2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Continuous-Time Averaged Models of Discrete-Time Stochastic Systems: Survey and Open Problems

A.L.Fradkov

A survey of a number of papers devoted to continuous-time modeling of discrete-time stochastic systems is given. It is concluded that, although different approaches to averaged (approximate) models justifying are in use, the procedures of building the averaged (approximate) models are similar in different papers. In addition to the deterministic (ODE) model some stochastic continuous-time models described by SDE are introduced. A new result concerning evaluation of the ODE model accuracy over the infinite time interval under partial stability condition is presented. Applications in adaptation, optimization and control are discussed. [1]

1. Introduction

The employment of continuous-time models for analysis and synthesis of discrete-time stochastic systems has started in the 1970s. Hundreds of papers and a number of monographs [7, 20, 43, 44, 51, 11], concerning both application of the machinery and its justification have been published since then. However few authors attempt to compare and unify different lines of research. An additional problem is in that quite a number of the results were published in Russian, i.e. they are not well known in the West. An outstanding impact in the area was made by the celebrated paper by L.Ljung [48] that was later listed among 25 seminal papers of the 20th century in control [17]. Currently the paper [48] has got more than 650 citations. What is most impressive its citing rate is about 20 citations per year and it is not going to decrease till now.

In the present paper we survey several avenues of research in continuous-time modeling of discrete-time stochastic systems. Among them the method proposed in [18] and further developed in a few papers and in a monograph [20] has some peculiarities allowing to analyze algorithms with the gain sequence not tending to zero.

[1] A.L. Fradkov is with Institute of Problems in Mechanical Engineering, Russian Academy of Sciences, 61 Bolshoy ave V.O., St. Petersburg, 199178, Russia fradkov@mail.ru

The work was supported by the Russian Foundation for Basic Research (RFBR), grant 11-08-01218; by the program of basic research of OEMPPU RAS #2; by the Russian Federal Program "Cadres", contracts 16.740.11.0042, 14.740.11.0942.

2. Continuous-time model building

The method of averaging has a wide applicability in modern control system theory, dynamical systems theory, nonlinear mechanics, etc. [2, **?**]. The essence of the method is in separation of slow and fast components of system motion, followed by averaging out the fast motion effects. The formal analysis of the technique for continuous-time systems one can find e.g. in [**?**, 55] (for deterministic case) and in [55, 63] (for stochastic case).

A specific form of averaging for discrete-time stochastic systems was developed in [18, 20] and, independently in [48] and then applied to various problems in identification and adaptive control. Below the scheme of [18, 20, 48] is described.

Consider a discrete-time stochastic system

$$x_{k+1} = x_k + \gamma_k F(x_k, f_k), \qquad k = 0, 1, 2, \ldots, \quad (1)$$

where $x_k \in \mathscr{R}^n$ — state vector, $f_k \in \mathscr{R}^m$ — random disturbance vector, γ_k — gain parameter. Create the averaged continuous system (continuous model)

$$\frac{dx}{dt} = A(x), \quad (2)$$

where $A(x) = \lim_{k\to\infty} EF(x, f_k)$ (the existence of the limit is assumed). Typical relationships between the discrete-time system and its continuous model are as follows.

1. If the gains γ_k are sufficiently small $(\gamma_k \le \gamma)$ then the trajectories $\{x_k\}$ of (1) are close to the trajectories of (2) $\{x(t_k)\}$, where $t_k = \gamma_0 + \cdots + \gamma_{k-1}$.

2. If the gains γ_k tend to zero as $k \to \infty$ then some asymptotic properties of the solutions of (1) (e.g. stability, ultimate boundedness, etc.) may be similar to those of the solutions of the continuous model (2).

In the case of similarity between (1) and (2) in the above sense one can use simplified model (2) instead of (1) for the purposes of system analysis and design.

Such an approach was called *the method of continuous models* [18, 20], *the ODE approach* [48] or *the Derevitskii-Fradkov-Ljung (DFL) scheme* [31]. Below the term 'method of continuous models' will be used since it takes into account two aspects:

— averaging is not the only way of the model generating (in some cases there is a similarity between (1) and (2) even for nonstochastic disturbances f_k investigated in [20]);

— one can use different types of models (e.g., stochastic differential equations).

