S&M
Split and Merge Compression Algorithm

By
Abdullah Hashim

S&M algorithm: The Asymptotic Case

S&M algorithm: the asymptotic case
-The Practical Simulation-

Input: Action (ω)
$\Delta = \{\omega_1, \omega_2, \ldots, \omega_\Phi\}$
$A = \{\alpha_1, \alpha_2, \ldots, \alpha_N\}$
$F = (c_1, c_2, \ldots, c_{|F|})$

Environment:
$S = \{s_1, s_2, \ldots, s_{|S|}\}$
$D = \{w_1, w_2, \ldots, w_\Phi\}$
$P(S)$ and $P(D)$

Output:
(s_s, s_j & $s_k \in S$)

Discrete time interval: n

<u>Assumptions</u>: In the asymptotic case we assume that: $|s_i| >> |F|$, (i =1, …, $|S|$). Throughout the process, the value of $|s_i|$ remain greater than unity.

<u>Initial Conditions</u>: The initial $p[s_1(0)]$ is fixed to one of ten values in the range of $\{0.001 \leq p[s_1(0)] \leq 1\}$ and all other set's probabilities is made to be equal to:

$$\{1 - p[s_1(0)]\} / (1 - |S|).$$

<u>Results</u>: Since we assumed that, $|s_i| >> |F|$, then: ***D-Norms*** is not applicable. The behaviour of the algorithm is evaluated by ***S-Norms*** only, by plotting the average values of $p(s_1(n)$, $Q_S(n)$ and $H_S(n)$ over hundred **(100)** trials, for every one of the ten predetermined different set of initial probabilities. Where:

$$Q_S(n) = \sum_{i=1}^{|S|} p[s_i(n)]^2; \qquad H_S(n) = \sum_{i=1}^{|S|} p[s_i(n)] \log_2(1/p[s_i(n)])$$

S&M algorithm: the asymptotic case

-The Strategy-

Sets	s_1	s_2	s_3	s_4
Interval $n = 0$	P_1	P_2	P_3	P_4
Interval $n = 1$	$P_1/2$	P_2	$P_3 + P_4$	$P_1/2$
Interval $n = 2$	$P_1/2 + P_3 + P_4$	$P_2/2$	$P_2/2$	$P_1/2$
Interval $n = 3$	$P_1/2 + P_3 + P_4$	$P_2/2+p_1/2$	$P_2/4$	$P_2/4$

