



# Maximum likelihood array calibration using particle swarm optimisation

S. Wan P.-J. Chung B. Mulgrew

Institute for Digital Communications, School of Engineering, The University of Edinburgh, Edinburgh EH9 3JL, UK  
 E-mail: cqswan@gmail.com

**Abstract:** Calibration of array shape error is a key issue for most existing source localisation algorithms. In this study, the far-field self-calibration and near-field pilot-calibration are carried out using unconditional maximum likelihood (UML) estimator whose objective function is optimised by particle swarm optimisation (PSO). A new technique, decaying diagonal loading (DDL), is proposed to enhance the performance of PSO at high signal-to-noise ratio (SNR) by dynamically lowering it, based on the counter-intuitive observation that the global optimum of the UML objective function is more prominent at lower SNR. Numerical simulations demonstrate that the UML estimator optimised by PSO with DDL is robust to large shape errors, optimally accurate and free of the initialisation problem. In addition, the DDL technique can be coupled with different global optimisation algorithms for performance enhancement. Mathematical analysis indicates that the DDL is applicable to any array processing problem where the UML estimator is employed.

## 1 Introduction

Errors in sensor location can severely degrade the source localisation performance of a sensor array. A small perturbation in sensor location can often dominate all other causes of source localisation error [1, 2]. Therefore it is crucial to calibrate the array shape. Calibration methods can be classified into self-calibration and pilot-calibration, dependent on whether calibration sources are at known or unknown positions. The field region of an array can be divided into far field and near field, based on the distance from it. In this paper two physical scenarios are considered, the self-calibration using far-field sources and the pilot-calibration using near-field sources.

Rockah proposed a far-field self-calibration method that requires sources to be separable in the time or the frequency domain [3]. Weiss and Friedlander used objective functions based on conditional maximum likelihood (CML) [4] (termed WF1) and multiple signal classification (MUSIC) [5] (termed WF2) for the far-field self-calibration problem. In Weiss and Friedlander's methods, a first-order Taylor approximation is applied to the sensor location parameters to achieve an analytical optimisation, which reduces the computational complexity. However, the Taylor approximation requires that the perturbation is small, and exacerbates the MSE of the result. Flanagan and Bell [6] relieved the small perturbation constraint by preceding WF2 with a 'coarse calibration' procedure. Chung and Wan [7] used EM algorithm to optimise the CML objective function. This led to improved calibration accuracy under conditions of large perturbation and closely located sources, but lacks an initialisation mechanism.

The near-field problem emerges when a source is placed close to an array, for example when array calibration is

carried out in an anechoic chamber where the space is limited because of the cost of construction. The fundamental difference between near field and far field is whether the range information is neglected by the plane-wave approximation [8]. In the near field the range effect must be taken into account, which makes near-field array processing generally more complicated than far field. Previous work on near-field array processing has largely been on source localisation, where a popular technique for computational simplification is a second-order Taylor expansion, see [9] and the references therein. However, this technique is only applicable to the 'Fresnel region', which is the outer part of the near field, and the approximation error increases as the source comes closer to the array. Recently He *et al.* [10] proposed a non-parametric near-field calibration technique based on calibration matrix interpolation; however, it requires a single source at a large number of known positions that form a two-dimensional grid.

We propose a method that employs unconditional maximum likelihood (UML) estimator optimised by particle swarm optimisation (PSO) with decaying diagonal loading (DDL) for both far-field self-calibration and near-field pilot-calibration. The array shape is calibrated with an UML estimator, whose objective function is optimised by PSO. A new technique, DDL is proposed to enhance the performance of PSO at high signal-to-noise ratio (SNR) by dynamically lowering it, based on the counter-intuitive observation that the global optimum of the UML objective function is more prominent at lower SNR. Since no approximation is involved, the proposed method is applicable to the entire near field and large shape errors. Unlike local optimisation algorithms, PSO does not need accurate initialisation to reach the global optimum of the

objective function. Its failure at high SNR is prevented by the novel DDL technique that dynamically lowers the SNR. Thanks to the quality of the UML estimator, the calibration accuracy attains the Cramer-Rao bound (CRB). Finally, because of the parametric nature of the UML estimator the method requires only a few sources to calibrate an array.

The rest of this paper is organised as follows. The problem is formulated in Section 2, the PSO-DDL paradigm is laid out in Section 3, followed by simulation results in Section 4. Section 5 summarises and concludes the paper.

## 2 Problem formulation

Consider an array of  $N$  sensors receiving signals emitted by  $M$  sources. The sensors and the sources are assumed to be in a plane for the simplicity of exposition. The geometry of one sensor and one source is illustrated in Fig. 1. The sources are with bearings  $\theta = [\theta_1, \dots, \theta_M]^T$  and ranges  $\rho = [\rho_1, \dots, \rho_M]^T$ . The known nominal location of the  $n$ th sensor is given by  $[x_n, y_n]^T$ , which is perturbed by unknown error  $[\Delta x_n, \Delta y_n]^T$  to the actual location  $[x_n + \Delta x_n, y_n + \Delta y_n]^T$ . The perturbations are summarised in vector  $\delta = [\Delta x_1, \Delta y_1, \dots, \Delta x_N, \Delta y_N]^T$ , which is considered constant in a calibration process.

For narrow band signals, the  $K$  snapshots of the array output can be expressed in the Fourier domain as

$$z(k) = Hs(k) + n(k), \quad k = 1, \dots, K \quad (1)$$

The  $n, m$ th element of the array manifold  $H$  is given by

$$H_{nm} = a_{nm} \exp\{j\omega\tau_{nm}\}, \quad n = 1, \dots, N, \quad m = 1, \dots, M \quad (2)$$

where  $a_{nm}$  is the attenuation factor and  $\exp\{j\omega\tau_{nm}\}$  is the phase factor, both from the  $m$ th source to the  $n$ th sensor. Here  $\omega$  is the angular frequency of the signal wave and  $\tau_{nm}$  is the time delay from the  $m$ th source to the  $n$ th sensor. For a spherical wave travelling in free space, which is the case considered in this paper

$$a_{nm} = \frac{1}{d_{nm}}, \quad \tau_{nm} = \frac{d_{nm}}{c}, \quad \omega = \frac{2\pi}{\lambda}c \quad (3)$$

where  $c$  is the wave speed,  $\lambda$  is the wavelength and

$$d_{nm} = \{[\rho_m \sin \theta_m - (x_n + \Delta x_n)]^2 + [\rho_m \cos \theta_m - (y_n + \Delta y_n)]^2\}^{1/2}$$

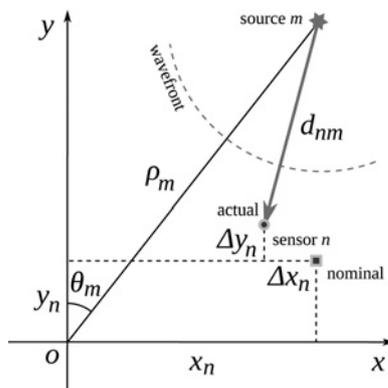


Fig. 1 Sensor–source geometry in near field

is the distance between the  $n$ th sensor and the  $m$ th source. Inserting (3) into (2),  $H_{nm}$  can be expressed as

$$H_{nm} = \frac{1}{d_{nm}} \exp\left\{j \frac{2\pi}{\lambda} d_{nm}\right\} \quad (4)$$

If the sources are far from the array,  $\rho_m > 2D_a^2/\lambda$  [8], where  $D_a$  is the dimension of the array, the far-field approximation [11] simplifies  $H_{nm}$  to

$$H_{nm} = \exp\left\{j \frac{2\pi}{\lambda} [(x_n + \Delta x_n) \sin \theta_m + (y_n + \Delta y_n) \cos \theta_m]\right\} \quad (5)$$

The  $H_{nm}$  for near field differs from that for far field in two aspects: one is the attenuation factor  $1/d_{nm}$ , which denotes the signal strength loss in free space; and the other is the phase shift  $(2\pi/\lambda)d_{nm}$ , which denotes the phase delay from the source to the sensor. They are both functions of source bearing and range. For the far-field case, the signal strength is considered the same across the array, thus the magnitude factor is absent; the plane wave approximation is applied, thus the phase difference between two sensors is simply determined by their distance projected to the source bearing. As a result the information of source range is lost and has to be recovered by other means.

