HERIOT WATT UNIVERSITY

Challenges in Imaging and Sensing in photon-starved regimes

Stephen McLaughlin

School of Engineering and Physical Sciences, Heriot-Watt University

Toulouse 2021

LEADERS IN IDEAS AND SOLUTIONS

EPSRC Engineering and Physical Sciences Research Council

[dstl]

UDRC

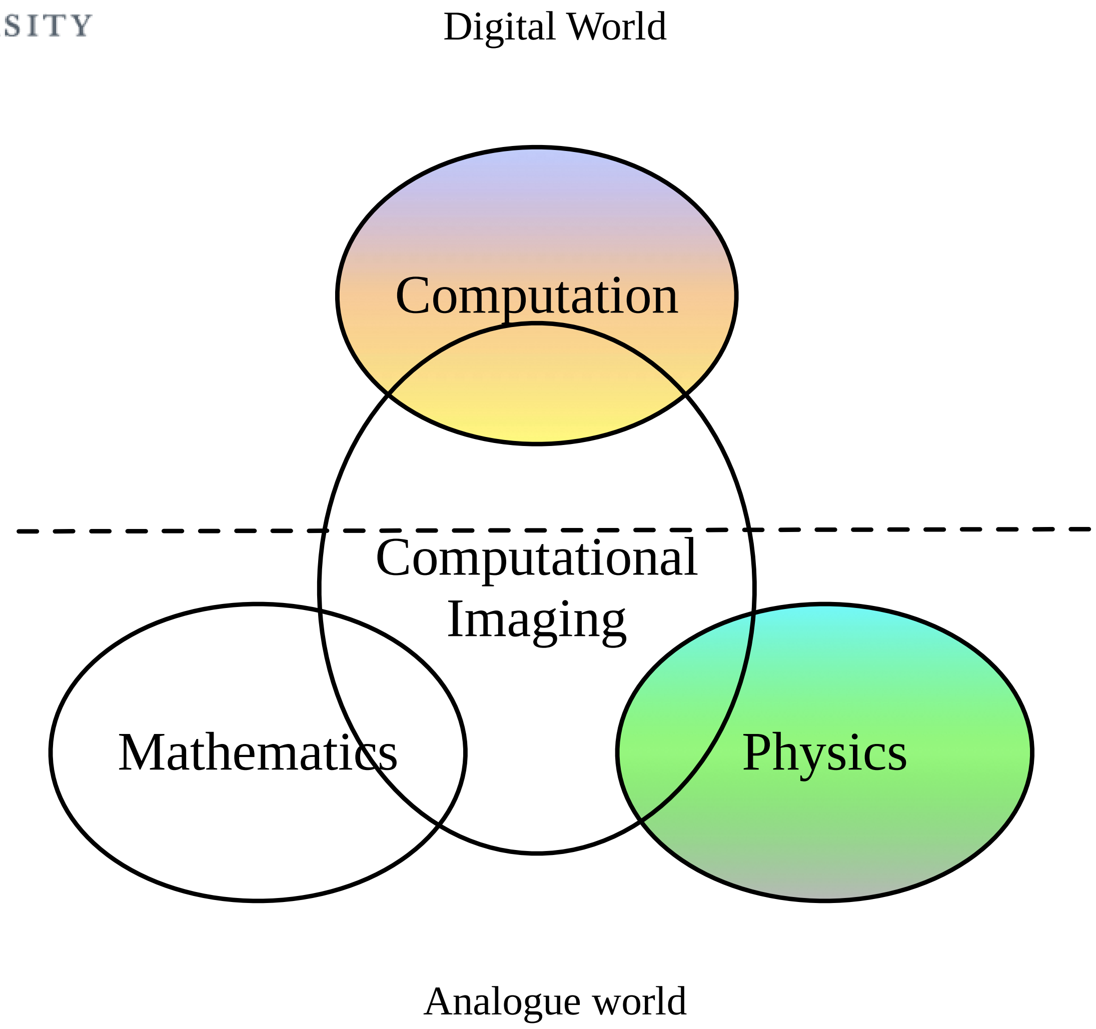
UK NATIONAL QUANTUM TECHNOLOGIES PROGRAMME

Royal Academy of Engineering

Acknowledgements

- Collaborators without whom it would not have been possible…
  - Signal Processing
    - **Yoann Altmann**, Julian Tachella, Quentin Legros, Abdullah Aziz, Mike Davies, Jean-Yves Tourneret, Abde Halimi, Vivek Goyal, Al Hero
  - Experimental
    - Gerald Buller, Aongus McCarthy, Aurore Maccarone, Rachael Tobin, Angela di Fulvio

Computational Imaging and Sensing



Imaging and sensing technologies are increasingly sophisticated with unprecedented levels of sensitivity and resolution achievable and at the same time sensors are reaching saturation levels

Single Photon Avalanche Diodes

- Single-Photon Avalanche Diode (SPAD) defines a class of photodetectors able to detect low intensity signals (down to the single photon) and to signal the time of the photon arrival with high temporal resolution (few tens of picoseconds).

- Latest Sony Chip available March 2022 IMX459 approximately 600 x 200 pixels (pixel size approx. $10\mu m$) for around 120 Euros

- This enables high-precision distance measuring at 15-centimeter range resolutions up to a distance of around 300 meters

Applications

- Defence

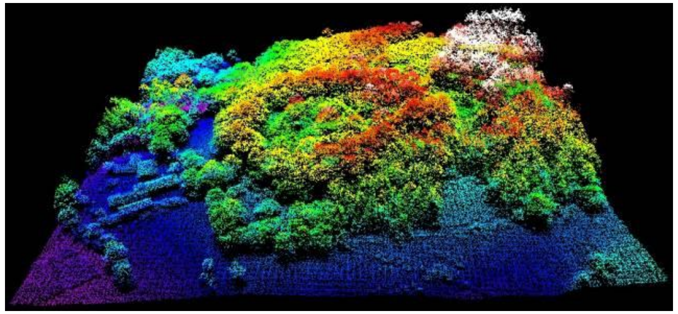
Long-range target identification

- Oil & gas, underwater imaging

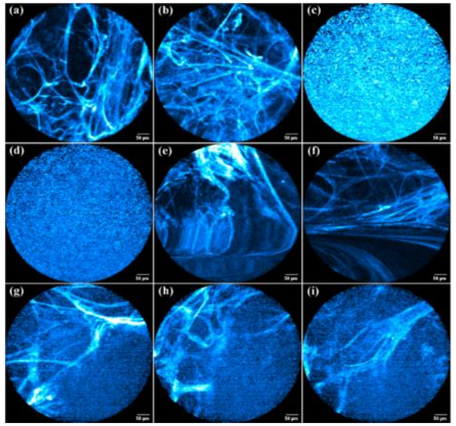
Pipeline inspection

- Environmental sciences

Forest canopy monitoring

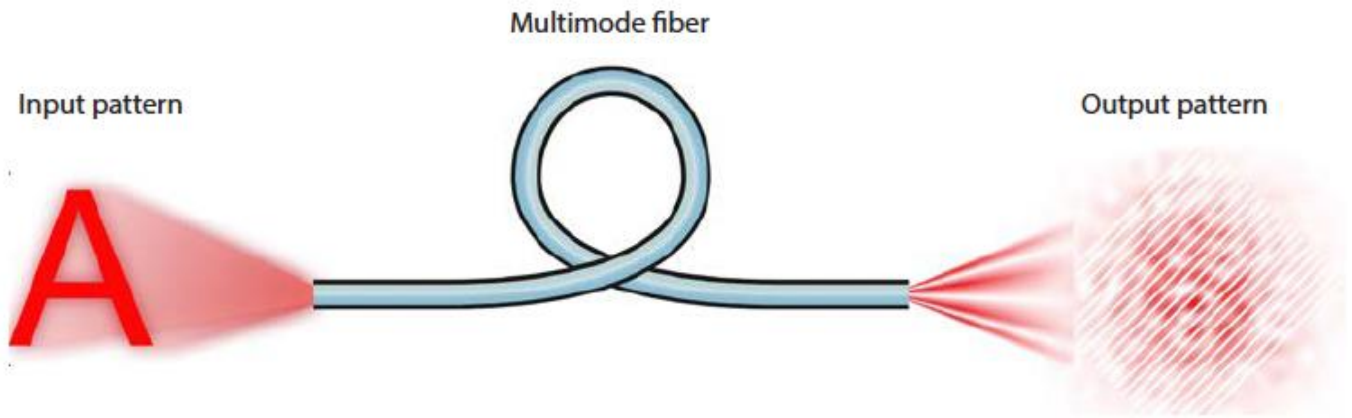


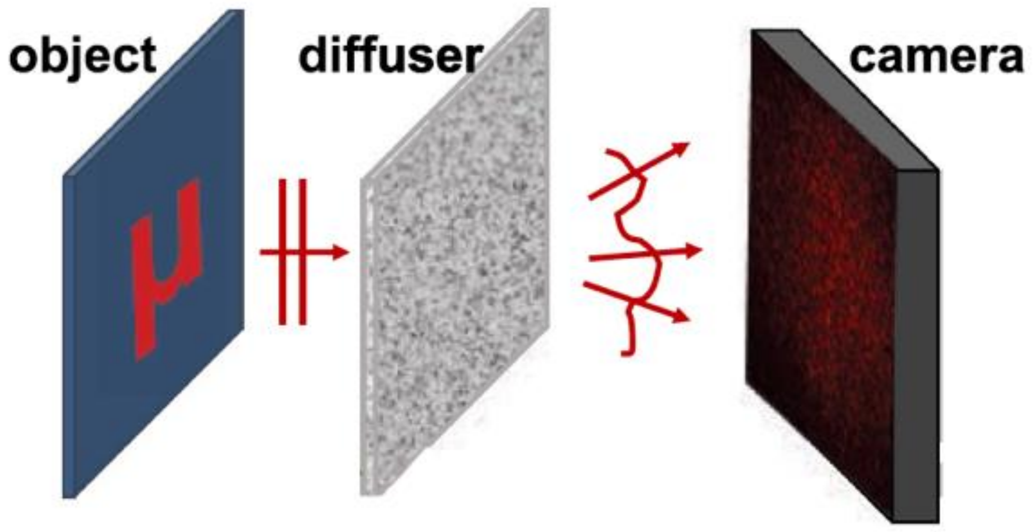
- Biomedical Imaging



Distal Lung Endomicroscopy

Speckle Imaging in Disordered Media

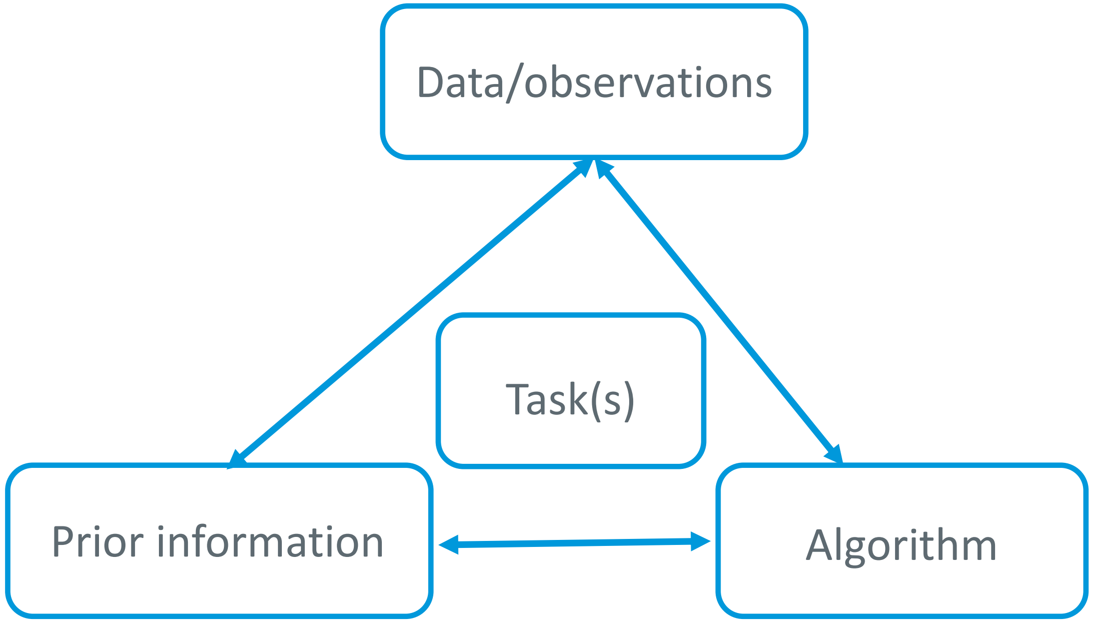




Bayesian methods

- Uncertainty management
  - Noisy/incomplete measurements
  - Prior information
  - Quality of output

- Algorithms
  - MAP/penalized MLE
  - Simulation/MCMC
  - Approximate methods



Bayesian modeling

- Observations $\boldsymbol{y}$ related to unknown parameters of interest $\boldsymbol{x}$ via stochastic process $\boldsymbol{f}(\boldsymbol{y}|\boldsymbol{x})$
- Exact model

$$f(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\theta}) = \frac{f(\boldsymbol{y}|\boldsymbol{x})f(\boldsymbol{x}|\boldsymbol{\theta})}{f(\boldsymbol{y}|\boldsymbol{\theta})}$$

- Approximating distribution

$$f(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\theta}) \approx q(\boldsymbol{x})$$

