

# Условия существования петель скрытой обратной связи в рекомендательных системах с учётом шума

А. А. Пилькевич<sup>1</sup>, А. С. Хританков<sup>2</sup>  
anton39reg@mail.ru; anton.khritankov@phystech.edu

В работе исследуются петли скрытой обратной связи в рекомендательных системах. Под положительной обратной связью подразумевается неограниченный рост интереса пользователя к предлагаемым объектам. Решается задача поиска условий возникновения положительной обратной связи. Учитывается наличие шума в выборе пользователя. Рекомендательная система использует алгоритм Thomson Sampling Multi-armed Bandit. В задачах без шума известно, что существуют условия неограниченного роста. Но отсутствие шума не реализуется в реальных системах. Экспериментально проверяются полученные условия в имитационной модели.

The paper explores hidden feedback loops in recommender systems. A positive feedback loop is an unlimited growth of user interest in proposed objects. The paper looks for conditions for positive feedback loops. The paper takes into account a noise in user responses. The recommender system uses Thompson Sampling Multi-armed Bandit algorithm. In noise-free problems other works improved that conditions for unlimited growth exist. But noise-free is not true in a practice. The paper carries out simulation experiments to check found conditions.

**Ключевые слова:** *machine learning, hidden feedback loops, filter bubble, thompson sampling*

**DOI:**

## 1 Введение

Рекомендательные системы являются важной составляющей социальных сетей, веб-поиска и других сфер [5,6,7,8,9]. Рассматриваются петли скрытой обратной связи, которые подразумевает рост качества предсказаний, как результат учёта принятых решений. Эффект петель скрытой обратной связи в реальных и модельных задачах в публикациях [7,8,9] описывается как нежелательное явление. Частные и часто рассматриваемые случаи скрытой обратной являются echo chamber и filter bubbles [1,9]. До сих пор нет строгой формализации условий возникновения этих эффектов при условиях приближенных к реальности [1,2,5,6].

Целью данной работы является нахождение условий существования петель обратной связи в рекомендательной системе с алгоритмом Thomson Sampling в условиях зашумлённости выбора пользователя. Зашумлённость выбора рассматривается, как смещение первоначального интереса к исходному объект или категории. Предлагается способ отыскание условий модели исходя из теоретических свойств алгоритма TS. Под условиями подразумеваются параметры шума и параметры рекомендательной системы. Для описания условий предлагается выражение для математического ожидания интереса. Также рассматривается вариант нахождения этих условий чисто из экспериментов. Целью является математическое описание искомых условий с дальнейшим экспериментальным подтверждением полученных условий. Для проверки результатов используется имитационная модель, использующая синтетические данные.

Ранее проблема изучалась с другой стороны - как преодолеть смещение распределения входных данных и сделать алгоритм лучше [5,6]. В этой работе важны изменения,

которые работа алгоритма привносит в данные. Важно, что источник изменений - сам алгоритм

Существует ранее описанная модель [1] петель в случае отсутствия шума в действиях пользователя. Подобное исследование проводилось в статье [1] на примере различных моделей ( Oracle, Optimal Oracle, UCB, TS ) в задаче многорукого бандита. Удалось показать условия существования неограниченного роста интереса пользователя. В работе [2] изучалась схожая постановка задачи и были получены условия возникновения, но рассматривалась линейная модель и градиентный бустинг. Но отсутствие шума в ответах пользователей в работах [1,2] не реализуется на практике. Важным отличием данной работы является факт рассмотрения более сложных условий модели, таких как шум в выборе пользователя и другой алгоритм рекомендательной системы.

В работе предлагается анализ роста интереса пользователя. Рассматривается математическое ожидание изменения интереса. Полученные условия проверяются в вычислительном эксперименте.

## 2 Петли скрытой обратной связи

Целью работы является теоретический анализ условий сходимости TS для различных параметров шума и экспериментальное подтверждение полученных соотношений. Также делается уточнение условий из [1].

### 2.1 Модель рекомендательной системы

Обозначим за  $t$  очередной момент выдачи рекомендаций. Рекомендательная система на шаге  $t$  выбирает элементы  $(a_t^1, \dots, a_t^l)$  из конечного набора  $M$ . Истинный *интерес* пользователя к элементу  $a \in M$  описывается неизвестным отображением  $\mu_t : M \rightarrow \mathbb{R}$ . При этом считается, что чем больше значение  $\mu_t(a)$ , тем заинтересованнее пользователь в рекомендации  $a$ .

