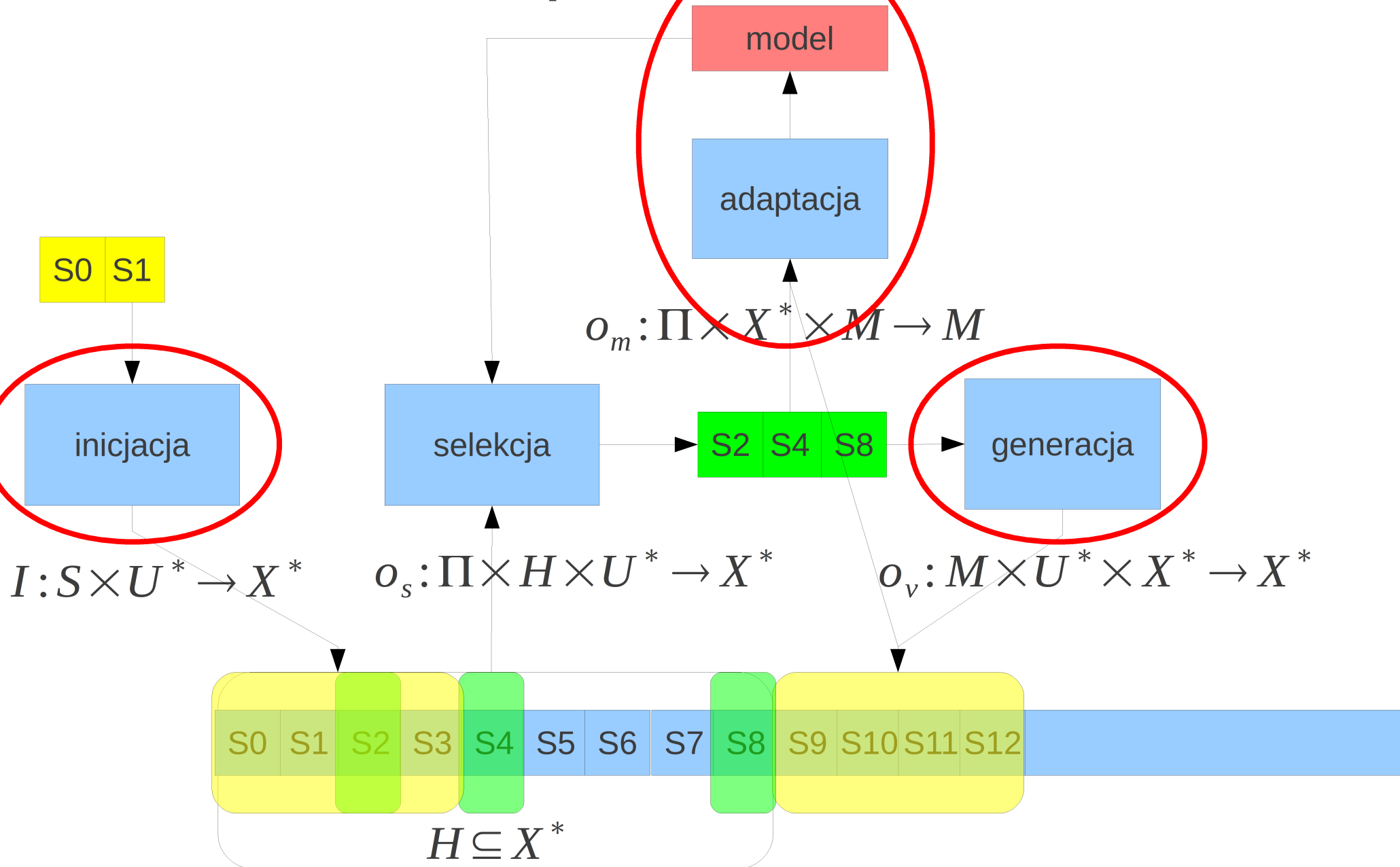


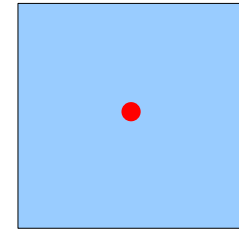
ALHE
Jarosław Arabas
Metaheurystyki w R^n
Ewolucja różnicowa
EDA
CMAES
Rój cząstek

Metoda przeszukiwania

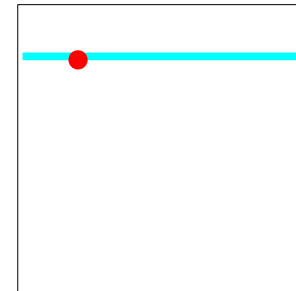


Typowe sposoby generacji sąsiadów

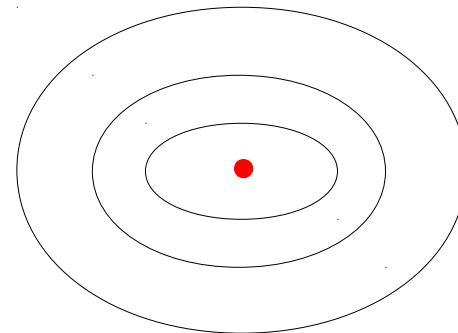
- Rozkład jednostajny w kostce



- Rozkład jednostajny w podprzestrzeni

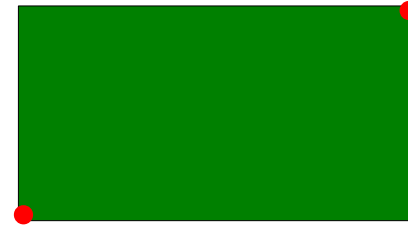


- Rozkład normalny

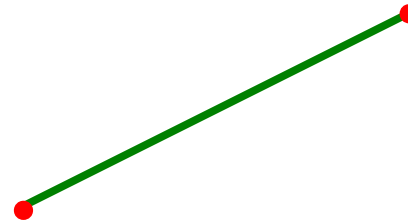


Typowe sposoby generacji punktów “pomiedzy”

- Rozkład jednostajny w kostce



- Rozkład jednostajny na odcinku łączącym



- Rozkład jednostajny na zbiorze narożników kostki

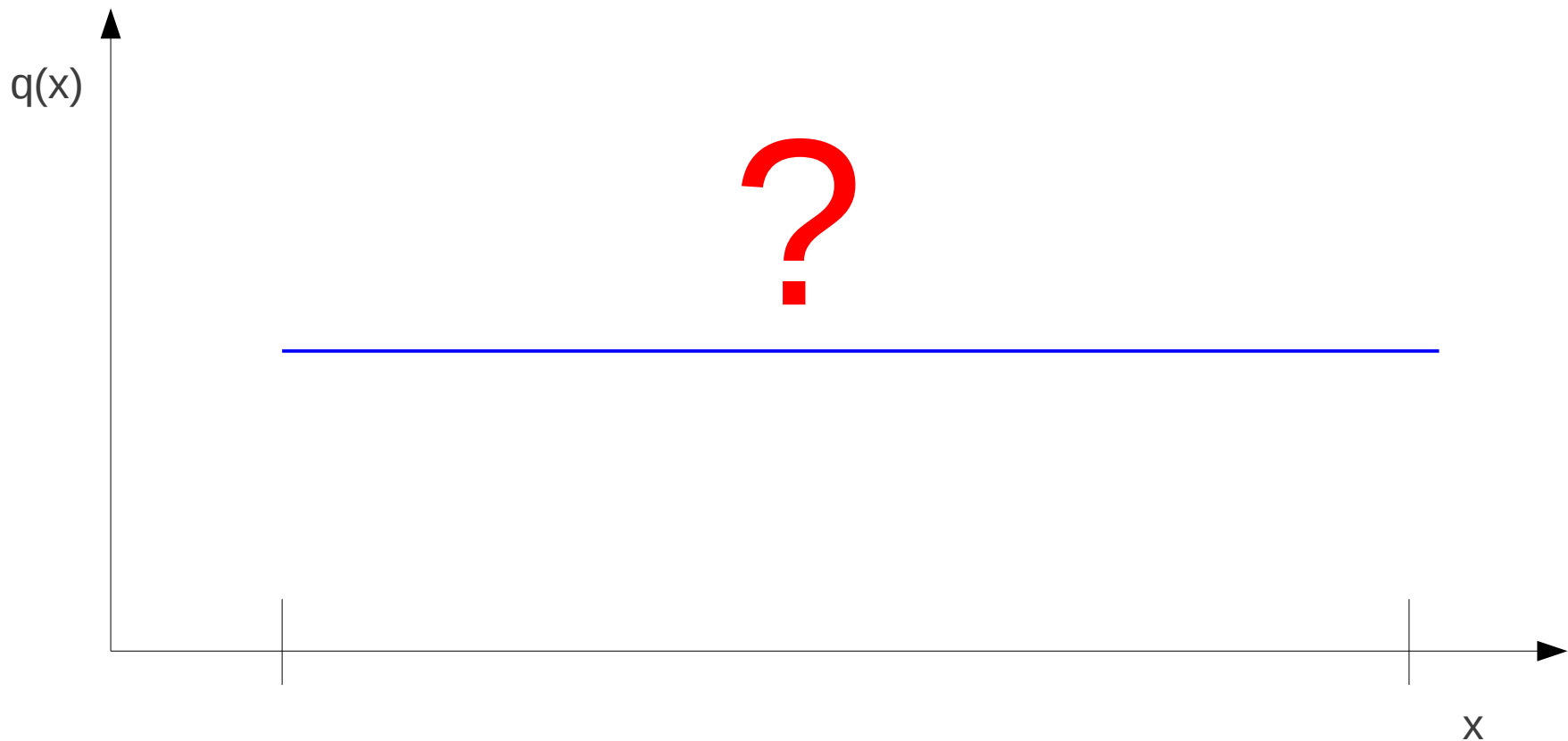


Rozkład próbkowania

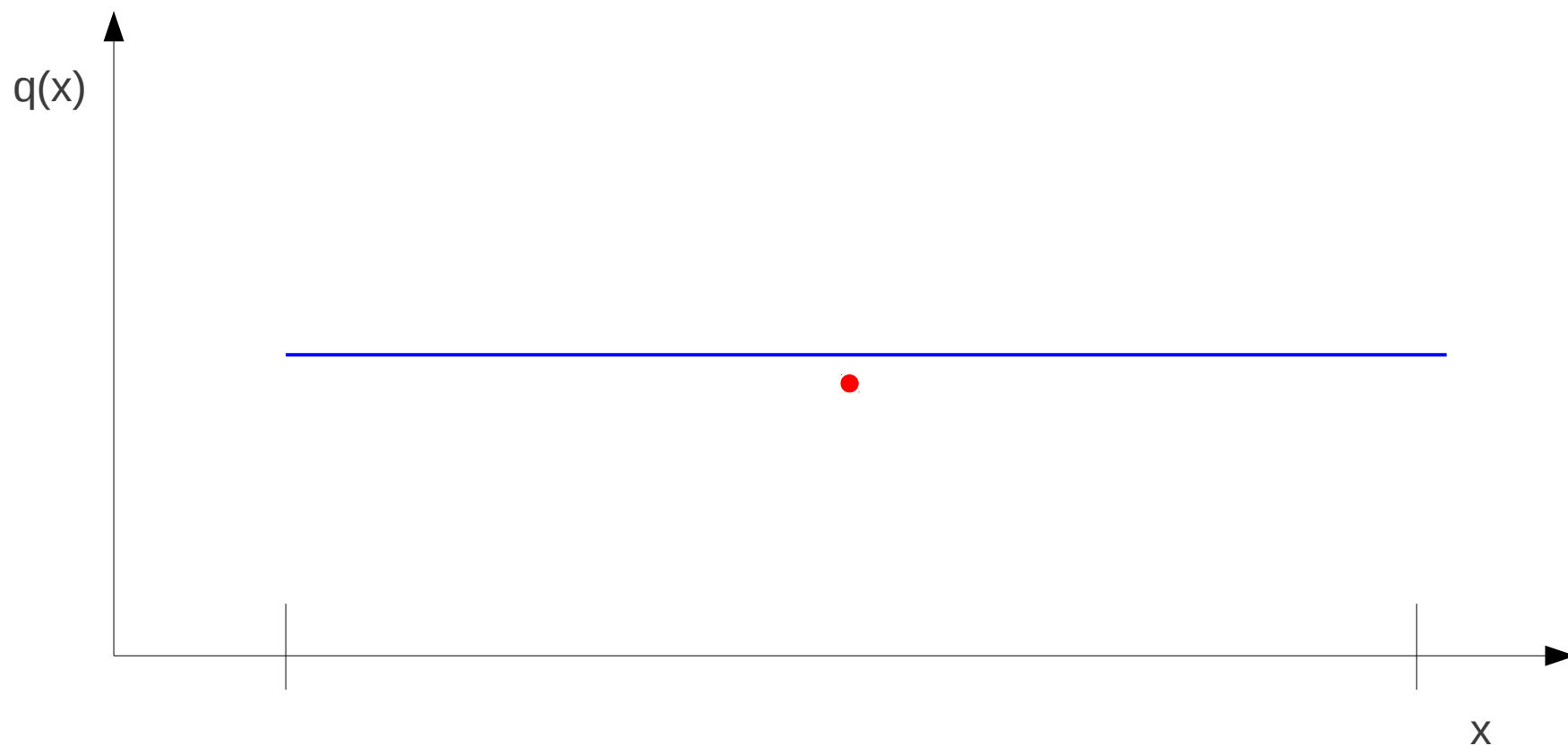
- Sposób tworzenia nowych punktów można opisać za pomocą zmiennej losowej, której rozkład da się analizować
- Jest to rozkład próbkowania
- Rozkład próbkowania może być założony z góry lub może się adaptować
- Przykłady?