In (1), the signal vector  $s(k)$  is considered to be stochastic. The noise vector  $n(k)$  is independent, identically complex normally distributed with zero mean and covariance matrix  $\nu I$ , where  $\nu$  is an unknown noise spectral parameter and  $I$  is an  $N \times N$  identity matrix. Given the observations  $\{z(k)\}_1^K$ , the problem of central interest is to estimate the parameters  $\alpha$  of  $H(\alpha)$ .  $H(\alpha)$  is  $H$  with its parameter vector  $\alpha$  explicitly expressed. The content of  $\alpha$  is determined by the physical scenario. For instance, for a self-calibration  $\alpha$  contains both the source and sensor positions, whereas for a pilot-calibration  $\alpha$  contains the sensor positions only. This will be specified in Section 4. Following [11], the UML estimator for  $\alpha$  can be expressed as

$$\hat{\alpha} = \arg \min_{\alpha} \left\{ \det \left[ \underbrace{P_H \hat{R} P_H + \frac{\text{tr}(P_H^\perp \hat{R} P_H^\perp)}{N - M}}_{F(\alpha)} \right] \right\} \quad (6)$$

where  $P_H = H(H^H H)^{-1} H^H$  is the projection matrix of  $H$ , and  $P_H^\perp = I - P_H$  is its orthogonal complement.  $\hat{R} = (1/K) \sum_{k=1}^K z(k)z^H(k)$  is the sample covariance matrix.  $F(\alpha)$  denotes the objective function, and is notationally simplified to  $F$  where appropriate.

The UML objective function in (6) is non-linear, multi-dimensional and multi-modal, which poses a difficult optimisation problem. Exhaustive search is computationally unfeasible for high-dimensional optimisation; local optimisation methods, such as the gradient algorithms and the EM algorithm, although computationally efficient, suffer from the initialisation problem. We thus propose stochastic global optimisation for the task, which does not need accurate initialisation, yet is computationally feasible.

In addition to the multimodality that fails the local optimisers, the UML objective function presents another difficulty to the global optimisers: modes of similar height at high SNR. Although the height of the global optimum compared with the local optimum is irrelevant to the success of a local optimiser, this height contrast is essential

to a global optimiser as a global optimiser searches the entire solution space and a prominent global optimum is easier to find. The height contrast between the global optimum and the local optima of  $F$  is visualised at the end of this section, and a mathematical analysis of the height contrast is provided in the Appendix. The difficulty caused by the similar height of the optima to the global optimiser is also shown by simulations in Section 4.

The visualisation and mathematical analysis are carried out after the concepts involved are established. It is a simple matter to establish that lowering the SNR by adding spatially white noise is asymptotically equivalent to diagonally loading the covariance matrix:  $\tilde{\mathbf{R}} = \mathbf{R} + \nu_l \mathbf{I}$ , where  $\mathbf{R} = \lim_{K \rightarrow \infty} \hat{\mathbf{R}}$  and  $\nu_l$  is the power of the loaded noise. Assuming normalised signal power, the original signal-to-noise ratio can be expressed as  $\text{SNR} = -10 \log \nu$ . We then define the effective signal-to-noise ratio after DL as

$$\text{SNR}_a = -10 \log(\nu + \nu_l) \quad (7)$$

and consequently the loaded signal-to-noise ratio as  $\text{SNR}_l = \text{SNR}_a - \text{SNR}$ .

Since it is impossible to visualise a high-dimensional  $F(\boldsymbol{\alpha})$ , we have to demonstrate the effect of DL using low-dimensional problems such as two-dimensional direction of arrival (DOA) estimation problem and calibration problem. Fig. 2 shows the  $F(\boldsymbol{\alpha})$  for a DOA estimation problem with the two source DOA  $\theta_1$  and  $\theta_2$  being the content of the parameter vector  $\boldsymbol{\alpha}$ . The DOA estimation problem can also be viewed as a degenerated form of a far-field self-calibration problem. Fig. 3 shows the  $F(\boldsymbol{\alpha})$  for a near-field calibration problem with one sensor's deviation from its true location,  $\delta x_1$  and  $\delta y_1$ , being the content of the parameter vector  $\boldsymbol{\alpha}$ .

Comparing Fig. 2a with b, and Figs. 3a with b, we see that diagonal loading of the sample covariance matrix  $\hat{\mathbf{R}}$ , which is asymptotically equivalent to reducing the SNR, makes prominent the global minimum of the objective function. This eases searching it for a global optimiser. In addition to the cases shown in Figs. 2 and 3, more visualisations reveal that the prominence phenomenon is independent of the

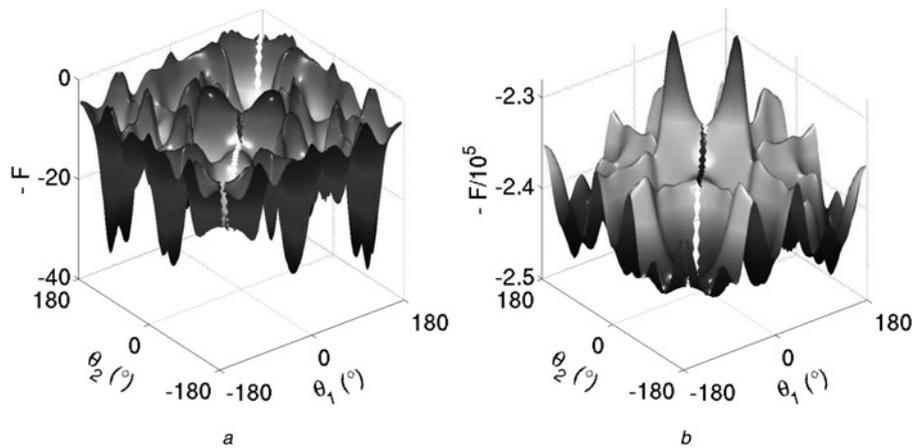


Fig. 2 Negative objective function

a Before diagonal loading  
b After diagonal loading

Far-field DOA estimation: 5-sensor uniform circular array, inter-sensor spacing  $\lambda/2$ ; 2 signals at  $[-35 \ 35]^\circ$ ,  $K = 1000$ ;  $\text{SNR} = 14$  dB,  $\text{SNR}_a = -10$  dB

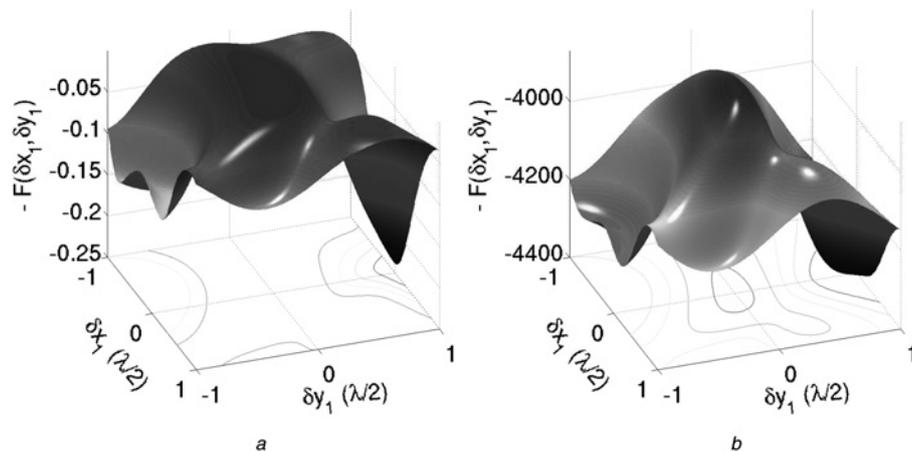


Fig. 3 Negative objective function

a Before diagonal loading  
b After diagonal loading

Near-field pilot-calibration: 6-sensor uniform circular array with  $a = \lambda/2$  inter-sensor spacing, perturbed by uniform random displacements with boundary  $b = 0.1a$ . three sources at bearings  $-120^\circ$ ,  $0^\circ$  and  $120^\circ$ , all with the same range  $r = 2a$  relative to the centre of the nominal array,  $K = 1000$ .  $\text{SNR} = 14$  dB,  $\text{SNR}_a = -5$  dB

number of sensors and sources, the geometry of the array, and the near-field and far-field scenarios. This SNR-varied prominence of the global minimum of  $F$  is also indicated by the mathematical analysis in the Appendix.

