Steganographie in Rechnernetzen

Tutorium “Sicherheit in Netzen” der
13. Arbeitstagung des DFN

Dr. Hannes Federrath
Technische Universität Dresden

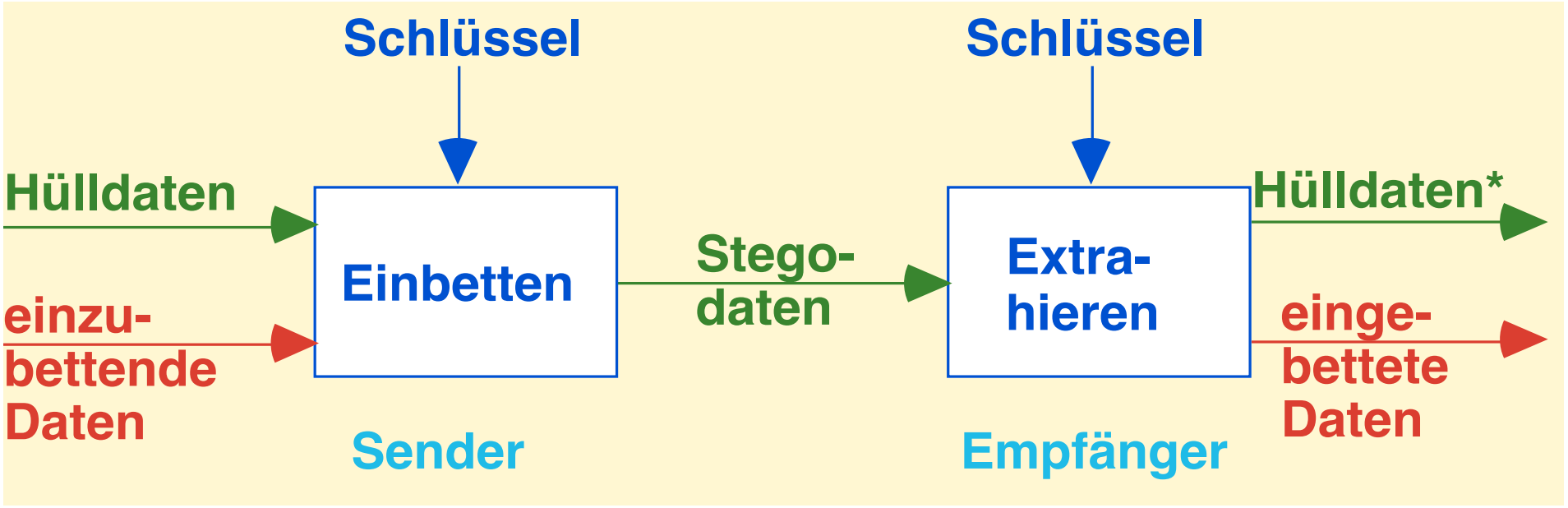
Was ist Steganographie?

- **Definition**
  - geheimzuhaltende Nachricht wird in eine Hülle eingebettet
  - minimale Veränderungen kaum bzw. nicht erkennbar
  - Veränderungen nicht mit Meßmethoden nachweisbar

- **Steganographie ist technologisch gesehen keine Verschlüsselung von Daten**
  - Kryptographie: Klartext $\xrightarrow{f}$ Schlüsseltext
  - Steganographie: Hülle $\xrightarrow{g}$ Hülle*

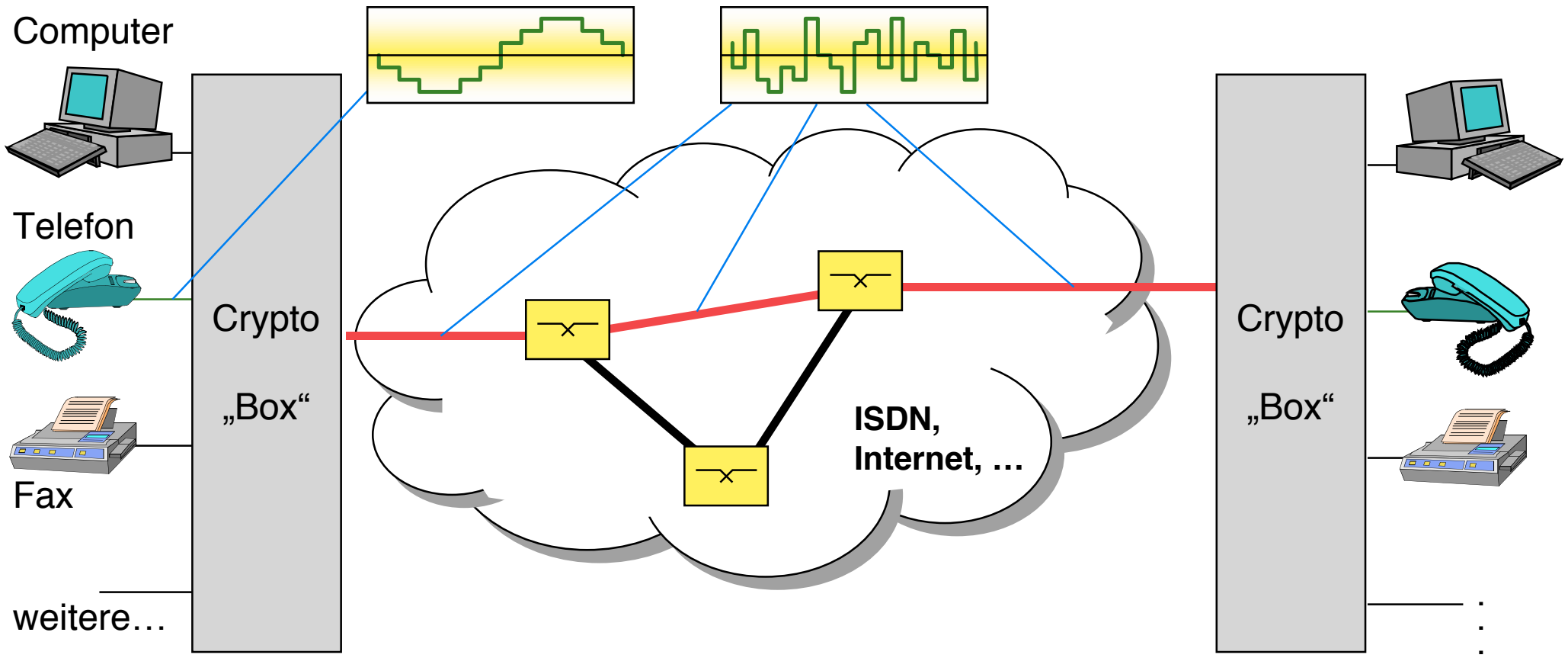
- **verschiedene Systeme für unterschiedliche Medien**

Aufbau eines Stegosystems



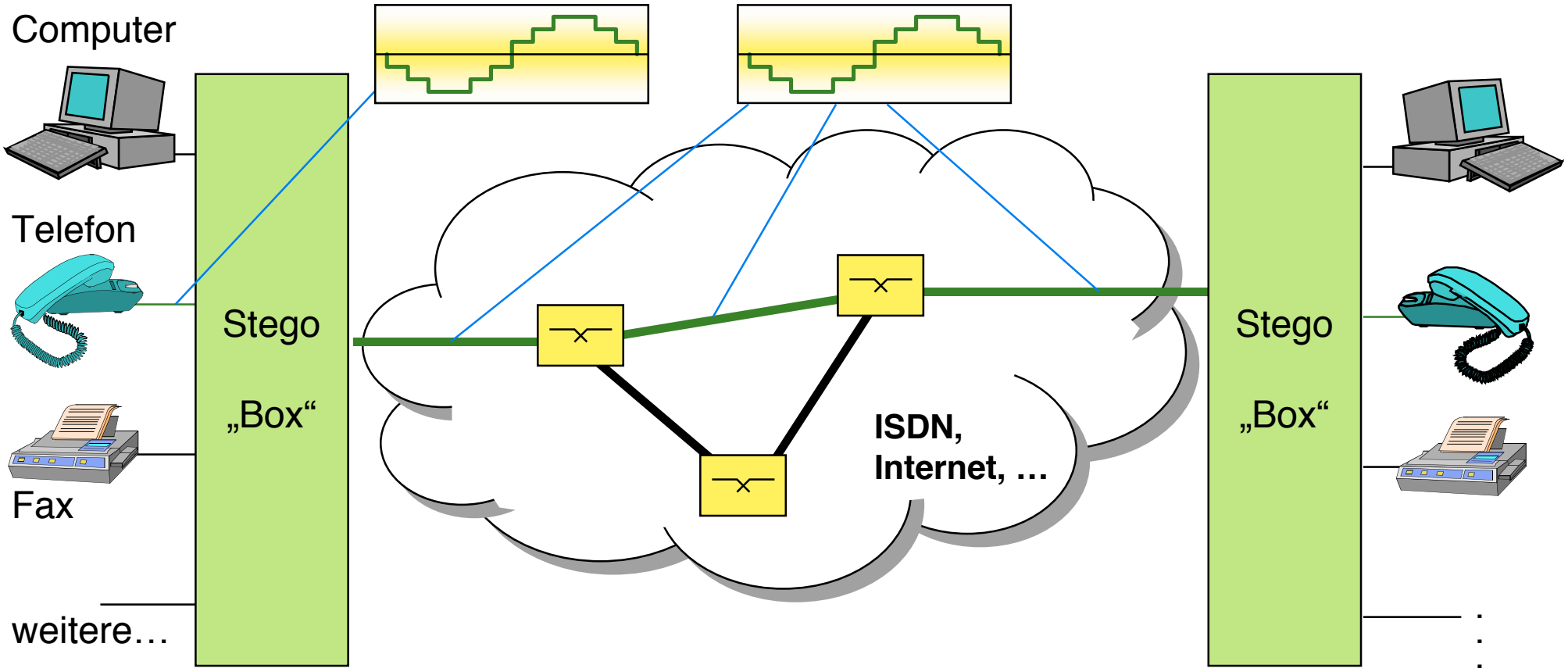
- sehr leistungsfähig durch technisch einfache Umsetzbarkeit
  - teilweise lediglich gezieltes Überschreiben niederwertiger Bits
- verschiedene Systeme für unterschiedliche Medien
- originale Hülldaten müssen unwiederbringlich vernichtet werden

