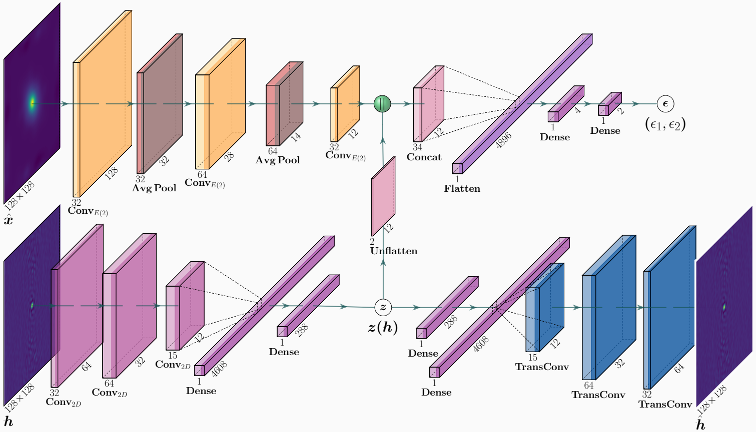
Radio Shape Measurement using Deep-Learning

TOSCA Update

Priyamvad Tripathi
08 November, 2024
Observatoire de la Côte d’Azur

Network Structure



- We used an E(2) equivariant convolutional neural network (CNN) [Weiler and Cesa, 2019] for our feature extraction layer

Image feature extraction

- We used an E(2) equivariant convolutional neural network (CNN) [Weiler and Cesa, 2019] for our feature extraction layer
- Equivariance is enforced in the structure by using convolution kernels expressed in a steerable basis of the $E(2)$ group:

$$k(\mathbf{x}|\mathbf{w}) = \sum_{\ell=1}^{8} w_\ell(r) Y_\ell(\alpha) \tag{1}$$

where $\mathbf{x} = (r, \alpha)$, $Y_\ell(\alpha) = e^{i\ell\alpha}$ are the basis vectors and the kernel weights $w_\ell(r)$ have a radial symmetry.

- We used an E(2) equivariant convolutional neural network (CNN) [Weiler and Cesa, 2019] for our feature extraction layer
- Equivariance is enforced in the structure by using convolution kernels expressed in a steerable basis of the $E(2)$ group:

$$k(\mathbf{x}|\mathbf{w}) = \sum_{\ell=1}^{8} w_\ell(r) Y_\ell(\alpha) \tag{1}$$

  where $\mathbf{x} = (r, \alpha)$, $Y_\ell(\alpha) = e^{i\ell\alpha}$ are the basis vectors and the kernel weights $w_\ell(r)$ have a radial symmetry.
- Produces a vector feature map that is equivariant to the actions of the $E(2)$ group:

$$\mathrm{C}_{E(2)}[G(\hat{x})] = G[\mathrm{C}_{E(2)}(\hat{x})] \tag{2}$$

PSF Encoding

- PSF can be computed using the uv sampling pattern and galaxy position

PSF Encoding

- PSF can be computed using the uv sampling pattern and galaxy position
- The reconstructed images contain artifacts from the corresponding PSFs

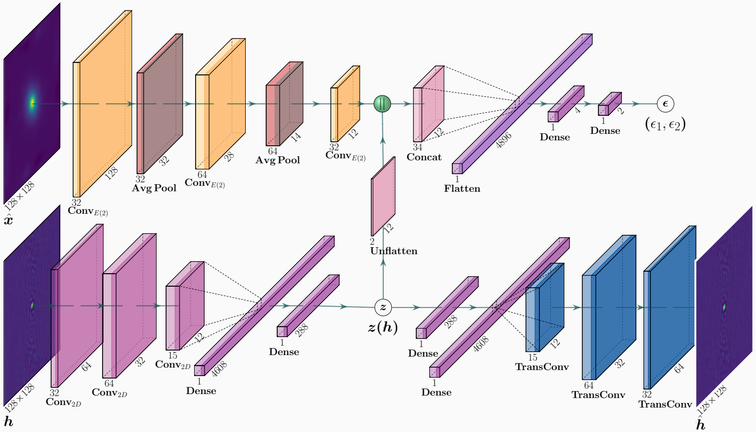
PSF Encoding

- PSF can be computed using the uv sampling pattern and galaxy position
- The reconstructed images contain artifacts from the corresponding PSFs
- We first train an autoencoder to encode the PSFs to a latent space