3. Continuous-time model justifying

A number of rigorous results are known justifying applicability of continuous models for sufficiently small gains γ_k. Small value of the gains is prerequisite of separation of motions in system. It implies that the disturbance f_k changes faster than the system state x_k. The standard condition of averaging is weak dependence of f_k and f_s for large $|k-s|$ (e.g. independence of f_k and f_s when $k \neq s$).

Probably the first results on justifying the averaging for discrete stochastic systems in control theory belong to Meerkov [52], who used discrete averaged model

$$z_{k+1} = z_k + \gamma_k A(z_k) \tag{3}$$

(replacing (3) by (2) creates no extra mathematical problems). The proofs in [52] are based on Krylov–Bogoliubov averaging method [55]. Similarly to the 1st and 2nd Bogoliubov theorems the convergence in probability of solutions of (1) and (3) on finite time interval and, under assumption of asymptotic stability of the model (2), the closeness of the trajectories on infinite interval were established for independent f_k.

Significant progress of the method was made by Ljung [47]–[51] who also used Krylov–Bogoliubov approach. In [48] the dependent f_k were treated generated by controlled Markov chain. Moreover, the case $\gamma_k \to 0$ was examined. It was demonstrated that in this case model (2) is responsible for the stability or instability of system (1).

Further development was made by Kul'chitsky [40]–[41] who studied the averaging for some functional of of the state vector rather then for the state vector as a whole. It allowed to weaken the restrictive boundedness condition of [48].

if the gain parameter goes to zero at a suitable rate similar in spirit results were obtained [4, 5] without requirements on the dynamics of the model employing a certain set-valued deterministic model.

Another series of results [18]–[22] is based on the machinery developed by S.N. Bernstein who introduced the concept of stochastic differential equation (SDE) as early as in 1934 [9] and established the conditions of the convergence in distribution (weak convergence) of trajectories of (1) either to ODE (2) or to some SDE [10]. In [18] the mean square bounds of the model [2] accuracy were obtained both for finite and for infinite time interval. E.g. it was shown (in [18] for independent f_k and in [**?**] for f_k satisfying strong mixing conditions) that under Lipschitz and growth conditions

$$||A(z) - A(z')|| \leq L_1||z - z'||, b(z) \leq L_2(1 + ||z||^2), \tag{4}$$

where $b(z) = E||F(z, f_k) - A(z)||^2$ the following inequality holds:

$$E \max_{0 \leq t_k \leq T} \|x_k - x(t_k)\|^2 \leq C_1 e^{C_2 T} \gamma, \tag{5}$$

where $\gamma = \max_{1 \leq k \leq N} \gamma_k$, $t_N \leq T$, $C_1 > 0$, $C_2 > 0$.

In the case when the continuous model (2) is exponentially stable it was additionally shown in [18, 20] that the accuracy of approximation over infinite time interval is of order γ^α for some $0 < \alpha < 1$. Namely, there exist $\bar{\gamma} > 0$, such that for $\leq \gamma_k \leq \gamma < \bar{\gamma}$ the following inequalities hold

$$E||z_k - z(t_k)||^2 \leq C_3 \gamma^\alpha, k = 1, 2, ..., \tag{6}$$

where numbers $C_3 > 0$, $\alpha > 0$ do not depend on γ.

Though the averaging scheme of [18]–[22] is similar to that of Ljung [48], the analytical results are different in that they allow to analyze dynamics of the systems over finite or infinite time intervals rather than convergence as $t \to \infty$. Moreover the results of [18]–[22] are applicable to the cases when the gain γ_k does not tend to zero which is important in many applications.

Finally an elegant approach was developed by Kushner [42]–[44] who used weak convergence theory for random functions. This framework however is convenient for the studying of asymptotics when $\gamma \to 0$ ($\gamma_k \equiv \gamma$) rather than for evaluating mean distance between the trajectories for finite values of γ.

