Constraining gradient-based inversion with a variational autoencoder to reproduce geological patterns

Jorge Lopez-Alvis[1,2], Eric Laloy[3], Thomas Hermans[2], Frédéric Nguyen[1]

[1]Applied Geophysics, University of Liège, Belgium

[2]Department of Geology, Ghent University, Belgium

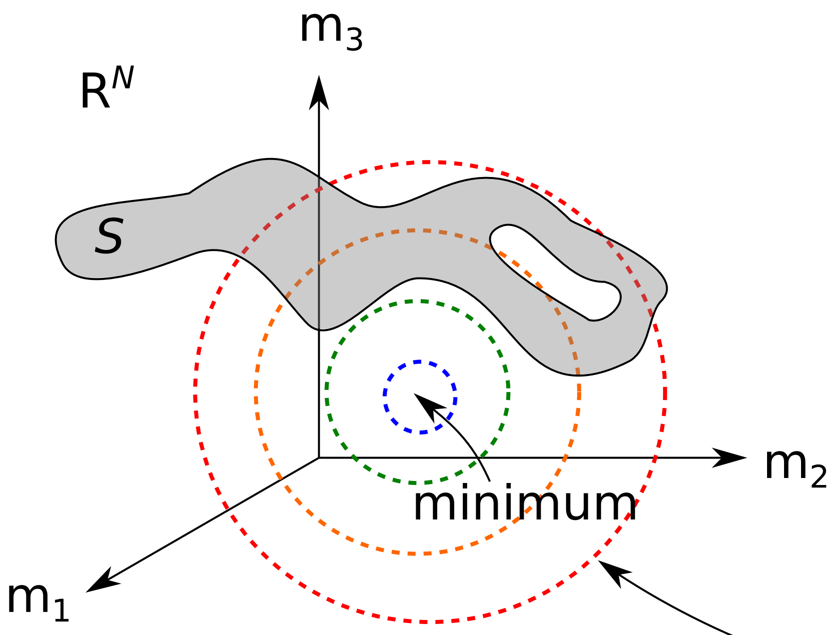
[3]SCK-CEN (Belgian Nuclear Research Center), Belgium

EGU2020: Sharing Geoscience Online

Inversion with structured models

Structured models lie in a *n*-dimesional manifold *S* embedded in high-dimensional model space R^N (e.g. *N* is the number of cells or pixels), where $n << N$.

A vector **m** is any possible model in R^N, but only vectors **m** lying in *S* exhibit the desired patterns (structures).



Then inversion problem is:

