TP2: Reconstruction of DOSY NMR signals

Dimitri Bouche
dimitri.bouche@ensae.fr
Vincent Plassier
vincent.plassier@ens-paris-saclay.fr

February 2018

1 Part I

1.1 Generation of synthetic data

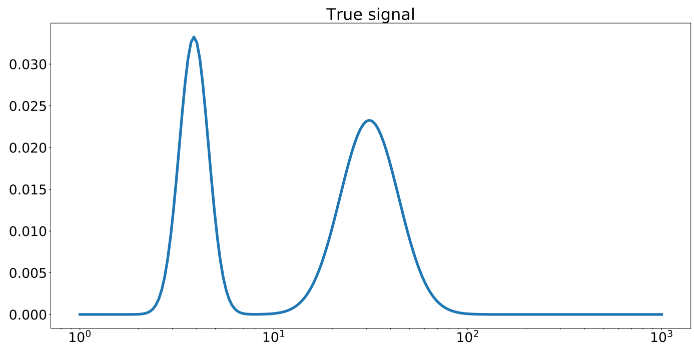

Figure 1: Original signal

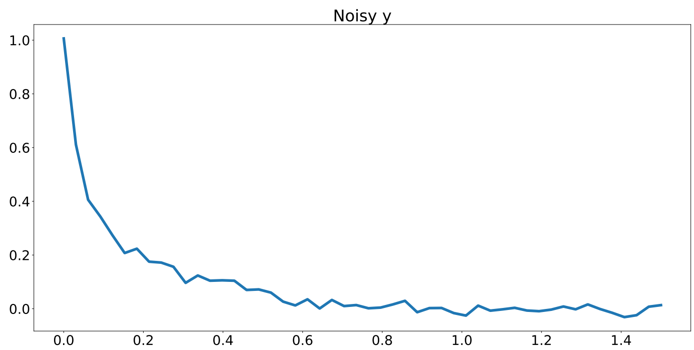

Figure 2: Noisy data

1.2 Comparison of regularization strategies

1.2.1 Smoothness prior

Discuss the existence and uniqueness of a solution The objective is the following:

$$f(x) = \frac{1}{2}||Kx - y||^2 + \beta\frac{1}{2}||Dx||^2$$

This function is twice differentiable. The gradient is:

$$\nabla f(x) = K^T Kx + \beta D^T Dx - Ky$$

And the Hessian is given by:

$$\nabla^2 f(x) = K^T K + \beta D^T D$$

K^TK and D^TD are positive as product of a matrix with its transpose, thus $\nabla^2 f(x)$ is positive. As a consequence f is convex.

Let us now show that it is also definite. $\mathbf{Ker}D$ is the span of the vector with only 1s: $x_1 = (1, ..., 1)$. Since K has only positive entries, $\mathbf{Ker}K \cap \mathbf{Ker}D = \{0\}$. As a consequence, $\nabla^2 f(x)$ is positive definite. We deduce from that that f is strictly convex and has a consequence there exist a unique solution to our optimization problem.

Propose an approach to solve the problem We can get a closed-form solution by solving $\nabla f(x) = 0$:

$$K^TKx + \beta D^TDx = Ky \Longleftrightarrow x = (K^TK + \beta D^TD)^{-1}Ky$$

Les us highlight that the matrix $K^TK + \beta D^TD$ is the Hessian from before and we have shown that it is positive definite, as a consequence, it is invertible.

Thus for this choice of prior we just solve the problem by solving the linear system above, no need for numerical algorithms.