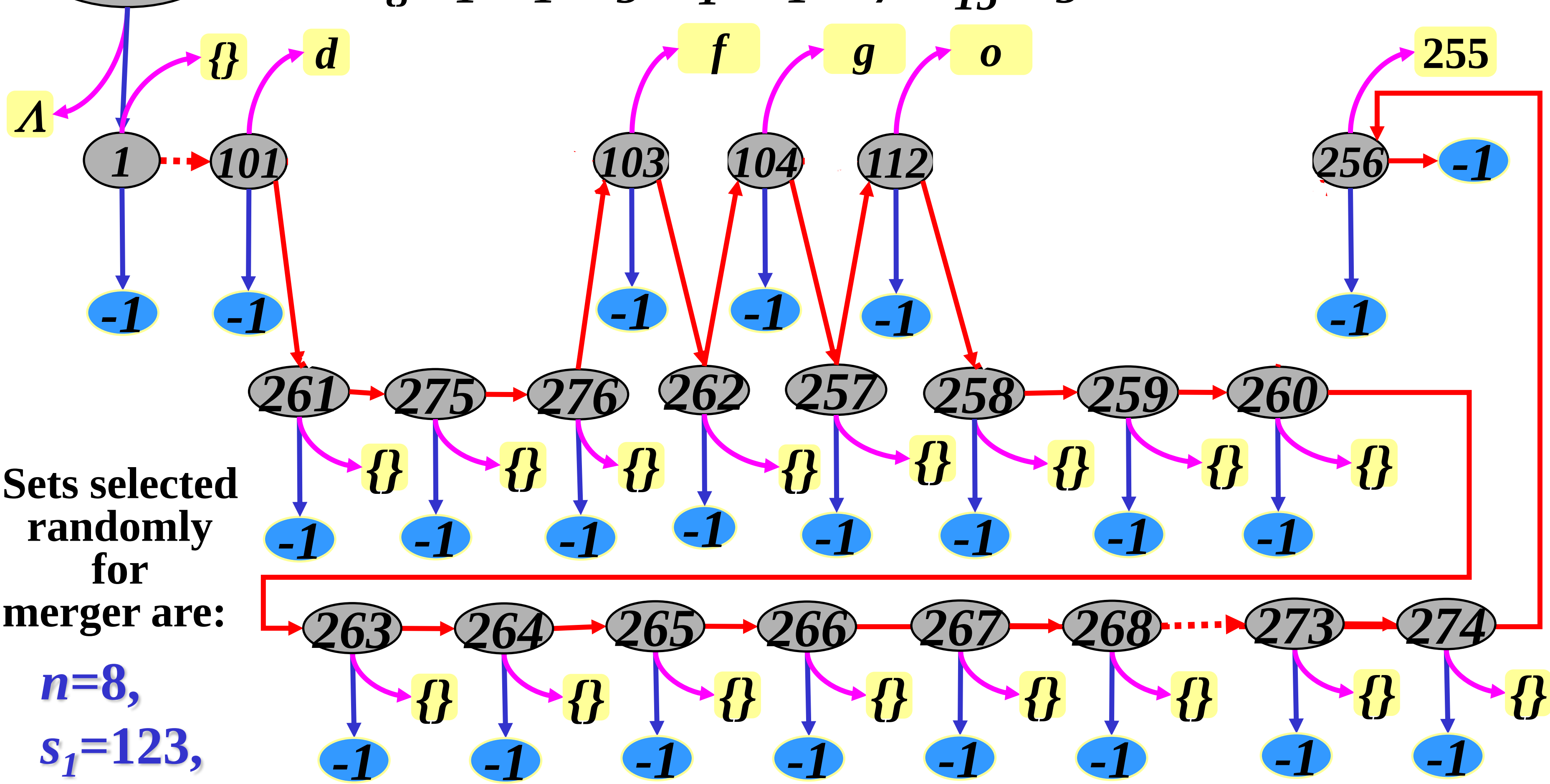
For large values of n, set probabilities will converge to $2/|S|$.

S&M algorithm: the finite case

input string <goodfood{}>

s_n	{}	g~~o~~	o~~o~~	o~~d~~	d~~f~~	f~~o~~	o~~d~~	o~~d~~	d~~{}~~	EOF
n	0	1	2	3	4	5	6	7	8	9
SBL	{}	1	1	3	1	1	7	15	3	

$$R = \frac{72}{65} = 1.1$$



Sets selected randomly for merger are:

$n=8$,
$s_1=123$,
$s_2=201$

s_9 = '{}' = 1 = 0 0 0 0 0 0 0 0

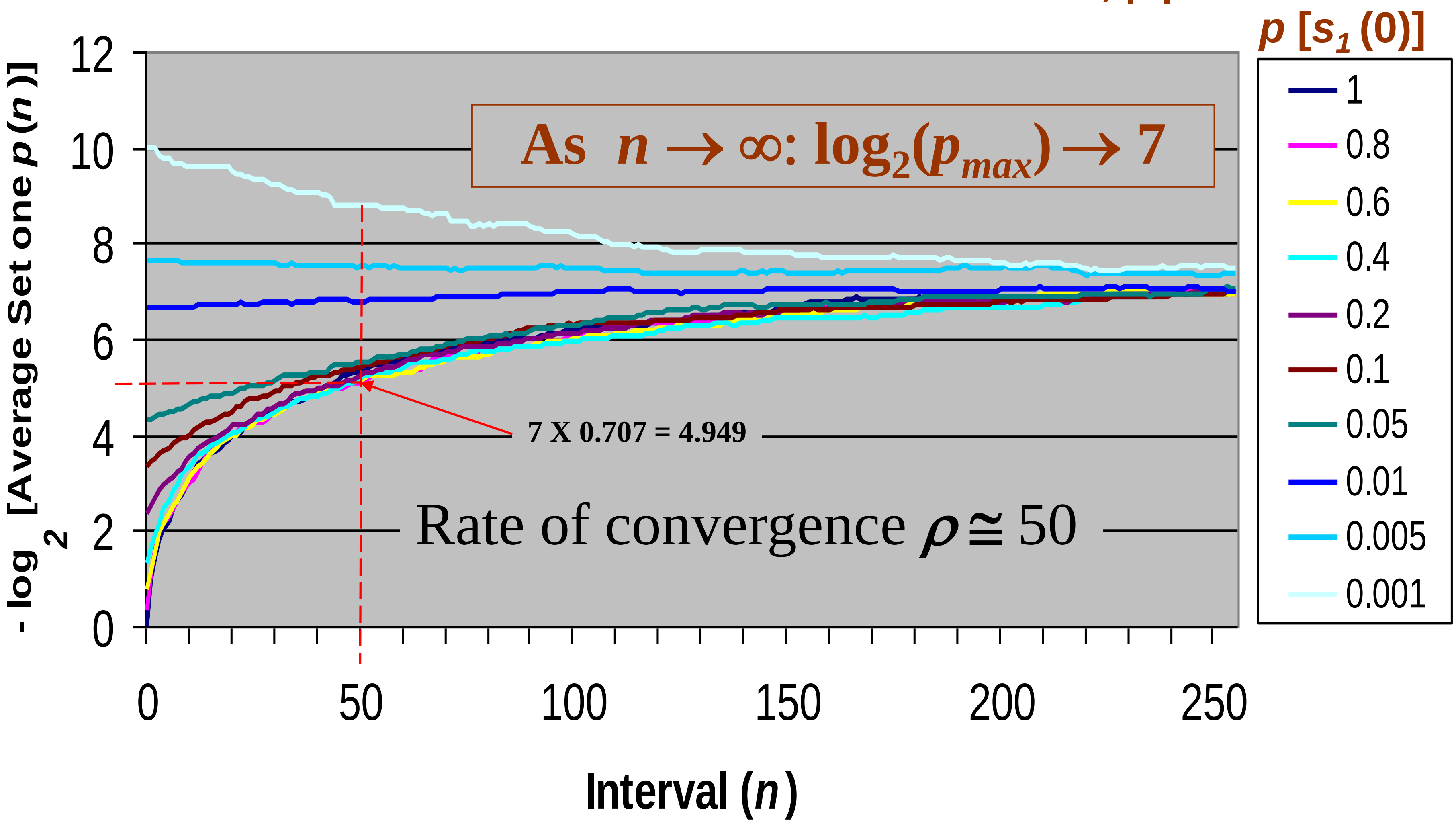
$L(w_{node9})$ = 8 bits

S&M algorithm: the asymptotic case

The S-Norms

