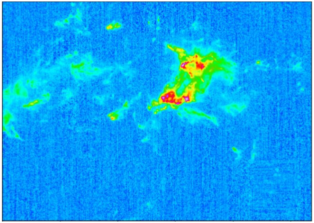
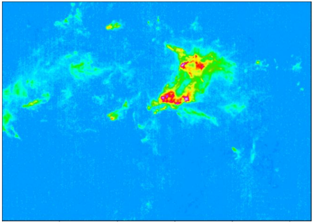
Deep learning denoising by dimension reduction

Application to the ORION-B line cubes

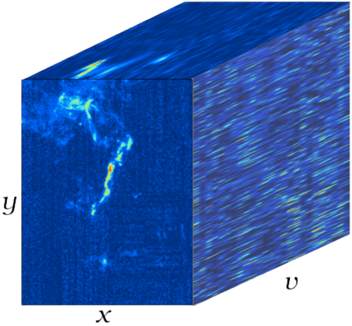
L. Einig, J. Pety, A. Roueff, P. Vandame, J. Chanussot, M. Gerin
and the ORION-B consortium





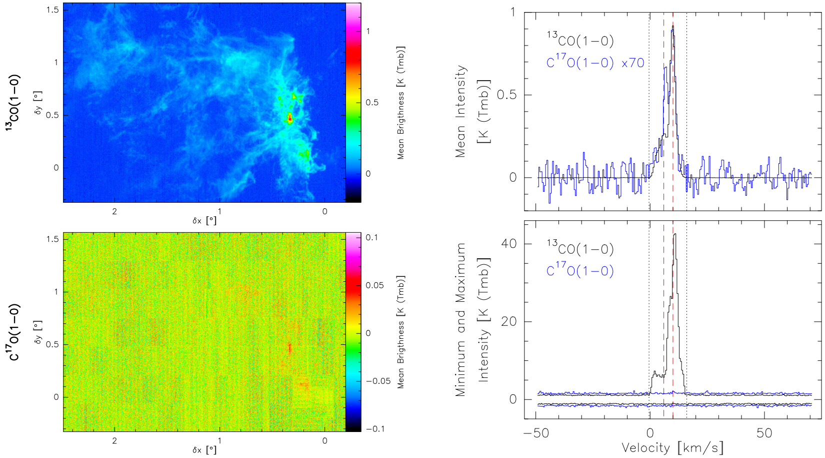
The ORION-B dataset

- Acquired by the wide-band receiver at the IRAM-30 m
- $\sim$ 30 molecular line cubes for J=1-0 transition
- Spatially and spectrally resolved
- 1074 $\times$ 758 profiles
- 240 velocity channels



^{13}CO (1-0) line cube (small part only).

The ORION-B dataset

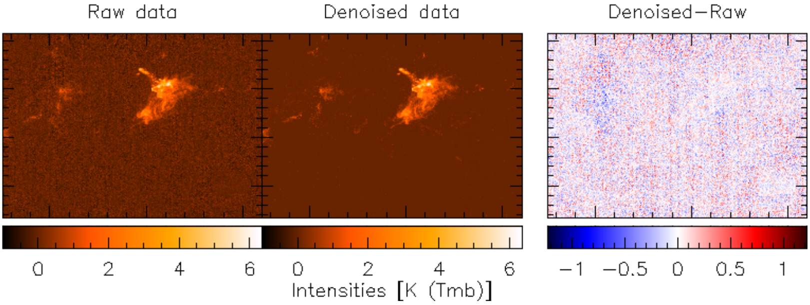


Integrated intensity **(left)** and mean, min and max spectra **(right)** of $^{13}\mathrm{CO}$ (1-0) and $\mathrm{C}^{17}\mathrm{O}$ (1-0).

About denoising

Interest of denoising

- Increasing the signal-to-noise ratio is an important step to lead to discoveries.
- Necessary to find statistical relations between certain lines and physical parameters (otherwise hidden by noise).