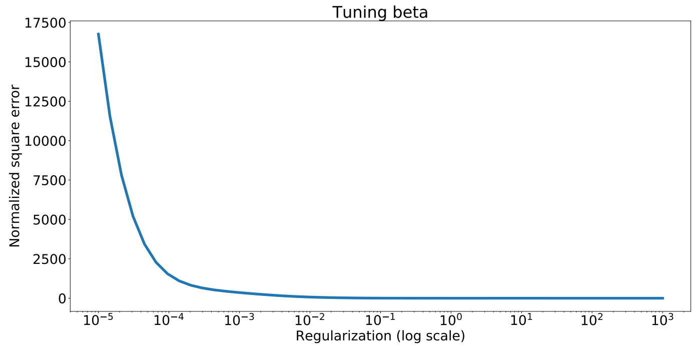

Figure 3: Tuning of regularization parameter for smoothness prior

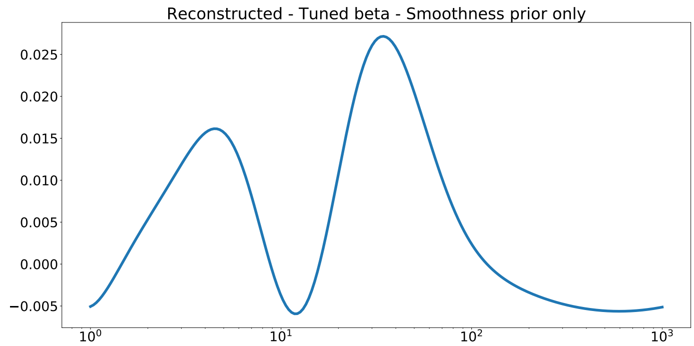

Figure 4: Reconstructed signal with optimal β for smoothness prior

With a tuned β we get a normalized squared error of 0.4056 for this reconstruction strategy.

1.2.2 Smoothness prior + contraints

Discuss the existence and uniqueness of a solution Now we want to minimize the function

$$f + g : x \longmapsto \frac{1}{2}\|Kx - y\|^2 + \frac{1}{2}\beta\|Dx\|^2 + \iota_{[x_{min}, x_{max}]^N}(x)$$

This function is convex as the sum of two convex functions. Moreover, the minimum is necessarily reached on the compact $[x_{\min}, x_{\max}]^N$ because of the indicator function. By continuity of $f + g$ on this compact we deduce the existence of a minimum. Moreover, the function is twice differential on the open $]x_{\min}, x_{\max}[^N$ with Hessian equal to $K^TK + \beta D^TD$. We deduce that the function is strictly convex on the convex $[x_{\min}, x_{\max}]^N$ and therefore admits a single minimum.

Propose an approach to solve the problem The function $f : x \longmapsto \frac{1}{2}\|Kx - y\|^2 \in \Gamma_0(\mathbb{R}^{\mathbb{N}})$ has a gradient $\nabla f(x) = K^T(Kx - y)$ which is $\|K\|^2$-Lipschitz and $g \in \Gamma_0(\mathbb{R}^{\mathbb{N}})$ where $\|K\|^2$ is the square of the spectral norm of K.

In order to solve this problem, we can implement a projected gradient algorithm. We chose $\gamma = \frac{1}{\|K\|^2}$

The projection operator P on the set $[x_{\min}, x_{\max}]^N$ is simply a coordinate-wise thresholding:

$$P(x) = \left(x_{\min}\mathbf{1}_{x_i<x_{\min}} + x_i\mathbf{1}_{[x_{\min},x_{\max}]}(x_i) + x_{\max}\mathbf{1}_{x_{\max}<x}\right)_i$$

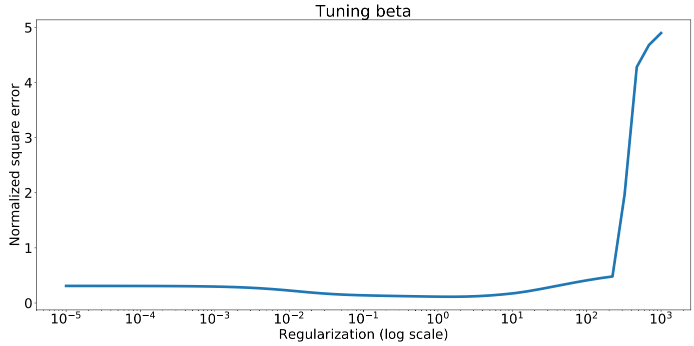

Figure 5: Tuning of regularization parameter for smoothness prior plus constraints

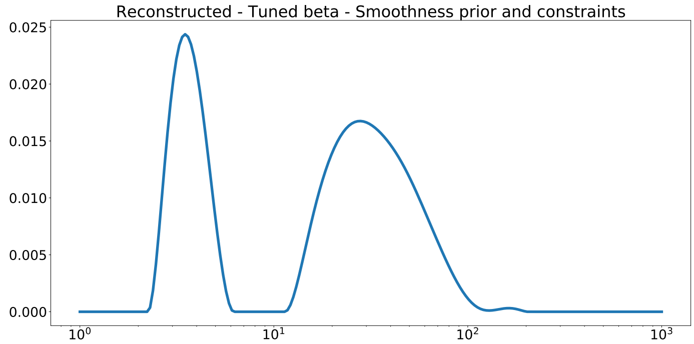

Figure 6: Reconstructed signal with optimal β for smoothness prior plus constraints

With a tuned β we get a normalized squared error of 0.1136 for this reconstruction strategy.