Kryptographie



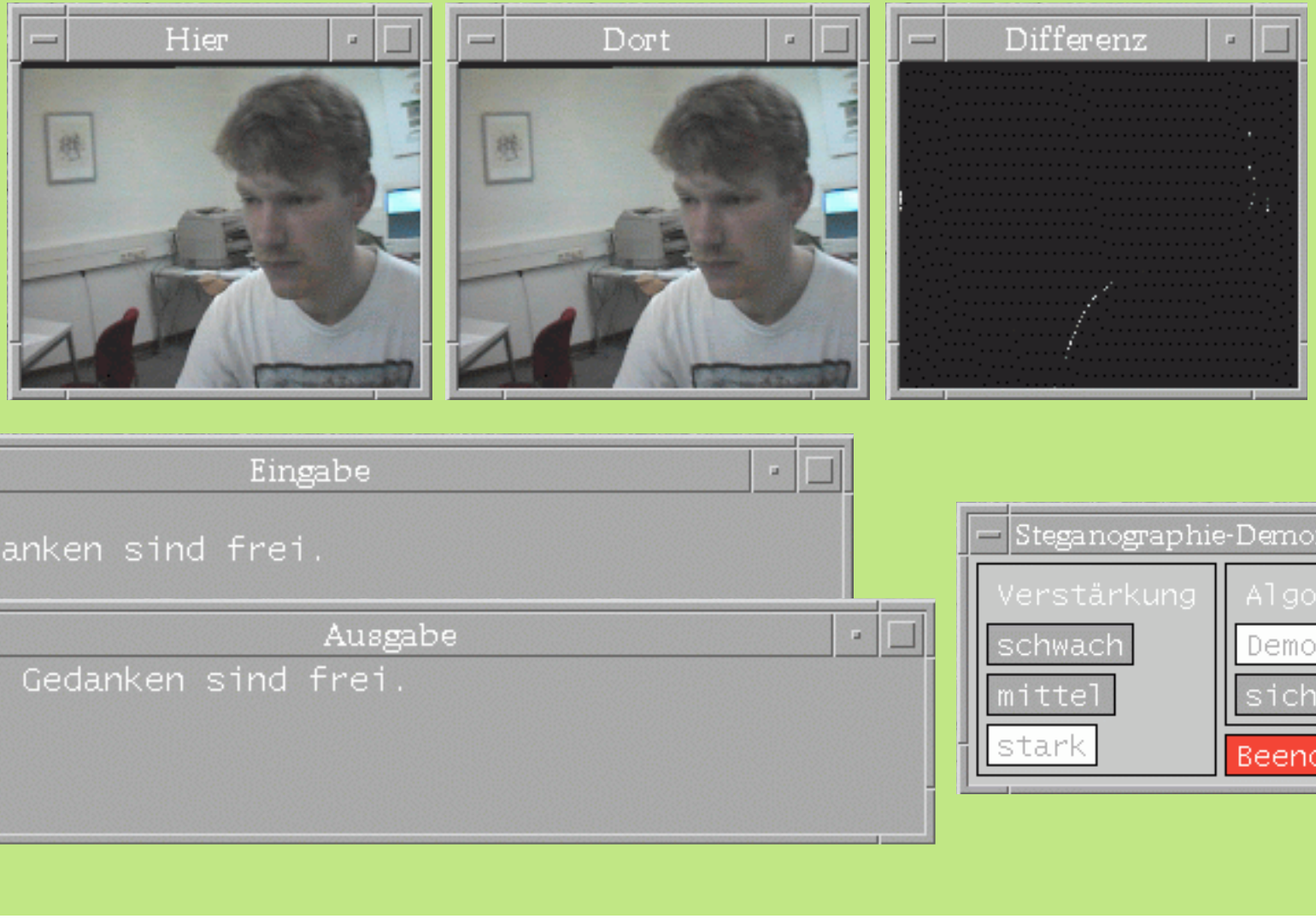
Verwendung von Kryptographie ist erkennbar

Multimediakommunikation —> Steganographie



Verwendung von Steganographie ist nicht erkennbar

Steganographie in Videokonferenzen



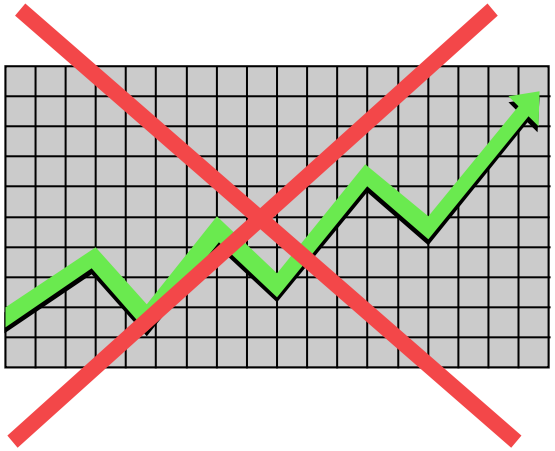
Güte steganographischer Systeme?

- **Heute verfügbare Systeme:**
  - meist schlecht (frei verfügbare)
  - Aufdecken von Schwächen führt zu deren Beseitigung
- **Was zeichnet gute steganographische Systeme aus?**
  - Algorithmus ist vollständig offengelegt
  - Parametrisierung durch steganographischen Schlüssel
  - Finden und Ausnutzen von „natürlichen Schmutzeffekten“
  - Brechen steganographischer Systeme ist zweistufig:
    - Erkennen, DASS etwas verändert/eingebettet wurde
    - Ermitteln, WAS eingebettet wurde
  - Beweis der Sicherheit eines Systems existiert bisher nicht

Randbedingungen für gute Systeme

- Original der Hülle unwiederbringlich vernichten!
- nie eine Hülle zweimal verwenden
- Vermeiden von unnatürlichen Prozessen, z.B. Einbetten in künstliche Computergrafiken

