The Parallel Asynchronous Differential Evolution Method as a Tool to Analyze Synchrotron Scattering Experimental Data from Vesicular Systems[*]

Evgeniya Zhabitskaya[1,2,a], Elena Zemlyanaya[1,2], Mikhail Kiselev[1,2], and Andrey Gruzinov[3]

[1]*Joint Institute for Nuclear Research, Joliot-Curie 6, 141980 Dubna, Russia*
[2]*Dubna State University, University str. 19, 141980 Dubna, Russia*
[3]*National Research Center «Kurchatov Institute», Moscow, Russia*

Abstract.
In this work we use an Asynchronous Differential Evolution (ADE) method to estimate parameters of the Separated Form Factor (SFF) model which is used to investigate a structure of drug delivery Phospholipid Transport Nano System (PTNS) unilamellar vesicles by experimental small angle synchrotron X-ray scattering spectra (SAXS). We compare the efficiency of different optimizing procedures (OP) for the search for the SFF-model parameters. It is shown that the probability to find the global solution of this problem by ADE-methods is significantly higher than that by either Nelder-Mead method or a Quasi-Newton method with Davidon-Fletcher-Powell formula. The parallel realization of ADE accelerates the calculations significantly. The speed-up obtained by the parallel realization of ADE and results of the model are presented.

1 Introduction

Differential Evolution (DE) [1] is an efficient algorithm to solve global optimization problems. Recent progress in the development of DE is summarized in the book [2] and in the review article [3]. The Asynchronous Differential Evolution (ADE) method [4] incorporates classical DE mutation, crossover and selection operations into a parallel steady-state strategy. It is well suited for parallel optimization. ADE with independent restart [5] automatically adapts the population size to the complexity of the problem. ADE with a restart and adaptive correlation matrix [6] (ADE-ACM) automatically adapts to the landscape of the optimized objective function.

In [7, 8] the method of separated form factors (SFF) is adapted for the analysis of experimental data on small angle synchrotron X-ray scattering (SAXS) to polydispersed population of dimirostoilphosphatidilholin (DMPC) unilamellar vesicles (ULVs), placed in an aqueous solution of sucrose. In this work we employ this method for the analysis of SAXS spectra from the drug delivery Phospholipid Transport Nano System (PTNS) [9]. Parameters of the SFF model are estimated with the help of ADE method. The calculations have been also made using other popular minimization algorithms in order to compare the efficiency of different approaches.

---

[*]The work has been performed under the grant of Russian Scientific Foundation (project No 14-12-00516)
[a]e-mail: Evgeniya.Zhabitskaya@jinr.ru

2 Materials and methods, mathematical model and numerical algorithm

The PTNS samples for the SAXS measurements were prepared via dilution of lyophilized PTNS in such a way that to have 20%, 25%, 30% and 35% of maltose concentration in water. The measurements were performed at room temperature at the DICSI station of the Kurchatov Synchrotron Radiation Source in Moscow, Russia. The scattering intensities $I^{\text{exp.}}(q)$ of photons from liophylized PTNS solutions in water are measured as a function of the scattering vector $q = 4\pi \sin\left(\frac{\theta}{2}\right)/\lambda$, where θ is the photon scattering angle. Experimental data are shown in Fig. 1 by light circles. The plot is the subtraction of the PTNS intensity and the corresponding buffer intensity from maltose solution.

