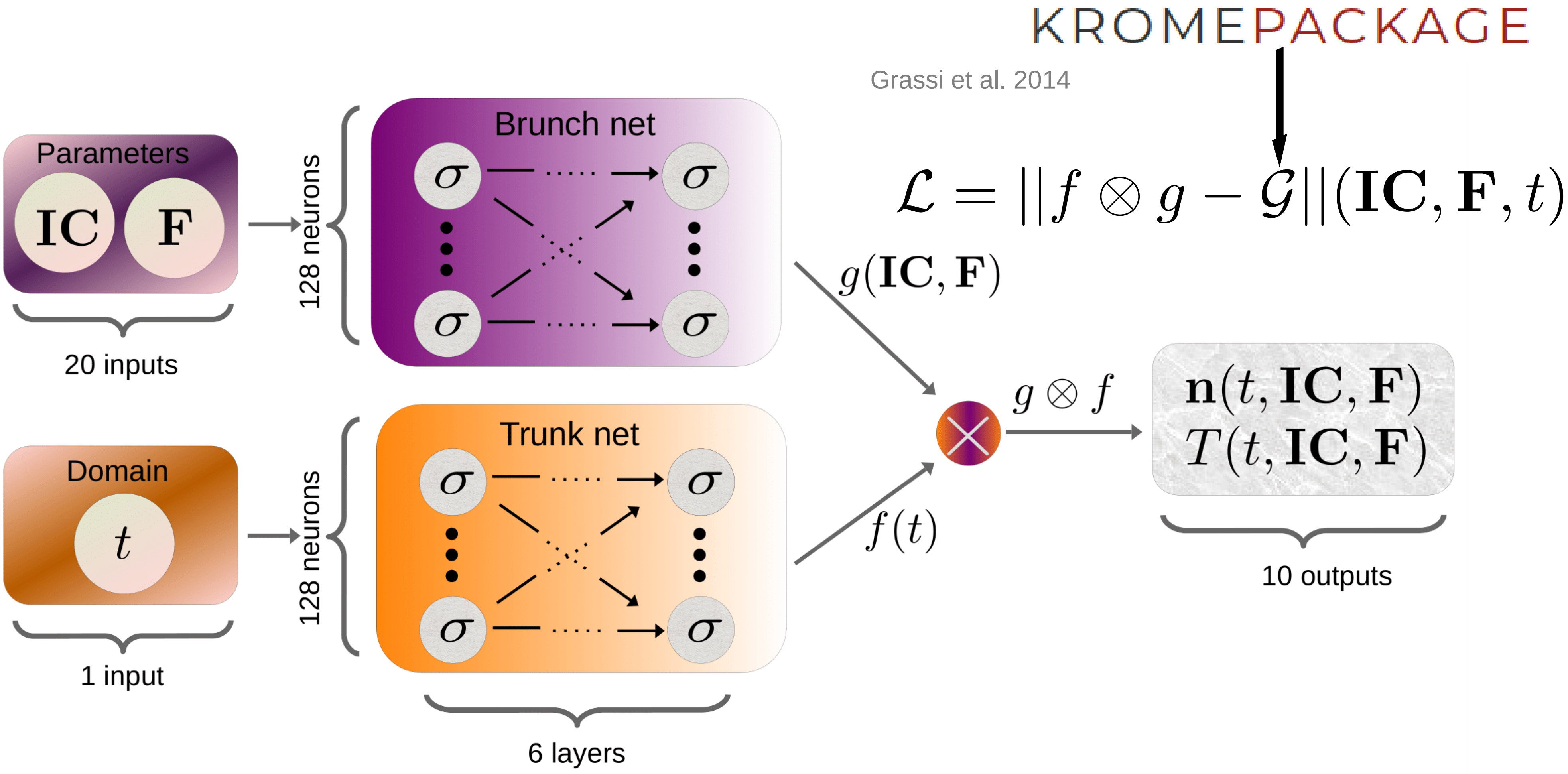
holds for all $u \in V$ and $y \in K_2$. Here, $C(K)$ is the Banach space of all continuous functions defined on K with norm $\|f\|_{C(K)} = \max_{x \in K} |f(x)|$.

$$||\mathcal{G}(u)(y) - f(u) \otimes g(y)|| < \epsilon$$

Deep Neural Operator Network



Deep Neural Operator Network



Data

quantity	variable	bins	min	max
gas density	$\log(n/\mathrm{cm}^{-3})$	64	-2	3.5
abundances	$\log(n_i/n)$	512	-6	0
temperature	$\log(T/\mathrm{K})$	64	log(20)	5.5
radiation	$\log(F_i/\mathrm{eV\,cm^{-2}\,s^{-1}\,Hz})$	64	-15	-5
time	t/kyr	16	0	1