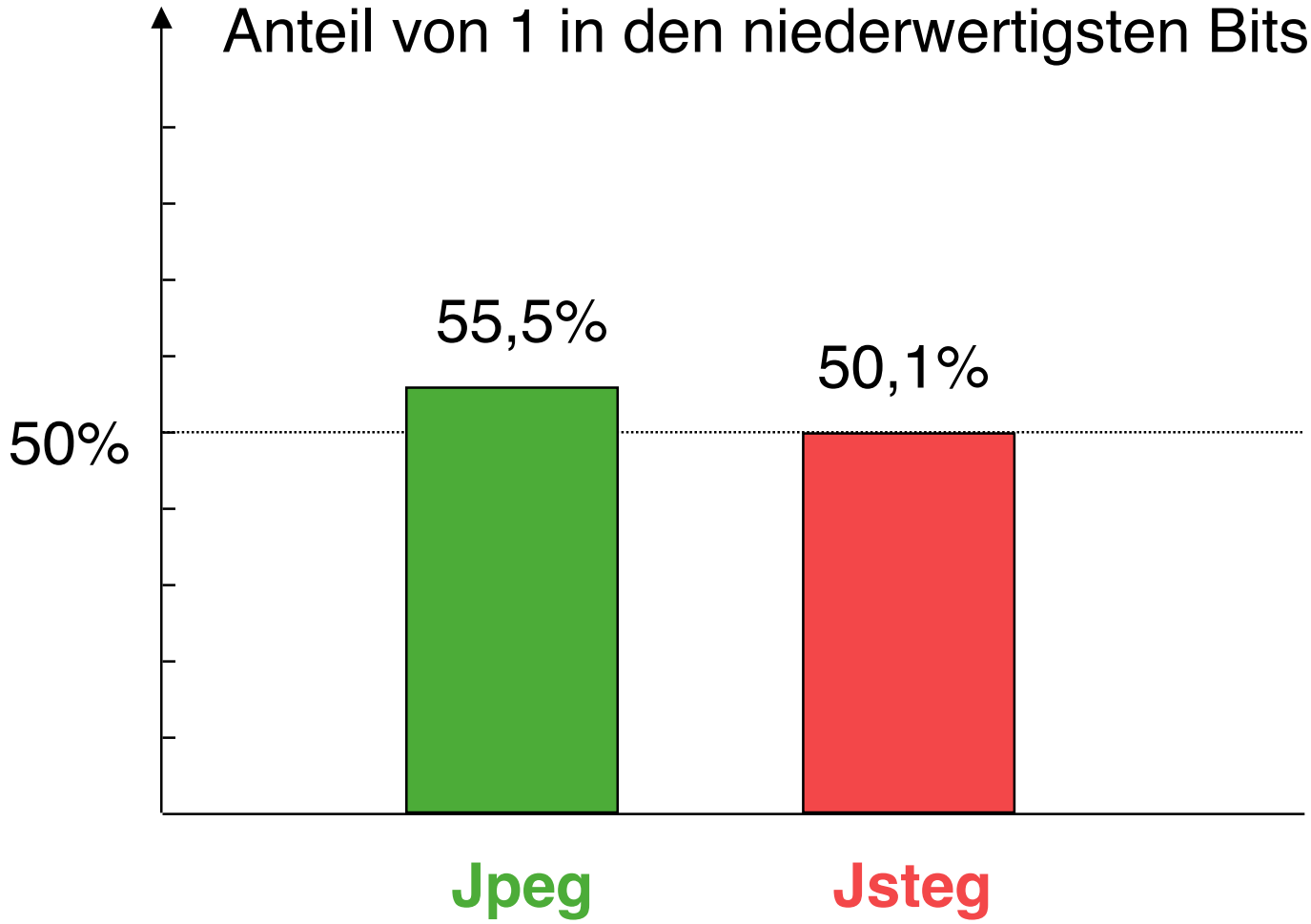
!



Beispielalgorithmus Jsteg

- **Algorithmus basiert auf Jpeg-Kompression**
- **Angriffe:**
  - visuelle Analyse:
    - liefert keine Anhaltspunkte
  - Verteilung der niederwertigsten Bits:
    - liefert in ungünstigen Fällen Verdacht
  - „Treppenangriff“:
    - deckt Verwendung von Steganographie auf
    - führt jedoch nicht unmittelbar zum Aufdecken der geheimen Nachricht