In the data analysis we employ an approach based on the SFF method [7]. In this framework the macroscopic cross-section of the polydispersed vesicular population is of the form

$$\frac{d\Sigma}{d\Omega}(q) = n\, I_0\, \hat{F}_s\, \hat{F}_b\, S_{\text{F}}, \tag{1}$$

where I_0 is the intensity of the incident beam, n is the number of vesicles in unit volume, $S_{\text{F}}(q; R, n)$ is the structural factor[1]. Considering the polydispersity of the vesicular radii like in [7, 8][2] and taking into account the fluctuations of the parameters of the internal structure of the bilayer[3], the form factor of the population of spherical vesicles $\hat{F}_s$ and the form-factor of the bilayer $\hat{F}_b$ are:

$$\hat{F}_s(q; \langle R\rangle, m) = \left.\int_{R_{\min}}^{R_{\max}} F_s(q, R) G(R, \langle R\rangle, m)\, dR \middle/ \int_{R_{\min}}^{R_{\max}} G(R, \langle R\rangle, m)\, dR\right., \tag{2}$$

$$\hat{F}_b(q, \langle\Theta_b\rangle, k_{\text{fl}}) = \left.\int_{-K}^{K} F_b(q, \Theta_b(\langle\Theta_b\rangle, k_{\text{fl}}, \xi))\, f(\xi)\, d\xi \middle/ \int_{-K}^{K} f(\xi)\, d\xi\right.. \tag{3}$$

The form factor of the spherical surface with radius R is $F_s(q; R) = (4\pi R^2 \sin(qR)/qR)^2$. The form factor of the symmetric lipid bilayer is $F_b(q, \Theta_b) = (\int_{-d/2}^{+d/2} \rho_c(x) \cos(qx)\, dx)^2$. The contrast function $\rho(x; \Theta_b)$ is described by the parameters of the internal structure: $\Theta_b = \{d, D, L, \rho_H\}$.

In view of the incoherent background (IB) and spectrometer resolution, the macroscopic cross section takes the form:

$$I(q) = \frac{d\Sigma}{d\Omega}(q) + \frac{1}{2}\Delta^2 \frac{d^2 \frac{d\Sigma}{d\Omega}(q)}{dq^2} + I_B, \tag{4}$$

where I_B is the parameter of incoherent background, $\Delta^2 = 3.6 \cdot 10^{-7}\ \text{cm}^2$ is the second momentum of the spectrometer resolution function.

The fit of the experimental data was carried out by minimizing the objective function describing the discrepancy between the theoretical and experimental values of the intensity

$$f(\Theta) = \sum_{j=1}^{N} \left(\frac{I(q_j, \Theta) - I^{\text{exp}}(q_j)}{\sigma^2(q_j)}\right)^2, \quad \Theta^* = \text{Arg}\min_{\Theta_i} f(\Theta), \tag{5}$$

---

[1]The structural factor considers inter-vesicular interaction. In the Debye approximation $S(q; R, n) = 1 - \frac{8V_\nu}{\nu}\left(\frac{\sin(2qR)}{2qR}\right)$, where $V_\nu = 4\pi R^3/3$ is the vesicle volume, $\nu = 1/n$ is the volume of the solvent per one vesicle.

[2]The polydispersity of vesicular population is described by the asymmetric Schulz distribution $G(R, \langle R\rangle, m) = R^m/m!(m+1/\langle R\rangle)^{m+1} \exp[-(m+1)R/\langle R\rangle]$, where $\langle R\rangle$ is the average radius of the vesicles, m is the polydispersity parameter. In the calculations below $R_{\min}=1$ nm, $R_{\max}=8.1$ nm.