Data

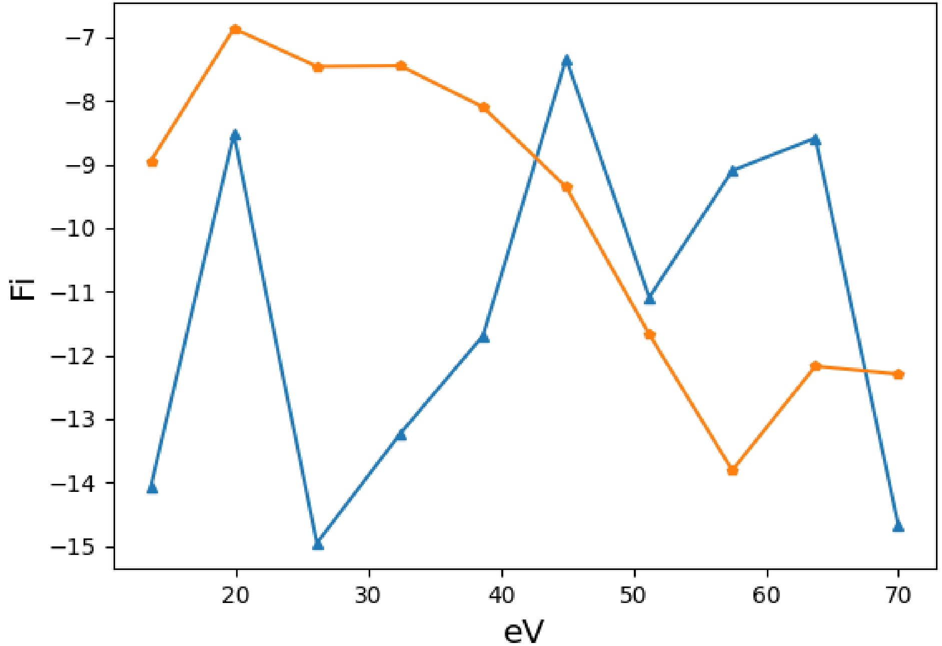
quantity	variable	bins	min	max
gas density	$\log(n/\mathrm{cm}^{-3})$	64	-2	3.5
abundances	$\log(n_i/n)$	512	-6	0
temperature	$\log(T/\mathrm{K})$	64	log(20)	5.5
radiation	$\log(F_i/\mathrm{eV\,cm^{-2}\,s^{-1}\,Hz})$	64	-15	-5
time	t/kyr	16	0	1

Photo-reaction			$h\nu_{\mathrm{min}}/\mathrm{eV}$
$H^- + \gamma$	$\rightarrow$	$H + e$	0.76
$H_2^+ + \gamma$	$\rightarrow$	$H^+ + H$	2.65
$H_2 + \gamma$	$\rightarrow$	$H + H$ (Solomon)	11.2
$H + \gamma$	$\rightarrow$	$H^+ + e$	13.6
$H_2 + \gamma$	$\rightarrow$	$H + H$ (direct)	14.2
$H_2 + \gamma$	$\rightarrow$	$H_2^+ + e$	15.4
$He + \gamma$	$\rightarrow$	$He^+ + e$	24.6
$H_2^+ + \gamma$	$\rightarrow$	$H^+ + H^+ + e$	30.0
$He^+ + \gamma$	$\rightarrow$	$He^{++} + e$	54.4

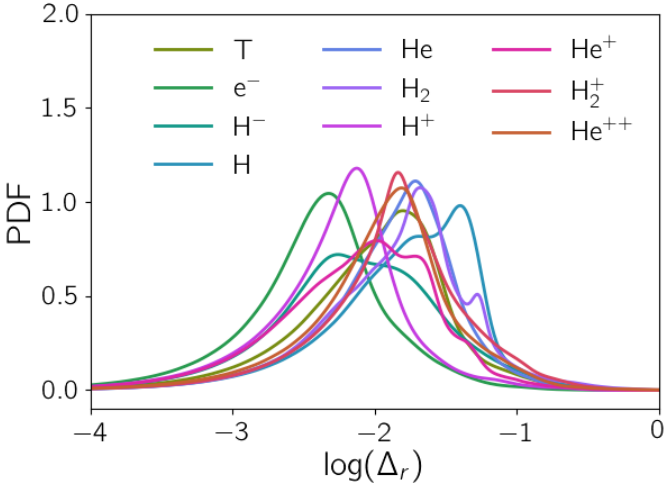
Data

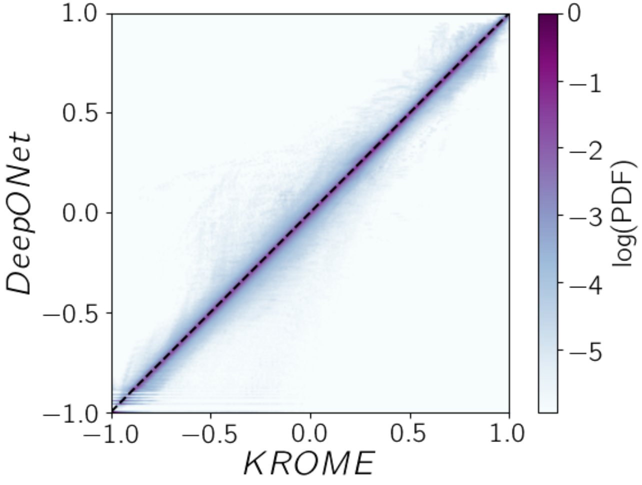
quantity	variable	bins	min	max
gas density	$\log(n/\mathrm{cm}^{-3})$	64	-2	3.5
abundances	$\log(n_i/n)$	512	-6	0
temperature	$\log(T/\mathrm{K})$	64	log(20)	5.5
radiation	$\log(F_i/\mathrm{eV\,cm^{-2}\,s^{-1}\,Hz})$	64	-15	-5
time	t/kyr	16	0	1

Photo-reaction			$h\nu_{\min}/\mathrm{eV}$
$H^- + \gamma$	$\rightarrow$	$H + e$	0.76
$H_2^+ + \gamma$	$\rightarrow$	$H^+ + H$	2.65
$H_2 + \gamma$	$\rightarrow$	$H + H$ (Solomon)	11.2
$H + \gamma$	$\rightarrow$	$H^+ + e$	13.6
$H_2 + \gamma$	$\rightarrow$	$H + H$ (direct)	14.2
$H_2 + \gamma$	$\rightarrow$	$H_2^+ + e$	15.4
$He + \gamma$	$\rightarrow$	$He^+ + e$	24.6
$H_2^+ + \gamma$	$\rightarrow$	$H^+ + H^+ + e$	30.0
$He^+ + \gamma$	$\rightarrow$	$He^{++} + e$	54.4

