Performance Evaluation of the MLSDA Decoders by Importance Sampling Techniques

Presenter: Jau-Yuan Weng

Advisor: Po-Ning Chen

Dept. of Communication Engineering

National Chiao Tung University

2001/6/14

1

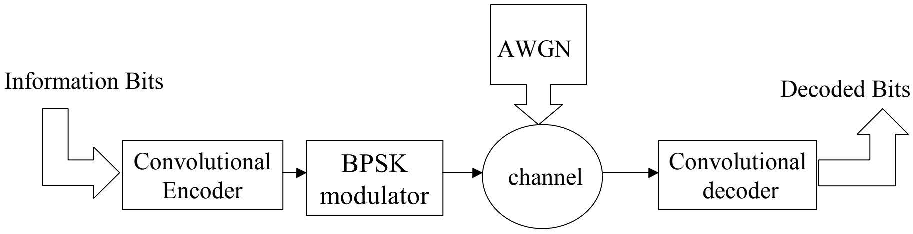
Outline

- Introduction
- A brief review of the MLSDA
- MC simulation method of the coded system
- Importance sampling for the simulation
- Simulation results
- Conclusions

2

Introduction

- A convolutional coded system with AWGN channel is considered.
- The convolutional code is a (2,1,6) code with generator 634,564 (octal). The length of the information bits is L=60.



3

Introduction

- Viterbi decoding algorithm and conventional sequential decoding algorithm vs. MLSDA
  - ➢ In spite of the good performance on the error correcting ability, *Viterbi decoding algorithm* becomes impractical when the *constraint length* grows.
  - ➢ Complexity of conventional *sequential decoding algorithm* is independent of the constraint length, but it does not always yields an ML decoder.
  - ➢ MLSDA is ML algorithm with complexity independent of the constraint length.
- Efficient methods for estimating the low bit error probability of the coded system
  - ➢ Stationary channel model: variance scaling is used
  - ➢ Non-stationary channel model: both variance and mean translation are used
  - Variance scaling: enlarging noise variance to increase error probability
  - Mean translation: noise with nonzero mean

4

A brief review of the MLSDA

- The trellis-based *maximum-likelihood soft-decision sequential decoding algorithm* (MLSDA) is a maximum likelihood decoder whose computational complexity is independent of the constraint length.
- The MLSDA is a modification of the conventional stack decoding algorithm where the *Fano* metric is replaced by a new metric:

$$M(x_{(\ell n-1)}) = \sum_{j=0}^{\ell n-1} M(x_j)$$

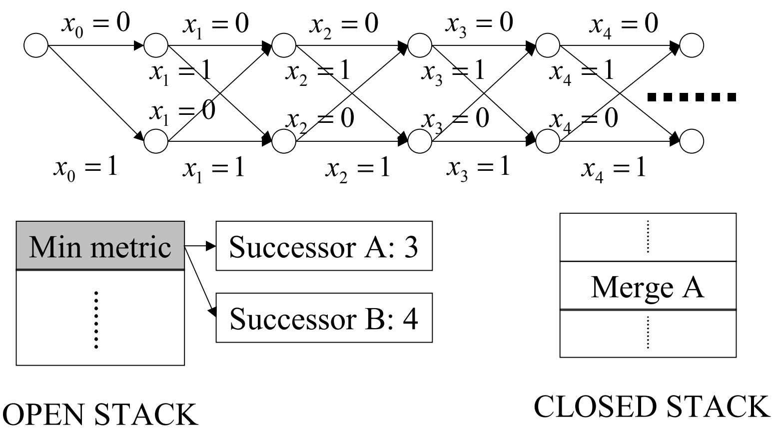
$$M(x_j) = (y_j \oplus x_j)\left|\ln \frac{\Pr(r_j \mid 0)}{\Pr(r_j \mid 1)}\right| \text{ where } y_j \text{ is the hard-decision result of the}$$

jth received bit r_j and x_j is the jth transmitted bit

5

A brief review of the MLSDA

- The property of maximum-likelihood is achieved by the second stack and the non-decreasing metric of the algorithm.



6

A brief review of the MLSDA

- The stack size constraint is not considered here (infinite stack size).
- The computational complexity is much less than Viterbi decoders when the SNR is high.
- In spite of the low computational complexity of the MLSDA decoders, the performance evaluation of the low bit error rate (BER) is still a difficult task and hence some more efficient method is desired.

