

RESEARCH

Open Access

Large-scale global optimization through consensus of opinions over complex networks

Omid Askari-Sichani and Mahdi Jalili*

* Correspondence: mjalili@sharif.edu
Department of Computer
Engineering, Sharif University of
Technology, Tehran, Iran

Abstract

Purpose: Large-scale optimization tasks have many applications in science and engineering. There are many algorithms to perform such optimization tasks. In this manuscript, we aim at using consensus in multi-agent systems as a tool for solving large-scale optimization tasks.

Method: The model is based on consensus of opinions among agents interacting over a complex networked structure. For each optimization task, a number of agents are considered, each with an opinion value. These agents interact over a networked structure and update their opinions based on their best-matching neighbor in the network. A neighbor with the best value of the objective function (of the optimization task) is referred to as the best-matching neighbor for an agent. We use structures such as pure random, small-world and scale-free networks as interaction graph. The optimization algorithm is applied on a number of benchmark problems and its performance is compared with a number of classic methods including genetic algorithms, differential evolution and particle swarm optimization.

Results: We show that the agents could solve various large-scale optimization tasks through collaborating with each other and getting into consensus in their opinions. Furthermore, we find pure random topology better than small-world and scale-free topologies in that it leads to faster convergence to the optimal solution. Our experiments show that the proposed consensus-based optimization method outperforms the classic optimization algorithms.

Conclusion: Consensus in multi-agents systems can be efficiently used for large-scale optimization problems. Connectivity structure of the consensus network is effective in the convergence to the optimum solution where random structures show better performance as compared to heterogeneous networks.

Keywords: Large-scale optimization, Complex networks, Continuous opinion formation, Consensus, Scale-free networks, Small-world networks

AMS subject classification: 15A04, 54A20, 60J20, 92D25

Background

Networks are everywhere and we confront many networks in our daily life; they are practically present where any kind of information is transmitted or exchanged. Networks such as the Internet, the World Wide Web, engineering, social, biological and economical networks have been subject to heavy studies in the last decade and many applications have been developed based on network science Albert & Barabasi (2002, 1999; Barabasi & Albert 1999; Boccaletti et al. 2006; Newman & Watts 2006; Newman

2003; Newman & Park 2003; Strogatz 2001; Watts & Strogatz 1998). The progress in network science accelerated after the seminal work of Watts & Strogatz (1998) on collective behavior of small-world networks Watts & Strogatz (1998) and Barabasi & Albert (1999) on scale-free graphs (Barabasi & Albert 1999). Watts and Strogatz discovered that many real-world networks have small-world property in that their characteristics path length scales logarithmically with network size (Watts & Strogatz 1998) – a property that is observed in random networks. At the same time, these networks show high levels of transitivity (clustering coefficient) (Watts & Strogatz 1998) – much higher than corresponding random networks. Furthermore, many real networks from different disciplines were shown to have a power-law degree distribution (Barabasi & Albert 1999); the probability of having a node with degree k is $k^{-\gamma}$ with γ being in the range 2–3. Real networks have been shown to have more complex properties such as motifs (Milo et al. 2002) and community structure (Girvan & Newman 2002). These structural features influence dynamics and functionality of networks. For example, synchronization and consensus properties of networks largely depend on their structure (Belykh et al. 2005; Lu et al. 2004).

The most striking pattern of networked structures appear when a number of agents (each with simple behavior) interact leading to complex behaviors as a result of collective motion. Synchronization of interacting agents - as the most striking form of collective behavior - has many applications in science and engineering. For example, techniques available in network theory can be used for efficient distributed inference in sensor networks (2007; Scutari et al. 2008). In this work we used tools available in network science to perform a numerical optimization task. Optimization is an approach that iteratively improves the performance of a system, which is formulated as a single standard measurement equation called cost (or objective) function.

In order to use network theory for solving an optimization task, we used the concept of consensus formation in the opinions of multi-agent systems. Let us consider a network of agents where a (discrete or continuous) opinion value is associated to each agent. Agents can influence each others' opinions through the connections existing between them, i.e., the edges of the network. Considering some simple update rules and if certain conditions are met, the agents can reach a consensus in their opinions through a number of opinion updates (Kozma & Barrat 2008; Carletti et al. 2006). In this work we considered the evolution of continuous opinions based on the modified version of bounded confidence model (Deffuant et al. 2001), which has been extensively studied in recent years (Gandica et al. 2010; Weisbuch 2004; Urbig et al. 2008). The previous studies of continuous opinion formation have been mainly performed considering uniform agents (Gandica et al. 2010; Weisbuch 2004; Urbig et al. 2008). However, in reality the agents are diverse in their wealth and social status, and hence, have diverse influence on others (Holyst et al. 2001; Lewenstein et al. 1992; Jalili 2013a; Jalili 2013b). Therefore, we associated specific weight for each agent resulting in faster consensus.

The paradigm proposed in this manuscript was applied on a number of benchmark problems. We first considered a simple function with many local optima and showed that the proposed optimization strategy could successfully find the optimum. We then applied the method on a number of benchmark problems from CEC 2010 competition benchmark set (Tang et al. 2009). We compared the performance of the proposed consensus-based optimization approach with that of a number of classic optimization methods including genetic algorithms, differential evolution and particle swarm optimization.

Methods

Optimization through consensus in the network

The paradigm we have proposed for large-scale optimization task is based on consensus in networked structures. In opinion formation models, there is a population of agents, each with a (discrete or continuous) opinion value representing its information about a subject (Deffuant et al. 2001; Gandica et al. 2010; Weisbuch 2004; Urbig et al. 2008). The term opinion is not easy to define in reality; however, it can be considered as a discrete or continuous value expressing the individuals' degree of desire or preference. This opinion is often represented as a real number when the model is unimodal or as a vector of real numbers when the model is multimodal. In this paper we aimed at optimizing an objective function, and therefore, each agent will have an opinion value containing all the input parameters of desired objective function, i.e., a multimodal model.

Opinion formation in multi-agent systems

The agents update their opinions as a result of interactions with their neighboring agents. Consider two neighboring agents i and j with opinions as x_i and x_j , respectively. Their opinions at time $n + 1$ will be a function of their previous opinions, i.e. $x_i(n + 1) = f_1(x_i(n), x_j(n))$, $x_j(n + 1) = f_2(x_i(n), x_j(n))$. If certain conditions are met, after a number of updates in these values, the agents can reach a consensus in their opinions (Kozma & Barrat 2008; Carletti et al. 2006). The collective behavior of the agents over complex networks largely depends on the structural properties of the networks (Amblard & Deffuant 2004), and minor modification in the structure of the network can have drastic effects on the behavior of opinion formation (Nardini et al. 2008).

There are a number of rules for modelling opinion formation in complex networks. For example, considering discrete opinions, in the voter model, randomly selected agents exchange their opinions by that of one of their neighbours (Krapivsky & Redner 2003). The agents might influence their neighbouring agents to change their opinions based on their strength and the neighbours' threshold (Leskovec et al. 2006). In the evolution of continuous opinions on a network, the opinions of two connected agents are updated if their difference is less than a threshold, i.e. the agents have evolving opinions (Deffuant et al. 2001; Amblard & Deffuant 2004; Lorenz 2007; Kurmyshev et al. 2011; Hegselmann & Krause 2002; Guo & Cai 2009).