4. Stochastic continuous model

The inequality (5) shows that the distance between trajectories of (1) and (2) is of order $(\gamma_k)^{1/2}$. This error arises in part due to random fluctuations. Therefore the model taking in account stochasticity potentially may have higher accuracy. Employing the framework of averaging for SDE [29, 63] yields the stochastic continuous model [6, 7, 37]

$$dy = A(x(t))dt + (\gamma(t)B(x(t))^{1/2}dw, \tag{7}$$

where $\gamma(t) \equiv \gamma_k$ for $t_k \leq t \leq t_{k+1}$, $B(x) = \lim_{k\to\infty} Mh(x,f_k)h(x,f_k)^{\mathrm{T}}$, $h(x,f_k) = F(x,f_k) - A(x)$, $w(t)$ — standard Wiener stochastic process, $x(t)$ — solution of deterministic model (2). In [18] the following stochastic model

$$dy = A(y(t))dt + (\gamma(t)B(y(t)))^{1/2}dw \tag{8}$$

was suggested. Model (8) does not use the solutions of deterministic model (2). It was shown that the conditional incremental covariances of solutions of both (7) and (8) coincide with corresponding characteristics of (1) with the accuracy of order γ_k^2. In [**?**, 20] the family of stochastic models having higher accuracy was introduced. E.g. the accuracy of model

$$dy = \left[I_n - \frac{1}{2}\gamma(t)\frac{\partial A(y)}{\partial y}\right]A(y)dt + (\gamma(t)B(y(t)))^{1/2}dw \tag{9}$$

in terms of the conditional incremental covariances is of order γ_k^3. Note that the model (9) is nothing but the Stratonovich version of the SDE (8).

5. Further results

A number of further results were aimed at extension of the approximation theorems and relaxing their conditions. The approximation bounds were extended to the systems under relaxed Lipschitz and growth conditions [19, 20], to the right hand sides depending on γ_k [19], to the hybrid (discrete-continuous) systems [3]. In the case when the model (2) possesses ultimate boundedness instead of asymptotic stability the ultimate boundedness of initial system (1) was established [19, 20].

New problems such as synchronization and control of networks have become popular during last decade. They demand for new approximation results. One of new demands is to study accuracy of continuous modes over infinite time interval under partial stability assumption for (2) instead of asymptotic stability. For such cases the following theorem can be useful.

Definition 1 *Let $\Omega, \Omega_0,\ \Omega \subseteq \Omega_0$ be closed subsets of $\mathscr{R}^n$ and Ω consists of equilibria of (2). The set Ω is called Ω_0-pointwise stable if it is Lyapunov stable and any solution starting from Ω_0 tends to a point from Ω when $t \to \infty$.*

Theorem 1 *Let Lipschitz and growth conditions (4) hold. Let there exist a smooth mapping $h : \mathscr{R}^n \to \mathscr{R}^l$ and a bounded set $\Omega_0 \subseteq \mathscr{R}^n$ such that* rank$\partial y/partial z = l$ *for $z \in \Omega = \{z \in \Omega_0 : h(z) = 0\}$ and the set Ω is Ω_0-pointwise stable. Let there exist a twice continuously differentiable function $V(z)$ and positive numbers $\varkappa_1, \varkappa_2, \varkappa_3$ such that*

$$\dot{V}(z) \leq -\varkappa_1 V(z), \tag{10}$$

$$\left|\frac{\partial^2 V(z)}{\partial z^{(i)}\partial z^{(j)}}\right| \leq \varkappa_3, V(z) \geq \varkappa_2 ||h(z)||^2. \tag{11}$$

Then there exist numbers $\bar{\gamma} > 0$, $K_2 > 0$, $0 < \alpha < 1$ such that for $0 \leq \gamma_k \leq \gamma < \bar{\gamma}$ the following inequalities hold

$$E||y_k - y(t_k)||^2 \leq K_2\gamma^\alpha, k = 1,2,..., \tag{12}$$

where $y_k = h(z_k),\ y(t_k) = h(z(t_k))$.

