Lecture 10 . Embedding with informed coder.

Using the code book in the case of 0-bit WM.

The code book:

With attacker's point of view

$$\pi_i(n),\ n=1,2,\ldots N,\ \ \pi_i(n)\in\{+1,-1\},\ \text{i.i.d} \qquad (15)$$
$$i=1,2,\ldots L$$

WM embedding

$$C_w(n)=C(n)+\alpha\pi_0(n),\ n=1,2\ldots N \ \text{where}\ \pi_0(n):\sum_{n=1}^{N}C(n)\pi_0(n)\geq\sum_{n=1}^{N}C(n)\pi_i(n),\ i\neq 0 \qquad (16)$$

Additive noise attack:

$$C'_w(n)=C_w(n)+\varepsilon(n),\ \ n=1,2,\ldots N \qquad (17)$$

Blind decoding:

$$\Lambda\geq\lambda \text{ - WM is present}$$
$$\Lambda<\lambda \text{ - WM is absent} \qquad (18)$$
$$\text{where } \lambda \text{ - is some threshold,}$$

$$\Lambda=\max_i\sum_{n=1}^{N}C'_w(n)\pi_i(n) \qquad (19)$$

The probability of incorrect WM detecting(P_m and P_{fa} [22])

$$P_m \le \frac{L}{(\sqrt{2\pi})^{L+1}} \int_{-\infty}^{\infty} e^{-y^2/2} F(y)^{L-1} F(\lambda - y\sqrt{\eta_w \eta_a /(\eta_w - \eta_a)} - \sqrt{N\eta_a /(\eta_w - \eta_a)})dy \qquad (20)$$

$$P_{fa} \le 1 - (\frac{1}{\sqrt{2\pi}}) F(\lambda \sqrt{(\eta_w - \eta_a)/(\eta_w \eta_a + \eta_w - \eta_a)})^L, \qquad (21)$$

where:

$$F(x) = \int_{-\infty}^{x} \exp(-t^2/2)dt$$

Numerical calculations show that the use of such informed encoder allows for $P_m = P_{fa} \approx 10^{-4}$ to decrease N in 2-10 times (depending on the values η_w , η_a) in comparison with non-informed encoder (see eq. (9) in the Lecture 9)

WM-ISS –based system [23]

The general idea for "Improved Spread Spectrum Signal (ISS) is – to reduce the impact of CO (as an interference) on the result of blind decoding.

Embedding:

$$C_w(n) = C(n) + \left(\beta(-1)^b - \lambda x\right)\pi'(n), n = 1,2 \ldots N \qquad (22)$$

$$\text{where } \beta, \lambda - \text{some constants}$$

$$x = (C, \pi') = \frac{1}{N\alpha^2}\sum_{n=1}^{N} C(n)\pi'(n), \pi'(n) = \alpha\pi(n) \qquad (23)$$

Particular case:

$$\lambda = 0, \beta = 1 \Rightarrow C_w(n) = C(n) + \alpha(-1)^b \pi(n) \text{(conventional SS - WM)}$$

Attack by additive noise:

$$C'_w(n) = C_w(n) + \varepsilon(n),$$

$$\text{where } E\{\varepsilon(n)\} = 0,\ Var\{\varepsilon(n)\} = \sigma_\varepsilon^2 \qquad (24)$$

Blind decoding rule:

$$\Lambda = \frac{1}{N\alpha^2}\sum_{n=1}^{N} C'_w(n)\pi'(n) \Rightarrow \begin{cases} b = 0, \text{if } \Lambda \geq 0 \\ b = 1, \text{if } \Lambda < 0 \end{cases} \qquad (25)$$

Substituting (22) and (24) in (25) we get:

$$\Lambda = x + \beta(-1)^b - \lambda x + y = \beta(-1)^b + (1-\lambda)x + y, \qquad (26)$$

where $y = \frac{1}{N\alpha^2}\sum_{n=1}^{N} \varepsilon(n)\pi'(n)$

If λ=1, then *C(n)* is absent as interference, but this does not mean that the value λ=1 is optimal one if take into account interferences after WM embedding (say additive noise).

