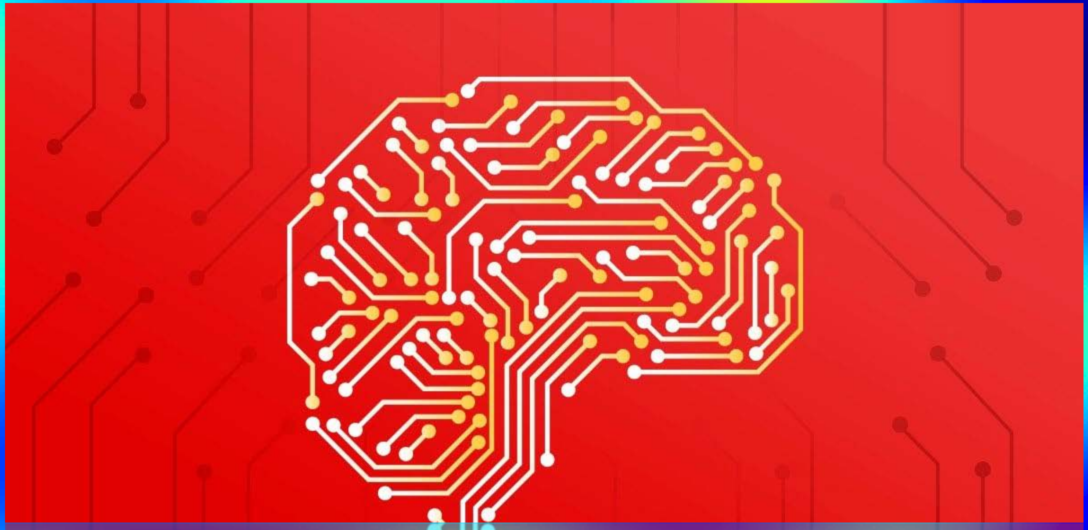


中山大學
SUN YAT-SEN UNIVERSITY

Map Reconstruction of Radio Observations with Conditional Invertible Neural Networks

Le Zhang (SYSU)

21 cm Cosmology Workshop 2023
Shenyang, 2023.07.17

In collaboration with S.F. Zuo and H.L. Zhang
arxiv:2306.09217; Research in Astronomy and Astrophysics

CONTENTS

1 The issue of sky map-making

2 Review the existing methods

3 Conditional invertible neural network(CINN)

4 Results of a FAST-like survey from cINN

5 Summary and Outlook

CONTENTS

1 The issue of sky map-making

2 Review the existing methods

3 Conditional invertible neural network(CINN)

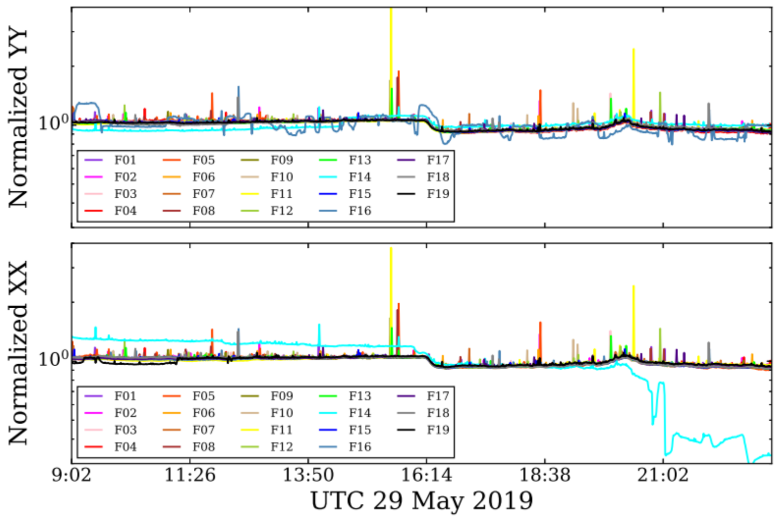
4 Results of a FAST-like survey from cINN

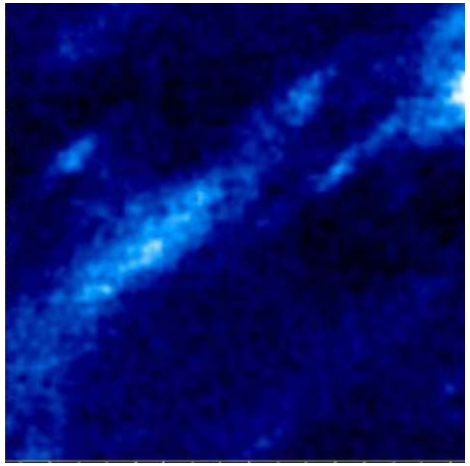
5 Summary and Outlook

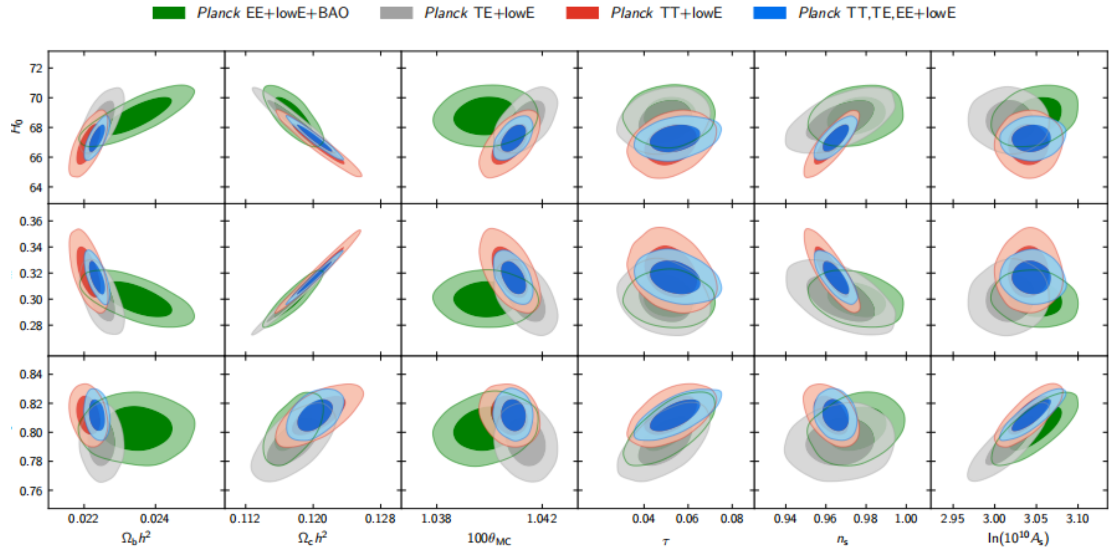
1.1 Necessity of map-making

- Map-making is a crucial step in radio observations, bridging the gap between the collected TODs and scientific analysis
- important to produce pixelized maps from TOD, with as much accuracy as possible