$$\min_{\mathbf{m}} \|f(\mathbf{m}) - \mathbf{d}\|^2$$

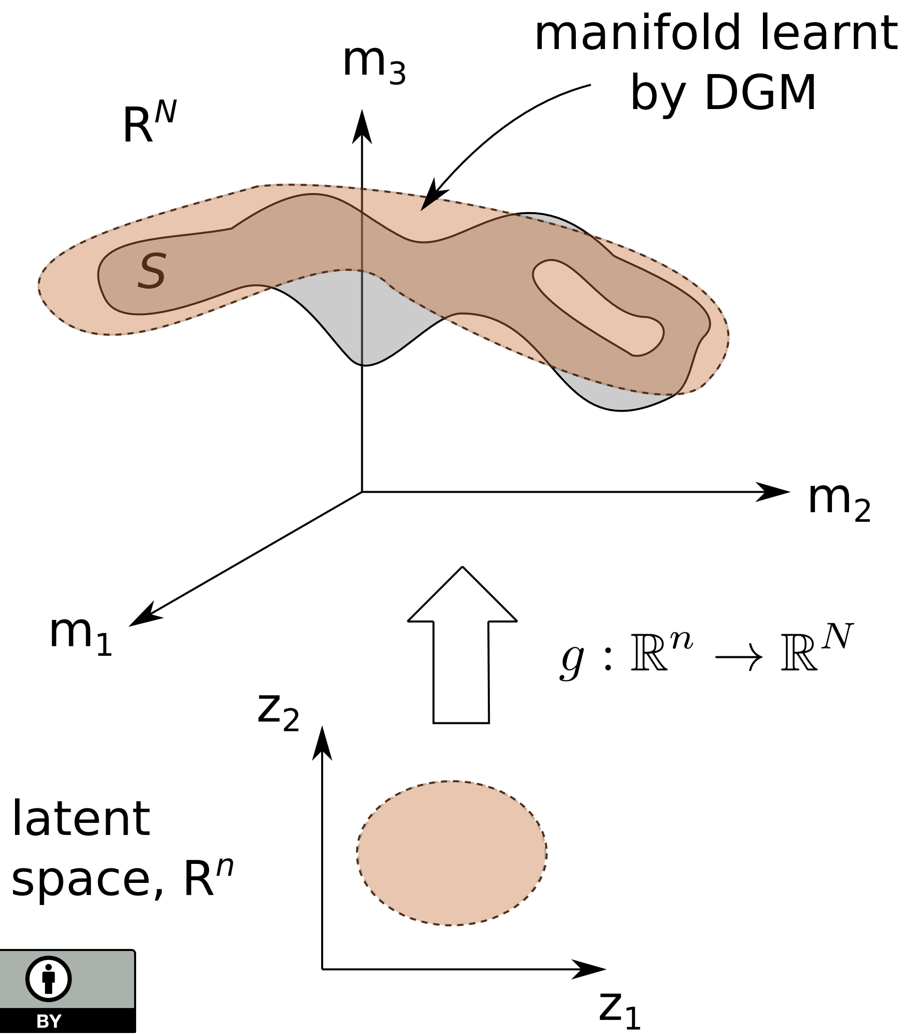
$$s.t.\ \mathbf{m} \in S$$

where f(**m**) is the forward operator and **d** is the measured data.

contours of objective function, defined in R^N.

Approximate the manifold

It is not easy to learn such manifolds for realistically structured models (e.g. geological media) but good approximations have been done recently with deep generative models (DGMs).



The setup is for DGMs is:

$$p(\mathbf{z}) = \mathcal{N}(\mathbf{0}, \mathbf{I}) \qquad p(\mathbf{m}|\mathbf{z}) = \mathcal{N}(g_{\boldsymbol{\theta}}(\mathbf{z}), \sigma^2_{\boldsymbol{\theta}}(\mathbf{z})\mathbf{I})$$

once trained, samples lying approximately on the manifold are generated by sampling first p(**z**) and then p(**m**|**z**).

p(**z**) is a chosen distribution from which it is easy to sample from.

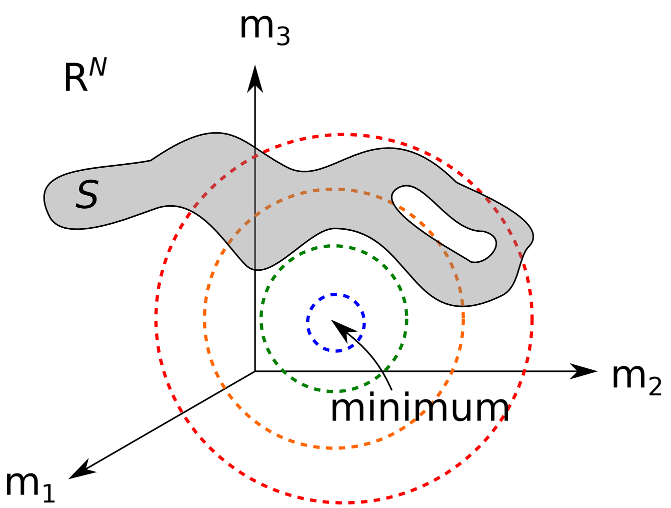
For this discussion, $\sigma^2_{\boldsymbol{\theta}}(\mathbf{z})$ can be assumed zero.