In this work we considered a specific form of continuous bounded confidence model in which each agent has an opinion in the range $[-1,1]$ - denoted by opinion space - and update its opinions based on a specific rule (Deffuant et al. 2001; Fortunato et al. 2005). First, each agent takes a random value from opinion space. Then, at each proceeding step, each agent finds its best-matching neighbour, i.e., the one that optimizes the objective function furthest among the neighbours, and then updates its opinion value with this best-matching adjacent. The update rule for agent i is as

$$\begin{cases} x_i(n+1) = x_i(n) + \mu [x_j(n) - x_i(n)] \\ \text{if } f(x_i(n)) > f(x_j(n)) \end{cases}, \quad i = 1, 2, \dots, N, \quad (1)$$

$$j = \arg \min_{k \in N_i} f(x_k(n)), \quad k \in N_i$$

where f is the desired cost (or objective) function to be optimized, N is the network size and N_i is the set of neighbours of agent i . μ is the convergence (or influence) parameter, which often takes a value between 0 and 1. This parameter controls the speed of convergence in such a way that small values of μ corresponds to slow but smooth convergence, while the large values of μ corresponds to faster but wavy convergence.

To some extent, the above model for opinion formation imitates the behavior of agents in real social networks. A person may know many individuals in the society; however, he/she is only influenced by his/her closest friends (i.e., neighbors in the network). In many cases, individuals get the maximum influence through their best (closest) friends and try to make themselves similar to them, i.e., making their opinion closer to their closest friends. People try to behave like their best friends for establishing and maintaining their friendships and they influenced by them more than the others in their life. Sometimes these changes will happen because people want to preserve their connections and friendships and they will act or behave like their close friends (Barry & Wentzel 2006). They project their own attitudes and habits to their friends. Furthermore, research showed that, in general, the influence of the very best friend approximately is equal or comparable to the influence of multiple friends (Berndt & Murphy 2003).

In our model, each agent finds one of its neighboring agents that have the best value in the objective function – which is denoted by best-matching neighbor. For example, if the objective function is an energy function; the neighbor with minimal energy function is selected. The agents then update their opinions using equation (1). It has been shown that considering proper connection weights can enhance the consensus properties of the network, i.e., the network reaches to consensus in a shorter time (Jalili 2013a; Jalili 2013b; Yang et al. 2009; Brunetti et al. 2012). Therefore, we also took proper weights while updating the opinions. The update equations read

$$\begin{cases} x_i(n+1) = x_i(n) + \mu \frac{f(x_j(n)) + \varepsilon}{f(x_i(n)) + \varepsilon} [x_j(n) - x_i(n)] \\ \text{if } f(x_i(n)) > f(x_j(n)) \end{cases}, \quad i = 1, 2, \dots, N, \quad (2)$$

$$j = \arg \min_{k \in N_i} f(x_k(n)), \quad k \in N_i$$

where ε is a small value (in order to make the denominator non-zero). The above weighted update rules can be justified as follows. Let us suppose that an objective function is to be minimized. As the best-matching neighbor is found for each agent, it influences the agent according to its fitness, i.e., its value in the objective function. To this end, the weight for the update rule of an agent gets as the fitness function at that agent divided by the fitness value of the best-matching neighbor, often resulting in a value in the range 0–1 (note that the opinions are updated only when the fitness of best-matching neighbor is better than that of the agent). It is worth mentioning that in some cases, the opinions are in multi dimensions, i.e., x is vector, in which the best matching agent is obtained separately for each dimension.

The method largely depends on the diffusion of good opinions (i.e., those that are good in terms of the objective function) in the network. Agents with opinion values close to the optimal objective function disseminate their opinions through communicating with their neighbors, i.e., getting into consensus with them. Indeed, influence of

opinions is in two folds. Indeed, closer opinions to the optimal value of the objective function have a better chance to be selected as best-matching neighbors.

The above rule for opinion formation is somehow inspired by communication in human societies. Our friends influence our behavior in daily life; however, we are usually affected only when our friends are better than us. Here, similarly, for each agent, first, the best matching agent is found, and then, its opinion is updated (using equation (2)) if the fitness of the best-matching neighbor is better (i.e., it results in a lower value in the objective function) than that of the agent.

It is worth mentioning that the consensus (or synchronization) properties of dynamical networks largely depend on their structure and some topologies are favored for fast consensus (Belykh et al. 2005; Ajdari Rad et al. 2008). Network topology plays also important role in the evolution of other dynamical phenomena over complex networks, such as evolution of cooperative behavior among interacting agents (Perc & Szolnoki 2010; Perc 2009).

A pseudo-code of the proposed consensus-based optimization algorithm is illustrated in section Pseudo-Code as follows.

Pseudo-Code for the proposed consensus-based optimization method

Function CBO

N : number of agents in the population (network size)

M : number of attributes of opinion vector

Boundaries: the range of the opinions

F : desired objective function which is needed to be optimized (minimized in this case)

Begin

- Initialize $N * M$ matrix X by a random normal distribution for the opinion values in *Boundaries*;
- net = Create a structured network;
- Repeat
- **for** each agent i in population **do**
- **for** each attribute a **do**
- $neighbors_opinion$ = mask other attributes of the opinions x in neighbors of agent i in network net by a dummy value;
- $self_opinion$ = mask other attributes of the opinions x agent i ;
- j = find the best agent in $neighbors_opinion$ resulting in the best value for F ;
- **if** $neighbor_opinion$ of agent j optimizes F better than $self_opinion$ **then**

$$weight = \frac{F(neighbors_opinion_j) + \varepsilon}{F(self_opinion) + \varepsilon};$$

$$x[i, a] = x[i, a] + \mu \cdot weight \cdot (x[j, a] - x[i, a]);$$

$$x[i, a] = \text{mod}(x[i, a], \text{Boundaries});$$

- **end if**
- **end for**
- **end for**
- Until stopping condition(s) has/have been met
- End

In the beginning of the process, the agents are initialized by some random values in a range acceptable by their opinion values. As indicated by Watts and Dodds (2007) “a minority of individuals who influence an exceptional number of their peers” (Watts & Dodds 2007), there is often a minority of agents that have a significant influence on others, which is mainly due to their specific position in the network. The hypothesis of influential agents demonstrated that initiating influential individuals will be explicitly different from initiating non-influential ones in the size and likelihood of a cascade (Watts & Dodds 2007). This means that initial opinions for influential agents probably would bias the result of the consensus. This phenomenon will not happen in proposed method, since CBO is not based on the bounded confidence model. Every agent selects his/her best-matching neighbor regardless of its great social power and degree.

Consensus of opinion values

In this section, we provide a mathematical proof that the update rule expressed in equation (2) leads the opinions to converge. To this end, let us rewrite it as follows (Hegselmann & Krause 2002):

$$x(t+1) = A(t, x(t)).x(t), \quad (3)$$

where $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]$ is the opinion vector at time t and A is a time-dependence state transition matrix that also depends on the opinion vector. We would like to verify that starting from an initial opinion values $x(0)$, whether or not all opinions converges to a single value, that is $\lim_{t \rightarrow \infty} x_i(t) = x^*$ for $i = 1, 2, \dots, N$. Let us define diameter d of the opinions as

$$d(x) = \max_{1 \leq i, j \leq N} (x_i - x_j), \quad (4)$$

Lemma 1 (Krause 2000): Consider a stochastic matrix A (i.e., a nonnegative matrix with row-sums equal to 1 is defined as a stochastic matrix), then, one has

$$\max_{i,j} |x_i(t+1) - x_j(t+1)| \leq d(A) \cdot \max_{i,j} |x_i(t) - x_j(t)|, \quad (5)$$

or equivalently,

$$d(x(t+1)) = d(A(t, x(t)).x(t)) \leq \left(1 - \min_{i,j \leq N} \sum_{k=1}^n \min\{a_{ik}, a_{jk}\}\right) d(x(t)), \quad (6)$$

The above lemma was proved in (Seneta 1981); however, we also give another proof using a simpler method.

