Hybrid Parallel Tempering and Levenberg-Marquardt method for efficient and stable fitting of noisy MRI dataset

Marco Palombo[1,2], Matthias Vandesquille[1,2], and Julien Valette[1,2]
[1]CEA/DSV/I2BM/MIRCen, Fontenay-aux-Roses, France, France, [2]CEA-CNRS URA 2210, Fontenay-aux-Roses, France, France

PURPOSE. During the last decade, new diffusion weighted NMR (DW-NMR) methods, based on non-Gaussian diffusion models, have been developed to measure physical parameters that are directly correlated with different properties of biological tissues, such as the diffusional kurtosis (DKI)[1] and the stretched exponential (STREMI)[2] imaging. These techniques allow increasing the sensitivity and specificity of the DW-NMR imaging for several pathological conditions[3] and image post-processing is based on non-linear data fitting. Regrettably, non-Gaussian DW-NMR requires high b-values acquisition schemes (b>2000 s/mm^2), that inevitably leads to more noisy measured signal. The higher noise fluctuations in experimental data can sometimes introduce many spurious local minima in the likelihood function, E, to be minimized in fitting routine. This implies that conventional fitting pipeline, usually based on Levenberg-Marquardt (LM) algorithm, often fail in finding the global minima. As a consequence, the estimated values for non-Gaussian diffusion parameters lack in stability and accuracy, leading to not accurate and grainy parametric maps. In this contribution we propose a hybrid scheme of Parallel Tempering (PT) and LM approaches (PT/LM), for optimization problems of non-linear models. The aim is to design an efficient and stable pipeline for non-linear fitting of noisy DW-NMR images.

METHODS. *Hybrid PT/LM method:* The PT method[4] was developed to achieve good sampling of complex objective function landscape with many local minima. In this context it will be used to find the global minima of the sum of squared model residuals, $E(x)$, where $x = \{p_1,p_2, \ldots,p_N\}$ represents a set of N model parameters. We track M sets of the system configurations, each at a different arbitrary "temperature" level, T_i. The different temperature levels form a temperature ladder and each configuration forms a subsystem at each level on the temperature ladder. A state of the composite system is specified by $X = \{x_1,x_2, \ldots,x_M\}$, where x_i is the subsystem at temperature level i. The equilibrium distribution of the ''macrostate'' X is $P(X)=\Pi_{i=1,M} exp[-\beta_i E(x_i)]/Z(T_i)$, where $\beta_i=1/T_i$ and $Z(T_i) = \Sigma_{j=1,M} exp[-\beta_i E(x_j)]$ is the partition function of the subsystem at temperature T_i. The hybrid PT/LM method is composed by a PT and a LM part. The goal of the PT part is to reduce the relaxation time of subsystems towards equilibrium state, i.e. the global minima of E for that subsystem within the composite system. The subsystems at highest temperature levels easily climb over all local minima in the likelihood function landscape, identifying the global minima; while the low temperature subsystems mainly explore the local minima.On the other hand, in the LM part, the target parameters (the N model parameters and the temperature that maximize P) are quickly estimated, using the subsystems global minima as initial conditions. *Cross-Validation:* to check the efficiency and the stability of the PT/LM with respect to the conventional LM and PT methods, we validate the model finding the global minimum of notorious benchmark functions in optimization: the Rosenbrock's function and a "rugged" funnel-like function. Due to the long narrow valley present in Rosenbrock's function, gradient-based methods have to spend a large number of iterations before identifying the global minimum. Moreover, the multiple local minima in the "rugged" funnel-like function introduce a strong dependence on the initial conditions, leading to a great instability of the solution (often wrong). *DW-NMR images acquisition*: an *ex vivo* healthy mouse brain, fixed in paraformaldehyde and stored in PBS, was scanned at 11.7 T (Bruker), to test the performance of the new procedure in non-Gaussian diffusion mapping.An imaging version of PGSTE sequence was performed with TE/TR=20/4200 ms, Δ/δ=400/3 ms, ; 10 sagittal slices of 0.5mm thickness, FOV=3.2cm, matrix 256x256, and 4 repetitions. Images were acquired for 10 b-values ranging from 100 to 5000s/mm^2 along direction (1,1,1) plus 2 b=0. *DW-NMR images analysis:* "low noise" dataset was obtained by averaging across the 4 repetitions and used as reference, while "high noise" dataset was obtained by taking just one repetition. Parametric maps of kurtosis metrics, were obtained by fitting on a voxel-by-voxel basis the following relationto the DW image signal intensities (for b≤3000s/mm^2): $S(b)/S(0)=\exp(-bD_{app}+1/6b^2D_{app}^2K_{app})+\eta$, where D_{app} and $K_{app}$are the apparent diffusion coefficient and kurtosis, respectively, and η is the background noise. STREMI parametric maps, i.e. γ-maps, were obtained by similar procedure, using the following relation: $S(b)/S(0)=\exp[-(D_{app}b)^{\gamma}] +\eta$, where γ is the stretching parameter, being between 0 and 1.

