Herbstworkshop des Graduiertenkollegs 466
Rügen, 2.-4. Nov. 2000

Digital Watermarking:
Theoretic Foundation and Applications

J.J. Eggers

Telecommunications Lab
Univ. of Erlangen-Nuremberg
eggers@LNT.de

Digital Watermarking in Erlangen

- Hartung, Girod `96:
  - first work on digital video watermarking
- Su, Eggers, Girod `98:
  - digital watermarking of multimedia documents
  - theoretic framework
  - part of the DFG digital library initiative V^3D^2
- Eggers, Bäuml, Huber `00/01:
  - V^3D^2 project continues ...

Overview

- Introduction into digital watermarking
- Theoretic framework and its application
- Example application: image watermarking

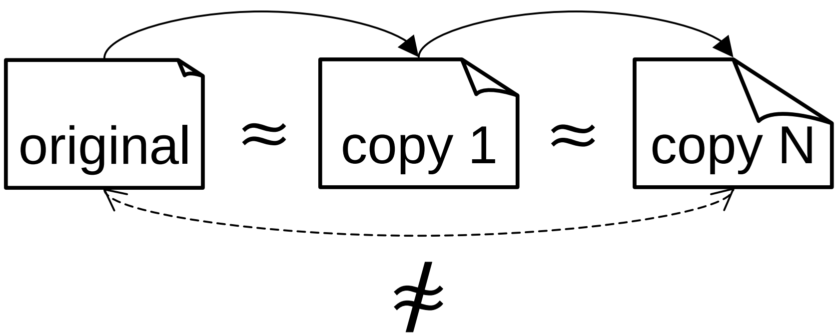
INTRODUCTION

- Motivation
- Definition of watermarking
- Desired properties
- Limitations
- Spread-Spectrum Watermarking

Analog and Digital Multimedia

Analog Media

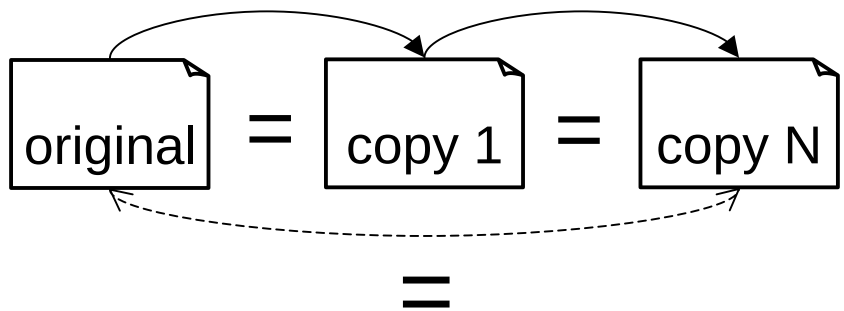
photocopies
audio cassettes
photographs
VHS videotapes



- "Built-in" protection against copying and redistribution
- Distribution net required

Digital Media

ASCII, PostScript, PDF
CDs, MP3 audio
JPEG images
DVDs, MPEG video



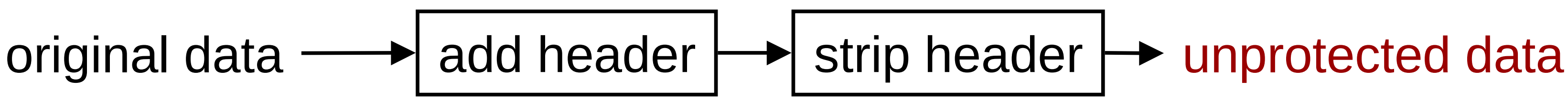
- **No inherent protection** against copying and redistribution
- **"Free" distribution net**: Internet

Unauthorized Use of Digital Data

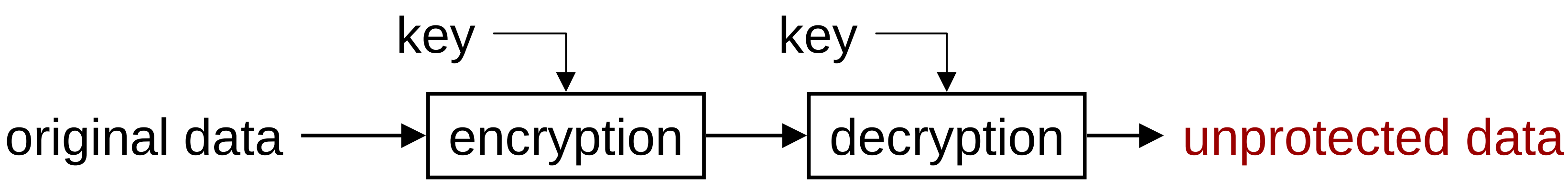
- Digital multimedia
  - can be stored, copied, and distributed easily, rapidly, and with no loss of fidelity
  - can be manipulated and edited easily and inexpensively
- Are these properties always advantageous?
  - Some Hollywood studios will not release DVDs unless copyright protection can be ensured
  - USA Today, Jan. 2000: Estimated lost revenue from digital audio piracy: US $8,500,000,000.00
  - Recent examples: MP3.com, Napster

Traditional Methods of Protecting Data

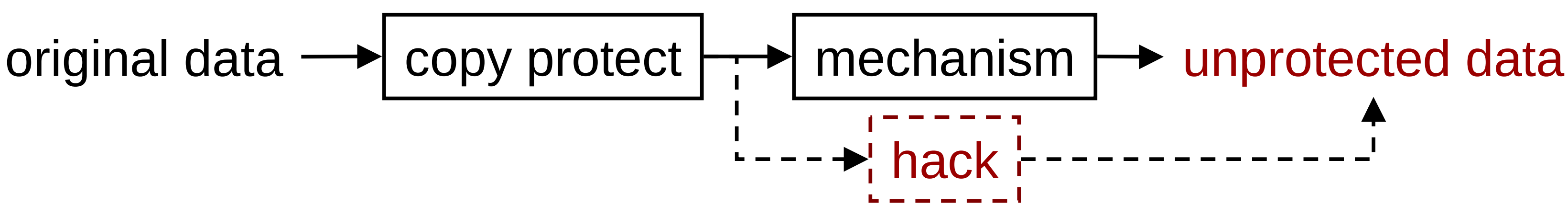
- Access-control headers: easily removed/altered