$g_{\boldsymbol{\theta}}(\mathbf{z})$ is the (mean) decoder or generator and for DGMs is a DNN (usually with convolutional layers).

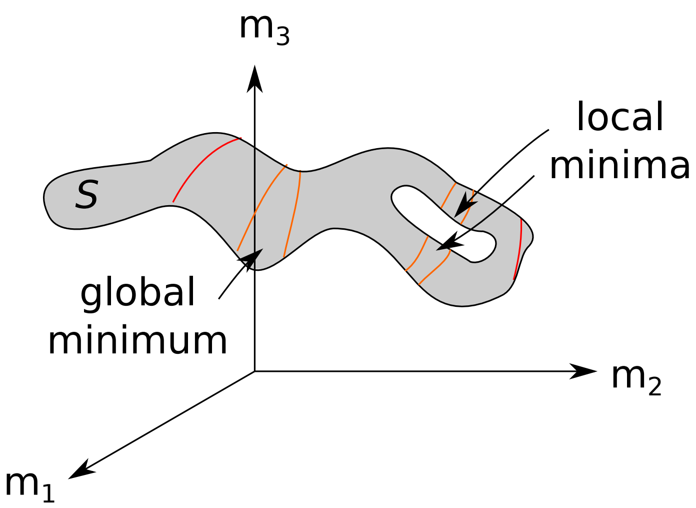
Therefore, $\boldsymbol{\theta}$ is the vector of parameters for such DNN and is estimated by different methods (e.g adversarial learning for GANs and variational inference for VAEs).

Objective function in the manifold

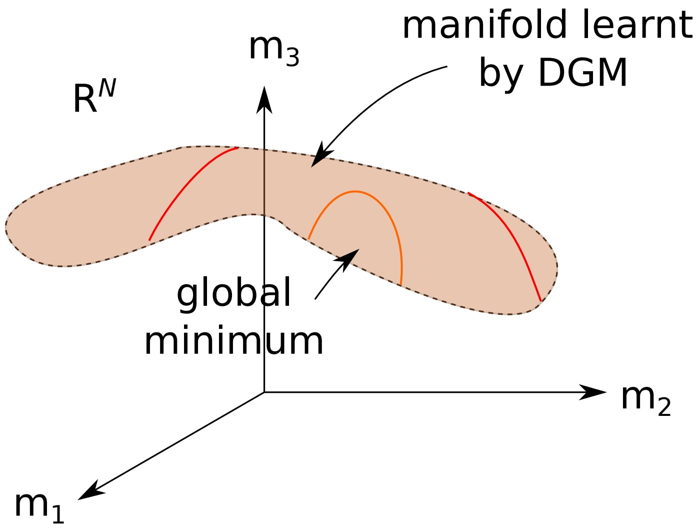
Constraining the inversion to the manifold may cause local minima even if the objective function in R^N is convex, e.g. $||f(\mathbf{m}) - \mathbf{d}||^2$ with a linear f.



Contours of convex objective function, defined in R^N.



Contours of objective function intersected by manifold *S*, notice the two local minima.



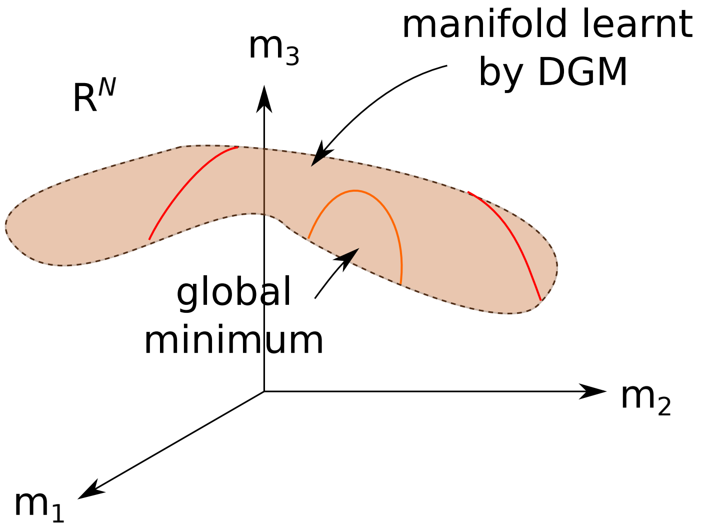
The situation improves when choosing an adequate approximate manifold, but one should be aware of error.