Distortion of CO just after WM embedding

$$\Delta = E\{\left(C_w(n) - C(n)\right)^2\} = E\left\{\left(\left(\beta(-1)^b - \lambda\frac{\tilde{x}}{\alpha^2}\right)\pi'(n)\right)^2\right\} = \alpha^2 E\left\{\left(\beta(-1)^b - \frac{\lambda\,\tilde{x}}{\alpha^2}\right)^2\right\} =$$

$$= \alpha^2 E\{\beta^2 - \frac{2\beta\lambda\tilde{x}(-1)^b}{\alpha^2} + \frac{\lambda^2}{\alpha^4}x^2\} = \alpha^2\left(\beta^2 + \frac{\lambda^2}{\alpha^4}E\{\tilde{x}^2\}\right), \qquad (27)$$

где $\tilde{x} = \frac{1}{N}\sum_{n=1}^{N} C(n)\pi'(n)$

Let us transform the last item in (27):

$$E\{\tilde{x}^2\} = E\{(\frac{1}{N}\sum_{n=1}^{N} C(n)\pi'(n))^2\} = \frac{1}{N^2}\sum_{n=1}^{N}\sum_{n'=1}^{N} E\{C(n)C(n')\pi'(n)\pi'(n')\} =$$

$$\frac{1}{N^2}\sum_{n=1}^{N}\sum_{n'=1}^{N} E\{C(n)C(n')\}E\{\pi(n)\pi'(n')\} = \frac{N}{N^2}\alpha^2\sigma_c^2 = \frac{\alpha^2\sigma_c^2}{N} \qquad (28)$$

Substituting (28) in (27) we get:

$$\Delta = \alpha^2\left(\beta^2 + \frac{\lambda^2\sigma_c^2}{N\alpha^2}\right) = \alpha^2\beta^2 + \frac{\lambda^2\sigma_c^2}{N} \qquad (29)$$

We want to find the parameter β for which CO distortion are equal to CO distortion in the case of conventional WM-SS-based embedding that gives $\Delta=\alpha^2$:

$$\alpha^2 = \alpha^2\beta^2 + \frac{\lambda^2\sigma_c^2}{N} \Rightarrow \beta = \sqrt{\frac{N\alpha^2 - \lambda^2\sigma_c^2}{N\alpha^2}} \qquad (30)$$

Finding of the probability of error for WM-ISS-based system :

$$p = Q\left(\frac{|E\{\Lambda\}|}{\sqrt{Var\{\Lambda\}}}\right) \qquad (31)$$

$$E\{\Lambda\{ = E\{\beta(-1)^b + (1-\lambda)x + y\} = \beta(-1)^b \qquad (32)$$

$$Var\{\Lambda\} = E\{((1-\lambda)x + y)^2\} = E\{(1-\lambda)^2x^2 + 2(1-\lambda)xy + y^2\} = (1-\lambda)^2E\{x^2\} + E\{y^2\} \qquad (33)$$

$$E\{x^2\} = \frac{\sigma_c^2}{\alpha^2 N} \qquad (34)$$

$$E\{y^2\} = \frac{\sigma_\varepsilon^2}{\alpha^2 N} \qquad (35)$$

Substituting (34),(35) in (33), we obtain:

$$Var\Lambda = \frac{(1-\lambda)^2\sigma_c^2 + \sigma_\varepsilon^2}{\alpha^2 N} \qquad (36)$$

Substituting (30) in (32) and (32), (36) into (31) we :

$$P = Q\left(\sqrt{\frac{N\alpha^2 - \lambda^2\sigma_c^2}{\left(1-\lambda^2\right)\sigma_c^2 + \sigma_\varepsilon^2}}\right) \tag{37}$$

In a particular case λ=0 (conventional WM-SS-based)we obtain :

$$\tilde{P} = Q\left(\sqrt{\frac{N\alpha^2}{\sigma_c^2 + \sigma_\varepsilon^2}}\right) = Q\left(\sqrt{\frac{N\eta_a}{\eta_a\eta_\omega + \eta_\omega - \eta_a}}\right) \approx Q\left(\sqrt{\frac{N}{\eta_\omega}}\right) \tag{38}$$

that coincides with (9) (see Lecture 9)

In order to minimize *P* by (37) the parameter λ should be optimized.