1.2.3 Sparsity prior

Discuss the existence and uniqueness of a solution Finally, we consider the objective function:

$$f + g : x \longmapsto \frac{1}{2}\|Kx - y\|^2 + \beta\|x\|_1$$

This function is convex as the sum of two convex functions. Since this function is coercive and continuous, we deduce the existence of a minimum. In addition, the function is differentiable on $\mathbb{R}^N \setminus \{0\}$ and its gradient is $\nabla(f+g) : x \longmapsto (K^TK + \beta Id)x - K^Ty$. Since $K^TK + \beta Id$ is an invertible matrix, the gradient is null only in

$$x^\star = (K^TK + \beta Id)^{-1}K^Ty \tag{1}$$

If $x^\star = 0$, the function $f+g$ is not differentiable in this point. However, we deduce the gradient is not null in $\mathbb{R}^N \setminus \{0\}$, thus by existence of the minimum we deduce the minimum is reached in 0. Now, suppose that $x^\star \neq 0$.
From 1, we deduce the uniqueness of the minimum on $\mathbb{R}^N \setminus \{0\}$. Moreover, suppose that $(f+g)(0) \leq (f+g)(x^\star)$, by convexity of $f+g$ on the segment $[0, x^\star]$, we deduce that $f+g$ is lower than $(f+g)(x^\star)$ on $]0, x^\star]$. This is absurd since $x^\star$ is the unique minimum of $f+g$ in $\mathbb{R}^N \setminus \{0\}$. In conclusion, $(f+g)(0) > (f+g)(x^\star)$, hence we have a unique minimum reached in $x^\star$.

Propose an approach to solve the problem For this problem, we can use a Forward Backward algorithm. Indeed the function $g = \beta\|x\|_1$ is non differentiable, however we know a simple close form for its proximity operator. Moreover, we know the function $f = \frac{1}{2}\|Kx - y\|^2$ is differentiable.

The formula for the proximity operator of g is:

$$\text{prox}_{\gamma g}(z) = \left(x_i \max\left(0, 1 - \frac{1}{\beta\gamma}\right)\right)_{i=1}^N$$

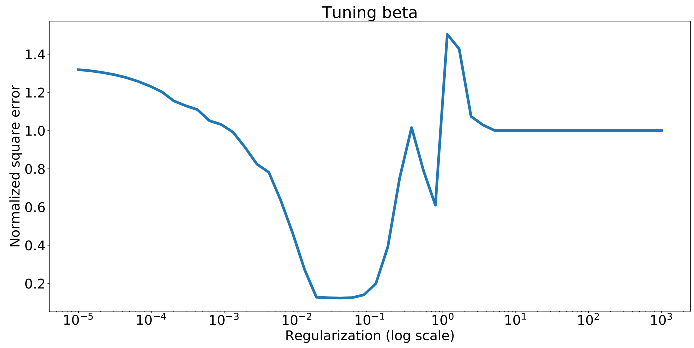

Figure 7: Tuning of regularization parameter for sparsity prior

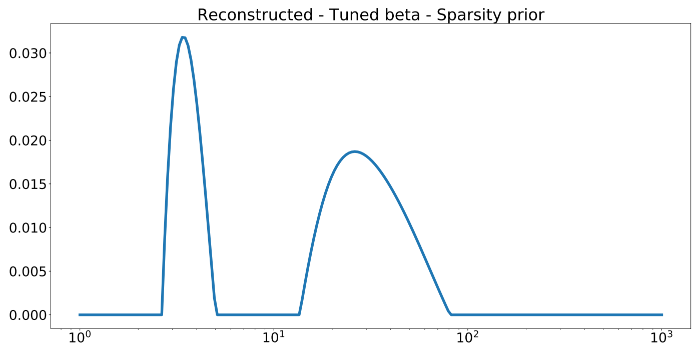

Figure 8: Reconstructed signal with optimal β for sparsity prior

With a tuned β we get a normalized squared error of 0.1234 for this reconstruction strategy. As expected, the resulting solution has many zero coefficients.

