



Факультет
Компьютерных Наук

Оптимизация для инженерного проектирования

Фомичева Н.В., Чжоу Ц., Рамазян Т.А.
НИУ ВШЭ



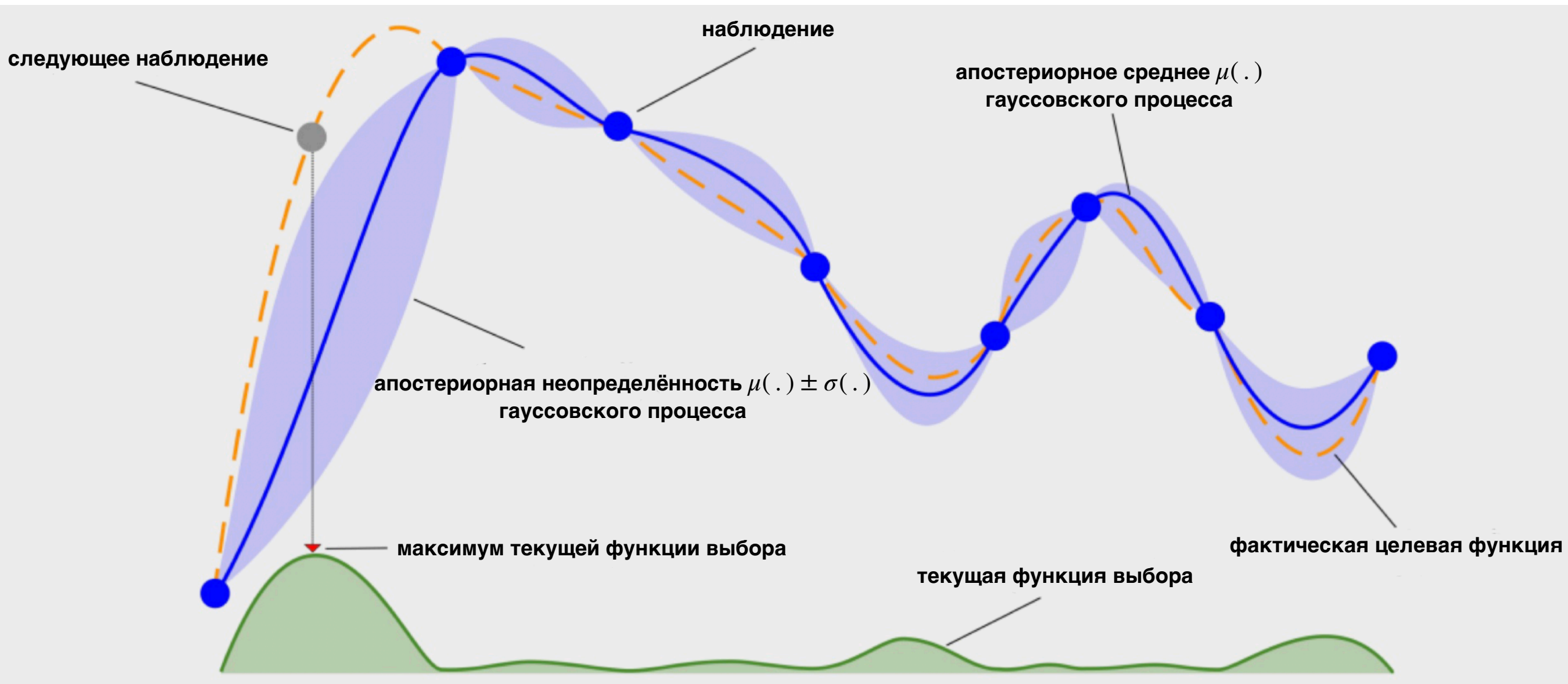


Мотивация

- Реальные эксперименты симулируются из первых принципов, что вычислительно дорого
- Симуляция необходима для поиска оптимального дизайна экспериментов
- Байесовская оптимизация максимально эффективно позволяет воспользоваться дорогостоящей симуляцией

Цель

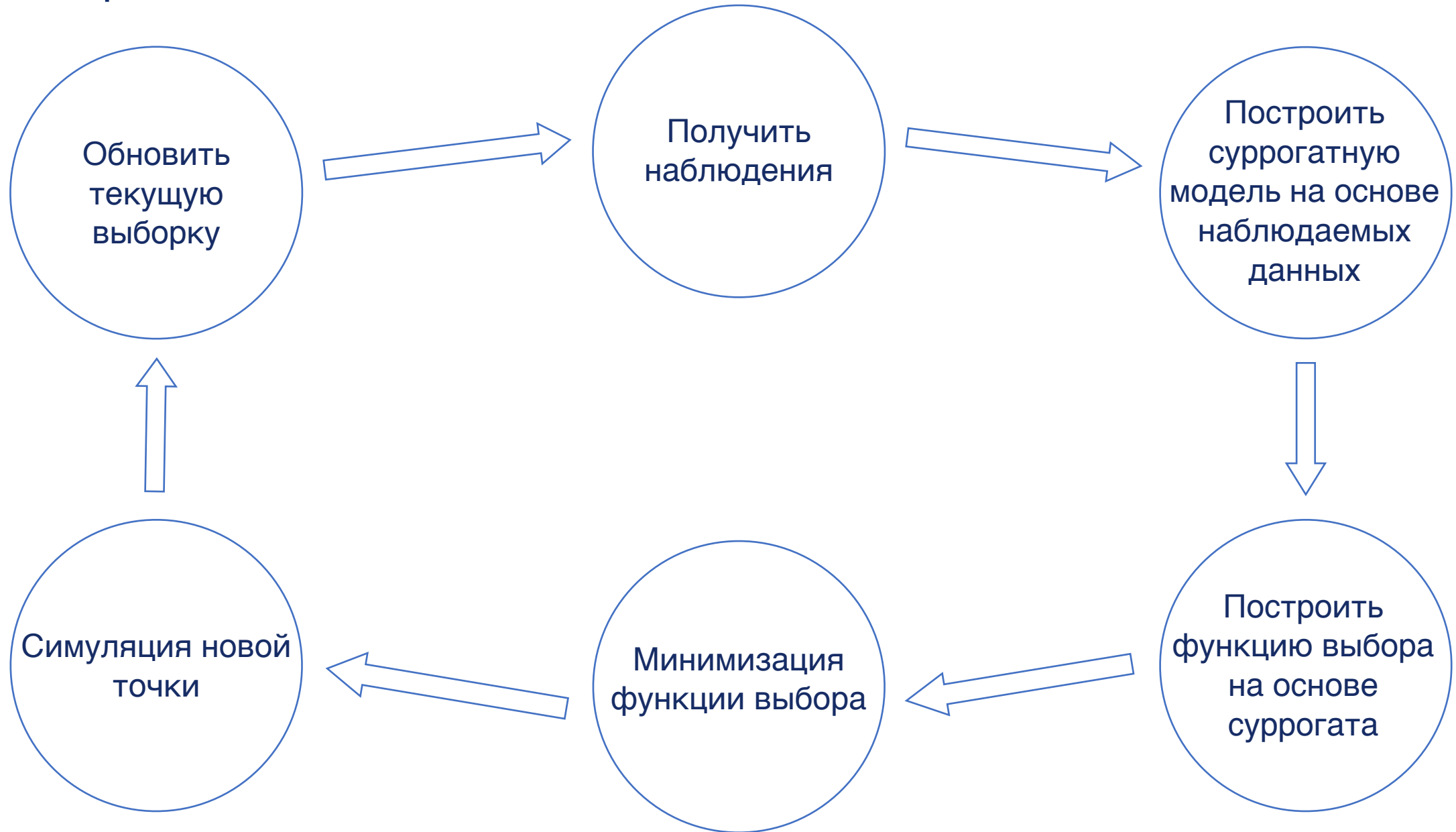
Проверить на инженерных задачах разные вариации Байесовской оптимизации для установления лучшего подхода.