После очередного набора рекомендаций  $a_t = (a_t^1, \dots, a_t^l)$  пользователь возвращает *отклик*  $c_t = (c_t^1, \dots, c_t^l)$ ,  $c_t^i \in \{0, 1\}$ . Предполагается, он выбирает элементы  $c_t^i$  случайно и независимо, пропорционально  $\mu_t(a)$ . Значит отклик имеет распределение Бернулли :

$$c_t^i \sim \text{Bern}(\sigma(\mu_t(a_t^i))), \text{ где } \sigma(x) = \frac{1}{1 + e^x} - \text{сигмоида}.$$

Предполагаем, что интерес пользователя во времени описывается как

$$\begin{cases} \mu_{t+1} \geq \mu_t, & \text{если } c_t = 1, \\ \mu_{t+1} < \mu_t, & \text{если } c_t = 0, \\ \mu_{t+1} = \mu_t, & \text{если элемент не попал в рекомендацию.} \end{cases}$$

Тогда петля обратной связи выражается как

$$\lim_{t \rightarrow \infty} \|\mu_t - \mu_0\|_2 = \infty. \quad (1)$$

Обновление интереса для элементов очередной рекомендации происходит по правилу:

$$\mu_{t+1} - \mu_t = \delta_t c_t - \delta_t (1 - c_t), \text{ где } \delta_t \sim U[0, 0.01]. \quad (2)$$

Оптимизационной задачей рекомендательной системы является задача минимизации потерь. Максимальная сумма наград :

$$\max_{c_t^i} \sum_{t=1}^T \sum_{i=1}^l c_t^i = T \cdot l.$$

Тогда задача ставится так :

$$T \cdot l - \sum_{t=1}^T \sum_{i=1}^l c_t^i \rightarrow \min_b,$$

где  $b$  — используемый алгоритм в рекомендательной системе.

## 2.2 Алгоритм рекомендательной системы

Задача многорукого бандита состоит из  $k$  бандитов и системы взаимодействующей с ними. Каждый бандит имеет собственное распределение неизвестное для системы. Система "дёргает" за ручки бандита и получает награду из соответствующего распределения бандита. Задачей системы является максимизации суммы наград или же минимизации потерь.

В данной задаче рекомендательная система использует алгоритм Thompson Sampling [3] для задачи бернуллиевского бандита. Бандитами являются отклики пользователя  $c_t^i$  на очередную рекомендацию. Средняя награда равна:  $\sigma(\mu_t(a_t^i))$ .

В начальный момент времени определены вероятности бернуллиевских случайных величин  $c_t^i$  для элементов  $M$  равные  $\pi_0(\theta_1), \dots, \pi_0(\theta_m)$ . Задаётся априорное распределение для  $\theta_i$  равное бэта-распределению  $Beta(1, 1) = U[0, 1]$ . Апостериорное распределение для элемента  $a^i \in M$  описывается бэта-распределением:  $Beta(\alpha_t^i, \beta_t^i)$ . Параметры после очередной рекомендации обновляются по закону :

$$\alpha_{t+1} = \alpha_t + c_t, \beta_{t+1} = \beta_t + 1 - c_t. \quad (3)$$

## 2.3 Учёт аддитивного шума в поведении пользователя

Шум откликов описывается следующим образом:

$$c_t^i \sim Bern(\sigma(\mu_t(a_t^i) + q_t^i)), \\ q_t^i \sim U[-w, w].$$

Наличие  $q_t^i$  позволяет описать несмещённый аддитивный шум, то есть отклонение от истинного интереса пользователь.

## 3 Теоретическое обоснование

Назовём *режимом работы TS с фиксированными лидерами* поведение алгоритма, в котором TS не меняются элементы рекомендаций.

**Утверждение 1.** Пусть  $p = 1$  и TS работает в режиме с фиксированными лидерами начиная с какого-то момента времени  $\tau$ . Тогда при  $w \geq 0$ :  $\lim_{t \rightarrow \infty} \|\mu_t - \mu_0\|_2 = \infty$ .

Или же при фиксированном  $p = 1$  и любых параметрах шума  $w$  возникает петля скрытой обратной связи.