Proof: Expression (6) can be written as

$$\begin{aligned}
 d(Ax) &= \max_{1 \leq i, j \leq N} (A_i x - A_j x) = \max_{1 \leq i, j \leq N} \sum_{k=1}^n (a_{ik} x_k - a_{jk} x_k) \\
 &= \max_{1 \leq i, j \leq N} \sum_{k=1}^n (a_{ik} - \min\{a_{ik}, a_{jk}\} + \min\{a_{ik}, a_{jk}\} - a_{jk}) x_k \\
 &= \max_{1 \leq i, j \leq N} \sum_{k=1}^n (a_{ik} - \min\{a_{ik}, a_{jk}\}) x_k - \min_{1 \leq i, j \leq N} \sum_{k=1}^n (a_{jk} - \min\{a_{ik}, a_{jk}\}) x_k \\
 &\leq \left(1 - \min_{1 \leq i, j \leq N} \sum_{k=1}^n \min\{a_{ik}, a_{jk}\} \right) \left(\max_{1 \leq i \leq N} x_i - \min_{1 \leq j \leq N} x_j \right) \\
 &= \left(1 - \min_{1 \leq i, j \leq N} \sum_{k=1}^n \min\{a_{ik}, a_{jk}\} \right) \max_{1 \leq i, j \leq N} (x_i - x_j) \\
 &= \left(1 - \min_{1 \leq i, j \leq N} \sum_{k=1}^n \min\{a_{ik}, a_{jk}\} \right) d(x)
 \end{aligned}$$

Our proposed weighted update rule for opinion formation, as expressed by equation (2), can be rewritten as

$$x_i(t+1) = x_i(t) + \mu w_j (x_j(t) - x_i(t)) = (1 - \mu w_j) x_i(t) + \mu w_j x_j(t) = A \cdot x(t), \quad (7)$$

It is clear that in the above representation, matrix A is a stochastic matrix.

Theorem 1: The product of two stochastic matrixes is a stochastic matrix.

Proof: Let A and B be two stochastic matrices, and $C = A.B$ is their product. Entries of C are multiplications of entries of A and B . Since A and B are stochastic, their elements are non-negative and, thus, the entries of C are also non-negative. The row-sum of C (C_i for $i = 1, 2, \dots, N$) can be obtained as

$$C_i = \sum_j C_{ij} = \sum_{j=1}^N \sum_{k=1}^N A_{ik} B_{kj} = \sum_{j=1}^N A_{ik} \sum_{k=1}^N B_{kj} = \sum_{j=1}^N A_{ik} \cdot 1 = 1, \quad (8)$$

Therefore, C is a matrix with nonnegative entries and row-sums of equal to 1, and thus, it is a stochastic matrix.

Let t_1 and t_2 represent time steps ($t_1 < t_2$) and $B(t_1, t_2) = A(t_1-1)A(t_1-2)A(t_1-3) \dots A(t_2)$, which models the accumulated weights between time t_1 and t_2 (Hegselmann & Krause 2002). It can be simply shown that for any $r \geq 0$, $1 - r \leq e^{-r}$.

Theorem 2 (Convergence Theorem): Considering opinion update rule (2), suppose we have a matrix $B(t_1, t_2) = [b_{ij}(t_1, t_2)]$, which is a stochastic matrix and models accumulated weights where b_{ij} is an element of matrix B the sequences $0 = t_0 < t_1 < t_2 < \dots \leq T$ and $\delta_1, \delta_2, \dots, \delta_i, \dots$ are such that $0 \leq \delta_t \leq 1$ and $\sum_{t=0}^{\infty} \delta_t = \infty$. If $\sum_{k=1}^{\infty} \min\{b_{ik}(t_m, t_{m-1}), b_{jk}(t_m, t_{m-1})\} \geq \delta_m$ for all $m \geq 1$ and $1 \leq i, j \leq N$, then for any initial condition, there exists a consensus, i.e., $\lim_{t \rightarrow \infty} x_i(t) = x^*$ for $i = 1, \dots, N$.

Proof: see the Appendix section.

Optimization tasks

We applied our optimization procedure on a number of benchmark problems and compared its performance with some well-known methods including genetic algorithms (GA), particle swarm optimization (PSO), differential evolution (DE) and distributed dual

averaging (DDA) algorithms. GA has been successfully applied to many optimization problems and was used as a basic paradigm in this work. It starts with a population of some random solutions – denoted by chromosomes. Therefore, the first step is to encode the initial solutions from phenotype to genotype. The objective function is then used for ranking the chromosomes. GA works iteratively and, in each step, uses some operators such as parent selection, recombination or cross over and mutation (Holland 1975). There are a number of parameters that should be tuned in order for a GA to work well. These include crossover probability, mutation probability, population model and parent selection models. Crossover probability P_c indicates the probability of creating a new chromosome from two parents. Mutation probability P_m indicates the portion of the population that undergoes mutation in each iteration of the algorithm.

DE is one of the best-performing evolutionary algorithms frequently used for optimization tasks, which often results in the optimal solution in shorter steps as compared to other optimization algorithms. DE uses the difference of a randomly selected pair of chromosomes – indicating diversity of the population – and adds it to one of the chromosomes in the population. Then, it uses crossover operators such as binomial and exponential crossover to combine the chromosomes (Storn & Price 1997). The parameters of the algorithm are as follows. β is a real value showing the coefficient of the difference between two selected chromosomes and controls the amplification of differential variation. P_r indicates the probability of using the mutant (trial) vector. N_v is an integer number indicating the number of couple chromosomes in calculating the mutant vector.

We also compared our algorithm with PSO that is a well-known optimization algorithm based on swarm intelligent (Kennedy & Eberhart 1995). In this algorithm, there is a population of agents called swarms (or particles) interacting with other agents – like our algorithm. PSO has two components: cognitive and social components. The cognitive component is the experience of each particle while the social component is the experience of the community the agents belong to. PSO has shown high degree of flexibility and acceptable speed in solving many optimization problems. Here we used one of the best extensions of PSO that is PSO with Inertia weights (Eberhart & Shi 2000). This feature plays an important role in balancing the powers of exploration and exploitation and making the algorithm more stable. PSO has a number of control parameters. Let us denote the parameters controlling the cognitive and social power of the algorithm as c_1 and c_2 , respectively.

Distributed dual averaging (DDA) algorithm – inspired by Nesterov's dual averaging algorithm (Xiao 2010; Nesterov 2009) – has been proposed for optimizing convex functions (Duchi & Wainwright 2012). Similar to CBO optimization algorithm, DDA is a network-based optimization method in which each node computes sub-differential of a local function while receiving information from its neighboring nodes. There is also a weight matrix to model the weighting process of the method. In any iteration, each node updates its solution vector by multiplying the stochastic weight matrix by the summation of its neighbors' parameters and the sub-gradient of the objective function. DDA is computationally efficient and the convergence time depends on properties of the objective function and underlying network topology. Expander graphs have been proposed as efficient connection topology for DDA optimization algorithm (Duchi & Wainwright 2012). Alternating direction method of multipliers (ADMM) is another

optimization method which uses properties of dual decomposition and augmented Lagrangian methods simultaneously (Boyd & Vandenberghe 2004). The Lagrange dual function is obtained by convex conjugate definition and the dual problem is solved using gradient ascent.