Proof. Let $\bar{z}(t)$ be the solution of the ODE (2) with the initial condition $\bar{z}(0) = z_0 \in \Omega_0$. Pointwise stability implies existence of $\bar{z}_* = \lim_{t\to\infty}\bar{z}(t) \in \Omega$. It follows from (10) that $A(z_*) = 0$. Let $\varepsilon > 0$ be chosen such that rank $\partial y/\partial z(z) = l$ for $z \in \Omega_\varepsilon$, where $\Omega_\varepsilon = \{z \in |\Omega_0 : dist(z,\Omega) < \varepsilon\}$ and $t(\varepsilon)$ is such that $||\bar{z}(t) - z_*|| < \varepsilon$ for $t > t(\varepsilon)$. Then that $||A(z)|| \leq$ (We use notation $K = const$, if K depends only on $L_1, L_2, \kappa_i, \varepsilon, z_0, n$, i.e. does not depend on γ_k, t_k, where $t_k = \sum_{i=0}^{k-1}\gamma_i$).

Proof relies on the following lemmas.

Lemma 1 *If the numbers $\mu_k \geq 0$ satisfy inequalities $\mu_k \leq (1 + r_1\gamma_k)\mu_{k-1} + r_2\gamma_k$, $k = 1,2,...,$ where $r_1, r_2 > 0$, then $\mu_k \leq (\mu_0 + r_2/r_1)\exp(r_1 t_k)$. If $r_1 < 0$ and $0 \leq \gamma_k \leq \gamma < -1/r_1$, then $\mu_k \leq -3r_2/r_1 + \mu_0\exp(r_1 t_k)$.*

Proof of Lemma 1 is standard and therefore omitted.

Lemma 2 *Let the conditions of the theorem hold and the vectors $d_k \in \mathscr{R}^n$ are defined as follows:*

$$d_{k+1} = d_k + \gamma_k A(d_k), \qquad k = 0,1,2,\ldots, \tag{13}$$

Then for $0 \leq \gamma_k \leq \gamma < \rho_1$ the following inequalities hold:

$$V(z + \gamma_k A(z)) \leq (1 - \rho_2\gamma_k)V(z), \tag{14}$$

$$\varkappa_2||h(d_k)||^2 \leq V(z_0)\exp(-\rho_2 t_k), \tag{15}$$

$$\varkappa_2 E||z_k||^2 \leq V(z_0)\exp(-\rho_2 t_k) + K_1\gamma, \tag{16}$$

where $\rho = \varkappa_1\varkappa_2/(n^2\varkappa_3(L_1^2 + L_2))$, $\rho_2 = \varkappa_1/2$.

Proof 1 *The relations (14) follow from the Taylor expansion and the inequalities (11):*

$$\begin{aligned} &V(z + \gamma_k A(z)) \leq \\ &V(z) + \gamma_k \nabla V(z)^T A(z) + \gamma_k^2 n^2 \varkappa_2 ||A(z)||^2/2 \leq \\ &(1 - \varkappa_1\gamma_k)V(z) + \gamma_k\gamma n^2\varkappa_2 L_1^2||z||^2/2 \leq \\ &[1 - \gamma_k(\varkappa_1 - \gamma n^2\varkappa_3 L_1^2/(2\varkappa_2))]V(z), \end{aligned} \tag{17}$$

i.e.(14) is fulfilled for $\gamma < 2\varkappa_1\varkappa_2/(n^2\varkappa_3 L_1^2)$ and so more for $\gamma < \rho_1$. Apparently (15) follows from (14). To prove (16) estimate the value of $V(z_{k+1})$, recursively applying Taylor expansion and the relation (14).

$$\begin{aligned}&V(z_{k+1}) = \\ &V(z_k + A(z_k)) + [V(z_k) - V(z_k + A(z_k))] \leq \\ &(1-\rho_2\gamma_k)V(z_k) + \gamma_k \nabla V(z_k + A(z_k))^T h_k + \\ &\gamma_k^2 n^2 \varkappa_3 h_k^2/2.\end{aligned} \quad (18)$$

Averaging for fixed z_k yields

$$\begin{aligned}&E\{V(z_{k+1})|z_k\} \leq \\ &(1-\rho_2\gamma_k)V(z_k) + \gamma_k \gamma n^2 \varkappa_3 L_2(1+||z_k||^2)/2 \leq \\ &[1-\gamma_k(\varkappa_1 - \rho_1 n^2 \varkappa_3 (L_1^2 + L_2)/(2\varkappa_2))]V(z_k) + \\ &\gamma_k \gamma n^2 \varkappa_3 L_2/2.\end{aligned} \quad (19)$$

Taking full averaging and applying Lemma 1 for $r_1 = -\rho_2 = -\varkappa/2$, $\mu_k = EV(z_k)$ we get (16). Lemma is proven.