Алгоритм Байесовской оптимизации

4



Подходы к оптимизации функции выбора

5

Использующие градиенты:

- Gradient Descent
- Adam

Эволюционные алгоритмы:

- Дифференциальная эволюция

Методы Ньютона:

- L-BFGS-B
- Внутренняя оптимизация BoTorch

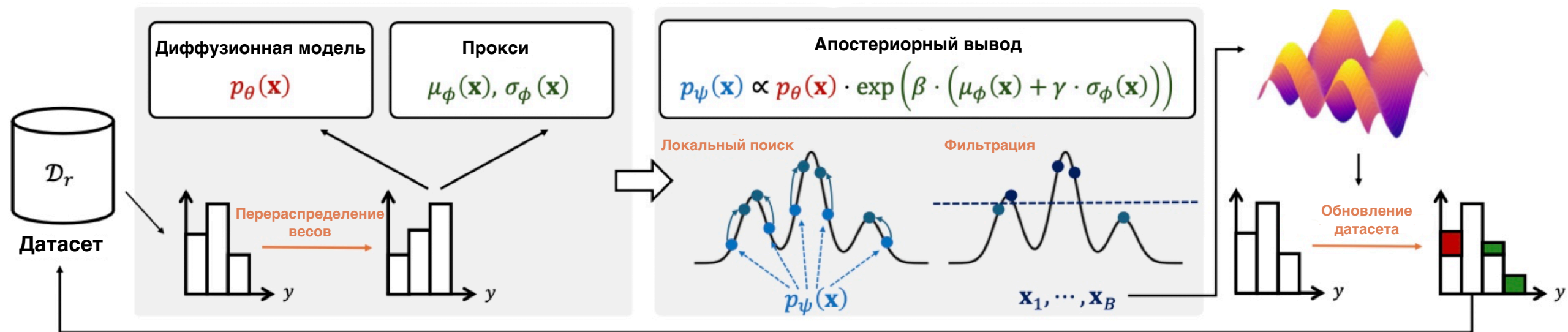
На основе диффузионных моделей:

- DiBo

Фаза 1. Обучение моделей

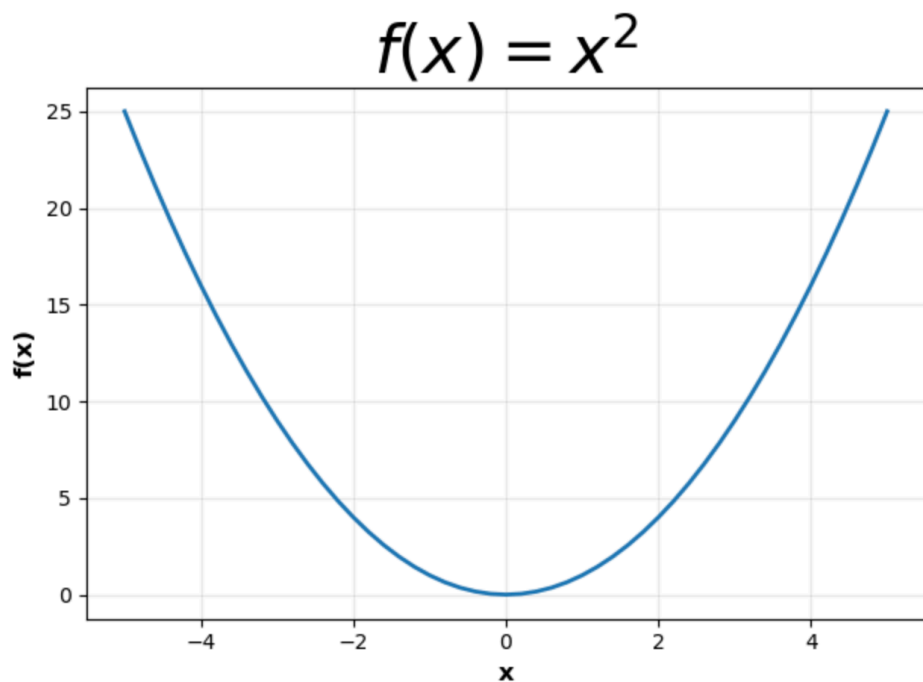
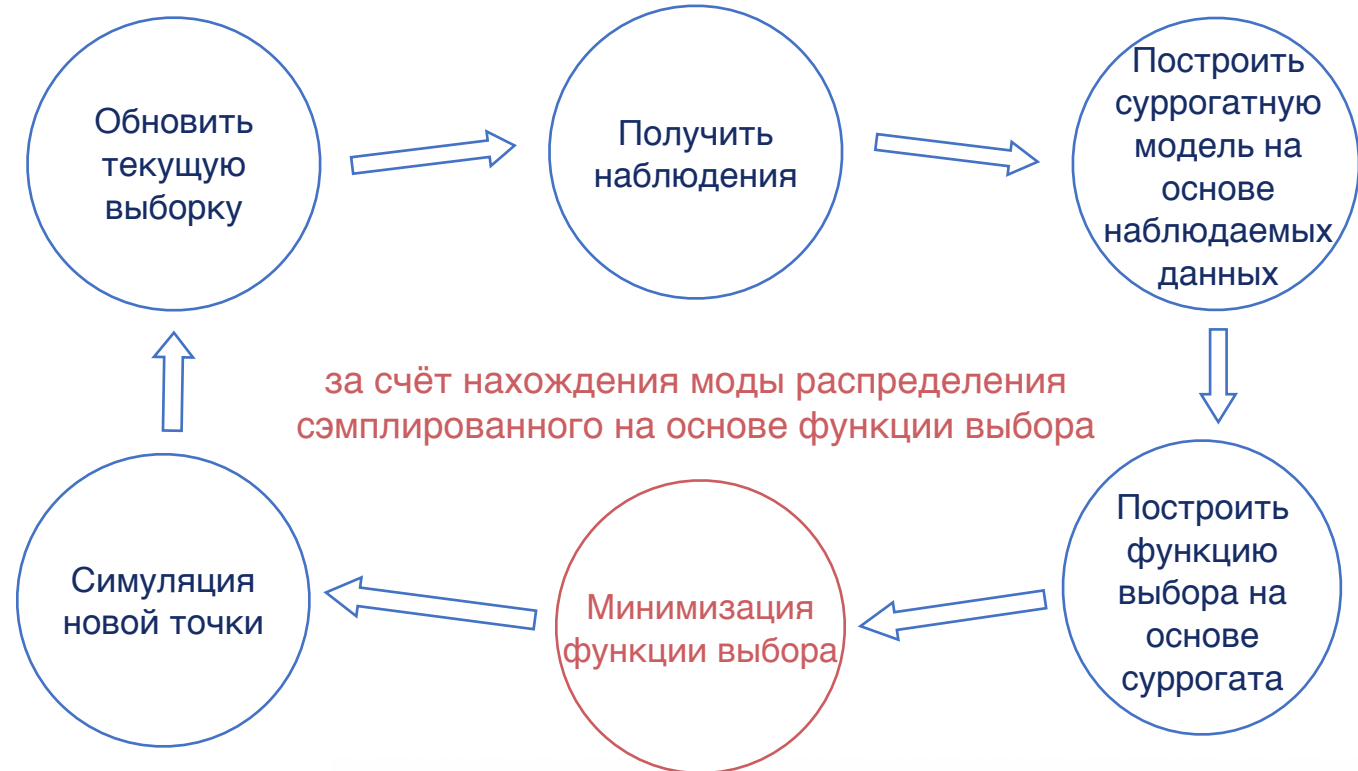
Фаза 2. Сэмплирование кандидатов

Симуляция новых точек

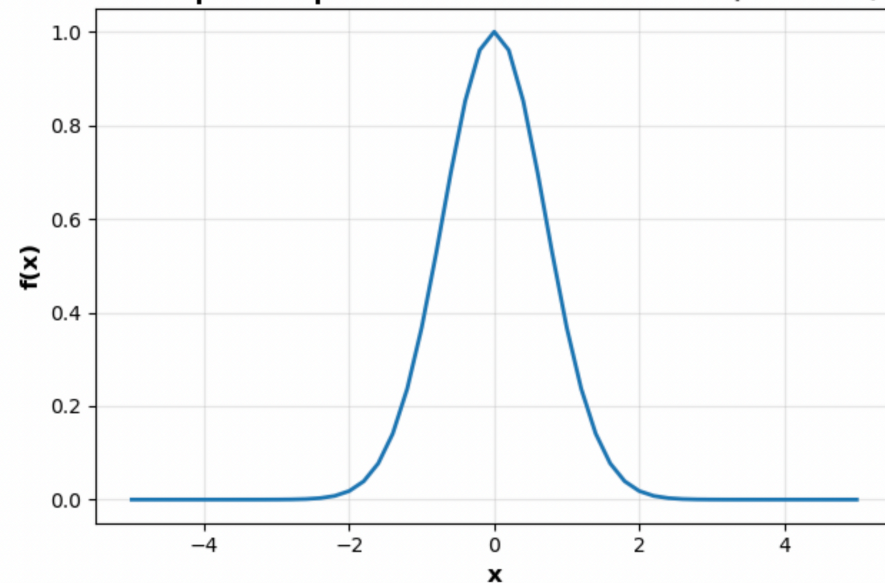


Использование MCMC

Рассматривается вариант **оптимизации функции выбора с помощью алгоритмов сэмплирования** вместо стандартных подходов поиска оптимума.