Optymalizacja jako proces adaptacji

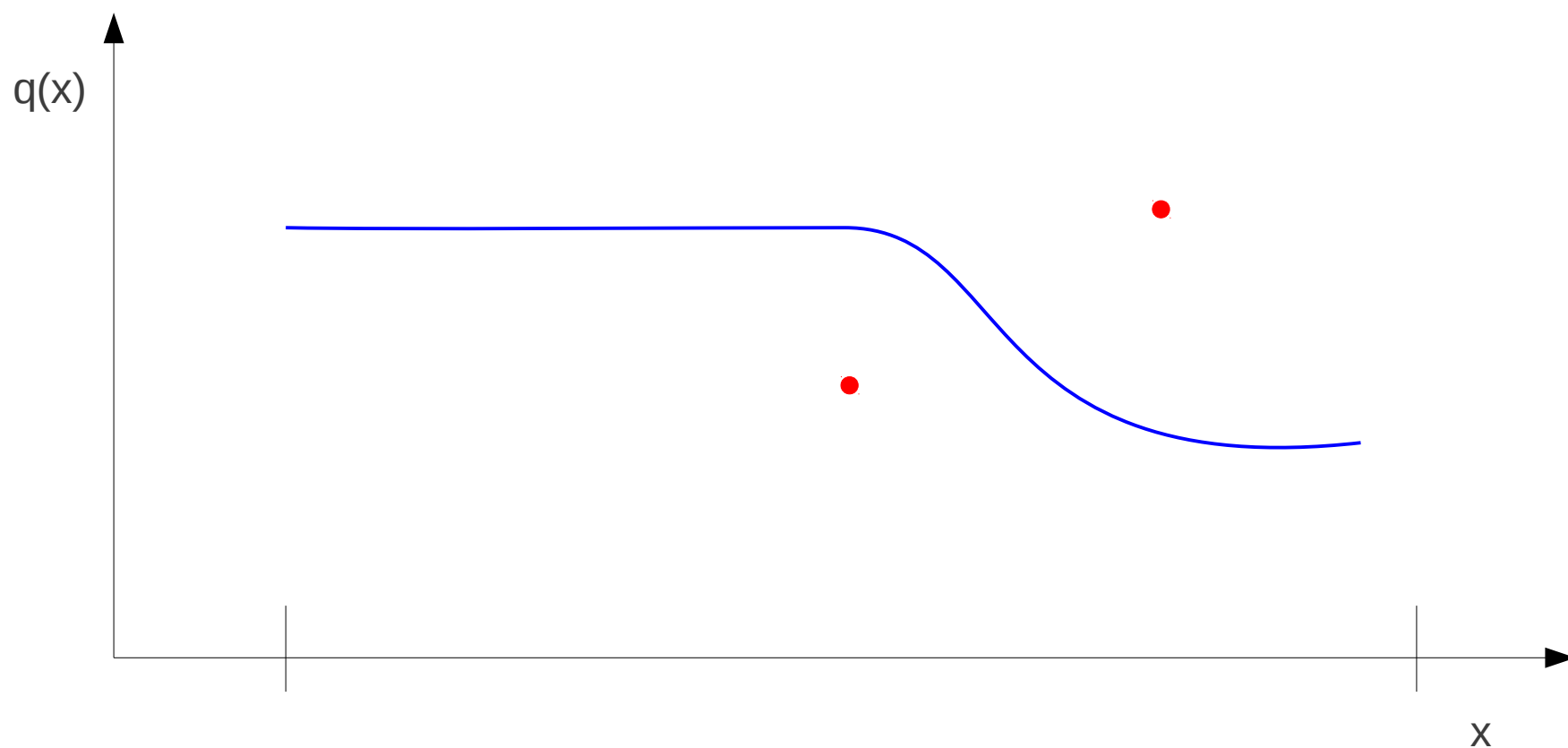
- Przekonanie o położeniu optimum



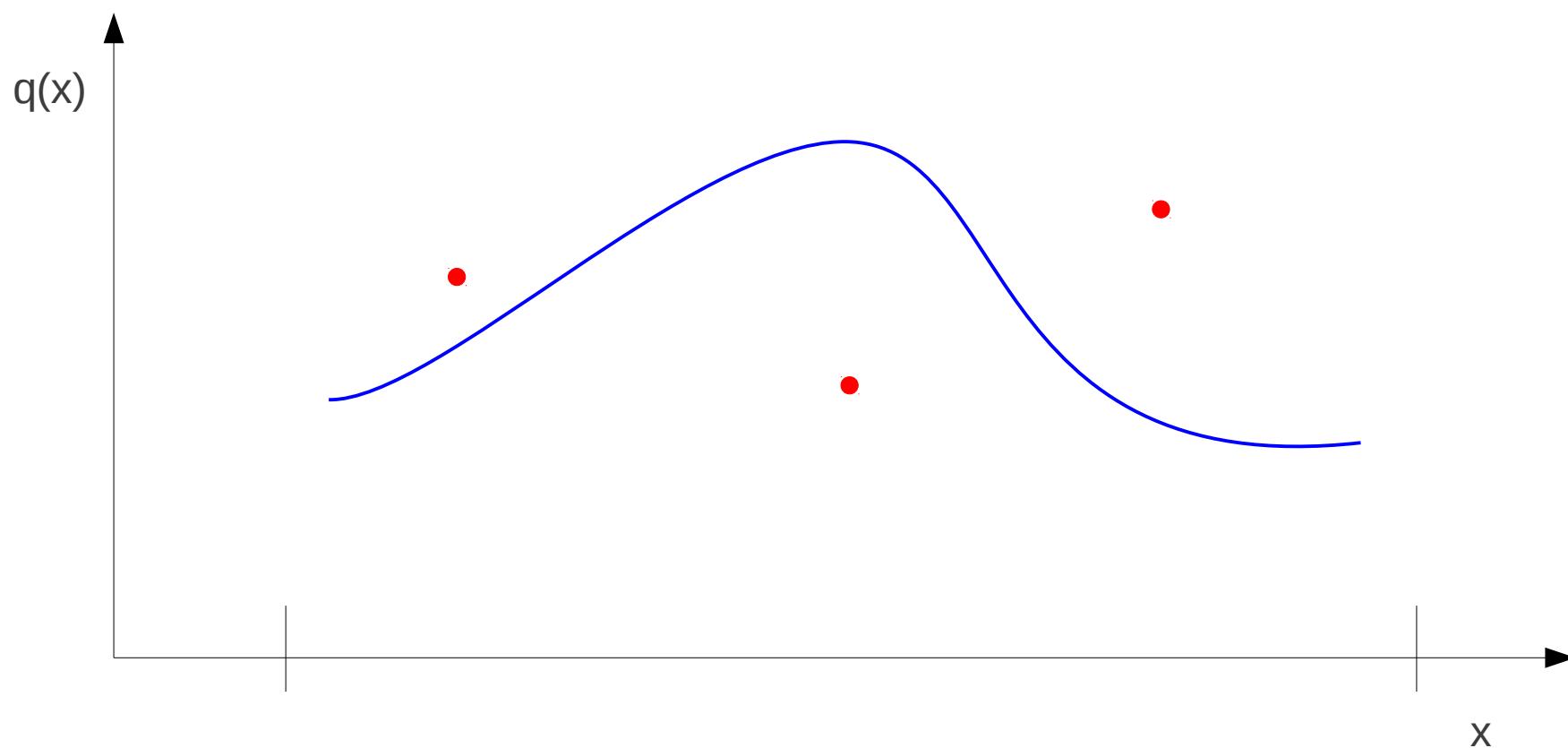
Optymalizacja jako proces adaptacji



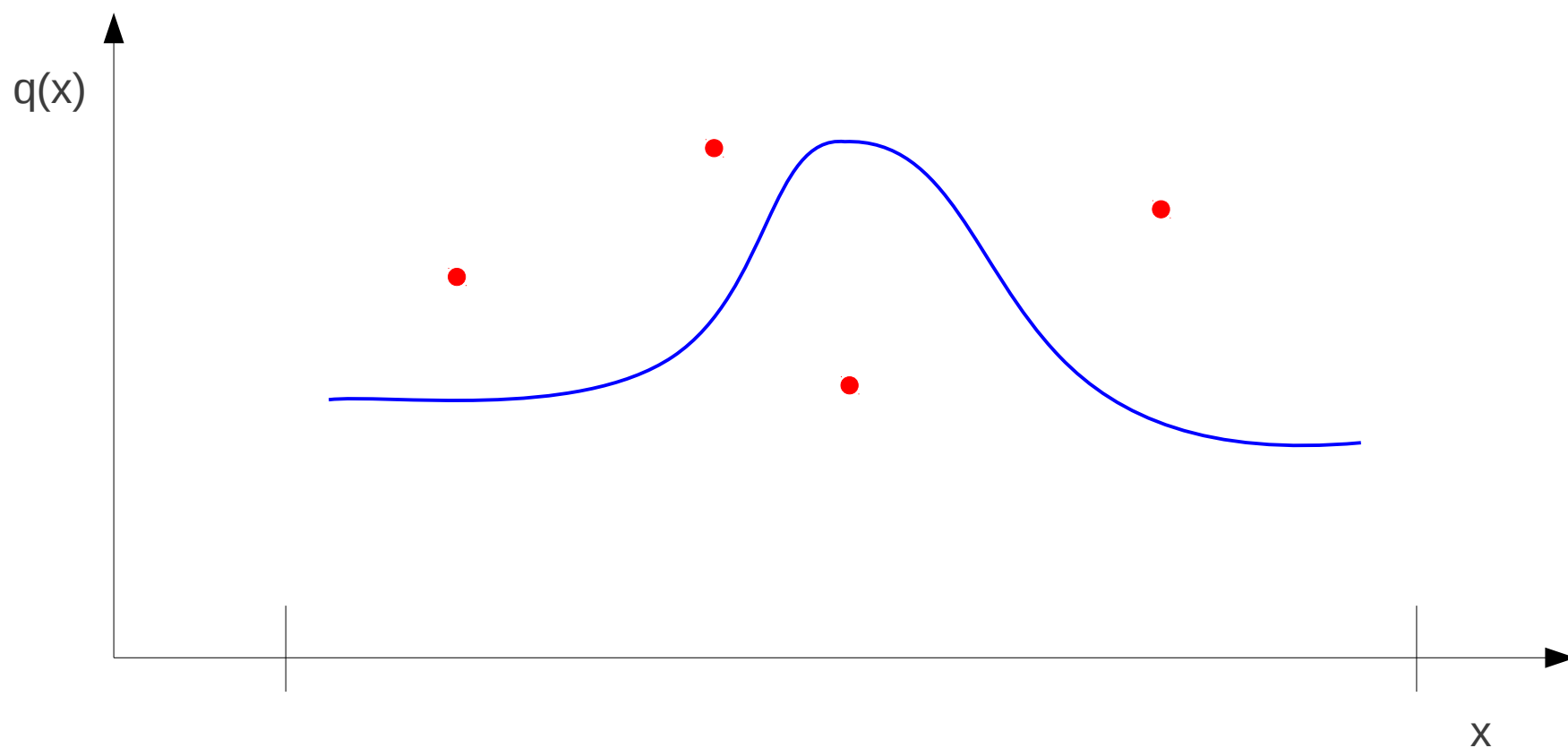
Optymalizacja jako proces adaptacji



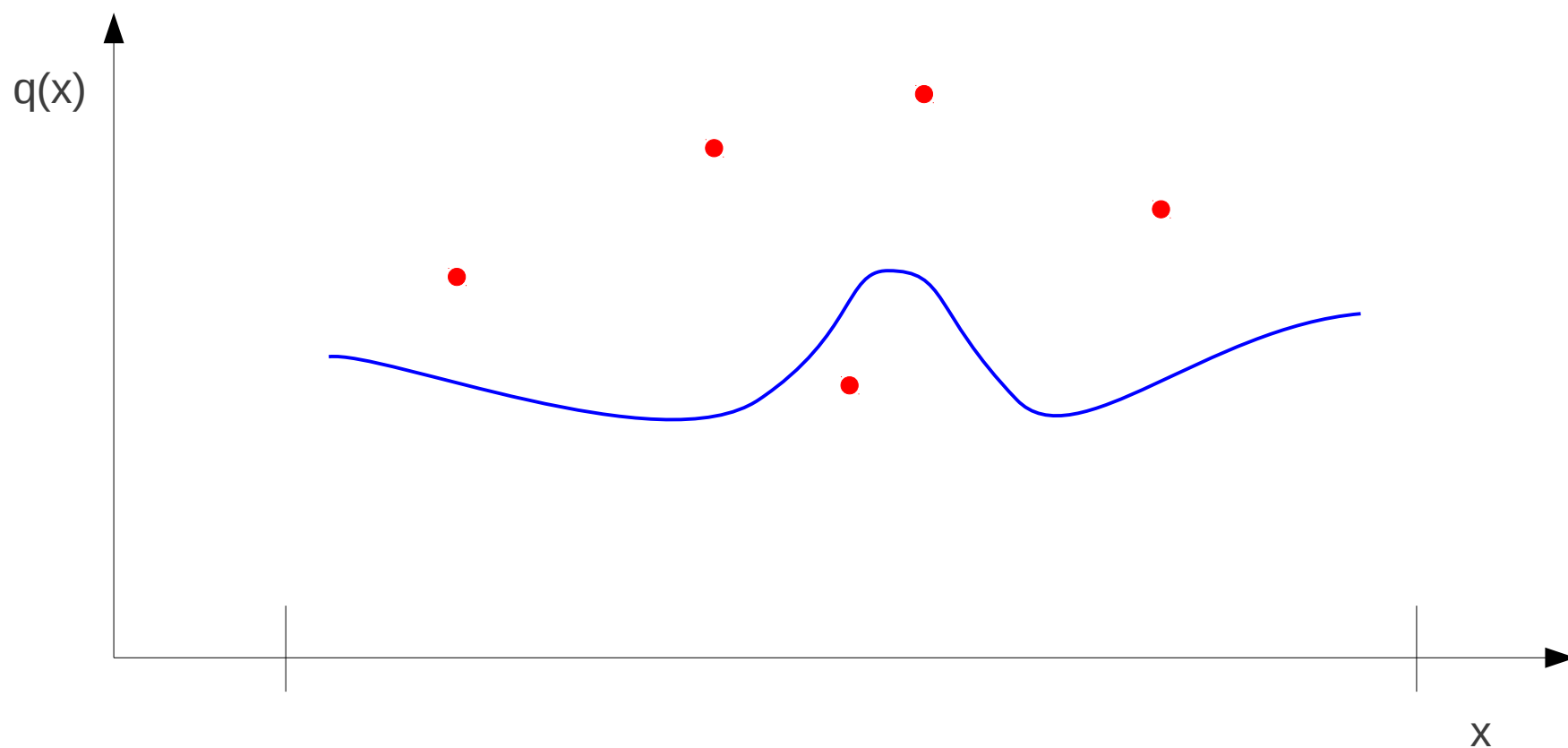
Optymalizacja jako proces adaptacji



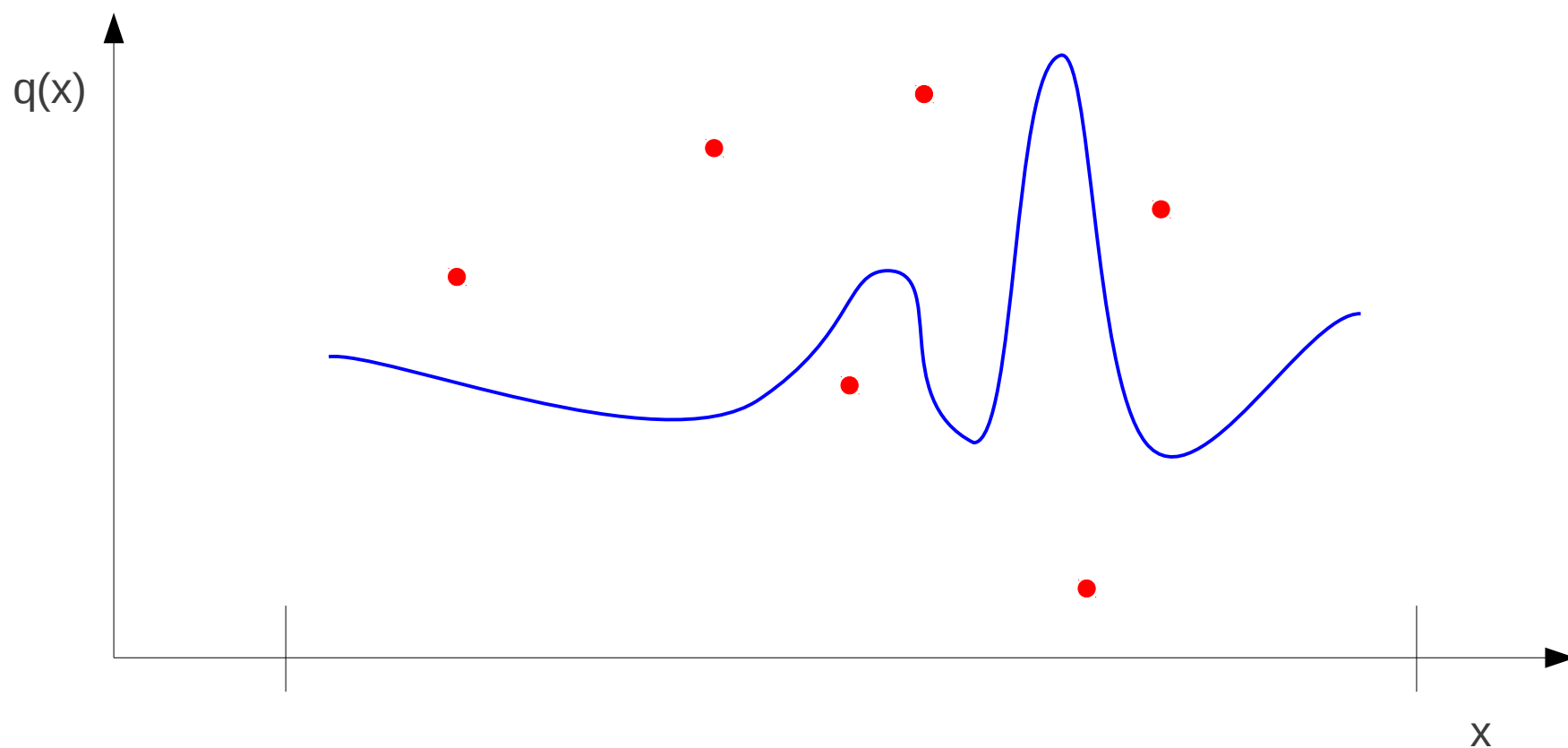
Optymalizacja jako proces adaptacji



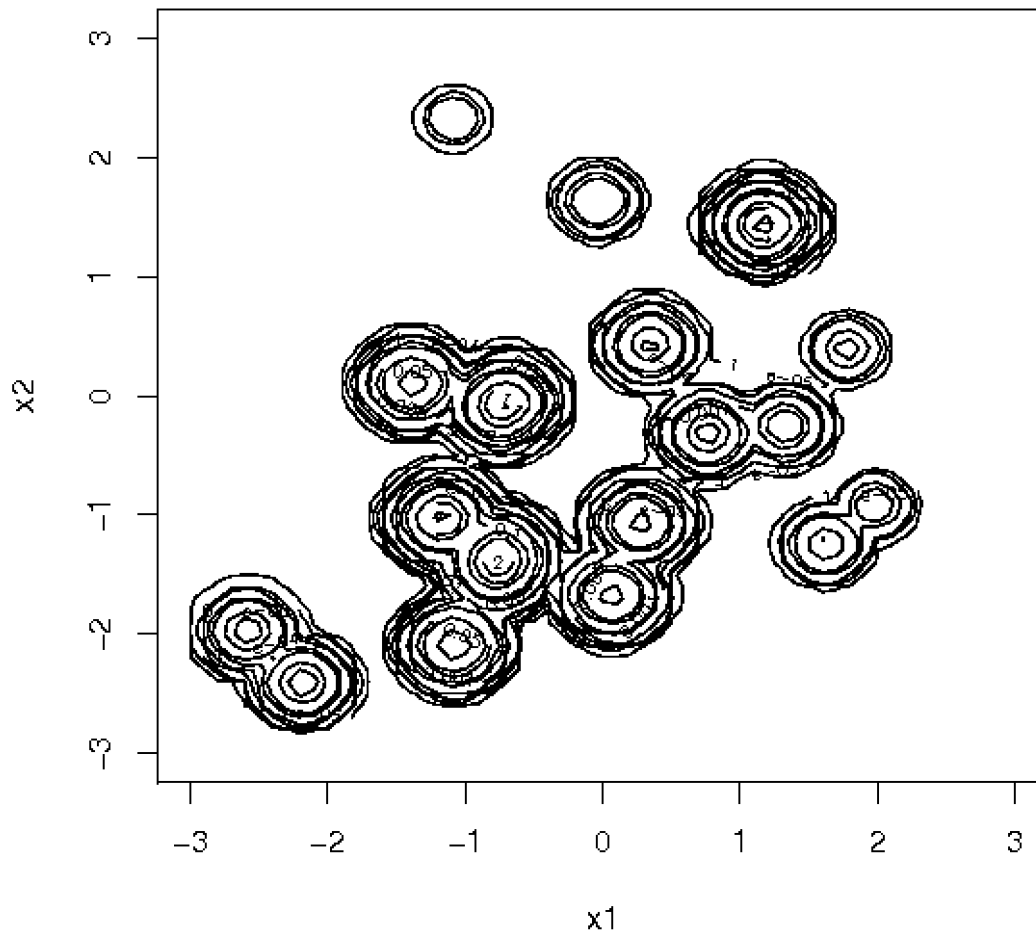
Optymalizacja jako proces adaptacji



Optymalizacja jako proces adaptacji

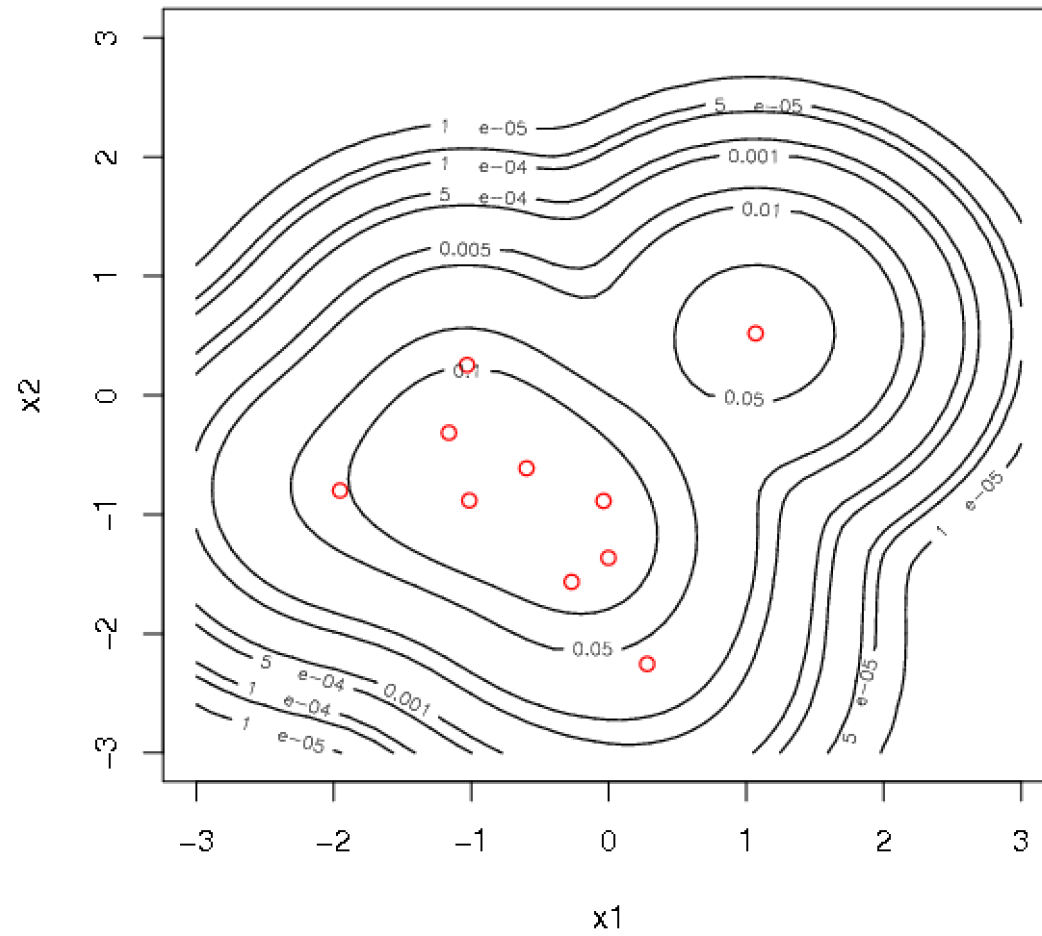


- Obserwujemy wartości funkcji celu w wielu punktach