RESULTS & DISCUSSION. Figure1 shows the results of the cross-validation. PT/LM algorithm allowed to correctly identify the global minimum of Rosenbrock's function (**Figure1a**) more efficiently than LM and PT. Indeed PT/LM is 2x faster than LM and 4x faster than PT. However, when the likelihood function has multiple local minima, such as the "rugged" funnel-like function in **Figure1b**, PT/LM is 2x faster than PT but 2x slower than LM. Nevertheless, PT/LM and PT show much higher accuracy and stability than LM in identifying the global minimum of "rugged" funnel-like function. Indeed, PT/LM always identify the global minimum (as PT does), while LM often fails, depending on the initial conditions (**Figure1b**). LM algorithm has almost the same performance of the PT/LM only if the initial conditions are opportunely chosen with an *a priori* knowledge of the global minimum location. Concerning non-Gaussian diffusion maps, those obtained using LM with "low noise" dataset was used as reference and compared to the ones obtained with the LM, PT and PT/LM approaches using the "high noise" dataset (**Figure2**). Using the PT/LM method, an increase in contrast-to-noise ratio, specially between white and gray matter regions, was found for both DKI and STREMI models (**Figure2**). Moreover, PT/LM maps are less grainy and more similar to the reference ones than the LM and PT maps. This demonstrates that the proposed approach allows a more stable and accurate estimation on non-Gaussian metrics when DW-NMR dataset is affected by "high noise".

CONCLUSION. In this contribution we demonstrate that a novel hybrid scheme of PT and LM approaches improves the efficiency and stability of conventional pipeline for non-linear fitting of noisy DW-NMR images.

REFERENCES. [1]Jensen JH, et al. *MagnReson Med* 2005; 53(6):1432-1440.[2]De Santis S, et al. *MagnReson Med* 2011; 65(4):1043-1052.[3]Grinberg F et al. *PloS one* 2014; 9(2):e89225.[4]HansmannUH, *ChemPhys Let*1997; 281(1): 140-150.

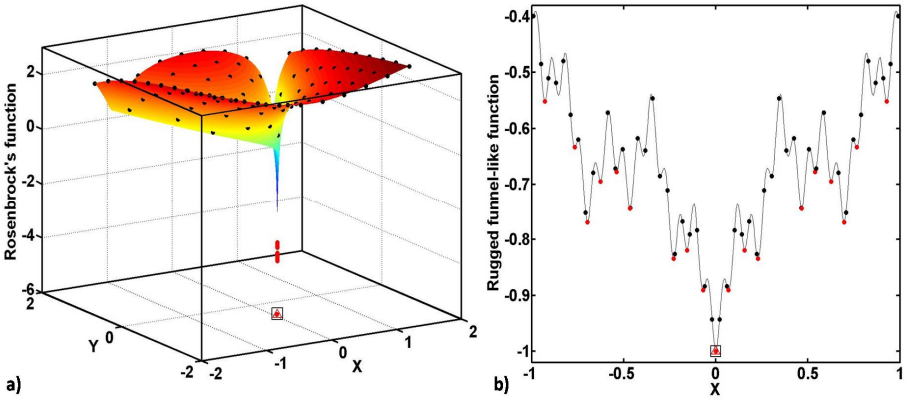

Figure1. *a) Rosenbrock's function surface plot. Red dots represent LM fitting results for different initial conditions (black dots), red empty triangle represents PT result, while PT/LM fitting result is represented as black box. b)"Rugged" funnel-like function, LM fitting results (red dots) for different initial conditions (black dots), PT (red open triangle) and PT/LM fitting result (black box).*

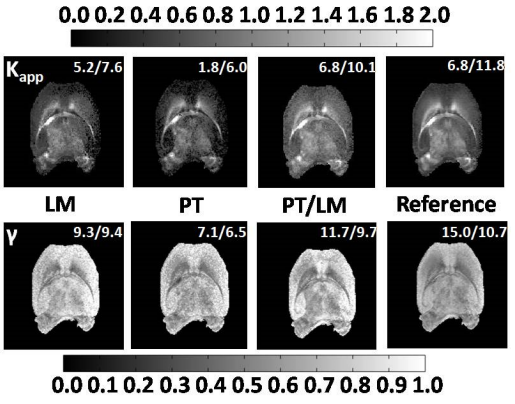

Figure2. *A representative slice of non-Gaussian diffusion maps (K_{app}- and γ-maps), obtained by using LM with the "low noise" dataset (reference) and LM, PT and PT/LM with the "high noise" dataset. Contrast-to-noise ratio for gray and white matter (gray/white) is reported in each map.*

ACKNOWLEDGEMENTS. This work was funded by the European Research Council ERC-336331-INCELL.