It is easy to see that if $\sigma_c^2 / \sigma_\varepsilon^2$ and *N* is large enough, we can let λ_{opt} ≈1

Then we obtain from (37)

$$P = Q\left(\sqrt{\frac{N\alpha^2 - \sigma_c^2}{\sigma_\varepsilon^2}}\right) = Q\left(\alpha\sqrt{\frac{N - \eta_\omega}{\sigma_\varepsilon^2}}\right) = Q\left(\sqrt{\frac{\eta_a\left(N - \eta_\omega\right)}{\eta_\omega - \eta_a}}\right) \tag{39}$$

Comparison WM-SS and WM-ISS

Let us transform (39)

$$P = Q\left(\sqrt{\frac{N-\eta_\omega}{\eta-1}}\right) \qquad (40)$$

где $\eta = \dfrac{\eta_\omega}{\eta_a}$

Compare (40) with relation of the probability of error for informed decoder (see (11) in previous lectures):

$$P = Q\left(\sqrt{\frac{N}{\eta-1}}\right) \qquad (41)$$

If $N \gg \eta_\omega$ we can see that (40) and (41) are close to one another that means that WM-ISS with blind decoder works similar to WM-SS with informed decoder.

Example: $\sigma_c = 50, \alpha = 5, \sigma_\varepsilon = 5, N = 1000$

then : $\eta_\omega = \dfrac{\sigma_c^2}{\alpha^2} = 100, \quad \eta_a = \dfrac{\sigma_c^2}{\alpha^2 + \sigma_\varepsilon^2} = 50 \; P = Q\left(\sqrt{\dfrac{N}{\eta_\omega}}\right) = Q(\sqrt{10}) \approx 3 \cdot 10^{-3}$

We can see that WM-ISS gives for blind decoder the same error probability as WM-SS with blind decoder if the length of sequences *N* is increased till 110 times .Thus WM-ISS is superior to WM-SS because its embedding rate for the same probability is 9 times more.

Concept of WM design different to modulation and demodulation principle typical for communication systems. (*Quantization projective modulation/demodulation – QPD [24]*)

Conventional (quantization index modulation - QIM)

Embedding:

$$C_w(n) = \begin{cases} Q_0(C(n)), \text{if } b = 0 \\ Q_1(C(n)), \text{if } b = 1 \end{cases} \tag{42}$$

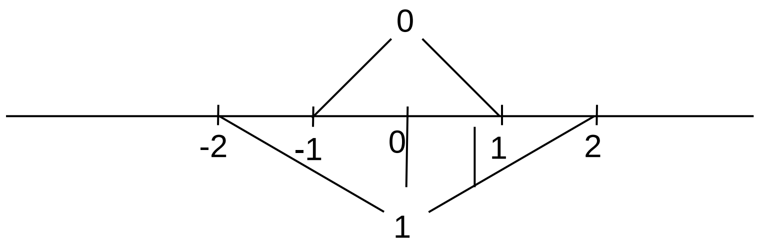
where $Q_i(...) - i^{th}$ type quantizer

Decoding:

$$\tilde{b} = argmin_b \left\| C'_w(n) - Q_b(C'_w(n)) \right\| \tag{43}$$

where || …|| - is a norm in Euclidean space

Example (scalar quantizer):



If $C'_w(n) = C_w(n)$ (no attacks on WM), then the embedded information can be extracted without errors.
If interference is additive white Gaussian noise $\varepsilon(n) \in N(0, \sigma_\varepsilon^2)$, then:

$$P = \sum_{n=-\infty}^{+\infty} \left(Q\left(\frac{\Delta(4n+1)}{\alpha\sqrt{\sigma_\varepsilon^2}} \right) - Q\left(\frac{\Delta(4n+3)}{\alpha\sqrt{\sigma_\varepsilon^2}} \right) \right), Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{\frac{-t^2}{2}} dt \quad (44)$$

, Δ - is a step of quantization.

