Latent Space Inference via Paired Autoencoders

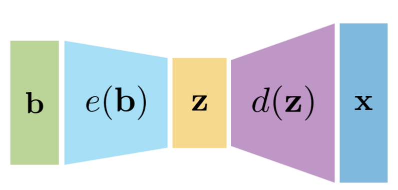
Emma Hart, Julianne Chung, Matthias Chung, Bas Peters
Department of Mathematics, Emory University

Background and Context

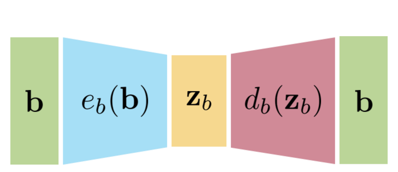
- **Inverse problems** involve determining unknown causes or parameters of a system based on observed outcomes

$$A(\mathbf{x}) + \boldsymbol{\epsilon} = \mathbf{b}$$

  - Forward process $A : \mathbb{R}^n \to \mathbb{R}^m$
  - Noise $\boldsymbol{\epsilon} \in \mathbb{R}^m$
  - Parameters $\mathbf{x} \in \mathbb{R}^n$
  - Observations $\mathbf{b} \in \mathbb{R}^m$
- **Machine learning** has been used to address many challenges in ill-posed and large-scale inverse problems, including full inversion, regularization, uncertainty quantification, and more
- **Encoder-Decoder networks** are one popular choice in many learning tasks, and can find a map directly from $\mathbf{b}$ to $\mathbf{x}$



  - $e : \mathbb{R}^m \to \mathbb{R}^r$ maps input $\mathbf{b}$ to latent variable $\mathbf{z}$
  - $d : \mathbb{R}^r \to \mathbb{R}^n$ maps from $\mathbf{z}$ to target $\mathbf{x}$
  - Find network parameters to minimize
$$\frac{1}{J}\sum_{j=1}^{J} \|d(e(\mathbf{b}_j)) - \mathbf{x}_j\| \text{ with samples } \{(\mathbf{b}_j, \mathbf{x}_j)\}_{j=1}^{J}$$
- **Autoencoder networks** are one kind of encoder-decoder network that map an input to itself, primarily used in denoising and dimensionality reduction applications

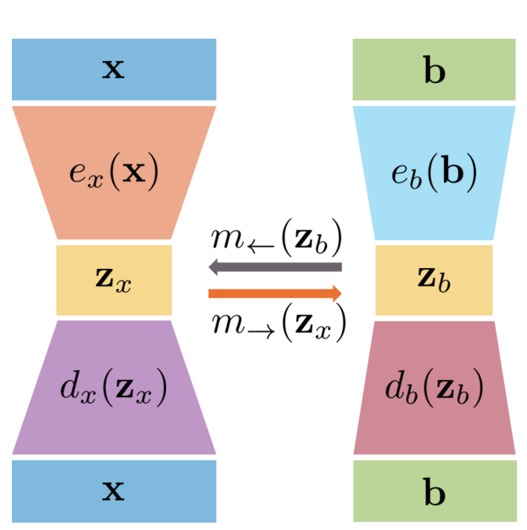


  - $e_b : \mathbb{R}^m \to \mathbb{R}^r$ maps input $\mathbf{b}$ to $\mathbf{z}_b$
  - $d_b : \mathbb{R}^r \to \mathbb{R}^m$ maps from $\mathbf{z}_b$ to same input $\mathbf{b}$
  - Find network parameters to minimize
$$\frac{1}{J'}\sum_{j=1}^{J'} \|d_b(e_b(\mathbf{b}_j)) - \mathbf{b}_j\| \text{ with samples } \{(\mathbf{b}_j)\}_{j=1}^{J'}$$
- **One key drawback** of many neural networks is that, when presented with discrepant or incomplete inputs $\mathbf{b}_{\text{sub}} = P(\mathbf{b})$, previously trained models on full data $\mathbf{b}$ become intractable or ineffective, due to the model's performance being highly sensitive to the characteristics of the training data

Paired Autoencoders

- **Aim** to leverage learned latent representations with **paired autoencoders** (PAIR) discussed in [1,2]

