Approximate Consensus in Stochastic Networks with Application to Load Balancing

Natalia Amelina, Alexander Fradkov, *Fellow, IEEE,* Yuming Jiang, *Member, IEEE,* and Dimitrios J. Vergados, *Member, IEEE*

***Abstract*—This paper is devoted to the approximate consensus problem for stochastic networks of nonlinear agents with switching topology, noisy and delayed information about agent states. A local voting protocol with nonvanishing (e.g., constant) step size is examined under time-varying environments of agents. To analyze dynamics of the closed loop system, the so-called method of averaged models is used. It allows us to reduce analysis complexity of the closed loop stochastic system. We derive the upper bounds for mean square distance between states of the initial stochastic system and its approximate averaged model. These upper bounds are used to obtain conditions for approximate consensus achievement.**

An application of general theoretical results to the load balancing problem in stochastic dynamic networks with incomplete information about the current states of agents and with changing set of communication links is considered. The conditions to achieve the optimal level of load balancing are established. The performance of the system is evaluated both analytically and by simulation.

***Index Terms*—Approximate consensus, stochastic discrete networks, distributed information systems, load balancing.**

I. Introduction

THE problems of distributed interaction in dynamical networks attracted much attention in the last decade. A number of survey papers [1, 2], monographs [3–5], special issues of journals [6, 7] and edited volumes [8] have been published in this area. This interest has been driven by applications in various fields, including, for example, information exchange in multiprocessor networks, transportation networks, production networks, sensor and wireless networks, coordinated motion for unmanned flying vehicles, submarines and mobile robots, distributed control systems for power networks, complex crystal lattices, and nanostructured plants [1–11].

Despite a large number of publications, satisfactory solutions have been obtained mostly for a restricted class of problems (see [1–8] and references therein). Factors such as nonlinearity of agent dynamics, switching topology, noisy and delayed measurements of agents' states may significantly complicate the solutions. Additional important factors are the limited transmission rate in the channel and discretization phenomenon. In the presence of various disruptive factors, asymptotically exact consensus may be hard to achieve, especially in a time-varying environment. For such cases, approximate consensus problems need to be examined.

In this paper, we investigate the approximate consensus problem in a multi-agent stochastic system with nonlinear dynamics, noisy and delayed information about agent states, and uncertainties in the topology and in the control protocol. Such a problem is important for the analysis and control of production networks, multiprocessors, sensor, wireless or multicomputer networks, *etc. To study the considered stochastic system we use the method of averaged models, which allows us to reduce the analysis complexity.* This method has been adopted to analyze various types of information systems (see, for example, [12–14]). *As an example, the load balancing system in a network with noisy and delayed information about the load and with switching topology is studied.* In contrast to the existing stochastic approximation-based control algorithms (protocols), local voting with **nonvanishing step size** is considered.

The consensus problem on graphs with noisy measurements of its neighbors states under general imperfect communications is considered in [15, 16], where stochastic approximation-type algorithms with *decreasing to zero step size* are used. Noisy convergence with nonvanishing step-size was studied in [17], but the step parameters were chosen differently for different agents and the network scenario considered is a specific one. The stochastic gradient-like (stochastic approximation) methods have also been used in the presence of stochastic uncertainties [15, 18–21]. For the linear case without feedback in stochastic network the problem of achieving an approximate consensus was considered in [22]. However, in the works [15, 17–21], the network scenarios considered are specific ones, much simpler than the more general network scenarios considered in this paper.

In [23], a stochastic approximation type algorithm was proposed for solving consensus problem and justified for the group of cooperating agents that communicate with imperfect information in discrete time, under the conditions of dynamic topology and delay. Under some general assumptions a necessary and sufficient condition was proved for the asymptotic mean square consensus when *step size tends to zero* and only

N. Amelina is with the Department of Telematics, Norwegian University of Science and Technology, NO-7491, Trondheim, Norway and also with Faculty of Mathematics and Mechanics, St. Petersburg State University, 198504, Universitetskii pr. 28, St. Petersburg, Russia, e-mail: natalia_amelina@mail.ru.

A. Fradkov is with the Institute of Problems in Mechanical Engineering, 199178, St. Petersburg, Russia and also with Faculty of Mathematics and Mechanics, St. Petersburg State University, 198504, Universitetskii pr. 28, St. Petersburg, Russia e-mail: fradkov@mail.ru.

Y. Jiang is with the Department of Telematics, Norwegian University of Science and Technology, NO-7491, Trondheim, Norway, e-mail: jiang@item.ntnu.no.

D. J. Vergados is with the Department of Telematics, Norwegian University of Science and Technology, NO-7491, Trondheim, Norway, e-mail: dimitrios.vergados@item.ntnu.no.

Manuscript received XX XX; revised XX XX, XXXX. Some of the results were presented at IEEE MSC Conference 2012 [36], and SNPD Conference 2012 [38].

with a simple dynamics when changes equal control actions. However, under time-varying environment (e.g., feeding new jobs) using step sizes that decrease to zero may greatly affect the convergence rate. In this paper, we focus on a more general case of nonlinear functions, which describe dynamics of agents, and nondecreasing to zero step size.

The importance of algorithms with nonvanishing step size (gain) was realized and first convergence results were obtained still in the 1970s [13, 24], see also references in [14]. In [25–27] the efficiency of stochastic approximation algorithms with constant step size was studied for some specific cases.

As for the load balancing problem, numerous articles are devoted to it (e.g., [28–34]), indicating the relevance of this topic. However, most of these articles do not consider noise or delays. While in a single computer this assumption could be rather realistic, if we consider networked systems, noise, delays and possible link-“breaks” need to be accounted for. The load balancing problem in centralized networks under uncertainties about agent productivities was analyzed in [35] where jobs among agents are redistributed by a load broker. However, in case when there is no agent which connects with every other agent, it is not possible to choose one of the agents as the load broker. In this case, it is necessary to consider decentralized networks. However, to the best of our knowledge, few results for load-balancing in such distributed networks are available.

In this paper, the results of our previous works [36–39] are summarized, extended and improved. In particular, we relax the assumption of the boundedness of weights of the control protocol, replacing it by the boundedness of its variances. In addition, new and much larger size simulation experiments were performed and results added.

The contributions of the paper are several-fold. First, the approximate consensus problem for a general network scenario is investigated, which is a network of nonlinear agents with switching topology, noisy and delayed measurements. Second, in this approximate consensus problem, we consider a nonlinear dynamics and the local voting protocol with step size α_t nondecreasing to zero. Third, to analyze the dynamics of the stochastic discrete systems, the method of averaged models (Derevitskii-Fradkov-Ljung (DFL)-scheme) [24, 40, 41] is adopted. Fourth, the approximate consensus conditions are obtained. In addition, to demonstrate the use of the obtained results, the load balancing problem in a distributed network is studied. Furthermore, simulation results validating the analysis are presented.

The rest of the paper is organized as follows. In Section II, the basic concepts of graph theory are introduced and the consensus problem is posed. In Section III, the main assumptions are described and the consensus conditions are derived. In Sections IV, the load balancing problem is considered, the analytical and simulation results are presented and discussed. Section V contains concluding remarks.

II. Preliminaries

A. Concepts of Graph Theory

First we present the notation used in this article. Consider a network as a set of agents (nodes) $N = \{1,2,\dots,n\}$.

A *directed graph (digraph)* $G = (N,E)$ consists of a set N and a set of directed edges E. Denote the *neighbour set* of node i as $N^i = \{j : (j,i) \in E\}$. (Here and later the agent index i is used as a superscript and not as an exponent.)

We associate a weight $a^{i,j} > 0$ with each edge $(j,i) \in E$. Matrix $A = [a^{i,j}]$ is called an *adjacency or connectivity matrix* of the graph. Denote $\mathscr{G}_A$ as the corresponding graph. The *in-degree* of node i is the number of edges having i as head. The *out-degree* of node i is the number of edges having i as tail. If the in-degree equals to the out-degree for all nodes $i \in N$ the graph is said to be *balanced*. Define the *weighted in-degree* of node i as the i-th row sum of A: $d^i(A) = \sum_{j=1}^n a^{i,j}$ and $D(A) = \text{diag}\{d^1(A), d^2(A),\dots,d^n(A)\}$ is the corresponding diagonal matrix. The symbol $\mathscr{L}(A) = D(A) - A$ stands for the *Laplacian* of graph $\mathscr{G}_A$.

A directed path from i_1 to i_s is a sequence of nodes $i_1,\dots,i_s$, $s \geq 2$, such that $(i_k, i_{k+1}) \in E, k \in \{1,2,\dots,s-1\}$. Node i is said to be *connected* to node j if a directed path from i to j exists. *The distance* from i to j is the length of the shortest path from i to j. The graph is said to be *strongly connected* if i and j are connected for all distinct nodes $i, j \in N$.

A *directed tree* is a digraph where each node i, except the roots, has exactly one parent node j so that $(j,i) \in E$. We call $\overline{\mathscr{G}_A} = (\overline{N}, \overline{E})$ a *subgraph* of $\mathscr{G}_A$ if $\overline{N} \subset N$ and $\overline{E} \subset E \cap (\overline{N} \times \overline{N})$. The digraph $\mathscr{G}_A$ is said to contain a *spanning tree* if there exists a directed tree $\mathscr{G}_{tr} = (N, E_{tr})$ as a subgraph of $\mathscr{G}_A$.

The symbol $d_{\max}(A)$ denotes a maximal in-degree of the graph $\mathscr{G}_A$. In correspondence with the Gershgorin Theorem [42], we can deduce another important property of the Laplacian: *all eigenvalues of the matrix $\mathscr{L}(A)$ have nonnegative real part and belong to the circle with center at the point $(0, d_{\max}(A))$ and with radius which equals to $d_{\max}(A)$.*

Let $\lambda_1,\dots,\lambda_n$ denote eigenvalues of the matrix $\mathscr{L}(A)$. We arrange them in ascending order of real parts: $0 \leq Re(\lambda_1) \leq Re(\lambda_2) \leq \dots \leq Re(\lambda_n)$. It is known, that if the graph has a spanning tree then $\lambda_1 = 0$ is a simple eigenvalue and all other eigenvalues of $\mathscr{L}$ are in the open right half of the complex plane.

The second eigenvalue λ_2 of matrix $\mathscr{L}$ is important for analysis in many applications. It is usually called Fiedler eigenvalue. For undirected graphs it was shown in [3] that:

$$Re(\lambda_2) \leq \frac{n}{n-1} \min_{i \in N} d^i(A), \tag{1}$$

and for the connected undirected graph G_A

$$Re(\lambda_2) \geq \frac{1}{\text{diam}G_A \cdot \text{vol}G_A}, \tag{2}$$

where $\text{diam}G_A$ is the longest distance between two nodes and $\text{vol}G_A = \sum_{i \in N} d^i(A)$.