- $\log_2$[Average Set One $p(n)$] for 256 intervals over 100 trials

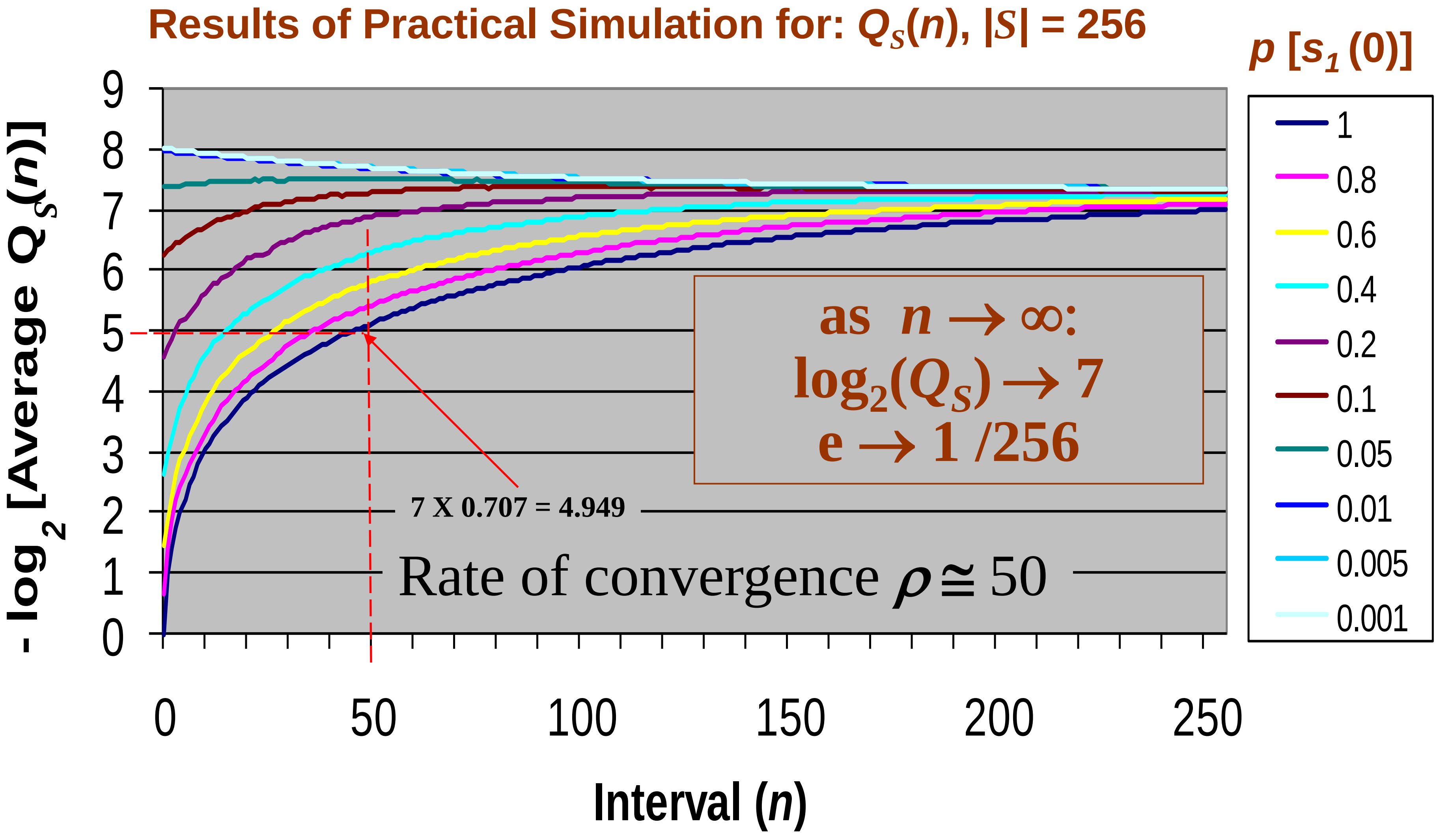
Results of Practical Simulation for: First Set Prob, $|S| = 256$