- Okolice każdego punktu są tym chętniej próbkowane, im jest on lepszy

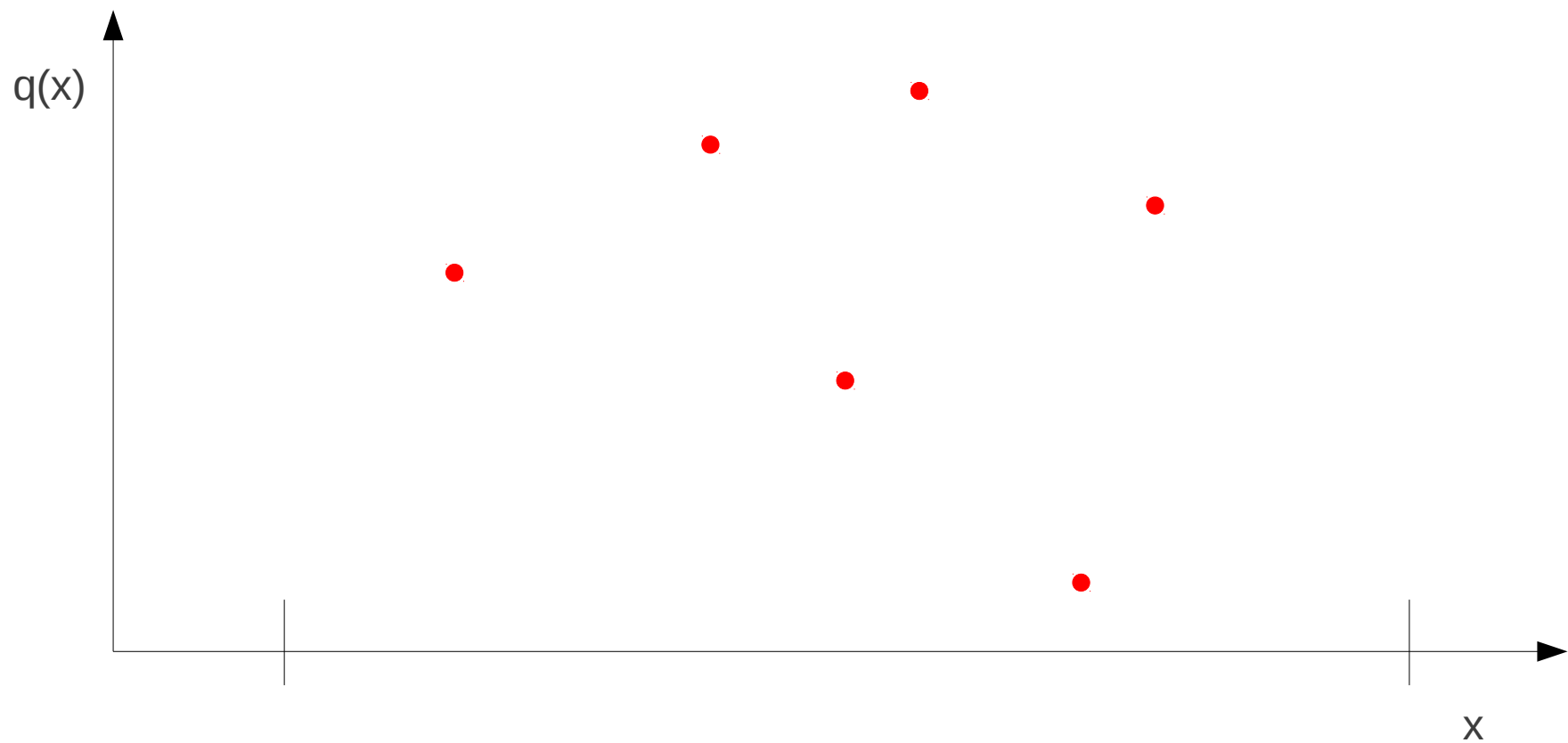


- Obserwujemy wartości funkcji celu w wielu punktach
- Okolice każdego punktu są tym chętniej próbkowane, im jest on lepszy

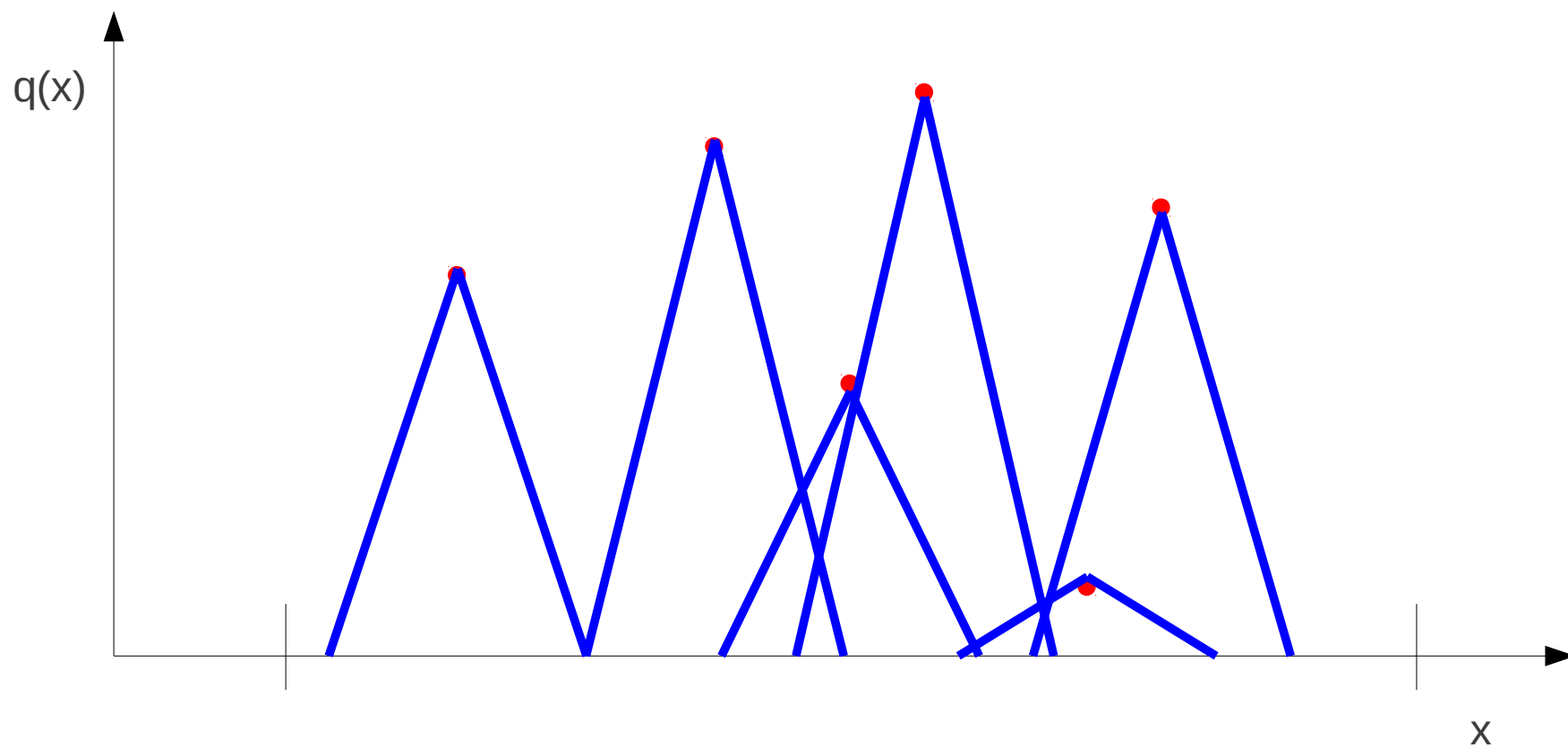
Po pewnym czasie



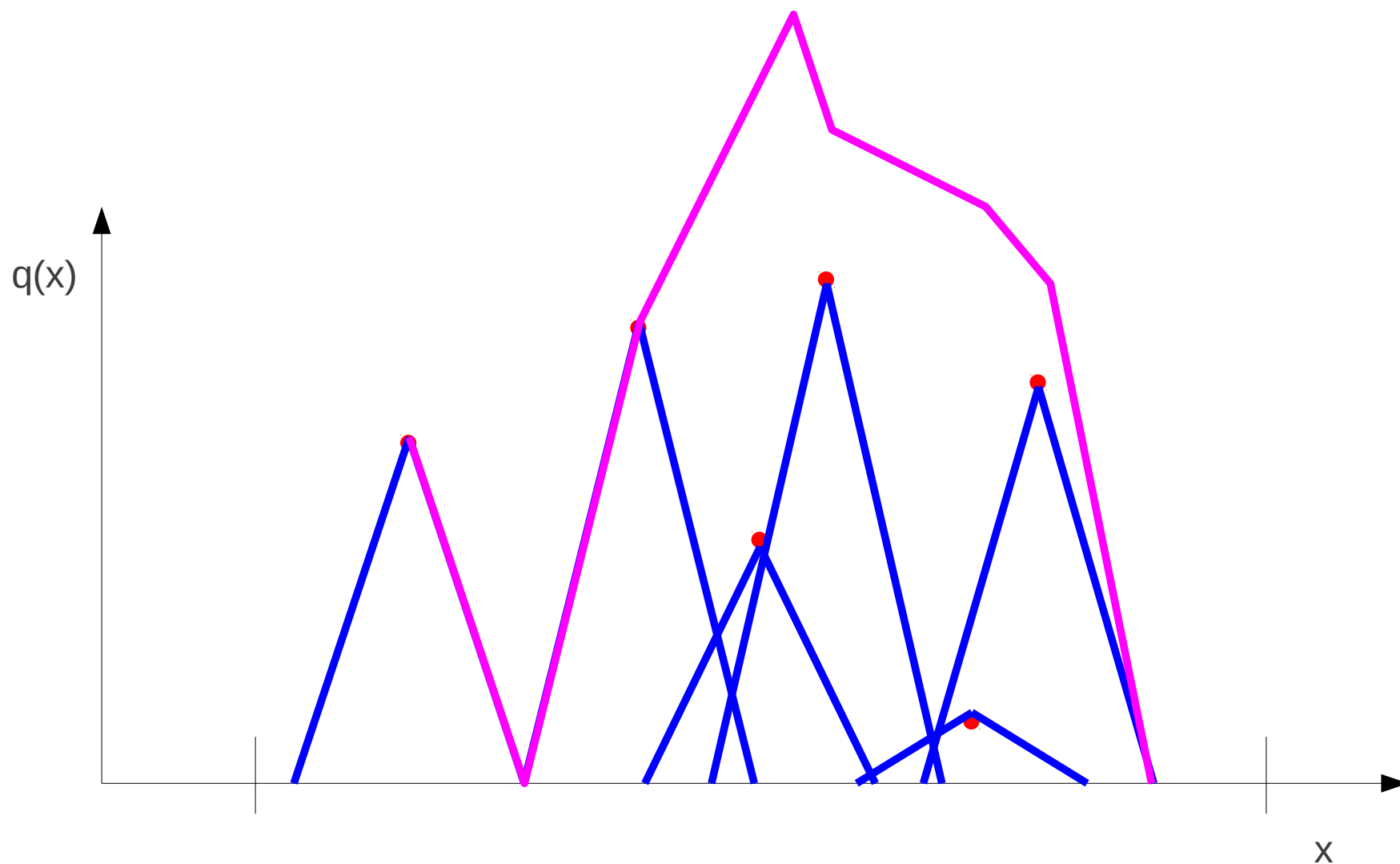
Mieszanina rozkładów p-stwa



Mieszanina rozkładów p-stwa



Mieszanina rozkładów p-stwa



Ewolucja różnicowa

algorytm differential evolution

inicjuj $P^0 \leftarrow \{P_1^0, P_2^0 \dots P_\mu^0\}$

$H \leftarrow P^0$

$t \leftarrow 0$

while ! stop

for ($i \in 1:\mu$)