For all vectors the ℓ_2-norm will be used, i.e., a square root of the sum of all its elements squares.

B. Problem Statement

1) The network model: Consider a dynamic network of n agents that collaborate to solve a problem that each cannot solve alone.

The concepts of graph theory will be used to describe the network topology. Let the dynamic network topology be modeled by a sequence of digraphs $\{(N,E_t)\}_{t\geq 0}$, where $E_t \subset E$ changes with time. The corresponding adjacency matrices are denoted as A_t. The maximal set of communication links is $E_{\max} = \{(j,i) : \sup_{t\geq 0} a_t^{i,j} > 0\}$.

We assume that a time-varying state variable $x_t^i \in \mathbb{R}$ corresponds to each agent $i \in N$ of the graph at time $t \in [0,T]$. Its dynamics are described for the discrete time case as

$$x_{t+1}^i = x_t^i + f^i(x_t^i, u_t^i),\ t = 0,1,2\ldots,T \qquad (3)$$

or for the continuous time case as

$$\dot{x}_t^i = f^i(x_t^i, u_t^i),\ t \in [0,T], \qquad (4)$$

with some functions $f^i(\cdot,\cdot) : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$, depending on states in the previous time x_t^i and control actions $u_t^i \in \mathbb{R}$.

Each agent uses its own state (possibly noisy) to form its update strategy

$$y_t^{i,i} = x_t^i + w_t^{i,i}, \qquad (5)$$

and if $N_t^i \neq \emptyset$, noisy and delayed measurements of its neighbors states

$$y_t^{i,j} = x^j_{t-d_t^{i,j}} + w_t^{i,j},\ j \in N_t^i, \qquad (6)$$

where $w_t^{i,i}, w_t^{i,j}$ are noises, $0 \leq d_t^{i,j} \leq \bar{d}$ is an integer-valued delay, and $\bar{d} \geq 0$ is a maximal delay.

If $(j,i) \in E_t$ then agent i receives information from agent j for the purposes of consensus seeking.

2) The local voting protocol:

Definition *1:* A feedback on observations

$$u_t^i = K_t^i(y_t^{i,j_1},\ldots,y_t^{i,j_{m_i}}), \qquad (7)$$

where $\{j_1,\ldots,j_{m_i}\} \in \{i\}\bigcup \overline{N}_t^i$, $\overline{N}_t^i \subseteq N_t^i$ is called a *protocol* (*update algorithm*) with topology (N,E_t).

In this paper, we consider the *local voting protocol*:

$$u_t^i = \alpha_t \sum_{j\in\overline{N}_t^i} b_t^{i,j}(y_t^{i,j} - y_t^{i,i}), \qquad (8)$$

where $\alpha_t > 0$ are step sizes of the protocol, $b_t^{i,j} > 0\ \forall j \in \bar{N}_t^i$. We set $b_t^{i,j} = 0$ for other pairs i,j and denote $B_t = [b_t^{i,j}]$ as the matrix of the protocol.

Note, that protocol (8) differs from a frequently used such protocol, where step parameters α vary for different agents $i \in N$ (for example, $\alpha^i = 1/d^i(B_t)$, see [17]).

3) Consensus concepts: The network is said to achieve ε-*consensus* at time t if there exists a variable $x^\star$ such that $||x_t^i - x^\star||^2 \leq \varepsilon$ for all $i \in N$.

The network is said to achieve *mean square ε-consensus* at time t if there exists a variable $x^\star$ such that $\mathrm{E}||x_t^i - x^\star||^2 \leq \varepsilon$ for all $i \in N$.

Definition *2:* The network is said to achieve *asymptotic mean square ε-consensus* if $\mathrm{E}||x_1^i||^2 < \infty$, $i \in N$ and there exists a variable $x^\star$ such that $\overline{\lim}_{t\to\infty} \mathrm{E}||x_t^i - x^\star||^2 \leq \varepsilon$ for all $i \in N$.

$T(\varepsilon)$ is called *time to ε-consensus*, if the network achieves ε-consensus for all $t \geq T(\varepsilon)$.

For reader's convenience, we provide a list of key notation used in this paper.

N	$\{1,2,\ldots,n\}$ — the set of nodes
E	$\{i,j\}$ — the set of edges, $i,j \in N$
$a^{i,j}$	weight of edge $(j,i) \in E$
(N,E)	digraph with nodes N and edges E
N^i	neighbour set of node i
A	adjacency or connectivity matrix
$\mathscr{G}_A$	graph defined by the adjacency matrix A
$d^i(A)$	$\sum_{j=1}^n a^{i,j}$ — weighted in-degree of node i (i-th row sum of A)
$d_{\max}(A)$	maximal in-degree of the graph $\mathscr{G}_A$
$D(A)$	$\mathrm{diag}\{d^1(A),d^2(A),\ldots,d^n(A)\}$ — diagonal matrix of weighted in-degree of A
$\mathscr{L}(A)$	$D(A)-A$ — Laplacian of graph $\mathscr{G}_A$
$\lambda_1,\ldots,\lambda_n$	eigenvalues of the matrix $\mathscr{L}(A)$
$\mathrm{diam}G_A$	diameter, the longest distance between two nodes
$\mathrm{vol}G_A$	$\sum_{i\in N} d^i(A)$ — volume, the sum of in-degrees
$E_{\max}$	$\{(j,i) : \sup_{t\geq 0} a_t^{i,j} > 0\}$ — maximal set of communication links
x_t^i	state of agent i at time t
$y_t^{i,j}$	noisy and delayed measurement agent i obtains from agent j at time t
$w_t^{i,j}$	noise in $y_t^{i,j}$ at time t
$d_t^{i,j}$	integer-valued delay in $y_t^{i,j}$ at time t
$\bar{d}$	maximal delay
u_t^i	control actions
K_t^i	protocol with topology (N,E_t)
$\bar{N}_t^i$	subset of N_t^i at time t
$\alpha_t > 0$	step sizes of the local voting protocol
$b_t^{i,j}$	weight parameter of the local voting protocol
B_t	matrix of the local voting protocol
$\bar{x}_t$	$[x_t^1;\ldots;x_t^n]$
$\bar{u}_t$	$[u_t^1;\ldots;u_t^n]$
I	identity matrix of size $n\times n$
$\underline{1}$	vector consisting of 1's
$\bar{z}_1$	left eigenvector of matrix P: $\bar{z}_1 = [z^1,\ldots,z^n]$
$T(\varepsilon)$	time to ε-consensus
$x^\star$	consensus value
E	mathematical expectation
E_x	conditional expectation under initial condition x
$\mathrm{E}_{\mathscr{F}}$	conditional expectation with respect to σ-algebra $\mathscr{F}$
$b^{i,j}$	$\mathrm{E}b_t^{i,j}$
$p_k^{i,j}$	probability that the delay $d_t^{i,j}$ equals k
$\bar{n}$	$n(\bar{d}+1)$
$A_{\max}$	adjacency matrix of the averaged system
τ_t	$\alpha_0+\alpha_1+\ldots+\alpha_{t-1}$
$\tau_{\max}$	$\sum_{t=0}^T \alpha_t$
$\bar{X}_t$	$[\bar{x}_t,\bar{x}_{t-1},\ldots,\bar{x}_{t-\bar{d}}]$

q_t^i	queue length of the atomic elementary jobs of the agent i at time t
p_t^i	productivity of the agent i at time t
z_t^i	new job received by agent i at time t
T_t	implementation time of jobs at time t
$\frac{q_t^i}{p_t^i}$	load of agent i at time t
Err	$\sqrt{\sum_i \frac{(x_t^i-x^\star)^2}{n}}$ — average residual
$\lvert D(t)\rvert$	maximum deviation from the average load on the network

III. Main Results

In this section, we present the main results of this paper. All proofs are included in the Appendix.

A. Main Assumptions

Let $(\Omega,\mathscr{F},P)$ be the underlying probability space corresponding to the sample space, the collection of all events, and the probability measure respectively.

We assume that the following conditions are satisfied:

A1. $\forall i\in N$ functions $f^i(x,u)$ are Lipschitz in x and u: $|f^i(x,u)-f^i(x',u')|\le L_1(L_x|x-x'|+|u-u'|)$, and for any fixed x the function $f^i(x,\cdot)$ is such that $\mathrm{E}_x f^i(x,u)=f^i(x,\mathrm{E}_x u)$. The last part of Assumption **A1** is satisfied if the system is almost surely affine in the control.

From this Lipschitz condition it follows that the growth rate is bounded: $|f^i(x,u)|^2\le L_2(L_c+L_x|x|^2+|u|^2)$.

A2. **a)** $\forall i\in N, j\in N^i_{\max}$ the noises $w_t^{i,j}$ are centered, independent and have bounded variance $\mathrm{E}(w_t^{i,j})^2\le\sigma_w^2$.

b) $\forall i\in N, j\in N^i_{\max}$ appearances of variable edges (j,i) in graph $\mathscr{G}_{A_t}$ are independent random events.

c) $\forall i\in N, j\in N^i_{\max}$ weights $b_t^{i,j}$ in the protocol (8) are independent random variables with $b^{i,j}=\mathrm{E}b_t^{i,j}$, $\sigma_b^{i,j}=E(b_t^{i,j}-b^{i,j})^2<\infty$.

d) $\forall i\in N, j\in N^i$ there exists a finite quantity $\bar d\in\mathbb{N}$: $d_t^{i,j}\le\bar d$ with probability 1 and integer-valued delays $d_t^{i,j}$ are independent, identically distributed random variables taking values $k=0,\dots,\bar d$ with probabilities $p_k^{i,j}$.

Moreover, all these random variables and matrices are mutually independent.

The next assumption is for a matrix $A_{\max}$ constructed as follows. Specifically, if $\bar d>0$, we add new "fictitious" agents whose states at time t equal to the corresponding states of the "real" agents at the previous $\bar d$ time: $t-1,t-2,\dots,t-\bar d$. Then, $A_{\max}$ is a matrix of size $\bar n\times\bar n$, where $\bar n=n\times(\bar d+1)$, with

$$a^{i,j}_{\max}=p^{i,j\,\mathrm{mod}\,\bar d}_{j\,\mathrm{div}\,\bar d}\,b^{i,j\,\mathrm{mod}\,\bar d},\ i\in N,\ j=1,2,\dots,\bar n, \quad (9)$$

$$a^{i,j}_{\max}=0,\ i=n+1,n+2,\dots,\bar n,\ j=1,2,\dots,\bar n.$$

Here, the operation *mod* is a remainder of division, and *div* is a division without remainder.