time-ordered data(TOD) → Sky Map → Data Analysis





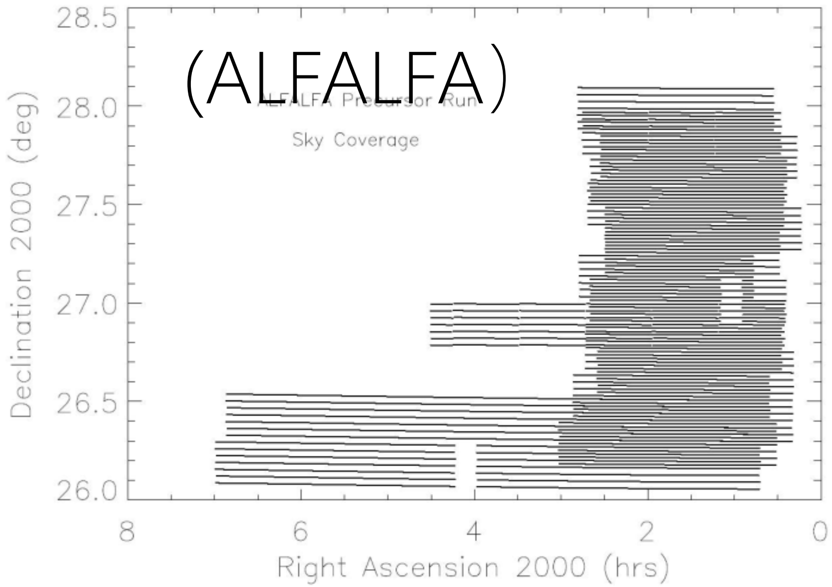


Hu 2021

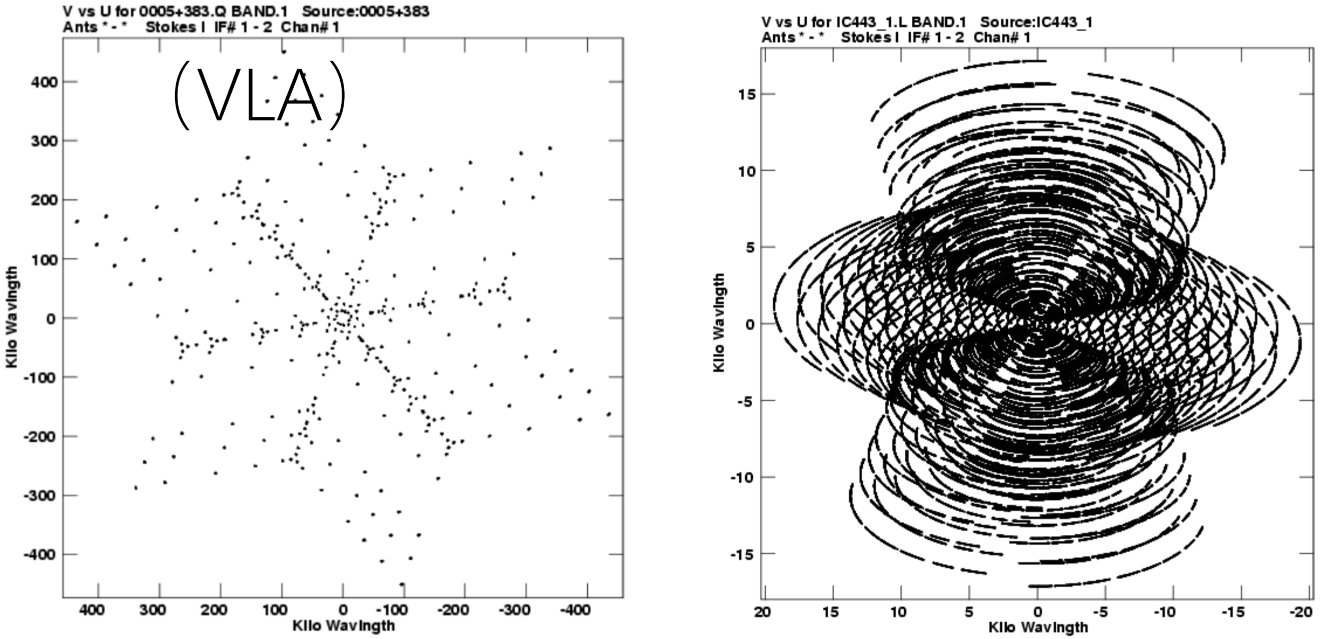
Planck 2018

1.2 Issues in map-making

- the TOD's sampling is uneven and irregular in practice
- **Map-making usually is an ill-posed inverse problem**—observational effects such as scan strategies, noise, complex geometry of the field and data excision like RFI flagging
- unbiased estimate with minimal variance is a big challenge



Scanning pattern of a single-aperture telescope



uv-coverage from interferometric observation

CONTENTS

1 The issue of sky map reconstruction

2 Review the existing methods

3 Conditional invertible neural network(CINN)

4 Results of a FAST-like survey from cINN

5 Summary and Outlook

2.1 Traditional estimator

- **The concept of Map-making is to apply a constructed linear operator W on data**
- $\vec{d} = A\vec{m} + \vec{n}$:

$$\hat{m} = W\vec{d}$$

MAP-MAKING METHODS

Number	Method	Specification
1	Generalized *COBE*	$\mathbf{W} = [\mathbf{A}^t\mathbf{M}\mathbf{A}]^{-1}\mathbf{A}^t\mathbf{M}$
2	Bin averaging	$\mathbf{W} = [\mathbf{A}^t\mathbf{A}]^{-1}\mathbf{A}^t$
3	*COBE*	$\mathbf{W} = [\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}]^{-1}\mathbf{A}^t\mathbf{N}^{-1}$
4	Wiener 1	$\mathbf{W} = \mathbf{S}\mathbf{A}^t[\mathbf{A}\mathbf{S}\mathbf{A}^t + \mathbf{N}]^{-1}$
5	Wiener 2	$\mathbf{W} = [\mathbf{S}^{-1} + \mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}]^{-1}\mathbf{A}^t\mathbf{N}^{-1}$
6	Saskatoon	$\mathbf{W} = [\eta\mathbf{S}^{-1} + \mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}]^{-1}\mathbf{A}^t\mathbf{N}^{-1}$
7	TE96	$\mathbf{W} = \Lambda\mathbf{S}\mathbf{A}^t[\mathbf{A}\mathbf{S}\mathbf{A}^t + \mathbf{N}]^{-1}$, $(\mathbf{W}\mathbf{A})_{ii} = 1$
8	TE97	$\mathbf{W} = \Lambda[\eta\mathbf{S}^{-1} + \mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}]^{-1}\mathbf{A}^t\mathbf{N}^{-1}$, $(\mathbf{W}\mathbf{A})_{ii} = 1$
9	Maximum probability	Nonlinear method if non-Gaussian
10	Maximum entropy	Nonlinear method

Tegmark, M (1997)

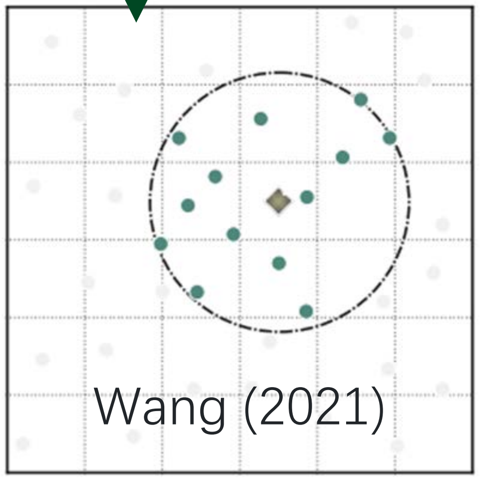
- The inverse of the matrix m may not exist, and therefore W may not exist
- Even if W exists, the computational complexity of the inverse operation can reach $O(N^3)$; computationally intractable if $N_p \sim 10^6$
- RFI and ⋯ subtractions in data preprocessing, leading to a degeneracy in the estimated map
- An accurate estimate of the noise is required, i.e., the covariance matrix N of the noise needs to be known
- The error estimate for each pixel is rather difficult

2.2 Traditional estimator: gridding approach

- The gridding: based on a direct weighted interpolation of TOD

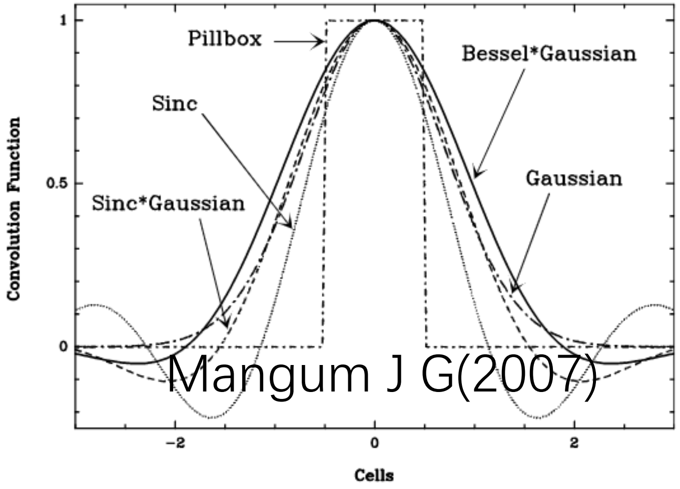
Convolution of TOD in a certain region

$$R_{i,j}\left(\alpha_{i,j},\delta_{i,j}\right)=\frac{1}{W_{i,j}}\sum_{n}R_n\left(\alpha_n,\delta_n\right)\mathrm{w}\left(\alpha_{i,j},\delta_{i,j};\alpha_n,\delta_n\right)$$