### 3 PSO-DDL technique

PSO [12] is chosen as the global optimiser to address the multi-modality, multi-dimensionality and non-linearity of the UML objective function, because of its optimisation competence and implementation simplicity. A novel technique, DDL, is proposed to solve the problem of modes of similar height, based on the observation that the global optimum of  $F$  is more prominent at lower SNR.

#### 3.1 PSO for array shape calibration

PSO [12] is a stochastic global optimisation technique that is simple in its kernel and robust in its control parameters. It is reported that PSO is computationally less intensive than the GA for a number of array processing and other problems [13]. In the following, a basic PSO is introduced, which is adequate for the problems considered. Other PSO variants that are also applicable can be found in [14] and the references therein. Firstly the problem-independent PSO kernel is described, then the problem-specific strategies follow.

The problem-independent PSO kernel is a constriction PSO [15]. In it, a swarm particle is a point in the  $D$ -dimensional solution space of the optimisation problem, whose coordinate is  $\xi_i(t) \in \mathbb{R}^D$ , where  $t$  is the time index and  $i$  denotes the  $i$ th particle,  $i = 1, 2, \dots, P$ .  $\xi_i(t)$  consists of the parameters to be estimated. For example, in a far-field self-calibration problem  $\xi_i(t)$  consists of the source DOA  $[\theta_1, \dots, \theta_M]^T$  and the sensor location errors  $[\Delta x_1, \Delta y_1, \dots, \Delta x_N, \Delta y_N]^T$ , thus  $D = M + 2N$ .  $P$  particles constitute a swarm, wherein each particle is propelled by random movement in the solution space but is guided towards the global optimum of the objective function by swarm intelligence – the processing of shared information of the objective function.

At the  $t$ th time step, the velocity of the  $i$ th particle  $\psi_i(t)$  is updated according to

$$\begin{aligned} \psi_i(t+1) = & \chi(\psi_i(t) + \varphi_1 \mathbf{r}_1 \odot [\mathbf{p}_i(t) - \xi_i(t)] \\ & + \varphi_2 \mathbf{r}_2 \odot [\mathbf{g}_i(t) - \xi_i(t)]), \quad i = 1, \dots, P \end{aligned} \quad (8)$$

where  $\odot$  is the element-wise product operator. The three terms in the brackets of (8) account for inertia, individual experience and group information, sequentially. Specifically,  $\mathbf{r}_1$  and  $\mathbf{r}_2$  are random vectors uniformly distributed in  $[0,1]$  that propel the particles,  $\mathbf{p}_i(t)$  is the particle best location of its own experience from time step 1 to  $t$ ,  $\mathbf{g}_i(t)$  is the group best location at current time step  $t$ . ‘Best’ is in the sense of the lowest value of the cost function to be minimised. The scalar  $\chi = 2/|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}|$  is the constriction factor that progressively constricts the harmful particle oscillation resulting from the random movement. Here  $\varphi = \varphi_1 + \varphi_2$ , in which  $\varphi_1$  is the acceleration constant towards the particle best location and thus a large value of it encourages exploration of the solution space; and  $\varphi_2$  is the acceleration constant towards the group best location and thus a large value of it encourages exploitation of potential districts of global optimum.  $\varphi$  should be greater than 4 to prevent particle oscillation [15], and in practice  $\varphi = 4.1$  is a viable value [14]. For unit time step, the position of the  $i$ th particle

is updated as

$$\xi_i(t+1) = \xi_i(t) + \psi_i(t+1), \quad i = 1, \dots, P \quad (9)$$

In order to prevent the particle from flying too fast and thus overshooting the target or exceeding the boundary of the solution space, the particle velocity is limited by a maximal value  $\Psi$ . If the velocity in any dimension  $d$  exceeds  $\Psi$ , it is clamped by

$$\psi_{id} = \Psi \frac{\psi_{id}}{|\psi_{id}|}, \quad d = 1, \dots, D \quad (10)$$

where  $\psi_{id}$  is the  $d$ th component of  $\psi_i$ . If a particle flies out of the boundary of the solution space in any dimension, it is reset to a random position in that dimension. In the case that the parameters to be optimised have different boundaries, one may map the feasible ranges into  $[0,1]$ . As a result the velocity can be updated uniformly, and  $\Psi$  is the same for all the dimensions.

Two strategies specific to the array calibration problem are stated below.

1. If at any PSO iteration, two DOA's happened to be closer than a pre-specified value  $\delta$ , one of them is assigned a random value [13]. This is to prevent  $\mathbf{H}$  from rank-deficiency.
2. The group best location  $\mathbf{g}_i(t)$  in (8) is the coordinate of the best particle in the  $i$ th group which is a subset of the particle swarm. The definition of ‘group’ determines the topology of the swarm. For instance, a swarm is of ‘star topology’ if one of the particles compares its objective function value with all the other’s and each of the other particles compares with the one only. Swarm topologies with less comparisons between the particles have a higher success probability of reaching the global optimum, at the expense of more iterations [16]. We therefore design a ‘string topology’ in which a particle only compares with its two immediate neighbours or one immediate neighbour if it is at the either end, forming a ‘string’

$$\text{Particles of the } i\text{th group} \begin{cases} i, i+1, & i = 1 \\ i-1, i, i+1, & i = 2, \dots, P-1 \\ i-1, i, & i = P \end{cases}$$

#### 3.2 Decaying diagonal loading

When the PSO is used to optimise the objective function  $F$  in (6), it often fails to reach the global minimum of  $F$  at high SNR. Visualisation of  $F$  in Section 2 reveals the reason being that the global minimum of  $F$  is more prominent at lower SNR than at higher SNR. An analysis of the first partial derivative of  $F$  in the Appendix indicates that this observation is a general feature of  $F$  regardless of the structure of  $\mathbf{H}(\boldsymbol{\alpha})$  and the content of the parameter vector  $\boldsymbol{\alpha}$ . Therefore, a new technique, DDL is proposed to make the global minimum more prominent and consequently enhance the performance of PSO.

DL is a common technique for performance improvement in adaptive beamforming, see [17] and the references therein. A constant amount of DL is added to the data covariance matrix, permanently changing the objective function, and an ‘optimal’ amount has to be determined to control the adverse effects [17]. However, observing that the particles of PSO congregate as the algorithm iterates on [14], which is a common

characteristic of population-based stochastic optimisation algorithms such as genetic, simulated annealing and differential evolution (DE) algorithms, we propose a DDL, whose amount reduces to zero at the end of the optimisation to avoid any adverse effect DL may cause to the objective function.

In DDL, the amount of DL is sufficiently large in the beginning to make the global minimum prominent and thus guides the PSO. It reduces to zero in the end and thus completely recovers the original  $F(\alpha)$ . DDL differs from existing DL techniques for improving performance in that it does not permanently alter the objective function. The objective function eventually optimised is the original one, therefore the quality of the estimator, such as precision and resolution, is preserved.

In DDL, the amount of loading exponentially decays (see (11))

where  $r_l \in (0, 1)$  is the ratio of loaded iteration,  $T$  is the maximum iteration and  $[\cdot]$  stands for the rounding operator. PSO-DDL is not sensitive to  $\text{SNR}_l$  as indicated by simulation. Therefore, in the case that the original SNR is unknown, one can set the SNR in (11) to the highest possible value and load  $\hat{R}$  in large amount without affecting the results. [18] contains a DDL schedule that is independent of the SNR based on this principle.