  – Proximal MCMC: Moreau envelope

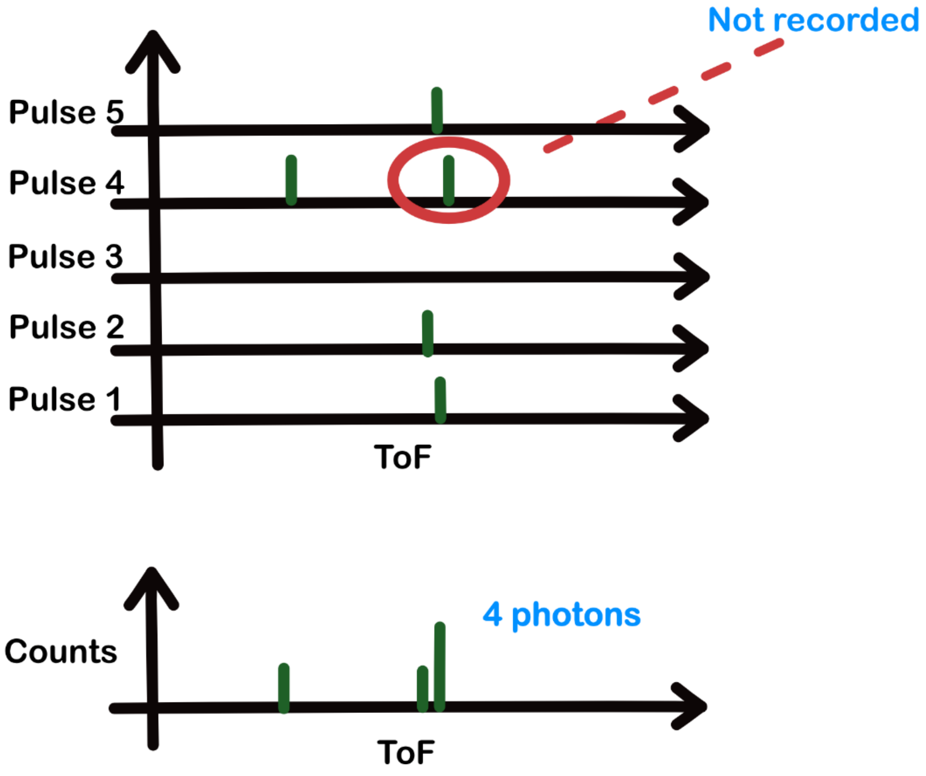
  – VB/EP: Divergence-based

Generative models

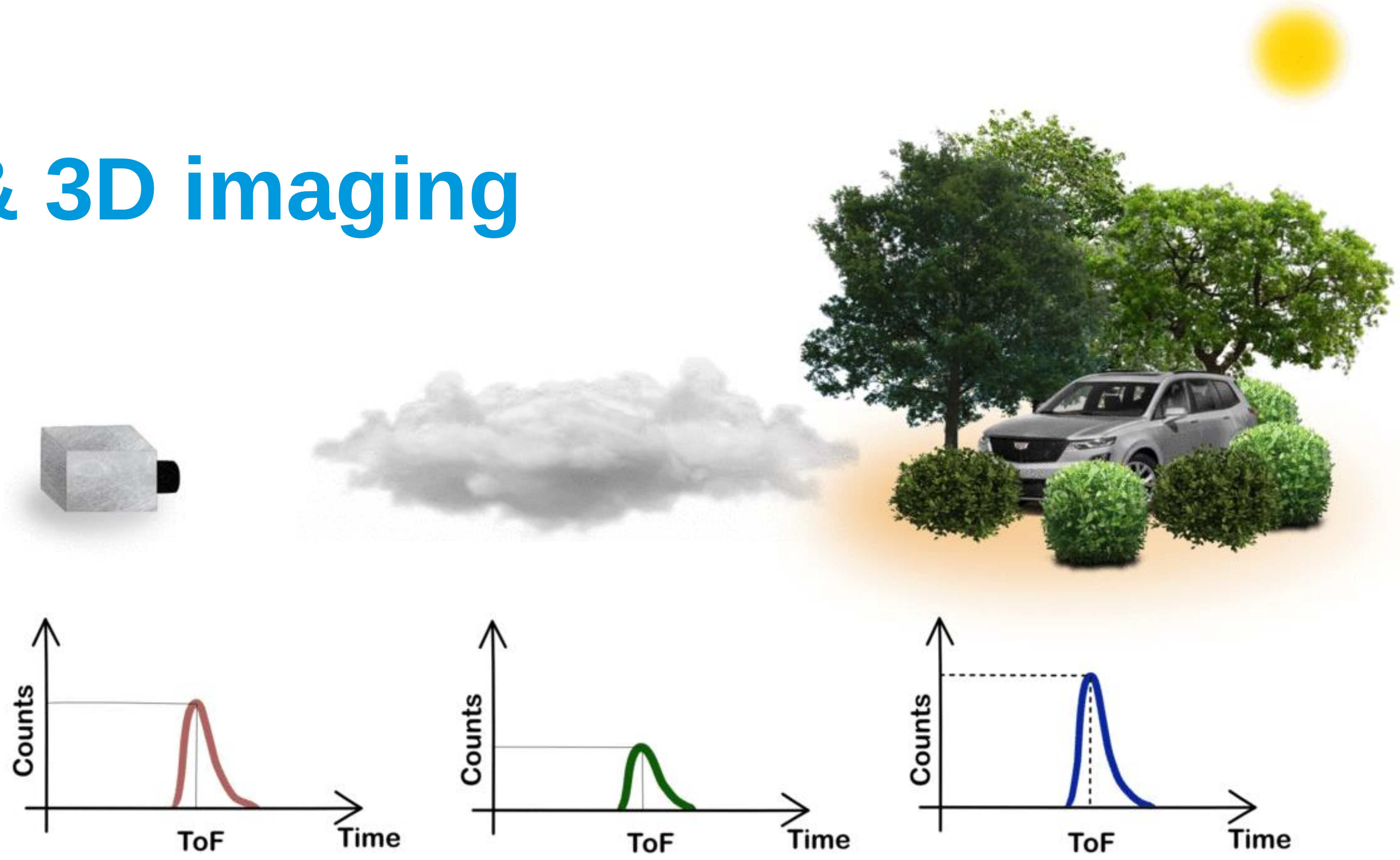
$$\boldsymbol{y} = g(\boldsymbol{A}\boldsymbol{x}) + \boldsymbol{n}$$

- Impact on uncertainty
  – Linear/nonlinear forward model
  – Non-Gaussian/i.i.d. noise
  – Outliers

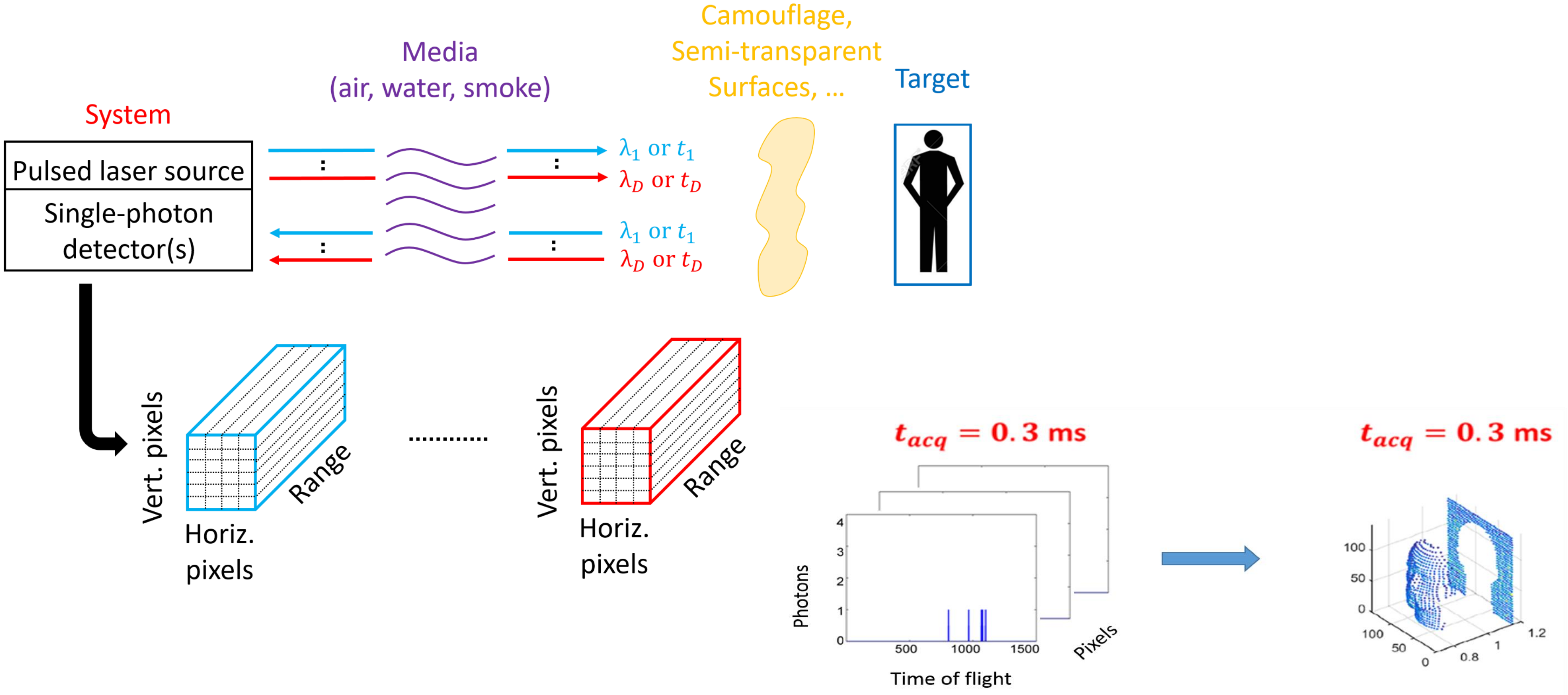
Single-photon Lidar



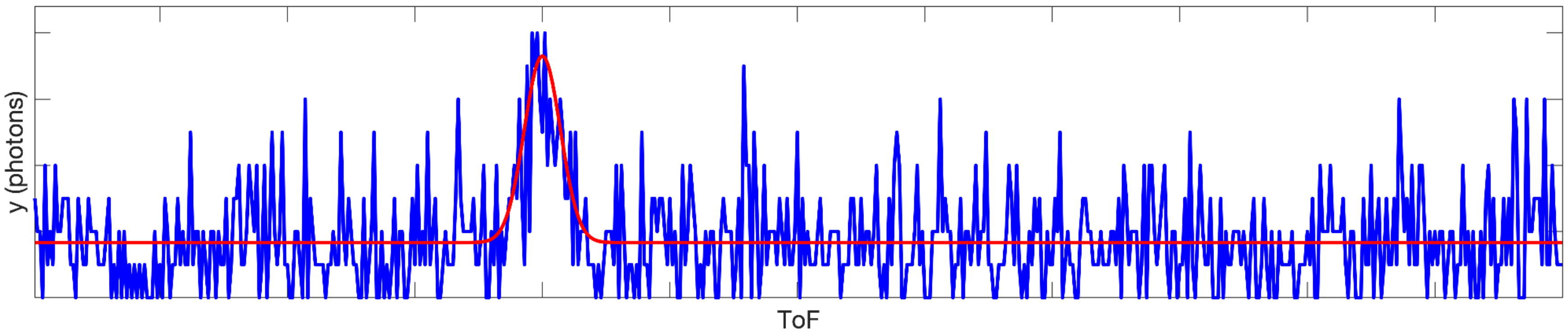
Color & 3D imaging



Lidar System

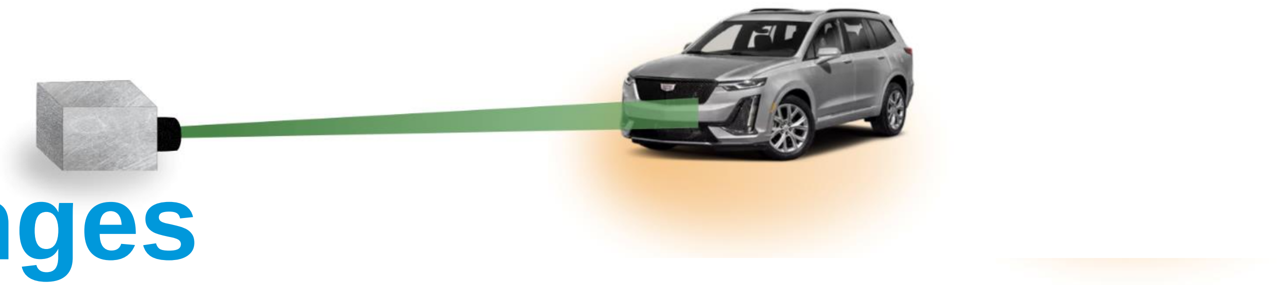


Observation model



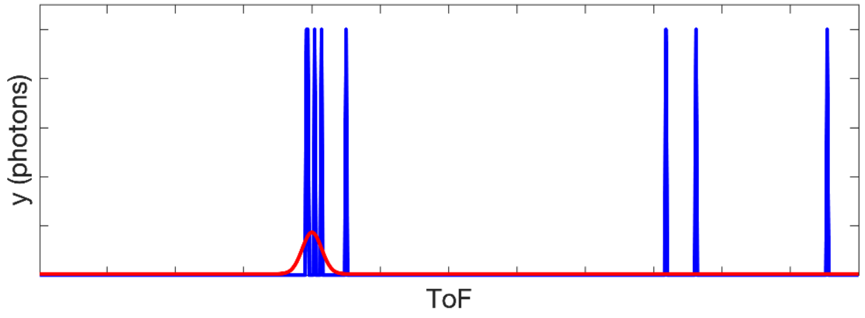
$$y_{n,t} \sim \mathcal{P}oisson(r_n g_0(t - t_n) + b_n), \qquad t \in \{1, \dots, T\}, n \in \{1, \dots, N\}$$

- $y_{n,t}$: photon count in t^{th} bin and n^{th} pixel
- $g_0(\cdot)$: instrumental response
- T : Histogram length
- b_n: background level
- r_n: target reflectivity
- t_n: Time-of-flight (ToF)

