Equivariant splitting: self-supervised learning from incomplete data (ICLR'26)

Victor Sechaud[*1], Jérémy Scanvic[*1,2], Quentin Barthélemy[2], Patrice Abry[1], **Julián Tachella**[1]

[1]CNRS, ENS de Lyon
[2]Hirsch, Lyon, France

2026-06-04

[*]Equal contribution

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\varepsilon}$$

Introduction

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\varepsilon}$$

$\boldsymbol{x}$ → $\mathcal{X} \subset \mathbb{R}^n$

Introduction

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\varepsilon}$$

$\boldsymbol{x} \rightarrow \mathcal{X} \subset \mathbb{R}^n$

$\boldsymbol{A} \rightarrow \mathbb{R}^{m \times n}$

$$\mathcal{X} \subset \mathbb{R}^n$$

$$\mathcal{Y} \subseteq \mathbb{R}^m \quad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\varepsilon}$$

$$\mathbb{R}^{m \times n}$$

Introduction

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\varepsilon}$$

$\boldsymbol{y} \in \mathcal{Y} \subseteq \mathbb{R}^m$

$\boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^n$

$\boldsymbol{A} \in \mathbb{R}^{m \times n}$

$\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_n)$

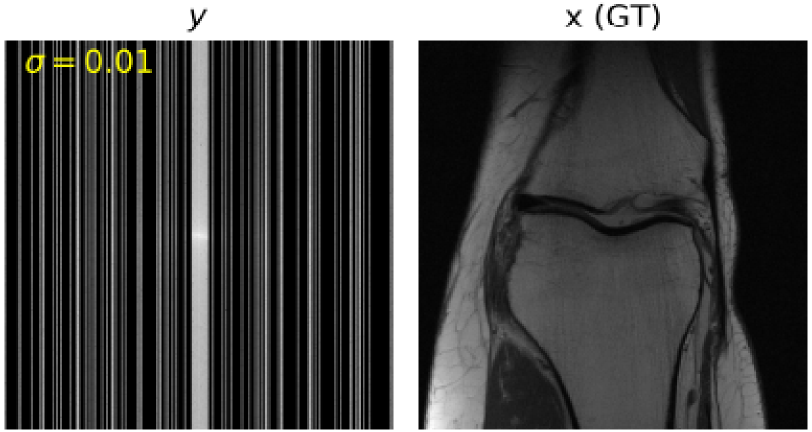
$$y = Ax + \varepsilon$$

$y \subseteq \mathbb{R}^m$

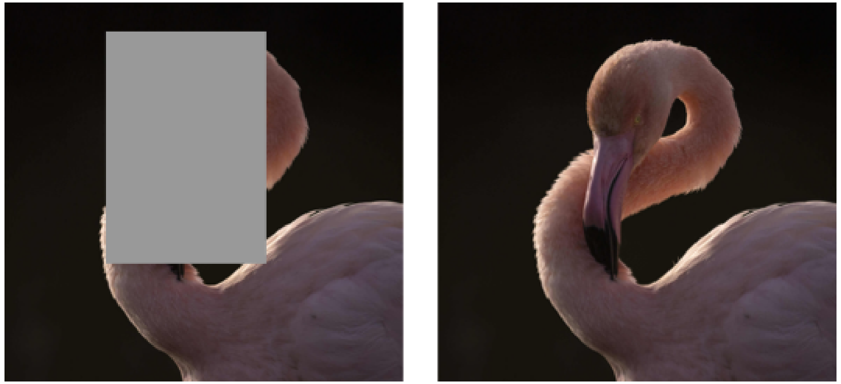
$A \in \mathbb{R}^{m \times n}$

$x \in \mathcal{X} \subset \mathbb{R}^n$

$\varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$



Accelerated MRI



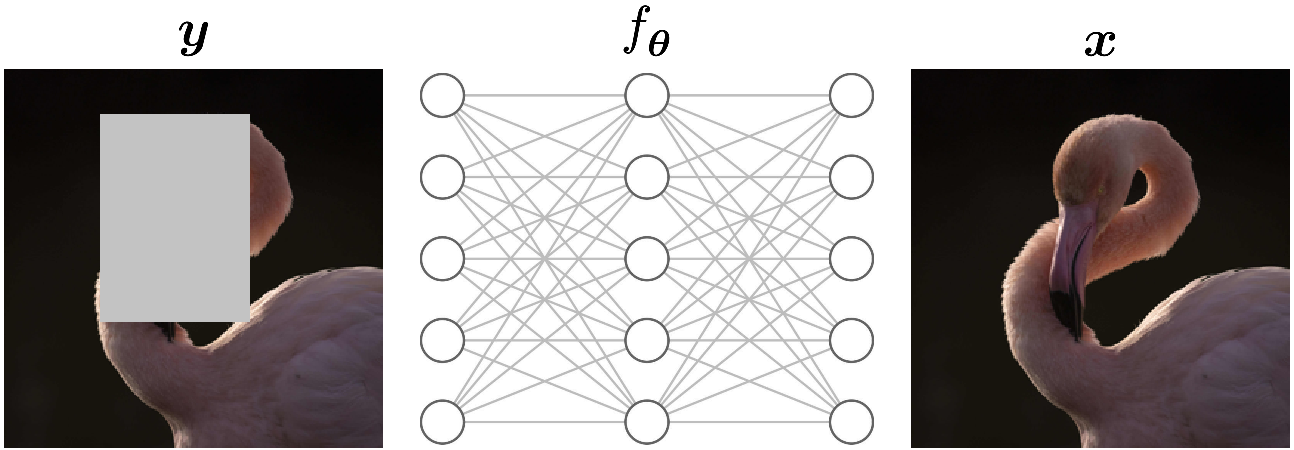
Inpainting

Deep learning approches

- Deep learning is now the state of the art.

Objective

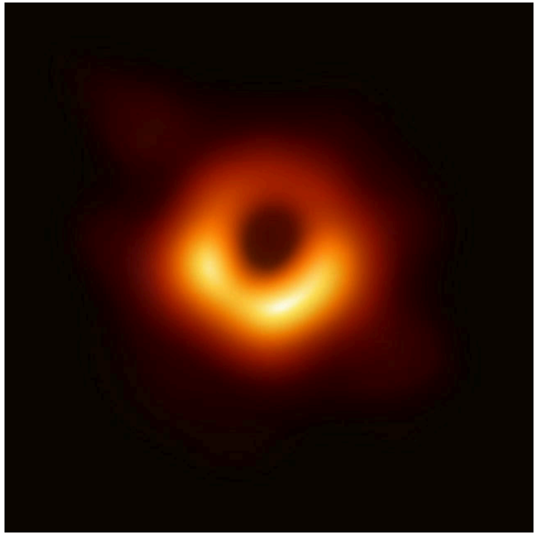
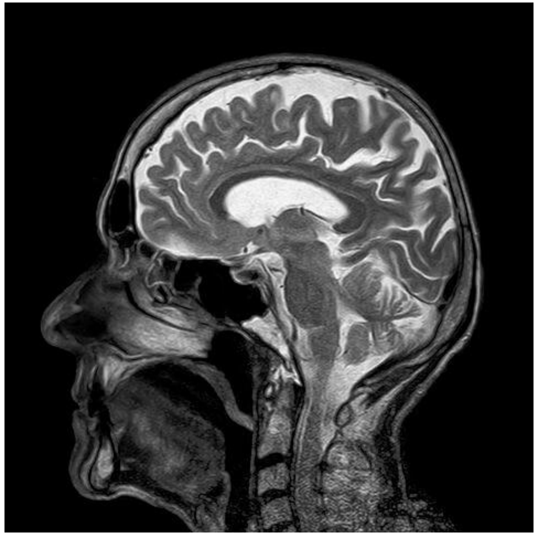
- **Learn** inverse forward operator of $\boldsymbol{A}$, $f_\theta(\boldsymbol{y}, \boldsymbol{A}) \approx \boldsymbol{x}$.
- **Using** a dataset $\{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_{i \in I}$ + a neural network.



$$\underset{\theta}{\operatorname{argmin}} \sum_{i \in I} \|f_\theta(\boldsymbol{y}_i, \boldsymbol{A}) - \boldsymbol{x}_i\|^2$$

Limits of the method

- Training and testing datasets can be very different.
- We need a large set of $\{\boldsymbol{x}_i\}_{i \in I}$ (ground-truth).