Note that if $\bar d=0$, this definition of extended network topology (of matrix $A_{\max}$ of size $n\times n$) is reduced to

$$a^{i,j}_{\max}=b^{i,j},\ i\in N,\ j\in N. \quad (10)$$

Also note that we have defined a matrix $A_{\max}$ in such a way that $\mathrm{E}_{\bar x_t}\bar u_t=-\alpha_t\mathscr{L}(A_{\max})\bar x_t$. We assume that the following condition is satisfied for this network topology matrix:

A3. Graph $(N,E_{\max})$ has a spanning tree, and for any edge $(j,i)\in E_{\max}$ among the elements $a^{i,j}_{\max},a^{i,j+n}_{\max},\dots,a^{i,j+\bar dn}_{\max}$ of the matrix $A_{\max}$, there exists at least one non-zero element.

B. Analysis of the Closed Loop System Dynamics

Denote $\bar X_t\in\mathbb{R}^{n\bar d}$ as the extended state vector $\bar X_t=[\bar x_t,\bar x_{t-1},\dots,\bar x_{t-\bar d}]$, where $\bar x_s\equiv[x_s^1;\dots;x_s^n]$ and, for $-\bar d\le s<0$, let $\bar x_s=[0;\dots;0]$.

Rewrite the dynamics of the agents in vector-matrix form:

$$\bar X_{t+1}=U\bar X_t+F(\alpha_t,\bar X_t,\bar w_t), \quad (11)$$

where U is the following matrix of size $\bar n\times\bar n$:

$$U=\begin{pmatrix} I & 0 & 0 & \dots & 0\\ I & 0 & 0 & \dots & 0\\ 0 & I & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \dots & I & 0\end{pmatrix}, \quad (12)$$

where I is the identity matrix of size $n\times n$, and $F(\alpha_t,\bar X_t,\bar w_t):\mathbb{R}\times\mathbb{R}^{\bar n}\times\mathbb{R}^{n^2}\to\mathbb{R}^{\bar n}$ — vector function of the arguments:

$$F(\alpha_t,\bar X_t,\bar w_t)=$$

$$=\begin{pmatrix} \dots\\ f^i(x_t^i,\alpha_t\sum_{j\in\bar N_t^i}b_t^{i,j}((x^j_{t-d_t^{i,j}}-x_t^i)+(w_t^{i,j}-w_t^{i,i})))\\ \dots\\ 0_{n\bar d}\end{pmatrix}, \quad (13)$$

containing non-zero components only in the first n places.

Consider the averaged discrete model corresponding to (11):

$$\bar Z_{t+1}=U\bar Z_t+G(\alpha_t,\bar Z_t),\ \bar Z_0=\bar X_0, \quad (14)$$

where

$$G(\alpha,\bar Z)=G\begin{pmatrix} & z^1\\ \alpha, & \vdots\\ & z^{n(\bar d+1)}\end{pmatrix}=\begin{pmatrix}\dots\\ f^i(z^i,\alpha s^i(\bar Z))\\ \dots\\ 0_{n\bar d}\end{pmatrix}, \quad (15)$$

$$s^i(\bar Z)=\sum_{j\in N^i}b^{i,j}((\sum_{k=0}^{\bar d}p_k^{i,j}z^{j+kn})-z^i)= \quad (16)$$

$$=-d^i(A_{\max})z^i+\sum_{j=1}^{\bar n}a^{i,j}_{\max}z^j,\ i\in N.$$

It turns out that the trajectory of solutions of the initial system $\{\bar X_t\}$ from (11) at time t is close in the mean square sense to the average trajectory of the discrete system (14).

In the following theorem the upper bounds for mean square distance between the solutions to the initial system and solutions to its averaged discrete model will be given.

Theorem *1:* If conditions **A1**, **A2** are satisfied, $0<\alpha_t\le\bar\alpha$, **then** there exists $\tilde\alpha$ such that for $\bar\alpha<\tilde\alpha$, the following inequality holds:

$$\mathrm{E}\max_{0\le t\le T}||\bar X_t-\bar Z_t||^2\le c_1\tau_T e^{c_2\tau_T^2}\bar\alpha, \quad (17)$$

where $\tau_T = 2^{\bar{d}}(\alpha_0 + \alpha_1 + \ldots + \alpha_{T-1})$, $c_1, c_2 > 0$ are some constants:

$$c_1 = 8n\left(\tilde{c} + \hat{c}(\frac{nL_2L_c + \bar{\alpha}^2\tilde{c}}{c_3} + ||\bar{X}_0||^2)e^{T\ln(c_3+1)}\right), \quad (18)$$

$$c_2 = 2^{1-\bar{d}}L_1^2(\frac{L_x}{\underline{\alpha}} + 2\bar{\alpha}^2||\mathscr{L}(A_{\max})||_2^2), \quad (19)$$

where

$$\tilde{c} = nL_1^2\sigma_w^2\bar{b},\ c_3 = \tilde{d} + L_x(2^{1+\tilde{d}/2}L_1 + L_2) + \bar{\alpha}c',\ \hat{c} = 2L_1^2n\bar{b}, \quad (20)$$

$$c' = 2^{1+\tilde{d}/2}L_1||\mathscr{L}(A_{\max})||_2 + \bar{\alpha}(L_2||\mathscr{L}(A_{\max})||_2^2 + \hat{c}), \quad (21)$$

$$\bar{b} = \max_i \sum_{j=1}^{n}(b^{i,j})^2 + \sigma_b^{i,j}, \quad (22)$$

$\underline{\alpha} = \min_{1\le t\le T}\alpha_t$, $\tilde{d} = 0$ if $\bar{d} = 0$, or $\tilde{d} = 1$ if $\bar{d} > 0$.

Theorem *2:* Let the conditions **A1**, **A2** be satisfied, $0 < \alpha_t \le \bar{\alpha}$, in the averaged discrete system (14) the $\frac{\varepsilon}{4}$-consensus is achieved for time T, and for constants c_1, c_2 from Theorem 1 the following estimate holds

$$c_1\tau_T e^{c_2\tau_T^2}\bar{\alpha} \le \frac{\varepsilon}{4}, \quad (23)$$

then the mean square ε-consensus in the stochastic discrete system (11) at time T is achieved.

Consider the important case where $\forall i \in N\ f^i(x,u) = u$ and $\alpha_t = \alpha = const$. In this case the discrete averaged system (14) has the form:

$$\bar{Z}_{t+1} = (I - ((I-U) - \mathscr{L}(\alpha A_{\max})))\bar{Z}_t. \quad (24)$$

Theorem *3:* If conditions **A2**, **A3** are satisfied, $\alpha_t = \alpha > 0$, $f^i(x,u) = u$ for any $i \in N$, and condition $\alpha < \frac{1}{d_{\max}}$ for matrix $A_{\max}$ is satisfied, **then** the asymptotic mean square ε-consensus in the averaged discrete system (24) is achieved.

In addition, if the $\frac{\varepsilon}{4}$-consensus is achieved for the time $T(\frac{\varepsilon}{4})$ in the averaged discrete system (24) and there exists $T_0 > T(\frac{\varepsilon}{4})$ for which the parameter α provides the condition

$$\bar{C}_1 e^{\bar{C}_2}\alpha \le \frac{\varepsilon}{4}, \quad (25)$$

where

$$\bar{C}_1 \equiv 8n\left(\tilde{c} + \hat{c}(\frac{\alpha^2\tilde{c}}{c_3} + ||\bar{X}_0||^2)e^{T\ln(c_3+1)}\right)\tau_t,$$

$$T(\frac{\varepsilon}{4}) < T < T_0$$

$$\bar{C}_2 \equiv 2^{2-\bar{d}}\alpha^2||\mathscr{L}(A_{\max})||_2^2,\ \tilde{c} = n^2\bar{b}^2\sigma_w^2,\ \hat{c} = 2n(n-1)\bar{b}^2\tau_t^2,$$

$$c_3 = 2^{1+\tilde{d}} + 2\alpha^2(||\mathscr{L}(A_{\max})||_2^2 + \hat{c}),$$

where $\tilde{d} = 0$ if $\bar{d} = 0$ or $\tilde{d} = 1$ if $\bar{d} > 0$, **then** the mean square ε-consensus at time $t: T(\frac{\varepsilon}{4}) \le t \le T$ in the stochastic discrete system (11) is achieved.

Note that in [23], under certain assumptions similar to the conditions of Theorem 3, the necessary and sufficient condition for achieving the mean square consensus in case when the step sizes α_t tend to zero and the second term of (3) has a simple form: $f^i(x_t^i, u_t^i) = u_t^i$ were proved. However, in the analysis above, the more general case of the form of functions $f^i(x_t^i, u_t^i)$ and step sizes α_t nondecreasing to zero has been considered.

IV. The Load Balancing Problem

To demonstrate the application of the results derived in the previous section, the load balancing problem is considered in this section.

A. Problem Statement

In recent years, distributed parallel computing systems have been increasingly used. For such systems the problem of separating a package of jobs among several computing devices is important. Similar problems arise also in transport, logistics and production networks.

We consider a system that separates the same type of jobs among different agents, for parallel computing or production with feedback. Let $N = \{1, \ldots, n\}$ be a set of intelligent agents, each of which serves the incoming requests using a first-in-first-out queue. Jobs may be received at different times and by different agents.

At any time instant t, the state of agent i, $i \in N$, is described by two characteristics:

- q_t^i is the queue length of the atomic elementary jobs of the agent i at time t;
- p_t^i is the productivity of the agent i at time t.

The dynamics of each agent are described by

$$q_{t+1}^i = q_t^i - p_t^i + z_t^i + u_t^i;\ i \in N,\ t = 0, 1, \ldots, T, \quad (26)$$

where z_t^i is the new job received by agent i at time t, u_t^i is the result of jobs redistribution between agents, which is obtained by using the selected protocol of jobs redistribution. In the dynamics we assume that $\sum_i u_t^i = 0$, $t = 0, 1, 2, \ldots$.

