Algorithmic collusion

1 unit of buyers value > 1

~~set~~ 2 sellers $\left[\frac{1}{2k}, \frac{2}{2k}, \dots, \frac{2k}{2k}\right]$

Bertrand competition, p_1 & p_2 two prices

$$u_1(p_1, p_2) = \begin{cases} 0 & \text{if } p_1 > p_2 \\ \frac{1}{2}p_1 & \text{if } p_1 = p_2 \\ p_1 & \text{if } p_1 < p_2 \end{cases}$$

Recall: Nash $\Rightarrow$ all prices $\frac{1}{2k}$ or $\frac{2}{2k}$

Steckelberg leader does much better
& so does follower
both get $\Omega(1)$

Question 1: outcome if both are no-regret learners

Can no-regret/swap learner use price 1

user i uses price 1 with prob c_i

value for player 1 when using 1

value — u — using $\frac{2k-1}{2k}$ instead

value when bidding 1:

$$(1-c_2)\cdot 0 + c_2 \cdot \frac{1}{2} p$$

p = max possibl price

if doing $\frac{2k-1}{2k}$ instead

$$\geq c_2 \boxed{\frac{2k-1}{2k}}$$

becomes $p - \frac{1}{2k}$

once 1 is not played

$\frac{2k-1}{2k}$ strictly dominated

Continue by induction...