Astronomical and medical imaging.

Measurement consistency is not enough

Problem: What can be done with a dataset of measurements only?

$$\mathcal{D} = (\boldsymbol{y}_i)_{i \in I} \neq (\boldsymbol{x}_i, \boldsymbol{y}_i)_{i \in I}$$

Measurement consistency is not enough

Problem: What can be done with a dataset of measurements only?

$$\mathcal{D} = (\boldsymbol{y}_i)_{i \in I} \neq (\boldsymbol{x}_i, \boldsymbol{y}_i)_{i \in I}$$

Idea: Enforce measurement consistency or maximize likelihood

$$\min_f \mathbb{E}_{\boldsymbol{y}, \boldsymbol{A}} \left\{ \| \boldsymbol{A} f(\boldsymbol{y}, \boldsymbol{A}) - \boldsymbol{y} \|^2 \right\}$$

Measurement consistency is not enough

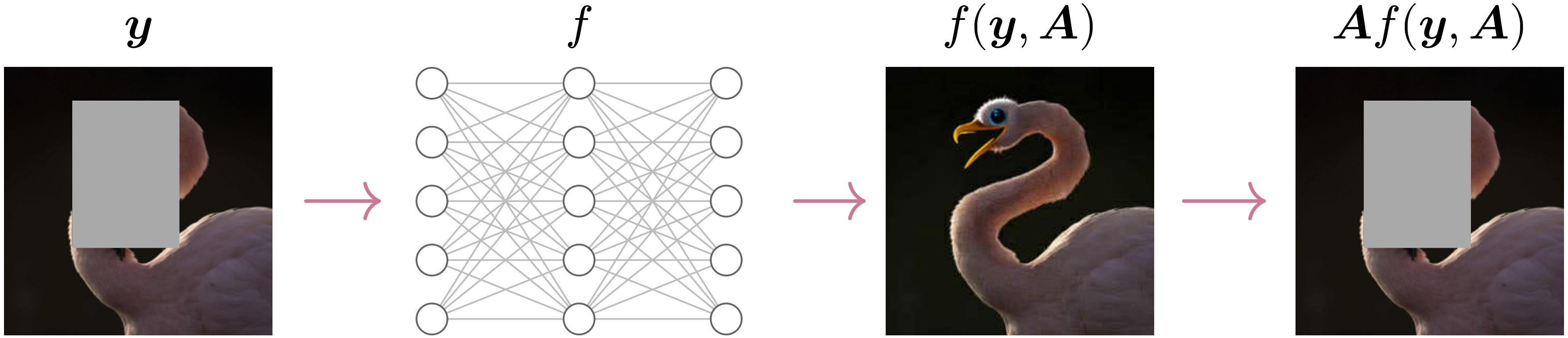
Problem: What can be done with a dataset of measurements only?

$$\mathcal{D} = (\boldsymbol{y}_i)_{i \in I} \neq (\boldsymbol{x}_i, \boldsymbol{y}_i)_{i \in I}$$

Idea: Enforce measurement consistency or maximize likelihood

$$\min_{f} \mathbb{E}_{\boldsymbol{y},\boldsymbol{A}} \left\{ \|\boldsymbol{A} f(\boldsymbol{y}, \boldsymbol{A}) - \boldsymbol{y}\|^2 \right\}$$

Limitation: It does not recover the information lost in the measurement process



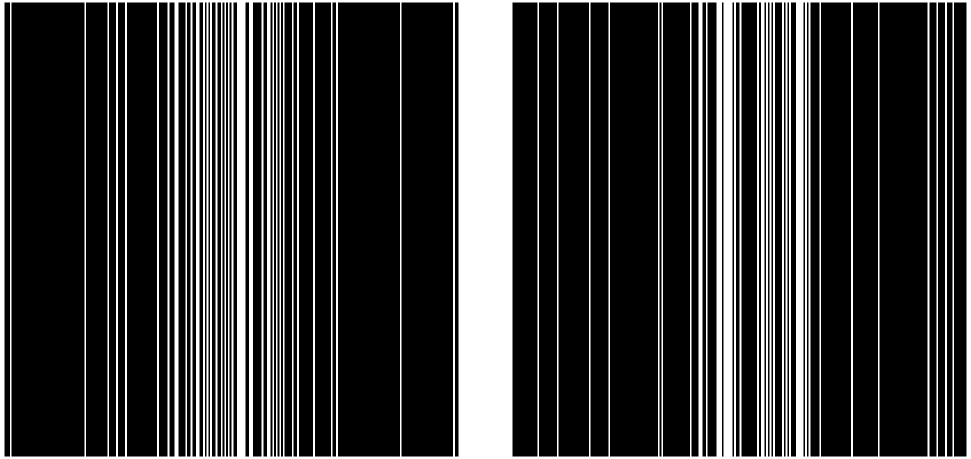
Multiple operators

- Leveraging **multiple** operators: $\boldsymbol{A} \sim p(\boldsymbol{A})$.
- **Differ** for each measurements $\boldsymbol{y} \sim p(\boldsymbol{y} \mid \boldsymbol{A}\boldsymbol{x})$.

Multiple operators

- Leveraging **multiple** operators: $\boldsymbol{A} \sim p(\boldsymbol{A})$.
- **Differ** for each measurements $\boldsymbol{y} \sim p(\boldsymbol{y} \mid \boldsymbol{A}\boldsymbol{x})$.

Inpainting

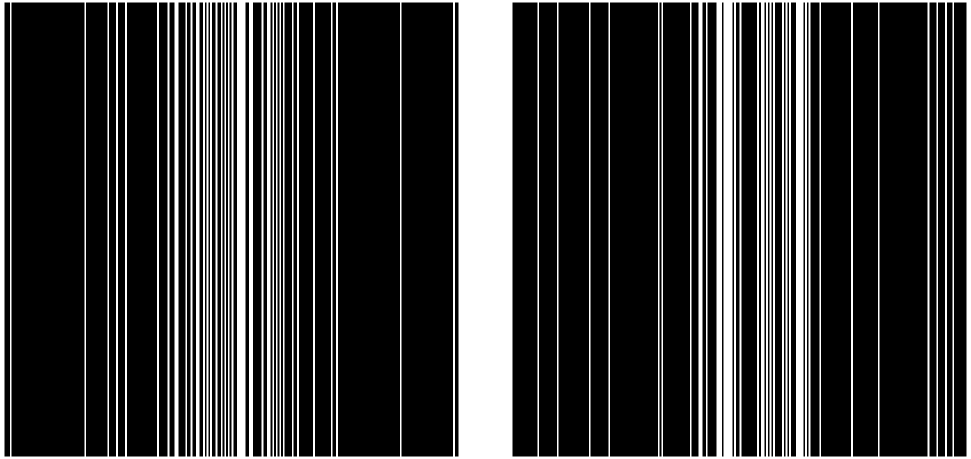


MRI

Multiple operators

- Leveraging **multiple** operators: $\boldsymbol{A} \sim p(\boldsymbol{A})$.
- **Differ** for each measurements $\boldsymbol{y} \sim p(\boldsymbol{y} \mid \boldsymbol{A}\boldsymbol{x})$.