We assume, that each agent $i \in N$ at time t can receive the following information to form the update strategy:

- noisy observations about its queue length

$$y_t^{i,i} = q_t^i + w_t^{i,i}, \quad (27)$$

- noisy and delayed observations about its neighbors queue length, if $N_t^i \ne \emptyset$

$$y_t^{i,j} = q_{t-d_t^{i,j}}^j + w_t^{i,j},\ j \in N_t^i, \quad (28)$$

  where $w_t^{i,j}$ are noises, $0 \le d_t^{i,j} \le \bar{d}$ is an integer-valued delay, and $\bar{d}$ is a maximal delay,
- information about its productivity p_t^i and about its neighbors productivities p_t^j, $j \in N_t^i$.

Let the fraction $\frac{q_t^i}{p_t^i}$ denote *the load* of agent i at time t, and T_t denote the overall implementation time of jobs at time t, where

$$T_t = \max_{i\in N}\frac{q_t^i}{p_t^i}. \quad (29)$$

The objective is to balance the load such that the overall implementation time can be minimized.

B. Analytical Results

To achieve the goal it is natural to use a redistribution protocol for jobs over time. Let's consider a stationary case where all jobs come to different agents at the initial time and no new job is received later. For this case, we have the following results.

Lemma *1: **(about the optimal update strategy)*** For the stationary case, among all possible options for redistributing jobs, the minimum completion time is achieved when

$$q_t^i/p_t^i = q_t^j/p_t^j,\ \forall i,j\ \in N. \tag{30}$$

Corollary *1:* If we take $x_t^i = q_t^i/p_t^i$ as the state of agent i in a dynamic network, then the goal — to achieve consensus in the network — will correspond to the optimal job redistribution between agents in the stationary case.
Note, that the optimality is understood in sense that if no new tasks arrives, all agents will finish at the same time.

These above results imply that the load balancing problem can essentially be treated as a consensus problem, i.e., how to keep the load equal among all agents in the network. We highlight that for this special case the model (26) corresponds to the difference equation (3).

Based on this intuition, we extend to the more general case where new jobs may arrive to any of the n agents at any time t. Specifically, consider the protocol (8), where $\forall\ i \in N,\ \forall\ t$ denote $\bar{N}_t^i = N_t^i$ and $b_t^{i,j} = p_t^j/p_t^i,\ j \in N_t^i$. Here, we assume that $p_t^i \neq 0\ \forall\ i$. Then, the dynamics of the load-balancing system (26) with local voting protocol (8) is as follows:

$$x_{t+1}^i = x_t^i - 1 + z_t^i/p_t^i + \alpha_t \sum_{j \in N_t^i} b_t^{i,j}(y_t^{i,j}/p_t^j - y_t^{i,i}/p_t^i). \tag{31}$$

where α_t are step sizes of the protocol, $y_t^{i,j}$ are noisy and delayed observation about j-th agents queue length, z_t^i is the new job received by agent i at time t.

It is important to say that we suggest the protocol which defines the intentions of agents, but in practice, when the protocol is implemented, it performs through a local coordination between all agents. As a result of this coordination, packages will not be lost. Specifically, the use of protocol (8) is justified in practice if additional assumptions are satisfied. First of all, it is assumed that the corresponding input data are exchanged instantaneously. In particular, in practice for the problem of load balancing in decentralized network, the protocol (8) requires additional coordinations of the sizes of the packages transmitted between the agents. Neither “overconsumption” nor “underconsumption” is permitted when relocating resources (or tasks) between the nodes. Additionally, one has to verify coordination of package forwarding since various collisions could occur due to the delays and noise. Additional checks and coordinations between the neighbors should be executed to satisfy this condition. With the use of the local voting protocol, each node determines how many tasks it can “give away” or “receive”. Then the nodes, which are ready to accept tasks, send requests to their neighbors about the amounts they are ready to give at the given time instant. Each “receiving” node sends in response to these requests a confirmation of how much task it can accept from one or another node and coordinates this amount with it (generally, each node orients to its current values $\tilde{u}_t^i$ recommended by the local voting protocol). It is assumed that the procedure of task coordination and transmission needs much less time than one cycle of the dynamic system. For example, in the problem of order allocation for cargo transportation the tomorrows tasks are coordinated at night upon completion of the current workday. In Section IV-C the simulation results were implemented as described above.

If the graph is balanced, for the general setting with random uncertainties in the measurements, in the network topology, and in the protocol (8), Theorem 3 allows to reduce the study of the dynamics of the load balancing system to the investigation of the corresponding averaged discrete model.

Theorem *4:* If $\alpha_t = \alpha = const$ is sufficiently small, the productivities stabilize over time: $\exists\ \mathrm{E}p_t^i = \bar{p}^i > 0,\ \forall\ i \in N$, $\mathrm{E}\frac{(w_t^{i,j})^2}{(p_t^i)^2} \leq \bar{\sigma}_w^2$, conditions **A2**, **A3** and the following condition for matrix $A_{\max}$ is satisfied

$$\alpha < \frac{1}{d_{\max}}, \tag{32}$$

then in the averaged discrete system the $\frac{\varepsilon}{4}$-consensus is achieved for the time $T(\frac{\varepsilon}{4})$, and there exists $T_0 > T(\frac{\varepsilon}{4})$ for which the parameter α ensures the condition

$$\bar{C}_1 e^{\bar{C}_2} \alpha \leq \frac{\varepsilon}{4}, \tag{33}$$

where

$$\bar{C}_1 \equiv 8n\left(\tilde{c} + \hat{c}(\frac{\alpha^2\tilde{c}}{c_3} + ||\bar{X}_0||^2)e^{T\ln(c_3+1)}\right)\tau_t,$$

$$\bar{C}_2 \equiv 2^{2-\tilde{d}}\alpha^2||\mathscr{L}(A_{\max})||_2^2,\ \tilde{c} = n(\sigma_w/\bar{p}^i)^2\bar{b},\ \hat{c} = 2n\bar{b}\tau_t,$$

$$c_3 = 2^{1+\tilde{d}} + 2\alpha^2(||\mathscr{L}(A_{\max})||_2^2 + \hat{c}),$$

$\tilde{d} = 0$ if $\bar{d} = 0$, or $\tilde{d} = 1$ if $\bar{d} > 0$, **then** in the stochastic discrete system for the n agents at time $t:\ T(\frac{\varepsilon}{4}) \leq t \leq T$, the ε-consensus is achieved.

We remark, that in Theorem 4, the conditions for productivities of agents are rather general. They hold for an adaptive problem statement, when information about the actual productivities is specified over time. In addition, due to the fact that the step sizes α_t of the protocol (8) do not tend to zero, the protocol considered shows good performance in the more general case. In a number of similar cases the validity of applying stochastic approximation update strategies with non-decreasing to zero step sizes in nonstationary problems could be theoretically proved (see, e.g., [25]).

C. Simulation Results

1) The six-node case: To show the convergence to consensus and to compare the initial stochastic system with the averaged model, we give an example of simulation for a computer network consisting of six computing agents.

Fig. 1 (left) shows the network, indicating the possible communication links, some of which may be “closed” and “opened up” over time. The network topology is random at any time t, and particularly, Link 1-3 or 1-2 appears with probability $1/2$ (Fig. 1 (right)).

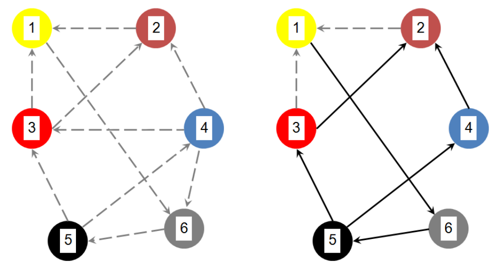


Fig. 1. Maximal set of communication links $E_{\max}$ (left); Network topology at time t (right).

In the case of uniformly distributed delays in the measurements, where the integer-valued delay d_t^{ij} equals 0 or 1 with probability 1/2, $\bar{d} = 1, p_0^{ij} = p_1^{ij} = 1/2$, we extend the state space:

$$\bar{X}_t = [x_t^1, \ldots, x_t^n, x_{t-1}^1, \ldots x_{t-1}^n] \in \mathbb{R}^{2n}. \quad (34)$$

Matrix G of the corresponding averaged discrete model (14) is as follows:

$$G = \begin{pmatrix} \frac{1}{2}H\alpha & \frac{1}{2}H\alpha \\ 0 & 0 \end{pmatrix}, \quad (35)$$

where

$$H = \begin{pmatrix} 0 & \frac{1}{2}\frac{p^2}{p^1} & \frac{1}{2}\frac{p^3}{p^1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{p^4}{p^2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{p^5}{p^3} & 0 \\ 0 & 0 & 0 & 0 & \frac{p^5}{p^4} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{p^6}{p^5} \\ \frac{p^1}{p^6} & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (36)$$

We set the initial queue lengths and the productivities of agents, and assume that the productivities of nodes do not change over time. In addition, we highlight that the information about the queue lengths is measured with random noise and delays.

We consider two cases, the special case and the general case, as discussed in the previous subsection. We use constant step size $\alpha_t = \alpha = 0.1$. The dynamics of the agents x_t^i with local voting protocol (31) is shown in Fig. 2 and 4.(a) respectively.

Fig. 2 shows how the system operates in the special case when there are no new incoming jobs during the system work (only the initial load). Each line, corresponding to one node, indicates how the load x_t^i evolves over time. These lines also show how the system evolve to reach load-balancing or consensus.

Now we estimate the time to consensus. We calculate eigenvalues and obtain that $|Re(\lambda_2)| = 0.7737$. By known formula

$$T(\varepsilon) = \frac{1}{2Re(\lambda_2)} \ln\left(\frac{(n-1)||x_0 - x^\star \underline{1}||^2}{\varepsilon}\right), \quad (37)$$

we can calculate $T(\varepsilon)$ for continuous system. If $\varepsilon = 0.1$ then $T(\varepsilon) = 12.8883$. If $\varepsilon = 1$ then $T(\varepsilon) = 11.4003$. The corresponding values are marked on Fig. 2.

