

# Стохастические системы

© 2023 г. А.С. КОПОСОВ (koposov96@yandex.ru),  
П.В. ПАКШИН, д-р физ.-мат. наук (pakshinpv@gmail.com)  
(Арзамасский политехнический институт (филиал)  
Нижегородского государственного технического  
университета им. Р.Е. Алексева)

## УПРАВЛЕНИЕ С ИТЕРАТИВНЫМ ОБУЧЕНИЕМ СТОХАСТИЧЕСКИМИ МУЛЬТИАГЕНТНЫМИ СИСТЕМАМИ С ИЗМЕНЯЕМОЙ ЖЕЛАЕМОЙ ТРАЕКТОРИЕЙ И ТОПОЛОГИЕЙ<sup>1</sup>

В современном интеллектуальном производстве роботы часто связаны сетью, и их задача может меняться по заранее заданной программе. Для роботов, выполняющих высокоточные операции, широко используется управление с итеративным обучением (УИО). В условиях сети, в случае смены программы, эффективность алгоритмов УИО может снижаться. Например, при изменении желаемой траектории ошибка обучения может временно увеличиться до недопустимого значения. В данной работе рассматривается сетевая система, в которой желаемая траектория и параметры изменяются между повторениями по известной программе, агенты подвергаются случайным возмущениям, а измерения производятся с шумами. Кроме того, топология сети также изменяется за счет выхода из сети некоторых агентов и подключения новых агентов по заданной программе. Предлагается метод синтеза распределенного УИО на основе векторных функций Ляпунова для повторяющихся процессов в сочетании с фильтрацией Калмана. Метод обеспечивает сходимость ошибки обучения и снижает ее переходные возрастания, вызванные изменением желаемой траектории и топологии сети. Приведен пример, подтверждающий эффективность предлагаемого метода.

*Ключевые слова:* управление с итеративным обучением, мультиагентная система, изменяемая топология, случайные возмущения, повторяющиеся процессы, устойчивость, стабилизация, векторная функция Ляпунова, линейные матричные неравенства.

DOI: 10.31857/S0005231023060053, EDN: CSGCPH

### 1. Введение

Новой парадигмой современной промышленности является интеллектуальное производство (ИП), которое часто называют четвертой промышленной революцией (Industry 4.0). Системы ИП интегрируют физические ком-

---

<sup>1</sup> Работа выполнена при поддержке Российского научного фонда, грант № 22-21-00612, <https://rscf.ru/project/22-21-00612>.

поненты производства с киберпространством. Эти системы появились благодаря разработке и использованию информационных технологий, таких как промышленный Интернет вещей, искусственный интеллект (ИИ) и облачные вычисления, в сочетании со значительно возросшей производительностью современных компьютеров. Создание интеллектуальных производств стало мощным движущим фактором в развитии машинного обучения (МО) и сетевых мультиагентных систем. МО и сетевые структуры обеспечивают гибкую адаптацию к сокращению сроков выполнения заказов, жестким допускам на параметры продукции, ограничениям по стоимости, частым изменениям спроса и постоянному совершенствованию технологий; см. [1, 2] и ссылки в них.

Для динамических систем в технике концепция МО была определена Я.З. Цыпкиным [3] еще в начале 60-х годов прошлого века как процесс получения в системе необходимой реакции на внешние сигналы путем многократных воздействий на нее и внешней корректировки. УИО полностью вписывается в эту концепцию, поскольку оно предназначено для повторяющихся процессов, и на каждом повторении корректируется на основе информации с предыдущих повторений. За счет этого последовательно улучшается тот или иной показатель качества, например, точность воспроизведения желаемой траектории движения. Начиная с новаторского исследования [4] УИО стало активно развивающейся областью исследований и находит многочисленные практические приложения, прежде всего в робототехнике. Для отправной точки в литературе можно рекомендовать обзорные статьи [5, 6].

В стандартной постановке УИО желаемая траектория не изменяется от повторения к повторению, а после очередного повторения система всякий раз возвращается в одно и то же начальное состояние. Эти допущения ограничивают возможности УИО во многих приложениях ИП. В современном ИП желаемые траектории и цели управления могут изменяться по определенным правилам, в связи с чем управление должно соответствующим образом перенастраиваться. Изменение желаемой траектории порождает переходную ошибку, которая часто снижает точность ниже допустимого уровня в течение нескольких повторений. Следовательно, необходимо разработать новые алгоритмы УИО для компенсации переходной ошибки. В [2] рассмотрен сценарий реконфигурации управления при заранее определенном изменении желаемой траектории между повторениями. Стохастическая версия того же сценария была рассмотрена в [7, 8]. В цитируемых работах переходная ошибка компенсировалась алгоритмами УИО на основе соответствующих правил переключения.

С момента публикации [9] распределенный (сетевой) закон УИО рассматривался в ряде статей; см. [10–15] и ссылки в них. Хорошие обзоры предоставлены в [16] и в [15, 17, 18]. Анализ современного состояния задачи показывает следующее: в подавляющем большинстве исследований по распределенному УИО для линейных систем использовался алгоритм Аримото или его дискретный аналог, а математический аппарат базировался на методе супервек-

торов [19]. Для нелинейных систем тот же алгоритм Аримото был объединен с техникой априорных оценок, что давало консервативные результаты. Как правило, степень консерватизма оценить невозможно: приведенные чисто иллюстративные примеры не отражают возможных приложений.

Работа [20] была одной из первых, где эффективность предложенных сетевых алгоритмов УИО была подтверждена как моделированием, так и экспериментами с группой квадрокоптеров. Результаты [20] получили дальнейшее развитие в [16]. В [21] эффективность предложенных сетевых алгоритмов УИО также подтверждена моделированием и экспериментами с группой мобильных роботов. При этом использовались дискретные модели с дискретными версиями алгоритмов Аримото.

В [17] была предложена конструкция УИО для мультиагентных систем на основе двумерных дискретных моделей (линейные повторяющиеся процессы). Вычислительная сложность алгоритма проектирования была уменьшена за счет использования двумерного аналога теоремы Факса–Мюррея, хорошо известного в теории сетевых систем. Подход [17] был расширен в [18] на стохастические мультиагентные системы. Значительно улучшены точность и скорость сходимости ошибки слежения по сравнению с известными алгоритмами; теоретические результаты были наглядно подтверждены примерами. В [22] рассмотрена задача синтеза УИО для мультиагентных неопределенных систем в детерминированной постановке с изменяемой топологией сети. Был предложен специальный закон переключения для уменьшения переходной ошибки, вызванной изменением топологии. В [23] рассмотрена задача синтеза УИО для мультиагентных стохастических систем с изменяемой желаемой траекторией и фиксированной топологией сети.

В отличие от известных работ, в данной статье впервые в литературе решается задача синтеза распределенного УИО для мультиагентной сетевой системы в стохастической постановке, когда желаемая траектория и топология сети изменяются на заранее известном и ограниченном интервале повторений по некоторой внешней программе.

## 2. Постановка задачи

Рассмотрим сетевую систему, состоящую из  $N$  линейных подсистем (агентов), задачей которых является выполнение одной и той же многократно повторяющейся операции. В процессе функционирования системы топология сети может изменяться. Динамика агента  $i$  описывается следующей дискретной моделью в пространстве состояний:

$$(2.1) \quad x_i(k, p+1) = A_{\sigma_i(k)}x_i(k, p) + B_{\sigma_i(k)}u_i(k, p) + D_{\sigma_i(k)}\omega_i(k, p),$$

$$(2.2) \quad y_i(k, p) = Cx_i(k, p),$$

$$(2.3) \quad y_{vi}(k, p) = y_i(k, p) + G_{\sigma_i(k)}\nu_i(k, p), \quad i \in \mathcal{I}, \quad k \geq 0, \quad 0 \leq p \leq T-1,$$

где  $x_i(k, p) \in \mathbb{R}^{n_x}$  — вектор состояния,  $u_i(k, p) \in \mathbb{R}^1$  — скалярная переменная управления,  $\omega_i(k, p) \in \mathbb{R}^{n_\omega}$  — вектор возмущений, действующих на объект,  $y_i(k, p) \in \mathbb{R}^1$  — скалярная выходная переменная (профиль повторения), недоступная наблюдению,  $y_{vi}(k, p) \in \mathbb{R}^1$  — наблюдаемая (измеряемая) выходная переменная,  $\nu_i(k, p) \in \mathbb{R}^1$  — шум измерений,  $k$  — номер повторения,  $p$  — дискретное время на повторении  $k$ ,  $T$  — продолжительность повторения, одинаковая для всех  $k$ ,  $\mathcal{I} = \{1, 2, \dots, N\}$  — множество всех агентов. Начальные условия  $x_i(k, 0)$  и  $u_i(0, p)$  одинаковы для всех агентов.

Предполагается, что возмущения  $\omega_i(k, p)$  и шумы измерений  $\nu_i(k, p)$  представляют собой независимые гауссовские белые шумы с нулевым математическим ожиданием и с ковариациями

$$S_{i\omega} = \mathbb{E} \left[ \omega_i(k, p) \omega_i^\top(k, p) \right],$$

$$S_{i\nu} = \mathbb{E} \left[ \nu_i(k, p)^2 \right],$$

где  $\mathbb{E}$  — символ математического ожидания. Также предположим, что  $\omega_i(k, p)$  не зависит от начального вектора состояния.

