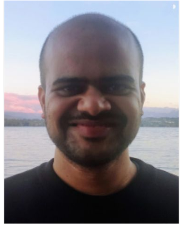
stochastic information processing

UNIVERSITÉ DE GENÈVE

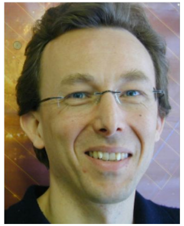
Radio-astronomical Image Reconstruction with Conditional Denoising Diffusion Models

M. Drozdova, V. Kinakh, O. Bait, O. Taran, E. Lastufka, M. Dessauges-Zavadsky, T. Holotyak, D. Schaerer, and S. Voloshynovskiy








contact: mariia.drozdova@unige.ch

Agenda

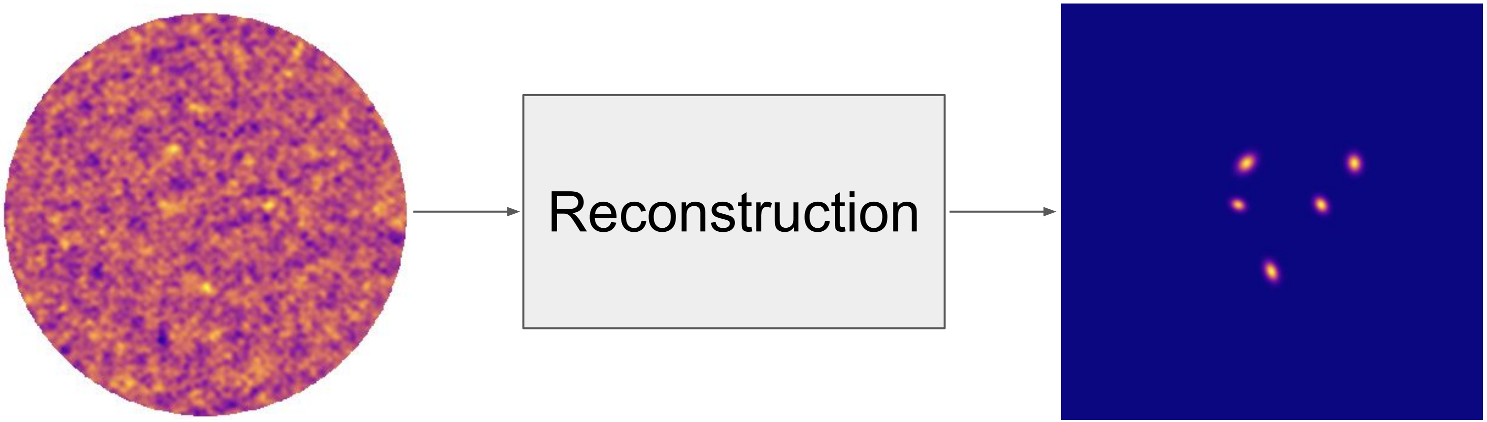
1. Problem formulation
2. Generative processes
3. Dataset
4. Pre-processing
5. Conditional DDPMs in our problem
6. Evaluation metrics
7. Results (aggregation strategies, comparative results)
8. Flux estimation
9. Conclusion

Problem Formulation

Problem: Localization and characterization of radio sources

Solution: Reconstruction of the sky model from dirty images

Possible tools: CLEAN, bayesian methods, sparse optimization algorithms, deep learning*