Wang (2021)

w: a convolution kernel

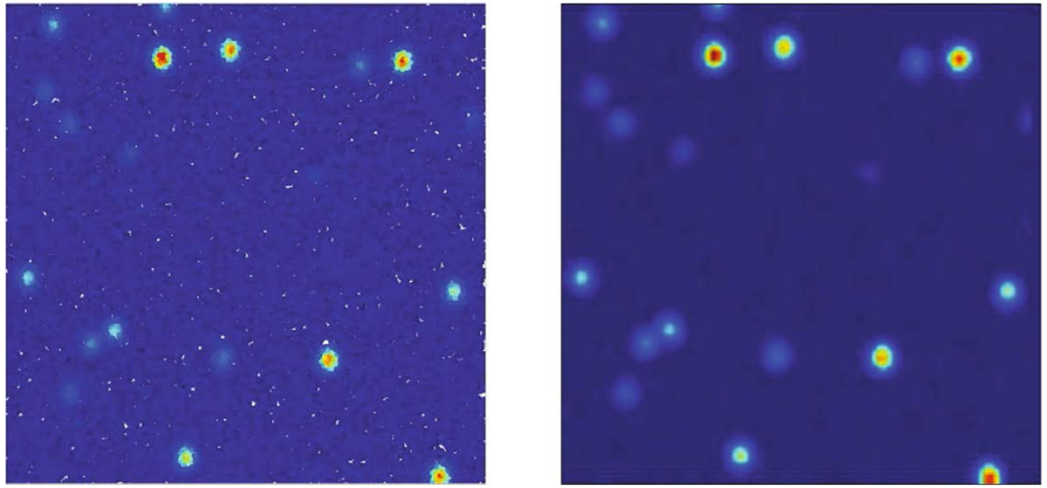


Mangum J G(2007)

Drawback:

- pixel-level-error estimate is difficult
- Gridding will further reduce the map resolution

$$\sigma_{gridded}=\sqrt{\sigma_{kernel}^2+\sigma_{data}^2}$$



Before and after gridding Luo(2018)

CONTENTS

1 The issue of sky map reconstruction

2 Review the existing methods

3 Conditional invertible neural network(CINN)

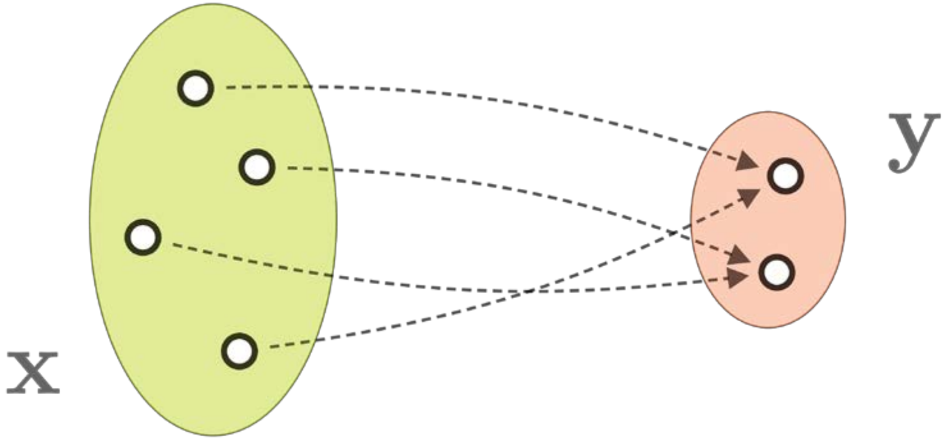
4 Results of a FAST-like survey from cINN

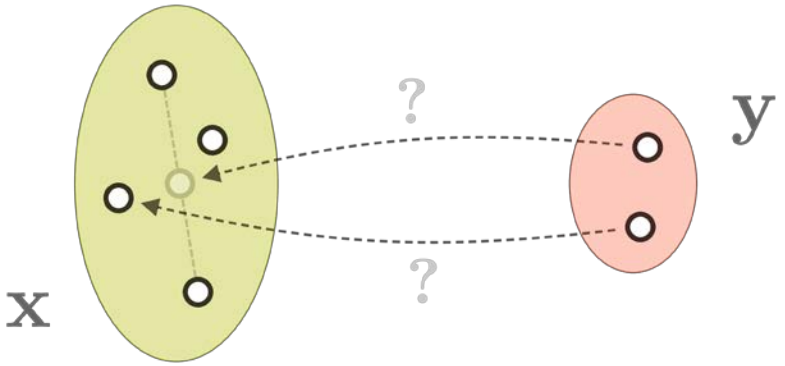
5 Summary and Outlook

3.1 Inverse problem

- Inverse problems are common in scientific research, where observations are utilized to infer physical parameters (i.e., map here)
- The forward modeling process (map → TOD) is well understood
- TOD → map is of an inverse problem

Ambiguous Inverse Problems (ill-posed)

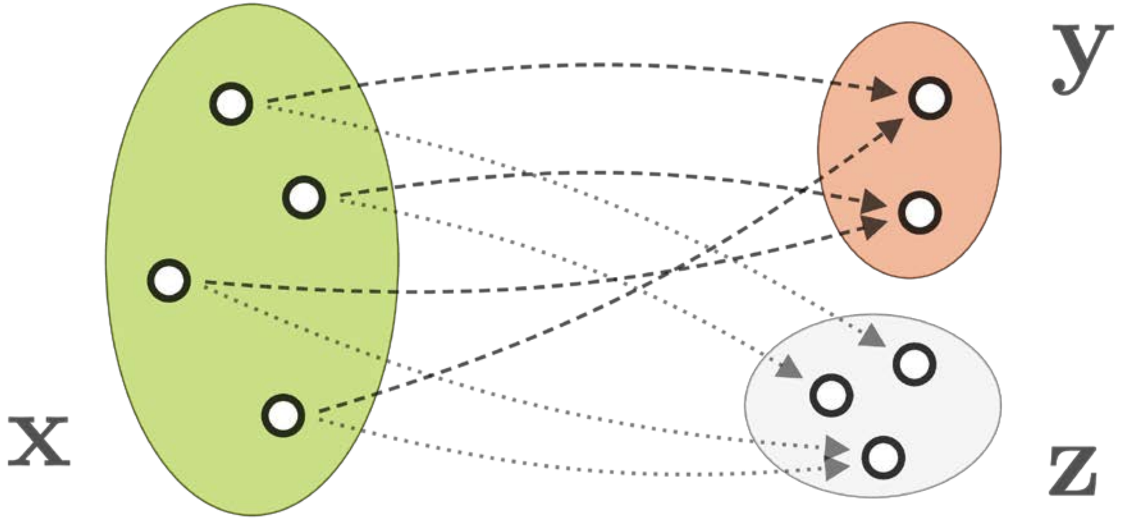




https://hci.iwr.uniheidelberg.de/vislearn/inverse-problems-invertible-neural-networks/

- one may apply statistical inference techniques to express the ambiguities in form of conditional probabilities $p(x|y)$
- classical Bayesian methods become very expensive even for moderate real-world problems

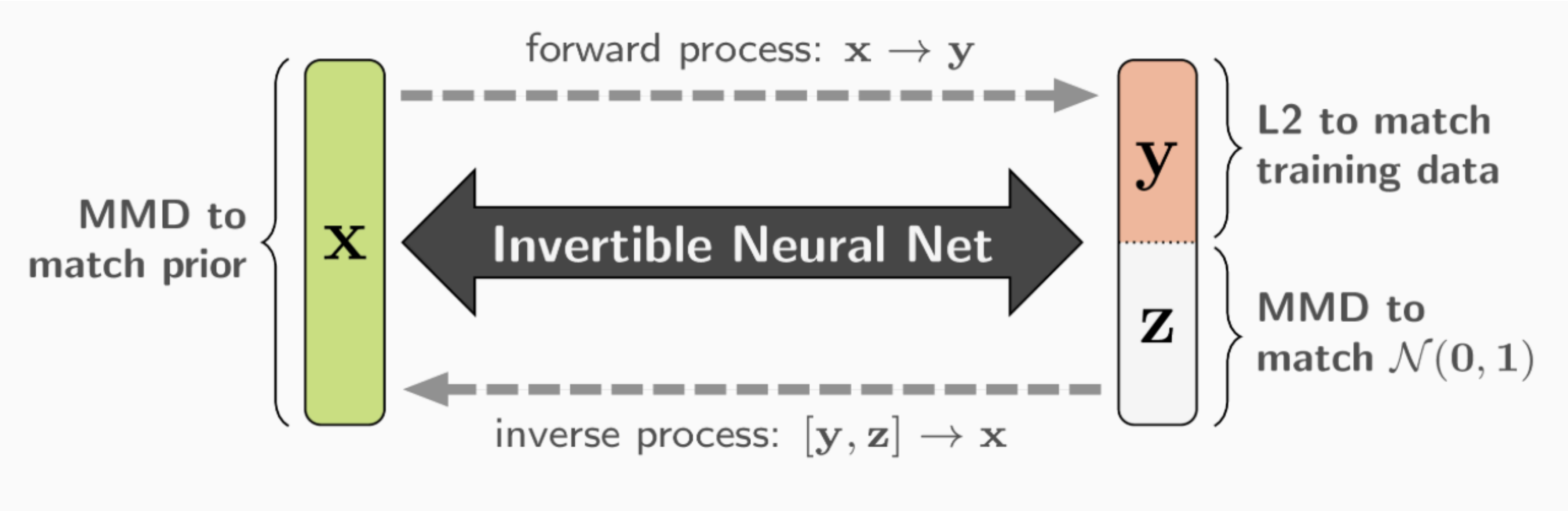
3.1 Resolving the ambiguity



Intuitively, the ambiguity of inverse mapping is transformed into $p(z)$.