## 4 Вычислительный эксперимент

Целью эксперимента является подтверждение существования петель скрытой обратной связи для произвольных параметров шума  $w$ . Важной частью эксперимента является сравнения поведения рекомендательной системы с шумом в ответах пользователя и без.

### 4.1 Описание данных и работы модели

Перед началом эксперимента фиксируются следующие параметры:  $T$  — число итераций рекомендательной системы,  $|M|$  — число рассматриваемых объектов для рекомендации,

94  $l$  — число элементов в одной выдаче. Также фиксируются параметры шума  $p, w, u$ . Да-  
 95 лее случайным образом сэмплируются начальные значения интереса  $\{\mu_0^i\}_{i=1}^{|M|}$ . Параметры  
 96 априорного распределения  $\{\alpha_0^i, \beta_0^i\}_{i=1}^{|M|}$  также семплируются случайно.

97 Генерация элементов очередной рекомендации производится на основе текущего апо-  
 98 стериорного распределения. Выбираются элементы с наибольшим значением. Получение  
 99 отклика от пользователя заключается в генерации случайных величин на основе рекомен-  
 100 дации. Обновление параметров апостериорного распределения происходит по правилу (3).  
 101 Интерес обновляется согласно (2).

102 Также рассматривается вариант эксперимента, когда используется случайная модель  
 103 генерации рекомендации. В этом случае  $l$  элементов для очередной рекомендации выби-  
 104 раются случайным образом.

105 В каждый момент выдачи  $t$  фиксируются значения интереса  $\mu_t^i$ , сумма откликов  $c_t^i$  и  
 106 параметры апостериорного распределения. По полученным данным строятся графики для  
 107 определения наличия петель (1) скрытой обратной связи (см. рис. 1).

## 108 4.2 Псевдокод проведения эксперимента

109 **Вход:**  $M, l, T, w, p$

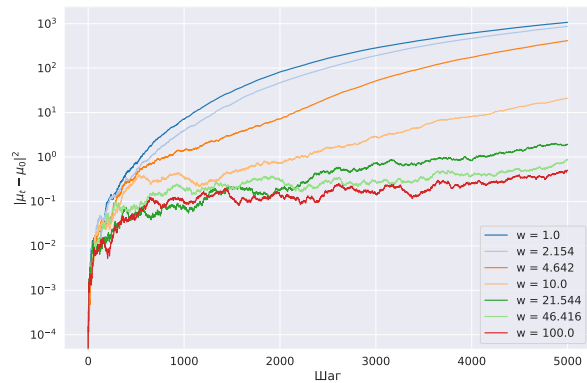
```
110 BanditLoopExperiment.prepare()
111 для  $t$  от 1 до  $T$ 
112    $r_t \leftarrow \text{TSBandit.predict}()$ 
113    $c_t \leftarrow \text{make\_response\_noise}(r_t, w, p)$ 
114    $\text{TSBandit.update}(c_t)$ 
115    $\text{Model.interest\_update}(c_t)$ 
116    $\text{save\_iter}(t, c_t, \mu_t)$ 
```

## 117 5 Результаты

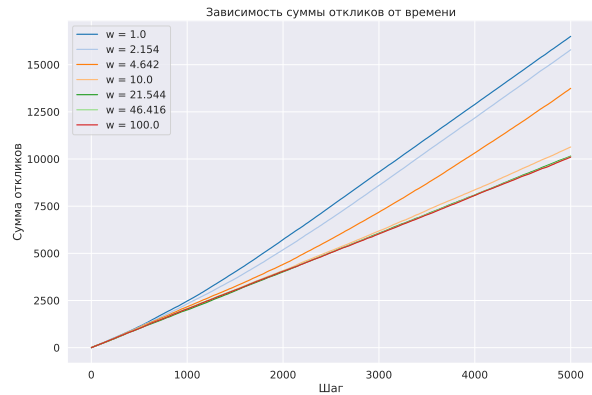
118 На рис. 1 изображена зависимость нормы разности начального значения интереса и  
 119 интереса в момент времени  $0 \leq t \leq 5000$ , используется логарифмический масштаб На  
 120 рис. 2 изображена сумма наград  $c_t$ . Рассматриваются различные параметры аддитивно-  
 121 го шума  $w$ . Видно, что наблюдается эффект неограниченного роста интереса даже для  
 122 больших значений шума. Причём величина шума никак не ограничивает рост интереса, а  
 123 лишь замедляет его, что согласуется с определением петли и утверждением 1.