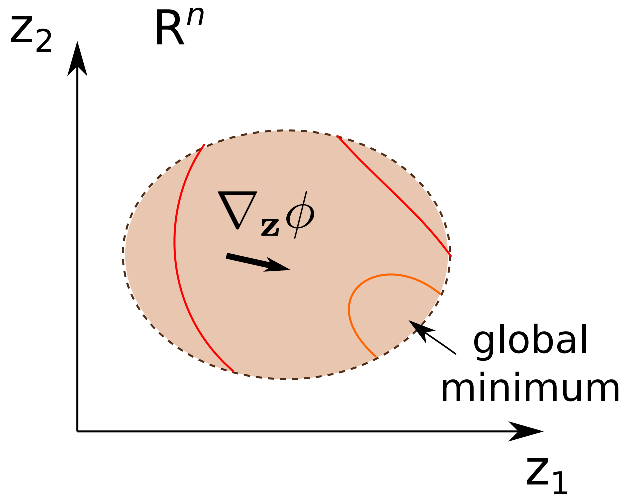
There is a tradeoff between the number of local minima and the accuracy of the patterns (i.e. staying close to the *real* manifold).

The issue: standard gradient descent methods usually converge to local minima.

Gradient descent on the manifold

In order to stay on the learnt manifold, gradient descent can be performed in the latent space (see Laloy et al. 2019, *Computers and Geosciences*).





The objective function is:

$$\phi(\mathbf{m}) = \|f(\mathbf{m}) - \mathbf{d}\|^2$$

$$\phi(\mathbf{z}) = \|f(g_{\boldsymbol{\theta}}(\mathbf{z})) - \mathbf{d}\|^2$$

Latent space may be seen as an Euclidian space R^n to where the manifold is embedded.

Then a gradient descent step is performed as:

$$\mathbf{z}_k = \mathbf{z}_{k-1} - \ell \cdot \nabla_{\mathbf{z}}\phi$$

where k is the iteration number and ℓ is the step size (or learning rate).