Benchmark problems

We evaluated the performance of the proposed optimization strategy on a number of benchmark problems. As the first problem, we considered the following cost function

$$F_1(x) = -e^{-2(\ln(2))\left(\frac{x-0.1}{0.8}\right)^2} \cdot \sin^6(5\pi x), \quad (9)$$

which is a function with many local optima. The optimal point for which the minimum 1 is achieved for this function is at $x^* = 0.1$.

We used a number of multiple competitive functions which have been introduced as benchmarks in optimization problems (Tang et al. 2009). The first function of this type is Shifted Rastrigin's function that is defined as

$$F_2(x) = F_{\text{rastrigin}}(x) = \sum_{i=1}^D [x_i^2 - 10 \cos(2\pi x_i) + 10], \quad (10)$$

which is a multimodal, shifted, separable and scalable function. The other function of this type considered here is Shifted Ackley's function, which is defined as

$$\begin{aligned} F_3(x) &= F_{\text{ackley}}(x) \\ &= -20 \exp \left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2} \right) - \exp \left(\frac{1}{D} \sum_{i=1}^D \cos(2\pi x_i) \right) + 20 + e. \end{aligned} \quad (11)$$

We also considered Shifted Schwefel's function, defined as

$$F_4(x) = F_{\text{schwefel}}(x) = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2, \quad (12)$$

and Shifted Elliptic function, defined as

$$F_5(x) = F_{\text{elliptic}}(x) = \sum_{i=1}^D (10^6)^{\frac{i-1}{D-1}} x_i^2. \quad (13)$$

In all above functions except F_1 , $x \in [-5, 5]^D$. Furthermore, the global optimum - which is $F_2^* = F_3^* = F_4^* = F_5^* = 0$ - is achieved at x^* which is a random and different vector of real numbers in each run.

Network structures

One of the key ingredients of the proposed optimization algorithm is the graph structure used for connecting the agents, which is kept unchanged during the optimization process. In other words, the set of neighbours are not changed for the agents. In this work, we used a number of well-known graph structures including, Erdős-Rényi random, Watts-Strogatz small-world and Barabasi-Albert scale-free networks.

We used the model introduced by Erdős and Rényi for construction of pure random networks (Erdős & Rényi 1960). In this model, N nodes are considered and each pair is

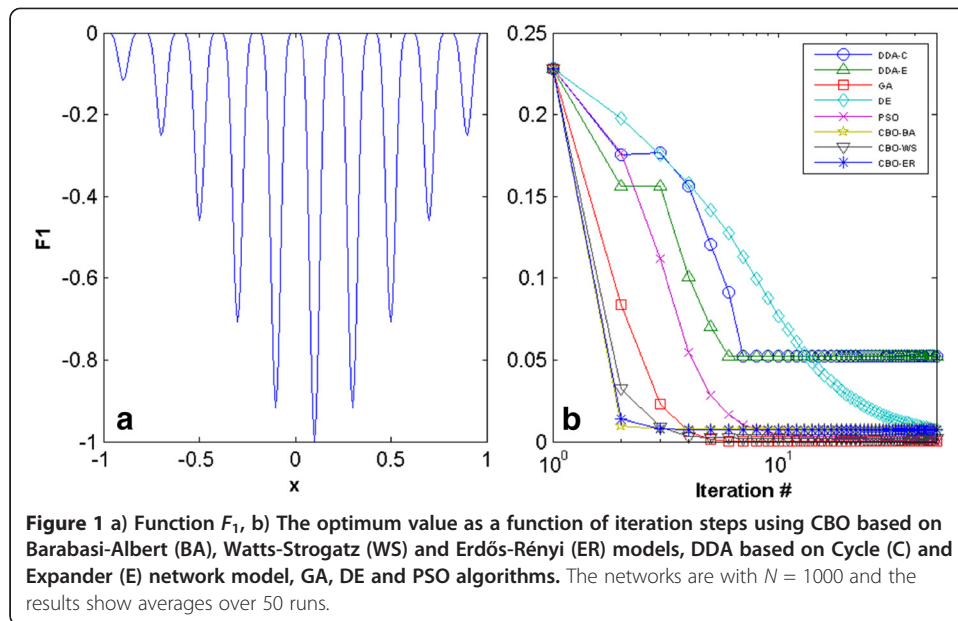
connected with probability P . Research showed that real networks are neither random nor regular but somewhere in between; they are indeed small-world. In order to construct small-world networks, we used the original model proposed by Watts and Strogatz, as follows (Watts & Strogatz 1998). Starting with a regular ring graph in which each node is connected to its k -nearest neighbours, each edge is rewired with probability P , provided that self-loops and duplication of edges are prohibited. They showed that for some intermediate values of the rewiring probability P , we obtain a network with low characteristic path length, comparable to that of random networks, and high clustering coefficient (i.e., transitivity) that is much higher than corresponding random networks.

Erdős-Rényi and Watts-Strogatz models result in networks with almost homogeneous degree distribution. However, it was shown that many real networks have heterogeneous degree distribution; there are many low-degree nodes in the network, while a few nodes are hubs with high degrees (Albert & Barabasi 2002; Barabasi & Albert 1999; Barabási 2009). Barabasi and Albert proposed a preferential attachment growth model for constructing such networks, which is used in this work (Barabasi & Albert 1999). The model starts with a $k + 1$ all-to-all connected nodes. In each step, a new node with k links is added to the network. This node tips to the old nodes with probability that is proportional to their degree, i.e., the higher is the degree of an old node in the network, the higher the probability of the making connection with the new node. The model results in scale-free networks whose degree distribution obeys a power-law (Barabasi & Albert 1999).