- Encryption: decrypted data unprotected



- Copy protection: susceptible to hacking

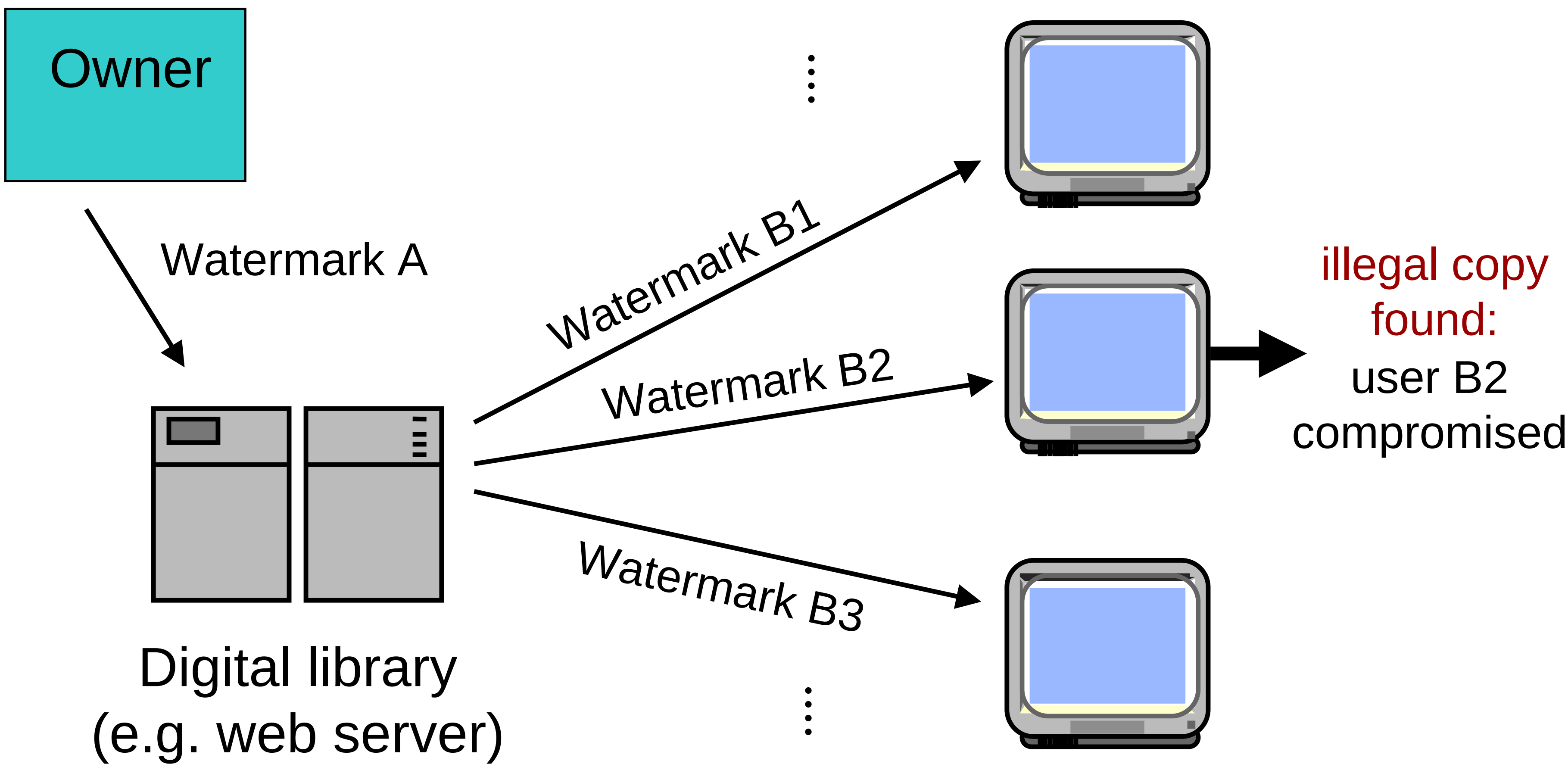


Motivation for Digital Watermarking

- Principle: Embed information that travels with the watermarked data, wherever it goes

- “last line of defense”
- loosely analogous to watermarks in paper

Example: Distribution from a Library

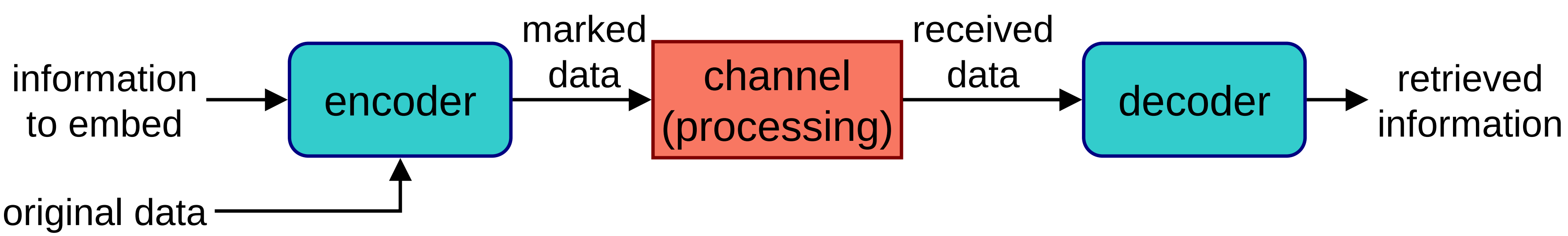


Watermarking Applications

- Access control
  - playback, copy-generation control (DVD)
  - copyright protection, proof of ownership
- Distribution tracing
  - fingerprinting
  - identification of compromised parties
- Broadcast monitoring
- Media authentication (fragile watermarking)
- Covert communication (steganography)
- Added value via meta-information
  - e.g., SmartImages by Digimarc Corp. [Alattar 2000]

“What is digital watermarking?”

- Digital Watermarking:
  - The *imperceptible, robust, secure communication* of information by embedding it in and retrieving it from other digital data.
- Watermarked data is likely to be processed
  - view processing as a communications channel