7

MC simulation of the coded system

- Monte Carlo (MC) simulation techniques are popular when estimating the performance of the coded system because of their generality and simplicity.

$$P(\boldsymbol{x}' \mid \boldsymbol{x}) = \int \cdots \int_{D(\boldsymbol{x}')} f(\boldsymbol{r} \mid \boldsymbol{x}) d\boldsymbol{r} = \int \cdots \int_{\Re^{n(L+m)}} I_{\boldsymbol{x}'}(\boldsymbol{r}) f(\boldsymbol{r} \mid \boldsymbol{x}) d\boldsymbol{y}$$

$$\text{where } I_{\boldsymbol{x}'}(\boldsymbol{r}) = \begin{cases} 1 & \text{for } \boldsymbol{r} \in D(\boldsymbol{x}') \ (\boldsymbol{x}' \text{ decoded}) \\ 0 & \text{for } \boldsymbol{r} \notin D(\boldsymbol{x}') \ (\boldsymbol{x}' \text{ not decoded}) \end{cases}$$

$$\hat{P}_{N_{MC}}(\boldsymbol{x}' \mid \boldsymbol{x}) = \frac{1}{N_{MC}} \sum_{l=1}^{N_{MC}} I_{\boldsymbol{x}'}(\boldsymbol{r}^{(l)})$$

$$\hat{P}_b = \frac{1}{L} E[N_b \mid \boldsymbol{x} \text{ transmitted}] = \frac{1}{L} \sum_{\boldsymbol{x}' \in C} n_b(\boldsymbol{x}', \boldsymbol{x}) \hat{P}_{N_{MC}}(\boldsymbol{x}' \mid \boldsymbol{x})$$

where $n_b(\boldsymbol{x}', \boldsymbol{x})$ is the number of error bits of decoding $\boldsymbol{x}'$ instead of $\boldsymbol{x}$ and L is the length of the information bits

8

MC simulation of the coded system

$$\mathrm{var}[\hat{P}_{N_{MC}}(\boldsymbol{x}'\,|\,\boldsymbol{x})] = \frac{1}{N_{MC}}\mathrm{var}[I_{\boldsymbol{x}'}(\boldsymbol{r})]$$

- The variance of estimated BER will be large because the values of $I_{\boldsymbol{x}'}(\boldsymbol{r})$ are 1 or 0.
- Hence, MC simulations suffer from the tremendous time consumption while the errors occur very rarely if precise estimation is desired.

9

Importance sampling for the simulation

- When the bit error probability of the system is low, it the MC method simulation is very time consuming.
- In order to save time, "important" error events are generated more often by biasing the noise.
  - Stationary channel model: Biasing the channel for each codeword bit to enhance the probability of the error events.
  - Non-stationary channel model: Biasing the channel for the codeword bits of the important error events (other close codewords).

$$P(\boldsymbol{x}'\,|\,\boldsymbol{x}) = \int\cdots\int_{D(x')} \frac{f(\boldsymbol{r}\,|\,\boldsymbol{x})}{f^*(\boldsymbol{r}\,|\,\boldsymbol{x})} I_{\boldsymbol{x}'}(\boldsymbol{r}) f^*(\boldsymbol{r}\,|\,\boldsymbol{x})d\boldsymbol{y}$$

$$\hat{P}^*_{N_{IS}}(\boldsymbol{x}'\,|\,\boldsymbol{x}) = \frac{1}{N_{IS}}\sum_{l=1}^{N_{IS}} w(\boldsymbol{r}^{(l)}\,|\,\boldsymbol{x}) I_{\boldsymbol{x}'}(\boldsymbol{r}^{(l)})$$

$$w(\boldsymbol{r}\,|\,\boldsymbol{x}) = \frac{f(\boldsymbol{r}\,|\,\boldsymbol{x})}{f^*(\boldsymbol{r}\,|\,\boldsymbol{x})}$$

10

Importance sampling for the simulation

$$\mathrm{var}[\hat{P}_{N_{IS}}(\boldsymbol{x}' \mid \boldsymbol{x})] = \frac{1}{N_{IS}} \mathrm{var}^*[w(\boldsymbol{r} \mid \boldsymbol{x}) I_{\boldsymbol{x}'}(\boldsymbol{r})]$$

- The original channel is first biased to a more erroneous channel to generate more errors and the calculated data are then weighted by the weight function to recover the real BER.
- As above, the expect value of the estimated BER is just the same as the real BER and therefore the method is practical.
- If the biased function is properly chosen, the distribution of $w(\boldsymbol{r}|\boldsymbol{x})I_{\boldsymbol{x}'}(\boldsymbol{r})$ is not as discrete as the MC simulations so the variance will be smaller.

Importance sampling for the simulation

- If the biased function is not suitable for the simulated channel, the weight will be too small and hence the simulation will fail.
  - Weight too small: the variance will converge fast but the estimated BER will be far away from the real value (too small).
- Although improper biased functions can be used to get the real BER, it wastes much more simulation time than ordinary MC trials.

Importance sampling for the simulation

- The optimal importance sampling density:

$$f_{\mathbf{x}'}^{*}(\mathbf{r} \mid \mathbf{x}) = \frac{f(\mathbf{r} \mid \mathbf{x}) I_{\mathbf{x}'}(\mathbf{r})}{P(\mathbf{x}' \mid \mathbf{x})}$$

- However, the known $P(\mathbf{x}'|\mathbf{x})$ makes the density impractical but it gives a guidance of finding "good" densities.
- The BER is calculated as follows:

$$\hat{P}_b^{*} = \frac{1}{L} E[N_b \mid \mathbf{x} \text{ transmitted}] = \frac{1}{L} \sum_{\mathbf{x}' \in C} n_b(\mathbf{x}', \mathbf{x}) \hat{P}_{N_{IS}}^{*}(\mathbf{x}' \mid \mathbf{x})$$

$$\hat{P}_{N_{IS}}^{*}(\mathbf{x}' \mid \mathbf{x}) = \frac{1}{N_{IS}} \sum_{l=1}^{N_{IS}} w(\mathbf{r}^{(l)} \mid \mathbf{x}) I_{\mathbf{x}'}(\mathbf{r}^{(l)})$$

13

Importance sampling for the simulation

- The RPE (relative precision estimate) is the standard deviation of the estimator as a percentage $\hat{P}_b^{*}$ is

$$\frac{\sqrt{\operatorname{var}[\hat{P}_b^{*}]}}{\hat{P}_b^{*}} \times 100\%$$

- With smaller RPE, the estimated BER is more accurate.