Lemma 3 *Under conditions of Lemma 2 the following is true $E||z_k - d_k||^2 \leq \gamma_k K_2 \exp(\rho_3 t_k)$, where $\rho_3 = 2L_1 + \rho_1 L_1^2$.*

Proof 2 *Comparing (2) and (13) yields $E\{||z_{k+1} - d_k||^2|z_k\} \leq (1+\gamma_k\rho_3)||z_k - d_k||^2 + \gamma_k^2 L_2(1+||z||^2)$. Averaging and application of the lemmas 1 and 2 finalize the proof of the Lemma 3.*

Continue the proof of the theorem.

Proof 3 *Let $\gamma_k \leq \gamma < \rho_1$. Choose α from the condition $1 - \alpha\rho_3/\rho_2 = \alpha$, or $\alpha = \rho_2/\rho_3 + \rho_2 > 0$, and put $t_\gamma = \rho_2^{-1} ln(V(z_0)/\gamma^\alpha)$. It follows from (10) that $V(z(t)) \leq \gamma^\alpha$ for $t \geq t_\gamma$. Lemma 2 implies that $EV(z_k) \leq \gamma^\alpha + K_1\gamma$ for $t_k \geq t_\gamma$. Therefore for $t_k \geq t_\gamma$ the following inequality holds*

$$\begin{aligned}&E||z_k - z(t_k)||^2 \leq \\ &2EV(z_k)/\varkappa_2 + 2V(z(t_k))/\varkappa_2 \leq K_3\gamma^\alpha,\end{aligned} \quad (20)$$

where $K_3 = const$. In the case $t_k \leq t_\gamma$ it follows from Lemma 3 that

$$\begin{aligned}&E||z_k - z(t_k)||^2 \leq \\ &2E||z_k - d_k||^2 + 2||d_k - z(t_k)||^2 \leq \\ &2\gamma K_2 \exp(\rho_3 t_\gamma) + 2\gamma^2 K_4 exp(2\alpha_1 t_\gamma) K_5 \gamma \exp(\rho_3 t_\gamma) \leq \\ &K_6 \gamma^{1-\alpha}\rho_3/\rho_2 = K_6\gamma^\alpha,\end{aligned} \quad (21)$$

where $K_4, K_5, K_6 = const$.

Therefore, $E||z_k - z(t_k)||^2 \leq K\gamma^\alpha, k = 1, 2, \ldots$, where $K = max\{K_3, K_6\}$. □

The theorem provides an upper mean square bound for the distance between the current state and the limit manifold $\Omega = \{z \in \Omega_0 : h(z) = 0\}$. An open problem is relaxation of the pointwise stability condition. Another problem is extending the results to discontinuous models important for economic games, and pattern recognition (some special cases are considered in [33, 46]).

6. Applications of Continuous Models

There are three stages of continuous model using: a) model building; b) model justifying; c) model analyzing (either analytic or numerical). The stage b) including checking the conditions of appropriate theorems sometimes happens to be rather involved. In many cases the theorems serve as "moral support" [48] of the designer's intuition.

Continuous models were used for the analyzing of algorithms of identification [6, 7, 8, 18, 20, 24, 42, 43, 44, 49, 51, 60], optimization [13, 20, 22, 26, 37, 44, 16], filtering [6, 7, 41, **?**, 54] and adaptive control [3, 19, **?**, 20, 21, 23, 27, 40, 50, 56, 62]; stochastic eigenvalue seeking [58, 65]; games solving [52, 61]; pattern recognition [18, 46, 49]; learning of neural networks [39]. A number of recent works open new networks related application areas: analyzing convergence of learning algorithms for coverage control of mobile sensing agents [14], distributed learning and cooperative control for multi-agent systems [15, 34], distributed topology control of wireless networks [12], etc.