[3]In the calculations presented below, it is assumed that the difference in the values of parameters of the internal structure is due to the fluctuation of the thickness of the membrane. Then the components of Θ_b in the integral in Eq. (3) are, respectively: $d=\langle d\rangle(1+k_{\text{fl}}\,\xi)$, $D=\langle D\rangle\left(1+\frac{\langle d\rangle}{\langle D\rangle}k_{\text{fl}}\,\xi\right)$, $L=\langle L\rangle\left(1-\frac{\langle d\rangle}{\langle L\rangle}k_{\text{fl}}\,\xi\right)$, $\rho_H\equiv\langle\rho_H\rangle$, where $\langle d\rangle$, $\langle D\rangle$ and $\langle L\rangle$ are the average values of the parameters of the internal structure of the bilayer; k_{fl} is the fluctuation parameter; $f(\xi)$ is the standard normal distribution.
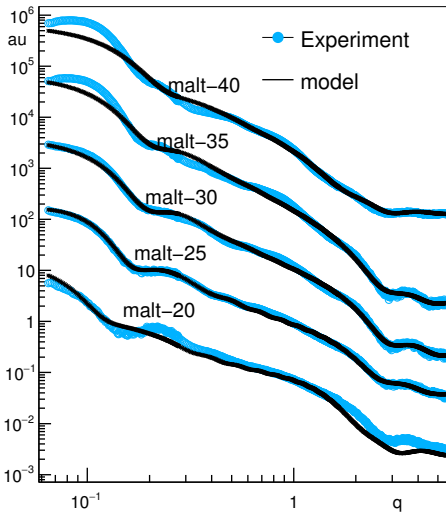
Figure 1. SAXS spectra from the PTNS vesicles in 20–40%(w/w) maltose solution in water. Circles — experiment, solid line — theory.
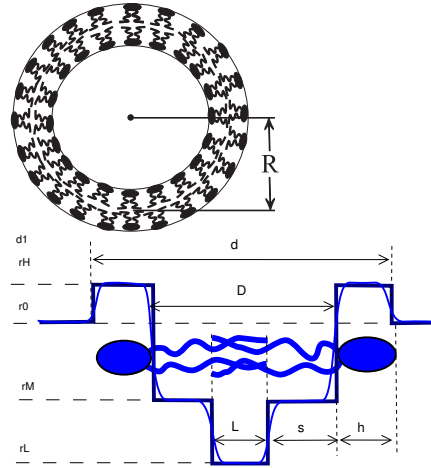
Figure 2. Structure of a phospholipid unilamellar vesicle with radius R and the "two-step" scattering length density (contrast function) $\rho(x; \Theta_b)$ across bilayer; $\Theta_b = \{d, D, L, \rho_H\}$; parameters ρ_0, ρ_M, ρ_L are fixed.

Table 1. The boundaries of the search area and the coordinates of the minima for the two step SFF model. Accuracy in the table is equal to the last significant digit. The number of data points is $N = 1000$.

Model	R, nm	m	$\langle l \rangle$, nm	$\langle s \rangle$, nm	$\langle h \rangle$, nm	k_{fl}	ρ_H, $\times 10^{10}$ cm^{-1}	nI_0, $\times 10^{18}$ cm^{-3}	I_B, $\times 10^{-3}$ a.u.	$\frac{V_\nu}{\nu}$	$\frac{f^*_{min}(\Theta^*)}{N}$
Θ_i^{min}	8.0	5.0	0.01	0.0	0.8	10^{-5}	10.1	10^{-4}	-100	0.02	
Θ_i^{max}	25.0	30.0	1.50	3.50	1.0	0.2	14.0	10^{-1}	100	0.08	
PTNS-20	21.9	12.9	0.322	1.33	0.900	0.2	12.7	0.0024	2.34	0.07	1.008
PTNS-25	17.1	20.3	0.100	1.14	1.00	0.147	11.4	0.0172	3.76	0.08	0.318
PTNS-30	14.3	16.9	0.403	1.07	0.820	0.106	11.4	0.0264	2.15	0.02	0.531
PTNS-35	13.1	14.1	0.000	1.09	0.819	0.095	11.3	0.0421	1.99	0.02	1.70
PTNS-40	10.6	9.34	0.184	1.71	0.977	0.2	12.6	0.0267	1.27	0.02	1.80

where N is the number of experimental points, σ is the error estimations, the set of the varied parameters is $\Theta = \{R, m, \langle l \rangle, \langle s \rangle, \langle h \rangle, k_{fl}, \rho_H, n_x I_0, I_B, V_\nu / \nu\}$.