14

Simulation models

- The simulation was operated on AWGN channels.
- The codewords are antipodally transmitted and the received vectors are

$$r_j = (-1)^{x_j}\sqrt{\varepsilon} + \lambda_j \text{ where } \varepsilon \text{ is the signal energy of each codeword bit}$$
$$\lambda_j \text{ is the jth noise sample}$$

- The SNR per information bit is thus

$$SNR = E_s / N_o = \frac{N}{L}\left(\frac{\varepsilon}{N_o}\right) \text{ where } \begin{cases} \text{N : the total length of the codeword} \\ \text{L : the block length of the information bits} \end{cases}$$

- Two kinds of channel are used: stationary and non-stationary.

15

Simulation models

➢ *Stationary* channel model: ***each*** codeword bit suffers from noise with *zero mean* and *larger variance*

transmitted signal	-1	-1	-1	-1	-1	-1	-1	-1	-1	......
	+	+	+	+	+	+	+	+	+	......
	N_{IS}	N_{IS}	N_{IS}	N_{IS}	N_{IS}	N_{IS}	N_{IS}	N_{IS}	N_{IS}	......

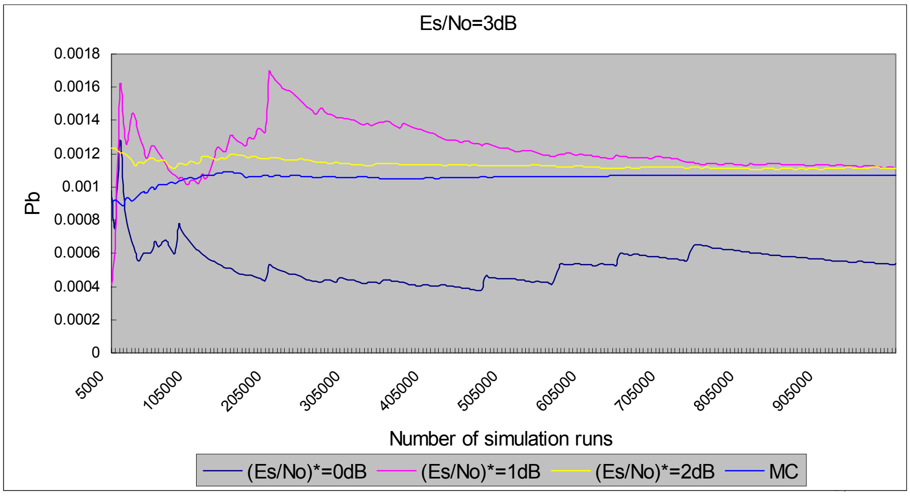
➢ *Non-stationary* channel model: ***free distance codeword position*** suffers from noise with nonzero mean

transmitted signal	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	......
	+	+	+	+	+	+	+	+	+	+	+	+	+	+	......
	N'	N	N'	N	N	N	N	N'	N	N	N'	N	N	N'	......
a free distance codeword	1	0	1	0	0	0	0	1	0	0	1	0	0	1	......

16

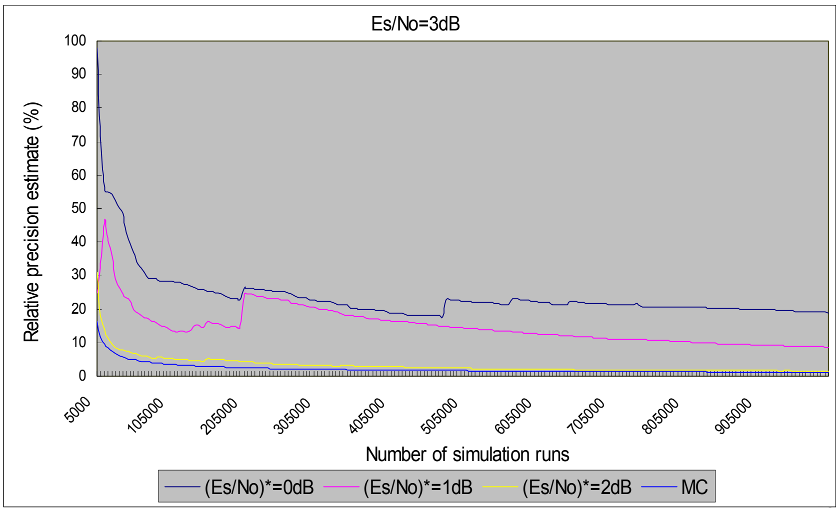
Simulation results *biased stationary channels*