Сигнал  $\sigma_i(k)$ , переключающий режим работы агента  $i$ , представляет собой кусочно-постоянную функцию, отображающую  $\mathbb{Z}_+$  в  $\{1, \dots, m\}$  (где  $m$  — количество режимов), точки разрыва которой будем называть *моментами переключения режима*. В каждом режиме задана своя желаемая траектория выходного сигнала  $y_{\sigma_i(k)}^{ref}(p)$  и матрицы  $A_{\sigma_i(k)} \in \{A_1, \dots, A_m\}$ ,  $B_{\sigma_i(k)} \in \{B_1, \dots, B_m\}$ ,  $D_{\sigma_i(k)} \in \{D_1, \dots, D_m\}$  и  $G_{\sigma_i(k)} \in \{G_1, \dots, G_m\}$  агентов, при этом тройки  $(A_{\sigma_i(k)}, B_{\sigma_i(k)}, C)$  полностью управляемы и наблюдаемы, а  $CB_{\sigma_i(k)} \neq 0$ .

Сигнал  $\rho(k)$ , переключающий топологию сети, представляет собой кусочно-постоянную функцию, отображающую  $\mathbb{Z}_+$  в  $\{1, \dots, c\}$  (где  $c$  — количество топологий), точки разрыва которой будем называть *моментами переключения топологии*. Каждая топология определяется множеством функционирующих агентов, которое представляется в виде набора их номеров  $\mathcal{I}_{\rho(k)} = \{i_n\}_{n=1}^{N_{\rho(k)}} \subseteq \mathcal{I}$ , где  $N_{\rho(k)} \leq N$  — количество агентов, функционирующих при топологии  $\rho(k)$ , и связями между ними, которые представляются в виде направленного графа  $\mathcal{G}_{\rho(k)} = (\mathcal{I}_{\rho(k)}, \mathcal{E}_{\rho(k)})$ , где  $\mathcal{E}_{\rho(k)} \subseteq \mathcal{I}_{\rho(k)} \times \mathcal{I}_{\rho(k)}$  — ребра графа. Возможность доступа агента  $i$  к выходным данным агента  $j$  ( $i, j \in \mathcal{I}_{\rho(k)}$ ) задается ребром, направленным от вершины  $j$  к вершине  $i$ , и обозначается упорядоченной парой  $(j, i) \in \mathcal{E}_{\rho(k)}$ . Элементы матрицы смежности

$$S_\rho(\mathcal{G}_\rho) = \begin{bmatrix} s_{i_1 i_1} & s_{i_1 i_2} & \cdots & s_{i_1 i_{N_\rho}} \\ s_{i_2 i_1} & s_{i_2 i_2} & \cdots & s_{i_2 i_{N_\rho}} \\ \vdots & \vdots & \ddots & \vdots \\ s_{i_{N_\rho} i_1} & s_{i_{N_\rho} i_2} & \cdots & s_{i_{N_\rho} i_{N_\rho}} \end{bmatrix}, \quad \rho = \rho(k)$$

задаются следующим образом:  $s_{ij} > 0$ , если  $(j, i) \in \mathcal{E}_{\rho(k)}$ ,  $s_{ij} = 0$  в противном случае,  $s_{ii} = 0$ . Матрица Лапласа графа  $\mathcal{G}_{\rho(k)}$  задается выражением

$$\mathcal{L}_{\rho}(\mathcal{G}_{\rho}) = \begin{bmatrix} \sum_{j \in \mathcal{I}_{\rho}} s_{i_1 j} & -s_{i_1 i_2} & \cdots & -s_{i_1 i_{N_{\rho}}} \\ -s_{i_2 i_1} & \sum_{j \in \mathcal{I}_{\rho}} s_{i_2 j} & \cdots & -s_{i_2 i_{N_{\rho}}} \\ \vdots & \vdots & \ddots & \vdots \\ -s_{i_{N_{\rho}} i_1} & -s_{i_{N_{\rho}} i_2} & \cdots & \sum_{j \in \mathcal{I}_{\rho}} s_{i_{N_{\rho}} j} \end{bmatrix}, \quad \rho = \rho(k).$$

Будем считать, что непосредственный доступ к  $y_{\sigma_i(k)}^{ref}(p)$  может иметь только некоторая подгруппа агентов, состав которой может меняться в зависимости от топологии сети. Возможность получения агентами информации о желаемой траектории выходного сигнала определяется матрицей  $\mathcal{R}_{\rho(k)} = \text{diag}[r_i]_{i=1}^{N_{\rho(k)}}$ , где  $r_i = 1$ , если  $i$  имеет доступ к  $y_{\sigma_i(k)}^{ref}(p)$ , и  $r_i = 0$  в противном случае.

Агентов, для которых  $r_i = 1$ , будем именовать *глобальными лидерами*. Остальные могут получать информацию о выходном сигнале либо от глобальных лидеров, либо от любых других агентов. Агента  $i$ , для которого  $r_i = 0$  и  $\exists j : s_{ij} > 0$ , будем называть *ведомым*, где агент  $j$  — один из агентов, передающих информацию агенту  $i$ , которых будем именовать *локальными лидерами* ведомого агента. Предполагается, что каждый ведомый агент не может передавать данные своим локальным лидерам.

Будем предполагать, что общее число переключений режима агентов и топологии сети  $N_{\sigma\rho}$  конечно. Такой сценарий возникает, например, когда агенты представляют собой порталные роботы, функционирующие по некоторой программе в интеллектуальном производстве.

Ошибка обучения

$$e_i(k, p) = y_{\sigma_i(k)}^{ref}(p) - y_i(k, p)$$

недоступна для измерения и последующего формирования управления, поэтому в дальнейшем вместо нее будем использовать

$$(2.4) \quad \hat{e}_i(k, p) = y_{\sigma_i(k)}^{ref}(p) - \hat{y}_i(k, p),$$

где  $\hat{y}_i(k, p) = C\hat{x}_i(k, p)$ ,  $\hat{x}_i(k, p)$  — оценка состояния агента  $i$ , полученная с использованием фильтра Калмана

$$(2.5) \quad \hat{x}_i(k, p+1) = A_{\sigma_i(k)}\hat{x}_i(k, p) + B_{\sigma_i(k)}u_i(k, p) + F_{i\sigma_i(k)}(y_{\nu i}(k, p) - C\hat{x}_i(k, p))$$

с начальным условием  $\hat{x}_i(k, 0) = F_{i\sigma_i(k)}y_{\nu i}(k, 0)$ ,  $i \in \mathcal{I}_{\rho(k)}$ , где  $F_{i\sigma_i(k)} = A_{\sigma_i(k)}S_{i\sigma_i(k)}C^{\top} \left[ CS_{i\sigma_i(k)}C^{\top} + G_{\sigma_i(k)}S_{i\nu}G_{\sigma_i(k)}^{\top} \right]^{-1}$ ,  $S_{i\sigma_i(k)}$  — решение алгеб-

раического уравнения Риккати

$$S_{i\sigma_i(k)} = A_{\sigma_i(k)} S_{i\sigma_i(k)} A_{\sigma_i(k)}^\top - \\ - A_{\sigma_i(k)} S_{i\sigma_i(k)} C^\top \left[ C S_{i\sigma_i(k)} C^\top + G_{\sigma_i(k)} S_{i\nu} G_{\sigma_i(k)}^\top \right]^{-1} C S_{i\sigma_i(k)} A_{\sigma_i(k)}^\top + \\ + D_{\sigma_i(k)} S_{i\nu} D_{\sigma_i(k)}^\top, \quad i \in \mathcal{I}_{\rho(k)}.$$

Задача заключается в нахождении такого управления (протокола)  $u_i(k, p)$ , который позволит достичь консенсуса в следующем смысле:

$$(2.6) \quad \mathbb{E} [|\hat{e}_i(k, p)|^2] \leq \kappa \varrho^k + \delta, \quad \kappa > 0, \quad 0 < \varrho < 1, \quad \delta > 0,$$

$$(2.7) \quad \lim_{k \rightarrow \infty} \mathbb{E} [|u_i(k, p)|^2] = \mathbb{E} [|u_i(\infty, p)|^2] < \infty, \quad i \in \mathcal{I}, \quad 0 \leq p \leq T - 1.$$

Предельное значение  $u_i(\infty, p)$  часто называют обученным управлением.

### 3. Построение 2D модели в приращениях переменных

#### 3.1. Фиксированный режим агентов и фиксированная топология сети

Вначале рассмотрим случай, когда режим работы агентов и топология сети не меняются, т.е. рассмотрим интервал вдоль повторений, на котором сигналы  $\sigma_i(k)$  и  $\rho(k)$  не имеют точек разрыва и их значения равны для всех  $i$  и  $k$ . Далее при решении задачи в подобных случаях будем использовать обозначения  $\sigma = \sigma_i(k)$  и  $\rho = \rho(k)$ . Зададим закон УИО как

$$(3.1) \quad u_i(k+1, p-1) = u_i(k, p-1) + \Delta u_i(k+1, p-1),$$

для которого корректирующую (обновляющую) поправку  $\Delta u_i(k+1, p-1)$  будем формировать в виде, аналогичном принятому в [18, 22]:

$$(3.2) \quad \Delta u_i(k+1, p-1) = K_{1\sigma\rho} (\hat{x}_i(k+1, p-1) - \hat{x}_i(k, p-1)) + \\ + K_{2\sigma\rho} \left( \sum_{j \in N_{\rho i}} s_{ij} (\hat{y}_j(k, p) - \hat{y}_i(k, p)) + r_i (y_\sigma^{ref}(p) - \hat{y}_i(k, p)) \right),$$

где  $K_{1\sigma\rho}$  и  $K_{2\sigma\rho}$  — матрицы протокола в режиме  $\sigma$  и топологии  $\rho$ ,  $N_{\rho i} = \{j \in \mathcal{I}_\rho | (j, i) \in \mathcal{E}_\rho\}$  — множество агентов, информация о которых доступна агенту  $i$  в топологии  $\rho$ ,  $s_{ij}$  и  $r_i$  — элементы матриц  $\mathcal{S}_\rho$  и  $\mathcal{R}_\rho$  соответственно.