- Bijective mapping: introducing additional latent variables, z, to preserve the information that would otherwise be lost during the forward process
- The mapping $x \leftrightarrow [y, z]$ becomes a one-to-one correspondence (well posed)
- In other words, $p(x|y)$ has been reparametrized into a deterministic function $x = f(y, z)$ with adding variable z

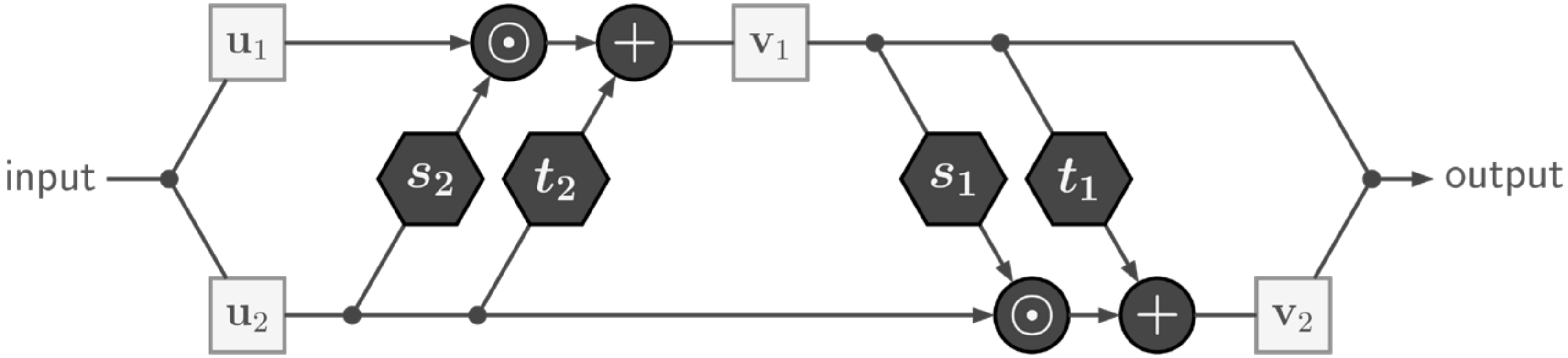
3.2 Training process



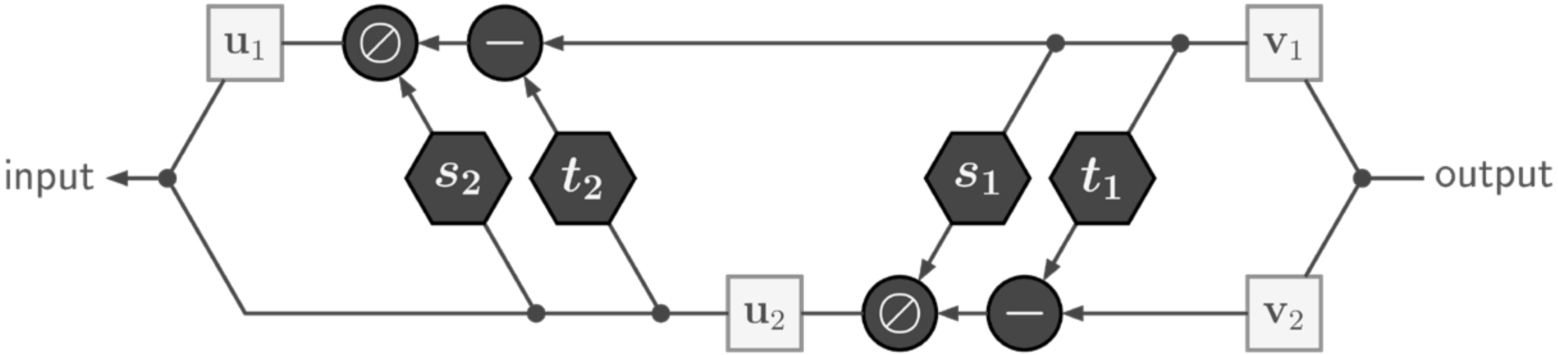
- train INN to solve the well-posed forward process $x \to y$ in a supervised manner, instead of the ill-posed inverse process
- the latent variables z to be independent of y, and to follow an easy-to-sample-from distribution, like $\mathcal{N}(0,1)$.
- use L2 (to match the data)+ a Maximum Mean Discrepancy (MMD) loss (to match the normal distribution) for training INN

3.3 Basic building block of INN

the *affine coupling layer* popularized by the **Real NVP** model.



Inverse of the whole affine coupling layer: recover $[\mathbf{u}_1, \mathbf{u}_2]$ from $[\mathbf{v}_1, \mathbf{v}_2]$



To construct deep invertible networks, just chain these layers like **ResNet blocks.**

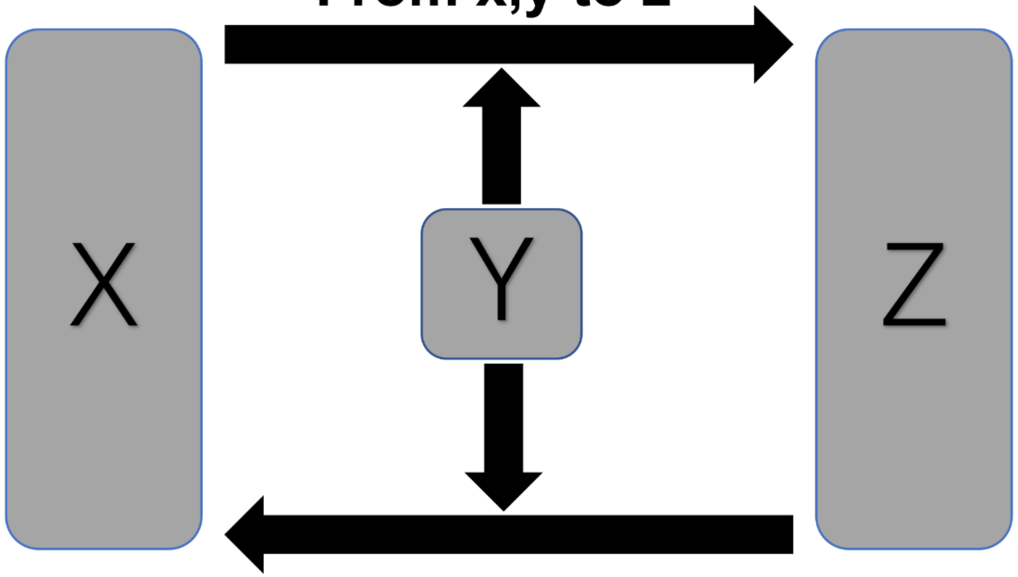
3.4 Conditional Invertible Neural Networks