2 Part II

2.1 Question 1

Let's show that φ is convex. For this, we remark φ is twice differentiable. On $\mathbb{R}_+^*$ we have $\varphi''(u) = \frac{1}{u} > 0$, thus φ is convex on $\mathbb{R}_+$. Be $u_1, u_2 \in \mathbb{R}$ and $\lambda \in [0,1]$. If $u_1 < 0$ or $u_2 < 0$, then

$$\varphi(\lambda u_1 + (1-\lambda)u_2) \leq \lambda\varphi(u_1) + (1-\lambda)\varphi(u_2) = +\infty$$

Thus φ is convex on $\mathbb{R}$.
Moreover,

$$\psi_{(i)} = \left\{ \begin{array}{l} \mathbb{R}^N \longrightarrow \mathbb{R} \\ x \longmapsto x^{(i)} \end{array} \right. \text{ is affine.}$$

By precomposition, $\varphi \circ \psi_{(i)}$ is convex. By sommation, *ent* is convex.
Since $ent(0_N) = 0$, *ent* is proper.
Now, let's show that φ is lower-semicontinuous. Be $(u_n)_n \in \mathbb{R}^{\mathbb{N}}$ such that $u_n \to u$. If $u \in \mathbb{R}_+^*$, by continuity of φ on $\mathbb{R}_+^*$,

$$\underline{\lim}\varphi(u_n) = u$$

If $u = 0$, then

$$\underline{\lim}\varphi(u_n) \geq \inf \varphi = 0 = \varphi(u)$$

If $u < 0$, from a certain rank $(u_n)_n$ is included in $\mathbb{R}-^*$. Thus $\underline{\lim}\varphi(u_n) = +\infty \geq +\infty = \varphi(u)$.
In all cases, $\underline{\lim}\varphi(u_n) \geq \varphi(u)$. Since $\forall i, \psi_{(i)}$ is continuous, we have $\varphi \circ \psi_{(i)}$ is lower-semicontinuous by precomposition by a continuous function. By sommation, *ent* is lower-semicontinuous.
However, *ent* is not differentiable since by noting e_1 the first vector of the canonical basis, we have $\lim_n \mathrm{ent}\left(-\frac{e_1}{n}\right) - \mathrm{ent}(0) = +\infty$.

2.2 Question 2

The function $f : x \longmapsto \|Kx - y\|^2 + \mathrm{ent}(x)$ is coercive, convex, thus there exists a minimum. Moreover, this minimum is reached in $\mathbb{R}_+$ on which f is strictly convex. We deduce the unicity of the minimum.

2.3 Question 3

We are trying to solve the problem :

$$\min_{z>0} \frac{1}{2}||z-x||^2 + \nu z \ln(z)$$

$$\iff \frac{z-x}{\nu} + 1 + \ln(z) = \frac{x}{\nu} - 1 - \ln(\nu) \iff \frac{z}{\nu}\exp\left(\frac{z}{\nu}\right) = \exp\left(\frac{x}{\nu} - 1 - \ln(\nu)\right)$$

Applying the Lambert function (W):

$$\frac{z}{\nu} = W\left(\exp\left(\frac{x}{\nu} - 1 - \ln(\nu)\right)\right)$$

Thus the proximal operator of the entropy is given by : Thus

$$\mathbf{prox}_{\nu\varphi}(x) = \frac{1}{\nu} W\left(\exp\left(\frac{x}{\nu} - 1 - \ln(\nu)\right)\right)$$

2.4 Question 4

2.4.1 Forward backward

We already have the gradient of $f : x \longmapsto \|Kx - y\|^2$ which is $K^TKx_n - K^Ty$. We can then use the proximity operator for the entropy that we have computed in the previous question and we have all the components to implement a forward backward algorithm.