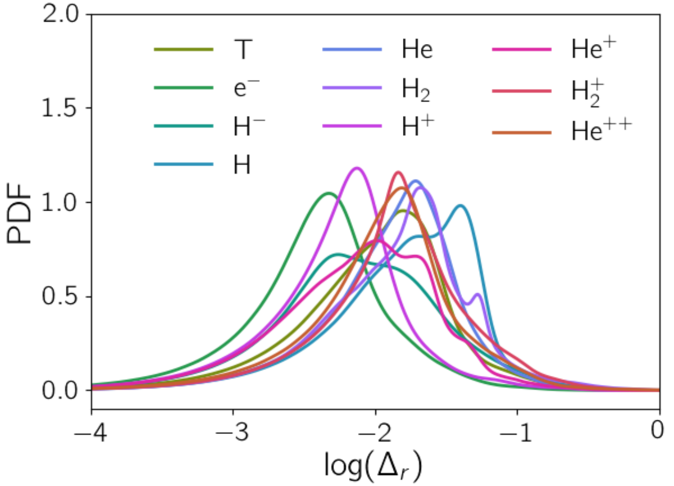


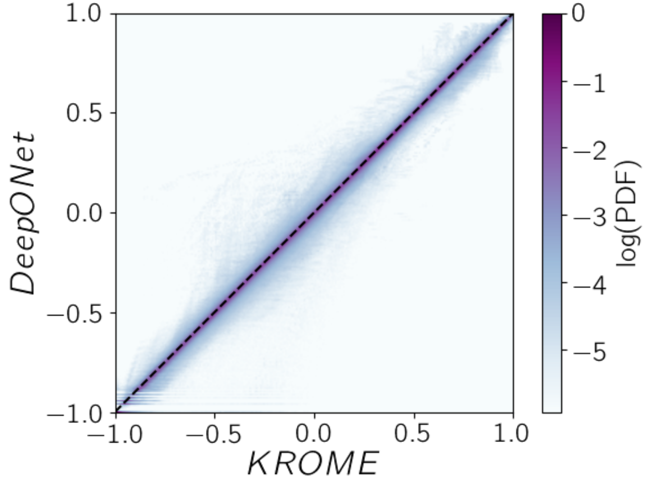
Results


2.0
1.5
1.0
0.5
0.0
PDF
−4
−3
−2
−1
0
log(Δ$_r$)
T
e$^-$
H$^-$
H
He
H$_2$
H$^+$
He$^+$
H$_2^+$
He^{++}


1.0
0.5
0.0
−0.5
−1.0
DeepONet
KROME
0
−1
−2
−3
−4
−5
log(PDF)

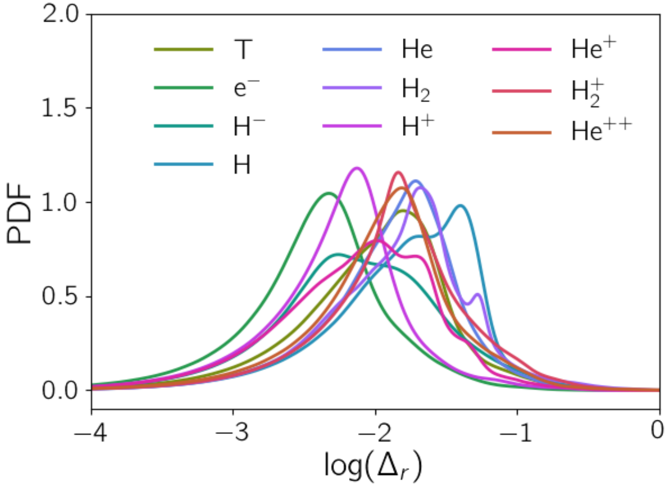
Results

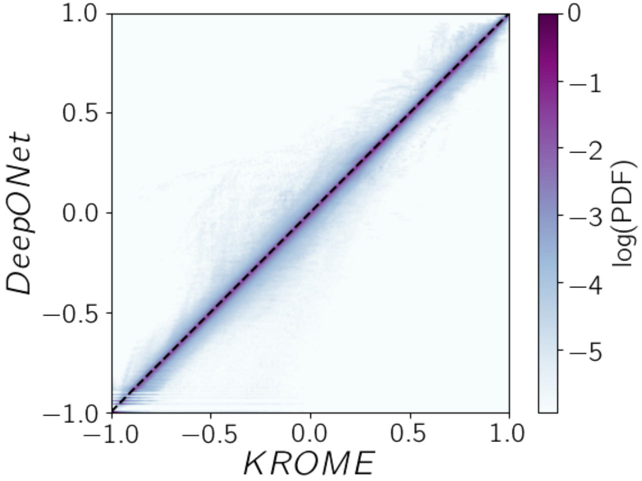




input	MRE	50%	75%	90%
T	0.0179	0.0114	0.0211	0.0353
e^-	0.0074	0.0044	0.0078	0.0151
H^-	0.0176	0.0076	0.0183	0.0398
H	0.0269	0.0190	0.0375	0.0547
He	0.0213	0.0150	0.0258	0.0422
H_2	0.0274	0.0167	0.0292	0.0548
H^+	0.0099	0.0060	0.0101	0.0178
He^+	0.0148	0.0082	0.0177	0.0314
H_2^+	0.0255	0.0146	0.0267	0.0552
He^{++}	0.0234	0.0125	0.0220	0.0436