$P_j^t \leftarrow \text{select}(P^t)$

$P_k^t, P_l^t \leftarrow \text{sample}(P^t)$

$M_i^t \leftarrow P_j^t + F(P_k^t - P_l^t)$

$O_i^t \leftarrow \text{crossover}(P_i^t, M_i^t)$

$H \leftarrow H \cup \{O_i^t\}$

$P_i^{t+1} \leftarrow \text{tournament}(P_i^t, O_i^t)$

$t \leftarrow t + 1$

sample jest procesem wyboru pary punktów z jednakowym p-stwem

crossover jest operacją krzyżowania wymieniającego

Typy ewolucji różnicowej - klasyka

- Typ selekcji
 - wybór losowego (rand)
 - wybór najlepszego w populacji (best)
- Typ krzyżowania
 - dwumianowe (bin)
 - wykładnicze (exp)
- Liczba par różnicowanych punktów – 1 albo 2
- Konwencja oznaczeń: DE/rand/1/bin

Typy krzyżowania

procedure binomial crossover

arguments: x, y

for ($i \in 1:n$)

if $a < c_r$

$z_i \leftarrow y_i$

else

$z_i \leftarrow x_i$

return z

procedure exponential crossover

arguments: x, y

$i \leftarrow 1$

while ($i \leq n$)

if $a < c_r$

$z_i \leftarrow y_i$

else break

while ($i \leq n$)

$z_i \leftarrow x_i$

return z

a jest zmienną losową o rozkładzie jednostajnym w $(0,1)$

c_r jest parametrem

Krzyżowanie wykładnicze a jednopunktowe

0.0759	0.062	-1.893	0.053
--------	-------	--------	-------

Rodzic 1

1	1	1	0
---	---	---	---

wagi

0.328	0.631	-0.299	0.194
-------	-------	--------	-------

Rodzic 2

0.328	0.631	-0.299	0.053
-------	-------	--------	-------

Potomek

Krzyżowanie wykładnicze a jednopunktowe

1	1	1	0	wagi
---	---	---	---	------

W krzyżowaniu **jednopunktowym** rozkład prawdopodobieństwa pojawienia się przejścia między jedynką a zerem jest rozkładem **jednostajnym**

1	1	1	0	
1/5	1/5	1/5	1/5	1/5

W krzyżowaniu wykładniczym rozkład ten jest rozkładem (prawie) wykładniczym

1	1	1	0	
1/2	1/4	1/8	1/16	1/16
p	p ²	p ³	p ⁴	dopełnienie do 1

Krzyżowanie równomierne a dwumianowe

0.0759	0.062	-1.893	0.053
--------	-------	--------	-------

Rodzic 1

0	1	1	0
---	---	---	---

wagi

0.328	0.631	-0.299	0.194
-------	-------	--------	-------

Rodzic 2

0.0759	0.631	-0.299	0.053
--------	-------	--------	-------

Potomek

Krzyżowanie równomierne a dwumianowe

0	1	1	0	wagi
---	---	---	---	------

W krzyżowaniu **równomiernym** prawdopodobieństwo pojawienia się jedynki i zera na każdej pozycji jest równe $1/2$

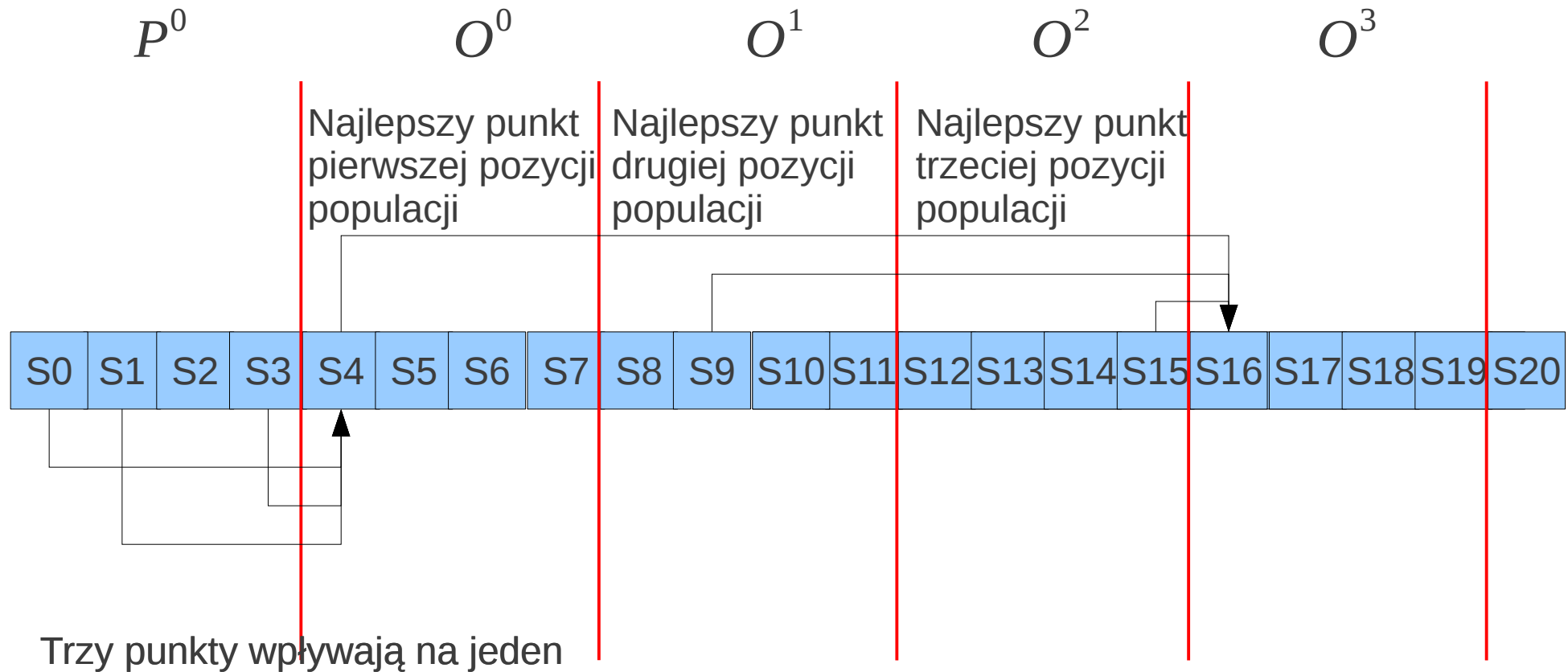
0	1	1	0
$1/2$	$1/2$	$1/2$	$1/2$

W krzyżowaniu **dwumianowym** te p-stwa są różne

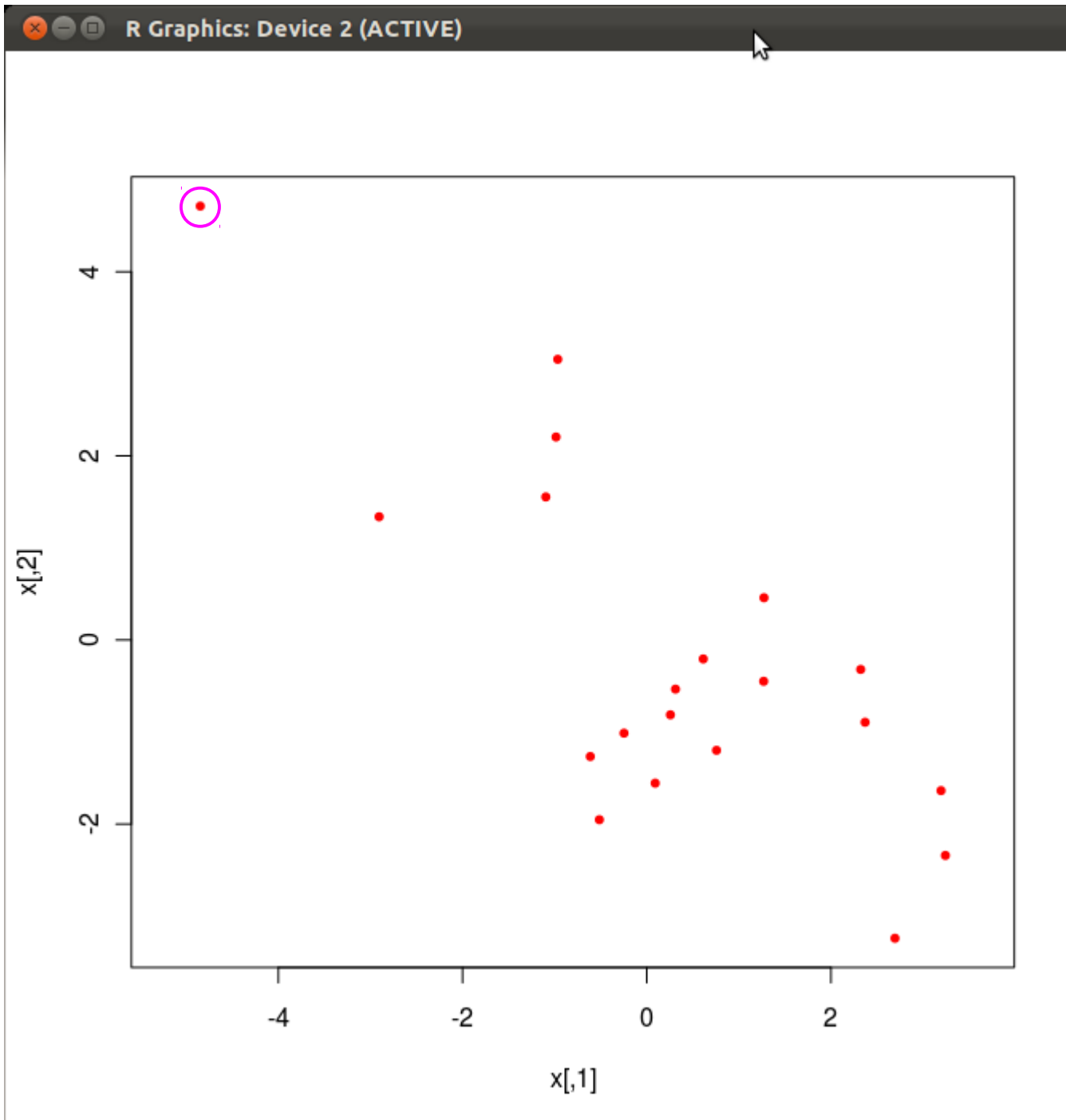
0	1	1	0
$1-p_e$	p_e	p_e	$1-p_e$

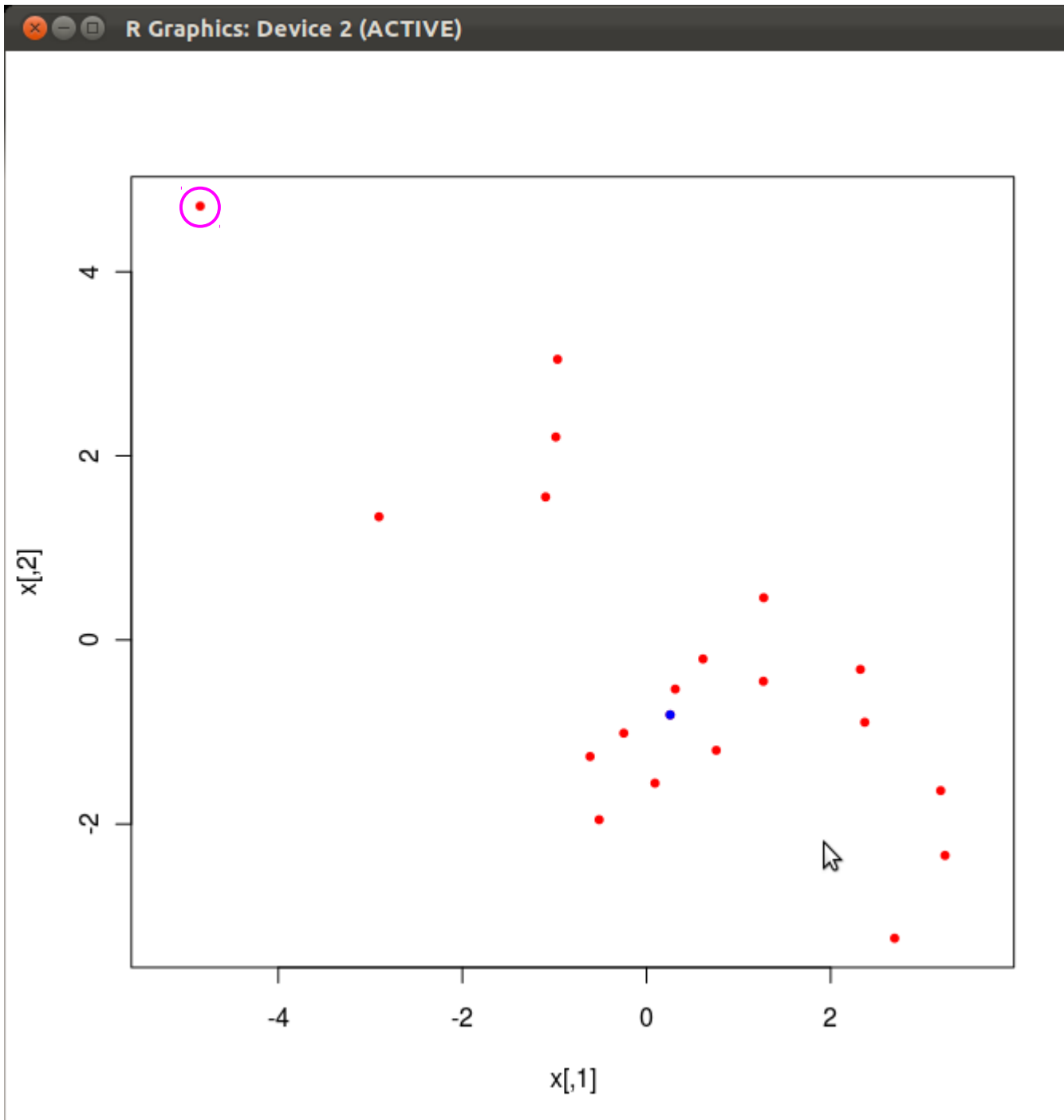
W obu przypadkach, rozkład p-stwa pojawienia się k jedynek i $n-k$ zer jest rozkładem Bernoulliego (wg angielskiej nomenklatury dwumianowym)

