О «серых» числах (или Стохастичность и вычисления)

Олег Николаевич Граничин

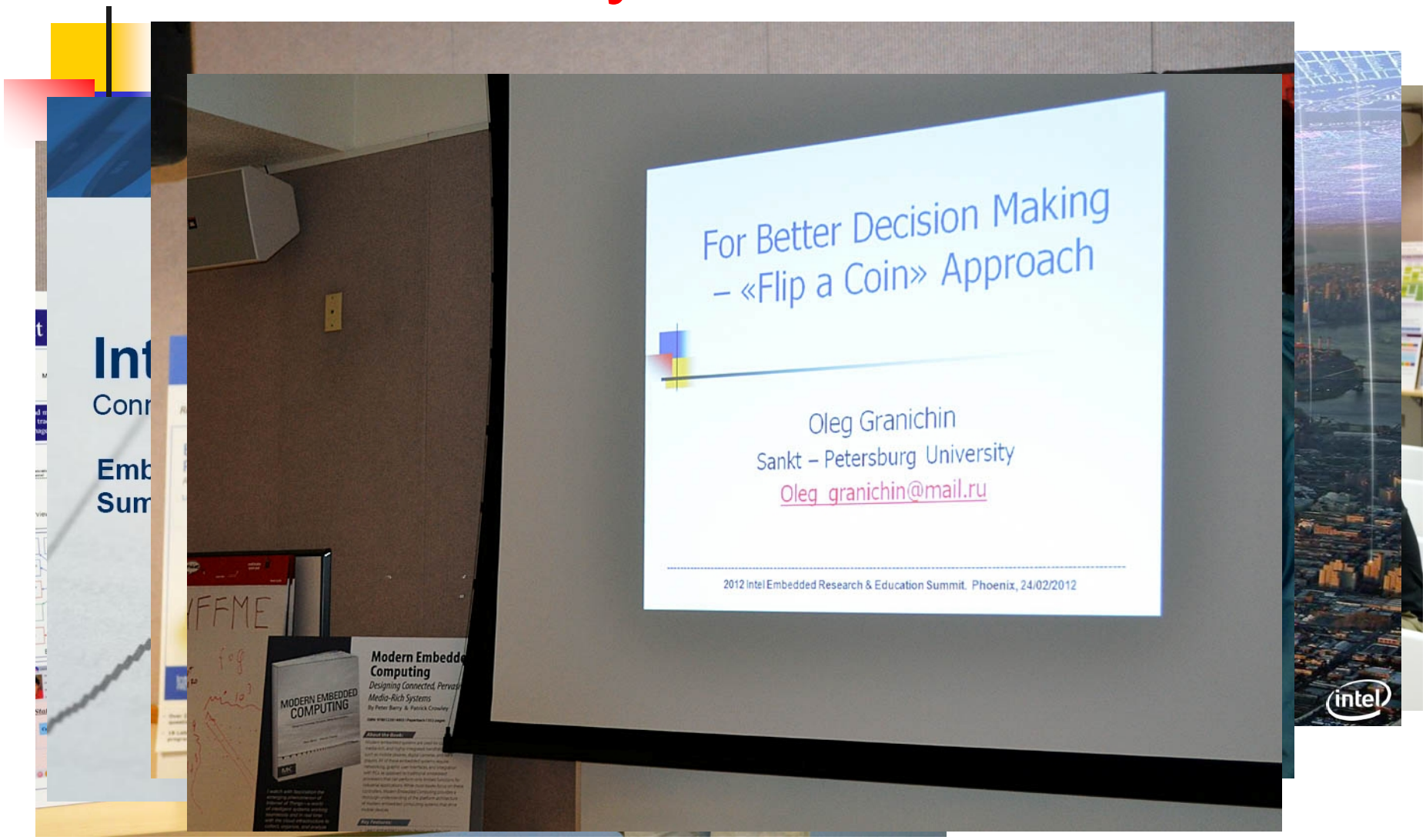
Санкт-Петербургский государственный университет

кафедра системного программирования
математико-Механического факультет

Oleg_granichin@mail.ru

СПбГУ, 26 апреля 2012 г.

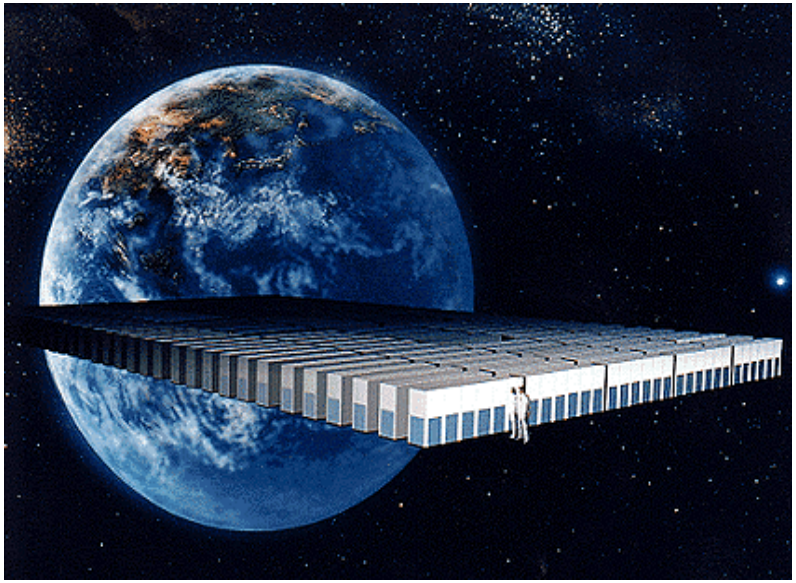
Почему это важно?



Что такое вычисления?

- Абак
- Компьютер

- Суперкомпьютер
- ...?



Информатика (Computer Science)

- Что может компьютер?
- Для чего нужен компьютер?

Уменьшение элементной базы

Основа современной ВТ – полупроводники (диоды и транзисторы)

- Закон Мура
- 2011 год – запущен завод, производящий микросхемы по технологии 22 нм
- Вся логика развития элементной базы ВТ ведет к уровню элементарных частиц, но в силу принципа неопределенности Гейзенберга снижается роль классической информации – битов {0} и {1}, неизменяющихся на протяжении такта ВУ

Понятие вычислимости

- Найти приближение для рассматриваемого объекта с любой наперед заданной точностью в классе вычислимых объектов
- Вычислимым объектом в смысле классической машины Тьюринга является множество конечных двоичных чисел
- Происходит редукция сложности изучаемого объекта. Например, иррациональные или трансцендентные числа заменяются существенно более простыми вычислимыми объектами – конечными двоичными числами

Сложность

- Чрезмерно сузив класс вычислимых объектов, получаем неразрешимые задачи
- Сложность вычислимого объекта должна быть эквивалентна сложности изучаемого реального объекта и адекватно отражать его свойства

Расширение класса вычислимых объектов

- Объект считаем вычислимым, если он арифметически связан с присутствующими в архитектуре вычислительного устройства вычислительными примитивами (функциональными элементами)

Модель вычислений

- Набор вычислительных примитивов (динамические модели H_i с параметрами Q)
- Память X – общее пространство состояний всех моделей
- Лента S – динамический граф с конечной битовой строкой s “включения” моделей в узлах
- Программа G – заданные на графе S правила (или цели) переключения “ленты” и параметров моделей при попадании на одном из “включенных” узлов пары (x, q) в множество переключения J
- Такт – промежуток времени между соседними моментами переключений
- Множество останова T

Процесс вычисления

$X \subset \mathbb{R}^\infty$ (или $\mathbb{C}^\infty$). Набор моделей $A = \{H_i\}_{i=1}^N$:

$$\dot{x}_i(t) = H_i(x_i, q_i, t), \; x_i \in X_i,$$

где $x_i \in X_i \subset X$, $q_i \in Q_i$ — параметры,
x_i, q_i — конечномерные.

Общая динамика системы между переключениями

$$\dot{x}(t) = \sum_{i=1}^{\infty} s_i(t) H_i(x_i, \theta_i, t), x \in X,$$

где $s_i(t) = 0$ или 1 — кусочно-постоянные.

Характерные черты

Вычислительное устройство — набор асинхронных моделей динамических систем, взаимодействующих между собой

Гибридность

Асинхронность

Кластерность

Стохастичность

……

Природа стохастичности

Неопределенности

- Моменты переключений
- Действия при переключениях
- (Задание Программы, Целей?..)

Стохастичность (στοχάζομαι)

- 1) целиться, метить
  ex. σ. τινος Xen., Isocr., Plat., Arst. — целиться во что (в кого)-л.
- 2) иметь в виду, стремиться
  ex. (τινος Plat., Arst. и πρός τι Plat.)
  σ. κριτῶν τῶν κρατίστων Xen. — стремиться к тому, чтобы судьями были самые влиятельные люди
- 3) применяться, приспособляться
  ex. (σ. τοῦ συμβουλευομένου Plat.)
  σ. τῆς τοῦ δήμου βουλήσεως Polyb. — применяться к воле народа
- 4) умозаключать, судить, догадываться, разгадывать
  ex. (τινος Isocr., Plat. и τι Xen.)
  σ. τὰ συμφέροντα Xen. — догадываться о том, что требуется;
  σ. ἔκ и διά τινος Polyb. — заключать на основании чего-л.;
  τῷ στοχάζεσθαι Plat. — путем догадок

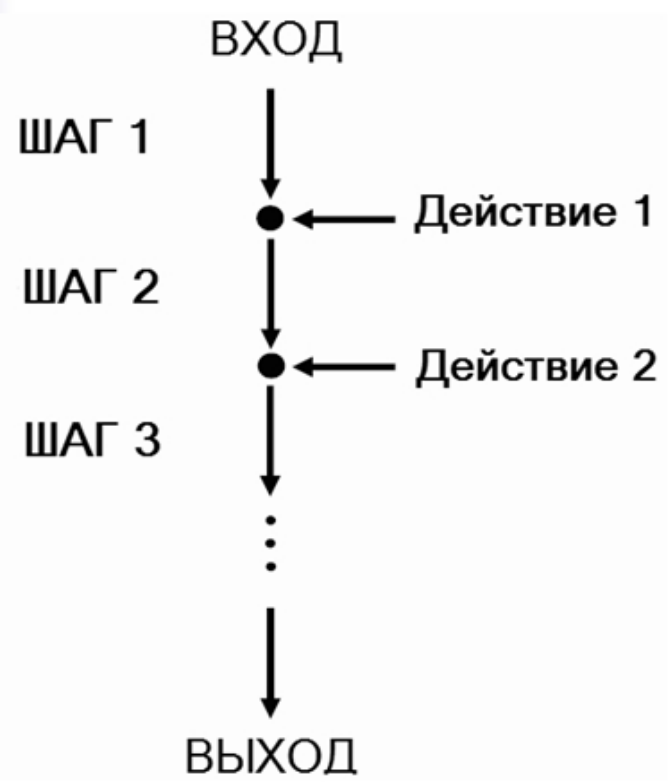
Стохастичность и вычисления

- «Кто нам мешает, тот нам поможет!»

Управление через обратные связи

- Обычно в условиях реального времени имеем:
  - ограничения в ресурсах
  - недостаточное количество данных с необходимым разнообразием

Детерминированный алгоритм



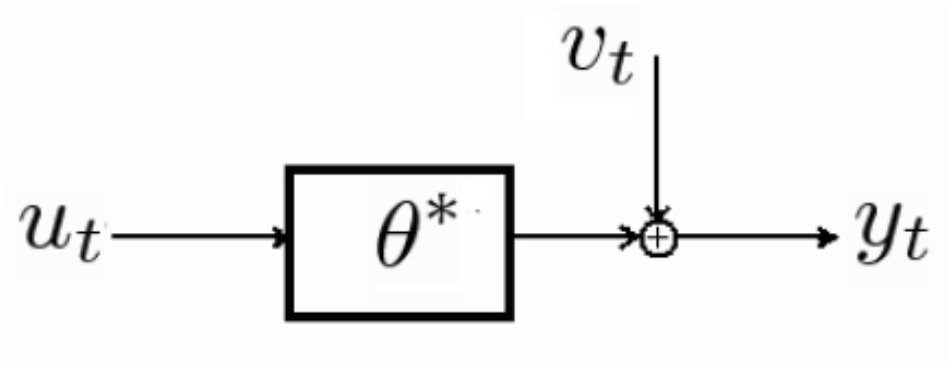
- Каждый шаг задается детерминированным правилом с использованием результатов предыдущих шагов, и полученная новая информация о системе (выход) возвращается для использования в последующих шагах алгоритма

Детерминированные подходы часто «проваливаются»

- В теории и практике возникает много трудностей, когда мы пытаемся исследовать “сложные” системы.

  Во многих практических приложениях традиционные детерминированные алгоритмы не дают результатов, когда система сложная. Часто это ведет к так называемым *NP*-сложным задачам

Пример: оценивание неизвестного параметра



$$y_t = \theta^\star \cdot u_t + v_t,$$

- можно выбирать входы (управления) u_t
- измерять выходы y_t , $t = 1, 2, ..., N$

При произвольных внешних помехах НЕТ РЕШЕНИЯ в классе детерминированных алгоритмов!