The calculation on the basis of SFF method with the "two-step" model of scattering length density distribution across bilayer (see Fig. 2) gives the values of the average radius and bilayer thicknesses presented in the Table 1. The respective curves $I(q)$ are plotted in Fig. 1 by solid lines. It is seen that the calculated intensity is in good agreement with the experimental data within the region $q < 5$ nm^{-1}.

3 Comparison of the effectiveness of ADE, ADE-ACM, Nelder-Mead simplex and quasi-Newtonian with Davidon-Fletcher-Powell formula optimization methods

To find the minimum (5), different methods of optimization have been used: in addition to the various strategies of ADE [4, 5] and ADE with Adaptive Correlation Matrix (ADE-ACM) [6] methods, there are the Nelder–Mead simplex method [10] and the quasi-Newtonian method with Davidon–Fletcher–Powell formula [11, 12]. As part of the ROOT-Minuit program [13], these methods are called SIMPLEX and MIGRAD, respectively.

The permissible search area for all the optimization procedures (OPs) has been set as: $\Theta_j \in [\Theta_j^{min}, \Theta_j^{max}]$. The Θ_j^{min} and Θ_j^{max} values are presented in 2nd and 3rd lines of the Table 1. The points for the initial approximation are uniformly distributed in this area for all methods tested. The ADE-method is used in two ways. (i) ADE *rand*/*rand*/1/*bin* strategy: in this case, the values have been set for the scaling factor as $F = 0.9$ and for crossover $C_r = 0.9$. (ii) ADE-ACM strategies *rand*/*rand*/1/*acm* and *linworst*/*current-to-pbest*/1/*acm*. In both cases, the initial population is $N_p^{init} = 32$, the restart criteria is $\Delta x_k = 10^{-12}$, the minimum required accuracy of localization (exit criteria) is $\Delta f = 10^{-2}$ (see [5] for details). In the settings of SIMPLEX and MIGRAD methods the maximum number of the function evaluations allowed were increased in comparison with the default settings[4]. The calculations were repeated by each method $n_t = 100$ times.

The minimal values of the function $f_{min}^*(\Theta^*)$ for each model have been found at the points Θ^*, the coordinates of which are given in lines $4-8$ of Table 1. Note that some of the SFF parameters are in significant correlation in the vicinity of the minimum. Moreover, each problem has several local minima. This complicates the process of minimization.

A comparison of the effectiveness of ADE, ADE-ACM, SIMPLEX and MIGRAD methods is shown in Fig. 3. On the horizontal axis the difference between the best value founded by ith optimizing procedure $f_i^*(\Theta)$ and the minimal obtained value $f_{min}^*(\Theta^*)$ is laid. On the vertical axis we give the percentage of attempts with the same or better result received by each of the methods used. Table 2 shows the summarized results obtained by each method. The probability of finding the solution by SIMPLEX method was found to be $1-2\%$ (see col. 3 of Table 2). Most often, this method diagnoses convergence at points far from the global minimum and stops further calculations. The probability of finding the solution[5] of the problem by MIGRAD is close to 40% .

At the same time, all the tested strategies of ADE converged to the global minimum with a probability of more than 70%, and the most successful (*DE*/*rand*/*rand*/1/*acm*) — with a probability of more than 90%. Note that this probability could be further enhanced by increasing the maximum number of function evaluations allowed. The use of the adaptive crossover can increase the probability of finding the global minimum significantly. The use of "fast track" strategy *DE*/*linworst*/*current-to-pbest*/1/*acm* reduces the average number of function evaluations approximately three times, but this strategy has a larger probability of ending calculations in one of the local minima.

Thus, in dealing with this problem, ADE are to be preferred to the classical ones for several reasons. First, ADE and ADE-ACM are robust global minimum search methods. They end the calculations with a larger probability at the minimum of the problem, while the search using SIMPLEX and MIGRAD ended at the minimum with a chance of $1\pm1\%$ and $43\pm7\%$, respectively. The probability of finding the global minimum by ADE-R depending on the chosen strategy varied from $p = 73\pm5\%$

---

[4]MaxFunctionsCalls = 300 000, which is enough for these programs to finish computations before the MaxFunctionsCalls is reached.