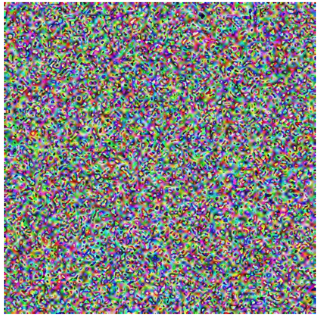
Dirty image

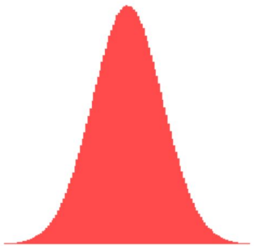
Sky model

*pix2pix, SRGAN, CycleGAN, turbo

Generative processes

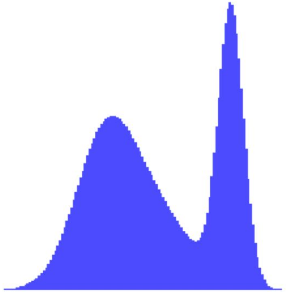
Model noise





one mode

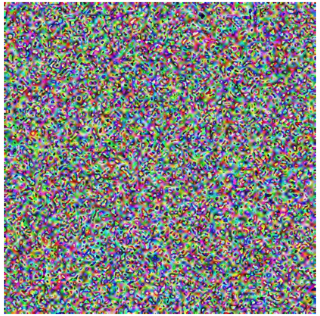
Image

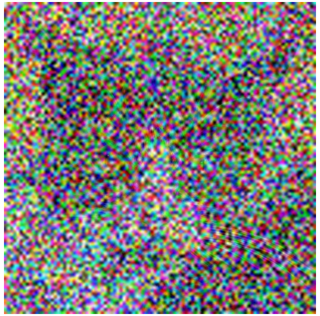


multiple modes

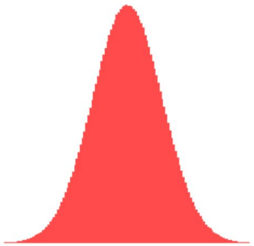
Generative processes

Model noise

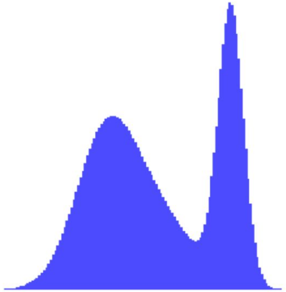




Image

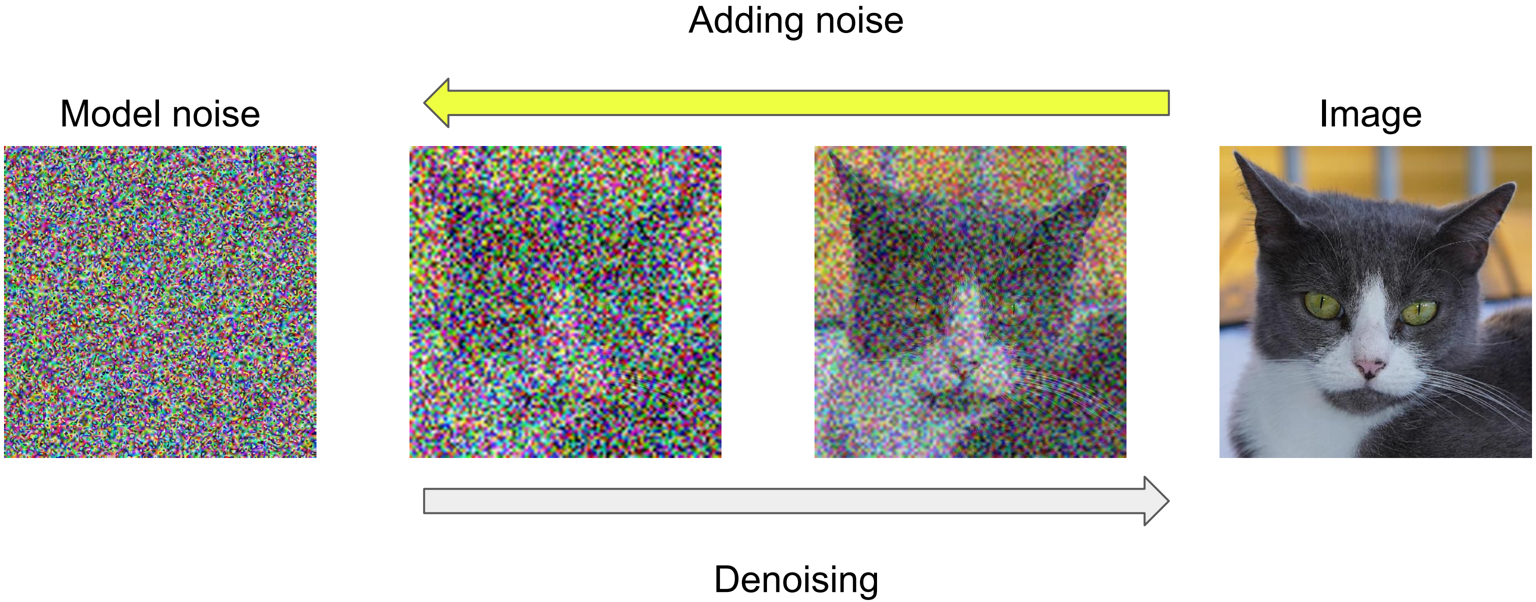


one mode

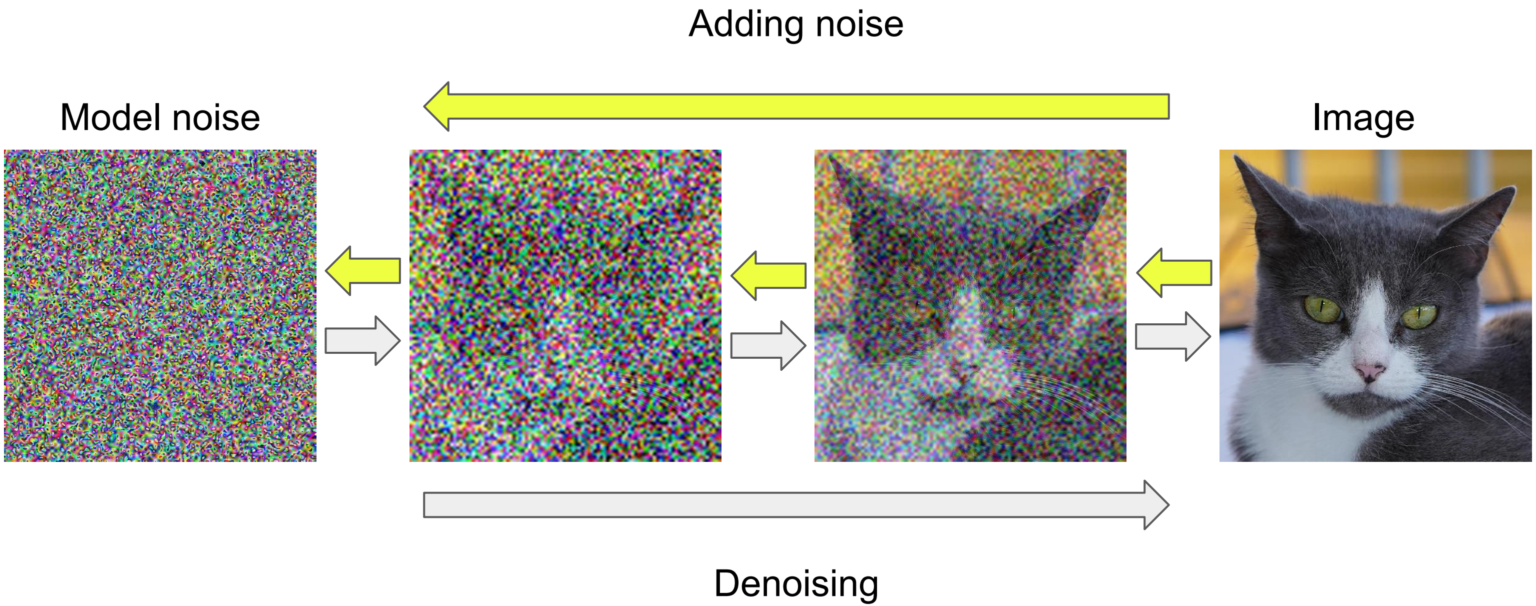


multiple modes

Generative processes

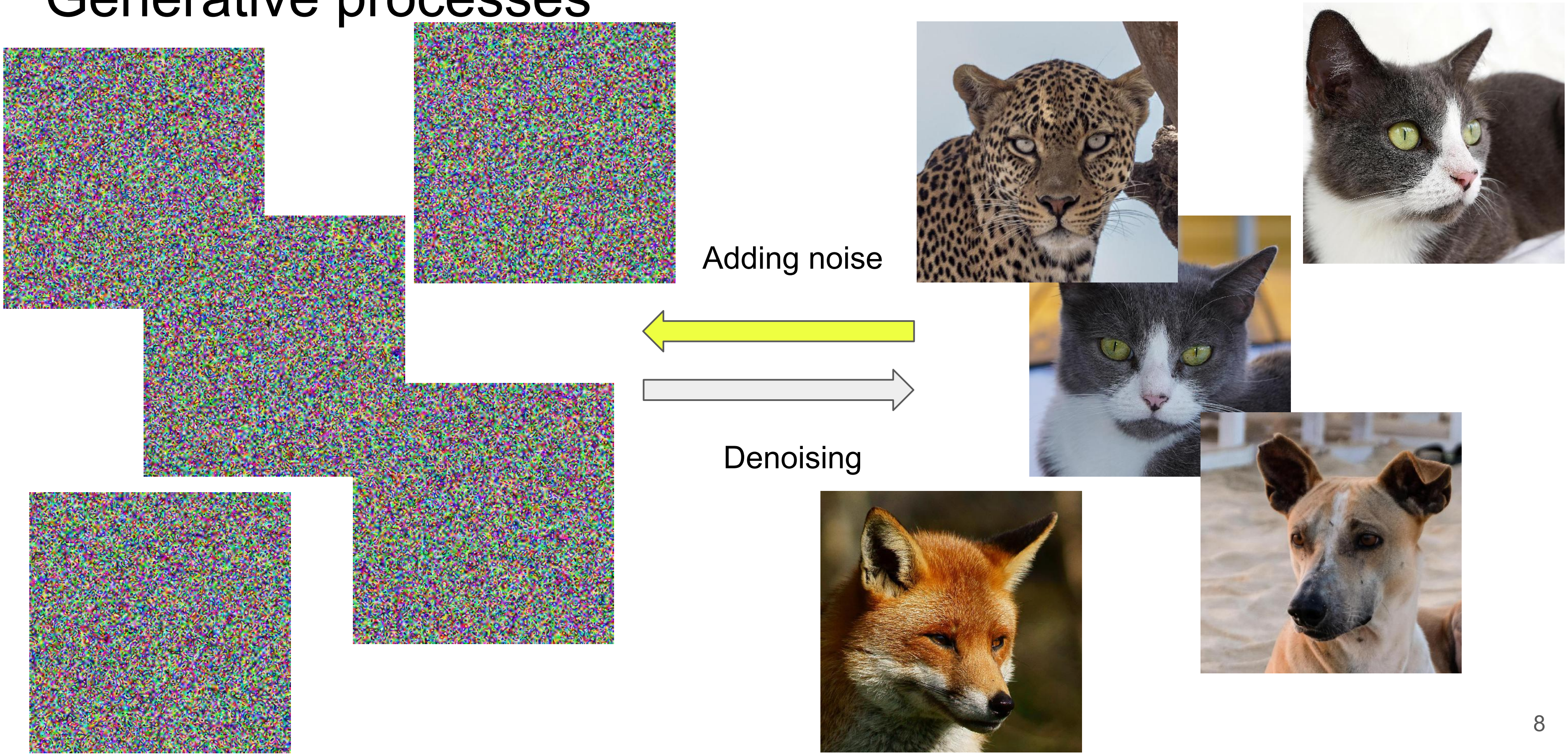


Generative processes

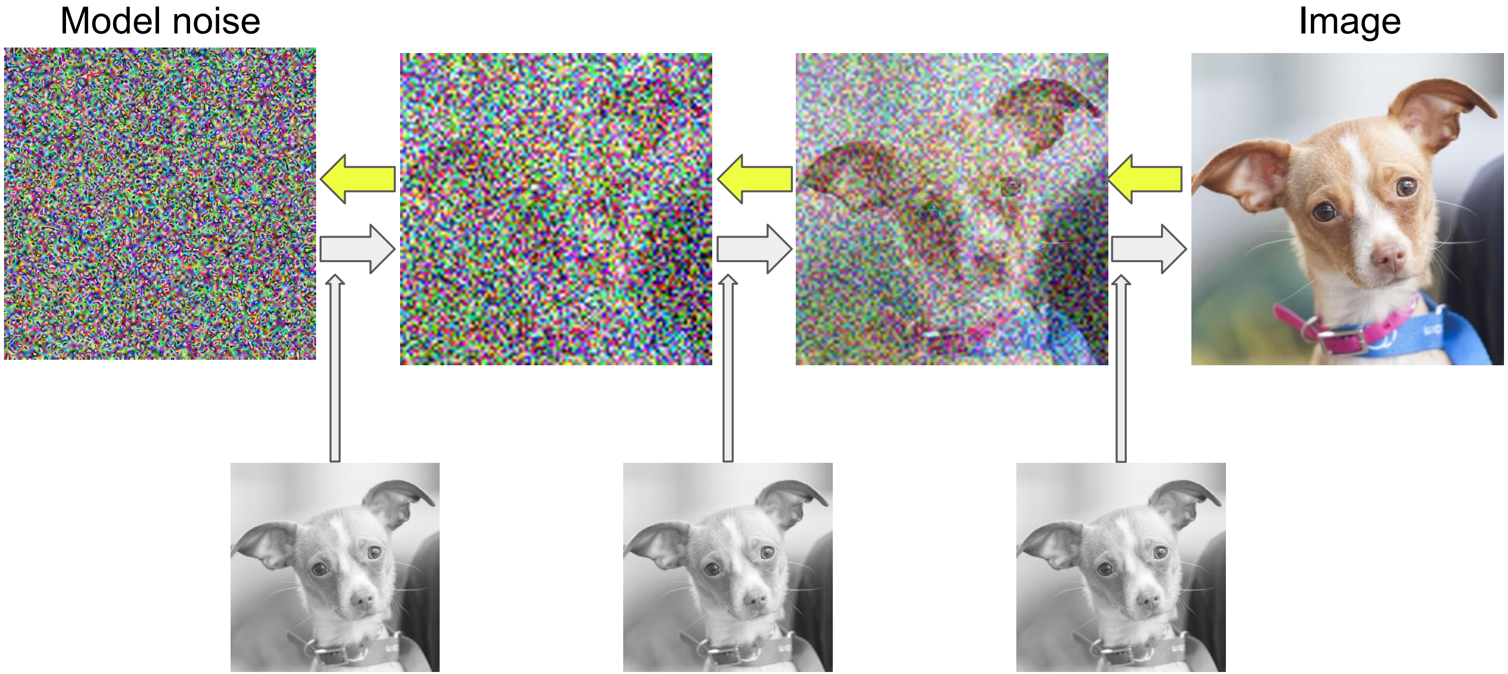


Denoising Diffusion Probabilistic Models (DDPMs)

Generative processes

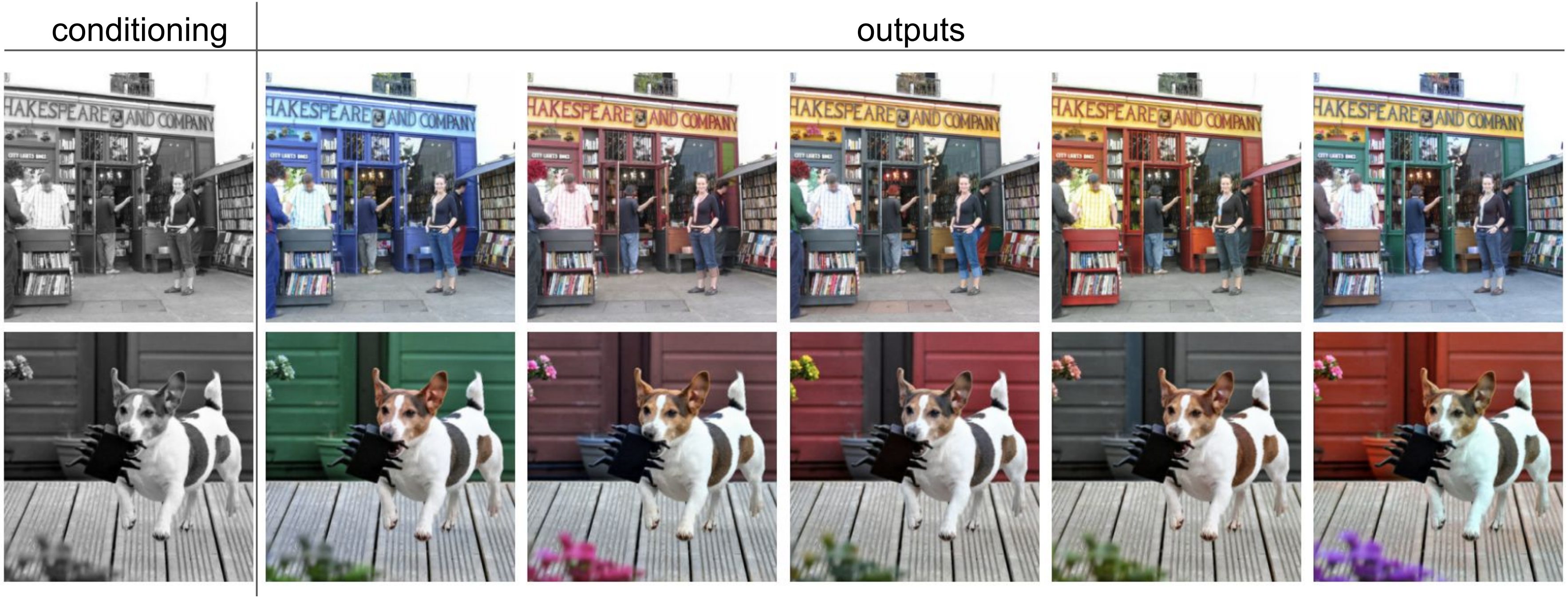


Generative processes



Generative processes

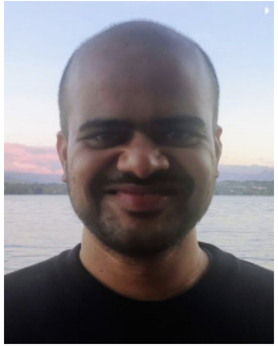
Conditional DDPMs are capable of handling ill-posed problem, as they are **stochastic**