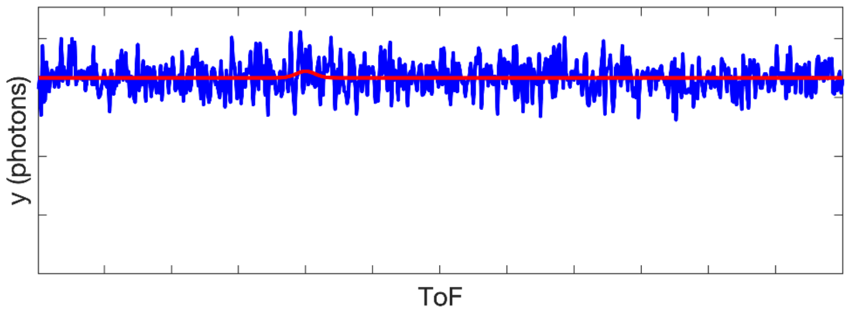


Detection Challenges

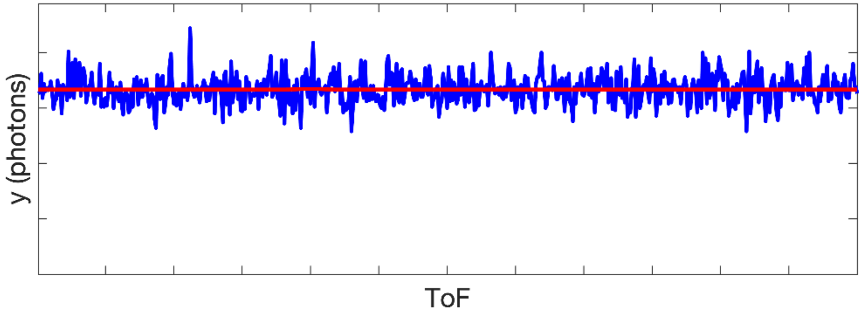
1. Few detected photons



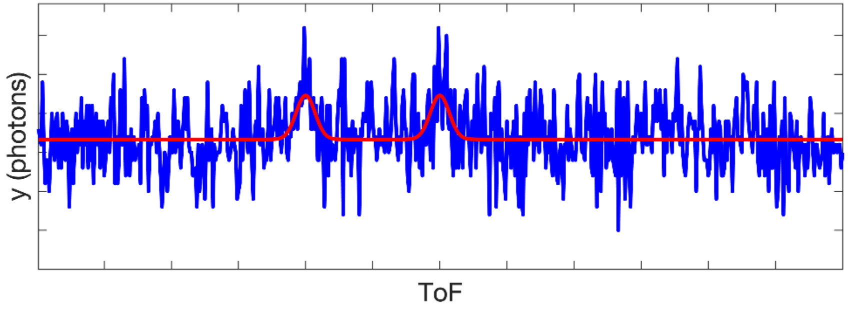
2. High background



3. No target



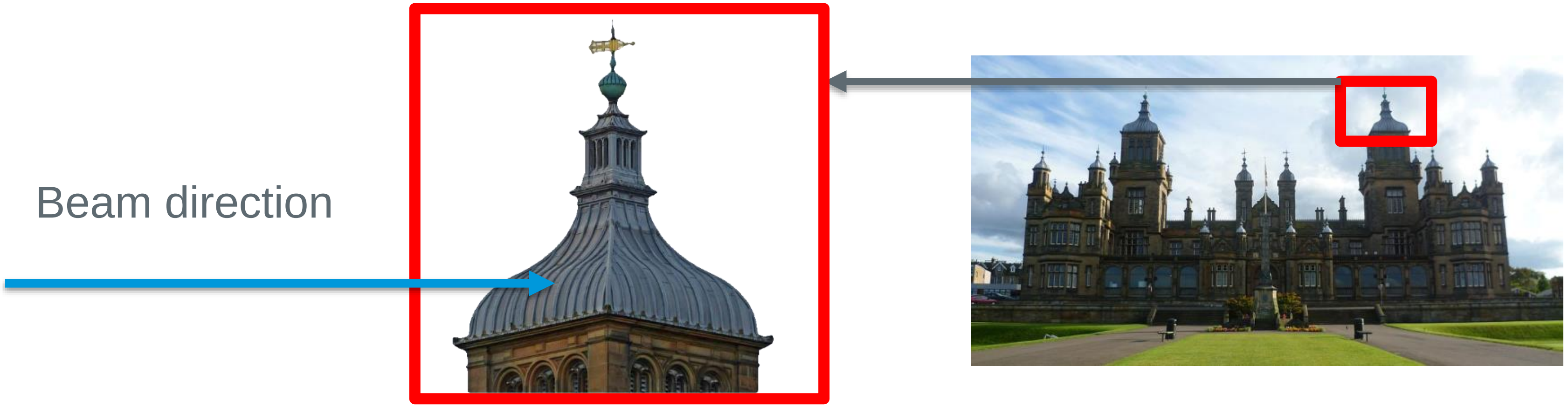
4. Multiple peaks

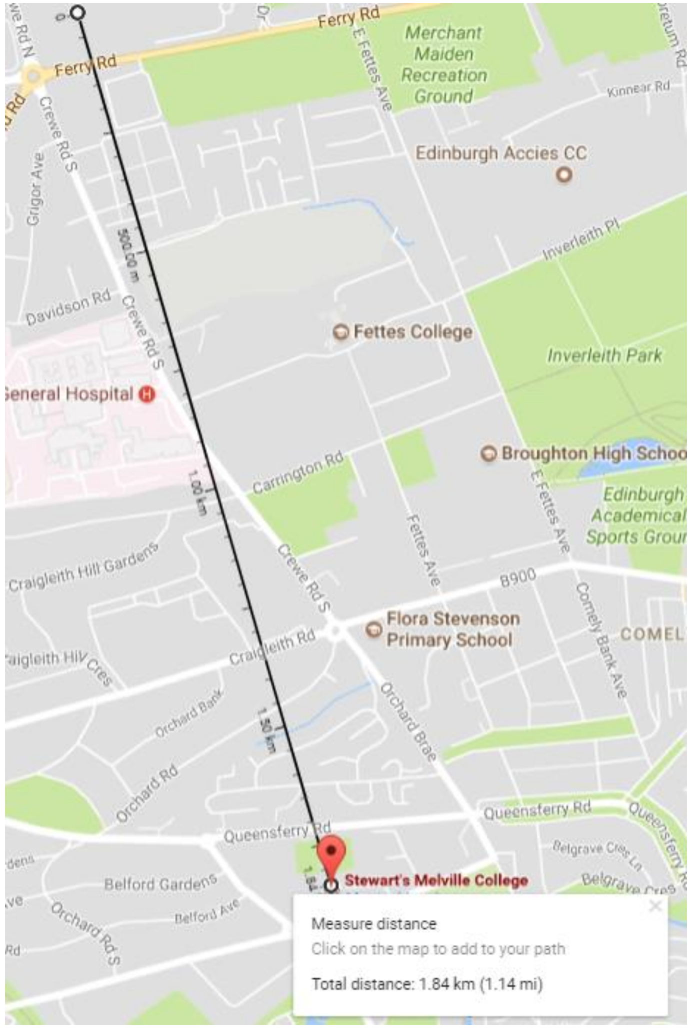


Reconstruction using MCMC

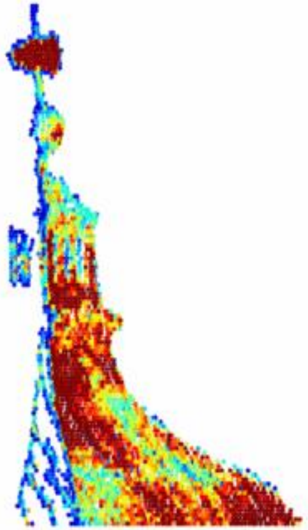
- Long range reconstruction (1.8 km)
- 123x96 pixels, 800 temporal bins
- Approx. 900 photons per pixel
- Signal-to-background-ratio: 1.64

Beam direction

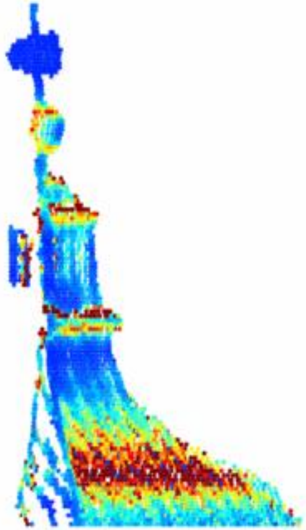




Reconstruction using MCMC



Intensity



Peak width

Execution time on a standard workstation: 195 s

J. Tachella et al., "Bayesian 3D reconstruction of complex scenes from single-photon Lidar data", SIAM J. Imaging Sci., 2019.
J. Tachella et al., "3D reconstruction using single-photon Lidar data exploiting the widths of the returns", ICASSP, 2019.

Towards real-time analysis

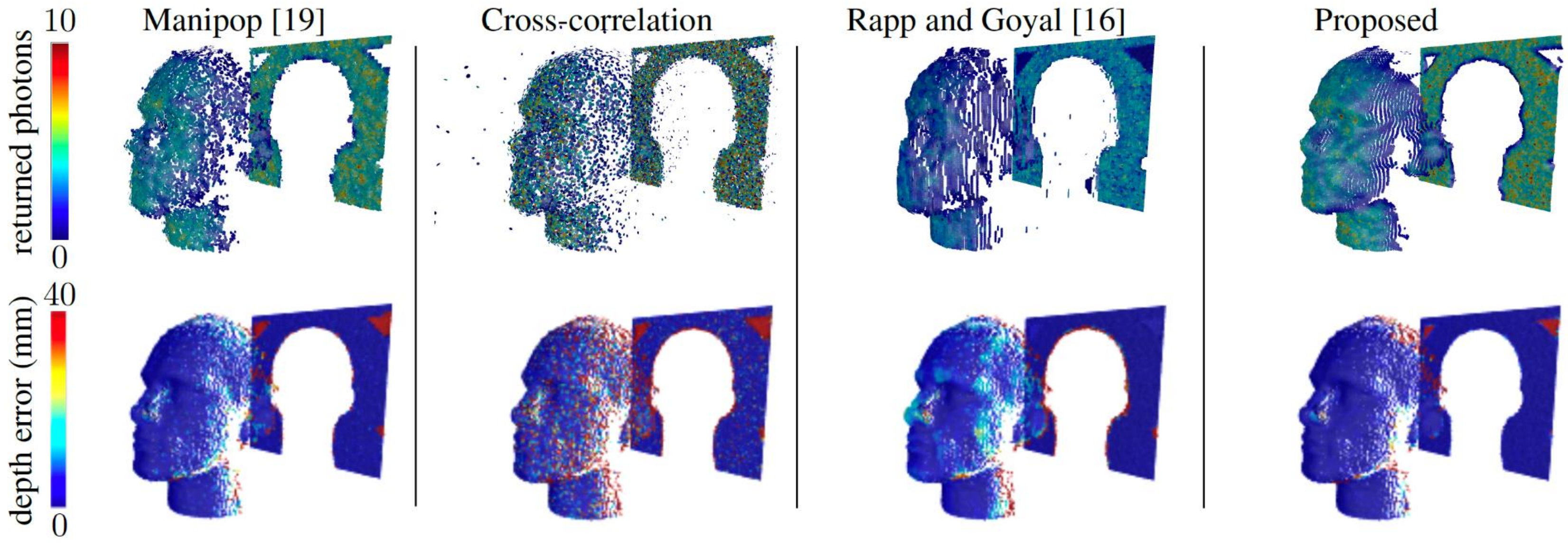
- Complex and difficult inference problems
- Data volume/array size
- Acquisition frame rate

- Optimization $\nRightarrow$ fast(er)
  – Dimensionality of the data/unknowns
  – Convergence speed
- Need to rethink the inference process

Towards real-time analysis

- Not just an implementation issue
- Parallel structures
  - Statistical models (single-photon)
  - Scalable denoisers
- Tools
  - Plug-and-play approaches
  - Point cloud denoiser (computer graphics)

Towards real-time analysis



Execution time: 201 s 1 ms 201 s 13 ms

Algorithm design

$$f(\boldsymbol{x}|\boldsymbol{y}) \propto f(\boldsymbol{y}|\boldsymbol{x})f(\boldsymbol{x})$$

- MAP estimation

$$\min_{\boldsymbol{x}} h_y(\boldsymbol{x}) + h_0(\boldsymbol{x})$$

- $h_y(\boldsymbol{x}) = -\log(f(\boldsymbol{y}|\boldsymbol{x}))$
- $h_0(\boldsymbol{x}) = -\log(f(\boldsymbol{x}))$
- Both can be challenging (e.g., non smooth, multimodal)

Algorithm design

$$\min_{\boldsymbol{x}} h_{\boldsymbol{y}}(\boldsymbol{x}) + h_0(\boldsymbol{x})$$

- Splitting strategy

$$\min_{\boldsymbol{x},\boldsymbol{u}} h_{\boldsymbol{y}}(\boldsymbol{x}) + h_0(\boldsymbol{u}), s.t.\ \boldsymbol{x} = \boldsymbol{u}$$

- Break down big problem into smaller, easier problems
- Smaller problems can often be seen as denoising problems
  - Dedicated denoisers can be used
- Plug-and-play (PnP) approaches possible
  - No analytical expression for $h_0(\boldsymbol{u})$

J. Tachella et al. "Real-time 3D reconstruction from single-photon lidar data using plug-and-play point cloud denoisers", Nature Comms., Nov 2019.

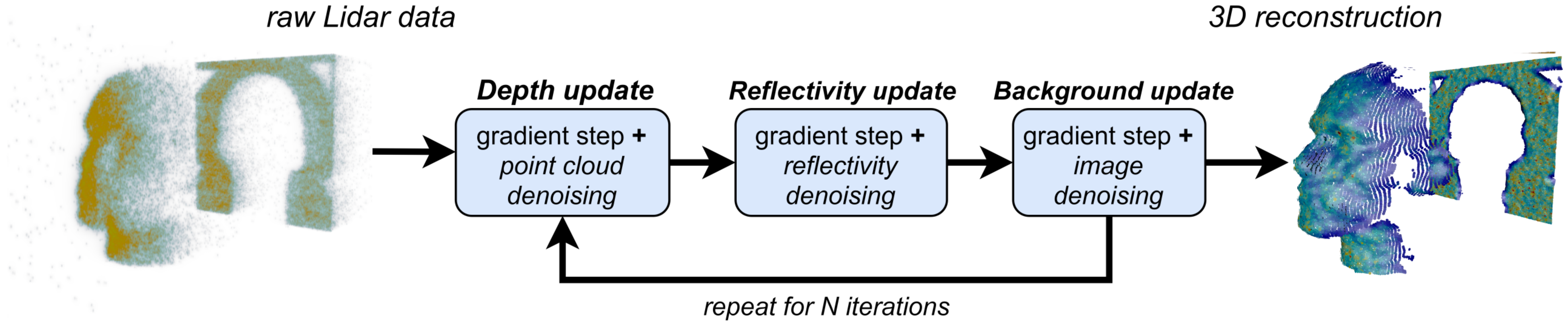
Algorithm design

$$\min_{\boldsymbol{x},\boldsymbol{u}} h_{\boldsymbol{y}}(\boldsymbol{x}) + h_0(\boldsymbol{u}), s.t.\ \boldsymbol{x} = \boldsymbol{u}$$

- PnP increasingly used for image restoration
- Here, $\boldsymbol{x}$ includes a 3D point cloud
  - 2D surfaces within a 3D volume
  - Voxel-based methods computationally intensive
- Proposed solution: point cloud denoisers from the computer graphics community (e.g. APSS)
  - More structured prior
  - Scalable

G. Guennebaud, & M. Gross, " Algebraic point set surfaces", ACM Trans. Graph. 26, 23 (2007).

RT3D algorithm



- Image-based PnP strategy for the reflectivity and background profiles
- Point cloud-based PnP using off-the-shelf point cloud denoisers (e.g. local sphere fitting)

J. Tachella et al. “Real-time 3D reconstruction from single-photon lidar data using plug-and-play point cloud denoisers”, Nature Comms., Nov 2019.

RT3D algorithm

Real-time (50 fps) 3D reconstruction from 32 x 32 pixels
Multiple surfaces per pixel
Distance: 300 m

Algorithm design

$$\min_{\boldsymbol{x},\boldsymbol{u}} h_{\boldsymbol{y}}(\boldsymbol{x}) + h_0(\boldsymbol{u}), s.t.\ \boldsymbol{x} = \boldsymbol{u}$$

- PnP generally used for better priors/denoising
- Can also be used for the data-fidelity term
  - Computational complexity
  - Robustness to model mismatch (e.g., outliers)
- How to design $h_{\boldsymbol{y}}(\boldsymbol{x})$ while keeping a Bayesian interpretation?

Pseudo Bayesian estimation

- MLE can be obtained by minimizing the KL divergence between the empirical data distribution $\hat{f}(y)$ and a parametric distribution $f_x(y)$

$$KL\left(\hat{f}(y)\,||f_x(y)\right)$$

  – Robust estimator can be obtained by changing the divergence

- $f(\boldsymbol{x}|\boldsymbol{y})$ can be obtained by solving a penalized KL divergence minimization problem

Q. Legros et al., “Robust depth imaging in adverse scenarios using single-photon Lidar and beta-divergences“, SSPD Conference 2020

Pseudo Bayesian estimation

- By changing the similarity measure we can obtain a pseudo-likelihood $\tilde{f}(\boldsymbol{y}|\boldsymbol{x})$ and a pseudo posterior

$$\tilde{f}(\boldsymbol{x}|\boldsymbol{y}) \propto \tilde{f}(\boldsymbol{y}|\boldsymbol{x}) f(\boldsymbol{x})$$

- Here we chose
  - Simple model which omits background
  - β-divergence which is robust to high background levels
- How to ensure scalability for dynamic scenes?
  - Variational inference

Q. Legros et al. “Robust 3D reconstruction of dynamic scenes from single-photon lidar using Beta-divergences”, to appear in IEEE Trans. Image Processing.