Inpainting



MRI

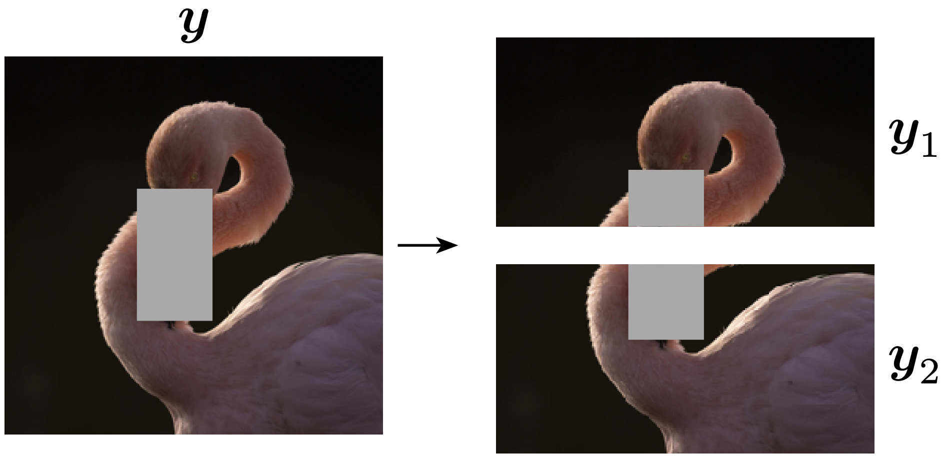
- **Measurement consistency** still learns the trivial reconstruction $f(\boldsymbol{y}, \boldsymbol{A}) = \boldsymbol{A}^{\dagger}\boldsymbol{y}$

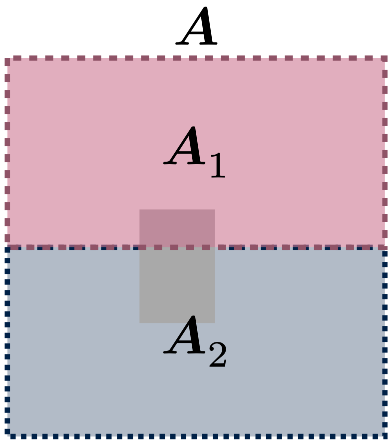
Measurements splitting

- Leveraging **multiple** operators: $\boldsymbol{y} \sim p(\boldsymbol{y} \mid \boldsymbol{A}\boldsymbol{x})$ with $\boldsymbol{A} \sim p(\boldsymbol{A})$.
- **Split** measurements $\boldsymbol{y} = [\boldsymbol{y}_1, \boldsymbol{y}_2]$ with **corresponding** operators $\boldsymbol{A} = [\boldsymbol{A}_1, \boldsymbol{A}_2]$.

$$f^* \in \underset{f}{\operatorname{argmin}} \, \mathbb{E}_{\boldsymbol{y},\boldsymbol{A}} \{ \mathcal{L}_{\text{SPLIT}}(\boldsymbol{y}, \boldsymbol{A}, f) \}$$

$$\text{with} \quad \mathcal{L}_{\text{SPLIT}}(\boldsymbol{y}, \boldsymbol{A}, f) = \mathbb{E}_{\boldsymbol{y}_1, \boldsymbol{A}_1 \mid \boldsymbol{y}, \boldsymbol{A}} \left[\| \boldsymbol{A} f(\boldsymbol{y}_1, \boldsymbol{A}_1) - \boldsymbol{y} \|^2 \right]$$



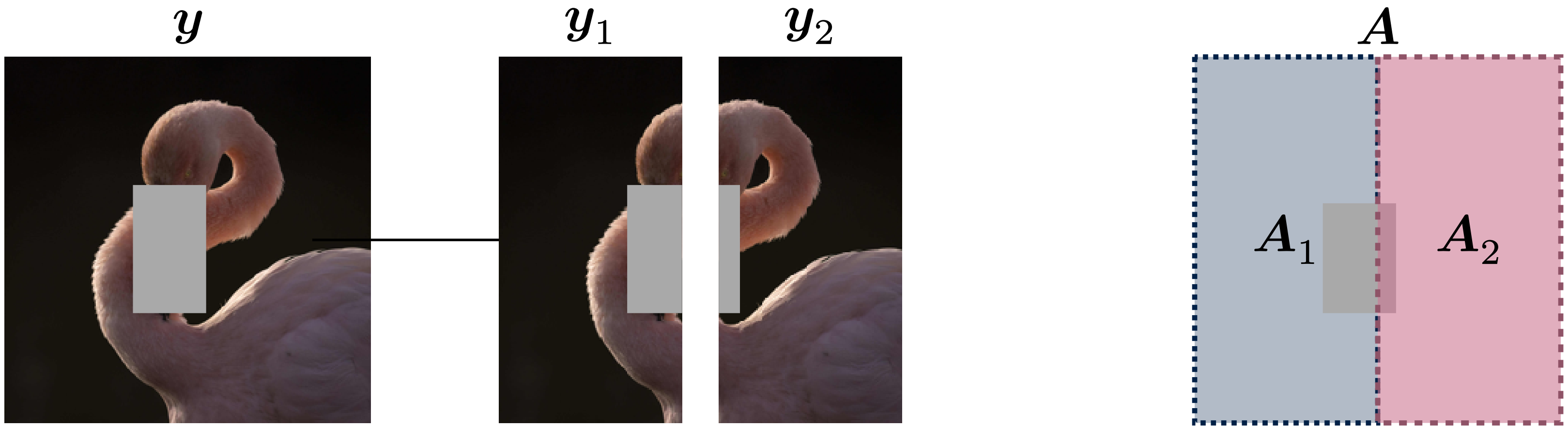


Measurements splitting

- Leveraging **multiple** operators: $\boldsymbol{y} \sim p(\boldsymbol{y} \mid \boldsymbol{A}\boldsymbol{x})$ with $\boldsymbol{A} \sim p(\boldsymbol{A})$.
- **Split** measurements $\boldsymbol{y} = [\boldsymbol{y}_1, \boldsymbol{y}_2]$ with **corresponding** operators $\boldsymbol{A} = [\boldsymbol{A}_1, \boldsymbol{A}_2]$.

$$f^* \in \operatorname*{argmin}_f \mathbb{E}_{\boldsymbol{y},\boldsymbol{A}} \{\mathcal{L}_{\mathrm{SPLIT}}(\boldsymbol{y}, \boldsymbol{A}, f)\}$$

$$\text{with} \quad \mathcal{L}_{\mathrm{SPLIT}}(\boldsymbol{y}, \boldsymbol{A}, f) = \mathbb{E}_{\boldsymbol{y}_1, \boldsymbol{A}_1 \mid \boldsymbol{y}, \boldsymbol{A}} \left[\| \boldsymbol{A} f(\boldsymbol{y}_1, \boldsymbol{A}_1) - \boldsymbol{y} \|^2 \right]$$



A visual interpretation of $\boldsymbol{Q}_{\boldsymbol{A}_1}$

Solution with $\boldsymbol{Q}_{\boldsymbol{A}_1} \triangleq \mathbb{E}_{\boldsymbol{A} \mid \boldsymbol{A}_1}\{\boldsymbol{A}^\top \boldsymbol{A}\}$ and v any function:

$$f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1} \underbrace{\mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}}_{\text{MMSE estimator}} + \underbrace{\left(\boldsymbol{I} - \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1}\right) v(\boldsymbol{y}_1)}_{\text{Nullspace Error}}$$

A visual interpretation of $\boldsymbol{Q}_{\boldsymbol{A}_1}$

Solution with $\boldsymbol{Q}_{\boldsymbol{A}_1} \triangleq \mathbb{E}_{\boldsymbol{A} \mid \boldsymbol{A}_1}\{\boldsymbol{A}^\top \boldsymbol{A}\}$ and v any function:

$$f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1} \underbrace{\mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}}_{\text{MMSE estimator}} + \underbrace{\left(\boldsymbol{I} - \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1}\right) v(\boldsymbol{y}_1)}_{\text{Nullspace Error}}$$

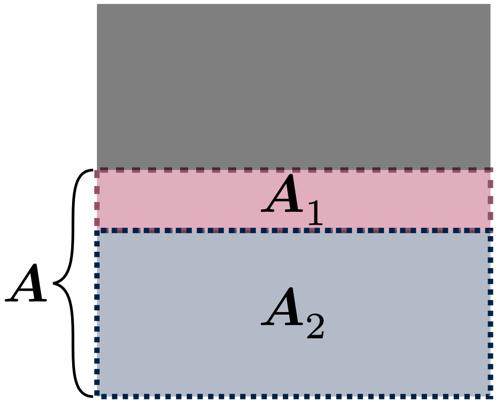
If $\boldsymbol{Q}_{\boldsymbol{A}_1}$ **invertible**: $f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}$