About denoising

Assumption

- There is a "true" signal s which can only be estimated through measurements giving data $d = f(s)$.
- **Example:** in the case of an additive noise, we have $d = s + n$, with n being randomly distributed.

About denoising

Assumption

- There is a "true" signal s which can only be estimated through measurements giving data $d = f(s)$.
- **Example:** in the case of an additive noise, we have $d = s + n$, with n being randomly distributed.

Hyperspectral cubes denoising

- Important topic in remote sensing and many algorithm has be developed since decades.
- Methods essentially developed with Earth images.
- What about molecular line cubes?

Denoising by dimension reduction

Intrinsic vs extrinsic dimension

- The **extrinsic dimension** n of a dataset is its apparent number of features.
- The **intrinsic dimension** m is the minimal number of features that can generate the dataset by mapping.
- If $m \ll n$, there is a lot of **redundancy** that can be exploited to "correct" the measurements.

Denoising by dimension reduction

Intrinsic vs extrinsic dimension

- The **extrinsic dimension** n of a dataset is its apparent number of features.
- The **intrinsic dimension** m is the minimal number of features that can generate the dataset by mapping.
- If $m \ll n$, there is a lot of **redundancy** that can be exploited to "correct" the measurements.

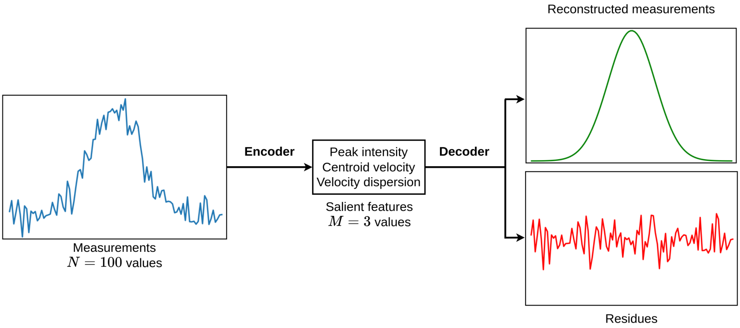
To be able to do this, we need:

- The intrinsic dimension of the measurements to be **much smaller** than the extrinsic dimension,
- To know the **relation** between features and measurements.

Denoising by dimension reduction

Illustration with the Gaussian fitting

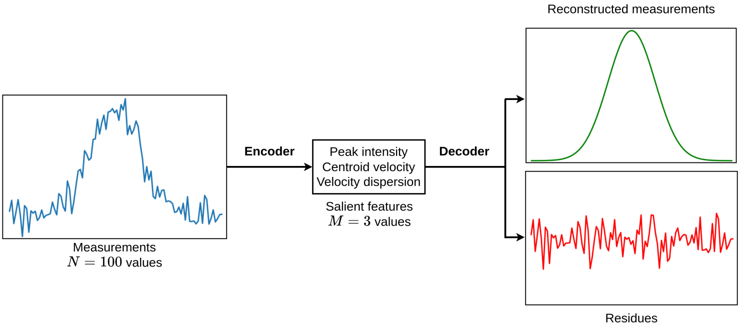
In the case of a Gaussian decomposition of a profile, we go from K velocity channels to only **3 values**:



Denoising by dimension reduction

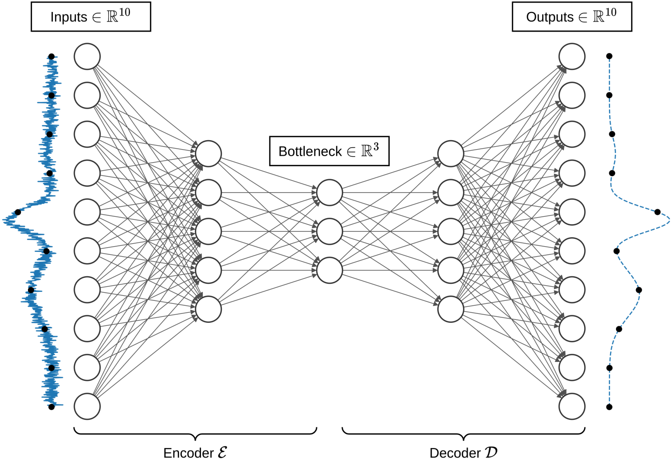
Illustration with the Gaussian fitting

In the case of a Gaussian decomposition of a profile, we go from K velocity channels to only **3 values**:



The encoder and decoder are set such that they minimize a "distance" between inputs and outputs. This is the **loss function**.

Autoencoder based denoising



Example of a autoencoder neural network with extrinsic and extrinsic dimensions of 10 and 3, respectively.

Autoencoder based denoising

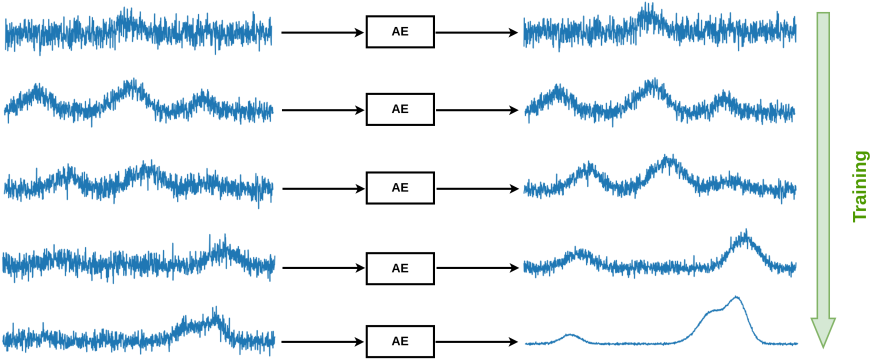
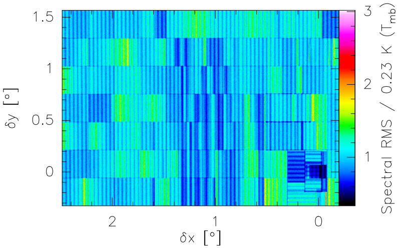
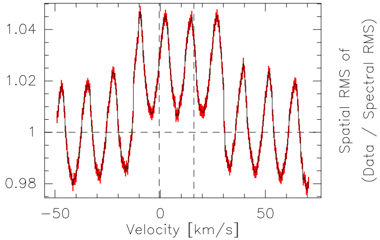


Illustration of AE training, profile by profile. The **same** $\mathcal{E}$ and $\mathcal{D}$ function is used for **every profiles**.

Noise model: noise RMS

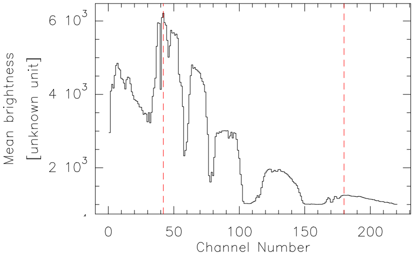


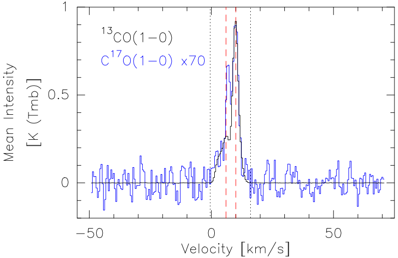


RMS of noise computed over the spectral axis **(left)** and over the spatial axes **(right)**