Online 3D reconstruction

- $\tilde{f}(\boldsymbol{x}|\boldsymbol{y})$ usually not standard
- Approximation by Gaussian distribution $q(\boldsymbol{x})$

$$q(\boldsymbol{x}) = \underset{p(\boldsymbol{x})}{\operatorname{argmin}}\, KL\left(\tilde{f}(\boldsymbol{x}|\boldsymbol{y}) \,||\, p(\boldsymbol{x})\right)$$

- Reduces to moment matching
- Assumed density filtering / Expectation-Propagation
- Only means and variances propagated over time

Expectation-Propagation

$In\ essence\ approximate\ a\ function\ by\ a\ simpler\ one$

$f(\boldsymbol{y},\boldsymbol{x}|\boldsymbol{\theta}) = f(\boldsymbol{y}|\boldsymbol{x})f(\boldsymbol{x}|\boldsymbol{\theta})$

$q(\boldsymbol{x}) \propto q_1(\boldsymbol{x})q_0(\boldsymbol{x})$: user-defined

- Based on the reverse KL-divergence

$$\min_{Z,\,q(\boldsymbol{x})} KL\big(f(\boldsymbol{y},\boldsymbol{x}|\boldsymbol{\theta}) \mid\mid Z_{y,\theta}\,q(\boldsymbol{x})\big)$$

$$Z_{\boldsymbol{y},\boldsymbol{\theta}} \approx f(\boldsymbol{y}|\boldsymbol{\theta})$$

EP: better at preserving the marginals

Online 3D reconstruction

- Projection:

  $$q_{(t-1)}(\boldsymbol{x}^{t-1}) \approx \tilde{f}(\boldsymbol{x}^{t-1}|\boldsymbol{y}^{t-1}, \ldots, \boldsymbol{y}^{0})$$ usually not standard

- GMM-based prediction:

  $$f(\boldsymbol{x}^{t}) = \int f(\boldsymbol{x}^{t}|\boldsymbol{x}^{t-1}) q_{(t-1)}(\boldsymbol{x}^{t-1}) \mathrm{d}\boldsymbol{x}^{t-1}$$

  – GMM allowing new surfaces and surfaces leaving the scene (birth/death processes)

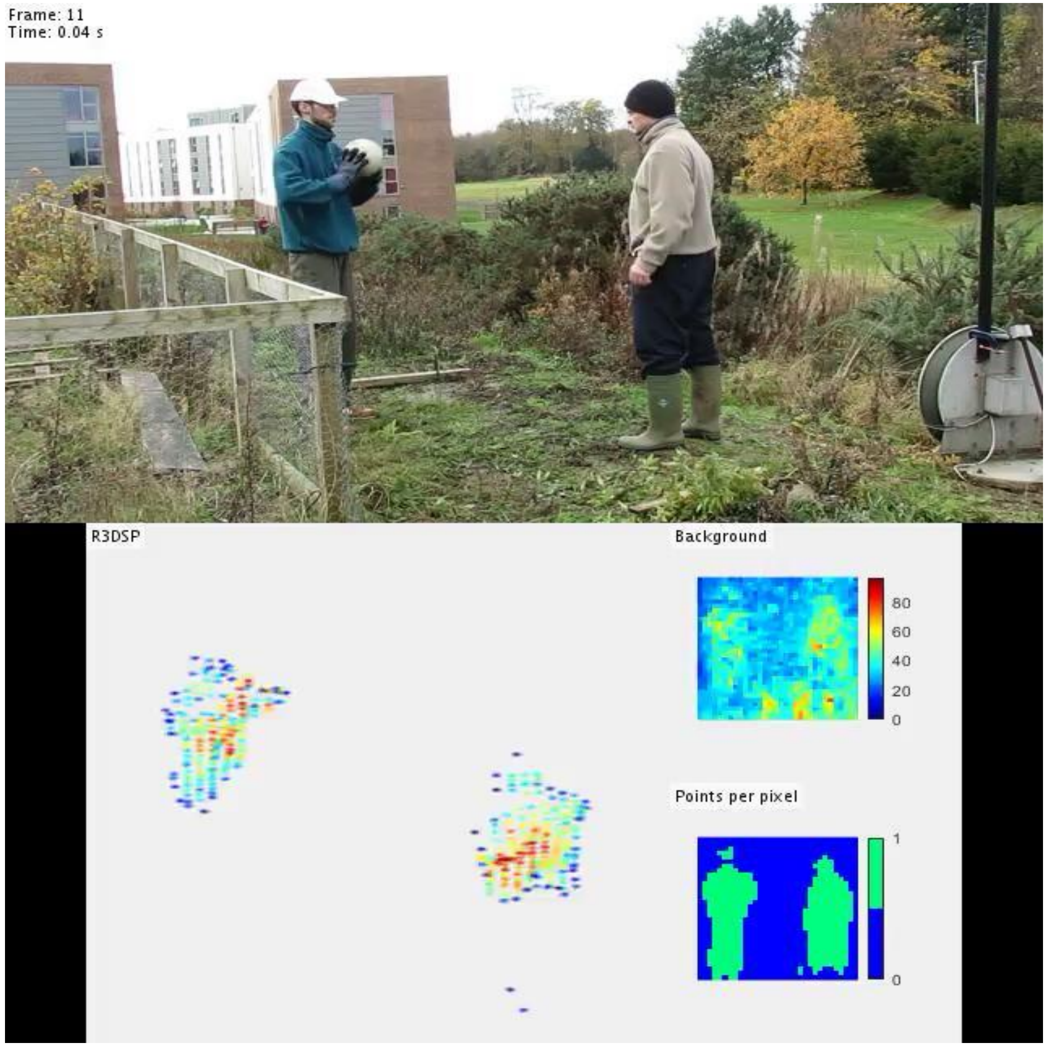
- Pseudo-posterior inference

  $$\tilde{f}(\boldsymbol{x}^{t}|\boldsymbol{y}^{t}, \ldots, \boldsymbol{y}^{0}) \propto \tilde{f}(\boldsymbol{y}^{t}|\boldsymbol{x}^{t}) f(\boldsymbol{x}^{t})$$

Y. Altmann et al. “Fast online 3D reconstruction of dynamic scenes from individual single-photon detection events”, IEEE Trans. Image Processing, 2019.

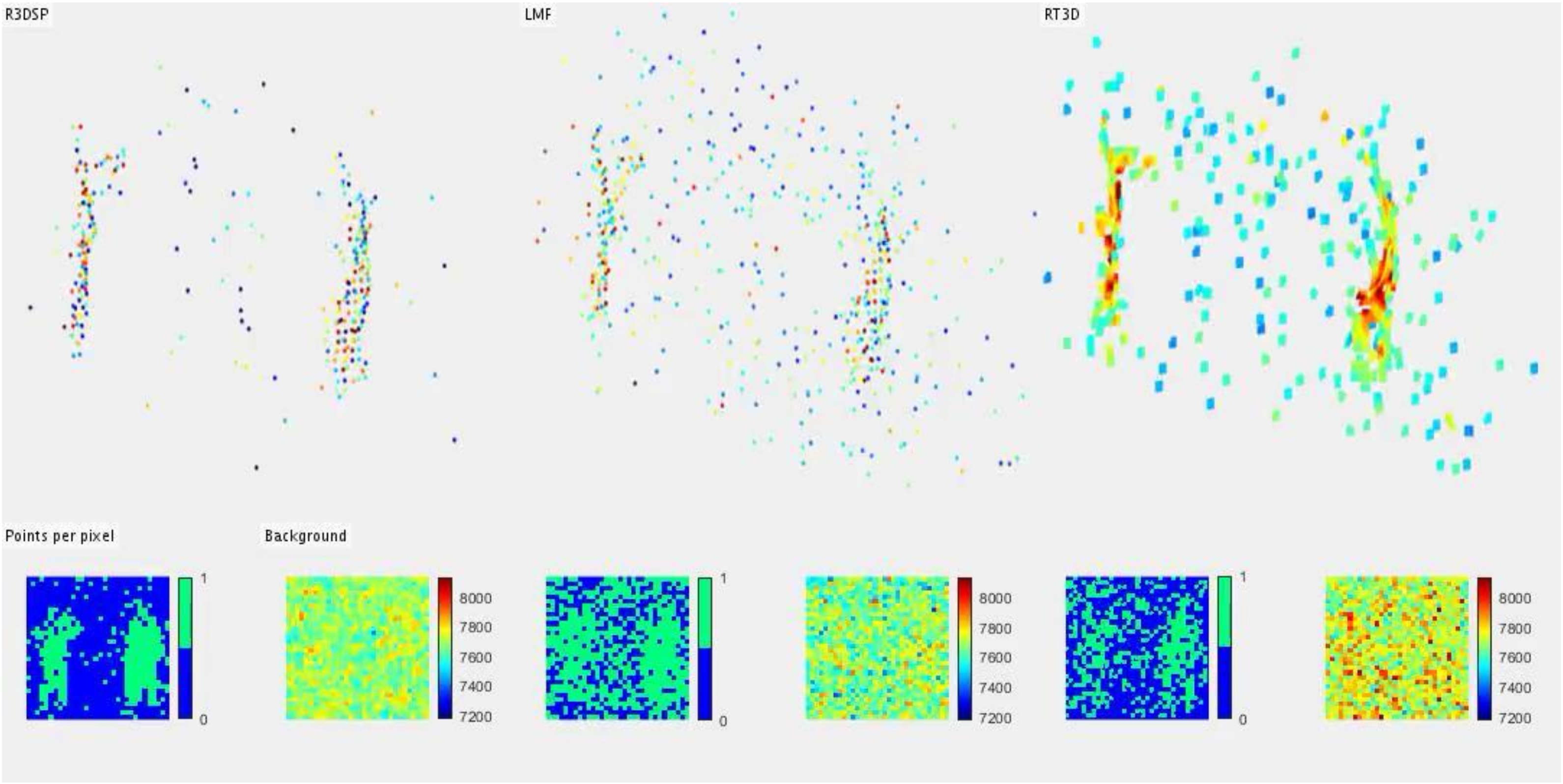
Robust online 3D reconstruction

- 3D reconstruction (5000 fps) from 32x32 pixels
- At most 1 surface per pixel



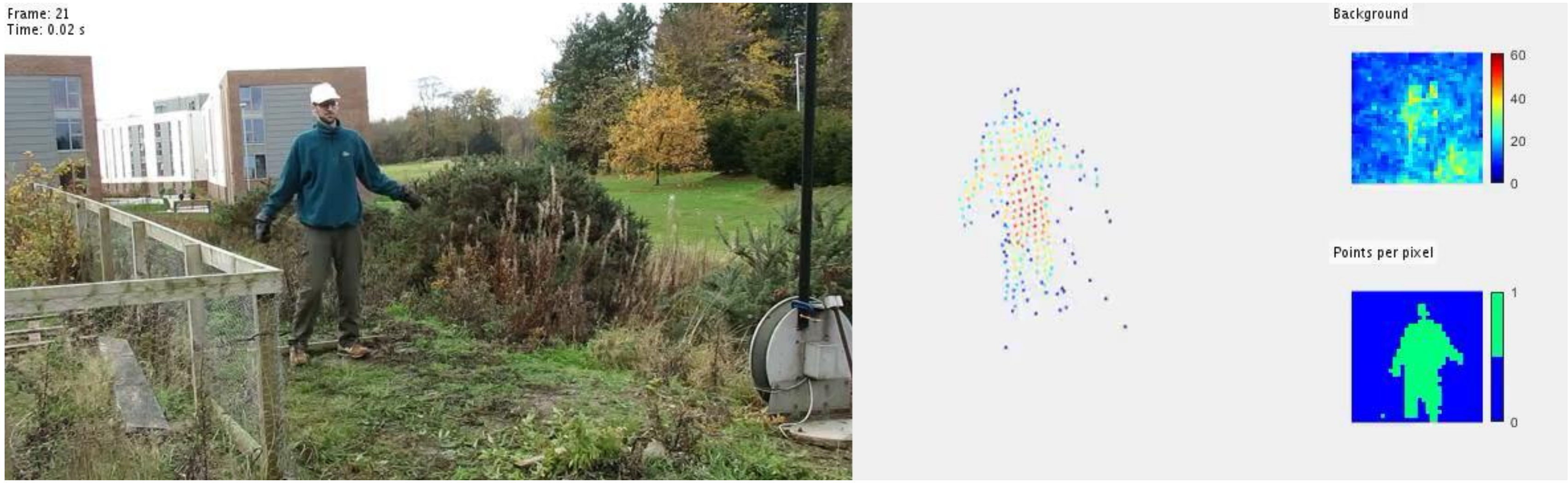
Q. Legros et al. “Robust 3D reconstruction of dynamic scenes from single-photon lidar using Beta-divergences”, IEEE Trans. Image Processing, 2020.

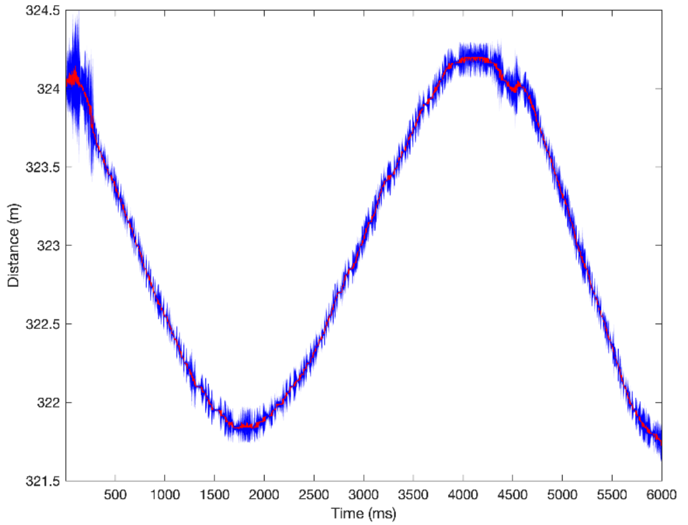
Robust online 3D reconstruction



Q. Legros et al. “Robust 3D reconstruction of dynamic scenes from single-photon lidar using Beta-divergences”, IEEE Trans. Image Processing, 2020.

Robust online 3D reconstruction





Red curve: depth posterior mean
Blue region: credible interval

Q. Legros et al. “Robust 3D reconstruction of dynamic scenes from single-photon lidar using Beta-divergences”, IEEE Trans. Image Processing, 2020.

Radionuclide detection

- Radiation portal monitors
  - Freight
  - Airports

- Challenges
  - Background
  - Shielding

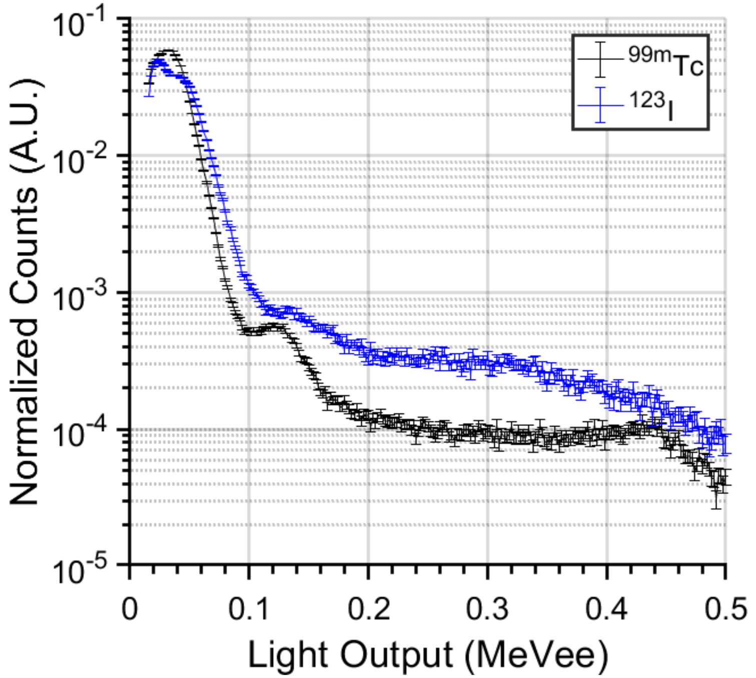
Source: https://en.wikipedia.org/wiki/Radiation_Portal_Monitor

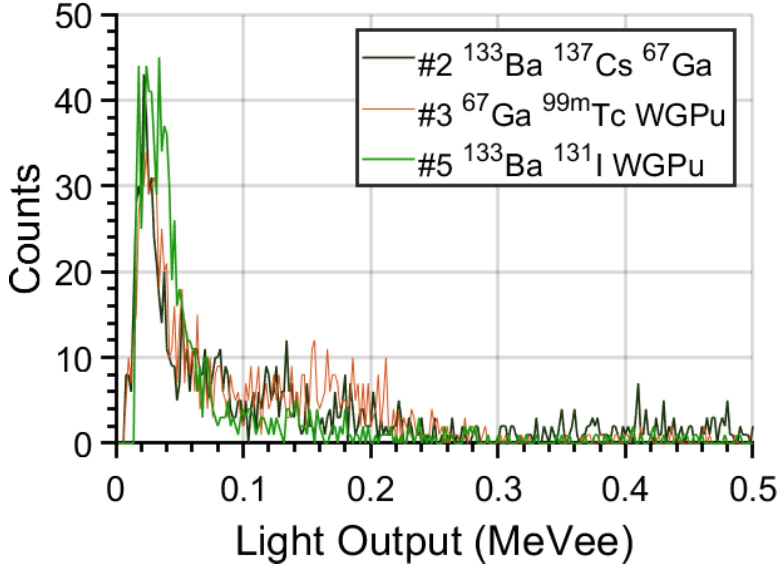
Source: www.airport-suppliers.com/supplier/arktis-radiation-detectors-ltd/