- Estimating the performance of the channel of Es/No=3dB



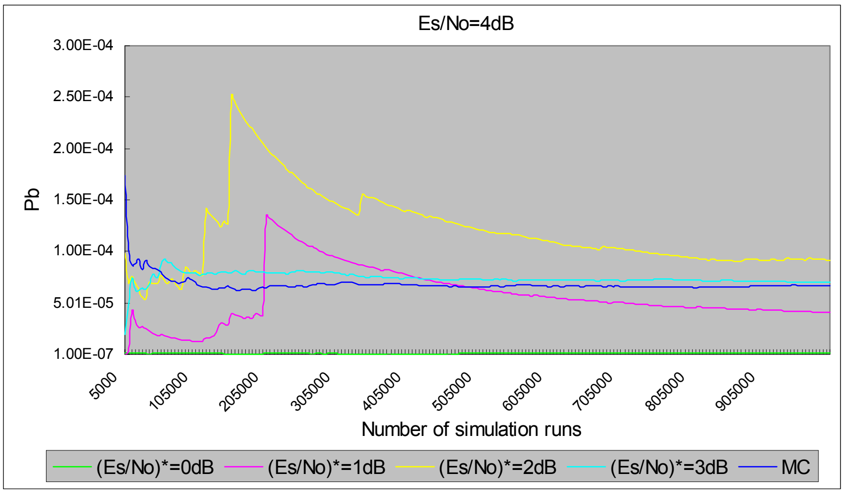
Simulation results *biased stationary channels*

- Estimating the performance of the channel of Es/No=3dB



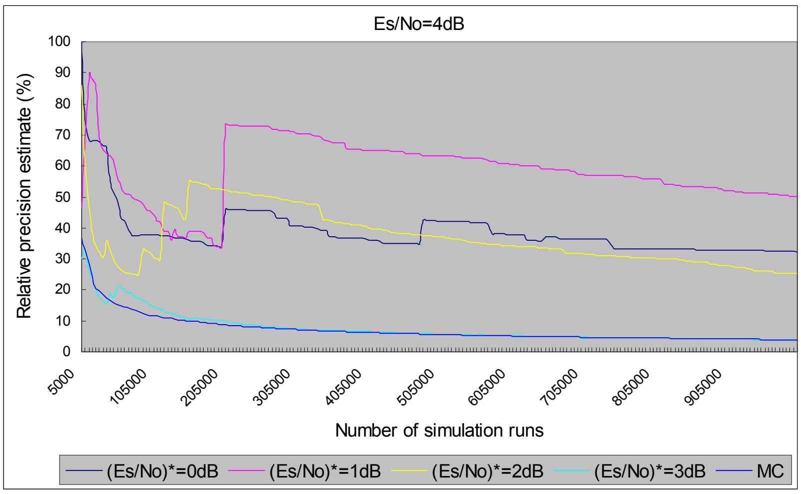
Simulation results *biased stationary channels*

•Estimating the performance of the channel of Es/No=4dB



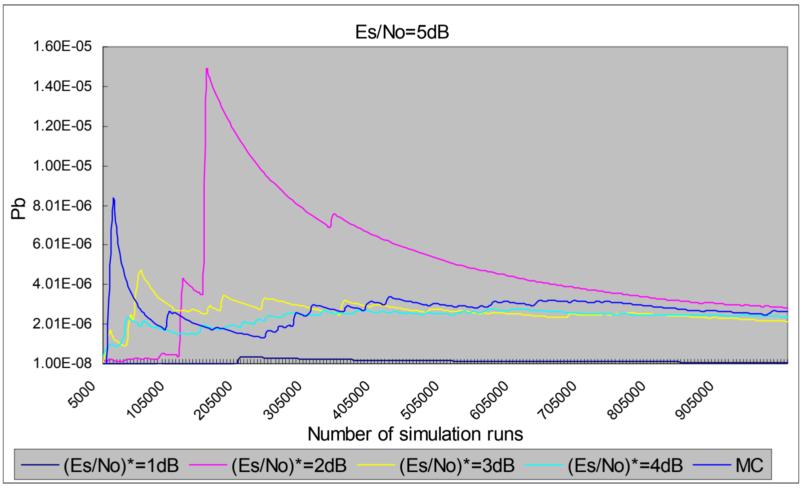
Simulation results *biased stationary channels*