7. Conclusion

Using the continuous models one can simplify the stability and performance analysis of adaptive systems and facilitate discrete-time system design by means of continuous-time design methods. Continuous models provide more detailed information about system behavior than, e.g., Lyapunov function. The main approaches to justifying averaging method for discrete-time stochastic systems are Krylov–Bogoliubov's approach [48, 52], Bernstein's approach [18]–[21] and weak convergence approach [42]. However the procedures of building the averaged (approximate) models are essentially the same. Basic conditions for applicability of averaging are stability of the system and mixing properties of disturbances.

A number of researches are devoted to analysis of the systems with constant or not tending to zero gain (learning rate). Perhaps, the first result of such kind was published in [18]. Unlike many other papers, in [18] approximation bounds were established for nonconstant not tending to zero gain. In this paper it was shown that the stability restriction of [18] can be relaxed to partial stability. Further relaxation is an avenue of further research.

References

[1] Ahmetgaliev T. Connection between stability of stochastic difference equations and stochastic differential equations. Differential equations, 1965, No 8.

[2] Anderson B.D.O., Bitmead R.R., Johnson Jr. C.R., Kokotovic P.V., Kosut R.L., Mareels I.M.Y., Praly L., Riedle B.D. Stability of adaptive systems: Passivity and averaging analysis. MIT Press, 1986.

[3] Andrievsky B.R., Blazhkin A.T., Derevitsky D.P., Fradkov A.L. The method of investigation of dynamics of digital adaptive systems of flight control. Proceedings of 6th IFAC Symposium on Automatic Control in the Space. Yerevan, 1974.

[4] Benaim M. A dynamical system approach to stochastic approximations *SIAM J. Control Optim.* 1996, V.34 (2), pp. 437–472.

[5] Benaim M, Hofbauer J, Sorin S. Stochastic approximations and differential inclusions. *SIAM J. Control Optim.* 2005, V.44 (1), pp. 328–348.

[6] Benveniste A. Design of adaptive algorithms for the tracking of time-varying systems. Int. J. Adapt. Control Sign. Proc. 1987, No 1, p. 3–31.

[7] Benveniste A., Metivier M., Priouret P. Algorithmes Adaptatifs et Approximations Stochastiques. Techniques Stochastiques, Masson, Paris, 1987 (English translation – Springer-Verlag, 1990).

[8] Benveniste A., Ruget G. A measure of the tracking capability of recursive stochastic algorithms with constant gains. IEEE Trans. Automatic Control AC-27, 1982, p. 639-649.

[9] Bernshtein S.N. Principes de la theorie des equationnes differentielles stochastiques. Proceedings of phys.-math. Inst. Ac. Sci. USSR; Math. div., v. 5, 1934, pp. 95–124.

[10] Bernshtein S.N. Theory of probabilities; 4th ed-n, 1946 app. 6. Stochastic equations in finite differences and stochastic differential equations, pp. 485–546 (in Russian).

[11] Borkar V.S. Stochastic Approximation: A Dynamical Systems Viewpoint. New Delhi: Hindustan Publishing Agency, and Cambridge, UK.

[12] Borkar V.S.,Manjunath D. Distributed topology control of wireless networks. Wireless Networks. V. 14, No 5, 671-682.

[13] Bozin A.S., Zarrop M.B. Self-tuning extremum optimizer — convergence and robustness properties. Proc. 1st ECC, Grenoble, 1991, pp. 672–678.

[14] Choi J., Horowitz R. Learning Coverage Control of Mobile Sensing Agents in One-Dimensional Stochastic Environments. IEEE Trans. Autom. Contr. 2010, V.55 (3), pp.804–809.

[15] Choi J, Oh S, Horowitz R. Distributed learning and cooperative control for multi-agent systems. Automatica, 2009, V.45(12), pp. 2802–2814.

[16] Coito F.J., Lemos J.M. Adaptive optimization with constraints: Convergence and Oscillatory Behaviour. Pattern Recognition and Image Analysis. LNCS, Springer, 2005, Volume 3523/2005, 335-366.