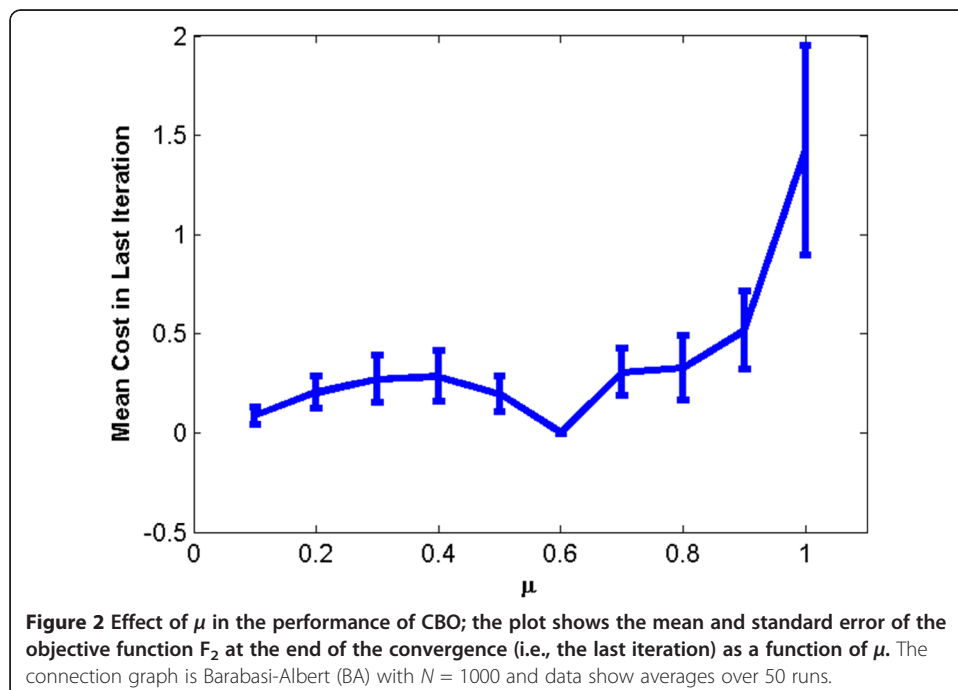
Results and discussion

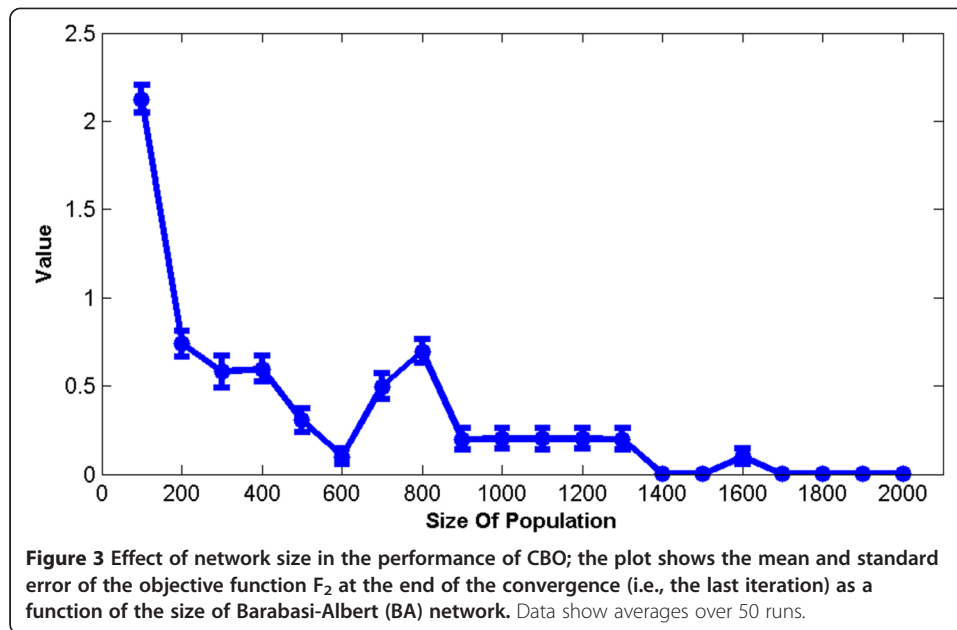
In this section, we report the performance of the proposed Consensus-Based Optimization (CBO) method as compared to other tools including DE, GA and PSO. In order to assess the performance of the methods, they were applied to five benchmark functions as expressed by equations (9)-(13). Function F_1 is a simple function consisting of many local optima and with a minimum of $F_1(x^*) = -1$ at $x^* = 0.1$ (Figure 1a). All classic optimization tools can solve such a simple task. We investigated whether or not the proposed CBO algorithm could find the optimum for this function. Figure 1b shows the performance of CBO when Barabasi-Albert (BA), Watts-Strogatz (WS) and Erdős-Rényi (ER) networks were used as interaction topology. We set the network parameters as $N = 1000$ and $k = 4$ for BA model; $N = 1000$, $k = 4$, and $P = 0.1$ for WS model; $N = 1000$ and $P = 0.1$ for ER model. As it is seen, the optimization strategy of CBO is efficient and could find the optimum value in all topologies. For this simple problem, GA resulted in the best performance followed by CBO method. However, in order to better compare the performance of the methods, we applied them on more difficult objective functions.

In order to assess the ability of CBO on solving more complex optimization tasks, we considered the connection graphs with structural properties as above. Figures 2 and 3 show the influence of parameters on the performance of CBO algorithm in optimizing the objective function F_2 . μ controls the speed of convergence; we obtained the value of $\mu = 0.6$ as optimal for the optimization task (Figure 2). Large values for μ result in zigzag convergence which is clearly indicated in large variance for such cases. Like



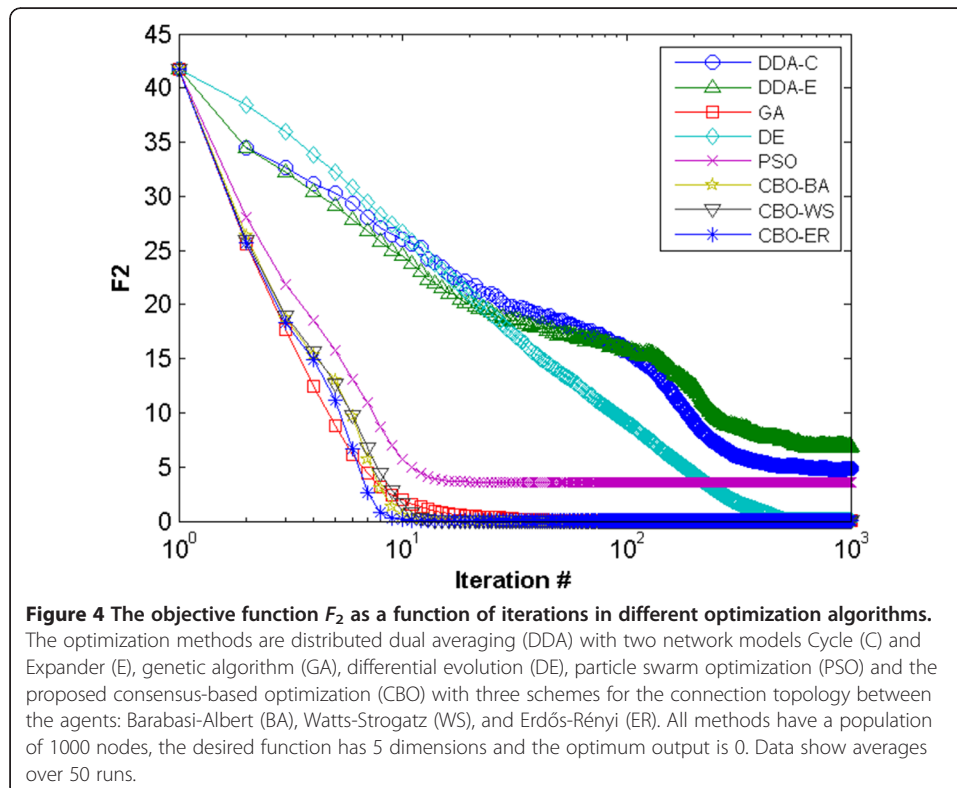
other optimization algorithms, CBO is sensitive to the number of agents (i.e., network size N), in which as the number of agents increases, the performance gets better (Figure 3). However, increasing the size of the network means increase in the computational complexity of the algorithm. For our simulations, we fixed the network size as $N = 1000$, since for larger networks than this size, the improvement is not significant. We also set the number of agents in other optimization algorithms at $N = 1000$; which makes their results comparable. For the parameters of GA, we set $P_c = 0.85$ and $P_m = 0.05$. Furthermore, we set steady-state form for population model and fitness proportionate