A visual interpretation of $\boldsymbol{Q}_{\boldsymbol{A}_1}$

Solution with $\boldsymbol{Q}_{\boldsymbol{A}_1} \triangleq \mathbb{E}_{\boldsymbol{A} \mid \boldsymbol{A}_1}\{\boldsymbol{A}^\top \boldsymbol{A}\}$ and v any function:

$$f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1} \underbrace{\mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}}_{\text{MMSE estimator}} + \underbrace{\left(\boldsymbol{I} - \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1}\right) v(\boldsymbol{y}_1)}_{\text{Nullspace Error}}$$

If $\boldsymbol{Q}_{\boldsymbol{A}_1}$ **invertible**: $f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}$

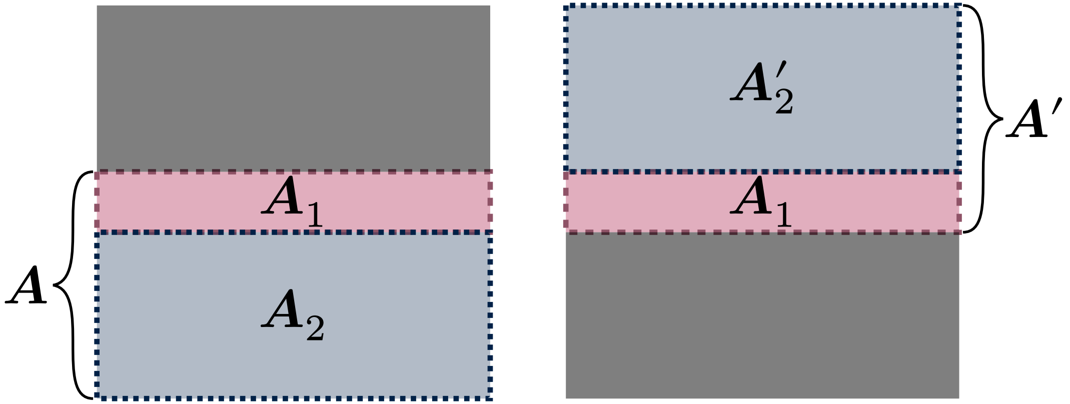


A visual interpretation of $\boldsymbol{Q}_{\boldsymbol{A}_1}$

Solution with $\boldsymbol{Q}_{\boldsymbol{A}_1} \triangleq \mathbb{E}_{\boldsymbol{A} \mid \boldsymbol{A}_1}\{\boldsymbol{A}^\top \boldsymbol{A}\}$ and v any function:

$$f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1} \underbrace{\mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}}_{\text{MMSE estimator}} + \underbrace{\left(\boldsymbol{I} - \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1}\right) v(\boldsymbol{y}_1)}_{\text{Nullspace Error}}$$

If $\boldsymbol{Q}_{\boldsymbol{A}_1}$ **invertible**: $f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}$

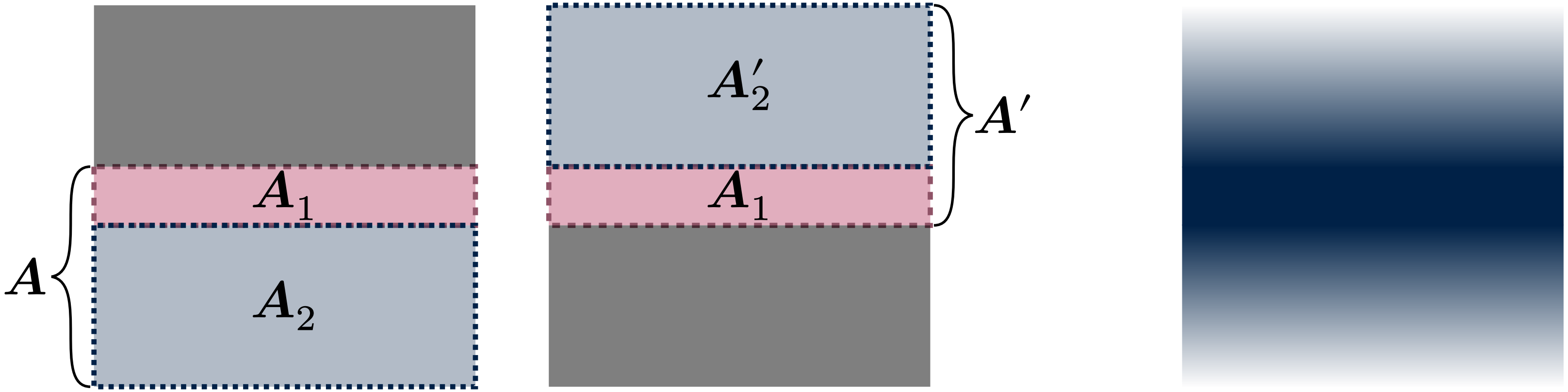


A visual interpretation of $\boldsymbol{Q}_{\boldsymbol{A}_1}$

Solution with $\boldsymbol{Q}_{\boldsymbol{A}_1} \triangleq \mathbb{E}_{\boldsymbol{A} \mid \boldsymbol{A}_1}\{\boldsymbol{A}^\top \boldsymbol{A}\}$ and v any function:

$$f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1} \underbrace{\mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}}_{\text{MMSE estimator}} + \underbrace{\left(\boldsymbol{I} - \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1}\right) v(\boldsymbol{y}_1)}_{\text{Nullspace Error}}$$

If $\boldsymbol{Q}_{\boldsymbol{A}_1}$ **invertible**: $\quad f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}$



$\mathbb{E}_{\boldsymbol{A} \mid \boldsymbol{A}_1}[\boldsymbol{A}^\top \boldsymbol{A}]$

Key ingredient: the invariances of image distributions

Can we learn if we have data from a **single operator** $\boldsymbol{A}$?

Insight: Information about the image distribution is needed

Key ingredient: the invariances of image distributions

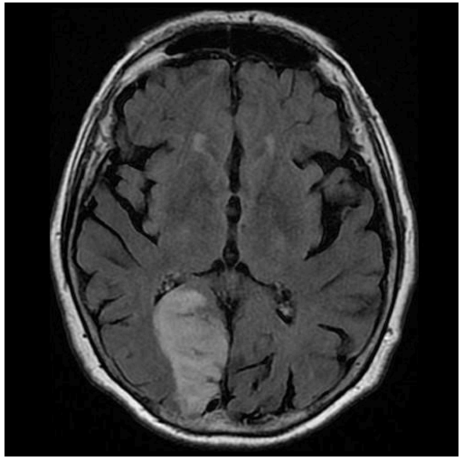
Can we learn if we have data from a **single operator $\boldsymbol{A}$**?

Insight: Information about the image distribution is needed

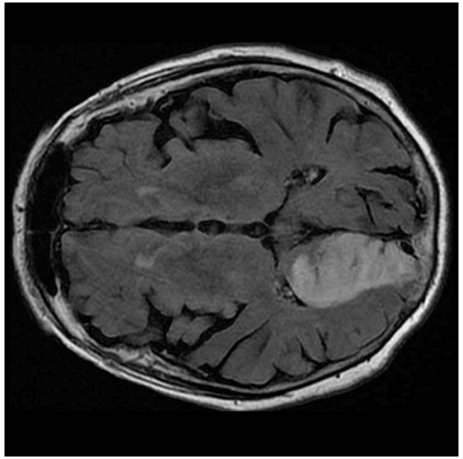
Assumption: The image distribution is invariant to certain transformations

$$p(\boldsymbol{T}_g \boldsymbol{x}) = p(\boldsymbol{x}), \quad \forall \boldsymbol{x}, g$$

$\boldsymbol{x}$



$\boldsymbol{T}_g \boldsymbol{x}$



- Equivariant Imaging (Chen, Tachella, Davies, 2021): **virtual operators** from invariance:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{T}_g \underbrace{\boldsymbol{T}_g^{-1}\boldsymbol{x}}_{\boldsymbol{x}' \in \mathcal{X}} = \boldsymbol{A}_g \boldsymbol{x}'$$