⟶ Noise intensity is pixel dependent **and** channel dependent

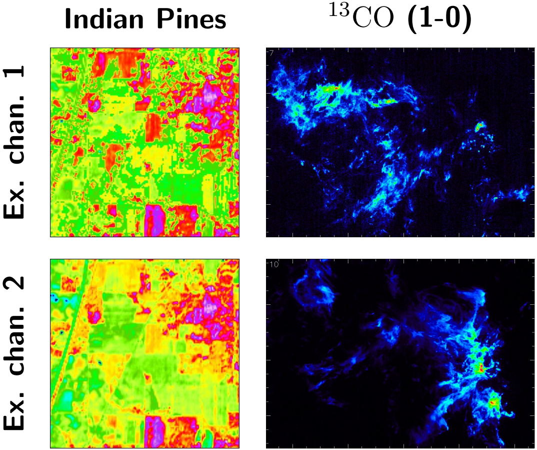
Molecular line cubes vs Earth RS images





Comparison between a hyperspectral Earth image profile named Indian Pines (continuum, **left**) and a line profile **(right).**

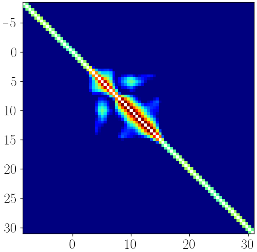
Molecular line cubes vs Earth RS images

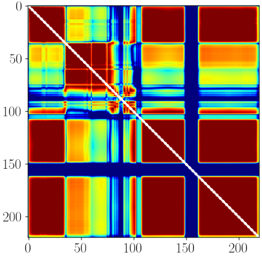


Examples of channels of Indian Pines **(left)** and $^{13}\mathrm{CO}$ (1-0) **(right)**. The latter seem to be independent.

Molecular line cubes vs Earth RS images

- Noise is pixel dependent, spectrally and spatially correlated
  $\longrightarrow$ false signal
- Unlike Earth remote sensing cubes, very low information redundancy
  $\longrightarrow$ need to make the best use of it





$^{13}\mathrm{CO}$ line and Earth image "Indian Pines" correlation matrices

A locally connected NN with prior information

Proposition 1: loss function to address the problem of sparcity

- The normal behavior of an autoencoder is **not** to put to exactly zero the signal free voxels.
- We want a loss function to enforce this behavior and to use the previous segmentation to help it to do this.
- To do so, we use **a prior** from a 3D segmentation method.

$$\mathcal{L}(\widehat{d}_{i,j}, d_{i,j}) = \frac{1}{K}\sum_{k=1}^{K}\begin{cases} \dfrac{(\widehat{d}_{i,j,k} - d_{i,j,k})^2}{\sigma_{i,j}} & \text{if probably signal + noise} \\ \left|\dfrac{\widehat{d}_{i,j,k}}{\sigma_{i,j}}\right|^q & \text{if probably only noise} \end{cases}$$

with $q \in]0,1]$ an hyperparameter than controls the sparcity.

A locally connected NN with prior information

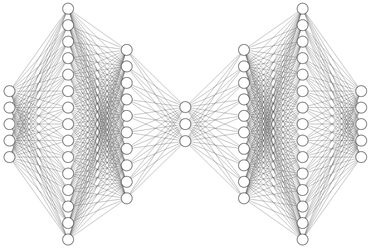
- Distant channels share almost no information
  $\longrightarrow$ most of the weights are useless, or even counter-productive

A locally connected NN with prior information

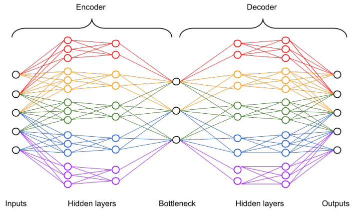
- Distant channels share almost no information
  $\longrightarrow$ most of the weights are useless, or even counter-productive

Proposition 2: locally connected architecture

We propose this kind of architecture where distant channels cannot be combined together.