To support the claim that we can use the averaged model to study our initial stochastic system, Fig. 3 is presented. The

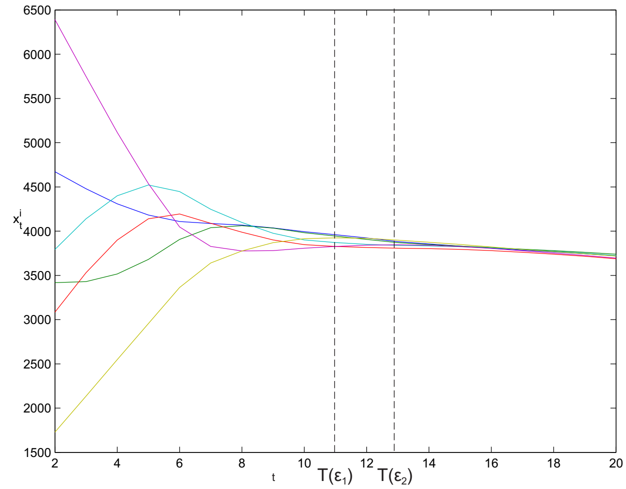


Fig. 2. Dynamics of the agents x_t^i at the start and time to consensus

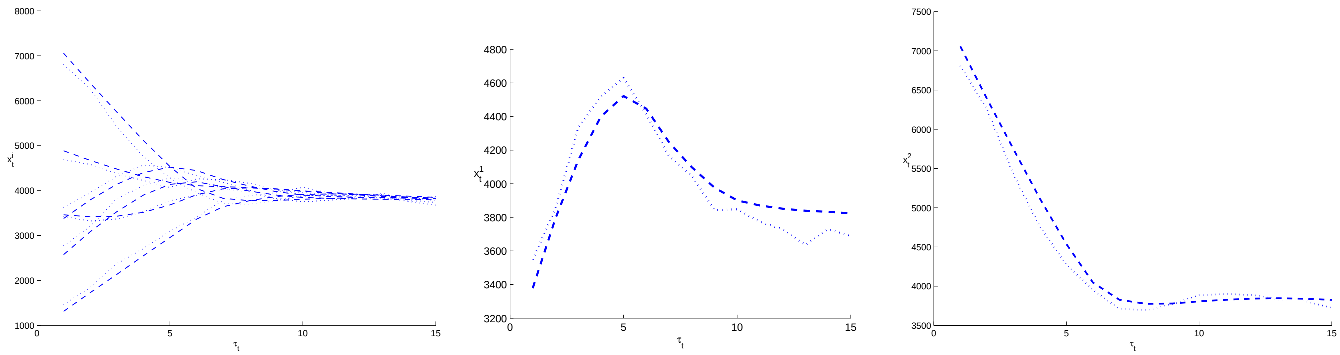
figure compares the dynamics of algorithm (8) and that of the averaged model described in Sec. III. Fig. 3 shows that trajectories of the stochastic discrete system (dotted lines) are close with the limiting trajectories of the average system (dashed lines).

To characterize the quality of the protocol (8) in terms of convergence of trajectories to consensus $x^\star$, we use the average residual, defined as $Err = \sqrt{\sum_i \frac{(x_t^i - x^\star)^2}{n}}$.

Fig. 4.(a) shows the dynamics of the system in a more general 6-node case where new jobs can come to different agents during the system work. New jobs arrive at a random node at random times. Specifically, Fig. 4.(a) indicates how the system tries to reach consensus using the local voting protocol (8) when there are new incoming jobs. In addition, the quality of the protocol (8) is indicated by Fig. 4.(b), where the corresponding evolvement of average residuals is displayed. It shows, how the average residual changes over time: it rapidly reduces and retains at low level until new jobs received, and then it reduces again. The simulation results shows the good performance of the protocol (8) in general case. This is explained by the properties of the stochastic approximation type algorithm with non-decreasing step, since each time instant when new jobs received might be considered as an initial time instant. In a number of similar cases the validity of applying stochastic approximation update strategies with non-decreasing to zero step sizes in nonstationary problems could be theoretically proved (see, e.g., [25]).

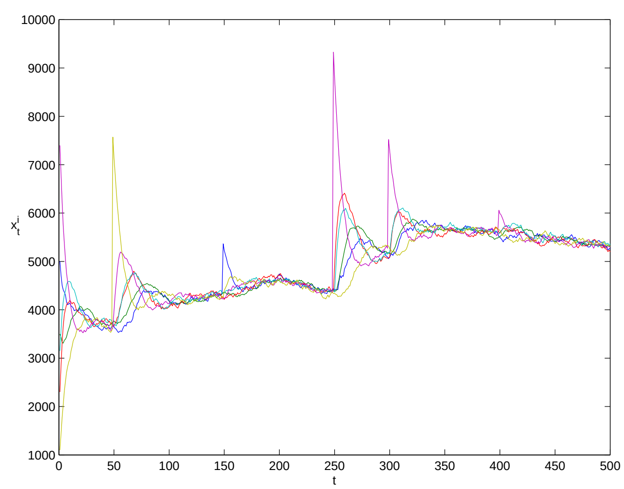
In Fig. 5 there are graphs for the average residuals with using of different parameters of step sizes α. In first four figures we used constant step sizes. It could be seen that if we increase the step size then the time to consensus will decrease until reaching a certain level. However, if we use the decreasing step size ($\alpha_t = 1/t$) as shown in the last figure in Fig. 5 then the convergence rate decreases with time.

Generally, when we reach the certain level of accuracy we do not need to run protocol any more. However, the simulation results show, that even if we run in beyond time T, the mean square error does not increase.

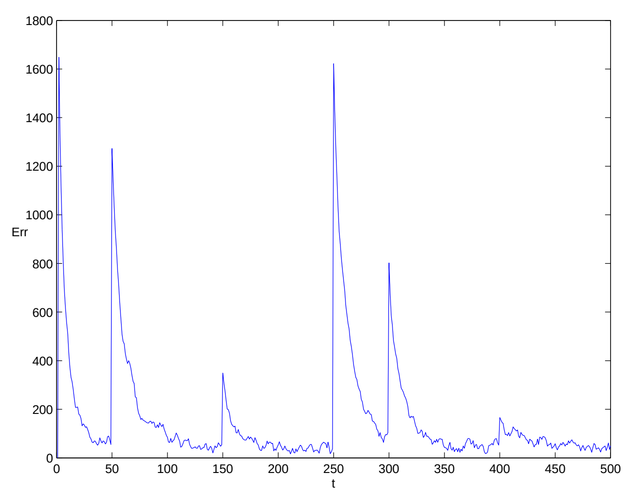


(a) All nodes (b) For the first node (c) For the second node

Fig. 3. Comparison of trajectories of the stochastic discrete system and its averaged model



(a) Dynamics of the agents x_t^i



(b) Convergence to consensus

Fig. 4. The 6-node general case

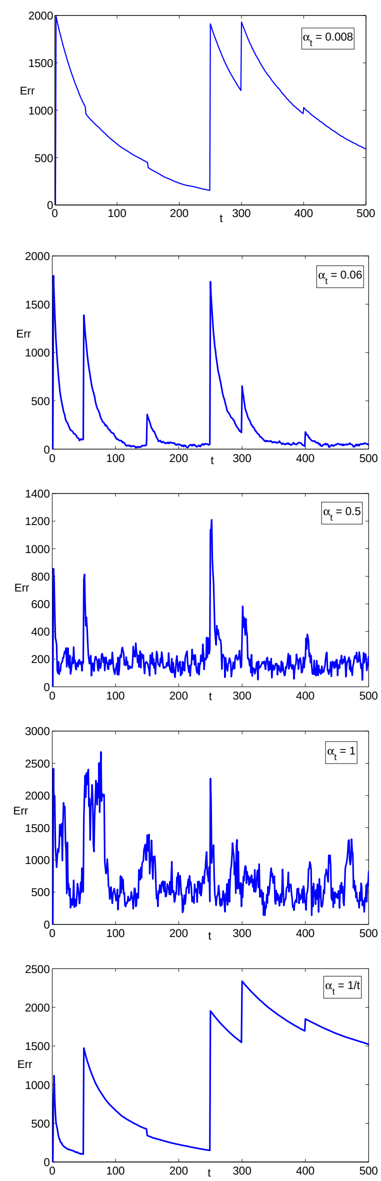


Fig. 5. Convergence to consensus with different step sizes.

2) The 1024-node case: To show how well the approach works to achieve load balancing and the advantage of redistribution of jobs in a larger network, we consider a network of 1024 agents. The focus here is to compare the performance of the system adopting the local voting protocol (8) to redistribute load with that without load-redistribution.

In the simulation, the time between events in the input stream is exponentially distributed with parameter $d_{in} = 1/3000$, and the normalized "complexities" of jobs are also exponentially distributed with parameter $d_p = 1$ (where, the normalized "complexity" of job is referred to as the time, required to perform the job on a single agent with productivity $p = 1$). The number of incoming jobs is 10^6. The choice of an agent, which receives the next job is performed randomly by the uniform distribution of 1024 agents.

Agents are connected in a circle. In addition, there are n random connections between agents on each iteration, that change over time. An example snapshot of the network is shown in Fig. 6.

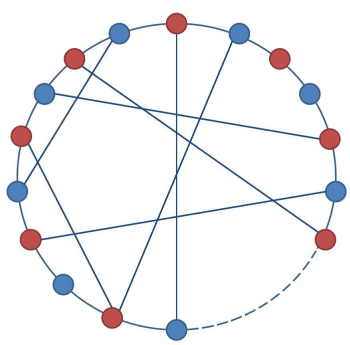
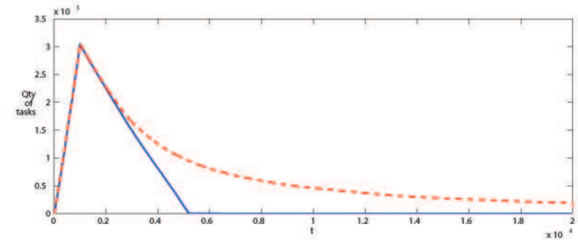
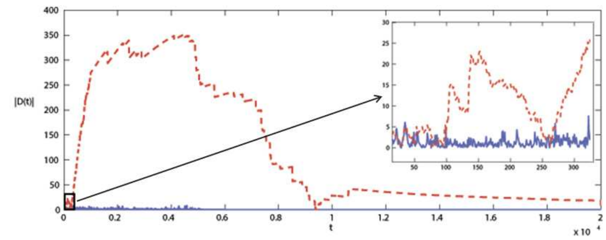


Fig. 6. An example of network topology.

Consider the case when all tasks arrive at different time instants in the interval from 1 to 2000. Fig. 7 shows typical results of simulations. In these figures, solid lines correspond to the case with redistribution of jobs by the local voting protocol, and dashed lines — to the case without redistribution, where symbol $|D(t)|$ stands for the maximum deviation from the average load on the network. Fig. 7 shows that the performance of the adaptive multi-agent strategy with the redistribution of jobs among "connected" neighbors is significantly better than the performance of the strategy without redistribution.



(a) The number of jobs in the queue in case, when all jobs arrive at different time instants.



(b) Maximum deviation from the average load on the network.