Results

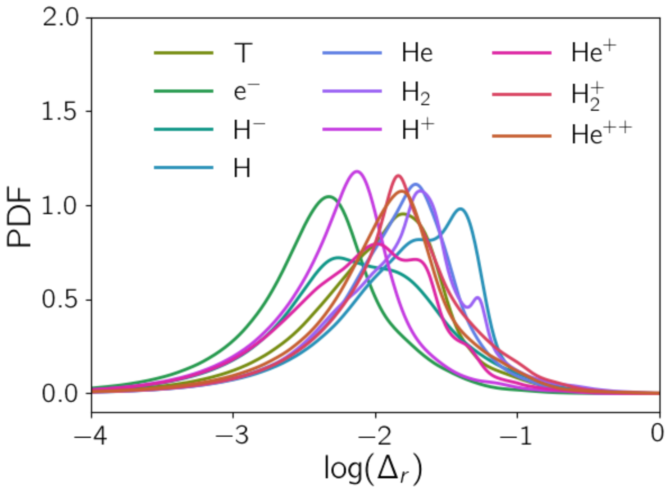


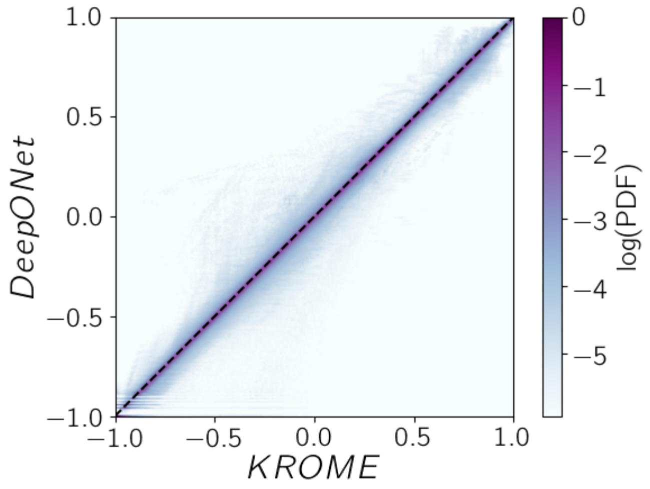


input	MRE	50%	75%	90%
T	0.0179	0.0114	0.0211	0.0353
e^-	0.0074	0.0044	0.0078	0.0151
H^-	0.0176	0.0076	0.0183	0.0398
H	0.0269	0.0190	0.0375	0.0547
He	0.0213	0.0150	0.0258	0.0422
H_2	0.0274	0.0167	0.0292	0.0548
H^+	0.0099	0.0060	0.0101	0.0178
He^+	0.0148	0.0082	0.0177	0.0314
H_2^+	0.0255	0.0146	0.0267	0.0552
He^{++}	0.0234	0.0125	0.0220	0.0436

- Training: ~40 GPUhrs

Results

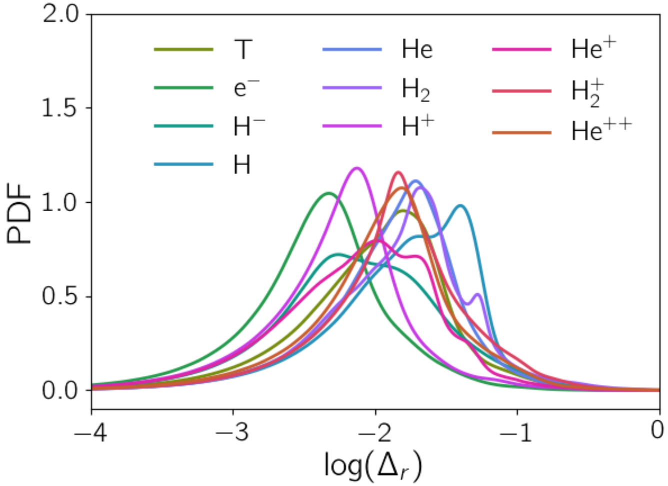


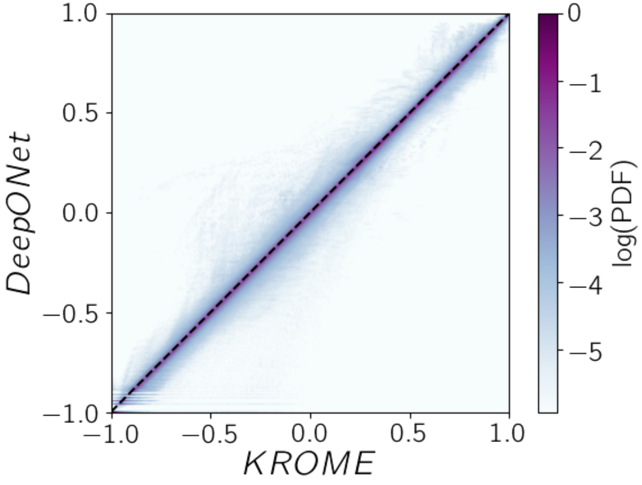


input	MRE	50%	75%	90%
T	0.0179	0.0114	0.0211	0.0353
e^-	0.0074	0.0044	0.0078	0.0151
H^-	0.0176	0.0076	0.0183	0.0398
H	0.0269	0.0190	0.0375	0.0547
He	0.0213	0.0150	0.0258	0.0422
H_2	0.0274	0.0167	0.0292	0.0548
H^+	0.0099	0.0060	0.0101	0.0178
He^+	0.0148	0.0082	0.0177	0.0314
H_2^+	0.0255	0.0146	0.0267	0.0552
He^{++}	0.0234	0.0125	0.0220	0.0436