- PSF can be computed using the uv sampling pattern and galaxy position
- The reconstructed images contain artifacts from the corresponding PSFs
- We first train an autoencoder to encode the PSFs to a latent space
- $$\{\hat{\boldsymbol{\theta}}_{\mathrm{E}}, \hat{\boldsymbol{\theta}}_{\mathrm{D}}\} = \underset{\{\boldsymbol{\theta}_{\mathrm{E}},\boldsymbol{\theta}_{\mathrm{D}}\}}{\operatorname{argmin}} \mathbb{E}_h[\|h - \hat{h}_{\boldsymbol{\theta}_{\mathrm{D}}}[\boldsymbol{z}_{\boldsymbol{\theta}_{\mathrm{E}}}(h)]\|^2] \tag{3}$$

  where $\hat{h}_{\boldsymbol{\theta}_{\mathrm{D}}}$ is the output from the decoder, $\boldsymbol{z}_{\boldsymbol{\theta}_{\mathrm{E}}}$ is the output from the encoder and $\{\boldsymbol{\theta}_{\mathrm{E}},\boldsymbol{\theta}_{\mathrm{D}}\}$ the encoder-decoder architecture parameters.

Network Structure



- **Population Model:** Star-Forming Galaxies (SFGs) catalogue from the Tiered-Radio Extragalactic Continuum Simulation (T- RECS) [Bonaldi et al., 2018]
- Isolated galaxies with galaxy center at antenna pointing: 128×128 pixels

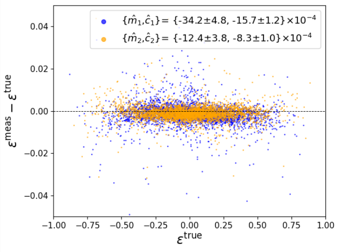
Simulating galaxy images

- **Population Model:** Star-Forming Galaxies (SFGs) catalogue from the Tiered-Radio Extragalactic Continuum Simulation (T- RECS) [Bonaldi et al., 2018]
- Isolated galaxies with galaxy center at antenna pointing: 128×128 pixels
- Visibilities based on SKA-MID configuration: 197 antennas, 154k unique visibility positions

Simulating galaxy images

- **Population Model:** Star-Forming Galaxies (SFGs) catalogue from the Tiered-Radio Extragalactic Continuum Simulation (T- RECS) [Bonaldi et al., 2018]
- Isolated galaxies with galaxy center at antenna pointing: 128×128 pixels
- Visibilities based on SKA-MID configuration: 197 antennas, 154k unique visibility positions
- Visibilities are gridded followed by Inverse FFT to get dirty image x^D

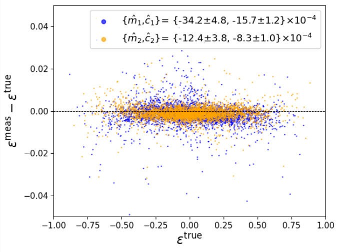
- **Population Model:** Star-Forming Galaxies (SFGs) catalogue from the Tiered-Radio Extragalactic Continuum Simulation (T- RECS) [Bonaldi et al., 2018]
- Isolated galaxies with galaxy center at antenna pointing: 128×128 pixels
- Visibilities based on SKA-MID configuration: 197 antennas, 154k unique visibility positions
- Visibilities are gridded followed by Inverse FFT to get dirty image x^D
- PSF h is then deconvolved from the dirty image using MS-Clean to get reconstructed image $\hat{x}$



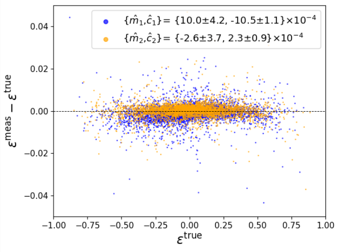
(a) Reconstructed Images: 1000 MS-Clean cycles

Figure 1: Trained/validated/tested on 16k/2k/2k galaxies-PSF pairs with varying size and intrinsic ellipticity. Autoencoder pretrained using 80k PSFs