•Estimating the performance of the channel of Es/No=4dB



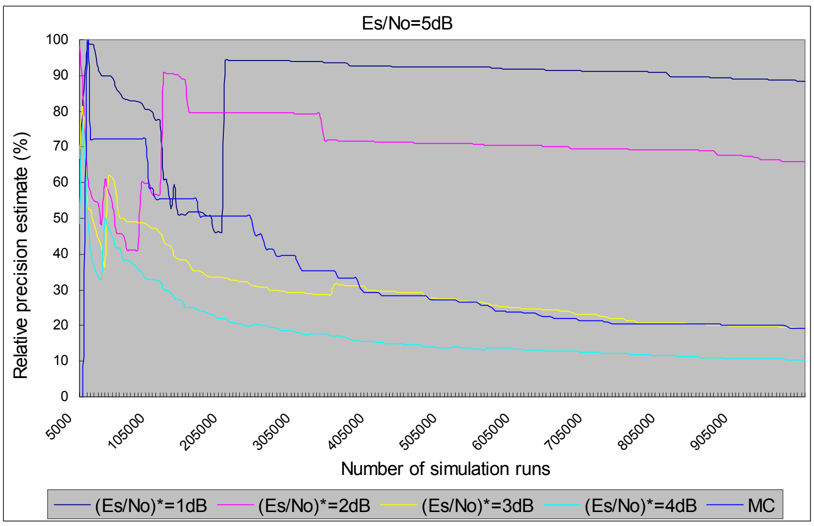
Simulation results *biased stationary channels*

•Estimating the performance of the channel of Es/No=5dB



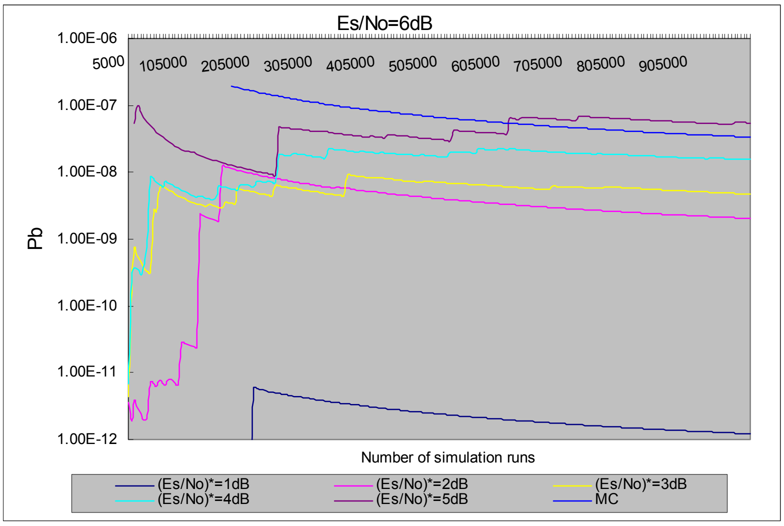
Simulation results *biased stationary channels*

•Estimating the performance of the channel of Es/No=5dB



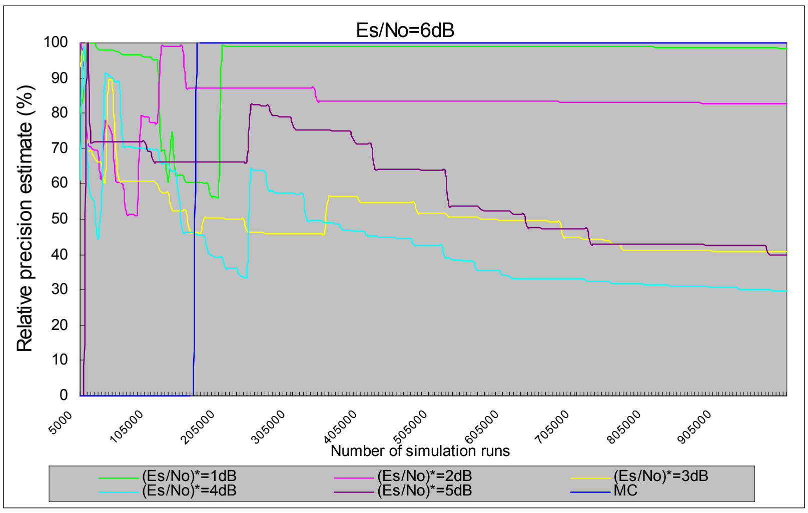
Simulation results *biased stationary channels*

•Estimating the performance of the channel of Es/No=6dB



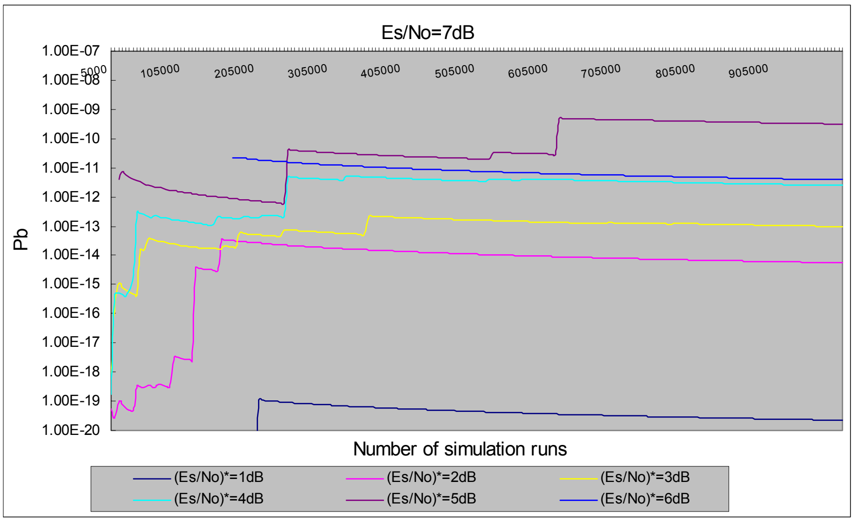
Simulation results *biased stationary channels*