Bayesian model

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

- $\boldsymbol{y}|\boldsymbol{\lambda} \sim \mathcal{P}oisson(\boldsymbol{\lambda})$, with $\boldsymbol{\lambda} = \boldsymbol{A}\boldsymbol{x}$
- $\boldsymbol{A}$: spectral library of radiation sources
- $\boldsymbol{x}$: Abundance vector
- Challenges
  - – Multiplicative noise
  - – Positivity of $\boldsymbol{x}$
  - – Sparsity of $\boldsymbol{x}$





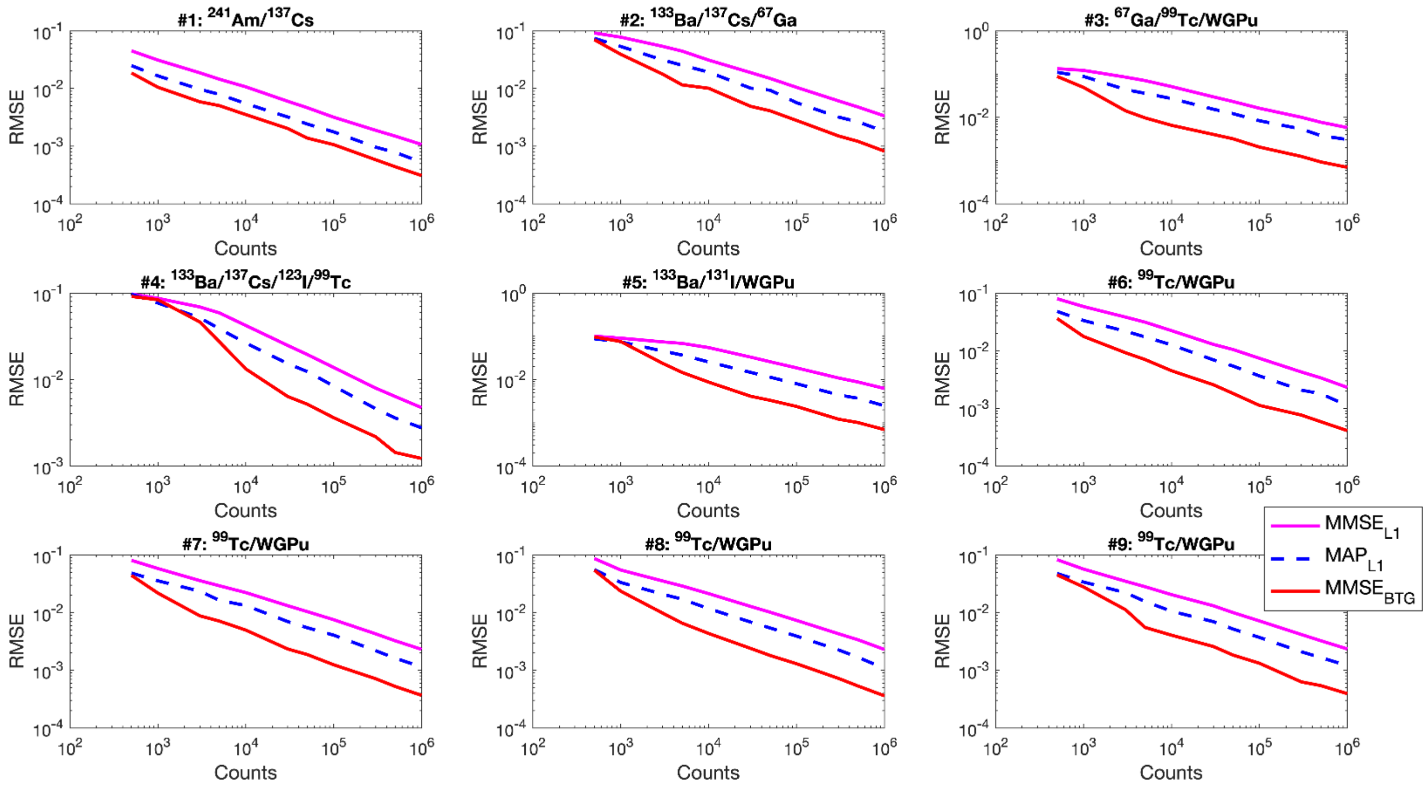
Approximate inference

- Spike and slab prior
  - Bernoulli-truncated Gaussian distributions
  - Multimodal posterior distribution
- Simulation-based estimation
  - Non-standard conditional distributions
- Variational inference using expectation-propagation

J. Hernández-Lobato, D. Hernández-Lobato, A. Suárez, “Expectation propagation in linear regression models with spike-and-slab priors“. Mach. Learn. 99, 2015.
Y. Altmann et al. “Expectation-propagation for weak radionuclide identification at radiation portal monitors “. Sci Rep 10, 6811, 2020.

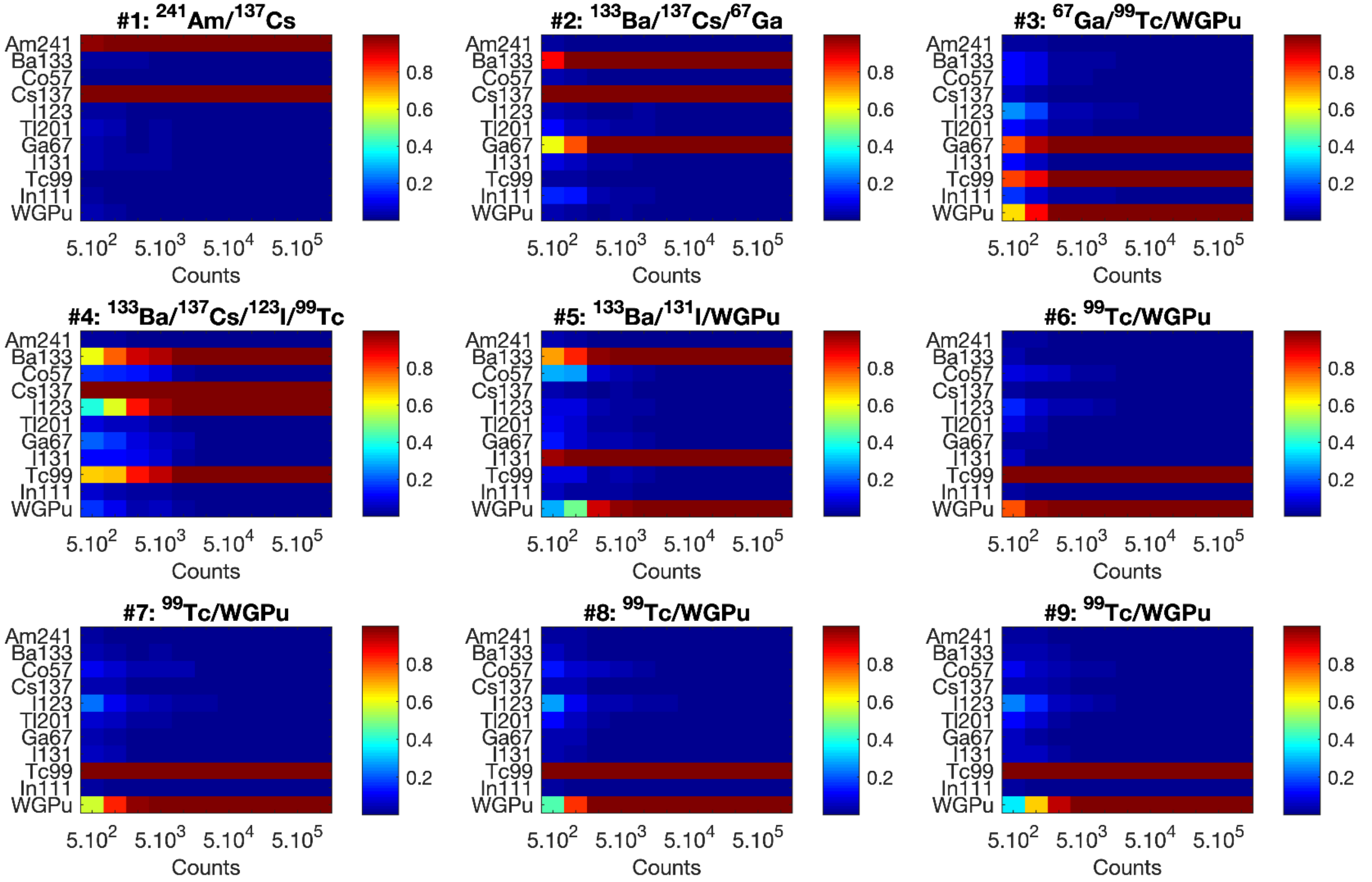
Results

- 9 mixtures
- Library of 11 sources
- Comparison with ℓ_1-norm based methods



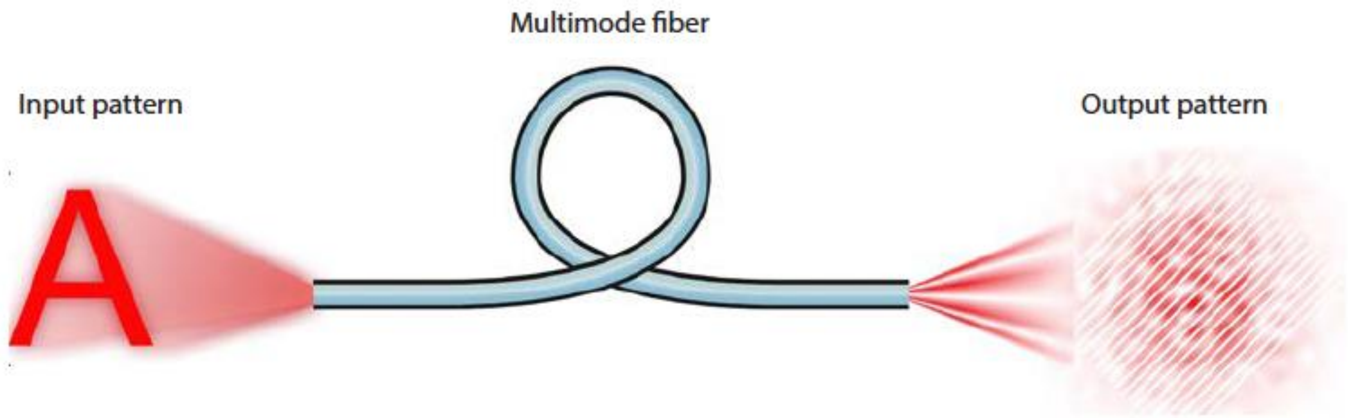
Abundance RMSEs

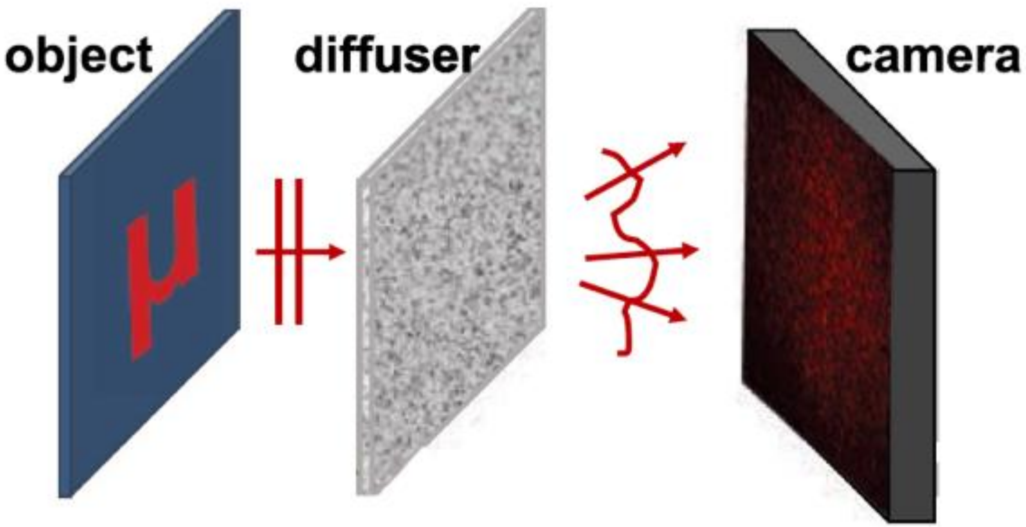
Results

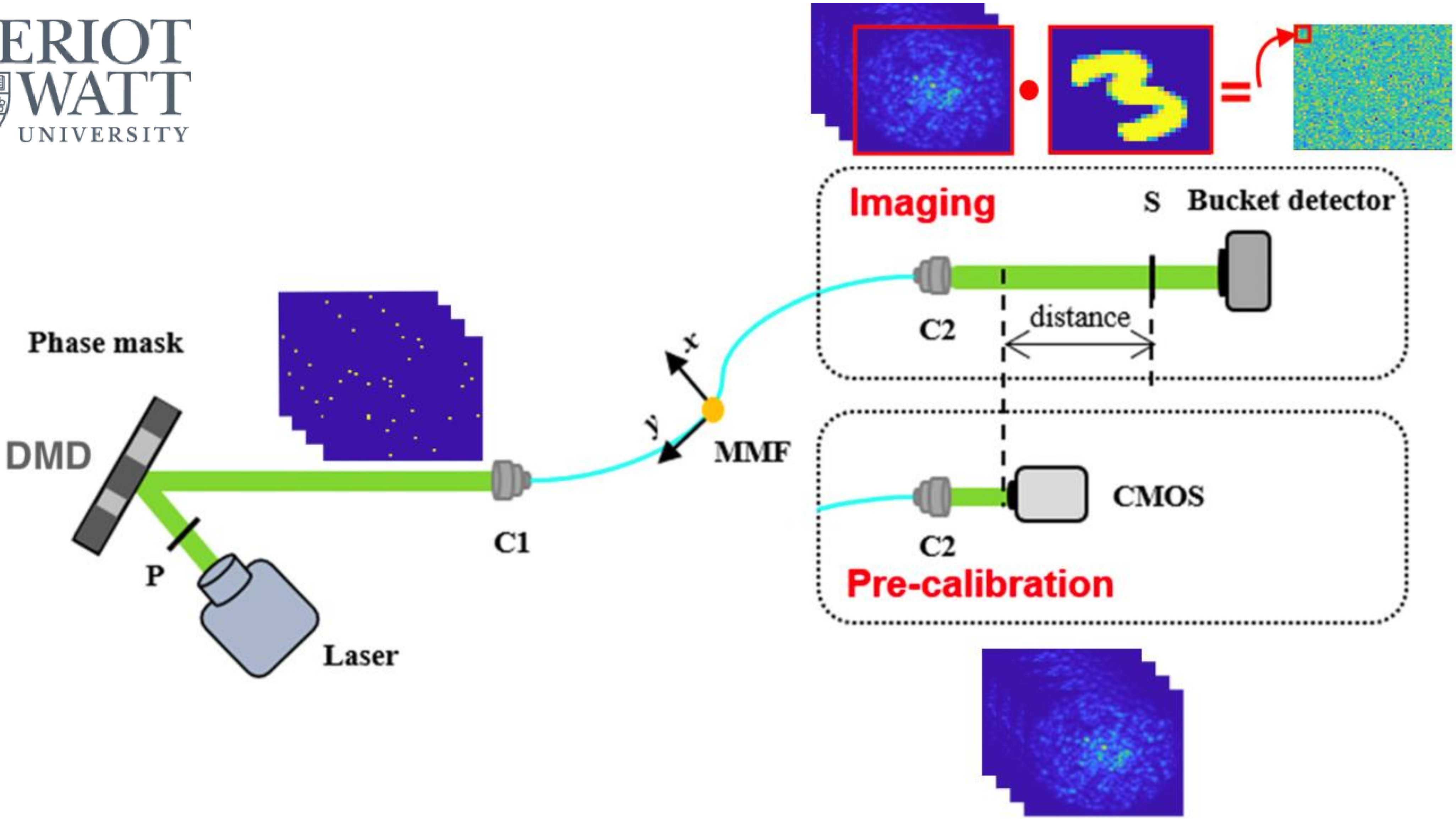


Approximate marginal probabilities of source presence

Speckle Imaging in Disordered Media







Experimental System

courtesy of Miles Padget and Peter Mekhail, University of Glasgow

Problem Definition

- Calibration: record a set of intensity speckle patterns to form a sensing matrix **A** **(**each row of **A** is an intensity speckle pattern**)**.
- Imaging: project the same speckle patterns (i.e. the sensing matrix **A**) onto the image of interest ***x*** and record the intensity measurements ***y***.
- Use Bayesian inference techniques to recover the image of interest from the noisy measurements.