Fig. 7. The 1024-node general case

V. Conclusion

In this paper, the approximate consensus problem statement of multi-agent stochastic system with nonlinear dynamics, noise, delays and switched topology was introduced. In contrast to the existing stochastic approximation-based update algorithms (protocols), the local voting protocol with nonvanishing step size was proposed. Nonvanishing (e.g., constant) step size ensures better transients in the time-invariant case and provides bounded error in the case of time-varying loads and agent states. The price to pay is replacement of the mean square convergence with an approximate one.

Analytic conditions for approximate consensus in stochastic network with noise, delays and switched topology were proposed. These conditions are based on the method of averaged models. This method allows to reduce the complexity of the closed loop system analysis. In this paper, upper bounds for the mean square distance between the initial system and its approximate average model were proposed. The proposed upper bounds were used to obtain conditions for approximate consensus achievement. In contrast to our previous works, we relaxed the assumption of the weights boundedness of the protocol replacing it by the boundedness of its variances.

The theoretical results were applied to the load balancing problem in a stochastic network. Theoretical results were confirmed analytically and by simulation. Large size simulation experiments were performed for a stochastic computer network. They showed that the performance of the adaptive multi-agent strategy with the redistribution of jobs among "connected" neighbors is significantly better than the performance of the strategy without redistribution.

Appendix A
Proof of Theorem 1

Proof: The following facts will be useful, for the remainder.

Proposition *1:* For $\bar{z} \in \mathbb{R}^n$ and matrix $A_{\max}$ the following inequality holds

$$\sum_{i=1}^{n} \Big(\sum_{j \in N^i_{\max}} a^{i,j}_{\max} z^j \Big)^2 \le ||A_{\max}||_2^2 ||\bar{z}||^2. \quad (38)$$

Proof: Using the Cauchy-Schwarz inequality we obtain

$$\sum_{i=1}^{n} \Big(\sum_{j \in N^i_{\max}} a^{i,j}_{\max} z^j \Big)^2 \le \sum_{i=1}^{n} \Big(\sum_{j \in N^i_{\max}} {a^{i,j}_{\max}}^2 \Big) \Big(\sum_{j \in N^i_{\max}} {z^j}^2 \Big) \le \quad (39)$$

$$\le \Big(\sum_{i=1}^{n} \sum_{j=1}^{n} {a^{i,j}_{\max}}^2 \Big) \Big(\sum_{j=1}^{n} {z^j}^2 \Big) \le ||A_{\max}||_2^2 ||\bar{z}||^2.$$

■

Proposition *2:*

$$||\bar{s}(\bar{z})||^2 \leq 2||\mathscr{L}(A_{\max})||_2^2||\bar{z}||^2. \quad (40)$$

Proof: Using the result of Proposition 1 yields

$$||\bar{s}(\bar{z})||^2 = \sum_{i=1}^{n}(\sum_{j=1}^{n} a_{\max}^{i,j}(z^j - z^i))^2 \leq \sum_{i=1}^{n}(d^i(A_{\max})|z^i| + \quad (41)$$

$$+|\sum_{j\in N_{\max}^i} a_{\max}^{i,j} z^j|)^2 \leq 2(\sum_{i=1}^{n} d^i(A_{\max})^2 + ||A_{\max}||_2^2)||\bar{z}||^2 =$$

$$2||\mathscr{L}(A_{\max})||_2^2||\bar{z}||^2.$$

■

Proposition *3:* If **A2** is satisfied then $s^i(\bar{x}) = \frac{1}{\alpha_t}\mathrm{E}_{\mathscr{F}_{t-1}} u_t^i$ and the following inequality holds

$$\frac{1}{\alpha_t^2}\mathrm{E}_{\mathscr{F}_{t-1}} {u_t^i}^2 \leq (||\bar{x}_t - x_t^i \underline{1}||^2 + 2\sigma_w^2)\bar{b},\ i \in N. \quad (42)$$

Proof: By the definition of the protocol (8)

$$\frac{1}{\alpha_t} u_t^i = \sum_{j\in\bar{N}_t^i} b_t^{i,j}((x^j - x^i) + (w_t^{i,j} - w_t^{i,i})). \quad (43)$$

It follows from conditions **A2** that $s^i(\bar{x}) = \frac{1}{\alpha_t}\mathrm{E}_{\mathscr{F}_{t-1}} u_t^i$.

By the centrality of observation noise (on condition **A2a**) we consecutively derive

$$\frac{1}{\alpha_t^2}\mathrm{E}_{\mathscr{F}_{t-1}} {u_t^i}^2 = \mathrm{E}_{\mathscr{F}_{t-1}}(\sum_{j\in\bar{N}_t^i} b_t^{i,j}((x_t^j - x_t^i) + (w_t^{i,j} - w_t^{i,i})))^2 = \quad (44)$$

$$= \mathrm{E}_{\mathscr{F}_{t-1}}(\sum_{j\in\bar{N}_t^i} b_t^{i,j}(x_t^j - x_t^i))^2 + \mathrm{E}_{\mathscr{F}_{t-1}}(\sum_{j\in\bar{N}_t^i} b_t^{i,j}(w_t^{i,j} - w_t^{i,i}))^2 \leq$$

$$\leq ||\bar{x}_t - x_t^i\underline{1}||^2 \mathrm{E}_{\mathscr{F}_{t-1}} \sum_{j\in\bar{N}_t^i}(b_t^{i,j})^2 + \mathrm{E}_{\mathscr{F}_{t-1}} \sum_{j\in\bar{N}_t^i} {b_t^{i,j}}^2({w_t^{i,j}}^2 +$$

$$+{w_t^{i,i}}^2) \leq (||\bar{x}_t - x_t^i\underline{1}||^2 + 2\sigma_w^2)\bar{b}.$$

■

Proposition *4:*

$$||U\bar{X}||^2 \leq 2^{\bar{d}}||\bar{X}||^2,\ \ldots,\ ||U^{\bar{d}}\bar{X}||^2 \leq 2^{\bar{d}}||\bar{X}||^2,\ \ldots,\ ||U^k\bar{X}||^2 \leq \quad (45)$$

$$\leq 2^{\bar{d}}||\bar{X}||^2,$$

Proof: By the definition of matrix U it is easy to obtain the first inequality, and the rest we get by induction on k and by the following equality

$$\forall k > \bar{d}\ \ U^k = U^{\bar{d}} = \begin{pmatrix} I & 0 & 0 & \ldots & 0 \\ I & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ I & 0 & 0 & \ldots & 0 \end{pmatrix}. \quad (46)$$

■

Denote

$$v_t = F(\alpha_t, \bar{X}_t, \bar{w}_t) - G(\alpha_t, \bar{X}_t). \quad (47)$$

Proposition *5:* By assumptions **A2** the following inequality holds

$$\mathrm{E}\max_{1\leq t\leq T}||\sum_{i=1}^{t} v_t||^2 \leq 4n\sum_{t=1}^{T}\mathrm{E}||v_t||^2. \quad (48)$$

Proof: Under the conditions **A2** random elements v_t are martingale differences, i.e., they are centered with respect to the conditional averaging of the background: $\mathrm{E}_{\mathscr{F}_{t-1}} v_t = 0$. So, Lemma 1 from section 3 of [43] is applicable. The dimension of vectors v_t is $n\bar{d}$, but since only the first n components of vectors v_t are nonzero, then it is possible to use in the estimation the value of n instead of $n\bar{d}$. ■

Proposition *6:* Let the sequence of numbers $\mu_t \geq 0$, $t = 0, 1, \ldots, T$ satisfies the inequalities

$$\mu_{t+1} \leq \bar{\alpha} c_1 \tau_t + c_2 2^{\bar{d}} \tau_t \sum_{k=1}^{t} \alpha_k \mu_k,\ c_1, c_2 \geq 0, \quad (49)$$

then

$$\mu_t \leq c_1 \tau_t e^{c_2 \tau_t^2}\bar{\alpha}. \quad (50)$$

Proof: Statement of Proposition follows directly from discrete Gronwall's inequality (or from the corresponding result in [44]). ■

Proposition *7:* By assumptions **A1**, **A2** yields

$$\mathrm{E}||\bar{X}_t||^2 \leq (\frac{2nL_2 + \bar{\alpha}^2\tilde{c}}{c_3} + ||\bar{X}_0||^2)e^{t\ln(c_3+1)}. \quad (51)$$

Proof: We write equation (11) as

$$\bar{X}_{t+1} = U\bar{X}_t + G(\alpha_t, \bar{X}_t) + v_t. \quad (52)$$

For the squared norm of $\bar{X}_{t+1}$ we have

$$||\bar{X}_{t+1}||^2 = ||U\bar{X}_t + G(\alpha_t,\bar{X}_t)||^2 + 2(U\bar{X}_t + G(\alpha_t,\bar{X}_t))^{\mathrm{T}} v_t + ||v_t||^2. \quad (53)$$

Taking the conditional expectation of both parts of (53) on σ-algebra $\mathscr{F}_{t-1}$ (i.e., for fixed $\bar{X}_t$) by the centrality of v_t we obtain

$$\mathrm{E}_{\mathscr{F}_{t-1}}||\bar{X}_{t+1}||^2 = ||U\bar{X}_t + G(\alpha_t,\bar{X}_t)||^2 + \mathrm{E}_{\mathscr{F}_{t-1}}||v_t||^2 \leq$$

$$\leq 2||U\bar{X}_t||^2 + 2||G(\alpha_t,\bar{X}_t)||^2 + \mathrm{E}_{\mathscr{F}_{t-1}}||v_t||^2. \quad (54)$$

By the form of v_t and Lipschitz property in u of functions $f^i(u)$ (by **A1**), for $||v_t||^2$ we have

$$||v_t||^2 = \sum_{i\in N}|f^i(x_t^i, \alpha_t \sum_{j\in\bar{N}_t^i} b_t^{i,j}(x^j_{t-d_t^{i,j}} - x_t^i + w_t^{i,j} - w_t^{i,i})) - \quad (55)$$

$$-f^i(x_t^i, \alpha_t s_t^i(\bar{X}_t))|^2 \leq L_1^2||\bar{u}_t - \alpha_t\bar{s}_t||^2.$$

Under the conditions **A2**, random variables $\mathrm{E}_{\mathscr{F}_{t-1}} u_t^i$, $i \in N$ satisfy the conditions of Proposition 3

$$\mathrm{E}_{\mathscr{F}_{t-1}}||v_t||^2 = \alpha_t^2 L_1^2(2n\bar{b}||\bar{X}_t||^2 + 2n^2\bar{b}\sigma_w^2). \quad (56)$$

Consistently evaluating all three summands on the right hand side of (54) and taking into account the results of Propositions 4, 2 and 3, we deduce

$$\mathrm{E}_{\mathscr{F}_t}||\bar{X}_{t+1}||^2 \leq 2^{\bar{d}}||\bar{X}_t||^2 + 2^{1+\bar{d}/2}||\bar{X}_t||L_1(L_x||\bar{X}_t|| + \alpha_t||\bar{s}||) + \quad (57)$$

$$+L_2(nL_c + L_x||\bar{X}_t||^2 + \alpha_t^2||\bar{s}||^2) + \alpha_t^2 L_1^2(2n\bar{b}||\bar{X}_t||^2 +$$

$$+2n\bar{b}\sigma_w^2) \leq (2^{\bar{d}} + 2^{1+\bar{d}/2}L_1L_x + L_2L_x + \alpha_t 2^{1+\bar{d}/2}L_1||\mathscr{L}(A_{\max})||_2 +$$

$$+\alpha_t^2(L_2||\mathscr{L}(A_{\max})||_2^2 + 2nL_1^2\bar{b}))||\bar{X}_t||^2 + nL_2L_c +$$

$$+2\alpha_t^2 nL_1^2\bar{b}\sigma_w^2 \leq \bar{c} + \bar{c}_3||\bar{X}_t||^2,$$

where $\bar{c} = nL_2L_c + \alpha_t^2\tilde{c}$, $\bar{c}_3 = c_3 + 1$.