[C. Saharia et al., 2022, Palette: Image-to-Image Diffusion Models]

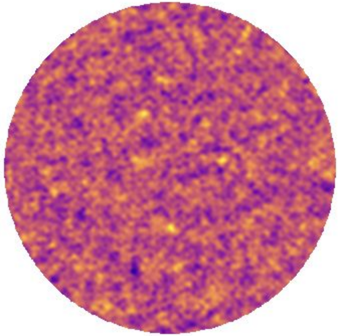
Dataset

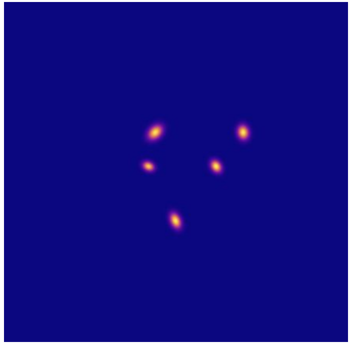
- simulated with CASA data processing software v6.2
- fixed ALMA configuration cycle 5.3
- image size 512x512 corresponding to 51.2'×51.2′ acros
- **9164** simulated sky models and the corresponding dirty images
- **1000** dirty images without sources
- a total of 27632 sources
- the root-mean-squared of the noise is 50 μJy



Dr Omkar Bait

omkar.bait@unige.ch

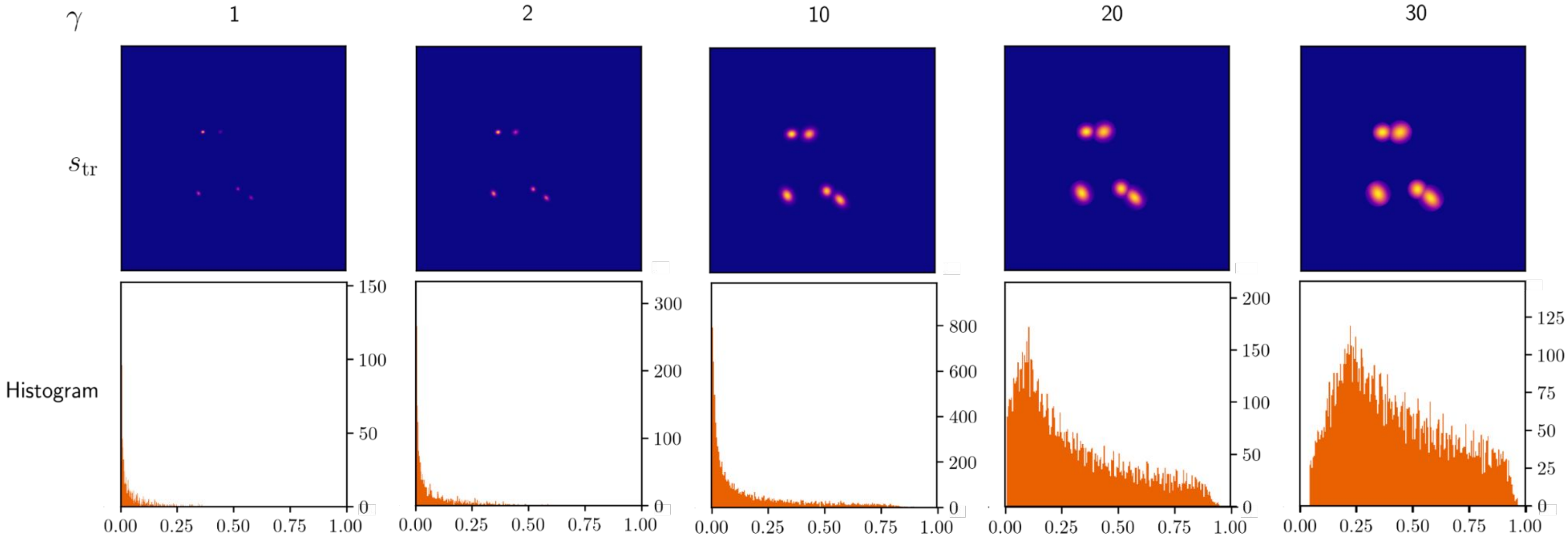




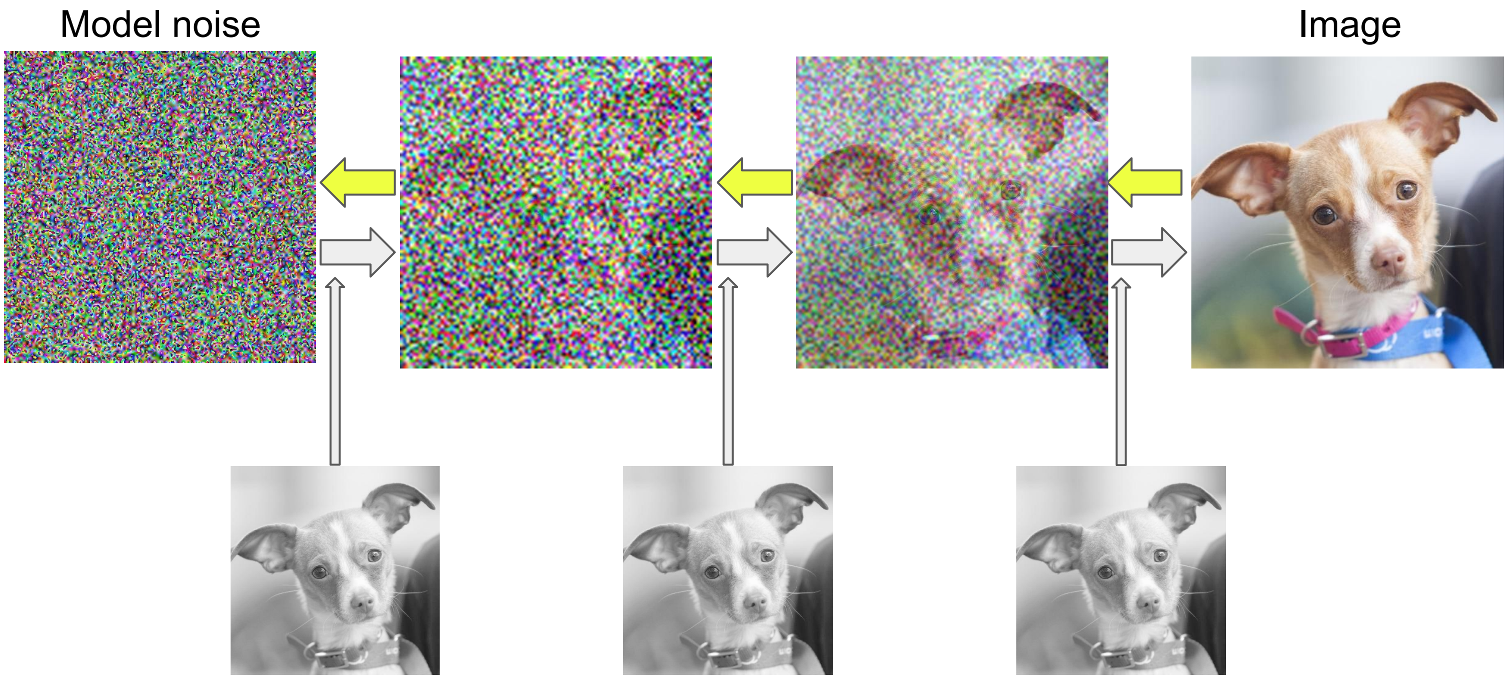
Pre-processing

Transformation applied to sky models:

$$\mathbf{s}_{\mathrm{tr}} = \left(\frac{\mathbf{s}}{c}\right)^{\frac{1}{\gamma}}$$