Key Ideas:
- *Pair two autoencoders*, one for the parameter space and one for the observation space by *connecting their latent spaces*
- Resulting network balances advantages of *end-to-end* and *representation learning*, admitting a *regularized inverse surrogate*, a *forward model surrogate*, and other *informative metrics*



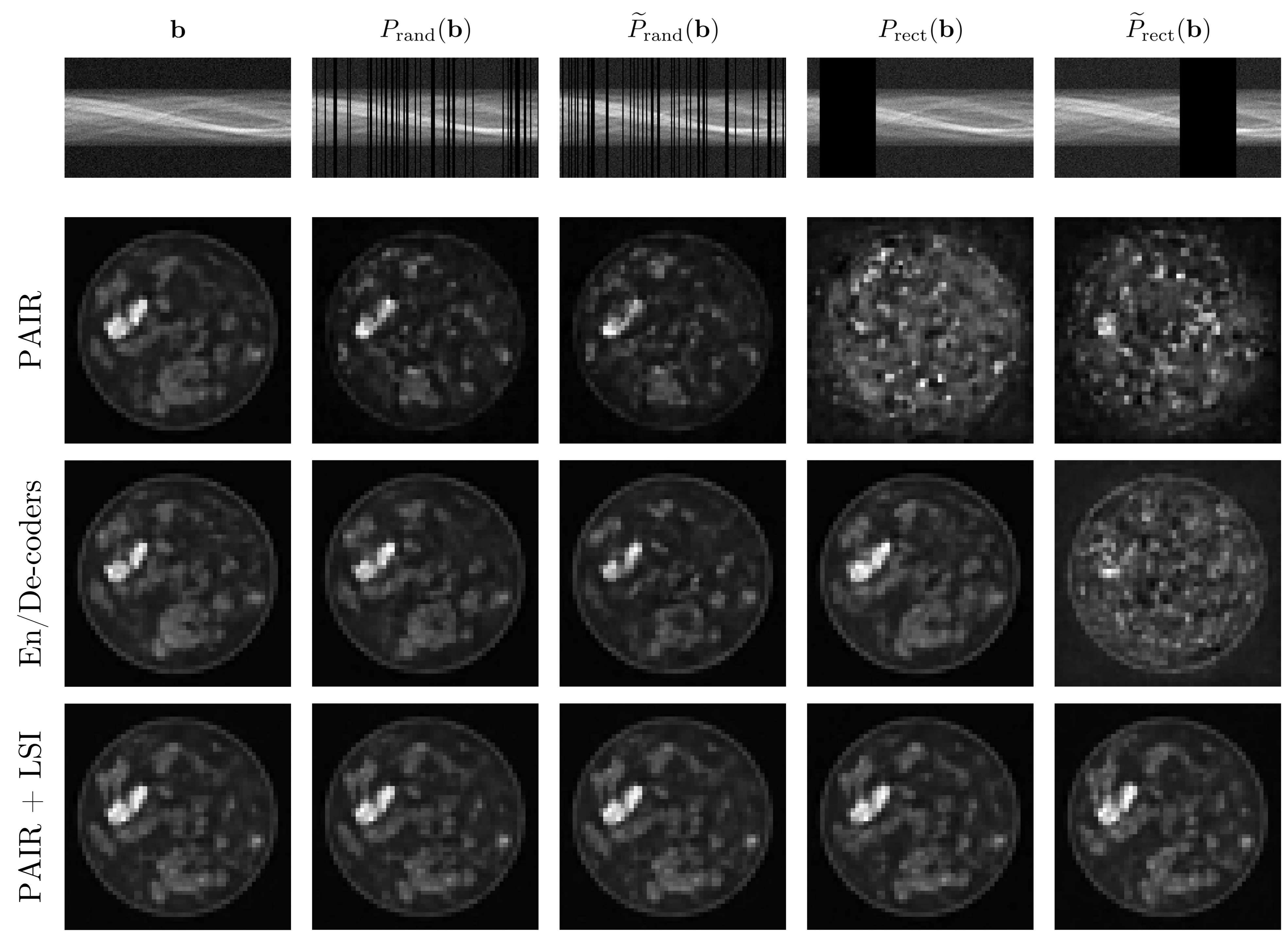
- Find network parameters with $\{(\mathbf{b}_j, \mathbf{x}_j)\}_{j=1}^{J}$ to minimize
$$\frac{1}{J}\sum_{j=1}^{J}\big(\|d_b(e_b(\mathbf{b}_j)) - \mathbf{b}_j\| + \|d_x(e_x(\mathbf{x}_j)) - \mathbf{x}_j\| + \|d_x(m_{\leftarrow}(e_b(\mathbf{b}_j))) - \mathbf{x}_j\| + \|d_b(m_{\rightarrow}(e_x(\mathbf{x}_j))) - \mathbf{b}_j\|\big)$$
- Low dimensional, informative latent spaces also offer ideal spaces for efficient optimization, where corruption or incompleteness in observations can be addressed through **latent space inference**