После преобразования Больцмана, e^{-x^2}



Использование динамики Ланжевена

7

ϵ **точность** в нахождении оптимума функции выбора при использовании **ULA** возможна за $\mathcal{O}(d/\epsilon)$, и за $\mathcal{O}(d^2 \ln(1/\epsilon))$, если использовать **MALA**, в то время как **остальные** алгоритмы сходятся за $\Omega((1/\epsilon)^d)$, где d - размерность эксперимента.

Предполагается, что эти результаты достижимы при использовании **Lower Confidence Bound (LCB)**

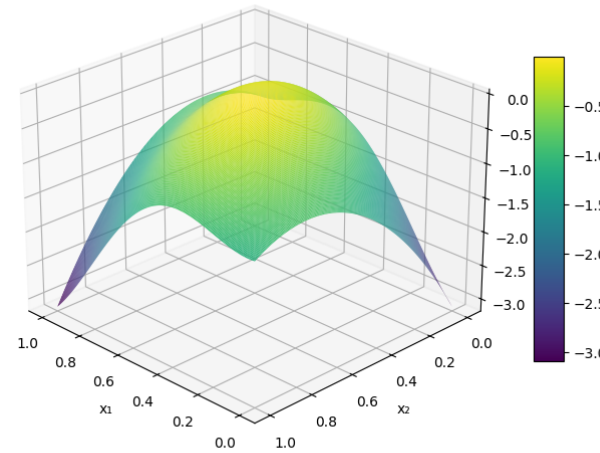
$$\alpha_t(x) = \mu_t(x) - \beta_t \sigma_t(x)$$

в качестве **функции выбора**,

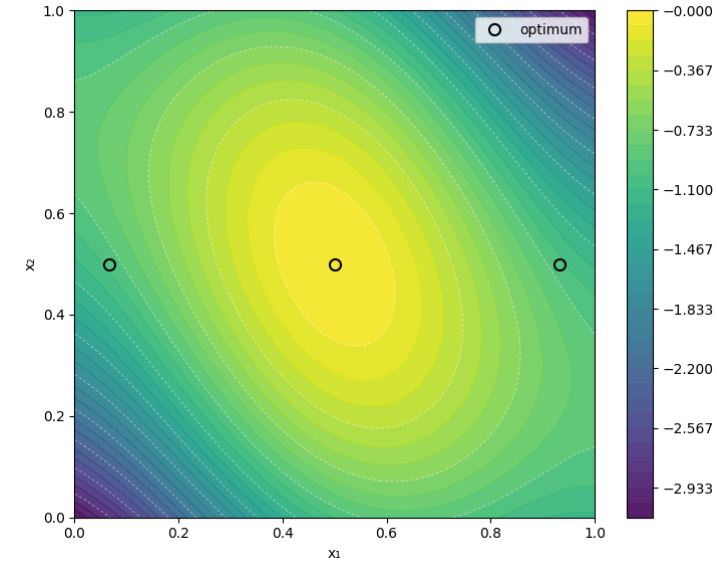
и **Гауссовского процесса**, определённого **RBF** ядром, в качестве суррогатной модели:

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{\|x - x'\|^2}{2\ell^2}\right)$$

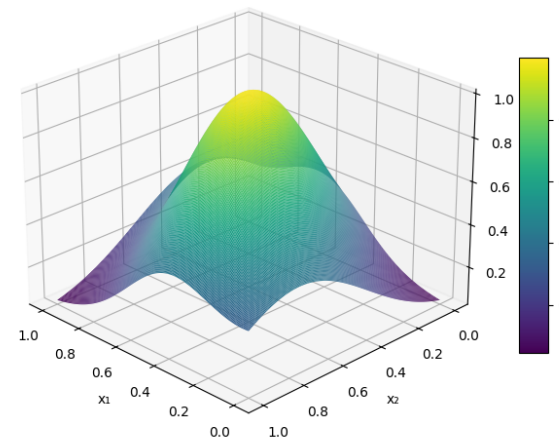
Функция трёхгорбого верблюда



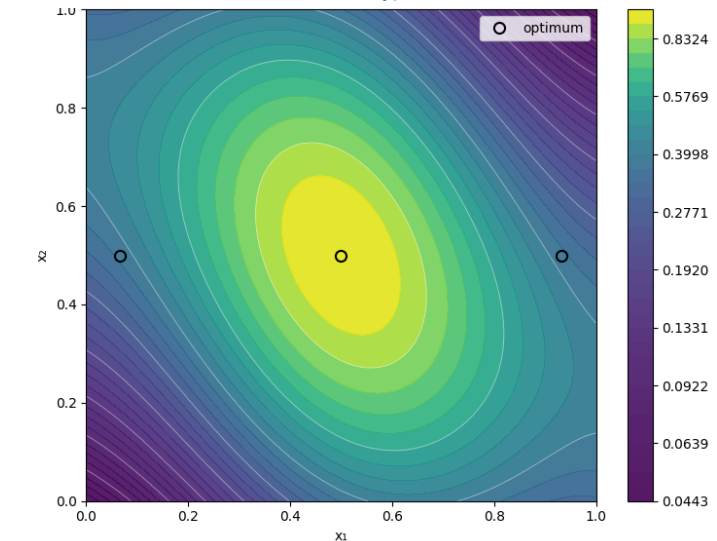
Вид контура



Функция трёхгорбого верблюда после преобразования Больцмана



Вид контура

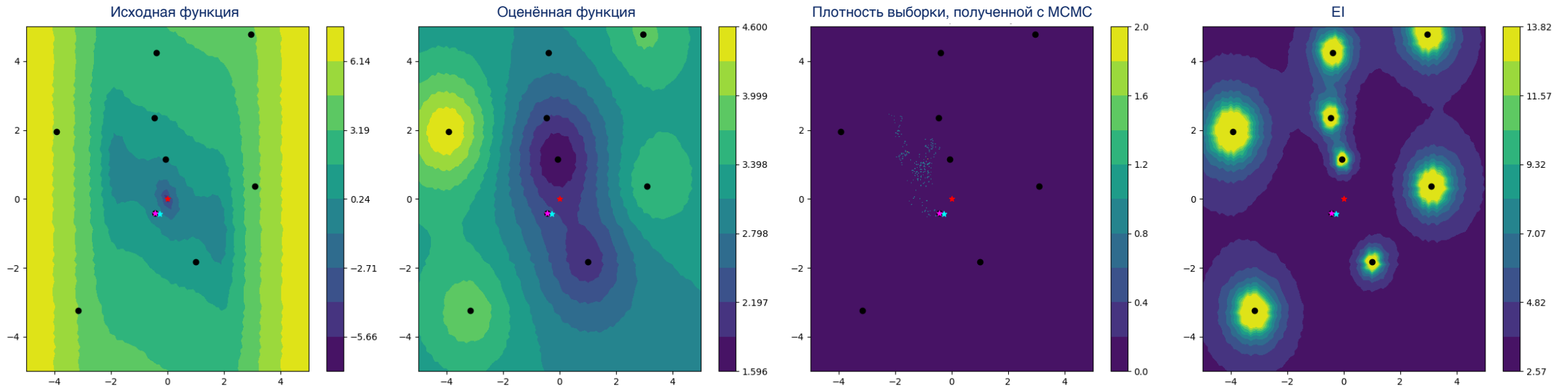


Марковские цепи Монте-Карло:

- ULA
 - MALA
 - ULMC
 - HMC
 - Gibbs
- } Динамика Ланжевена

Тесты для проектирования:

- Пружина растяжения-сжатия
- Сосуд под давлением
- Сварная балка
- Редуктор



Результаты 5й итерации Байесовской оптимизации с помощью HMC на 2D функции трёхгорбого верблюда, использующей 100 шагов для разогрева, 200 - на сэмплирование, шаг=1e-2

Р Описание экспериментов

Пружина растяжения-сжатия

Цель: минимизировать объём винтовой пружины при постоянном натяжении.