There also exist other transformations of objective function to enhance optimisation result, such as eliminating local

minima to make global minimum prominent [19], and partially convexifying the objective function [20], which differ from DDL in technique and effect on the objective function. In addition, they both permanently change the objective function.

To conclude Section 3, the PSO-DDL technique is summarised in Fig. 4.

## 4 Simulation results

Two physical scenarios are used to demonstrated the performance of PSO-DDL, the far-field self-calibration and near-field pilot-calibration. In addition, DDL is also coupled with the DE algorithm [21] to support the remark that the DL effect is independent of specific optimisation algorithm. For the far-field self-calibration, the array manifold is expressed by (5), and the content of the parameter vector  $\alpha$  is the source DOA's  $[\theta_1, \dots, \theta_M]^T$  and the sensor location perturbations  $\delta$ . For the near-field pilot-calibration, the array manifold is expressed by (4), and the content of the parameter vector  $\alpha$  is the sensor location perturbations  $\delta$ .

### 4.1 Far-field self-calibration

A 5-sensor 3-source geometry is considered. The nominal array is a uniform circular array with inter-sensor spacing  $a = (\lambda/2)$ , whose sensor locations are perturbed by  $\Delta x_n$  and

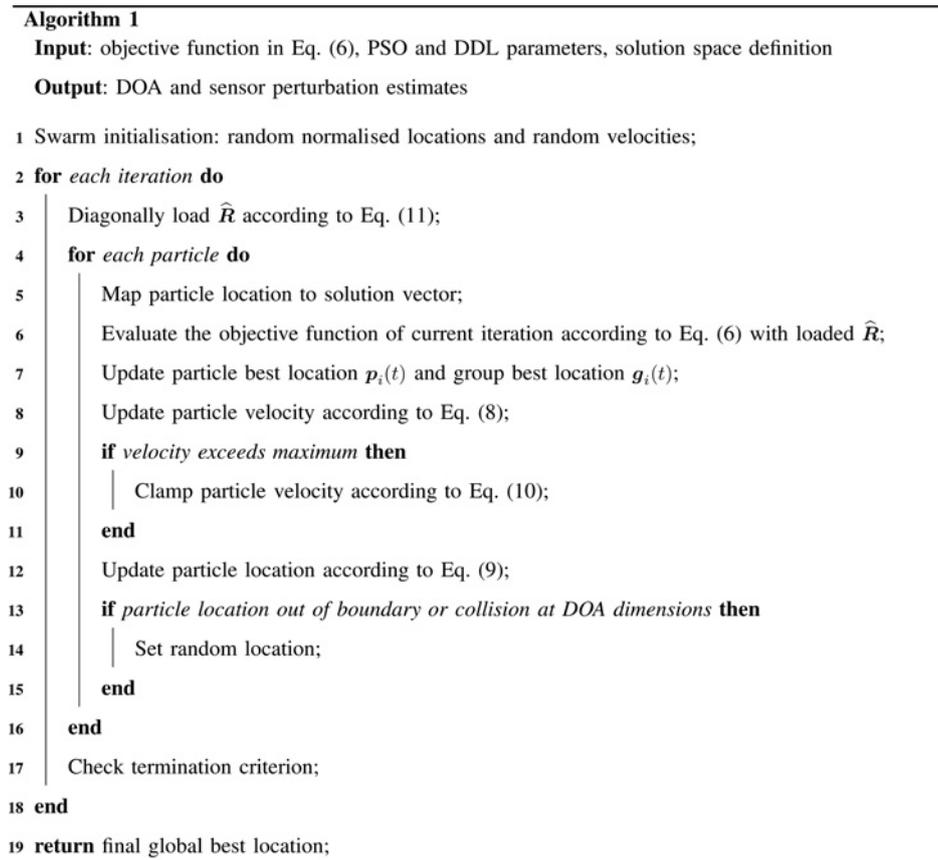


Fig. 4 Particle swarm optimisation-decaying diagonal loading

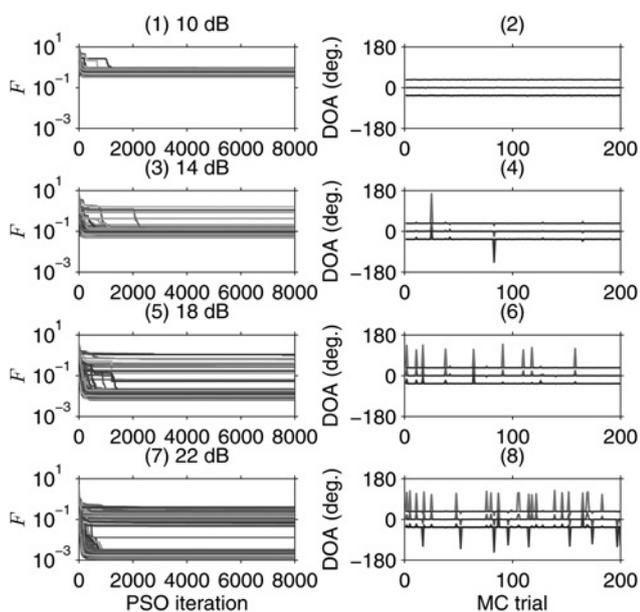
$$v_l(t) = \begin{cases} 10^{-(\text{SNR}/10)} \{ 10^{-(\text{SNR}_l/10)(1-\{(t-1)/[r_l T]\})} - 1 \}, & t = 1, \dots, [r_l T] \\ 0, & t = [r_l T] + 1, \dots, T \end{cases} \quad (11)$$

$\Delta y_n$  that are independently, uniformly distributed in  $[-b, b]$ , in which  $b$  is the sensor location tolerance.  $\Delta x_1$ ,  $\Delta y_1$  and  $\Delta y_2$  are set to zero according to [1] to fulfil the identifiability condition. Following [4, 5], the three sources are set at  $\theta_0 = [-35^\circ \ 0^\circ \ 35^\circ]^T$ , which however is unknown to the self-calibration procedure. The three sources emit equi-power, uncorrelated signals. The SNR simulated ranges from  $-18$  to  $38$  dB with  $4$  dB step size, and  $200$  Monte Carlo (MC) trials are performed for each SNR. The array shape is the same for the entire SNR range but is different in each MC trial. The error measure is the sum of squared error (SE sum) of the three signals,  $\|\hat{\theta} - \theta_0\|^2$ , averaged over the  $200$  trials. The control parameters for PSO and DDL are summarised in Table 1. PSO is terminated after the maximum number of iterations,  $T$ , is reached, which amounts to  $PT$  objective function evaluations per optimisation.

Firstly, the performances of PSO alone and PSO-DDL are compared. PSO starts to fail when the SNR is higher than a certain threshold, and the proportion of failure increases as the SNR increases. An example of the threshold and failure increase is shown in Fig. 5. The left-hand column shows the convergence of the cost function  $F$  over  $200$  trial runs at

**Table 1** PSO and DDL control parameters for far-field self-calibration

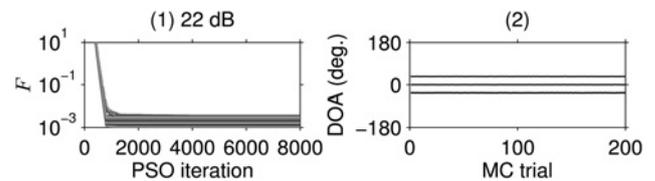
PSO parameters						DDL parameters	
$\varphi_1$	$\varphi_2$	$\Delta$	$P$	$\Psi$	$T$	SNR <sub>a</sub>	$r_l$
2.4	1.7	$10^{-4}$ rad	30	0.5	3000	$-15$ dB	0.13



**Fig. 5** PSO performance at critical SNR's

Left column  $F$ : convergence over  $200$  MC trials at 4 values of SNR  
Right column: Final converged value of the DOA estimates for each of the trials

Nominal array is a 5-sensor uniform circular array with inter-sensor spacing  $a = \lambda/2$ , perturbed to actual array shape by uniformly distributed random displacements with boundary  $b = 0.2a$ . Three sources of opportunity reside at DOA  $\theta_0 = [-35^\circ \ 0^\circ \ 35^\circ]^T$ , which are unknown a priori to the estimator