Requantization attack:

$$C'_w(n) = \begin{cases} C_w(n), \text{with the probability } 0{,}5 \\ C_w(n) \pm \Delta, \text{ with the probability } 0{,}5 \end{cases}$$

,then *p*=0,5 (WM is removed completely) (45)

Ditter QIM **(DM)**

Embedding:

$$C_w(n) = Q\big(C(n) + d(b,n)\big) - d(b,n) \quad (46)$$

Q(…)-quantizer with a step «*Δ*»

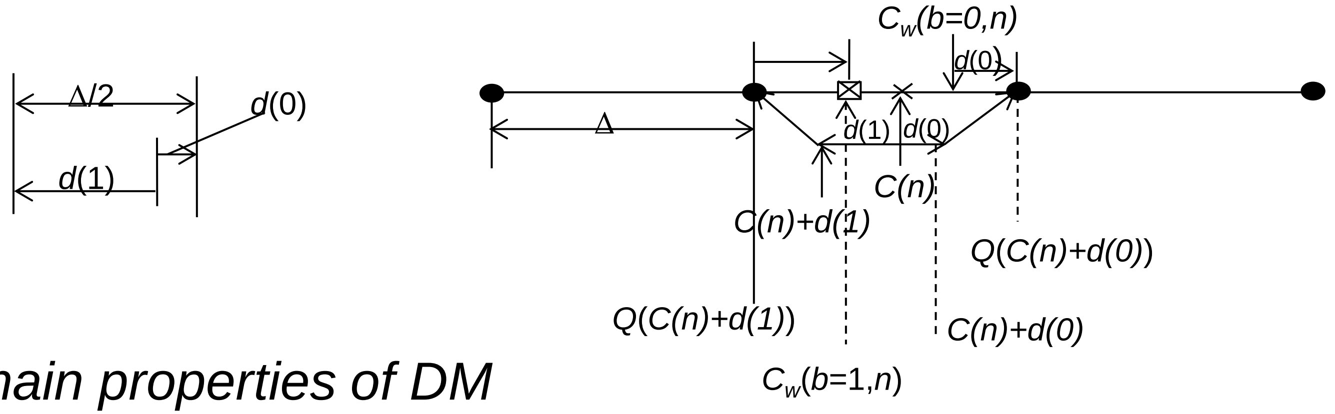
where *d(0,n)* – *i.i.d*., is uniformly independently and distributed on interval [-*Δ/2*,+*Δ/2*]

$$d(1,n) = \begin{cases} d(0,n) + \frac{\Delta}{2,} \text{ if } d(0,n) < 0 \\ d(0,n) - \frac{\Delta}{2,} \text{ if } d(0,n) \geq 0 \end{cases} \quad (47)$$

Decoder:

$$\tilde{b} = argmin_b \left\| C'_w(n) - Q\left(C'_w(n) + d(b,n)\right) + d(b,n) \right\| \quad (48)$$

Graphical interpretation for uniform scalar quantizer :



The main properties of DM

1. If $C'_w(n) = C_w(n)$, then $p = 0$
2. If $C'_w(n) = C_w(n) + \varepsilon(n)$, then $p =$ see (44)
3. If $C'_w(n) = C_w(n) + \tilde{\varepsilon}(n)$, and $|\tilde{\varepsilon}(n)| < \dfrac{\Delta}{4}$ then $p = 0$
4. Quantization errors do not depend on $C(n)$,

that improves comprehension

Vector QIM .

Scalar QIM practically coincides with LSB-WM and therefore it has all its defects .

In the case of vector QIM it is selected initially some code book (consisting from two volumes for embedding of one bit taken from each of volumes) :

$$C_{io}(n), n = 1,2...N, C_{i1}(n), n = 1,2...N, i = 1,2,...L$$

Embedding:

$$C_w(n) = \begin{cases} C_{\tilde{i}o}(n), \text{if } b = 0 \\ C_{\tilde{i}1}(n), \text{if } b = 1 \end{cases} \qquad (49)$$

where $C_{\tilde{i}1}(n) = argmin_i \left\| C_w(n) - C_{ib}(n) \right\|$

Decoding:

$$\tilde{b} = argmin_b min_i \left\| C'_w(n) - C_{ib}(n) \right\| \qquad (59)$$

Remark 1.

In order to get small CO distortions the code books have to be chosen in such a way that for any CO ПС *C(n),* $||C(_w(n)-C(n)||$,should be small in comparison with $||C(n)||$

Remark 2.

Such system can be used also as SG and it will be resistant to deliberate removal if the selection of code books is controlled by stegokey .

Remark 3.

This WM system is agreed with vector coding using in speech coders (*vocoders*).

Quantization projective modulation/demodulation (QPD)[24]

The reason to use QPD:

To provide a resistance WM against its deliberate removal by randomizing of quantization levels .

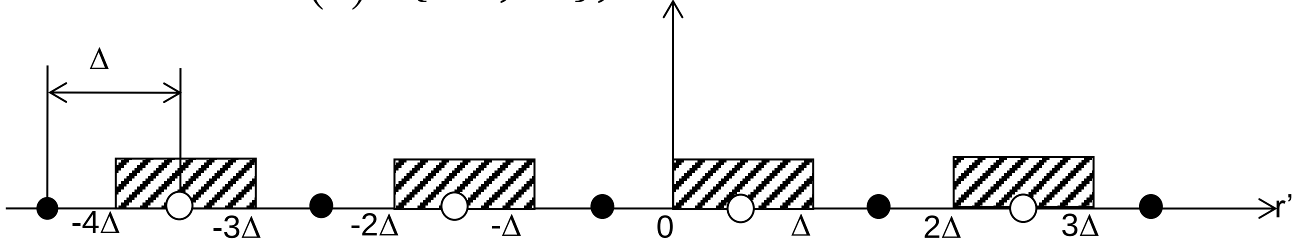
Embedding:

$$C_w(n) = C(n) + \frac{Q_b(r) - r}{N}\pi(n), n = 1,2,...,N \quad (51)$$

where $r = \sum_{n=1}^{N} C(n)\pi(n)$

$Q_b(\ldots)$ – uniform quantizer with step Δ, as for b=0 and for b=1 are taken alternating points (see Fig. below)

$$\pi(n) \in \{-1,+1\}, i.i.d.$$



● → b=1, ○ → b=0, shaded regions → 0 , non-shaded regions → 1

Fig1. Uniform quantizer with the step Δ

Additive noise attack:

$$C'_w(n) = C(n) + \varepsilon(n), \text{где } E\{\varepsilon(n)\} = 0, \qquad (52)$$