Параметры:

- кол-во витков N ($x_1 \in [2, 15]$)
- диаметр намотки D ($x_2 \in [0.25, 1.3]$)
- диаметр проволоки d ($x_3 \in [0.05, 2]$)

Объём и механические ограничения:

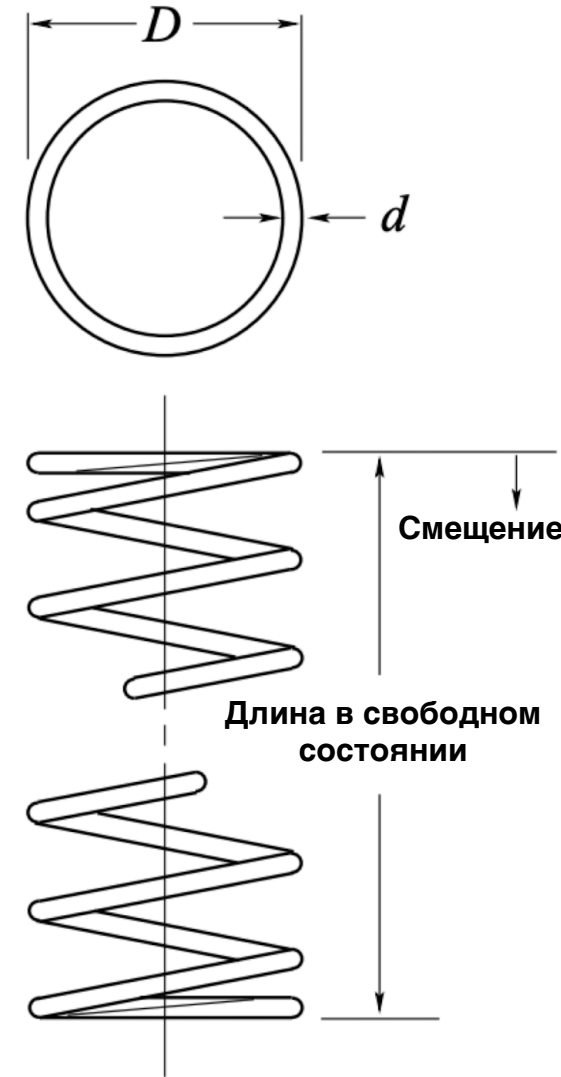
$$V(x) = (x_1 + 2)x_2x_3^2 \rightarrow \min_x$$

$$g_1(x) = 1 - \frac{x_2^3x_1}{71785x_3^4} \leq 0$$

$$g_2(x) = \frac{4x_2^2 - x_3x_2}{12566(x_2x_3^3 - x_3^4)} + \frac{1}{5108x_3^2} - 1 \leq 0$$

$$g_3(x) = 1 - \frac{140.45x_3}{x_2^2x_1} \leq 0$$

$$g_4(x) = \frac{x_2 + x_3}{1.5} - 1 \leq 0$$





Описание экспериментов

Сосуд под давлением

Цель: минимизировать вес цилиндрического сосуда под давлением с двумя сферическими крышками

Параметры:

- толщина сосуда $0.0625 \leq T_s$
- толщина крышки $T_h \leq 5$ (с шагом 0.0625)
- внутренний радиус сосуда $10 \leq R$
- длина цилиндрической составляющей $L \leq 200$

Вес и механические ограничения:

10

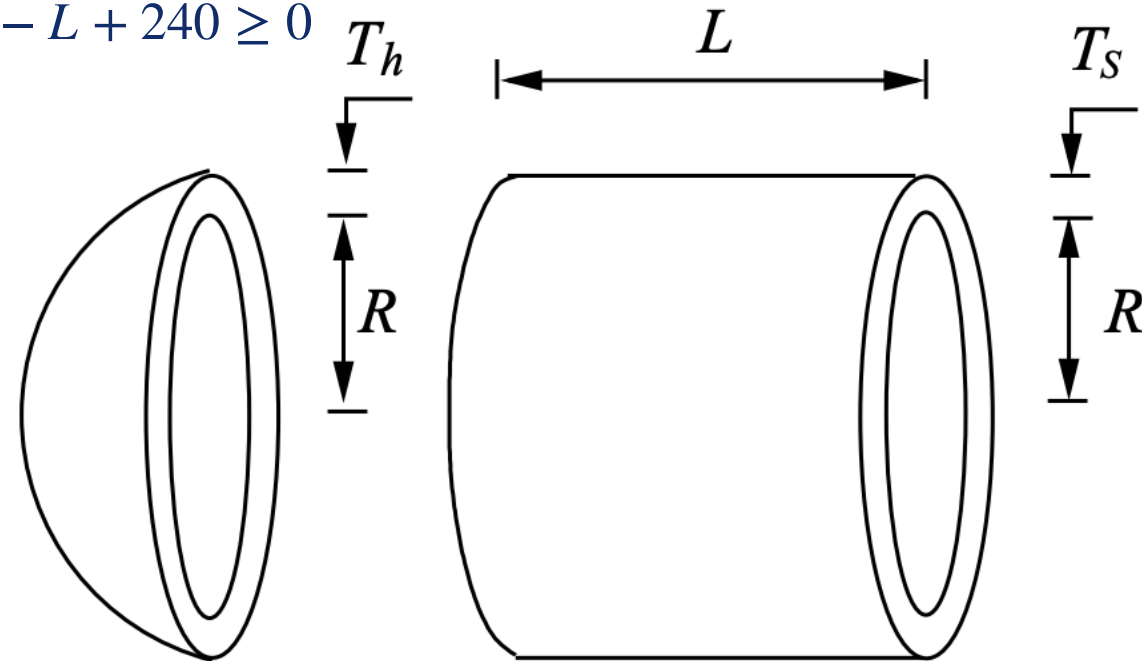
$$W(T_s, T_h, R, L) = 0.6224T_sT_hR + 1.7781T_hR^2 + 3.1661T_s^2L + 19.84T_s^2R \rightarrow \min_{T_s, T_h, R, L}$$

$$g_1(T_s, R) = T_s - 0.0193R \geq 0$$

$$g_2(T_h, R) = T_h - 0.00954R \geq 0$$

$$g_3(R, L) = \pi R^2L + \frac{4}{3}\pi R^3 - 1\,296\,000 \geq 0$$

$$g_4(L) = -L + 240 \geq 0$$





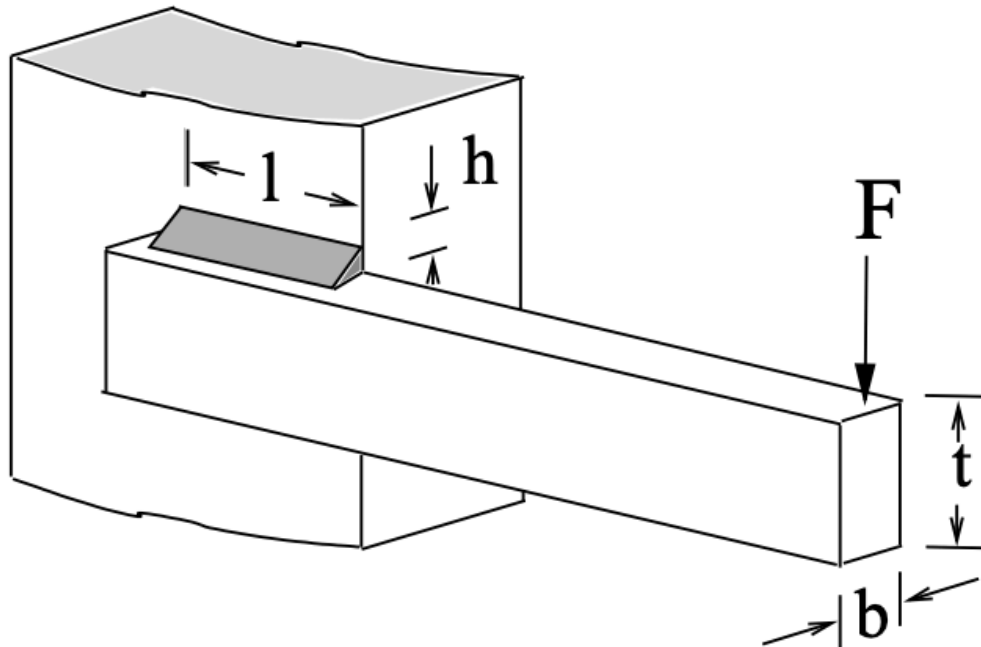
Описание экспериментов

Сварная балка

Цель: минимизировать стоимость изготовления балки.