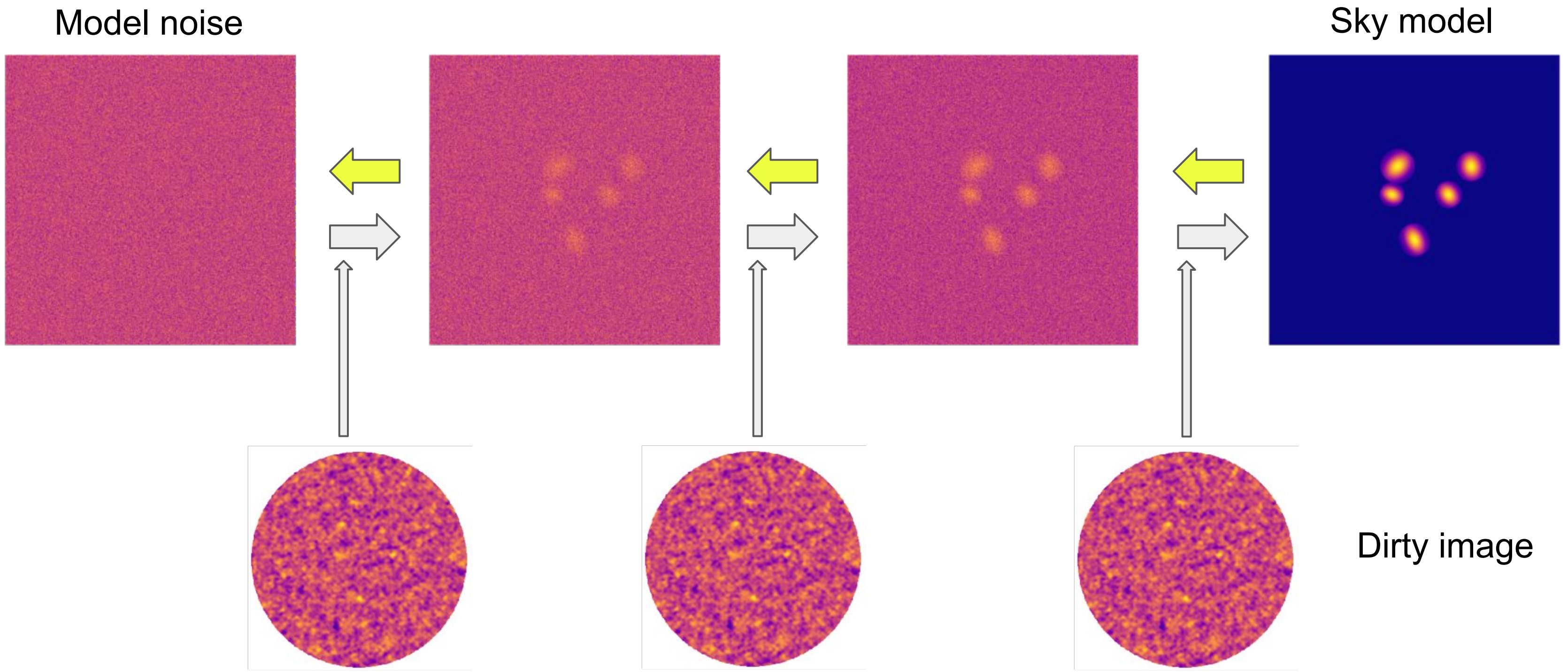


Generative processes



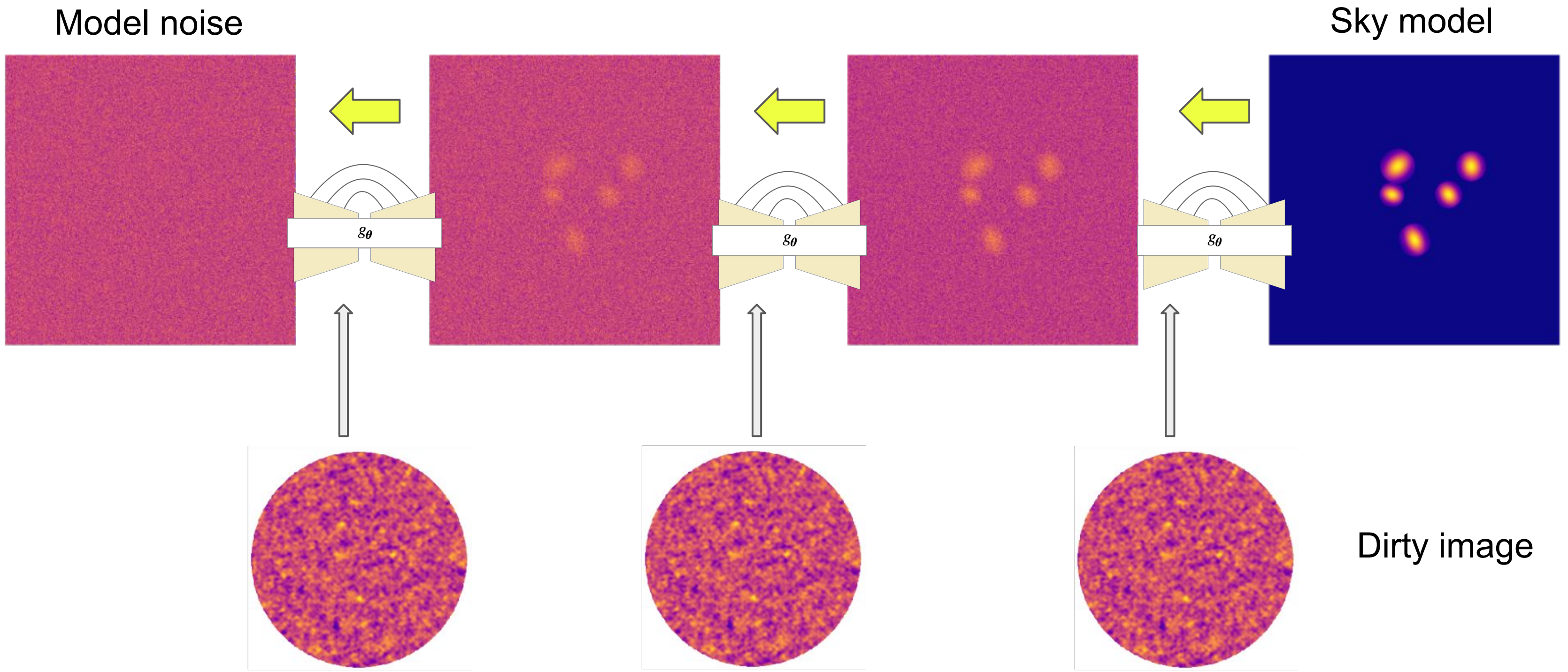
Conditional DDPMs in our problem

Full denoising step of conditional DDPMs consists of 1000 steps.

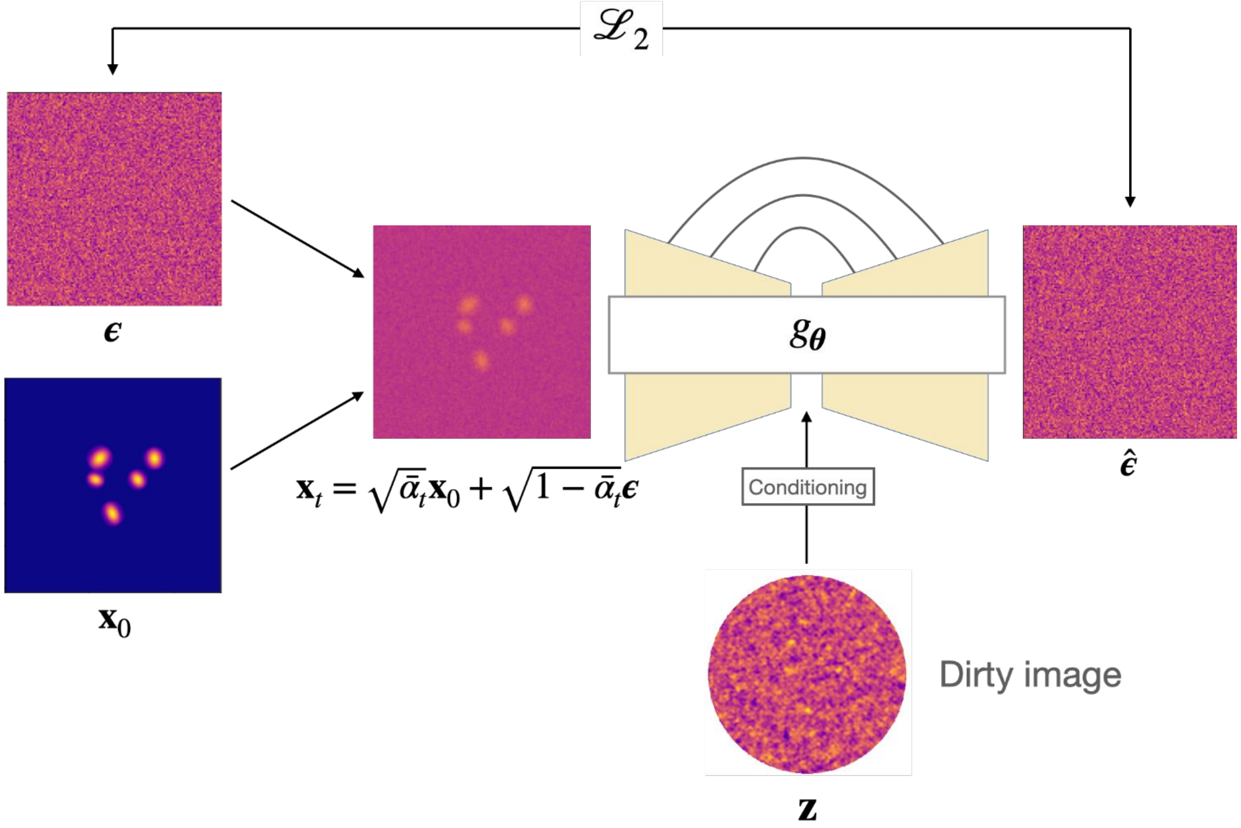


Conditional DDPMs in our problem

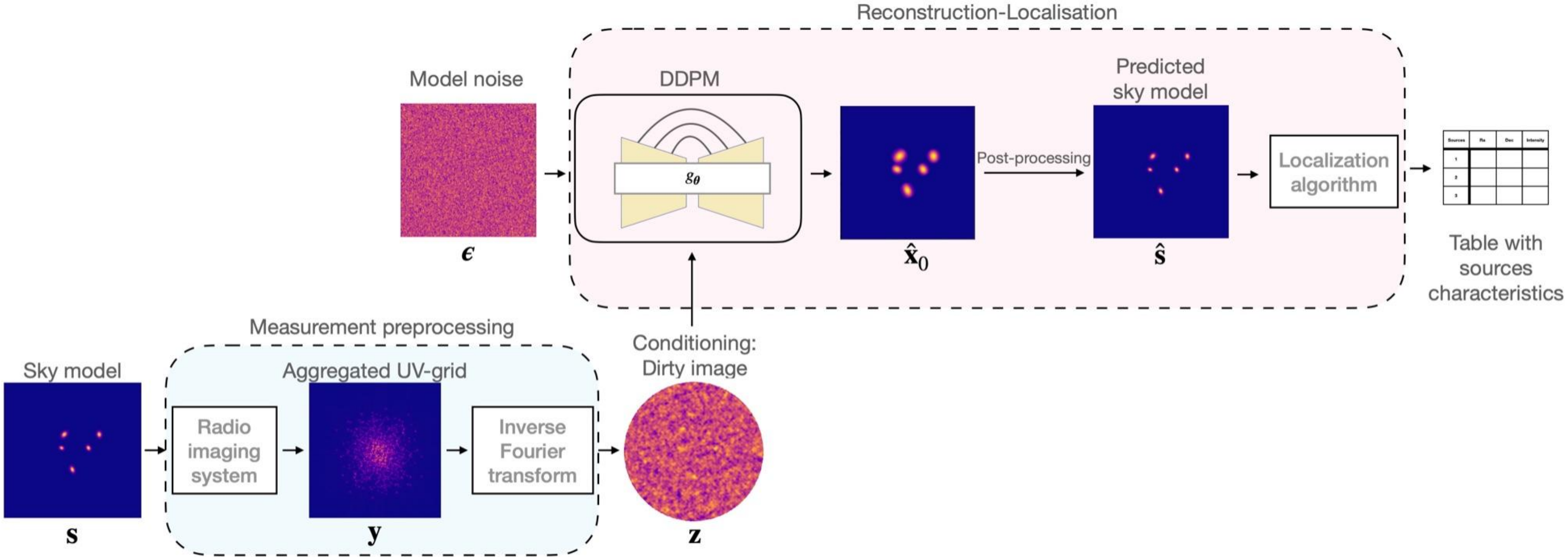
Full denoising step of conditional DDPMs consists of 1000 steps.