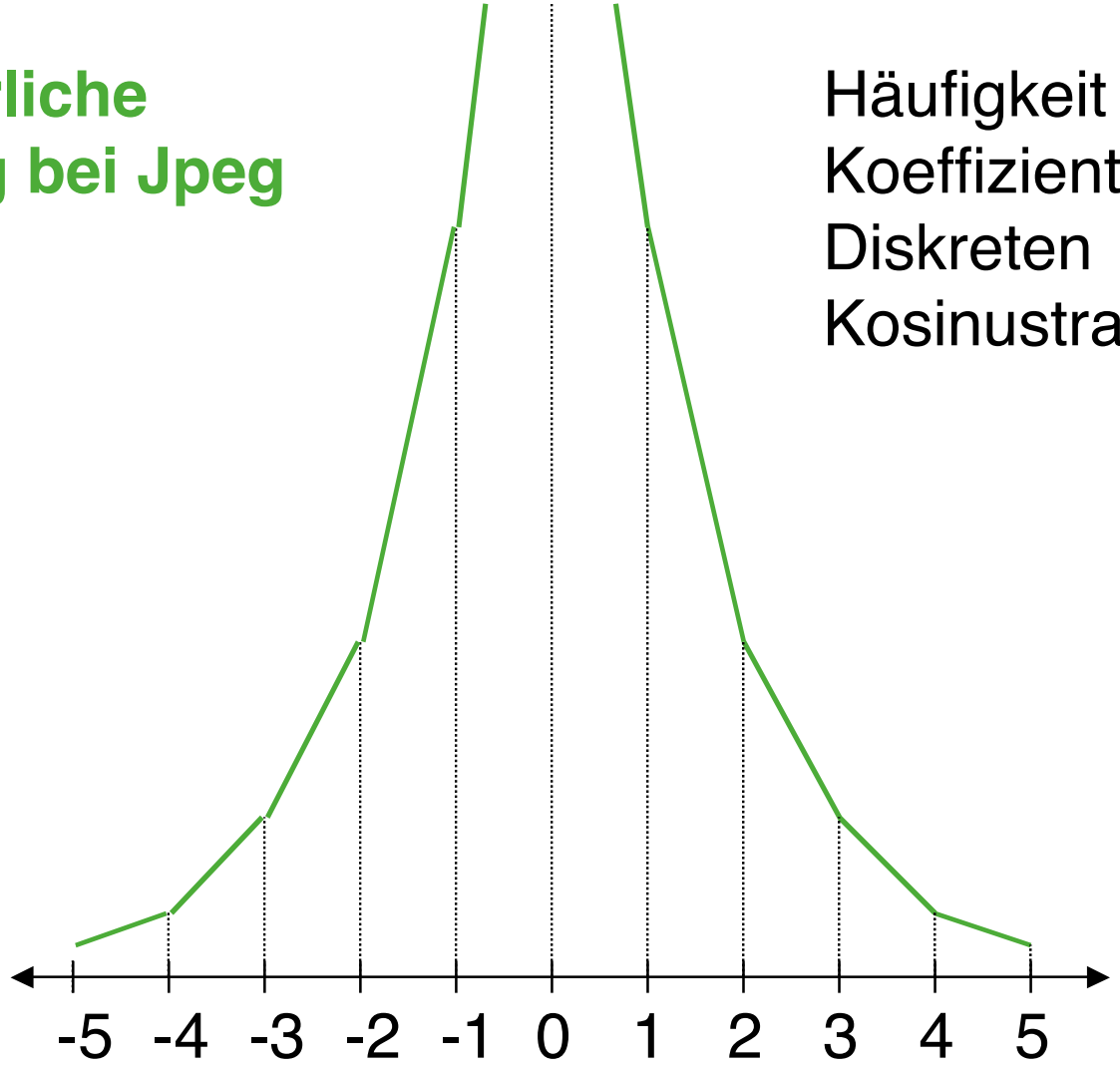
Verteilung der niederwertigsten Bits



Treppenangriff

Natürliche Verteilung bei Jpeg

Häufigkeit der Koeffizienten der Diskreten Kosinustransformation

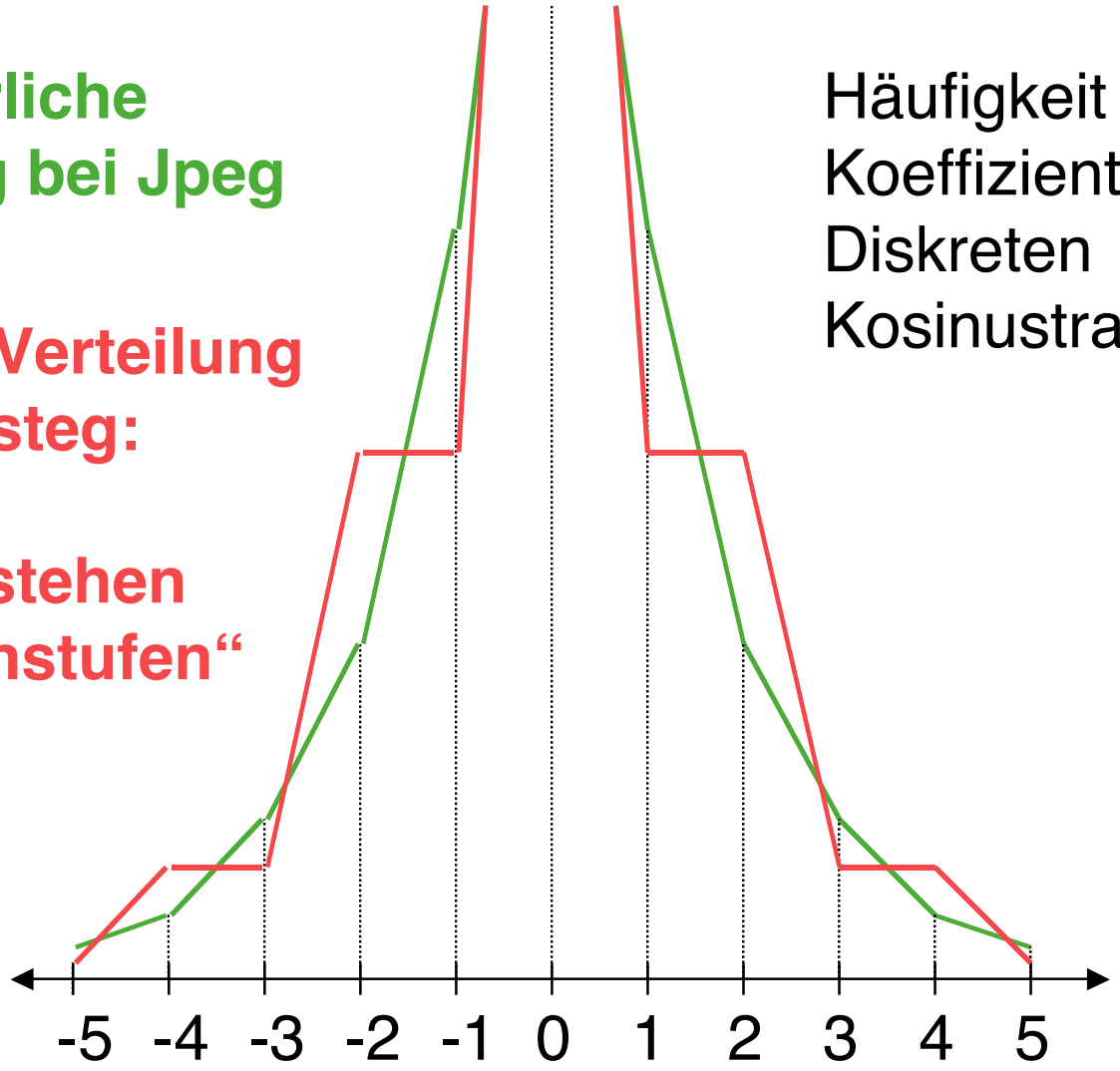


Treppenangriff

Natürliche Verteilung bei Jpeg

Entartete Verteilung bei Jsteg:

es entstehen „Treppenstufen“



Häufigkeit der Koeffizienten der Diskreten Kosinustransformation

Leistungsfähigkeit

- **Qualitativ**
  - bezogen auf Vertraulichkeit von Daten
    - Schutz von Nachrichteninhalten
    - Verbergen einer überlagerten Kommunikation
  - bezogen auf Integrität von Daten
    - Markieren von Daten zum Zwecke des Urheberschutzes
- **Quantitativ**
  - In Videoströmen hat ein komprimiertes Telefongespräch Platz (ca. 10 kbit/s)
  - In eingescannten Bildern ca. 1% des Datenmaterials
  - In ISDN-Telefongesprächen einige hundert Bit/s

Fazit

geheimer Nachrichtentyp	minimaler Trägertyp
E-Mail-Text	Digitale Sprache, ISDN, Bilder
Sprache	Digitales Video
Bilder, Video	?

- Keines der heute frei verfügbaren Systeme verwenden!
- Kryptoreglementierung wird auf Dauer zu besseren Stegosystemen führen
- Multimediakommunikation bietet hervorragende Basis für Steganographie

Steganographie

Ziel: vertrauliche Kommunikation

Schlüssel

Schlüssel

Hülldaten

einzubettende Daten

Einbetten

Stegodaten

Extrahieren

Hülldaten*

eingebettete Daten

Text

Sender

Angreifer

Empfänger

Text

keine Änderungen

- exakt gleich
- nicht feststellbar
- möglichst viel

Steganographie

Ziel: Urheberschaft digitaler Werke sichern

Schlüssel

Schlüssel

Hülldaten

einzubettende Daten

Einbetten

Stegodaten

Extrahieren

Hülldaten*

eingebettete Daten

Text

Sender

Angreifer

Empfänger

Text

u.U. starke Änderungen

- Korrelation genügt
- einige 100 Bit genügen

Watermarkingsysteme

Schlüssel

Original

Einbetten

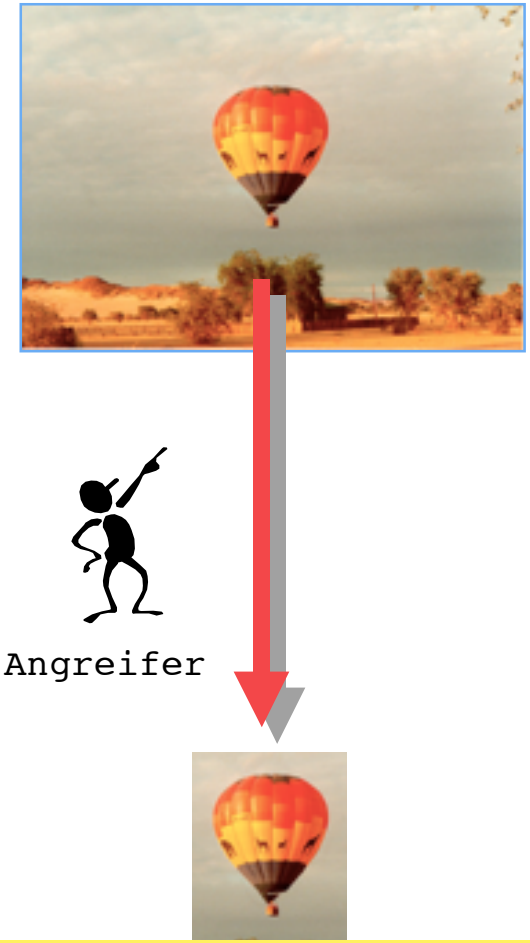
Sender

Watermark-Signal:

Distribution

Angreifer

Angriffe auf Watermarkingsysteme



- Digital-Analog-Wandlung
- Analog-Digital-Wandlung
- Re-Sampling
- Re-Quantisierung
- Kompression
- Dithering
- Rotation
- Translation
- Cropping
- Scaling

- Collution Attacks

Designkriterien für Watermarking

- Herkömmliche steganographische Systeme sind meist nicht in der Lage, solche Robustheitsanforderungen zu erfüllen.

- **Anforderungen allgemein:**

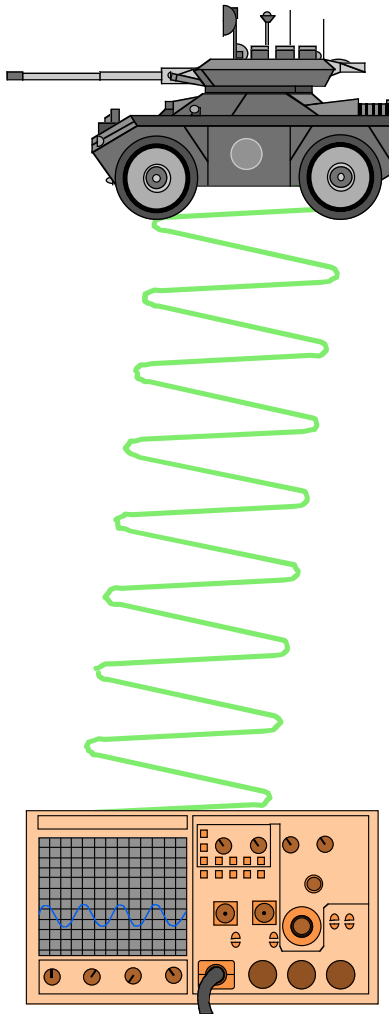
Watermarking zum Schutz von Rechten	**Steganographie** zur vertraulichen Kommunikation
Robustheit	**fehlerfreie Übertragung**
Beeinträchtigungslosigkeit	**Unauffälligkeit im Träger**
Nachweisbarkeit	**Nichtnachweisbarkeit**

→ Offenlegung des Schlüssels

Spread Spectrum Systems

Exkurs:

- Funktechnik
- insbesondere militärischer Bereich



- Funkkontakt zwischen verschiedenen militärischen Einheiten
- Sendung auf einer bestimmten Frequenz f_0 mit einer bestimmten Bandbreite *B*

Problem:

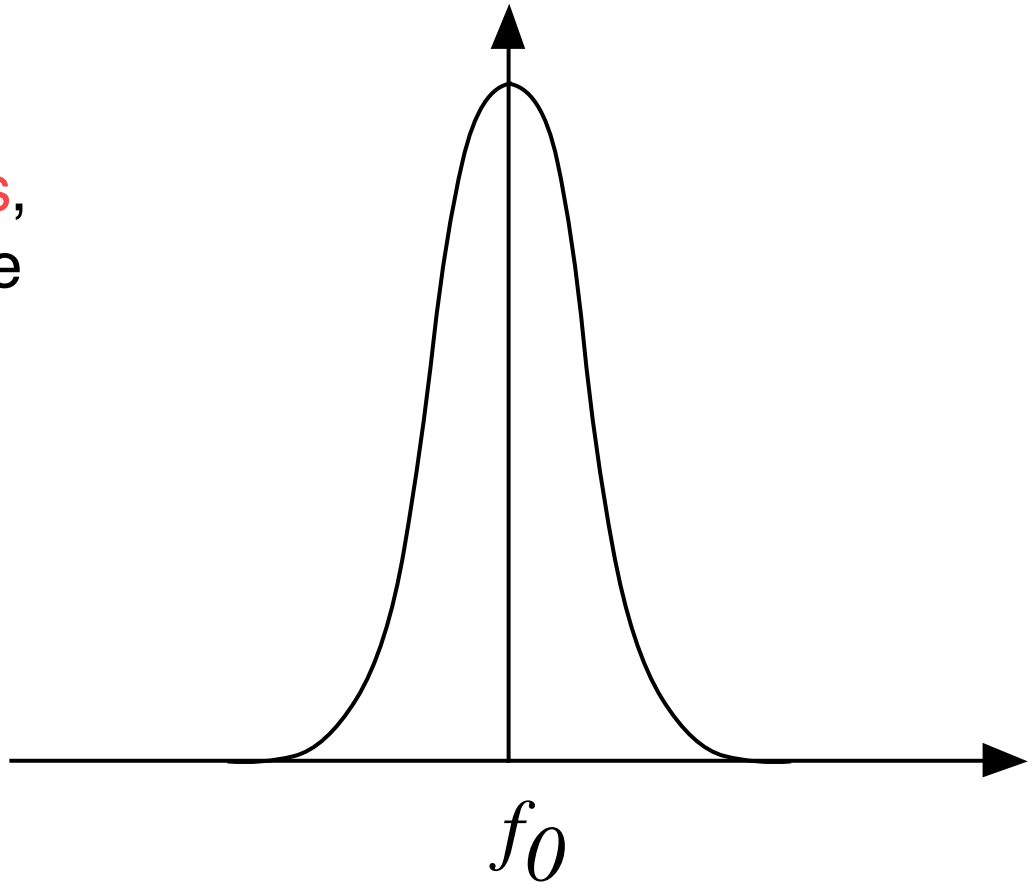
- deutliche Energiezunahme im Spektrum um f_0 herum

Schmalbandiges Senden

Folgen:

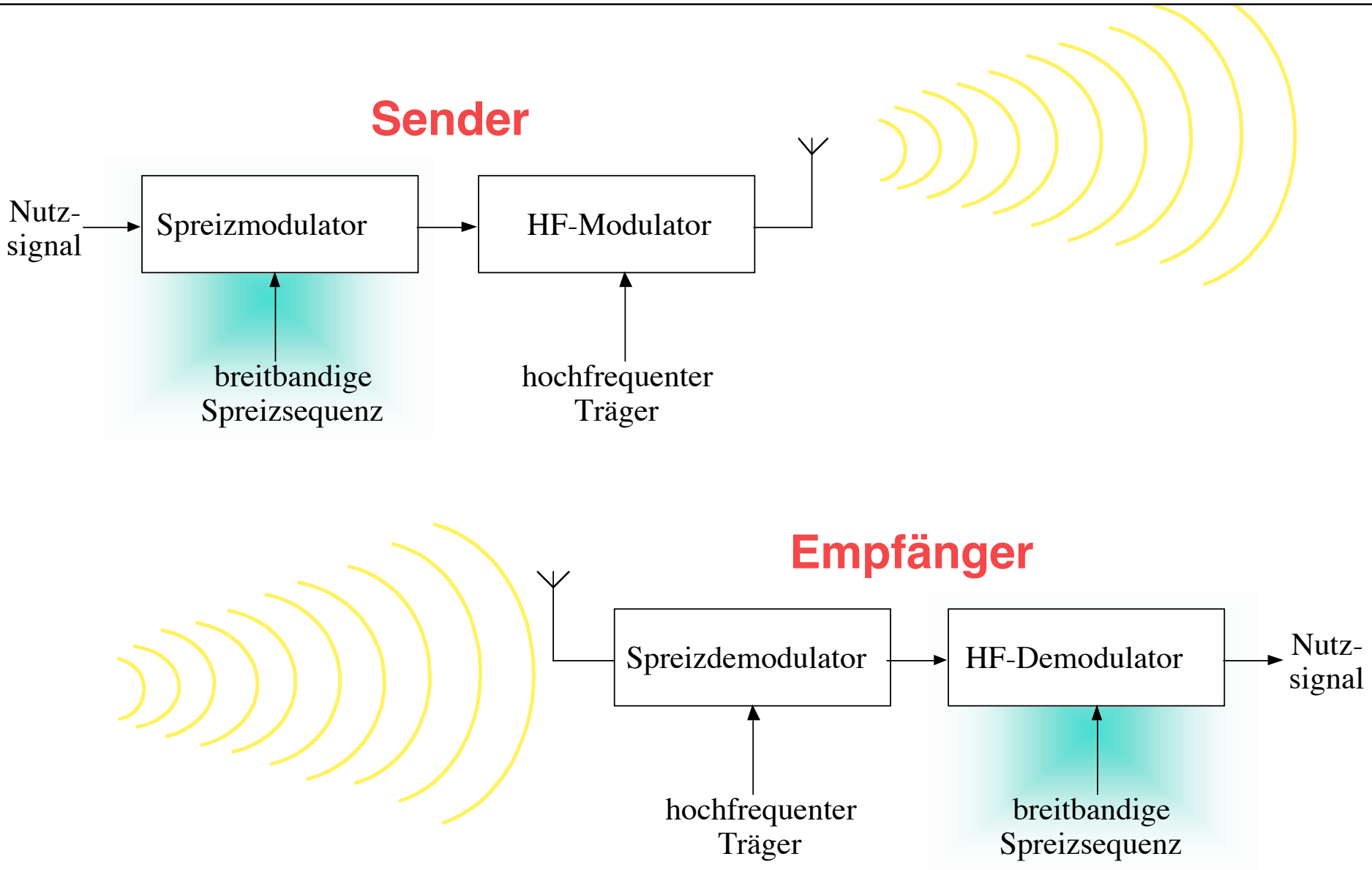
- Beobachtbarkeit des Sendens,
  da ein Spektrumanalysator die
  Energiezunahme registriert

- Peilbarkeit des Senders,
  da die elektromagnetischen
  Wellen Richtungsinformation
  in sich tragen