https://jaan.io/what-is-variational-autoencoder-vae-tutorial/

Problem Formulation

- The speckle imaging problem can be formulated as: $\boldsymbol{y} = \mathbf{A}\boldsymbol{x} + \boldsymbol{n}$
  - $\boldsymbol{y}$ is the measurement vector which represents the total intensity of the speckle patterns projected onto the image of interest $\boldsymbol{x}$
  - $\boldsymbol{n}$ is i.i.d. Gaussian noise with known variance σ^2
  - $\mathbf{A}$ is the sensing matrix where each row represents an intensity speckle pattern
- The Inverse imaging problem is then: $p(\boldsymbol{x}|\boldsymbol{y},\mathbf{A}) \propto p(\boldsymbol{y}|\boldsymbol{x},\mathbf{A})\, p(\boldsymbol{x})$
  - Likelihood is $\boldsymbol{y}|\, \boldsymbol{x},\mathbf{A} \sim N(\boldsymbol{y}|\mathbf{A}\boldsymbol{x},\sigma^2\mathbf{I})$
  - Prior (e.g. smoothness, sparsity)

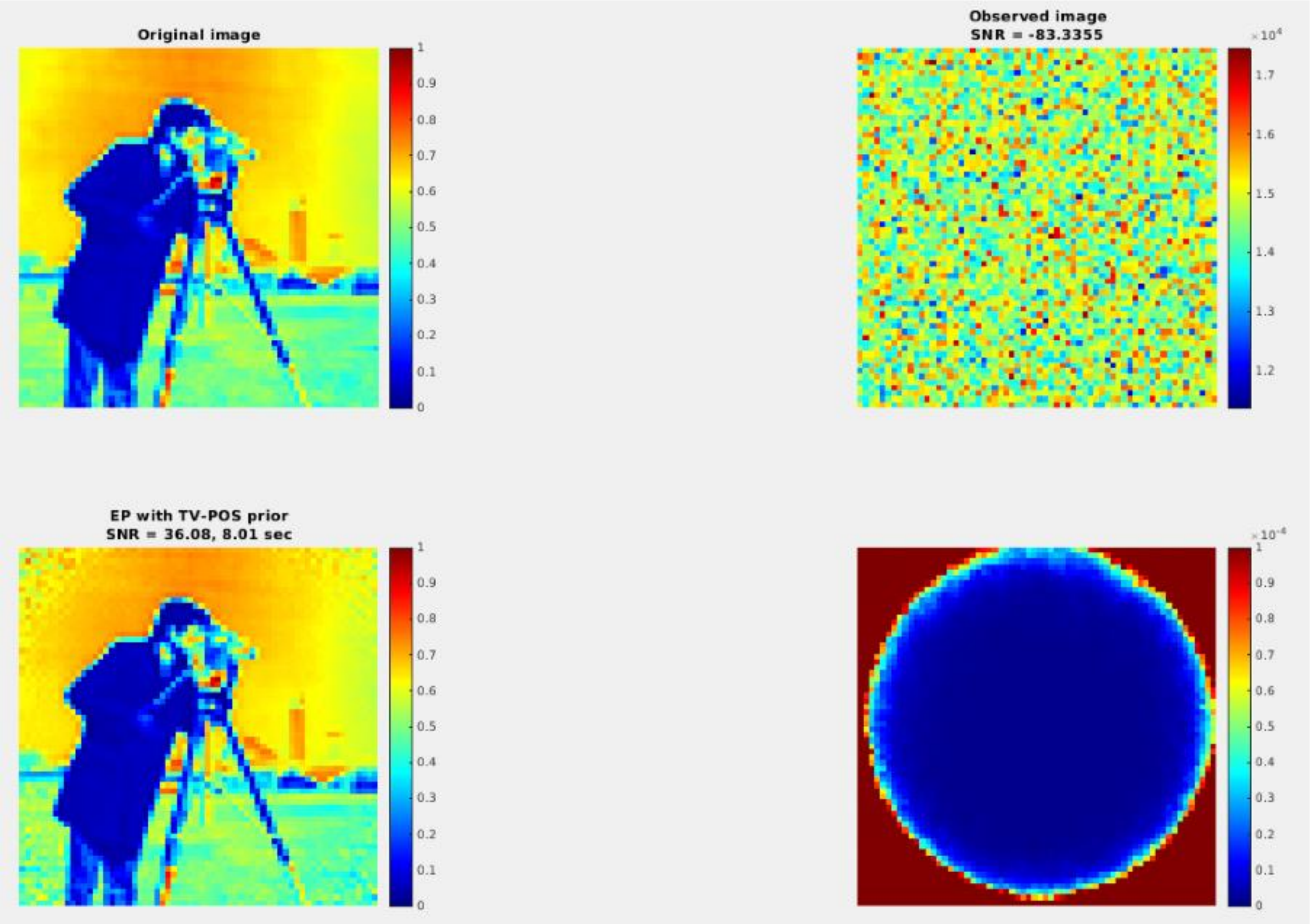
Algorithm Design

- Sampling methods (MCMC)
- MAP Estimation

$$\hat{\boldsymbol{x}} = argmin(-\log\{p(\boldsymbol{y}|\,\boldsymbol{x}, \boldsymbol{A})\} - \log\{p(\boldsymbol{x})\})$$

- Approximate Bayesian methods:
  – Variational Bayes (VB)
  – Expectation propagation (EP)

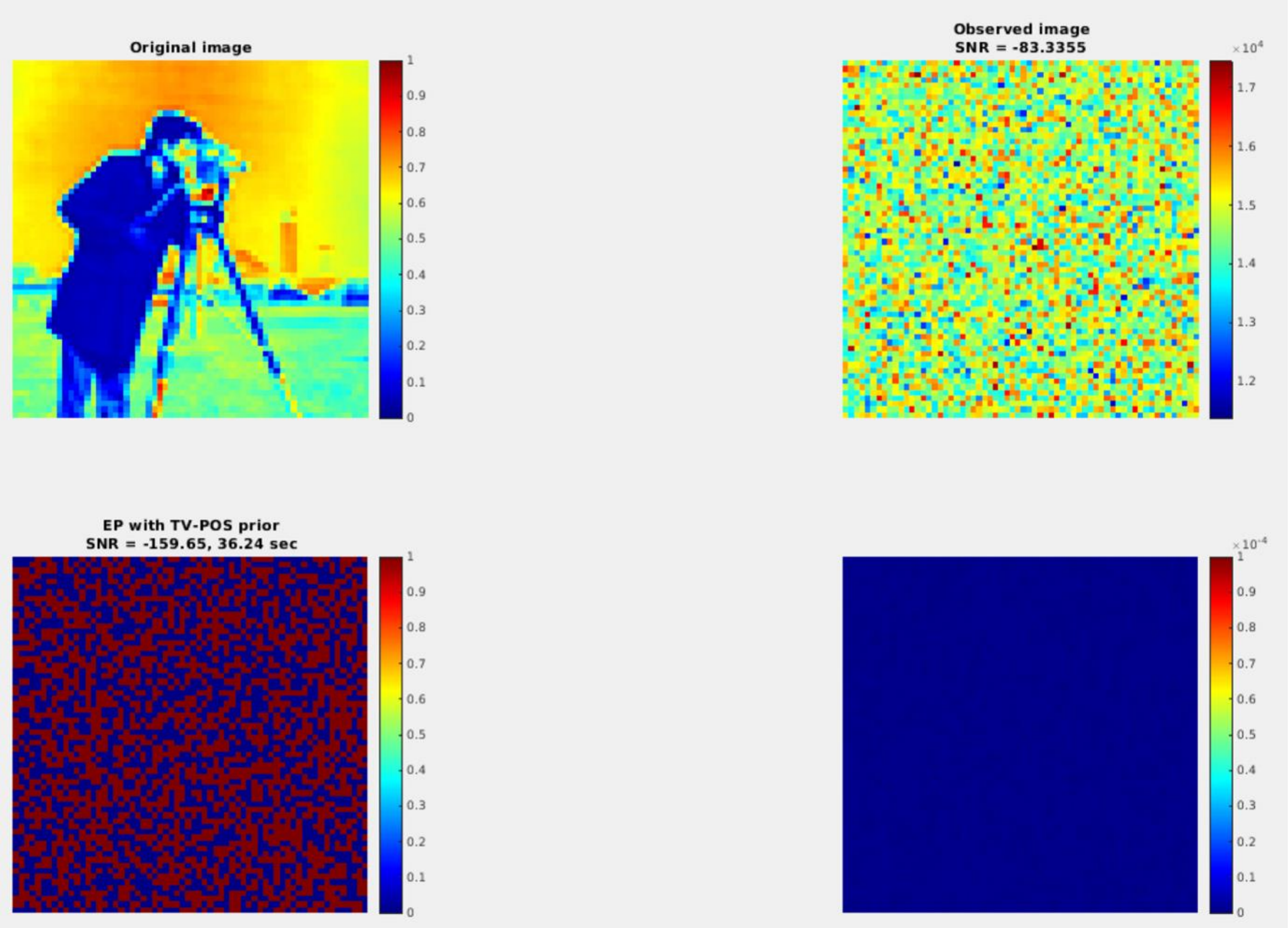
Imaging results with EP



Is our approach Robust in this context?

- What happens when the configuration of the fibre changes?
- New measurements ***y*** will no longer correspond to the pre-recorded sensing matrix **A**!
- What happens if we image with the **A** computed during calibration?

Imaging results with EP



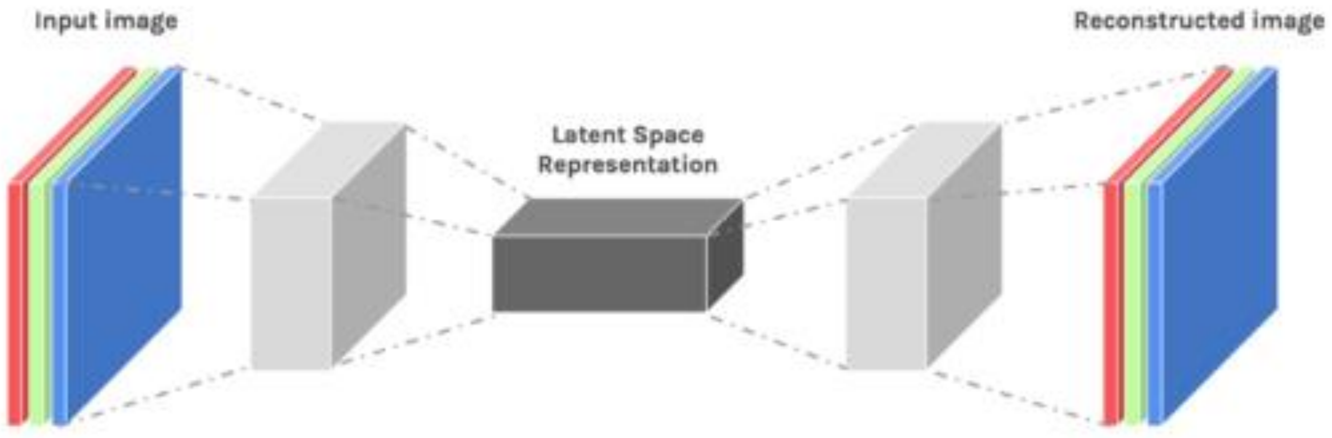
Our solution is not Robust!

- This implies we need to re-record the sensing matrix **A** for each new configuration of the fibre!!!

- Solution
  – Build an **A**gnostic imaging system using Autoencoder approach
  – Leverage a variational autoencoder with Gaussian mixture latent space

Autoencoders

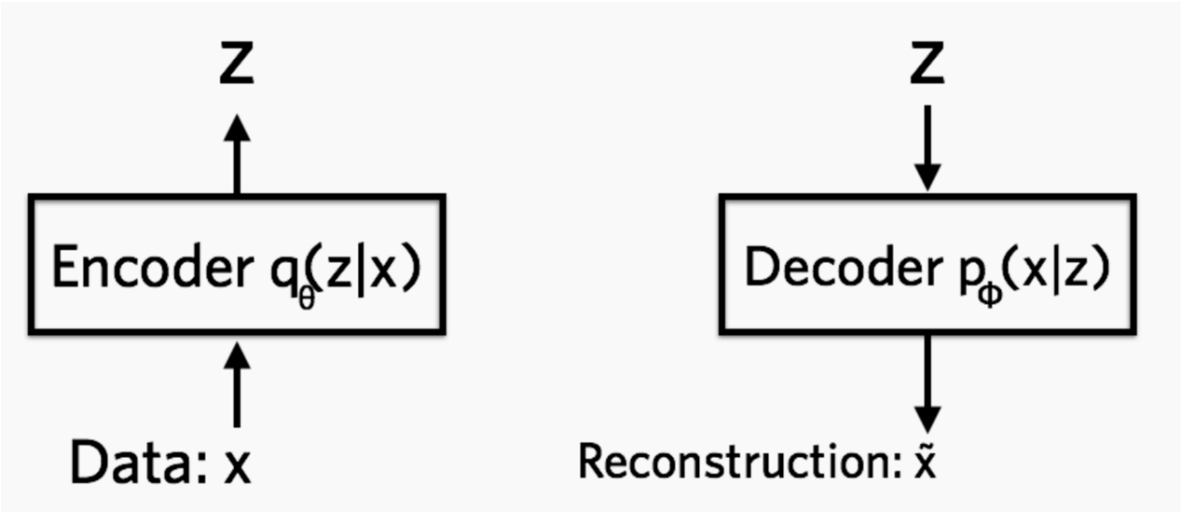
- Autoencoders are designed to reproduce their input, especially for images.
  - Key point is to reproduce the input from a learned encoding.



https://www.edureka.co/blog/autoencoders-tutorial/

Variational Autoencoder (VAE)

- Key idea: make both the encoder and the decoder probabilistic.
- The latent variables, z, are drawn from a probability distribution depending on the input, x, and the reconstruction is chosen probabilistically from z.



https://jaan.io/what-is-variational-autoencoder-vae-tutorial/

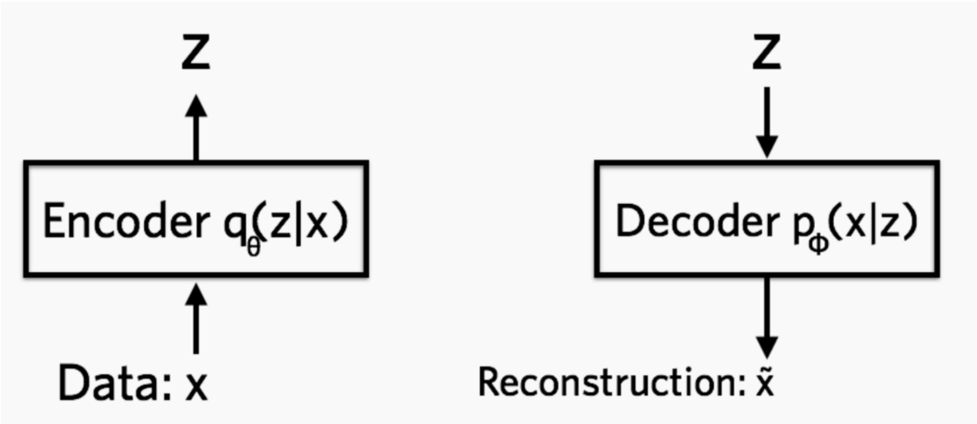
VAE Encoder

- The encoder takes input and returns parameters for a probability density (e.g. Gaussian) $q_\theta(z \mid x)$: gives the mean and co-variance matrix.
- We can sample from this distribution to get random values of the lower-dimensional representation *z*.
- Implemented via a neural network: each input *x* gives a vector mean and diagonal covariance matrix that determine the Gaussian density
- Parameters θ for the NN need to be learned – need to set up a loss function.

https://jaan.io/what-is-variational-autoencoder-vae-tutorial/

VAE Decoder

- The decoder takes latent variable *z* and returns parameters for a distribution. $p_\phi(x|z)$ gives the mean and variance for each pixel in the output.
- Reconstruction $\tilde{x}$ is produced by sampling.
- Implemented via neural network, the NN parameters ϕ are learned.