Введем вектор приращения оценки состояния

$$\hat{\eta}_i(k+1, p+1) = \hat{x}_i(k+1, p) - \hat{x}_i(k, p),$$

ошибку оценивания  $\tilde{x}_i(k, p) = x_i(k, p) - \hat{x}_i(k, p)$ , приращение ошибки оценивания

$$\tilde{\eta}_i(k+1, p+1) = \tilde{x}_i(k+1, p) - \tilde{x}_i(k, p),$$

и запишем систему (2.1)–(2.3) в терминах приращений и оценки ошибки обучения (2.4):

$$(3.3) \quad \begin{aligned} \hat{\eta}_i(k+1, p+1) &= A_\sigma \hat{\eta}_i(k+1, p) + F_{i\sigma} C \tilde{\eta}_i(k+1, p) + \\ &+ B_\sigma \Delta u_i(k+1, p-1) + F_{i\sigma} G_\sigma \Delta \nu_i(k+1, p-1), \end{aligned}$$

$$(3.4) \quad \begin{aligned} \tilde{\eta}_i(k+1, p+1) &= (A_\sigma - F_{i\sigma} C) \tilde{\eta}_i(k+1, p) + \\ &+ D_\sigma \Delta \omega_i(k+1, p-1) - F_{i\sigma} G_\sigma \Delta \nu_i(k+1, p-1), \end{aligned}$$

$$(3.5) \quad \begin{aligned} \hat{e}_i(k+1, p) &= -C A_\sigma \hat{\eta}_i(k+1, p) - C F_{i\sigma} C \tilde{\eta}_i(k+1, p) + \hat{e}_i(k, p) - \\ &- C B_\sigma \Delta u_i(k+1, p-1) - C F_{i\sigma} G_\sigma \Delta \nu_i(k+1, p-1), \end{aligned}$$

где  $\Delta \nu_i(k+1, p-1) = \nu_i(k+1, p-1) - \nu_i(k, p-1)$ ,  $\Delta \omega_i(k+1, p-1) = \omega_i(k+1, p-1) - \omega_i(k, p-1)$ . Второе уравнение в системе (3.3)–(3.5) не зависит от других, следовательно, справедлив хорошо известный для обычных линейных систем принцип разделения, согласно которому фильтр и регулятор можно синтезировать независимо. Поэтому  $\tilde{\eta}_i$  можно рассматривать как внешнюю ограниченную переменную и для дальнейшего синтеза управления использовать следующую систему:

$$(3.6) \quad \begin{aligned} \hat{\eta}_i(k+1, p+1) &= A_\sigma \hat{\eta}_i(k+1, p) + B_\sigma \Delta u_i(k+1, p-1) + \\ &+ F_{i\sigma} G_\sigma \Delta \nu_i(k+1, p-1), \end{aligned}$$

$$(3.7) \quad \begin{aligned} \hat{e}_i(k+1, p) &= -C A_\sigma \hat{\eta}_i(k+1, p) + \hat{e}_i(k, p) - C B_\sigma \Delta u_i(k+1, p-1) - \\ &- C F_{i\sigma} G_\sigma \Delta \nu_i(k+1, p-1). \end{aligned}$$

Введем расширенные векторы

$$\begin{aligned} \hat{x}(k, p) &= \left[ \hat{x}_{i_1}^\top(k, p) \quad \dots \quad \hat{x}_{i_{N_\rho}}^\top(k, p) \right]^\top, \\ \hat{\eta}(k, p) &= \left[ \hat{\eta}_{i_1}^\top(k, p) \quad \dots \quad \hat{\eta}_{i_{N_\rho}}^\top(k, p) \right]^\top, \\ \hat{e}(k, p) &= \left[ \hat{e}_{i_1}(k, p) \quad \dots \quad \hat{e}_{i_{N_\rho}}(k, p) \right]^\top, \\ \Delta \nu(k, p) &= \left[ \Delta \nu_{i_1}(k, p) \quad \dots \quad \Delta \nu_{i_{N_\rho}}(k, p) \right]^\top, \end{aligned}$$

и, с учетом (3.2), перепишем расширенную систему (3.6)–(3.7) в виде

$$(3.8) \quad \begin{aligned} \hat{\eta}(k+1, p+1) &= (\bar{A}_{11\sigma\rho} + \bar{B}_{1\sigma\rho} \bar{K}_{1\sigma\rho} \bar{\mathcal{H}}_{1\rho}) \hat{\eta}(k+1, p) + \\ &+ (\bar{A}_{12\sigma\rho} + \bar{B}_{1\sigma\rho} \bar{K}_{2\sigma\rho} \bar{\mathcal{H}}_{2\rho}) \hat{e}(k, p) + \bar{F}_{1\sigma\rho} \Delta \nu(k+1, p-1), \end{aligned}$$

$$(3.9) \quad \begin{aligned} \hat{e}(k+1, p) &= (\bar{A}_{21\sigma\rho} + \bar{B}_{2\sigma\rho} \bar{K}_{1\sigma\rho} \bar{\mathcal{H}}_{1\rho}) \hat{\eta}(k+1, p) + \\ &+ (\bar{A}_{22\sigma\rho} + \bar{B}_{2\sigma\rho} \bar{K}_{2\sigma\rho} \bar{\mathcal{H}}_{2\rho}) \hat{e}(k, p) + \bar{F}_{2\sigma\rho} \Delta \nu(k+1, p-1), \end{aligned}$$

где

$$\begin{aligned}
\bar{A}_{11\sigma\rho} &= I_{N_\rho} \otimes A_\sigma, \quad \bar{A}_{12\sigma\rho} = 0, \quad \bar{A}_{21\sigma\rho} = I_{N_\rho} \otimes (-CA_\sigma), \quad \bar{A}_{22\sigma\rho} = I_{N_\rho}, \\
\bar{B}_{1\sigma\rho} &= I_{N_\rho} \otimes B_\sigma, \quad \bar{B}_{2\sigma\rho} = I_{N_\rho} \otimes (-CB_\sigma), \\
\bar{K}_{1\sigma\rho} &= I_{N_\rho} \otimes K_{1\sigma\rho}, \quad \bar{K}_{2\sigma\rho} = I_{N_\rho} \otimes K_{2\sigma\rho}, \\
\bar{\mathcal{H}}_{1\rho} &= I_{N_\rho} \otimes \mathcal{H}_1, \quad \mathcal{H}_1 = I_{n_x}, \quad \bar{\mathcal{H}}_{2\rho} = (\mathcal{L}_\rho + \mathcal{R}_\rho) \otimes \mathcal{H}_2, \quad \mathcal{H}_2 = 1, \\
\bar{F}_{1\sigma\rho} &= \text{diag}[F_{i_n\sigma}G_\sigma]_{n=1}^{N_\rho}, \quad \bar{F}_{2\sigma\rho} = \text{diag}[-CF_{i_n\sigma}G_\sigma]_{n=1}^{N_\rho},
\end{aligned}$$

и  $\otimes$  — символ произведения Кронекера.

Система (3.8)–(3.9) записана в стандартной форме повторяющегося процесса и дальнейший вывод условий сходимости будет основан на теории устойчивости стохастических повторяющихся процессов с переключениями [24].

### 3.2. Переключения режимов агентов и топологии

Рассмотрим случай, когда происходит переключение режима работы агентов. Это может происходить при неизменной топологии и сигнал, переключающий топологию, для краткости обозначим за  $\rho$ , подразумевая, что он одинаков на всех рассматриваемых повторениях. Начнем с рассмотрения глобальных лидеров. Пусть  $k + 1$  — один из моментов переключения глобального лидера  $i$ . Его управляемая динамика описывается следующим образом:

$$(3.10) \quad \begin{aligned}
\hat{x}_i(k+1, p) &= A_{\sigma_i(k+1)}\hat{x}_i(k+1, p-1) + F_{i\sigma_i(k+1)}C\tilde{x}_i(k+1, p-1) + \\
&+ B_{\sigma_i(k+1)}u_i(k+1, p-1) + F_{i\sigma_i(k+1)}G_{\sigma_i(k+1)}\nu_i(k+1, p-1),
\end{aligned}$$

$$(3.11) \quad \begin{aligned}
\tilde{x}_i(k+1, p) &= (A_{\sigma_i(k+1)} - F_{i\sigma_i(k+1)}C)\tilde{x}_i(k+1, p-1) + \\
&+ D_{\sigma_i(k+1)}\omega_i(k+1, p-1) - F_{i\sigma_i(k+1)}G_{\sigma_i(k+1)}\nu_i(k+1, p-1),
\end{aligned}$$

$$(3.12) \quad \begin{aligned}
\hat{e}_i(k+1, p) &= -C(A_{\sigma_i(k+1)}\hat{x}_i(k+1, p-1) - A_{\sigma_i(k)}\hat{x}_i(k, p-1)) - \\
&- C(F_{i\sigma_i(k+1)}C\tilde{x}_i(k+1, p-1) - F_{i\sigma_i(k)}C\tilde{x}_i(k, p-1)) + \hat{e}_i(k, p) - \\
&- C(B_{\sigma_i(k+1)} - B_{\sigma_i(k)})u_i(k, p-1) - CB_{\sigma_i(k+1)}\Delta u_i(k+1, p-1) - \\
&- C(F_{i\sigma_i(k+1)}G_{\sigma_i(k+1)}\nu_i(k+1, p-1) - F_{i\sigma_i(k)}G_{\sigma_i(k)}\nu_i(k, p-1)) + \\
&+ \left( y_{\sigma_i(k+1)}^{ref}(p) - y_{\sigma_i(k)}^{ref}(p) \right).
\end{aligned}$$

Поскольку  $\tilde{x}_i(k+1, p)$  не зависит от других переменных и недоступна для измерения по причине, указанной в предыдущем разделе, при дальнейшем синтезе управления исключим ее из (3.10) и (3.12).