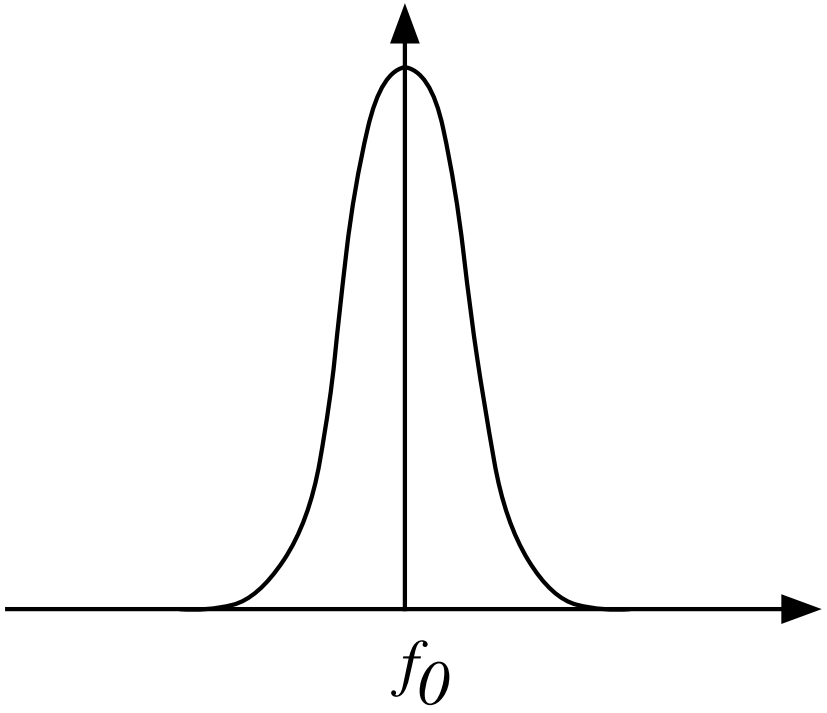


- Gegner kann Kommunikation mit Störsender verhindern

Übertragungsmodell beim Bandspreizverfahren

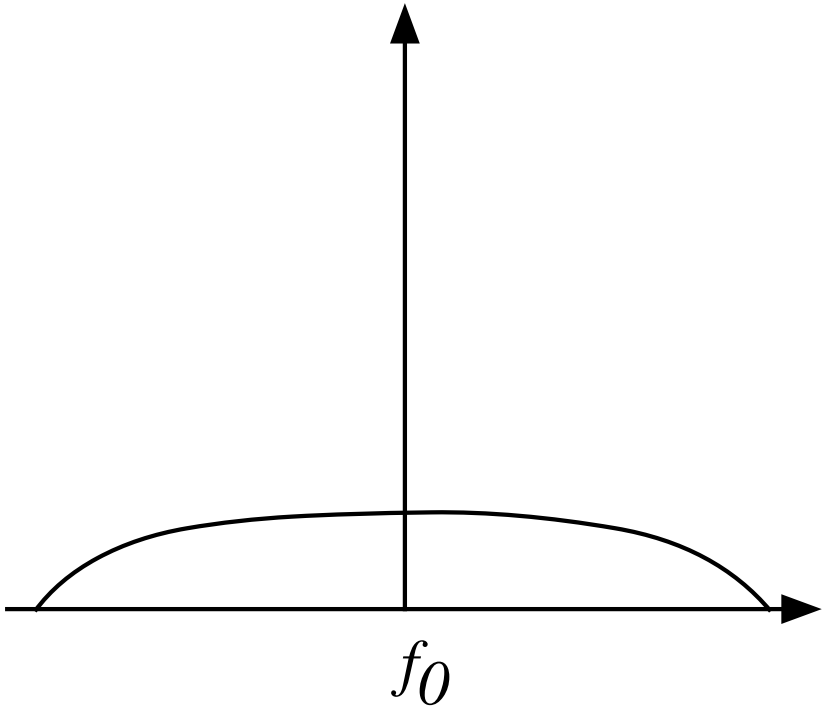


Spreizung



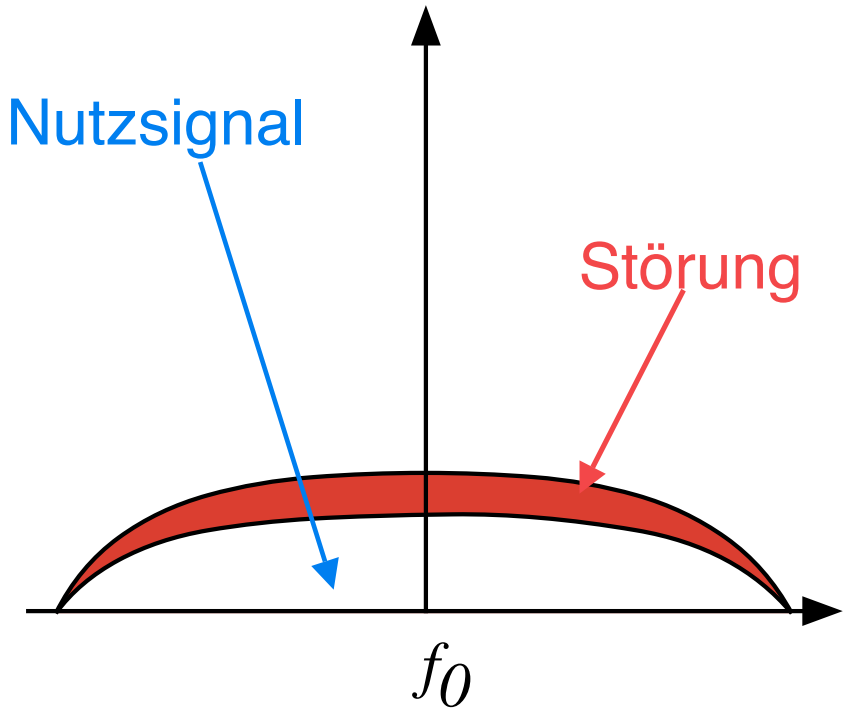
- Schmalbandiges Nutzsignal vor der Spreizung
- Modulation mit breitbandiger Spreizsequenz:
  - spezielle Funktionen (z.B. Walsh-Funktionen)
  - Pseudo-Noise-Sequence (PC-Code)

Spreizung



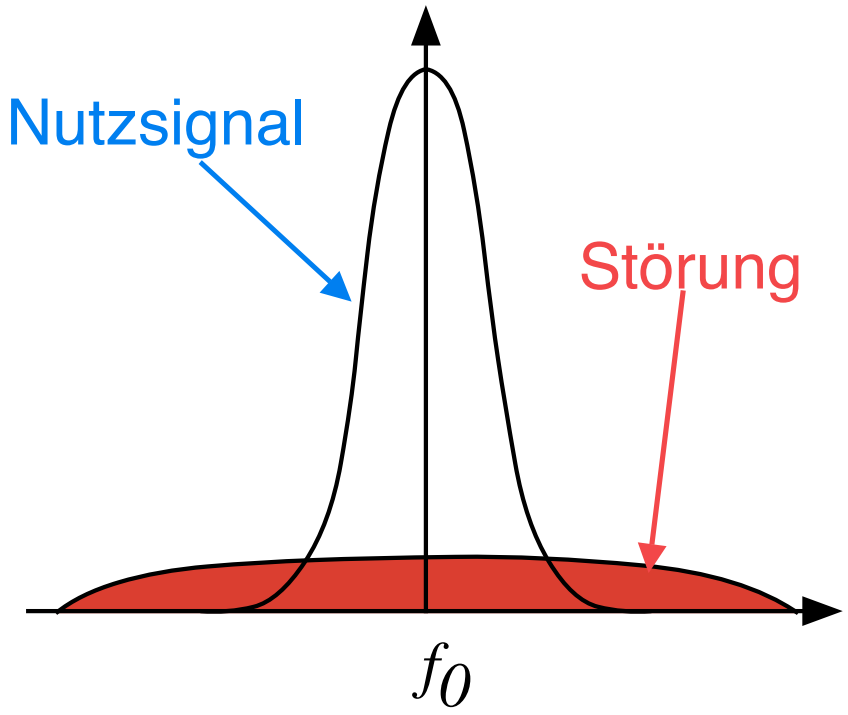
- Schmalbandiges Nutzsignal vor der Spreizung
- Modulation mit breitbandiger Spreizsequenz:
  - spezielle Funktionen (z.B. Walsh-Funktionen)
  - Pseudo-Noise-Sequence (PC-Code)
- Spektrale Spreizung
- Verteilung der Energie auf ein großes Frequenzspektrum

Despreizung



- gespreiztes Nutzsignal mit überlagerter Störung

Despreizung



- gespreiztes Nutzsignal mit überlagerter Störung

- Spektrale Spreizung der Störung
- despreiztes Nutzsignal

Prinzip des Spread Spectrum Watermarking

Document-ID: #A53-229D789
Author: J.Fitzgerald

Watermark-Signal

Frequenz-transformation

Einbetten

inverse Frequ.-transformation

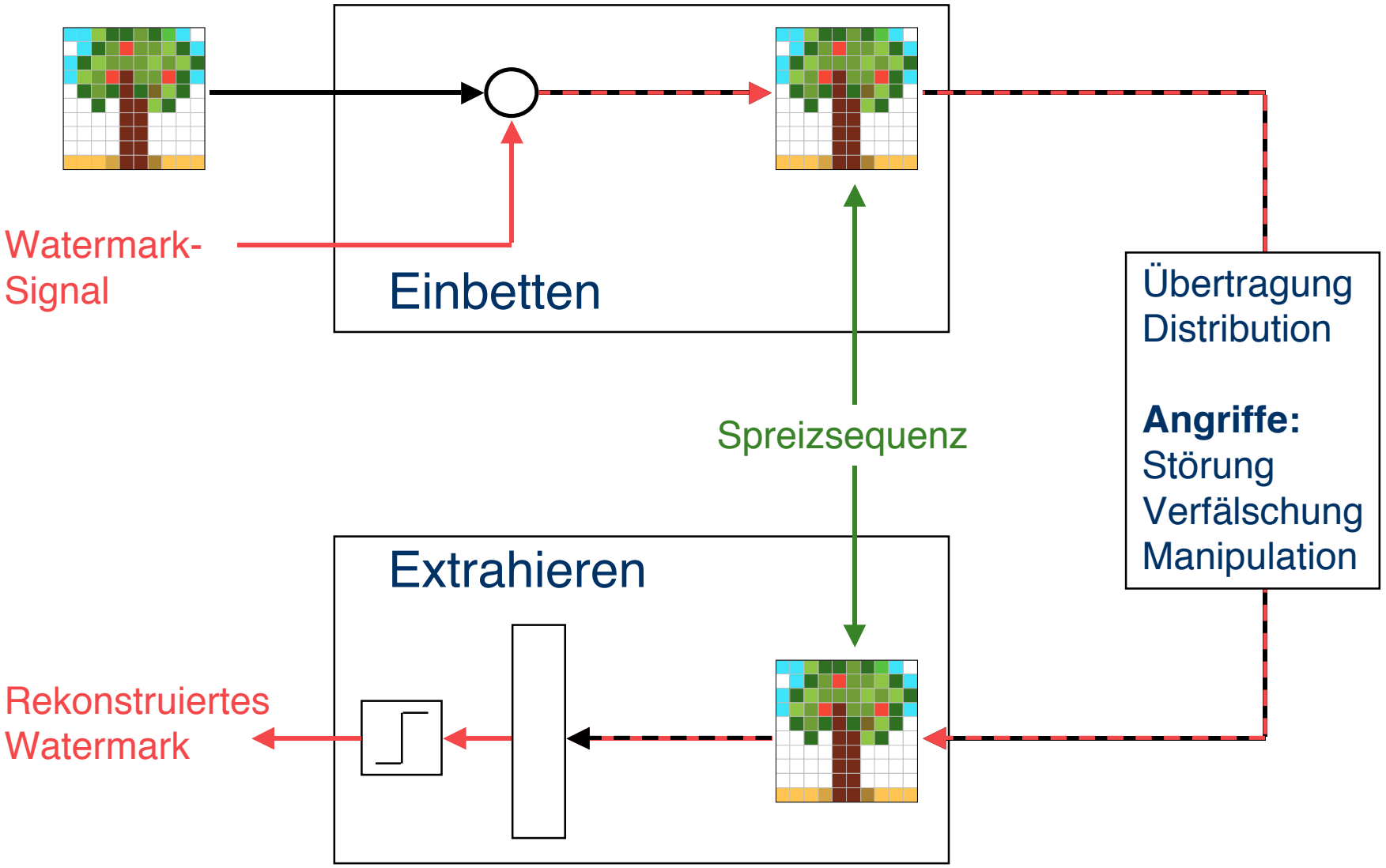
z.B. Diskrete Kosinustrans-formation (DCT)