Full Model

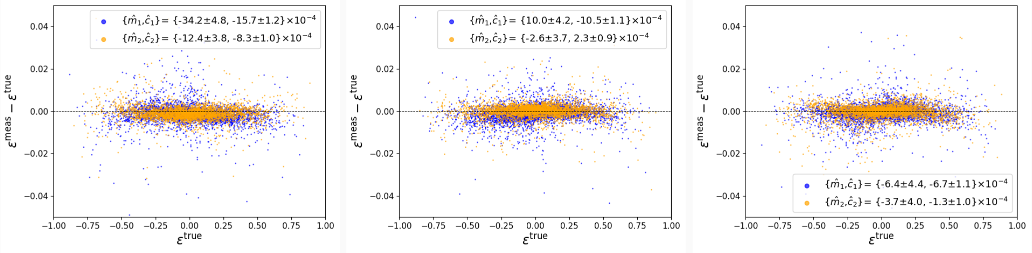


(a) Reconstructed Images: 1000 MS-Clean cycles



(b) Reconstructed Images: 500 MS-Clean cycles

Figure 1: Trained/validated/tested on 16k/2k/2k galaxies-PSF pairs with varying size and intrinsic ellipticity. Autoencoder pretrained using 80k PSFs



(a) Reconstructed Images: 1000 MS-Clean cycles

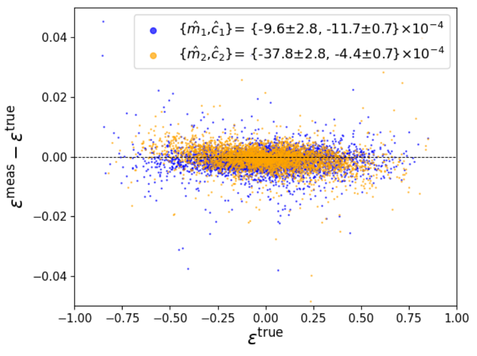
(b) Reconstructed Images: 500 MS-Clean cycles

(c) Dirty Images

Figure 1: Trained/validated/tested on 16k/2k/2k galaxies-PSF pairs with varying size and intrinsic ellipticity. Autoencoder pretrained using 80k PSFs

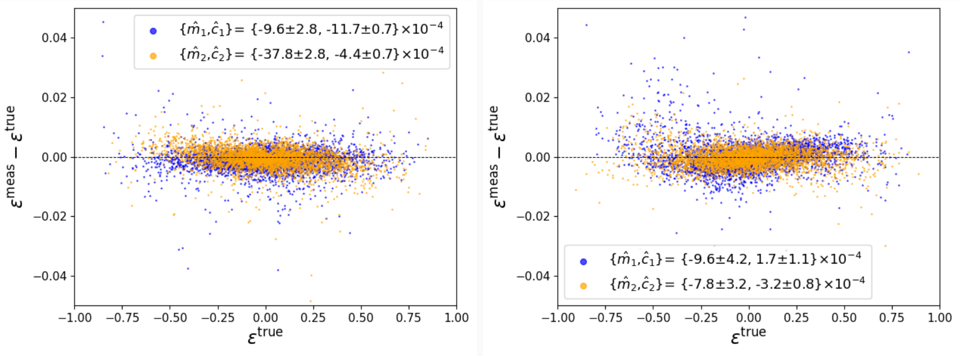
Model Bias

Figure 2: Galaxies following a Sérsic brightness profile: $I(r) = I_0 \exp[-(\frac{r}{r_0})^{\frac{1}{n}}]$ with index n drawn from $\mathcal{U}(1, 4)$



(a) Reconstructed Images

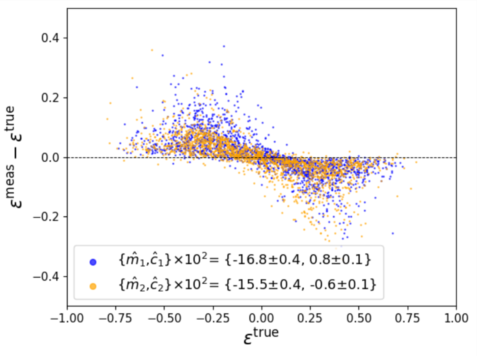
Figure 2: Galaxies following a Sérsic brightness profile: $I(r) = I_0 \exp[-(\frac{r}{r_0})^{\frac{1}{n}}]$ with index n drawn from $\mathcal{U}(1, 4)$