На шаге переключения, в отличие от раздела 3.1, в последнем уравнении (3.12) появляется возмущающее воздействие  $\left( y_{\sigma_i(k+1)}^{ref}(p) - y_{\sigma_i(k)}^{ref}(p) \right)$ , порождающее переходный процесс, за счет которого ошибка обучения может существенно возрасти, что является нежелательным эффектом, поэтому на



данном шаге закон управления целесообразно построить так, чтобы минимизировать влияние этого возмущения, а затем вернуться к исходному закону управления, гарантирующему сходимость. Такой закон управления можно найти из решения задачи минимизации отклонения выходного сигнала агента от доступного ему образа желаемой траектории. Аналогичная ситуация возникает при переключениях режима для ведомых агентов и при изменении топологии. Синтез управления для всех этих случаев будет описан далее.

#### 4. Теорема о сходимости

Для нахождения условий сходимости воспользуемся результатами [24]. В отличие от указанного исследования, в рассматриваемой задаче число переключений конечно и нет необходимости оценивать среднее время ожидания. Определим на траекториях системы (3.8)–(3.9) векторную функцию Ляпунова

$$(4.1) \quad V_{\sigma\rho}(\xi, \epsilon) = \begin{bmatrix} V_{1\sigma\rho}(\xi) \\ V_{2\sigma\rho}(\epsilon) \end{bmatrix},$$

где  $V_{1\sigma\rho}(\xi) > 0$ ,  $\xi \neq 0$ ,  $V_{2\sigma\rho}(\epsilon) > 0$ ,  $\epsilon \neq 0$  и  $V_{1\sigma\rho}(0) = 0$ ,  $V_{2\sigma\rho}(0) = 0$ , и дискретный аналог ее дивергенции

$$\begin{aligned} \mathcal{D}V_{\sigma\rho}(\xi, \epsilon) &= \mathbb{E}[V_{1\sigma\rho}(\hat{\eta}(k+1, p+1)) | \hat{\eta}(k+1, p) = \xi, \hat{\epsilon}(k, p) = \epsilon] - \\ &- V_{1\sigma\rho}(\xi) + \mathbb{E}[V_{2\sigma\rho}(\hat{\epsilon}(k+1, p)) | \hat{\eta}(k+1, p) = \xi, \hat{\epsilon}(k, p) = \epsilon] - V_{2\sigma\rho}(\epsilon). \end{aligned}$$

*Теорема 1.* Пусть существует векторная функция Ляпунова (4.1) и положительные скаляры  $c_1$ ,  $c_2$ ,  $c_3$  и  $\gamma$  такие, что на траекториях системы (3.8)–(3.9) для всех пар  $\sigma\rho$  выполняются неравенства

$$\begin{aligned} c_1 \|\xi\|^2 &\leq V_{1\sigma\rho}(\xi) \leq c_2 \|\xi\|^2, \\ c_1 |\epsilon|^2 &\leq V_{2\sigma\rho}(\epsilon) \leq c_2 |\epsilon|^2, \\ \mathcal{D}V_{\sigma\rho}(\xi, \epsilon) &\leq \gamma - c_3 (\|\xi\|^2 + |\epsilon|^2). \end{aligned}$$

Тогда закон управления (3.1) с корректирующей поправкой (3.2) обеспечивает условия сходимости (2.6).

*Доказательство.* Вычисляя  $\mathcal{D}V_{\sigma\rho}(\xi, \epsilon)$  в силу системы (3.8)–(3.9) и следуя доказательству теоремы 1 из [24], получим, что для всех  $k$ ,  $p$  справедлива оценка

$$(4.2) \quad \begin{aligned} \mathbb{E}[|\hat{\epsilon}(k, p-1)|^2] &\leq \mu^{N_{\sigma\rho}+1} \left[ \lambda^k \sum_{q=0}^{p-1} \lambda^{p-1-q} |\hat{\epsilon}(0, q)|^2 \right] + \\ &+ \frac{\gamma}{c_1(1-\lambda)^2}, \quad 0 < \lambda < 1, \end{aligned}$$

где  $\mu = c_2/c_1 \geq 1$ , откуда следует справедливость (2.6) при  $\varrho = \lambda$ ,  $\kappa = \frac{\mu^{N_{\sigma\rho}+1}|\bar{e}|^2}{1-\lambda}$ ,  $|\bar{e}|^2 = \max_q |e(0, q)|^2$ ,  $\delta = \frac{\gamma}{c_1(1-\lambda)^2}$ . Параметр  $\lambda$ , а следовательно, и  $\varrho$ , зависит от  $c_2$  и  $c_3$  и определяет скорость сходимости ошибки обучения, в то время как  $\mu$  зависит от  $c_1$  и  $c_2$  и определяет начальное значение оценки. Теорема доказана.

Заметим, что с учетом (4.2) из (2.4) для любого  $\sigma$  следует

$$(4.3) \quad \mathbb{E}[|C\hat{x}_i(k, p)|^2] = \mathbb{E}[|\hat{y}_i(k, p)|^2] \leq 2|y_{\sigma}^{ref}(p)|^2 + 2\mathbb{E}[|\hat{e}_i(k, p)|^2] < \infty.$$

Эта оценка будет использована в дальнейшем для доказательства условия (2.7).

## 5. Синтез управления

### 5.1. Управление при фиксированном режиме агентов и фиксированной топологии сети

Выберем компоненты векторной функции Ляпунова в виде квадратичных форм

$$\begin{aligned} V_{1\sigma\rho}(\xi) &= \xi^\top \bar{P}_{1\sigma\rho} \xi, \\ V_{2\sigma\rho}(\epsilon) &= \epsilon^\top \bar{P}_{2\sigma\rho} \epsilon, \end{aligned}$$

где  $\bar{P}_{1\sigma\rho} = I_{N_\rho} \otimes P_{1\sigma\rho}$ ,  $\bar{P}_{2\sigma\rho} = I_{N_\rho} \otimes P_{2\sigma\rho}$ , и потребуем выполнения неравенства

$$(5.1) \quad DV_{\sigma\rho}(\xi, \epsilon) \leq \gamma - \left( \xi^\top \bar{Q}_{1\rho} \xi + \epsilon^\top \bar{Q}_{2\rho} \epsilon + \Delta u^\top \bar{R}_\rho \Delta u \right),$$

где  $\bar{Q}_{1\rho} \succ 0$ ,  $\bar{Q}_{2\rho} \succ 0$ ,  $\bar{R}_\rho \succ 0$ ,  $\bar{Q}_{1\rho} = I_{N_\rho} \otimes Q_{1\rho}$ ,  $\bar{Q}_{2\rho} = I_{N_\rho} \otimes Q_{2\rho}$ ,  $\bar{R}_\rho = I_{N_\rho} \otimes R_\rho$ , и  $\Delta u = \bar{K}_{1\sigma\rho} \bar{\mathcal{H}}_{1\rho} \xi + \bar{K}_{2\sigma\rho} \bar{\mathcal{H}}_{2\rho} \epsilon$ .

Вычисляя аналог дивергенции векторной функции Ляпунова, получим

$$\begin{aligned} DV_{\sigma\rho}(\xi, \epsilon) &= \begin{bmatrix} \xi \\ \epsilon \end{bmatrix}^\top \left( (\bar{A}_{\sigma\rho} + \bar{B}_{\sigma\rho} \bar{K}_{\sigma\rho} \bar{\mathcal{H}}_\rho)^\top \bar{P}_{\sigma\rho} (\bar{A}_{\sigma\rho} + \bar{B}_{\sigma\rho} \bar{K}_{\sigma\rho} \bar{\mathcal{H}}_\rho) - \bar{P}_{\sigma\rho} + \right. \\ &\quad \left. + \bar{Q}_\rho + (\bar{K}_{\sigma\rho} \bar{\mathcal{H}}_\rho)^\top \bar{R}_\rho \bar{K}_{\sigma\rho} \bar{\mathcal{H}}_\rho \right) \begin{bmatrix} \xi \\ \epsilon \end{bmatrix} + 2 \left( \text{tr} [\bar{P}_{1\sigma\rho} S_{1\sigma}] + \text{tr} [\bar{P}_{2\sigma\rho} S_{2\sigma}] \right), \end{aligned}$$

где

$$\begin{aligned} \bar{A}_{\sigma\rho} &= \begin{bmatrix} \bar{A}_{11\sigma\rho} & \bar{A}_{12\sigma\rho} \\ \bar{A}_{21\sigma\rho} & \bar{A}_{22\sigma\rho} \end{bmatrix}, \quad \bar{B}_{\sigma\rho} = \begin{bmatrix} \bar{B}_{1\sigma\rho} \\ \bar{B}_{2\sigma\rho} \end{bmatrix}, \quad \bar{K}_{\sigma\rho} = [ \bar{K}_{1\sigma\rho} \quad \bar{K}_{2\sigma\rho} ], \\ \bar{Q}_\rho &= \text{diag} [ \bar{Q}_{1\rho} \quad \bar{Q}_{2\rho} ], \quad \bar{\mathcal{H}}_\rho = \text{diag} [ \bar{\mathcal{H}}_{1\rho} \quad \bar{\mathcal{H}}_{2\rho} ], \quad \bar{P}_{\sigma\rho} = \text{diag} [ \bar{P}_{1\sigma\rho} \quad \bar{P}_{2\sigma\rho} ], \\ S_{1\sigma} &= \text{diag} [ S_{1i_n\sigma} ]_{n=1}^{N_\rho}, \quad S_{2\sigma} = \text{diag} [ S_{2i_n\sigma} ]_{n=1}^{N_\rho}, \\ S_{1i\sigma} &= F_{i\sigma} G_\sigma S_{i\nu} G_\sigma^\top F_{i\sigma}^\top, \quad S_{2i\sigma} = C F_{i\sigma} G_\sigma S_{i\nu} G_\sigma^\top F_{i\sigma}^\top C^\top. \end{aligned}$$