Параметры:

- высота сварного шва $h \in [0.125, 10]$
- длина выступающей части балки $0.1 \leq l$
- толщина и ширина балки $t, b \leq 10$



Цена и механические ограничения:

11

$$C(h, l, t, b) = 1.1047h^2l + 0.04811tb(14.0 + l) \rightarrow \min_{h, l, t, b}$$

$$g_1(\tau) = 13\,600 - \sqrt{(\tau')^2 + (\tau'')^2 + l\tau'\tau''/\alpha} \geq 0$$

$$g_2(\sigma) = 30\,000 - 504\,000/(t^2b) \geq 0$$

$$g_3(b, h) = b - h \geq 0$$

$$g_4(P_c) = P_c - 6\,000 \geq 0$$

$$g_5(\delta) = 0.25 - 2.1952/(t^3b) \geq 0$$

$$\text{где } \alpha = \sqrt{0.25(l^2 + (h + t)^2)},$$

$$\tau' = \frac{6000}{\sqrt{2}hl}$$

$$\tau'' = \frac{6000(14 + 0.5l)\alpha}{2(0.707hl(l^2/12 + 0.25(h + t)^2))},$$

$$P_c = 64746.022(1 - 0.0282346t)tb^3$$

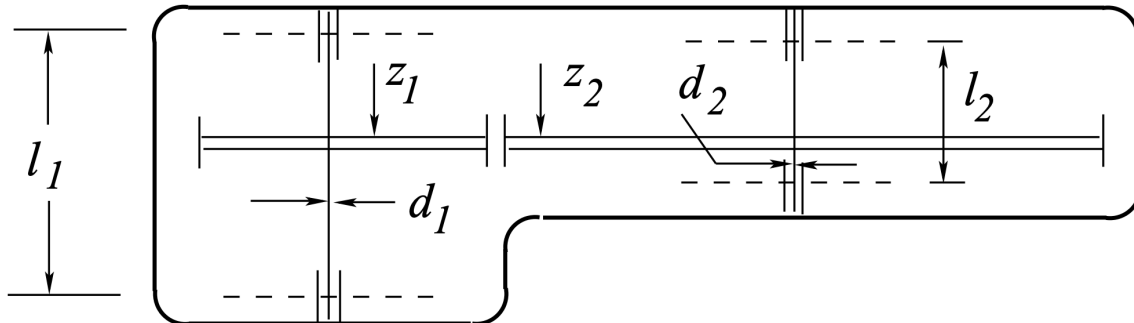
Описание экспериментов

Редуктор

Цель: минимизировать вес редуктора.

Параметры:

- ширина лицевой стороны b ($x_1 \in [2.6, 3.6]$)
- модуль зубьев t ($x_2 \in [0.7, 0.8]$)
- кол-во зубцов на шестерне n ($x_3 \in [17, 28]$)
- длина вала 1 между подшипниками l_1 ($x_4 \in [7.3, 8.3]$)
- длина вала 2 между подшипниками l_2 ($x_5 \in [7.8, 8.3]$)
- диаметр вала 1 d_1 ($x_6 \in [2.9, 3.9]$)
- диаметр вала 2 d_2 (x_7)



Вес и механические ограничения:

$$W(x) = 0.7854x_1x_2^2(3.3333x_3^2 + 14.9334x_3 - 43.0934) - 1.508x_1(x_6^2 + x_7^2) + 7.4777(x_6^3 + x_7^3) + 0.7854(x_4x_6^2 + x_5x_7^2) \rightarrow \min_x$$

$$g_1(x) = 27x_1^{-1}x_2^{-2}x_3^{-1} \leq 1, \quad g_2(x) = 397.5x_1^{-1}x_2^{-2}x_3^{-2} \leq 1$$

$$g_3(x) = 1.93x_2^{-1}x_3^{-1}x_4^3x_6^{-4} \leq 1, \quad g_4(x) = 1.93x_2^{-1}x_3^{-1}x_5^3x_7^{-4} \leq 1$$

$$g_5(x) = \frac{1}{0.1x_6^3} \left[\left(\frac{745x_4}{x_2x_3} \right)^2 + \{16.9\}10^6 \right]^{0.5} \leq 1100$$

$$g_6(x) = \frac{1}{0.1x_7^3} \left[\left(\frac{745x_5}{x_2x_3} \right)^2 + \{157.5\}10^6 \right]^{0.5} \leq 850$$

$$g_7(x) = x_2x_3 \leq 40, \quad g_8(x) = x_1/x_2 \geq 5$$

$$g_9(x) = x_1/x_2 \leq 12, \quad g_{10}(x) = (1.5x_6 + 1.9)x_4^{-1} \leq 1$$

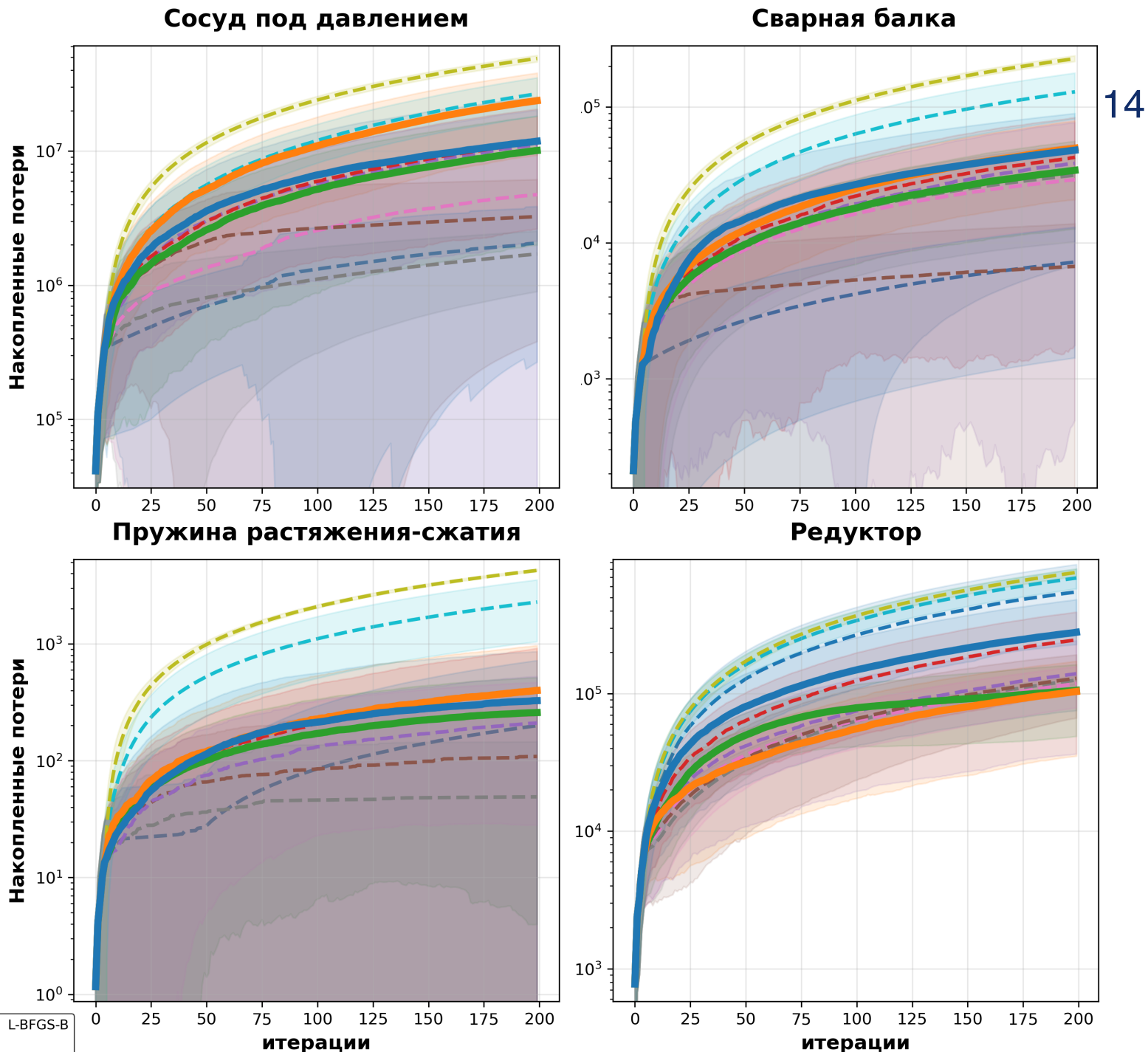
$$g_{11}(x) = (1.1x_7 + 1.9)x_5^{-1} \leq 1$$