By taking unconditional expectation of both parts of this inequality and consistently iterating on t, we obtain Proposition 7

$$\mathrm{E}||\bar{X}_t||^2 \le \bar{c} + \bar{c}_3\mathrm{E}||\bar{X}_{t-1}||^2 \le \bar{c} + \bar{c}\bar{c}_3 + \bar{c}_3^2\mathrm{E}||\bar{X}_{t-2}||^2 \le \quad (58)$$

$$\le \bar{c}(1 + \bar{c}_3 + \bar{c}_3^2 + \ldots + \bar{c}_3^{t-1}) + \bar{c}_3^t||\bar{X}_0||^2 \le \bar{c}\frac{\bar{c}_3^t - 1}{c_3} + \bar{c}_3^t||\bar{X}_0||^2 \le$$

$$\le \left(\frac{\bar{c}}{c_3} + ||\bar{X}_0||^2\right)\bar{c}_3^t \le (\bar{c}_4 + ||\bar{X}_0||^2)e^{t\ln\bar{c}_3},$$

where $\bar{c}_4 = \bar{c}/c_3$. ■

By condition **A2** averaging with respect to σ-algebras $\mathscr{F}_t^d$ and $\mathscr{F}_t$ yields $\mathrm{E}_{\mathscr{F}_t} v_t = 0$. By iterating equation (11) for $t, t-1, \ldots t-d+1$ we obtain

$$\bar{X}_{t+1} = U\bar{X}_t + G(\alpha_t, \bar{X}_t) + v_t =$$

$$= U^2\bar{X}_{t-1} + UG(\alpha_{t-1}, \bar{X}_{t-1}) + G(\alpha_t, \bar{X}_t) + Uv_{t-1} + v_t = \quad (59)$$

$$= \cdots = U^{t+1}\bar{X}_0 + \sum_{k=0}^{t} U^{t-k}G(\alpha_k, \bar{X}_k) + \sum_{k=0}^{t} U^{t-k}v_k.$$

Similarly we obtain

$$\bar{Z}_{t+1} = U^{t+1}\bar{X}_0 + \sum_{k=0}^{t} U^{t-k}G(\alpha_k, \bar{Z}_k). \quad (60)$$

Let us estimate $||\bar{X}_t - \bar{Z}_t||^2$, $t = 1, \ldots, T$. Denote $g_k = U^{t-k}(G(\alpha_k, \bar{X}_k) - G(\alpha_k, \bar{Z}_k))$. By subtracting (60) from (59) and squaring the result we obtain

$$||\bar{X}_t - \bar{Z}_t||^2 = ||\sum_{k=1}^{t} U^{t-k}v_k + \sum_{k=1}^{t} g_k||^2 \le$$

$$\le 2||\sum_{k=1}^{t} U^{t-k}v_k||^2 + 2||\sum_{k=1}^{t} g_k||^2 \le$$

$$\le 2||\sum_{k=1}^{t} U^{t-k}v_k||^2 + 2\frac{\tau_t}{2^{\bar{d}}}\sum_{k=1}^{t}\frac{1}{\alpha_t}||g_k||^2, \quad (61)$$

since

$$||\sum_{k=1}^{t}\sqrt{\alpha}_k\frac{1}{\sqrt{\alpha}_k}g_k||^2 \le \sum_{k=1}^{t}\alpha_k\sum_{k=1}^{t}\frac{1}{\alpha_k}||g_k||^2 = \frac{\tau_t}{2^{\bar{d}}}\sum_{k=1}^{t}\frac{1}{\alpha_t}||g_k||^2$$

For the summands in the second sum of (61) using Propositions 2, 4 and Lipschitz condition $f^i(\cdot,\cdot)$ (assumption **A1**) we obtain

$$||U^{t-k}(G(\alpha_k, \bar{X}_k) - G(\alpha_k, \bar{Z}_k))||^2 \le 2^{\bar{d}}L_1^2\sum_{i=1}^{n}(L_x|x_k^i - z_k^i| + \quad (62)$$

$$+\alpha_k|s(x_k^i) - s(z_k^i)|)^2 \le 2^{1+\bar{d}}L_1^2\sum_{i=1}^{n}L_x|x_k^i - z_k^i|^2 + \alpha_k^2 s(x_k^i - z_k^i)^2 \le$$

$$\le 2^{1+\bar{d}}L_1^2(L_x + 2\alpha_k^2||\mathscr{L}(A_{\max})||_2^2)||\bar{X}_k - \bar{Z}_k||^2$$

We take expectation of both parts of (61) and denote $\mu_T = \max_{0\le t\le T}\mathrm{E}||\bar{X}_t - \bar{Z}_t||^2$. By applying Proposition 5 to the first summand and obtained above estimate of the second summand we obtain

$$\mu_T \le 2^{3+\bar{d}}n\sum_{k=1}^{T}\mathrm{E}||v_k||^2 + 2\tau_T L_1^2\sum_{k=1}^{t}(\frac{L_x}{\underline{\alpha}} + 2\alpha_k||\mathscr{L}(A_{\max})||_2^2)\mu_k. \quad (63)$$

To estimate $\mathrm{E}||v_k||^2$ by using previously obtained relation (56) and the result of Proposition 7 we deduce

$$\mathrm{E}||v_k||^2 \le \alpha_k^2(\tilde{c} + \hat{c}(\bar{c}_4 + ||\bar{X}_0||^2)e^{k\ln(c_3+1)}) \quad (64)$$

and hence

$$2^{3+\bar{d}}n\sum_{k=1}^{T}\mathrm{E}||v_k||^2 \le \bar{\alpha}2^{\bar{d}}8n\tau_T(\tilde{c} + \hat{c}(\bar{c}_4 + ||\bar{X}_0||^2)e^{T\ln(c_3+1)}). \quad (65)$$

By the following relation $2^{\bar{d}}\sum_{k=1}^{t}\alpha_k^2 \le \bar{\alpha}2^{\bar{d}}\sum_{k=1}^{t}\alpha_k = 2^{\bar{d}}\bar{\alpha}\tau_t$, considering estimates (65) from (63), we have

$$\mathrm{E}\mu_T \le \bar{\alpha}c_1\tau_T + c_2\tau_T 2^{\bar{d}}\sum_{k=1}^{T}\alpha_k\mathrm{E}\mu_k. \quad (66)$$

From last inequality (66) by applying Proposition 6 we get the conclusion of Theorem 1. ■

APPENDIX B
PROOF OF THEOREM 2

Proof: Denote $x^\star$ as consensus value of discrete system (14). From the first group of conditions of Theorem 2 the conditions of Theorem 1 hold. From other conditions of Theorem 2 and the result of Theorem 1 we obtain

$$\mathrm{E}||\bar{X}_t - x^\star\underline{1}||^2 \le 2\mathrm{E}||\bar{X}_t - \bar{Z}_t||^2 + 2||\bar{Z}_t - x^\star\underline{1}||^2 \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \le \varepsilon. \quad (67)$$

■

APPENDIX C
PROOF OF THEOREM 3

Proof: The result of Theorem 3 is derived from Theorem 2.

All row sums of elements of the matrix $\bar{\mathscr{L}} = (I - U) - \mathscr{L}(\alpha A_{\max})$ are equal to zero and, moreover, all the diagonal elements are positive and equal to the absolute value of the sum of all the other elements in the row, which are negative. Hence the matrix $\bar{\mathscr{L}}$ is the Laplacian of a graph and a vector of 1's $\underline{1}$ is the right eigenvector corresponding to zero eigenvalue.

By condition **A3**, the graph corresponding to the Laplacian $\bar{\mathscr{L}}$ has a spanning tree. By condition **A3** graph of the first n nodes has a spanning tree. And units on $(n+1)$-th diagonal consistently connect $\bar{n}$-th node with $(\bar{n} - \bar{d})$-th node, $(\bar{n} - 1)$-th node with $(\bar{n} - \bar{d} - 1)$-th and so on. Hence the asymptotic consensus is achieved in such a discrete system since the condition $\alpha < \frac{1}{d_{\max}}$ holds by the assumptions of Theorem 3.

To satisfy the conditions of Theorem 2 it remains to show that the constants $\bar{C}_1$ and $\bar{C}_2$ are the same as the corresponding constants from Theorem 1. It follows from the fact that in this case $L_1 = L_2 = 1$, $L_x = L_c = 0$. ■

APPENDIX D
PROOF OF LEMMA 1

Proof: We take $x_t^i = q_t^i / p_t^i$ as the state of agent i.

We give the proof by contradiction. Let $x_t^1, \dots, x_t^n$ be the optimal redistribution among all possible options for the redistribution of jobs and denote T_t^{opt} as a corresponding minimum completion time. Assume that $k \in N$ is a maximizer, i.e. for some optimal strategy not all x_t^i are equal to each other, i.e., there is a agent with the number $k \in N$ and the subset of agents N_t' such that $x_t^k > x_t^j$, $\forall j \in N_t'$.

Denoted by $l = |N_t'|$ the number of agents in N_t'. The states of other $N\ N_t'$ agents equal x_t^k. Note, that $|N\ N_t'| = n - l$.