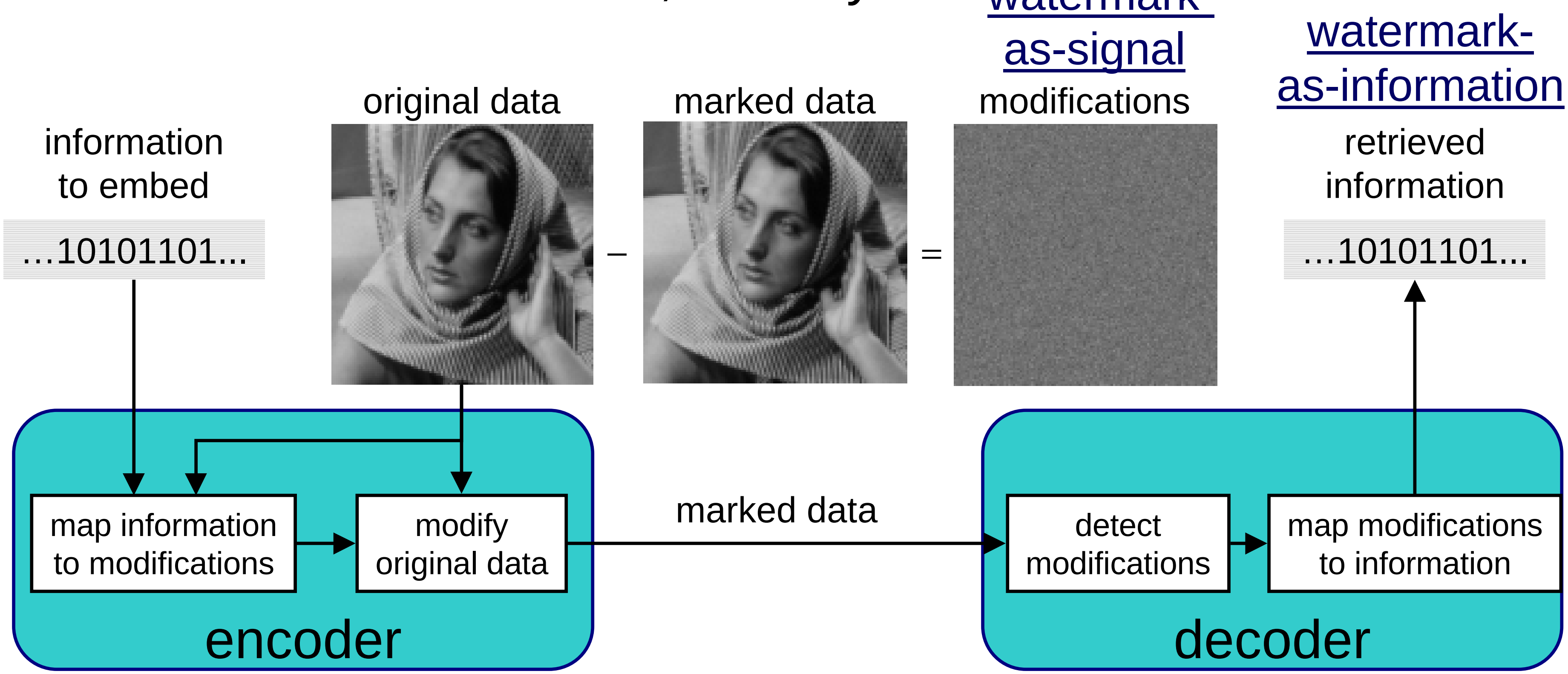


Desired Properties

- Imperceptibility:
  - Watermarked data and original data should be perceptually indistinguishable
- Robustness:
  - Processing of the watermarked data cannot damage the watermarks without rendering the processed data useless
- Security:
  - Watermarks cannot be detected, read, and/or modified by unauthorized parties
  - **Kerckhoff's principle**: Security resides in the secrecy of the key, <u>not</u> in the secrecy of the algorithm**.**

Two Basic Questions

- How can information be hidden in digital data?
- What is the watermark, actually?



Additional Aspects

- "Blind" watermarking
  - no reference to original data during decoding
  - possible interference from original data
- Multiple watermarks
  - one copy with several information streams
  - different information in different copies
- Compressed-domain processing
  - combined watermarking and compression
  - bit-rate constraint
- Implementation concerns
  - speed, computational load, footprint, cost

Limitations

- Digital watermarking does not prevent copying or distribution by itself
  - but embedded information remains in copied data
- Digital watermarking alone is not a complete solution for access/copy control or copyright protection!
- Digital watermarking is a part of a larger system for protecting digital data against unauthorized use

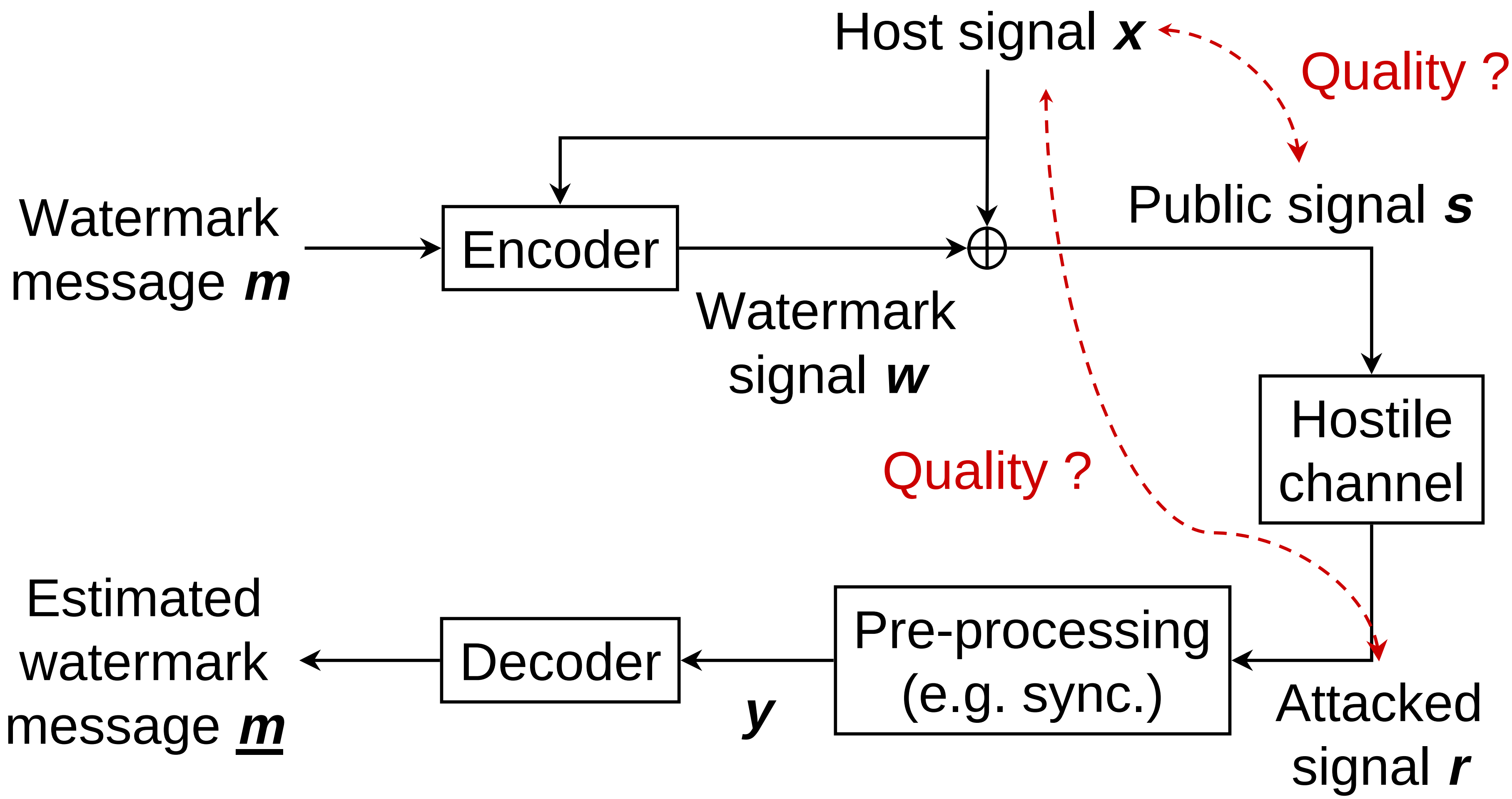
Spread-Spectrum Watermarking

- Most popular watermarking technique
  - add pseudo-noise sequence to the host data
  - detection: correlation of pseudo-noise sequence and received data
- Advantages
  - pseudo-noise sequence = secret key
  - for long signals correlation detection is very reliable
- Disadvantages
  - low rate of embedded watermark
  - detector must be synchronized

