ICLR 2021 Workshop
Deep Learning
for Simulation

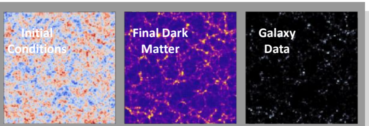
CosmicRIM : Reconstructing Early Universe by Combining Differentiable Simulations with Recurrent Inference Machines

Chirag Modi[a], François Lanusse[b], Uroš Seljak[c], David N. Spergel[a], Laurence Perreault-Levasseur[a,d]

(a) Center for Computational Astrophysics, Flatiron Institute (b) AIM, CEA, CNRS, Université Paris-Saclay (c) Department of Physics, University of California, Berkeley (d) Department of Physics, Université de Montréal, Montréal, Canada

FLATIRON INSTITUTE
Center for Computational Astrophysics

Motivation



Aim : To reconstruct the **Gaussian** initial conditions[1] at the beginning of the Universe from a sparse galaxy sample data

Challenge :

1. solve an inverse problem in high-dimensions (~10^8) to reconstruct density at all points in space
2. forward model is complex, *non-linear* and expensive N-body simulations

Approach : Combine *differentiable* cosmological N-body simulations, like **FlowPM**[2] with *learnt optimization* methods to learn inference schemes[3] and tackle these challenges efficiently.

Experiments

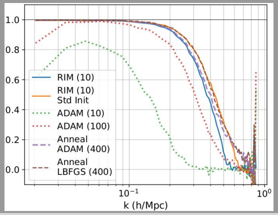
Forward model : 64^3 particles, 2 step PM sim. with 2^{nd} order bias model

CosmicRIM outperforms traditional optimization, even with *physically* motivated annealing at **40x** less cost

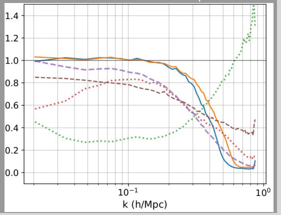
Informed initial position[5] $\mathbf{x}_0$ can improve reconstruction (see orange)

Comparing 2-point statistics between reconstructed & true initial conditions

Cross Correlation $r_c(k) = \frac{P_{ab}(k)}{\sqrt{P_a(k) \times P_b(k)}}$



Transfer Function $t_f(k) = \sqrt{\frac{P_a(k)}{P_b(k)}}$



CosmicRIM

Setup : To obtain the MAP estimate given the likelihood p of the observed data $\mathbf{y}$ for the initial conditions $\mathbf{x}$ and Gaussian prior p_θ on the initial conditions with known power spectrum

$$\max_{\mathbf{x}} \ln p(\mathbf{x}, \mathbf{y}) = \max_{\mathbf{x}} [\ln p(\mathbf{y}|\mathbf{x}) + \ln p_\theta(\mathbf{x})]$$

RIM (Recurrent inference machine)[4] : learn optimization by training a recurrent neural network (h_ϕ, g_ϕ) to learn the update equations at time step t, given state s_t

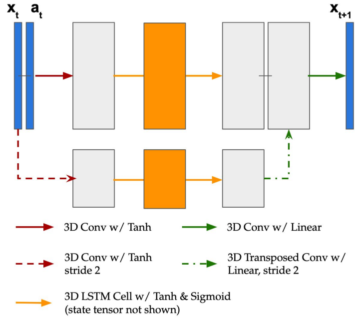
$$\mathbf{s}_{t+1} = \mathbf{s}_t + h_\phi(\mathbf{s}_t, \nabla_x \ln p(\mathbf{x}, \mathbf{y}), \mathbf{x}_t),$$
$$\mathbf{x}_{t+1} = \mathbf{x}_t + g_\phi(\mathbf{x}_t, \nabla_x \ln p(\mathbf{x}, \mathbf{y}), \mathbf{s}_{t+1})$$

Training loss function for a 10 step RIM

$$\mathcal{L} = \sum_{t=0}^{10} (\mathbf{x}_t(\phi) - \mathbf{x}_{true})^2$$

Architectural modifications

1. **Replace gradients** $\nabla_x \ln\ p(x,y)$ in the update functions (h,g) with update predicted by **ADAM** algorithm $\mathbf{a}_t$
2. **U-Net** architecture with different LSTM cells in **parallel** to learn updates for large and small scales separately

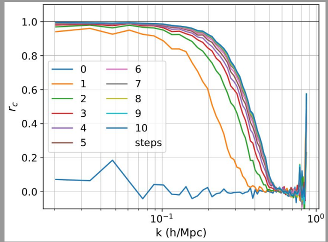


What CosmicRIM learns?

Physically motivated annealing : large scales have higher SNR & linear dynamics

This led [6] to develop annealing scheme that smooth the likelihood on small scales & reconstruct the large scales first.

RIM implicitly learns a similar path to reconstruction



(Local) Minima : i.e ADAM/L-BFGS starting from RIM output don't improve results

Outlook

- **First application** combining complex, non-linear albeit differentiable forward models with learnt optimization schemes for high-dimensional inference problems.
- **40x speed up** & better reconstruction of the initial conditions of the Universe with CosmicRIM over other approaches
- **Implicitly learning** the optimization path otherwise strictly imposed with physically motivated annealing schemes.
- **Challenge** : High memory requirements for training can be a bottleneck but potential solutions can be found in splitting the optimization path and sequential learning

References

[1] U. Seljak, G. Aslanyan, Y. Feng & C. Modi. arXiv:1706.06645
[2] M Andrychowicz et.al. arXiv:1606.0447
[3] C. Modi, F. Lanusse, U. Seljak. (*FlowPM*) arXiv:2010.11847
[4] P. Putzky & M. Welling. (*RIM*) arXiv:1706.04008
[5] D Eisenstein et.al.(Std. Reconstruction) ApJ,664(2):675–679
[6] C. Modi, Y. Feng, U. Seljak. arXiv: 1805.02247