- cINN: a modification of INN, enabling simpler training

- under the condition (y), the distribution of x is obtained by sampling z from Gaussian distribution; naturally obtain the error of x
- changed the role of y on the input side, while introducing the advantage of conditional input

$$z = f(x|y, \theta)$$

Forward Modeling:
From x,y to z



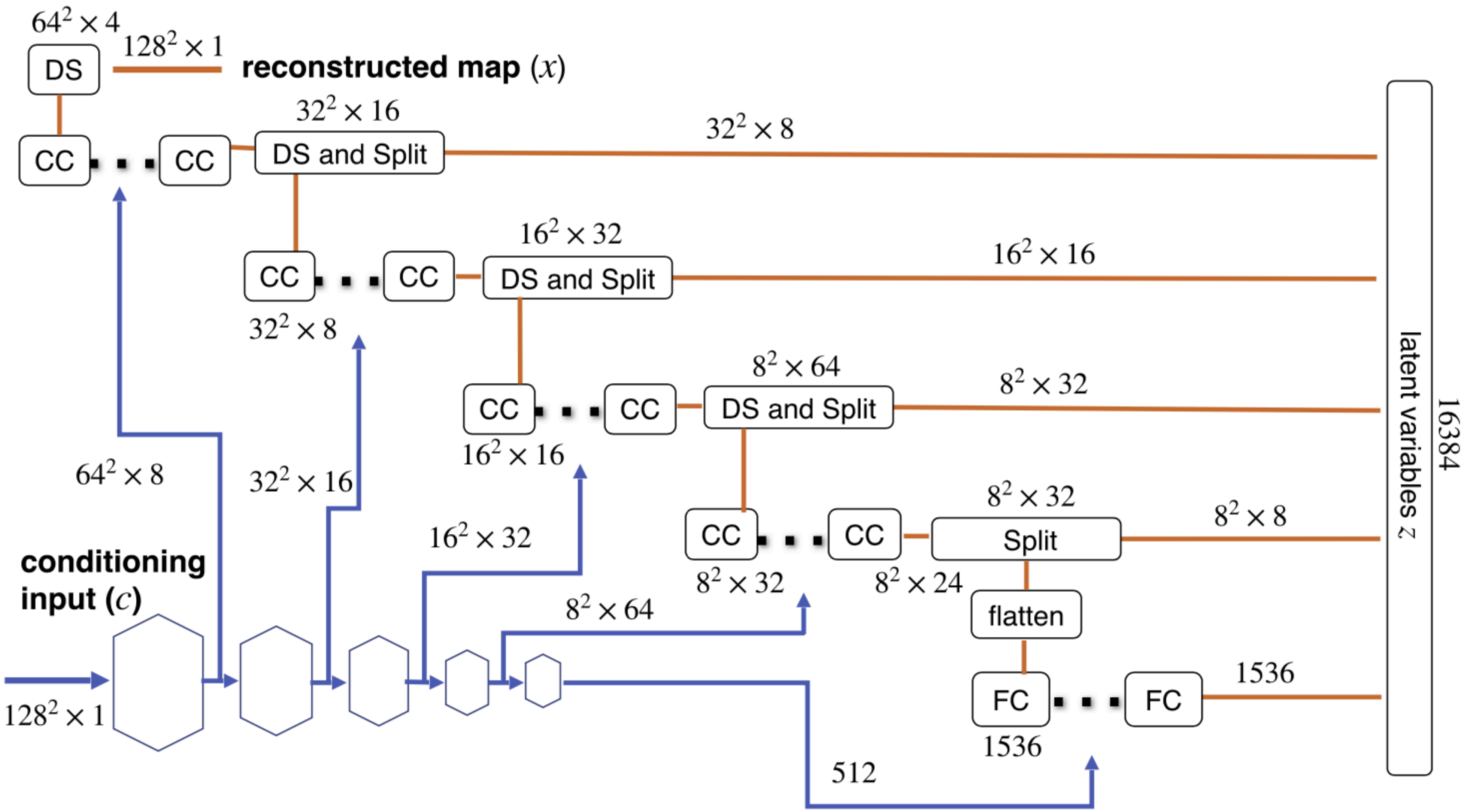
Backward Sampling:
From z,y to x

$$x = f^{-1}(z|y, \theta)$$

Ardizzone (2019)

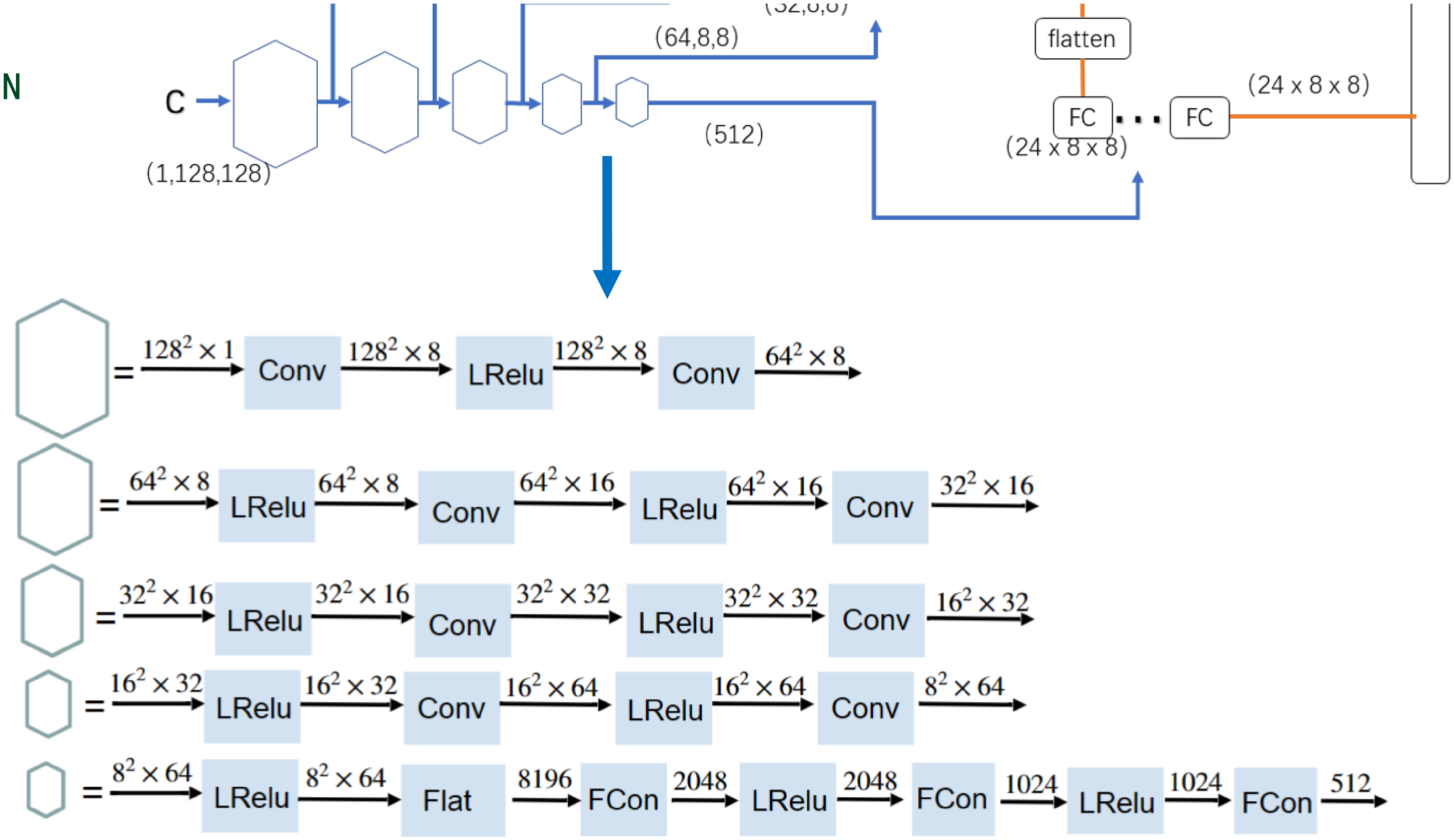
3.5 Structure of invertible neural network (I)

