Swarm Led Tomographic Reconstruction

Mohammad Majid al-Rifaie*
m.alrifaie@gre.ac.uk
School of Computing & Mathematical Sciences
University of Greenwich
London, United Kingdom

Tim Blackwell
t.blackwell.gold.ac.uk
Department of Computing
Goldsmiths College, University of London
London, United Kingdom

ABSTRACT

Image reconstruction from ray projections is a common technique in medical imaging. In particular, the few-view scenario, in which the number of projections is very limited, is important for cases where the patient is vulnerable to potentially damaging radiation. This paper considers swarm-based reconstruction where individuals, or particles, swarm in image space in an attempt to lower the reconstruction error. We compare several swarm algorithms with standard algebraic reconstruction techniques and filtered back-projection for five standard test phantoms viewed under reduced projections. We find that although swarm algorithms do not produce solutions with lower reconstruction errors, they generally find more accurate reconstructions; that is, swarm techniques furnish reconstructions that are more similar to the original phantom. A function profiling method suggests that the ability of the swarm to optimise these high dimensional problems can be attributed to a broad funnel leading to complex structure close to the optima. This finding is further exploited by optimising the parameters of the best performing swarm technique, and the results are compared against three unconstrained and boxed local search methods. The tomographic reconstruction-optimised swarm technique is shown to be superior to prominent algebraic reconstructions and local search algorithms.

CCS CONCEPTS

• **Computing methodologies** → *Continuous space search*; • **Theory of computation** → *Bio-inspired optimization.*

KEYWORDS

tomographic reconstruction, swarm optimisation, function profiling

ACM Reference Format:
Mohammad Majid al-Rifaie and Tim Blackwell. 2022. Swarm Led Tomographic Reconstruction. In *Genetic and Evolutionary Computation Conference (GECCO '22), July 9–13, 2022, Boston, MA, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3512290.3528737

1 INTRODUCTION

Tomographic Reconstruction (TR) is the inference of the internal structure of an object from the projected images cast by penetrating radiation [23]. TR plays an essential computational role in all medical imaging procedures (X-Ray CT, PET, MRI, Nuclear Medicine and ultrasound [20]). It also has wide applications in industry and science (data compression and data security [27], image processing [42], electron microscopy [11], crystal structure [7], angiography [12], nondestructive testing of homogeneous objects [22], seismic tomography [38], astronomy [9]) and geometric, combinatorial and recreational mathematics [6, 18].

The number of projections acquired in the imaging process is usually insufficient for a unique reconstruction. The problem is under-determined; furthermore, the random nature of the radiation, the operating characteristics of the detectors and, in medical applications, patient movement, mean that projection data is incomplete and noisy. The development of fast and accurate reconstruction algorithms that can identify structures from procedures with reduced patient radiation dose (by weakening the incident radiation and/or reducing acquisition time) is an important concern, enabling procedures that were once prohibited, and accelerating the throughput of patients. Diagnostic investigation would become possible in cases where it might have been withheld. These include imaging of young children, and where a patient would otherwise be required (for the requirements of a particularly clear image) to remain completely still, and to not even breathe for several minutes.

Various numerical reconstruction procedures have been developed. The Filtered Backprojection (FBP) algorithm is capable of fast reconstruction in a single iteration but requires a large number of projections [20]. Algebraic Reconstruction Techniques (ART) [27] are iterative algorithms based on Kaczmarz's method [42] for solving an underdetermined system of equations. ART demonstrably reduces image noise and holds the potential for few-view reconstructions in a medical context but is subject to overfitting and evidence of their worth in large patient populations is lacking [19]. Exact reconstruction is possible if the data is transformed to a sparse and therefore compressible representation. Compressed Sensing (CS) exploits this principle and shows promise for few-view reconstruction, but the technique is dependant on knowledge of the sparse representation, and on replacing the non-convex optimisation problem typically unsolvable by traditional methods, by a solvable convex minimisation [10].

Solution techniques based on iterative statistical methods are also popular. The idea is to maximise the likelihood of parameters of an underlying statistical model (e.g. maximum likelihood expectation maximisation algorithm or MLEM [15]). MLEM converges slowly – an accelerated form known as ordered subset expectation maximization (OSEM) [26] is faster – and images tend to become noisy. Noise is a symptom of overfitting and is a common problem with iterative techniques. Regularisation methods, such as maximum a posteriori (MAP) estimation, are often used to mitigates noise. Deep learning (DL) has revolutionised visual and natural language processing in recent years and there is a possibility that

*Corresponding author

algorithms can be transferred and adapted for TR; however a radical improvement over analytical methods for solving inverse problems in imaging is so far missing [35, 46].

Conventional methods (e.g. FPB, ART, OSEM, CS and DL) have the potential for improved imaging (clearer images and reconstructions from reduced projections) but this potential has yet to be realised.

TR can be reformulated as a (typically) non-convex optimisation task: find a solution that minimises the error between the reconstructed image and the measured projections. Population-based, metaheuristic algorithms are often employed in optimisation and especially in cases where exact methods fail. One such metaheuristic is exemplified by particle swarm optimisation (PSO) [40]. Particles are simple agents that inhabit the space of feasible solutions; they have a memory of the best, as measured by the objective function, previously visited position and are subject to forces that drive them towards their personal best the best attained positions of social neighbours. PSO, by virtue of its simplicity, malleability, and wide applicability, has been the subject of much experimentation since its inception in 1995. It has been applied to 26 different problem categories [40] and around 100 papers are listed in a recent report on the application of computational swarm intelligence to real-world problems [2].

The solution set of the typical reconstruction problem is not unique and, although many reconstructions have small or zero error, only a few might correspond to medically useful images. ART and other least squares techniques will converge on a minimum norm (the sum of the squares of each pixel value) solution. Solutions with many small values will therefore dominate solutions with high values and the reconstructed image may be diffuse and lack structure. Swarm algorithms make no assumptions about the nature of the solutions. They do not require any special condition such as sparsity and are not driven towards a least squares solution.

There has been some investigation of heuristic TR algorithms. Ouaddah [39] combined harmony and local search. Other work (e.g. [13, 28, 29]) report on tabu search, simulated annealing and a memetic reconstruction, respectively. Batenberg [6] considered an evolutionary framework. Several authors report on swarm intelligent TR, covering areas such as: binary reconstruction [36], geophysical reconstruction [43], electrical capacitance and impedance tomography [25, 47] and surface reconstruction from 3D data [16]. Additionally PART, based on the movement of particles, was proposed for binary reconstruction [4].