Equivariant imaging

- Equivariant Imaging (Chen, Tachella, Davies, 2021): **virtual operators** from invariance:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{T}_g \underbrace{\boldsymbol{T}_g^{-1}\boldsymbol{x}}_{\boldsymbol{x}' \in \mathcal{X}} = \boldsymbol{A}_g \boldsymbol{x}'$$

- This work: **split** measurements $\boldsymbol{y}$ for **different** operators $\boldsymbol{A}\boldsymbol{T}_g$:

$$\mathcal{L}_{\mathrm{ES}}(\boldsymbol{y}, \boldsymbol{A}, f) \triangleq \mathbb{E}_g \{ \mathcal{L}_{\mathrm{SPLIT}}(\boldsymbol{y}, \boldsymbol{A}\boldsymbol{T}_g, f) \}$$

A visual interpretation of $\boldsymbol{Q}_{\boldsymbol{A}_1}$

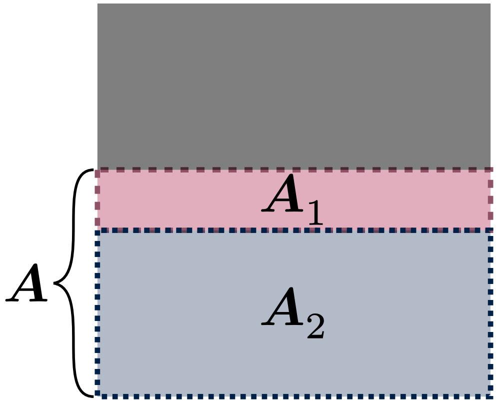
Solution with $\boldsymbol{Q}_{\boldsymbol{A}_1} \triangleq \mathbb{E}_{g \mid \boldsymbol{A}_1}\left\{\left(\boldsymbol{A}\boldsymbol{T}_g\right)^\top \boldsymbol{A}\boldsymbol{T}_g\right\}$:

$$f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1} \underbrace{\mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}}_{\text{MMSE estimator}} + \underbrace{\left(\boldsymbol{I} - \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1}\right) v(\boldsymbol{y}_1)}_{\text{Error}}$$

A visual interpretation of $\boldsymbol{Q}_{\boldsymbol{A}_1}$

Solution with $\boldsymbol{Q}_{\boldsymbol{A}_1} \triangleq \mathbb{E}_{g \mid \boldsymbol{A}_1}\left\{\left(\boldsymbol{A}\boldsymbol{T}_g\right)^\top \boldsymbol{A}\boldsymbol{T}_g\right\}$:

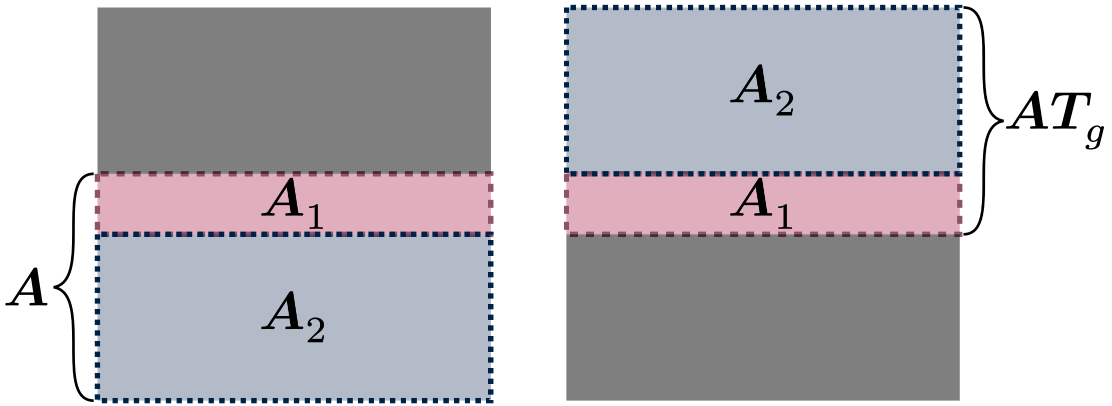
$$f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1} \underbrace{\mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}}_{\text{MMSE estimator}} + \underbrace{\left(\boldsymbol{I} - \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1}\right) v(\boldsymbol{y}_1)}_{\text{Error}}$$



A visual interpretation of $\boldsymbol{Q}_{\boldsymbol{A}_1}$

Solution with $\boldsymbol{Q}_{\boldsymbol{A}_1} \triangleq \mathbb{E}_{g \mid \boldsymbol{A}_1}\left\{(\boldsymbol{A}\boldsymbol{T}_g)^\top \boldsymbol{A}\boldsymbol{T}_g\right\}$:

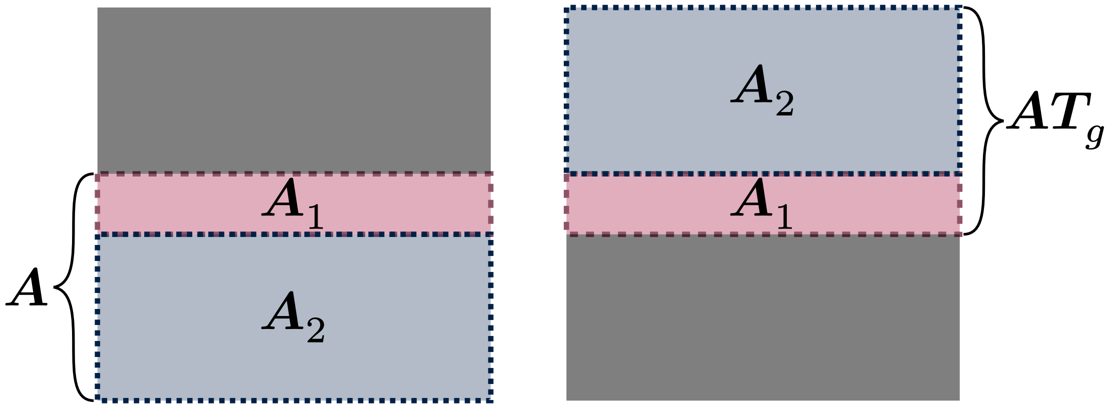
$$f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1} \underbrace{\mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}}_{\text{MMSE estimator}} + \underbrace{\left(\boldsymbol{I} - \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1}\right) v(\boldsymbol{y}_1)}_{\text{Error}}$$



A visual interpretation of $\boldsymbol{Q}_{\boldsymbol{A}_1}$

Solution with $\boldsymbol{Q}_{\boldsymbol{A}_1} \triangleq \mathbb{E}_{g \mid \boldsymbol{A}_1}\left\{\left(\boldsymbol{A}\boldsymbol{T}_g\right)^\top \boldsymbol{A}\boldsymbol{T}_g\right\}$:

$$f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1} \underbrace{\mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}}_{\text{MMSE estimator}} + \underbrace{\left(\boldsymbol{I} - \boldsymbol{Q}_{\boldsymbol{A}_1}^\dagger \boldsymbol{Q}_{\boldsymbol{A}_1}\right) v(\boldsymbol{y}_1)}_{\text{Error}}$$



$$\mathbb{E}_{g \mid \boldsymbol{A}_1}\left[\left(\boldsymbol{A}\boldsymbol{T}_g\right)^\top \boldsymbol{A}\boldsymbol{T}_g\right]$$

Inference

$$f^* \in \operatorname*{argmin}_f \mathbb{E}_{\boldsymbol{y}}\{\mathcal{L}_{\text{ES}}(\boldsymbol{y}, \boldsymbol{A}, f)\}$$

- Naive inference, sample a single split: $\overline{f}(\boldsymbol{y}, \boldsymbol{A}) = f^*(\boldsymbol{y}_1, \boldsymbol{A}_i)$
-
-

Inference

$$f^* \in \underset{f}{\operatorname{argmin}} \, \mathbb{E}_{\boldsymbol{y}} \{ \mathcal{L}_{\text{ES}}(\boldsymbol{y}, \boldsymbol{A}, f) \}$$