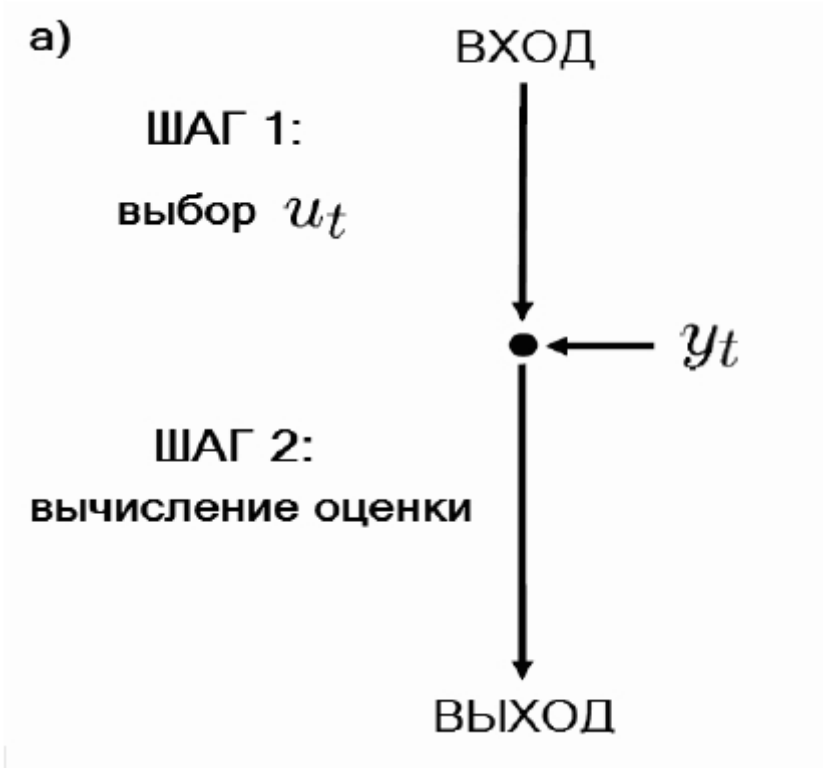
$\theta^\star = 3$

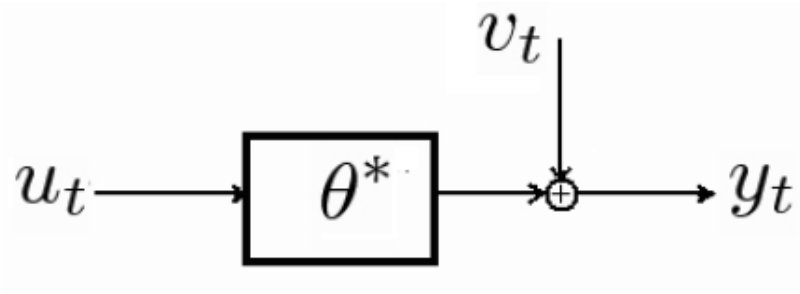
t	1	2	3	4	5	6	7
u_t	1	1	1	1	1	1	1
$v_t = rand() - 0.5$							
y_t	2.9	2.8	3.2	3.3	2.6	3.4	2.7
$\widehat{\theta}_t$	2.9	2.85	2.97	3.05	2.96	3.03	2.99
$v_t = rand() - 0.5 + m,\ m = 1$							
y_t	3.9	3.8	4.2	4.3	3.6	3.9	4.2
$\widehat{\theta}_t$	3.9	3.85	3.97	4.05	3.96	4.03	3.99

Рандомизация

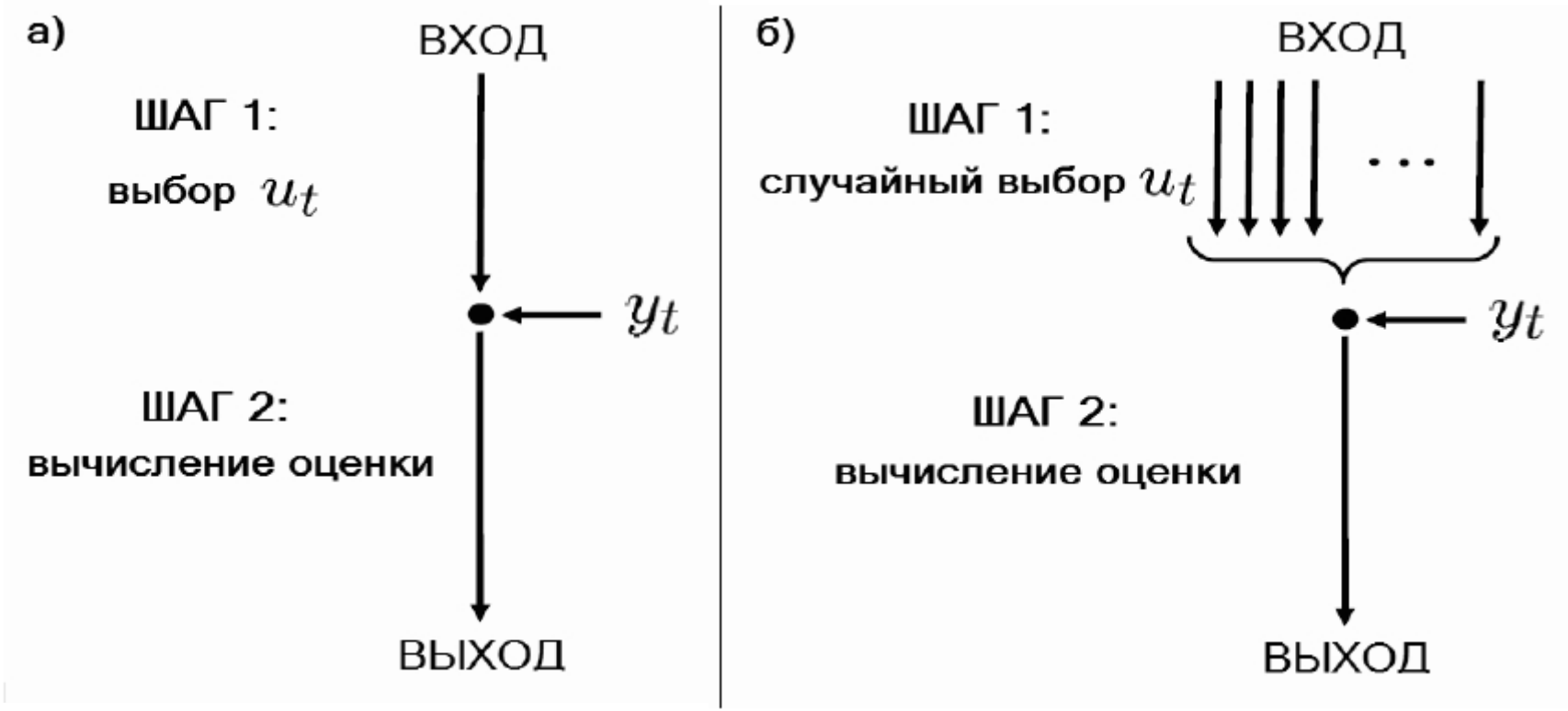
- Рандомизация может оказаться мощным средством для решения ряда проблем, которые кажутся нерешаемыми детерминированными методами
- В рандомизированном алгоритме выполнение одного или несколько шагов основано на случайном правиле (т. е. среди многих детерминированных правил одно выбирается случайно в соответствии с вероятностью P)

Два типа алгоритмов





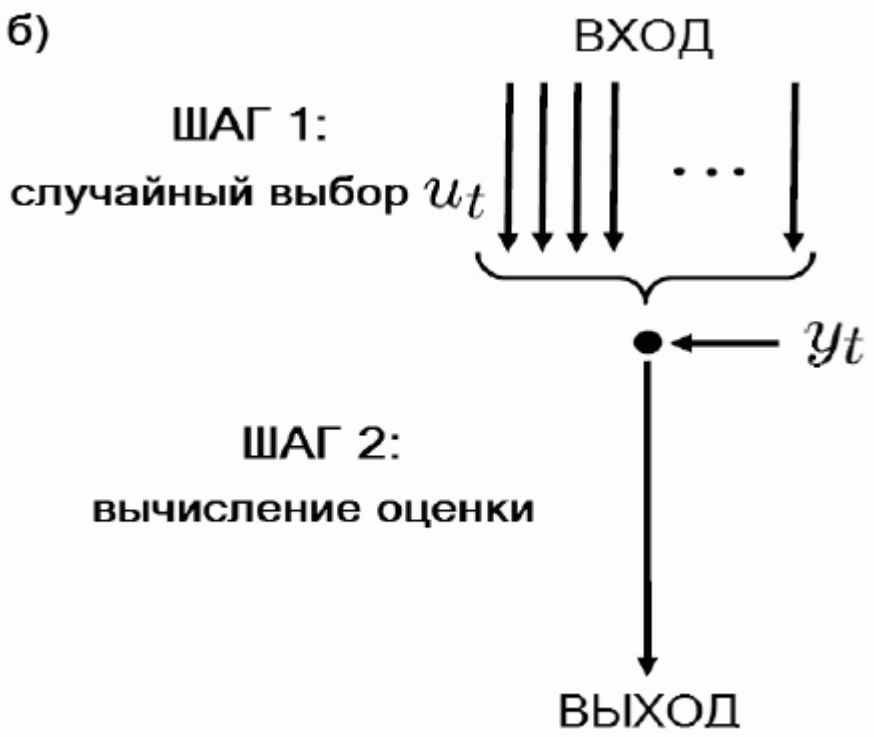
Два типа алгоритмов



Два типа алгоритмов

а)

б)



«Обогащенные» наблюдения

На первом шаге

$$u_t = \begin{cases} +1, \text{с вероятностью } \frac{1}{2}, \\ -1, \text{с вероятностью } \frac{1}{2}, \end{cases}$$

На втором шаге формируем величины

$$\bar{y}_t = u_t \cdot y_t.$$

Для “новой” последовательности наблюдений

$$\bar{y}_t = \theta^\star \cdot \bar{u}_t + \bar{v}_t, \quad \bar{u}_t = u_t^2 \text{ и } \bar{v}_t = u_t \cdot v_t.$$

Предварительный результат

$\theta^\star = 3$

t	1	2	3	4	5	6	7
u_t	-1	1	-1	1	1	1	-1
$v_t = rand() - 0.5 + m,\ m = 1$							
y_t	-2.1	3.8	-1.8	4.3	3.6	4.4	-2.3
$\bar{u}_t$	1	1	1	1	1	1	1
$\bar{y}_t$	2.1	3.8	1.8	4.3	3.6	4.4	2.3
$\widehat{\theta}_t$	2.1	2.95	2.57	3.00	3.12	3.33	3.19

Гарантированное множество

1. Пусть $M = 8$, выберем случайно семь $(= M - 1)$ разных групп по четыре индекса $T_1, \ldots, T_7$.
2. Вычислим семь частичных сумм $\bar{s}_i = \sum_{j \in T_i} \bar{y}_j$, $i = 1, \ldots, 7$.
3. Сформируем доверительный интервал

$$\widehat{\Theta} = [\min_{i \in 1:7} \bar{s}_i; \max_{i \in 1:7} \bar{s}_i],$$

содержащий $\theta^\star$ с вероятностью $p = 75\%$ $(= 1 - 2 \cdot 1/M)$.

Результат

i	T_i	$\bar{s}_i$
1	$\{2, 3, 4, 5\}$	3.375
2	$\{1, 3, 4, 6\}$	3.15
3	$\{2, 3, 5, 6\}$	3.4
4	$\{1, 2, 6, 7\}$	3.15
5	$\{1, 4, 5, 7\}$	3.075
6	$\{2, 3, 5, 7\}$	2.875
7	$\{1, 4, 6, 7\}$	3.275

- интервал $\widehat{\Theta} = [2.875; 3.4]$ содержит $\theta^\star$ с вероятностью $p = 75\%$,

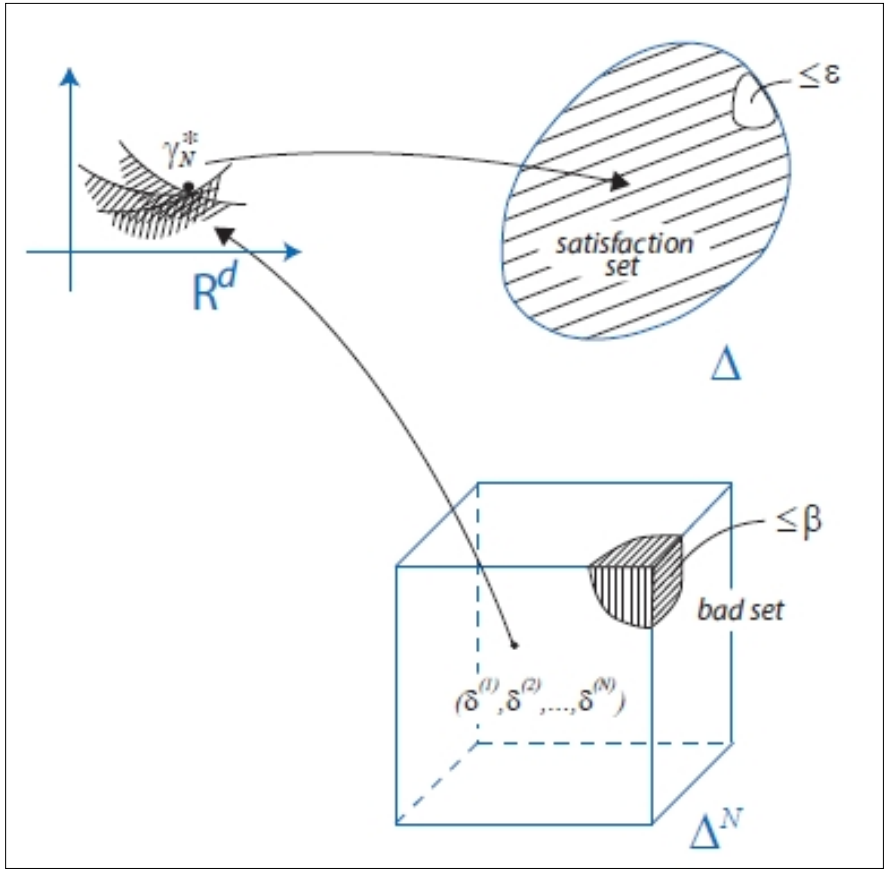
Вероятностно-успешный алгоритм

- Рандомизированный алгоритм называется вероятностно-успешным с вероятностью *p*, если вероятность его правильного результата не менее *p*
- Кроме обогащения данных наблюдений рандомизированный сценарный подход позволяет решить задачи «эффективно с высокой вероятностью» для почти всех ограничений, которых часто может быть очень много