**Fig. 6** PSO-DDL performance at SNR = 22 dB

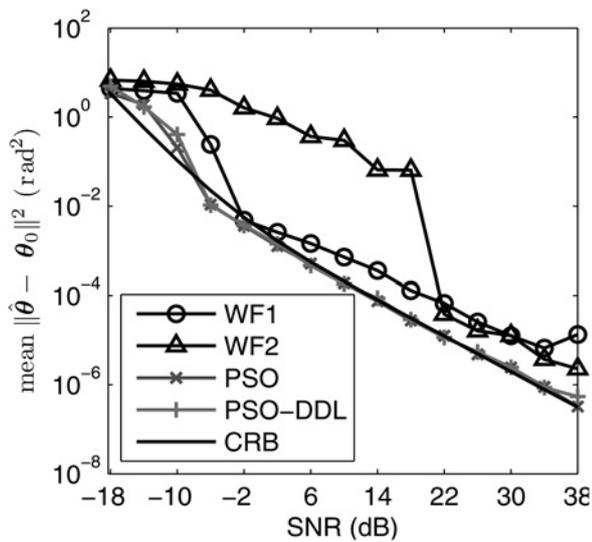
Nominal array is a 5-sensor uniform circular array with inter-sensor spacing  $a = \lambda/2$ , perturbed to actual array shape by uniformly distributed random displacements with boundary  $b = 0.2a$ . Three sources of opportunity reside at DOA  $\theta_0 = [-35^\circ \ 0^\circ \ 35^\circ]^T$ , which are unknown a priori to the estimator

4 values of SNR. The right-hand column shows the final converged value of the angle estimates for each of the trials. It can be seen that at SNR = 10 dB all the  $200$  DOA estimates are close to true value, but from SNR = 14 dB on, spiky estimates appear indicating that some runs produce poor estimates. The number of spikes increases as the SNR increases.

In contrast, Fig. 6(1) shows that the PSO-DDL optimised values of  $F$  are well grouped in one region instead of separated in two bundles. In addition, for the first  $800$  PSO iterations, which are diagonal loaded ones, the optimised  $F$  value follows an approximately exponential decline that expresses the effect of DDL; Fig. 6(2) shows that all the calibrated DOA estimates are close to true values rather than many an outlier straying far away in Fig. 5(8). The SNR shown in Fig. 6 is 22 dB, which is the highest SNR in Fig. 5. At the other SNR the improvements by DDL are similar to Fig. 6 and thus the results are omitted.

Secondly, the statistical results of PSO-DDL are compared with the WF1 and WF2 algorithms, and the CRB which is based on stochastic signal assumption (Section 8.11.2 of [11]). The  $K = 1000$  results are presented here. Cases with a small number of snapshots (e.g.  $K = 20$ ) can be found in [22] with similar results. WF2 is initialised at true DOA for convenience. The results are not compared with [6] as its final step is WF2 and are also not compared with [7] as it lacks the initialisation step. Although the coarse calibration steps of [6] can be applied to PSO-DDL for larger  $b$ , and existing DOA estimation algorithms can be used to initialise [7], a thorough investigation is out of the scope of this work as initialisers differ in sensitivity to array shape perturbation and consequently affect the self-calibration result.

In the first experiment the perturbation boundary  $b = 0.05a$  is considered. The DOA ranges from  $-180^\circ$  to  $180^\circ$ . One observes from Fig. 7 that both WF1 and WF2 succeed in calibrating the array shape at SNR's higher than 22 dB, with a similar MSE sum. Owing to the fact that WF2 is based on the eigen-structure objective function, it shows a threshold at 22 dB, below which the MSE sum increases dramatically. WF1 shows a much lower threshold at  $-2$  dB, because of its CML objective function. Nevertheless, neither of the two algorithms attains the CRB at an SNR higher than their thresholds. This results from the Taylor approximation that renders the bias the same order of magnitude as the STD, making the MSE considerably higher than the CRB despite the STD approaching the CRB [4–6]. PSO-DDL shows a threshold at  $-6$  dB which is lower than both WF1 and WF2. Above this SNR, its MSE sum attains the CRB. This MSE sum is apparently lower than that of both WF1 and WF2. This experiment shows that under small perturbation, PSO-DDL has optimal accuracy that is better than WF1 and WF2,

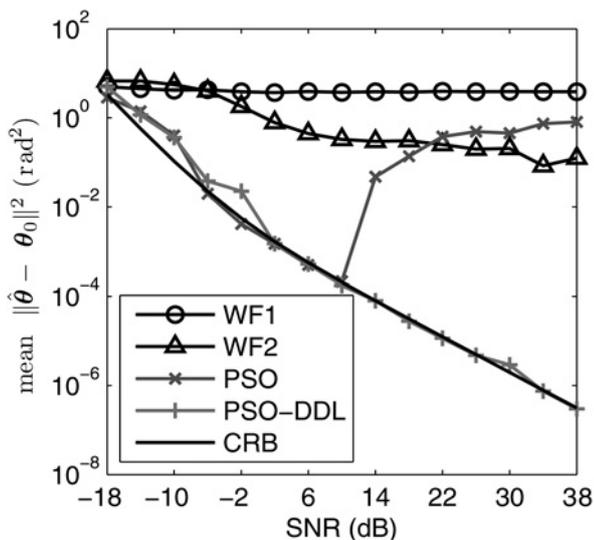


**Fig. 7** Comparison of self-calibration accuracy under small sensor location error

5-sensor uniform circular array with  $a = \lambda/2$  inter-sensor spacing, perturbed by uniform random displacements with boundary  $b = 0.05a$ . Three sources of opportunity reside at DOA  $\theta_0 = [-35^\circ \ 0^\circ \ 35^\circ]^T$ , which are unknown a priori to the estimator

and has a lower threshold SNR. One also observes from Fig. 7 that PSO-DDL has almost identical performance with PSO alone in this case, which reflects that DDL recovers the original objective function eventually.

In the second experiment the perturbation boundary  $b$  is enlarged to  $0.2a$ . Two of the PSO and DDL parameters are adjusted to  $T = 8000$  and  $r_l = 0.1$  for the more challenging case. Other settings are the same as in the first experiment. It is observed from Fig. 8 that PSO-DDL succeeds in calibrating the array with CRB-attaining MSE sum of DOA parameters at SNR's greater than 2 dB. Both WF1 and WF2 fail with estimation errors much higher than PSO-DDL and the CRB. The errors of WF1 and WF2 do not



**Fig. 8** Comparison of self-calibration accuracy under large sensor location error

5-sensor uniform circular array with  $a = \lambda/2$  inter-sensor spacing, perturbed by uniform random displacements with boundary  $b = 0.2a$ . Three sources of opportunity reside at DOA  $\theta_0 = [-35^\circ \ 0^\circ \ 35^\circ]^T$ , which are unknown a priori to the estimator

decrease as the SNR increases. We also note that without DDL, PSO alone would fail with large estimation error, comparable to WF1 and WF2, when the SNR is greater than 10 dB. This experiment shows that PSO-DDL has much higher robustness against large sensor location errors than both WF1 and WF2, and it retains optimal accuracy at large perturbation.