124 На рис. 3 изображён разброс значений нормы интереса для 30 запусков эксперимента.

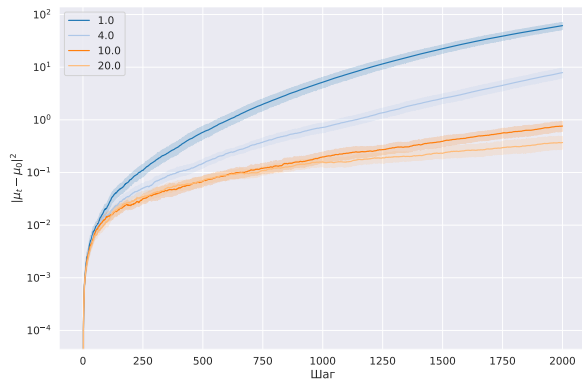
125 На рис. 4 сравниваются рекомендательные системы со случайным алгоритмом и TS.  
 126 Для случайной модели тоже наблюдается образование петли (см. рис. 4). Она более хао-  
 127 тична, но тренд неограниченного роста интереса всё равно присутствует.



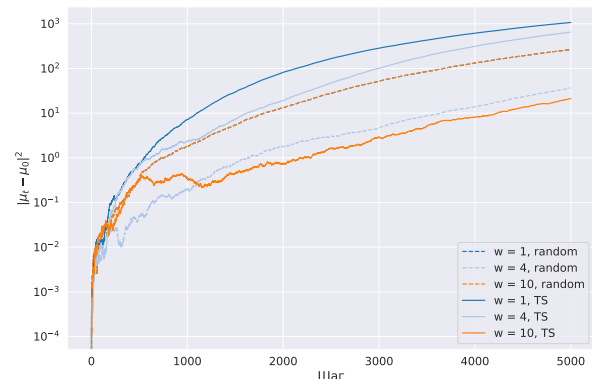
**Рис. 1** Логарифм нормы интереса на очередном шаге рекомендации.



**Рис. 2** Суммы наград на очередном шаге рекомендации.



**Рис. 3** Разброс логарифма нормы интереса от шага.



**Рис. 4** Сравнение случайного алгоритма и TS.

## 6 Заключение

Поставлена задача существования петель скрытой обратной связи при наличии шума в ответах пользователя. Для текущей модели шума было получено, что при любых параметрах возникают петли. Это также подтверждается в эксперименте.

В дальнейшем требуется проверить гипотезу о возникновении петель при любом несмещённом аддитивном шума. Также стоит рассмотреть другие модели шума. Так как после определённого значения интереса из-за сигмной любое влияние шума сводилось на нет.

## Литература

- [1] Ray Jiang, Silvia Chiappa, Tor Lattimore, András György, Pushmeet Kohli Degenerate Feedback Loops in Recommender Systems// CoRR, 2019, Vol. abs/1902.10730, URL: <https://arxiv.org/abs/1902.10730>.

- [2] *Khrutankov, Anton* Hidden Feedback Loops in Machine Learning Systems: A simulation Model and Preliminary Results// Springer, 2021, P. 54–65.
- [3] *Daniel Russo, Benjamin Van Roy, Abbas Kazerouni, Ian Osband* A Tutorial on Thompson Sampling// CoRR, 2017, Vol. abs/1707.02038, URL: <https://arxiv.org/abs/1707.02038>.
- [4] *Shipra Agrawal, Navin Goyal* Analysis of Thompson Sampling for the multi-armed// CoRR, 2011, Vol. abs/1111.1797, URL: <https://arxiv.org/abs/1111.1797>.
- [5] *Giuseppe Burtini, Jason L. Loepky, Ramon Lawrence* Improving Online Marketing Experiments with Drifting Multi-armed Bandits// SciTePress, 2018, P. 630–636.
- [6] *David Krueger and Tegan Maharaj and Jan Leike* Hidden Incentives for Auto-Induced Distributional Shift// CoRR, 2020, Vol. abs/2009.09153.
- [7] *Wilbert Samuel Rossi, Jan Willem Polderman, Paolo Frasca* The closed loop between opinion formation and personalised recommendations// CoRR, 2018, Vol. abs/1809.04644.
- [8] *Pedreschi, D. and Miliou, I. and European Parliament. Directorate-General for Internal Policies of the Union* Artificial Intelligence (AI): new developments and innovations applied to e-commerce// European Parliament, 2020.
- [9] *Dominic DiFranzo, Kristine Gloria-Garcia* Filter bubbles and fake news// XRDS, 2017.
- [10] *Пилькевич Антон, Хританков Антон* Условие существования петель скрытой обратной связи в рекомендательных системах с учётом шума// URL: [github.com/Intelligent-Systems-Phystech/2021-Project-74](https://github.com/Intelligent-Systems-Phystech/2021-Project-74).