[5]Assuming that the minimum has been founded if $\delta\Delta^2 = \left(f_i^*(\Theta) - f_{min}^*(\Theta^*)\right)^2 \leq 1$.

Table 2. Efficiency of the optimization procedures (OP), based on different optimization methods. $^{\dagger}N_p^{\text{init}} = 32$. The percentage of successful ($\delta\Delta = f_i^*(\Theta) - f_{\min}^*(\Theta^*) \leq 1$) attempts, average number of function evaluations $\langle N_{\text{FE}} \rangle$ and median number of function evaluations Med N_{FE} for each method.

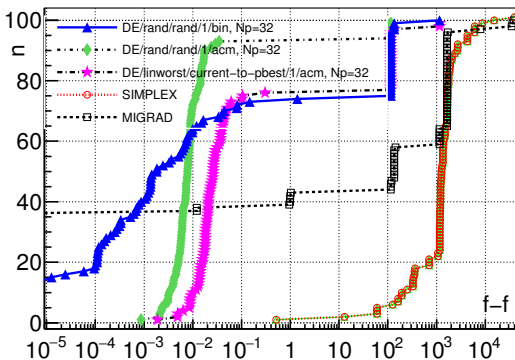| Method	MIGRAD	SIMPLEX	ADE† rand/rand/1/bin	ADE-ACM†	
				rand/rand/1/acm	linworst/curtopbest/1/acm
$\delta\Delta^2 \leq 1$	43%	1%	73%	93%	76%
$\langle N_{\text{FE}} \rangle$	3650 ±140	3500 ±170	29000 ±3000	89000 ±7300	20600 ±900
Med N_{FE}	3120 ±270	3650 ±370	16760 ±1500	50000 ±34000	18500 ±1100

Figure 3. Comparison of the effectiveness of various OP solving SFF for PTNS-30 problem. Horizontal axis: the difference between the best value, found by i^{th} OP $f_i^*(\Theta)$ and the global minimum $f_{\min}^*(\Theta^*)$; vertical axis: percentage of attempts with same or better result, received by the methods. *DE−rrb* means *DE/rand/rand/1/bin* method; *DE−rra* means *DE/rand/rand/1/acm*; *DE−lwa* means *DE/linworst/current-to-pbest/1/acm*; SIM. means SIMPLEX; MIG. means MIGRAD.
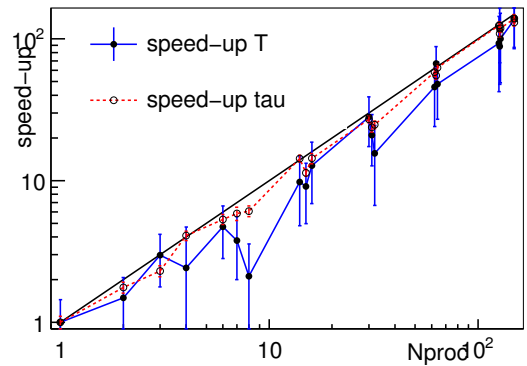

Figure 4. Acceleration of parallel computing (speed-up) depending on the number of calculating processors N_{proc} used. Solid line — resultant acceleration — decrease in mean whole calculation time $\langle T \rangle$; dashed line — acceleration for mean time of one calculation of the objective function $\langle \tau \rangle$.

(for *DE/rand/rand/1/bin*) to $p = 93\pm3\%$ (for *DE/rand/rand/acm*). On the other hand, if necessary, to find the secondary minima, preliminary ADE-produced results during calculations can be used.

As a direct (without the use of derivatives) method, ADE continues calculations and finds solutions in situations where quasi-Newton Davidon-Fletcher-Powell's method diagnoses inability of further searches, and the simplex method completes the calculations and does not come close to the minimum.