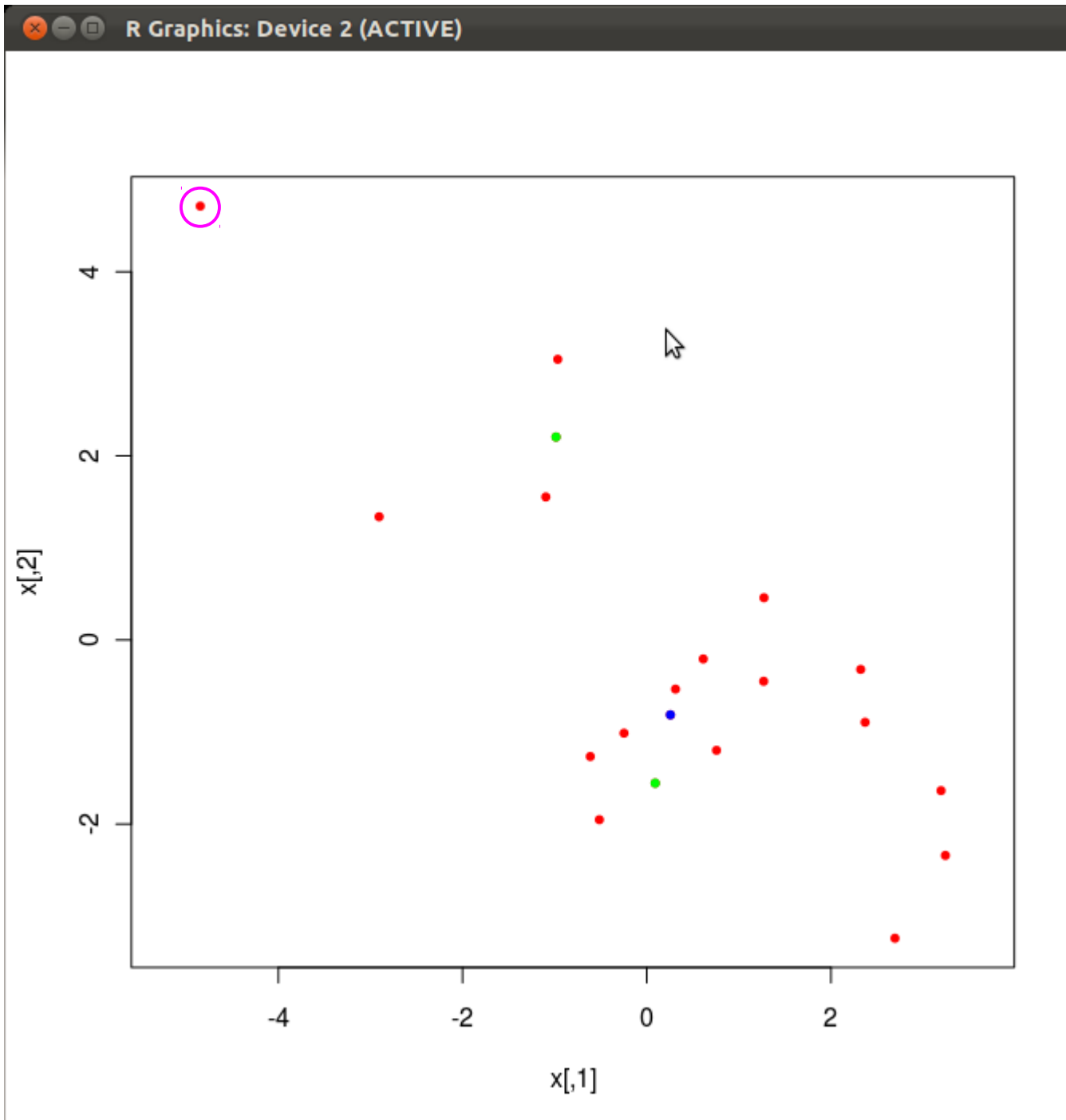
DE/rand/1

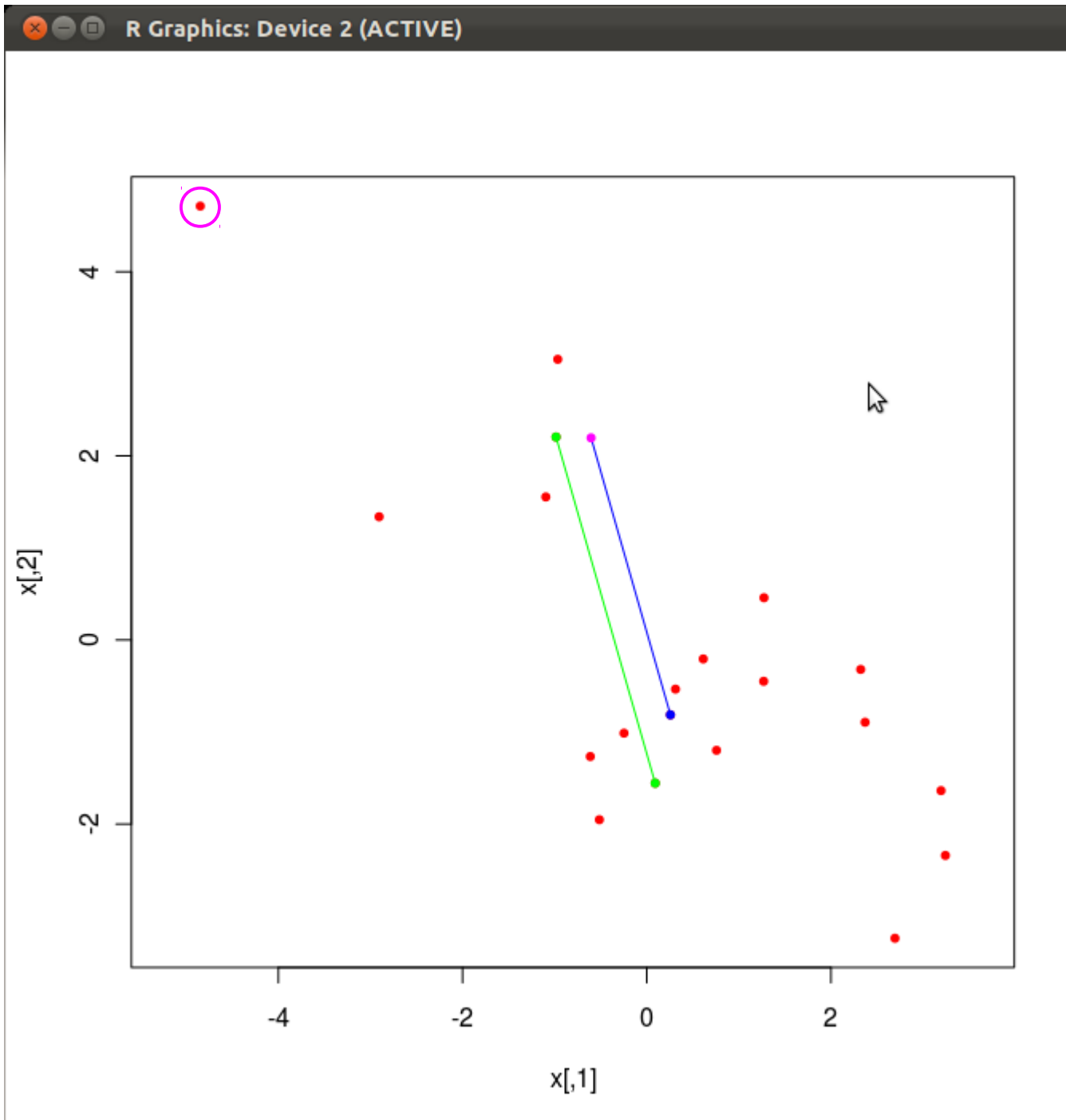


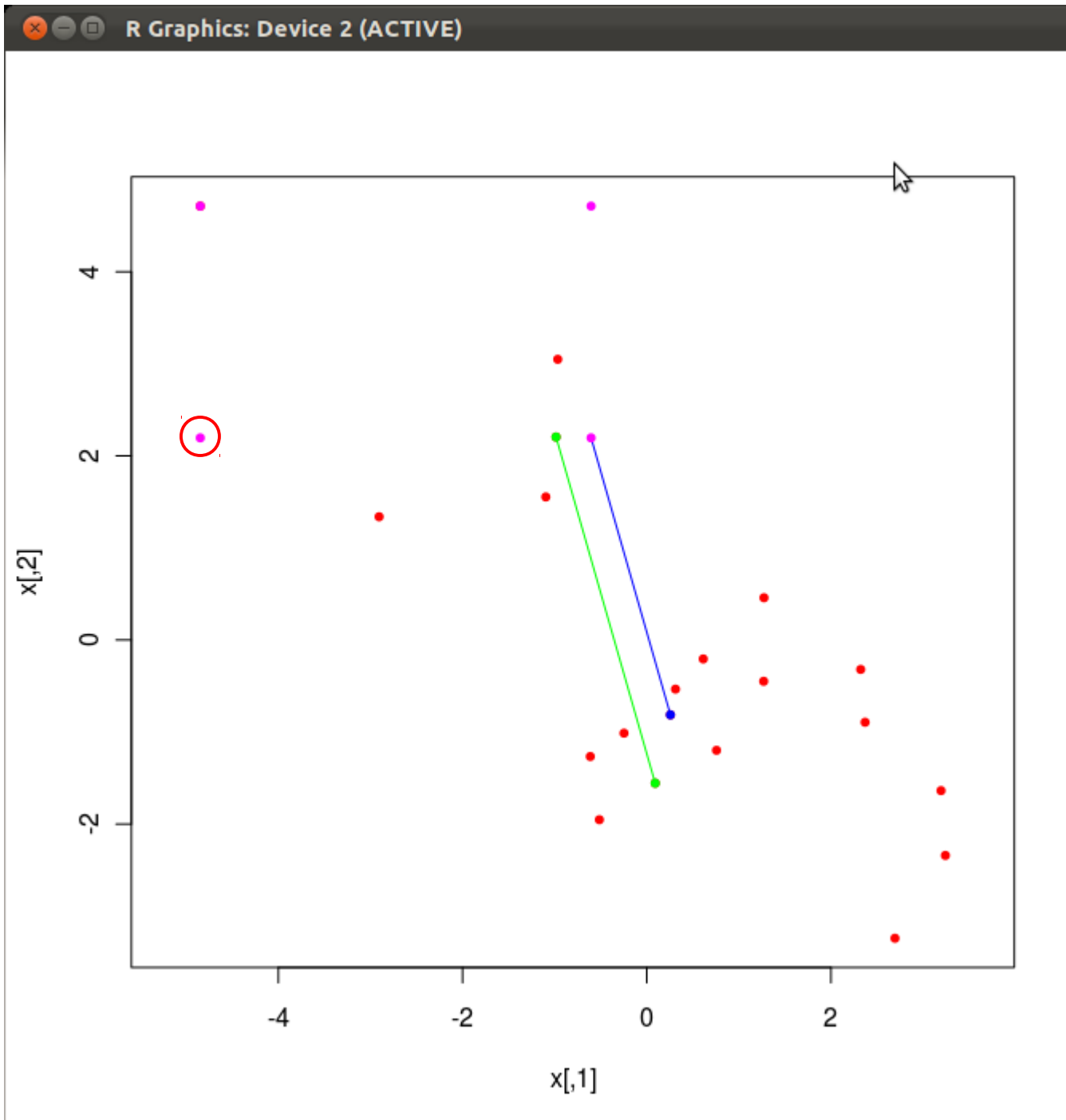
Strzałki między punktami S_x oraz S_y oznaczają, że punkt S_y jest lokalną modyfikacją punktu S_x

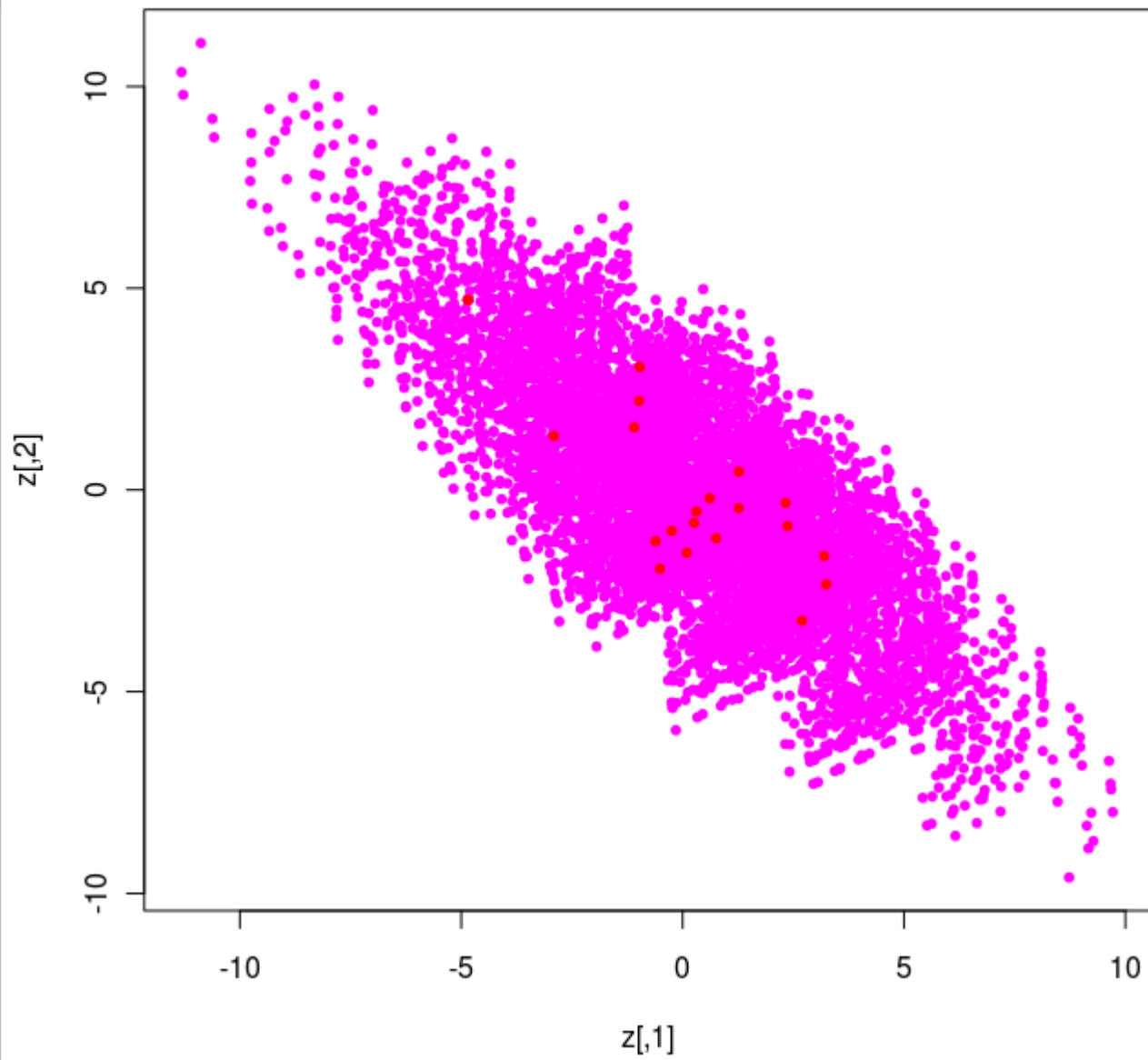


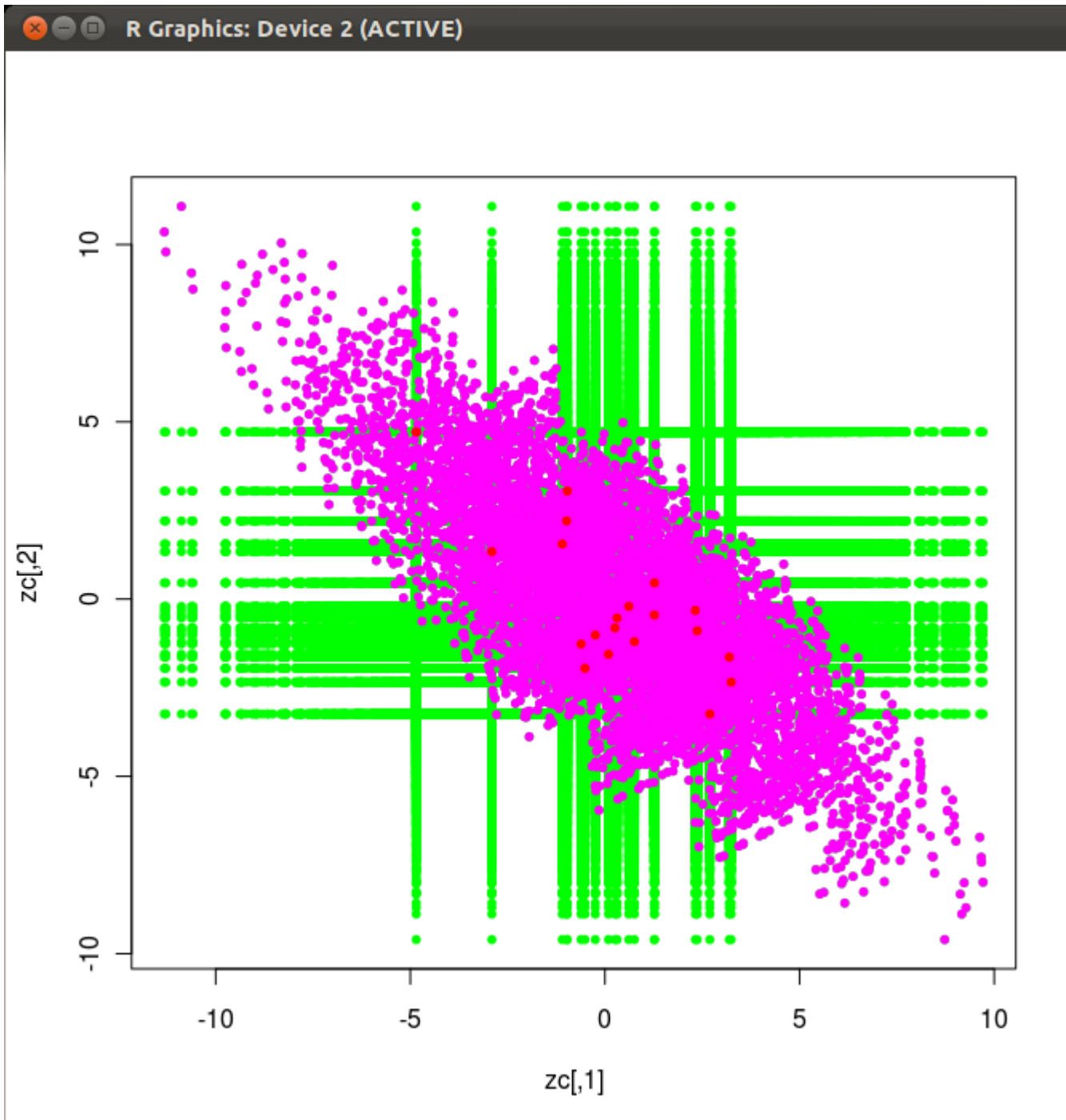












Ewolucja różnicowa

- Inne metody selekcji

- current-to-best $KP_i^t + (1 - K)P_{best}^t$

- current-to-rand $KP_i^t + (1 - K)P_j^t$

- rand-to-best $KP_{best}^t + (1 - K)P_j^t$

- Krzyżowanie uśredniające $z = KP_i^t + (1 - K)v$

- DE/either-or

$$z = \begin{cases} P_i^t + F(P_j^t - P_k^t) & \text{z } p\text{-stwem } p_F \\ KP_i^t + (1 - K)(P_j^t + P_k^t) & \text{z } p\text{-stwem } 1 - p_F \end{cases}$$