- Training: ~40 GPUhrs
- Speed-up: 128X

Results

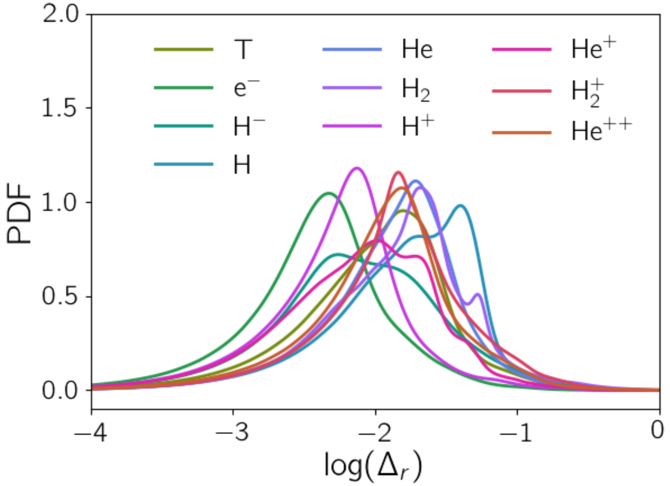


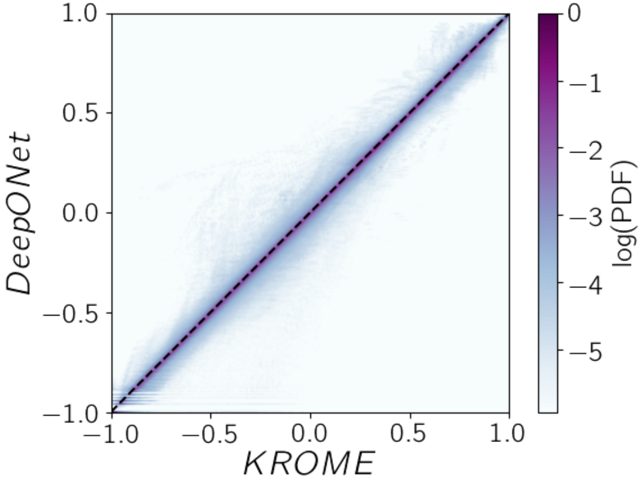


input	MRE	50%	75%	90%
T	0.0179	0.0114	0.0211	0.0353
e^-	0.0074	0.0044	0.0078	0.0151
H^-	0.0176	0.0076	0.0183	0.0398
H	0.0269	0.0190	0.0375	0.0547
He	0.0213	0.0150	0.0258	0.0422
H_2	0.0274	0.0167	0.0292	0.0548
H^+	0.0099	0.0060	0.0101	0.0178
He^+	0.0148	0.0082	0.0177	0.0314
H_2^+	0.0255	0.0146	0.0267	0.0552
He^{++}	0.0234	0.0125	0.0220	0.0436

- Training: ~40 GPUhrs
- Speed-up: 128X
- Accuracy: ~ 1%

Results

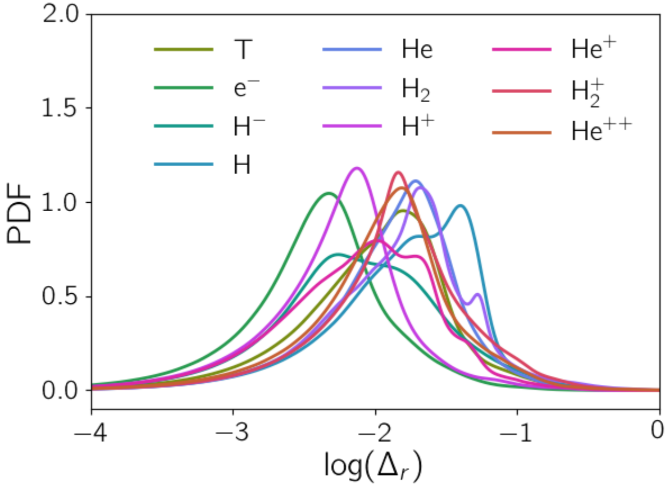


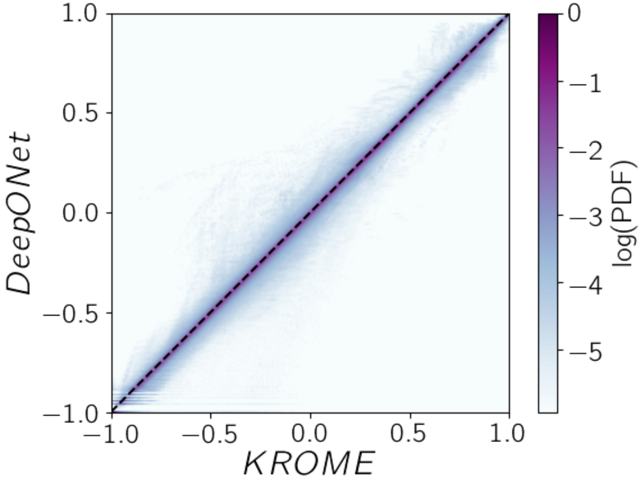


input	MRE	50%	75%	90%
T	0.0179	0.0114	0.0211	0.0353
e^-	0.0074	0.0044	0.0078	0.0151
H^-	0.0176	0.0076	0.0183	0.0398
H	0.0269	0.0190	0.0375	0.0547
He	0.0213	0.0150	0.0258	0.0422
H_2	0.0274	0.0167	0.0292	0.0548
H^+	0.0099	0.0060	0.0101	0.0178
He^+	0.0148	0.0082	0.0177	0.0314
H_2^+	0.0255	0.0146	0.0267	0.0552
He^{++}	0.0234	0.0125	0.0220	0.0436

- Training: ~40 GPUhrs
- Speed-up: 128X
- Accuracy: ~ 1%
- N ouliers: ~ 1/1.000.000