S&M algorithm: the asymptotic case

The S-Norms

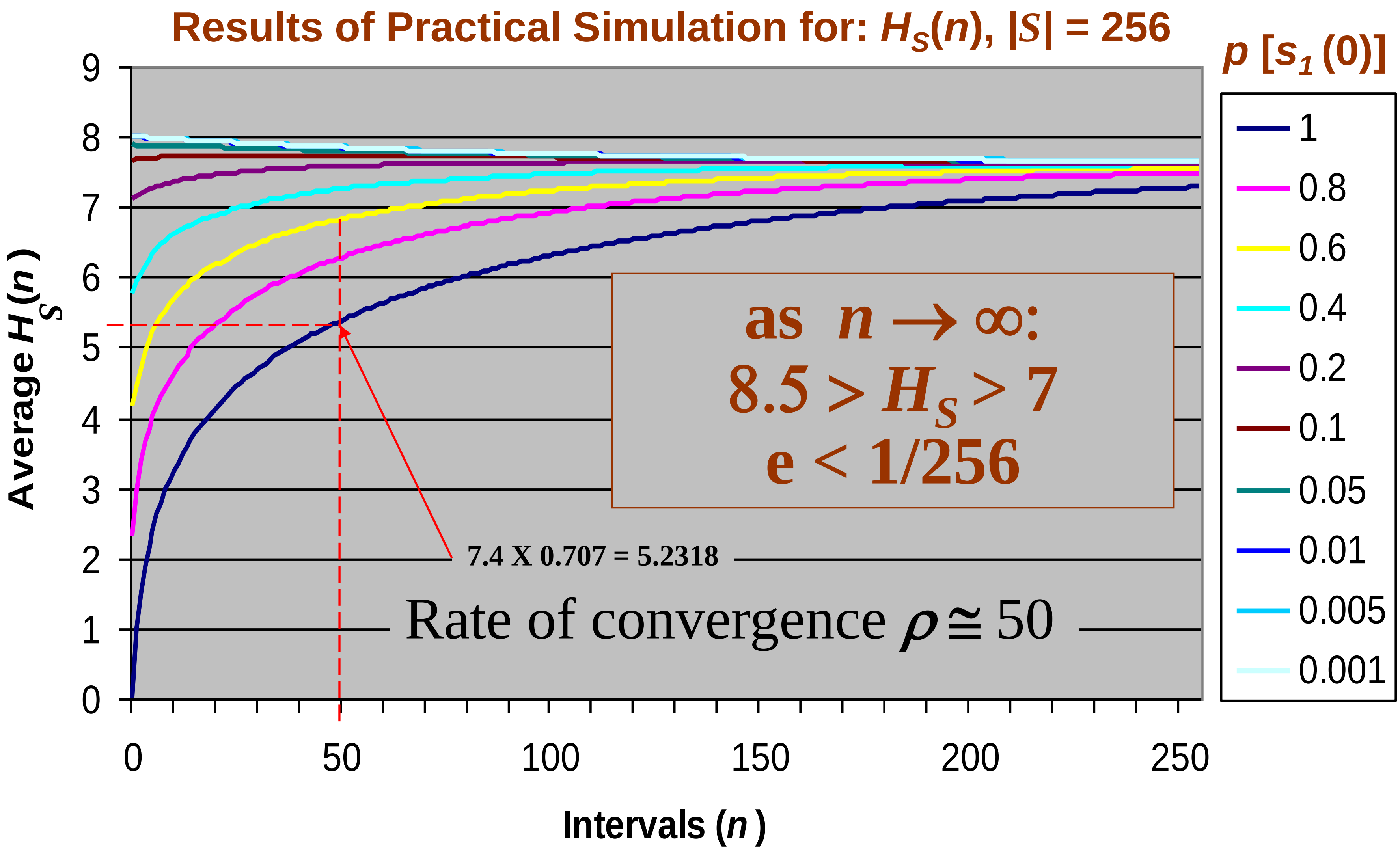
$-\log_2$ [Average $Q_S(n)$] for 256 intervals over 100 trials