Ewolucja różnicowa

- Poinformowanie T
- Determinizm N
- Typ modelu pamięciowy
- Wielkość modelu licznosc populacji
- Lokalność generacji T
- Miętkość selekcji N
- Okno historii nieskończone
- Zupełność N

Algorytm ewolucyjny wypukła funkcja celu

- Model populacji nieskończonej
- Dystrybuanta empiryczna punktów populacji (skokowa) → dystrybuanta rozkładu próbkowania (ciągła)

DE/rand/1

wypukła funkcja celu

- Wariancja punktów po selekcji v_P
- Wariancja punktów po mutacji

$$v_O = v_P + F^2(v_P + v_P) = v_P(1 + 2F^2)$$

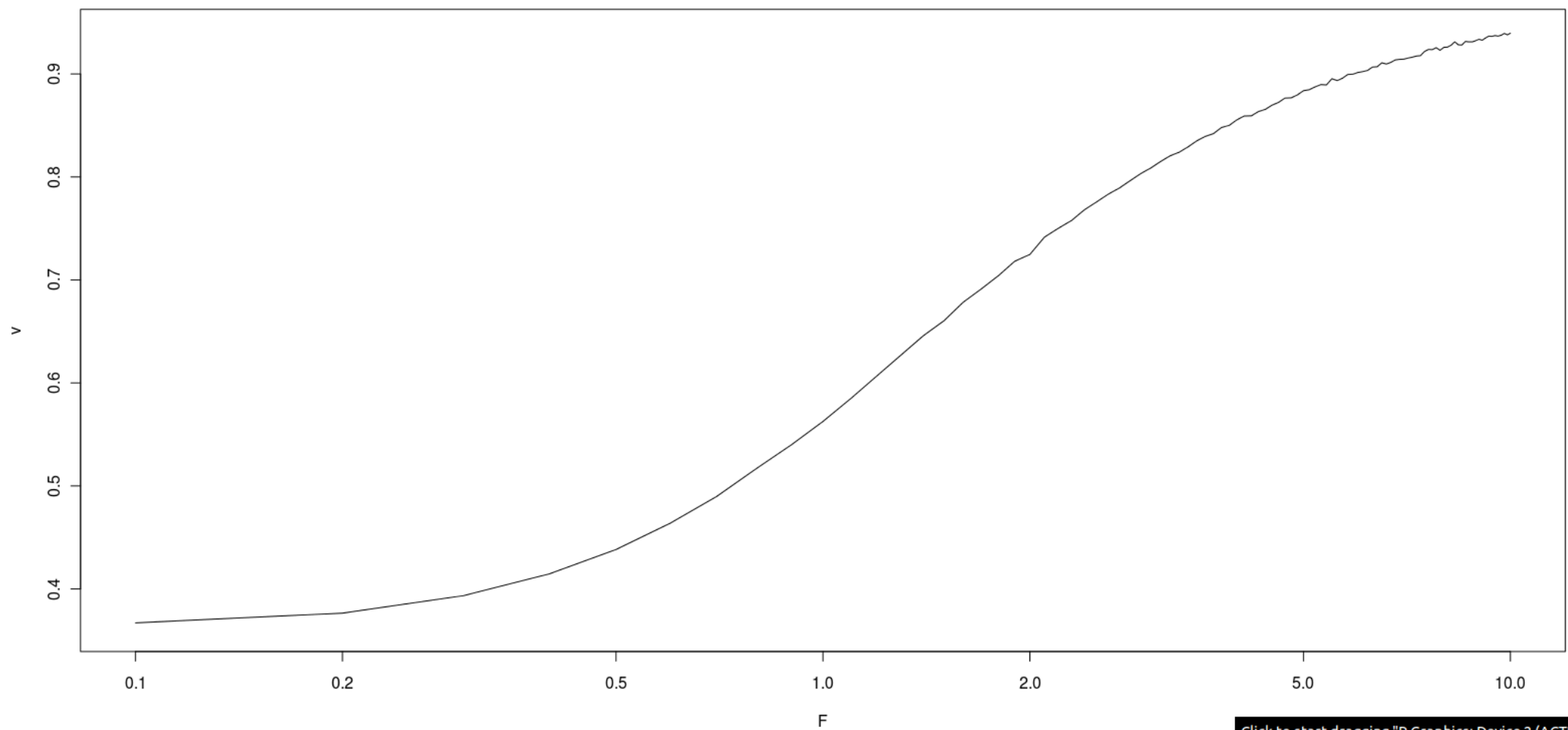
- Krzyżowanie zmienia wariancję (wzór dla bin)

$$v_C = (1 + 2F^2(1 - c_r))v_P$$

-

DE/rand/1

wypukła funkcja celu -
wariancja po sukcesji



DE/rand/1

wypukła funkcja celu

- Wariancja punktów po sukcesji

$$v_P(t+1) = k(F) v_P(t) \quad 0 < k < 1$$

- Równowagowa wariancja populacji:

$$v_P(\infty) = 0$$

- A dla alg. ewolucyjnego
(np. selekcja turniejowa, $s=2$, $pc=0$)

$$v_P(\infty) = \frac{\pi}{2} v_m$$

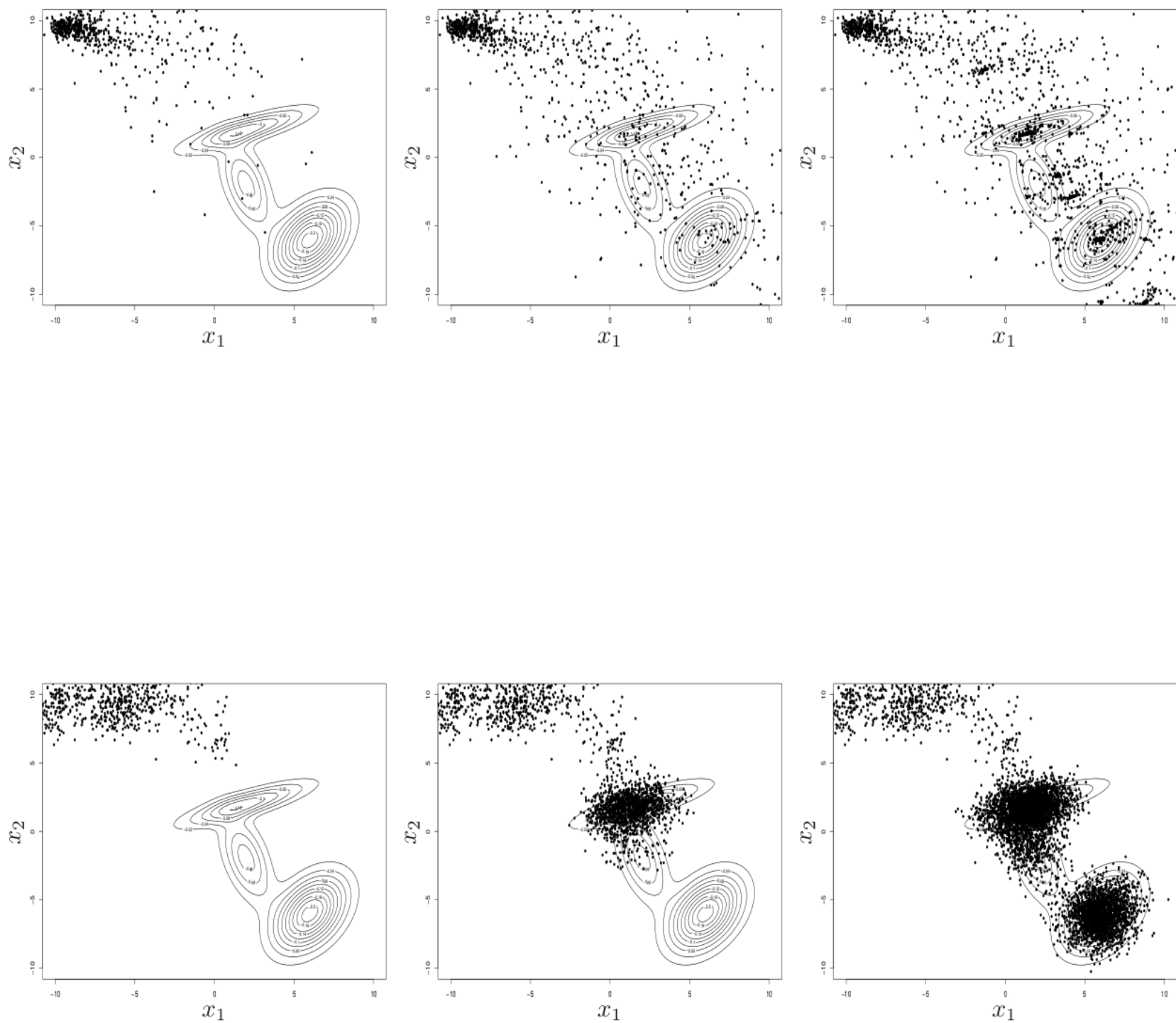
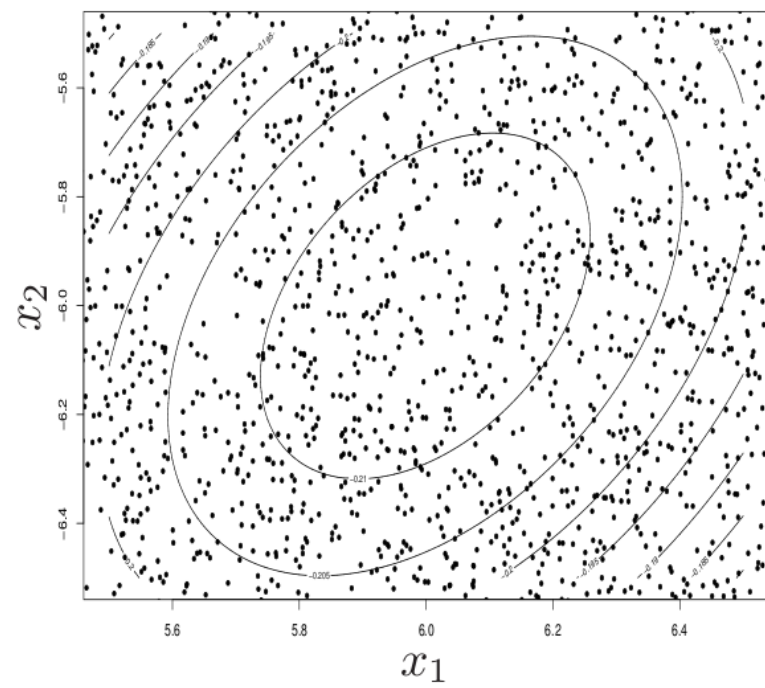
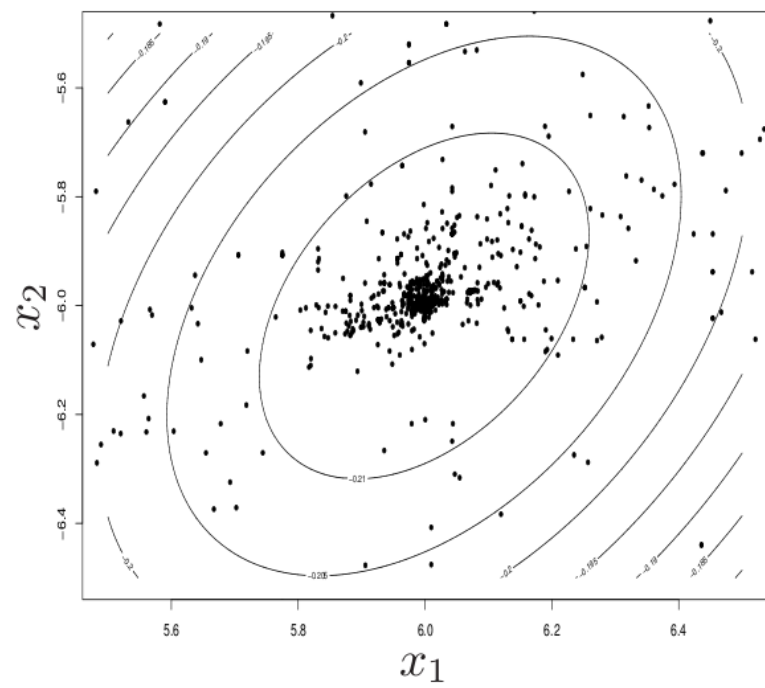


Fig. 6 Set of chromosomes generated by the GMEA algorithm with $\mu = 16$ points in population and mutation standard deviation equal to 0.6 : consecutive



Algorytm ewolucyjny

- Algorytm ewolucyjny jest techniką **adaptacji rozkładu populacji**
- Celem jest maksymalizacja wartości oczekiwanej jakości generowanych punktów
- Środek populacji – najlepszy estymator ekstremum lokalnego dla funkcji symetrycznej
- *Survival of the fittest vs. survival of the flattest*

Metoda EDA

Estimation of Distribution Algorithm

algorithm **EDA**

initialize($\mathbf{m}^0, \mathbf{C}^0$)

$H \leftarrow \emptyset$

$t \leftarrow 0$

while ! stop

$P^t \leftarrow \text{sample } N(\mathbf{m}^t, \mathbf{C}^t)$

$H \leftarrow H \cup P^t$

$O^t \leftarrow \text{select}(P^t)$

$(\mathbf{m}^{t+1}, \mathbf{C}^{t+1}) \leftarrow \text{update}(O^t, \mathbf{m}^t, \mathbf{C}^t)$

$t \leftarrow t + 1$

Wariant z rozkładem normalnym

Metoda EDA

Estimation of Distribution Algorithm

- UMDA (Univariate Marginal Distribution)
- Wartość oczekiwana i wariancja estymowana z próby jako

$$\mathbf{m}(t+1)_j \leftarrow \sum_{i=1}^{\mu} w(i) P_{ij}^t$$

$$\mathbf{C}(t+1)_{jj} \leftarrow \sum_{i=1}^{\mu} w(i) (P_{ij}^t - \mathbf{m}(t+1)_j)^2$$

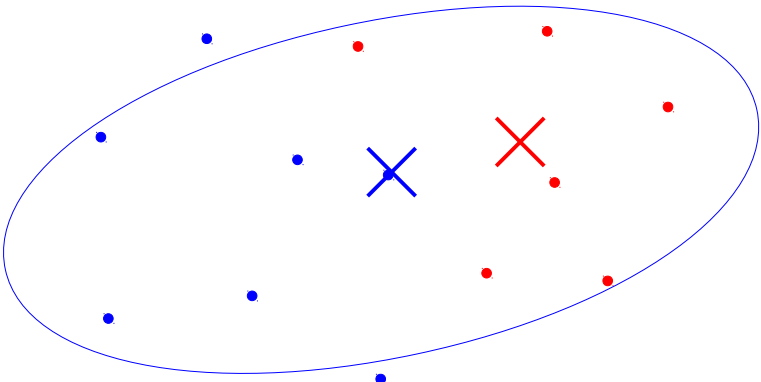
$$\mathbf{C}(t+1)_{ij} = 0 \quad i \neq j$$

$$w(i) = \frac{q(P_i^t)}{\sum q(P_i^t)}$$

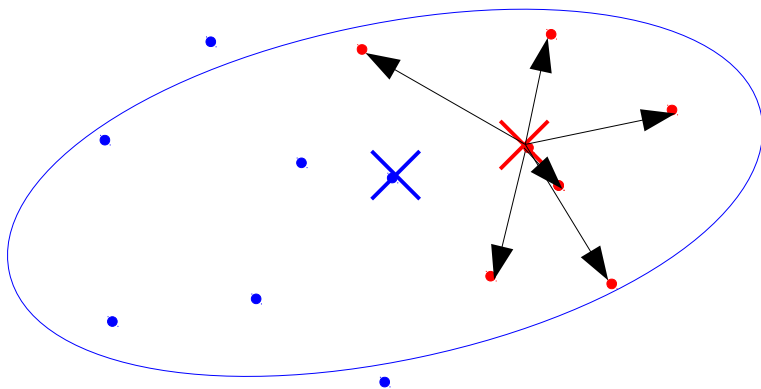
Metoda EDA

- Poinformowanie T
- Determinizm N
- Typ modelu zagregowany
- Wielkość modelu 1
- Lokalność generacji N
- Miętkość selekcji raczej N (selekcja?)
- Pamiętliwość tak, horyzont zależny od metody update
- Zupełność asymptotyczna