Полагая  $\gamma = 2 (\text{tr} [\bar{P}_{1\sigma\rho} S_{1\sigma}] + \text{tr} [\bar{P}_{2\sigma\rho} S_{2\sigma}])$ , получим, что условия теоремы 1 будут выполнены, если неравенство

$$\begin{aligned} \begin{bmatrix} \xi \\ \epsilon \end{bmatrix}^\top & \left( (\bar{A}_{\sigma\rho} + \bar{B}_{\sigma\rho} \bar{K}_{\sigma\rho} \bar{\mathcal{H}}_\rho)^\top \bar{P}_{\sigma\rho} (\bar{A}_{\sigma\rho} + \bar{B}_{\sigma\rho} \bar{K}_{\sigma\rho} \bar{\mathcal{H}}_\rho) - \bar{P}_{\sigma\rho} + \right. \\ & \left. + \bar{Q}_\rho + (\bar{K}_{\sigma\rho} \bar{\mathcal{H}}_\rho)^\top \bar{\mathcal{R}}_\rho \bar{K}_{\sigma\rho} \bar{\mathcal{H}}_\rho \right) \begin{bmatrix} \xi \\ \epsilon \end{bmatrix} \leq 0 \end{aligned}$$

разрешимо относительно положительно определенной матрицы  $\bar{P}_{\sigma\rho}$ . Это неравенство эквивалентно следующему матричному неравенству:

$$(5.2) \quad \begin{aligned} & (\bar{A}_{\sigma\rho} + \bar{B}_{\sigma\rho} \bar{K}_{\sigma\rho} \bar{\mathcal{H}}_\rho)^\top \bar{P}_{\sigma\rho} (\bar{A}_{\sigma\rho} + \bar{B}_{\sigma\rho} \bar{K}_{\sigma\rho} \bar{\mathcal{H}}_\rho) - \bar{P}_{\sigma\rho} + \bar{Q}_\rho + \\ & + (\bar{K}_{\sigma\rho} \bar{\mathcal{H}}_\rho)^\top \bar{R}_\rho \bar{K}_{\sigma\rho} \bar{\mathcal{H}}_\rho \preceq 0. \end{aligned}$$

Введем переменные

$$\begin{aligned} \bar{X}_{\sigma\rho} &= \text{diag} [\bar{X}_{1\sigma\rho} \quad \bar{X}_{2\sigma\rho}] = \bar{P}_{\sigma\rho}^{-1}, \quad \bar{X}_{1\sigma\rho} = I_{N_\rho} \otimes X_{1\sigma\rho}, \quad \bar{X}_{2\sigma\rho} = I_{N_\rho} \otimes X_{2\sigma\rho}, \\ \bar{Z}_{\sigma\rho} &= \text{diag} [\bar{Z}_{1\sigma\rho} \quad \bar{Z}_{2\sigma\rho}], \quad \bar{Z}_{1\sigma\rho} = I_{N_\rho} \otimes Z_{1\sigma\rho}, \quad \bar{Z}_{2\sigma\rho} = I_{N_\rho} \otimes Z_{2\sigma\rho}, \\ \bar{Y}_{\sigma\rho} &= [\bar{Y}_{1\sigma\rho} \quad \bar{Y}_{2\sigma\rho}] = \bar{K}_{\sigma\rho} \bar{Z}_{\sigma\rho}, \quad \bar{Y}_{1\sigma\rho} = I_{N_\rho} \otimes Y_{1\sigma\rho}, \quad \bar{Y}_{2\sigma\rho} = I_{N_\rho} \otimes Y_{2\sigma\rho}, \end{aligned}$$

где  $\bar{Z}_{\sigma\rho}$  является решением уравнения

$$\bar{Z}_{\sigma\rho} \bar{\mathcal{H}}_\rho = \bar{\mathcal{H}}_\rho \bar{X}_{\sigma\rho}.$$

Умножая (5.2) слева и справа на  $\bar{P}_{\sigma\rho}^{-1}$  и применяя лемму Шура о дополнении, получим следующую систему матричных уравнений и неравенств:

$$(5.3) \quad \begin{bmatrix} \bar{X}_{\sigma\rho} & (\bar{A}_{\sigma\rho} \bar{X}_{\sigma\rho} + \bar{B}_{\sigma\rho} \bar{Y}_{\sigma\rho} \bar{\mathcal{H}}_\rho)^\top & \bar{X}_{\sigma\rho} & (\bar{Y}_{\sigma\rho} \bar{\mathcal{H}}_\rho)^\top \\ \bar{A}_{\sigma\rho} \bar{X}_{\sigma\rho} + \bar{B}_{\sigma\rho} \bar{Y}_{\sigma\rho} \bar{\mathcal{H}}_\rho & \bar{X}_{\sigma\rho} & 0 & 0 \\ \bar{X}_{\sigma\rho} & 0 & \bar{Q}_\rho^{-1} & 0 \\ \bar{Y}_{\sigma\rho} \bar{\mathcal{H}}_\rho & 0 & 0 & \bar{R}_\rho^{-1} \end{bmatrix} \succcurlyeq 0,$$

$$\bar{Z}_{\sigma\rho} \bar{\mathcal{H}}_\rho = \bar{\mathcal{H}}_\rho \bar{X}_{\sigma\rho}, \quad \bar{X}_{\sigma\rho} \succ 0.$$

Таким образом, в рассмотренном случае закон УИО (3.1) с корректирующей поправкой (3.2) и матрицами протокола

$$K_{1\sigma\rho} = Y_{1\sigma\rho} Z_{1\sigma\rho}^{-1}, \quad K_{2\sigma\rho} = Y_{2\sigma\rho} Z_{2\sigma\rho}^{-1},$$

где  $Z_{1\sigma\rho}$ ,  $Z_{2\sigma\rho}$ ,  $Y_{1\sigma\rho}$  и  $Y_{2\sigma\rho}$  находятся из решения системы (5.3), обеспечивает выполнение условий сходимости (2.6). Матрицы  $\bar{Q}_\rho$  и  $\bar{R}_\rho$  здесь аналогичны весовым матрицам в теории линейно-квадратичного регулятора, изменяя которые можно корректировать управляющий сигнал и достичь желаемых характеристик.

## 5.2. Управление при переключениях режима

Как уже отмечалось, переключение режима приводит к появлению возмущающего воздействия, которое порождает переходный процесс. Этот процесс увеличивает достигнутое значение ошибки обучения на некоторое число шагов и замедляет сходимость. Поэтому на шаге переключения корректирующую поправку будем искать из условия минимизации ошибки. В случае с глобальным лидером введем функционал

$$J_{ml} = E[|\hat{\epsilon}_i(k+1, p)|^2 | (*)]$$

и будем искать корректирующую поправку из условия его минимума при ограничении, вытекающем из (3.10) и (3.12):

$$(5.4) \quad \begin{aligned} \hat{\epsilon}_i(k+1, p) = & -C \left( A_{\sigma_i(k+1)} \hat{x}_i(k+1, p-1) - A_{\sigma_i(k)} \hat{x}_i(k, p-1) \right) + \hat{\epsilon}_i(k, p) - \\ & - C \left( B_{\sigma_i(k+1)} - B_{\sigma_i(k)} \right) u_i(k, p-1) - C B_{\sigma_i(k+1)} \Delta u_i(k+1, p-1) - \\ & - C \left( F_{i\sigma_i(k+1)} G_{\sigma_i(k+1)} \nu_i(k+1, p-1) - F_{i\sigma_i(k)} G_{\sigma_i(k)} \nu_i(k, p-1) \right) + \\ & + \left( y_{\sigma_i(k+1)}^{ref}(p) - y_{\sigma_i(k)}^{ref}(p) \right), \end{aligned}$$

где символ (\*) означает, что математическое ожидание вычисляется при фиксированных значениях переменных состояния и управления в правой части (5.4). Полученная таким образом корректирующая поправка имеет вид

$$(5.5) \quad \begin{aligned} \Delta u_i(k+1, p-1) = & (C B_{\sigma_i(k+1)})^{-1} \hat{\epsilon}_i(k, p) - \\ & - (C B_{\sigma_i(k+1)})^{-1} C \left( A_{\sigma_i(k+1)} \hat{x}_i(k+1, p-1) - A_{\sigma_i(k)} \hat{x}_i(k, p-1) \right) - \\ & - (C B_{\sigma_i(k+1)})^{-1} C \left( B_{\sigma_i(k+1)} - B_{\sigma_i(k)} \right) u_i(k, p-1) + \\ & + (C B_{\sigma_i(k+1)})^{-1} \left( y_{\sigma_i(k+1)}^{ref}(p) - y_{\sigma_i(k)}^{ref}(p) \right) \end{aligned}$$

для закона УИО (3.1).