потеря_{*i*} = оценённый минимум_{*i*} — реальный минимум

$$\text{накопленные потери}_T = \sum_{i=1}^T \text{потеря}_i$$

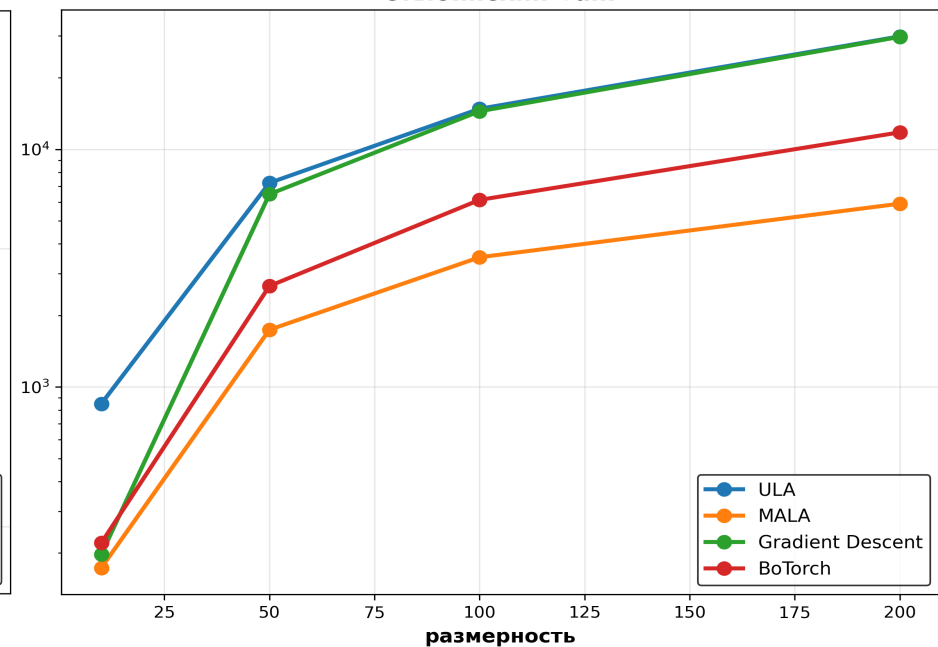
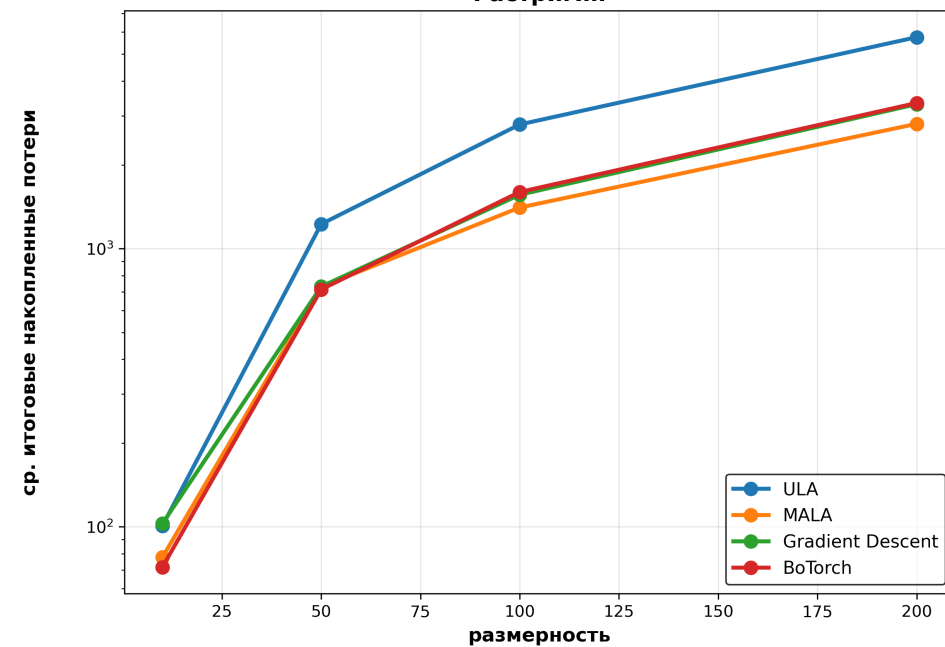
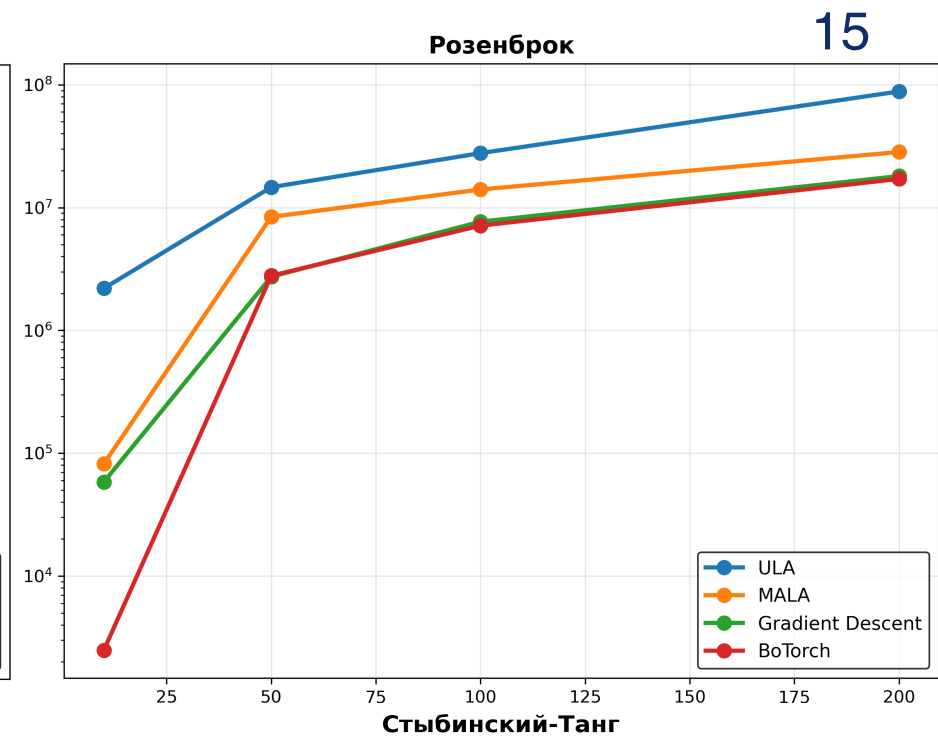
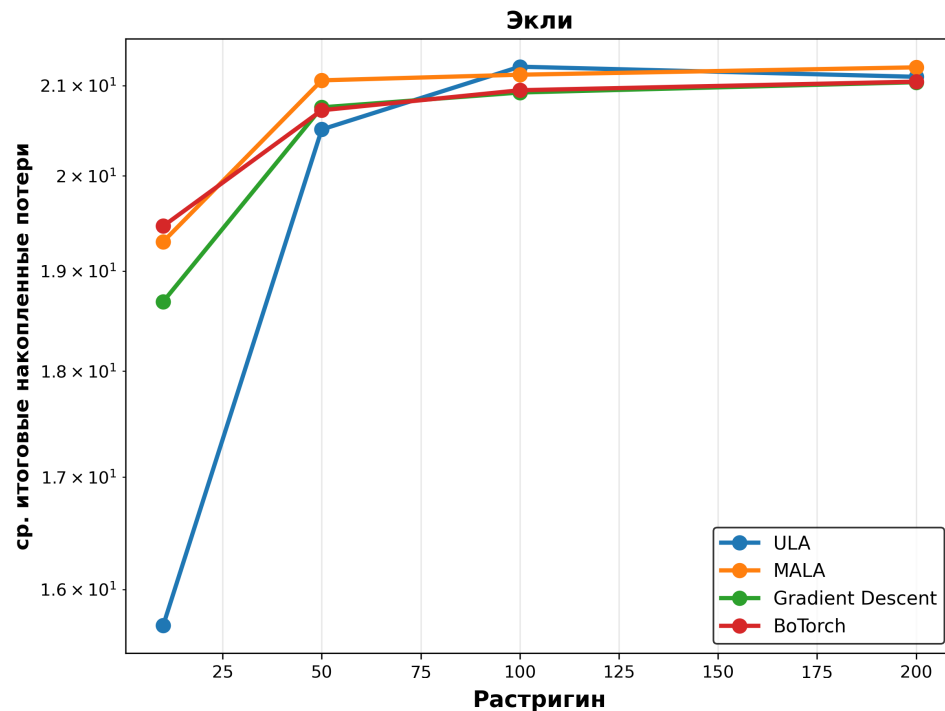
Результаты

Динамика Ланжевена может быть эффективнее других алгоритмов, но алгоритмы, основанные на L-BFGS-B и счёте градиента, могут быть лучше.



Результаты

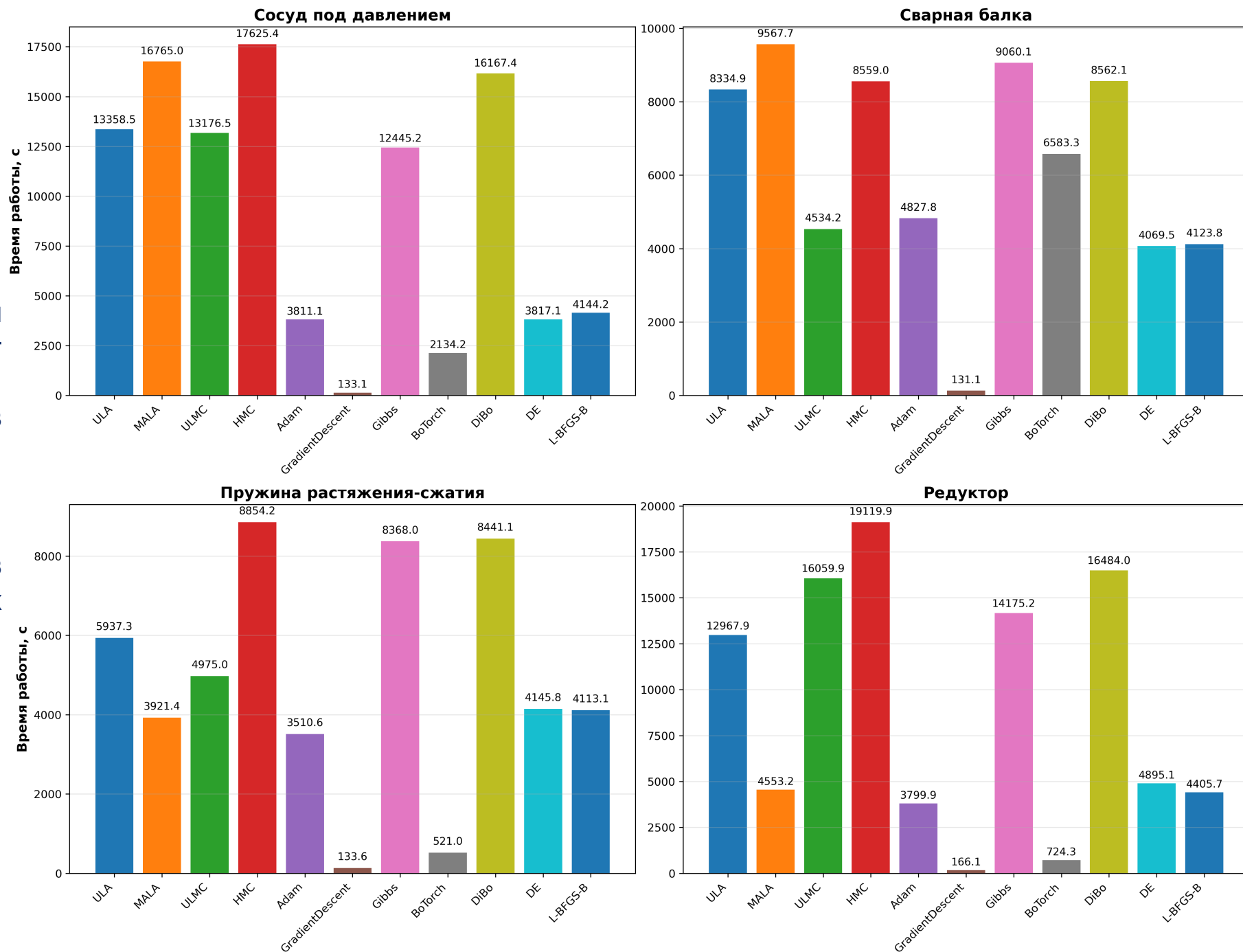
На синтетических тестах при увеличении размерности экспериментов повышается эффективность динамики Ланжевена относительно других алгоритмов.



Результаты

По времени вычислений динамика Ланжевена имеет **тот же порядок величины**, что и большинство методов оптимизации.

Динамика Ланжевена **в среднем быстрее** других MCMC алгоритмов.



- Мы рассмотрели задачу внутренней оптимизации Байесовской оптимизации на инженерных задачах.
- Было предложено заменить вычислительно дорогую оптимизацию невыпуклой функции алгоритмами Марковских цепей Монте-Карло.
- Динамика Ланжевена может быть эффективнее, особенно при больших размерностях экспериментов.



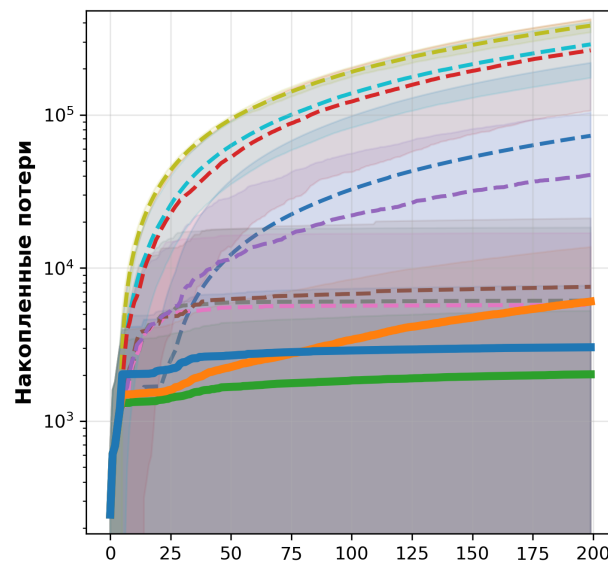


Аппендикс: результаты

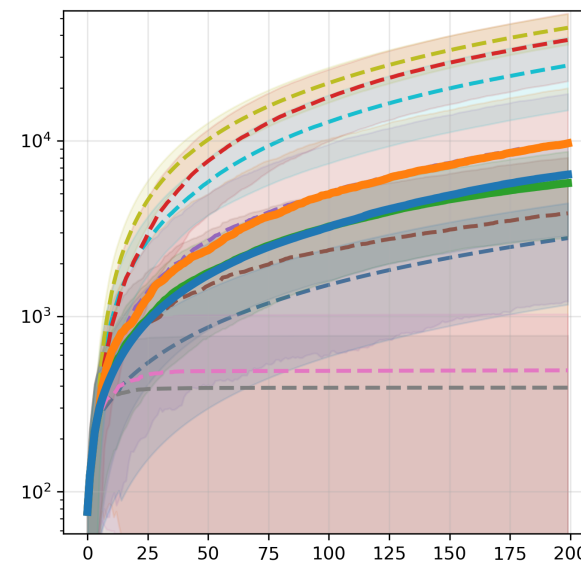
19



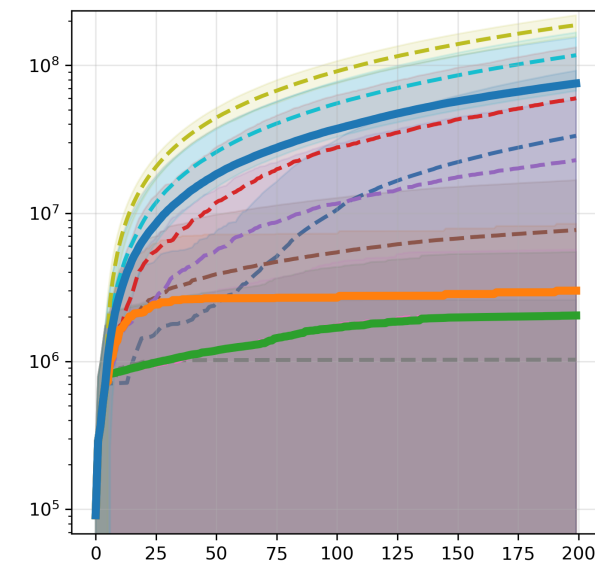
Трехгорбый верблюд



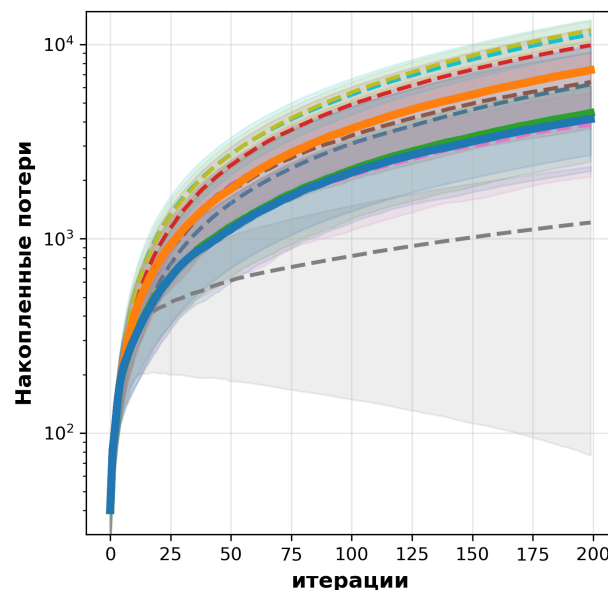
Бранин



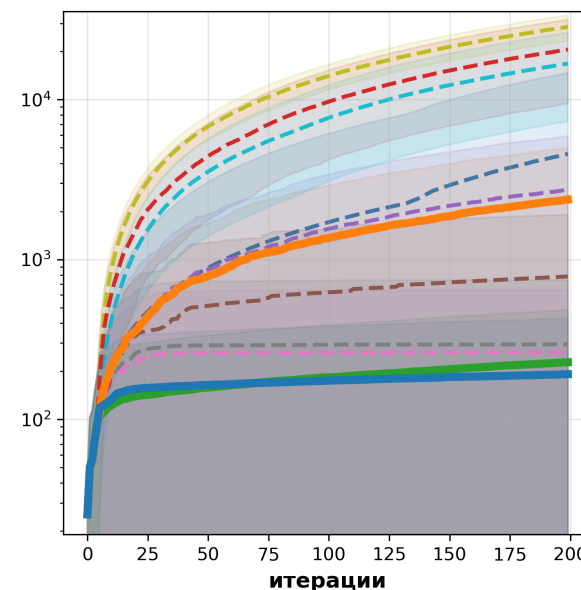
Розенброк



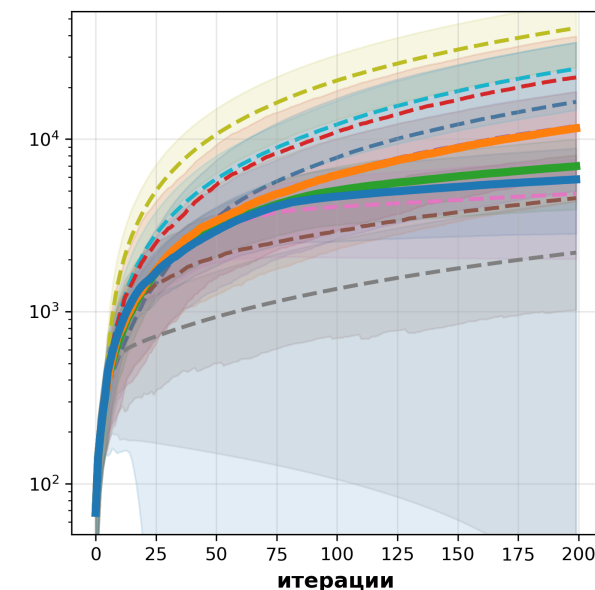
Растринг



Шестигорбый верблюд



Стыбинский-Танг

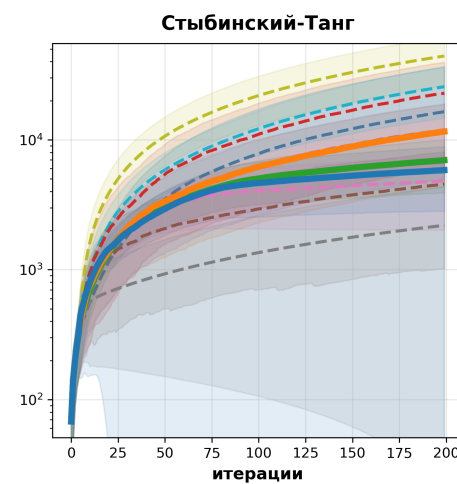
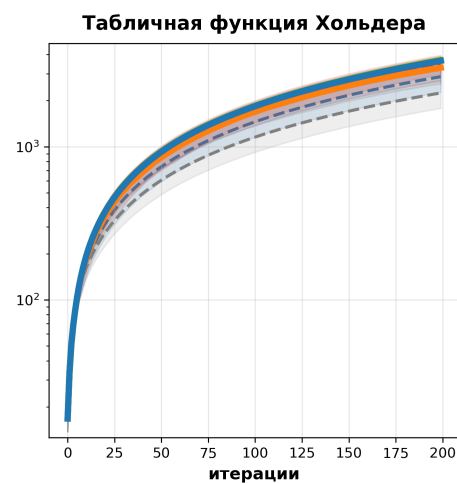
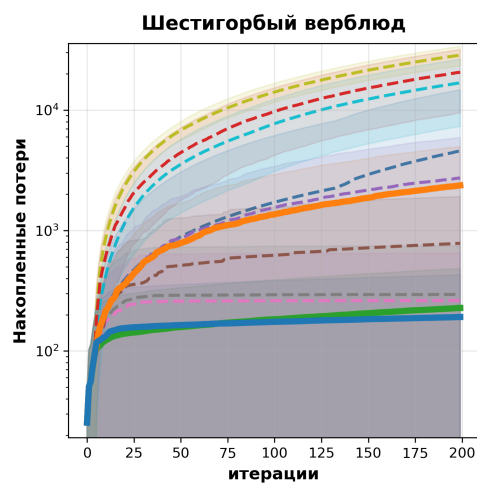
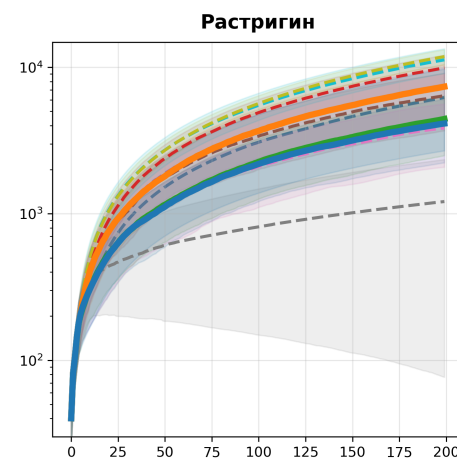
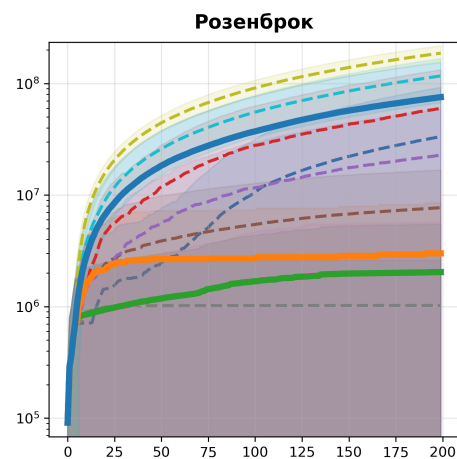
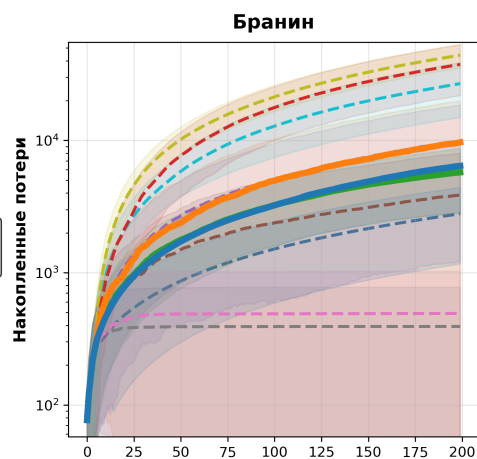
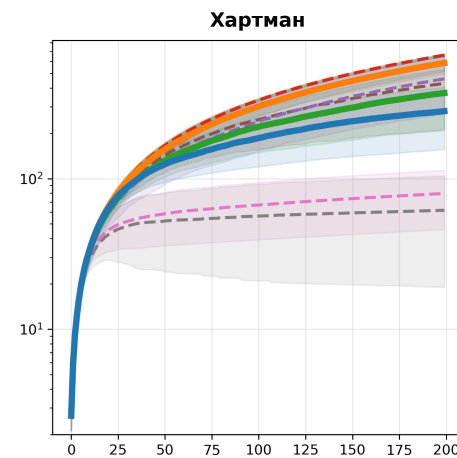
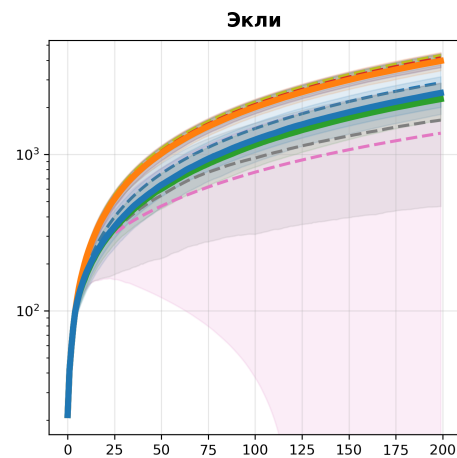
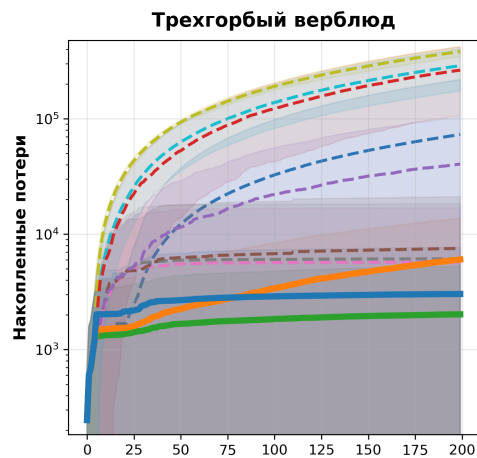
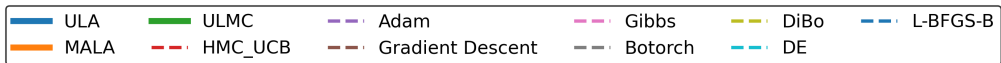


На синтетических тестах динамика Ланжевена не превосходит L-BFGS-B, особенно при низких размерностях.



Аппендикс: результаты

20



Результаты