gewichtete DCT-Basisbilder

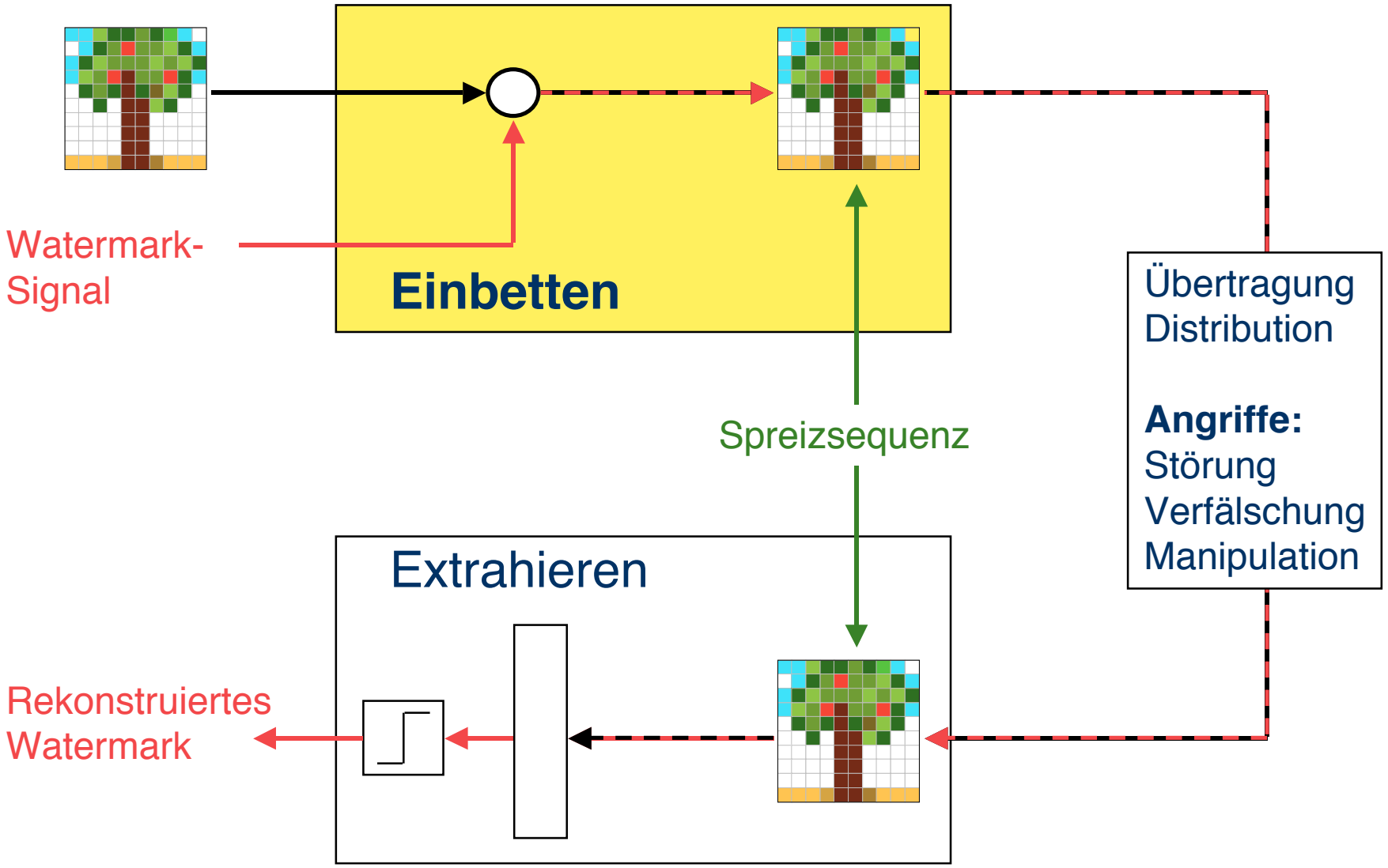
Spreizsequenz

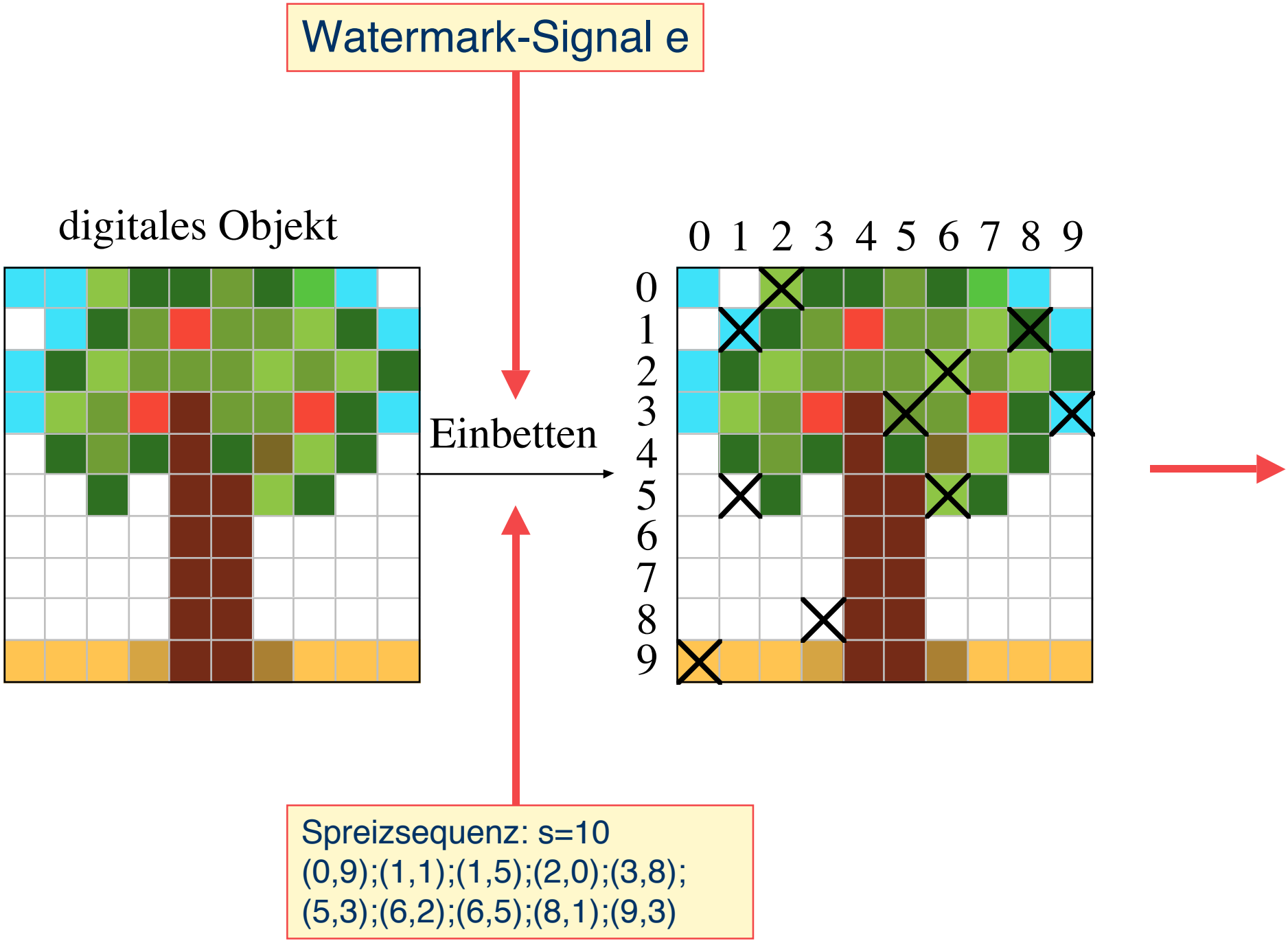
Zufallszahlen:
01101000101110101011010101010010100010000101010100010101010100010