Theoretic Framework

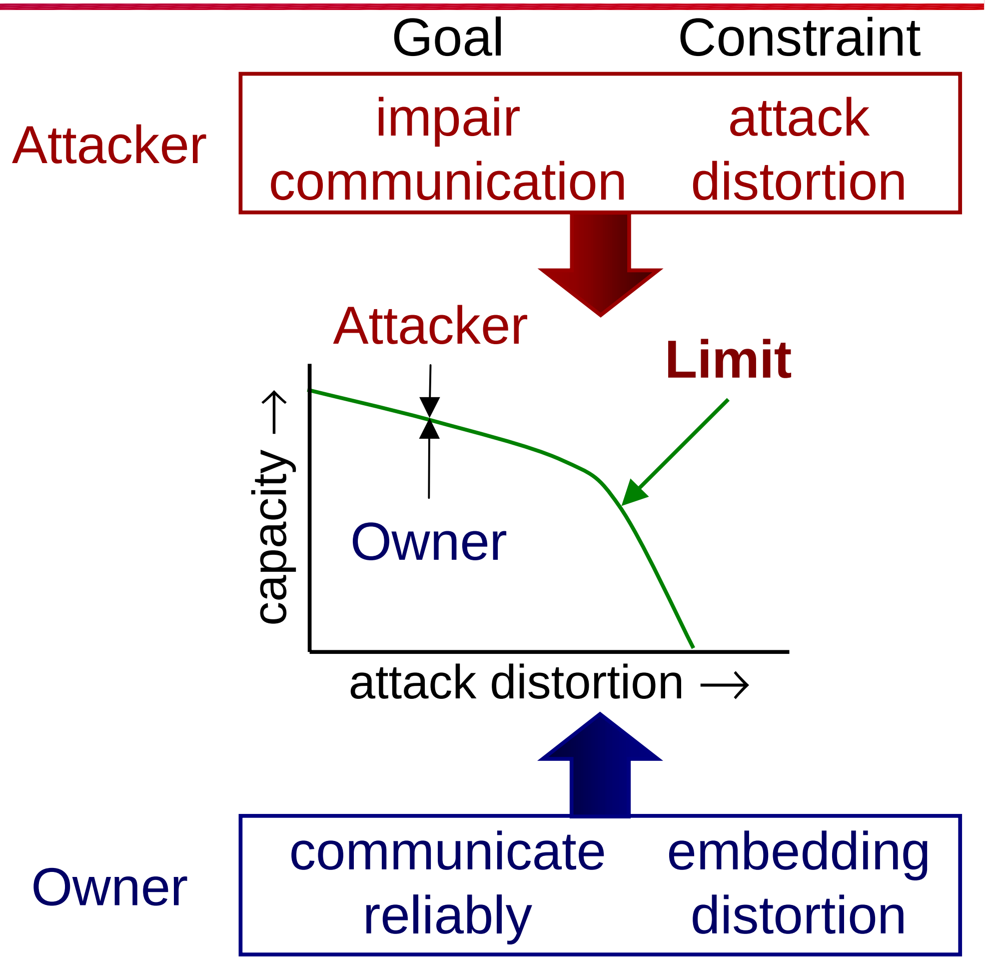
- General model of blind watermarking
- Theoretic results for white host signals
- Performance of practical schemes
- Colored signals

Model for Blind Watermarking



Watermarking as a Game

- Game between embedder and attacker
  - payoff: communication rate
  - penalty: embedding and attack distortion
- Information-theoretic limits
  - results apply to all data that satisfy assumptions
  - robustness well-defined
  - provable watermarking guidelines



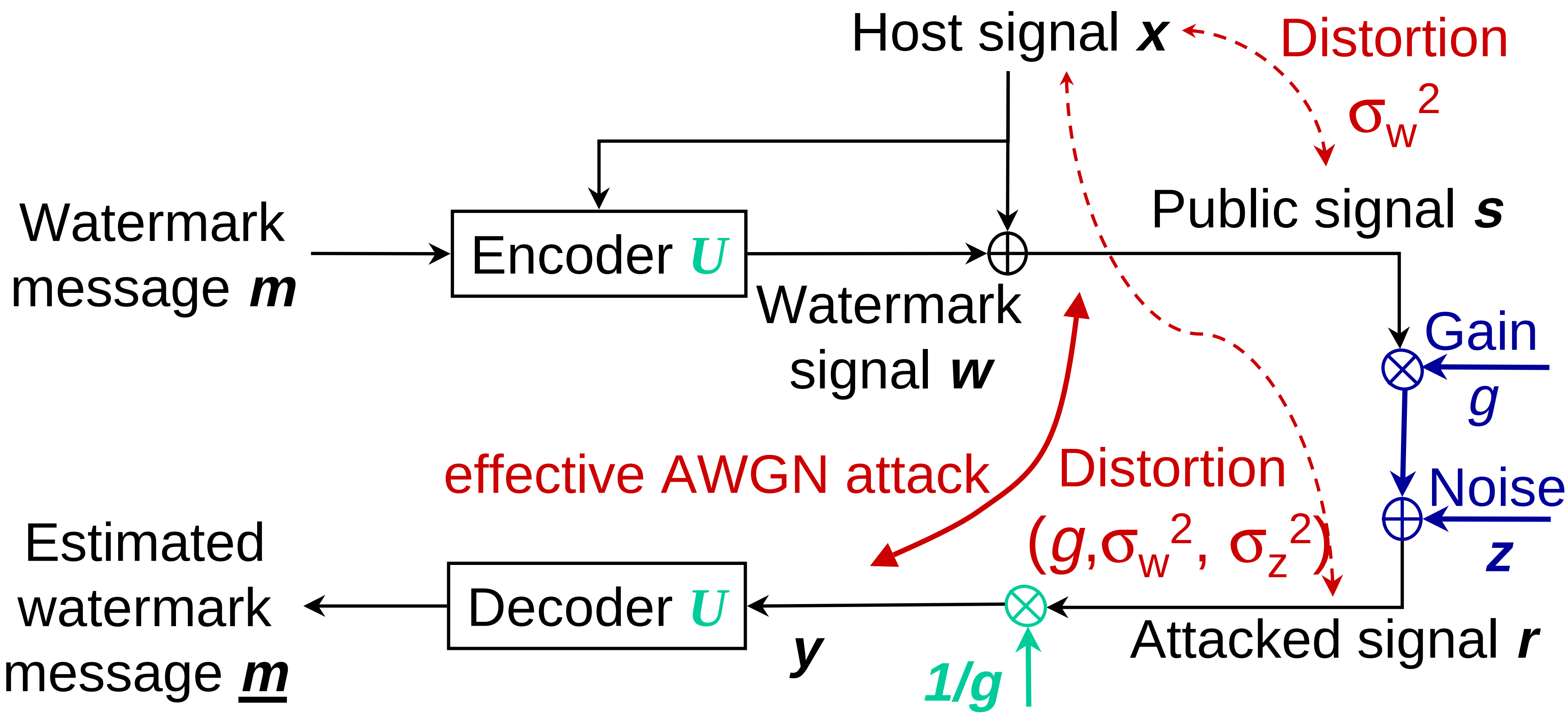
White Gaussian Host Signals

- Assumption
  - mean-free, white Gaussian host ***x***
  - mean squared-error distortion measurement
- Information theoretic result [Moulin, O'Sullivan '99]
  - optimal attack: **Gaussian test channel (GTC)**

  - capacity:

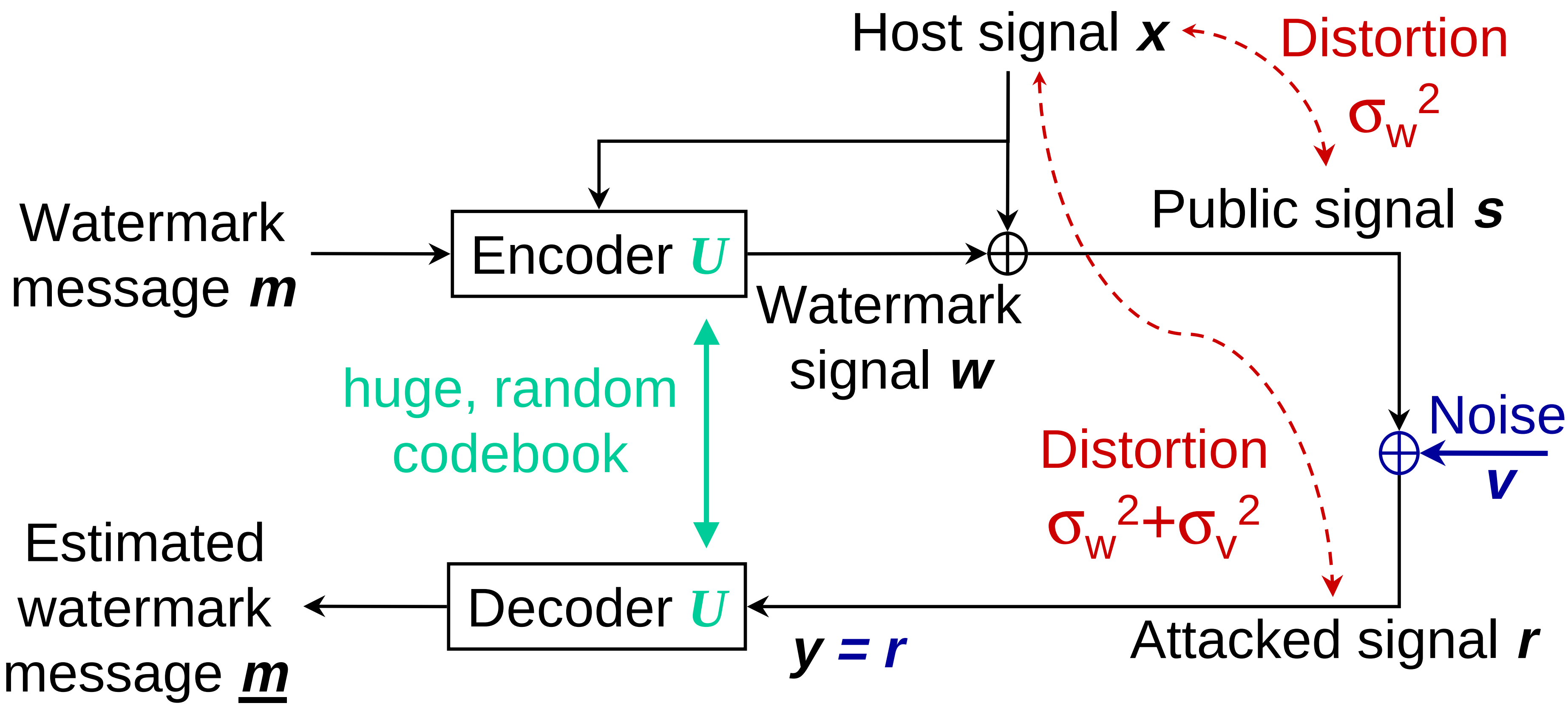
$$C = \begin{cases} 0.5 \log\left(1 + \dfrac{D(\boldsymbol{x},\boldsymbol{s})}{\beta D(\boldsymbol{x},\boldsymbol{r})}\right) & \text{if} \quad D(\boldsymbol{x},\boldsymbol{r}) < \sigma_{\boldsymbol{x}}^2 + D(\boldsymbol{x},\boldsymbol{s}) \\ 0 & \text{else} \end{cases}$$

GTC = Effective AWGN Channel



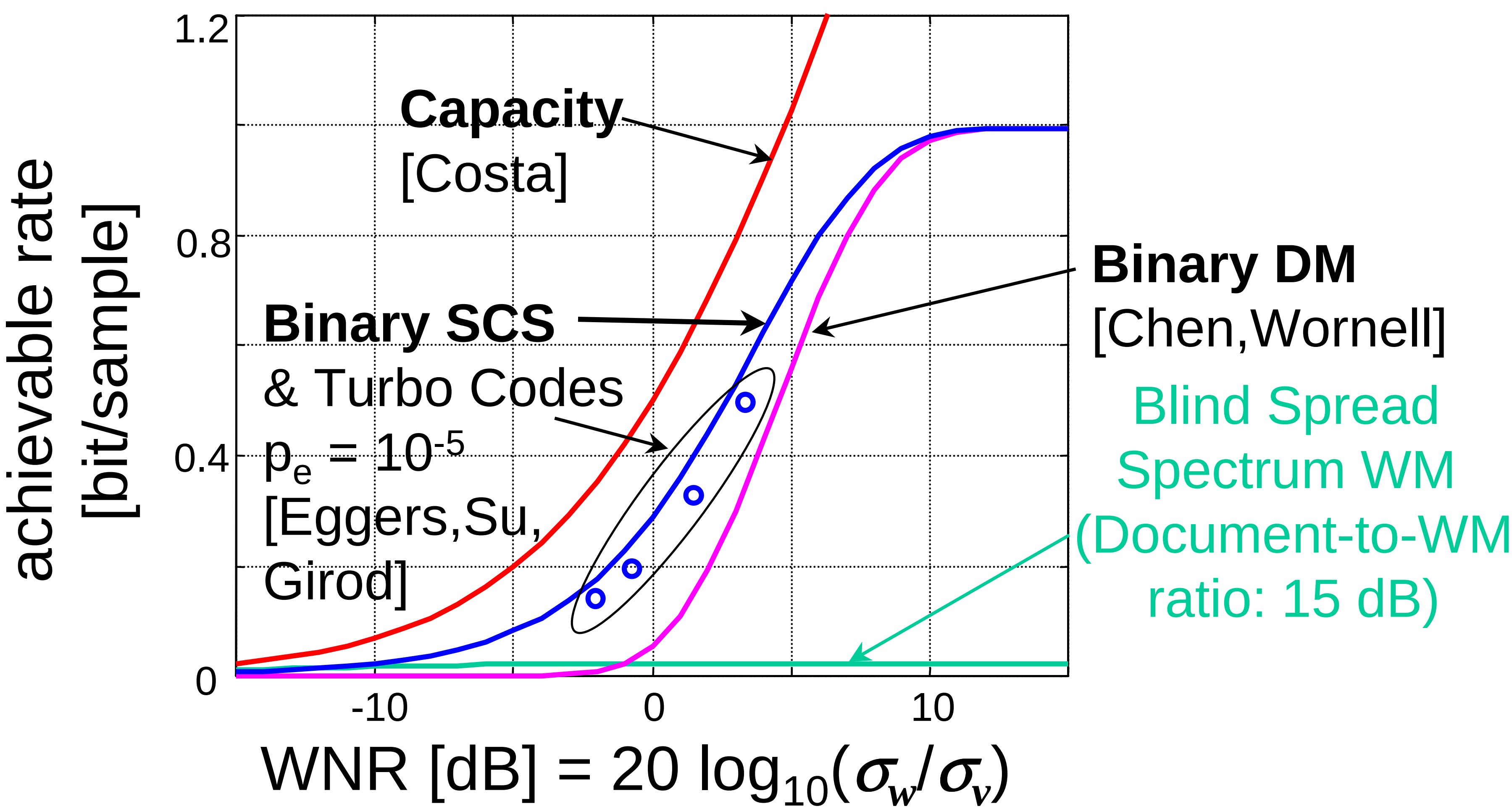
Design U for effective channel noise $\boldsymbol{v} = \boldsymbol{z}/g$!

White Signals & AWGN Attack



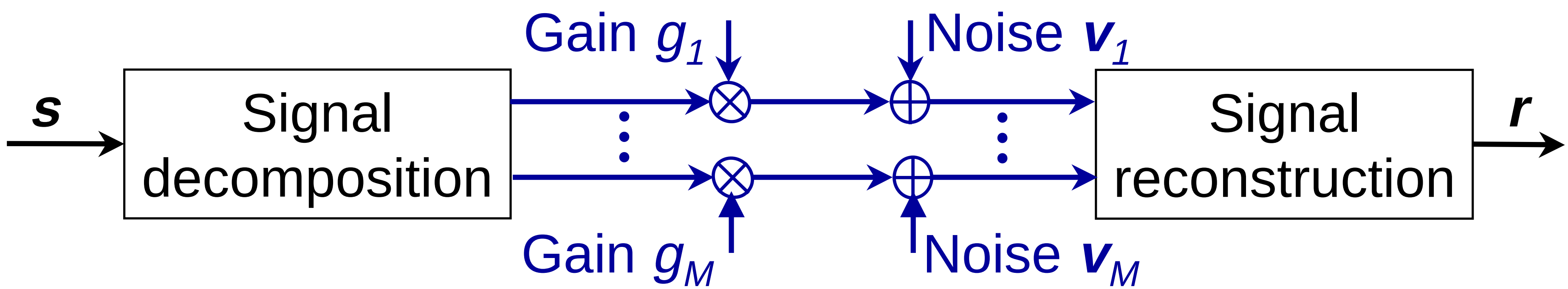
Costa (´83): Capacity is independent from host ***x*** !

Achievable Rate of Blind Schemes



Colored Host Signals
Linear Filtering & Additive Noise

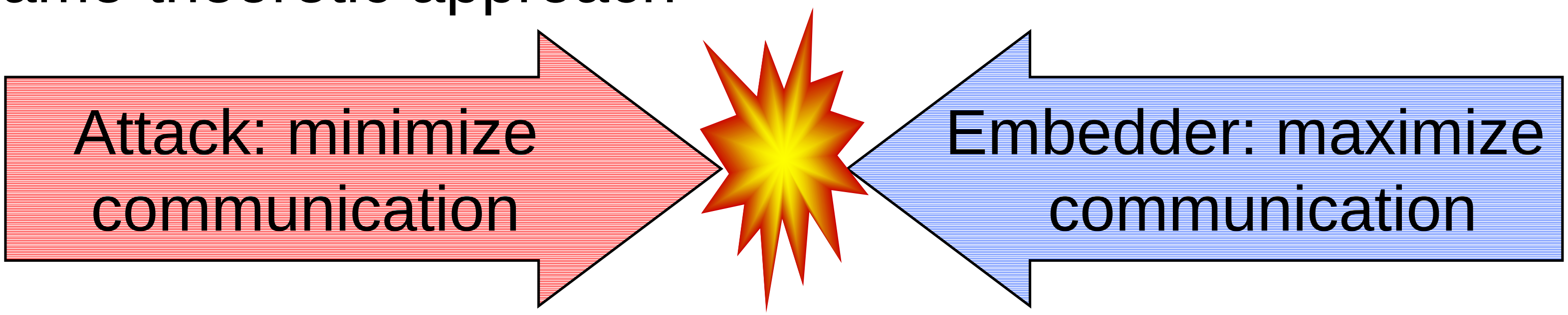
- Decompose host signal
  - *M* approximately independent sub-channels
  - white signal statistics within sub-channel
- Linear filtering & additive noise attack



- Watermark communication over parallel channels

Optimized Embedding and Attack

- Game-theoretic approach



Embedder: allocation of watermark power?

Attacker: allocation of attack distortion (filter, noise)?

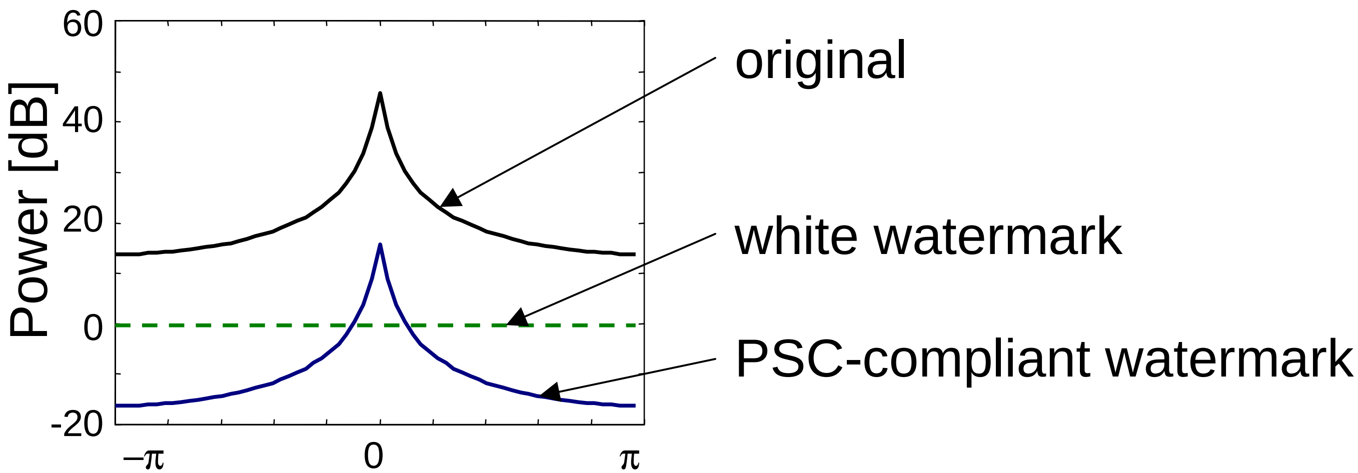
- Apply Kerckhoff's Principle

Attacker and embedder know their opponent's behavior

- Note: Waterfilling rule does not apply here!

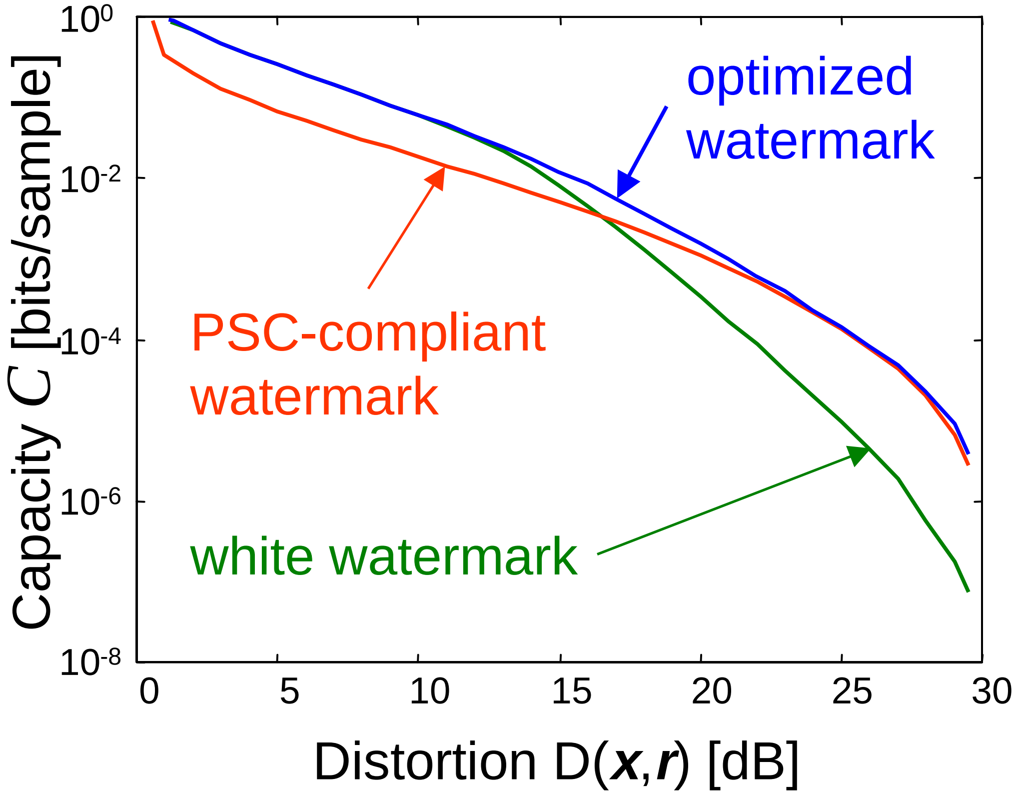
Optimum Power Allocation

- Optimum attack: complicated equations
- Optimized defense: iterative numerical optimization
- Example Analysis with autoregressive signal model
  - original data: lowpass
  - various watermarks (white, PSC-compliant, optimized)



Rule-of-Thumb for Robustness

- **No unique solution over entire distortion range!**
- Low distortion: **white**
  - attack ~ "add noise"
  - force attack to spread its power over all channels
- High distortion: **PSC**
  - attack ~ "throw away"
  - attack cannot discard watermark without also destroying original

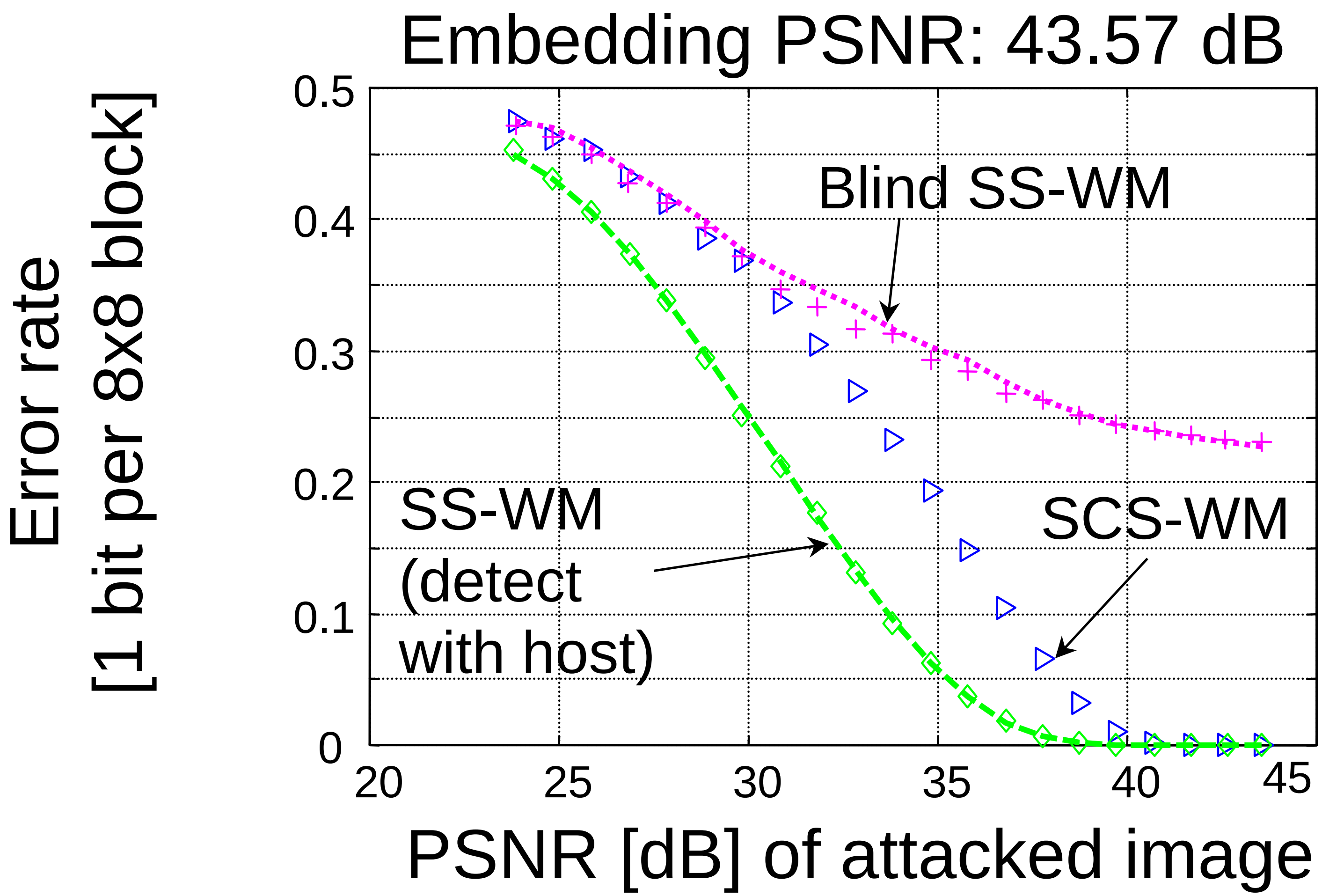


Example: Image Watermarking

Blind Image Watermarking

- Decomposition
  - 8x8 block DCT (64 sub-channels)
- Moderately strong filtering attack
  - white watermark power allocation over first 21 DCT coefficients in zig-zag-scan
- Simulations with
  - blind spread spectrum watermarking (SS-WM)
  - blind SCS watermarking
  - spread spectrum watermarking (detect with host)

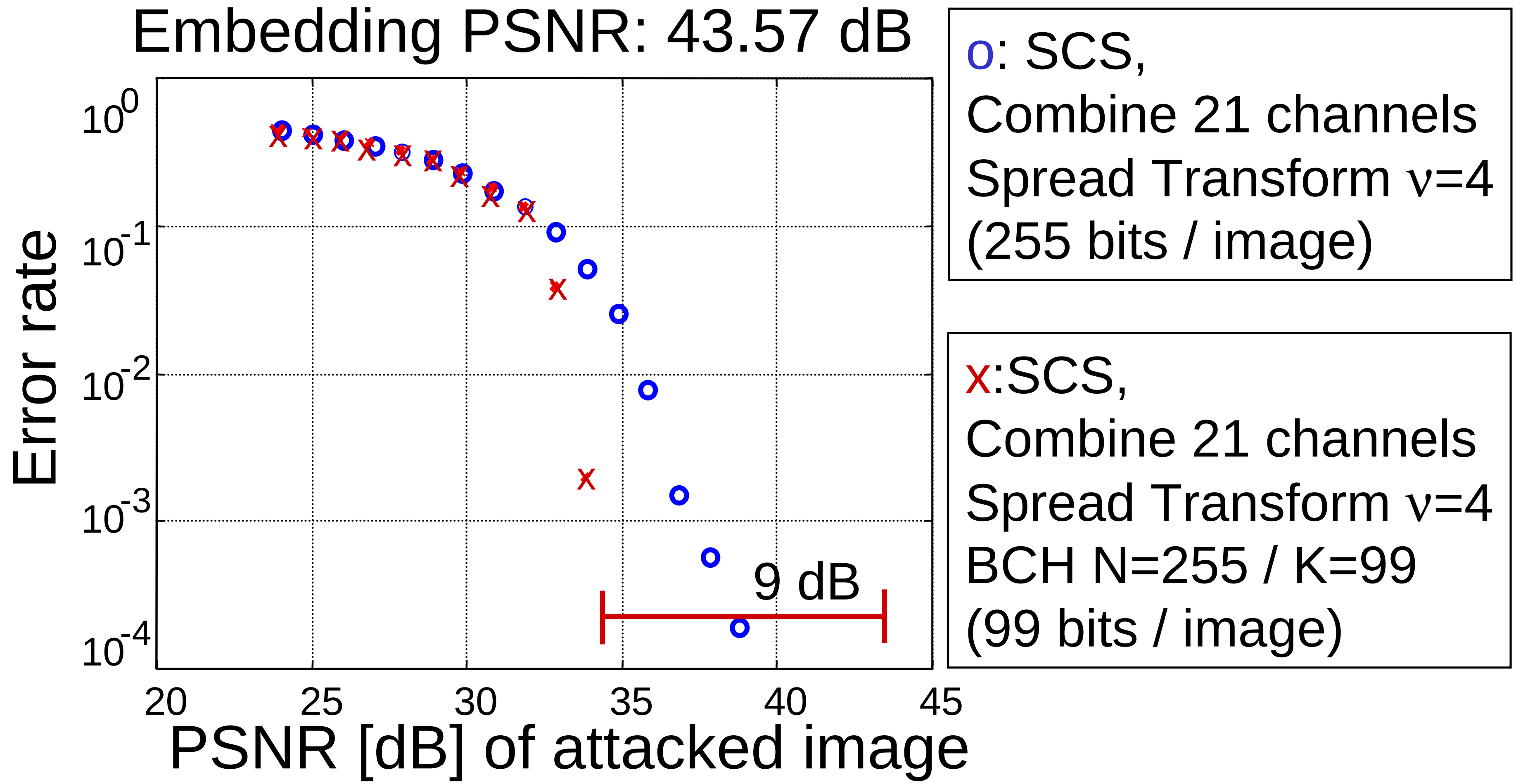
“Uncoded“ Image Watermarking



Test image
Lenna
grayscale
256 x 256

Attack
optimized
filtering &
additive noise

Coded SCS Image Watermarking



Summary

- Motivation for digital watermarking
  - illegal use of digital data
  - added value
- Theoretical framework for watermarking emerges
  - new blind watermarking technology
  - allocation of embedding/attack distortion
- Open problems
  - efficient capacity achieving watermarking
  - efficient synchronization algorithms
  - robustness dependent on host PDF