The dedicated network structure was found to reconstruct the sky map effectively



3.6 Structure of invertible neural network (II)

CINN



CONTENTS

1 The issue of sky map reconstruction

2 Review the existing methods

3 Conditional invertible neural network(CINN)

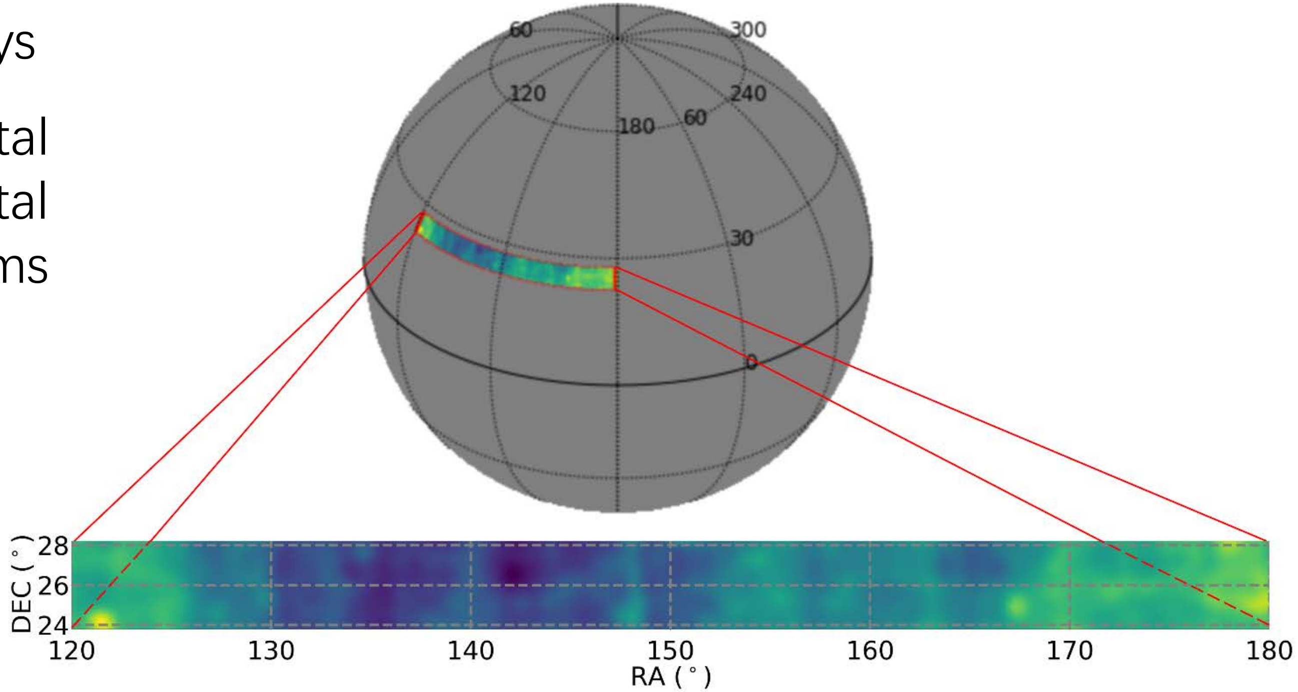
4 Results of a FAST-like survey from cINN

5 Summary and Outlook

4.1 TOD generation and training samples

- **survey simulation**: based on the FAST configuration
- **drift scan:** consisting of a 19-beam receiver in the frequency range of 1100–1120 MHz, 20 channels
- **coverage:** a sky area of over 300 square degrees
- **observation time:** 2020.5.4–5.28, totaling 25 days
- an integration time of 1 s per beam and a total observation time of 14400 s/day, the total number of time samples for all 19 beams amounts to 25×14400×20×19 ~ 10^8
- **thermal noise** (T_{sys} in 0-25 K) added

- Data pre-possessing: due to the TODs' varying length and large data size, preprocessing is needed before feeding into the network
- TODs are gridded onto a 2D flat-sky maps, each having an area of $4.3° \times 4.3°$ and a resolution of 128 × 128



4.2 Evaluation Metrics

- the mean square error (MSE) : mean distance between two maps

$$\mathrm{MSE}\left(x_{\mathrm{true}}, x_{\mathrm{rec}}\right) = \frac{1}{N} \sum_{k=1}^{N} \left(x_{\mathrm{true}}^{k} - x_{\mathrm{rec}}^{k}\right)^{2}$$

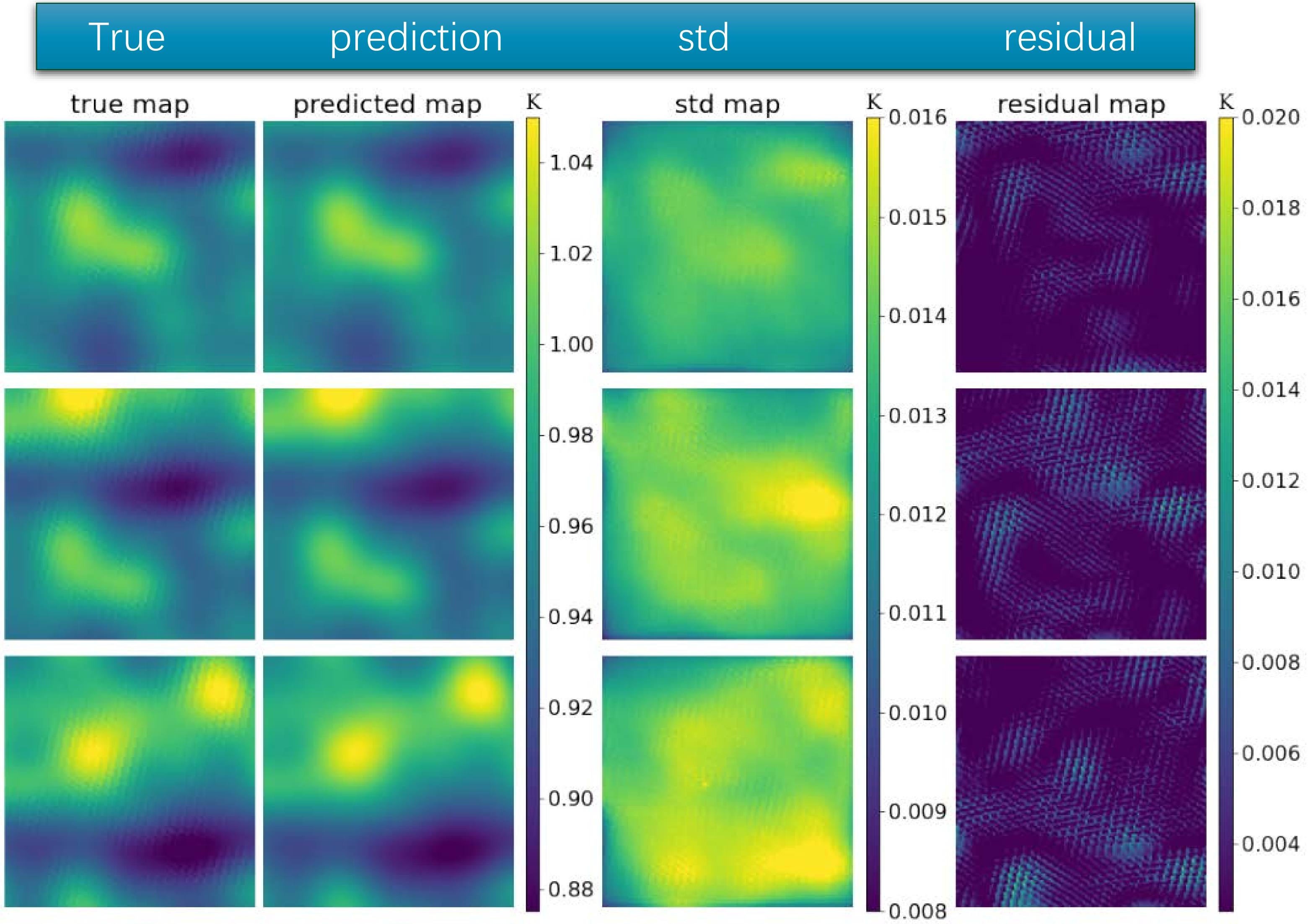
- the Peak Signal-to-Noise Ratio (PSNR): evaluating the reconstruction quality in dB

$$\mathrm{PSNR}(x_{\mathrm{true}}, x_{\mathrm{rec}}) = 10 \log_{10} \left(\frac{L^2}{\mathrm{MSE}\left(x_{\mathrm{true}}, x_{\mathrm{rec}}\right)} \right)$$

- the structural similarity index measure (SSIM)