4 ADE speed-up for parallel computation

The average number of function evaluations $\langle N_{\text{FE}} \rangle$ for ADE and ADE-ACM larger than that for classical methods. However, the use of parallel computations significantly reduces the actual time-out of the ADE methods [14] (see Fig. 4). The effectiveness of MPI-based[6] parallel implementation of the ADE algorithm have been tested on the multi-processor cluster CICC (LIT, JINR Dubna).

---

[6]Message Passing Interface (MPI). Standards documents are on http://www.mpi-forum.org/.

The calculations of the SFF model for PTNS-30 with two-step density approximation have been repeatedly run for different numbers of computational nodes: $N_{proc} \in [2, 150]$.

Since the ADE algorithm is non-deterministic, the number of function evaluations required to achieve a certain accuracy is varying and its average value is slightly increasing with the increase of the number of parallel processors used.

Fig. 4 shows nearly linear speed-up up to $N_{proc} = 150$ nodes involved in parallel computing. The solid line shows the resultant acceleration. The dotted line shows speed-up per one calculation of the objective function. This acceleration is associated with max N_{FE} stopping criteria.

5 Conclusions

The parameters of a polydispersed population of unilamellar vesicles PTNS, including the internal structure of the lipid bilayer of vesicles membrane have been estimated by the ADE-method. The optimization procedures implemented on the basis of ADE and ADE-ACM methods were compared with the OPs, implemented on the basis of commonly used methods SIMPLEX and MIGRAD. It is shown that the probability of obtaining a correct solution to the optimization problem by the ADE methods considerably exceeds the probabilities for the SIMPLEX and MIGRAD methods. The speed up in parallel MPI-based calculations has been estimated. Thus, it is shown that ADE is well suited for the parallel computation of the vesicular systems parameter estimations.

References

[1] Price K. V. and Storn, R.V., J. of Global Optimization **11**, 341–359 (1997)

[2] Price K.V., Storn R.M., Lampinen J.A *Differential Evolution: A Practical Approach to Global Optimization*. Springer (2005)

[3] Das S., Suganthan P.N., IEEE Trans. Evol. Comput. **15**, 4–31 (2011)

[4] Zhabitskaya E.I. and Zhabitsky M.V., Lecture Notes in Computer Science **7125**, 328–333 (2012)

[5] Zhabitskaya E. and Zhabitsky M., Lecture Notes in Computer Science **8236**, 555–561 (2013)

[6] Zhabitskaya E.I. and Zhabitsky M.V., *Proceeding of the 15th Annual Conference on Genetic and Evolutionary Computation*, (USA, New York, 2013) 455–462

[7] Kiselev M.A., Zemlianaia E.V., Zhabitskaia E.I. and Aksenov V.L., Crystallography Reports **60**, 143–147 (2015)

[8] Zhabitskaia E.I., Zemlianaia E.V. and Kiselev M.A., Vestnik Rossijskogo universiteta druzhby narodov, seriia: Matematika. Informatika. Fizika **2**, 253–259 (2014)

[9] Kiselev M.A., Zemlyanaya E.V., Ipatova O.M., Gruzinov A.Yu., Ermakova E.V., Zabelin A.V., Zhabitskaya E.I., Druzhilovskaya O.S. and Aksenov V.L., Journal of Pharmaceutical and Biomedical Analysis **114**, 288–291 (2015)

[10] Nelder J.A. and Mead R., Computer J. **7**, 308–313 (1965)

[11] Davidon W.C., *A.E.C. Res. and Develop. Report ANL-5990* (Argonne National Laboratory. Argonne; Illinois, 1959) p. 21

[12] Fletcher R. and Powell M.J.D., Comput. J. **6**, 163–168 (1963)

[13] James F. and Roos M., Computer Physics Communications **10**, 6, 343–446 (1975)

[14] Zhabitskaya E.I. and Zhabitsky M.V., Mathematical Modelling **24**, 12, 33–37 (2012)