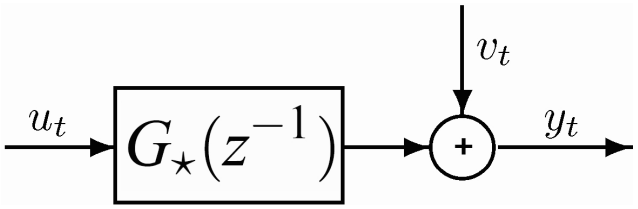
Сценарный подход

- Выбрать a-priori длину выборки *N*
- Решение задачи с *Prob=1-b*

  для почти всех условий



Идентификация параметров



$$u_{sn+i} = \begin{cases} \Delta_n + \bar{u}_{sn}, & \text{при } i = 0, \\ \bar{u}_{sn+i}, & \text{при } i = 1, 2, \ldots, s-1, \end{cases}$$

$n = 0, \ldots, N_\Delta - 1,$

$\bar{u}_t = \mathcal{U}_t(y_t, y_{t-1}, \ldots, \bar{u}_{t-1}, \ldots),\ t \geq 0$

1. $\hat{y}_{sn+k-1}(\theta) = \Delta_n \theta_k + \sum_{i=1}^{k-1} \theta_{k-i} \bar{u}_{sn+k-l-i},$
   $k = 1, \ldots, s.$
2. $\epsilon_t(\theta) = y_t - \hat{y}_t(\theta),\ t = 1, \ldots, N.$
3. $f_{sn+k-1}(\theta) = \Delta_n \epsilon_{sn+k-1}(\theta),\ k = 1, \ldots, s.$
4. Выберем $M > 2s$. $N_\Delta - 1$
   Вычислим $g_i^k(\theta) = \sum_{n=0} h_{i,ns+k} \cdot f_{ns+k-1}(\theta),$
   $i = 1, \ldots, M-1$
   $(h_{i,1}, \ldots, h_{i,N})$ из нулей и единиц .
5. Выберем q из $[1; M/2s]$.
   $\widehat{\Theta}^k$ не менее q из $g_i^k(\theta)$ строго $>$ нуля и
   не менее q строго меньше нуля.

$$\widehat{\Theta} = \bigcap_{k=1}^{s} \widehat{\Theta}^k.$$

$$Prob(\theta_\star \in \widehat{\Theta}) \geq 1 - 2sq/M$$

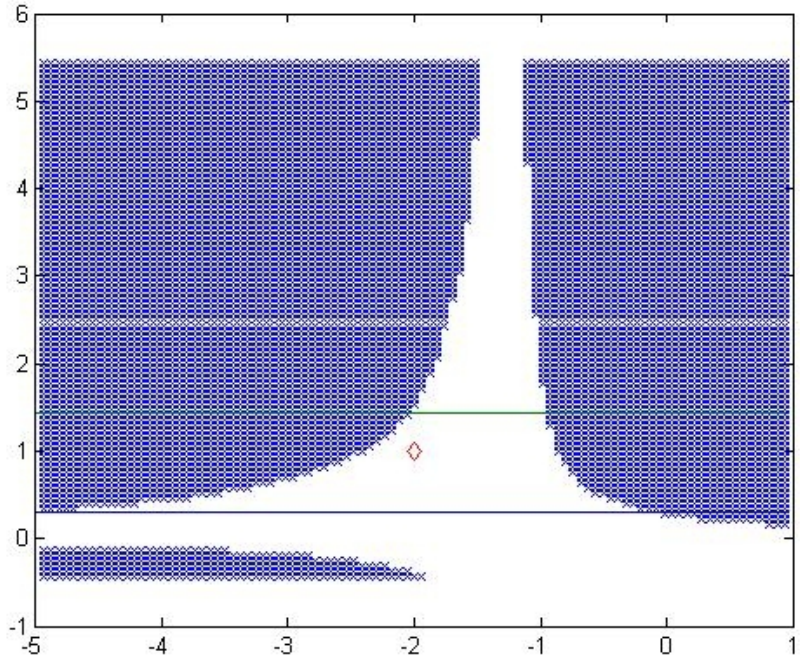
Пример

$$y_t + a_\star^1 y_{t-1} + y_{t-2} = b_\star^1 u_{t-1} + 1.6 u_{t-2} + v_t,\ t = 1, 2, \ldots, N$$

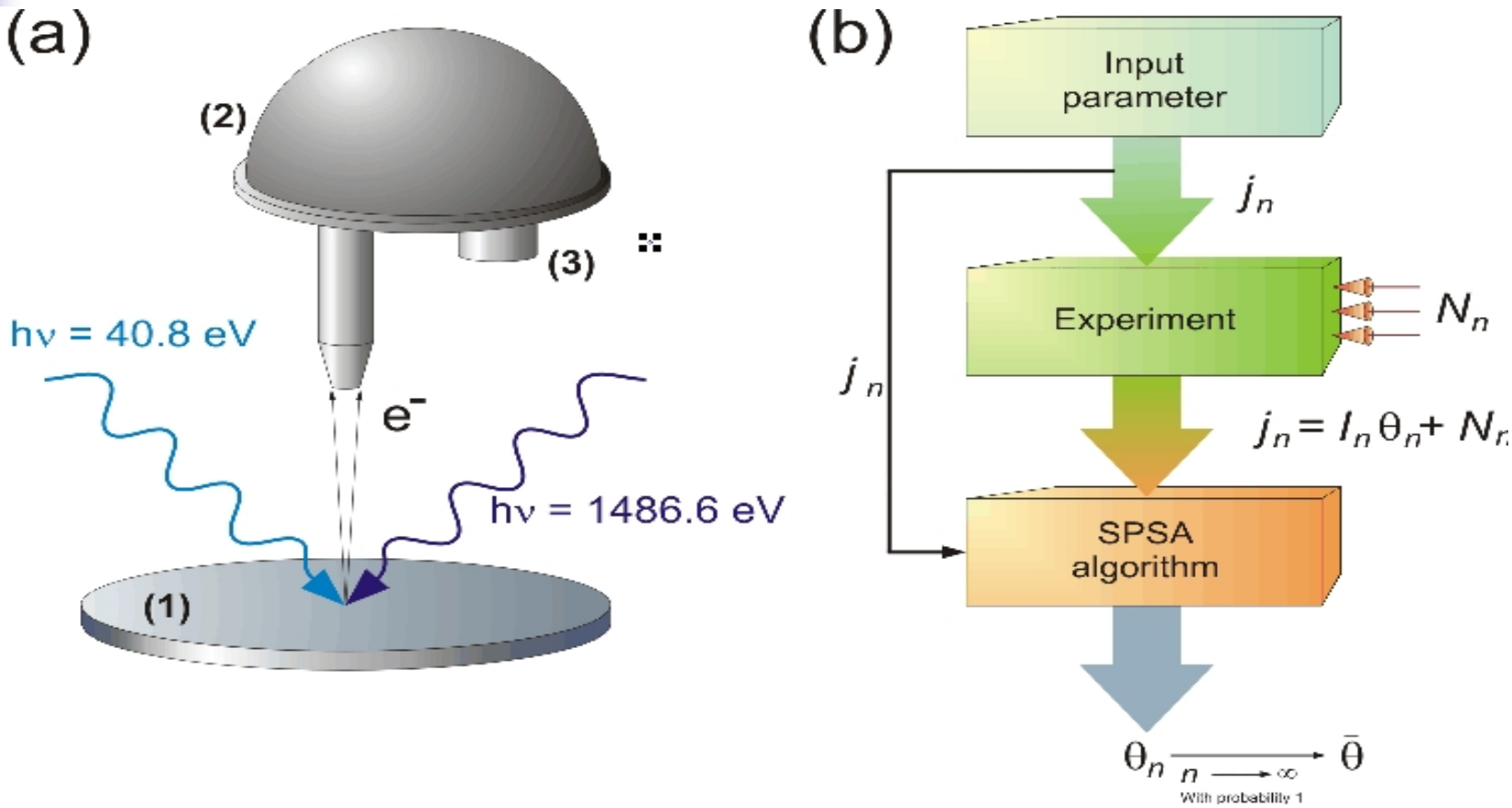
$$Prob(\theta_\star \in \widehat{\Theta}) \geq 1 - 2sq/M$$

$$95\% = 1 - 2 \cdot 2 \cdot 6/480$$

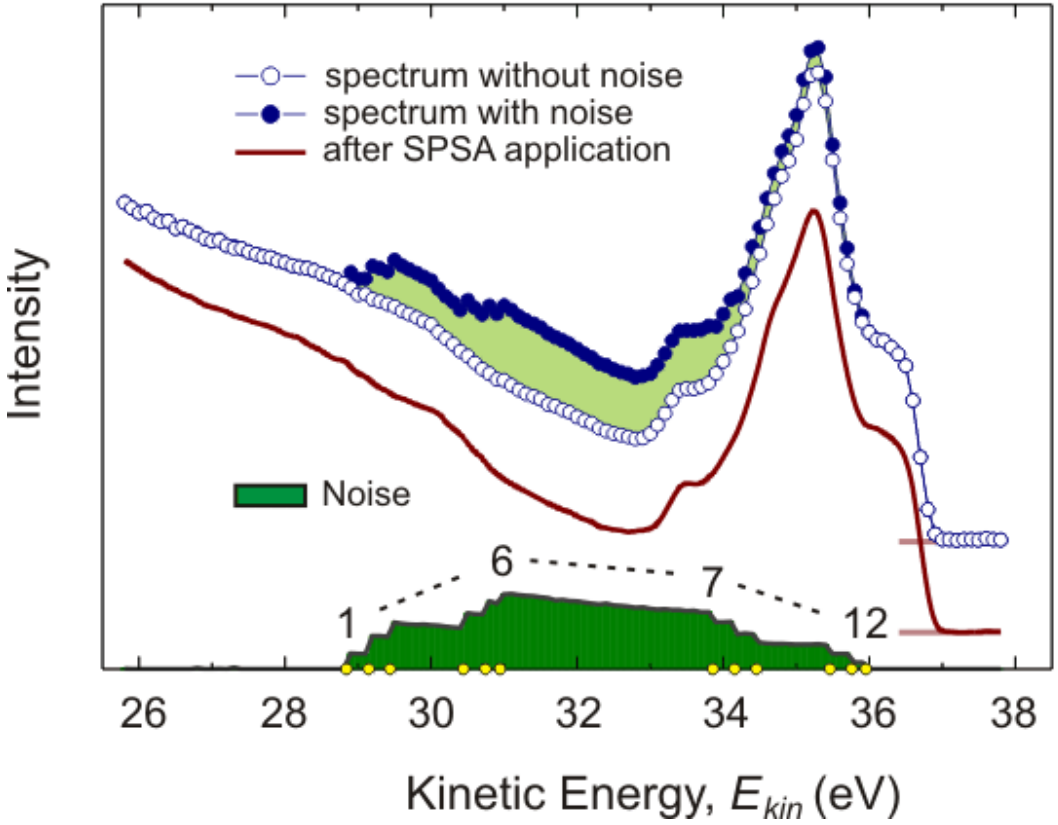
- Граничин О.Н. *Неасимптотическое доверительное множество для параметров линейного объекта управления при почти произвольных помехах* // АиТ, 2012.
- Campi M.C., Weyer E. *Non-asymptotic confidence sets for the parameters of li transfer functions* // IEEE TAC, 2010.



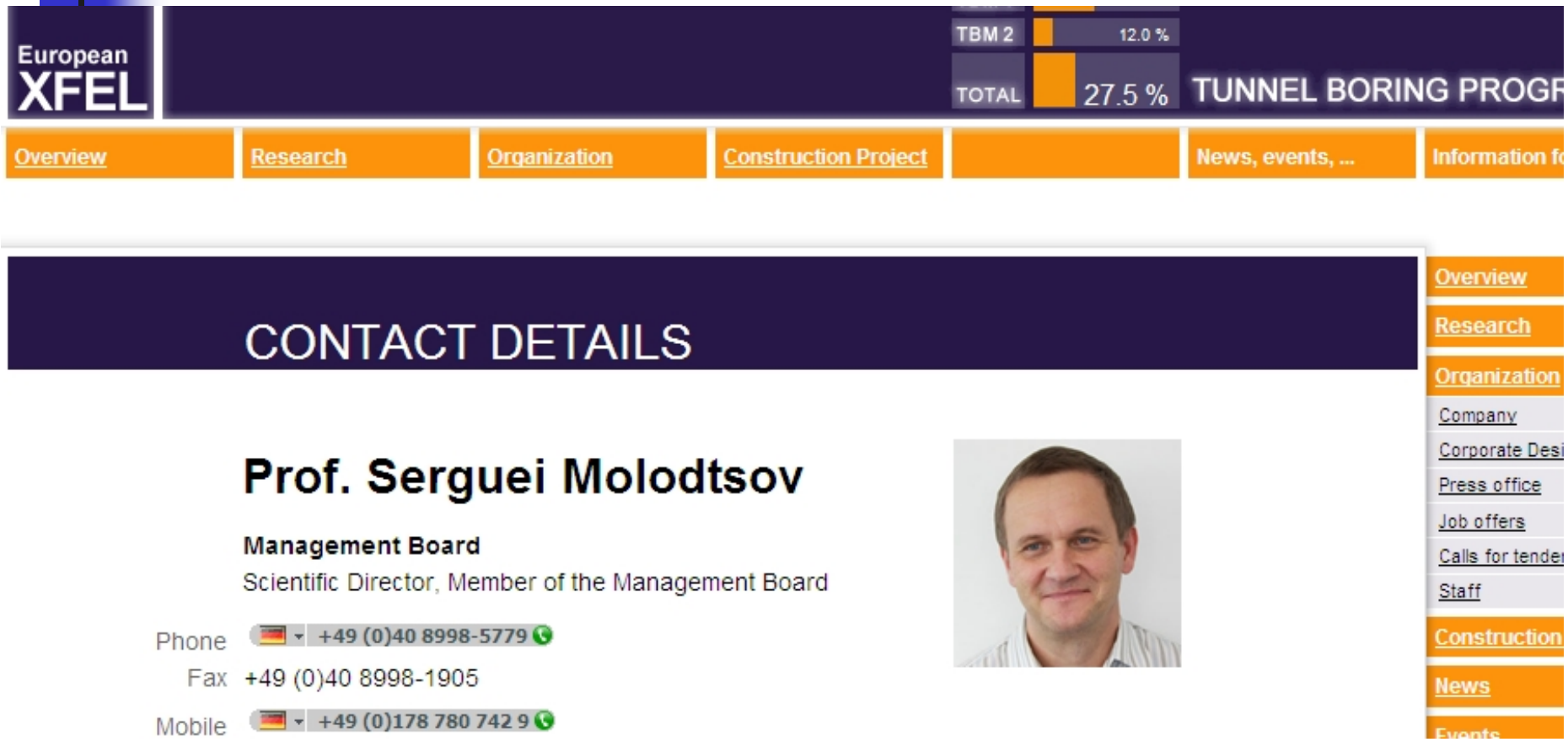
Рандомизация в фотоэмиссионном эксперименте