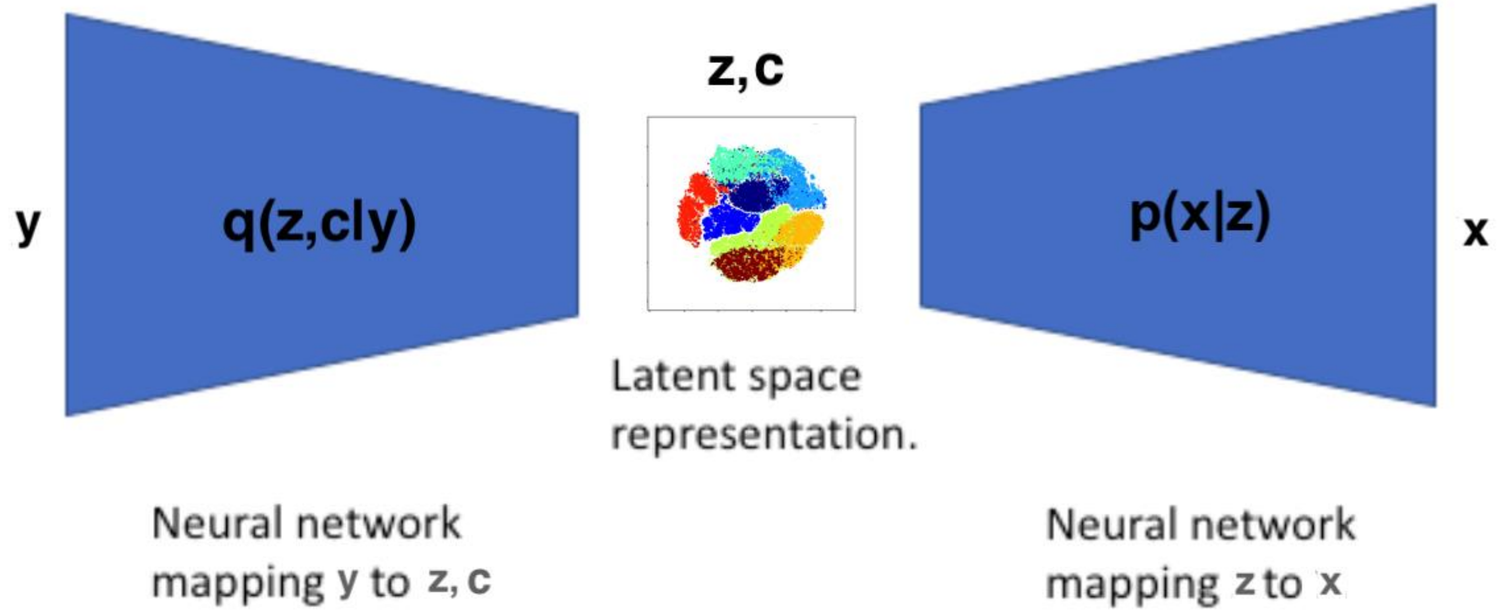


https://jaan.io/what-is-variational-autoencoder-vae-tutorial/

Gaussian Mixture Variational Auto Encoder (GMVAE)

$\boldsymbol{c} \sim \mathrm{p}(\boldsymbol{c})$, $\boldsymbol{z} \sim \mathrm{p}(\boldsymbol{z}|\boldsymbol{c})$, $\boldsymbol{x} \sim \mathrm{p}(\boldsymbol{x}|\boldsymbol{z})$

Where ***c*** follows a categorical distribution with *k* classes. p(***z***| ***c***) is a Gaussian conditioned on ***c***, and marginally p(***z***) follows a Gaussian mixture distribution with *k* classes.



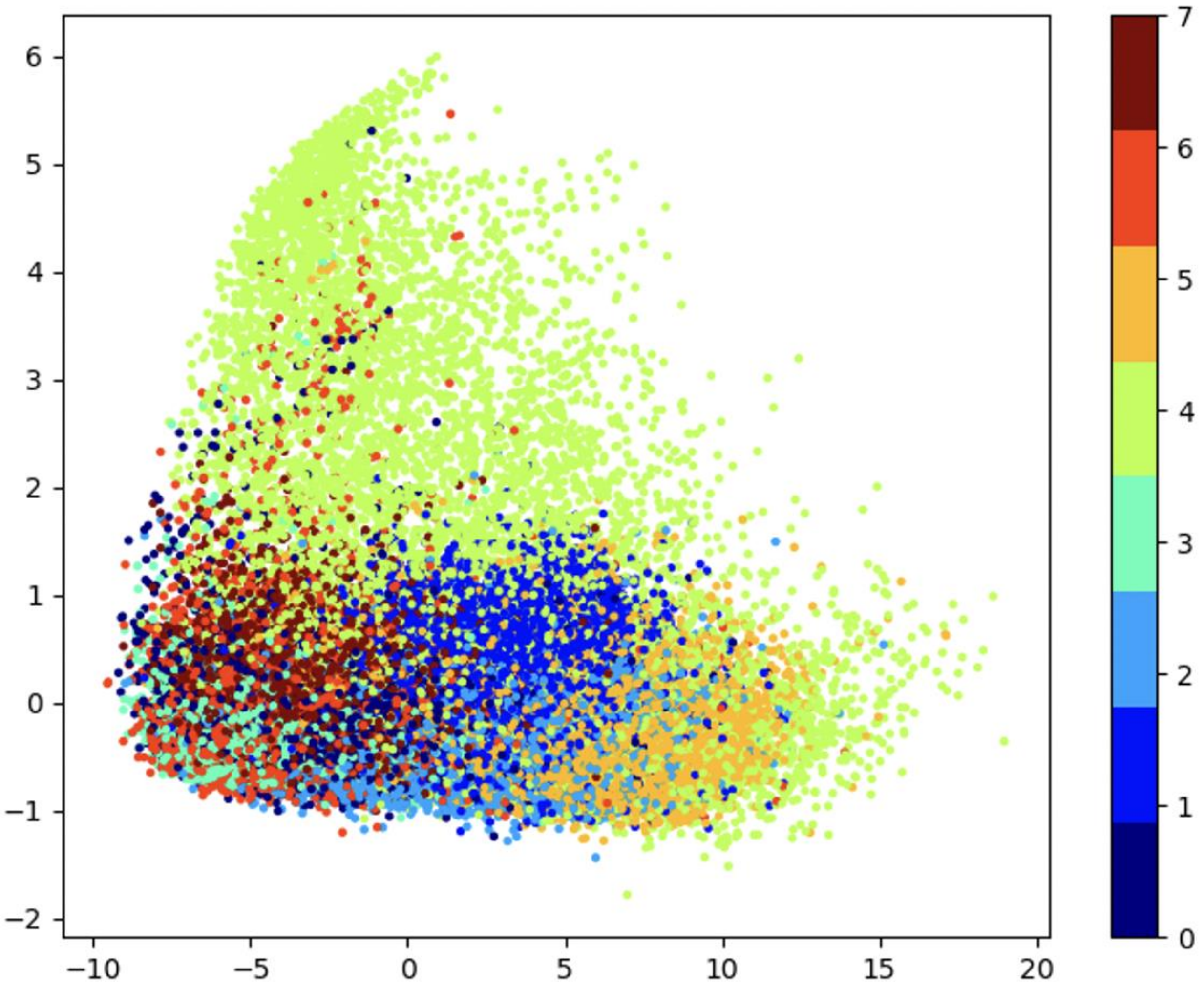
Loss = reconstruction error+KL(q(***z***|***y***, ***c***)||p(***z***|***c***))+KL(q(***c***|***y***)||p(***c***))

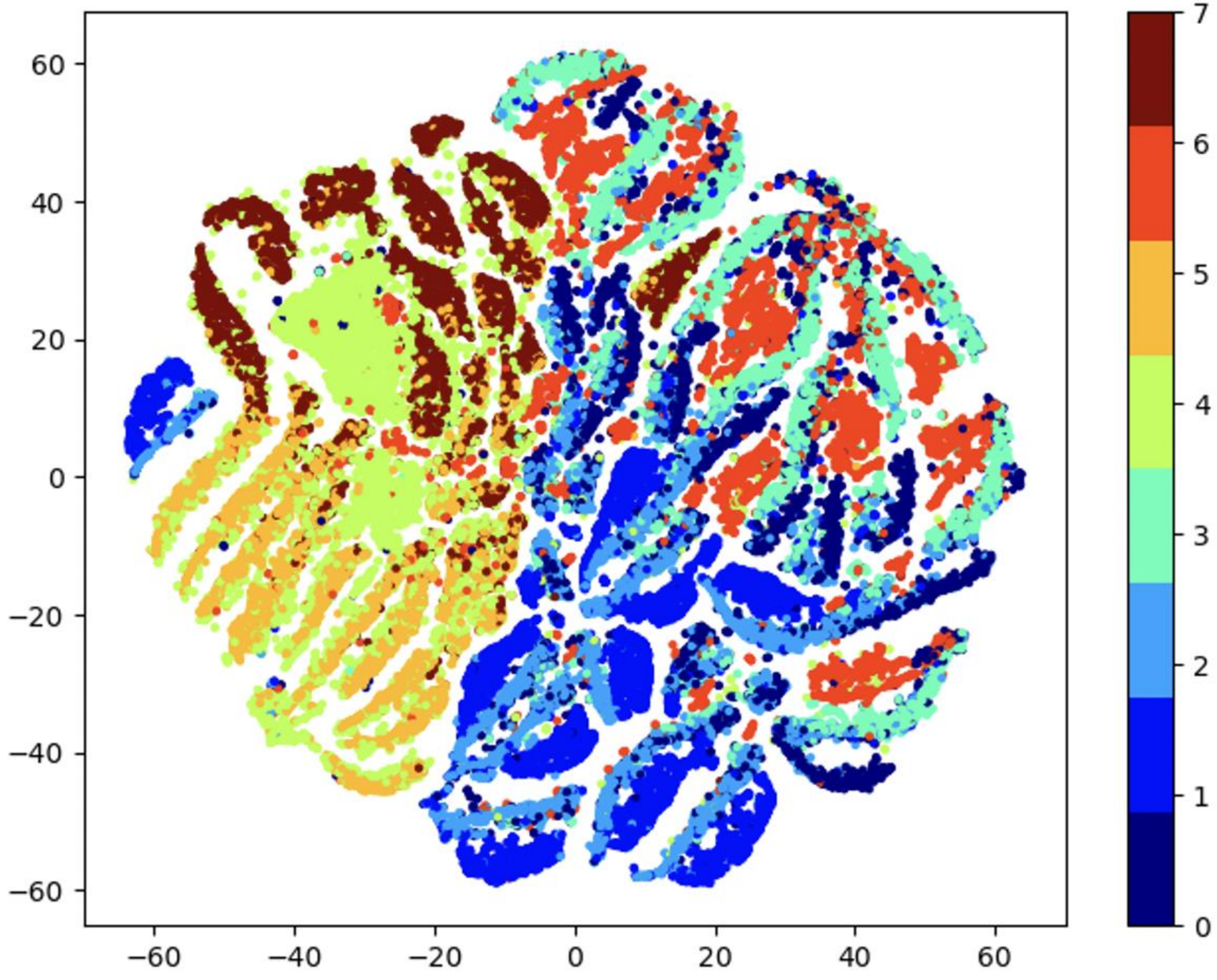
Experiment

- 24000 images of size 64x64 pixels from 8 classes of the fashion-MNIST dataset (3000 images per class).
- Record 4096 measurements for each image for 6 different displacements of a MMF (the displacements are 50, 60, 70, 80, 90 and 100 degrees). The measurements are reshaped to size 64x64.
- Training dataset is of size 144000x64x64.
- Test on measurements corresponding to a new displacement (75 degrees).
- Testing dataset is of size 10000x64x64; measurements corresponding to 1000 test image for 10 classes of the fashion-MNIST dataset (8 trained-on classes and 2 new classes).

Results

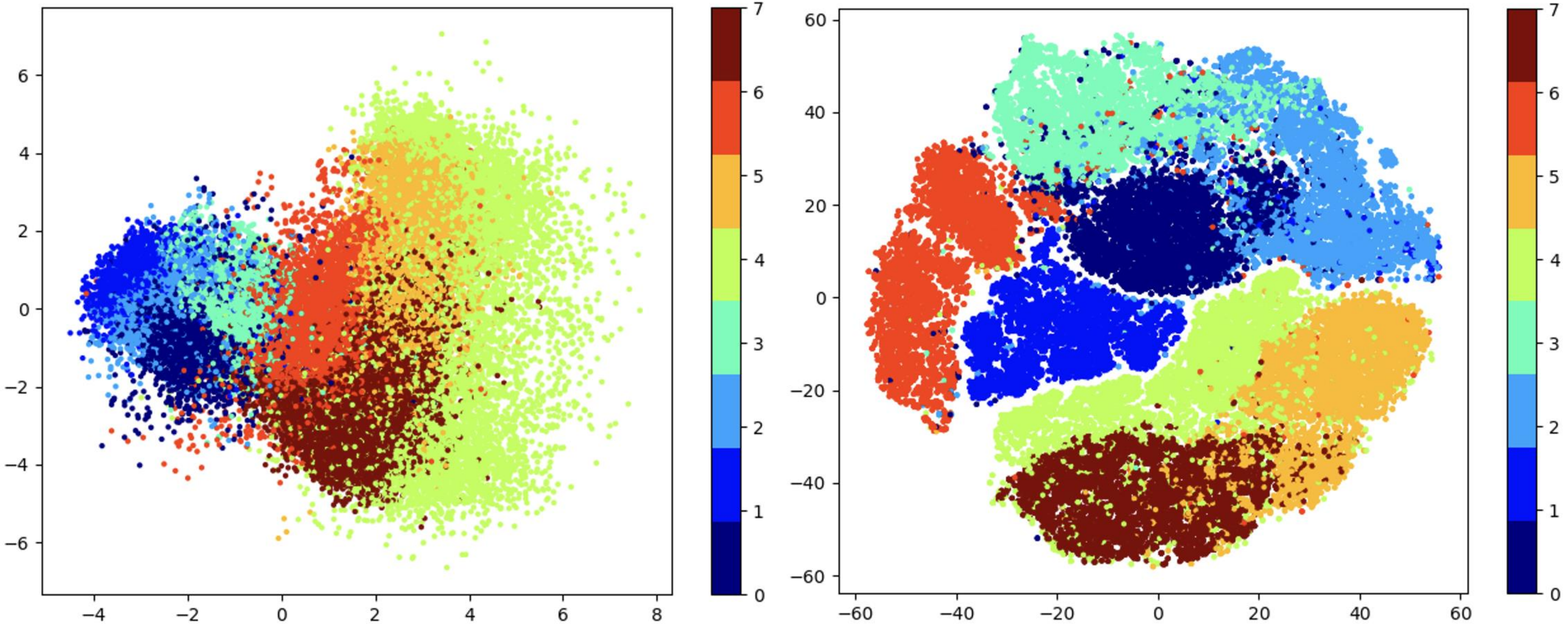
2D projections of training data - PCA (left) and t-SNE (right):