•Estimating the performance of the channel of Es/No=6dB



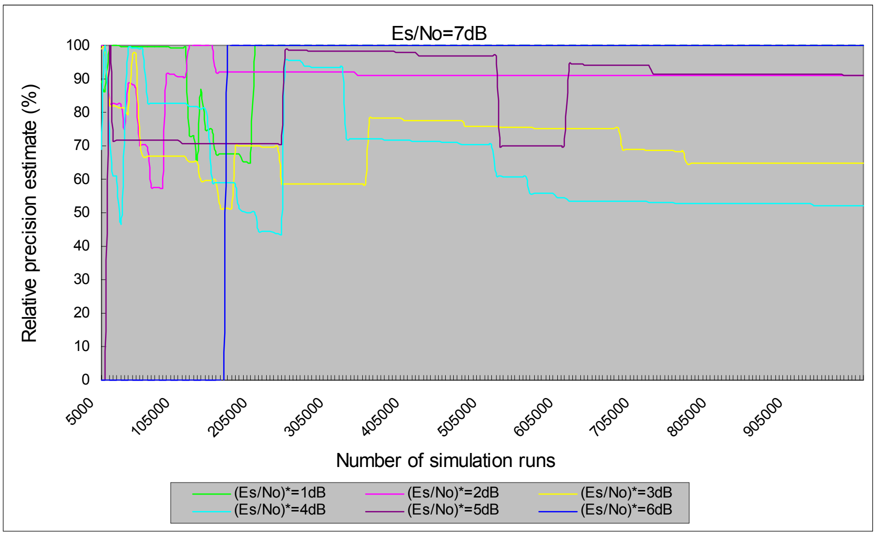
Simulation results *biased stationary channels*

•Estimating the performance of the channel of Es/No=7dB



Simulation results *biased stationary channels*

•Estimating the performance of the channel of Es/No=7dB



Simulation results *biased stationary channels*

- Lists of the required computational complexity for 10% relative precision estimate for Es/No=3dB~4dB.

Es/No=3dB				
	pb	rpe	complexity	IS-complexity/MC-complexity
(Es/No)*=2dB	1.19E-03	9.99%	6.26E+07	3.846251182
(Es/No)*=2.5dB	8.86E-04	9.93%	1.40E+07	0.860821386
MC	8.65E-04	9.98%	1.63E+07	1

Es/No=4dB				
	pb	rpe	complexity	IS-complexity/MC-complexity
(Es/No)*=2dB	6.56E-05	10.00%	1.12E+10	201.4669211
(Es/No)*=3dB	8.03E-05	9.99%	2.18E+08	3.930713075
(Es/No)*=3.5dB	6.10E-05	9.98%	4.92E+07	0.888036214
MC	6.69E-05	9.97%	5.54E+07	1

27

Simulation results *biased stationary channels*

- Lists of the required computational complexity for 10% relative precision estimate for Es/No=5dB~6dB.

Es/No=5dB				
	pb	rpe	complexity	IS-complexity/MC-complexity
(Es/No)*=3dB	1.76E-06	10.00%	6.76E+09	11.71058356
(Es/No)*=4dB	2.24E-06	9.99%	4.11E+08	0.71176908
(Es/No)*=4.5dB	2.70E-06	9.99%	3.54E+08	0.61317836
MC	2.68E-06	9.98%	5.77E+08	1

Es/No=6dB				
	pb	rpe	complexity	IS-complexity/MC-complexity
(Es/No)*=4dB	3.87E-08	9.98%	3.30E+10	1.467250732
(Es/No)*=5dB	4.22E-08	9.99%	6.83E+09	0.303572365
(Es/No)*=5.5dB	3.93E-08	9.99%	6.90E+09	0.306531639
MC	4.49E-08	9.97%	2.25E+10	1

28

Simulation results *biased stationary channels*

- The lists above shows that if the density is not properly chosen, the estimate will be inaccurate and waste on computation will be caused.
- More computational efforts saving will be gained with higher SNR.

29

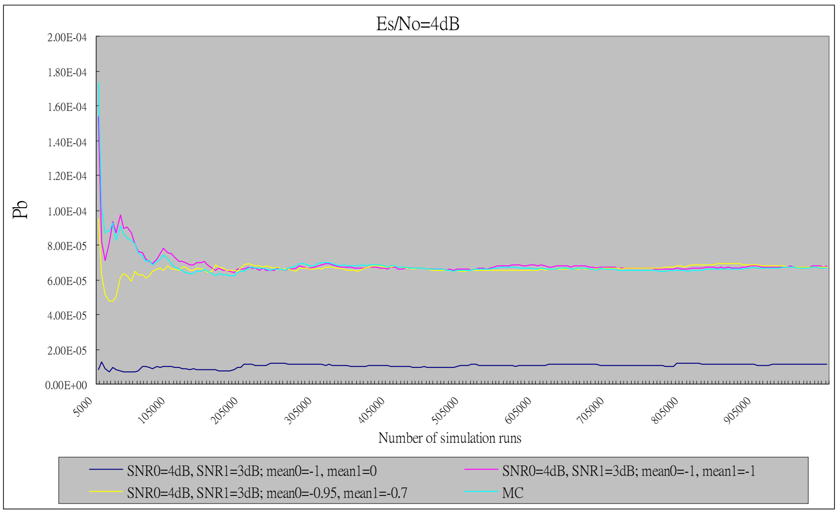
Simulation results *biased non-stationary channels*