Latent Space Inference with Computed Tomography

- Let $\mathbf{b}_{\text{sub}} = P(\mathbf{b})$, where P is an operation that introduces errors or data gaps in $\mathbf{b}_{\text{sub}}$
- Latent Space Inference (LSI) addresses corruption in $\mathbf{b}_{\text{sub}}$ by reconstructing:

$$\hat{\mathbf{z}}_b \in \underset{\mathbf{z}_b}{\arg\min}\, \|(P \circ d_b)(\mathbf{z}_b) - \mathbf{b}_{\text{sub}}\| \quad \text{with } \hat{\mathbf{x}} = (d_x \circ m_{\leftarrow})(\hat{\mathbf{z}}_b)$$

- We demonstrate LSI with the 2DeteCT dataset of trail mix, curated for machine learning training in [3], and we compare our proposed PAIR + LSI method to:
  - PAIR alone, with $\mathbf{x}_{\text{pred}} = d_x(m_{\leftarrow}(e_b(\mathbf{b}_{\text{sub}})))$
  - End-to-end encoder-decoder networks, trained with samples $\{\mathbf{b}_{\text{sub},j}, \mathbf{x}_j\}_{j=1}^{J}$
- We consider two types of missing angle scenarios: one in which randomly selected angles across the full range are missing (P_{rand}), and one in which a block of angles are missing (P_{rect})
- Tildes ($\widetilde{P}_{\text{rand}}$, $\widetilde{P}_{\text{rect}}$) denote that the specifically selected missing angles are different than those that were missing in end-to-end encoder/decoder training data



- When full observations $\mathbf{b}$ are available, the PAIR method alone outperforms the end-to-end encoder-decoder network and PAIR + LSI
- The benefit of PAIR + LSI is evident when data are missing, most significantly when missing input angles are different than missing angles in encoder/decoder training

References

[1] M. Chung, E. Hart, J. Chung, B. Peters, and E. Haber. "Paired Autoencoders for Likelihood-Free Estimation in inverse problems". In: *Machine Learning: Science and Technology 5.4* (2024)