- Naive inference, sample a single split: $\overline{f}(\boldsymbol{y}, \boldsymbol{A}) = f^*(\boldsymbol{y}_1, \boldsymbol{A}_i)$
- Better idea, average splits: $\overline{f}(\boldsymbol{y}, \boldsymbol{A}) = \frac{1}{J} \sum_{j=1}^{J} f^* \left(\boldsymbol{y}_1^{(j)}, \boldsymbol{A}_i^{(j)} \right)$
-

Inference

$$f^* \in \operatorname*{argmin}_f \mathbb{E}_{\boldsymbol{y}}\{\mathcal{L}_{\text{ES}}(\boldsymbol{y}, \boldsymbol{A}, f)\}$$

- Naive inference, sample a single split: $\overline{f}(\boldsymbol{y}, \boldsymbol{A}) = f^*(\boldsymbol{y}_1, \boldsymbol{A}_i)$
- Better idea, average splits: $\overline{f}(\boldsymbol{y}, \boldsymbol{A}) = \frac{1}{J}\sum_{j=1}^{J} f^*\left(\boldsymbol{y}_1^{(j)}, \boldsymbol{A}_i^{(j)}\right)$
- Even better: if $\overline{\boldsymbol{Q}}_{\boldsymbol{A}} \triangleq \mathbb{E}_{\boldsymbol{A}_1 \mid \boldsymbol{A}}\left\{\boldsymbol{Q}_{\boldsymbol{A}_1}\right\}$ **is invertible**:

$$\begin{aligned}\overline{f}(\boldsymbol{y}, \boldsymbol{A}) &\triangleq \mathbb{E}_{\boldsymbol{y}_1, \boldsymbol{A}_1 \mid \boldsymbol{y}, \boldsymbol{A}}\left\{\overline{\boldsymbol{Q}}_{\boldsymbol{A}}^{-1} \boldsymbol{Q}_{\boldsymbol{A}_1} f^*(\boldsymbol{y}_1, \boldsymbol{A}_1)\right\} \\ &= \underbrace{\mathbb{E}_{\boldsymbol{y}_1, \boldsymbol{A}_1 \mid \boldsymbol{y}, \boldsymbol{A}}\left\{\overline{\boldsymbol{Q}}_{\boldsymbol{A}}^{-1} \boldsymbol{Q}_{\boldsymbol{A}_1} \mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}\right\}}_{\text{Convex combination of MMSE estimators}}.\end{aligned}$$

Computing the loss

Problem: we need to average over the group of transformations

- Solution 1: Compute **random transformations** per batch during training
- Solution 2: Train **equivariant reconstructor** f

Equivariant reconstructors

Definition: A reconstructor $f(\boldsymbol{y}, \boldsymbol{A})$ is equivariant if it respects the equivalence

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\varepsilon} \quad \Longleftrightarrow \quad \boldsymbol{y} = \left(\boldsymbol{A}\boldsymbol{T}_g\right)\left(\boldsymbol{T}_g^{-1}\boldsymbol{x}\right) + \boldsymbol{\varepsilon}$$

that is, if

$$f\left(\boldsymbol{y}, \boldsymbol{A}\boldsymbol{T}_g\right) = \boldsymbol{T}_g^{-1} f(\boldsymbol{y}, \boldsymbol{A}), \quad \forall \boldsymbol{y}, \boldsymbol{A}, g$$

Moreover, if $p(\boldsymbol{x})$ is invariant, then **MMSE and MAP** are equivariant reconstructors.

Artifact removal

$$f(\boldsymbol{y}, \boldsymbol{A}) = \phi(\boldsymbol{A}^{+}\boldsymbol{y}) \quad \text{with } \phi(\cdot) \text{ satisfying} \quad \phi\big(\boldsymbol{T}_g x\big) = \boldsymbol{T}_g \phi(x)$$

Artifact removal

$$f(\boldsymbol{y}, \boldsymbol{A}) = \phi(\boldsymbol{A}^{+}\boldsymbol{y}) \quad \text{with } \phi(\cdot) \text{ satisfying} \quad \phi\big(\boldsymbol{T}_g x\big) = \boldsymbol{T}_g \phi(x)$$

Unrolled architecture

$$x_{k+1} = \phi\left(\boldsymbol{x}_k - \gamma \nabla_{\boldsymbol{x}_k}\left\{\|\boldsymbol{A}\boldsymbol{x}_k - \boldsymbol{y}\|^2\right\}\right) \quad \text{with } \phi(\cdot) \text{ satisfying} \quad \phi\big(\boldsymbol{T}_g x\big) = \boldsymbol{T}_g \phi(x)$$

Artifact removal

$$f(\boldsymbol{y}, \boldsymbol{A}) = \phi(\boldsymbol{A}^{+}\boldsymbol{y}) \quad \text{with } \phi(\cdot) \text{ satisfying} \quad \phi\big(\boldsymbol{T}_g x\big) = \boldsymbol{T}_g \phi(x)$$

Unrolled architecture

$$x_{k+1} = \phi\Big(\boldsymbol{x}_k - \gamma \nabla_{\boldsymbol{x}_k}\Big\{\|\boldsymbol{A}\boldsymbol{x}_k - \boldsymbol{y}\|^2\Big\}\Big) \quad \text{with } \phi(\cdot) \text{ satisfying} \quad \phi\big(\boldsymbol{T}_g x\big) = \boldsymbol{T}_g \phi(x)$$

Reynolds averaging

$$f(\boldsymbol{y}, \boldsymbol{A}) = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \boldsymbol{T}_g \tilde{f}\big(\boldsymbol{y}, \boldsymbol{A}\boldsymbol{T}_g\big) \quad \text{with } \tilde{f}(\boldsymbol{y}, \boldsymbol{A}) \text{ any reconstructor}$$

Equivariance for faster compute

Theorem. If $f(\boldsymbol{y}, \boldsymbol{A})$ is an equivariant reconstructor, i.e.,

$$f(\boldsymbol{y}, \boldsymbol{A}\boldsymbol{T}_g) = \boldsymbol{T}_g^{-1} f(\boldsymbol{y}, \boldsymbol{A}), \quad \forall \boldsymbol{y}, \boldsymbol{A}, g,$$

then

$$\mathcal{L}_{\text{ES}}(\boldsymbol{y}, \boldsymbol{A}, f) = \mathcal{L}_{\text{SPLIT}}(\boldsymbol{y}, \boldsymbol{A}, f).$$

- If we use an equivariant architecture -> transformations are computed implicitly!

Noisy measurements

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\varepsilon} \quad \text{with} \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I)$$

- **Divide** the $\mathcal{L}_{\text{ES}}$ into 2 terms:

$$\mathcal{L}_{\text{ES}}(\boldsymbol{y}, \boldsymbol{A}, f) = \mathbb{E}_g \mathbb{E}_{\boldsymbol{y}_1, \boldsymbol{A}_1 \mid \boldsymbol{y}, \boldsymbol{A}\boldsymbol{T}_g} \Big\{ \underbrace{\|\boldsymbol{A}_1 f(\boldsymbol{y}_1, \boldsymbol{A}_1) - \boldsymbol{y}_1\|^2}_{\text{Measurement consistency (MC)}} + \underbrace{\|\boldsymbol{A}_2 f(\boldsymbol{y}_1, \boldsymbol{A}_1) - \boldsymbol{y}_2\|^2}_{\text{prediction accuracy}} \Big\}.$$

-
-

Noisy measurements

$$y = Ax + \varepsilon \quad \text{with} \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I)$$

- **Divide** the $\mathcal{L}_{\text{ES}}$ into 2 terms:

$$\mathcal{L}_{\text{ES}}(\boldsymbol{y}, \boldsymbol{A}, f) = \mathbb{E}_g \mathbb{E}_{\boldsymbol{y}_1, \boldsymbol{A}_1 \mid \boldsymbol{y}, \boldsymbol{A}\boldsymbol{T}_g} \Big\{ \underbrace{\|\boldsymbol{A}_1 f(\boldsymbol{y}_1, \boldsymbol{A}_1) - \boldsymbol{y}_1\|^2}_{\text{Measurement consistency (MC)}} + \underbrace{\|\boldsymbol{A}_2 f(\boldsymbol{y}_1, \boldsymbol{A}_1) - \boldsymbol{y}_2\|^2}_{\text{prediction accuracy}} \Big\}.$$