Поступила в редакцию

## 7 Приложение

**Утверждение 2.** Пусть  $p = 1$  и TS работает в особом режиме начиная с какого-то момента времени  $\tau$ . Тогда при  $w \geq 0$ :  $\lim_{t \rightarrow \infty} \|\mu_t - \mu_0\|_2 = \infty$ .

Или же при фиксированном  $p = 1$  и любых параметрах шума  $w$  возникает петля скрытой обратной связи.

**Доказательство.** Так как алгоритм работает в особом режиме, то при  $t \geq \tau$  известно какие объекты он будет рекомендовать. Для случая нормы интересов:

$$\|\mu_t - \mu_0\|_2^2 = \sum_{i=1}^M (\mu_t^i - \mu_0^i)^2,$$

с ростом  $t$  основной вклад будут давать только  $l < M$  объектов попавших в рекомендацию. Причём эти объекты известны и не меняются для очередного шага.

Рассмотрим изменение интереса для произвольного  $a \in M$ . Обновление интереса происходит согласно:  $\mu_t - \mu_{t-1} = \delta_t c_t - \delta(1 - c_t)$ . Случайные величины  $\delta_t, c_t$  независимы, поэтому:

$$\mathbb{E} \delta_t c_t = \mathbb{E} \delta_t \mathbb{E} c_t.$$

Для удобства будем считать, что у нас  $c_t \sim \text{Bern}_{\pm}(\sigma(s_t \cdot \mu_t(a_t) + q_t))$  Тогда:

$$\begin{aligned} \mathbb{E}(c_t | s_t = x, q_t = y) &= 2\sigma(x \cdot \mathbb{E} \mu_{t-1} + y) - 1, \\ \mathbb{E}(\mathbb{E}(c_t | s_t, q_t = y)) &= p \cdot (2\sigma(\mathbb{E} \mu_{t-1} + y) - 1) + (1 - p) \cdot (2\sigma(-\mathbb{E} \mu_{t-1} + y) - 1). \end{aligned}$$

170 В случае  $E(E(c_t|s_t, q_t)) > 0$  петля будет возникать, так как рост интереса в среднем  
171 положителен.

Далее для простоты считается, что  $\sigma(x) \approx \left(\frac{x}{4} + \frac{1}{2}\right) \cdot I[-2, 2] + I[2, \infty]$  и  $p = 1$ . Задача в этом случае записывается так:

$$E(E(c_t|s_t, q_t = y)) \approx 2 \left( \frac{E\mu_{t-1} + y}{4} + \frac{1}{2} \right) - 1.$$

172 Теперь петля возникает при условии:  $E\sigma(x) > \frac{1}{2}$ .

Тогда остаётся посчитать:

$$\begin{aligned} E\sigma(\mu_t) &\approx \int_{-\infty}^{\infty} \left( \frac{E\mu_t + y}{4} + \frac{1}{2} \right) I\{-2 < E\mu_t + y < 2\} f(y) dy \\ &\quad + \int_{-\infty}^{\infty} I\{2 < E\mu_t + y\} f(y) dy = \\ &\quad \int_{-2}^2 \left( \frac{z}{4} + \frac{1}{2} \right) f_s(z) dz + \int_2^{\infty} f_s(z) dz, \end{aligned}$$

173 где  $f_s(z)$  плотность  $U[E\mu_t - w, E\mu_t + w]$ . Таким образом у нас возникает 6 случаев.

1.  $E\mu_t + w < -2$ . Тогда, очевидно:

$$E\sigma(\mu_t) = 0 \rightarrow E(E(c_t|s_t, q_t)) = -1.$$

174 В этом случае интерес бесконечно убывает. Так как рассматривается норма интересов,  
175 то всё равно  $(\mu_t - \mu_0)^2 \rightarrow \infty$  при  $t \rightarrow \infty$ .