[17] Control Theory: Twenty-Five Seminal Papers (19321981). Wiley, 2000.

[18] Derevitskii D.P., Fradkov A.L. Two models for analysing dynamics of adaptation algorithms. Automation and Remote Control, 1974, V. 35 Is.1, pp. 59–67.

[19] Derevitskii D.P., Fradkov A.L. Investigation of discrete adaptive control systems using continuous models. Izvestia AN SSSR. Tekhn. Kibernetika, 1975 (5).

[20] Derevitskii D.P., Fradkov A.L. Applied theory of discrete adaptive control systems. Moscow: Nauka, 1981 (in Russian).

[21] Derevitskii D.P., Fradkov A.L. Averaging method for discrete stochastic systems and its application in adaptive control. Preprints of International Conference "Stochastic Optimization". Part 1. Kiev: Inst. of Cybernetics Ac.Sci. Ukraine, 1984, pp. 74–76.

[22] Derevitskii D.P., Ripa K.K., Fradkov A.L. Investigation of dynamics of random search algorithms. In: Problems of random search, Ed. by L.A. Rastrigin, Is 4, Riga: Zinatne, 1975, pp. 32–47.

[23] Dragan V., Halanay A. Preservation of exponential stability for discrete-time control implementation applied to adaptive stabilization. Siberian Math. J., 1990, (6).

[24] Dugard L., Landau I.D. Recursive output error identification algorithms. — Theory and evalution. Automatica, 1980, v. 16, No 5, p. 443–462.

[25] Egardt B. Stability of adaptive controllers. NY: Springer–Verlag, 1979.

[26] Ermoljev Yu.M., Kaniovsky Yu.M. Asymptotic properties of some stochastic programming methods with constant gain. Computational mathematics and mathematical physics, 1979, No 2, p. 356–366.

[27] Fradkov A.L. Adaptive control in complex systems. Moscow: Nauka, 1990 (in Russian).

[28] Fradkov A.L. Separation of motions in adaptive control systems. In: Voprosy Kibernetiki. Teoria i praktika adaptivnogo upravlenia. Moscow: AN SSSR, 1985, pp. 71–82.

[29] Freidlin M.I. The averaging principle and theorems on large deviations. Russian mathematical surveys, 1978, 33, No 5, pp.117-176.

[30] Geraschenko E.I., Geraschenko S.M. Method of separation of motions and optimization of nonlinear systems. Moscow: Nauka (in Russian).

[31] Gerencser L. A representation theorem for the error of recursive estimators, *SIAM J. Control Optim.* V. 44 (6), 2006, pp.2123–2188.

[32] Gikhman I.I., Skorokhod A.V. Stochastic differential equations. Kiev: Naukova Dumka, 1968 (Engl. transl. Berlin, New York, Springer-Verlag, 1972).

[33] Gorodeisky Z. Deterministic approximation of best-response dynamics for the Matching Pennies game Source: Games And Economic Behavior. 2009, Vol. 66,