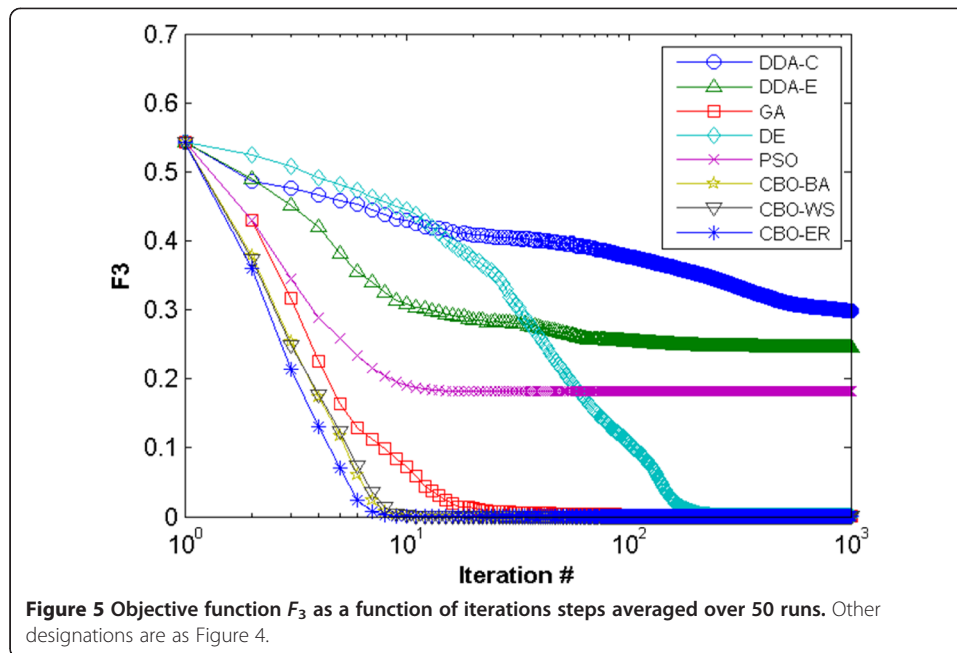




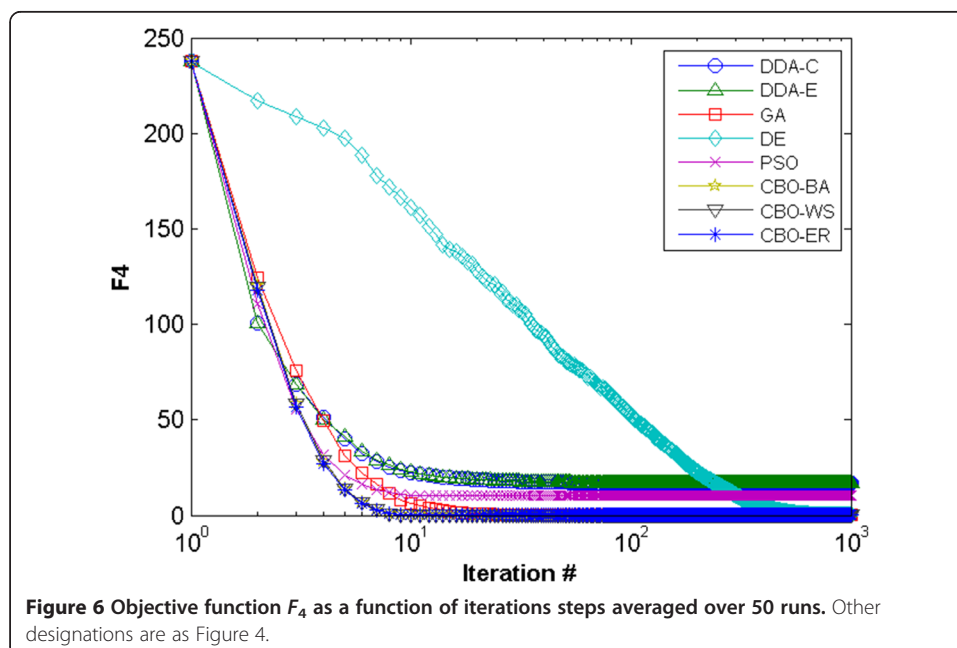
selection policy for parent selection model. The parameters of DE algorithm was set as $\beta = 0.9$, $P_r = 0.9$ and $N_v = 1$ with binomial crossover and random parent selection policy. We set $c_1 = 0.1$ and $c_2 = 0.9$ for PSO resulting in balanced framework for both exploration and exploitation power.

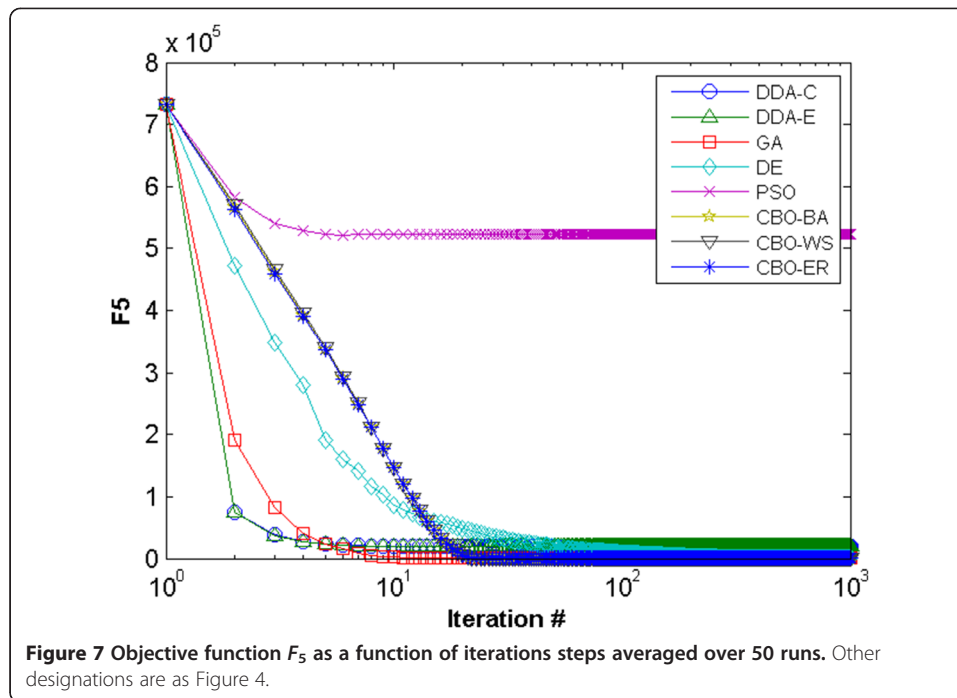
Figures 4,5,6 and 7 show the performance of CBO algorithm using various graph topologies as compared to DE, GA, PSO and DDA methods for optimizing F_2 - F_5 ,





respectively. The results were averaged over 50 runs. Note that some algorithms could not find the optimum solution in the considered number of iterations in some runs, resulting in a drift from the optimum value in their reported profile (which shows average over 50 runs). Majority of the algorithms could find the optimum; however, they showed significantly different performance. CBO with BA and ER as connection structure among the agents, showed the best performance (i.e., finding the optimum in the least iteration steps) as compared to other algorithms for F_2 (Figure 4). They could find the optimum solution in about 16 steps, which is much less than the steps of GA





and DE needed to converge. PSO and CBO with WS topology could not find the solution in some runs resulting in a drift in their performance. Except for F_2 , DDA based on Expander network worked better than or equals the one based on Cyclic one. Although DDA has been shown to be well-performing optimizer for convex functions, it did not result in good performance for the objective functions considered in this work. This is due to the fact that these functions are widely non-convex and a method specially designed for convex functions might not properly work on them.

All algorithms could find the optimum solution for objective function F_3 as expressed by equation (11) with the results shown in Figure 5. However, CBO algorithms showed much faster convergence than others. In terms of network topology in CBO, BA and ER topologies worked better than the case when WS model used for constructing interaction topology. CBO was also better than GA, DE, PSO and DDA in optimizing the objective functions F_4 as expressed by equation (12) and F_5 as expressed by equation (13) and the results are shown in Figure 6 and 7, respectively. While, ER topology resulted in a bit faster convergence than BA and WS for F_4 , their performance was almost the same for F_5 . It is expected that consensus on random networks should be the fastest as compared to the one in BA and WS networks. This is mainly due to the fact that random networks often have shorter average path length compared to other models.