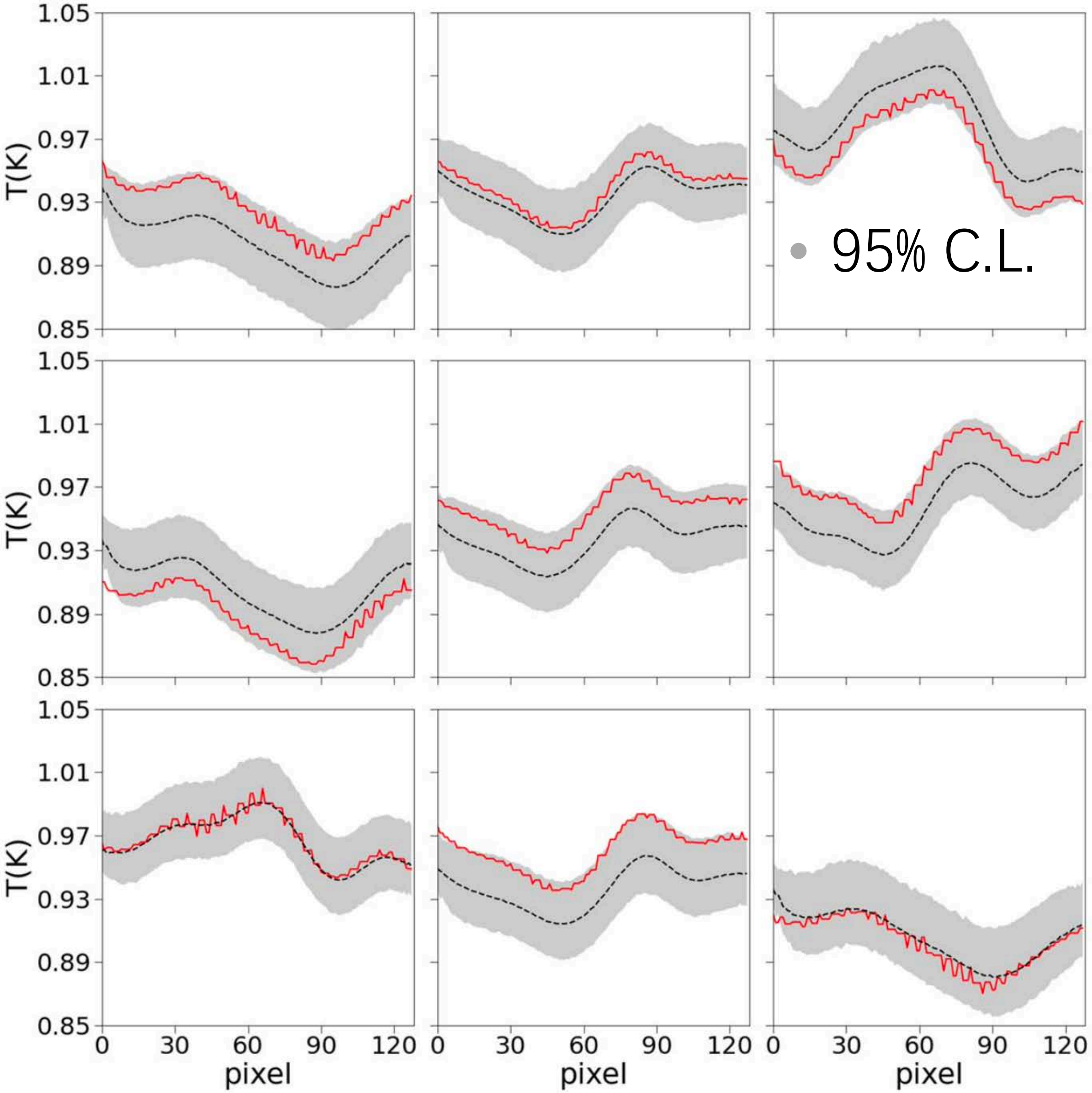
$$\mathrm{SSIM}(x_{\mathrm{true}}, x_{\mathrm{rec}}) = \frac{(2\mu_i\mu_j + C_1)(2\Sigma_{ij} + C_2)}{(\mu_i^2 + \mu_j^2 + C_1)(\sigma_i^2 + \sigma_j^2 + C_2)}$$

4.3 Results of map reconstruction



- $P(x|y)$ estimated by 200 reconstructed maps through drawing latent variables z from Gaussian
- std and residuals are about 0.01 K，1% level of true map
- good reconstruction quality

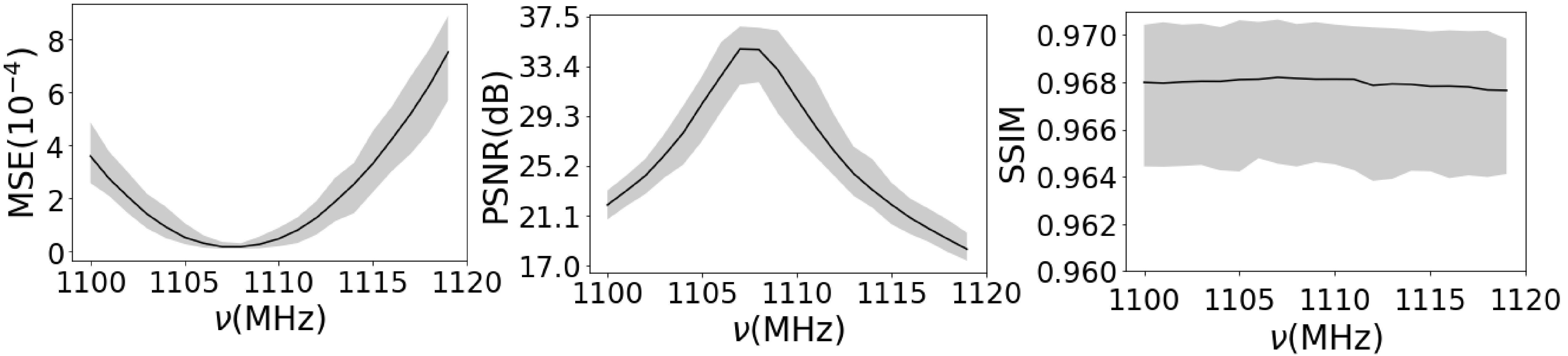
4.3 Results of map reconstruction



- randomly selected a row of map
- the mean vs. the true
- pixel error: 200 realizations of the latent variables z from normal distribution

Reconstruction mean and errors for each pixel can be precisely quantified.

4.4 MSE, PSNR, SSIM vs. frequency



over all test samples:

Performance	MSE ($\times 10^{-4}$)	SSIM	PSNR
	2.6 ± 5.0	0.95 ± 0.003	25.37 ± 4.21

- good performance in all three metrics across frequency

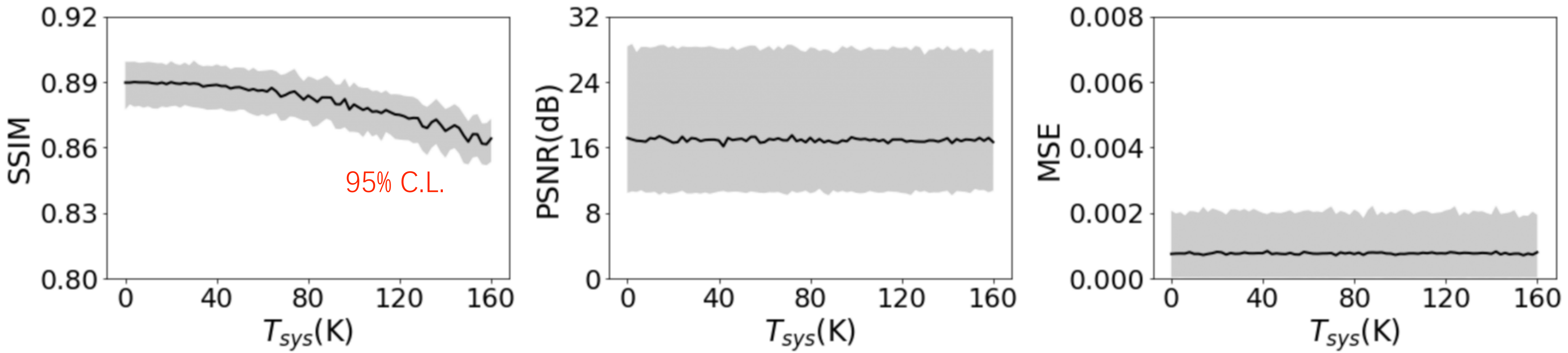
5. Summary and Outlook

- First introduce cINN to solve the map-making problem
- Good performance in reconstruction has been achieved
- cINN framework has the potential to tackle ill-posed problems in astronomy, like radio interferometric observations, where imaging can be particularly challenging due to sparse uv coverage