#### 4.2 Near-field pilot-calibration

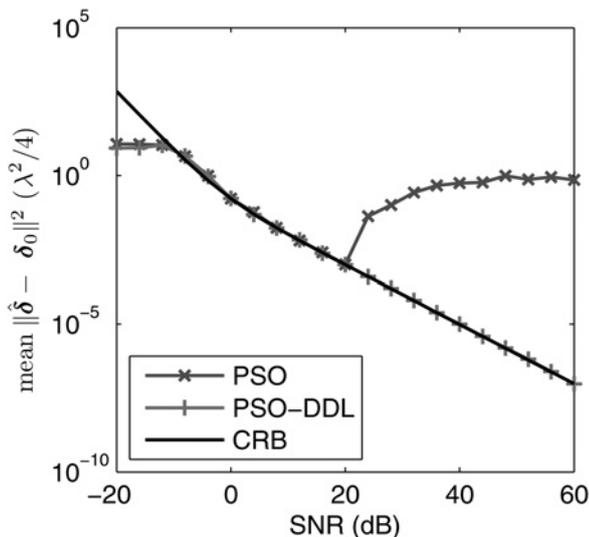
Two array geometries, the circular array (CA) and the linear array (LA), are considered. The first case is the calibration of a 20-sensor CA using three near-field sources. Specifically, the nominal array is a 20-sensor CA with  $a = \lambda/2$  inter-sensor spacing, perturbed by  $\Delta x_n$  and  $\Delta y_n$  that are independently, uniformly distributed in  $[-b, b]$ , in which  $b = 0.65a$  is the sensor location tolerance. The perturbed array shape is the same for all the MC trials. The three sources reside at bearings  $-120^\circ, 0^\circ$  and  $120^\circ$ , all with range  $5a$  relative to the centre of the nominal array, emitting equipower, uncorrelated signals. The source locations are known to the pilot-calibration procedure. The SNR simulated varies from  $-20$  to  $60$  dB with 4 dB step size, and 200 MC trials are performed for each SNR. The error measure is the sum of squared errors for the 20 sensors,  $\|\hat{\delta} - \delta_0\|^2$ , averaged over the 200 trials.  $\delta_0$  denotes the true perturbations. The control parameters for PSO and DDL are specified in Table 2. As in pilot-calibration the source locations are not parameters to be estimated, the ‘collision avoidance’ step (line 13–15) in Fig. 4 can be omitted and the PSO control parameter  $\delta$  is absent in Table 2. Compared with the parameters for far-field self-calibration in Table 1, the PSO is enhanced in search ability to accommodate the solution space of higher dimensions and a larger range in each dimension, at the cost of more computation load which is proportional to  $PT$ .

The identifiability condition is given in [23], and is satisfied here. The near field [8] is defined as range  $\rho < 2D_a^2/\lambda$ , where  $D_a$  is the dimension of the array. In this case  $D_a = 3.20\lambda$  is the diameter of the array, hence  $\rho < 20.43\lambda$  is the near field. The Fresnel region [8] is defined as  $0.62\sqrt{D_a^3/\lambda} < \rho < 2D_a^2/\lambda$ , which is  $3.54\lambda < \rho < 20.43\lambda$  in this case. The range of the sources  $\rho_m = 2.5\lambda$  falls in the inner part of the near field, beyond the Fresnel region.

The calibration result is illustrated in Fig. 9. The mean squared error (MSE) sums of PSO and PSO-DDL are compared with each other and with the CRB. The technique of [10] is not compared as it is not a parametric method that explicitly estimates the sensor locations. For  $\text{SNR} \leq 20$  dB, PSO yields accurate results that attain the CRB. However, it shows a threshold at  $\text{SNR} = 20$  dB, above that SNR the MSE sum increases and eventually levels off at a high value that is comparable to that of  $\text{SNR} = -4$  dB. In contrast, the MSE sum of PSO-DDL result attains the CRB in the entire SNR range simulated. Its level-off in  $-12$  to  $-20$  dB is a result of the assumption that the sensor location error is limited to  $[-0.65a, 0.65a]$ . It is well known that CRB is an asymptotic bound that is invalid at low SNR: when the

**Table 2** PSO and DDL control parameters for near-field pilot-calibration

PSO parameters					DDL parameters	
$\varphi_1$	$\varphi_2$	$P$	$\Psi$	$T$	$\text{SNR}_a$	$r_l$
2.7	1.4	100	0.01	16000	-15 dB	0.4



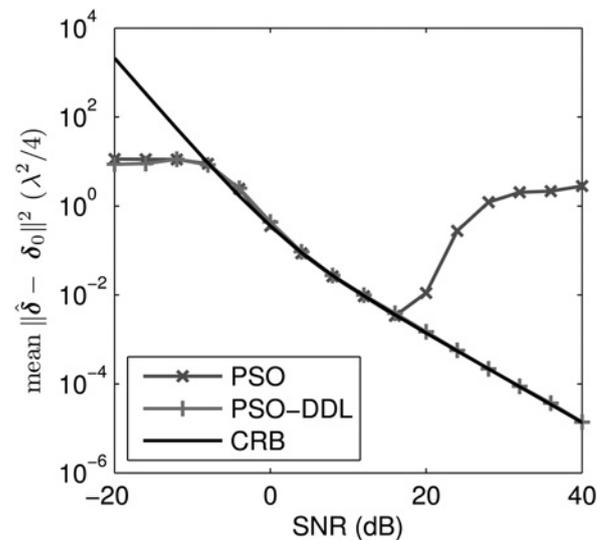
**Fig. 9** Comparison of PSO and PSO-DDL's MSE's and the CRB for pilot-calibration

Nominal array is a 20-sensor circular array with  $a = \lambda/2$  inter-sensor spacing, perturbed by uniform random displacements with boundary  $b = 0.65a$ . The three sources reside at bearings  $-120^\circ$ ,  $0^\circ$  and  $120^\circ$ , all with range  $5a$  relative to the centre of the nominal array. The source locations are known a priori to the estimator

SNR falls to  $-\infty$  the CRB goes to  $\infty$ , so the MSE sum is bound to be lower than the CRB at low SNR. Another observation is that the MSE sums of neither PSO nor PSO-DDL demonstrate a threshold at low SNR, typically about 0 dB, where usually the MSE of DOA estimation and self-calibration sharply rises above the CRB.

The second case is the calibration of a 20-sensor LA using three near-field sources. Contrary to far-field self-calibration where LA is unidentifiable for geometrical reasons, in near-field pilot-calibration LA is identifiable under mild conditions [23], which are satisfied here. In the simulation, except for the array geometry and the SNR range, the other settings of the LA calibration are the same as those of the CA calibration in this subsection. The nominal LA is of 20 sensor, with half-wavelength inter-sensor spacing. It lies on the  $x$ -axis of the Cartesian coordinates and its centroid coincides with the origin. The sensor locations are displaced to the actual locations by the same perturbation as to the CA in this subsection. All the other simulation settings, such as calibration source placement, signal and noise environment, are also the same as the CA simulation in this subsection.

The calibration result is illustrated in Fig. 10. The MSE sums of PSO and PSO-DDL are compared with each other and with the CRB. The SNR simulated varies from  $-20$  to  $40$  dB with 4 dB step size. For  $\text{SNR} \leq 16$  dB, PSO yields accurate results that attain the CRB. However, it shows a threshold at  $\text{SNR} = 16$  dB, above that SNR the MSE sum increases and eventually levels off at a high value that is comparable to that of  $\text{SNR} = -4$  dB. In contrast, the MSE sum of PSO-DDL result attains the CRB in the entire SNR range simulated. Its level-off in  $-8$  to  $-20$  dB is a result of the assumption that the sensor location error is limited to  $[-0.65a, 0.65a]$ . It is well known that CRB is an asymptotic bound that is invalid at low SNR: when the SNR falls to  $-\infty$  the CRB goes to  $\infty$ , so the MSE sum is bound to be lower than the CRB at low SNR. Another observation is that the MSE sums of neither PSO nor PSO-DDL demonstrate a threshold at low SNR, typically about



**Fig. 10** Comparison of PSO and PSO-DDL's MSE's and the CRB for pilot-calibration

Nominal array is a 20-sensor linear array with  $a = \lambda/2$  inter-sensor spacing. It lies on the  $x$ -axis of the Cartesian coordinates and its centroid coincides with the origin. The sensor locations are displaced to the actual locations by the same perturbation as to the circular array in this subsection. The three sources reside at bearings  $-120^\circ$ ,  $0^\circ$  and  $120^\circ$  relative to the  $y$ -axis, all with range  $5a$  relative to the origin. The source locations are known a priori to the estimator

0 dB, where usually the MSE of DOA estimation and self-calibration sharply rises above the CRB.