Results



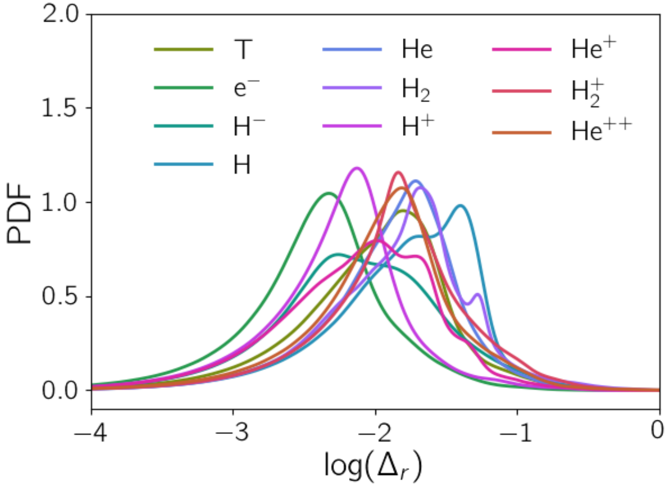


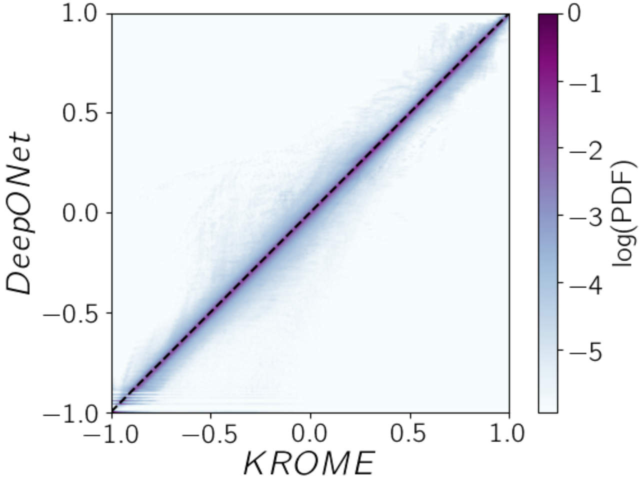
input	MRE	50%	75%	90%
T	0.0179	0.0114	0.0211	0.0353
e^-	0.0074	0.0044	0.0078	0.0151
H^-	0.0176	0.0076	0.0183	0.0398
H	0.0269	0.0190	0.0375	0.0547
He	0.0213	0.0150	0.0258	0.0422
H_2	0.0274	0.0167	0.0292	0.0548
H^+	0.0099	0.0060	0.0101	0.0178
He^+	0.0148	0.0082	0.0177	0.0314
H_2^+	0.0255	0.0146	0.0267	0.0552
He^{++}	0.0234	0.0125	0.0220	0.0436

- Training: ~40 GPUhrs
- Speed-up: 128X
- Accuracy: ~ 1%
- N ouliers: ~ 1/1.000.000

FIRST TIME WITH RADIATION FIELD COUPLING!!!!!!

Results



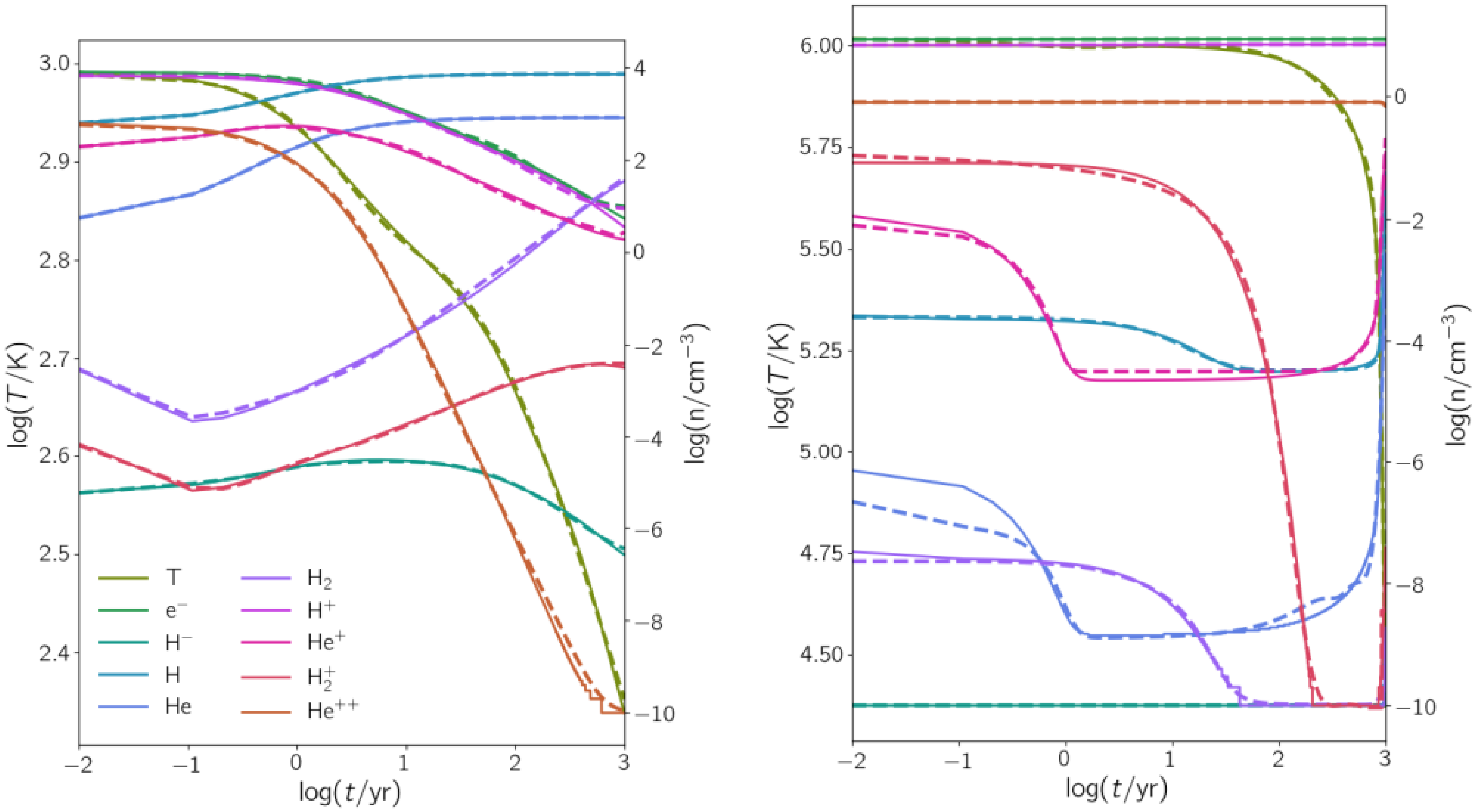


input	MRE	50%	75%	90%
T	0.0179	0.0114	0.0211	0.0353
e^-	0.0074	0.0044	0.0078	0.0151
H^-	0.0176	0.0076	0.0183	0.0398
H	0.0269	0.0190	0.0375	0.0547
He	0.0213	0.0150	0.0258	0.0422
H_2	0.0274	0.0167	0.0292	0.0548
H^+	0.0099	0.0060	0.0101	0.0178
He^+	0.0148	0.0082	0.0177	0.0314
H_2^+	0.0255	0.0146	0.0267	0.0552
He^{++}	0.0234	0.0125	0.0220	0.0436