(a) Reconstructed Images

(b) Dirty Images

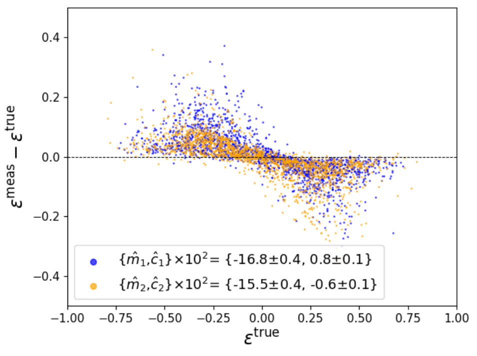
Figure 2: Galaxies following a Sérsic brightness profile: $I(r) = I_0 \exp[-(\frac{r}{r_0})^{\frac{1}{n}}]$ with index n drawn from $\mathcal{U}(1,4)$



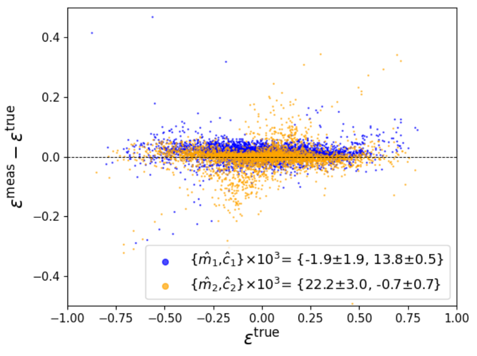
(a) ShapeNet Deconvolution

Figure 3: Tested on same test case (exp profile). ShapeNet has been trained/validated on same dataset.

Comparison with Other works



(a) ShapeNet Deconvolution



(b) SuperCLASS Calibration

Figure 3: Tested on same test case (exp profile). ShapeNet has been trained/validated on same dataset.

Table 1: Linear Bias in Ellipticity estimates (at the order of 10^{-3})

		m_1	c_1	m_2	c_2
Exp Profile	Recon 500	1.0 ± 0.4	-1.1 ± 0.1	$\mathbf{-0.3 \pm 0.4}$	0.2 ± 0.1
	Recon 1000	-3.4 ± 0.5	-1.6 ± 0.1	-1.2 ± 0.4	-0.8 ± 0.1
	Dirty	$\mathbf{-0.6 \pm 0.4}$	$\mathbf{-0.7 \pm 0.1}$	-0.4 ± 0.4	$\mathbf{-0.1 \pm 0.1}$
	Shapenet Decon	-168.1 ± 4.0	8.3 ± 1.4	-155.3 ± 4.1	-6.8 ± 1.0
	SuperCLASS Calib	-1.9 ± 1.9	13.8 ± 0.5	22.2 ± 3.0	-0.7 ± 0.7
Sersic Profile	Recon	$\mathbf{1.0 \pm 0.3}$	-1.2 ± 0.1	-3.8 ± 0.3	-0.4 ± 0.1
	Dirty	-1.0 ± 0.4	$\mathbf{0.2 \pm 0.1}$	$\mathbf{-0.8 \pm 0.3}$	$\mathbf{-0.3 \pm 0.1}$

Realistic Noise

- Initial Approach: $V_g^{obs} = V_g^{calc} + \mathcal{N}(0, \sigma_g/20)$ where $\sigma_g = \sigma(V_g^{calc})$ for each galaxy g

Realistic Noise

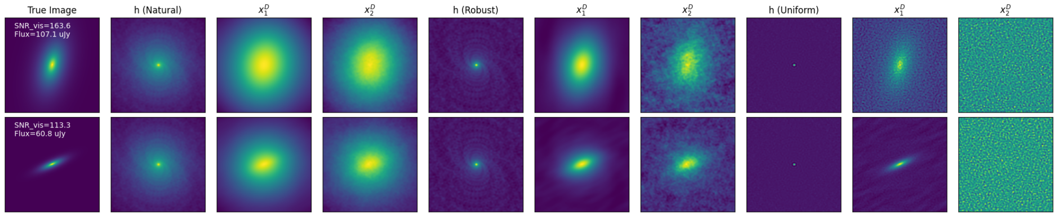
- Initial Approach: $V_g^{\text{obs}} = V_g^{\text{calc}} + \mathcal{N}(0, \sigma_g/20)$ where $\sigma_g = \sigma(V_g^{\text{calc}})$ for each galaxy g
- Realistic Situation: $V_g^{\text{obs}} = V_g^{\text{calc}} + \mathcal{N}(0, \sigma)$ where $\sigma = \frac{2\kappa_B T_{\text{sys}}}{A_{\text{eff}}} \times \frac{1}{n_s \sqrt{2\Delta\nu\tau_{\text{int}}}}$