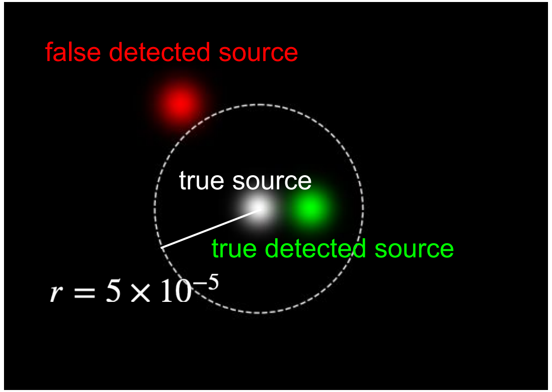
Conditional DDPMs in our problem: training



Conditional DDPMs in our problem: entire pipeline



Evaluation metrics



Reconstruction:

- $\text{MSE} = \frac{1}{H \times W} \sum_{h=1}^{H} \sum_{w=1}^{W} \left(\hat{\mathbf{x}}_0(h,w) - \mathbf{x}_0(h,w) \right)^2$
- $\text{PSNR} = 20 \cdot \log_{10}(\text{MAX}) - 10 \cdot \log_{10}(\text{MSE})$

Sources Localization:

- Purity = fraction of true sources among detected sources
- Completeness = fraction of true sources which are detected

$$\text{Purity} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Completeness} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

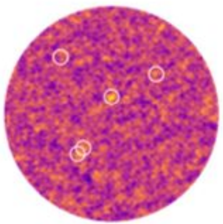
Flux estimation:

- Fraction of sources with flux estimates within noise amplitude of the true value.

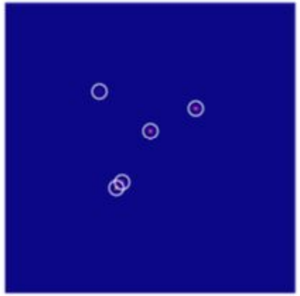
Results: reconstructions

Due to stochasticity of DDPMs we can have multiple reconstruction for the same dirty noisy image:

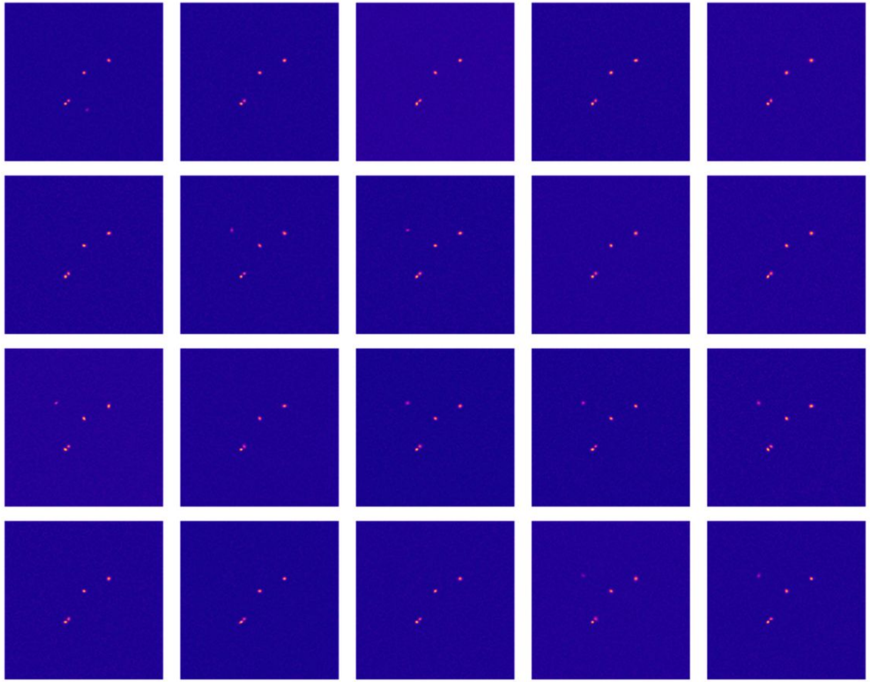
Dirty image



Sky model



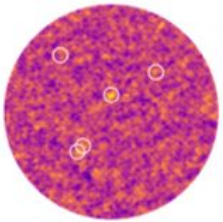
Stochastic reconstructions



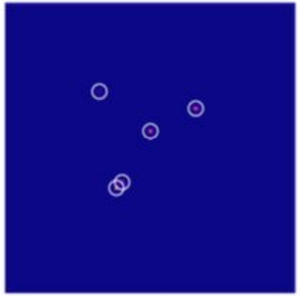
Results: reconstructions

Due to stochasticity of DDPMs we can have multiple reconstruction for the same dirty noisy image:

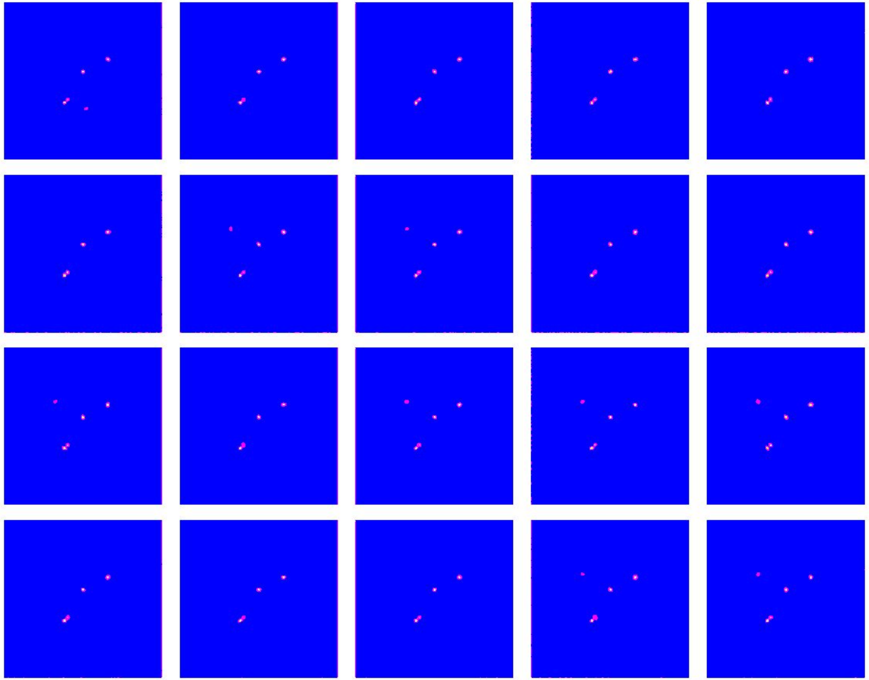
Dirty image



Sky model



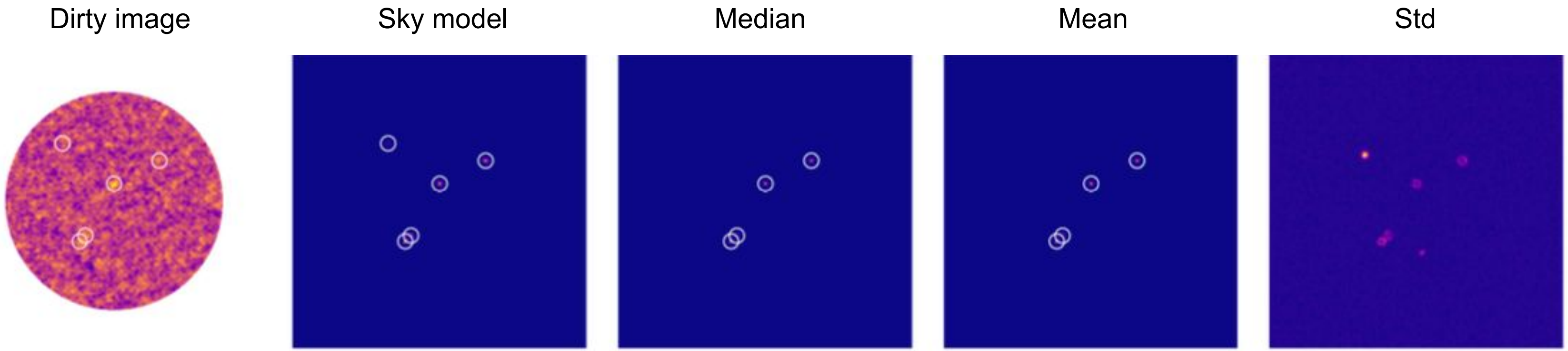
Stochastic reconstructions



Results: aggregations strategies

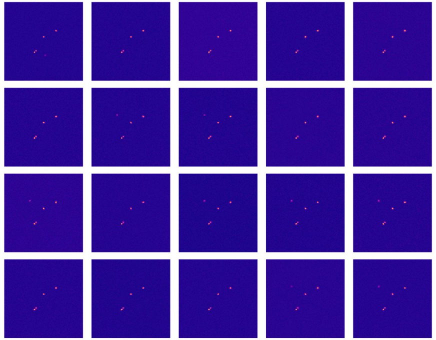
- **Aggregate-detect**:
  - Image-based aggregation (mean, median)
  - Followed by detection
  - Uncertainty image is the standard deviation across 20 outputs

This uncertainty can be used to estimate the robustness of the predictions, as well as identifying possible missed sources.



Results: aggregations strategies

- **Detect-aggregate**:
  - Detect sources in each image. Put all sources in one list.
  - Sources within distance r are merged as identical.
  - Merged sources are given mean and standard deviation for coordinates and fluxes.
  - Reliability score determined by the ratio of nb of detections to total nb of reconstructions.



