



Eidgenössische Technische Hochschule Zürich  
Swiss Federal Institute of Technology Zurich



# Information Theory

Lecture Notes

Stefan M. Moser

© Copyright Stefan M. Moser

Signal and Information Processing Lab  
ETH Zürich  
Zurich, Switzerland

Department of Electrical and Computer Engineering  
National Chiao Tung University (NCTU)  
Hsinchu, Taiwan

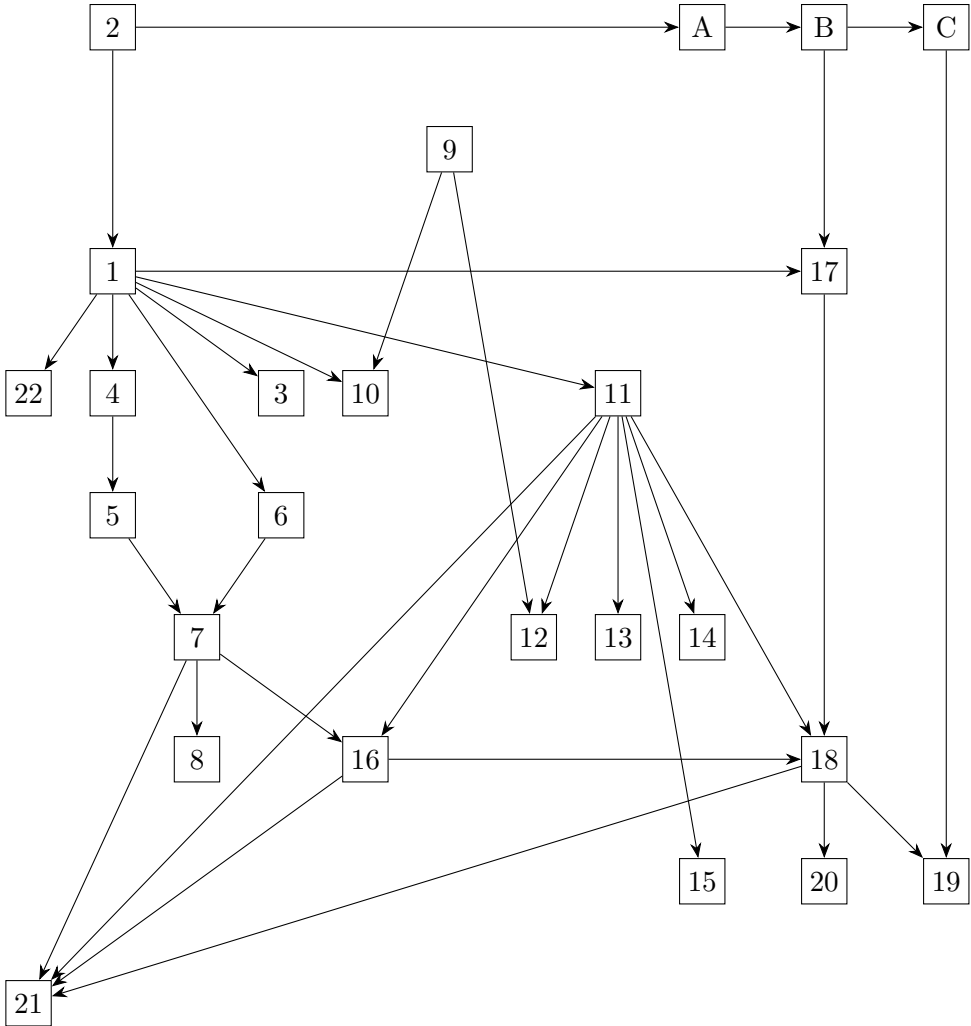
You are welcome to use these lecture notes for yourself, for teaching, or for any other noncommercial purpose. If you use extracts from these lecture notes, please make sure that their origin is shown. The author assumes no liability or responsibility for any errors or omissions.

5<sup>th</sup> Edition — 2017.

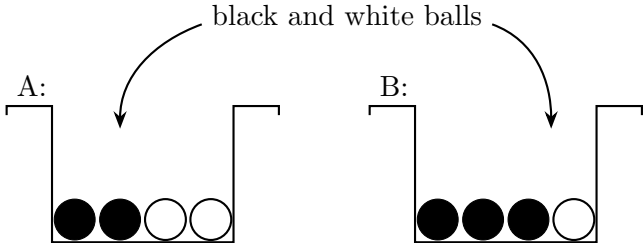
Version 5.3.

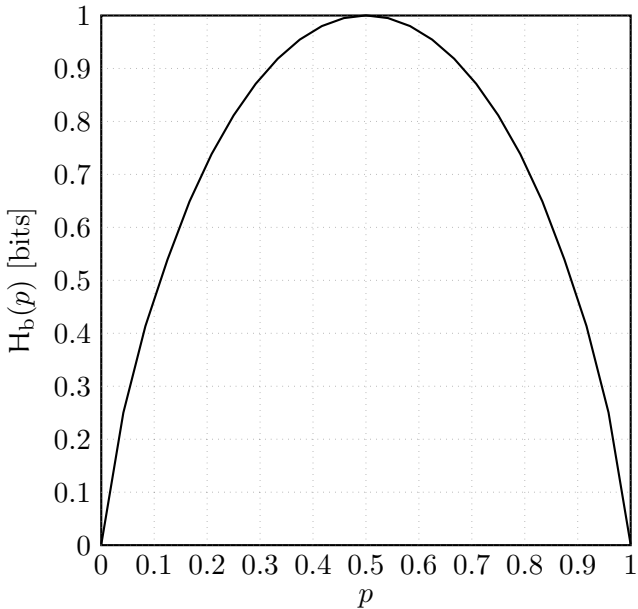
Compiled on 17 March 2017.

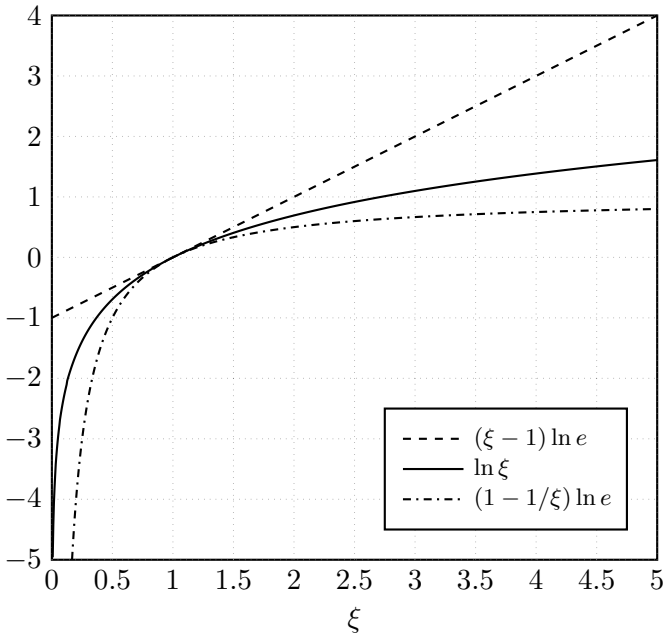
For the latest version see <http://moser-isi.ethz.ch/scripts.html>

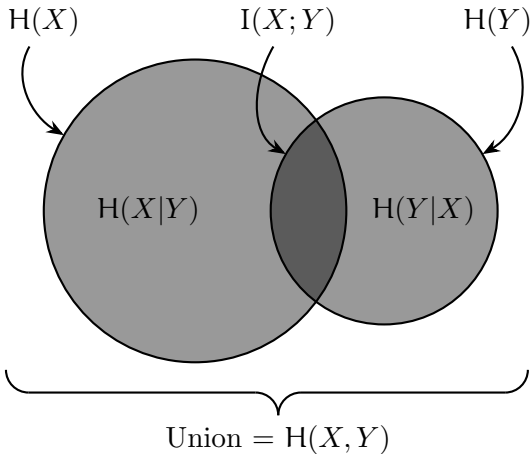


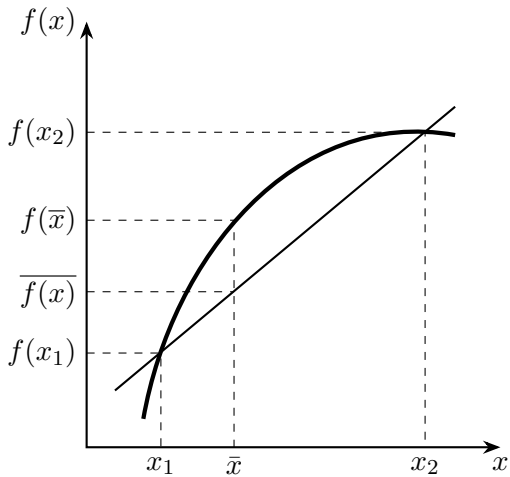
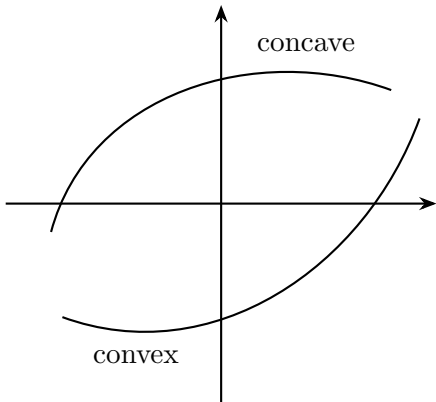
Dependency chart. An arrow has the meaning of “is required for”.

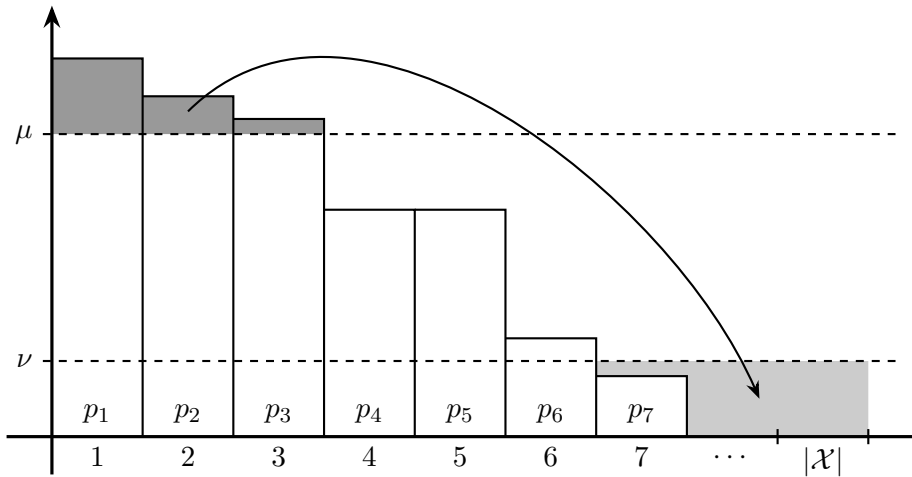




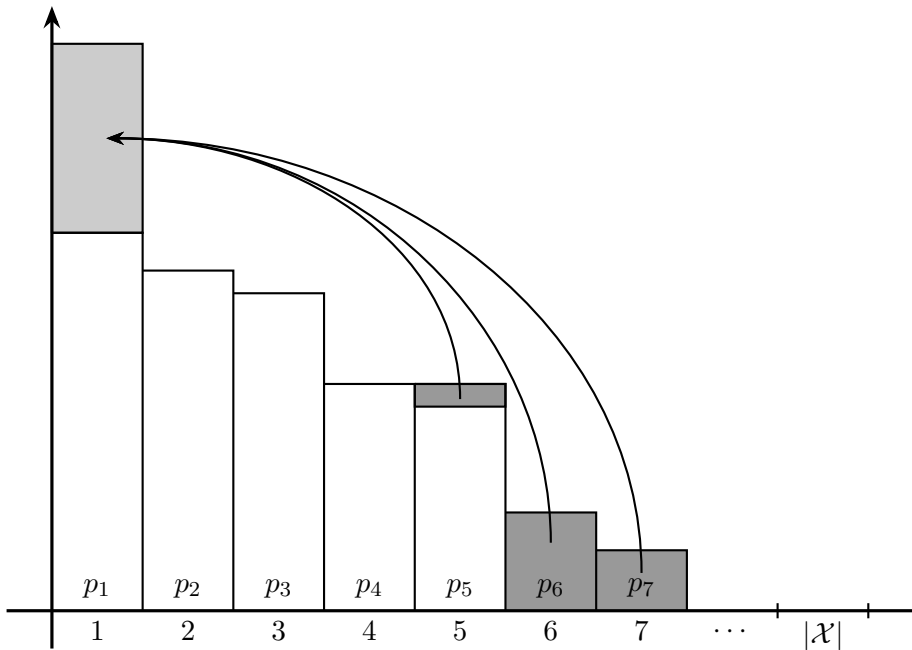




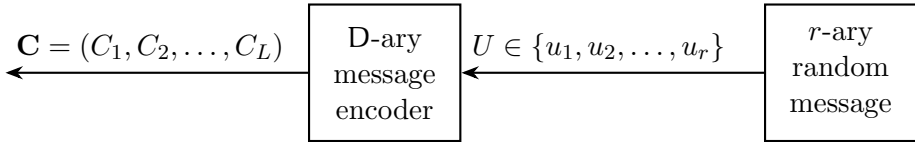




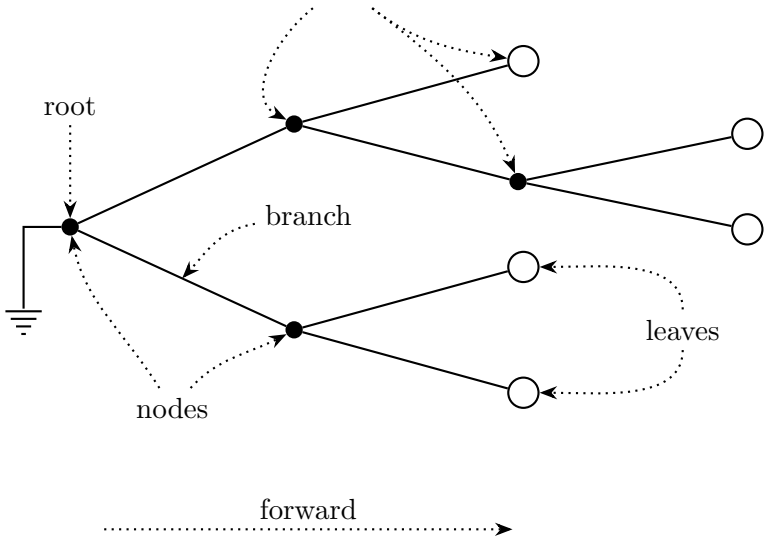


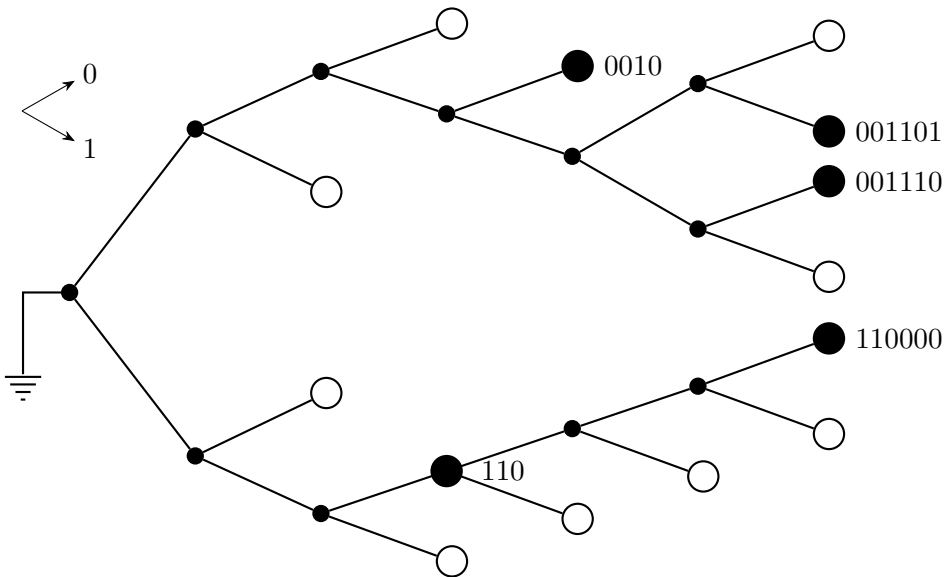


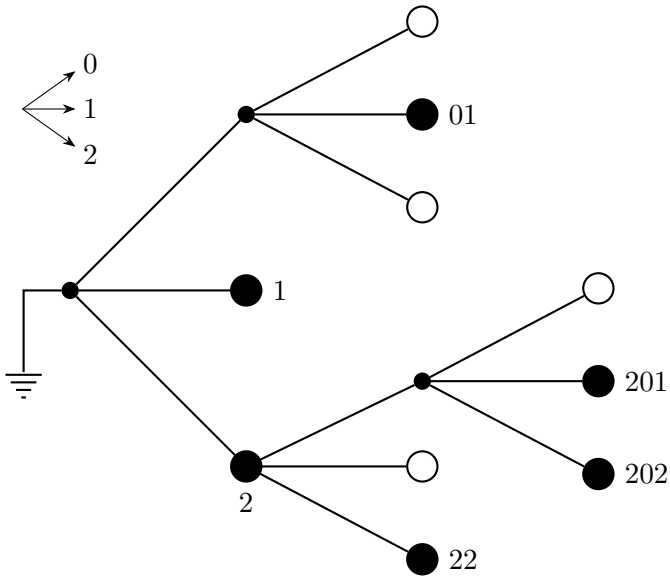
Friend	Probability	Sets of Phone Numbers					
		(i)	(ii)	(iii)	(iv)	(v)	(vi)
Alice	$\frac{1}{4}$	0011	001101	0	00	0	10
Bob	$\frac{1}{2}$	0011	001110	1	11	11	0
Carol	$\frac{1}{4}$	1100	110000	10	10	10	11

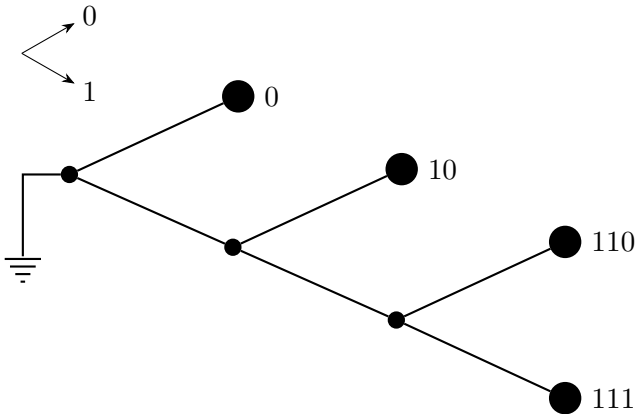


parent node with two children

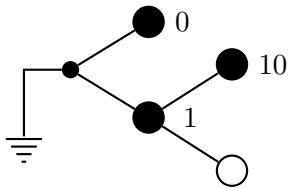






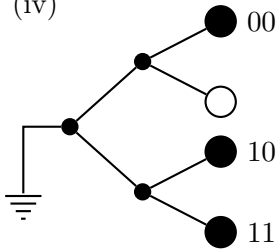


(iii)



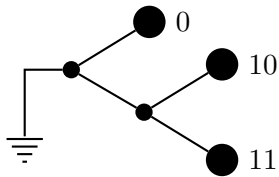
not prefix-free

(iv)



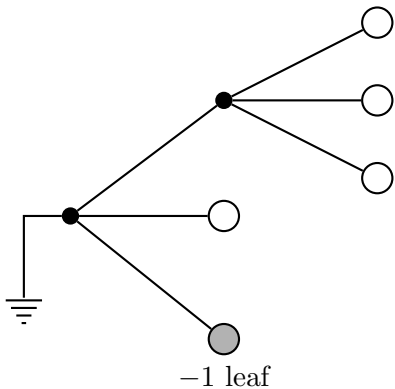
prefix-free

(v)–(vi)