- **Replace** MC by a self-supervised **denoising** loss with $\mathcal{N}(0, \sigma^2 I)$:

$$\mathbb{E}_g \mathbb{E}_{\boldsymbol{y}_1, \boldsymbol{A}_1, \boldsymbol{\omega} \mid \boldsymbol{y}, \boldsymbol{A}\boldsymbol{T}_g} \Bigg\{ \underbrace{\left\| \boldsymbol{A}_1 f(\boldsymbol{y}_1 + \alpha\boldsymbol{\omega}, \boldsymbol{A}_1) - \left(\boldsymbol{y}_1 - \frac{\boldsymbol{\omega}}{\alpha} \right) \right\|^2}_{\text{R2R (denoising) loss}} + \|\boldsymbol{A}_2 f(\boldsymbol{y}_1 + \alpha\boldsymbol{\omega}, \boldsymbol{A}_1) - \boldsymbol{y}_2\|^2 \Bigg\}.$$

-

Noisy measurements

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\varepsilon} \quad \text{with} \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I)$$

- **Divide** the $\mathcal{L}_{\text{ES}}$ into 2 terms:

$$\mathcal{L}_{\text{ES}}(\boldsymbol{y}, \boldsymbol{A}, f) = \mathbb{E}_g \mathbb{E}_{\boldsymbol{y}_1, \boldsymbol{A}_1 \mid \boldsymbol{y}, \boldsymbol{A}\boldsymbol{T}_g} \Big\{ \underbrace{\|\boldsymbol{A}_1 f(\boldsymbol{y}_1, \boldsymbol{A}_1) - \boldsymbol{y}_1\|^2}_{\text{Measurement consistency (MC)}} + \underbrace{\|\boldsymbol{A}_2 f(\boldsymbol{y}_1, \boldsymbol{A}_1) - \boldsymbol{y}_2\|^2}_{\text{prediction accuracy}} \Big\}.$$

- **Replace** MC by a self-supervised **denoising** loss with $\mathcal{N}(0, \sigma^2 I)$:

$$\mathbb{E}_g \mathbb{E}_{\boldsymbol{y}_1, \boldsymbol{A}_1, \boldsymbol{\omega} \mid \boldsymbol{y}, \boldsymbol{A}\boldsymbol{T}_g} \Big\{ \underbrace{\left\| \boldsymbol{A}_1 f(\boldsymbol{y}_1 + \alpha\boldsymbol{\omega}, \boldsymbol{A}_1) - \left(\boldsymbol{y}_1 - \frac{\boldsymbol{\omega}}{\alpha} \right) \right\|^2}_{\text{R2R (denoising) loss}} + \|\boldsymbol{A}_2 f(\boldsymbol{y}_1 + \alpha\boldsymbol{\omega}, \boldsymbol{A}_1) - \boldsymbol{y}_2\|^2 \Big\}.$$

- $\boldsymbol{Q}_{\boldsymbol{A}_1}$ invertible $\Rightarrow f^*(\boldsymbol{y}_1, \boldsymbol{A}_1) = \mathbb{E}_{\boldsymbol{x} \mid \boldsymbol{y}_1, \boldsymbol{A}_1}\{\boldsymbol{x}\}$.

Overview of the experiments

Modality	**Network arch. (UNet)**	**Dataset**
Compressive sensing	Shift-Equivariant (Unrolled)	MNIST
Image inpainting	Shift-Equivariant	DIV2K
MRI	Rotoflip-Equivariant (Unrolled)	FastMRI
CT	Rotoflip-Equivariant (Unrolled)	LIDC-IDRI

- We compare with **Equivariant Imaging (EI)** (Chen, Davies & T., CVPR'22), which was the only self-supervised alternative in the single operator setting.

ES is close to the supervised gold standard

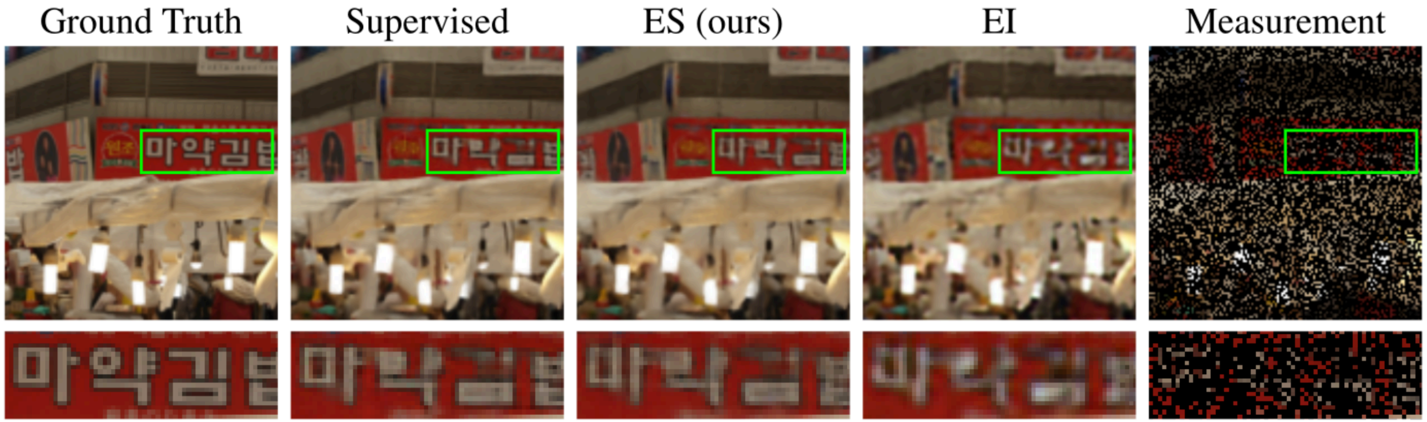
Reconstruction performance (PSNR in dB).

Method	Inpainting 30% rate	MRI ×8 accel.	CT 50 views
Supervised	28.5	28.8	34.0
ES	27.5	28.3	32.6

ES is close to the supervised gold standard

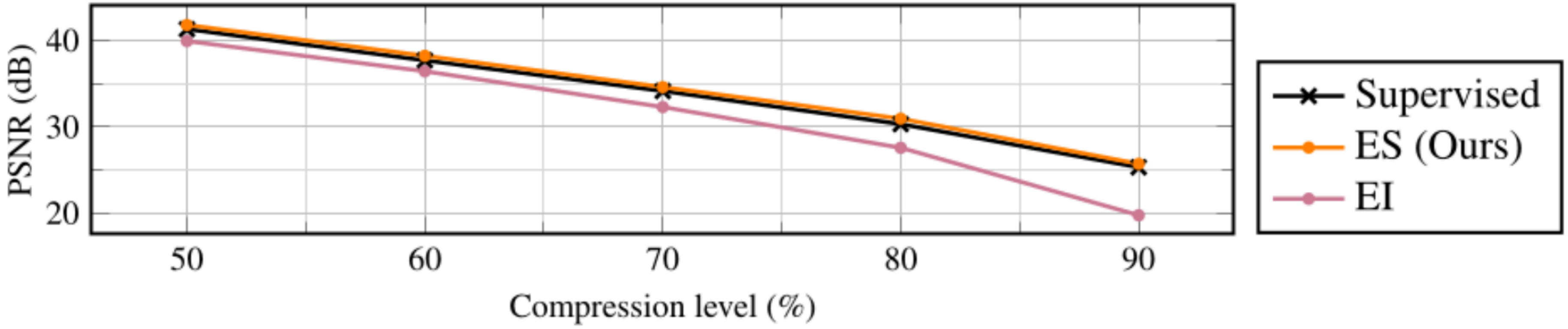
Reconstruction performance (PSNR in dB).

Method	Inpainting 30% rate	MRI ×8 accel.	CT 50 views
Supervised	28.5	28.8	34.0
ES	27.5	28.3	32.6



,

ES vs EI for highly compressed measures



Compressive sensing results.