This paper reports on a comparative study of swarm optimised TR, classic ART and FBP, for five standard benchmarks in the medical domain. We find that although the two population metaheuristics trialled (swarm optimisers and differential evolution) generally find solutions with higher reconstruction error than the best classical technique, they exhibit lower reproduction error i.e. the found solutions are closer to the original image. The observation that metaheuristics are unencumbered by the need to find a least squares solution might play a part in this finding; otherwise, a profiling study suggests that the objective function has a single broad funnel at large scales leading to complex multimodality close to the global optimal. This observation is further investigated by tailoring the swarm optimiser parameters to the TR context. The optimised swarm technique outperforms a number of local search methods and maintains its leading achievements against the best performing algebraic reconstruction approaches.

It is shown the tomographic reconstruction-optimised swarm technique demonstrates a lead in performance against the leading algebraic reconstruction approaches and the local search algorithms.

2 TOMOGRAPHY AND ALGEBRAIC RECONSTRUCTION

The imaging process proceeds by shining radiation on an object and collecting emergent radiation with a bank of detectors. The radiation source is swung around in a series of projection angles and emergent levels are recorded for each projection. The task is to infer interior structures (which absorb more or less radiation) from the projected images. The physical situation is exactly modelled by Radon transforms but transform inversions are infeasible. Instead, an approximate model (the 'forward' model) of the physical measurement must be built in order to formalise the mathematical reconstruction problem.

2.1 Problem statement

Incident beams are typically modelled by parallel rays. Each ray is incident on the centre of each detector or projection bin. The imaging process is approximated by a projection matrix $A \in \mathbb{R}^{m \times n}_{\geq 0}$ where m is the total number of rays collected (equal to the number of rays at each projection angle multiplied by the number of projection angles) and n is the number of pixels in the reconstructed image. If $b \in \mathbb{R}^m$ is a vector of detector values, the continuous/discrete reconstruction problem can be stated as:

$$\text{find } x \begin{cases} \in \mathbb{R}^n \\ \in \{0, 1, \ldots, k-1\}^n, k > 1 \end{cases} \quad \text{such that } Ax = b. \tag{1}$$

The binary problem is $k = 2$ i.e. with $x \in \{0, 1\}^n$.

Since the equation $Ax = b$ is, in general, underdetermined, it cannot be inverted. Instead an approximate solution y must be obtained (e.g. by filtered back projection, or by algebraic reconstruction). This trial solution is forward projected according to the measurement model:

$$Ay = c$$

with an associated reconstruction error

$$e_1(y) = ||b - c||_1$$

An iterative scheme will produce a sequence of candidate solutions, $y^{(k)}, k = 1, 2, \ldots$, of non-increasing error.

A zero projection error might yield a reconstructed solution y that is not identical to the original object x^*. This is due to underdetermination. However, in cases where the reference image is known, the proximity of y to x^* offers a second measure of algorithm performance:

$$e_2 \quad = \quad ||y - x^*||_1 \tag{2}$$

A zero value of e_1 solves the problem $Ay = b$ but does not guarantee reconstruction fidelity. e_2, a reproduction error, provides a check, where a value of zero corresponds to a reconstructed image that is the exact replica (modulo the approximation that constitutes the forward model) of the original.

2.2 Algebraic reconstruction algorithms

The classical back projection [24, 32] technique, although a relatively quick and effective reconstruction procedure, suffers from high frequency 'blurring' which is only partly ameliorated by filters (FBP). However, increasing computation power means that algebraic reconstruction techniques (algebraic-RT or ART) are gaining prominence. This is due to ART's potential for greater accuracy, albeit at increased time of execution.

The first ART was a rediscovery [21] of the Kaczmarz method for solving linear equations [31]; a diagrammatic explanation of the principle is provided in [33]. ART reconstructions suffer from salt and pepper noise, an artefact partially due to successive updates to components of x during an iteration changing the results of previously tuned values. This pathology is mitigated by the more slowly converging simultaneous iterative reconstruction technique (SIRT), where the updates are not applied immediately but are averaged and applied at the end of an entire iteration. In a further development, SART [5] improves upon ART and SIRT with the employment of a more sophisticated forward model. Good quality reconstructions are often obtained in a single iteration [33]. SART remains popular to this day and has been the subject of mathematical analysis (e.g. [30]).

3 SWARM OPTIMISATION

The essence of an optimisation swarm for continuous problems is a population of individuals or particles that move in a real-valued D-dimensional search space $X \subset \mathbb{R}^D$. Particles respond to each others position and value as determined by the objective function $f : X \rightarrow \mathbb{R}$. The swarm moves through X, typically clustering around promising areas before converging on a putative solution to the problem: find $\arg\min f(x)$.

The precise way that a particle alters its position in response to its neighbours, and any internal structure that it might possess, depends on the particular variety of swarm. For example, particles might carry a memory of a previously visited location; they might possess a velocity and they might interact with other particles in a pre-defined network, or with a random selection of neighbours chosen at each iteration.

The following outlines three swarm varieties: canonical particle swarm optimisation (PSO), a PSO-variant known as DFO and differential evolution (DE). Although DE is generally considered to be an evolutionary algorithm, it fits into the broad view of an optimisation swarm as outlined above.

3.1 PSO

Particles i in a canonical PSO swarm [34**? ?**] of N particles have dynamical variables x_i, v_i, representing position and velocity in the search space $X \subset \mathbb{R}^D$, and an internal 'memory' or *pbest* (personal best), p_i, of the best position achieved so far in the run, as measured by the objective function f.

Dynamical variables are updated according to the rule

$$
\begin{aligned}
v_i^{t+1} &= wv_i^t + cu_1 \circ (n_i^{t+1} - x_i^t) \\
&\quad + \ cu_2 \circ (p_i^{t+1} - x_i^t) \\
x_i^{t+1} &= x_i^t + v_i^{t+1}
\end{aligned}
\tag{3}
$$

where $u_{1,2} \sim U(0,1)$ are uniform random variables in $[0,1]^D$ and $\circ$ is the Hadamard (entry-wise) product. n_i is the pbest of the best neighbour in i's social network, the inertial weight, w, and acceleration coefficients c, are two arbitrary (but constrained) positive real parameters chosen to balance convergence and exploration and t labels iteration. In synchronous updating, iteration $t+1$ begins by updating all pbests:

$$p_i^{t+1} = \arg\min^* \left(f(x_i^t), f(p_i^t) \right)$$

where min* returns the first member of the list in the case of non-uniqueness. The iteration is completed by updating all N positions and velocities according to Eq. 3.

Two communication schemes are in common use. Particles in the global-best PSO (GPSO) network have access to all memories: the social network is maximally connected; on the other hand, particles in a local-best PSO (LPSO) network can only access memories in a restricted network. Networks do not include self. In the commonly chosen ring LPSO, particles communicate with 'right' and 'left' neighbours. LPSO, by virtue of a slower information transfer that inhibits convergence and favours early exploration, is generally better at more complex multi-modal problems [8].

3.2 DE

Differential evolution exists in a wide variety of forms; we specify the DE/best/1 version which is considered competitive and robust [14].