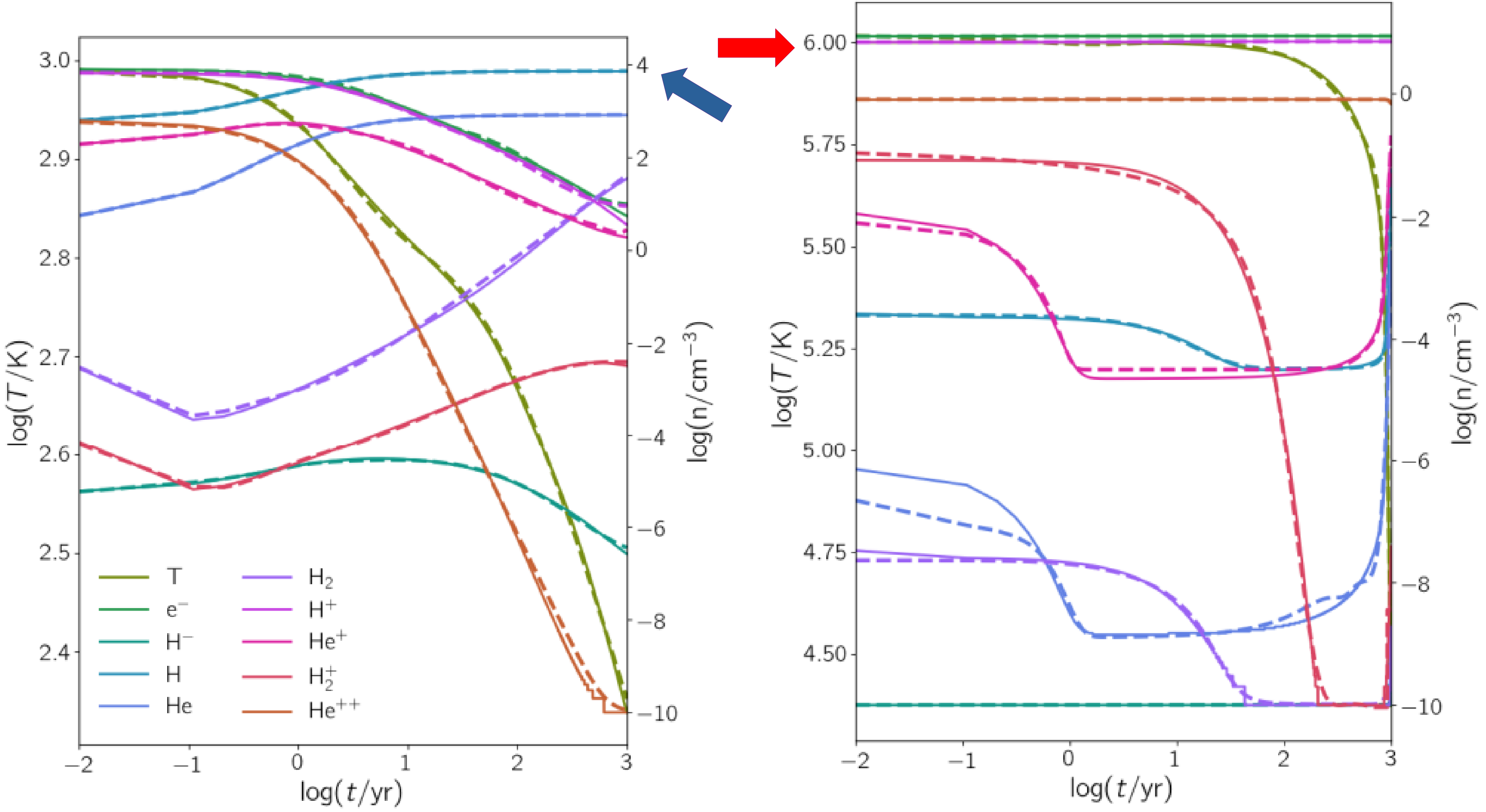
- Training: ~40 GPUhrs
- Speed-up: 128X
- Accuracy: ~ 1%
- N ouliers: ~ 1/1.000.000

FIRST TIME WITH RADIATION FIELD COUPLING!!!!!!