We also compared the computational complexity of these methods by calculating their convergence times for the considered objective functions (Table 1). CBO showed the best performance (i.e., the least number of iterations until convergence) followed by PSO and GA. DE and DDA were the slowest methods among these optimization algorithms.

In sum, our experiments showed that performing an optimization task with a simple consensus network provides the solution with a better performance than a number of

Table 1 Iteration count averaged over 50 runs till convergence happened in optimization methods include distributed dual averaging (DDA), genetic algorithm (GA), differential evolution (DE), particle swarm optimization (PSO) and the proposed consensus-based optimization (CBO)

	DDA	GA	DE	PSO	CBO
F ₁	6	6	26	9	3
F ₂	63	44	535	26	12
F ₃	13	18	87	13	9
F ₄	134	49	365	21	13
F ₅	123	55	121	24	31

classic optimization tools including GA, DE, PSO and DDA. Furthermore, we found pure random topologies constructed by Erdős-Rényi model more effective than small-world topologies constructed by Watts-Strogatz model and scale-free topologies obtained through Barabasi-Albert preferential attachment model.

Conclusion

In this paper we introduced a novel application for consensus phenomenon in complex networks. Consensus in networked structures has many applications ranging from engineering (e.g., sensor networks) to sociology (e.g., opinion formation in social networks). In this manuscript, we used network consensus to solve optimization tasks. We considered a number of agents interacting over a networked structure with topology as random, small-world or scale-free. Furthermore, each agent was associated with an opinion value which could change in collaboration with neighboring agents. The agents worked collectively with their friends (which was defined based on the considered network topology and was kept unchanged during the optimization process) to solve an optimization task. To this end, each agent adapted its opinion value based on the best-matching neighbor, i.e., the neighbor with the best value in the objective function. The proposed consensus-based optimization (CBO) method was applied on a number of benchmark problems and its performance was compared with that of a number of classic optimization tools such as genetic algorithms, differential evolution and particle swarm optimization. Our experiments showed that CBO could always find the optimal solution faster and more reliable. We also found Erdős-Rényi random topology better than Watts-Strogatz small-world and Barabasi-Albert scale-free topologies for which it could solve the optimization task faster when used in CBO as connection graph.

Appendix

Proof of the Convergence Theorem: Based on the assumption made in the theorem

$$\sum_{k=1}^N \min_{1 \leq i, j \leq N, m \geq 1} \{b_{ik}(t_m, t_{m-1}), b_{jk}(t_m, t_{m-1})\} \geq \delta_m, \quad (14)$$

We have

$$\left(1 - \sum_{k=1}^N \min_{1 \leq i, j \leq N, m \geq 1} \{b_{ik}(t_m, t_{m-1}), b_{jk}(t_m, t_{m-1})\}\right) \cdot d(x(m)) \leq (1 - \delta_m) \cdot d(x(m)). \quad (15)$$

Having (15) and the results of Lemma 1, one concludes that

$$\begin{aligned} d(x(m+1)) &= d(B \cdot x(m)) \\ &\leq \left(1 - \sum_{k=1}^N \min_{1 \leq i, j \leq n, m \geq 1} \{b_{ik}(t_m, t_{m-1}), b_{jk}(t_m, t_{m-1})\}\right) \cdot d(x(m)), \\ &\leq (1 - \delta_m) \cdot d(x(m)) \end{aligned} \quad (16)$$

That is

$$d(B \cdot x(m)) \leq (1 - \delta_m) \cdot d(x(m)). \quad (17)$$

Having (17) and using the results of Theorem 1, one concludes

$$(1 - \delta_m) \cdot d(x(m)) \leq e^{-\delta_m} d(x(m)). \quad (18)$$

We have

$$\begin{aligned} e^{-\delta_m} d(x(m)) &\leq e^{-\delta_m} e^{-\delta_{m-1}} d(x(m)) d(x(m-1)) \leq \\ &\leq e^{-\delta_m} e^{-\delta_{m-1}} \dots e^{-\delta_0} d(x(0)) = e^{-\sum_{i=0}^m \delta_i} d(x(0)). \end{aligned} \quad (19)$$

Using (17)-(19) and the fact that $\sum_{t=0}^{\infty} \delta_t = \infty$, we conclude that

$$d(x(m+1)) = d(B \cdot x(m)) \leq e^{-\sum_{i=0}^m \delta_i} d(x(0)) = 0. \quad (20)$$

Or, in other words

$$\lim_{t \rightarrow \infty} d(x(t)) = 0. \quad (21)$$

Since B is a stochastic matrix and $x(t+1) = B(x(t), t) \cdot x(t)$, we have

$$\min_j x_j(t) \leq x_i(t+1) \leq \max_j x_j(t), \quad i = 1, \dots, N, \quad (22)$$

Which becomes the following when time approaches to infinity

$$\lim_{t \rightarrow \infty} \min_j x_j(t) \leq \lim_{t \rightarrow \infty} x_i(t+1) \leq \lim_{t \rightarrow \infty} \max_j x_j(t). \quad (23)$$

Let us denote the lower bound of the opinions by p and their upper bound by q . Then, the above expression becomes

$$p \leq \lim_{t \rightarrow \infty} x_i(t+1) \leq q. \quad (24)$$

Now, using equations (21) and (4) we have

$$\lim_{t \rightarrow \infty} \max_i x_i(t) - \lim_{t \rightarrow \infty} \min_i x_i(t) = p - q = 0 \rightarrow p = q. \quad (25)$$

Applying *Squeeze Theorem* (Steinhaus 1938), we have

$$\lim_{t \rightarrow \infty} x_i(t+1) = \lim_{t \rightarrow \infty} B(x(t), t) \cdot x_i(t) = x^*, \quad i = 1, \dots, N. \quad (26)$$

And this completes the proof.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

OAS and MJ designed the research and developed the methods. OAS did the numerical simulations. OAS and MJ wrote the manuscript. Both authors read and approved the final manuscript.

Acknowledgment

The authors would like to thank Dr. Hamid Beigy for his insightful comments and discussion on the topic.