Multiple reconstructions

source 1 source 2 source 3 source 4	source 1 source 2 source 3 source 4	source 1 source 2 source 3 source 4	source 1 source 2 source 3 source 4	source 1 source 2 source 3 source 4
source 1 source 2 source 3 source 4	source 1 source 2 source 3 source 4 source 5	source 1 source 2 source 3 source 4 source 5	source 1 source 2 source 3 source 4	source 1 source 2 source 3 source 4
source 1 source 2 source 3 source 4 source 5	source 1 source 2 source 3 source 4	source 1 source 2 source 3 source 4 source 5	source 1 source 2 source 3 source 4 source 5	source 1 source 2 source 3 source 4 source 5
source 1 source 2 source 3 source 4	source 1 source 2 source 3 source 4	source 1 source 2 source 3 source 4	source 1 source 2 source 3 source 4	source 1 source 2 source 3 source 4 source 5

Source list extracted from each reconstructed sky model

Merge sources

source 1 - 20
source 2 - 20
source 3 - 20
source 4 - 20
source 5 - 7

Final catalog

Results: comparative results

Median aggregation:

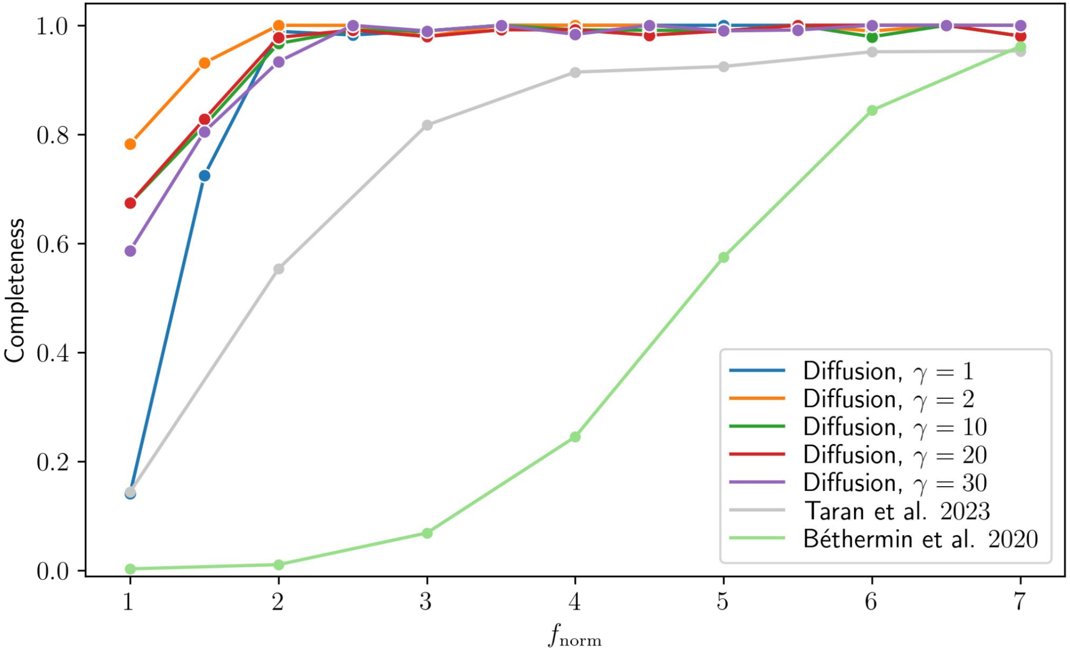
- Outstanding performance in low SNR regions.

Total localisation metrics:

- Our purity: 0.99
- Our completeness: 0.97

Previous state-of-the-art [3]:
- Purity: 0.91
- Completeness: 0.74

Completeness vs normalized SNR

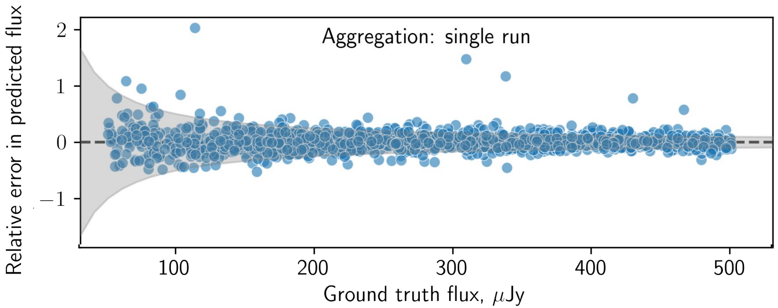


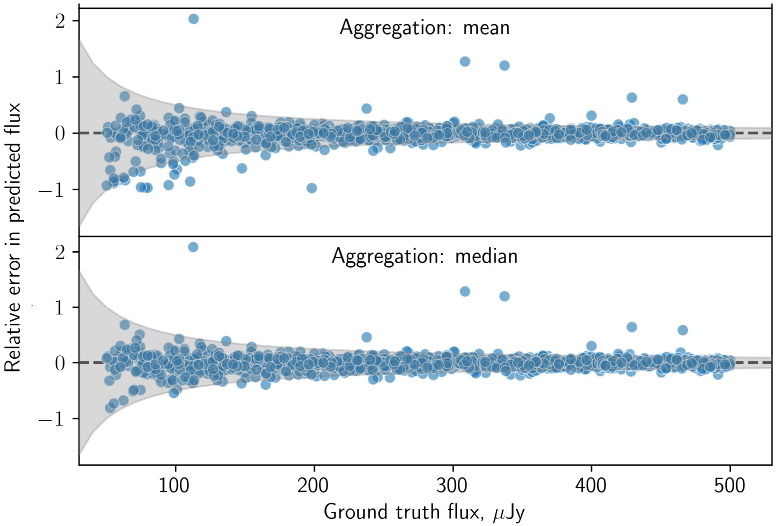
[O.Taran et al., 2023](), [Challenging interferometric imaging: Machine learning-based source localization from uv-plane observations]()

[M.Bethermin et al., 2020, The ALPINE-ALMA [CII] Survey: data processing, catalogs, and statistical source properties]()

Results: flux estimation

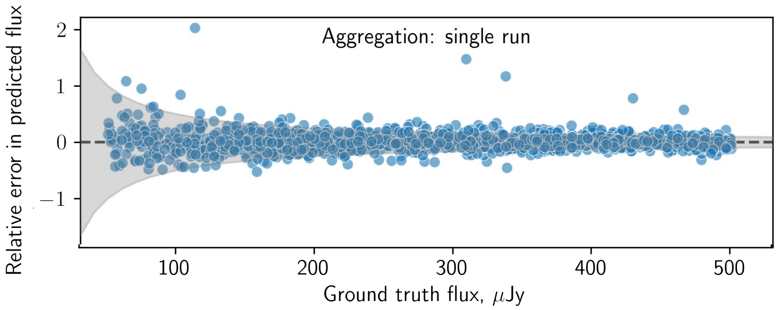
- Flux Estimation Methods: **aggregate-detect (mean, median)** vs. detect-aggregate.
- Gray area - sources with flux estimates within the noise amplitude of the actual value.

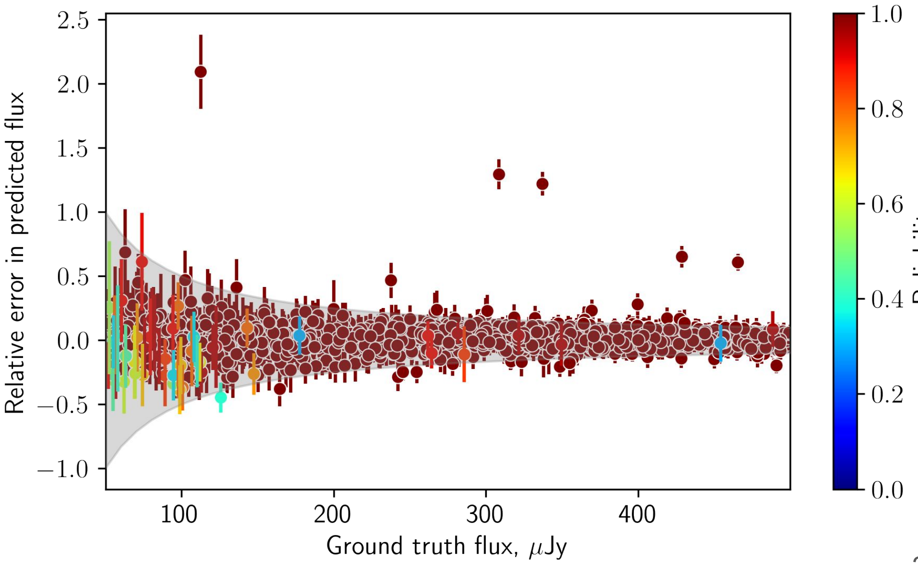




Results: flux estimation

- Flux Estimation Methods: aggregate-detect (mean, median) vs. **detect-aggregate**.
- Gray area - sources with flux estimates within the noise amplitude of the actual value.





Results: flux estimation

The Table shows the percentage of sources with flux estimates within the noise amplitude of the actual value (fraction).

Model	Normalization	Aggregation			
		Mean	Median	Single run	Detect-aggregate
PyBDSF	-			0.57	
	$\gamma = 1$	0.85	0.66	0.048	0.015
	$\gamma = 2$	**0.96**	**0.97**	**0.88**	**0.97**
	$\gamma = 10$	0.92	0.93	0.71	0.89
Diffusion	$\gamma = 20$	0.82	0.83	0.53	0.72
	$\gamma = 30$	0.74	0.75	0.47	0.66

Conclusions

- ➢ Significant improvement in source localization using our DDPMs-based approach over prior state-of-the-art.
- ➢ At SNR=2, our model achieves 0.7 completeness without normalization, surpassing the previous best of 0.55.
- ➢ Introduced a reliability estimation for predicted sources leveraging DDPMs' stochastic nature.
- ➢ Outperformed CLEAN+PyBDSF by 0.4 in estimating fluxes, specifically within noise level.
- ➢ Role of normalization power for localization and flux estimation.

Thank you!

Sources

1) Lil'Log
2) Palette: Image-to-Image Diffusion Models, C. Saharia et al., 2022
3) Challenging interferometric imaging: Machine learning-based source localization from uv-plane observations, O.Taran et al., 2023
4) AFHQ dataset

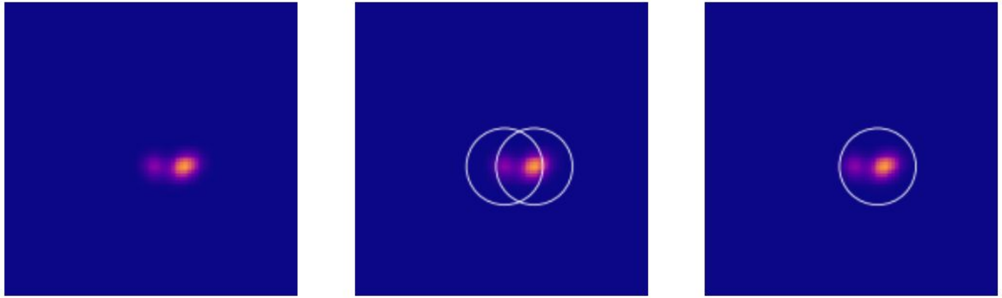
Backup

Final scores comparison

Model	Normalization	Input	Aggregation	Metrics		
				Purity	Completeness	F1
Diffusion	$\gamma = 1$	dirty image	single run	100.00	44.97	62.04
			detect-aggregate	100.00	53.52	69.72
			mean	**99.60**	**93.11**	**96.24**
			median	99.83	90.84	95.12
Diffusion	$\gamma = 2$	dirty image	single run	97.45	96.71	97.08
			detect-aggregate	98.33	97.88	98.10
			mean	97.88	97.80	97.84
			median	**99.30**	**97.05**	**98.16**
Diffusion	$\gamma = 10$	dirty image	single run	98.66	94.66	96.62
			detect-aggregate	**98.90**	**95.69**	**97.27**
			mean	98.51	95.23	96.84
			median	99.29	94.78	96.98
Diffusion	$\gamma = 20$	dirty image	single run	98.26	94.43	95.81
			detect-aggregate	**99.52**	**94.70**	**97.05**
			mean	99.52	93.64	96.49
			median	99.52	93.79	96.57
Diffusion	$\gamma = 30$	dirty image	single run	98.58	93.46	94.93
			detect-aggregate	**99.20**	**94.10**	**96.58**
			mean	98.80	93.41	96.03
			median	98.81	94.25	96.47
PyBDSF	-	dirty image	-	72.18	20.82	32.31
PyBDSF	-	dirty noiseless image	-	99.04	77.82	87.16
Taran et al. (2023)	-	reduced uv-samples	-	91.02	74.14	81.72
Photutils localization	-	sky model	-	99.70	99.10	99.40