Compared with the CA calibration in this subsection, PSO for LA calibration shows a threshold 4 dB lower than that for CA, and the start of level-off at low SNR for both PSO and PSO-DDL is 4 dB higher for LA calibration than CA calibration. These indicate that the calibration of LA is slightly more difficult than CA in this configuration.

#### 4.3 Differential evolution (DE) results

In support of the remark that DDL is a universal technique that can be coupled with other algorithms to improve performance, a different stochastic global optimisation algorithm, DE [21] has also been employed to optimise the UML objective function for both far-field self-calibration and near-field pilot-calibration. The details of DE with DDL for far-field self-calibration can be found in [24]. The result of DE with DDL for CA near-field pilot-calibration as that in Section 4.2 is summarised as follows, whereas the simulation details can be found in [22].

Fig. 11 shows the simulation results of the same DE with and without the aid of DDL. The array-source settings are exactly the same as the CA near-field pilot-calibration in Section 4.2, or in other words  $F$  is the same. The SNR shown is 26 dB. Fig. 11a shows that a significant proportion of the 200 MC trials performed by DE alone are trapped in local minima, ending up with optimised  $F$  values grouped higher than the global minimum bundle, and correspondingly Fig. 11b shows a significant proportion of spiky false estimates of sensor location. In contrast, Fig. 11c shows that DE-DDL optimises  $F$  all to the global minimum, the slope between the 1st and the 400th DE iterations reflecting the effect of DDL, and correspondingly Fig. 11d shows that all the sensor location estimates are close to the true values. Note that this is achieved in one-tenth of the number of iterations of DE alone.

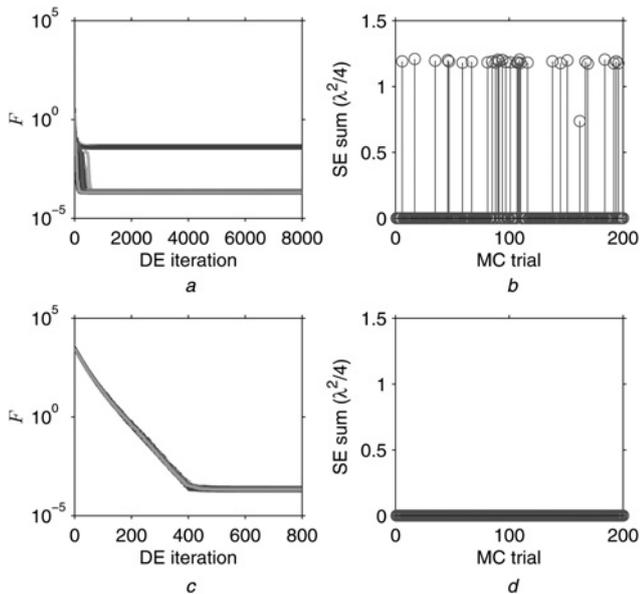


Fig. 11 DE performance at SNR = 22 dB

a and b DE

c and d DE-DDL

Near-field pilot-calibration. The nominal array is a 20-sensor circular array with  $a = \lambda/2$  inter-sensor spacing, perturbed by uniform random displacements with boundary  $b = 0.65a$ . The three sources reside at bearings  $-120^\circ$ ,  $0^\circ$  and  $120^\circ$ , all with range  $5a$  relative to the centre of the nominal array. The source locations are known a priori to the estimator

## 5 Conclusions

The far-field self-calibration and near-field pilot-calibration are carried out using UML estimator whose objective function is optimised by PSO. A new technique, DDL is proposed to enhance the performance of PSO at high SNR by dynamically lowering it. Numerical simulations demonstrate that the UML estimator optimised by PSO with DDL is optimally accurate, robust to large shape errors, and free of the initialisation problem.

Furthermore, the development process of the DDL technique introduces the ‘prominence’ concept of the global optimum. This is different from the precision of it, which improves as the SNR increases. The counter-intuitive behaviour of the UML objective function that its global optimum stands more prominent at lower SNR is harnessed by the DDL technique to enhance the performance of the stochastic global optimisation algorithms. Mathematical analysis indicates that the DDL technique is applicable to a wide range of array processing problems where the UML estimator is employed.

## 6 References

- 1 Rockah, Y., Schultheiss, P.: ‘Array shape calibration using sources in unknown locations – part I: far-field sources’, *IEEE Trans. Acoust. Speech Signal Process.*, 1987, **35**, (3), pp. 286–299
- 2 Rockah, Y., Schultheiss, P.: ‘Array shape calibration using sources in unknown locations – part II: near-field sources and estimator implementation’, *IEEE Trans. Acoust. Speech Signal Process.*, 1987, **35**, (6), pp. 724–735
- 3 Rockah, Y.: ‘Array processing in the presence of uncertainty’. PhD dissertation, Yale University, 1986
- 4 Weiss, A.J., Friedlander, B.: ‘Array shape calibration using sources in unknown locations – a maximum likelihood approach’, *IEEE Trans. Acoust. Speech Signal Process.*, 1989, **37**, (12), pp. 1958–1966
- 5 Weiss, A.J., Friedlander, B.: ‘Array shape calibration using eigenstructure methods’, *Signal Process.*, 1991, **22**, (3), pp. 251–258
- 6 Flanagan, B.P., Bell, K.L.: ‘Array self-calibration with large sensor position errors’, *Signal Process.*, 2001, **81**, (10), pp. 2201–2214

- 7 Chung, P.-J., Wan, S.: ‘Array self-calibration using SAGE algorithm’. Proc. Fifth IEEE Sensor Array and Multichannel Signal Processing, SAM 2008, Darmstadt, Germany, July 2008, pp. 165–169
- 8 Stutzman, W., Thiele, G.: ‘Antenna theory and design’ (Wiley, 1998)
- 9 Grosicki, E., Abed-Meraim, K., Hua, Y.-B.: ‘A weighted linear prediction method for near-field source localization’, *IEEE Trans. Signal Process.*, 2005, **53**, (10), pp. 3651–3660
- 10 He, H., Wang, Y., Saillard, J.: ‘Interpolation-based calibration for near-field source localization’. Proc. 17th European Signal Processing Conf., EUSIPCO 2009, Glasgow, UK, August 2009, pp. 1131–1135
- 11 Van Trees, H.L.: ‘Optimum array processing’ (Wiley-Interscience, New York, 2002)
- 12 Eberhart, R., Kennedy, J.: ‘A new optimizer using particle swarm theory’. Proc. Sixth Int. Symp. Micro Machine and Human Science, 1995. MHS’95, Nagoya, Japan, October 1995, pp. 39–43
- 13 Li, M.-H., Lu, Y.-L.: ‘Maximum likelihood DOA estimation in unknown colored noise fields’, *IEEE Trans. Aerosp. Electron. Syst.*, 2008, **44**, (3), pp. 1079–1090
- 14 del Valle, Y., Venayagamoorthy, G.K., Mohagheghi, S., Hernandez, J.-C., Harley, R.G.: ‘Particle swarm optimization: basic concepts, variants and applications in power systems’, *IEEE Trans. Evol. Comput.*, 2008, **12**, (2), pp. 171–195
- 15 Eberhart, R., Shi, Y.: ‘Comparing inertia weights and constriction factors in particle swarm optimization’. Proc. 2000 Congress on Evolutionary Computation, CEC’00, La Jolla, CA, USA, July 2000, vol. 1, pp. 84–88
- 16 Kennedy, J., Mendes, R.: ‘Population structure and particle swarm performance’. Proc. 2002 Congress on Evolutionary Computation, CEC’02, Honolulu, HI, USA, May 2002, vol. 2, pp. 1671–1676
- 17 Li, J., Stoica, P., Wang, Z.: ‘On robust capon beamforming and diagonal loading’, *IEEE Trans. Signal Process.*, 2003, **51**, (7), pp. 1702–1715
- 18 Wan, S., Chung, P.-J., Mulgrew, B.: ‘Near-field array shape calibration’. Proc. 36th Int. Conf. Acoustics, Speech, and Signal Processing, ICASSP’2011, Prague, Czech Republic, May 2011, pp. 2564–2567
- 19 Parsopoulos, K., Vrahatis, M.: ‘On the computation of all global minimizers through particle swarm optimization’, *IEEE Trans. Evol. Comput.*, 2004, **8**, (3), pp. 211–224
- 20 Lo, J.: ‘Minimization through convexitization in training neural networks’. Proc. 2002 Int. Joint Conf. on Neural Networks, 2002. IJCNN’02, Honolulu, HI, USA, May 2002, vol. 2, pp. 1889–1894
- 21 Storn, R., Price, K.: ‘Minimizing the real functions of the ICEC’96 contest by differential evolution’. Proc. IEEE Int. Conf. on Evolutionary Computation, Nagoya, Japan, May 1996, pp. 842–844
- 22 Wan, S.: ‘Parametric array calibration’. PhD dissertation, University of Edinburgh, 2011
- 23 Levi, M., Messer, H.: ‘Sufficient conditions for array calibration using sources of mixed types’. Proc. Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP’1990, Albuquerque, NM, USA, April 1990, vol. 5, pp. 2943–2946
- 24 Wan, S., Chung, P.-J., Mulgrew, B.: ‘Array shape self-calibration using particle swarm optimization and decaying diagonal loading’. Proc. Sensor Signal Processing for Defence, London, UK, 2010, (in press)
- 25 Magnus, J.R., Neudecker, H.: ‘Matrix differential calculus with applications in statistics and econometrics’ (Wiley, 1999)
- 26 Stoica, P., Nehorai, A.: ‘Performance study of conditional and unconditional direction-of-arrival estimation’, *IEEE Trans. Acoust. Speech Signal Process.*, 1990, **38**, (10), pp. 1783–1795