Each iteration begins with a determination of the current position, g of the best particle. Then, for each particle i, indices j and k are selected such that $i \neq j \neq k$. A random component $r \in \{1, 2 \ldots D\}$ is also selected.

Then for each component d of x_i:

$$
\begin{aligned}
&\text{if } u \sim U(0,1) < C_R \text{ or } d == r \\
&\quad y_d = g_d + F(x_{jd}^t - x_{kd}^t) \\
&\text{else} \\
&\quad y_d = x_{id}^t
\end{aligned}
\tag{4}
$$

where y is a trial position and the parameters $C_R \in [0,1]$ and $F \in [0,2]$ are known as the 'cross-over rate' and the 'differential weight'.

Then, after each component of y has been set, i is conditionally moved:

$$x^{t+1} = \arg\min^* \left(f(y), f(x^t) \right)$$

3.3 DFO

DFO, 'dispersive flies optimisation' [1], is a slimmed-down PSO variant that abolishes particle memory and velocity in favour of updates based on instantaneous, rather than historical, position; the algorithm's exploration and exploitation behaviour is studied in [3]. In addition, it incorporates component-wise particle jumps [**?**]. The iteration starts by determining the best overall position g^{t+1}, if unique, and positions of all best ring neighbours, n_i^{t+1} of each particle (except for the current swarm best particle, which is not updated). An arbitrary choice of g is made if there is a tie for the best position. Position component d of all particles i, (other than

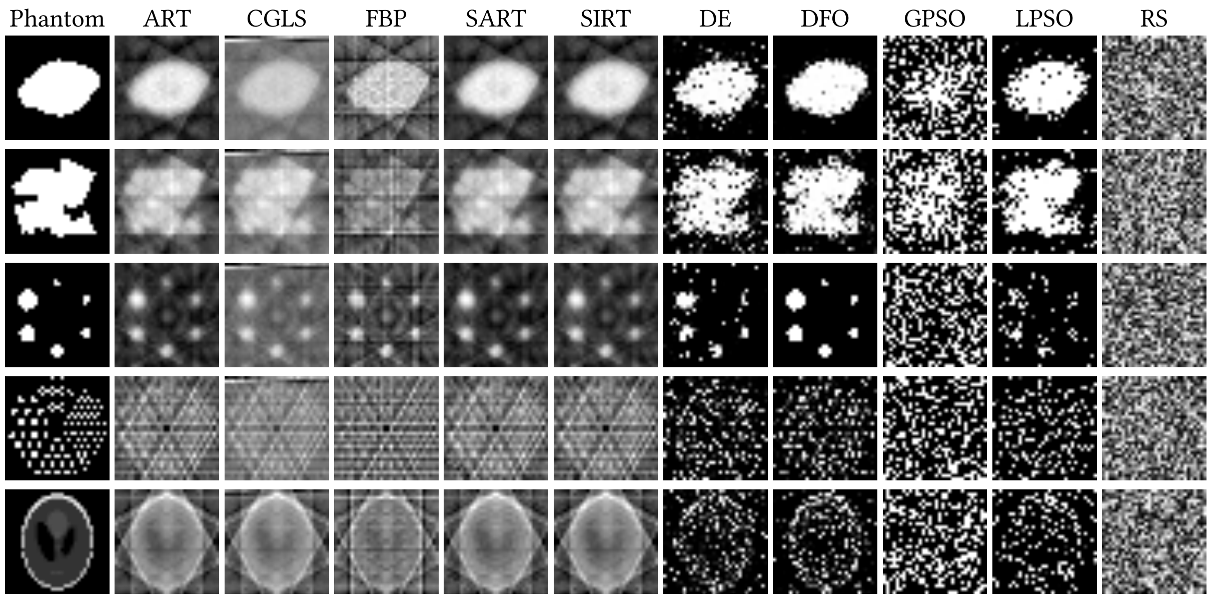


Figure 1: Phantom size: 32×32**, projections:** 6

the swarm best) updates according to

$$\begin{aligned} &\text{if } u \sim U(0,1) < \Delta \\ &\quad x_{id}^{t+1} \sim U(X_d) \\ &\text{else} \\ &\quad x_{id}^{t+1} = n_{id}^{t+1} + \phi u_1 (g_d^{t+1} - x_{id}^t) \end{aligned} \tag{5}$$

where Δ is a preset jump probability and $U(X_d)$ is the uniform distribution along axis d of the search space X, $u_1 \sim U(0,1)$ and $\phi \in [0, \sqrt{3}]$ (the constraint on ϕ is derived from a convergence analysis for stochastic difference equations [**?**]). ϕ is invariably set to 1 in published studies. The algorithm employs global and local strategies and has three arbitrary parameters: N, Δ and ϕ.

In a formal sense, the DFO algorithm, with its reliance on instantaneous position, abandonment of particle memory and retention of a static communication network, stands between PSO and DE. The comparison is only formal however, and does not imply intermediate performance on any particular problem.

4 EXPERIMENTS AND RESULTS

The phantoms are depicted in Fig. 1, where phantoms 1 - 4 are binary reconstruction problems and phantom 5, the Shepp-Logan phantom, is a discrete problem with six pixel value levels [41]. Two sizes, 32×32 and 64×64, were trialled and, in order to test few-view conditions, the number of projections, α, was set to 6, 8, 16 and 32. Phantom imaging was conducted by the ASTRA toolbox [45] using parallel geometry with the number of rays set to 32 and 64 for the the 32×32 and 64×64 phantoms respectively.

Five reference algorithms from the ASTRA toolbox were selected: filtered backprojection (FBP), the algebraic algorithms ART, SIRT and SART, and a gradient descent reconstruction procedure, CGLS. The potential of swarm reconstruction was tested with three swarm algorithms, PSO (in two varieties, GPSO and ring LPSO), DE and DFO; finally, since the swarm algorithms rely on extensive sampling, random search (RS) was also tested as a control.

The swarm algorithms and RS were run for 100,000 function evaluations. ART, CGLS, FBP, SART and SIRT perform the reconstruction in $\mathbb{R}^D$, where $D = 32 \times 32$ or 64×64. Reconstructions were scaled to $X = [0, 255]^D$ for the purpose of computing the reproduction error, e_2.

A swarm size of $N = 100$ was chosen for G/LPSO, DE and DFO. Particles were initialised in X with the uniform distribution and G/LPSO velocities were set to zero. Particles in all three swarms were clamped to the search box: any particle attempting to leave X was placed on the boundary.

The DFO jump probability was set to 0.001; G/LPSO was run with $w = 0.729844$ and $c = 1.49618$ and the DE/best/1 parameters F and C_R were both set to 0.5.

All algorithms with randomisation were run 30 times on each of the 40 problems (5 phantoms, 4 projection types (6, 8, 16, 32) and 2 sizes, 32×32 and 64×64).