Let the difference between the state of k-th agent and the biggest of the state value of agents from the set N_t' be equal to ε_t, i.e.,

$$\varepsilon_t = x_t^k - \max_{j \in N_t'} x_t^j. \tag{68}$$

Let's consider the new strategy of job redistribution. Let the amount of redistributed load be equal to $u_t^i = -\frac{\varepsilon_t}{2(n-l)}$ for all $i \in N\ N_t'$ and $u_t^j = \frac{\varepsilon_t}{2}$ for some $j \in N_t'$. For the new obtained set of loads $\tilde{x}_t^1, \dots, \tilde{x}_t^n$ we have that the overall implementation time of jobs $\tilde{T}_t$ from (29) is less then T_t^{opt} on $\frac{\varepsilon_t}{2(n-l)}$, i.e. less then the minimum T_t^{opt} by the assumption. We get a contradiction. Hence, for optimal control strategy all x_t^i should be equal to each other.

■

APPENDIX E
PROOF OF THEOREM 4

Proof: You should verify if the conditions **A1**, **A2** for the considered protocol and functions $f^i(\cdot,\cdot)$ are satisfied. If they are satisfied, then all the conditions of Theorem 2 are satisfied and the result is valid for this case.

The condition **A1** holds since the function $f^i(x,u) = -1 + u$ is linear in u. The condition **A2** holds because of the formation rules for the weighting coefficients in the protocol and stabilization conditions for p_t^i.

■

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments. This work was supported by

REFERENCES

[1] R. Olfati-Saber, J. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[2] W. Ren, R. Beard, and E. Atkins, "Information consensus in multivehicle cooperative control," *Control Systems, IEEE*, vol. 27, no. 2, pp. 71–82, 2007.

[3] C. Wu, *Synchronization in complex networks of nonlinear dynamical systems*. World Scientific Publishing Company Incorporated, 2007.

[4] W. Ren and R. Beard, *Distributed consensus in multi-vehicle cooperative control: theory and applications*. Springer, 2007.

[5] F. Bullo, J. Cortés, and S. Martinez, *Distributed control of robotic networks: a mathematical approach to motion coordination algorithms*. Princeton University Press, 2009.

[6] P. Antsaklis and J. Baillieul, "Guest editorial special issue on networked control systems," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1421–1423, 2004.

[7] C. Abdallah and H. Tanner, "Complex networked control systems: introduction to the special section," *Control Systems, IEEE*, vol. 27, no. 4, pp. 30–32, 2007.

[8] D. Armbruster, A. Mikhailov, K. Kaneko *et al.*, *Networks of Interacting Machines: Production Organization in Complex Industrial Systems and Biological Cells*. World Scientific. Singapore, 2005.

[9] O. Granichin, P. Skobelev, A. Lada, I. Mayorov, and A. Tsarev, "Comparing adaptive and non-adaptive models of cargo transportation in multi-agent system for real time truck scheduling," in *Proc. 4th International Conference on Computational Intelligence. Evolutionary Computation Theory and Applications (ECTA2012)*, Barcelona, Spain, 2012, pp. 282–285.

[10] L. Zheng, Y. Yao, M. Deng, and S. S. Yau, "Decentralized detection in ad hoc sensor networks with low data rate inter sensor communication," *Information Theory, IEEE Transactions on*, vol. 58, no. 5, pp. 3215–3224, 2012.

[11] S. Kar, J. M. Moura, and K. Ramanan, "Distributed parameter estimation in sensor networks: Nonlinear observation models and imperfect communication," *Information Theory, IEEE Transactions on*, vol. 58, no. 6, pp. 3575–3605, 2012.

[12] H. Pezeshki-Esfahani and A. J. Heunis, "Strong diffusion approximations for recursive stochastic algorithms," *Information Theory, IEEE Transactions on*, vol. 43, no. 2, pp. 512–523, 1997.

[13] A. Weiss and D. Mitra, "Digital adaptive filters: Conditions for convergence, rates of convergence, effects of noise and errors arising from the implementation," *Information Theory, IEEE Transactions on*, vol. 25, no. 6, pp. 637–652, 1979.

[14] A. Benveniste, M. Métivier, and P. Priouret, *Adaptive algorithms and stochastic approximations*. Springer Publishing Company, Incorporated, 2012.

[15] M. Huang and J. Manton, "Coordination and consensus of networked agents with noisy measurements: stochastic algorithms and asymptotic behavior," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 134–161, 2009.

[16] R. Rajagopal and M. Wainwright, "Network-based consensus averaging with general noisy channels," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 373–385, 2011.

[17] L. Wang, Z. Liu, and L. Guo, "Robust consensus of multi-agent systems with noise," in *Proc. of the 26th Chinese Control Conference*, Zhangjiajie, Hunan, China, 2007, pp. 737–741.

[18] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *Automatic Control, IEEE Transactions on*, vol. 49, no. 9, pp. 1520–1533, 2004.

[19] W. Ren and R. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Transactions on Automatic Control*, vol. 50, no. 5, pp. 655–661, 2005.

[20] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Transactions on Automatic Control*, vol. 31, no. 9, pp. 803–812, 1986.

[21] T. Li and J. Zhang, "Mean square average-consensus under measurement noises and fixed topologies: Necessary and sufficient conditions," *Automatica*, vol. 45, no. 8, pp. 1929–1936, 2009.

[22] T. C. Aysal and K. E. Barner, "Convergence of consensus models with stochastic disturbances," *Information Theory, IEEE Transactions on*, vol. 56, no. 8, pp. 4101–4113, 2010.

[23] M. Huang, "Stochastic approximation for consensus: a new approach via ergodic backward products," *IEEE Transactions on Automatic Control*, vol. 57, no. 12, pp. 2994–3008, 2012.

[24] D. Derevitskii and A. Fradkov, "Two models for analysis the dynamics of adaptation algorithms," *Automation and Remote Control*, no. 1, pp. 59–67, 1974.

[25] A. Vakhitov, O. Granichin, and L. Gurevich, "Algorithm for stochastic approximation with trial input perturbation in the nonstationary problem of optimization," *Automation and Remote Control*, vol. 70, no. 11, pp. 1827–1835, 2009.

[26] V. Borkar, *Stochastic approximation: a dynamical systems viewpoint*. Cambridge University Press Cambridge, 2008.

[27] G. Yin, V. Krishnamurthy, and C. Ion, "Regime switching stochastic approximation algorithms with application to adaptive discrete stochastic optimization," *SIAM Journal on Optimization*, vol. 14, no. 4, pp. 1187–1215, 2004.

[28] T. Friedrich, T. Sauerwald, and D. Vilenchik, "Smoothed analysis of balancing networks," *Random Structures & Algorithms*, vol. 39, no. 1, pp. 115–138, 2010.

[29] H. Li, "Load balancing algorithm for heterogeneous p2p systems based on mobile agent," in *Electric Information and Control Engineering*

(ICEICE), 2011 International Conference on, Wuhan, China, 2011, pp. 1446–1449.

[30] M. Kechadi, I. Savvas *et al.*, "Dynamic task scheduling for irregular network topologies," *Parallel Computing*, vol. 31, no. 7, pp. 757–776, 2005.

[31] G. Cybenko, "Dynamic load balancing for distributed memory multiprocessors," *Journal of parallel and distributed computing*, vol. 7, no. 2, pp. 279–301, 1989.

[32] Y. Li and Z. Lan, "A survey of load balancing in grid computing," *Computational and Information science*, pp. 280–285, 2005.

[33] K. Gilly, C. Juiz, and R. Puigjaner, "An up-to-date survey in web load balancing," *World Wide Web*, vol. 14, no. 2, pp. 105–131, 2011.

[34] K. Sim and W. Sun, "Ant colony optimization for routing and load-balancing: survey and new directions," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 33, no. 5, pp. 560–572, 2003.

[35] O. Granichin and N. Amelina, "Simultaneous perturbation stochastic approximation for tracking under unknown but bounded disturbances," *IEEE Transactions on Automatic Control*, vol. 59.

[36] N. Amelina, A. Fradkov, and K. Amelin, "Approximate consensus in multi-agent stochastic systems with switched topology and noise," in *Proc. IEEE 2012 Multiconference on Systems and Control (MSC2012)*, Dubrovnik, Croatia, 2012, pp. 445–450.

[37] N. Amelina and A. Fradkov, "Approximate consensus in the dynamic stochastic network with incomplete information and measurement delays," *Automation and Remote Control*, vol. 73, no. 11, pp. 1765–1783, 2012.

[38] K. Amelin, N. Amelina, O. Granichin, and O. Granichina, "Multi-agent stochastic systems with switched topology and noise," in *Proc. 2012 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel & Distributed Computing (SNPD)*, Kyoto, Japan, 2012, pp. 438–443.

[39] N. Amelina, "Scheduling networks with variable topology in the presence of noise and delays in measurements," *Vestnik St. Petersburg University: Mathematics*, vol. 45, no. 2, pp. 56–60, 2012.

[40] L. Ljung, "Analysis of recursive stochastic algorithms," *IEEE Transactions on Automatic Control*, vol. 22, no. 4, pp. 551–575, 1977.

[41] H. Kushner, "Convergence of recursive adaptive and identification procedures via weak convergence theory," *IEEE Transactions on Automatic Control*, vol. 22, no. 6, pp. 921–930, 1977.

[42] S. Gershgorin, "On bounding the eigenvalues of a matrix," *Izv. Akad. Nauk. SSSR Otd Mat. Estest 1*, no. 6, pp. 749–754 749–754, 1931.

[43] I. Gihman and A. Skorohkod, *Stochastic Differential Equations*. Kiev: Nauk. dumka, 1968.

[44] S. Bernstein, "Stochastic difference equations and stochastic differential equations," *Sobr. soch. in 4-th Vol.*, vol. 4, pp. 484–542, 1964.

Natalia Amelina Natalia Amelina graduated from Department of Informatics, St. Petersburg State University, Russia in 2009, and received the Ph.D. degree from the Department of Theoretical Cybernetics, St. Petersburg State University, Russia, in 2012, in the area of network systems and control. She is currently a Postdoctoral Researcher of St. Petersburg State University, Russia. Her research interests include distributed control of multiagent systems, consensus problem, load balancing, cooperative control of UAVs, sensor and wireless networks.

PLACE PHOTO HERE

Alexander Fradkov Biography text here.

PLACE PHOTO HERE

Yuming Jiang Biography text here.

PLACE PHOTO HERE

Dimitrios J. Vergados Biography text here.