$$Var\{\varepsilon(n)\} = \sigma_\varepsilon^2$$

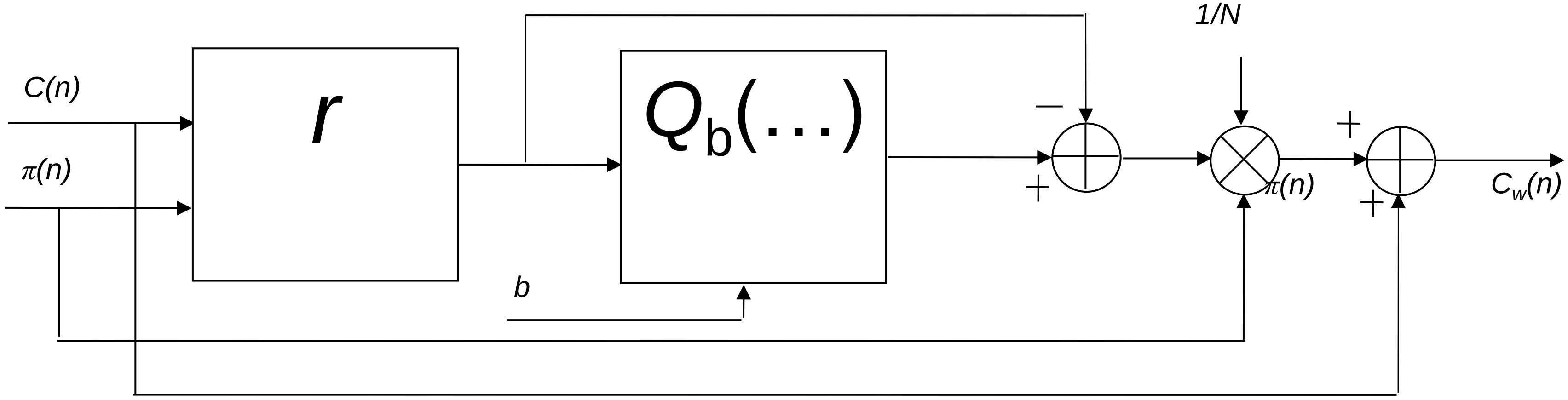
Decoder :

$$\tilde{b} = argmin_b \left\| r' - Q_b(r') \right\|^2, b \in \{0,1\} \qquad (53)$$

where:

$$r' = \sum_{n=1}^{N} C'_w(n)\pi(n)$$

Fig. 2. WM embedding scheme:



Recovering «b» under attack absence:

Let us *b*=0, then we get from (52) :

$$r' = \sum_{n=1}^{N} C'_w(n)\pi(n) = \sum_{n=1}^{N}\left(C(n) + \frac{\rho_0}{N}\pi(n)\right)\pi(n) = \sum_{n=1}^{N}\left(C(n)\pi(n) + \frac{\rho_0}{N}\right) =$$

$$\sum_{n=1}^{N} C(n)\pi(n) + \sum_{n=1}^{N}\left(\frac{Q_0\left(\sum_{n=1}^{N} C(n)\pi(n)\right) - \sum_{n=1}^{N} C(n)\pi(n)}{N}\right) = \tag{54}$$

$$\sum_{n=1}^{N} C(n)\pi(n) + Q_0\left(\sum_{n=1}^{N} C(n)\pi(n)\right) - \sum_{n=1}^{N} C(n)\pi(n) = Q_0\left(\sum_{n=1}^{N} C(n)\pi(n)\right)$$

$$Q_0(r) - r' = Q_0\left(Q_0\left(\sum_{n=1}^{N} C(n)\pi(n)\right)\right) - Q_0\left(\sum_{n=1}^{N} C(n)\pi(n)\right) = 0$$

$$Q_1(r) - r' = Q_1\left(Q_0\left(\sum_{n=1}^{N} C(n)\pi(n)\right)\right) - Q_0\left(\sum_{n=1}^{N} C(n)\pi(n)\right) = \Delta \tag{55}$$

For b=1 , we get in a similar manner that $Q_0(r') - r' = \Delta, Q_1(r') - r' = 0$

Conclusion:

If attack is absent then decoder gives no errors

Calculation of the probability of error for decoding of the bit b under additive noise attack .

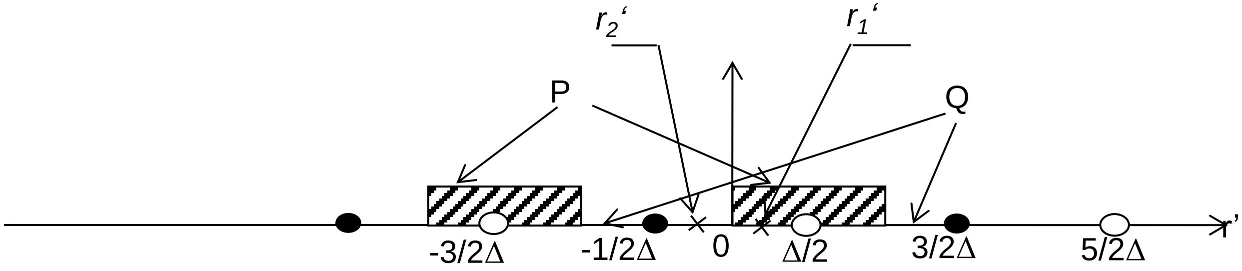
Let us b=0. then we get from (51), (52) :

$$r' = \sum_{n=1}^{N} C'_w(n)\pi(n) = \sum_{n=1}^{N}\left(C(n) + \frac{\rho_0}{N}\pi(n) + \varepsilon(n)\right)\pi(n) = \sum_{n=1}^{N}\left(C(n)\pi(n) + \varepsilon(n)\pi(n) + \frac{\rho_0}{N}\right) =$$

$$\sum_{n=1}^{N} C(n)\pi(n) + \sum_{n=1}^{N}\varepsilon(n)\pi(n) + \sum_{n=1}^{N}\left(\frac{Q_0\left(\sum_{n=1}^{N} C(n)\pi(n)\right) - \sum_{n=1}^{N} C(n)\pi(n)}{N} \right) =$$

$$\sum_{n=1}^{N} C(n)\pi(n) + \sum_{n=1}^{N}\varepsilon(n)\pi(n) + Q_0\left(\sum_{n=1}^{N} C(n)\pi(n)\right) - \sum_{n=1}^{N} C(n)\pi(n) = Q_0\left(\sum_{n=1}^{N} C(n)\pi(n)\right) + \sum_{n=1}^{N}\varepsilon(n)\pi(n) \quad (56)$$

Consider the region where is taking a decision about symbol *b:*



Decoding algorithm given values r'_1, r'_2 *is*

$$\left|Q_0(r')-r'\right| < \left|Q_1(r')-r'\right| \Rightarrow b=0$$

$$\left|Q_0(r')-r'\right| \geq \left|Q_1(r')-r'\right| \Rightarrow b=1$$

$$r'_1 : Q_0(r'_1) = \Delta/2, Q_1(r'_1) = -\Delta/2, \left|\Delta/2 - r'_1\right| < \left|-\Delta/2 - r'_1\right| \Rightarrow b=0$$

$$r'_2 : Q_0(r'_2) = \Delta/2, Q_1(r'_2) = -\Delta/2, \left|\Delta/2 - r'_2\right| > \left|-\Delta/2 - r'_2\right| \Rightarrow b=1$$

Conclusion:

Shaded regions (*P*) corresponds to decision *b*=0,

whereas non-shaded- *b*=1

thus if *b*=0 is embedded then

$$P = Pr\{r' \notin P\} = Pr\left\{ r' \notin \bigcup_{i=-\infty}^{\infty} \left(2\Delta i, \Delta(2i+1)\right) \right\} \qquad (57)$$

$$r' = Q_0\left(\sum_{n=1}^{N} C(n)\pi(n)\right) + \sum_{n=1}^{N} \varepsilon(n)\pi(n)$$

$$\left.\begin{array}{l} Q_0\left(\sum_{n=1}^{N} C(n)\pi(n)\right) \in \bigcup_{i=-\infty}^{\infty} \Delta(2i+1/2) \\ r' \notin P \end{array}\right\} \Rightarrow \Lambda = \sum_{n=1}^{N} \varepsilon(n)\pi(n) \in \bigcup_{i=-\infty}^{\infty} (2\Delta i + \Delta/2, 2\Delta i + 3\Delta/2) \Rightarrow$$

$$\Lambda \in \bigcup_{i=-\infty}^{\infty} (2\Delta i + \Delta/2, 2\Delta i + 3\Delta/2) \qquad (58)$$

$$\Lambda \in N\left(0, N\sigma_\varepsilon^2\right) \qquad (59)$$

$$(58),(59) \Rightarrow P = \sum_{i=-\infty}^{\infty}\left(Q\left(\Delta\frac{(2i+1/2)}{\sqrt{N\sigma_\varepsilon^2}}\right) - Q\left(\Delta\frac{(2i+3/2)}{\sqrt{N\sigma_\varepsilon^2}}\right)\right) =$$

$$\sum_{i=-\infty}^{\infty} Q\left(\Delta\frac{(4i+1)}{2\sqrt{N\sigma_\varepsilon^2}}\right) - Q\left(\Delta\frac{(4i+3)}{2\sqrt{N\sigma_\varepsilon^2}}\right) \qquad (60)$$