Granichin O. N., Molodtsov S. L. *и др.* ***Method of measurements with random perturbation: Application in photoemission experiments*** ***// Review of Scientific Instruments.*** *79,* ***036103. 2008.***



Granichin O. N., Molodtsov S. L. *и др.* ***Method of measurements with random perturbation: Application in photoemission experiments*** ***// Review of Scientific Instruments.*** *79,* ***036103. 2008.***



Granichin O. N., Molodtsov S. L. *и др.* ***Method of measurements with random photoemission experiments*** *Instruments. 79, 036103. 20*

European XFEL

Overview | Research | Organization | Construction

CONTACT DETAILS

Prof. Serguei Molodtsov

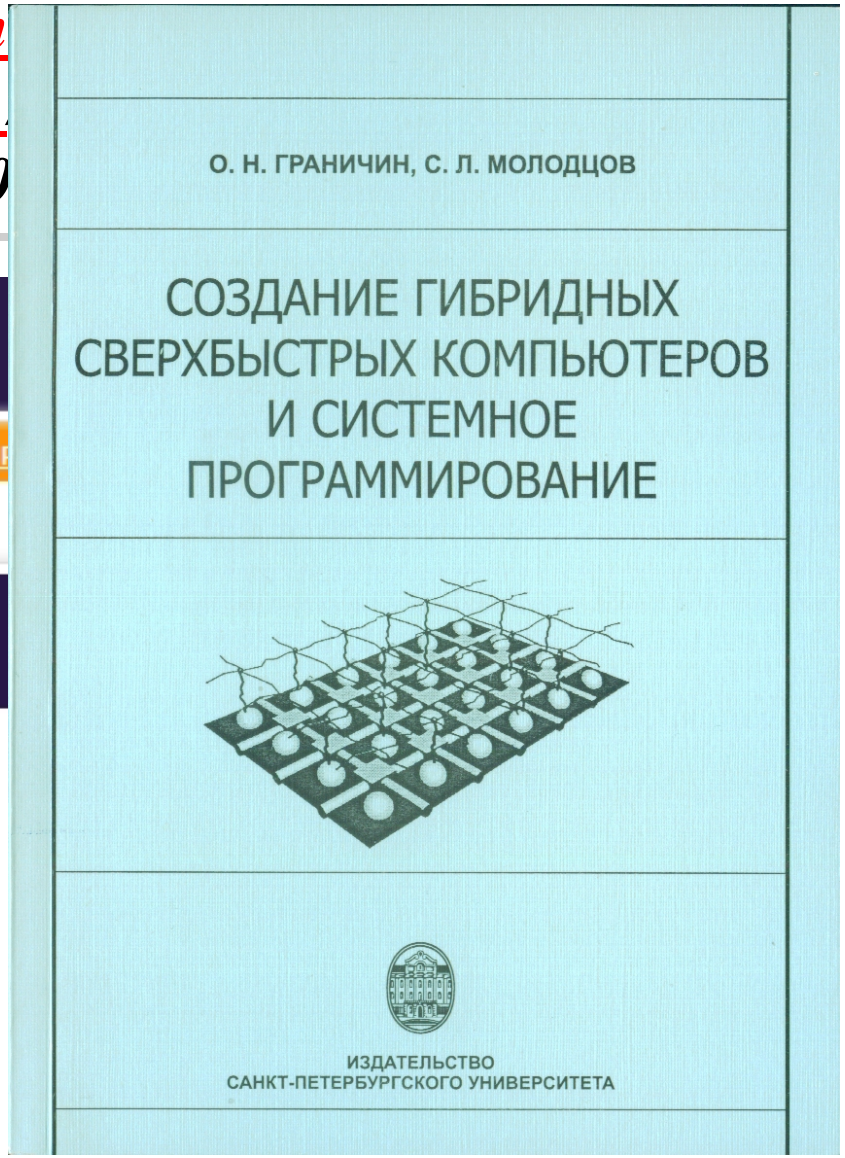
Management Board

Scientific Director, Member of the Management Board

Phone +49 (0)40 8998-5779

Fax +49 (0)40 8998-1905

Mobile +49 (0)178 780 742 9



Рандомизация...

1930…

- Fisher (remove bias)

1950…

- Metropolis, Ulam (method Monte-Carlo)
- Растригин, Ермаков и др. (случайный поиск)
- Kirkpatrick, Holland и др. (метод отжига, генетические алгоритмы)

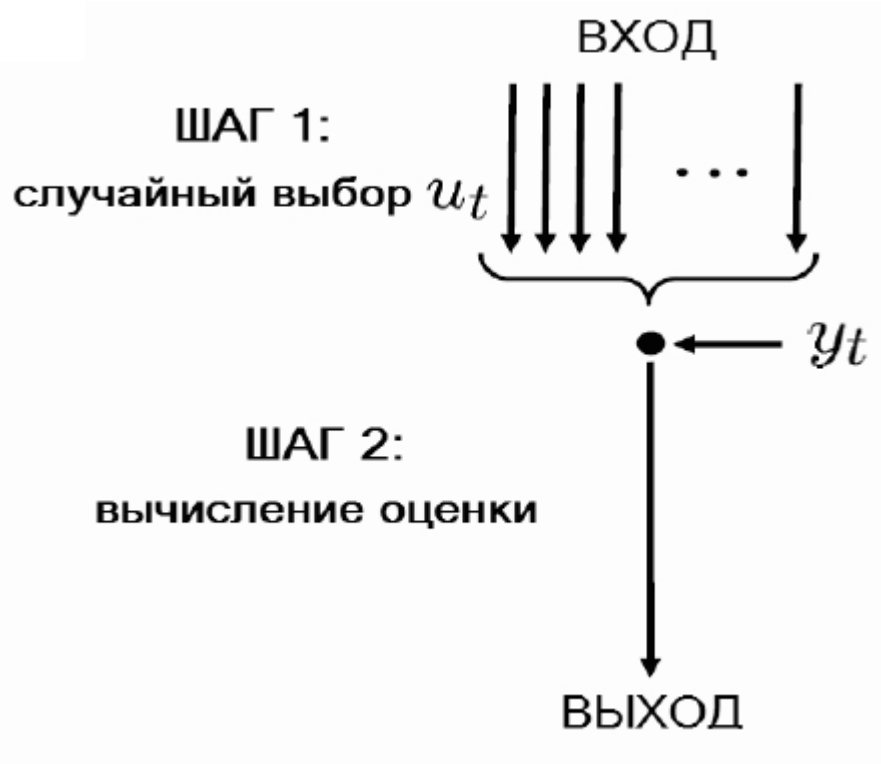
1980-90

- Поляк, Цыбаков, Luing, Guffi, Spall (fast algorithms)
- Granichin (arbitrary noise)
- Vadiyasagar (Randomized Learning Theory)

2000….

- Tempo, Campi, Calafiore, Sherbakov *etc.*
- …

Рандомизированный алгоритм



Квантовые вычисления

Многомерная оптимизация

$x_1, x_2, \ldots$ — точки наблюдений $\in \mathbb{R}^d$

$y_1, y_2, \ldots$ — наблюдения $\in \mathbb{R}^1$

$$y_n = F(x_n, w_n) + v_n$$

$w_1, w_2, \ldots$ — неконтролируемые случайные возмущения с неизвестным распределением $\mathsf{P}(\cdot)$

$v_1, v_2, \ldots$ — неизв., но ограниченные (неслучайн.)

Цель

$$f(x) = \int F(x, w)\mathsf{P}(dw) \to \min$$

Алгоритм SPSA

$\Delta_1, \Delta_2, \ldots$ — probing sampling with Bernoulli distribution (simultaneous perturbation) $\Delta_n = \begin{pmatrix} \pm 1 \\ \pm 1 \\ \vdots \\ \pm 1 \end{pmatrix} \in \mathbb{R}^d$

$$\widehat{\theta}_0 \in \mathbb{R}^d$$

$$x_n^{\pm} = \widehat{\theta}_{n-1} \pm \beta_n^{\pm} \Delta_n$$

$$\widehat{\theta}_n = \widehat{\theta}_{n-1} - \frac{\alpha_n}{\beta_n^+ + \beta_n^-} \Delta_n (y_n^+ - y_n^-)$$

Алгоритм с одним измерением

$$x_n = \widehat{\theta}_{n-1} - \beta_n \Delta_n$$

$$\widehat{\theta}_n = \widehat{\theta}_{n-1} - \frac{\alpha_n}{\beta_n} \Delta_n y_n$$

Алгоритм с одним измерением

$$x_n = \widehat{\theta}_{n-1} - \beta_n \Delta_n$$

$$\widehat{\theta}_n = \widehat{\theta}_{n-1} - \frac{\alpha_n}{\beta_n} \Delta_n y_n$$

- Сокращение числа наблюдений до 1 или 2 вместо $2d$

Преимущества SPSA

- Асимптотически оптимальная скорость сходимости
- Min число измерений на итерации
- Состоятельность при почти произвольных помехах
- Работоспособность в нестационарных задачах
- Легко может быть реализован на квантовом компьютере

Обоснование SPSA

$$E(\theta_n - \frac{\alpha_n}{\beta_n}\Delta_n f(\theta_n + \beta_n \Delta_n)) \approx$$

$$\theta_n - \frac{\alpha_n}{\beta_n} E(\Delta_n f(\theta_n) + \frac{\beta_n \Delta_n \Delta_n \nabla f(\theta_n)}{2}) =$$

$$\theta_n - \frac{\alpha_n}{2} \nabla f(\theta_n).$$

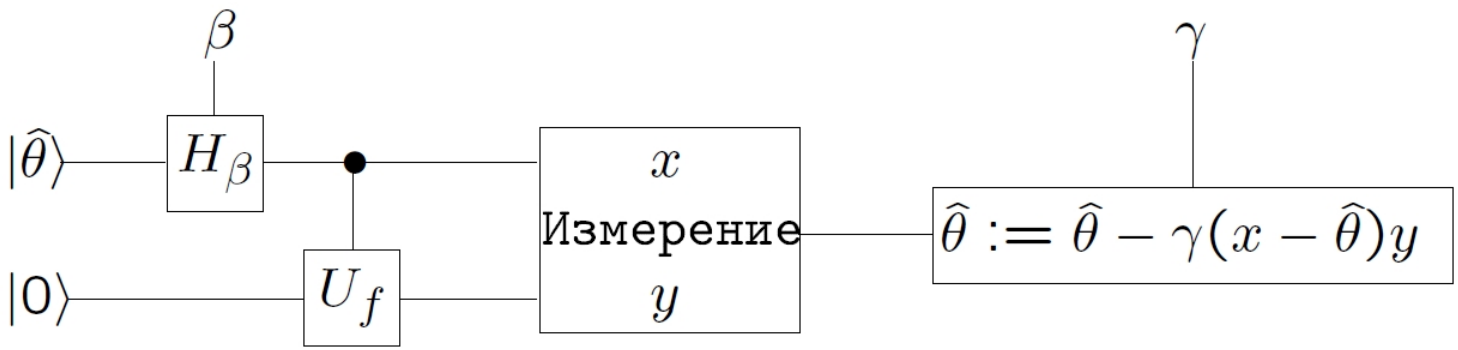
Преимущества SPSA

- Асимптотически оптимальная скорость сходимости
- Min число измерений на итерации
- Состоятельность при почти произвольных помехах
- Работоспособность в нестационарных задачах
- Легко может быть реализован на квантовом компьютере

Вычисление градиента «за такт»

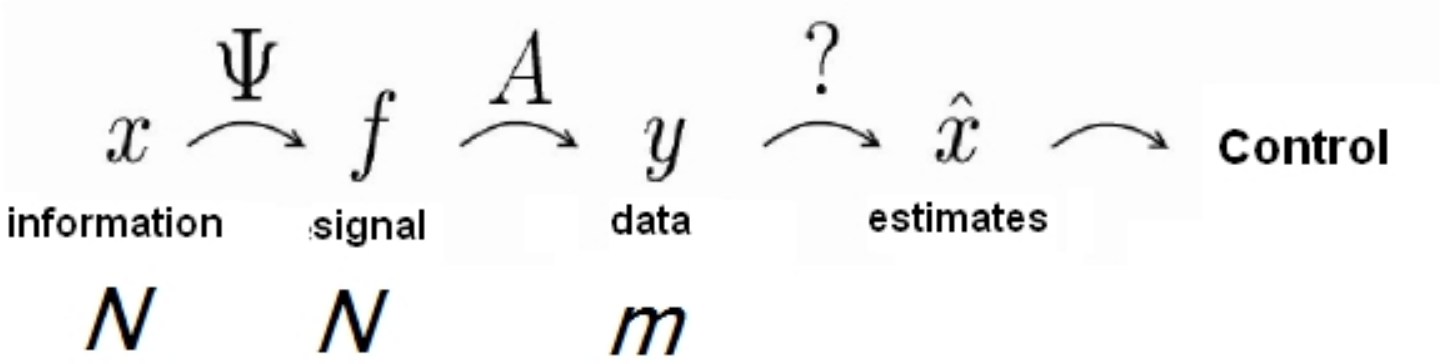
$$\hat{\theta}_k = \hat{\theta}_{k-1} - \gamma_k (x_k - \hat{\theta}_{k-1}) y_k, \quad x_k = \hat{\theta}_{k-1} + \beta_k \Delta_k, \quad y_k = f(x_k) + v_k, \quad \gamma_k = \frac{\alpha_k}{\beta_k^2}$$

$$U_f : |x\rangle |z\rangle \to |x\rangle |z \oplus f(x)\rangle, \quad x_k = H_\beta |\hat{\theta}_{k-1}\rangle = \frac{1}{2^{\frac{d}{2}}} \sum_{\Delta_i \in \{-1,+1\}^d} |\hat{\theta}_{k-1} + \beta_k \Delta_i\rangle$$



$$|\hat{\theta}_{k-1} + \beta_k \Delta_i\rangle |f(\hat{\theta}_{k-1} + \beta_k \Delta_i)\rangle, \quad \Delta_i \in \{-1,+1\}^d.$$

Compressive Sensing



$$y = A f = A \Psi x$$

s-sparse vector

$x = col(0,\ldots,0,1,0,\ldots,7,0,\ldots,0,3,0,\ldots)$ – *3-sparse*

Возможно ли оценить x при $m << N$?

Задачи

- Как выбрать универсальную матрицу *A?*
- Каков алгоритм восстановления сигнала?

Candes E., Romberg J., Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information // IEEE Trans. Inform. Theory. Vol. **52**. No. 2. Feb. 2006. P. 489–509.

Donoho D. Compressed sensing // IEEE Trans. Inform. Theory. Vol. **52**. No. 4. Apr. 2006. P. 1289–1306.

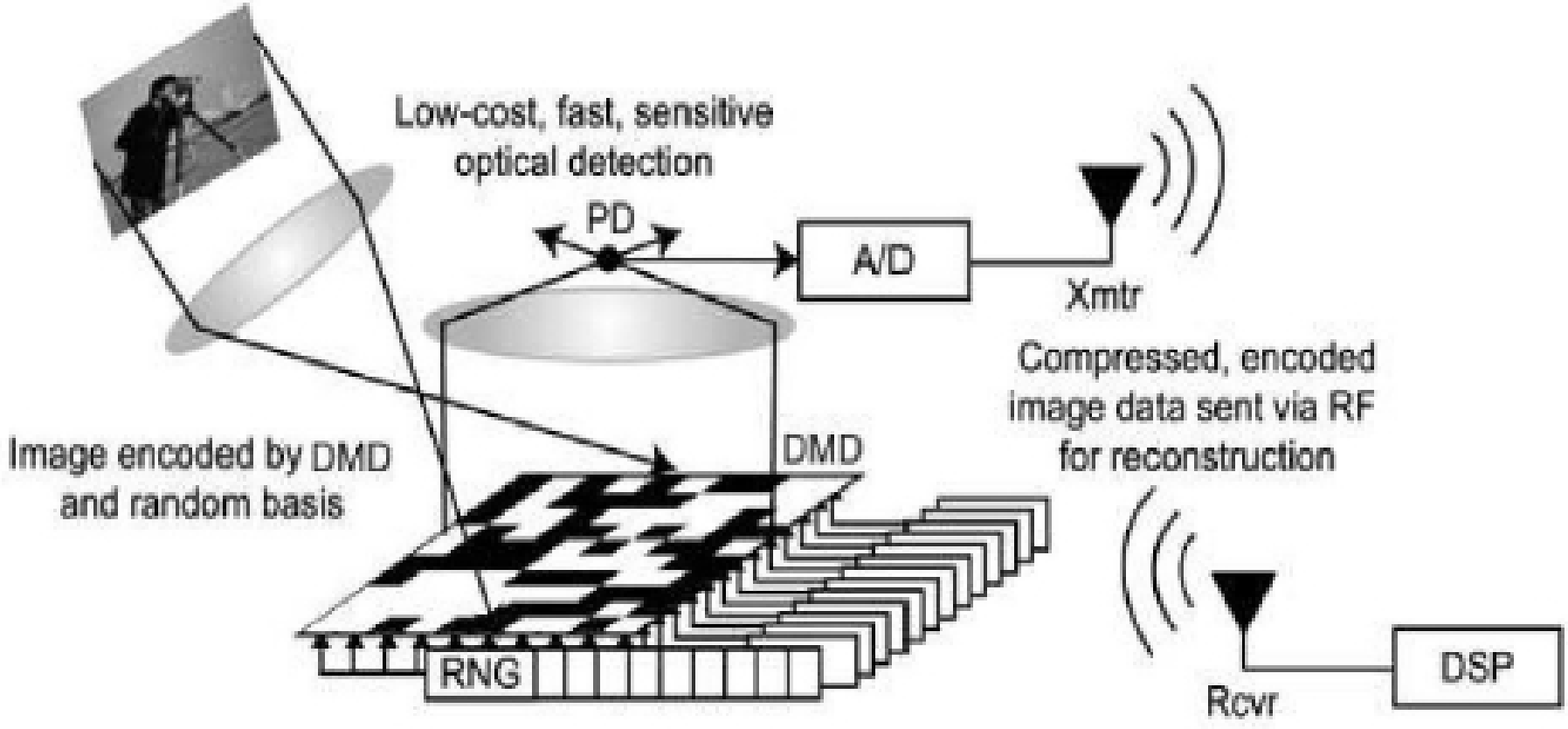
Универсальная случайная матрица *A*

- Gaussian *N(0,1/m)*
- Bernoulli $\mathrm{P}(a[i,j] = \pm 1/\sqrt{m}) = \frac{1}{2}$

  или $a[i,j] \in \{0, \pm\sqrt{3/m}\}$

$$m \geq c\, s \log(N/s)$$

Однопиксельная камера

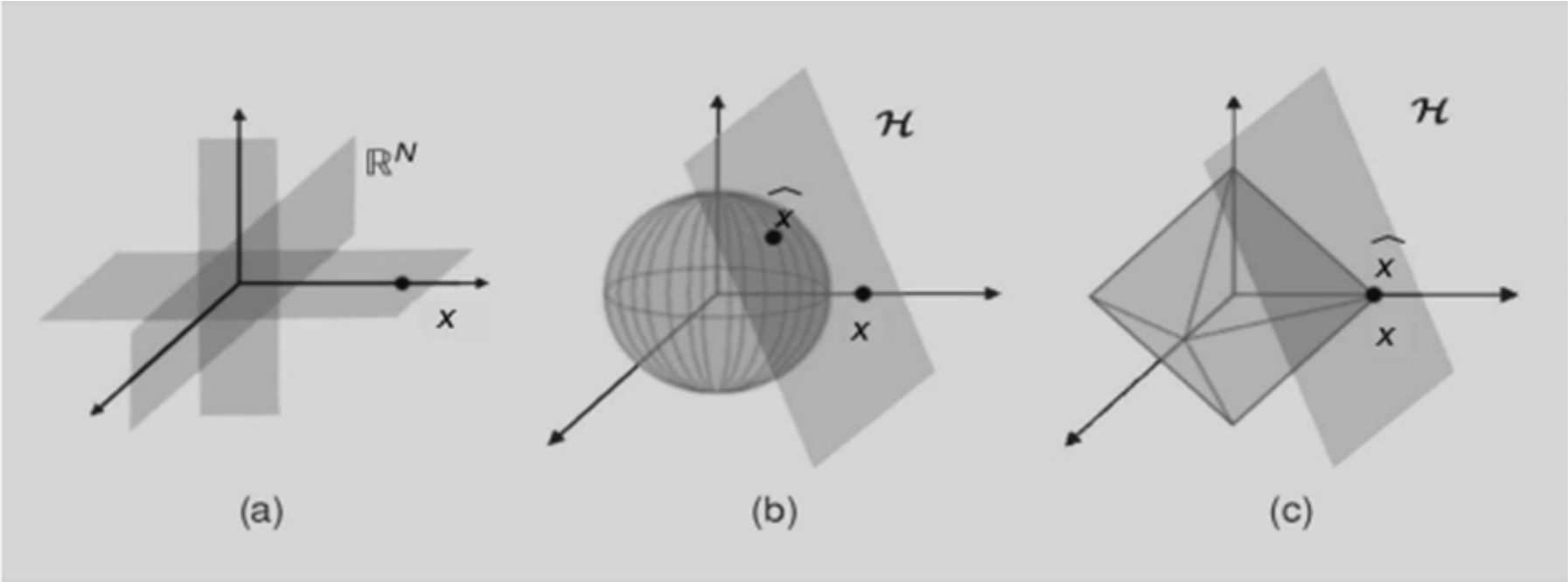


Алгоритм восстановления сигнала

Если выполнено условие RIP (2s) или RIP(3s), то достаточно решить задачу ℓ_1-оптимизации

$$||x||_1 = \sum_j |x[j]| \to \min \ : \quad y = \Phi x$$

ℓ_1- оптимизация



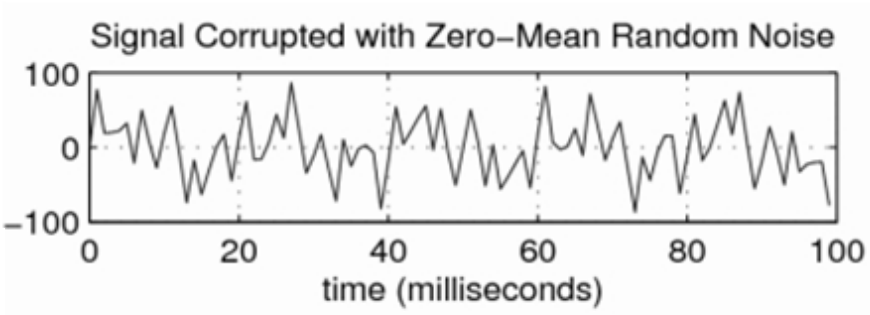
Пример

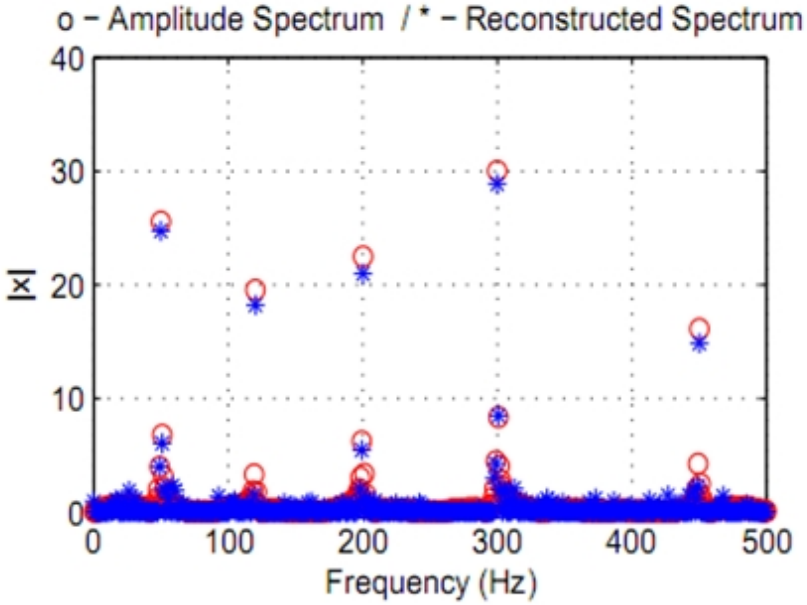
$$f(t) = \sum_{j=0}^{N} x[j] \exp^{i2\pi jt/N},$$

$$t = 0, \ldots, N-1,$$

$$s << N$$

$$s = 5$$





Определение частот и амплитуд 5-редкого сигнала.

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Благодарю за внимание!

Вопросы?

СПбГУ, 26 апреля 2012 г.

Литература

- Амелин К.С., Граничин О.Н. *Новые рандомизированные алгоритмы в управлении и обработке данных* // Стохастическая оптимизация в информатике. 2011. 7. С. 3-68.
- Граничин О.Н., Павленко Д.В. Рандомизация получения данных и ℓ1-оптимизация (опознание со сжатием) (обзор) // Автоматика и телемеханика, 2010, № 11. С. 3 -28.
- Вахитов А.Т., Граничин О.Н., Гуревич Л.С. (2009) Алгоритм стохастической аппроксимации с пробным возмущением на входе в нестационарной задаче оптимизации // Автоматика и телемеханика, № 11, с. 70-79.
- Вахитов А.Т., Граничин О.Н., Сысоев С.С., (2006) Точность оценивания рандомизированного алгоритма стохастической оптимизации. *Автоматика и телемеханика*, № 4 , с. 86-96.
- Граничин О.Н., (1989) Об одной стохастической рекуррентной процедуре при зависимых помехах в наблюдении, использующей на входе пробные возмущения. *Вестник Ленингр. ун-та*, сер. 1, вып. 1, с. 19--21.
- Граничин О.Н., Васильев В.И., (2010) Гибридная модель процесса вычислений: обобщение концепции машины Тьюринга. *Нейрокомпьютеры: разработка, применение*, № 6, с. 51 -58.
- Граничин О.Н., Молодцов С.Л., (2006) *Создание гибридных сверхбыстрых компьютеров и системное программирование*. СПб, 108 с.
- Граничин О.Н., Поляк Б.Т., (2003) *Рандомизированные алгоритмы оценивания и оптимизации при почти произвольных помехах*. М.: Наука, 291с.
- Поляк Б.Т., Цыбаков А.Б., (1990) Оптимальные порядки точности поисковых
- алгоритмов стохастической аппроксимации. *Проблемы передачи информации*, № 2, с. 45-53.
- Shor, P., (1997) Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput. 26*, pp. 1484-1509.
- Tien, D.K., (2003) Computing the non-computable, *Contemporary Physics, 44,* pp. 51-71.
- Granichin O.N., Vasil'ev V.I., (2010) Computational model based on evolutionary primitives. Turing machine generalization. *International Journal of Nanotechnology and Molecular Computation*, v. 2, no. 2, pp. 30-43.
- *Statistics*, Rustagi J.S. ed.. New York: Academic Press, pp. 233-257.
- Spall J.C., (1992) Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. IEEE Transactions on Automatic Control, vol. 37, pp. 332-341.

Поколения компьютеров

1

2

3

4

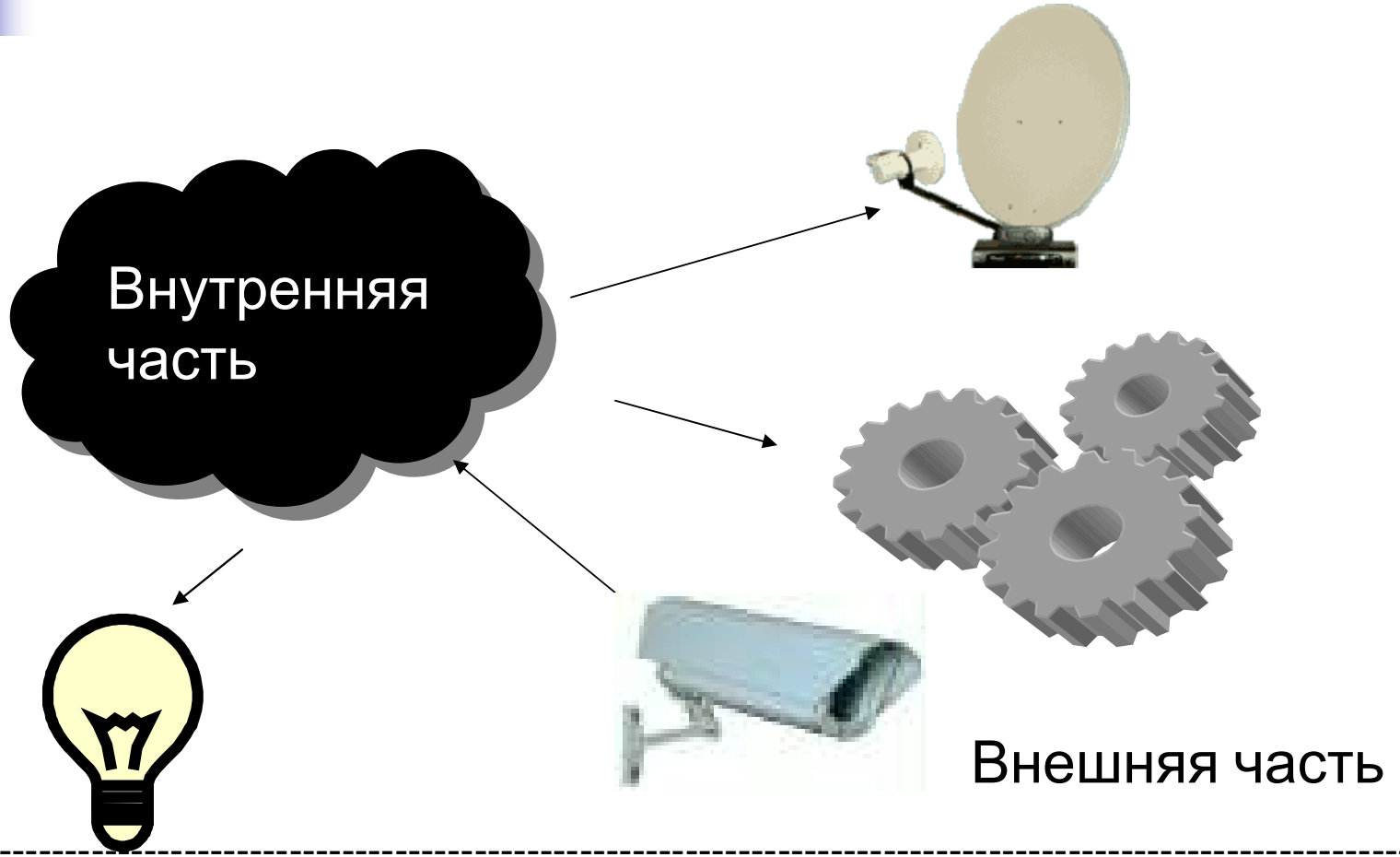
5

?

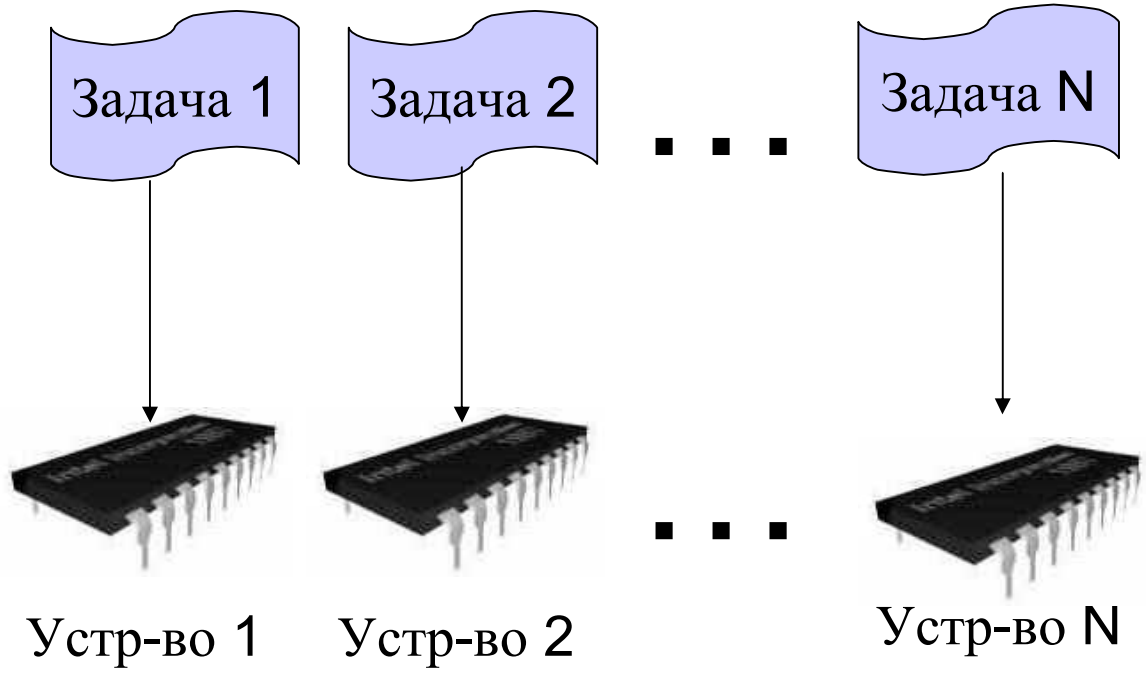
Что дальше?

The Intelligent System

Как создать искусственный интеллект ?



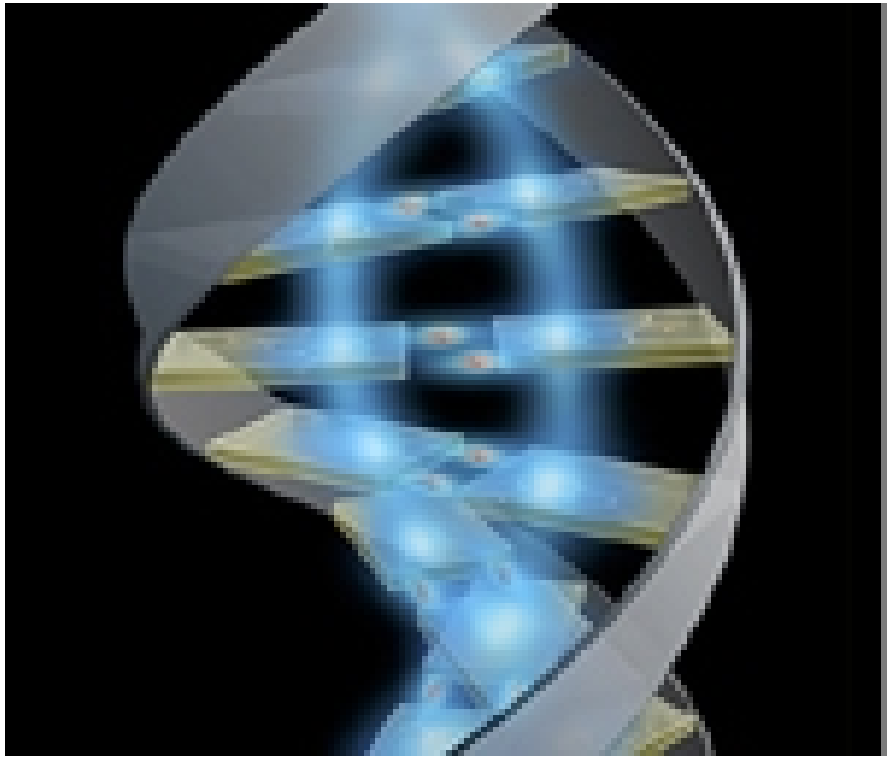
Интеллектуальный выбор

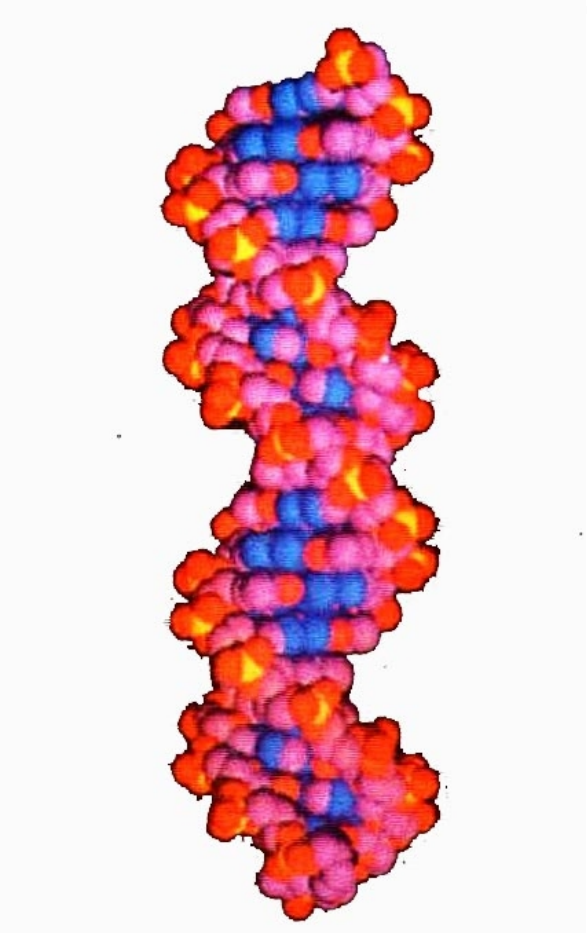


Условия для внутренней части ИС

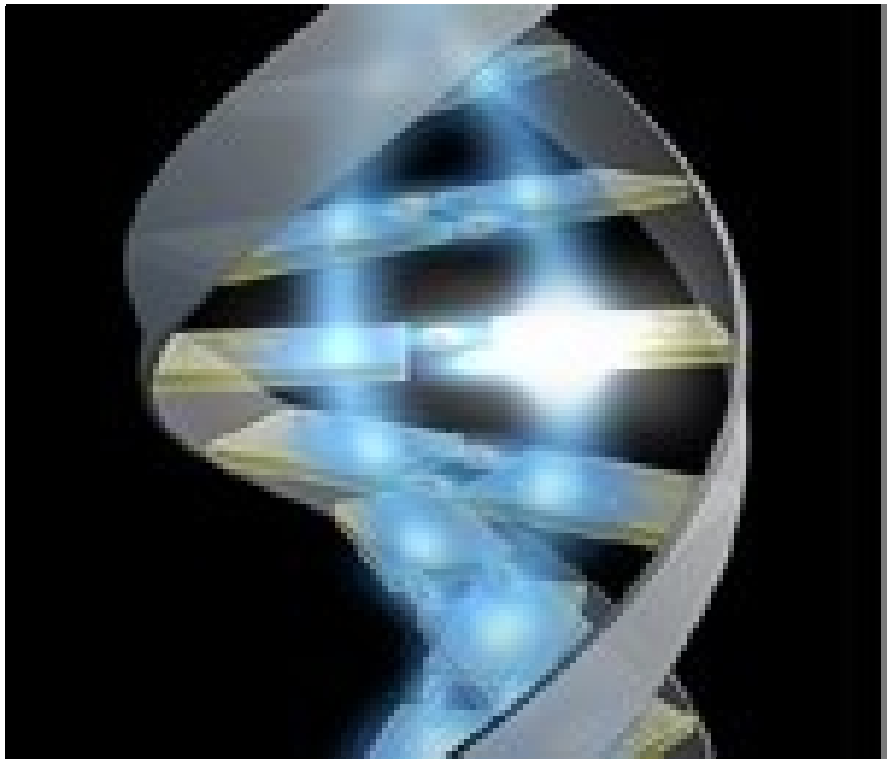
- Для любой задачи в системе есть специальное устройство для ее решения
- Информация из внешнего мира «доставляется» ко всем таким устройства одновременно

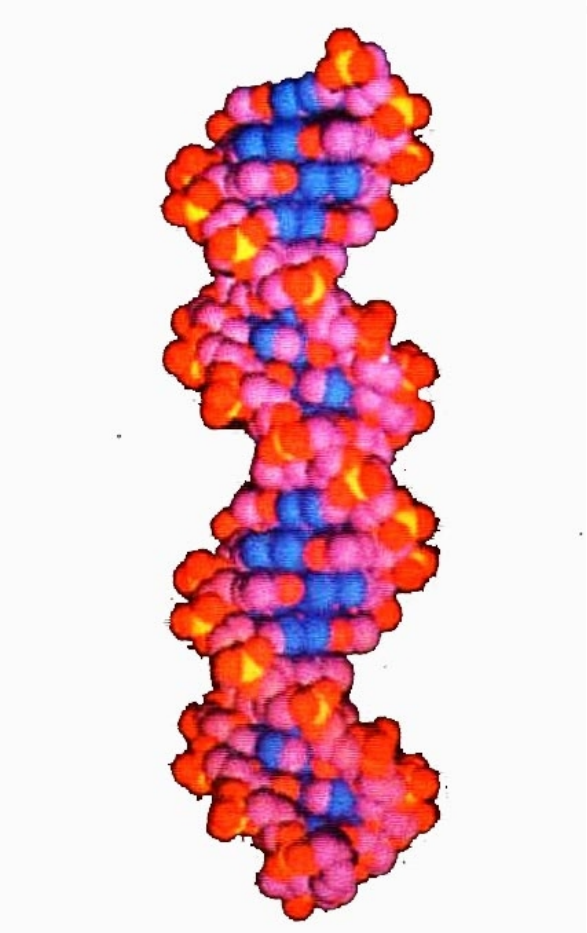
Информационный резонанс



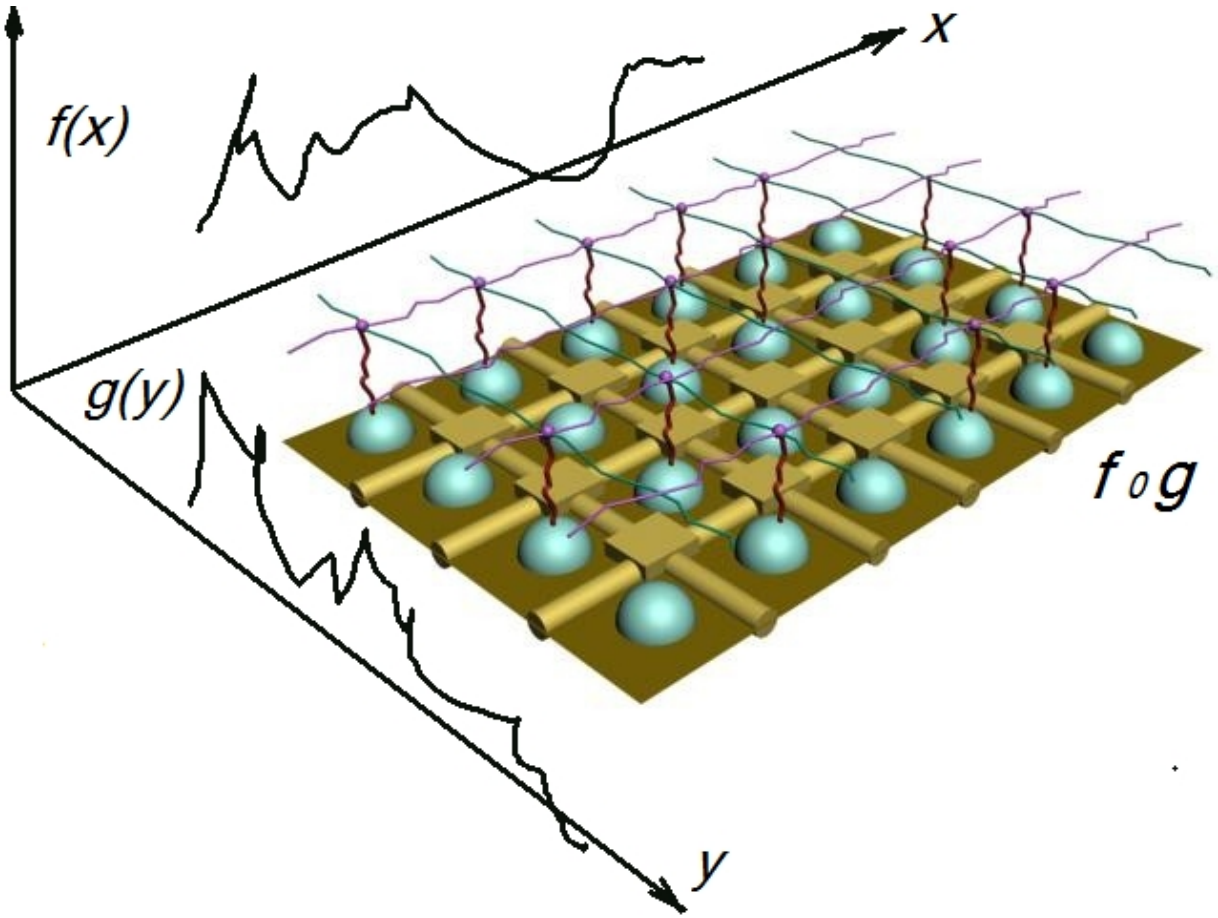


Информационный резонанс





Вычисление свертки



Electronic structure of pristine and activated bacterial S-layers (*Bacillus spharicus*)

S.L. Molodtsov *Inst. of Solid State Physics TU Dresden*

* Supported by SFB463

Cooperation

Institut für Werkstoffwissenschaft

Professur für Materialwissenschaft und Nanotechnik

Inhaber: Wolfgang Pompe

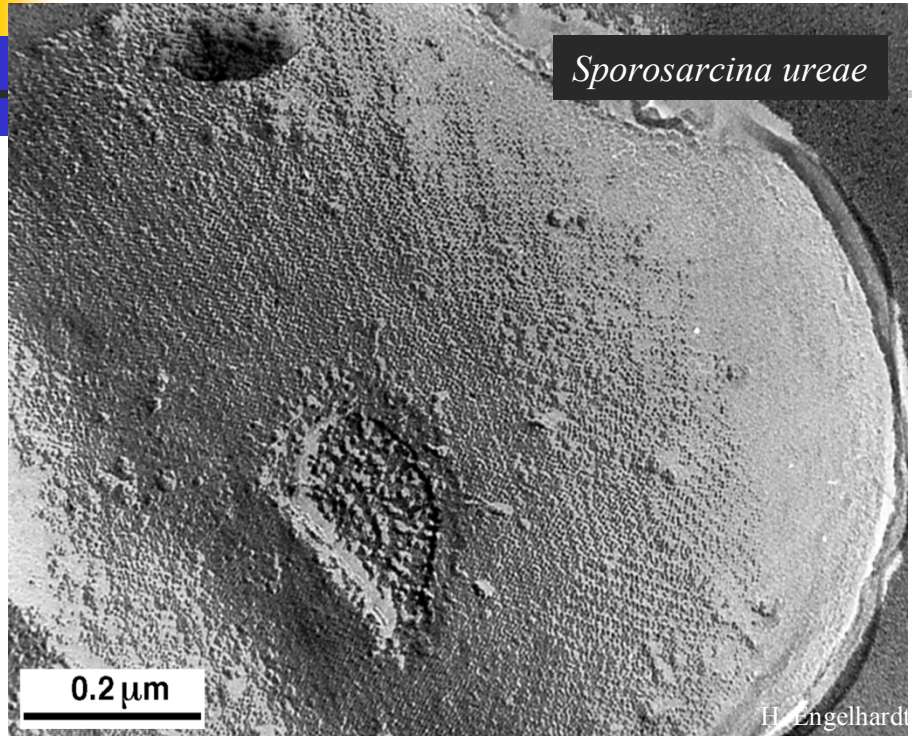
IFP Institut für Festkörperphysik

A. Kirchner, M. Mertig
W. Pompe

D.V. Vyalikh, S. Danzenbächer, M. Besold
Yu.S. Dedkov, S.L. Molodtsov

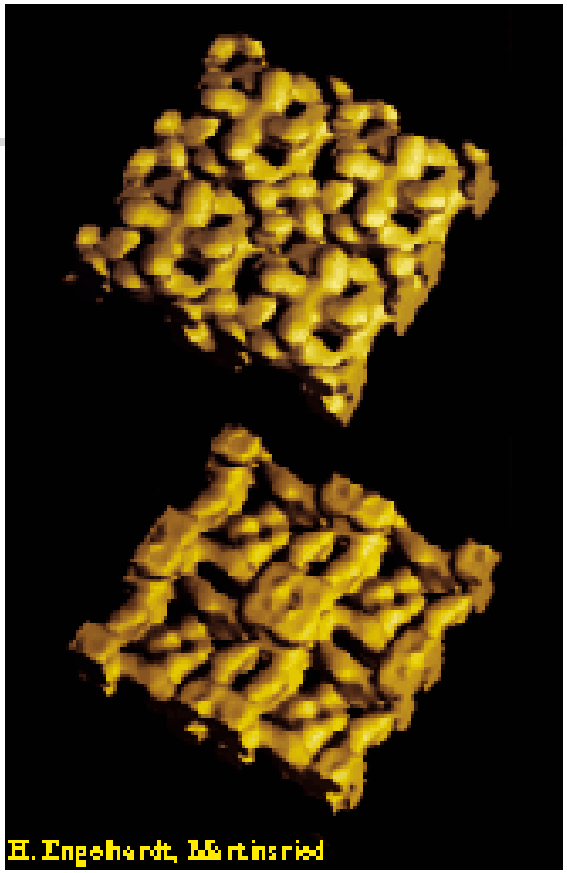
S layer: Structural properties

H. Eberhardt *et al.*, J. Bacteriol. 168 (1986) 309



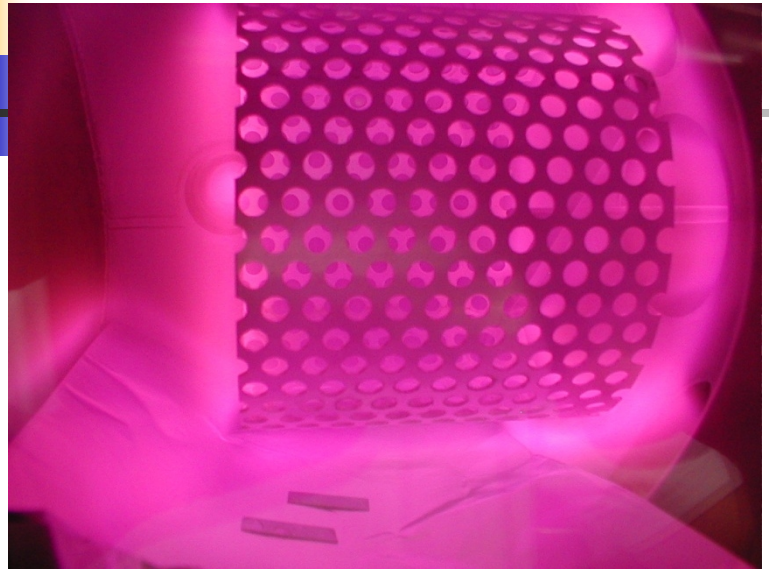
H. Eberhardt, NATO ASI Series (Springer-Verlag,1991)