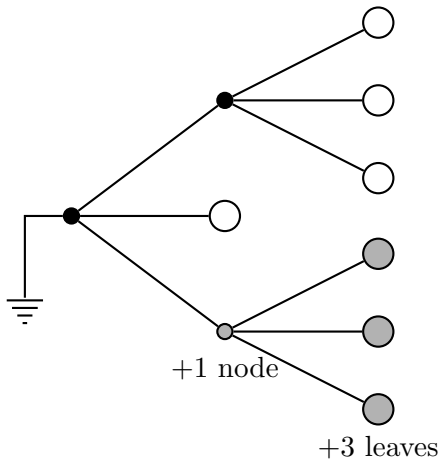


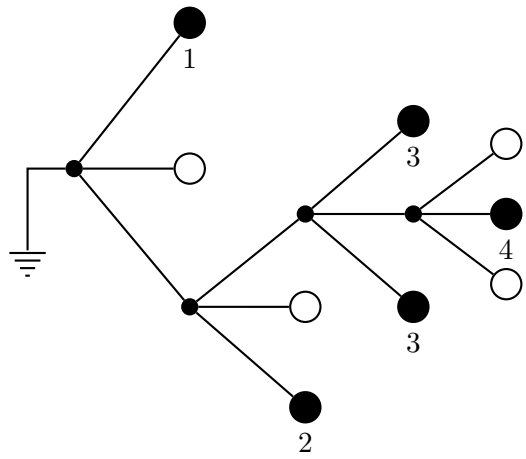
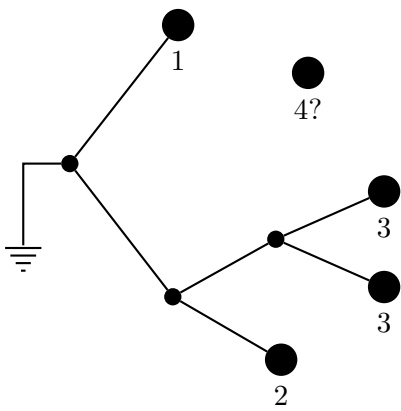
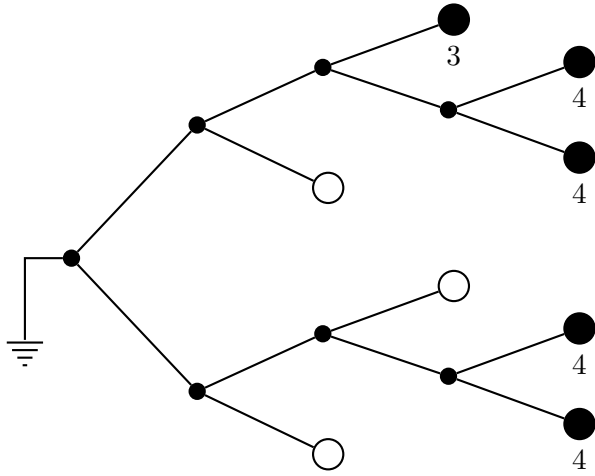
prefix-free

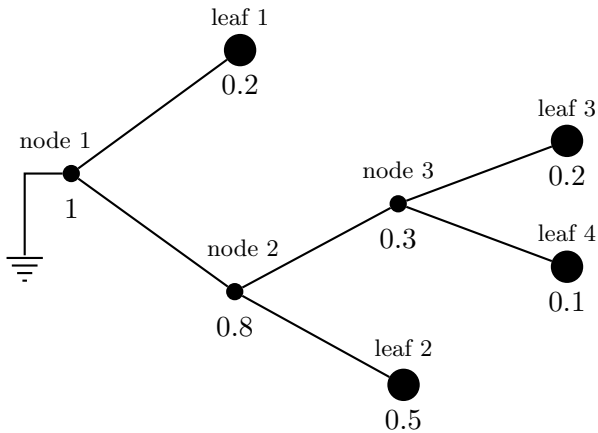


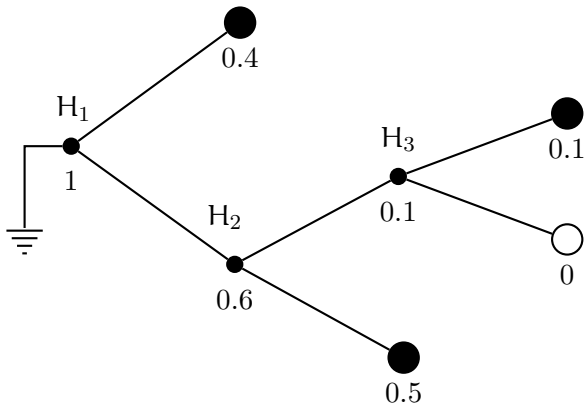


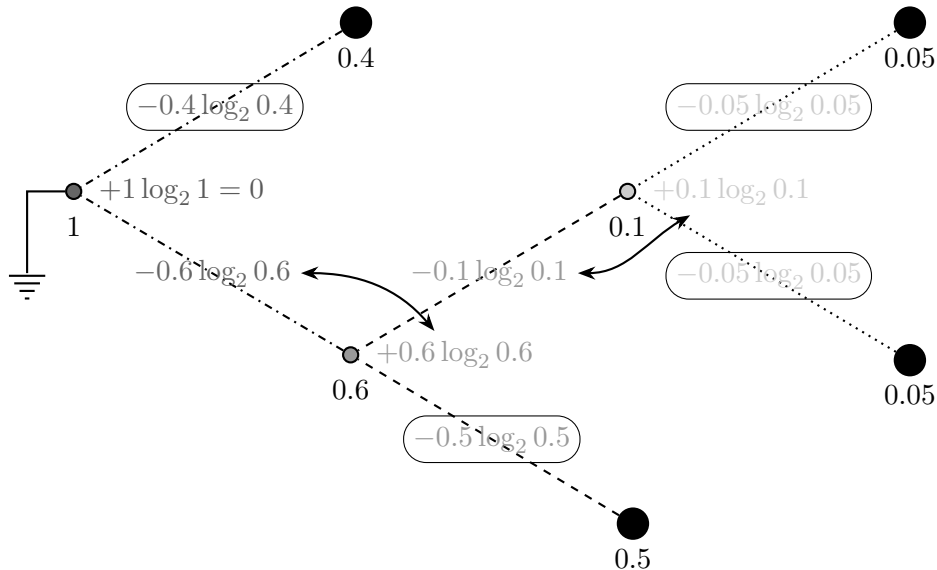
$\Rightarrow$











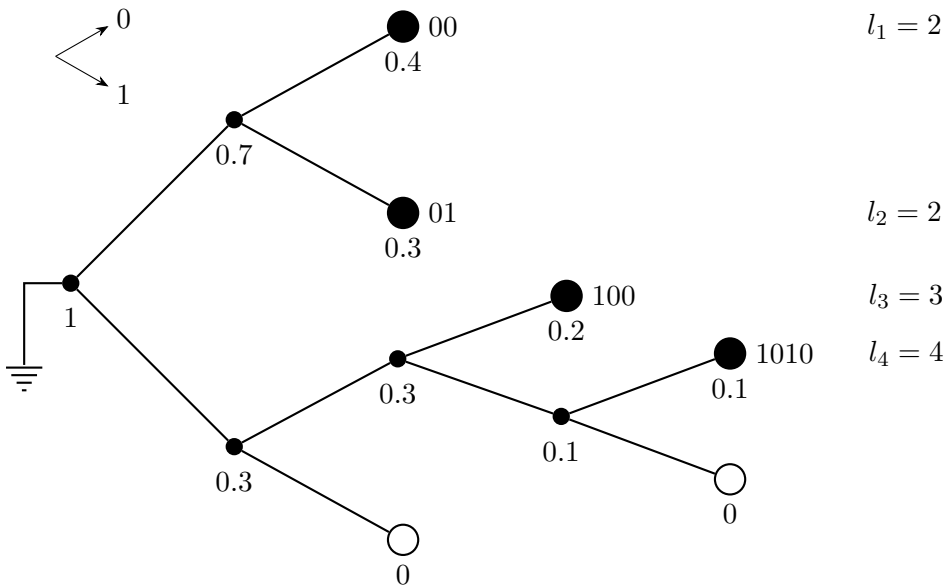
● ..... contribution of  $P_1 H_1$

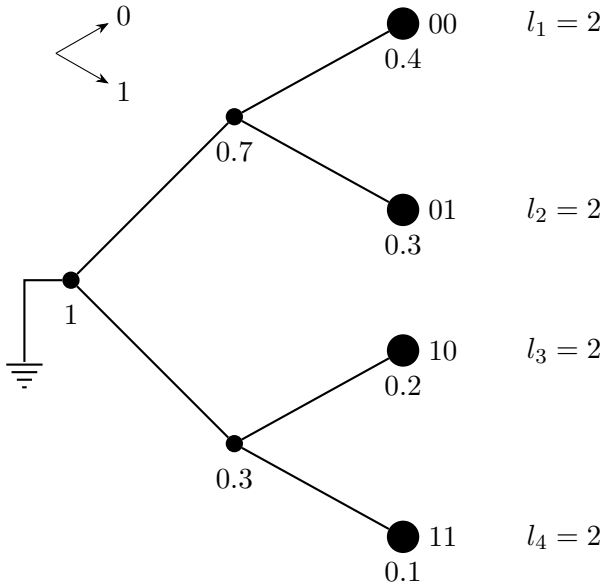
● ..... contribution of  $P_2 H_2$

○ ..... contribution of  $P_3 H_3$

↔ terms that cancel each other

○ contributions to leaf entropy





	$u_1$	$u_2$	$u_3$	$u_4$
$p_i$	0.4	0.3	0.2	0.1
$F_i$	0	0.4	0.7	0.9
binary representation	0.0	0.01100...	0.10110...	0.11100...
$\left\lceil \log_2 \frac{1}{p_i} \right\rceil$	2	2	3	4
shortened representation	0.00	0.01	0.101	0.1110
$\mathbf{c}_i$	00	01	101	1110



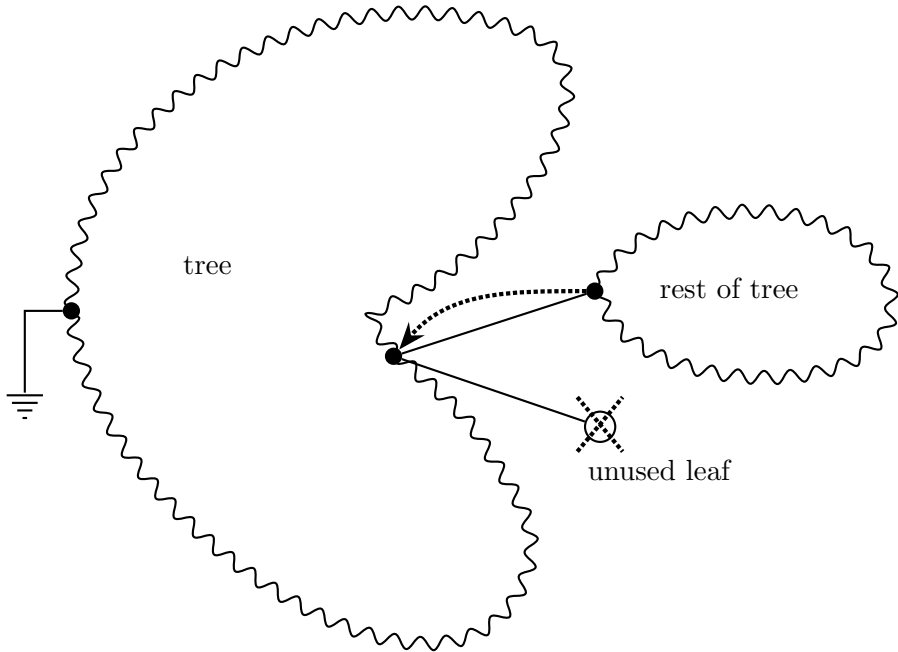
	$u_1$	$u_2$	$u_3$	$u_4$
$p_i$	0.4	0.3	0.2	0.1
$F_i$	0	0.4	0.7	0.9
ternary representation	0.0	0.10121...	0.20022...	0.22002...
$\left\lceil \log_3 \frac{1}{p_i} \right\rceil$	1	2	2	3
shortened representation	0.0	0.10	0.20	0.220
$\mathbf{c}_i$	0	10	20	220

$p_1$	$p_2$	$p_3$	$p_4$	$p_5$
0.35	0.25	0.15	0.15	0.1
0.6		0.4		
0		1		
0.35	0.25	0.15	0.15	0.1
0		0.15	0.25	
1		0	1	
		0.15		0.1
		0		1
<b>00</b>	<b>01</b>	<b>10</b>	<b>110</b>	<b>111</b>

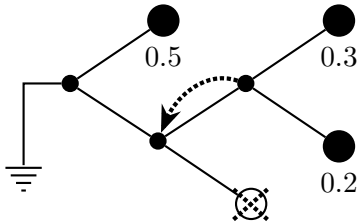
$p_1$	$p_2$	$p_3$	$p_4$	$p_5$
0.35	0.25	0.15	0.15	0.1
0.35	0.25	0.4		
<b>0</b>	<b>1</b>	<b>2</b>		
		0.15	0.15	0.1
		<b>0</b>	<b>1</b>	<b>2</b>
<b>0</b>	<b>1</b>	<b>20</b>	<b>21</b>	<b>22</b>

$p_1$	$p_2$	$p_3$	$p_4$	$p_5$	$p_6$	$p_7$	
0.35	0.3	0.15	0.05	0.05	0.05	0.05	
0.65 <b>0</b>		0.35 <b>1</b>					
0.35	0.3	0.15	0.05	0.05	0.05	0.05	
<b>0</b>	<b>1</b>	0.2 <b>0</b>		0.15 <b>1</b>			
		0.15	0.05	0.05	0.05	0.05	
		<b>0</b>	<b>1</b>	0.1 <b>0</b>		0.05	0.05
				0.05	0.05		
<b>0</b>	<b>1</b>						
<b>00</b>	<b>01</b>	<b>100</b>	<b>101</b>	<b>1100</b>	<b>1101</b>	<b>111</b>	

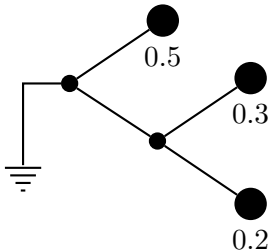
$p_1$	$p_2$	$p_3$	$p_4$	$p_5$	$p_6$	$p_7$
0.35	0.3	0.15	0.05	0.05	0.05	0.05
0.35 0	0.65 1					
	0.3	0.15	0.05	0.05	0.05	0.05
0.3 0	0.35 1					
	0.15	0.05	0.05	0.05	0.05	0.05
0.15 0	0.2 1					
	0.05	0.05	0.05	0.05	0.05	0.05
0.1 0	0.1 1					
	0.05	0.05	0.05	0.05	0.05	0.05
0	1	0	1	0	1	0
0	10	110	11100	11101	11110	11111



(i)

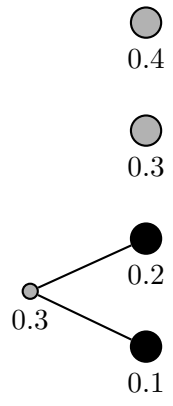


(ii)

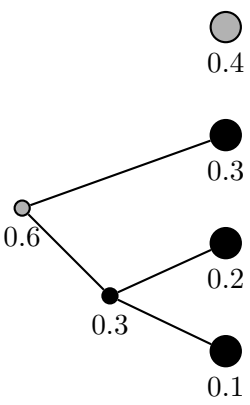




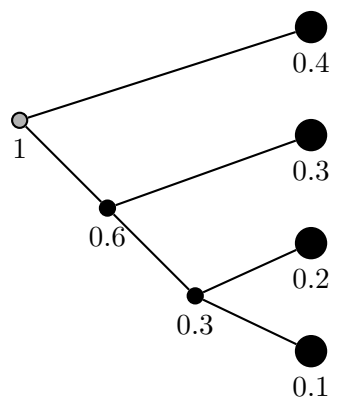
Step 1



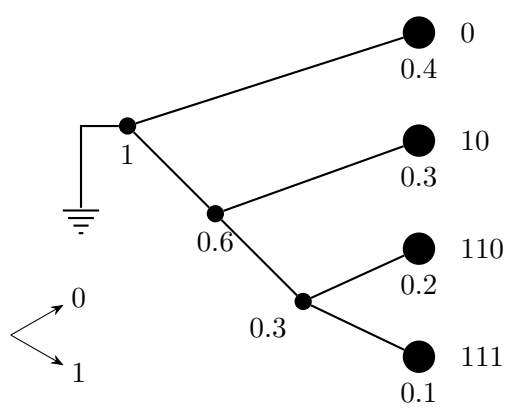
Step 2, first time



Step 2, second time

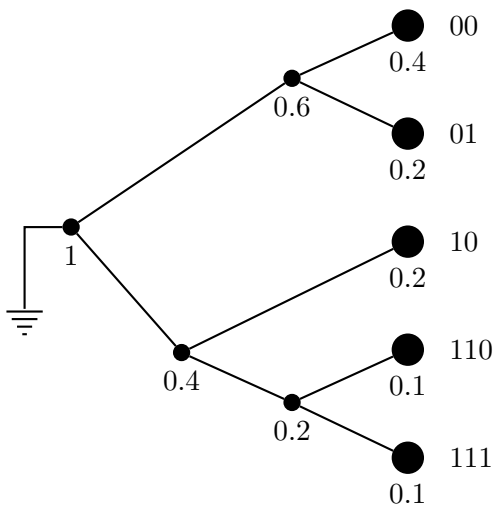
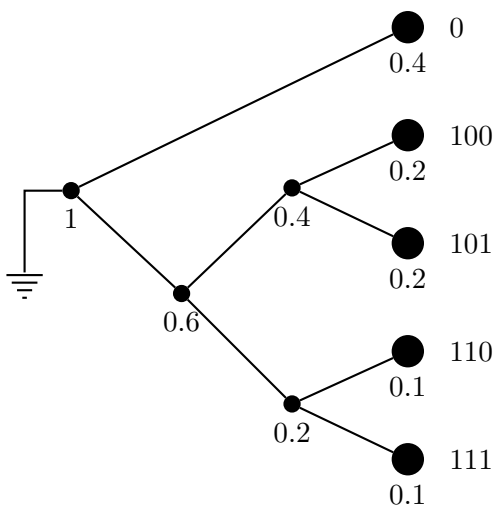
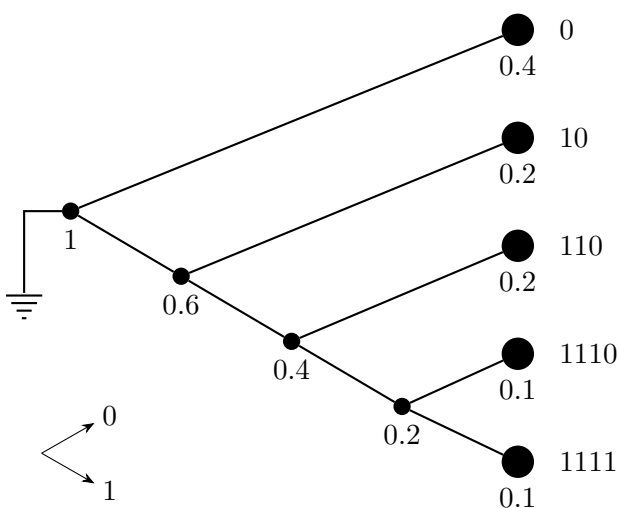


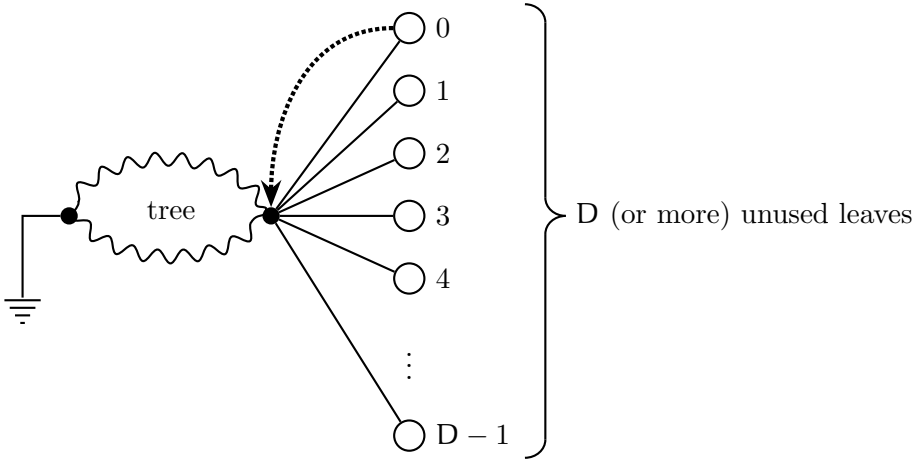
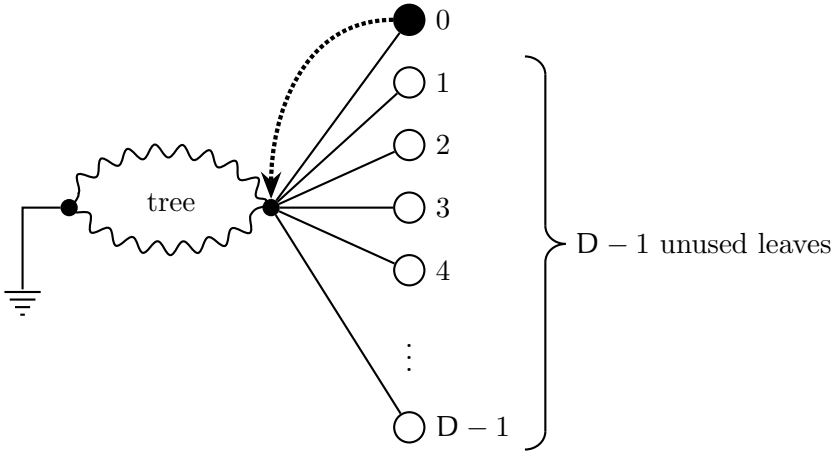
Step 2, third time

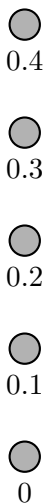


Step 3

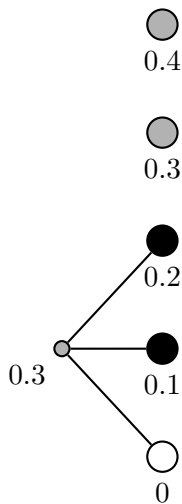




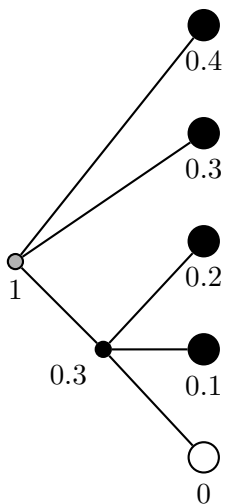




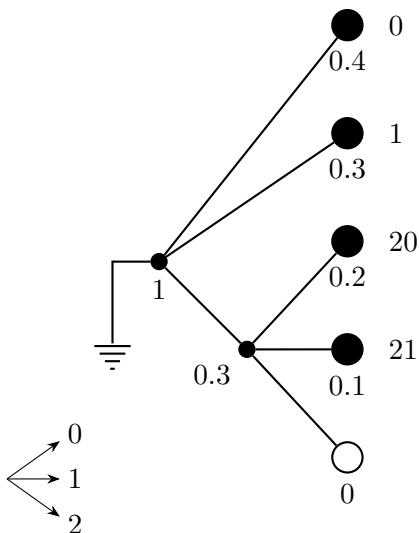
Step 1



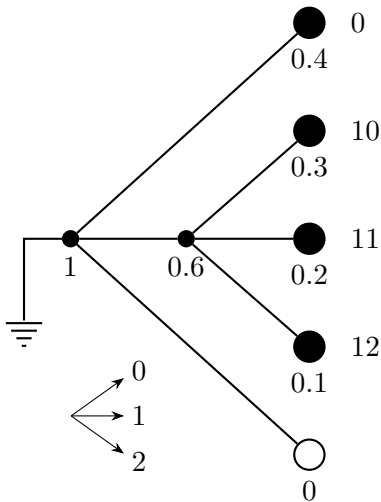
Step 2, first time



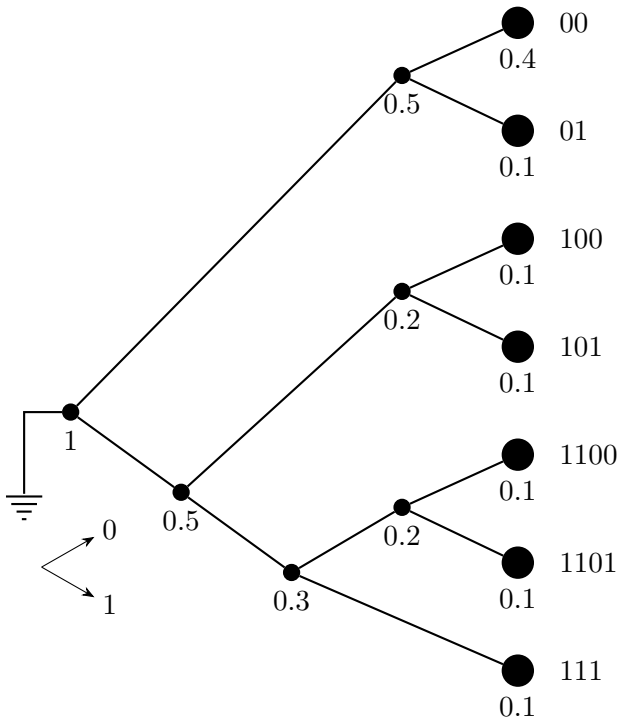
Step 2, second time



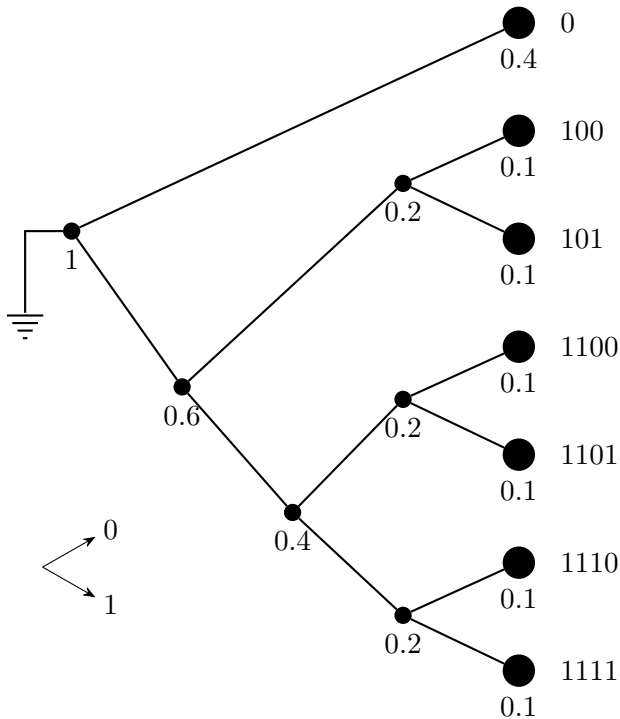
Step 3



$p_1$	$p_2$	$p_3$	$p_4$	$p_5$	$p_6$	$p_7$
0.4	0.1	0.1	0.1	0.1	0.1	0.1
0.5 0		0.5 1				
0.4	0.1	0.1	0.1	0.1	0.1	0.1
0	1	0.2 0		0.3 1		
		0.1	0.1	0.1	0.1	0.1
		0	1	0.2 0		0.1 1
				0.1	0.1	
				0	1	
00	01	100	101	1100	1101	111



	$u_1$	$u_2$	$u_3$	$u_4$	$u_5$	$u_6$	$u_7$
$p_i$	0.4	0.1	0.1	0.1	0.1	0.1	0.1
$F_i$	0	0.4	0.5	0.6	0.7	0.8	0.9
binary representation	0.0	0.01100	0.1	0.10011	0.10110	0.11001	0.11100
$\left\lceil \log_2 \frac{1}{p_i} \right\rceil$	2	4	4	4	4	4	4
shortened representation	0.00	0.0110	0.1000	0.1001	0.1011	0.1100	0.1110
$\mathbf{c}_i$	00	0110	1000	1001	1011	1100	1110





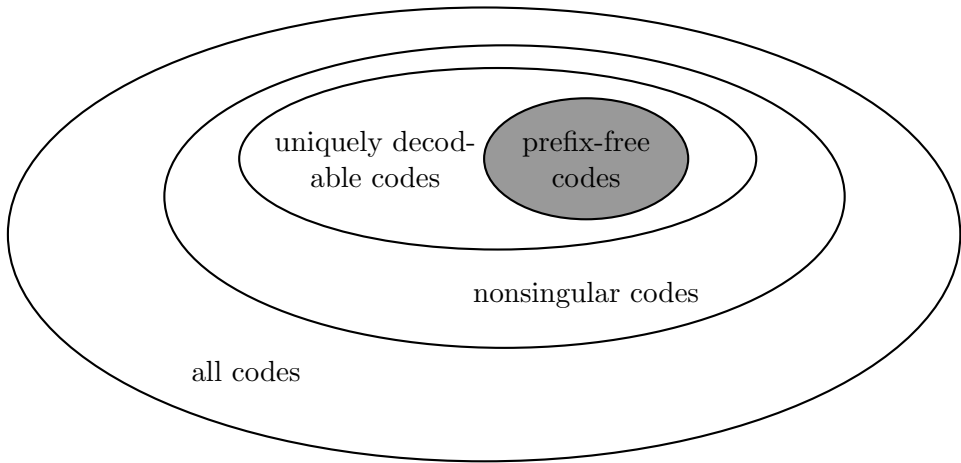
---

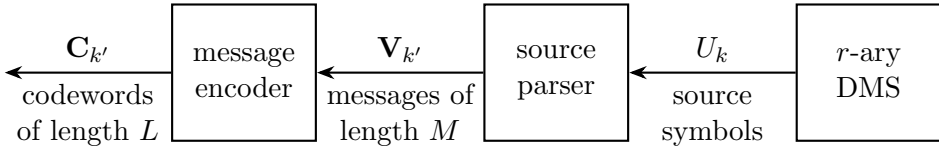
---

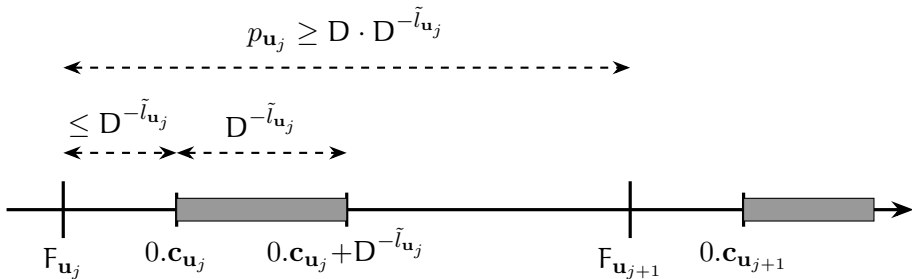
$U$	Code (i)	Code (ii)	Code (iii)	Code (iv)
$a$	0	0	10	0
$b$	0	010	00	10
$c$	1	01	11	110
$d$	1	10	110	111

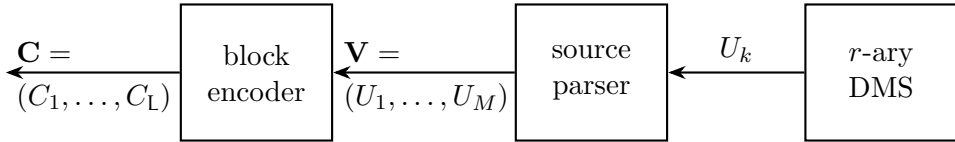
---

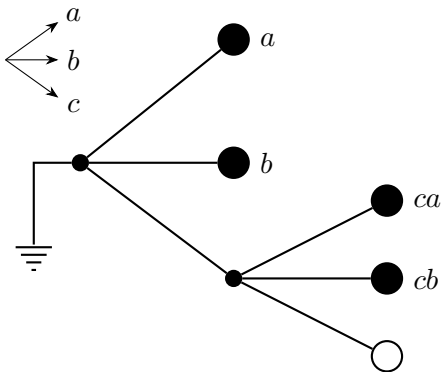
---



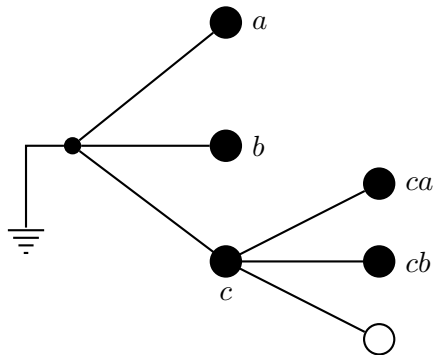




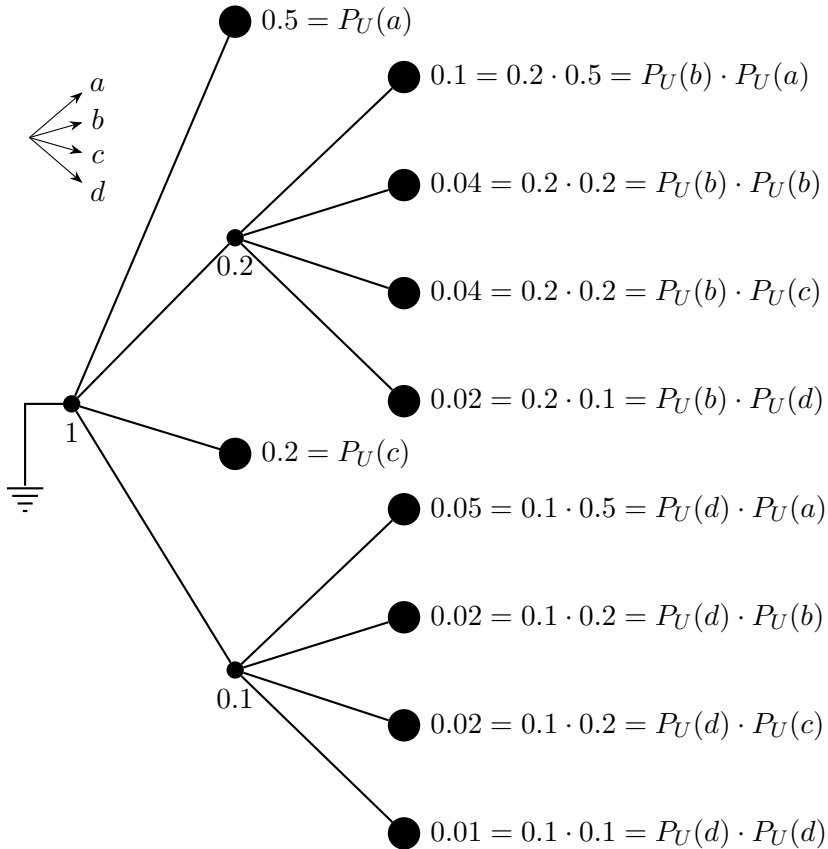


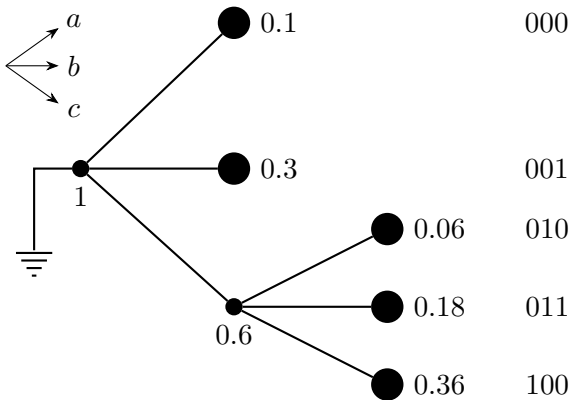


“illegal”

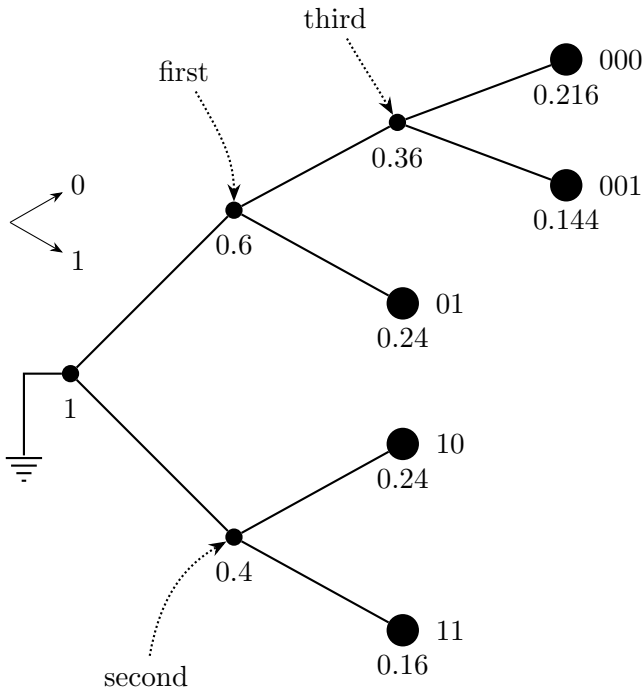


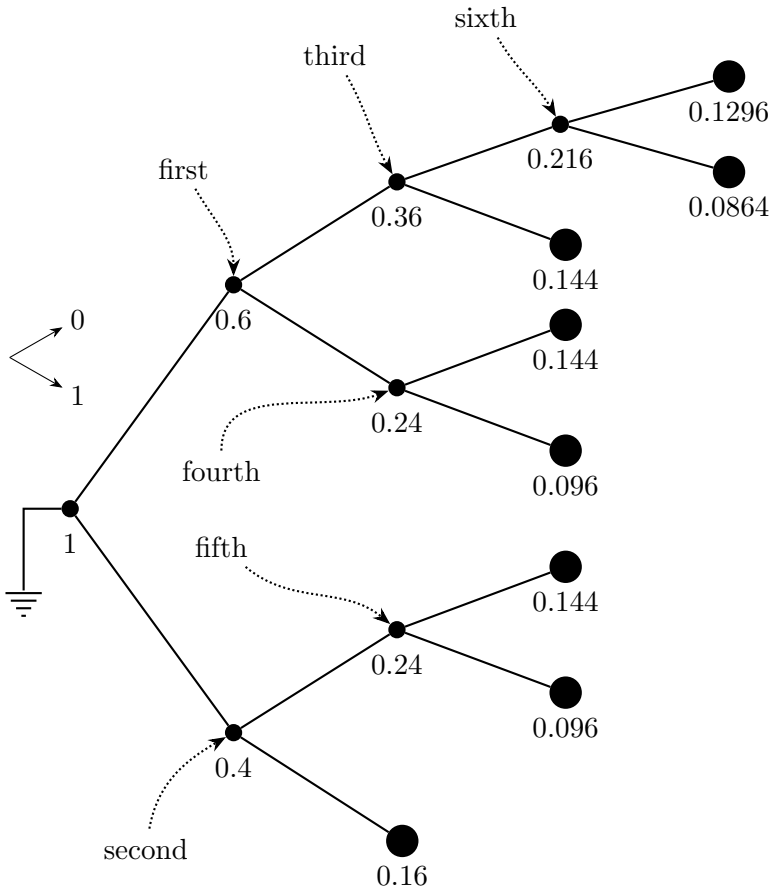
“legal”, but not proper





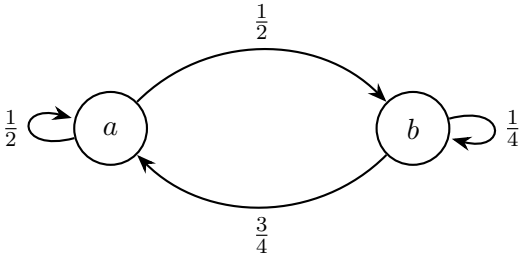


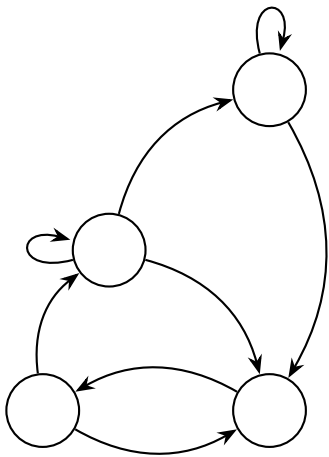




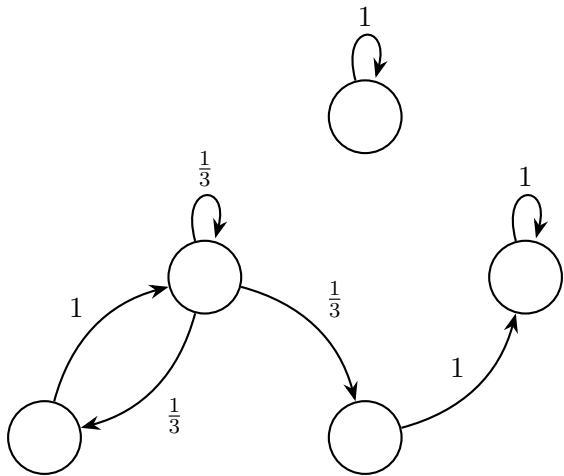
Message	Codewords
0000	000
0001	001
001	010
010	011
011	100
100	101
101	110
11	111



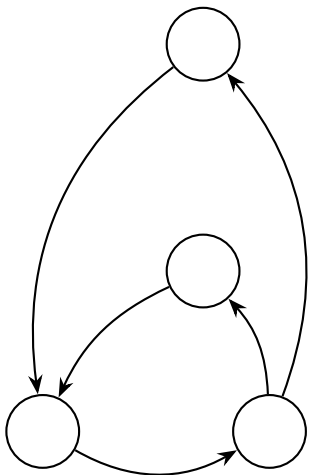




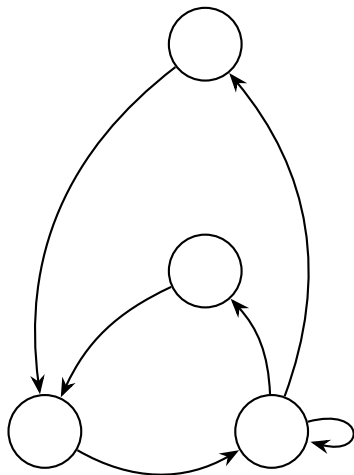
irreducible



reducible



periodic



aperiodic

---



---

$P_{X_k}(\cdot)$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
$\Pr[X_k = a]$	1	$\frac{1}{2} = 0.5$	$\frac{5}{8} = 0.625$	$\frac{19}{32} = 0.59375$
$\Pr[X_k = b]$	0	$\frac{1}{2} = 0.5$	$\frac{3}{8} = 0.375$	$\frac{13}{32} = 0.40625$

---



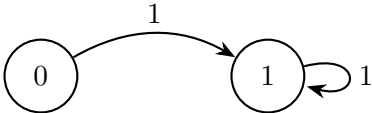
---

$P_{X_k}(\cdot)$	$k = 5$	$\dots$	$k = \infty$
$\Pr[X_k = a]$	$\frac{77}{128} = 0.6015625$	$\dots$	$\frac{3}{5} = 0.6$
$\Pr[X_k = b]$	$\frac{51}{128} = 0.3984375$	$\dots$	$\frac{2}{5} = 0.4$

---

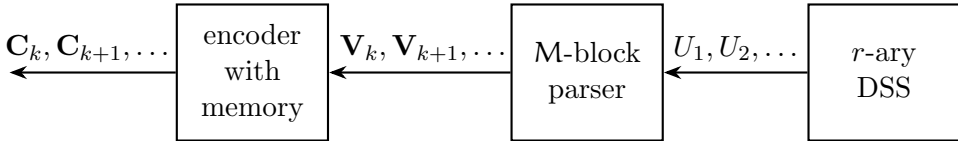


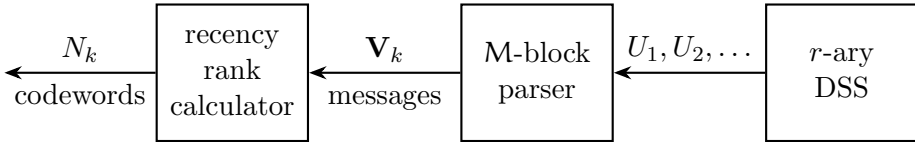
---



Stefan M. Moser, *Information Theory*, Version 5.3.







---

---

Message	Recency Rank at Time $k$
---------	--------------------------

---

00	4
----	---

01	1
----	---

10	3
----	---

11	2
----	---

---

---

---

---

Message	Recency Rank at Time $k + 1$
---------	------------------------------

---

00

4

01

2

10

1

11

3

---

---

---

Message	Recency Rank at Time 1
	(by definition)

---

00	1
----	---

01	2
----	---

10	3
----	---

11	4
----	---

---

---

---



---

$n$	1	2	3	4	5	6	7	8	9	10	$\dots$
$B_0(n)$	1	10	11	100	101	110	111	1000	1001	1010	$\dots$

---

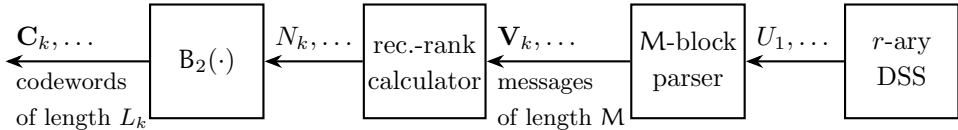


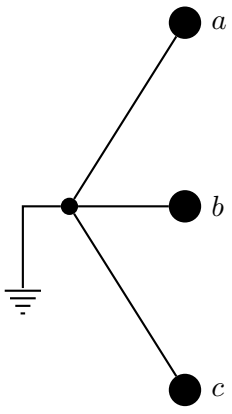
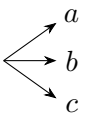
---

$n$	1	2	3	4	5	6
$B_0(n)$	1	10	11	100	101	110
$L_{B_0}(n) - 1$	0	1	1	2	2	2
$B_1(n)$	1	010	011	00100	00101	00110
$n$	7	8	9	10	11	...
$B_0(n)$	111	1000	1001	1010	1011	...
$L_{B_0}(n) - 1$	2	3	3	3	3	...
$B_1(n)$	00111	0001000	0001001	0001010	0001011	...

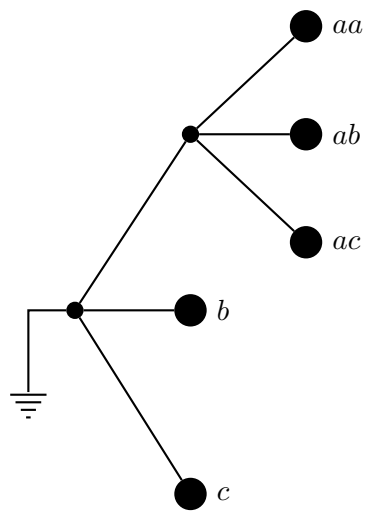
$n$	1	2	3	4	5	6
$B_0(n)$	1	10	11	100	101	110
$L_{B_0}(n)$	1	2	2	3	3	3
$B_1(L_{B_0}(n))$	1	010	010	011	011	011
$B_2(n)$	1	0100	0101	01100	01101	01110
$n$	7	8	9	10	11	...
$B_0(n)$	111	1000	1001	1010	1011	...
$L_{B_0}(n)$	3	4	4	4	4	...
$B_1(L_{B_0}(n))$	011	00100	00100	00100	00100	...
$B_2(n)$	01111	00100000	00100001	00100010	00100011	...



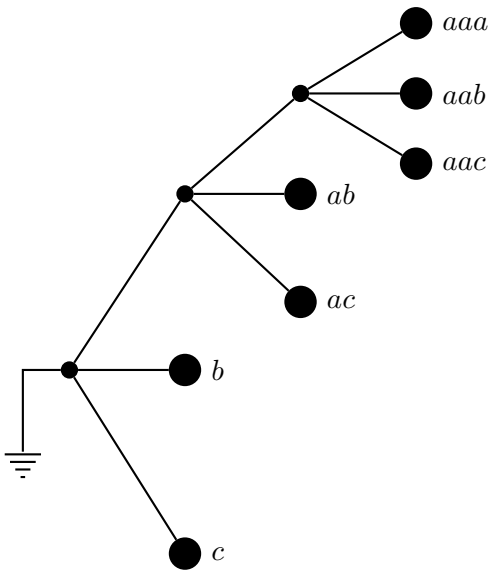




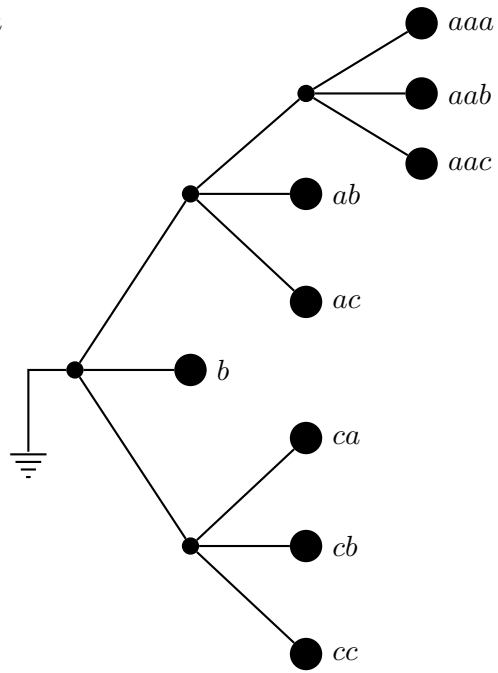
(a)



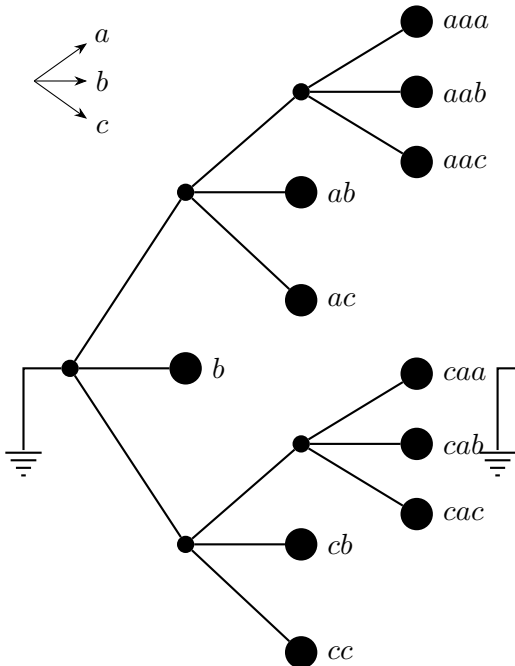
(b)



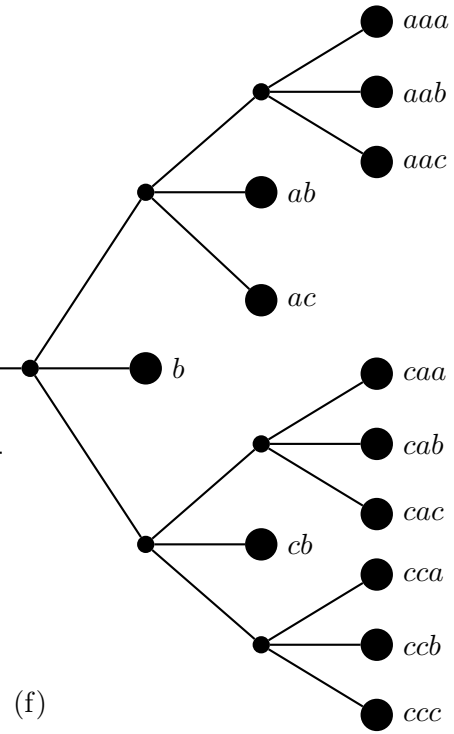
(c)



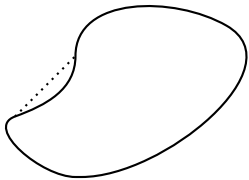
(d)



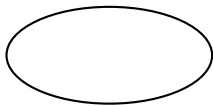
(e)



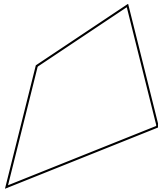
(f)



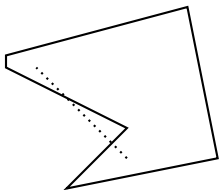
nonconvex



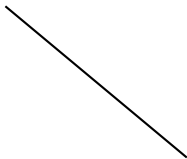
convex



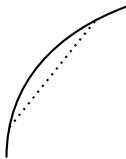
convex



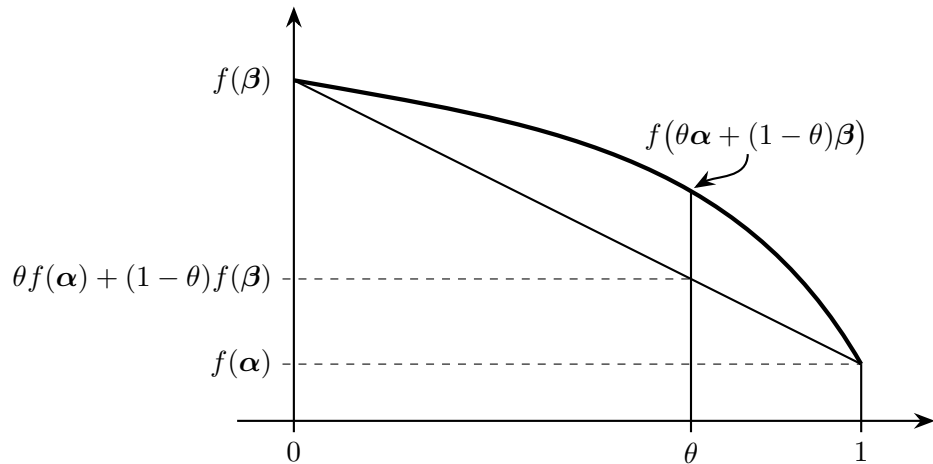
nonconvex

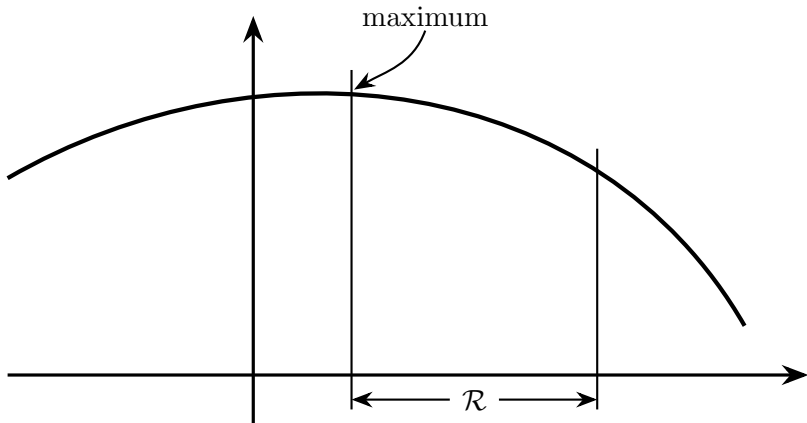


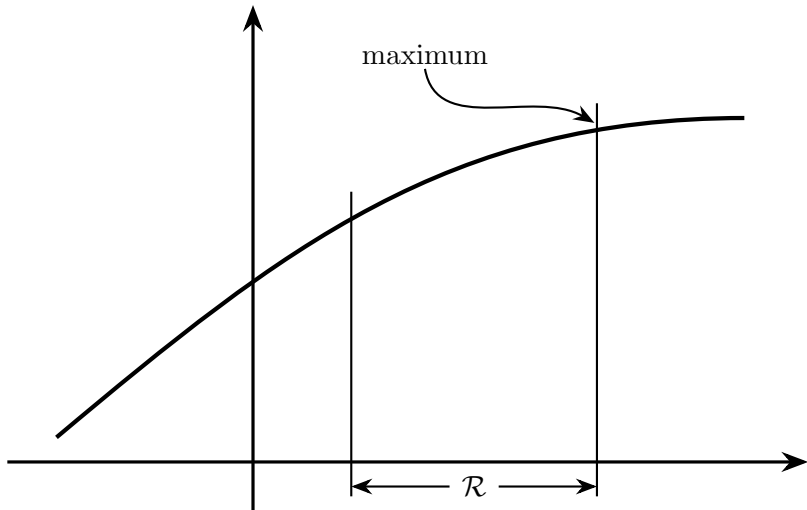
convex



nonconvex



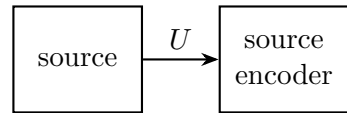




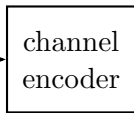
original name	2	3	1
new name ( $t$ )	1	2	3
$p_i$	0.2	0.1	0.7
$o_i$	6	10	1.2
$p_i o_i$	1.2	1.0	0.84
	$\vee$	$\vee$	$\wedge$
$\beta_t$	1	0.96	0.955



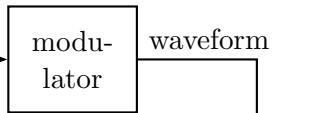
uniform memoryless source



$M$

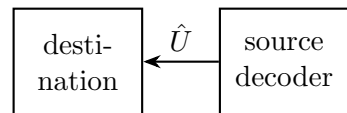


discrete-time channel

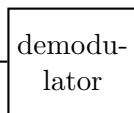
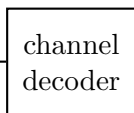


noise

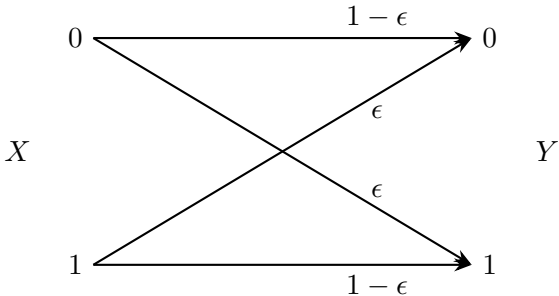
noisy  
channel

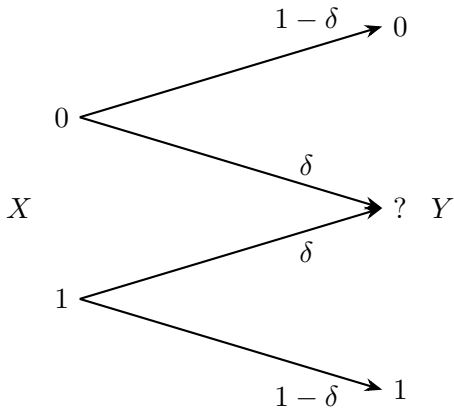


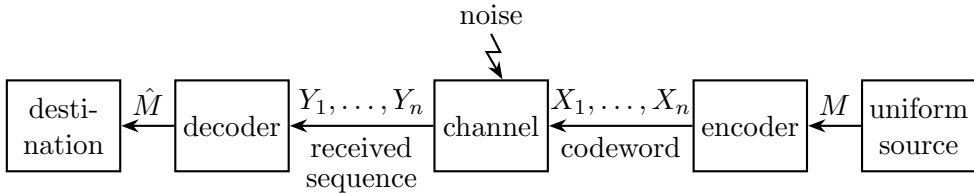
$\hat{M}$

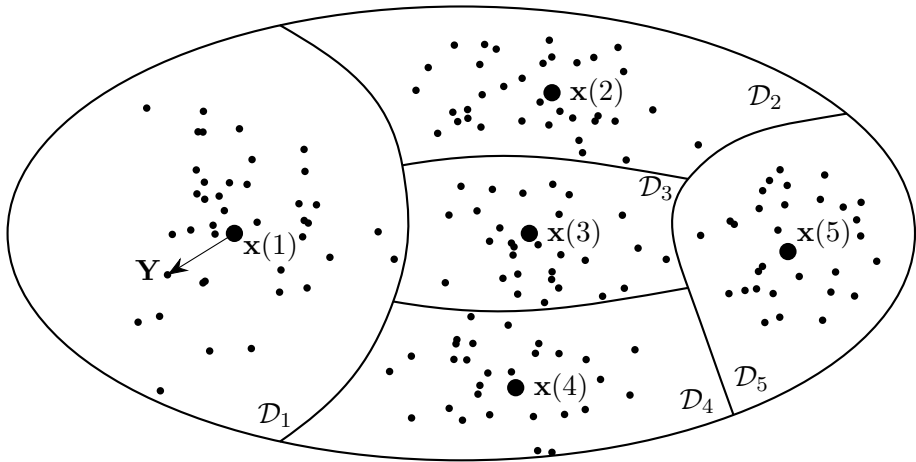


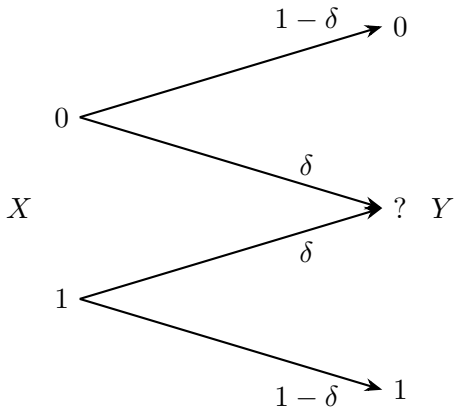
dest. of compressed data

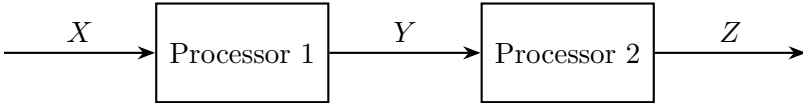












generate  $\mathbf{X}(1)$



send it  
through DMC



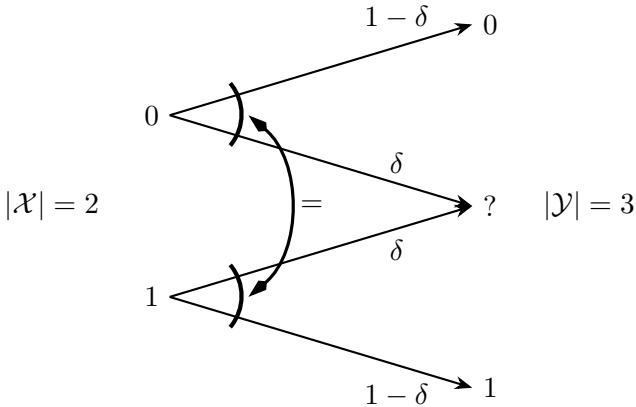
receive  $\mathbf{Y}$

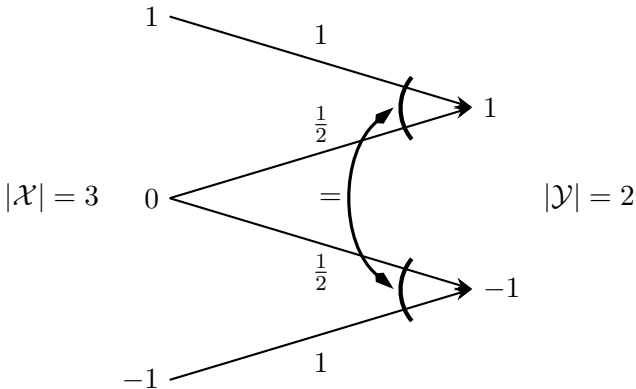
← independent →

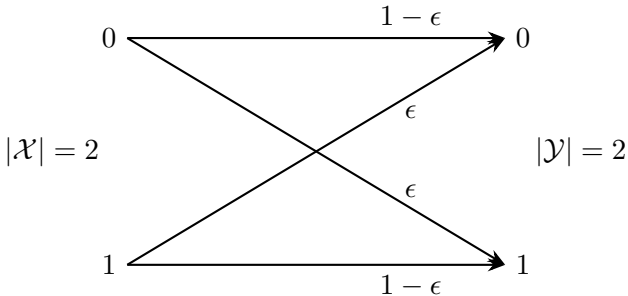
← independent →

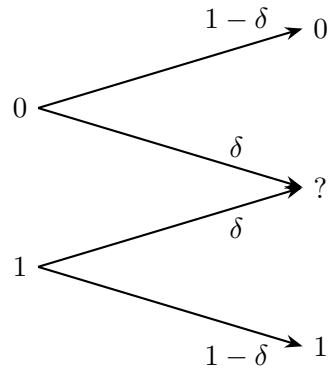
generate other  
codewords  $\mathbf{X}(m)$  using  
the same  $P_X(\cdot)$ ,  
but independently











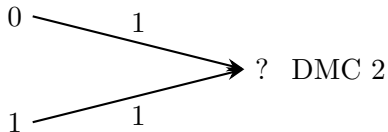
$$q_1 = 1 - \delta$$

$$q_2 = \delta$$

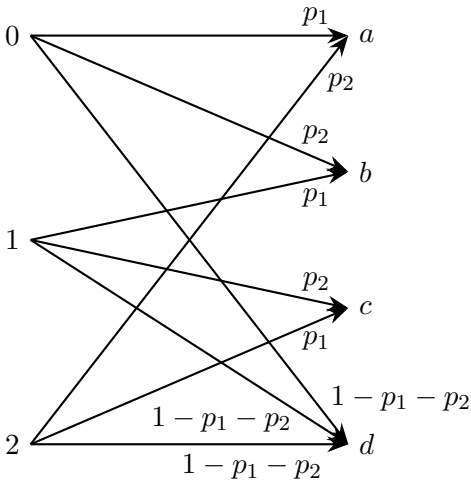
$$0 \xrightarrow{1} 0$$

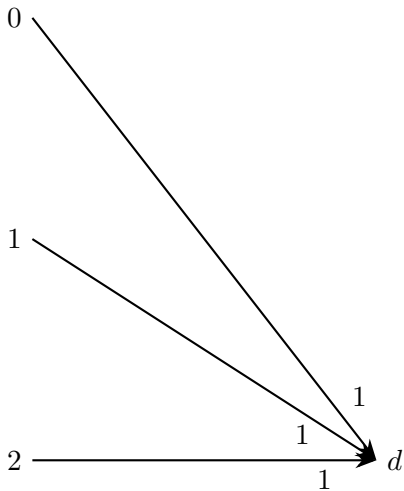
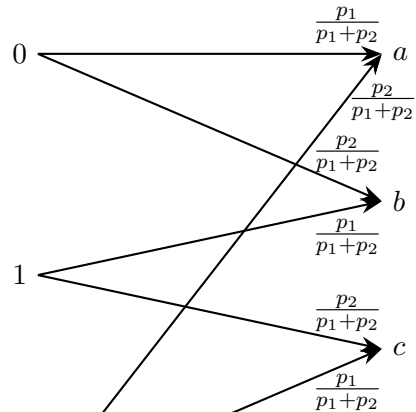
$$1 \xrightarrow{1} 1$$

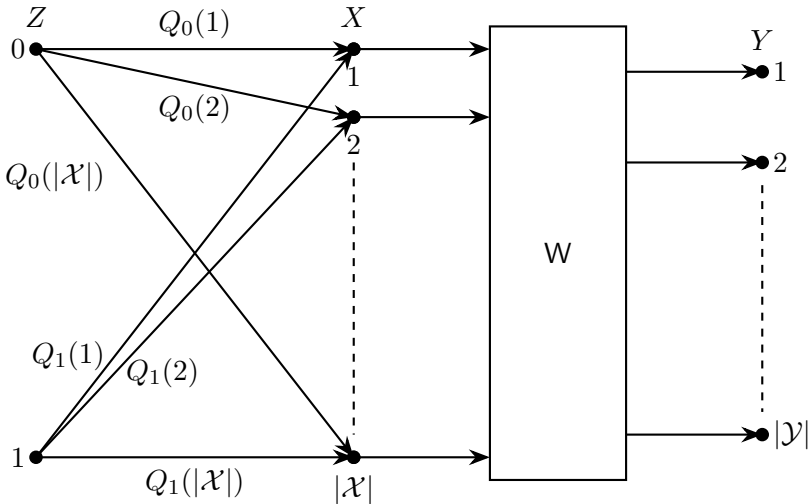
DMC 1



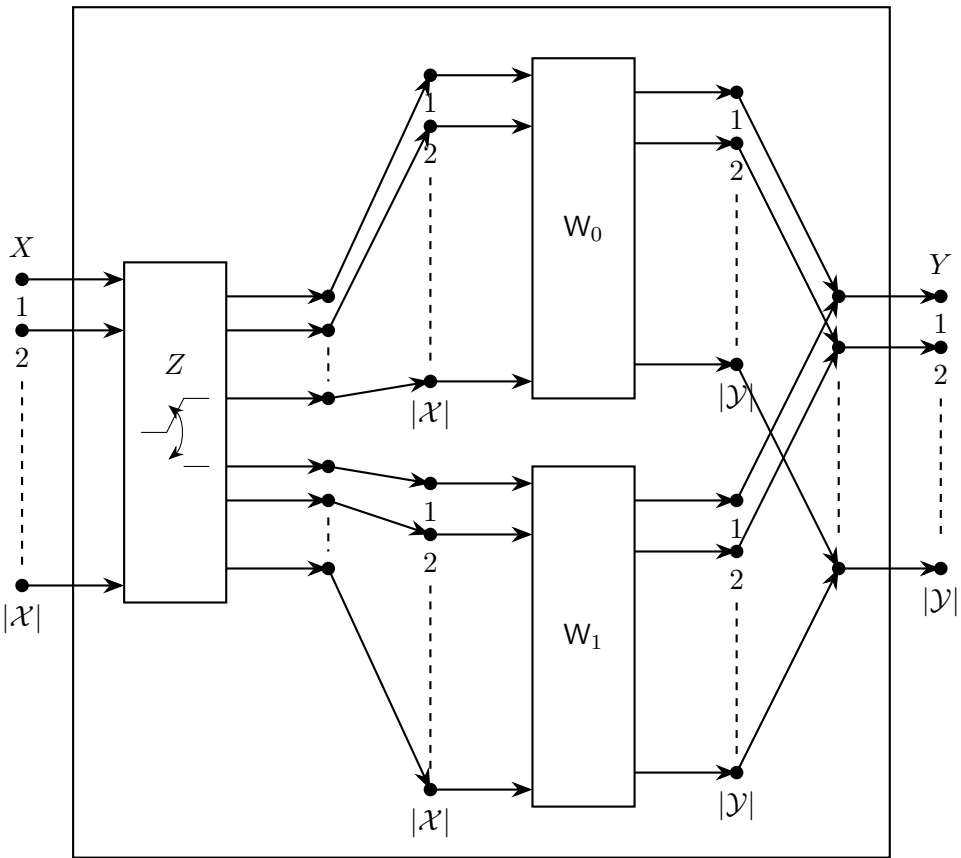
DMC 2



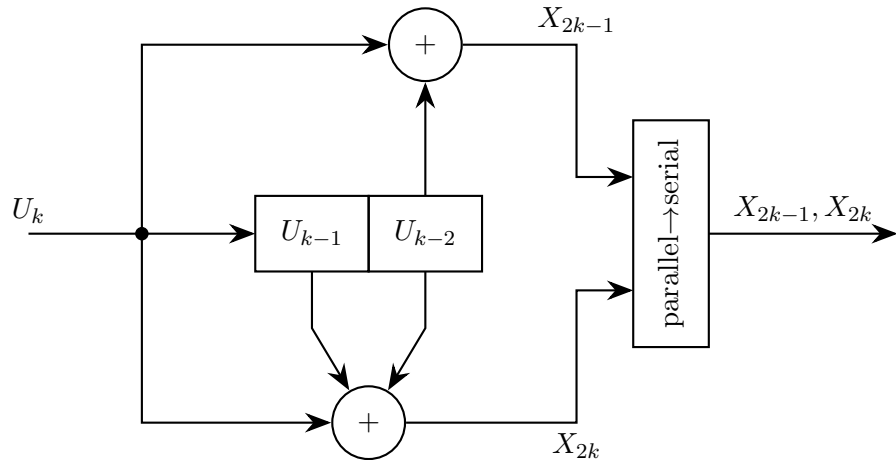


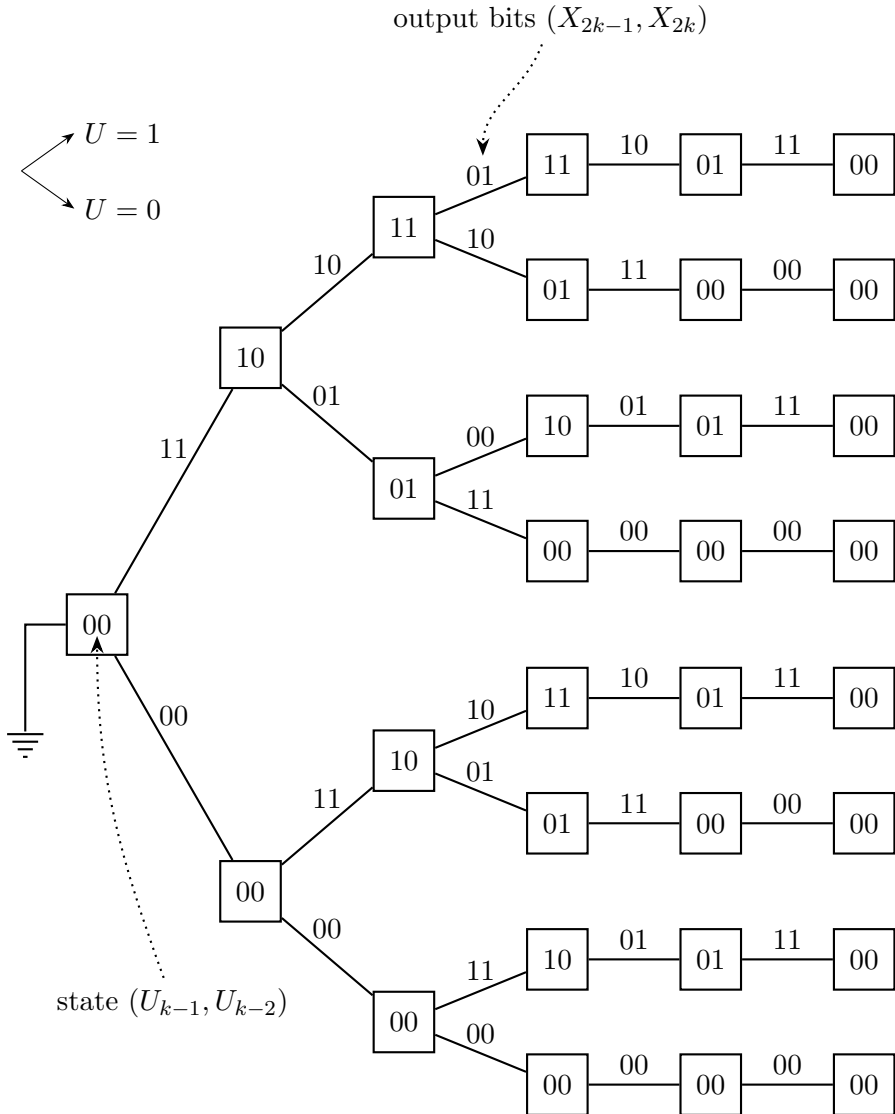


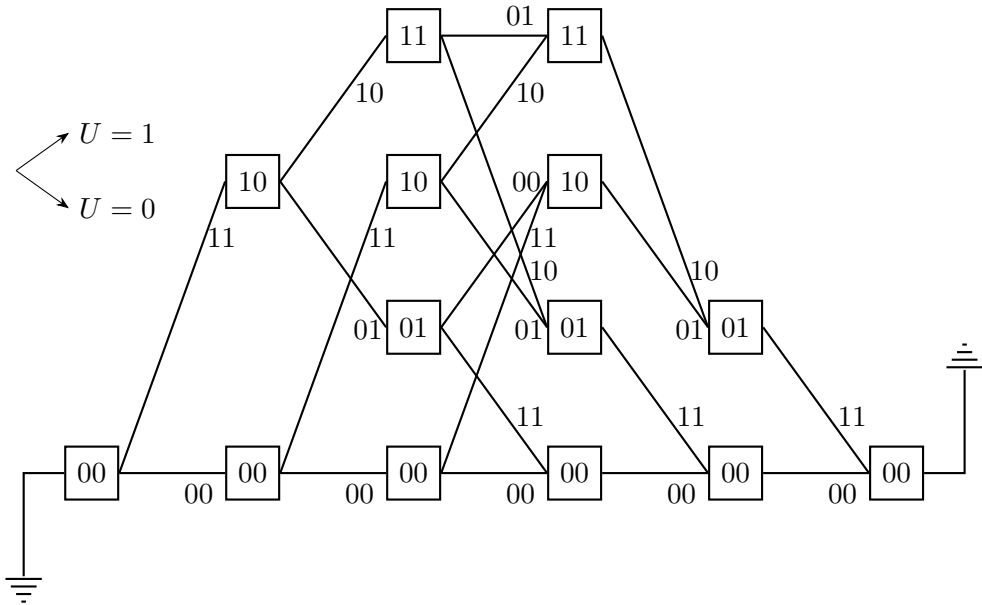
W

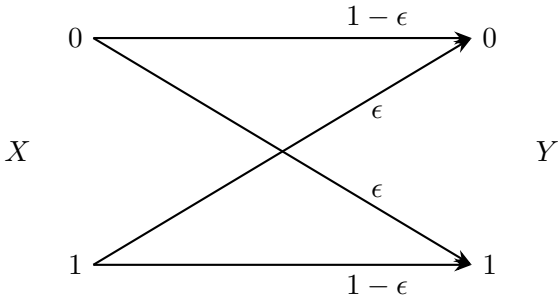


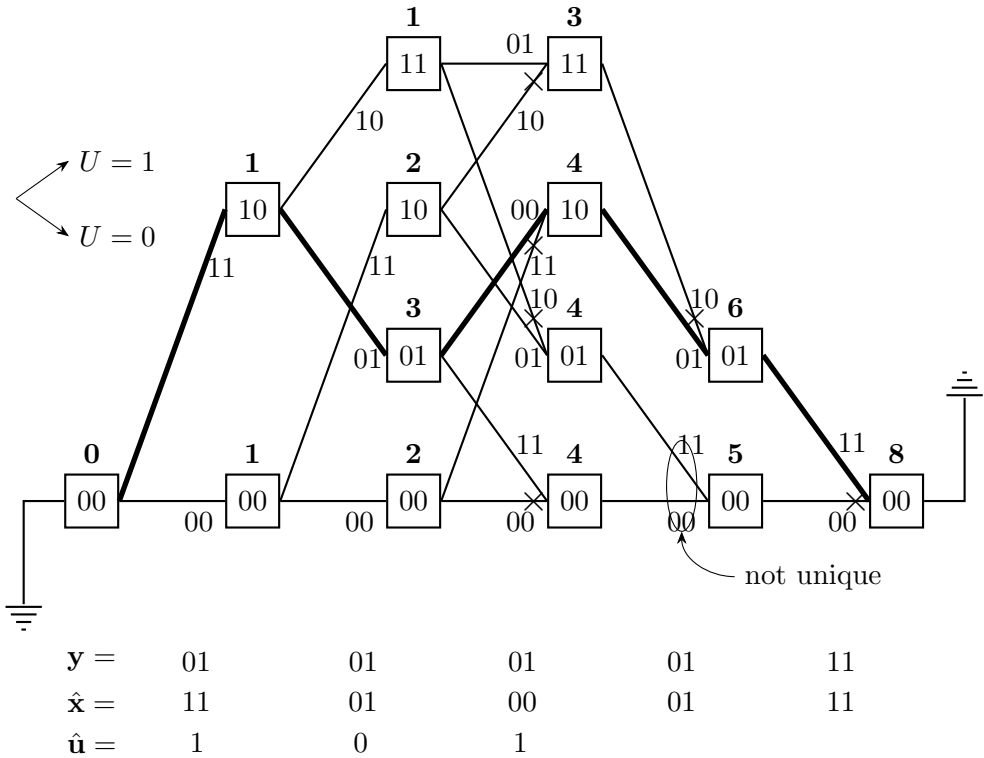


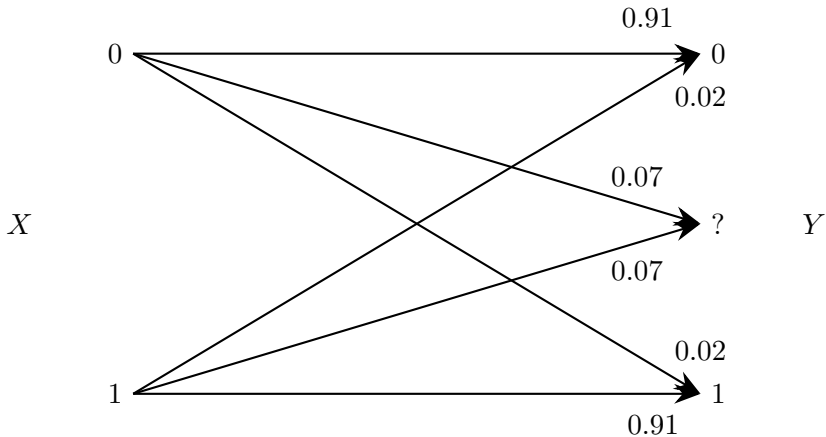












	$y = 0$	$y = ?$	$y = 1$
$x = 0$	$\log 0.91 - \log 0.02$ $\approx 5.50$	$\log 0.07 - \log 0.07$ $= 0$	$\log 0.02 - \log 0.02$ $= 0$
$x = 1$	$\log 0.02 - \log 0.02$ $= 0$	$\log 0.07 - \log 0.07$ $= 0$	$\log 0.91 - \log 0.02$ $\approx 5.50$

---

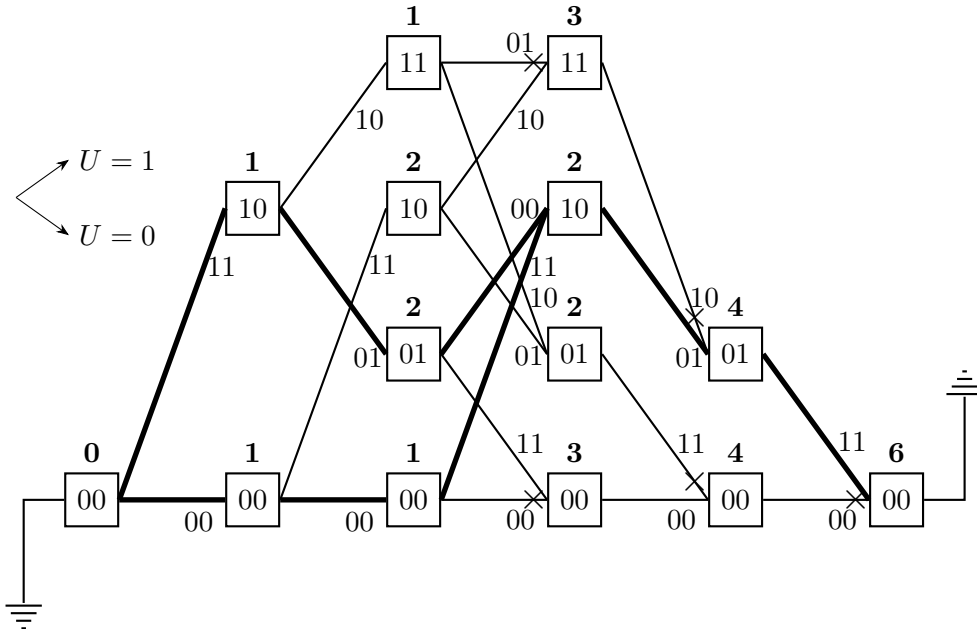
---

	$y = 0$	$y = ?$	$y = 1$
$x = 0$	1	0	0
$x = 1$	0	0	1

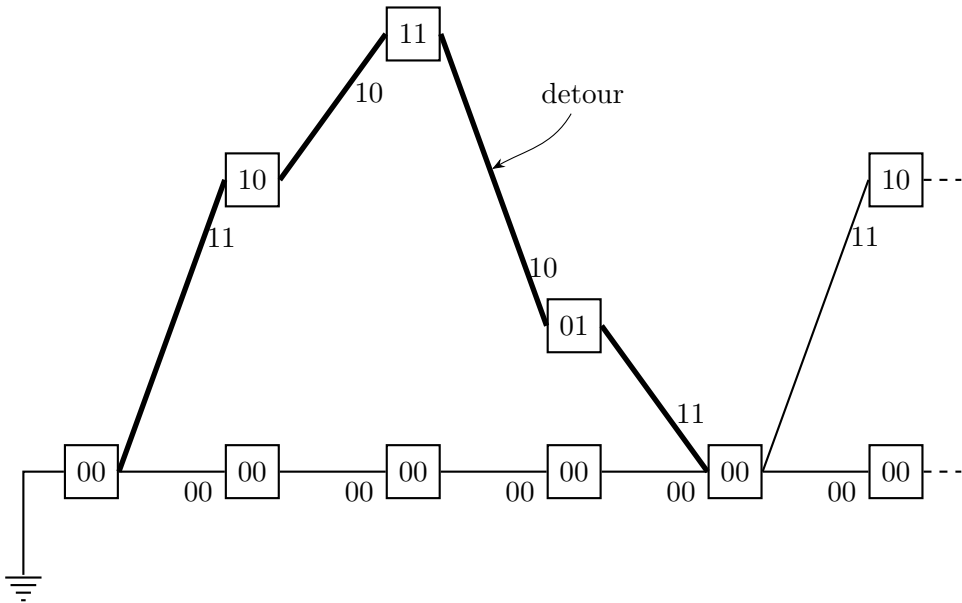
---

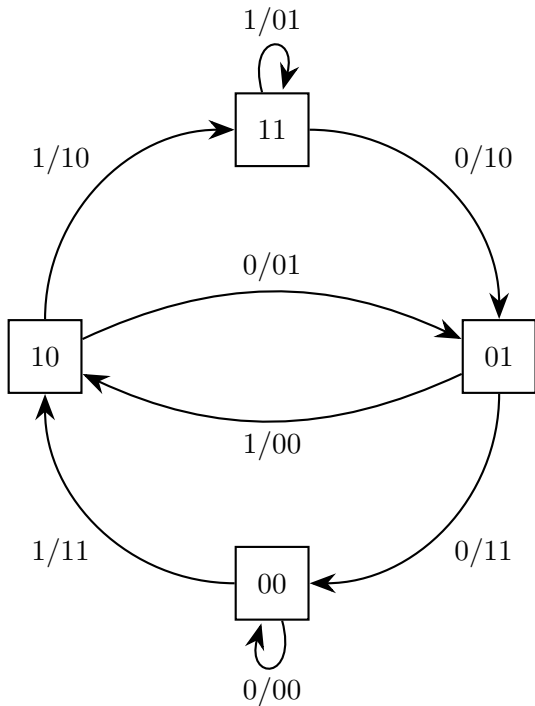
---

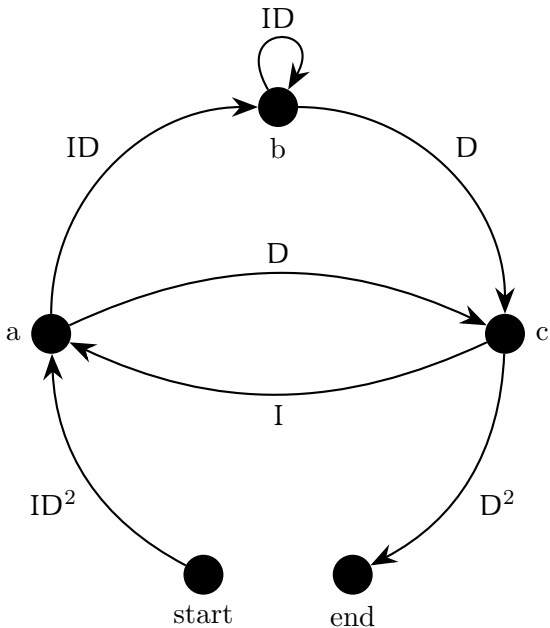


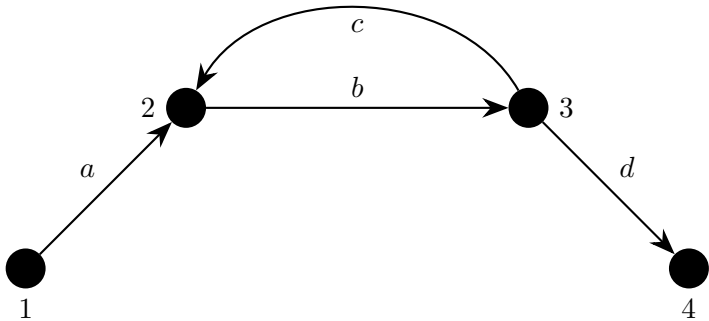


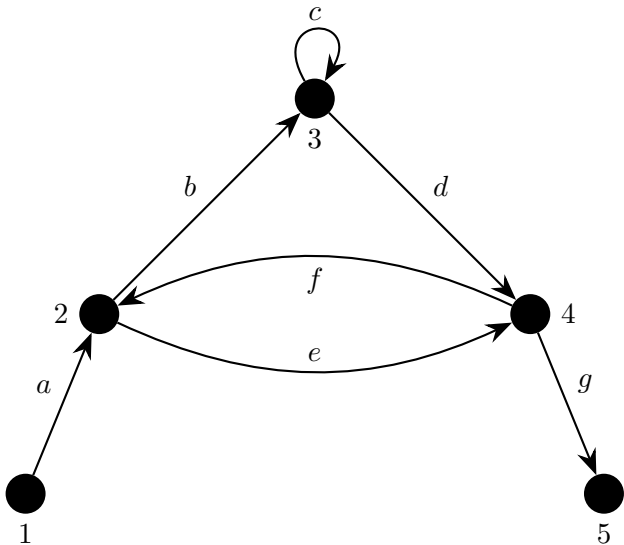
$\mathbf{y} =$	01	?1	1?	01	11
$\hat{\mathbf{x}} =$	11	01	00	01	11
	00	00	11	01	11
$\hat{\mathbf{u}} =$	1	0	1		
	0	0	1		