Backup

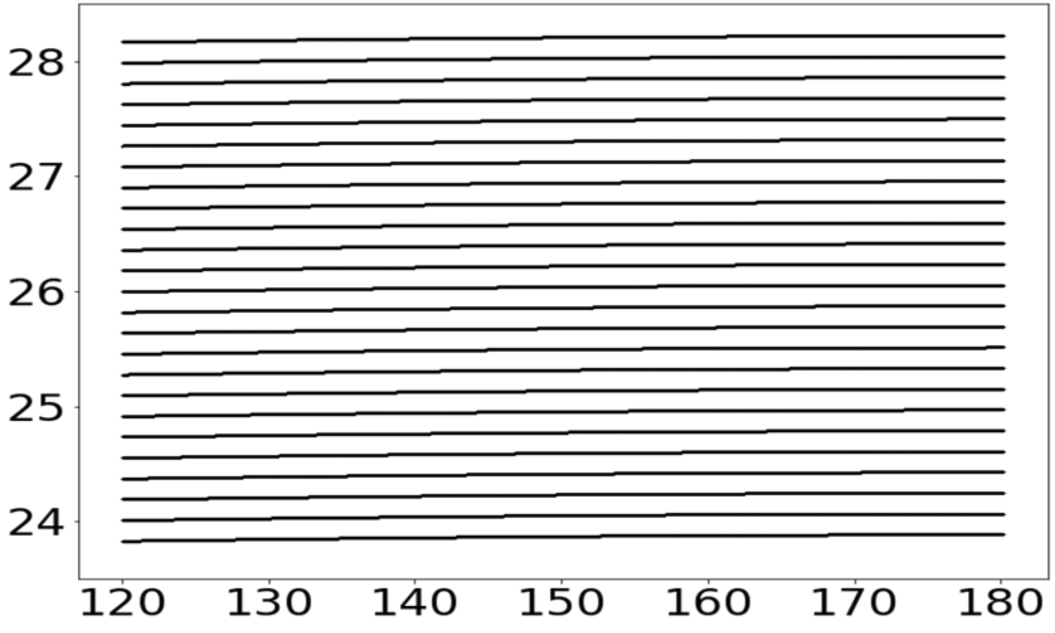
4.5 Performance against noise level

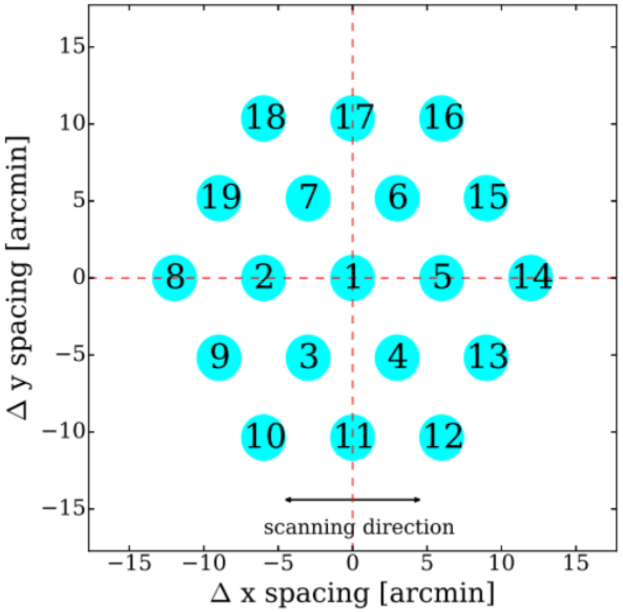
- T_{sys} in training： 0-25K; T_{sys} in generalization tests： 0-160K
- As the noise level increases, the SSIM value decreases from 0.89 to 0.85
- MSE and PSNR remain essentially constant with increasing noise level



2.1 Modeling of sky map reconstruction

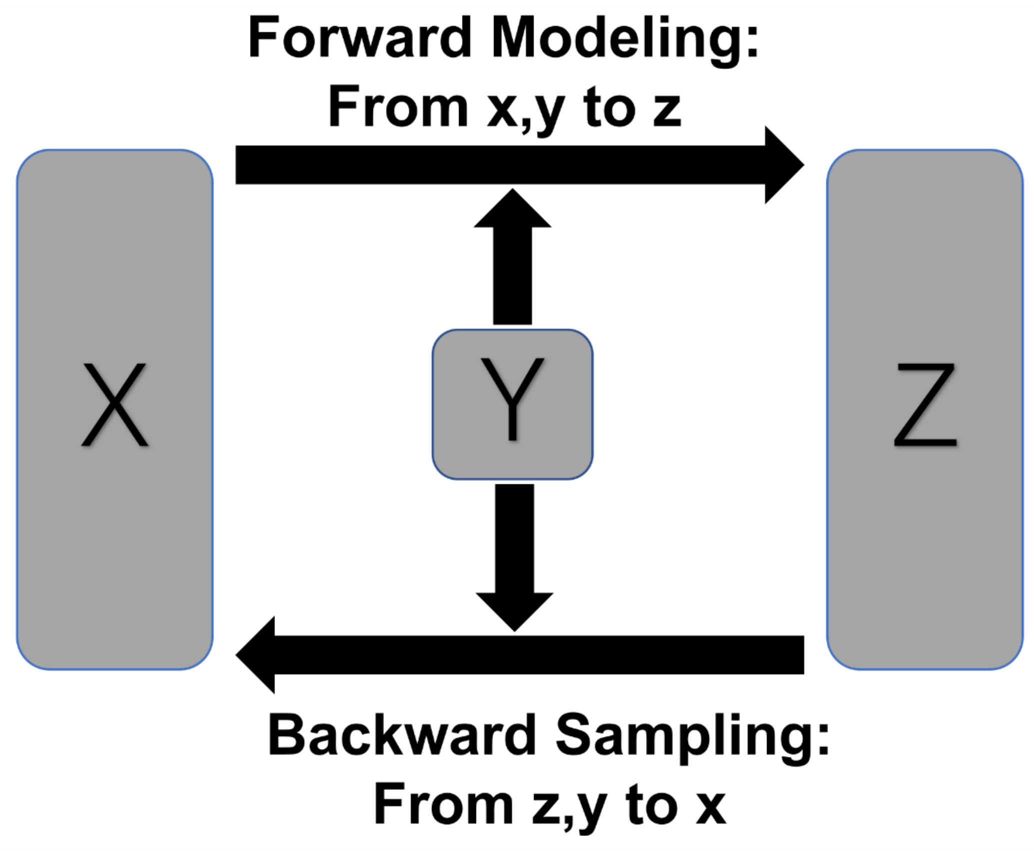
- **Observation equation:** $\vec{d} = A\vec{m} + \vec{n}$
  - m denotes sky map, d is for TOD, n is for noise
  - Dimension: d has N_t, m has N_m, A has $N_t \times N_m$, n has N_t
  - The Eq. can be extended to the multi-frequency and multi-beam case



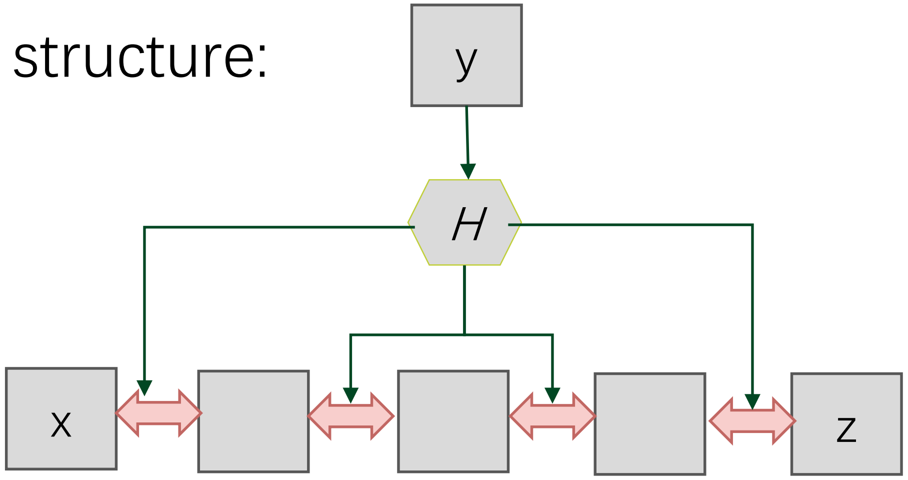


Hu (2021)

Main structure



Fully connected structure:



Multi-level structure: