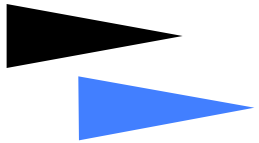


Tatouage robuste par étalement de spectre avec prise en compte de l'information adjacente

Gaëtan Le Guelvouit

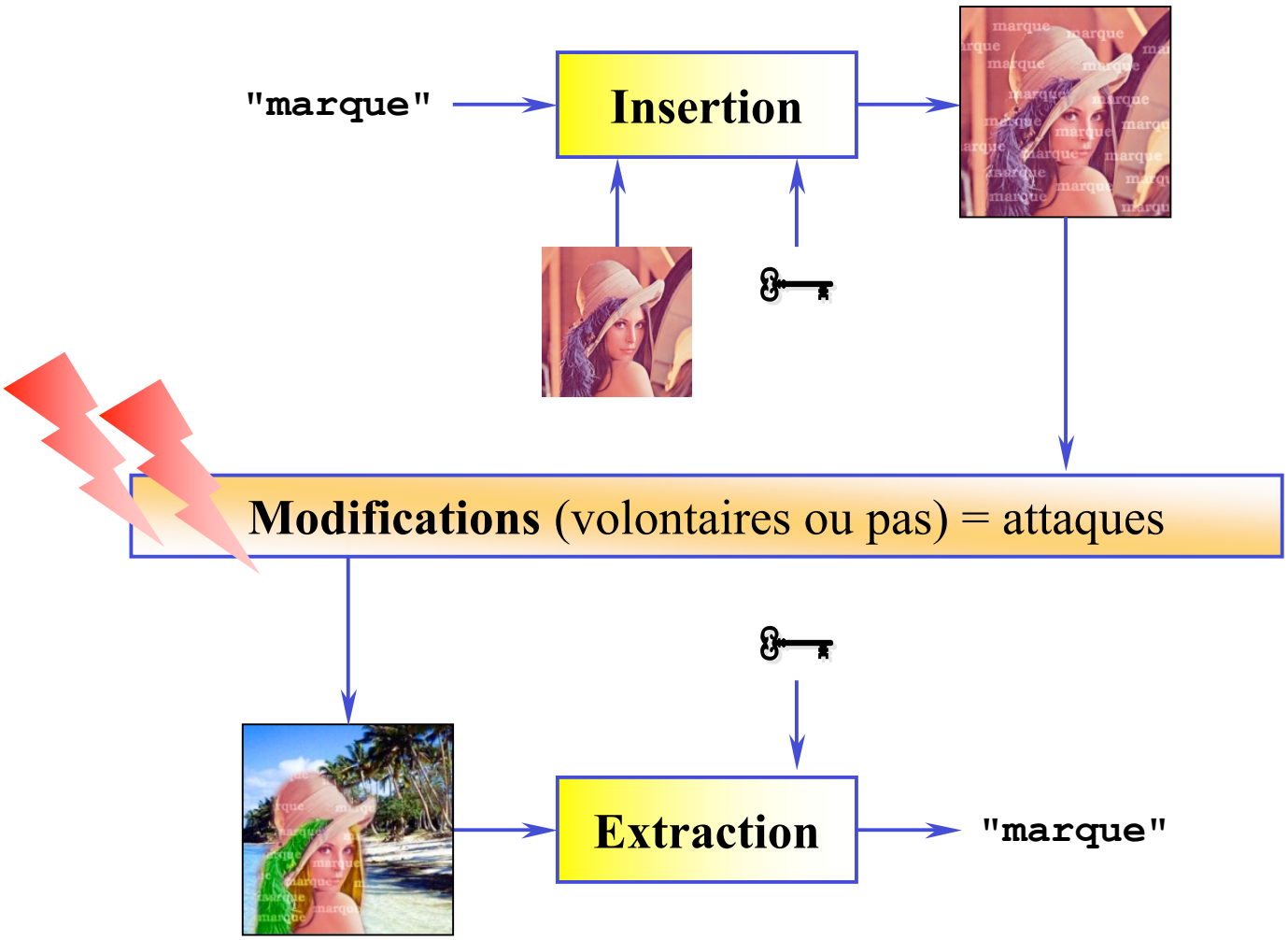
Sous la direction de C. Guillemot et S. Pateux

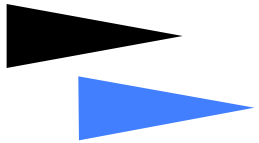
Temics - IRISA

Sécurité des documents numériques

- De plus en plus de documents numériques (CD audio, DVD, photo numérique, …)
  - Facilité de stockage, manipulation et transmission
  - Mais problèmes de piratage, authentification et droits d'auteur
- Solution depuis une dizaine d'années : tatouage numérique
  - Insertion d'une marque au sein d'un document
  - Ne doit pas gêner l'exploitation
  - Notion de robustesse (≠ stéganographie)

Tatouage robuste : principe

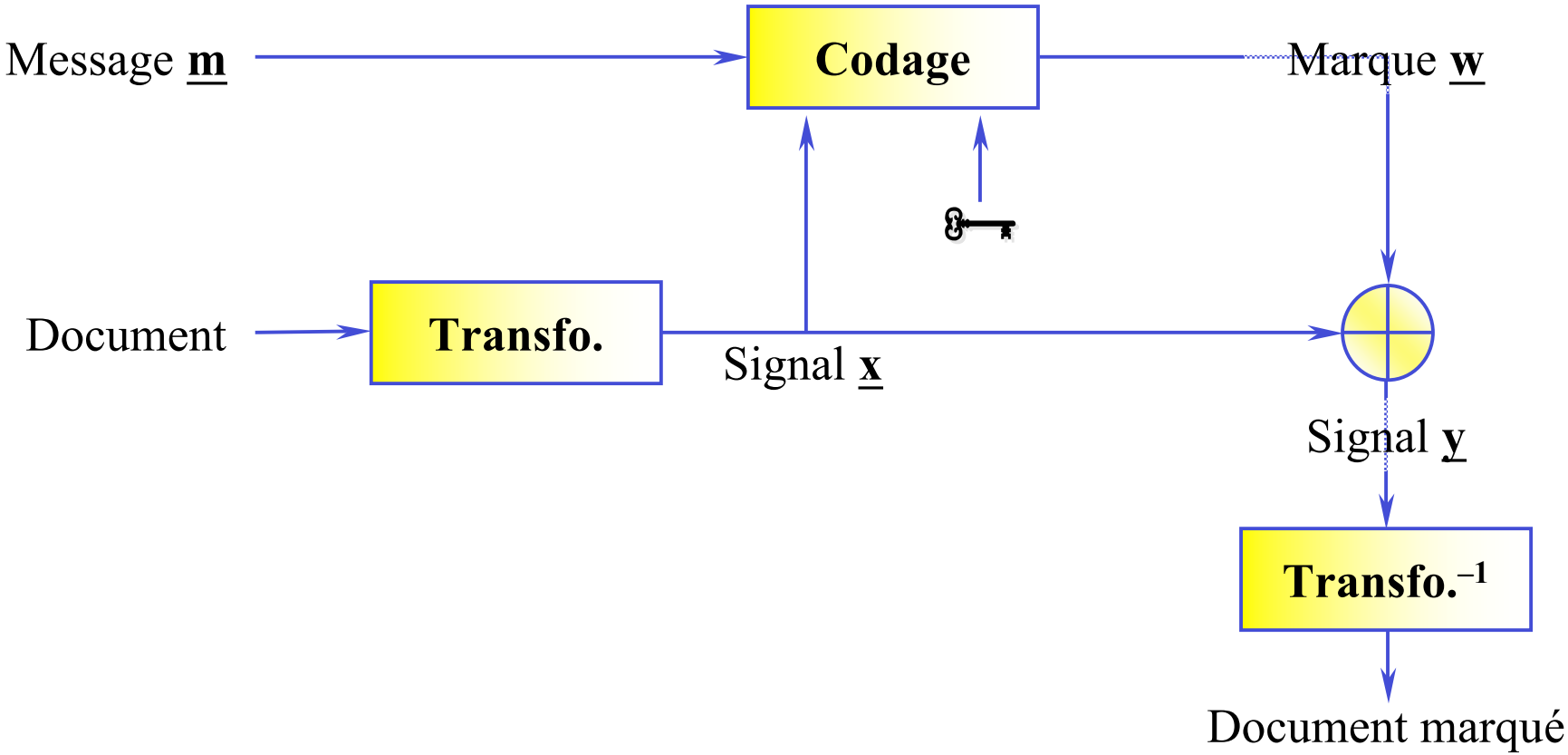




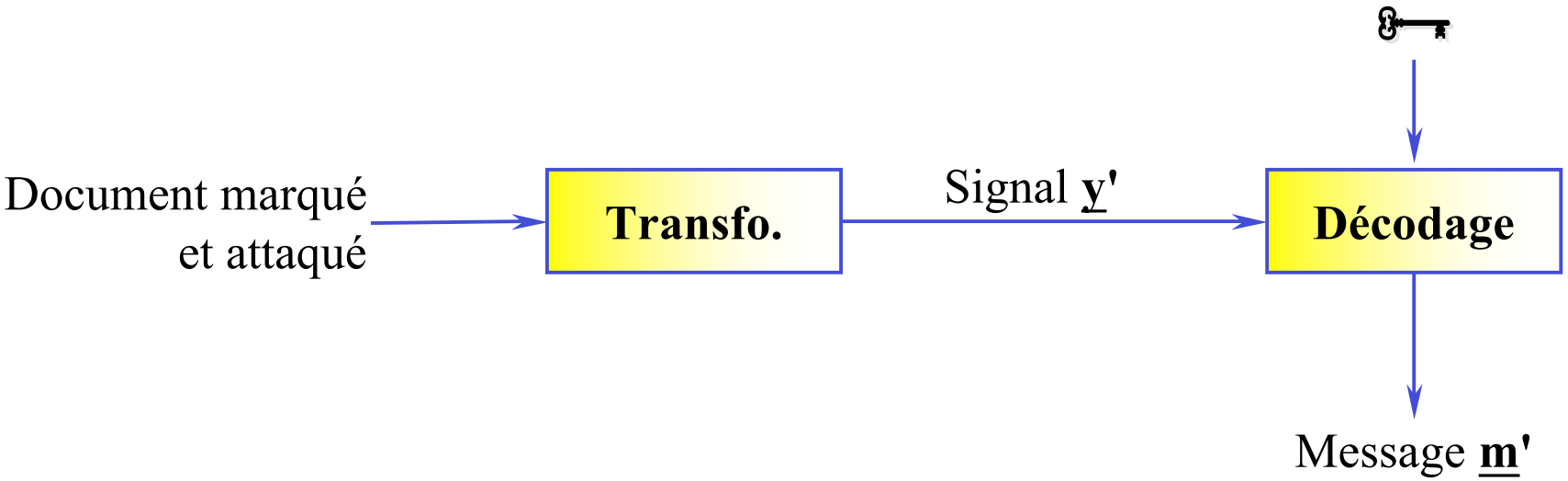
Deux principales techniques

- Le tatouage substitutif [Turner89, Koch05, Puate96]
  - Modification du document afin de correspondre à une marque (ex : quantification [Eggers00])
  - Bonnes performances pour des attaques simples (bruit)
  - Performances médiocres sur attaques évoluées
- Le tatouage additif [Bender95, Cox97]
  - Ajout d'une marque au document
  - Le document est une source d'interférence
    ⇒ performances a priori moins bonnes
  - Contrôle aisée de la distorsion introduite
  - Robustesse aux attaques évoluées (filtrage, compression, ...)

Tatouage additif
Insertion

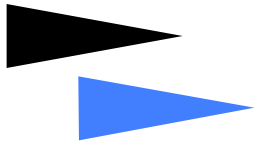


Tatouage additif
Extraction



Problématique

- $\underline{\mathbf{y}} = \underline{\mathbf{x}} + \underline{\mathbf{w}} = \underline{\mathbf{x}} + \text{codage}(\underline{\mathbf{m}}, \underline{\mathbf{x}}, \text{🗝})$
- Comment définir les fonctions $\text{codage}(\underline{\mathbf{m}}, \underline{\mathbf{x}}, \text{🗝})$ et $\text{décodage}(\underline{\mathbf{y}}', \text{🗝})$?
  - Pour être le plus robuste possible
  - Pour respecter une distorsion maximale donnée
- Décomposable en deux parties
  1. *Mise en forme* : passage de k bits (<1000) à un signal de dimension m ($>10^5$): $\underline{\mathbf{m}} \rightarrow \underline{\mathbf{w}}^0$
  2. *Adaptation* : adapter $\underline{\mathbf{w}}^0$ au document à marquer : $\underline{\mathbf{w}}^0 \rightarrow \underline{\mathbf{w}}$

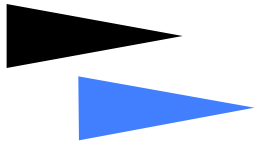


Etat de l'art

Techniques empiriques

- Mise en forme $\underline{\mathbf{m}} \rightarrow \underline{\mathbf{w}}^0$
  - Patchwork [Bender95]
  - Etalement de spectre [Cox97, ...]
  - Puis ajout de codes correcteurs
    - Simple répétition
    - Codes BCH, codes convolutifs
    - Combinaison de répétition + codes

- Adaptation $\underline{\mathbf{w}}^0 \rightarrow \underline{\mathbf{w}}$
  - Pas d'adaptation à $\underline{\mathbf{x}}$: ajout d'un bruit blanc
  - Marquage uniquement de certaines fréquences
  - Critère perceptuel [Swanson96, Piva97, Podilchuk98, ...]



Etat de l'art

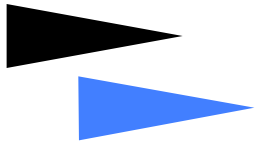
Début de théorisation

- Mise en forme $\underline{\mathbf{m}} \rightarrow \underline{\mathbf{w}}^0$
  - Tatouage = problème de communication
    - Justification de l'utilisation de codes correcteurs
    - Avec information adjacente (une partie du bruit est connue) [Cox99] ⇒ redécouverte de [Costa83]
- Adaptation $\underline{\mathbf{w}}^0 \rightarrow \underline{\mathbf{w}}$
  - Prise en compte des attaques
    - Concentrer la marque sur les échantillons perceptuellement importants [Cox97]
    - Marque similaire au signal hôte (PSC [Su99])
    - tatouage/attaque = jeu (max. robustesse *vs* min. robustesse) [Moulin98]

Objectif

- Comment exploiter les bases théoriques évoquées dans un schéma **pratique** de tatouage ?
  - Exhiber un canal à partir d'un schéma de tatouage additif
  - Répartir la marque que nous ajoutons
  - Coder le message

- Deux axes d'études :
  1. Définition et optimisation d'un canal de communication
  2. Proposition d'une technique de codage adaptée

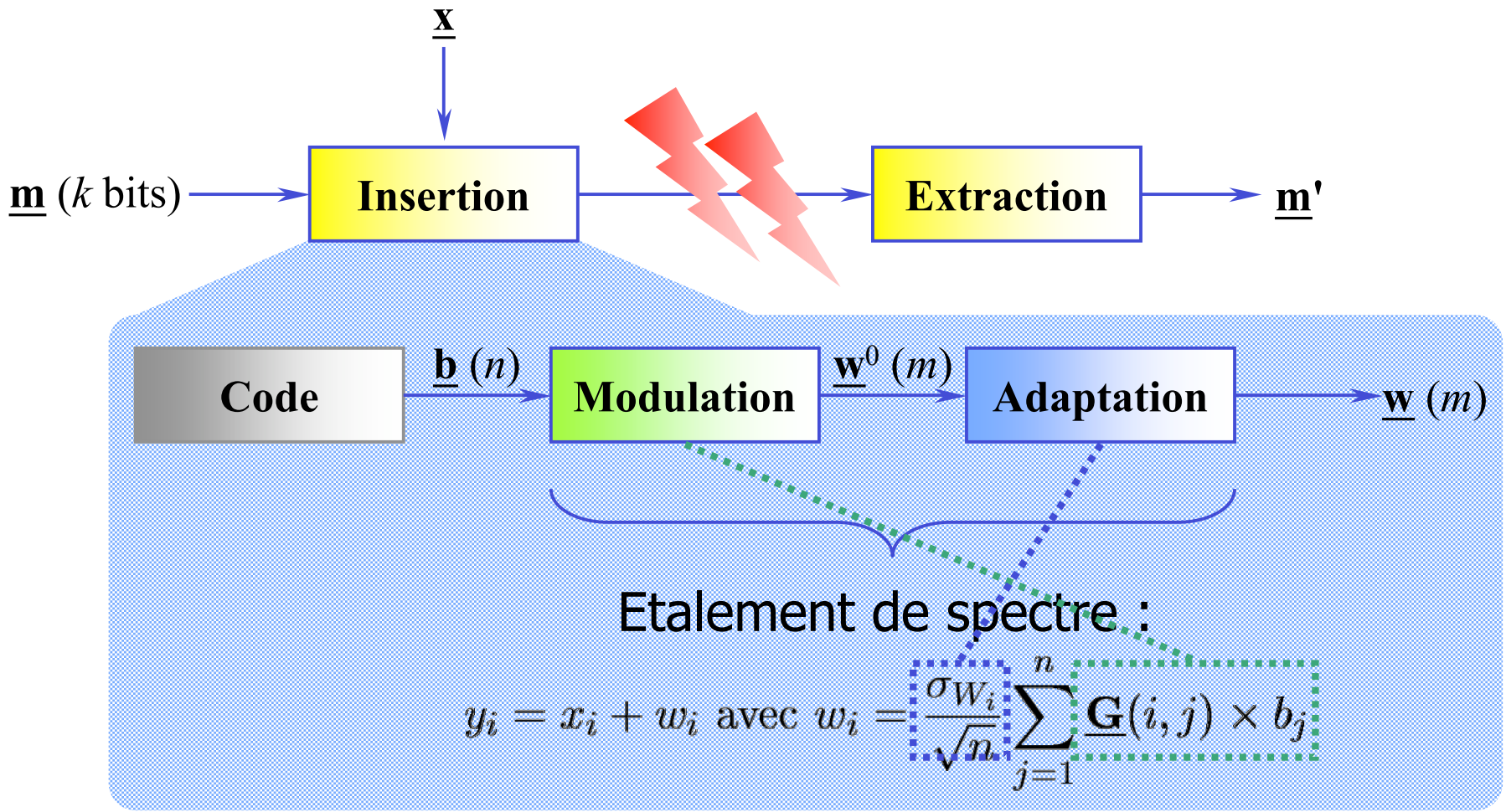


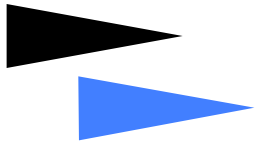
Plan

- Tatouage = problème de communication
  - Définition du canal
  - Optimisation par théorie des jeux
  - Résultats
- Codage du message
  - Rappel du schéma de Costa
  - Dictionnaire partitionné par codes poinçonnés
  - Construction de la marque
- Application pratique : tatouage d'images

Définition du canal

Etalement de spectre





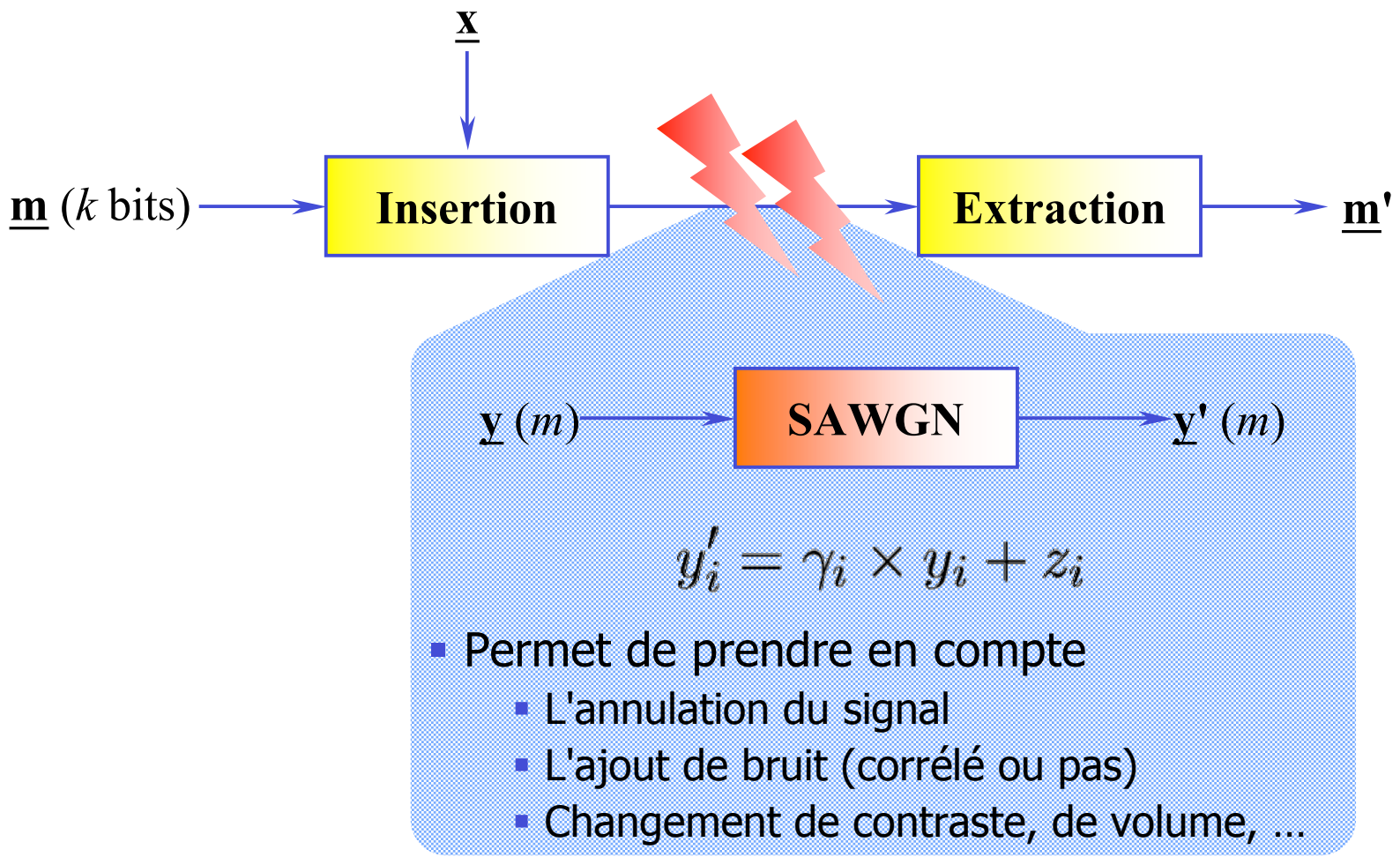
Définition du canal

Etalement de spectre

- Principe : construit une marque de grande dimension par *"étalement"* de symboles
  $\Rightarrow$ passage de n bits à un signal de $\underline{\mathbf{w}}^0$ de dimension m
- Adaptée aux canaux fortement bruités
- Notion de sécurité grâce à des porteuses pseudo-aléatoires
- Définit un canal gaussien [Piva97]
  $\Rightarrow$ même si $\underline{\mathbf{x}}$ ne suit pas une loi Normale, on se ramène à un canal connu

Définition du canal

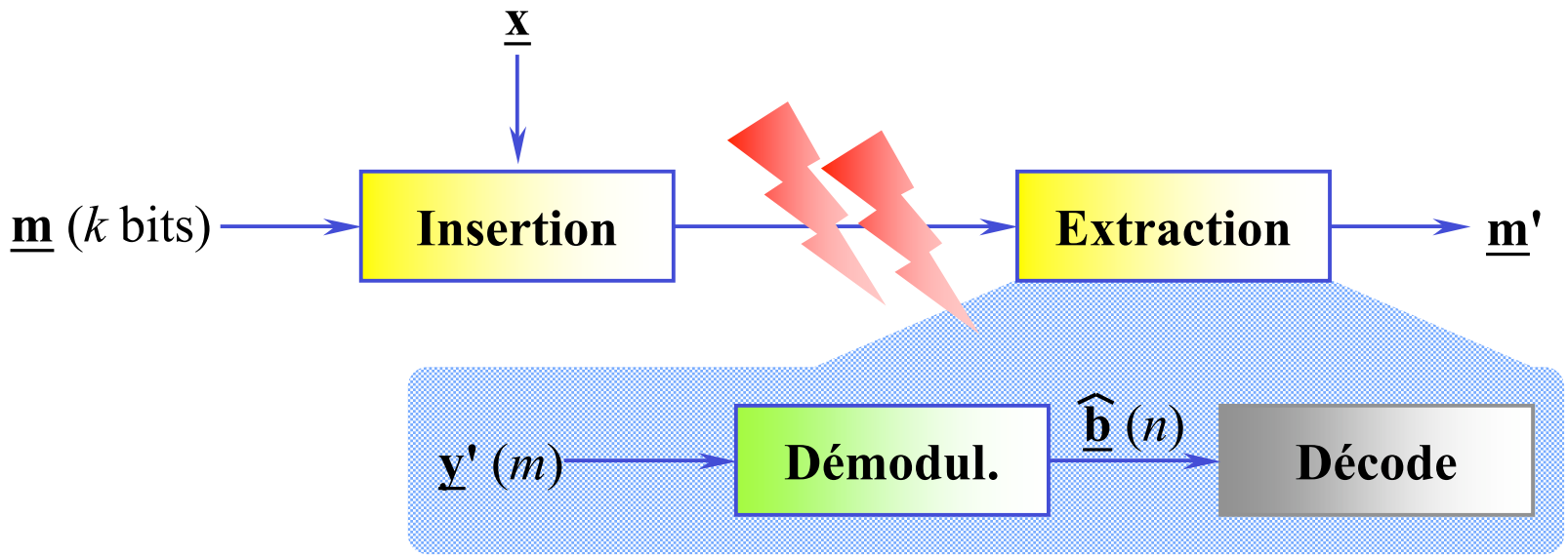
Modélisation des attaques



- Permet de prendre en compte
  - L'annulation du signal
  - L'ajout de bruit (corrélé ou pas)
  - Changement de contraste, de volume, …

Définition du canal

Extraction

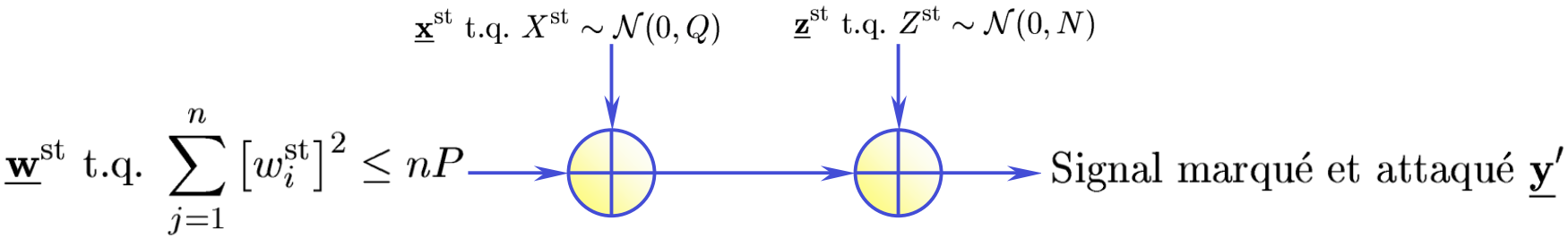


Définition du canal

Canal gaussien…

- n porteuses → sous-espace linéaire
  - Signal hôte $\underline{\mathbf{x}} \rightarrow \underline{\mathbf{x}}^{\mathrm{st}}$
  - Bruit d'attaque $\underline{\mathbf{z}} \rightarrow \underline{\mathbf{z}}^{\mathrm{st}}$
  - En posant $\underline{\mathbf{w}}^{\mathrm{st}} = \underline{\mathbf{b}}$, on a $\underline{\mathbf{y}}^{\mathrm{st}} = \underline{\mathbf{w}}^{\mathrm{st}} + \underline{\mathbf{x}}^{\mathrm{st}} + \underline{\mathbf{z}}^{\mathrm{st}}$
  - $\underline{\mathbf{x}}^{\mathrm{st}}$ et $\underline{\mathbf{z}}^{\mathrm{st}}$ signaux gaussiens i.i.d.

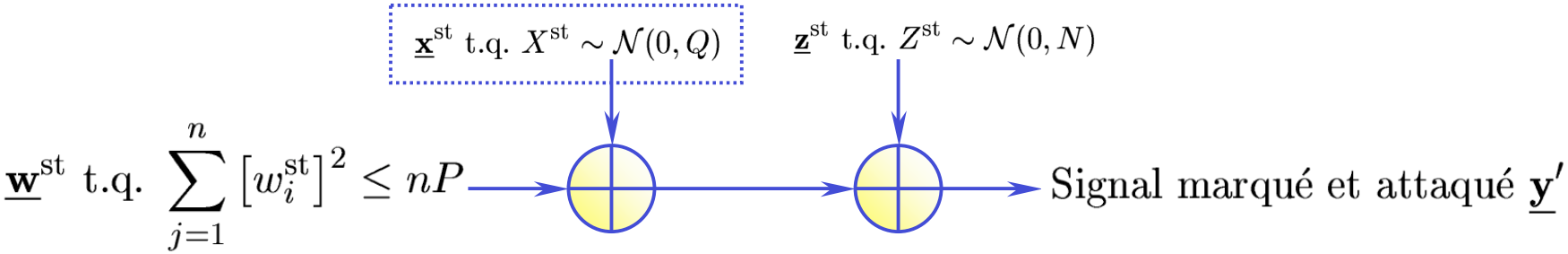
➔ Canal gaussien



➔ Comment maximiser cette performance en jouant sur les σ_W ?

Définition du canal

… Avec info. adjacente



- $\underline{\mathbf{x}}^{\mathrm{st}}$ parfaitement connu à l'insertion (signal à marquer)

➔ Canal avec information adjacente

- Capacité [Costa83] :

$$C = \frac{1}{2}\log_2\left[1 + \frac{P}{N}\right] \quad \text{avec} \quad \frac{P}{N} = \sum_{i=1}^{m} \frac{\gamma_i \sigma_{W_i}}{\sigma_{Z_i}^2}$$

➔ P/N = mesure de performance à optimiser

Optimisation par théorie des jeux

Principe du max-min

- Soient la distorsion d'insertion D_{e} et d'attaque D_{a}

1. Pour une stratégie d'insertion $\underline{\mathbf{e}}$ donnée, l'attaque optimale est

$$\underline{\mathbf{a}}^{\star} = \arg \min_{\underline{\mathbf{a}} \in \mathcal{A}(D_{\mathrm{a}})} \{\mathrm{perf}(\underline{\mathbf{e}}, \underline{\mathbf{a}})\}$$

2. La meilleure défense face à cette attaque est

$$\begin{aligned} \underline{\mathbf{e}}^{\star} &= \arg \max_{\underline{\mathbf{e}} \in \mathcal{E}(D_{\mathrm{e}})} \{\mathrm{perf}(\underline{\mathbf{e}}, \underline{\mathbf{a}}^{\star})\} \\ &= \arg \boxed{\max_{\underline{\mathbf{e}} \in \mathcal{E}(D_{\mathrm{e}})} \left\{ \min_{\underline{\mathbf{a}} \in \mathcal{A}(D_{\mathrm{a}})} \{\mathrm{perf}(\underline{\mathbf{e}}, \underline{\mathbf{a}})\} \right\}} \end{aligned}$$

- **Performance minimale garantie** pour attaques de distorsion $< D_{\mathrm{a}}$

Optimisation par théorie des jeux

1. Attaque optimale

- Distorsion d'attaque = EQM pondérée entre $\underline{\mathbf{x}}$ et $\underline{\mathbf{y}}'$:

$$D_{\mathrm{a}} = D_{\underline{\mathbf{x}}\underline{\mathbf{y}}'} = \mathbb{E}\left[\varphi_i^2 \left(x_i - y_i'\right)^2\right]$$

Pondération perceptuelle

- Formulation lagrangienne

$$\left(\gamma_i^\star, \sigma_{Z_i}^\star\right) = \arg \min_{\gamma_i, \sigma_{Z_i} \geq 0} \left\{P/N + \lambda D_{\underline{\mathbf{x}}\underline{\mathbf{y}}'}\right\}$$

➔ Résolution = deux stratégies d'attaques

- Annulation de l'échantillon
- Ajout de bruit + filtrage de Wiener (réduction de D_{a})

Optimisation par théorie des jeux

2. Stratégie de défense

- Distorsion d'insertion :

$$D_{\mathrm{e}} = D_{\underline{\mathbf{xy}}} = \mathbb{E}\left[\varphi_i^2 \left(x_i - y_i\right)^2\right]$$

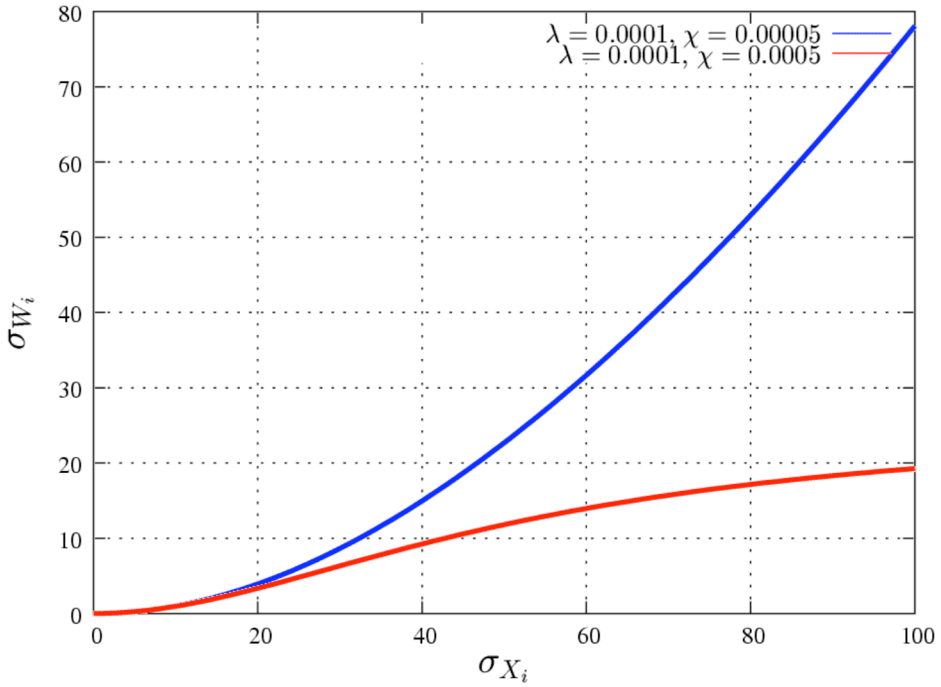
- Wiener à l'attaque = réduction de D_{a} sans impact sur P/N ➔ **auto-application** à l'insertion
- Encore une formulation lagrangienne

$$\sigma^{\star}_{W_i} = \arg \max_{\sigma_{W_i} \geq 0} \left\{ P/N + \lambda D_{\underline{\mathbf{xy}'}} - \chi D_{\underline{\mathbf{xy}}} \right\}$$

➔ Résolution pour les 2 stratégies d'attaques

Optimisation par théorie des jeux

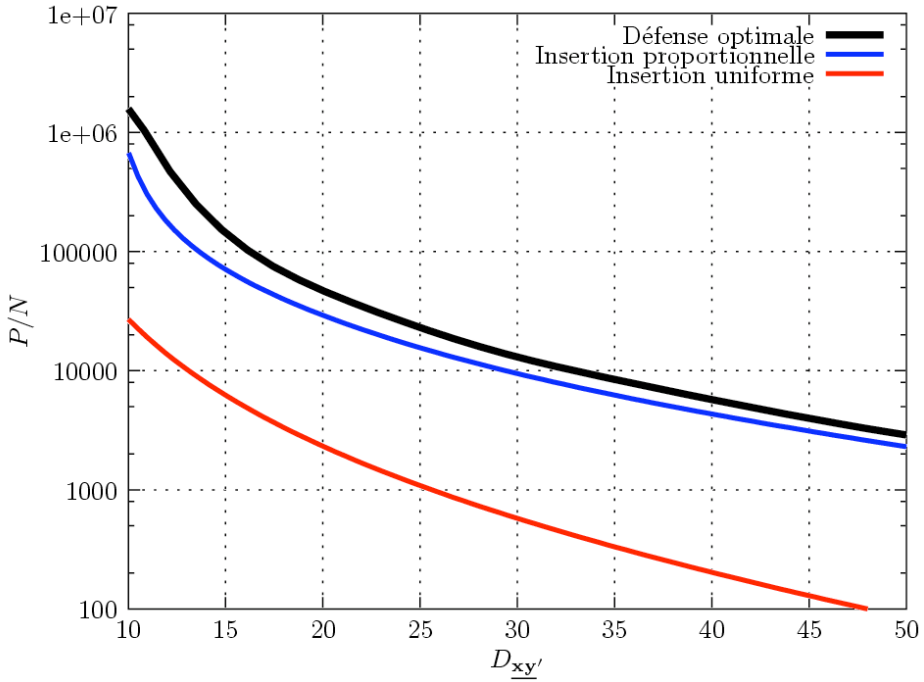
2. Stratégie de défense



- Deux stratégies suivant l'attaque visée
  - Si attaque > tatouage ➔ Concentration dans les fortes énergies
  - Si attaque < tatouage ➔ Répartition équilibrée

Résultats

Face à l'attaque optimale

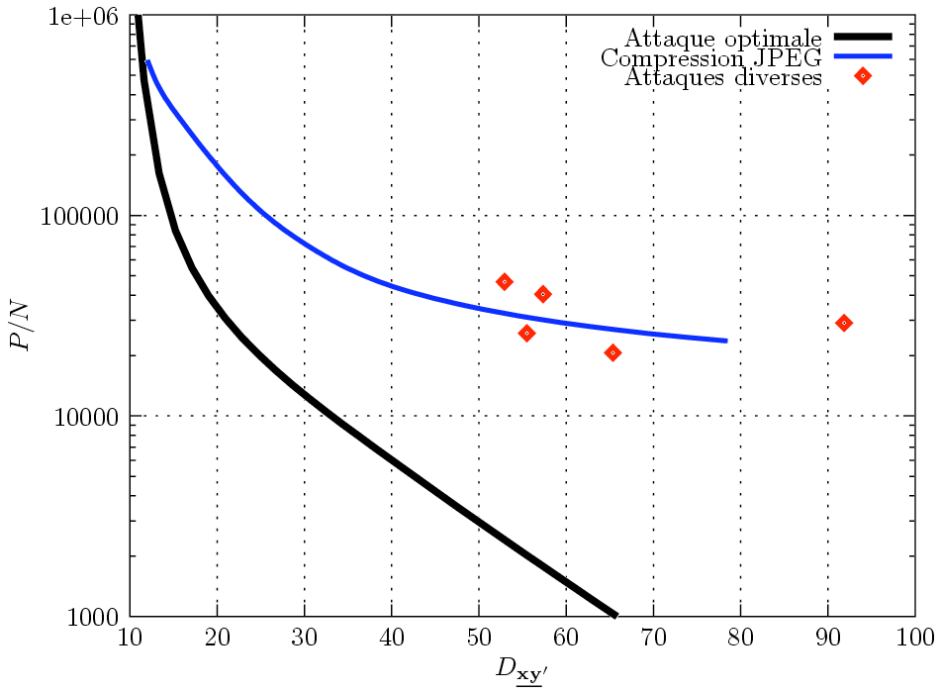


- Signal $\underline{\mathbf{x}}$ obtenu par DWT 3 niveaux de *Lena* 512×512
- Tatouage avec $D_e = 10$ (PSNR ≈ 38 dB)
- Application de l'attaque optimale
- Comparaison face à
  - Insertion uniforme
  - Proportionnelle [Piva97, Su99]

➔ Face à la pire des attaques, notre stratégie est la meilleure

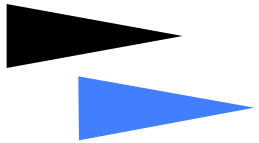
Résultats

Face à des attaques réelles



- Insertion réglée pour une attaque $D_a = 40$
- Attaques de Stirmark [Petitcolas00]
  - Compression JPEG
  - Filtrages (médian, flou…)

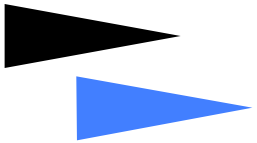
➔ Attaques réelles très sous-optimales, les performances sont encore meilleures



Première partie

Résumé

- Identification du canal de tatouage
  - L'étalement de spectre définit un sous-espace linéaire
  - Canal gaussien **unique** avec information adjacente (≠ [Moulin01])
- Optimisation du signal-à-bruit P/N
  - Utilisation de distorsions pondérées
  - Définition d'une attaque SAWGN optimale
  - Calcul d'une stratégie d'insertion assurant une performance minimale pour toute attaque

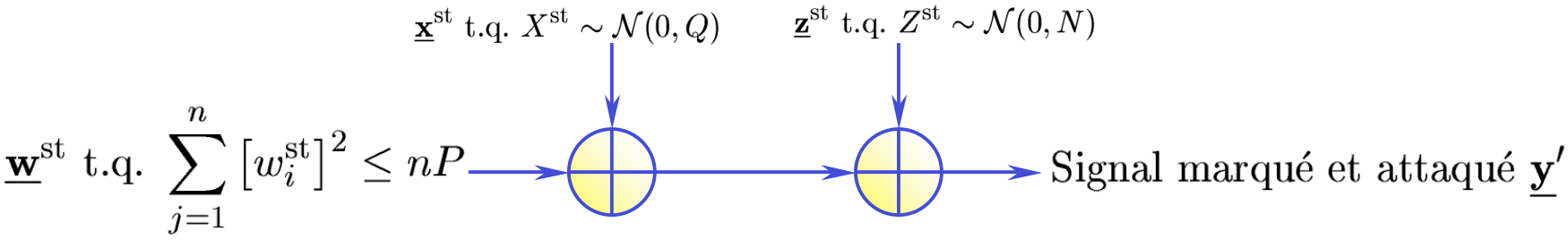


Plan de l'exposé

- Tatouage = problème de communication
  - Définition du canal
  - Optimisation par théorie des jeux
  - Résultats
- Codage du message
  - Rappel du schéma de Costa
  - Dictionnaire partitionné par codes poinçonnés
  - Construction de la marque
- Application pratique : tatouage d'images

Schéma de Costa

Canal de tatouage : rappel



- Transmission d'un signal $\underline{\mathbf{w}}^{\mathrm{st}}$ d'énergie max. P
- Canal bruité par $\underline{\mathbf{x}}^{\mathrm{st}}$ et $\underline{\mathbf{z}}^{\mathrm{st}}$ (énergies Q et N)
- Source $\underline{\mathbf{x}}^{\mathrm{st}}$ connue = information adjacente
- Capacité [Costa83] :

$$C = \frac{1}{2} \log_2 \left[1 + \frac{P}{\cancel{Q} + N} \right]$$

Schéma de Costa
Principe

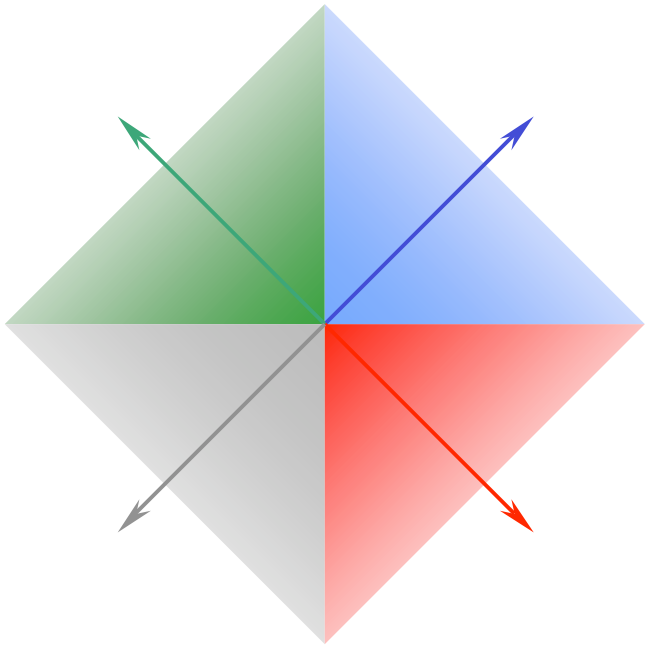
- *Ideal Costa Scheme* : dictionnaire de 2^{nC+i} éléments
  - Dictionnaire $\mathcal{U}$ partitionné en nC sous-dictionnaires $\mathcal{U}_{\underline{\mathbf{m}}}$
  - Message $\underline{\mathbf{m}}$ associé à $\mathcal{U}_{\underline{\mathbf{m}}}$
  - Dimension des $\mathcal{U}_{\underline{\mathbf{m}}} = 2^i$ mots de codes avec

$$i = \frac{n}{2}\log_2\left[1 + \frac{PQ}{(P+N)^2}\right]$$

➔ Partionnement doit être adapté au signal à marquer

Schéma de Costa

Intérêt de la multiplicité



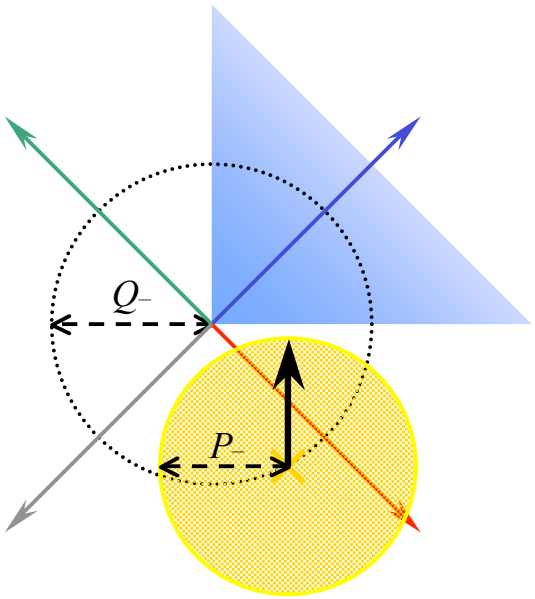
$\mathcal{M} = \{00, 01, 10, \underline{11}\}$

$i = 0$

- Mot de code → zone de robustesse

Schéma de Costa

Intérêt de la multiplicité

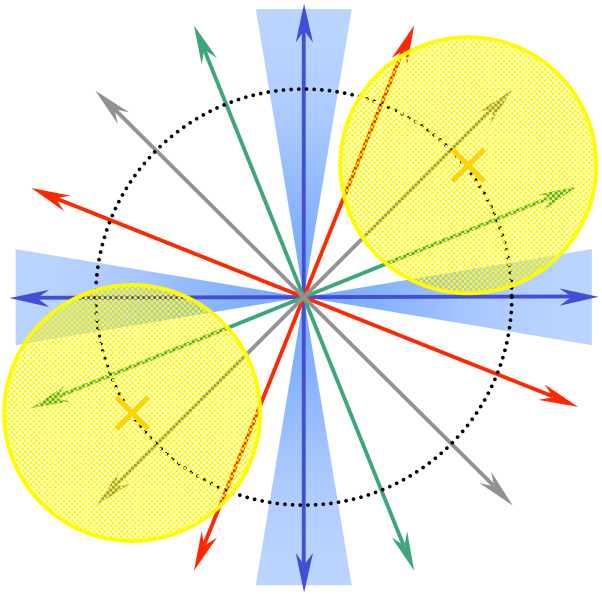


$\mathcal{M} = \{00, 01, 10, \underline{11}\}$

$i = 0$

- Mot de code → zone de robustesse
- Si $Q > P$, on peut ne pas atteindre la zone
  ⇒ erreur de décodage même sans d'attaque

Schéma de Costa :
Intérêt de la multiplicité



$\mathcal{M} = \{00, 01, 10, \underline{11}\}$

$i = 2$

- Mot de code → zone de robustesse
- Si $Q > P$, on peut ne pas atteindre la zone
  ⇒ erreur de décodage même sans d'attaque
- Grâce au dictionnaire partitionné, on est sûr d'atteindre la bonne zone
- i doit être adapté en fonction de Q

Comment construire $\mathcal{V}$?

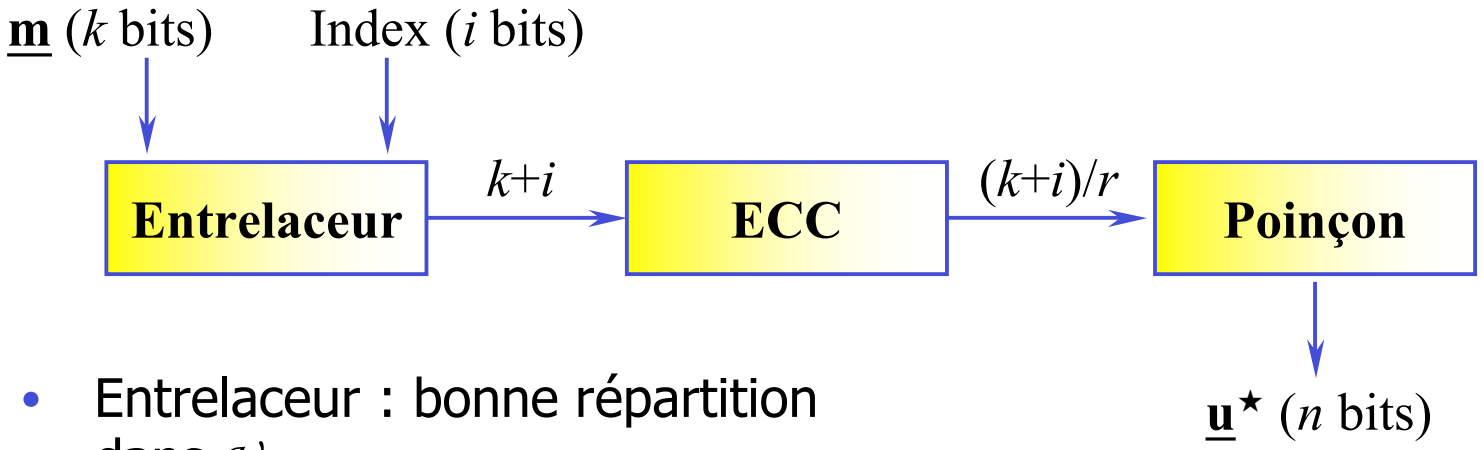
- Propriétés nécessaires
  - Bonne répartition des mots de code au sein de $\mathcal{V}$ et des sous-dictionnaires $\mathcal{V}_{\underline{\mathbf{m}}}$
  - Changement de la structure (valeur de i) aisée pour s'adapter à tout $\underline{\mathbf{x}}^{st}$
- Approches possibles
  - Schéma de Costa : mots de code aléatoires
    $\Rightarrow$ inutilisable en pratique
  - Techniques déjà proposées basées sur ECC : syndromes [Chou01], treillis à multiples chemins [Miller02]
    $\Rightarrow$ mais partitionnement du dictionnaire fixé

➔ Approche par codes correcteurs poinçonnés

Proposition de dictionnaire

Index et codes poinçonnés

- 2^i mots de code par message
  $\Rightarrow$ introduction de i bits d'index supplémentaires

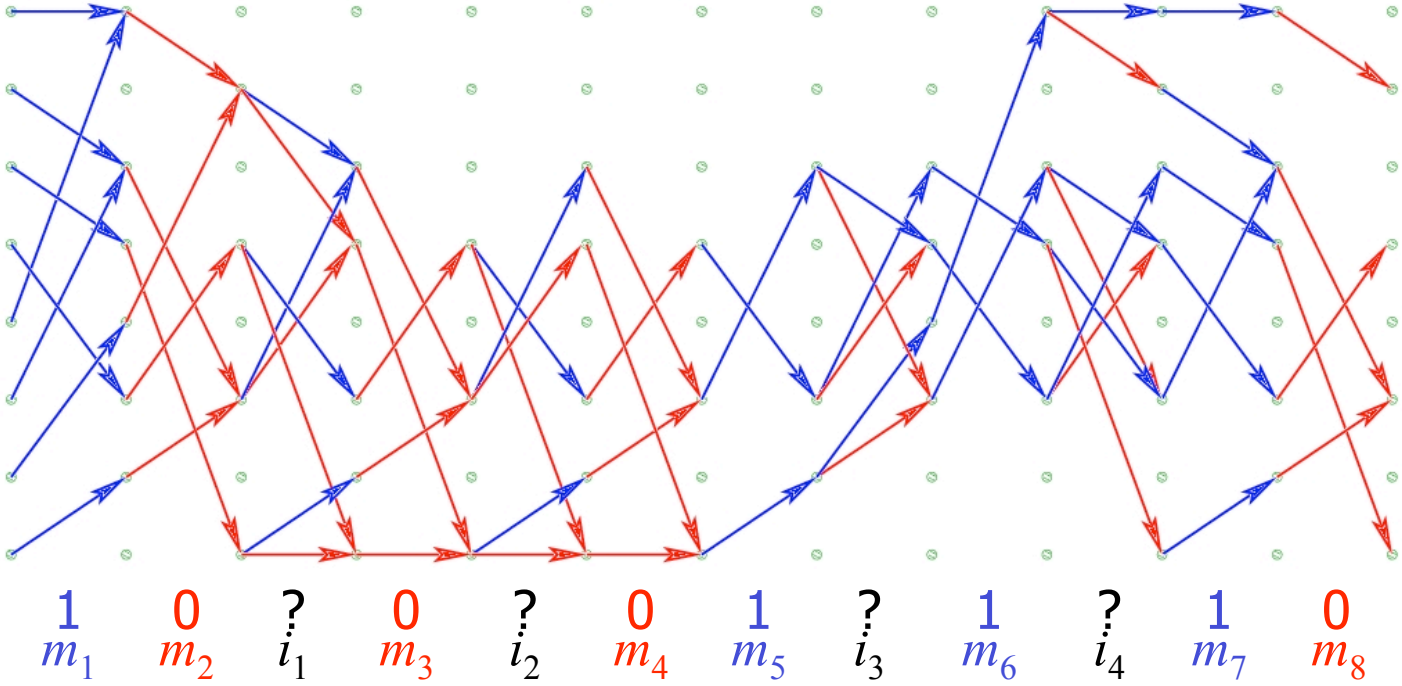


- Entrelaceur : bonne répartition dans $\mathcal{V}_{\underline{\mathbf{m}}}$
- ECC convolutif : bon pouvoir de correction (répartition dans $\mathcal{V}$)
  + décodage rapide par algorithme de Viterbi
- Poinçonnage t.q. k/n constant pour tout i

Proposition de dictionnaire

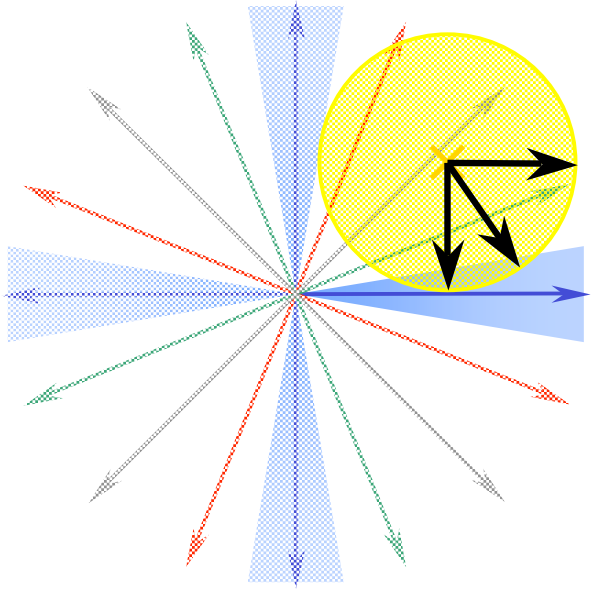
Technique d'encodage

- A l'encodage
  - Recherche de $\underline{\mathbf{u}}^{\star} \in \mathcal{V}_{\mathbf{m}}$ le plus proche de $\underline{\mathbf{x}}^{\mathrm{st}}$
    $\Rightarrow$ trouver le meilleur index
  - Décodage de $\underline{\mathbf{x}}^{\mathrm{st}}$ avec a priori $\underline{\mathbf{u}}^{\star} \in \mathcal{V}_{\mathbf{m}}$



Construction de la marque

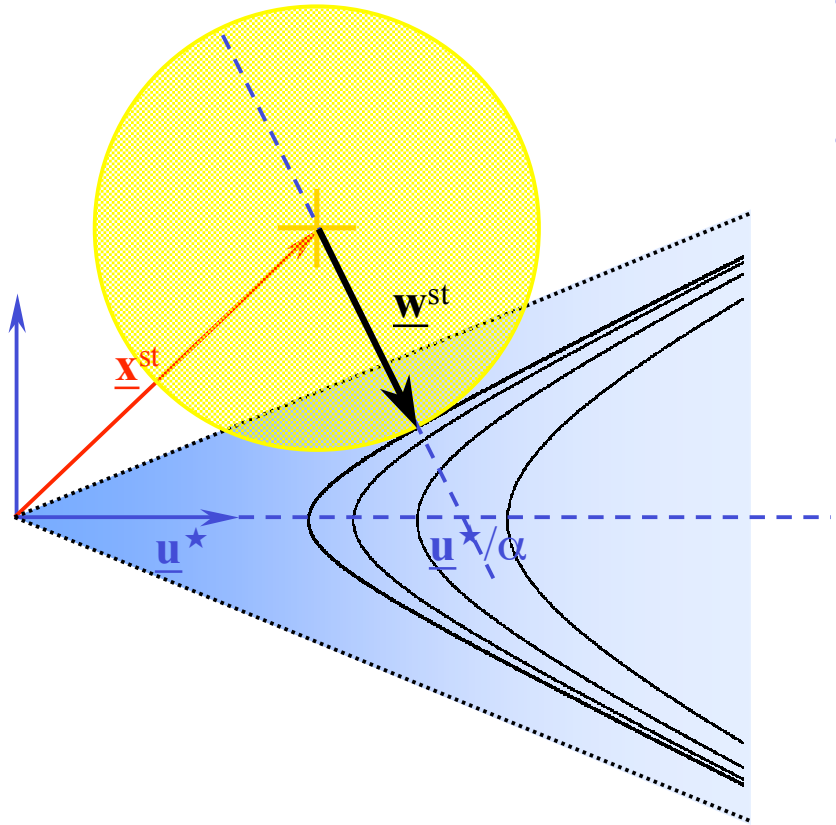
Passage de $\underline{\mathbf{u}}^{\star}$ à $\underline{\mathbf{w}}^{st}$?



- On sait construire $\underline{\mathbf{u}}^{\star}$, reste à trouver le meilleur $\underline{\mathbf{w}}^{st}$

Construction de la marque

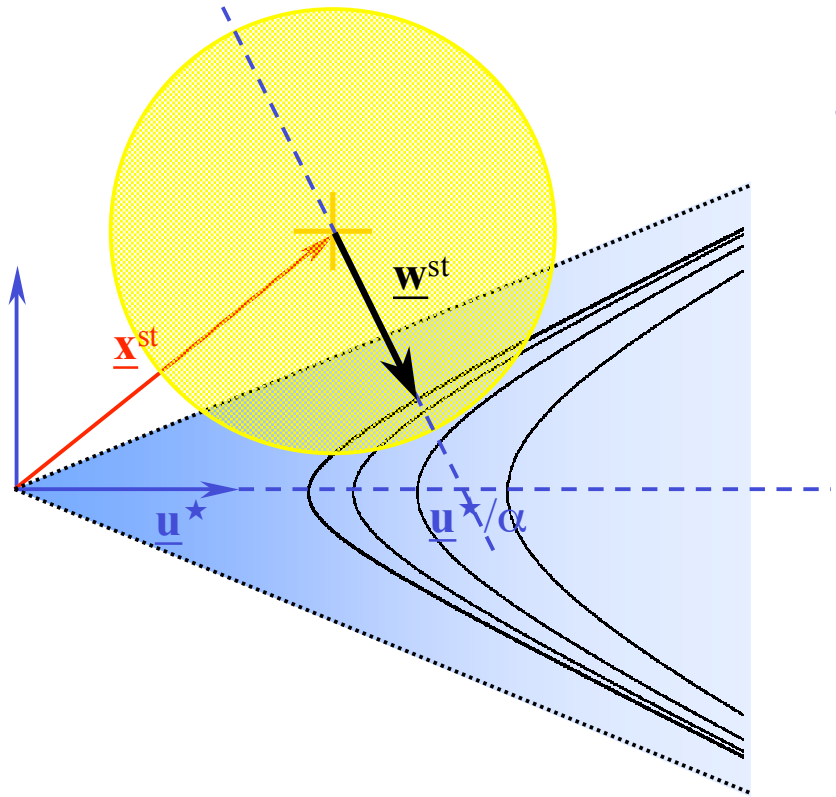
Rappel de Costa



- On sait construire $\underline{\mathbf{u}}^{\star}$, reste à trouver le meilleur $\underline{\mathbf{w}}^{st}$
- Costa pose $\underline{\mathbf{w}}^{st} = \underline{\mathbf{u}}^{\star} - \alpha \underline{\mathbf{x}}^{st}$ avec $\alpha = P/(P+N)$
  $\Rightarrow$ optimal dans le pire cas, assure d'être robuste à un bruit d'énergie N

Construction de la marque

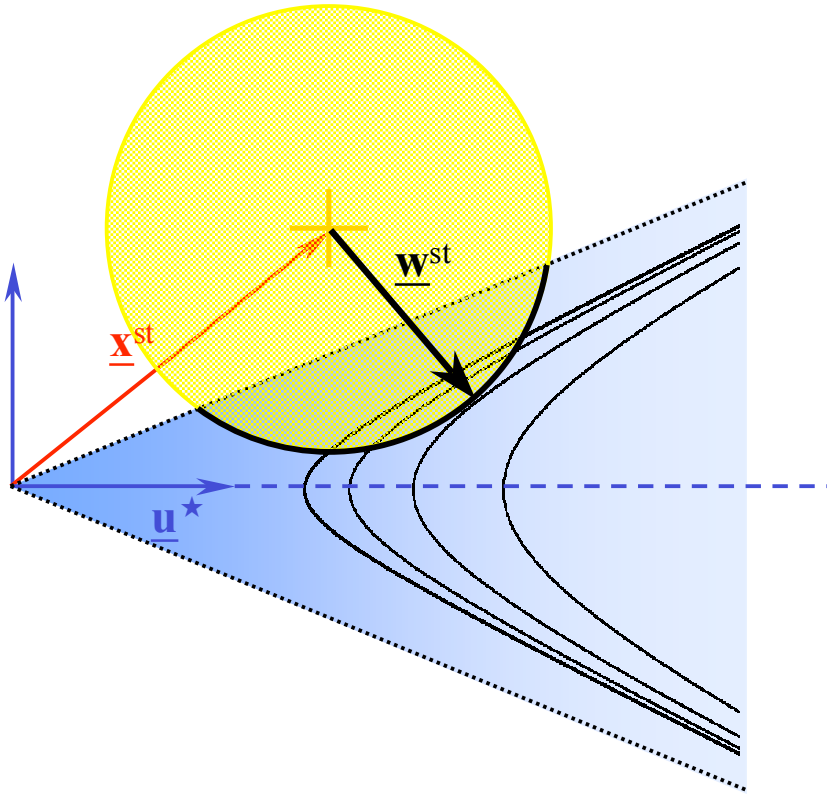
Rappel de Costa



- On sait construire $\underline{\mathbf{u}}^{\star}$, reste à trouver le meilleur $\underline{\mathbf{w}}^{st}$
- Costa pose $\underline{\mathbf{w}}^{st} = \underline{\mathbf{u}}^{\star} - \alpha\underline{\mathbf{x}}^{st}$ avec $\alpha = P/(P+N)$
  $\Rightarrow$ optimal dans le pire cas, assure d'être robuste à un bruit d'énergie N

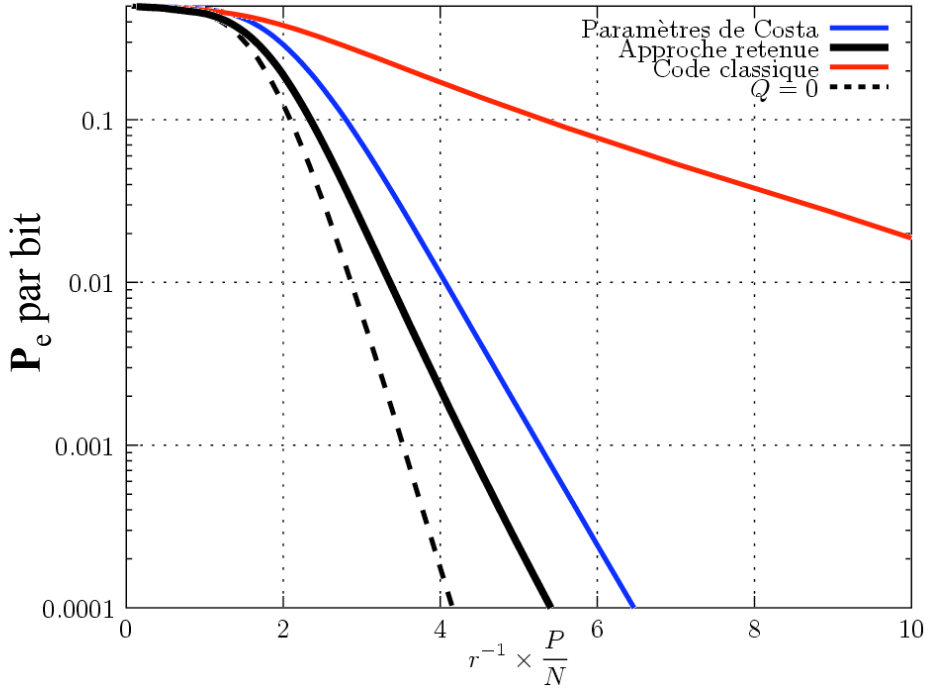
Construction de la marque

Maximiser la robustesse



- On sait construire $\underline{\mathbf{u}}^{\star}$, reste à trouver le meilleur $\underline{\mathbf{w}}^{st}$
- Costa pose $\underline{\mathbf{w}}^{st} = \underline{\mathbf{u}}^{\star} - \alpha \underline{\mathbf{x}}^{st}$ avec $\alpha = P/(P+N)$
  $\Rightarrow$ optimal dans le pire cas, assure d'être robuste à un bruit d'énergie N
- Recherche du point qui maximise la robustesse (généralisation de [Miller00])
- *(Prise en compte de l'interférence inter-symboles)*

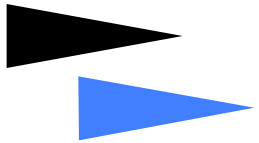
Performances du code proposé



$r = 1/2,\ P = 1$

- Code proposé bien supérieur à un code classique
- Maximisation de la robustesse meilleure que $\alpha = P/(P+N)$ défini par Costa

➔ Performances proches de l'idéal atteignable



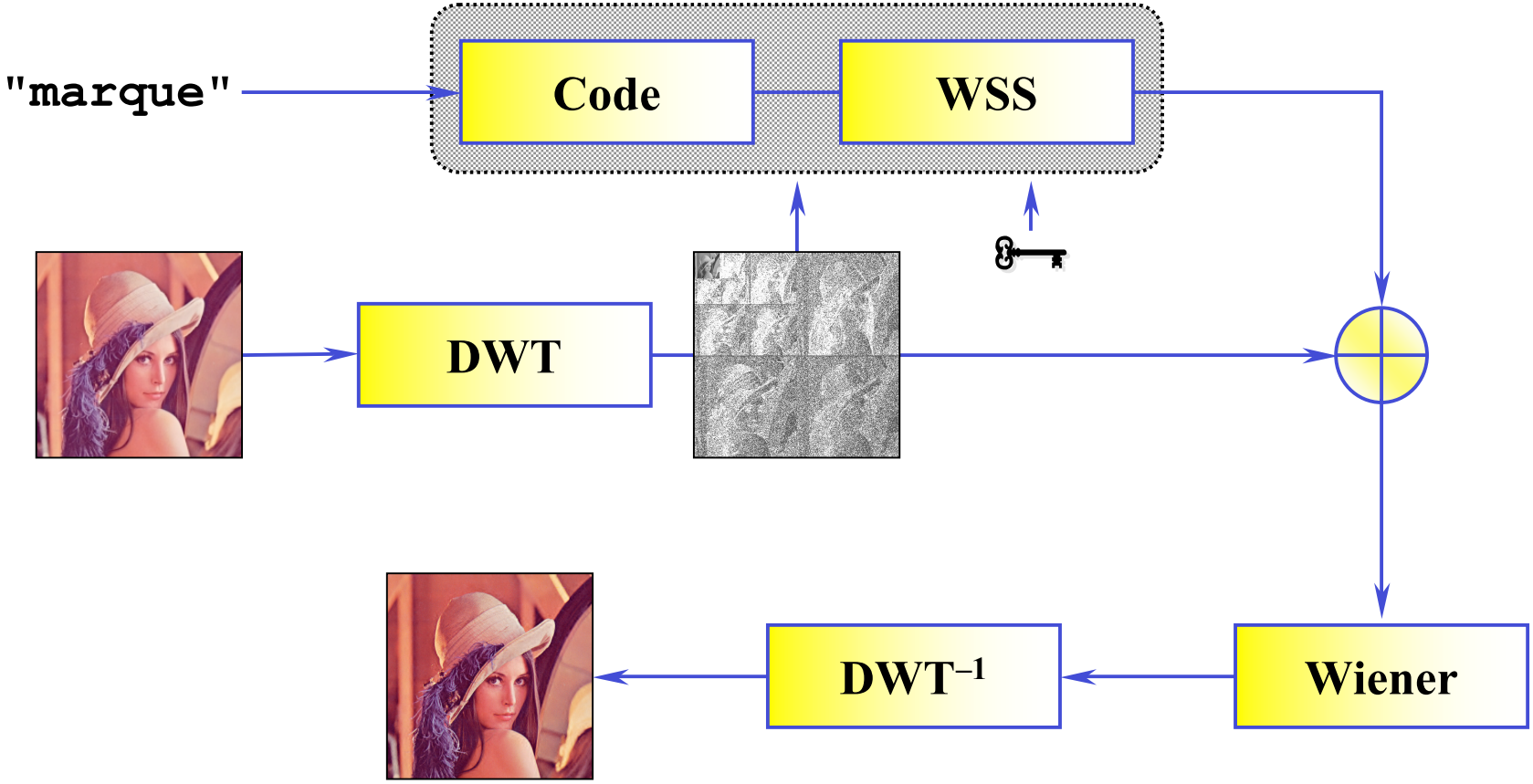
Plan de l'exposé

- Tatouage = problème de communication
  - Définition du canal
  - Optimisation par théorie des jeux
  - Résultats
- Codage du message
  - Rappel du schéma de Costa
  - Dictionnaire partitionné par codes poinçonnés
  - Construction de la marque
- Application pratique : tatouage d'images

Exploitation de nos techniques

- Application au tatouage d'image
- Définition du signal $\underline{\mathbf{x}}$ depuis une image ?
  - Tatouage direct des pixels (domaine spatial)
  - Transformée fréquentielle (Fourier, DCT, ...)
- Quelles performances espérer ?
  - Robustesse et probabilité d'erreur
  - Capacité
  - Qualité visuelle (distorsion introduite)

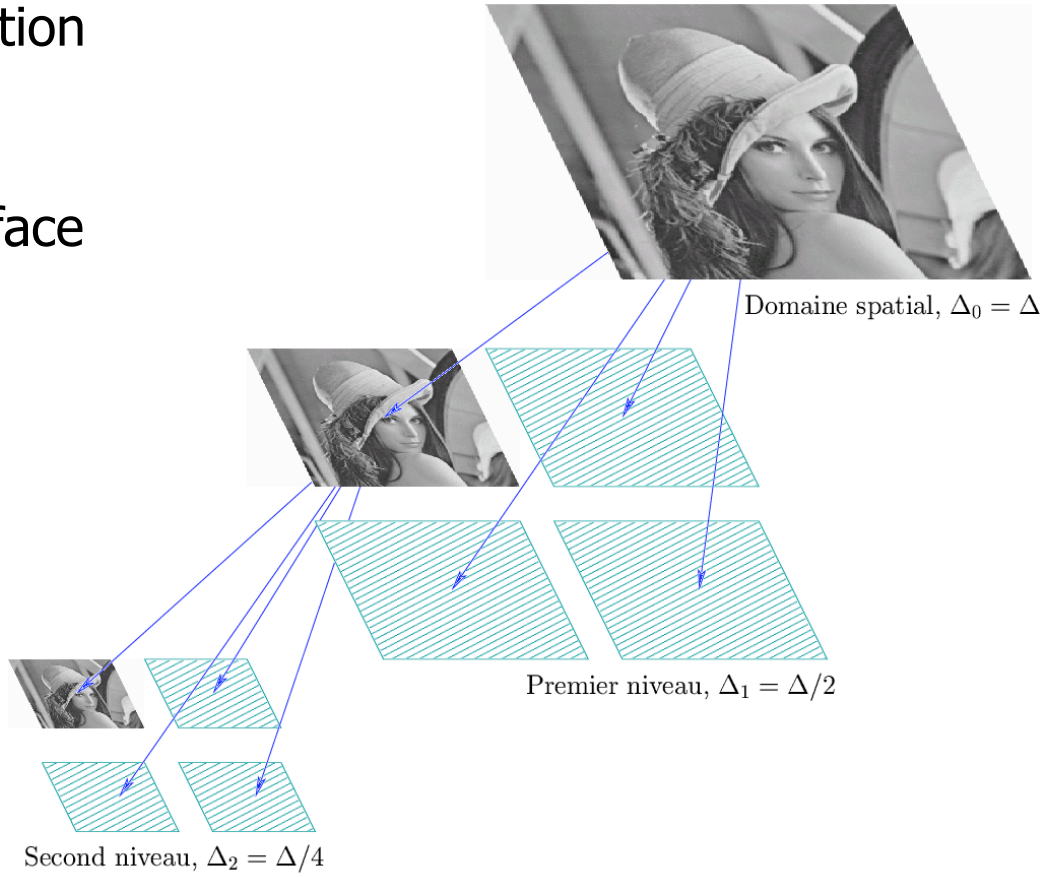
Schéma choisi



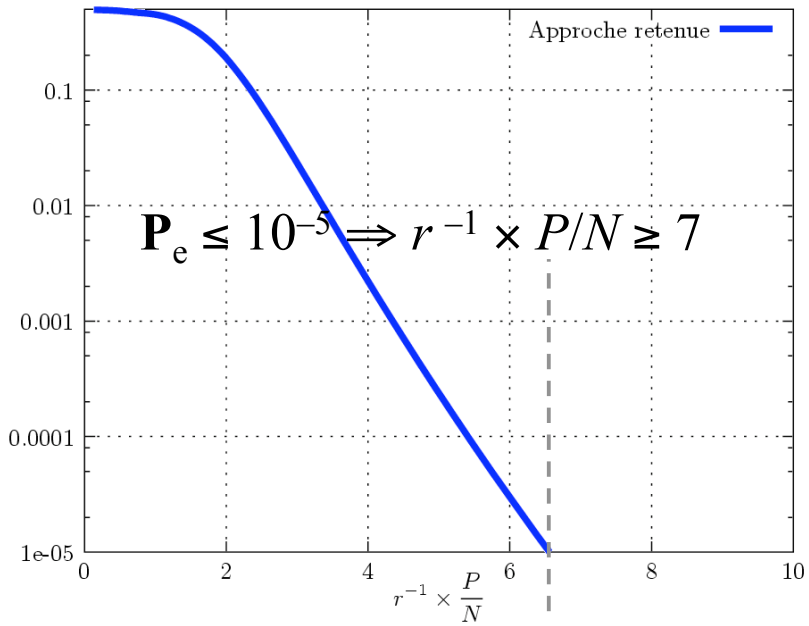
Pourquoi la transformée en ondelettes ?

- Bon pouvoir de décorrélation
- Modèles psycho-visuels adaptés [Taubman00]
- Propriétés intéressantes face aux desynchronisations géométriques
  - 1 pixel dans le domaine spatial ⇒ 1/2 pixel au premier niveau de décomposition
  - …

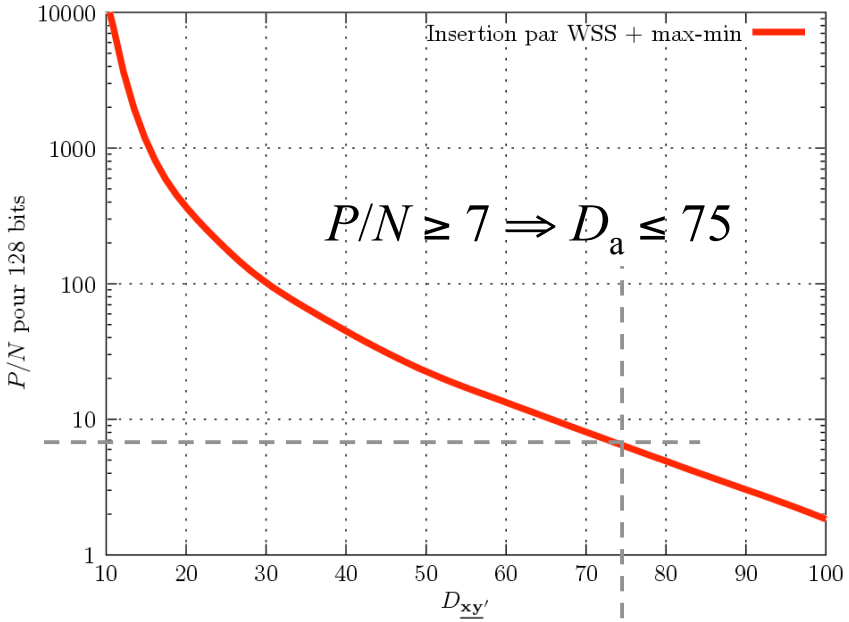
➔ Possibilité d'amélioration du jeu



Point de fonctionnement

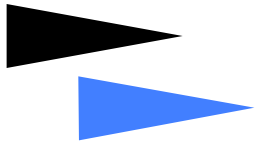


Performance du code



Canal obtenu par max-min (*Lena*)

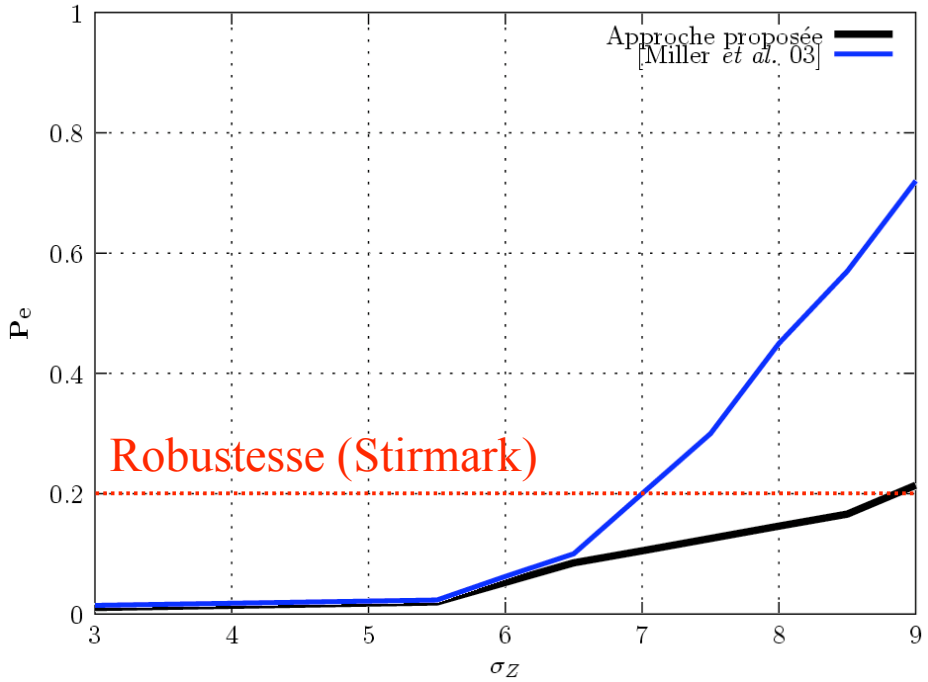
➔ Pour toute attaque de distorsion ≤ 29.4 dB, on assure la transmission de 128 bits avec $\mathbf{P}_e \leq 10^{-5}$



Impact sur les applications

- Gestion de droits
  - capacité moyenne, grande robustesse
  - 128 bits et $D_e = 38$ dB $\Rightarrow$ $D_a < 29.4$ dB (*Lena* 512×512)
- Meta-données
  - forte capacité, robustesse faible
  - $D_e = 40$ dB et $D_a < 35$ dB $\Rightarrow$ 2760 bits
- Détection d'une marque (réponse oui/non)
  - faible probabilité de fausse alarme, grande robustesse
  - $D_e = 38$ dB + $\mathbf{P}_f < 10^{-7}$ (20 bits) $\Rightarrow$ $D_a < 27.9$ dB

Performance face à l'état de l'art



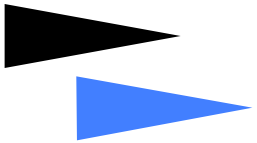
- [Miller03] : tatouage robuste de haute capacité
  - Insertion et codage informés
  - Domaine DCT
- Images 240×368 pixels, insertion de 1380 bits
- Contrainte de distorsion par distance de Watson
- Attaque par ajout de bruit gaussien i.i.d.

Qualité visuelle

- Avec $D_e = 38$ dB et en visant $D_a = 35$ dB
  - Marque bien étalée sur toute l'image
  - Légers rebonds près des contours

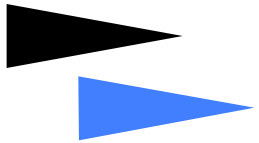
Qualité visuelle

- Avec $D_e = 38$ dB et en visant $D_a = 35$ dB
  - Marque bien étalée sur toute l'image
  - Légers rebonds près des contours
- Avec $D_e = 38$ dB et en visant $D_a = 28$ dB
  - Concentration dans les basses fréquences
  - Rebonds plus visibles



Conclusion

- Un schéma de tatouage…
  - Chaque étape guidée par des bases théoriques
    - Stratégie d'insertion → théorie des jeux
    - Technique de codage → adaptation du schéma de Costa
  - Aux performances potentiellement proches des limites théoriques définies par l'état de l'art
  - Répercussion pratique immédiate
    - Application aux images [RNRT Diphonet]

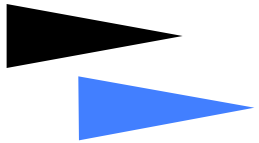


Contributions

Aspect canal

- Optimisation de l'étalement de spectre par théorie des jeux
  - Canal gaussien unique dans un sous-espace linéaire
    ⇒ équilibre du jeu max-min (≠ [Moulin])
  - Filtrage de Wiener à l'insertion
- Extensions inédites du jeu
  - Impact des désynchronisations
  - Prise en compte de la réalisation du signal (attaque *informée*)

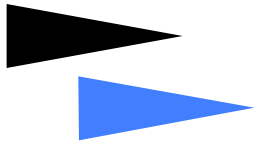
- Permet de justifier plusieurs techniques empiriques
  - Attaque par ajout de bruit et annulation [Su01]
  - Marquage des échantillons perceptuellement importants [Cox97]
  - Privilégier les basses fréquences pour les attaques géométriques [Kalker01]



Contributions

Aspect codage

- Codes partitionnés par poinçonnage et bits d'index
  - Très grande flexibilité (≠ [Pradhan99])
  - Encodage et décodage aisé et rapide
  - Performances très supérieures à un code classique, proches de l'idéal de Costa
  - Adaptable à d'autres types de code
- Maximisation de la robustesse
  - Généralisation de [Miller00]
  - Gain de performance supplémentaire
- Suppression de l'interférence inter-symboles
  - ISI = information adjacente
    ⇒ Gain de performance important



Perspectives

- Extension du principe de structuration à des codes plus performants
- Comment transmettre i, paramètre de structuration
  - Marque supplémentaire ?
  - Introduction dans l'a priori lors de l'encodage ?
- Mesure de distorsion moyenne $\Rightarrow$ possibilité de concentration de la marque et donc perceptibilité
  - Ajout d'une mesure JND (*just noticeable difference*)
- Développement/intégration d'un schéma de resynchronisation afin de résister aux attaques géométriques