Transmission electron microscopy image



3D reconstruction of S layer viewed from outer (top) and inner (bottom) surface

Growth of S layer on SiO_x/Si substrates



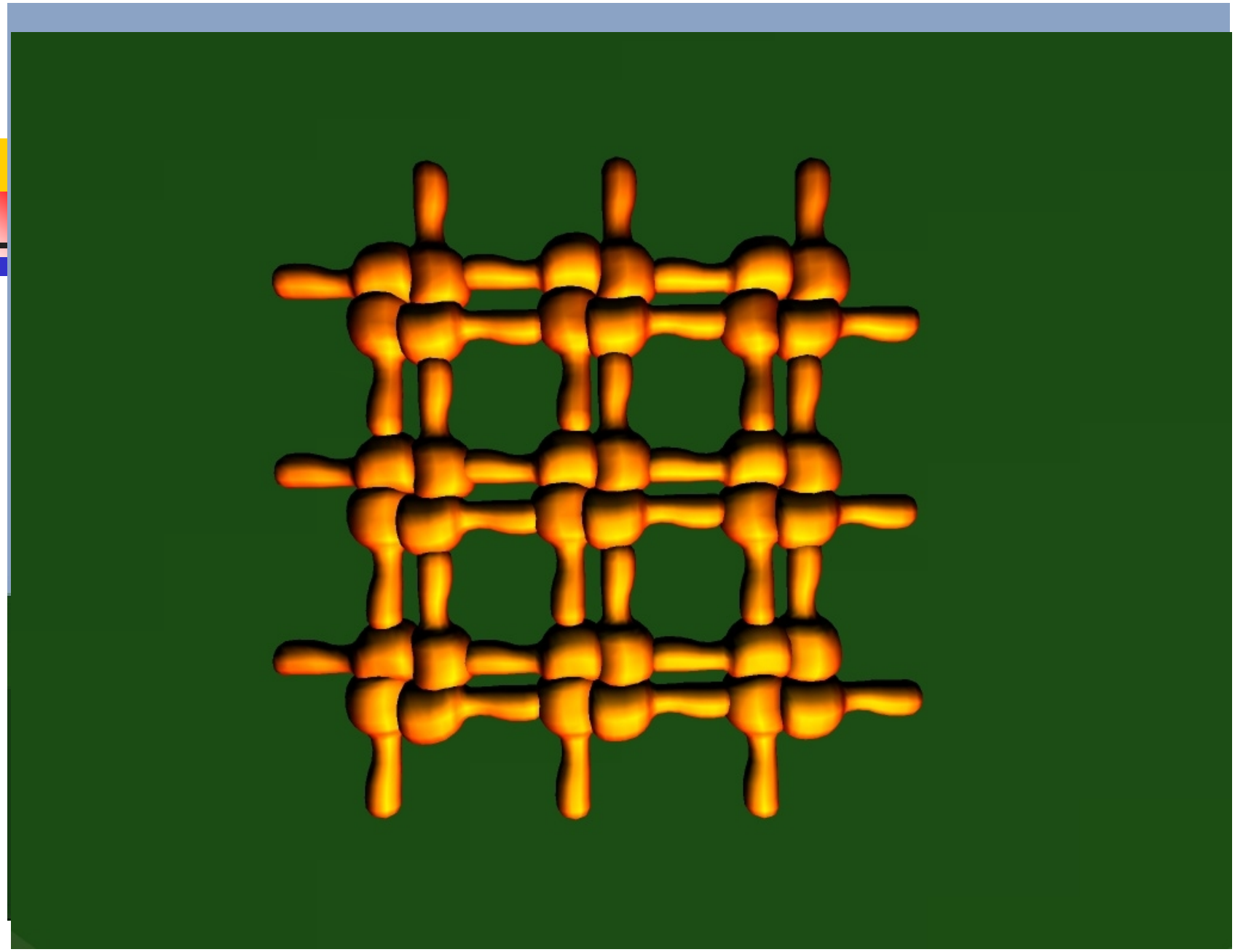
Plasma treatment of substrates

without plasma treatment

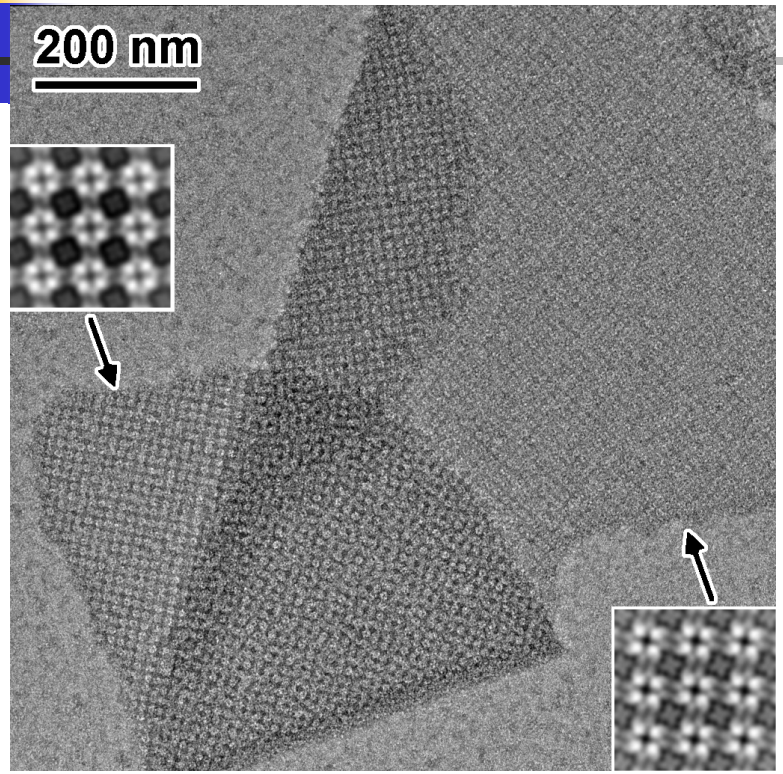
Solution of sodium phosphate, $MgCl_2$ and NaN_3

with plasma treatment

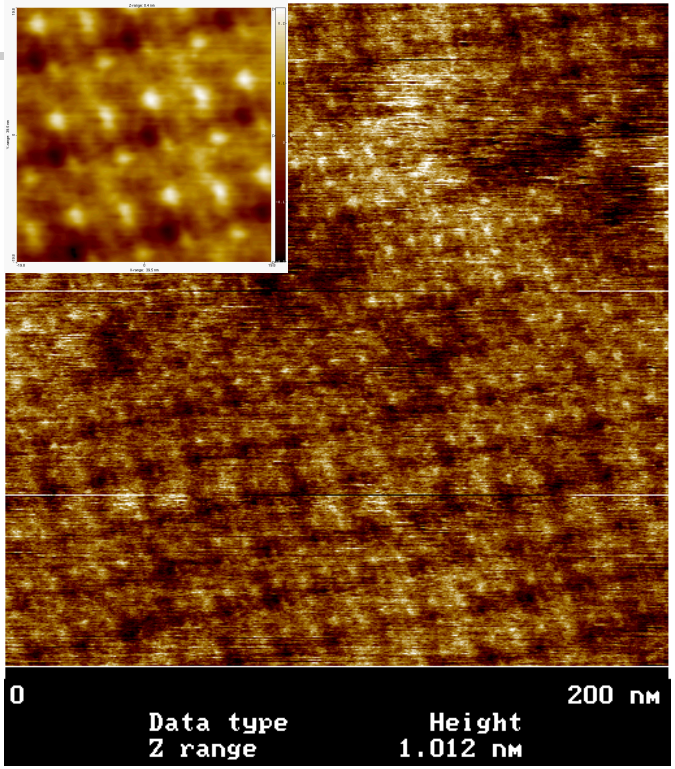
Protein subunits are composed of 20 different amino acids. Their atomic composition is $C_{4912}H_{7860}N_{1324}O_{1615}S_5$



S layer: Sample characterization

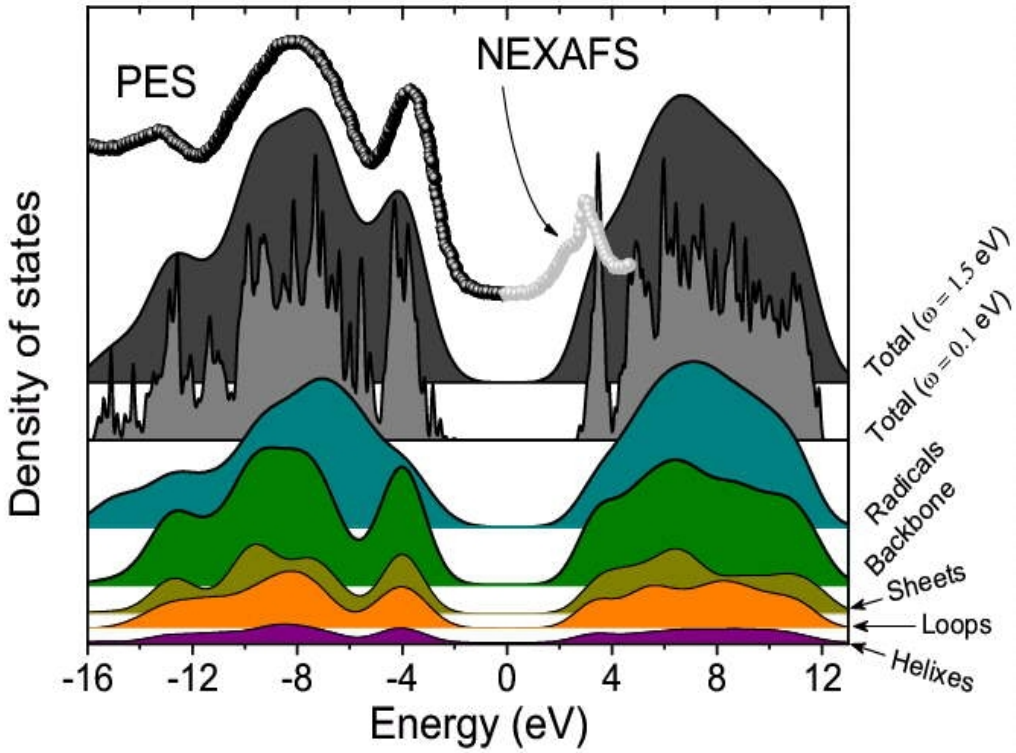


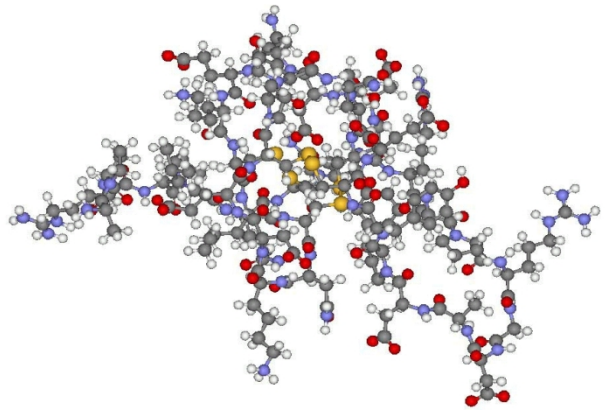
TEM micrograph image

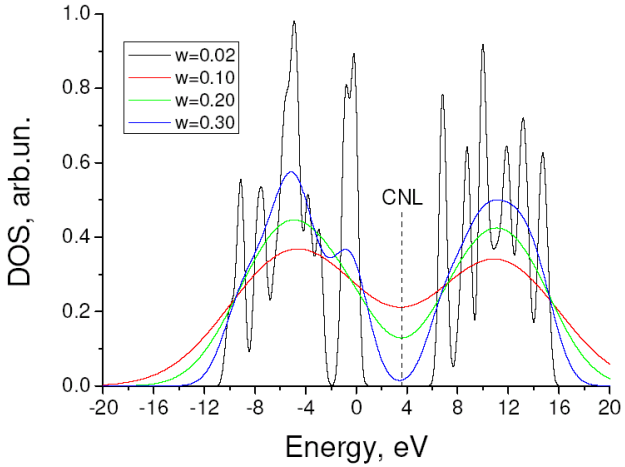


AFM image

Electronic structure of S-layer: LDA, building-block model



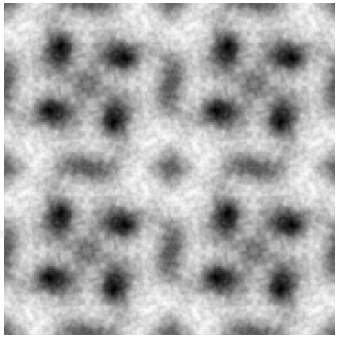




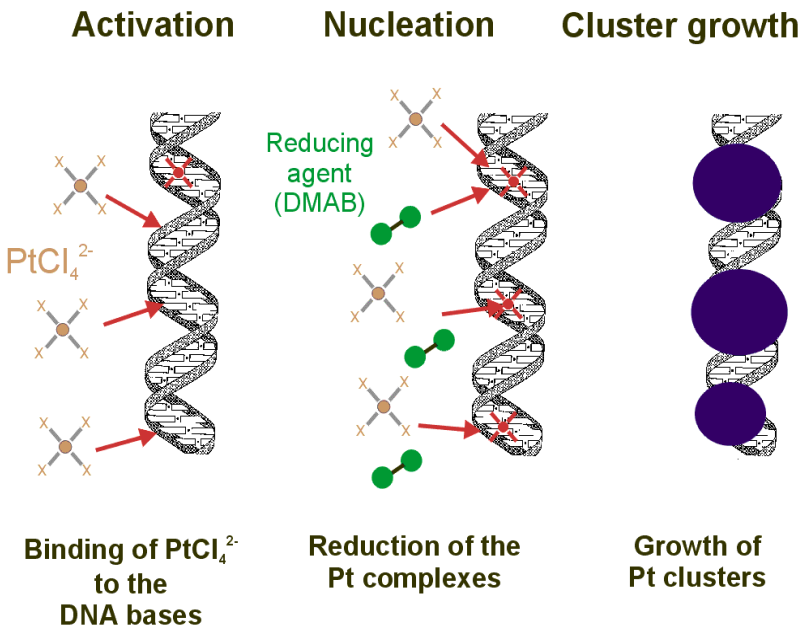
Theory: broadened with exp. resolution

Charge-neutrality level model

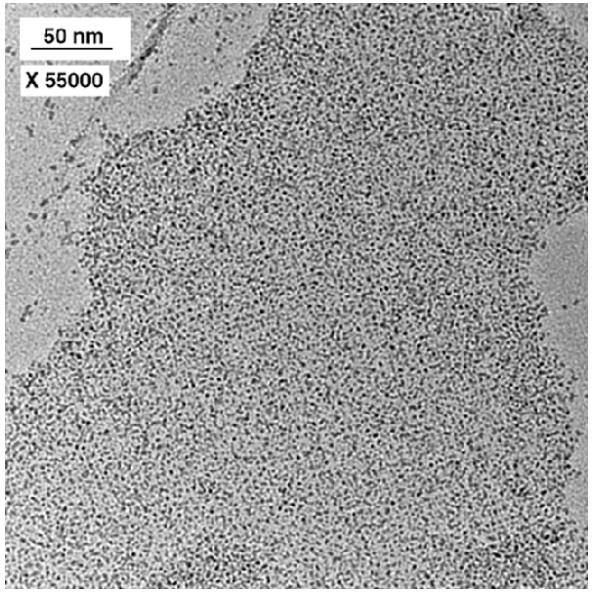
Electronic structure of activated (K_2PtCl_4) and reduced (DMAB) Pt/S-layer



Averaged symmetrized unit cells



Sample preparation



Regular metal cluster arrays