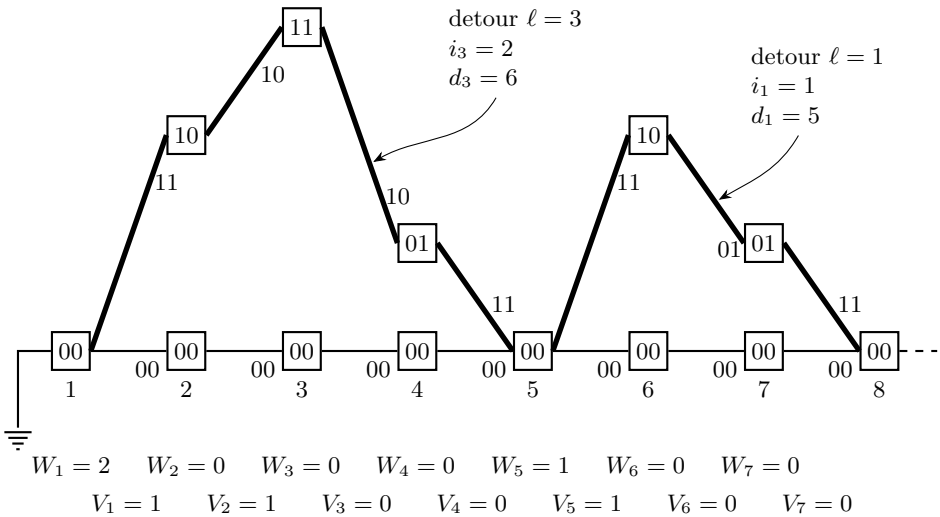


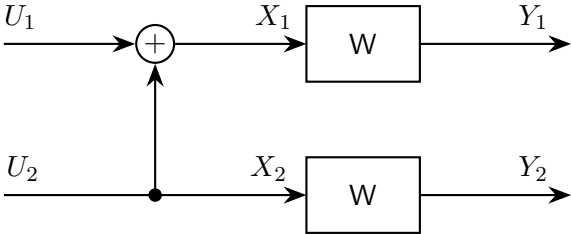




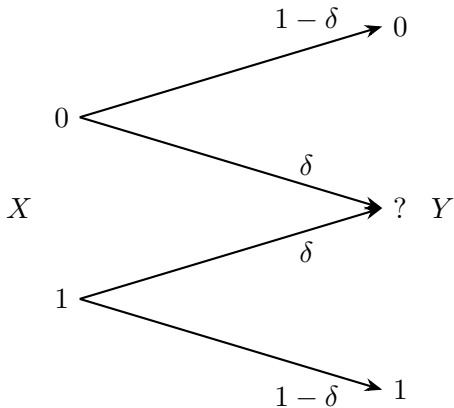




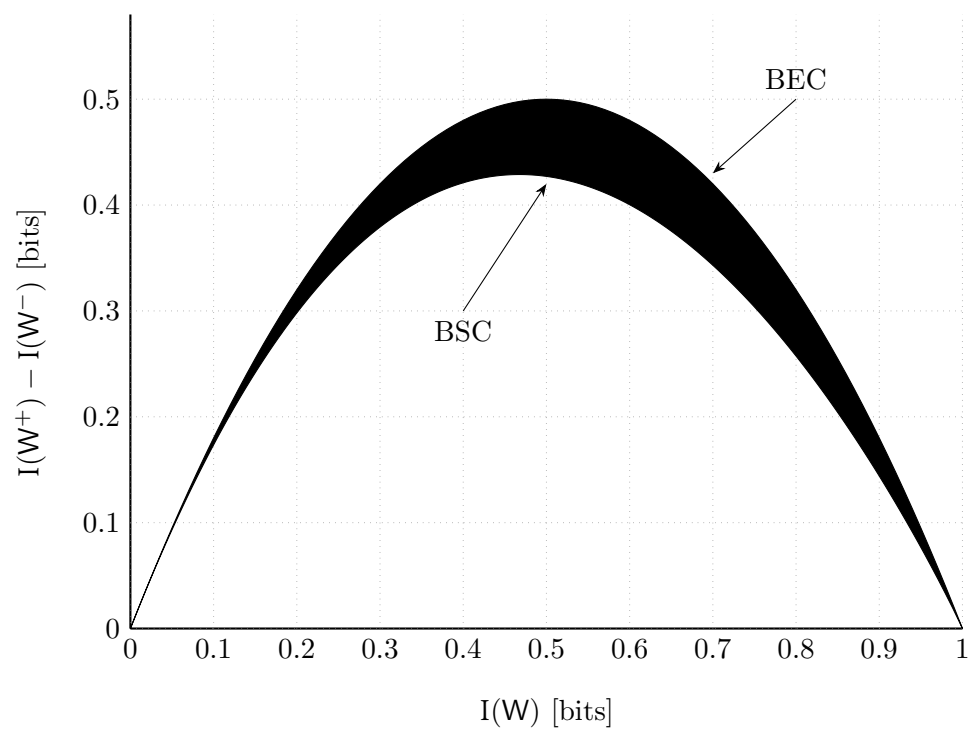


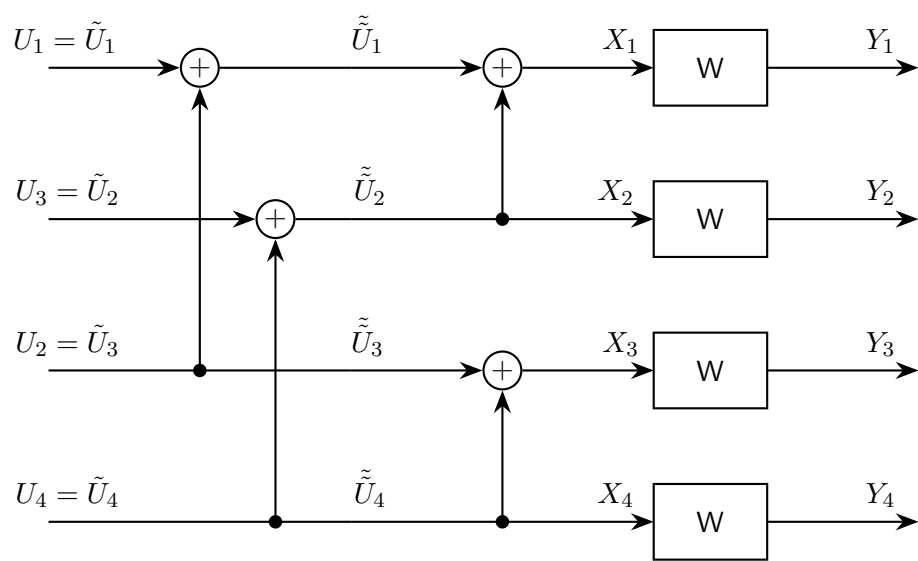


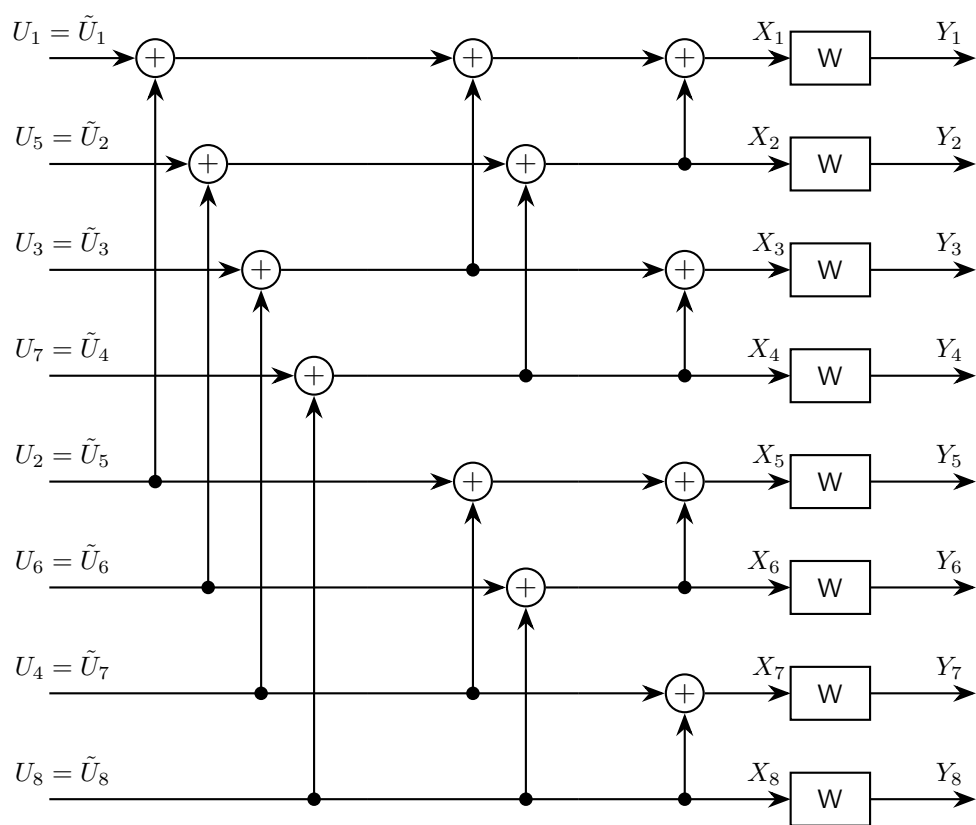


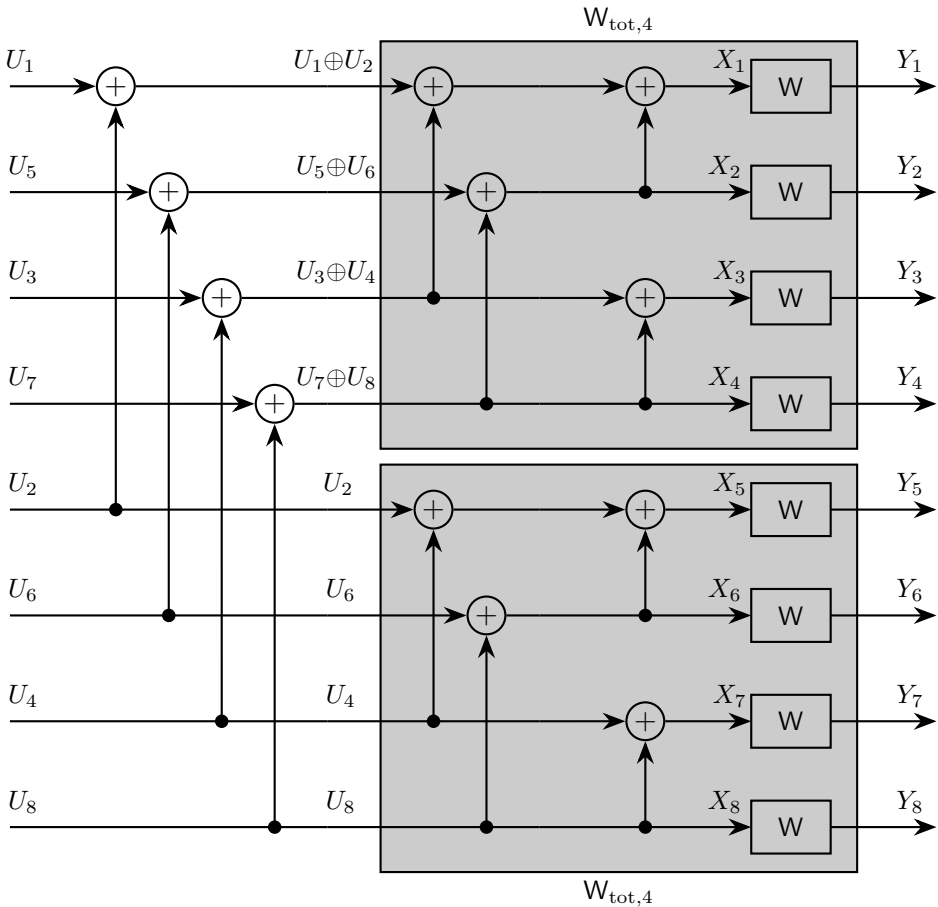


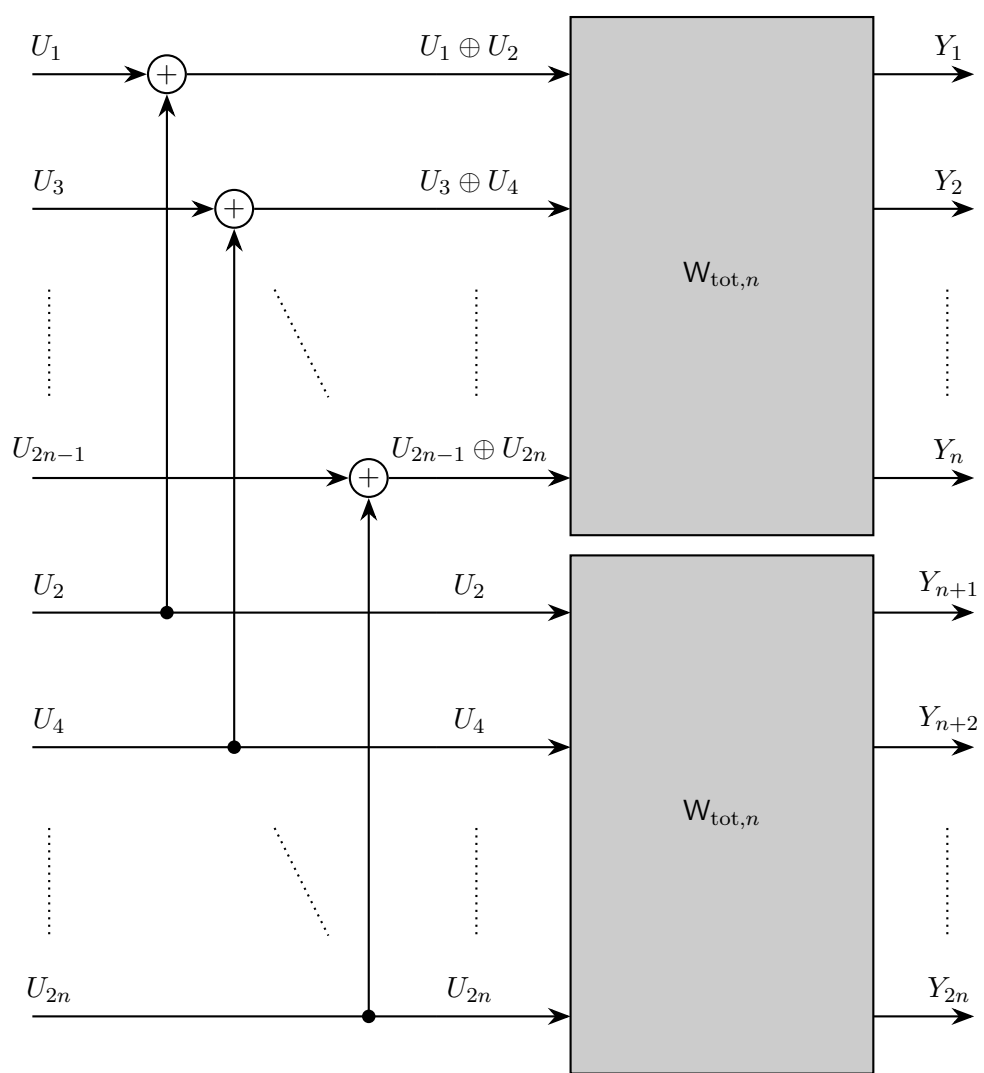
$y_1$	$y_2$	$W^-(y_1, y_2 0)$	$W^-(y_1, y_2 1)$
0	0	$(1 - \delta)^2/2$	0
0	?	$\delta(1 - \delta)/2$	$\delta(1 - \delta)/2$
0	1	0	$(1 - \delta)^2/2$
?	0	$\delta(1 - \delta)/2$	$\delta(1 - \delta)/2$
?	?	$\delta^2$	$\delta^2$
?	1	$\delta(1 - \delta)/2$	$\delta(1 - \delta)/2$
1	0	0	$(1 - \delta)^2/2$
1	?	$\delta(1 - \delta)/2$	$\delta(1 - \delta)/2$
1	1	$(1 - \delta)^2/2$	0

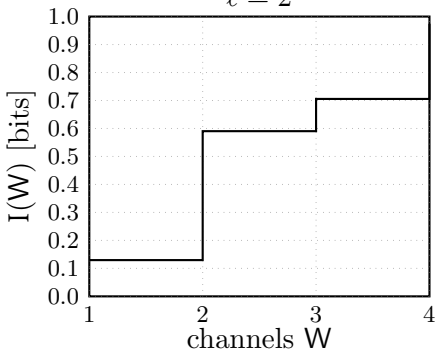
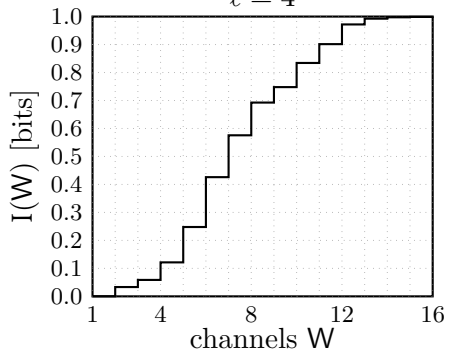
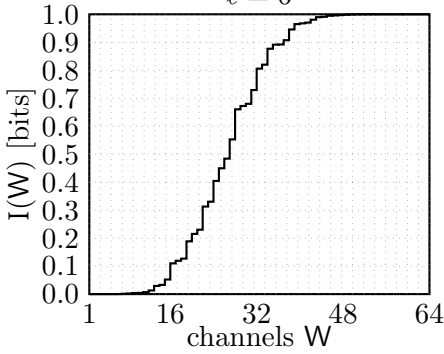
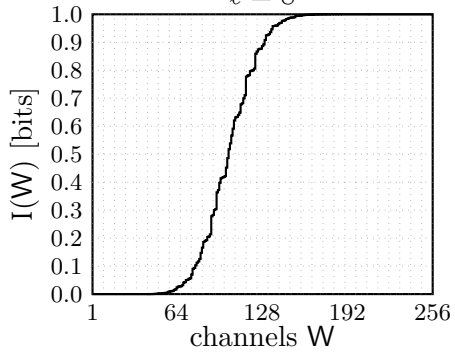
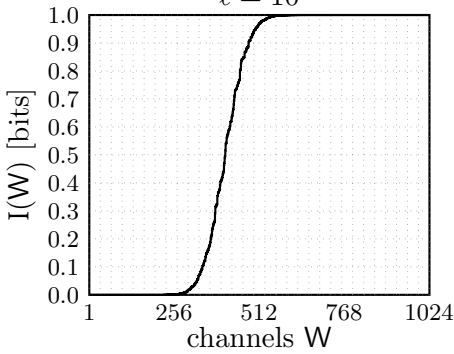
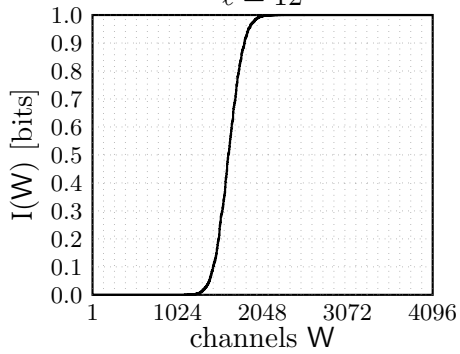
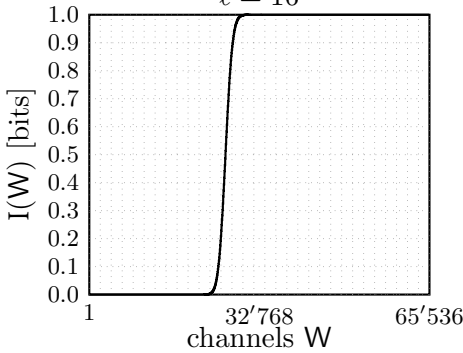
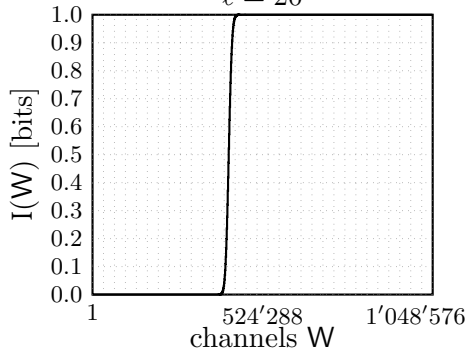




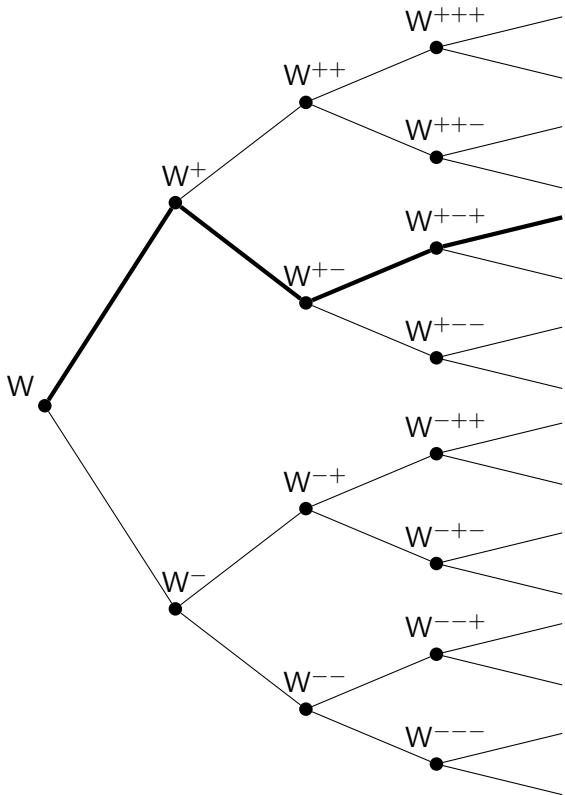


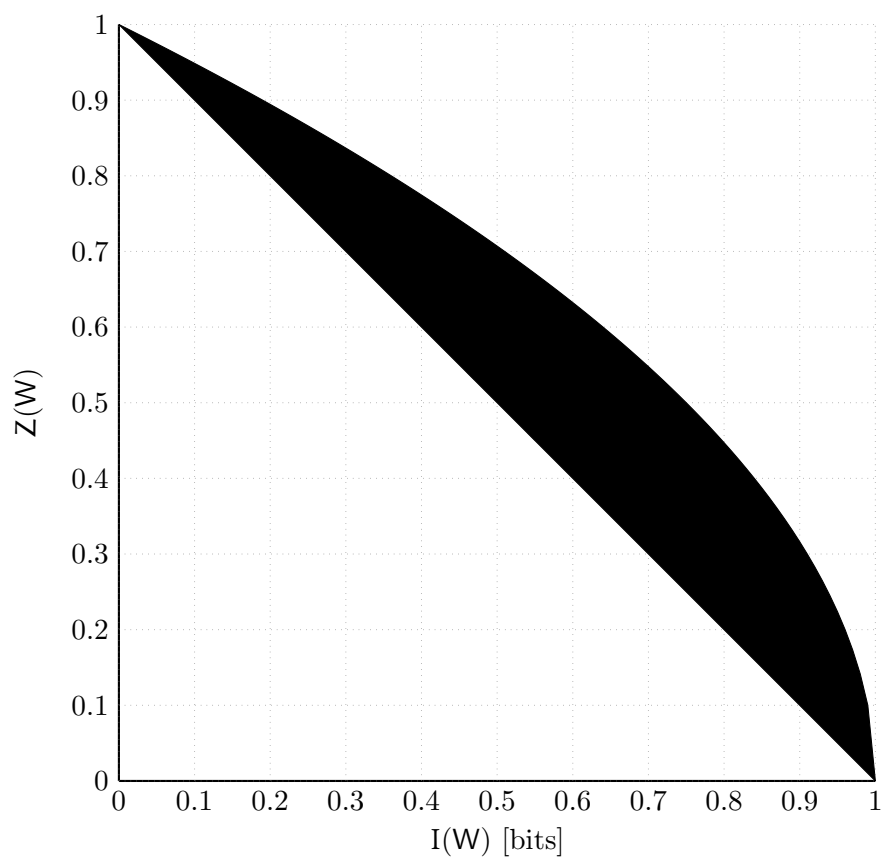


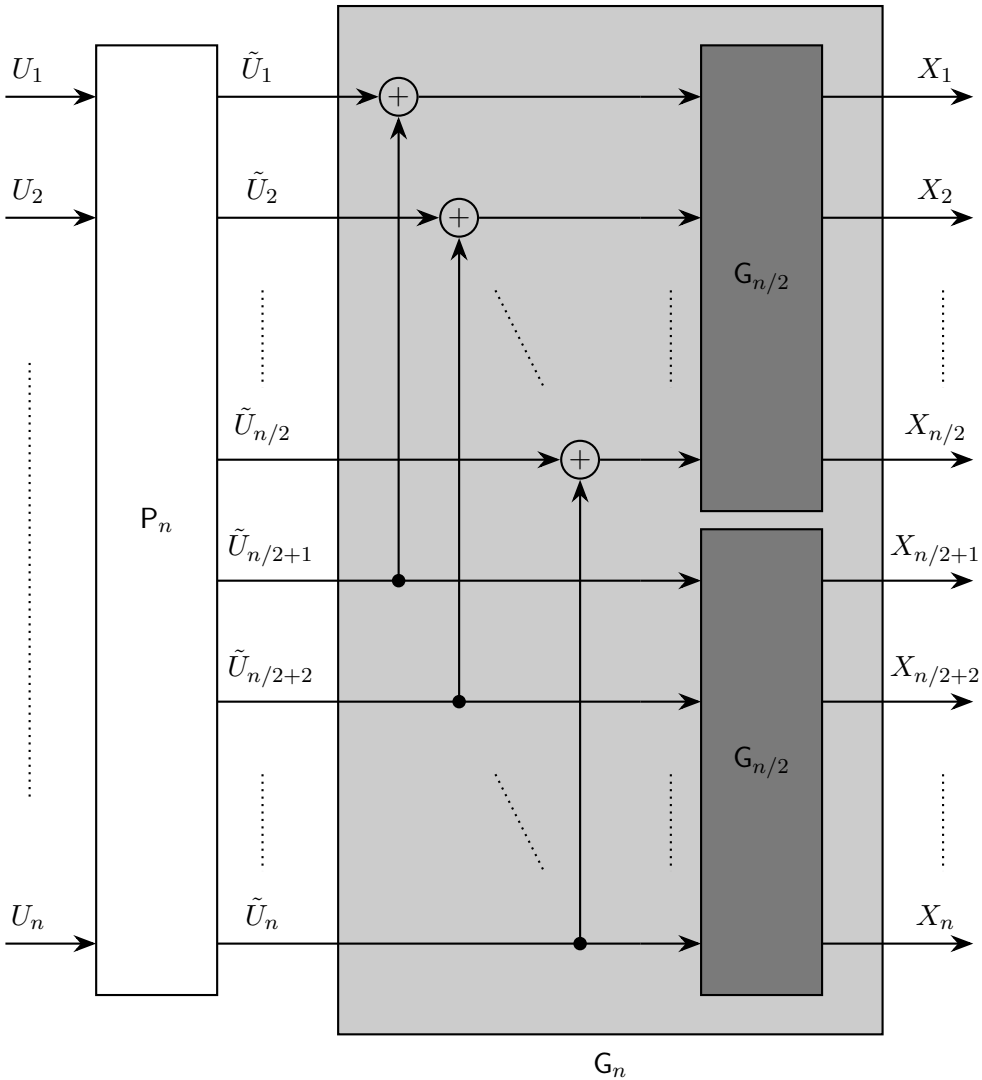


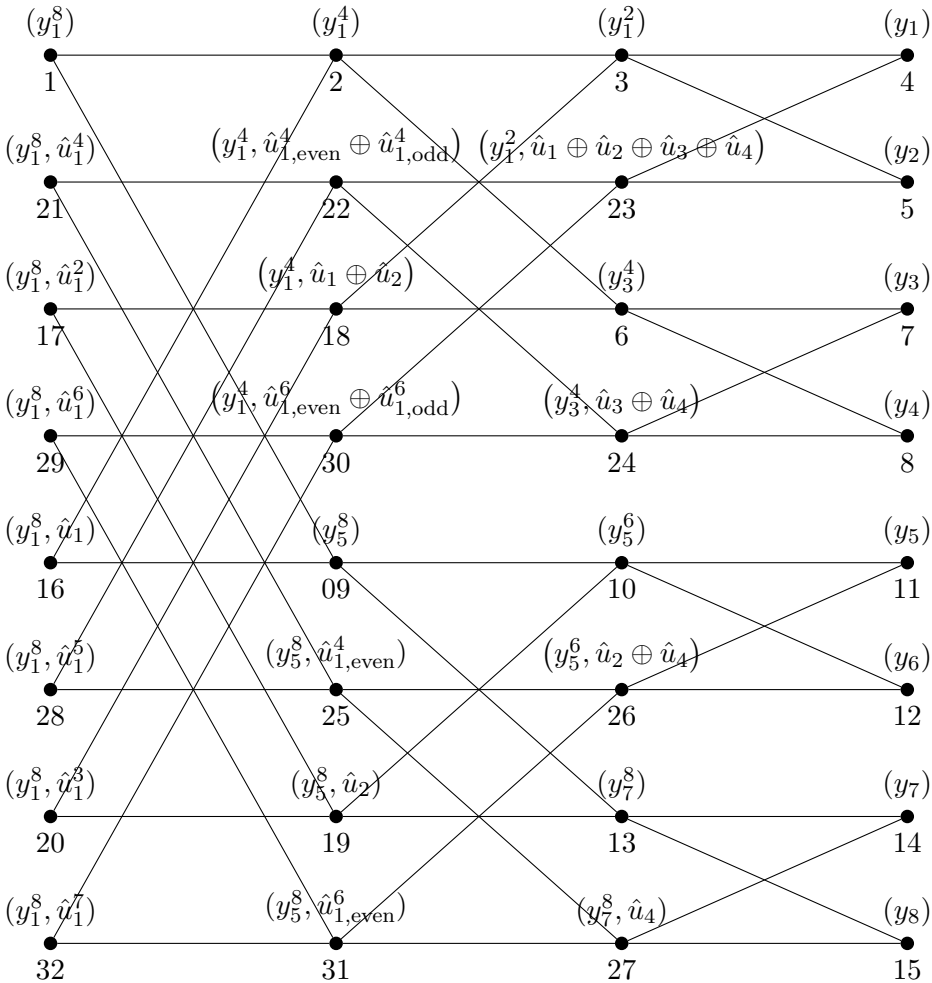
$\ell = 2$  $\ell = 4$  $\ell = 6$  $\ell = 8$  $\ell = 10$  $\ell = 12$  $\ell = 16$  $\ell = 20$ 