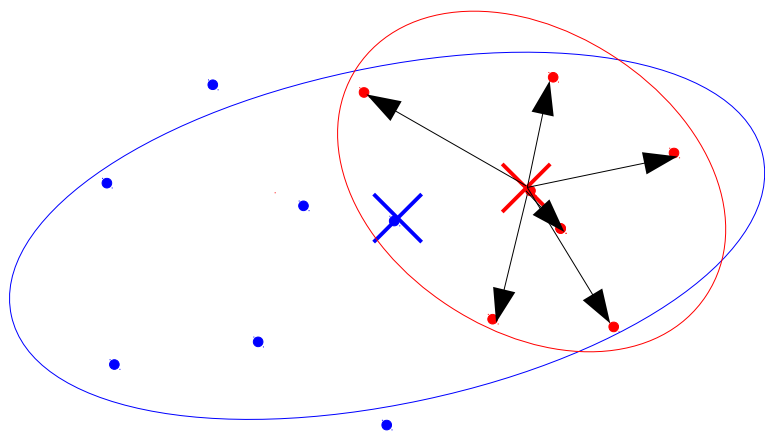
EDA



EDA

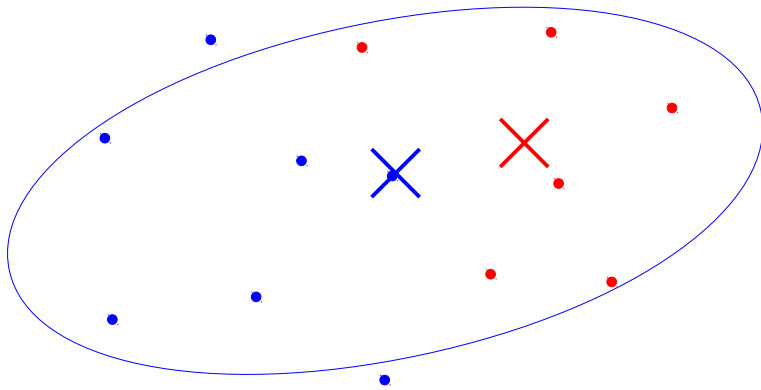


EDA



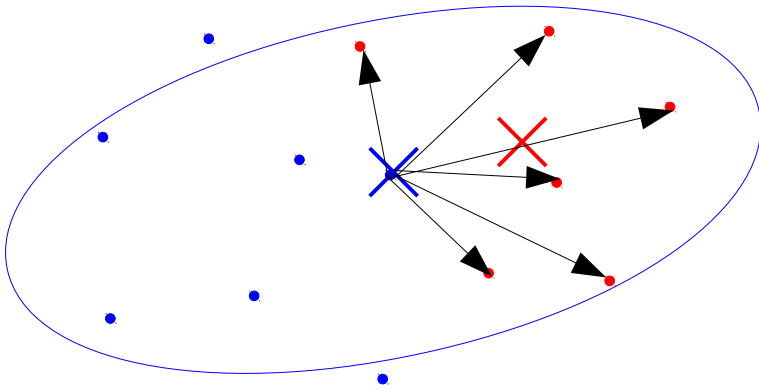
CMAES

Covariance Matrix Adaptation Evolution Strategy



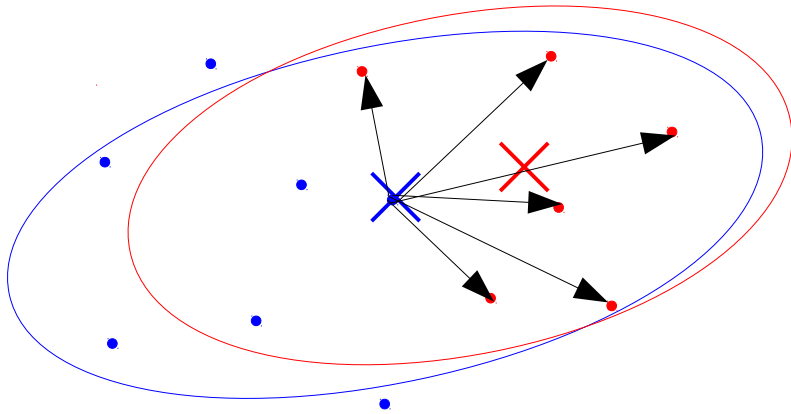
CMAES

Covariance Matrix Adaptation Evolution Strategy



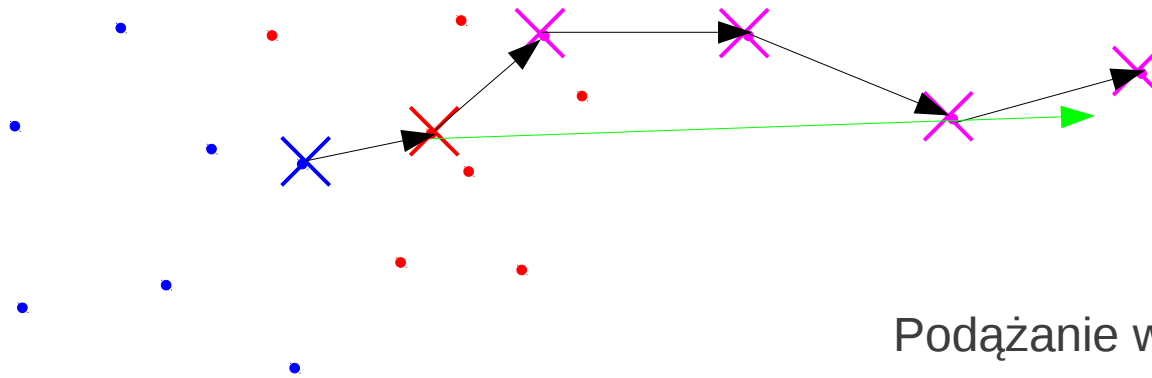
CMAES

Covariance Matrix Adaptation Evolution Strategy

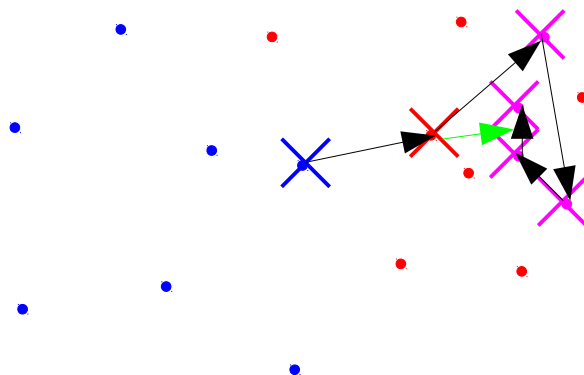


CMAES

evolution path



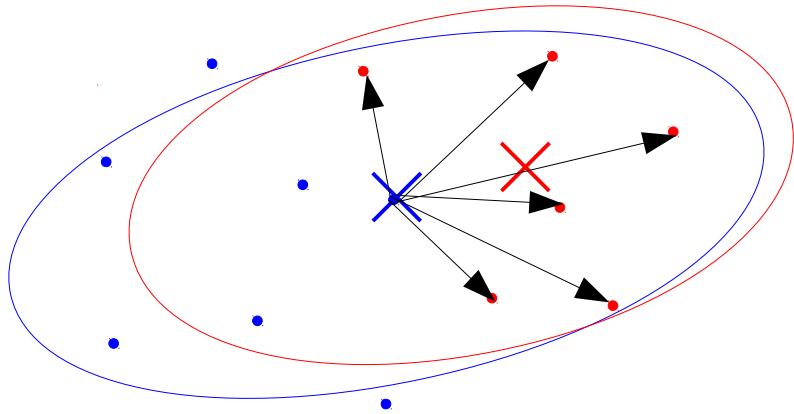
Podążanie w +/- zgodnych kierunkach



Populacja fluktuuje w jednym obszarze

CMAES

Covariance Matrix Adaptation Evolution Strategy



Na podstawie selekcji adaptuje się
kształt macierzy kowariancji

Jej skala zależy od ścieżki ewolucji

$$+ \text{[Diagram of a population with a new mean and search space]} = \text{CMAES}$$

A diagram illustrating the combination of a population (blue dots) and a new mean (blue 'X') to form the CMAES algorithm. The diagram shows a population of points (blue dots) and a new mean (blue 'X'). A red ellipse represents the current search space. A new mean (red 'X') is shown, and several black arrows indicate the movement of points towards the new mean, illustrating the adaptation of the search space.

Przeszukiwanie rojem cząstek

algorytm *particle swarm*

inicjuj $P^0 \leftarrow \{P_1^0, P_2^0 \dots P_\mu^0\}$

inicjuj $V^0 \leftarrow \{V_1^0, V_2^0 \dots V_\mu^0\}$

$H \leftarrow P^0$

$t \leftarrow 0$

while ! stop

$\mathbf{g}(t) \leftarrow \arg \max_{i,t} q(P_i^t)$

for ($i \in 1:\mu$)

$\mathbf{b}_i(t) \leftarrow \arg \max_t q(P_i^t)$

$V_i^{t+1} \leftarrow a(V_i^t + c(r_g(\mathbf{g}(t) - P_i^t) + r_l(\mathbf{b}_i(t) - P_i^t)))$