Realistic Noise

- Initial Approach: $V_g^{\text{obs}} = V_g^{\text{calc}} + \mathcal{N}(0, \sigma_g/20)$ where $\sigma_g = \sigma(V_g^{\text{calc}})$ for each galaxy g
- Realistic Situation: $V_g^{\text{obs}} = V_g^{\text{calc}} + \mathcal{N}(0, \sigma)$ where $\sigma = \frac{2\kappa_B T_{\text{sys}}}{A_{\text{eff}}} \times \frac{1}{n_s\sqrt{2\Delta\nu\tau_{\text{int}}}}$
- More careful consideration of galaxy properties like flux and size, and visibility weights

Realistic Noise

- Initial Approach: $V_g^{obs} = V_g^{calc} + \mathcal{N}(0, \sigma_g/20)$ where $\sigma_g = \sigma(V_g^{calc})$ for each galaxy g
- Realistic Situation: $V_g^{obs} = V_g^{calc} + \mathcal{N}(0, \sigma)$ where $\sigma = \frac{2\kappa_B T_{sys}}{A_{eff}} \times \frac{1}{n_s\sqrt{2\Delta\nu\tau_{int}}}$
- More careful consideration of galaxy properties like flux and size, and visibility weights



RadioLensfit [Rivi et al., 2016]

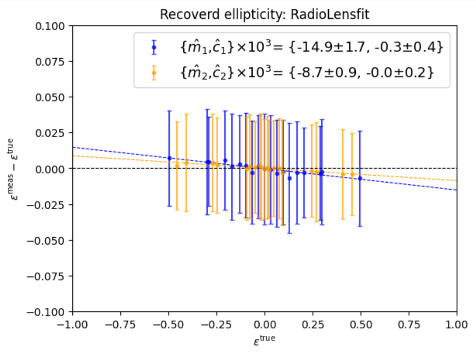
- Works using visibilities
- Galaxy brightness profile: $I(r) = I_0 \exp(-r/\alpha)$,
- Transformation matrix $\mathbf{A}$ with ellipticity parameters $\mathbf{e} = (e_1, e_2)$ such that:

$$\begin{pmatrix} l_r \\ m_r \end{pmatrix} = \mathbf{A}x = \begin{pmatrix} 1 - e_1 & -e_2 \\ -e_2 & 1 + e_1 \end{pmatrix} \times \begin{pmatrix} l \\ m \end{pmatrix}$$

- Observed visibility due to a galaxy at point $\mathbf{k} = (u, v)$ can be given by:

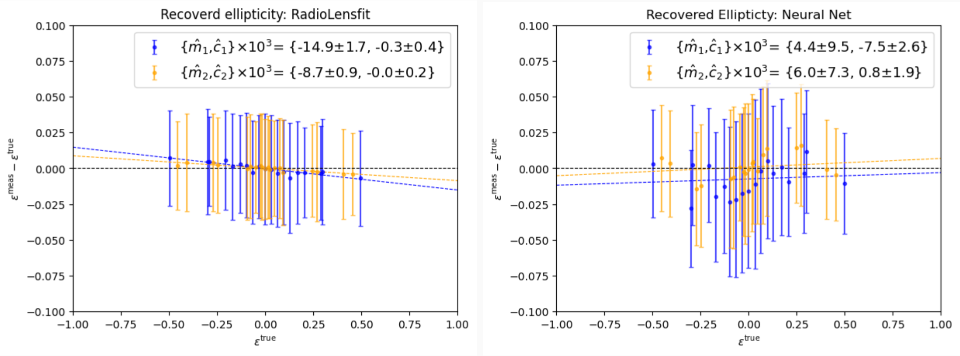
$$V_s(u, v) = \frac{2\pi\alpha^2 I_0}{|\mathbf{A}|(1 + 4\pi^2\alpha^2|\mathbf{A}^{-\intercal}k|)^{3/2}} \times \exp 2\pi i \mathbf{k}^\intercal \mathbf{x_0} \tag{4}$$