stops at $\frac{2}{2k}$

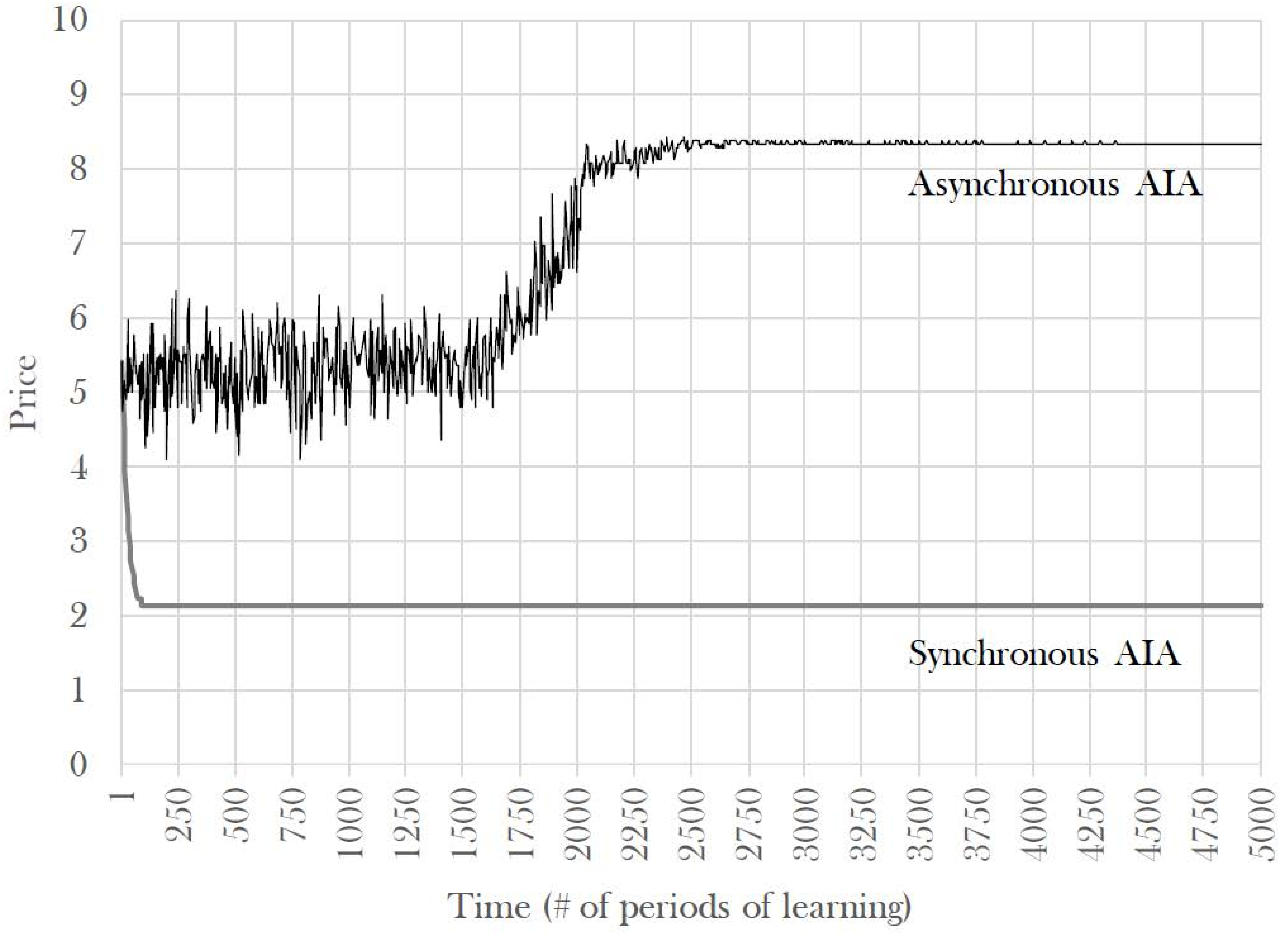
Example 2 prisoner's dilemma

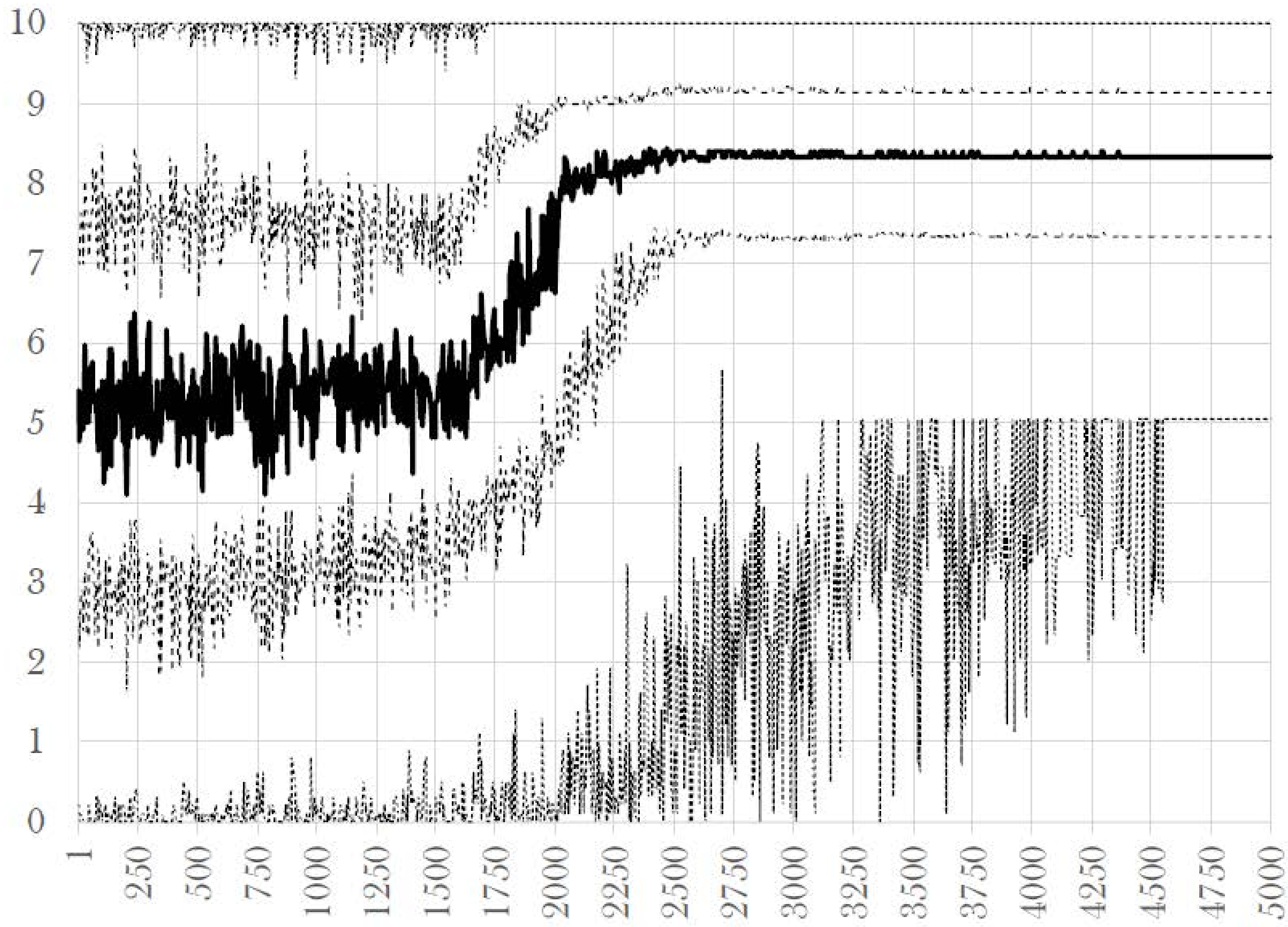
	C	D
C	5, 5	0, 6
D	6, 0	1, 1

$u_C = 0$ initial estimate

$u_D = 0$

Figure 1: Price paths with different algorithm designs





10
9
8
7
6
5
4
3
2
1
0
1
250
500
750
1000
1250
1500
1750
2000
2250
2500
2750
3000
3250
3500
3750
4000
4250
4500
4750
5000

first two coins (D,D) & (C,C)

$$u_C = 5, \quad u_D = 1$$

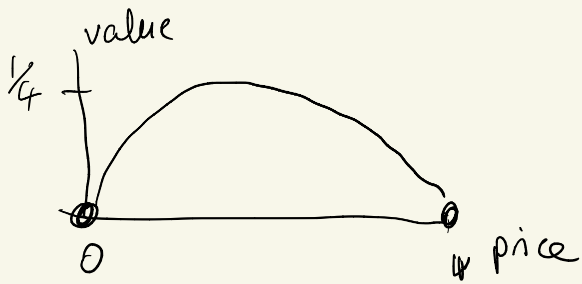
More general Q-learning update rule

$$u_C \leftarrow (1-\alpha)\, u_C + \alpha\, r_C$$

new value — old value — reward is current net

Uniform random price to start:

$$\text{value } p \longrightarrow p \cdot 0 + (1-p)p = p - p^2$$



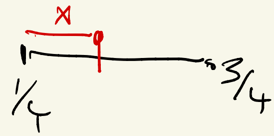
May stop playing $p \sim 0$ $p \sim 1$

Example:

uniform price $\left[\frac{1}{4}, \frac{3}{4}\right]$

value

$p < \frac{1}{4}$	p
$p > 3/4$	0



$\frac{1}{4} \leq p \leq 3/4 \longrightarrow$

$p = \frac{1}{4} + x$: $2x \cdot 0 + (1-2x)\left(\frac{1}{4} + x\right)$

$$= \frac{1 + 2x - 8x^2}{4}$$