Neglecting by «side petals» in (60), we get:

$$p \approx 2\mathrm{Q}\left(\frac{\Delta}{2\sqrt{N\sigma_\varepsilon^2}}\right) \qquad (61)$$

Distortion evaluation of CO under WM embedding and additive noise attack :

$$\eta_\omega = \frac{\sigma_c^2}{E\{\left(C_w(n)-C(n)\right)^2\}} = \frac{\sigma_c^2 N^2}{E\{(Q_b(r)-r)^2\}} \qquad (62)$$

$$where: r = \sum_{n=1}^{N} C(n)\pi(n)$$

We can see from(62) that η_ω depends not only from the current value *C(n)* but also from adjacent samples *C(n), n*=1,2,…*N*, and by non-linear manner.

However the following bound holds $\left|Q_b(r)-r\right| \le \Delta$ and therefore we get :

$$\eta_\omega \ge \frac{\sigma_c^2 N^2}{\Delta^2} \qquad (63)$$

$$C'_w(n) = C_w(n) + \varepsilon(n), Var\{\varepsilon(n)\} = \sigma_\varepsilon^2 \Rightarrow$$

$$\eta_a = \frac{\sigma_c^2}{E\{\left(C_w(n)-C(n)\right)^2\} + \sigma_\varepsilon^2} = \frac{\sigma_c^2}{\Delta^2 / N^2 + \sigma_\varepsilon^2} \qquad (64)$$

If we let equality in (63), (64), we get from them

$$\frac{\sigma_\varepsilon^2 N}{\Delta^2} = \frac{\eta_\omega}{N\eta_a} - 1/N = 1/N\left(\frac{\eta_\omega}{N\eta_a} - 1\right) \tag{65}$$

Substituting (63) and (65) in (61), we obtain

$$P \le 2\mathrm{Q}\left(\frac{1}{2}\sqrt{\frac{N\eta_a}{\eta_\omega - \eta_a}}\right) = 2\mathrm{Q}\left(\frac{1}{2}\sqrt{\frac{N}{\eta - 1}}\right) = 2Q\left(\sqrt{\frac{N}{4(\eta-1)}}\right) \tag{66}$$

ãäå $\eta = \dfrac{\eta_\omega}{\eta_a}$

For more precise evaluation of distortion for QPD, we get the bound[24]: $P \le 2Q\left(\sqrt{\frac{0{,}75N}{(\eta-1)}}\right)$

Conclusion:

If we compare (66) with (41) that gives the probability of error for informed decoder with the use of SS-based WM we can see that for the same reliability the length of pseudorandom sequence N has to be increased for QPD (with blind decoder) at 1.3 times approximately.This is some sacrifice on altar of “blind decoding”.

Parameter optimization of QPD-WM

We fix the following values:

$P, \sigma^2_{c.}\ \eta_a$. It is necessary to find such parameters Δ и N, that maximize η_ω.

Remark 1.

Equations (63), (64), (66) are approximate ones and therefore they should be specified by simulation .

Remark 2.

QPD-WM, (similar as ISS-WM) and in contrast to SS-WM produces correlated errors of CO on the interval of the length N samples (pixels).

Remark 3.

If we compare QPD with ISS(see for that eq.(40)

$$P = Q\left(\sqrt{\frac{\eta_a\left(N-\eta_\omega\right)}{\eta_\omega-\eta_a}}\right) = Q\left(\sqrt{\frac{N-\eta_\omega}{\eta-1}}\right)$$

we can conclude that for large N, ISS is superior to QPD, but there may be another situation (see lecture 15).