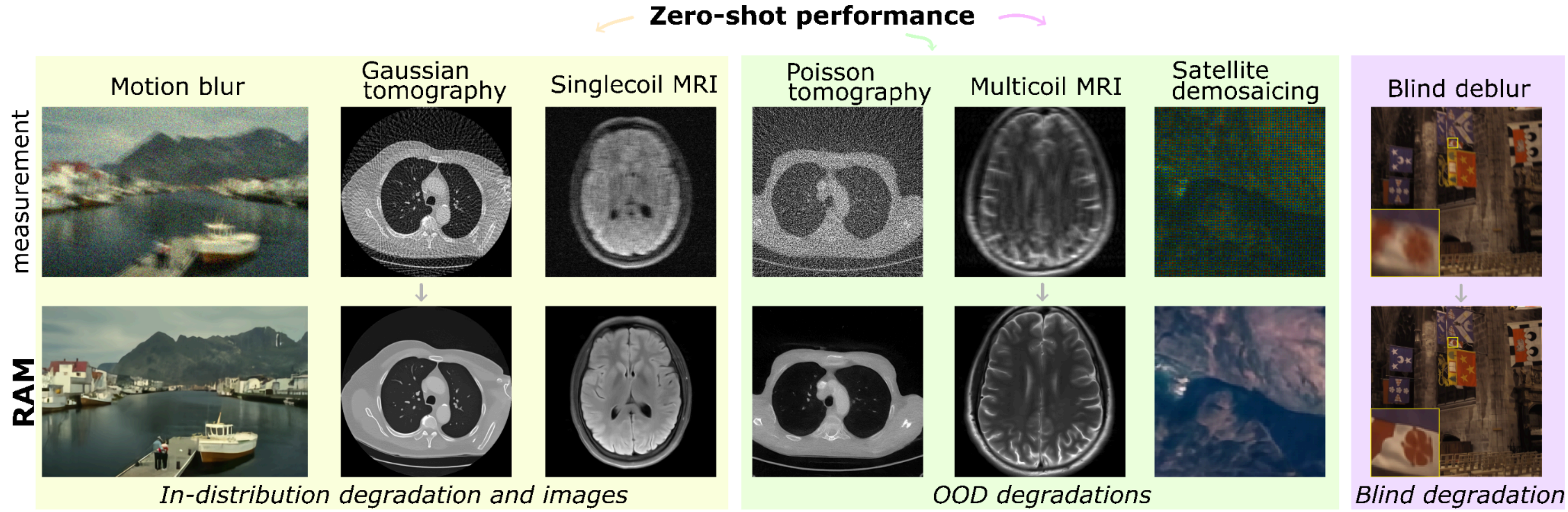
Empirical benefit of equivariance

Impact of equivariance (best in bold).

	Inpainting 30% rate		MRI ×8 accel.	
Equiv. arch.	PSNR ↑	EQUIV ↑	PSNR ↑	EQUIV ↑
✅	**27.5**	**27.5**	**28.5**	**31.5**
❌	27.2	26.5	28.2	27.3

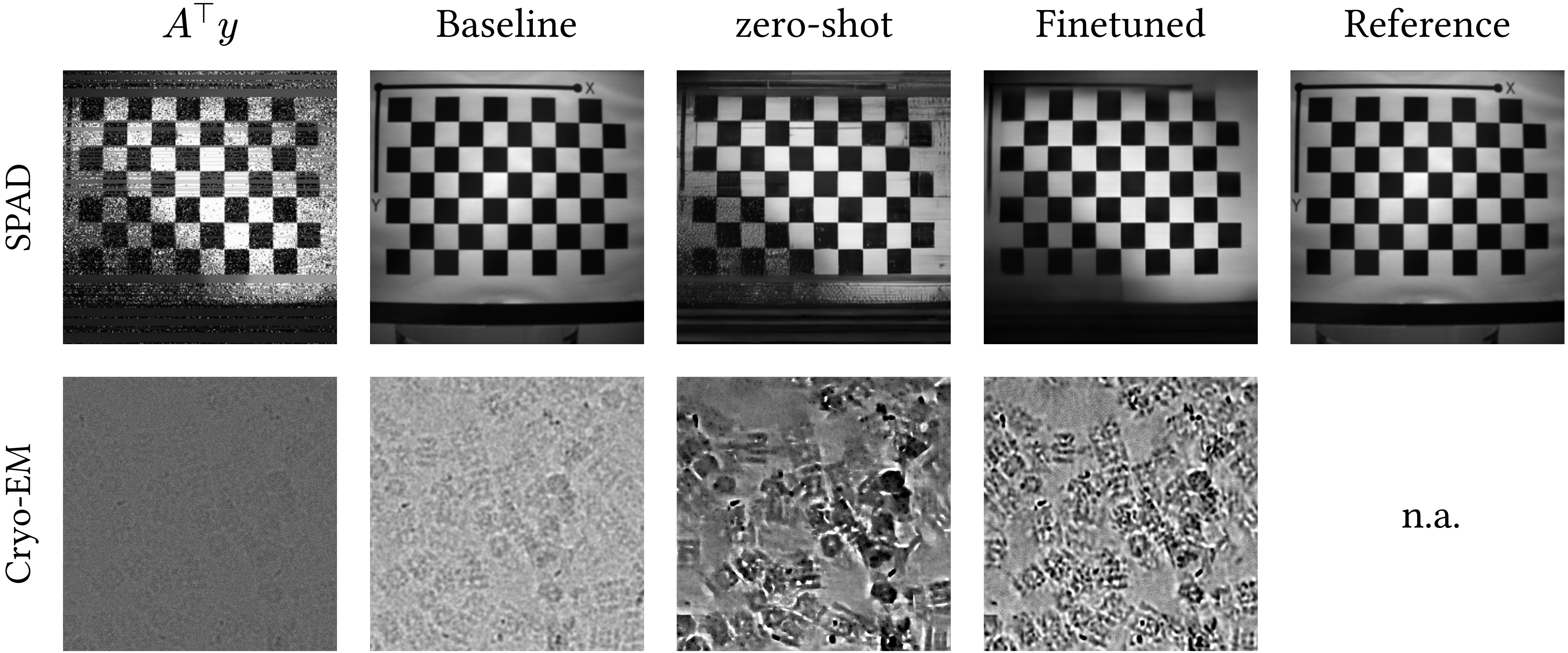
Fine-tuning foundation models

- Self-supervised losses can be used to fine-tune pretrained models!
- For example, our recent Reconstruct Anything Model (ICLR'26):



Fine-tuning foundation models

- Finetuning with 9 images (SPAD), 4 large images (cryoEM).



Conclusion

- Self-supervised loss for learning from a **single incomplete** operator
- We show that **equivariant architectures** can enable self-supervised learning
- Theoretical guarantees on the learned estimator
- Self-supervision + foundation models
- **Bonus:** Precise definition of equivariance in inverse problems

- "Self-Supervised Learning from Noisy and Incomplete Data", Foundations & Trends in SP
- Three-hour tutorial available in YouTube https://youtu.be/gf-WCHXAdfk?/

Foundations and Trends® in Signal Processing

Self-Supervised Learning from Noisy and Incomplete Data

Towards knowledge discovery with AI

Suggested Citation: Julián Tachella and Mike Davies (2025), "Self-Supervised Learning from Noisy and Incomplete Data", Foundations and Trends® in Signal Processing: Vol. xx, No. xx, pp 1–xx. DOI: 10.1561/XXXXXXXXX.

Julián Tachella
CNRS, ENS Lyon
julian.tachella@cnrs.fr

Mike Davies
University of Edinburgh
mike.davies@ed.ac.uk

Code & acknowledgment

-**Code:** https://github.com/vsechaud/Equivariant-Splitting

-**MRI**: https://github.com/Andrewwango/ssibench

-**Photographs of flamingos:** Émile Sechaud

DeepInverse

- Open-source library for computational imaging
- State-of-the-art reconstructors (PnP, diffusion, self-supervised, unrolling, etc)
- Annual hackathons and events