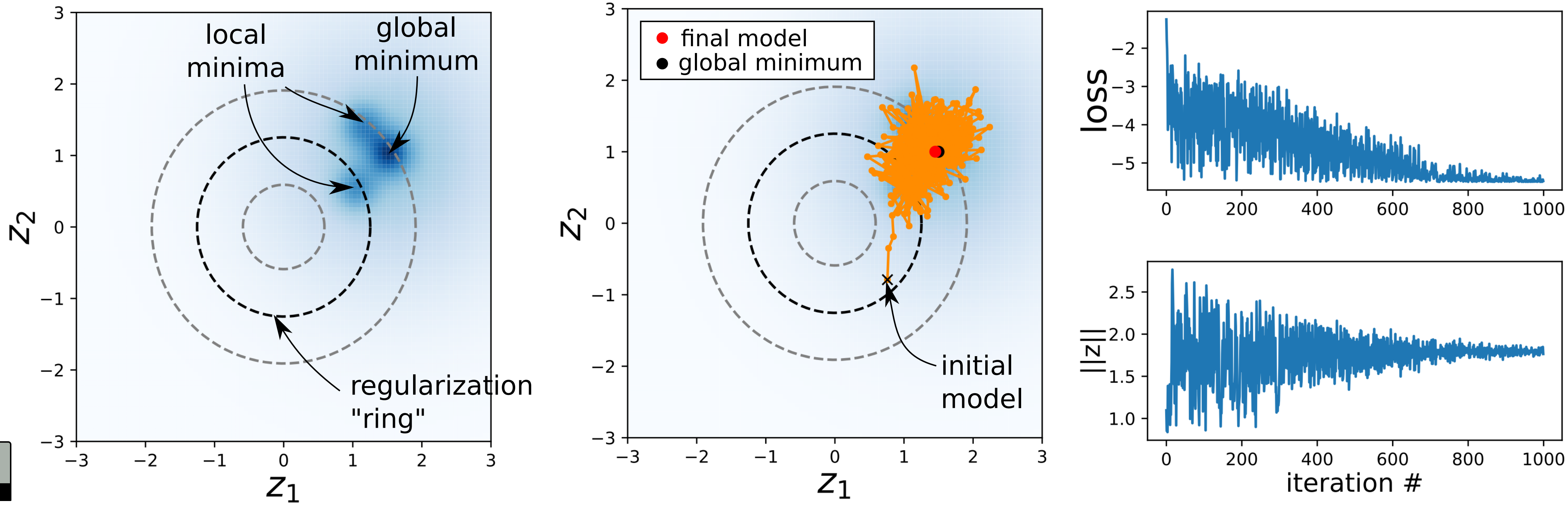
Stochastic gradient descent on the manifold

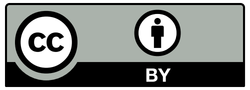
Even when we choose a manifold that does not produce too much local minima, it will usually not result in a single global minimum.

Stochastic gradient descent (SGD) is less likely to get trapped in local minima.

We propose to use a *decreasing* learning rate such that SGD will perform good when a global "basin of attraction" is present.

In order to stabilize this SGD in the initial phase (when steps are large), we propose a *decreasing* regularization to the regions of higher density of samples in the latent space (i.e. as enforced by p(**z**)).





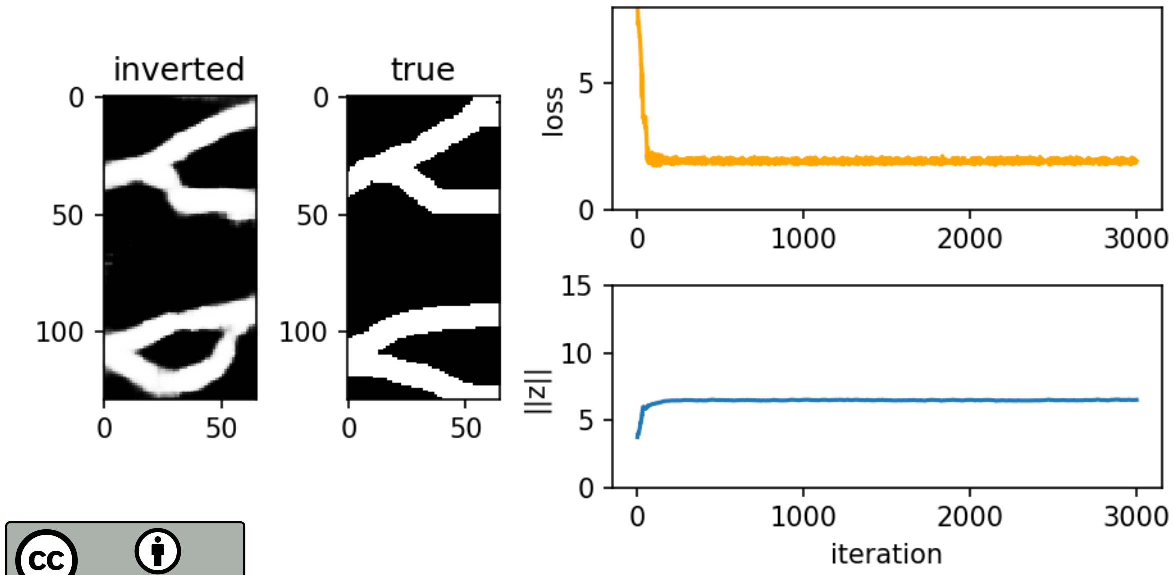
Stochastic gradient descent on a realistic manifold

A realistic case was considered with *channelized* patterns and a linear forward model (linear traveltime tomography) for which the objective function is convex in R^N but not in the manifold (in this case N=8325 and n=20).