4.1 Reconstruction results

Table 1 reports on Wilcoxon statistical significance tests on the reconstruction error for algorithm pairs for the 40 problem instances at a significance level of 0.05. The rows show the number of instances in which the row algorithm performed better than the column algorithm. For example, reading along the first row, ART gave a significantly smaller reconstruction error, e_1, than SART on 9 of the 40 trials.

Algebraic reconstruction algorithms ART, SART and SIRT and gradient descent, CGLS, uniformly outperform DE, GPSO and LPSO and are better than DFO in 32 trials. The result is not surprising;

Table 1: Algorithms comparison based on e_1. The numbers indicate statistically significant wins for the algorithm in the left hand column versus the algorithm in the top row.

e_1	ART	CGLS	FBP	SART	SIRT	DE	DFO	GPSO	LPSO	RS
ART	NA	0	40	9	0	40	32	40	40	40
CGLS	40	NA	40	30	21	40	32	40	40	40
FBP	0	0	NA	0	0	16	9	38	13	40
SART	31	10	40	NA	0	40	32	40	40	40
SIRT	40	19	40	40	NA	40	32	40	40	40
DE	0	0	23	0	0	NA	0	40	4	40
DFO	6	6	30	6	6	40	NA	40	32	40
GPSO	0	0	2	0	0	0	0	NA	0	40
LPSO	0	0	27	0	0	28	8	40	NA	40
RS	0	0	0	0	0	0	0	0	0	NA

Table 2: Algorithms comparison based on e_2. The numbers indicate statistically significant wins for the algorithm in the left hand column versus the algorithm in the top row.

e_2	ART	CGLS	FBP	SART	SIRT	DE	DFO	GPSO	LPSO	RS
ART	NA	31	37	29	28	5	3	38	3	40
CGLS	9	NA	22	15	13	7	3	25	6	37
FBP	3	18	NA	16	4	1	0	32	1	40
SART	11	25	24	NA	18	0	0	29	0	38
SIRT	12	27	36	22	NA	5	1	38	3	40
DE	34	31	39	40	35	NA	0	40	8	40
DFO	37	36	40	40	38	40	NA	40	32	40
GPSO	2	13	4	11	2	0	0	NA	0	40
LPSO	36	34	39	40	36	28	8	40	NA	40
RS	0	2	0	2	0	0	0	0	0	NA

these toolbox algorithms have been specifically developed for reconstruction whereas the swarm algorithms are off-the-shelf multipurpose low-dimensional optimisers that have not been tuned to the reconstruction task.

The performance of swarm algorithms LPSO and DE is comparable to the classic filtered backprojection method; GPSO is notably worse and DFO has a slight advantage.

All algorithms are better than randoms search which shows that the performance of the swarm algorithms is not merely due to repeated sampling.

Table 2 compares algorithm pairs when rated according to the reproduction error, e_2. DE, LPSO and DFO consistently find better reproductions than any of the toolbox algorithms, and of the swarm algorithms, DFO exhibits a better performance. GPSO infrequently finds a lower e_2. DE and LPSO, although producing worse reconstructions when judged by the projection error, e_1, than the toolbox algorithms, nevertheless produce images that are closer to the original. However the visual comparison in Fig. 1 shows that low reproduction error is not always a reliable measure on how a clinician might interpret an image; the toolbox algorithms produce more recognisable, albeit blurred, reconstructions of phantom 4 and the Shepp Logan phantom. DFO produce sharper images of phantoms 1-3. The blurriness seen in the toolbox reconstructions seems to have been traded for noise.

Table 3 presents a more detailed account of the e_2 results. The cells display the median reproduction error in the set of 40 trials, for each phantom; lighter shading indicates lower e_2. DFO produces good reproductions of phantoms 1-3, especially for higher numbers of projections: indeed it achieves perfect reproduction in seven cases. CGLS tends to have a lower median than the other toolbox algorithms, specially with higher number of projections, and SIRT is more consistent than SART. However, no toolbox algorithm is capable of producing an exact reconstruction of the original phantom.

Fig. 2 shows e_1 convergence plots for sample runs. The plots do not show evidence of stagnation, a characteristic morbidity of swarm algorithms, and indicate that extended runs would have reduced the errors still further.

4.2 Function profiles

A surprising feature of the experiments is the ability of the swarm algorithms, as exemplified by the convergence plots of Fig. 2, to make any progress given the high dimensionality and perceived difficulty of the reconstruction problems.

It is worth noting that the swarm algorithms of these trials have been developed for comparably low-dimensional problems (typically in 30 dimensions) whereas the reconstructions here, at 1024 and 4096 dimensions, are considered high in the global optimisation community and are subject to special methods.

In order to test the ability of a general swarm optimiser i.e. one that has not been enhanced for high dimensional problems, trial runs of DFO on a blank phantom (in which all pixel values are zero) and on the unimodal Sphere problem in 1024 dimensions were performed. Fig. 5, a representative convergence plot, shows the swarm does indeed make progress and does not suffer any periods of stagnation. The shape of the plots differs from the phantom convergence plots of Fig. 2 in which the initial fast convergence is followed by a slowing down (the plots are convex rather than concave), hinting that the problem is unimodal at large scales but has a more complex multimodal structure on small scales.

The nature of the task was investigated further by measuring function profiles along 'adaptive' walks. An adaptive walk is a sequence of steps in X such that each step lowers or equals the function value of its predecessor. We define an *adaptive walk profile of granularity* λ as a plot of objective function values taken at $\frac{1}{\lambda}$ points in a straight line between steps in an adaptive walk.

Fig. 3 shows adaptive walks profiles at iterations $t = 200$ to $t = 210$. The adaptive walk is the trace of the swarm best position at each iteration. The 'anchor steps' (actual values attained by the swarm) are marked with black blobs. The continuous black line follows e_1 values at $\frac{1}{\lambda} = 100$ intermediate points. In all five cases we see that the profile never rises above the value at the end-points, indicating that the function, as seen by the swarm, is unimodal at this scale, and particles do no encounter ridges between optima and the swarm is contained within a single funnel.

Fig. 4 shows profiling at later iterations. The objective functions appears to lose their apparent unimodality close to the optima at

Table 3: Reproduction error, e_2, for each problem and each algorithm. Lighter shading indicates the proximity of the reconstructions to the phantoms. The largest errors in phantoms of sizes 32^2 and 64^2 are 255×32^2 and 255×64^2 respectively.

