# A Data Science Perspective on Deconvolution

**Presentation of work originally published in the Proc. of the 5th Int. Conf. on Data Science and Advanced Analytics (DSAA)**

Mirko Bunse[1] Nico Piatkowski[1] Tim Ruhe[2] Katharina Morik[1] Wolfgang Rhode[2]

**Keywords:** deconvolution; unfolding; supervised learning; transductive learning; density estimation; Cherenkov astronomy; regularization

Deconvolution problems arise when the probability density function (pdf) of a quantity $Y$ is estimated even though $Y$ cannot be measured directly. In this scenario, the pdf of $Y$ has to be inferred from related quantities $X_1, X_2, \ldots$ which are measured instead. Several algorithms solving this task have been proposed in particle physics, a field where deconvolution problems frequently arise. In this extended abstract, we summarize our findings made from a data science perspective [Bu18] and our on-going work on deconvolution.

The term "de-convolution" (also known as "unfolding") is motivated by the traditional formalization of the problem, which models the pdf $g : \mathcal{X} \rightarrow \mathbb{R}$ of the observed quantities as a convolution of the sought-after pdf $f : \mathcal{Y} \rightarrow \mathbb{R}$ of $Y$ with another function $R : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$. In this model, $R$ represents a conditional probability function which is learned from a set of training data. The goal is to infer $f$ from given $g$ and $R$.

$$g(\vec{x}) \;=\; \int_{\mathcal{Y}} R\,(\vec{x} \mid y) \cdot f\,(y)\;\mathrm{d}y \tag{1}$$

Traditional approaches maximize the likelihood [Bl85] or employ Bayes' theorem [D'95] in a discrete variant of this formalization. Thus, they estimate the probability $\mathbb{P}(Y \equiv i)$ of each discrete state $i$ of $Y$. Unfortunately, previous publications only present these methods as single monolithic instances. Two of our contributions are the unification of traditional algorithms and the identification of theoretic similarities between them.

Furthermore, we advocate a more recent approach [Ru16] based on supervised machine learning. It recasts deconvolution as a classification task, providing a modular framework in which the learning method is exchangeable. The idea is to recover each $\mathbb{P}(Y \equiv i)$ from a classifier's confidence $c_{\mathcal{M}}\,(i \mid \vec{x})$, which is interpreted as a probability conditioned on each observation. This reconstruction is then repeated in an expectation maximization (EM) procedure. While the original algorithm exhibits a diverging behavior, we propose several improvements which lead to a robust and also accelerated algorithm.

[1] TU Dortmund, Artificial Intelligence Group, D-44227 Dortmund, `firstname.lastname@tu-dortmund.de`
[2] TU Dortmund, Astroparticle Physics Group, D-44227 Dortmund, `firstname.lastname@tu-dortmund.de`
This work has been supported by Deutsche Forschungsgemeinschaft (DFG) within the Collaborative Research Center SFB 876 "Providing Information by Resource-Constrained Data Analysis", projects C3 and A1. `https://sfb876.tu-dortmund.de`.

 doi:10.18420/inf2019_43

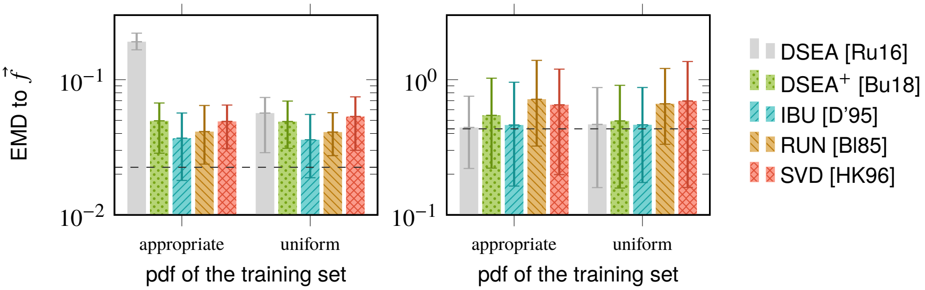


Fig. 1: The reconstruction quality of five methods is assessed in terms of the Earth Mover's Distance (EMD) between their estimates and the true discrete solution $\vec{f} \in \mathbb{R}^d$. On the left, we present an experiment with many observations [Bu18], while on the right only few observations are available for deconvolution (on-going work). Our improved method, DSEA$^+$, performs well throughout.

$$\hat{\mathbb{P}}(Y \equiv i) \;=\; \sum_{\vec{x} \in \mathcal{X}} \hat{\mathbb{P}}\,(Y \equiv i | X = \vec{x}) \cdot \hat{\mathbb{P}}(X = \vec{x}) \;=\; \sum_{n} c_{\mathcal{M}}\,(i \,|\, \vec{x}_n) \cdot \frac{1}{N} \tag{2}$$

Finally, we evaluate the traditional approaches and the learning-based method in comparative experiments. The essence of our findings, as indicated by Fig. 1, is that all methods are able to obtain results of a similarly high quality, given that they are provided enough observations. One notable exception to this end is the learning-based approach *without* our improvements, which produces less accurate results.

In our on-going work, we are investigating relations between deconvolution and other tasks in machine learning. For example, we establish a connection to transductive learning. Also, we apply our methods to text corpora of political manifestos, thus demonstrating that deconvolution is a general data science problem by far not limited to particle physics.

## References

[Bl85] Blobel, V.: Unfolding methods in high-energy physics experiments. Technical report, CERN, 1985.

[Bu18] Bunse, Mirko; Piatkowski, Nico; Ruhe, Tim; Rhode, Wolfgang; Morik, Katharina: Unification of Deconvolution Algorithms for Cherenkov Astronomy. In: Proc. of the 5th DSAA. IEEE, pp. 21–30, 2018.

[D'95] D'Agostini, G.: A multidimensional unfolding method based on Bayes' theorem. Nucl. Instrum. Methods Phys. Res. A, 362(2-3):487–498, 1995.

[HK96] Hoecker, A.; Kartvelishvili, V.: SVD approach to data unfolding. Nucl. Instrum. Methods Phys. Res. A, 372(3):469–481, 1996.

[Ru16] Ruhe, T.; Börner, M.; Wornowizki, M. et al.: Mining for Spectra – The Dortmund Spectrum Estimation Algorithm. In: Proc. of the ADASS XXVI. 2016. in press.