## 7 Appendix

An analysis of the first-order partial derivative sheds some light on the shape of the UML objective function in (6). Jacobi’s formula [25] (Part Three, Section 8.3) states that

$$(\det X)' = (\det X) \cdot \text{tr}(X^{-1} X')$$
(12)

where  $X$  is an invertible matrix.

Denote

$$X = P_H \hat{R} P_H + \frac{\text{tr}(P_H \hat{R} P_H)}{N - M}$$
(13)

and use (B.3) and (B.15) in [26], it follows that

$$\text{tr}(X^{-1} X') = 2\text{Re} \left( \text{tr} \left\{ \left[ (H^H \hat{R} H)^{-1} - \frac{1}{\hat{v}} (H^H H)^{-1} \right] H^H \hat{R} P_H H' \right\} \right)$$
(14)

where

$$\hat{\nu} = \frac{\text{tr}(\mathbf{P}_H^\perp \hat{\mathbf{R}})}{N - M} \quad (15)$$

Inserting (13) and (14) into (12) we obtain the first partial derivative of  $F$  as

$$F' = 2 \det[\mathbf{P}_H \hat{\mathbf{R}} \mathbf{P}_H + \hat{\nu} \mathbf{P}_H^\perp] \text{Re} \left( \text{tr} \left\{ \left[ (\mathbf{H}^H \hat{\mathbf{R}} \mathbf{H})^{-1} - \frac{1}{\hat{\nu}} (\mathbf{H}^H \mathbf{H})^{-1} \right] \mathbf{H}^H \hat{\mathbf{R}} \mathbf{P}_H^\perp \mathbf{H}' \right\} \right) \quad (16)$$

where  $(\cdot)'$  is the first partial derivative with respect to any entry of  $\boldsymbol{\alpha}$ .

Replace  $\hat{\mathbf{R}}$  with  $\hat{\mathbf{R}} + \nu_l \mathbf{I}$  in (15) and (16), note that  $\text{tr}(\mathbf{P}_H^\perp) = N - M$  and  $\mathbf{H}^H \mathbf{P}_H^\perp = \mathbf{0}$ , then the first derivative after diagonal loading,  $F'_l$ , is given by

$$F'_l = 2 \det[\mathbf{P}_H (\hat{\mathbf{R}} + \nu_l \mathbf{I}) \mathbf{P}_H + (\hat{\nu} + \nu_l) \mathbf{P}_H^\perp] \text{Re} \left( \text{tr} \left\{ \left[ (\mathbf{H}^H (\hat{\mathbf{R}} + \nu_l \mathbf{I}) \mathbf{H})^{-1} - \frac{1}{\hat{\nu} + \nu_l} (\mathbf{H}^H \mathbf{H})^{-1} \right] \mathbf{H}^H \hat{\mathbf{R}} \mathbf{P}_H^\perp \mathbf{H}' \right\} \right) \quad (17)$$

Two asymptotic ( $K \rightarrow \infty$ ) properties of DL can be derived from (16) and (17):

1. An extremum of  $F$  remains at  $\boldsymbol{\alpha}_0$  after DL.

*Proof:* At the true parameters  $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$

$$\mathbf{R} = \mathbf{H}_0 \mathbf{S} \mathbf{H}_0^H + \nu \mathbf{I}$$

where  $\mathbf{H}_0$  stands for  $\mathbf{H}(\boldsymbol{\alpha}_0)$  and  $\mathbf{S}$  the signal covariance matrix  $E(\mathbf{s}\mathbf{s}^H)$ . The factor  $\mathbf{H}_0^H \mathbf{R} \mathbf{P}_{H_0}^\perp$  in (16) and (17) satisfies

$$\mathbf{H}_0^H \mathbf{R} \mathbf{P}_{H_0}^\perp = \mathbf{H}_0^H (\mathbf{H}_0 \mathbf{S} \mathbf{H}_0^H + \nu \mathbf{I}) \mathbf{P}_{H_0}^\perp = \mathbf{0}$$

Therefore  $F'_l|_{\boldsymbol{\alpha}_0} = F'|_{\boldsymbol{\alpha}_0} = 0$ , an extremum remains at  $\boldsymbol{\alpha}_0$  after DL.

2.  $|F'_l| > |F'|$  when  $\nu_l$  is large. □

*Proof:* When  $\nu_l$  is large,  $\mathbf{P}_H \mathbf{R} \mathbf{P}_H + \hat{\nu} \mathbf{P}_H^\perp + \nu_l \mathbf{I} \simeq \nu_l \mathbf{I}$ , therefore

$$\det[\mathbf{P}_H \mathbf{R} \mathbf{P}_H + \hat{\nu} \mathbf{P}_H^\perp + \nu_l \mathbf{I}] \simeq \nu_l^N$$

and

$$\begin{aligned} [\mathbf{H}^H (\mathbf{R} + \nu_l \mathbf{I}) \mathbf{H}]^{-1} &= (\nu_l^{-1} - \nu \nu_l^{-2}) (\mathbf{H}^H \mathbf{H})^{-1} \\ &\quad - \nu_l^{-2} (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{H}_0 \mathbf{S} \mathbf{H}_0^H \mathbf{H} (\mathbf{H}^H \mathbf{H})^{-1} \\ &\quad + O(\nu_l^{-3}) \end{aligned}$$

therefore

$$\begin{aligned} (\mathbf{H}^H (\mathbf{R} + \nu_l \mathbf{I}) \mathbf{H})^{-1} - \frac{1}{\hat{\nu} + \nu_l} (\mathbf{H}^H \mathbf{H})^{-1} \\ \simeq \nu_l^{-2} (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{H}_0 \mathbf{S} \mathbf{H}_0^H \mathbf{H} (\mathbf{H}^H \mathbf{H})^{-1} \end{aligned}$$

As a result

$$F'_l = O(\nu_l^{N-2})$$

Because  $F'$  is independent of  $\nu_l$ ,  $|F'_l| > |F'|$  when  $\nu_l$  is large. □

Assuming  $F$  is sufficiently smooth in the vicinity of the global minimum, properties (1) and (2) show that DL makes  $F$  more pointed when  $\nu_l$  is large, which suggests that the global minimum is made more prominent. A complete proof of the prominence of global minimum requires analytical information of the local minima of  $F$ , which appears to be a difficult problem. It is remarked that during the derivation no assumption was made on the structure of the array manifold or type of parameters, thus the DL effect is independent of them; and DL takes effect on the objective function, thus its effect is independent of the specific global optimiser.