	ART	CGLS	FBP	SART	SIRT	DE	DFO	GPSO	LPSO	RS
Phantom 1, size = 32^2, $\alpha = 6$	51836	103862	82141	52578	52254	24360	8925	90896	17006	123172
$\alpha = 8$	51730	53772	87373	52374	53452	18306	1825	88231	12444	122473
$\alpha = 16$	38732	59710	55565	56237	42305	10800	0	87185	8898	123530
$\alpha = 32$	37764	7104	35628	110013	27998	8646	0	87993	8540	122754
Phantom 1, size = 64^2, $\alpha = 6$	215811	416802	426382	214874	246054	210940	196817	435764	154123	506638
$\alpha = 8$	227777	352616	386348	225147	241197	193099	179152	432352	137348	507789
$\alpha = 16$	163363	473576	263650	193959	179971	172757	161105	430594	117964	507951
$\alpha = 32$	132912	130092	183064	452604	150775	163515	147693	430219	109999	507091
Phantom 2, size = 32^2, $\alpha = 6$	75422	93088	96273	74995	75948	39667	26212	93150	28901	122240
$\alpha = 8$	66259	66446	91778	70501	71291	35209	19449	92225	22997	122429
$\alpha = 16$	53855	85828	67687	61974	52369	17269	0	90415	8816	122056
$\alpha = 32$	36254	16990	45938	117980	30212	11701	0	87261	7921	122628
Phantom 2, size = 64^2, $\alpha = 6$	306104	481477	397097	300757	308809	251333	231673	438343	190386	507088
$\alpha = 8$	271607	413516	408564	275983	280526	237651	220855	436555	170015	506604
$\alpha = 16$	231209	244594	319779	249123	241244	189399	179313	428432	120831	504912
$\alpha = 32$	178477	179468	218892	379754	181168	171271	160561	430698	109442	504911
Phantom 3, size = 32^2, $\alpha = 6$	53974	104471	75166	53029	61049	14213	0	84787	17259	122377
$\alpha = 8$	69150	118941	87030	70117	65759	17187	23	88102	20912	121866
$\alpha = 16$	43576	88594	57915	81110	40385	12000	0	86317	18689	122053
$\alpha = 32$	33432	10015	36126	110555	31715	10536	0	85807	18903	121947
Phantom 3, size = 64^2, $\alpha = 6$	239967	352404	397112	236259	243334	173616	150138	425212	156488	504370
$\alpha = 8$	275341	477491	408459	271367	297473	181518	155322	423937	159573	504085
$\alpha = 16$	208053	472483	315616	199170	219134	178099	144857	421260	160890	505120
$\alpha = 32$	141146	142068	192379	474858	141821	175900	142731	424702	149612	504338
Phantom 4, size = 32^2, $\alpha = 6$	106304	123937	106080	107887	106043	54862	48542	93915	52286	123660
$\alpha = 8$	99110	101757	112841	97892	101792	58459	52578	94987	56592	122876
$\alpha = 16$	78390	96935	93572	119100	90408	48886	32917	95115	44677	123858
$\alpha = 32$	61533	18604	75447	148600	60032	40083	12384	94222	38095	123279
Phantom 4, size = 64^2, $\alpha = 6$	359961	453232	341694	351839	369251	251265	225083	437839	244786	472486
$\alpha = 8$	370563	482330	397720	378974	402474	261236	237621	443354	253149	471607
$\alpha = 16$	327407	340777	352633	335292	378158	252098	219382	442305	244071	473273
$\alpha = 32$	260784	241559	315892	471062	250803	242980	201774	440890	231698	473312
Phantom 5, size = 32^2, $\alpha = 6$	81771	89546	84742	82790	80748	48691	43853	93344	52980	103306
$\alpha = 8$	69001	80591	93809	69563	71061	46981	40419	92871	51388	103155
$\alpha = 16$	45268	122465	66107	80417	46304	42407	31422	93687	45457	103397
$\alpha = 32$	26132	1688	45887	87035	15953	38162	23964	93373	42364	103167
Phantom 5, size = 64^2, $\alpha = 6$	289161	417457	325745	287476	292780	231047	200976	438982	229764	442217
$\alpha = 8$	283867	279289	379831	287028	288081	229070	197188	434856	229850	442038
$\alpha = 16$	225719	225655	292590	249716	224916	226062	186232	435355	221121	441738
$\alpha = 32$	144916	159077	178347	465039	184713	222332	177855	436841	219545	441961

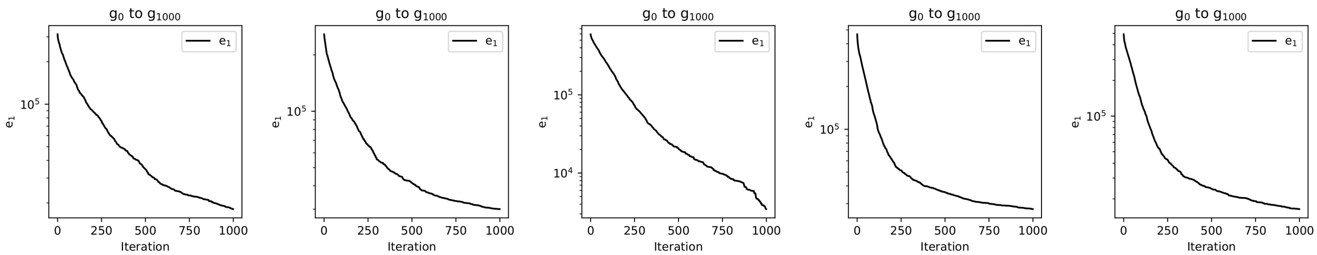
$e_1 = 0$. The profile shows a complex structure of ridges where intermediate points exceed the values at the end-points and the profiles are very jagged. At this stage in the optimisation, the swarm has to negotiate micro-basins that separate sub-optimal minima.

The emergent picture of e_1 as a single broad funnel leading to a highly modal surface close to the global optimum provides some explanation of the relative success of any swarm algorithm to make progress in this high dimensionality, and of superior performance of LPSO over GPSO, given LPSO is known to have better multi-modal characteristics [8]. DFO, which is uniformly better than DE and GPSO (see Table 1) is also more successful than DE on these reconstruction problems. Given the (formal) intermediate nature of DFO between PSO and DE, it seems that the removal of velocity and history, and the retention of a mixed (local and global) communication network is a superior strategy.

4.3 Parameter fine-tuning

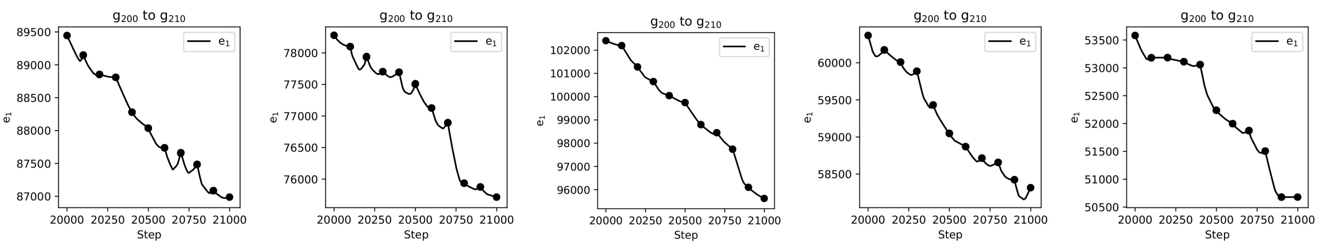
The parameters of the leading swarm method were optimised in order to better gauge its performance in the context of the problems.