2.  $E\mu_t - w < -2 < E\mu_t + w < 2$ . Тогда:

$$\begin{aligned} E\sigma(\mu_t) &= \frac{1}{16w} (y + 2)^2 \Big|_{-2}^{E\mu_t + w} = \frac{1}{16w} (E\mu_t + w + 2)^2 < \frac{1}{2}, \\ (E\mu_t + w + 2)^2 &< 8w, \\ \begin{cases} E\mu_t < -w - 2 + \sqrt{8w}, \\ E\mu_t > -w - 2 - \sqrt{8w}, \end{cases} &\rightarrow \text{рост.} \end{aligned}$$

176 В случае  $E\sigma(\mu_t) > \frac{1}{2}$  система будет несовместна.

3.  $E\mu_t - w < -2, E\mu_t + w > 2$ .

$$\begin{aligned} E\sigma(\mu_t) &= \frac{1}{16w} (y + 2)^2 \Big|_{-2}^2 + \frac{1}{2w} (E\mu_t + w - 2) = \\ \frac{1}{w} + \frac{E\mu_t + w}{2w} - \frac{1}{w} &= \frac{E\mu_t + w}{2w} > \frac{1}{2} \Rightarrow \\ E\mu_t > 0, w > 2 &\rightarrow \text{рост.} \end{aligned}$$

4.  $E\mu_t - w > -2, E\mu_t + w < 2$ . Тогда:

$$\begin{aligned} E\sigma(\mu_t) &= \frac{1}{16w} (y + 2)^2 \Big|_{E\mu_t - w}^{E\mu_t + w} > \frac{1}{2}, \\ (E\mu_t + w + 2)^2 - (E\mu_t - w + 2)^2 &> 8w, \\ (2E\mu_t + 4) \cdot 2w &> 8w, \\ E\mu_t > 0 &\rightarrow \text{рост.} \end{aligned}$$

5.  $E\mu_t - w > -2, E\mu_t + w > 2$ . Тогда:

$$\begin{aligned}
 E\sigma(\mu_t) &= \frac{1}{16w} (y+2)^2 \Big|_{E\mu_t-w}^2 + \frac{1}{2w} \Big|_{E\mu_t-w}^{E\mu_t+w} = \\
 &= \frac{1}{16w} (16 - (E\mu_t - w + 2)^2) + \frac{1}{2w} (E\mu_t + w - 2) = \\
 &= \frac{1}{w} - \frac{(E\mu_t - w + 2)^2}{16w} + \frac{E\mu_t + w}{2w} - \frac{1}{w} = \\
 &= -\frac{1}{16w} (E^2\mu_t - 2(w-2)E\mu_t + (w-2)^2) + \frac{E\mu_t + w}{2w} > \frac{1}{2} \Rightarrow \\
 &E^2\mu_t - 2(w-2)E\mu_t + (w-2)^2 - 8(E\mu_t + w) + 8w < 0, \\
 &E^2\mu_t - 2(w+2)E\mu_t + (w-2)^2 < 0, \\
 &(E\mu_t - (w+2))^2 - (w+2)^2 + (w-2)^2 < 0, \\
 &(E\mu_t - (w+2))^2 - 8w < 0, \\
 &\begin{cases} E\mu_t < w+2 + \sqrt{8w}, \\ E\mu_t > w+2 - \sqrt{8w}, \end{cases} \rightarrow \text{рост.}
 \end{aligned}$$

6.  $E\mu_t - w > 2$ . Тогда:

$$E\sigma(\mu_t) = 1 > \frac{1}{2}.$$

Во всех случаях при заданных условиях удалось отделить от нуля изменение интереса:

$$E(\mu_t - \mu_{t-1}) = E\delta_t \cdot E(E(c_t|s_t, q_t)) > 0.$$

Поэтому существует  $m > 0$  такое, что

$$E(\mu_t - \mu_{t-1}) = E\delta_t \cdot E(E(c_t|s_t, q_t)) > m > 0.$$

Тогда

$$E^2(\mu_t - \mu_0) > t^2 \cdot m^2 \rightarrow \infty \text{ при } t \rightarrow \infty.$$

Раз одно слагаемое стремится к бесконечности, то:

$$E\|\mu_t - \mu_0\|_2^2 \rightarrow \infty \text{ при } t \rightarrow \infty.$$