Корректирующая поправка для ведомого агента находится аналогичным способом, однако поскольку такие агенты не имеют непосредственный доступ к информации о желаемой траектории, в качестве ее образа будет выступать взвешенная сумма оценок выходных сигналов соответствующих локальных лидеров с предыдущего повторения. В связи с этим вместо (2.4) введем отклонение оценки выходного сигнала ведомого агента от оценок выходных сигналов его локальных лидеров

$$\hat{\epsilon}_i(k+1, p) = \sum_{j \in N_{\rho i}} s_{ij} (\hat{y}_j(k, p) - \hat{y}_i(k+1, p)),$$

где  $s_{ij}$  — элемент матрицы  $\mathcal{S}_\rho$ , и корректирующую поправку будем искать, решая задачу минимизации функционала

$$J_{mf} = E \left[ |\hat{\epsilon}_i(k+1, p)|^2 | (**) \right]$$

при ограничении

$$\begin{aligned}
 \hat{\varepsilon}_i(k+1, p) = & -\ell_{ii}C \left( A_{\sigma_i(k+1)}\hat{x}_i(k+1, p-1) - A_{\sigma_i(k)}\hat{x}_i(k, p-1) \right) - \\
 & -\ell_{ii}C \left( B_{\sigma_i(k+1)} - B_{\sigma_i(k)} \right) u_i(k, p-1) - \ell_{ii}CB_{\sigma_i(k+1)}\Delta u_i(k+1, p-1) + \\
 (5.6) \quad & + \sum_{j \in N_{\rho i}} s_{ij} (\hat{y}_j(k, p) - \hat{y}_i(k, p)) - \ell_{ii}C \left( F_{i\sigma_i(k+1)}G_{\sigma_i(k+1)}\nu_i(k+1, p-1) - \right. \\
 & \left. - F_{i\sigma_i(k)}G_{\sigma_i(k)}\nu_i(k, p-1) \right),
 \end{aligned}$$

где  $\ell_{ii}$  — элемент матрицы  $\mathcal{L}_\rho$ , и символ  $(**)$  означает, что математическое ожидание вычисляется при фиксированных значениях оценки состояния и управления в правой части (5.6). В результате получим

$$\begin{aligned}
 \Delta u_i(k+1, p-1) = & \ell_{ii}^{-1} \left( CB_{\sigma_i(k+1)} \right)^{-1} \sum_{j \in N_{\rho i}} s_{ij} (\hat{y}_j(k, p) - \hat{y}_i(k, p)) - \\
 (5.7) \quad & - \left( CB_{\sigma_i(k+1)} \right)^{-1} C \left( A_{\sigma_i(k+1)}\hat{x}_i(k+1, p-1) - A_{\sigma_i(k)}\hat{x}_i(k, p-1) \right) - \\
 & - \left( CB_{\sigma_i(k+1)} \right)^{-1} C \left( B_{\sigma_i(k+1)} - B_{\sigma_i(k)} \right) u_i(k, p-1).
 \end{aligned}$$

Поскольку локальные лидеры передают данные, полученные ими на предыдущем повторении, переключение ведомого агента системы должно происходить с запаздыванием относительно переключения его локального лидера, чтобы информация о желаемой траектории соответствовала новому режиму. По аналогии с сигналом  $\sigma_i(k)$ , который теперь будем именовать *локальным сигналом, переключающим режим*, введем *глобальный сигнал*  $\sigma(k)$ , который запускает процесс переключения режимов агентов. Моменты переключения  $\sigma_i(k)$  глобальных лидеров совпадают с моментами переключения  $\sigma(k)$ , т.е.  $\sigma_i(k) = \sigma(k) \forall i : r_i = 1$ . Для ведомых агентов локальный сигнал будет запаздывать относительно глобального сигнала:  $\sigma_i(k) = \sigma_j(k-1) \forall i, j : r_i = 0, j \in N_{\rho i}$ .

Таким образом, в момент переключения режима будем применять закон УИО (3.1) с корректирующими поправками (5.5) и (5.7) для глобальных лидеров и ведомых агентов соответственно. При переключении режима будем учитывать изложенное правило.

### 5.3. Управление при переключении топологии сети

Рассмотрим последний случай, когда происходит переключение топологии сети. При этом все агенты работают в одном и том же режиме на всех рассматриваемых повторениях, поэтому для краткости сигналы, переключающие режимы, обозначим через  $\sigma$ .

В момент переключения топологии агент может выполнить одно из следующих действий: подключиться к сети путем подключения к функционирующим агентам, отключиться от сети или сменить локальных лидеров. Во

втором и третьем случаях управление агентами производится согласно закону УИО (3.1) с корректирующей поправкой (3.2) с матрицами протокола, соответствующими новой топологии. В первом случае ошибка подключаемого агента при использовании этого же закона может существенно отличаться от достигнутой другими агентами в течение нескольких повторений после подключения.

Пусть  $k + 1$  — момент переключения топологии с подключением агента  $i$ . С целью снизить ошибку агента при подключении, для поиска закона управления применим тот же подход, что и в предыдущем разделе, однако в этом случае проще найти не корректирующую поправку, а непосредственно управление. Введем функционал

$$J_c = E \left[ \left| \hat{\varepsilon}_i(k+1, p) \right|^2 | (***) \right]$$

и будем искать управление для подключаемого агента из условия его минимума при ограничении

$$(5.8) \quad \begin{aligned} \hat{\varepsilon}_i(k+1, p) = & \sum_{j \in N_{\rho(k+1)i}} s_{ij} \hat{y}_j(k, p) - \ell_{ii} C A_{\sigma_i(k+1)} \hat{x}_i(k+1, p-1) - \\ & - \ell_{ii} C B_{\sigma_i(k+1)} u_i(k+1, p-1) - \ell_{ii} C F_{i\sigma_i(k+1)} G_{\sigma_i(k+1)} \nu_i(k+1, p-1), \end{aligned}$$

где  $s_{ij}$  и  $\ell_{ii}$  — элементы матриц  $\mathcal{S}_{\rho(k+1)}$  и  $\mathcal{L}_{\rho(k+1)}$  соответственно и символ  $(***)$  означает, что математическое ожидание вычисляется при фиксированных значениях оценки состояния и управления в правой части (5.8). Полученный таким образом закон управления для подключаемой системы имеет вид

$$(5.9) \quad \begin{aligned} u_i(k+1, p-1) = & \ell_{ii}^{-1} (C B_{\sigma_i(k+1)})^{-1} \sum_{j \in N_{\rho(k+1)i}} s_{ij} \hat{y}_j(k, p) - \\ & - (C B_{\sigma_i(k+1)})^{-1} C A_{\sigma_i(k+1)} \hat{x}_i(k+1, p-1). \end{aligned}$$

#### 5.4. Общий закон управления

Объединим полученные результаты в виде следующих правил переключения режима работы агентов, топологии информационной сети и закона управления.

Переключения инициируются сигналами  $\sigma(k)$  и  $\rho(k)$ . Сигнал  $\sigma(k)$ , именуемый глобальным сигналом, переключающим режим, запускает процесс переключения режимов агентов. Переключение режима глобальных лидеров происходит в момент запуска процесса переключения глобальным сигналом, т.е.  $\sigma_i(k) = \sigma(k) \forall i : r_i = 1$ . Переключение остальных агентов происходит с запаздыванием в одно повторение после переключения их локальных лидеров, т.е.  $\sigma_i(k) = \sigma_j(k-1) \forall i, j : r_i = 0, j \in N_{\rho(k)i}$ . Сигнал  $\rho(k)$  переключает топологию сети.

Управляющий сигнал на повторении  $k + 1$  формируется в виде (3.1) с корректирующей поправкой (3.2), матрицы протокола которой находятся из решения системы (5.3), если  $\sigma_i(k + 1) = \sigma_i(k)$  и  $i \in \mathcal{I}_{\rho(k)}$ , с корректирующей поправкой (5.5), если  $r_i = 1$ ,  $\sigma_i(k + 1) \neq \sigma_i(k)$  и  $i \in \mathcal{I}_{\rho(k)}$ , и поправкой (5.7), если  $r_i = 0$ ,  $\sigma_i(k + 1) \neq \sigma_i(k)$  и  $i \in \mathcal{I}_{\rho(k)}$ , где  $r_i$  — элемент матрицы  $\mathcal{R}_{\rho(k+1)}$ . Если  $i \notin \mathcal{I}_{\rho(k)}$ , то управляющий сигнал на повторении  $k + 1$  формируется в виде (5.9).

Остается доказать условие ограниченности управления в среднем квадратичном (2.7). Рассмотрим интервал повторений до первого переключения. Из (2.5) имеем

$$(5.10) \quad u_i(k, p-1) = (CB_\sigma)^{-1} \left[ CA\hat{x}_i(k, p) - CA_\sigma^p \hat{x}_i(k, 0) - \sum_{q=0}^{p-2} CA_\sigma^{p-1-q} B_\sigma u_i(k, p) - \sum_{q=0}^{p-1} CA_\sigma^{p-1-q} (F_\sigma C \tilde{x}_i(k, q) + F_\sigma G_\sigma \nu_i(k, p)) \right].$$

В частности, при  $p = 2$

$$u_i(k, 1) = (CB_\sigma)^{-1} \left[ CA\hat{x}_i(k, 2) - CA_\sigma^2 \hat{x}_i(k, 0) - CA_\sigma B_\sigma u_i(k, 0) - \sum_{q=0}^1 CA_\sigma^{p-1-q} (F_\sigma C \tilde{x}_i(k, q) + F_\sigma G_\sigma \nu_i(k, p)) \right].$$

Возведем обе части последнего равенства в квадрат и оценим правую часть в соответствии с известным неравенством между квадратом суммы и суммой квадратов  $\left( (\sum_{i=1}^n a_i)^2 \leq n \sum_{i=1}^n a_i^2 \right)$ . Применяя к полученному неравенству оператор математического ожидания с учетом того, что в силу (4.3)  $E[|C\hat{x}_i(k, 2)|^2] < \infty$ , и величины  $\|\hat{x}_i(k, 0)\|^2$ ,  $|u_i(k, 0)|^2$ ,  $\|\tilde{x}_i(k, q)\|^2$  и  $|\nu_i(k, q)|^2$  имеют ограниченные математические ожидания, получим, что  $E[|u_i(k, 1)|^2] < \infty$ . Последовательно изменяя в (5.10)  $p$  в пределах от 3 до  $T$ , получим, что  $E[|u_i(k, p)|^2] < \infty$  для всех  $k$  до первого переключения. Последовательно рассматривая интервалы между переключениями, с учетом того, что в соответствии с подходом к выбору управления в моменты переключений в начале каждого интервала справедливо (4.3) аналогично получим, что  $E[|u_i(k, p)|^2] < \infty$  для всех  $k$  до последнего переключения на шаге  $k_f$ . Повторяя аналогичную процедуру для  $k \geq k_f$  и устремляя  $k$  к  $\infty$ , убеждаемся в справедливости (2.7). Таким образом, полученное управление обеспечивает консенсус в смысле выполнения условий (2.6), (2.7).