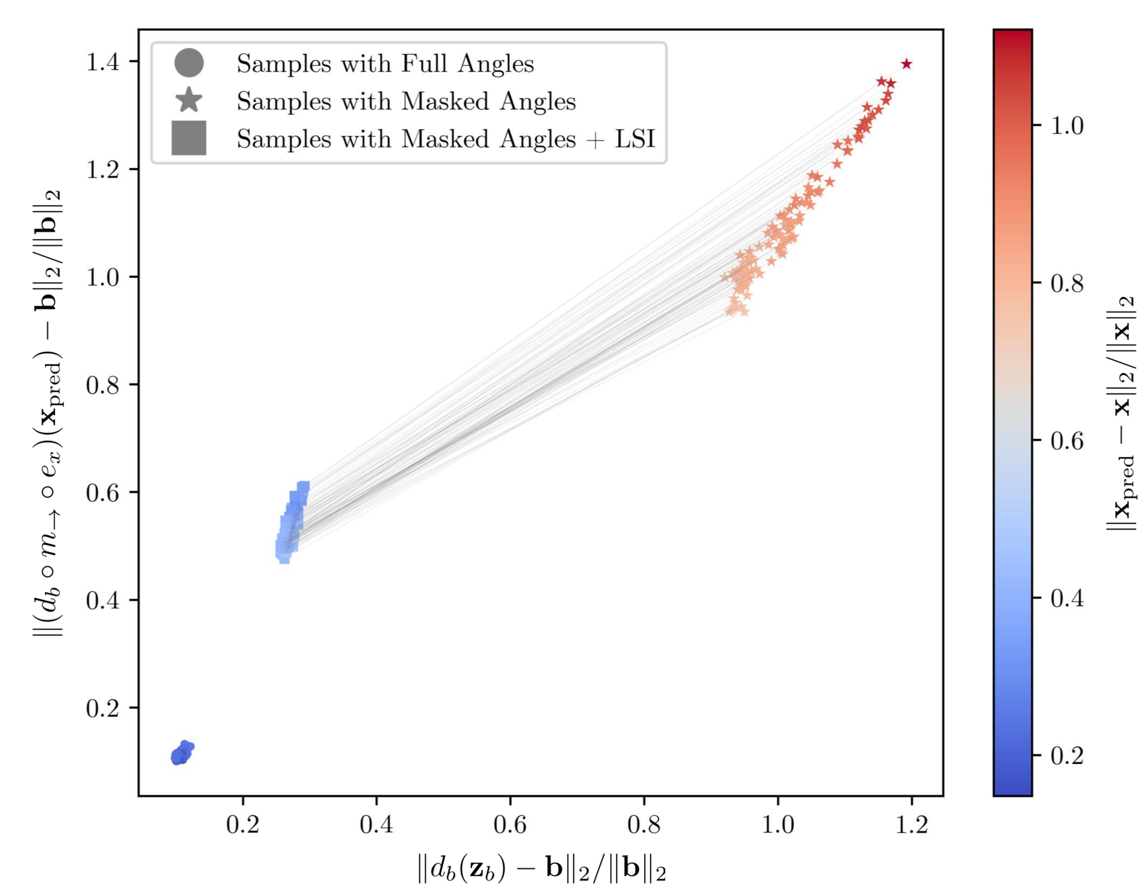
[2] E. Hart, J. Chung, and M. Chung. "A Paired Autoencoder Framework for Inverse Problems via Bayes Risk Minimization". In: *arXiv preprint arXiv:2501.14636* (2025)

[3] M. B. Kiss, S. B. Coban, K. J. Batenburg, T. van Leeuwen, and F. Lucka. "2DeteCT - A large2D expandable, trainable, experimental Computed Tomography dataset for machine learning". In: *Scientific Data 10.576* (2023)

Out of Distribution Detection

- The PAIR framework offers several cheaply computable metrics that can serve as proxies to indicate reconstruction quality, and can also be used to assess whether a new observation lies within distribution of training data, see [2]
- If the PAIR metrics for the new sample are similar to the metrics observed for the training samples, then we have some indication we can trust the prediction from the PAIR network, and if not, further investigation and refinement (for example, PAIR + LSI) may be required
- We use two of these metrics—the relative residual estimate and relative autoencoded data difference—to illustrate the ability to detect corrupted data (in this case, missing angles) and to improve reconstructions using PAIR + LSI

PAIR Metrics for Out of Distribution Detection



- We observe three well-separated clusters in PAIR metrics, and although it does not achieve the same accuracy as in the full angle case, we can see that PAIR + LSI reduces reconstruction errors associated with missing angles

Conclusions

- By exploiting a paired autoencoder framework and performing LSI in the latent space, more accurate reconstructions can be obtained for problems with missing observational data than PAIR or comparable end-to-end methods alone
- Future work includes exploring the use of LSI + PAIR for optimal experimental design (e.g., identifying the important angles to take measurements, or sensors for sensor placement)
- Other future work involves investigating LSI for timely solves in hyperparameter estimation or in scenarios with the forward model or noise model is different from assumptions

Acknowledgements

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Department of Energy Computational Science Graduate Fellowship under Award Number DE-SC0024386.