- Perform a Bayesian marginalization of the likelihood over I_0, α and source centroid position $\mathbf{x_0} = (l_0, m_0) \Rightarrow P(\mathbf{A}|D)$



(a) RadioLensfit Measurement
MAE=4.1×10^{-3}, RMSE=0.5×10^{-2}

Figure 4: Measurements made on common test (25×500) set with exponential profile, Flux $50-200\ \mu$Jy

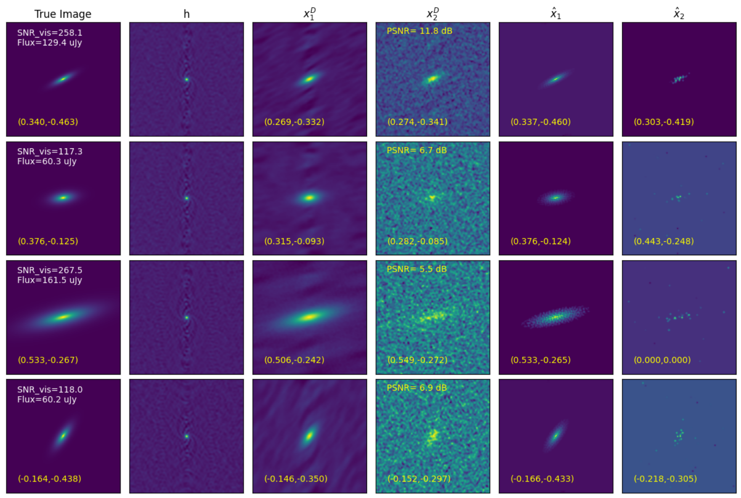


(a) RadioLensfit Measurement
MAE=4.1×10^{-3}, RMSE=0.5×10^{-2}

(b) Neural Net trained with 200k $x^D - h$ pairs
MAE=9.7×10^{-3}, RMSE=1.5×10^{-2}

Figure 4: Measurements made on common test (25×500) set with exponential profile, Flux $50 - 200$ μJy

Original Approach: MS-Clean deconvolution



HQS-PnP algorithm [Zhang et al., 2017]

- Problem: $\mathbf{y} = \mathbf{H}\mathbf{x} + \epsilon$
- HQS sol: $\mathcal{L}_\mu(\mathbf{x}, \mathbf{z}) = \frac{1}{2}\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \lambda\Phi(\mathbf{z}) + \frac{\mu}{2}\|\mathbf{x} - \mathbf{z}\|_2^2$
- Iterative Sol:

$$\mathbf{x}_{k+1} = \arg\min_x \frac{1}{2}\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \mu\|\mathbf{x} - \mathbf{z}_k\|_2^2$$

$$\mathbf{z}_{k+1} = \arg\min_z \frac{\mu}{2}\|\mathbf{x}_{k+1} - \mathbf{z}\|_2^2 + \lambda\Phi(\mathbf{z})$$

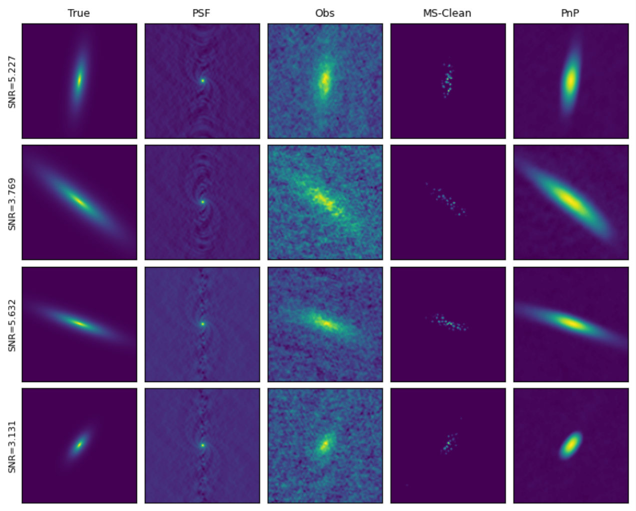
- Sol:

$$\mathbf{x}_{k+1} = (\mathbf{H}^T\mathbf{H} + \mu\mathcal{I})^{-1}(\mathbf{H}^T\mathbf{y} + \mu\mathbf{z}_k)$$

$$\mathbf{z}_{k+1} = \arg\min_z \frac{1}{2(\sqrt{\lambda/\mu})^2}\|\mathbf{x}_{k+1} - \mathbf{z}\|_2^2 + \lambda\Phi(\mathbf{z})$$