We obtain the following reconstruction for $\beta = 10^{-2}$:

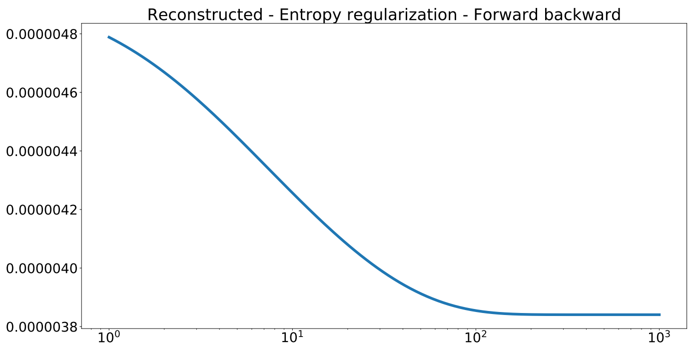

Figure 9: Reconstructed - Forward backward - Entropy regularization - $\beta = 10^{-2}$

2.4.2 Douglas-Rachford

So as to implement the Douglas Rachford algorithm, we need to compute the proximity operator of $f : z \longmapsto \frac{1}{2}||Kz - y||^2$. We thus want to solve the problem : $\min_z \frac{1}{2}||Kz-y||^2 + \frac{1}{2\nu}||z-x||^2$

This is a convex quadratic form, thus it admits a unique minimizer which we can recover by setting the gradient to zero:

$$K^TKz - K^Ty + \frac{1}{\nu}(z-x) = 0 \iff z = \left(K^TK + \frac{1}{\nu}I\right)^{-1}\left(K^TT + \frac{1}{\nu}x\right)$$

As a consequence:

$$\mathbf{prox}_{\nu f}(x) = \left(K^TK + \frac{1}{\nu}I\right)^{-1}\left(K^TT + \frac{1}{\nu}x\right)$$

We now have all the ingredients to implement the Douglas Rachford algorithm using the proximity operator of the entropy from the previous question.

2.5 Question 5

We obtain the following reconstruction for $\beta = 10^{-2}$ with Forward Backward :

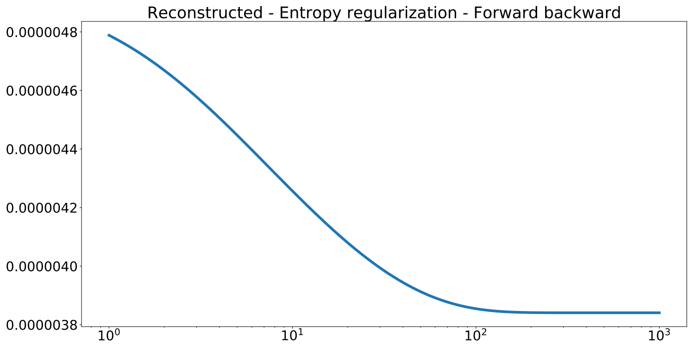

Figure 10: Reconstructed - Forward backward - Entropy regularization - $\beta = 10^{-2}$

We obtain the following reconstruction with $\beta = 10^{-2}$ with Douglas Rachford:

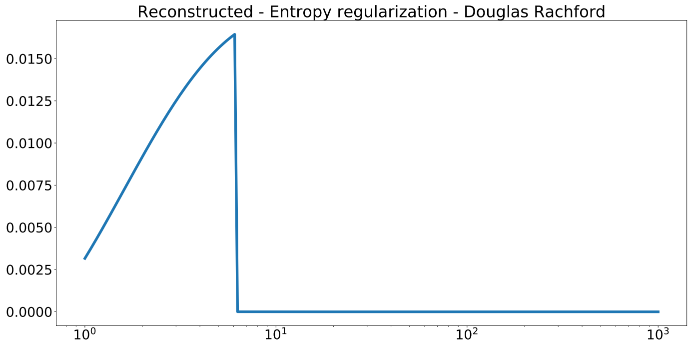

Figure 11: Reconstructed - Douglas Rachford - Entropy regularization - $\beta = 10^{-2}$

We observe that for $\beta = 10^{-2}$ we do not get the same reconstruction with Forward backward and Douglas Rachford whereas they should converge to the same solution. However, Forward Backward did not converge. If we take lower values of β we do get the same reconstruction for both algorithms, in fact taking $\beta = 10^{-3}$ or a lower β results exactly in the same reconstruction. Thus we did not tune the parameter β and took $\beta = 10^{-3}$.