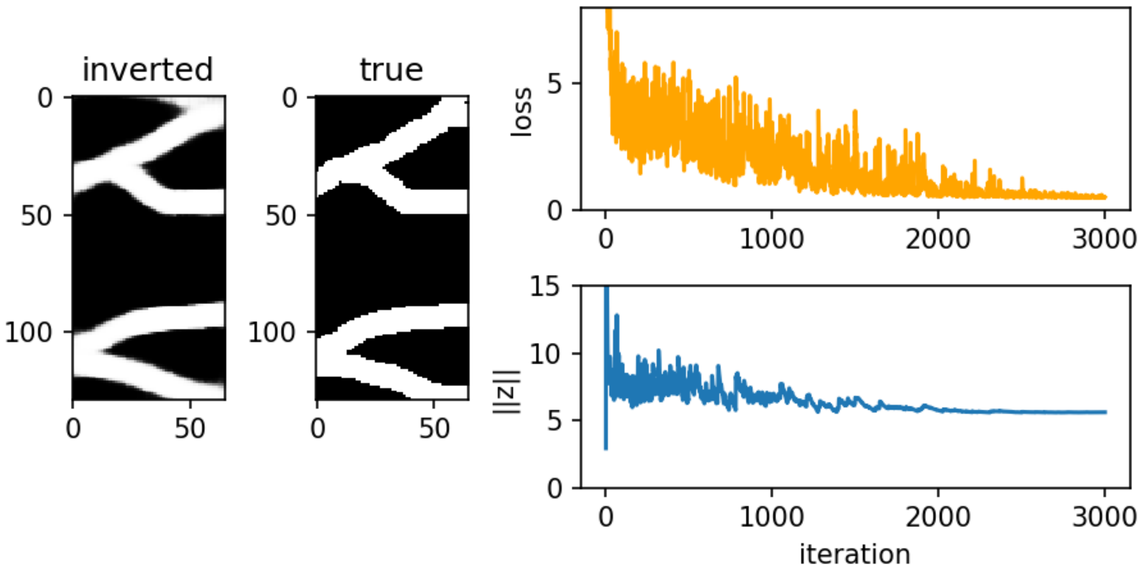
As previously suggested, to handle the problem with local minima we propose to use:

1) VAE as DGM
2) SGD with decreasing learning rate and regularization on the latent space.

With regular GD, the optimization sometimes gets trapped in local minima.