## 6. Пример

Рассмотрим сетевую систему, состоящую из одинаковых манипуляторов с поворотным гибким звеном, управление которой построим в соответствии с полученными результатами. Динамика манипулятора с учетом случайных возмущений описывается следующей моделью в пространстве состояний с непрерывным временем:

$$(6.1) \quad \dot{x}_i(k, t) = A_{\sigma_i(k)}^{cont} x_i(k, t) + B^{cont} (u_i(k, t) + \mu_i(k, t)),$$

$$(6.2) \quad y_i(k, t) = C x_i(k, t) + \rho_i(k, t),$$

где  $x = [\theta \ \alpha \ \dot{\theta} \ \dot{\alpha}]^\top$ ,  $\theta$  — угол поворота сервопривода,  $\alpha$  — угол отклонения звена,  $u = \tau$  — крутящий момент на силовом редукторе сервопривода, приводящем звено в движение,  $\mu$  и  $\rho$  — независимые непрерывные гауссовские белые шумы с постоянными интенсивностями  $Q_n$  и  $R_n$  соответственно. Матрицы системы задаются следующим образом:

$$A_{\sigma}^{cont} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & K_s/J_{eq} & -B_{eq}/J_{eq} & 0 \\ 0 & -K_s(J_l(\sigma) + J_{eq})/J_{eq}J_l(\sigma) & B_{eq}/J_{eq} & 0 \end{bmatrix},$$

$$B^{cont} = [0 \ 0 \ 1/J_{eq} \ -1/J_{eq}]^\top, \quad C = [1 \ 0 \ 0 \ 0],$$

где  $K_s$  — жесткость звена,  $J_{eq}$  — момент инерции сервопривода,  $B_{eq}$  — коэффициент вязкого трения сервопривода,  $J_l(\sigma)$  — момент инерции звена относительно центра масс,  $\sigma = \sigma_i(k)$ .

Были приняты следующие параметры манипуляторов с гибким звеном:  $K_s = 1,3 \text{ н} \cdot \text{м/рад}$ ,  $J_{eq} = 2,08 \cdot 10^{-3} \text{ кг} \cdot \text{м}^2$ ,  $B_{eq} = 0,004 \text{ н} \cdot \text{м}/(\text{рад/с})$  [25],  $Q_n = 5 \cdot 10^{-5}$ ,  $R_n = 10^{-6}$ .

Динамика агентов имеет два режима, в каждом из которых определена своя желаемая траектория выходного сигнала

$$y_{\sigma_i(k)}^{ref}(t) = \begin{cases} \pi \left( \frac{t^2}{6} - \frac{t^3}{27} \right), & \sigma_i(k) = 1, \\ \frac{\pi}{2} \sin \frac{\pi t}{6}, & \sigma_i(k) = 2, \end{cases}$$

кроме того, режимы отличаются моментом инерции звена

$$J_l(\sigma_i(k)) = \begin{cases} 0,0038 \text{ кг} \cdot \text{м}^2, & \sigma_i(k) = 1, \\ 0,008 \text{ кг} \cdot \text{м}^2, & \sigma_i(k) = 2. \end{cases}$$



Продолжительность повторения составляет 3 с. Перейдем от модели (6.1)–(6.2) к эквивалентной дискретной модели вида (2.1)–(2.3), в которой матрицы определяются соотношениями

$$A_{\sigma_i(k)} = \exp A_{\sigma_i(k)}^{cont} T_s, \quad B_{\sigma_i(k)} = \int_0^{T_s} \exp \left( A_{\sigma_i(k)}^{cont} \tau \right) B^{cont} d\tau,$$

$$D_{\sigma_i(k)} = \left[ \int_0^{T_s} \exp \left( A_{\sigma_i(k)}^{cont} \tau \right) B^{cont} Q_n (B^{cont})^\top \left( \exp \left( A_{\sigma_i(k)}^{cont} \tau \right) \right)^\top d\tau \right]^{\frac{1}{2}},$$

$$G_{\sigma_i(k)} = \left( \frac{R_n}{T_s} \right)^{\frac{1}{2}},$$

и шумы  $\omega_i(k, p)$  и  $\nu_i(k, p)$  имеют единичные ковариации  $S_{i\omega} = I_{n_x}$  и  $S_{i\nu} = 1$ . Здесь  $T_s$  — период дискретизации, который выбран равным 0,01 с.

Рассмотрим сетевую систему из трех манипуляторов, среди которых выделен один глобальный лидер, и сначала работает только он, затем к глобальному лидеру подключается первый ведомый агент, и далее к первому ведомому агенту подключается второй. Такой сценарий соответствует производственной системе с изменением объема производства. Когда объем производства увеличивается, происходит подключение новых агентов, при уменьшении — отключение лишних. Соответствующая этой программе изменяемая топология сети описывается следующим образом:

$$\mathcal{I}_{\rho(k)} = \{1\}, \quad \mathcal{L}_{\rho(k)} = 0, \quad \mathcal{R}_{\rho(k)} = 1 \text{ при } \rho(k) = 1,$$

$$\mathcal{I}_{\rho(k)} = \{1, 2\}, \quad \mathcal{L}_{\rho(k)} = \begin{bmatrix} 0 & 0 \\ -1 & 1 \end{bmatrix}, \quad \mathcal{R}_{\rho(k)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ при } \rho(k) = 2,$$

$$\mathcal{I}_{\rho(k)} = \{1, 2, 3\}, \quad \mathcal{L}_{\rho(k)} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}, \quad \mathcal{R}_{\rho(k)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ при } \rho(k) = 3.$$

Для фильтра Калмана были получены следующие матрицы усиления:

$$F_{i\sigma_i(k)} = \begin{bmatrix} 0,7106 & -0,5711 & 16,0723 & -12,8866 \end{bmatrix}^\top \text{ при } \sigma_i(k) = 1,$$

$$F_{i\sigma_i(k)} = \begin{bmatrix} 0,7038 & -0,6084 & 15,5888 & -13,9268 \end{bmatrix}^\top \text{ при } \sigma_i(k) = 2.$$

Решая (5.3) с весовыми матрицами

$$Q_{1\rho(k)} = \begin{cases} 10I_{n_x}, & \rho(k) = 1, \\ 10^{-4}I_{n_x}, & \rho(k) \neq 1, \end{cases}$$

$$Q_{2\rho(k)} = 10^5, \quad R_{\rho(k)} = 10^{-3},$$

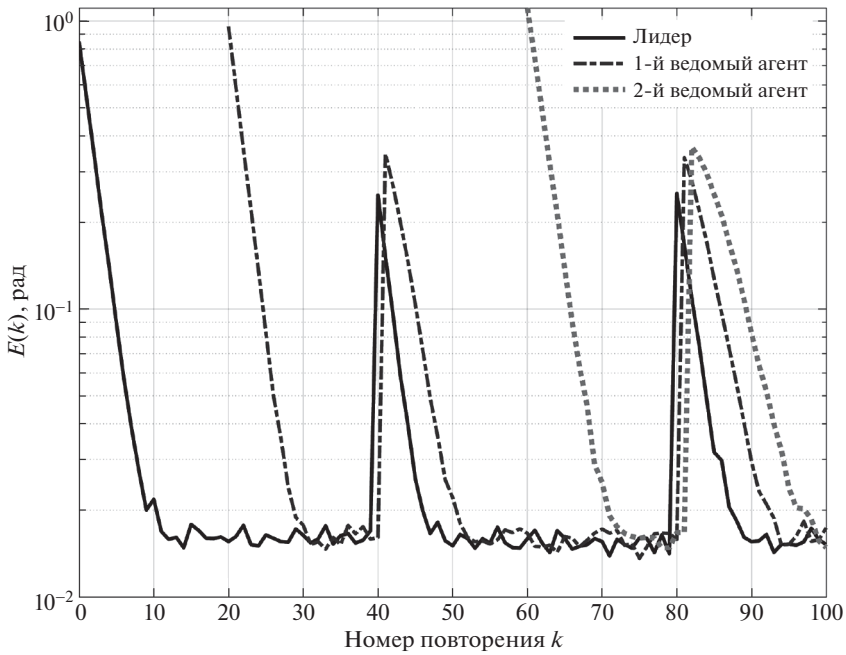


Рис. 1. Среднеквадратические ошибки обучения агентов без переключения закона управления.

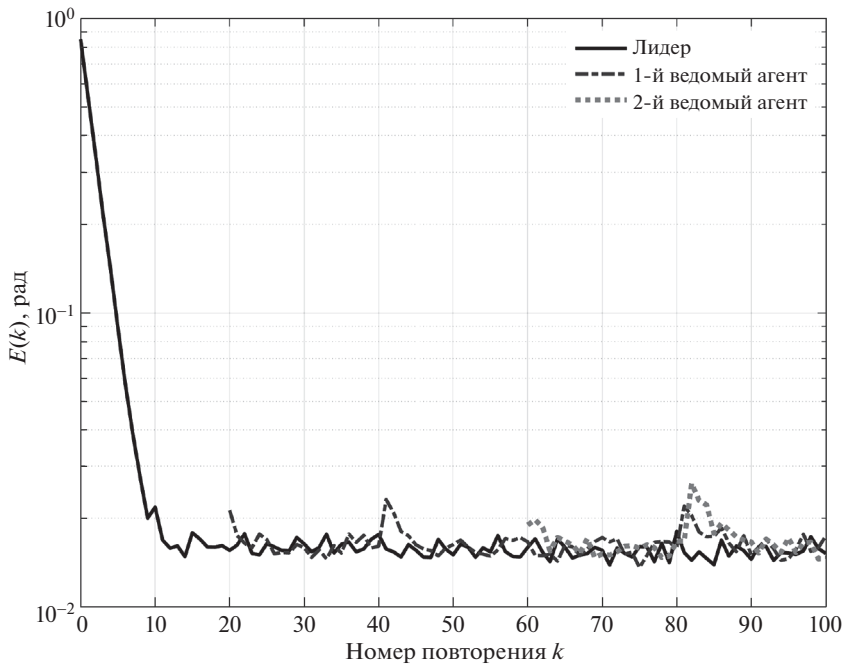


Рис. 2. Среднеквадратические ошибки обучения агентов с переключением закона управления.