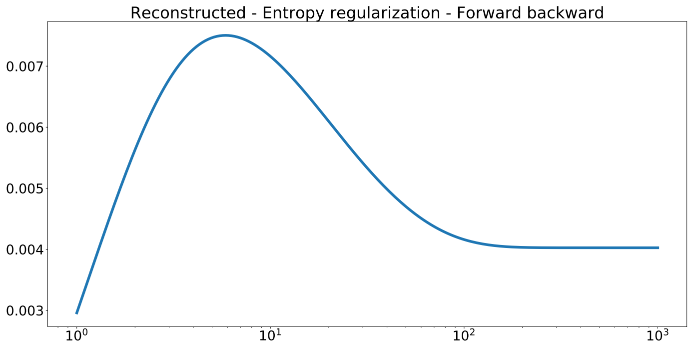

Figure 12: Reconstructed - Forward backward - Entropy regularization - $\beta = 10^{-3}$

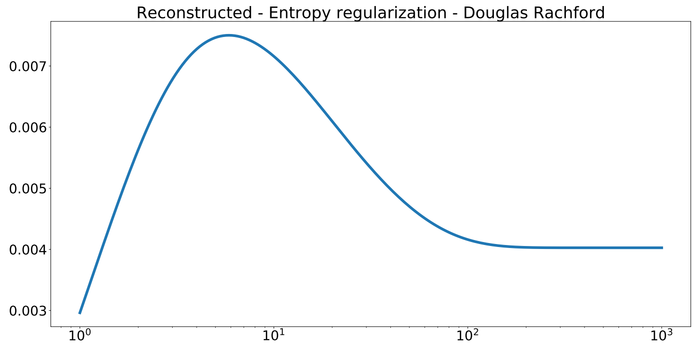

Figure 13: Reconstructed - Forward backward - Entropy regularization - $\beta = 10^{-3}$

As a consequence, Douglas-Rachford is more robust to numerical instability which can be caused by a choice of "limit" parameters. However it takes longer to run per iteration in that case since we need to invert a matrix to compute one of the proximal operators. This shows that it can be interesting to use Douglas Rachford even if we have the possibility to use gradient based algorithms to ensure a more robust convergence.

2.6 Question 6

As stated in the previous question, we found that as long as β is small enough($\beta \leq 10^{-3}$), we get the same reconstruction. So we took $\beta = 10^{-3}$ which results in a reconstruction error of 0.6675

2.7 Question 7

	Smoothness	Smoothness and constraints	Sparsity	Entropy
Reconstruction error	0.4056	0.1136	0.1234	0.6675

The reconstruction error is higher with entropy regularization than with all other regularization we tested. It seems that entropy favors too much solutions that are too simple for the original signal which results in not so good reconstruction errors.

2.8 Question 8

In order to denoise the image without adjusting a constant β, we consider a slightly different minimization problem:

$$\underset{x \in \mathbb{R}^N}{\text{minimize}}\ \text{ent}(x) \quad \text{subject to} \quad \|Kx - y\|^2 \leq \eta M \sigma^2$$

To solve this problem, we considered a parallel form of the Douglas-Rachford algorithm. With the course notations, we consider $L_1 = \text{Id}, L_2 = K$ and the functions $g_1 = \text{ent}, g_2 = \iota_{B(y,\eta M\sigma^2)}$.
For this algorithm, we only need the proximal operators of g_1 and g_2. Thanks to the third question and the course, we get

$$\underset{\gamma g_1}{\text{prox}}(x) := \left(\frac{1}{\gamma} W\left(\exp\left(\frac{x}{\gamma} - 1 - \ln(\gamma)\right)\right)\right)_{i=1}^{N} \quad \text{and} \quad \underset{\gamma g_2}{\text{prox}} = P_{B(y,\eta M\sigma^2)}$$

With 5 iterations and $\gamma = 10$, we get 14.

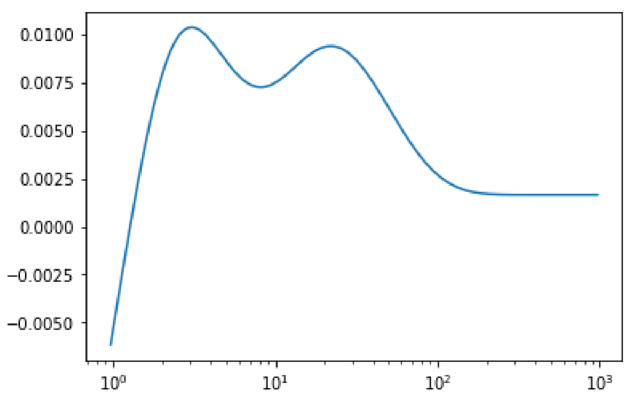

Figure 14: Reconstructed - PPXA - $\eta = 1$