Received: 31 October 2012 Accepted: 28 March 2013

Published: 24 April 2013

References

- Ajdari Rad A, Jalili M, Hasler M: Efficient rewirings for enhancing synchronizability of dynamical networks. *Chaos* 2008, **18**:037104.
- Albert HJR, Barabasi A-L: "Diameter of the World Wide Web". *Nature* 1999, **401**:130–131.
- Albert R, Barabasi A-L: Statistical mechanics of complex networks. *Rev Modern Physics* 2002, **74**:47–97.
- Amblard F, Deffuant G: The role of network topology on extremism propagation with the relative agreement opinion dynamics. *Physica A* 2004, **343**:725–738.
- Barabási AL: Scale-free networks: a decade and beyond. *Science* 2009, **325**:412–413.
- Barabasi A-L, Albert R: Emergence of scaling in random networks. *Science* 1999, **286**:5009–5012.
- Barbarossa S, Scutari G: Decentralized maximum likelihood estimation for sensor networks composed of nonlinearly coupled dynamical systems. *IEEE Trans Signal Process* 2007, **55**:3456–3470.
- Barry CMN, Wentzel KR: Friend influence on prosocial behavior: The role of motivational factors and friendship characteristics. *Dev Psychol* 2006, **42**:153–162.
- Belykh I, Hasler M, Laurent M, Nijmeijer H: Synchronization and graph topology. *Int J Bifurcation Chaos* 2005, **15**:3423–3433.
- Berndt TJ, Murphy LM: Influences of friends and friendships: Myths, truths, and research recommendations. *Adv Child Dev Behav* 2003, **30**:275–310.
- Boccaletti VLS, Moreno Y, Chavez M, Hwang DU: "Complex networks: structure and dynamics". *Physics Reports* 2006, **424**:175–308.
- Boyd S, Vandenberghe L: *Convex optimization*. Cambridge University Press; 2004.
- Brunetti S, Cordasco G, Gargano L, Lodi E, Quattrociocchi W: "Minimum Weight Dynamo and Fast Opinion Spreading". In *38th international conference on Graph-Theoretic Concepts in Computer Science*. 2012:249–261.
- Carletti T, Fanelli D, Grolli S, Guarino A: How to make an efficient propaganda. *Europhysics Letters* 2006, **74**:222–228.
- Deffuant G, Neau D, Amblard F, Weisbuch G: Mixing beliefs among interacting agents. *Adv Comp Sys* 2001, **3**:87–98.
- Duchi AAJC, Wainwright MJ: Dual averaging for distributed optimization: convergence analysis and network scaling. *IEEE Trans Automatic Con* 2012, **57**:592–606.
- Eberhart RC, Shi Y: Comparing inertia weights and constriction factors in particle swarm optimization. In *Proceedings of the Congress on Evolutionary Computation*. 2000, 1:84–88.
- Erdős P, Rényi A: On the evolution of random graphs. *Pub Mathematic Ins Hungarica Acad Sci* 1960, **5**:17–61.
- Fortunato S, Latora V, Pluchino A, Rapisarda A: "Vector opinion dynamics in a bounded confidence consensus model". *Int J Modern Physics C* 2005, **16**:1535–1553.
- Gandica Y, Del Castillo-Mussot M, Vázquez GJ, Rojas S: Continuous opinion model in small-world directed networks. *Physica A* 2010, **389**:5864–5870.
- Girvan M, Newman MEJ: Community structure in social and biological networks. *Proc Nat Acad Sci USA* 2002, **99**:7821–7826.
- Guo L, Cai X: Continuous opinion dynamics in complex networks. *Com Computational Physics* 2009, **5**:1045–1053.
- Hegselmann R, Krause U: Opinion dynamics and bounded confidence models, analysis, and simulation. *J Art Soc Social Simulation* 2002:5, no. 3.
- Holland JH: *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: The University of Michigan Press; 1975.
- Holyst JA, Kacperski K, Schweitzer F: Social impact models of opinion dynamics. *Ann Rev Computatio Physics* 2001, **20**:531–535.
- Jalili M: "Effects of leaders and social power on opinion formation in complex networks". In *Simulation - Transactions of the Society for Modeling and Simulation International*. 2013a. doi:10.1177/0037549712462621.
- Jalili M: Social power and opinion formation in complex networks. *Physica A* 2013b, **392**:959–966.
- Kennedy J, Eberhart R: "Particle swarm optimization". In *Proceedings IEEE International Conference Neural Networks* 1995, **4**:1942–1948.
- Kozma B, Barrat A: Consensus formation on adaptive networks. *Physical Review E* 2008, **77**:016102.
- Krapivsky PL, Redner S: Dynamics of majority rule in two-state interacting spin systems. *Phys Rev Let* 2003, **90**:238701.
- Krause U: A discrete nonlinear and non-autonomous model of consensus formation. *Com Diff Eq* 2000, **2000**:227–236.
- Kurmyshev E, Juárez HA, González-Silva RA: Dynamics of bounded confidence opinion in heterogeneous social networks: concord against partial antagonism. *Physica A* 2011, **390**:2945–2955.
- Leskovec J, Adamic LA, Huberman BA: *The dynamics of viral marketing*. ACM Conference on Electronic Commerce. Ann Arbor: Michigan, USA; 2006.
- Lewenstein M, Nowak A, Latane B: Statistical mechanics of social impact. *Physical Rev A* 1992, **45**:703–716.
- Lorenz J: Continuous opinion dynamics under bounded confidence: a survey. *Int J Modern Physics C* 2007, **18**:1819–1838.

- Lu JH, Yu XH, Chen GR, Cheng DZ: **Characterizing the synchronizability of small-world dynamical networks.** *IEEE Tran Cir Syst I: Fundamental Theory Ad App* 2004, **51**:787–796.
- Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U: **Network motifs: simple building blocks of complex networks.** *Science* 2002, **298**:824–827.
- Nardini C, Kozma B, Barrat A: **Who's talking first? consensus or lack thereof in coevolving opinion formation models.** *Phys Rev Let* 2008, **100**:158701.
- Nesterov Y: **Primal-dual subgradient methods for convex problems.** *Mathematical Programming* 2009, **120**:221–259.
- Newman MEJ: **The structure and function of complex networks.** *SIAM Review* 2003, **45**:167–256.
- Newman MEJ, Park J: **Why social networks are different from other types of networks.** *Physical Review E* 2003, **68**:036122.
- Newman A-LBM, Watts DJ: *The structure and dynamics of networks. 1.* Princeton University Press; 2006.
- Perc M: **Evolution of cooperation on scale-free networks subject to error and attack.** *New J Physics* 2009, **11**:033027.
- Perc M, Szolnoki A: **Coevolutionary games-A mini review.** *Biosystems* 2010, **99**:109–125.
- Scutari G, Barbarossa S, Pescosolido L: **Distributed decision through self-synchronizing sensor networks in the presence of propagation delays and asymmetric channels.** *IEEE Trans Signal Process* 2008, **56**:1667–1684.
- Seneta E: *Non-negative matrices and Markov chains. 2nd edition.* Berlin: Springer Verlag, New York; 1981.
- Steinhaus H: **A note on the ham sandwich theorem.** *Mathesis Polska* 1938, **9**:26–28.
- Storn R, Price K: **Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces.** *J Global Optimization* 1997, **11**:341–359.
- Strogatz SH: **Exploring complex networks.** *Nature* 2001, **410**:268–276.
- Tang K, Xiaodong L, Suganthan N, Yang PZ, Weise T: **"Benchmark functions for the cec'2010 special session and competition on large-scale global optimization".** *Technical report, University of Science and Technology of China* 2009, **1**:1–23.
- Urbig D, Lorenz J, Herzberg H: **Opinion dynamics: the effect of the number of peers met at once.** *J Artif Soc Social Simulation* 2008, **11**:4.
- Watts DJ, Dodds PS: **Influentials, networks, and public opinion formation.** *J Consumer Res* 2007, **34**:441–458.
- Watts DJ, Strogatz SH: **Collective dynamics of 'small-world' networks.** *Nature* 1998, **393**:440–442.
- Weisbuch G: **Bounded confidence and social networks.** *Eur Phys J B* 2004, **38**:339–343.
- Xiao L: **Dual averaging methods for regularized stochastic learning and online optimization.** *J Machine Lear Res* 2010, **11**:2543–2596.
- Yang H-X, Wu Z-X, Zhou C, Zhou T, Wang B-H: **Effects of social diversity on the emergence of global consensus in opinion dynamics.** *Physical Review E* 2009, **80**:046108.

doi:10.1186/2194-3206-1-11

Cite this article as: Askari-Sichani and Jalili: Large-scale global optimization through consensus of opinions over complex networks. *Complex Adaptive Systems Modeling* 2013 **1**:11.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com