$P_i^{t+1} \leftarrow P_i^t + V_i^{t+1}$

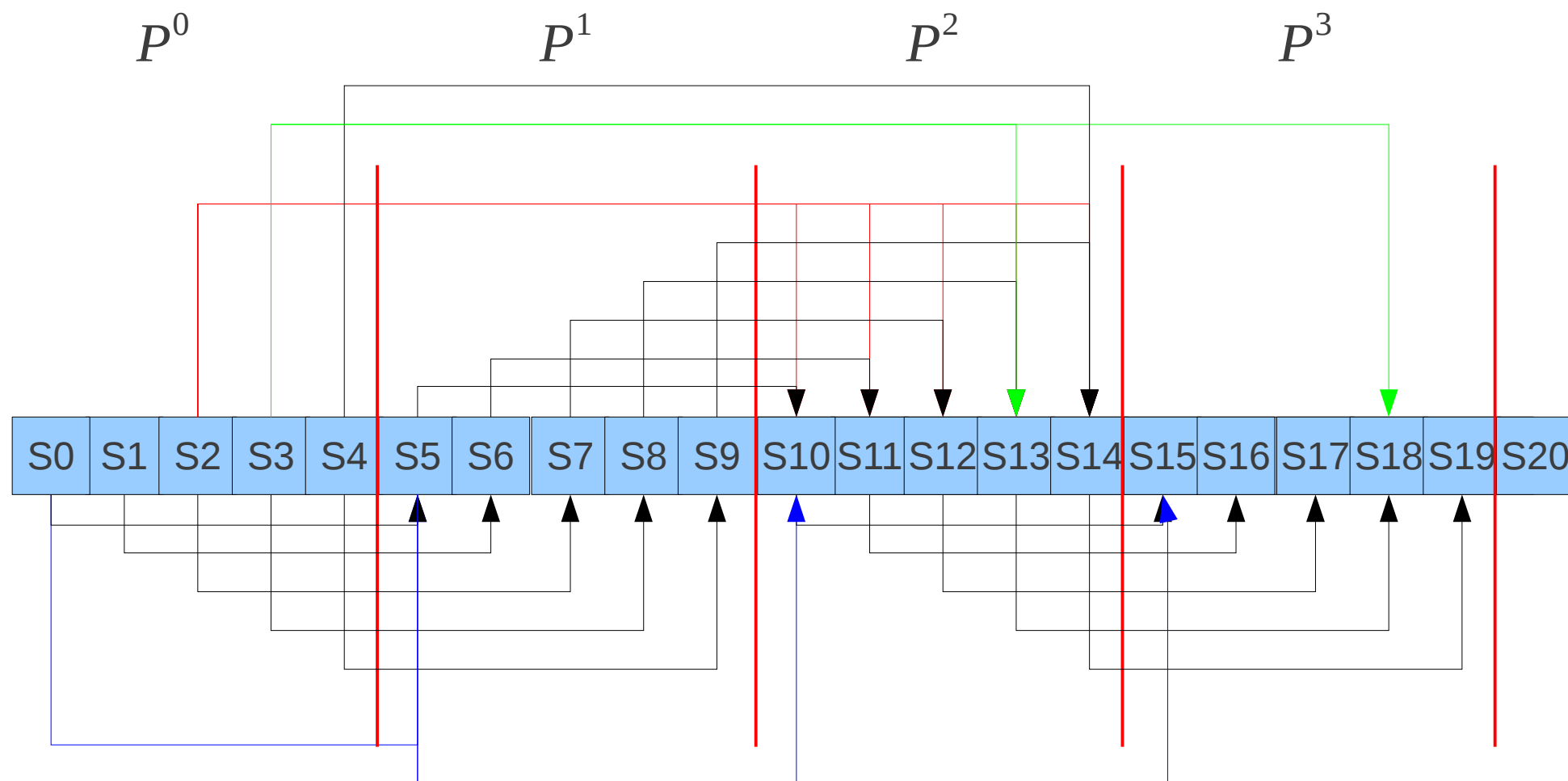
$H \leftarrow H \cup P^{t+1}$

$t \leftarrow t + 1$

a, c są parametrami
typowo a=0.73, c=2.05

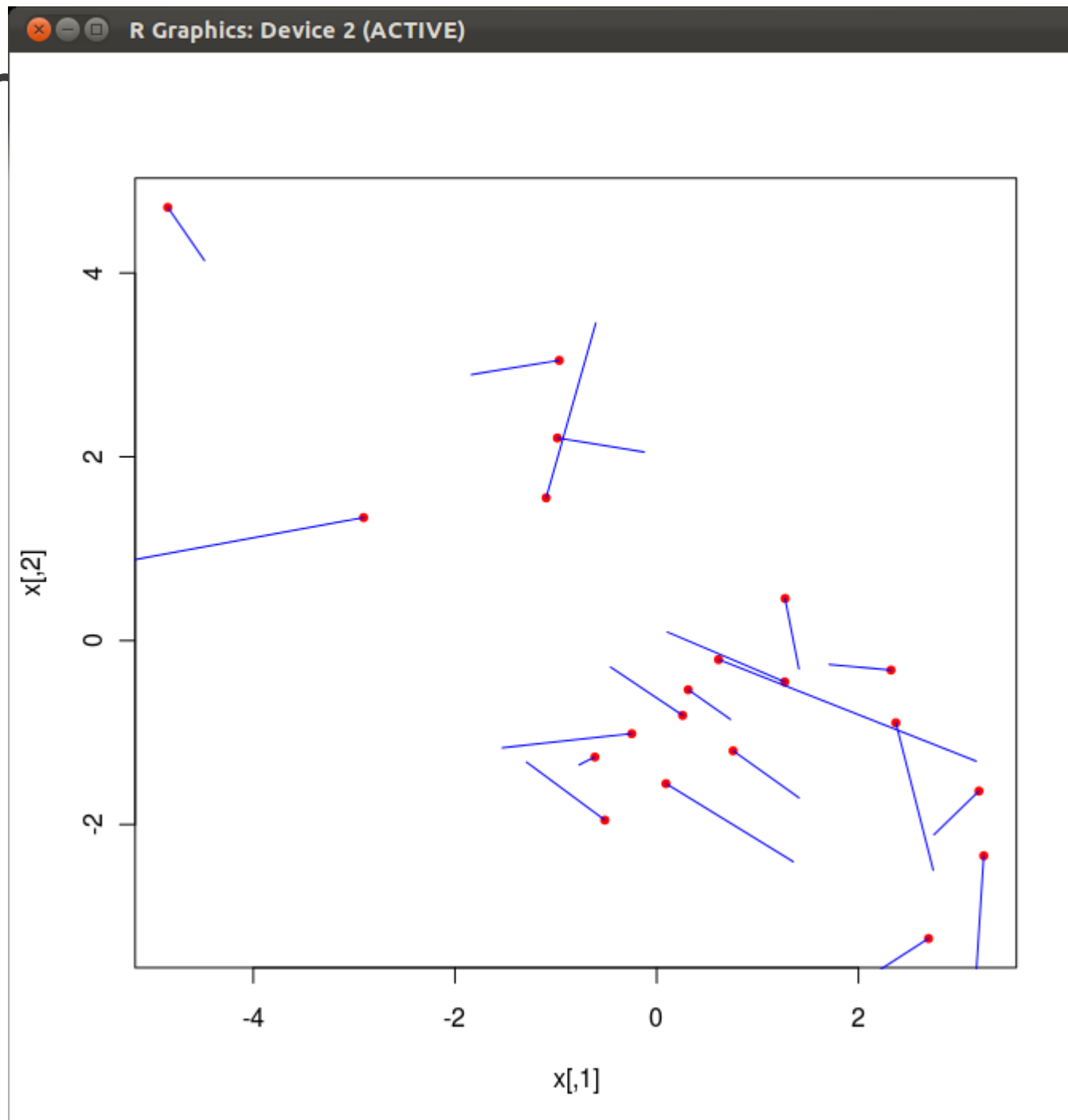
$r_g, r_l \sim U(0,1)$

Przeszukiwanie rojem cząstek



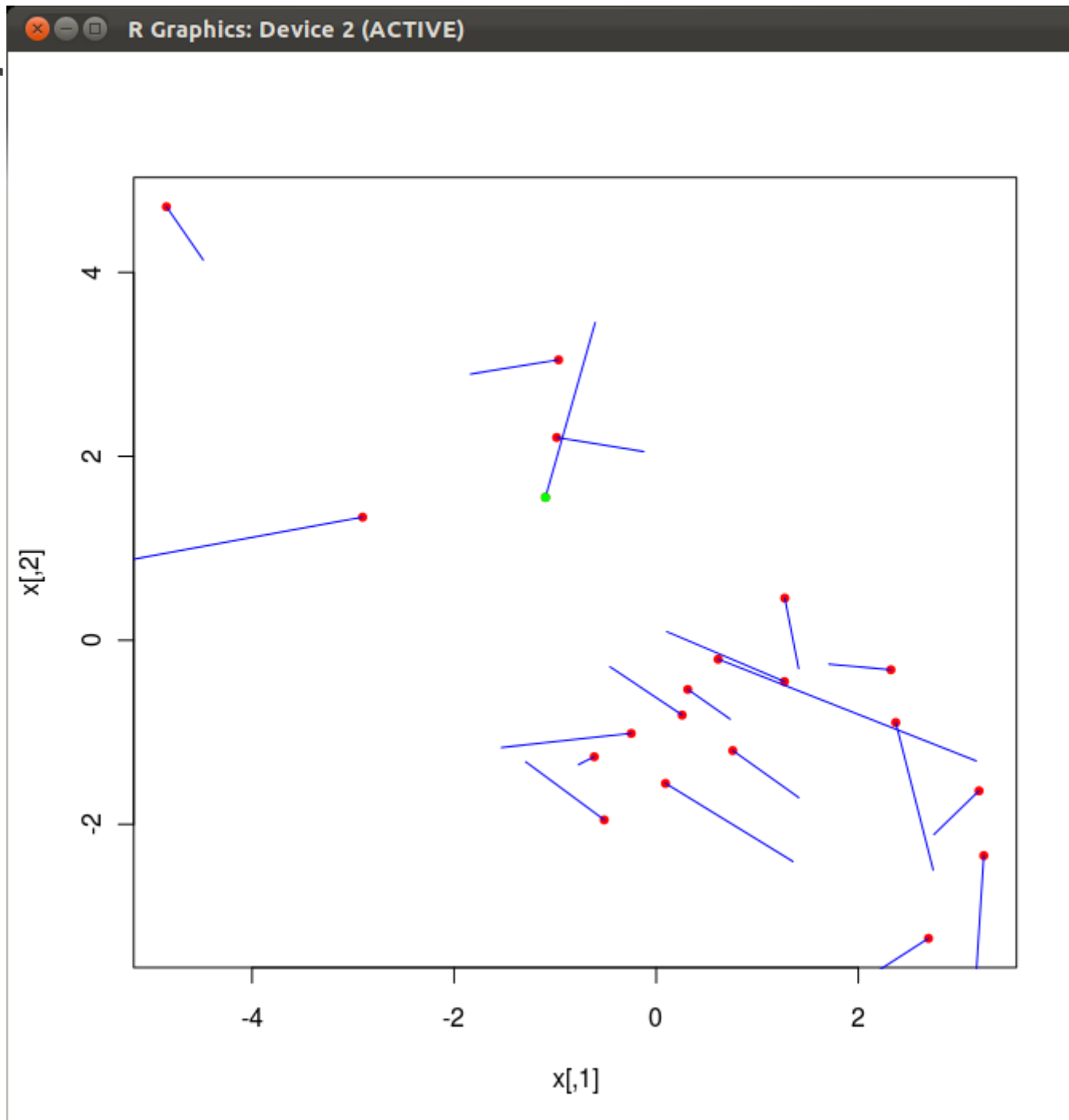
Pr

ek

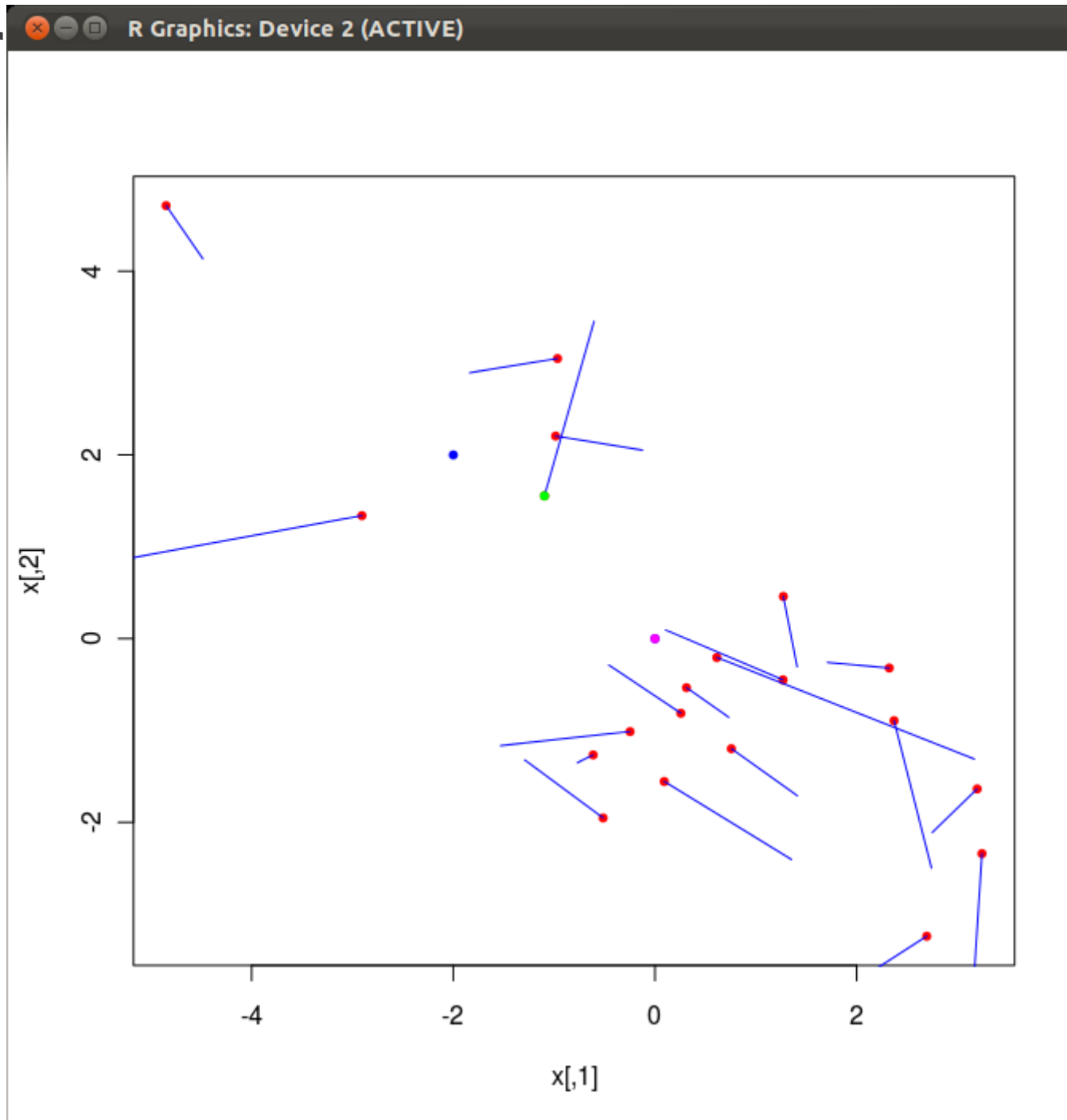


Pr

ek

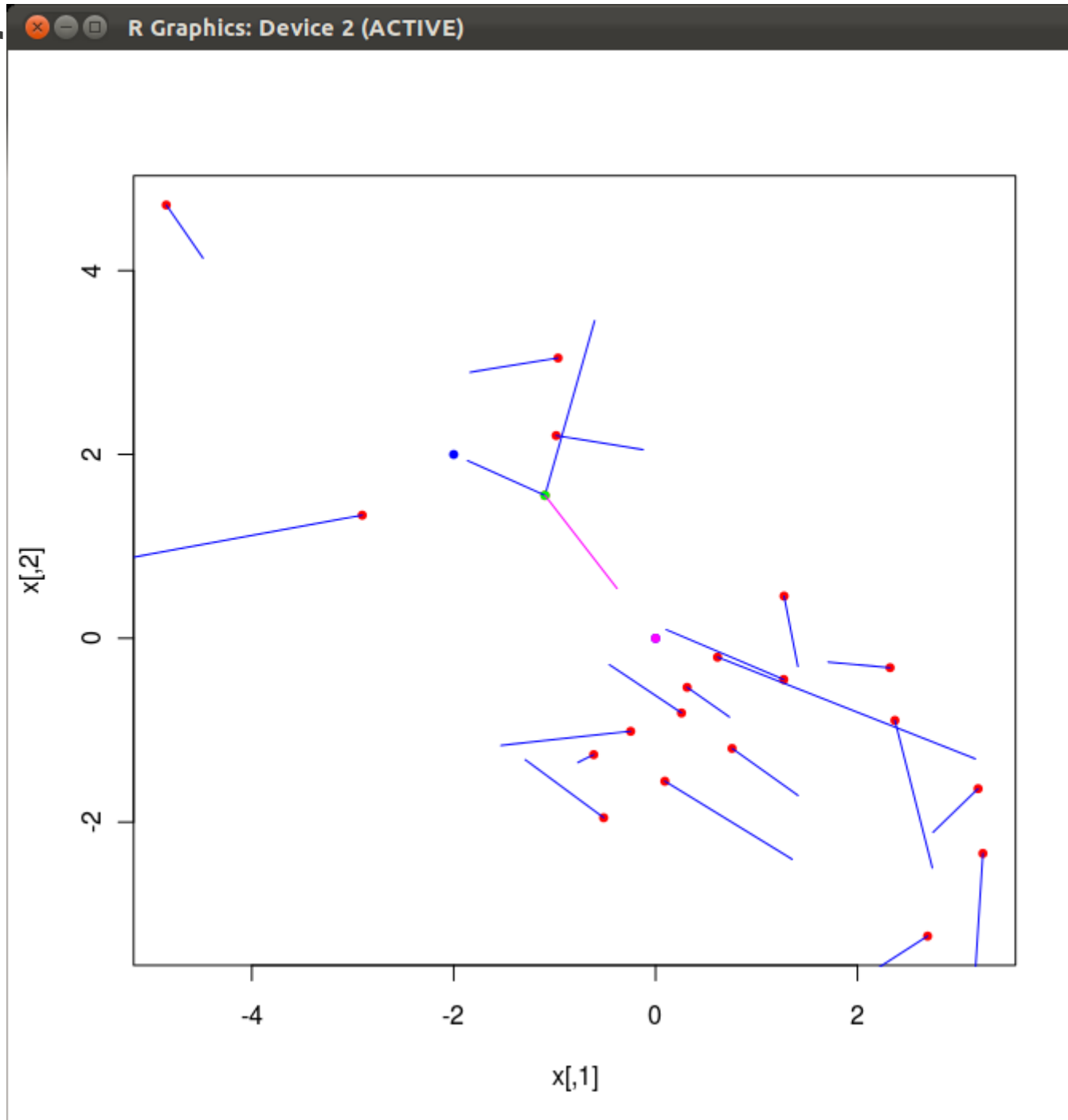


Pr



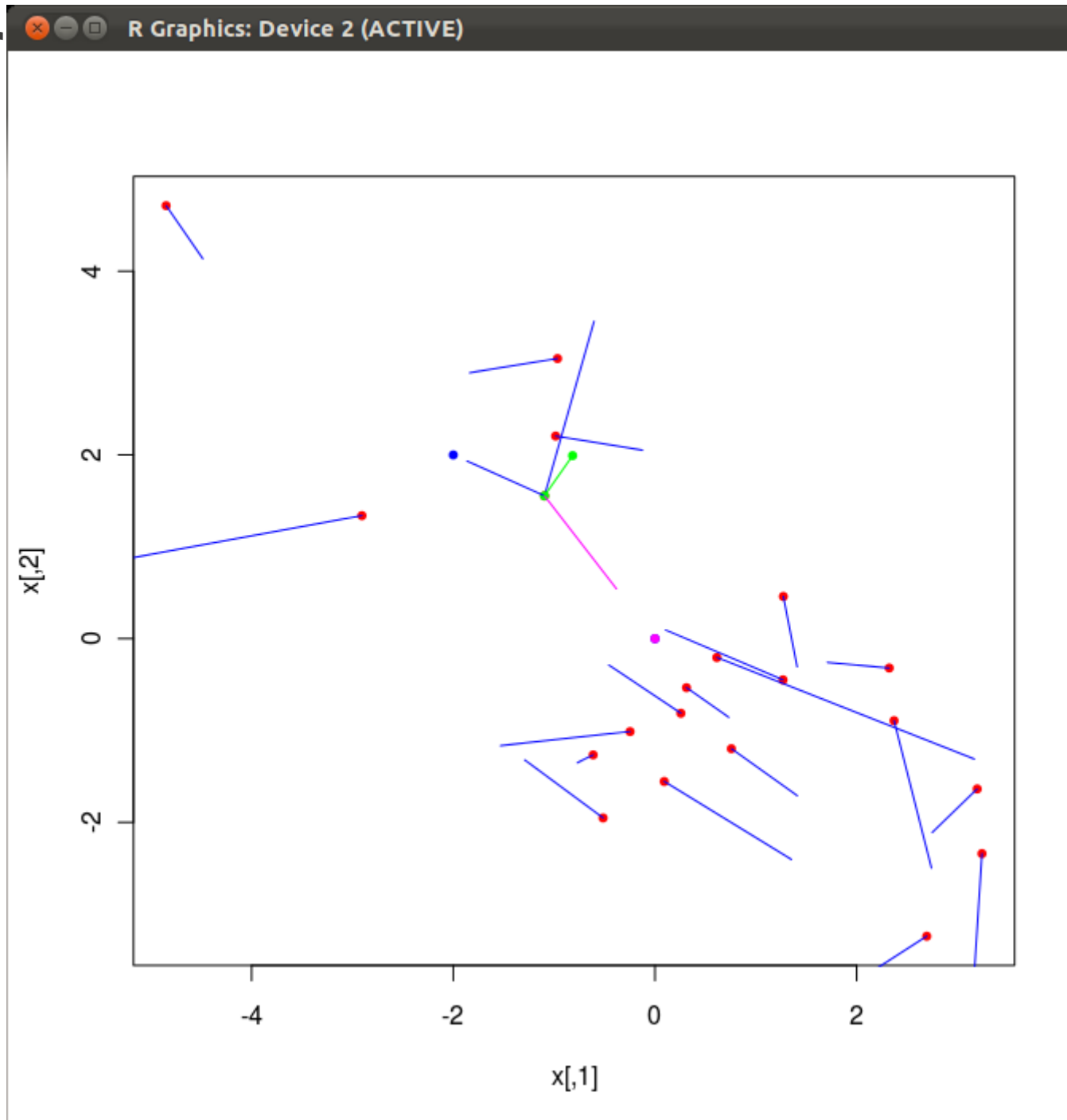
ek

Pr



ek

Pr



ek

Przeszukiwanie rojem cząstek

- Poinformowanie T
- Determinizm N
- Typ modelu zagregowany
- Wielkość modelu liczność populacji
- Lokalność generacji T
- Miętkość selekcji T
- Okno historii nieskończone
- Zupełność ?

Binarna ewolucja różnicowa

- Jak obliczyć różnicę między wektorami binarnymi?

$x_1 = 01000011110101$

$x_2 = 11001100110011$

$y = x_2 - x_1$?

$y = 10001111000110$

Binarny rój cząstek

- Jak wyrazić prędkość dla wektora binarnego?

$x=01000011110101$

$v=11001100110011$

$y=x+v$

$y=10001111000110$

Ewolucja różnicowa i rój cząstek w przestrzeni ścieżek w grafie