Example of fully connected AE



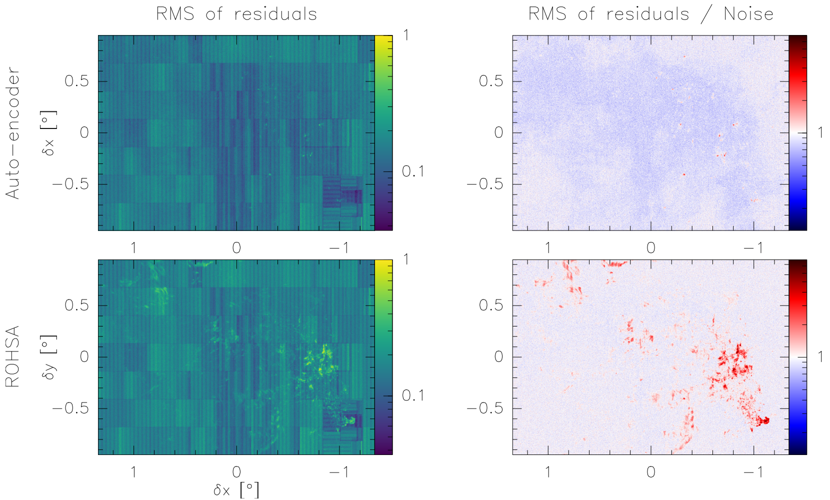
Example of locally connected AE

Gaussian fitting method ROHSA

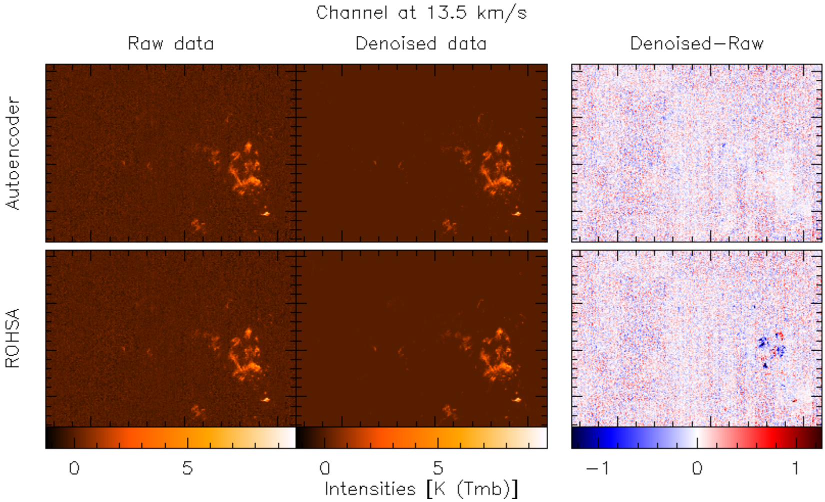
We compare our results with ROHSA gaussian fitting methods:

- **ROHSA**: Regularized Optimization for Hyper-Spectral Analysis (Marchal+2019).
- Spatially constrained **Gaussian decomposition** of profiles.
- **Goal**: To extract the multiphase structure of the ISM.
- The reconstruction after the decomposition can be seen as a **denoising**.

Denoising performances: RMS of residuals



Denoising performances: residuals



Conclusion

Take-home messages

- Usual dimension reduction based denoising methods are poorly adapted to line cubes.
- We have developed a deep learning method based on an **in-depth data analysis**.
- This method is able to denoise cubes with a **SNR-dependent behavior** in order not to distort the signal.

- **Paper:** Einig et al., A&A, 677, A158 (2023).
- **GitHub:** `einigl/line-cubes-denoising` – can be used for other datasets.

References

F. Le Petit et al. (2006)
A model for atomic and molecular interstellar gas: The Meudon PDR code
The Astrophysical Journal Supplement Series, 164, p. 506

R. Wu et al (2018)
Constraining physical conditions for the PDR of Trumpler 14 in the Carina Nebula
Astronomy & Astrophysics, 618, p. A53

J. Pety et al. (2017)
The anatomy of the Orion B giant molecular cloud: A local template for studies of nearby galaxies
Astronomy & Astrophysics, 599, p. A98