- Only codewords of free distance from the transmitted codeword will be most probably decoded as errors when the SNR is high.
- Thus the distance spectrum decides the usefulness of the non-stationary channel model.
- The free distance is 10 and the number of codes with weight 10 is 668.
- When SNR is larger than 4dB, not all positions of the free distance codeword are mean translated because the weight will be too small and thus the calculated BER will be quite different from the real value.

30

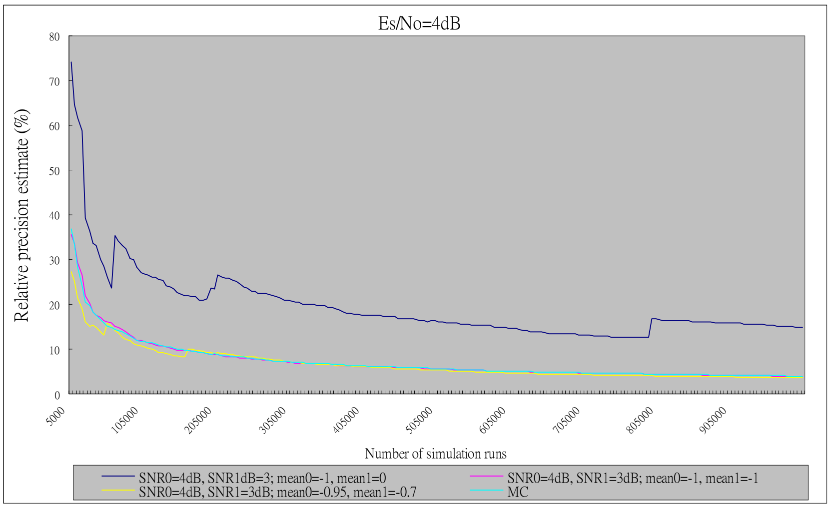
Simulation results *biased non-stationary channels*

•Estimating the performance of the channel of Es/No=4dB



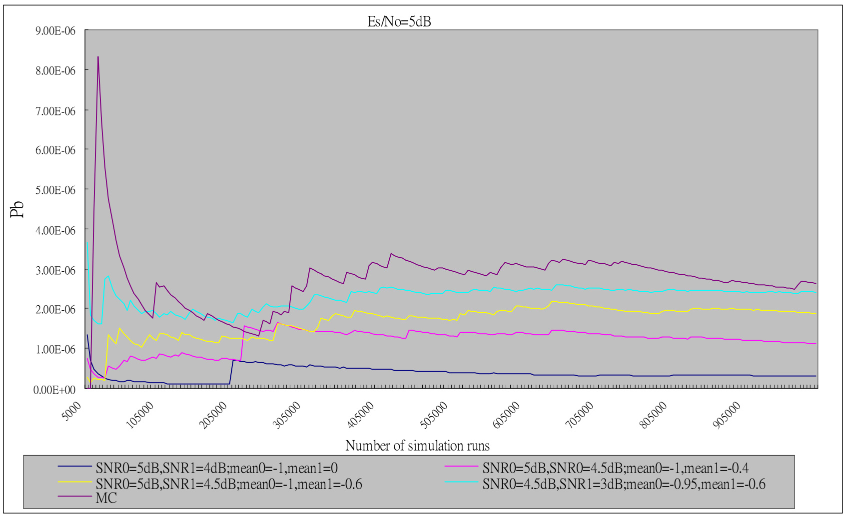
Simulation results *biased non-stationary channels*

•Estimating the performance of the channel of Es/No=4dB



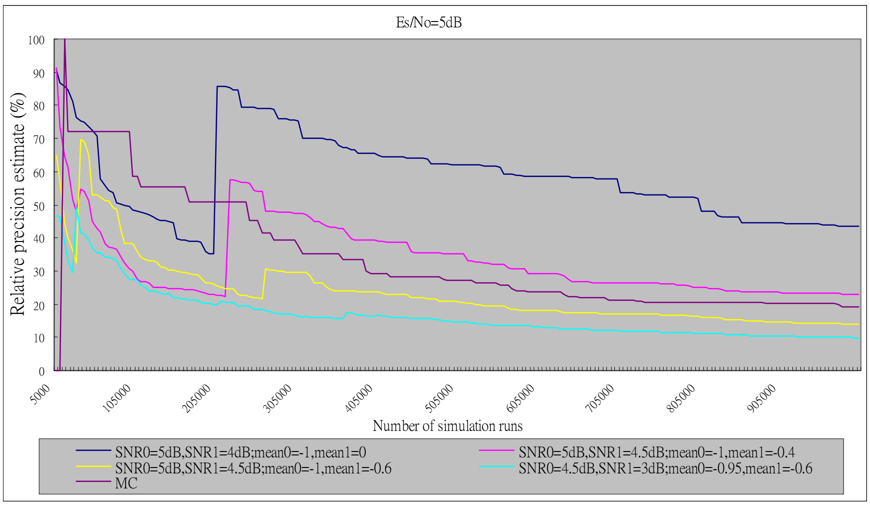
Simulation results *biased non-stationary channels*