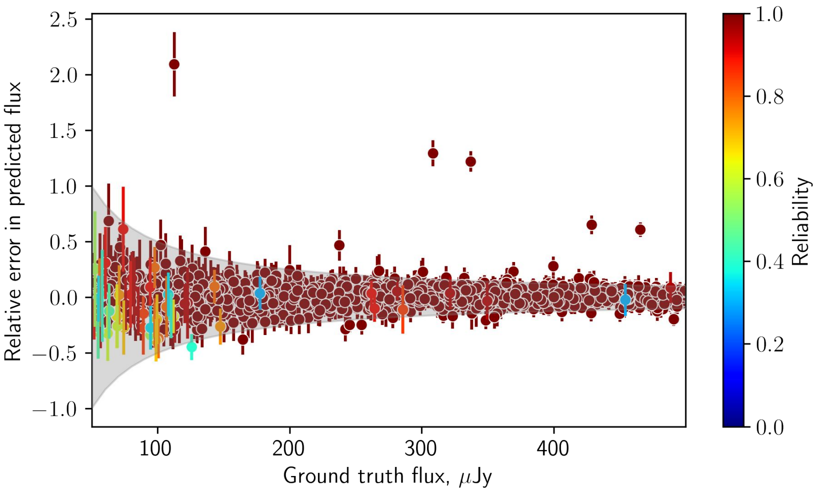
Photutils error

- Employed the photutils algorithm for localizing and characterizing sources from sky models.
- Parameter selection based on true sky models.
- Less than 1% error rate, primarily with closely situated sources, as illustrated in the example.



Flux estimation: errors

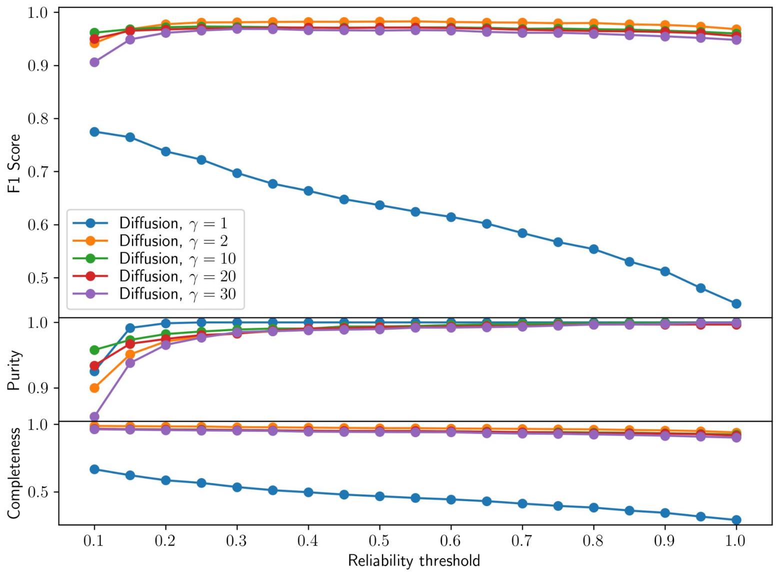
- Examples of incorrect flux estimations.
- Two cases involve degenerated scenarios with sources at the same location: one large-intensity source estimated instead of two smaller ones.
- Another example with two proximate sources, resulting in a significant overestimation of one flux.



Data Type	Ra	Dec	Flux (μJy)
Real	150.192944	3.995500	309
Real	150.192944	3.995500	429
Predicted	150.192945	3.995501	738
Real	149.655167	3.643278	238
Real	149.655167	3.643278	113
Predicted	149.655165	3.643277	346
Real	148.163222	1.370722	126
Real	148.163222	1.370444	400
Predicted	148.163220	1.370729	399
Predicted	148.163223	1.370444	436

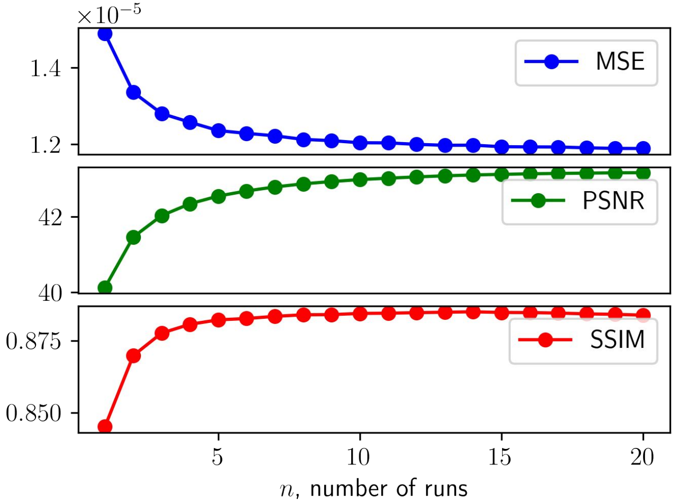
Score vs "reliability" threshold

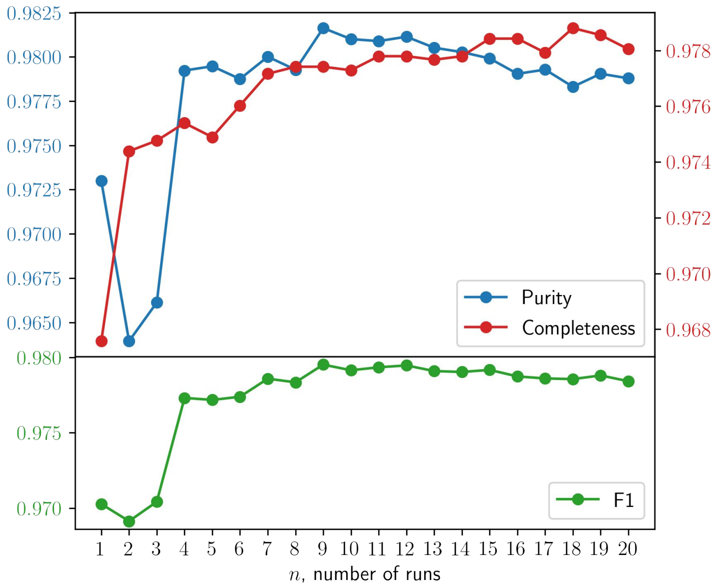
- Increasing reliability value makes our model stricter, leading to missed faint sources.
- A low value accepts all sources, resulting in potential false positives.
- These false positives don't impact completeness, but can affect accuracy.



Reconstruction and localisation metrics - the number of runs

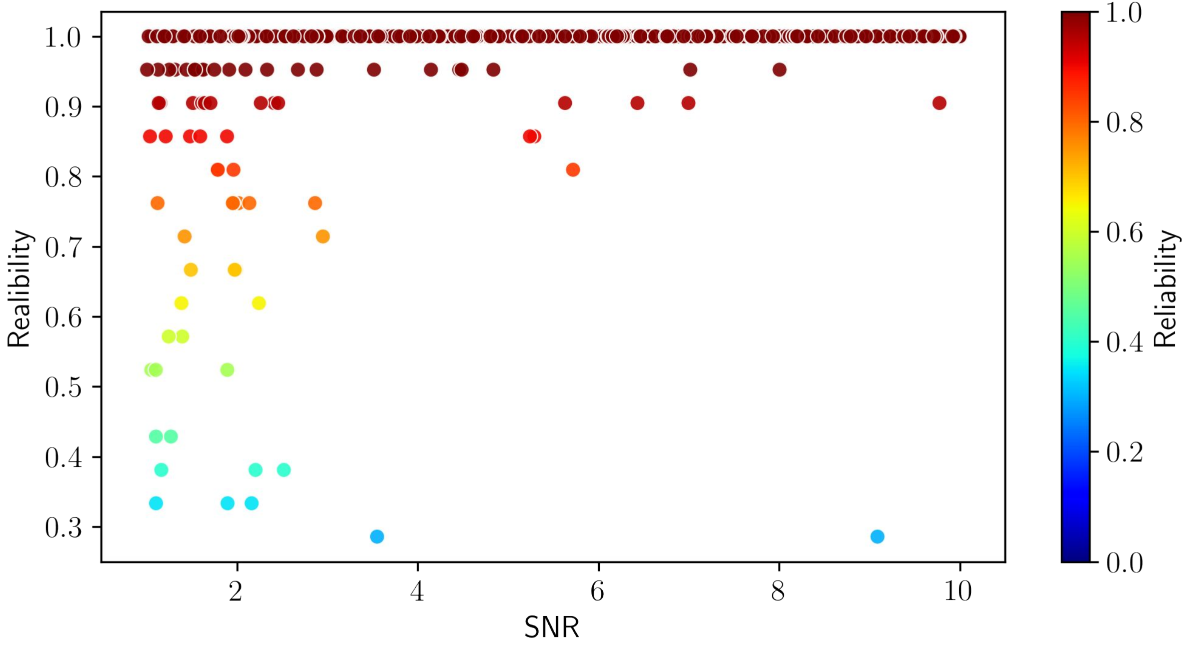
- Results aggregated with median show improved metrics with as we perform more runs of DDPM.
- Reconstruction metrics improve gradually.
- Localization metrics plateau quickly; even 5 runs is often enough.





Reliability vs SNR

- Observed correlation between introduced reliability and SNR.
- Faint sources appear in fewer runs, while bright ones consistently show in all.
- All experiments conducted over 20 runs.