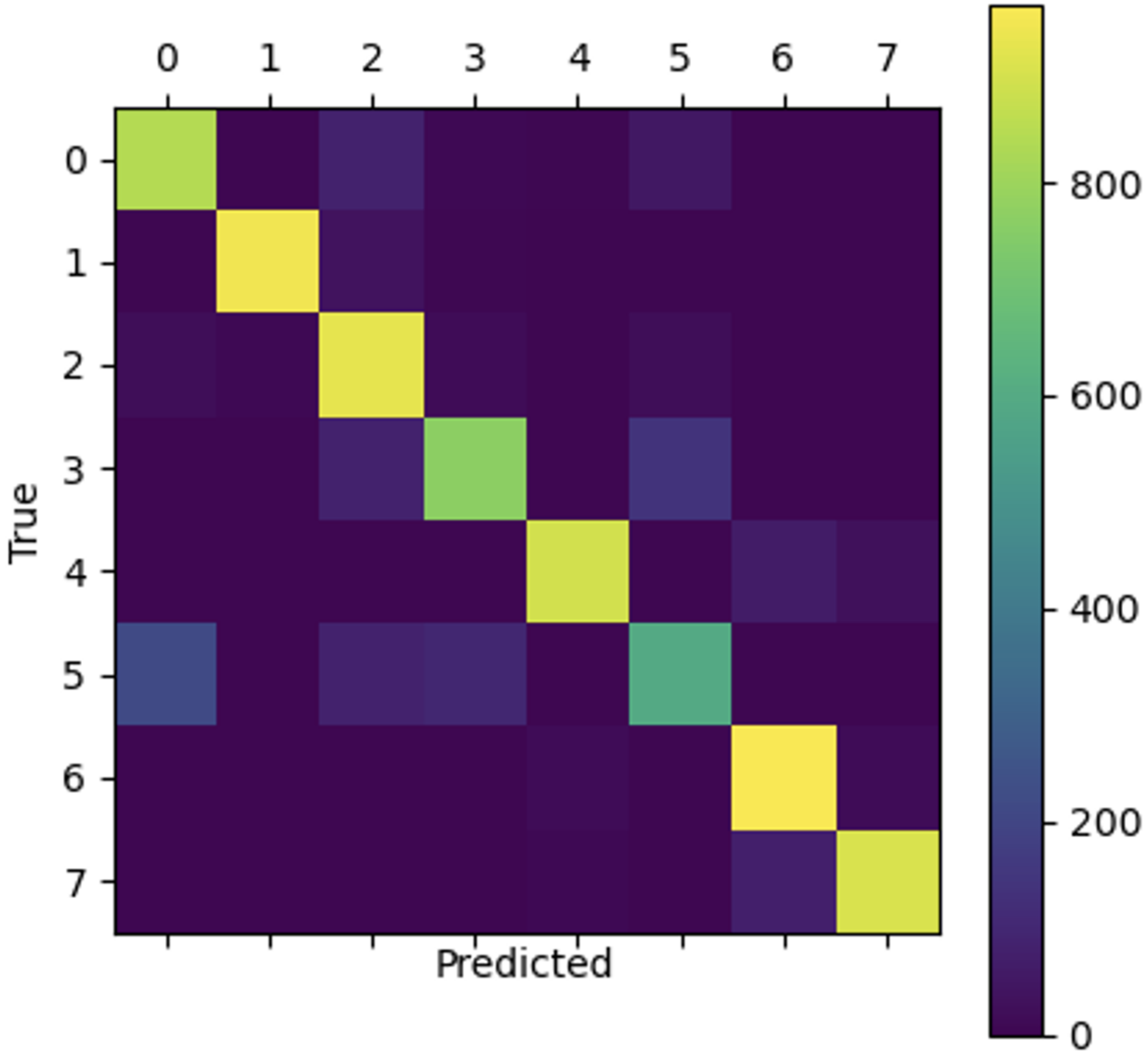
Results

2D projections of latent space of training data - PCA (left) and t-SNE (right):



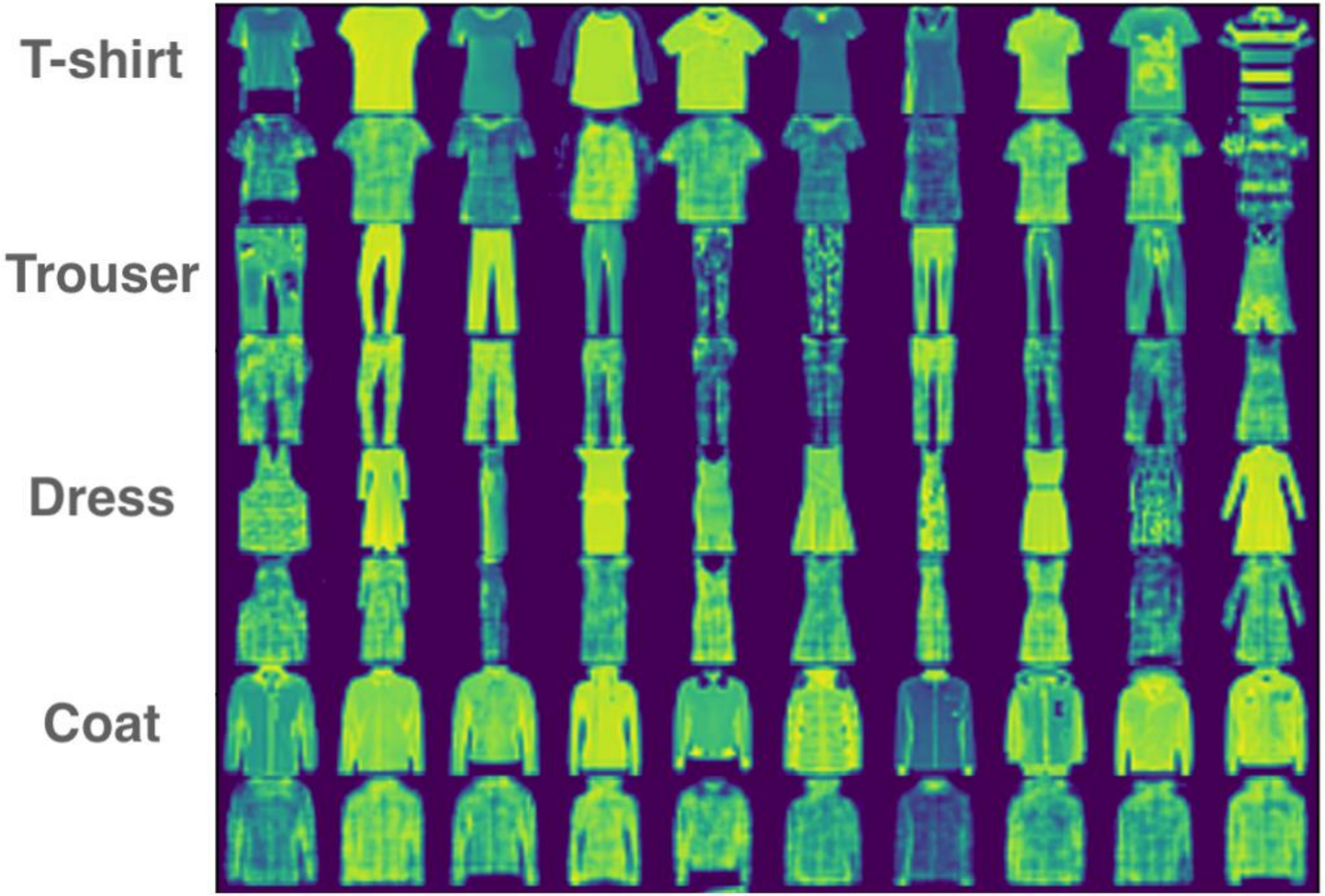
Results

Classification accuracy = 85.8%, computational time = 0.1sec for 1000 images



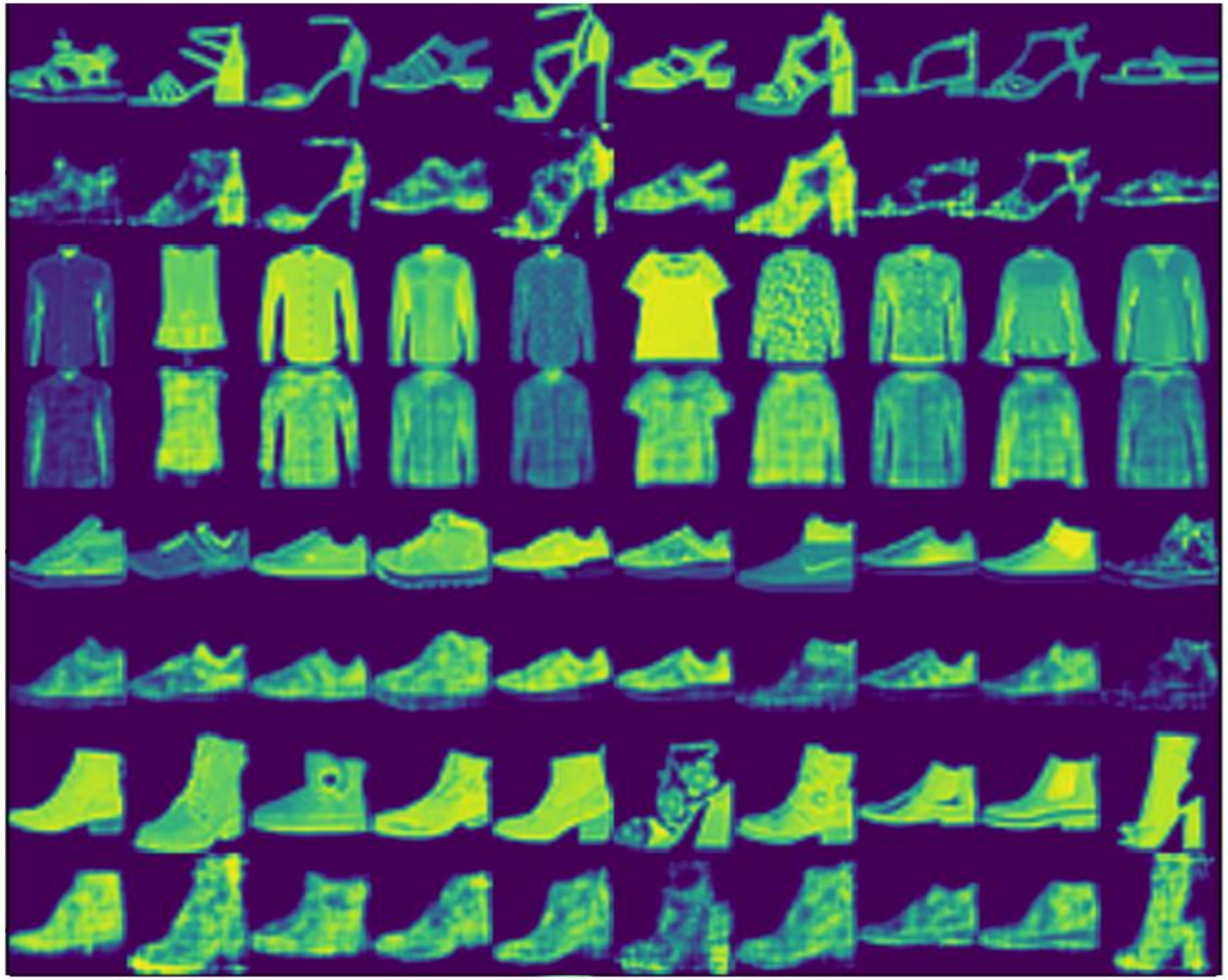
Results

Reconstruction results for new displacement (75) - test images from trained-on classes
- first row of each class: ground-truth; second row of each class: estimated



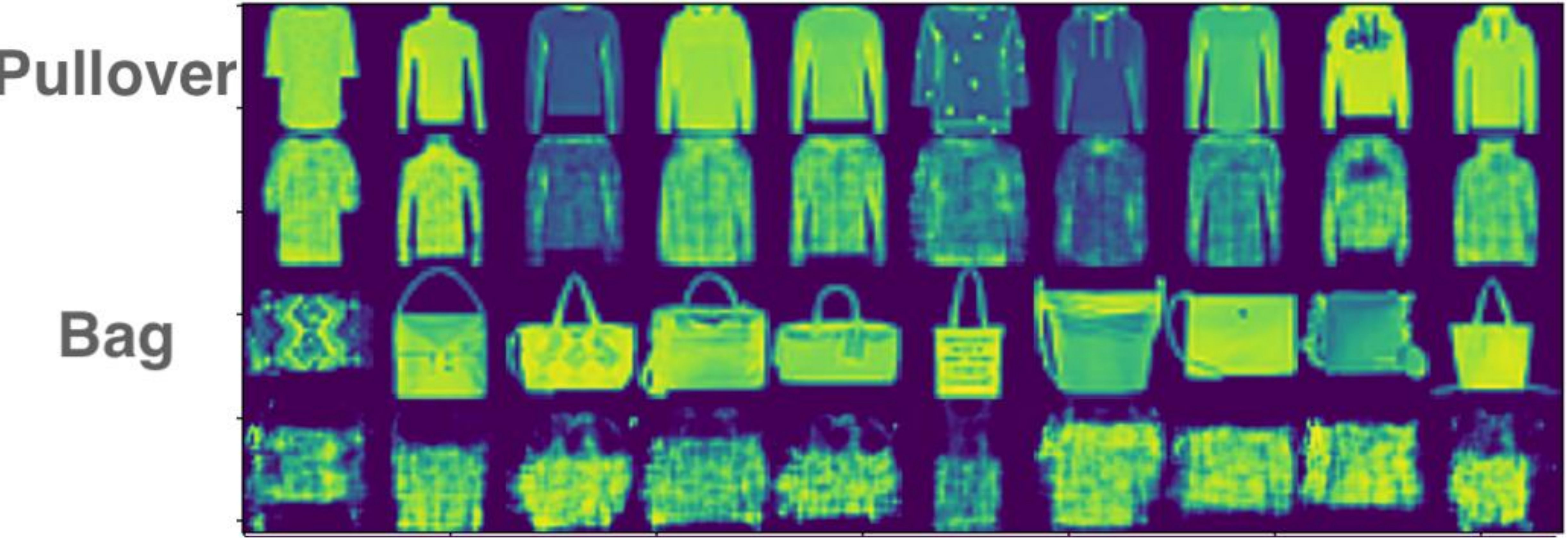
Results

Reconstruction results for new displacement (75) - test images from trained-on classes
- first row of each class: ground-truth; second row of each class: estimated



Results

Reconstruction results for new displacement (75) - test images from trained-on classes
- first row of each class: ground-truth; second row of each class: estimated



Observations….

- Low illumination regime
  - Large uncertainties
  - Non-Gaussian regime
- Variational methods for faster inference
  - Tractable approximations
  - PnP for scalability
- Tradeoff accuracy/robustness
  - Model mismatch (e.g. shielding, turbulence)
- Guarantees with PnP approaches...

Conclusions

To fully exploit technological advances means we need to go beyond the traditional decoupled imaging pipeline requiring separate consideration of

- device physics
- signal processing, and
- end-user application

and rethink imaging as an integrated sensing and inference model

- Some open questions…
  – How can we best combine data driven and (physical) model-based approaches?
  – How to adapt to nonstandard acquisition systems, e.g. event-based cameras?
  – What can we do with little or no ground truth data?
  – What are the limits?

Thanks for your attention!

Contact: S.Mclaughlin@hw.ac.uk,Y.Altmann@hw.ac.uk