•Estimating the performance of the channel of Es/No=5dB



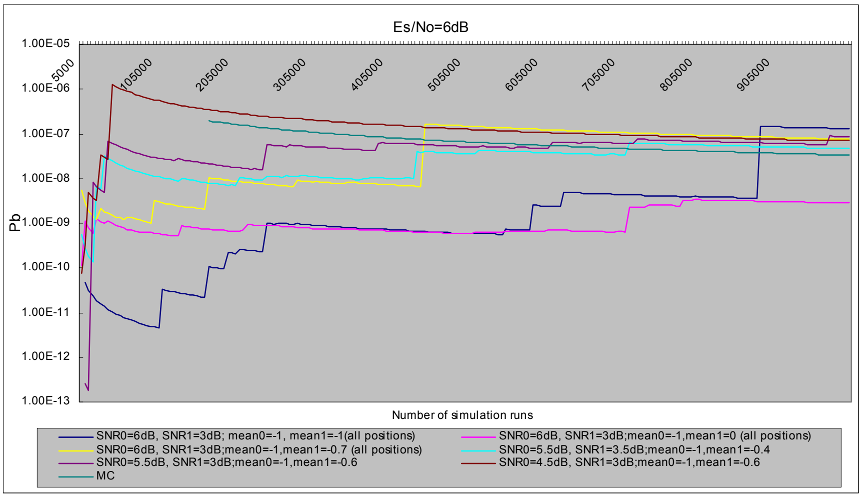
Simulation results *biased non-stationary channels*

•Estimating the performance of the channel of Es/No=5dB



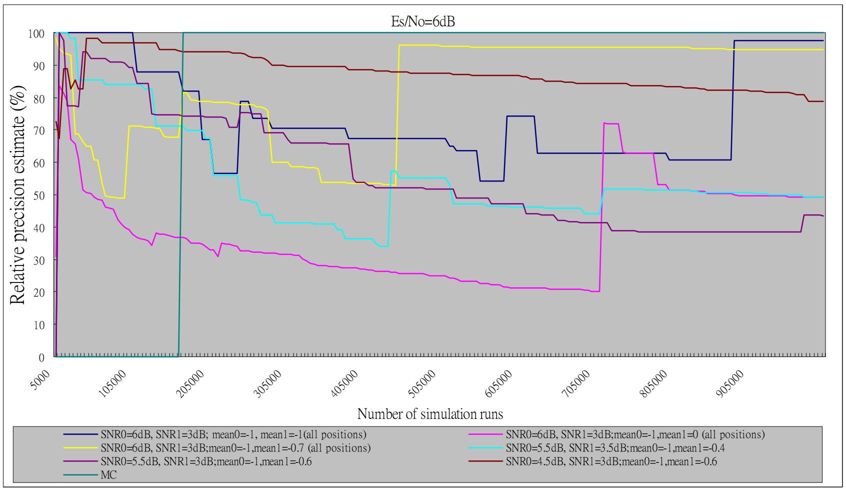
Simulation results *biased non-stationary channels*

•Estimating the performance of the channel of Es/No=6dB



Simulation results *biased non-stationary channels*

•Estimating the performance of the channel of Es/No=6dB



Simulation results *biased non-stationary channels*

- Lists of the required computational complexity for 10% relative precision estimate for Es/No=5dB~6dB. Only 3 out of 10 positions are biased by SNR1 and mean1.

Es/No=5dB				
	pb	rpe	complexity	IS-complexity/MC-complexity
SNR0=4.5dB, SNR1=3dB; mean0=-0.95, mean1=-0.6	2.54E-06	9.99%	3.41E+08	0.59087876
MC	2.68E-06	9.98%	5.77E+08	1

Es/No=6dB				
	pb	rpe	complexity	IS-complexity/MC-complexity
SNR0=5.5dB, SNR1=3dB; mean0=-0.95, mean1=-0.6	4.27E-08	10.00%	4.80E+09	0.213042397
MC	4.49E-08	9.97%	2.25E+10	1

37

Conclusions

- The Monte Carlo simulation needs prohibitive simulation run times when Es/No is large.
- The use of the importance sampling provides much time-saving simulations with acceptable relative precision while appropriate channel model used.
- The choice of the biasing function decides the efficiency of the simulation.
- The biased stationary channel model can induce more efficient simulations than ordinary MC simulations.
- More time savings will be gained when non-stationary channel model is used with the importance sampling density function carefully chosen.

38