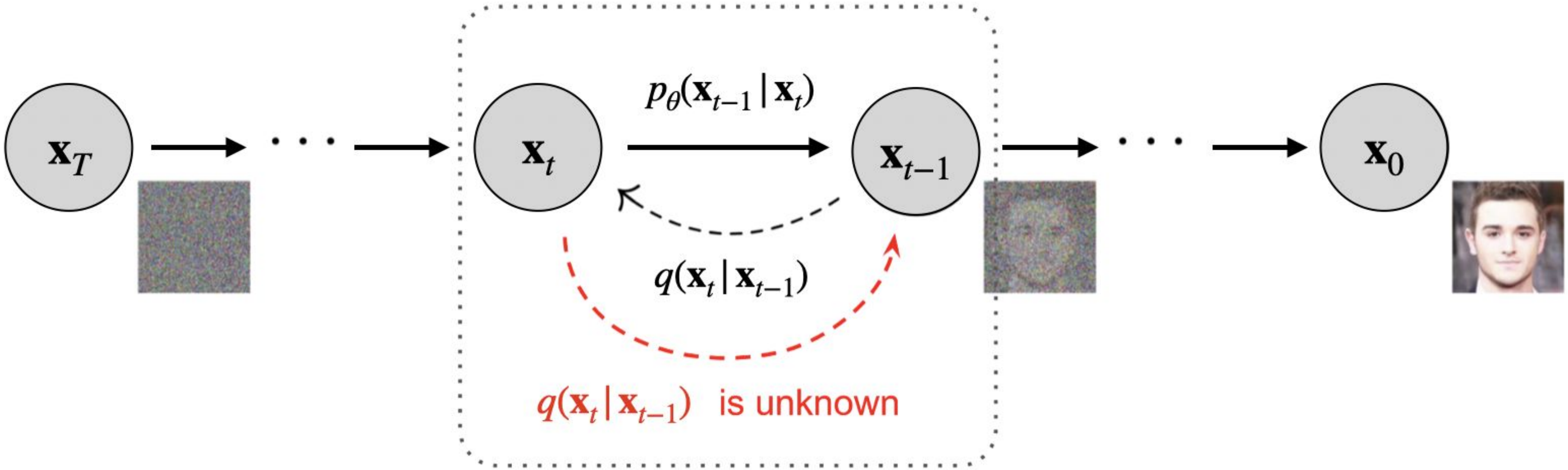


Denoising Diffusion Probabilistic Models

Two passes:

$q(\mathbf{x}_t \mid \mathbf{x}_{t-1})$ forward (progressively adding noise)

$q(\mathbf{x}_t \mid \mathbf{x}_{t-1})$ backward (progressively denoising)



*scheme from Lil'Log

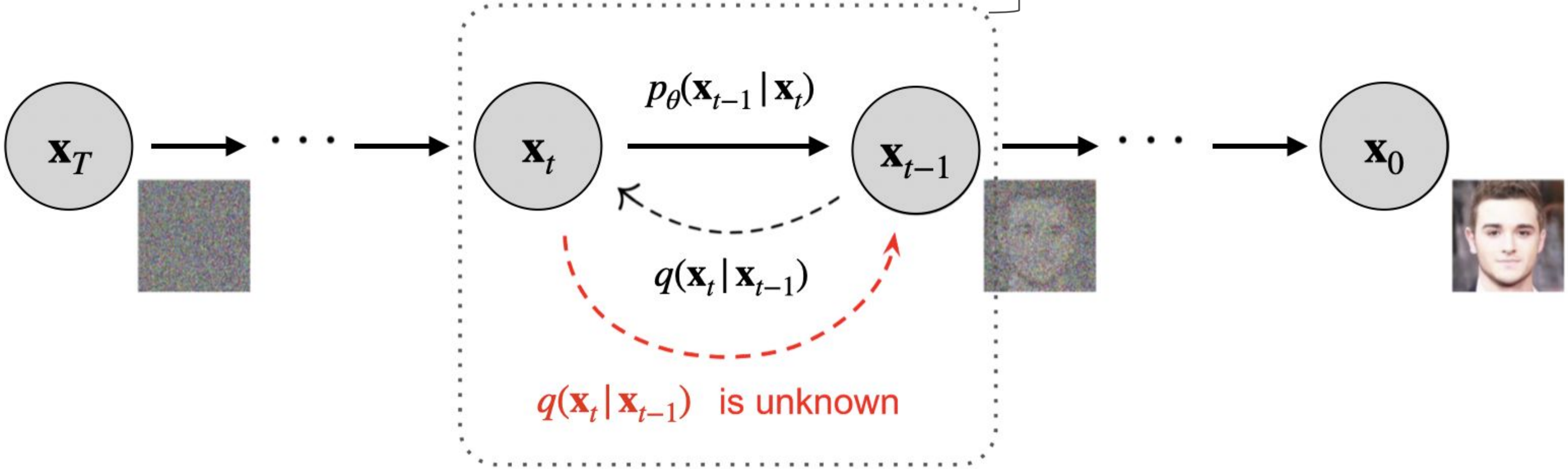
Denoising Diffusion Probabilistic Models

Two passes:

$q(\mathbf{x}_t | \mathbf{x}_{t-1})$ forward (progressively adding noise)

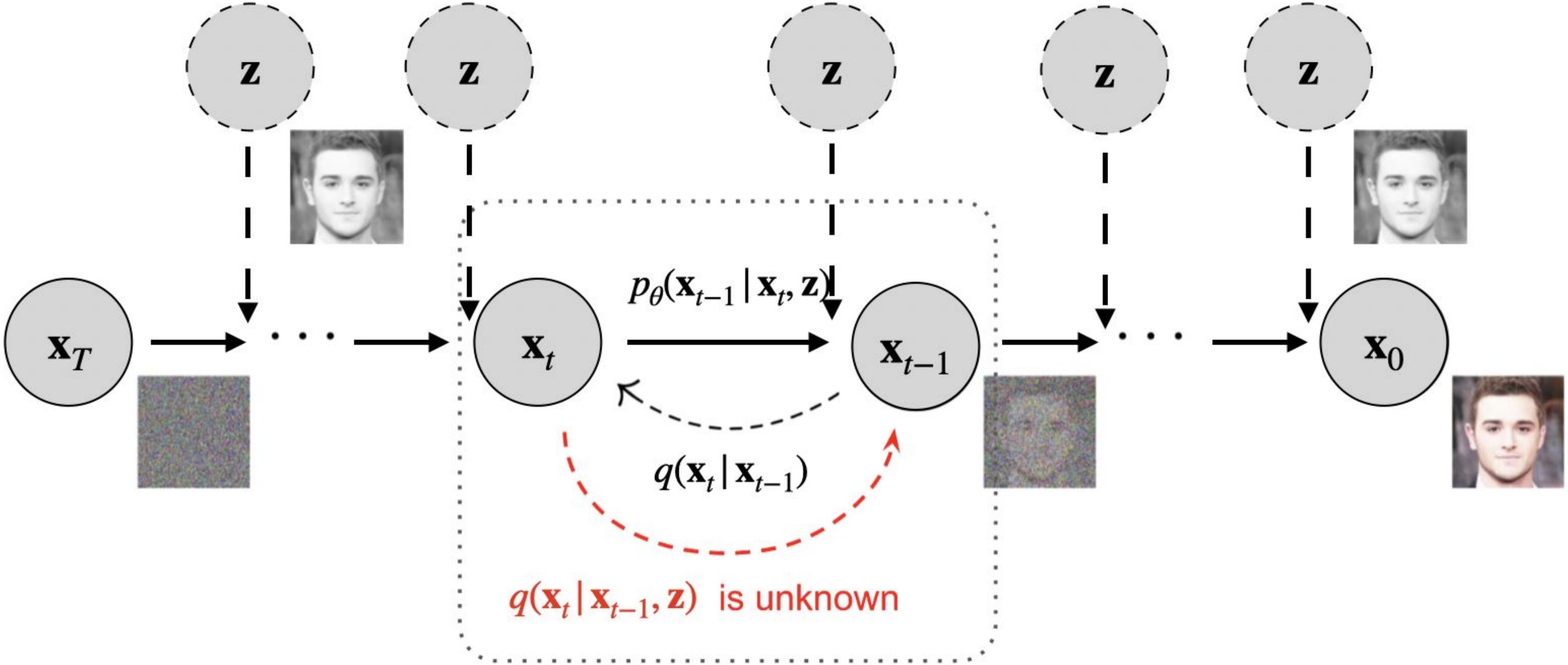
$q(\mathbf{x}_t | \mathbf{x}_{t-1})$ backward (progressively denoising)

both processes are stochastic, what if we need guidance?



*scheme from Lil'Log

Conditional Denoising Diffusion Probabilistic Models!

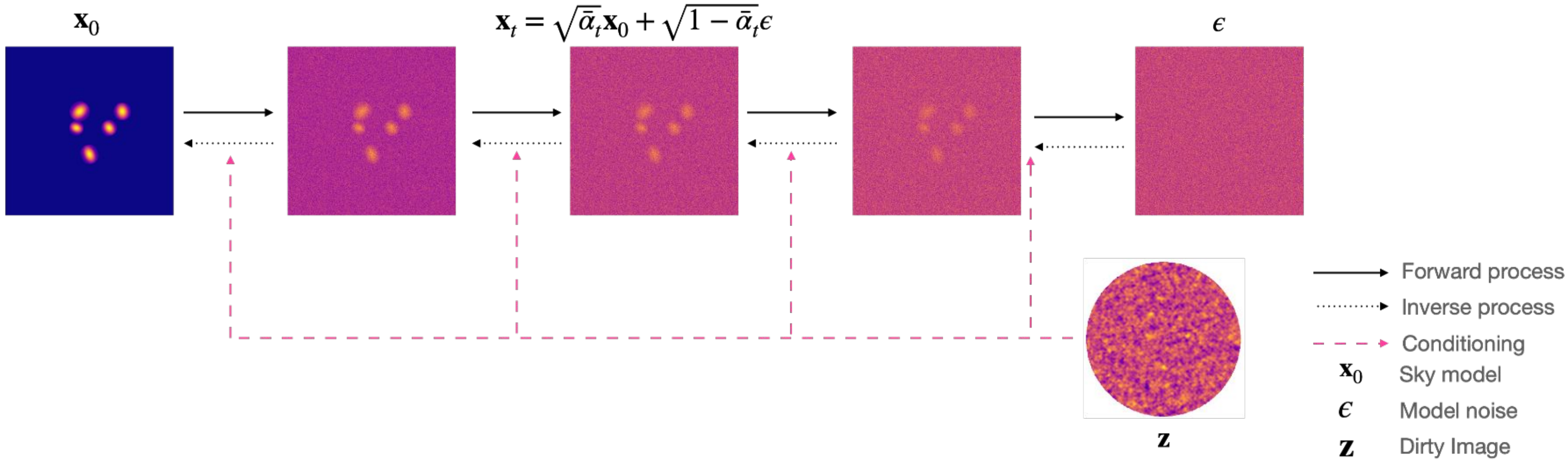


*scheme from Lil'Log

Conditional DDPM in our problem

CASA simulations for ALMA: dirty images (condition) and sky models (output)

Noise of DDPM goes through conditioned inverse process 250 times to get the sky model.



Limitations

- Performance declines as amount of water vapor increases.
- With water vapor triple the training value, completeness drops from 95% to 75%.
- Indicates fine-tuning may be beneficial.