получим следующие матрицы протокола для (3.2):

$$K_{1\sigma\rho} = \begin{bmatrix} -25,3429 & -1,2471 & -0,3469 & -0,0181 \end{bmatrix},$$

$$K_{2\sigma\rho} = 9,2965 \text{ при } \sigma = 1 \text{ и } \rho = 1,$$

$$K_{1\sigma\rho} = \begin{bmatrix} -41,8954 & -1,2959 & -0,416 & -0,0044 \end{bmatrix},$$

$$K_{2\sigma\rho} = 16,2678 \text{ при } \sigma = 1 \text{ и } \rho = 2,$$

$$K_{1\sigma\rho} = \begin{bmatrix} -41,9283 & -1,296 & -0,4161 & -0,0044 \end{bmatrix},$$

$$K_{2\sigma\rho} = 14,3248 \text{ при } \sigma = 1 \text{ и } \rho = 3,$$

$$K_{1\sigma\rho} = \begin{bmatrix} -25,3841 & -1,2813 & -0,3428 & -0,0139 \end{bmatrix},$$

$$K_{2\sigma\rho} = 9,2971 \text{ при } \sigma = 2 \text{ и } \rho = 1,$$

$$K_{1\sigma\rho} = \begin{bmatrix} -41,889 & -1,2981 & -0,4159 & -0,0044 \end{bmatrix},$$

$$K_{2\sigma\rho} = 16,4272 \text{ при } \sigma = 2 \text{ и } \rho = 2,$$

$$K_{1\sigma\rho} = \begin{bmatrix} -41,9326 & -1,2982 & -0,4161 & -0,0044 \end{bmatrix},$$

$$K_{2\sigma\rho} = 14,2538 \text{ при } \sigma = 2 \text{ и } \rho = 3,$$

где  $\sigma = \sigma_i(k)$  и  $\rho = \rho(k)$ . Переключающие сигналы зададим в виде

$$\sigma(k) = \begin{cases} 1, & k < 40, \\ 2, & 40 \leq k < 80, \\ 1, & k \geq 80, \end{cases} \quad \rho(k) = \begin{cases} 1, & k < 20, \\ 2, & 20 \leq k < 60, \\ 3, & k \geq 60. \end{cases}$$

Для оценки эффективности алгоритма введем среднеквадратическую ошибку обучения

$$E_i(k) = \sqrt{\frac{1}{T} \sum_{p=0}^{T-1} |e_i(k, p)|^2}.$$

На рис. 1 представлены графики среднеквадратических ошибок обучения агентов без учета переключения управления при изменении режима работы и топологии сети, т.е. управление происходит согласно закону (3.1) с корректирующей поправкой (3.2) на всем протяжении функционирования системы, а на рис. 2 — с учетом переключения управления. Данные результаты демонстрируют, что разработанный закон управления действительно позволяет снизить величину переходной ошибки в моменты переключения.

## 7. Заключение

Предложенный алгоритм УИО позволяет снизить величину переходной ошибки в моменты переключения режима и подключения новых агентов, однако он накладывает некоторые ограничения на топологию сети. Во-первых, невозможен взаимный обмен информацией между агентами. В самом деле,

для переключения режима работы ведомого агента необходимо, чтобы режим работы его локальных лидеров на предыдущем повторении соответствовал тому, на который происходит переключение. По этой причине при взаимном обмене агенты будут ожидать смены режима друг друга, и в итоге оно не произойдет ни для одного из них. В рассмотренном примере такая ситуация возникла бы, если первый ведомый агент передавал свой выходной сигнал второму, а второй — первому. В этом случае первый ведомый агент не смог бы произвести переключения из-за ожидания переключения второго, а второй из-за ожидания первого. Также по этой причине невозможно реализовать замкнутую цепь информационного обмена. Во-вторых, в случае применения данного алгоритма к системе с последовательным соединением агентов процесс переключения всей сетевой системы на новый режим работы может занять недопустимо длительное время в связи с задержками переключения ведомых агентов.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Saez M.A., Maturana F.P., Barton K., Tilbury D.M.* Context-sensitive modeling and analysis of cyber-physical manufacturing systems for anomaly detection and diagnosis // *IEEE Transaction on Automation Science and Engineering*. 2020. V. 17. No. 1. P. 29–40.
2. *Balta E.C., Tilbury D.M., Barton K.* Switch Based Iterative Learning Control for Tracking Iteration Varying References // *IFAC PapersOnLine*. 2020. V 53. No. 2. P. 1493–1498.
3. *Tsytkin Ya.Z.* Adaptation and Learning in Automatic Systems. New York: Academic Press, 1971.
4. *Arimoto S., Kawamura S., Miyazaki F.* Bettering operation of robots by learning // *Journal of Robotic Systems*. 1984. V. 1. No. 2. P. 123–140.
5. *Bristow D.A., Tharayil M., Alleyne A.* A survey of iterative learning control // *IEEE Control Systems Magazine*. 2006. V. 26. No. 3. P. 96–114.
6. *Ahn H.S., Chen Y.Q., Moore K.L.* Iterative learning control: Brief survey and categorization // *IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews*. 2007. V. 37. No. 6. P. 1099–1121.
7. *Pakshin P., Emelianova J., Emelianov M.* Iterative learning control of stochastic linear systems under switching of the reference trajectory and parameters // *Proceedings of the 29th Mediterranean Conference on Control and Automation*. 2021. P. 1311–1316. 2021. Bari, Puglia, Italy.
8. *Pakshin P., Emelianova J., Rogers E., Galkowski K.* Iterative learning control of stochastic linear systems with reference trajectory switching // *Proceedings of the 60th IEEE Conference on Decision and Control*. 2021. P. 6565–6570.
9. *Ahn H.S., Chen Y.Q.* Iterative learning control for multi-agent formation // *In Proc. ICROS-SICE Int. Joint Conf.* 2009. P. 3111–3116.
10. *Liu Q., Bristow D.A.* An iteration-domain filter for controlling transient growth in iterative learning control // *Proc. 2010 Amer. Control Conf.* 2010. P. 2039–2044.
11. *Liu Y., Jia Y.* An iterative learning approach to formation control of multi-agent systems // *Systems & Control Letters*. 2012. V. 61. P. 148–154.

12. *Yang S., Xu J.X., Huang D., Tan Y.* Optimal iterative learning control design for multi-agent systems consensus tracking // *Systems & Control Letters*. 2014. V. 69 P. 80–89.
13. *Li J., Li J.* Adaptive iterative learning control for coordination of second-order multi-agent systems // *Int. J. Robust Nonlinear Control*. 2014. V. 24. P. 3282–3299.
14. *Meng D., Du W., Jia Y.* Data-driven consensus control for networked agents: an iterative learning control-motivated approach // *IET Control Theory & Applications*. 2015. V. 9. P. 2084–2096.
15. *Yu X., Hou Z., Polycarpou M.M.* Distributed Data-Driven Iterative Learning Consensus Tracking for Nonlinear Discrete-Time Multiagent Systems // *IEEE Transactions on Automatic Control*. 2022. V. 67. No. 7. P. 3670–3677.
16. *Hock A., Schoellig A.* Distributed iterative learning control for multi-agent systems // *Autonomous Robots*. 2019. V. 43. P. 1989–2010.
17. *Pakshin P.V., Emelianova J.P., Emelianov M.A.* (2018). Iterative learning control design for multiagent systems based on 2D models // *Automation and Remote Control*. 2018. V. 79. No. 6. P. 1040–1056.
18. *Pakshin P.V., Koposov A.S., Emelianova J.P.* Iterative learning control of a multiagent system under random perturbations // *Automation and Remote Control*, (2020). 81(3), 483–502.
19. *Ahn H.S., Moore K.L., Chen Y.Q.* Iterative Learning Control. Robustness and Monotonic Convergence for Interval Systems. *Lecture Notes in Control and Information Sciences*. Springer-Verlag: London, 2007.
20. *Hock A., Schoellig A.* Distributed iterative learning control for a team of quadrotors // *Proceedings of the 55th IEEE Conference on Decision and Control*. 2016. P. 4640–4646.
21. *Sun S., Endo T., Matsuno F.* Iterative learning control based robust distributed algorithm for non-holonomic mobile robots formation // *IEEE Access*. 2018. V. 6. P. 61904–61917.
22. *Koposov A., Emelianova J., Pakshin P.* Iterative learning control of multi-agent systems under changing network configuration // *IFAC PapersOnLine*, 2021. V. 54. No. 20. P. 669–674.
23. *Koposov A., Emelianova J., Pakshin P.* Iterative learning control of multi-agent systems under changing reference trajectory // *IFAC PapersOnLine*, 2022. V. 55. No. 12. P. 759–764.
24. *Pakshin P., Emelianova J.* Iterative learning control design for discrete-time stochastic switched systems // *Autom. Remote Control*. 2020. V. 81. No. 11. P. 2011–2025.
25. *Apkarian J., Karam P., Levis M.* Workbook on Flexible Link Experiment for MatlabR/Simulink Users. Quanser, 2011.

*Статья представлена к публикации членом редколлегии А.И. Кибзуном.*

Поступила в редакцию 22.11.2022

После доработки 07.02.2023

Принята к публикации 30.03.2023