Let's look at the solutions

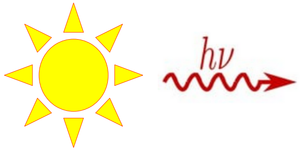


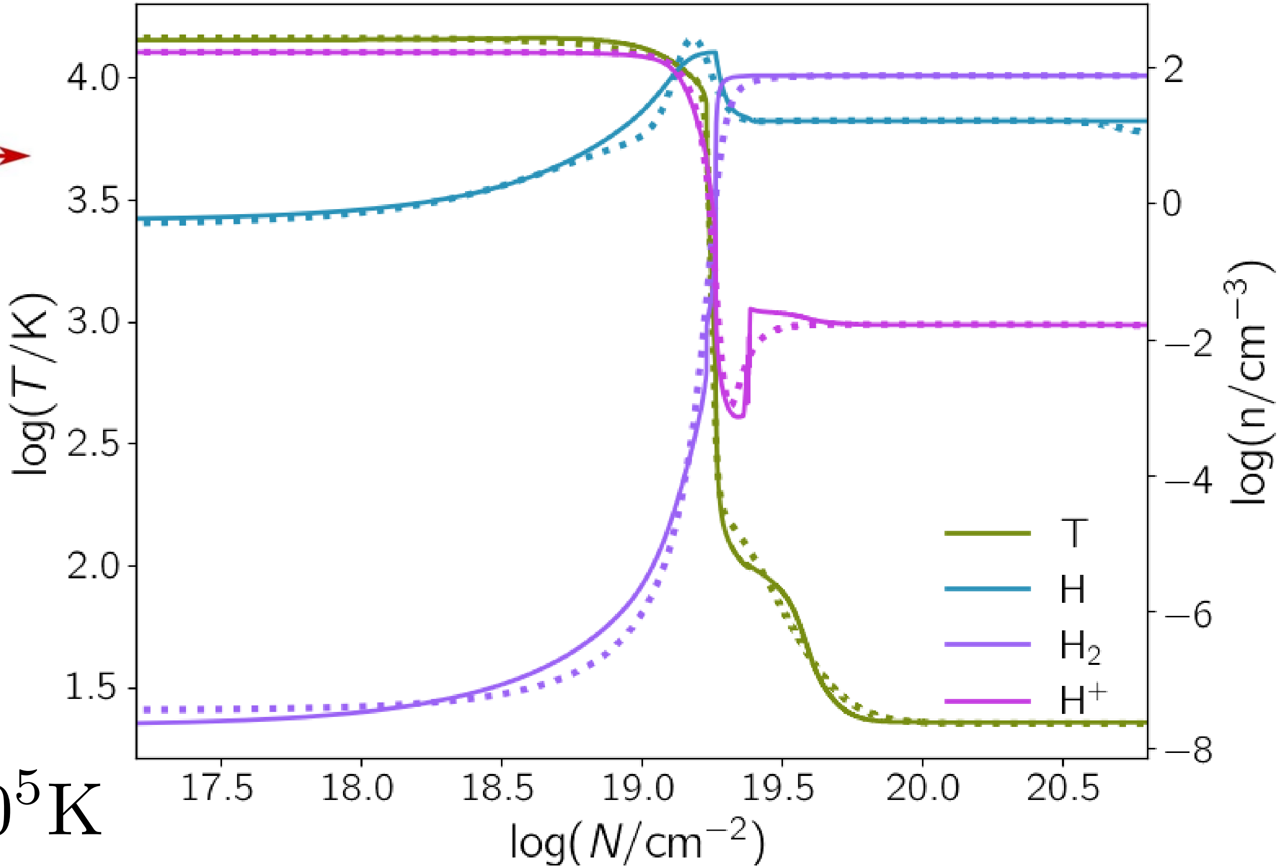
Let's look at the solutions



PDR test

$G/G_0 = 10$





$T_{bb} = 10^5$K

$R_c \simeq 16$pc

$n_{tot} = 10^2 \text{cm}^{-3}$

$t_{end} = 1$Kyr

$T_{in} = 10^3$K

$n_{ion}/n \simeq 10^{-4}$

$n_{H_2}/n \simeq 0.4$

Some considerations

Some considerations

Coupling with simulations:

Some considerations

Coupling with simulations:

Outliers (rel_err>50%) appear every 10^6 inferences

Some considerations

Coupling with simulations:

Outliers (rel_err>50%) appear every 10^6 inferences

↓

Not clear how to manage the error propagation

Some considerations

Coupling with simulations:

Outliers (rel_err>50%) appear every 10^6 inferences

↓

Not clear how to manage the error propagation

More chemical species:

Some considerations

Coupling with simulations:

Outliers (rel_err>50%) appear every 10^6 inferences

↓

Not clear how to manage the error propagation

More chemical species:

↓

Dataset became huge!

Some considerations

Coupling with simulations:

Outliers (rel_err>50%) appear every 10^6 inferences

↓

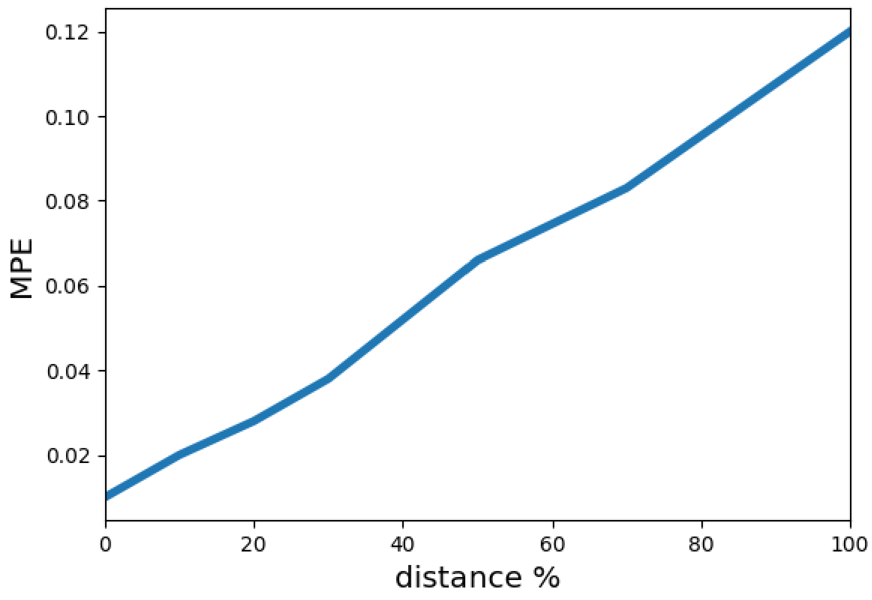
Not clear how to manage the error propagation

More chemical species:

↓

Dataset became huge!

Generalization: operator are better than functions.
However…



Summary

- ISM (photo-thermo) chemistry represents a current challenge in astrophysical simulations
- We developed a Neural Operator based emulator to replace procedural solvers
- Accuracy around 1%
- For the first time we include radiation field

Thanks!