- $\mathbf{z}_{k+1} = \text{Denoiser}\left(\mathbf{x}_{k+1}, \sqrt{\lambda/\mu}\right)$

Results



Thank you for your time

ShapeNet Deconvolution [Nammour et al., 2022]

- Tikhonov solution: $\tilde{x} = (H^T H + \lambda \Gamma^T \Gamma)^{-1} H^T x^D$ where H corresponds to the PSF operator, Γ corresponds to Tikhonov linear filter and λ is the regularisation weight.
- A UNET architecture is then trained to learn the mapping b/w the Tikhonov output and the true image.
- The network is trained to minimize the following loss function: $l(\hat{x}) = \|\hat{x} - x\|^2 + \gamma M(\hat{x})$
- $M(\hat{x}) = \sum_{i=1}^{6} \omega_i \langle \hat{x} - x, u_i \rangle$ is a shape constraint with $\{\omega_i\}$ and $\{u_i\}$ are constant scalar weights and images respectively

SuperCLASS calibration [Harrison et al., 2020]

- Reconstruct image by deconvolving the PSF from the dirty image and estimate ellipticity $\boldsymbol{\epsilon}_k^{\mathrm{calc}}$
- In the residual image, inject model sources with the same size and flux properties, but known ellipticity $\epsilon_i^{\mathrm{inp}} = \{0,\ \pm 0.2375,\ \pm 0.475,\ \pm 0.7125,\ \pm 0.95\}$
- Simulate visibilities $\Rightarrow$ Dirty Image $\Rightarrow$ Reconstructed Image $\Rightarrow$ Measure ellipticity $\boldsymbol{\epsilon}^{\mathrm{obs}}$
- Fit second order 2D polynomial $b_k(\epsilon_1^{\mathrm{inp}}, \epsilon_2^{\mathrm{inp}}) = \epsilon_1^{\mathrm{obs}} - \epsilon_1^{\mathrm{inp}}$
- Calibrate observed ellipticities using $\epsilon_{1,k}^{\mathrm{SC}} = \epsilon_{1,k}^{\mathrm{calc}} - b_k(\epsilon_{1,k}^{\mathrm{calc}}, \epsilon_{2,k}^{\mathrm{calc}})$
- Repeat for ϵ_2

Bonaldi, A., Bonato, M., Galluzzi, V., Harrison, I., Massardi, M., Kay, S., De Zotti, G., and Brown, M. L. (2018).
The tiered radio extragalactic continuum simulation (t-recs).
MNRAS, 482(1):2–19.

Harrison, I., Brown, M. L., Tunbridge, B., Thomas, D. B., Hillier, T., Thomson, A. P., Whittaker, L., Abdalla, F. B., Battye, R. A., Bonaldi, A., Camera, S., Casey, C. M., Demetroullas, C., Hales, C. A., Jackson, N. J., Kay, S. T., Manning, S. M., Peters, A., Riseley, C. J., and Watson, R. A. (2020).
Superclass – iii. weak lensing from radio and optical observations in data release 1.
MNRAS, 495(2):1737–1759.

Nammour, F., Akhaury, U., Girard, J. N., Lanusse, F., Sureau, F., Ben Ali, C., and Starck, J. L. (2022).
ShapeNet: Shape constraint for galaxy image deconvolution.
A&A, 663:A69.

Rivi, M., Miller, L., Makhathini, S., and Abdalla, F. B. (2016).
Radio weak lensing shear measurement in the visibility domain – i. methodology.
MNRAS, 463(2):1881–1890.

Weiler, M. and Cesa, G. (2019).
General e (2)-equivariant steerable cnns.
Advances in neural information processing systems, 32.

Zhang, K., Zuo, W., Gu, S., and Zhang, L. (2017).
Learning deep cnn denoiser prior for image restoration.