Phantom 1, with size 32×32 and $\alpha = 6$, was used as a benchmark for the sweep through parameter space. The three DFO parameters were selected from the sets $N = \{2, 3, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$, $\phi = \{0.173, 0.346, 0.520, 0.693, 0.866, 1.039, 1.212, 1.386, 1.559, 1.732\}$ and $\Delta = \{0.0, 0.00001, 0.00005, 0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1\}$. The search was first applied to find the optimum N and ϕ setting. Each parameter combination was run 10 times. The best values found were $N = \{2, 3\}$ and $\phi = \{1.386, 1.559, \sqrt{3}\}$, with $N = 2$ having a small lead (i.e. performing statistically significantly better in {118,117,117} cases for each ϕ values out of the total of 130 algorithms, as apposed to $N = 3$, with {117,111,113} cases). Subsequently, using $N = 2$, and $\phi = 1.386$ the search for the optimum restart thresholds resulted in an equal performance for $\Delta = \{0.0005, 0.001\}$.



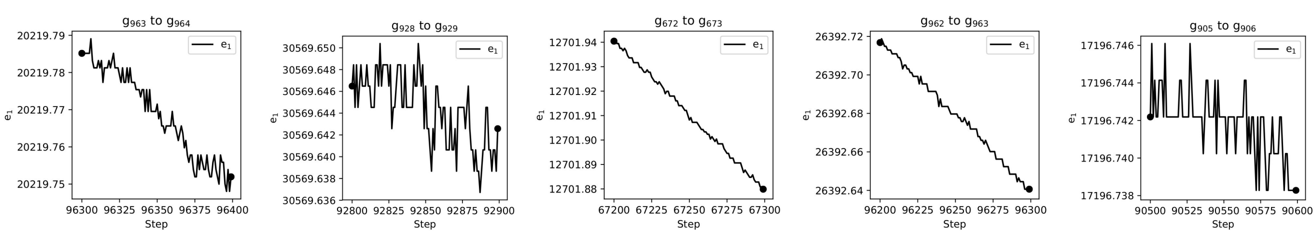
Phantom 1 Phantom 2 Phantom 3 Phantom 4 Phantom 5

Figure 2: Reconstruction error, e_1, during 1,000 iterations in sample runs for the phantoms.



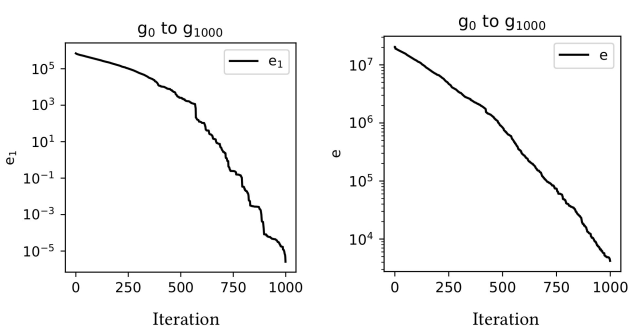
Phantom 1 Phantom 2 Phantom 3 Phantom 4 Phantom 5

Figure 3: Unimodal function profiles



Phantom 1 Phantom 2 Phantom 3 Phantom 4 Phantom 5

Figure 4: The revelation of complex function profiles at later iterations



Blank Phantom Sphere function

Figure 5: DFO error in 1000 **iterations in a sample run for a blank phantom** (32×32) **where all pixels are set to zero, and the Sphere function,** $e = \sum_{d=1}^{D} X_d^2$**, with** X **constrained to** $[0, 255]^D$ **and** $D = 1024$**.**

Grid-based search for G/LPSO with $N \in \{2, 3..., 100\}$ and DE with $N \in \{3, 4..., 100\}$, resulted in $N = 100$ in both cases which indicates that small populations are not favoured by these algorithms.

To have a more precise picture of the optimal parameter values for N, ϕ and Δ, DFO was used as a hyper-heuristic. The experiment was run 30 times for the 3 dimensional problem, with the population size of 10 and the termination criterion set to 100 iterations (1000 function evaluations). In this experiment, the elitism mechanism was relaxed to allow for the re-evaluation of the current best parameter set in each iteration, followed by a best individual update if necessary. The median values of the parameters found in the trials are then extracted. The optimum values found were $N = 2$, $\phi = 1.7320508 \approx \sqrt{3}$ and $\Delta = 0.0011245 \approx 0.001$, in agreement with the sweet spot in the grid search.

While the small value of optimal N might come as a surprise, it confirms the function profiling which demonstrated the presence of a single broad funnel leading to small scale complex structure close to global optima, in effect rendering the task into a largely

Table 4: Algorithms comparison based on e_1 and e_2. The numbers indicate statistically significant wins for the algorithm in the left hand column versus the algorithm in the top row.

e_1	SIRT	DFO-TR	MTS-LS1	L-BFGS-B	SNMS	ANMS
SIRT	NA	31	40	40	40	40
DFO-TR	8	NA	37	40	40	40
MTS-LS1	0	1	NA	40	40	40
L-BFGS-B	0	0	0	NA	0	0
SNMS	0	0	0	21	NA	2
ANMS	0	0	0	23	8	NA

e_2	SIRT	DFO-TR	MTS-LS1	L-BFGS-B	SNMS	ANMS
SIRT	NA	1	1	40	40	40
DFO-TR	39	NA	38	40	40	40
MTS-LS1	39	2	NA	40	40	40
L-BFGS-B	0	0	0	NA	4	4
SNMS	0	0	0	2	NA	3
ANMS	0	0	0	3	6	NA

unimodal problem. It also indicates that although DFO is acting as a 'swarm-inspired local search' and the collective presence of a large communication network is unnecessary in this context.

4.4 Local search and algorithm comparison

To further explore this finding, the tomographic reconstruction-optimised DFO (DFO-TR), with $N = 2$, $\phi = \sqrt{3}$ and $\Delta = 0.001$, is compared against three local search algorithms (Nelder-Mead [37], L-BFGS-B [48], MTS-LS1 [44]) and the best performing toolbox algorithm, SIRT. Nelder-Mead is unconstrained while L-BFGS-B and MTS-LS1 are boxed in the feasible space. MTS-LS1 is especially designed for large scale global optimisation, and the classic L-BFGS-B uses an approximation of the gradient to improve the search. MTS-LS1 is appropriated for separable problems, but is sensitive to rotations. On the other hand, L-BFGS-B is less powerful but is less sensitive to rotations. Nelder-Mead is trialled in two flavours, the Standard Nelder-Mead Simplex (SNMS) [37] and the Adaptive Nelder-Mead Simplex (ANMS) [17]. ANMS has been proposed to deal with the inefficiency of SNMS in high dimensions.