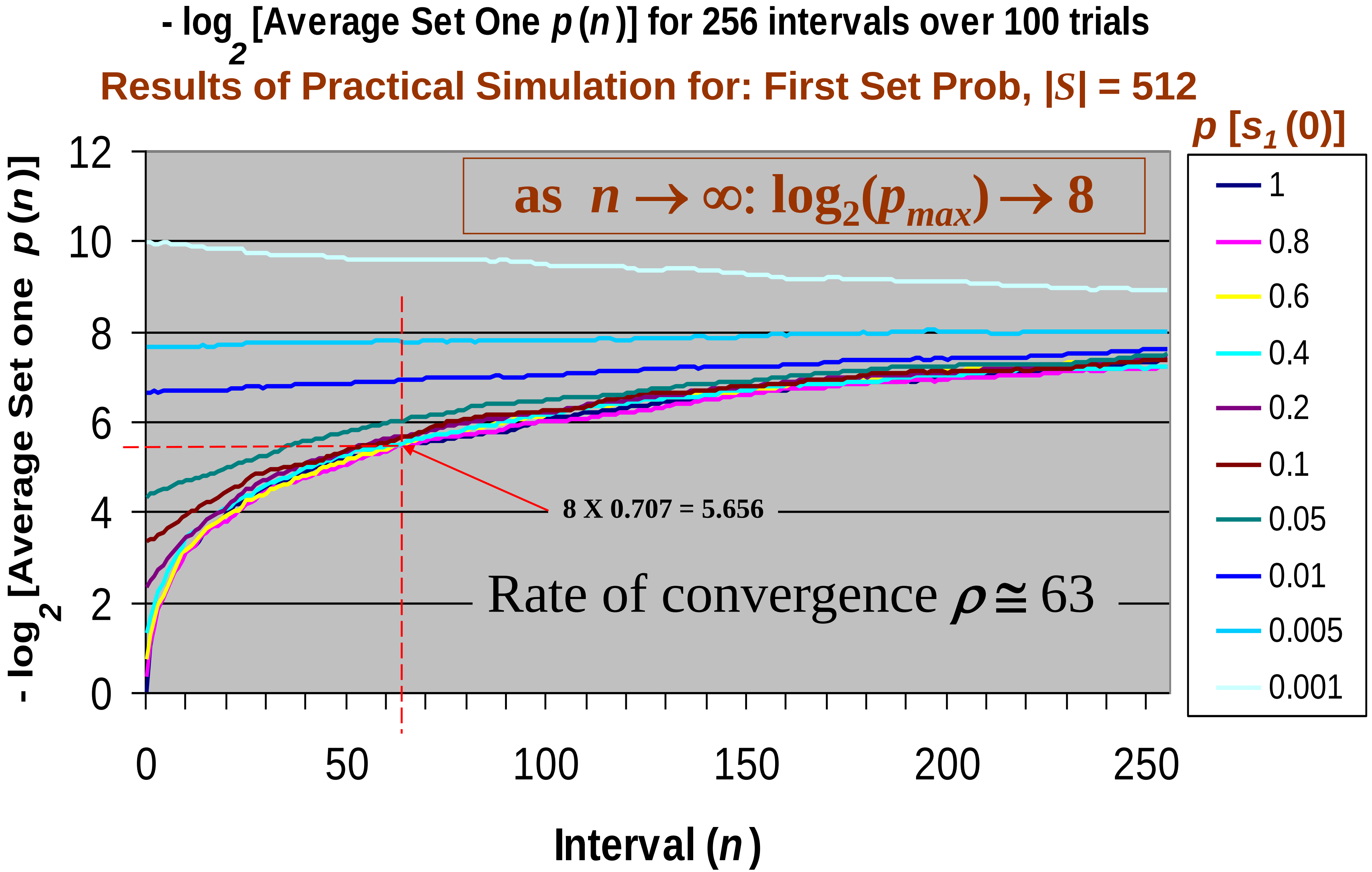
S&M algorithm: the asymptotic case

The S-Norms

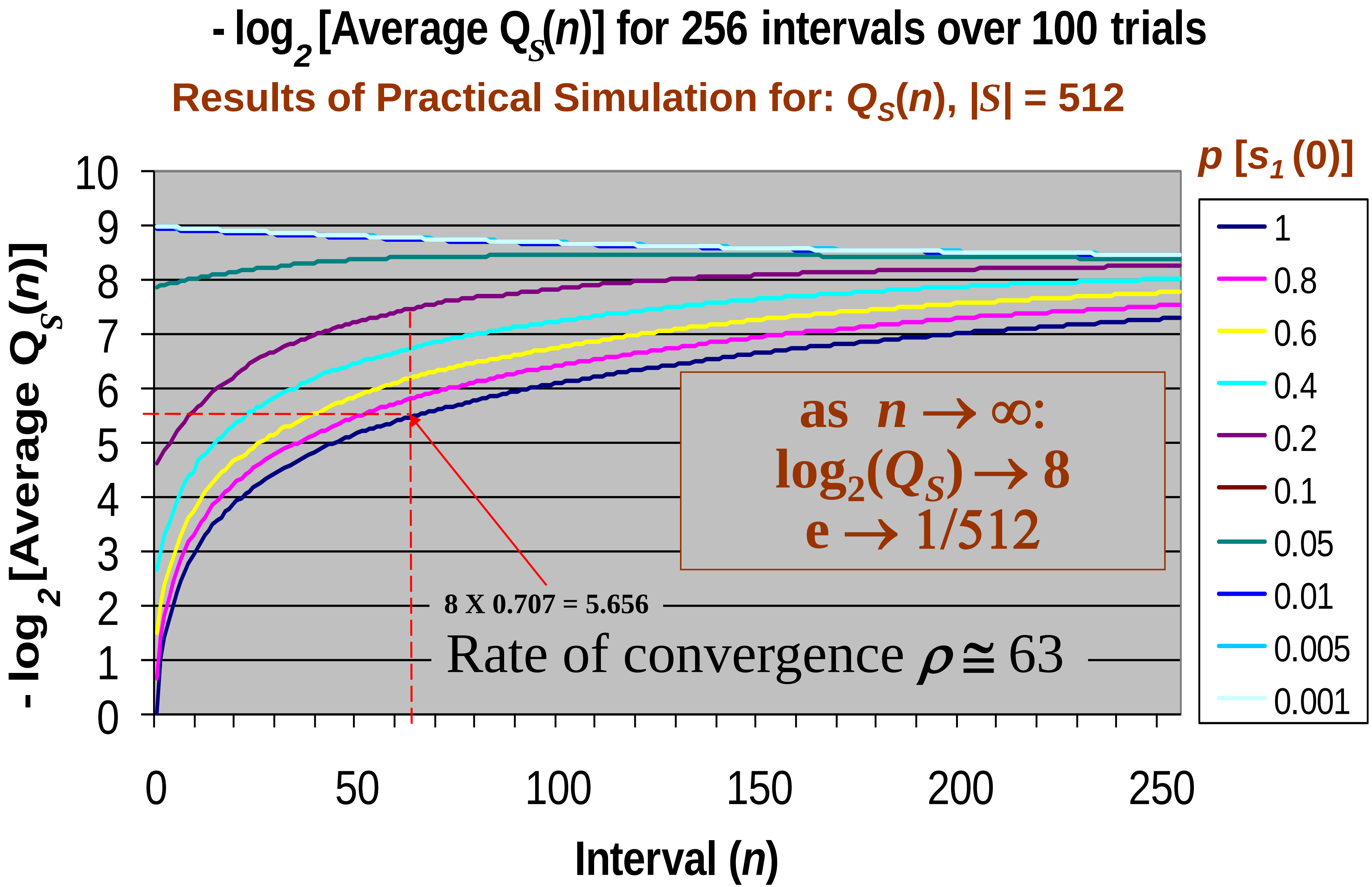
Average $H_S(n)$ for 256 intervals over 100 trials



S&M algorithm: the asymptotic case
The S-Norms



S&M algorithm: the asymptotic case
The S-Norms



S&M algorithm: the asymptotic case
The S-Norms

Average $H_S(n)$ for 256 intervals over 100 trials