Is.1, pp. 191–201.
[34] Huang M., Manton J.H. Coordination and consensus of networked agents with noisy measurements: stochastic algorithms and asymptotic behavior. *SIAM J. Control Optim.* 2009, V.48(1), pp. 134–161.
[35] Ibragimov I.A., Linnik Yu.V. Independent and stationary sequences of random variables, Wolters-Noordhoff, 1971.
[36] Ioannou P., Kokotovic P.V. Adaptive systems with reduced models. NY: Springer–Verlag, 1983.
[37] Kaniovky Yu.M., Knopov P.S., Nekrylova Z.V. Limiting theorems for processes of stochastic programming. Kiev: Naukova Dumka, 1980 (in Russian).
[38] Korostelev A.P. Convergence of recursive stochastic algorithms under gaussian disturbances. Kibernetika, 1979, No 4, pp. 93–98.
[39] Kuan C.M., Hornik K. Convergence of learning algorithms with constant learning rates. IEEE Trans. on Neural Networks, 1991, No 5, p. 484–489.
[40] Kulchitskii O.Yu. Algorithms of stochastic approximation in adaptation loop of linear dynamic system. Aut. Remote Contr., 1983, No 9; 1984, No 3.
[41] Kulchitskii O.Yu. The method of convergence analysis of adaptive filtration algorithms. Problems of Information Transmission, 1985, 21:4, pp. 285297.
[42] Kushner H.J. Convergence of recursive adaptive and identification procedures via weak convergence theory. IEEE Trans. Aut. Control, 1977, No 6, pp. 921–930.
[43] Kushner H.J. Approximation and weak convergence methods for random processes, with application to stochastic systems theory. MIT Press, Boston, 1984.
[44] Kushner H.J., Clark D.S. Stochastic approximation methods for constrained and unconstrained systems. NY: Springer–Verlag, 1978.
[45] Kushner H.J., Huang T.S. Asymptotic properties of stochastic approximations with constant coefficients. SIAM J. Control Opt., 1981, No 19, p. 87–105.
[46] Levitan V.D. Analysis of dynamics of discrete adaptation processes having discontinuous stochastic model. In: Voprosy Kibernetiki. Adaptivnye sistemy upravlenia. Moscow, AN SSSR, 1977, p. 127–129.
[47] Ljung L. Convergence of recursive stochastic algorithms. Proc. IFAC Symp. Stochastic Control. Budapest, Sept. 25-27, 1974.
[48] Ljung L. Analysis of recursive stochastic algorithms. IEEE Trans. Aut. Control, 1977, No 4, p. 551–575.
[49] Ljung L. On positive real transfer functions and the convergence of some recursive schemes. IEEE Trans. Aut. Control, 1977, No 4, p. 539–551.
[50] Ljung L. The ODE approach to the analysis of adaptive control systems. — Possibilities and limitations. Techn. Report LiTH–ISU–0371, Linköping University, 1980.
[51] Ljung L., Soderstrom T. Theory and practice of recursive identification. MIT Press, Boston, 1983.
[52] Meerkov S.M. On simplification of slow markovian walks description. Aut. Remote Contr., 1972, No 3, No 5.
[53] Meerkov S.M. Averaging of the trajectories of slow dynamic systems. Differential Equations, 1973, 9, pp.1239-1245.
[54] Metivier M., Priouret P. Application of a lemma of Kushner and Clark in general classes of stochastic algorithms. IEEE Trans. IT–30, 1984, pp. 140–150.
[55] Bogoliubov N. N., Mitropolsky Y. A. Asymptotic Methods in the Theory of Non-Linear Oscillations. New York, Gordon and Breach, 1961.
[56] Mosca E., Lemos J.M., Mendonca T.F., Nistri P. Input variance constrained adaptive control and singularly perturbed ODE's. Proceedings of the 1st ECC. Grenoble, 1991, p. 2176–2180.
[57] Nevelson M.B., Khasminskii R.Z. Stochastic approximation and recursive estimation. Moscow: Nauka, 1972.
[58] Oja E., Karhunen J. On stochastic approximation of the eigenvectors and the eigenvalues of the expectation of a random matrix. J. Math. Anal. and Appl., v. 106, 1985, pp. 69–84.
[59] Polyak B.T. Convergence and convergence rate of iterative stochastic algorithms. Autom. Rem. Control, 1976 (12).
[60] Solo V. The convergence of AML. IEEE Trans. Aut. Contr., 1979, No 10, p. 958–962.
[61] Stanković M.S., Johansson K.H., Stipanović D.M. Distributed Seeking of Nash Equilibria in Mobile Sensor Networks. Proc. IEEE CDC 2010, pp. 5598–5603.
[62] Tsykunov A.M. Adaptive control of plants with lags. Moscow: Nauka, 1984 (in Russian).
[63] Ventsel A.D., Freidlin M.I. Fluctuations in dynamic systems under small random disturbances. Moscow: Nauka, 1979. (Translated by Kluwer Publishers).
[64] Yoshihara K. Moment inequalities for mixing conditions. Kodai Math. Journal, v. 1, No 2, p. 316–327.
[65] Zhdanov A.I. Recursive estimations of minimal eigenvalues of informational matrices. Aut. Remote Control, 1987, No 4, p. 26–34.