All algorithms were run 30 times on each of the 40 problems, with each trial running for 100,000 function evaluations.

Table 4 compares these algorithms based on the reconstruction (e_1), and reproduction (e_2) errors. L-BFGS-B and Nelder-Mead algorithms show no promise in the context of these high dimensional problems. SIRT is the best algorithm in terms of the reconstruction error as shown in Table 4-top, followed by DFO-TR, and then MTS-LS1. However, in terms of the reproduction error (see Table 4-bottom), DFO-TR is the leading method, followed by MST-LS1 and then SIRT. The low reproduction error of DFO-TR, in comparison to the other algorithms (toolbox and local search techniques), places this algorithm in a promising spot for dealing with reconstruction problems. More work is required to establish if there are factors other than clamping, restart and local search which contribute to the performance of the investigated methods.

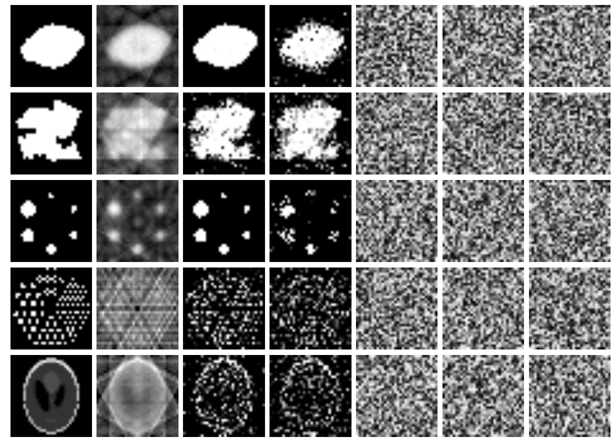


Figure 6: Phantom size: 32×32**, projections:** 6

5 CONCLUSION

This paper reports on swarm optimised tomographic reconstruction of five standard phantoms in the few-view regime. Three representative swarm algorithms were tested and compared with algebraic reconstruction algorithms and filtered backprojection, along with two unconstrained and two boxed local search algorithms. We find that DFO – an algorithm that formally interpolates between the widely studied and applied differential evolution and particle swarm optimisation techniques – is the most competitive swarm method. The standard toolbox techniques provide lower reconstruction error and blurred final images, whereas swarm algorithms produce lower reproduction error (sometimes even producing exact reconstructions) and sharper images peppered with noise.

An adaptive walk function profiling technique suggests that these reconstruction problems consist of a single broad funnel leading to small scale complex structure close to global optima. This large-scale unimodality would enable swarms to make rapid progress towards the optima. Convergence plots confirm that the swarms do not suffer from stagnation.

Following the profiling and evident large-scale unimodality, the TR-optimised DFO or DFO-TR is compared against a number of local search methods. The results demonstrate the better performance of this small-swarm algorithm against both the local search methods and the best performing toolbox algorithm. More detailed studies are required to establish the reasons behind the outcome which could lie in the use of clamping, the restart mechanism, and the overwhelming occupancy of the optima on the edges.

Furthermore, a hypbridisation of classical toolbox techniques (with their low reconstruction error) and swarm optimisation (returning low reproduction error) could produce a powerful algorithm capable of fast reconstructions in the few view regime.

REFERENCES

[1] Mohammad Majid Al-Rifaie. 2014. Dispersive Flies Optimisation. In *Proceedings of the 2014 Federated Conference on Computer Science and Information Systems (Annals of Computer Science and Information Systems, Vol. 2)*, M. Paprzycki M. Ganzha, L. Maciaszek (Ed.). IEEE, pages 529–538. https://doi.org/10.15439/2014F142

[2] Mohammad Majid al-Rifaie. 2018. Swarm Intelligence and Evolutionary Techniques for Real World Applications. https://gala.gre.ac.uk/id/eprint/21045/. [Online; accessed 13-Aug-2019].

[3] Mohammad Majid al-Rifaie. 2021. Investigating Knowledge-Based Exploration-Exploitation Balance in a Minimalist Swarm Optimiser. In *IEEE Congress on Evolutionary Computation, 2021. CEC'21*. IEEE.

[4] Mohammad Majid al-Rifaie and Tim Blackwell. 2016. Binary tomography reconstruction by particle aggregation. In *European Conference on the Applications of Evolutionary Computation*. Springer, 754–769.

[5] A Andersen and Kac A. 1984. Simultaneous Algebraic Reconstruction Technique (SART): a superior implementation of the ART algorithm. *Ultrason. Img.* 6 (1984), 81 – 94.

[6] Kees Joost Batenburg and Walter A Kosters. 2009. Solving Nonograms by combining relaxations. *Pattern Recognition* 42, 8 (2009), 1672–1683.

[7] Kees Joost Batenburg and Willem Jan Palenstijn. 2004. On the reconstruction of crystals through discrete tomography. In *International Workshop on Combinatorial Image Analysis*. Springer, 23–37.

[8] T. Blackwell and J. Kennedy. 2019. Impact of Communication Topology in Particle Swarm Optimization. *IEEE Transactions on Evolutionary Computation* 23, 4 (2019), 689–702.

[9] MD Butala, RJ Hewett, RA Frazin, and F Kamalabadi. 2010. Dynamic three-dimensional tomography of the solar corona. *Solar Physics* 262, 2 (2010), 495–509.

[10] Emmanuel J Candes, Justin K Romberg, and Terence Tao. 2006. Stable signal recovery from incomplete and inaccurate measurements. *Communications on pure and applied mathematics* 59, 8 (2006), 1207–1223.

[11] José-Maria Carazo, CO Sorzano, Eicke Rietzel, R Schröder, and Roberto Marabini. 1999. Discrete tomography in electron microscopy. In *Discrete Tomography*. Springer, 405–416.

[12] Bruno Carvalho, Gabor Herman, Samuel Matej, Claudia Salzberg, and Eilat Vardi. 1999. Binary tomography for triplane cardiography. In *Information Processing in Medical Imaging*. Springer, 29–41.

[13] Marco Cipolla, Giosuè Lo Bosco, Filippo Millonzi, and Cesare Valenti. 2014. An island strategy for memetic discrete tomography reconstruction. *Information Sciences* 257 (2014), 357–368.

[14] Swagatam Das and Ponnuthurai Nagaratnam Suganthan. 2011. Differential Evolution: A Survey of the State-of-the-Art. *IEEE Transactions on Evolutionary Computation* 15, 1 (2011), 4–31. https://doi.org/10.1109/TEVC.2010.2059031

[15] Arthur P Dempster, Nan M Laird, and Donald B Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* 39, 1 (1977), 1–22.