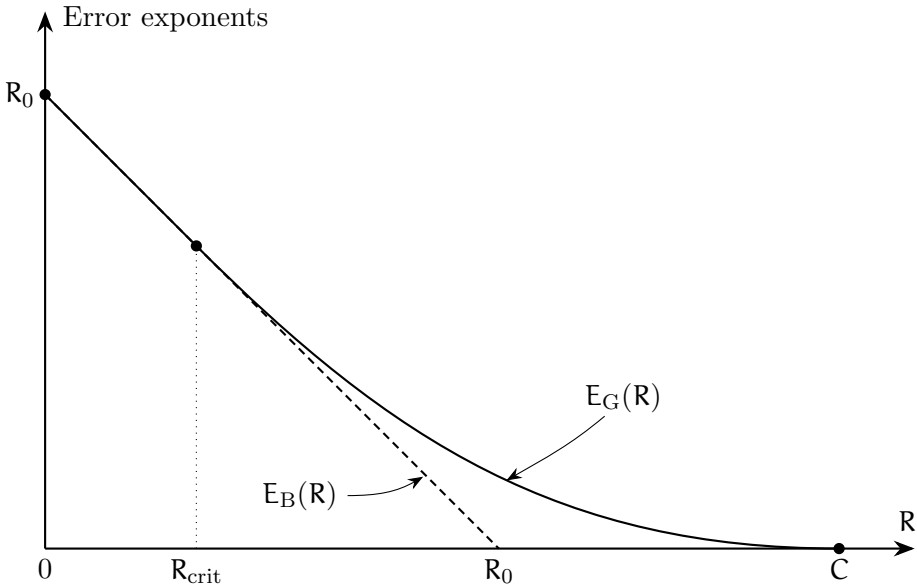


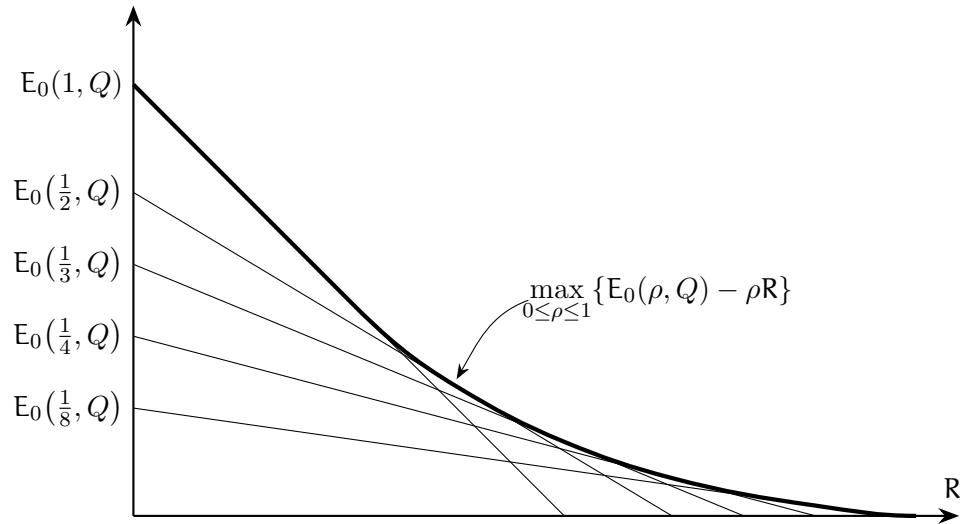


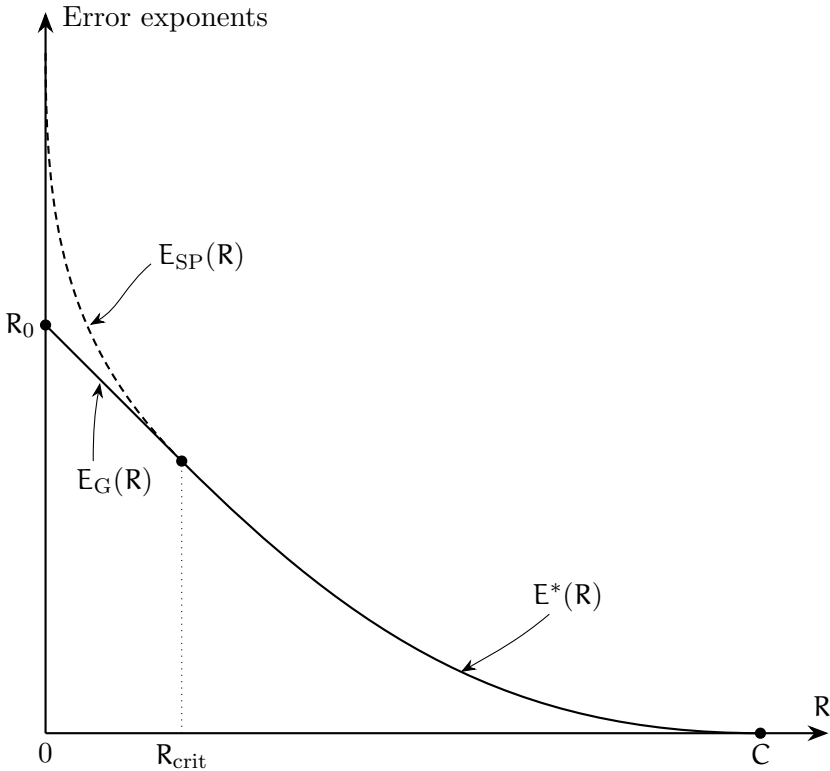


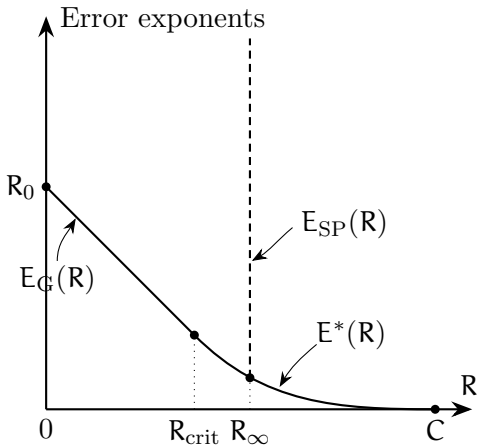
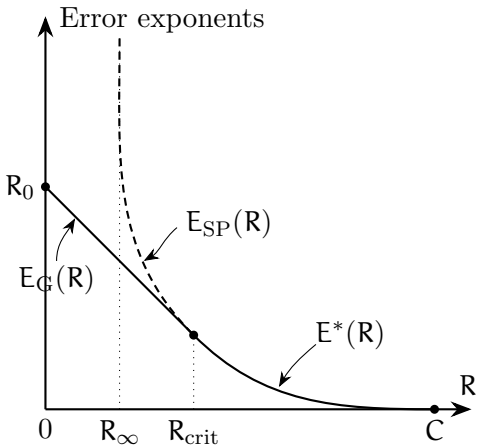




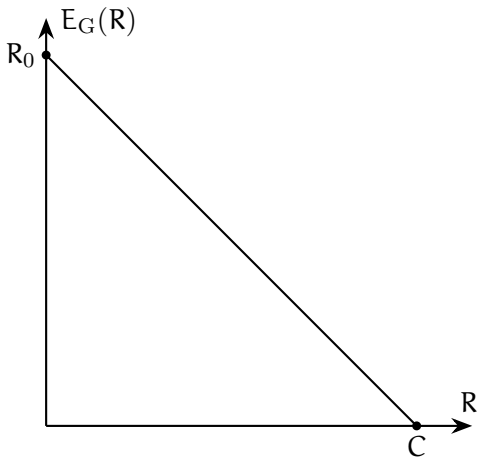
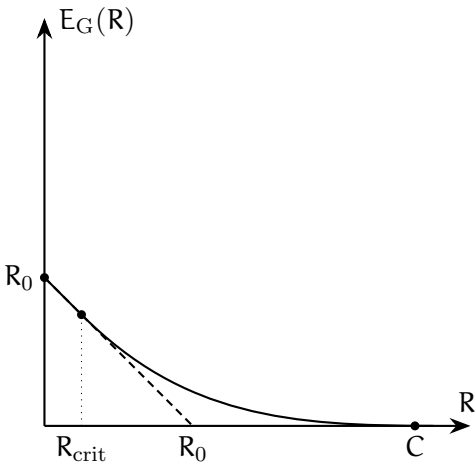




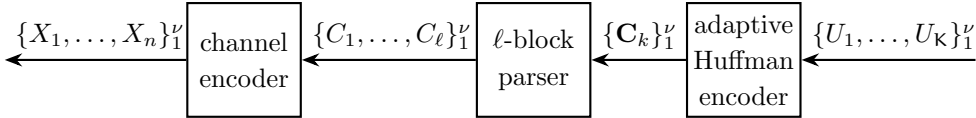


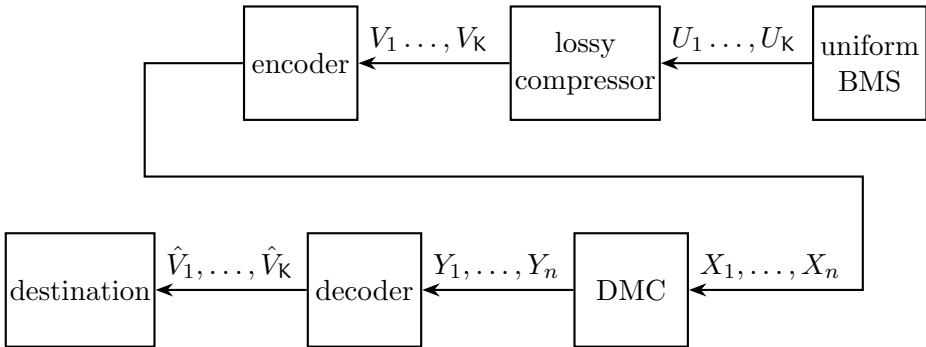


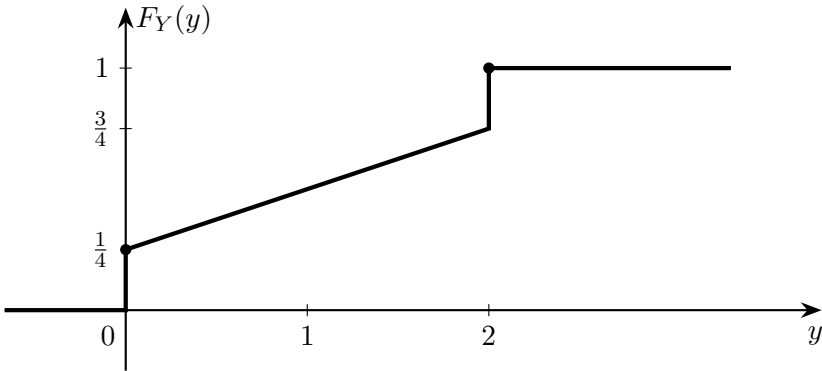


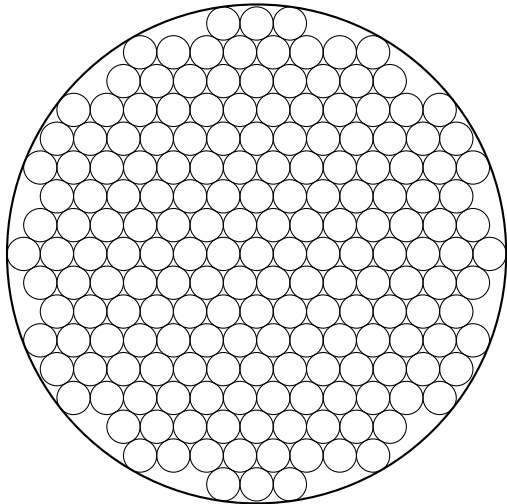




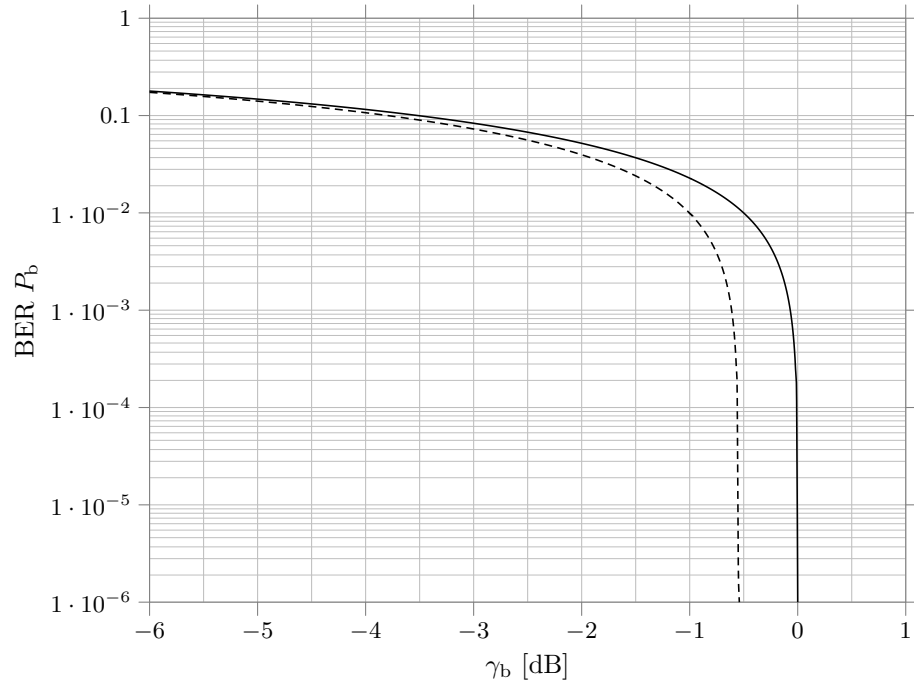






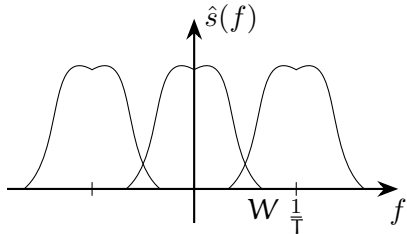
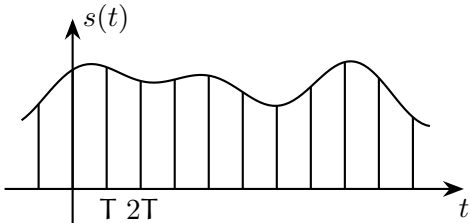


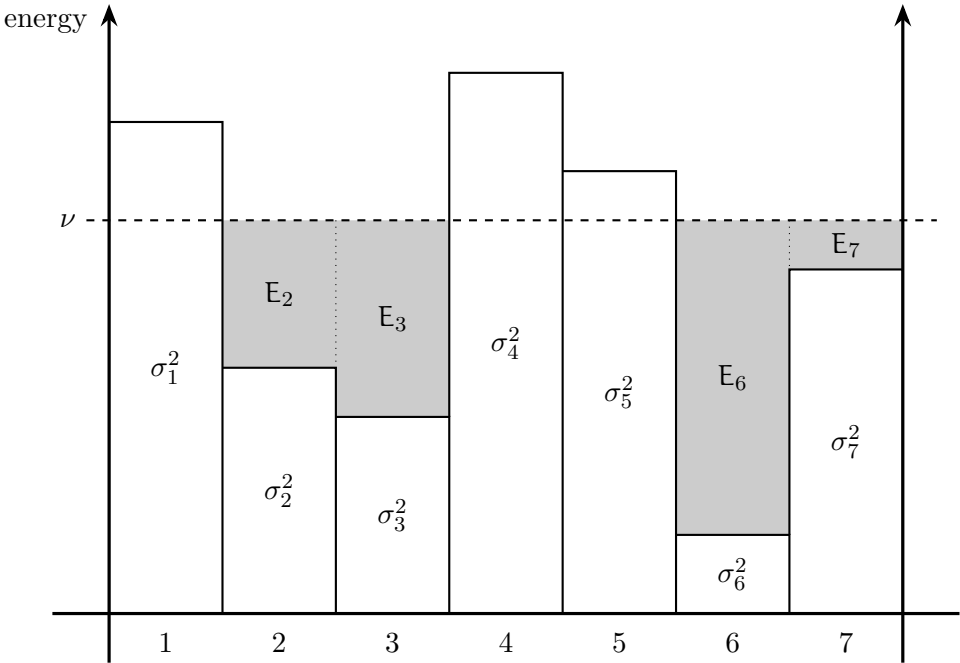


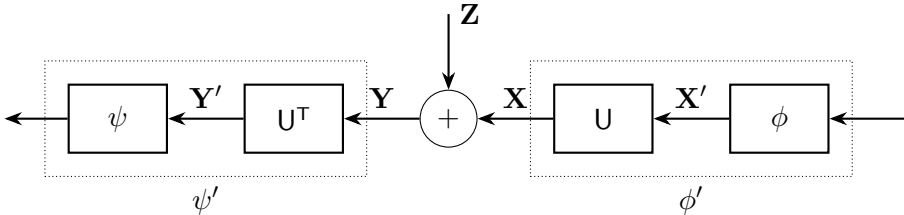


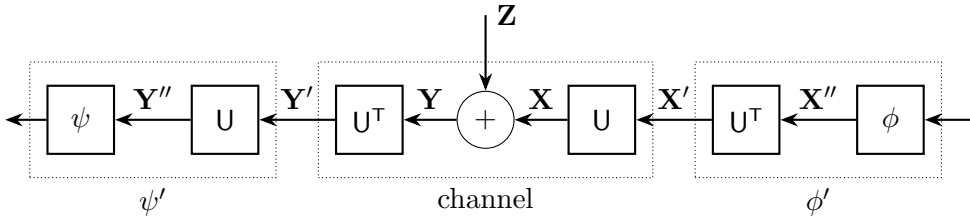
—  $R = 1/2$     - - -  $R = 1/3$

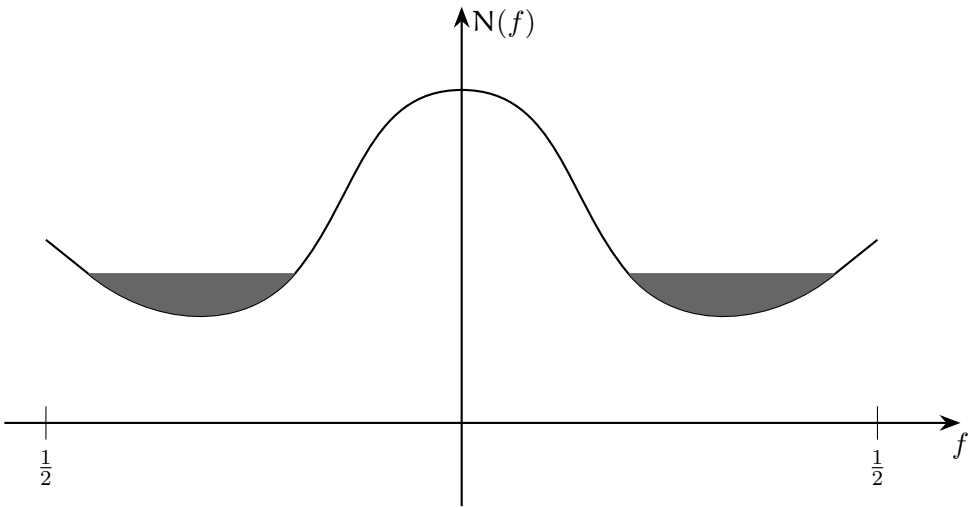


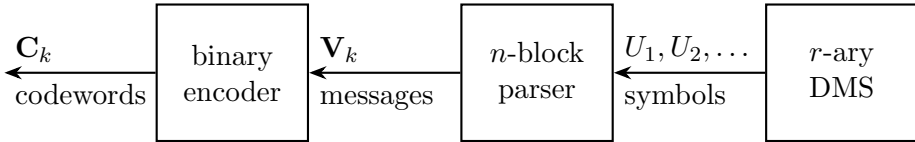












		Y			
$P_{X,Y}$		0	1	2	$P_X$
X	0	$\frac{1}{2}$	0	$\frac{1}{4}$	$\frac{3}{4}$
	1	0	$\frac{1}{4}$	0	$\frac{1}{4}$
$P_Y$		$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	