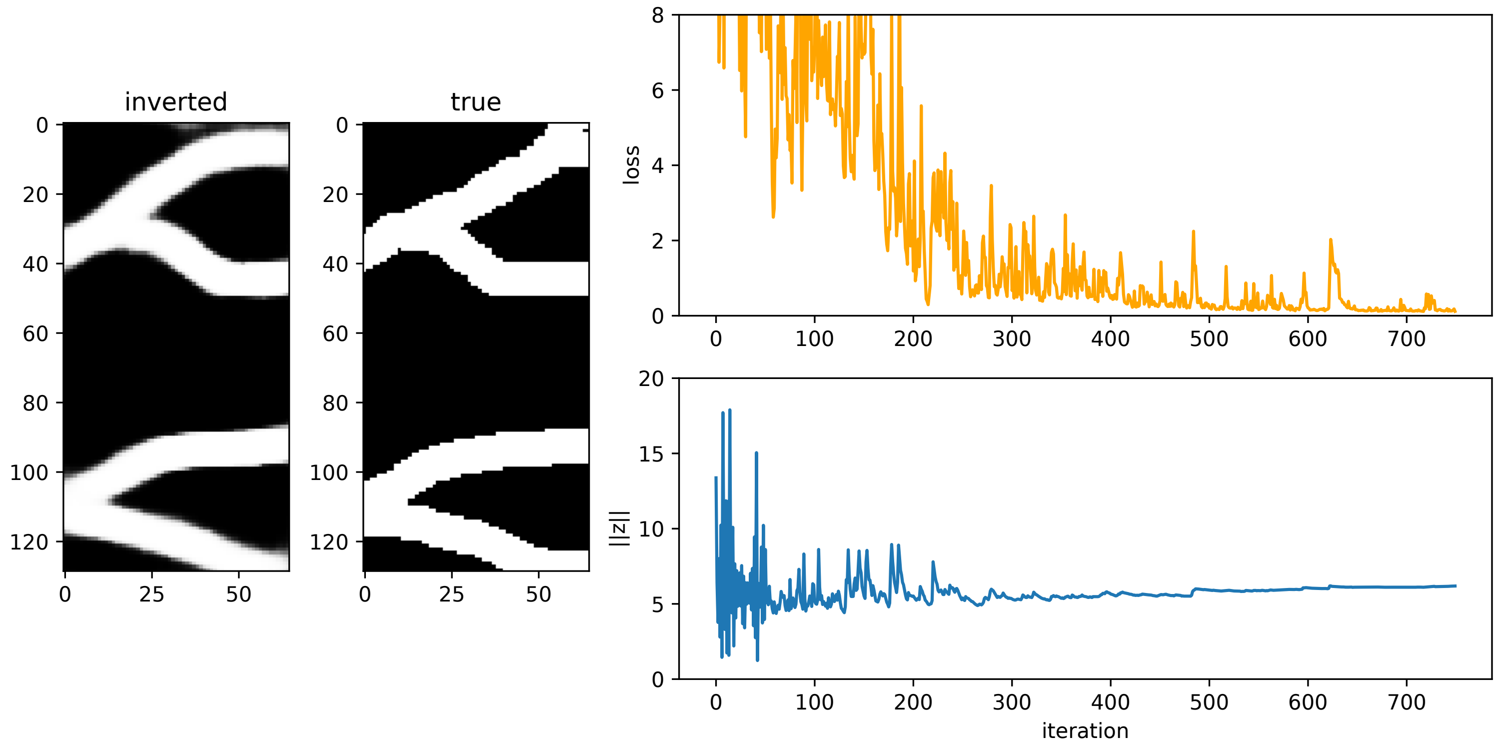


With the proposed SGD, the optimization converges to the global minimum with high probability.



Stochastic gradient descent on a realistic manifold (nonlinear forward)

Even with a nonlinear forward model (shortest path method), where the nonlinearity usually introduces further local minima, the proposed method performs good.



Given the higher computational demand of the nonlinear forward model, we reduced the number of iterations and adjust the decreasing factors for the learning rate and the regularization accordingly (this may increase the probability of converging to local minima).

Conclusions

Inversion with structured models may be pursued by using gradient-descent in the latent space of a deep generative model (DGM). However, local minima may arise when optimization is constrained to the manifold learnt by the DGM.

To deal with such local minima, we propose to consider a variational autoencoder (VAE) latent space and stochastic gradient descent (SGD) with a decreasing learning rate and regularization in the latent space.

A realistic case with channelized patterns shows that the proposed method converges to the global minimum with high probability for both linear and nonlinear forward models.

The same case shows that errors incurred by the *imperfect* pattern reproduction of the VAE decoder are minor with respect to the overall reconstruction of the real model by inversion.

ITN Enigma

This work has received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement number 722028 (ENIGMA ITN).