[16] Akemi Gálvez and Andrés Iglesias. 2012. Particle swarm optimization for non-uniform rational B-spline surface reconstruction from clouds of 3D data points. *Information Sciences* 192 (2012), 174–192.

[17] Fuchang Gao and Lixing Han. 2012. Implementing the Nelder-Mead simplex algorithm with adaptive parameters. *Computational Optimization and Applications* 51, 1 (2012), 259–277.

[18] Richard J Gardner. 1995. *Geometric tomography*. Vol. 1. Cambridge University Press Cambridge.

[19] Lucas L Geyer, U Joseph Schoepf, Felix G Meinel, John W Nance Jr, Gorka Bastarrika, Jonathon A Leipsic, Narinder S Paul, Marco Rengo, Andrea Laghi, and Carlo N De Cecco. 2015. State of the art: iterative CT reconstruction techniques. *Radiology* 276, 2 (2015), 339–357.

[20] A. Giussani and C. Hoeschen. 2013. *Imaging in nuclear medicine*. Springer.

[21] Richard Gordon, Robert Bender, and Gabor T. Herman. 1970. Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography. *Journal of Theoretical Biology* 29, 3 (1970), 471 – 481. https://doi.org/10.1016/0022-5193(70)90109-8

[22] U Hampel, A Bieberle, D Hoppe, J Kronenberg, E Schleicher, T Sühnel, F Zimmermann, and C Zippe. 2007. High resolution gamma ray tomography scanner for flow measurement and non-destructive testing applications. *Review of scientific instruments* 78, 10 (2007), 103704.

[23] G. T. Herman. 2009. *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*. Springer.

[24] Gabor T Herman. 2009. *Fundamentals of computerized tomography: image reconstruction from projections*. Springer Science & Business Media.

[25] Gang Hu, Min-you Chen, Wei He, and Jin-qian Zhai. 2011. Clustering-based particle swarm optimization for electrical impedance imaging. *Advances in Swarm Intelligence* (2011), 165–171.

[26] H Malcolm Hudson and Richard S Larkin. 1994. Accelerated image reconstruction using ordered subsets of projection data. *IEEE transactions on medical imaging* 13, 4 (1994), 601–609.

[27] R.W. Irving and M.R. Jerrum. 1994. Three-dimensional data security problems. *SIAM J. Comput.* 23 (1994), 170–184.

[28] Fethi Jarray and Ghassen Tlig. 2010. A simulated annealing for reconstructing hv-convex binary matrices. *Electronic Notes in Discrete Mathematics* 36 (2010), 447–454.

[29] Fethi Jarray, G Tlig, and A Dakhli. 2010. Reconstructing hv-convex images by tabu research approach. In *International Conference on Metaheuristics and Nature Inspired Computing*. 3.

[30] Min Jiang and Ge Wang. 2001. Convergence of the simultaneous algebraic reconstruction technique (SART). In *Signals, Systems and Computers, 2001. Conference Record of the Thirty-Fifth Asilomar Conference on*, Vol. 1. 360–364 vol.1. https://doi.org/10.1109/ACSSC.2001.986951

[31] Stefan Kaczmarz. 1937. Angenaherte Auflosung von Systemen linearer Gleichungen. *Bull. Internat. Acad. Polon.Sci. Lettres A* (1937), 335 – 357.

[32] Avinash C.. Kak and Malcolm Slaney. 2001. *Principles of computerized tomographic imaging*. Society for Industrial and Applied Mathematics.

[33] Avinash C Kak and Malcolm Slaney. 2001. *Principles of computerized tomographic imaging*. SIAM.

[34] J. Kennedy. 1999. Small worlds and mega-minds: effects of neighborhood topology on particle swarm performance. In *In: Proceedings of the 1999, Congress of Evolutionary Computation*, Vol. 3. IEEE Press, 1931–1938.

[35] A. Lucas, M. Iliadis, R. Molina, and A. K. Katsaggelos. 2018. Using Deep Neural Networks for Inverse Problems in Imaging: Beyond Analytical Methods. *IEEE Signal Processing Magazine* 35, 1 (2018), 20–36.

[36] Póth Miklós. 2014. Particle swarm optimization approach to discrete tomography reconstruction problems of binary matrices. In *Intelligent Systems and Informatics (SISY), 2014 IEEE 12th International Symposium on*. IEEE, 321–324.

[37] John A Nelder and Roger Mead. 1965. A simplex method for function minimization. *The computer journal* 7, 4 (1965), 308–313.

[38] Guust Nolet et al. 2008. A breviary of seismic tomography. *Imaging the Interior* (2008).

[39] Ahlem Ouaddah and Dalila Boughaci. 2014. Improving reconstructed images using hybridization between local search and harmony search meta-heuristics. In *Proceedings of the Companion Publication of the 2014 Annual Conference on Genetic and Evolutionary Computation*. ACM, 1475–1476.

[40] Riccardo Poli, James Kennedy, and Tim Blackwell. 2007. Particle swarm optimization. *Swarm intelligence* 1, 1 (2007), 33–57.

[41] Lawrence A Shepp and Benjamin F Logan. 1974. The Fourier reconstruction of a head section. *IEEE Transactions on nuclear science* 21, 3 (1974), 21–43.

[42] Abe R. Shliferstein and YT Chien. 1977. Some properties of image-processing operations on projection sets obtained from digital pictures. *IEEE Trans. Comput.* 10 (1977), 958–970.

[43] Jens Tronicke, Hendrik Paasche, and Urs Böniger. 2012. Crosshole traveltime tomography using particle swarm optimization: A near-surface field example. *Geophysics* 77, 1 (2012), R19–R32.

[44] Lin-Yu Tseng and Chun Chen. 2008. Multiple trajectory search for large scale global optimization. In *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*. IEEE, 3052–3059.

[45] Wim Van Aarle, Willem Jan Palenstijn, Jeroen Cant, Eline Janssens, Folkert Bleichrodt, Andrei Dabravolski, Jan De Beenhouwer, K Joost Batenburg, and Jan Sijbers. 2016. Fast and flexible X-ray tomography using the ASTRA toolbox. *Optics express* 24, 22 (2016), 25129–25147.

[46] Ge Wang, Jong Chul Ye, and De Man Bruno. 2020. Deep learning for tomographic image reconstruction. *Nature Machine Learning* 2 (2020), 737–748.

[47] P Wang, JS Lin, and M Wang. 2015. An image reconstruction algorithm for electrical capacitance tomography based on simulated annealing particle swarm optimization. *Journal of applied research and technology* 13, 2 (2015), 197–204.

[48] Ciyou Zhu, Richard H Byrd, Peihuang Lu, and Jorge Nocedal. 1997. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on mathematical software (TOMS)* 23, 4 (1997), 550–560.