Ein vereinfachtes Beispiel

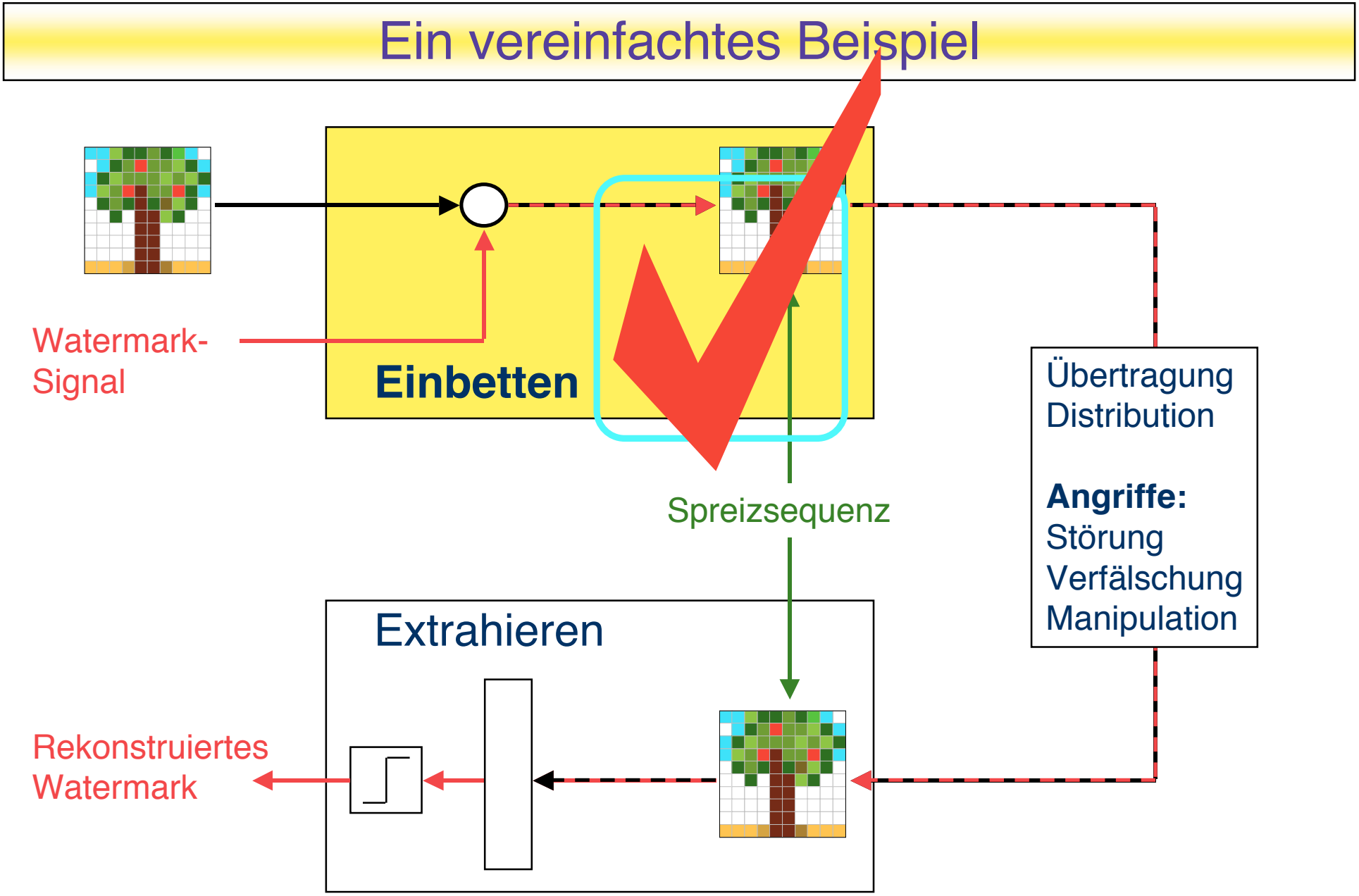


Ein vereinfachtes Beispiel

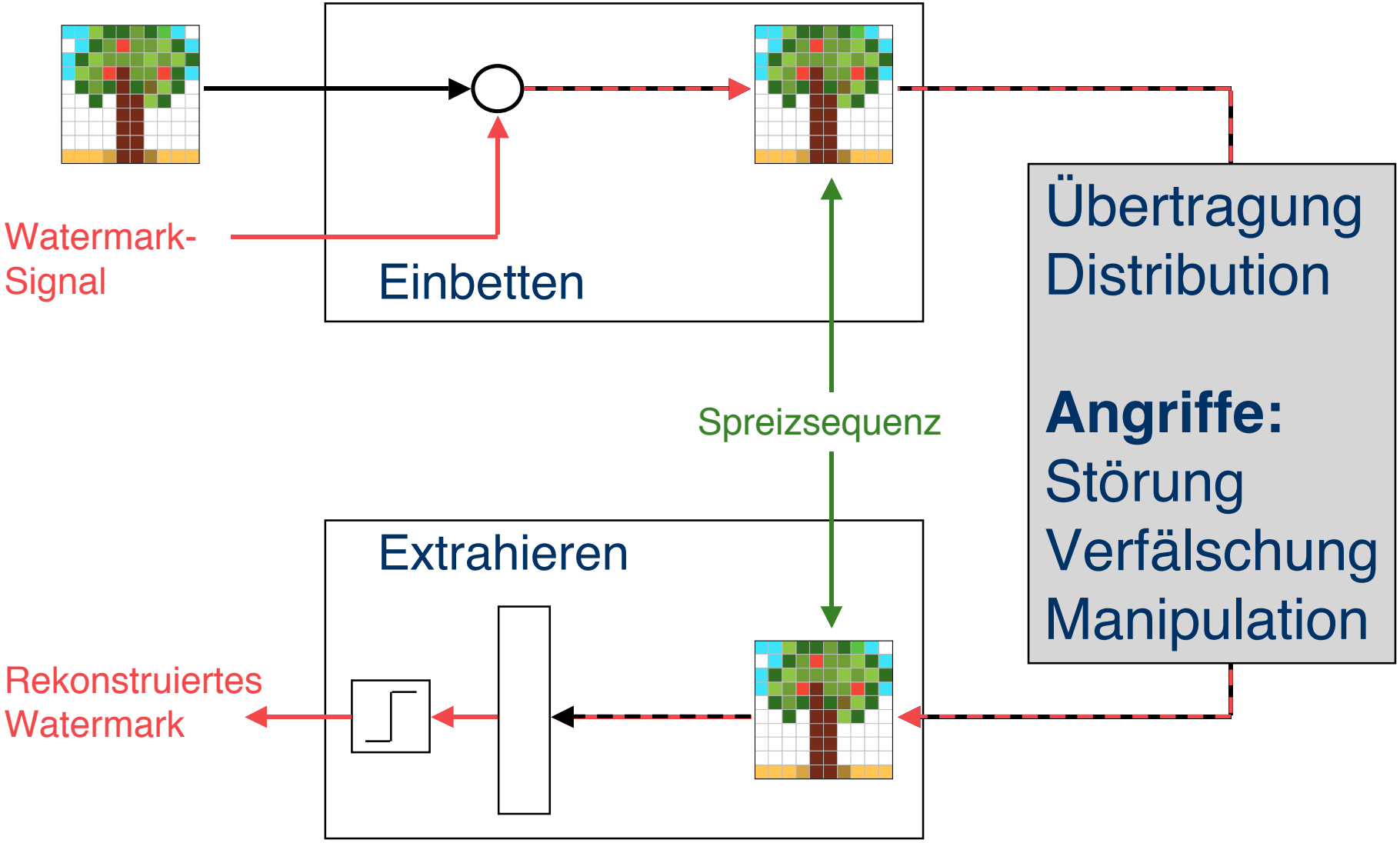



Watermark-Signal e
digitales Objekt
Einbetten
0 1 2 3 4 5 6 7 8 9
0
1
2
3
4
5
6
7
8
9
Spreizsequenz: s=10
(0,9);(1,1);(1,5);(2,0);(3,8);
(5,3);(6,2);(6,5);(8,1);(9,3)

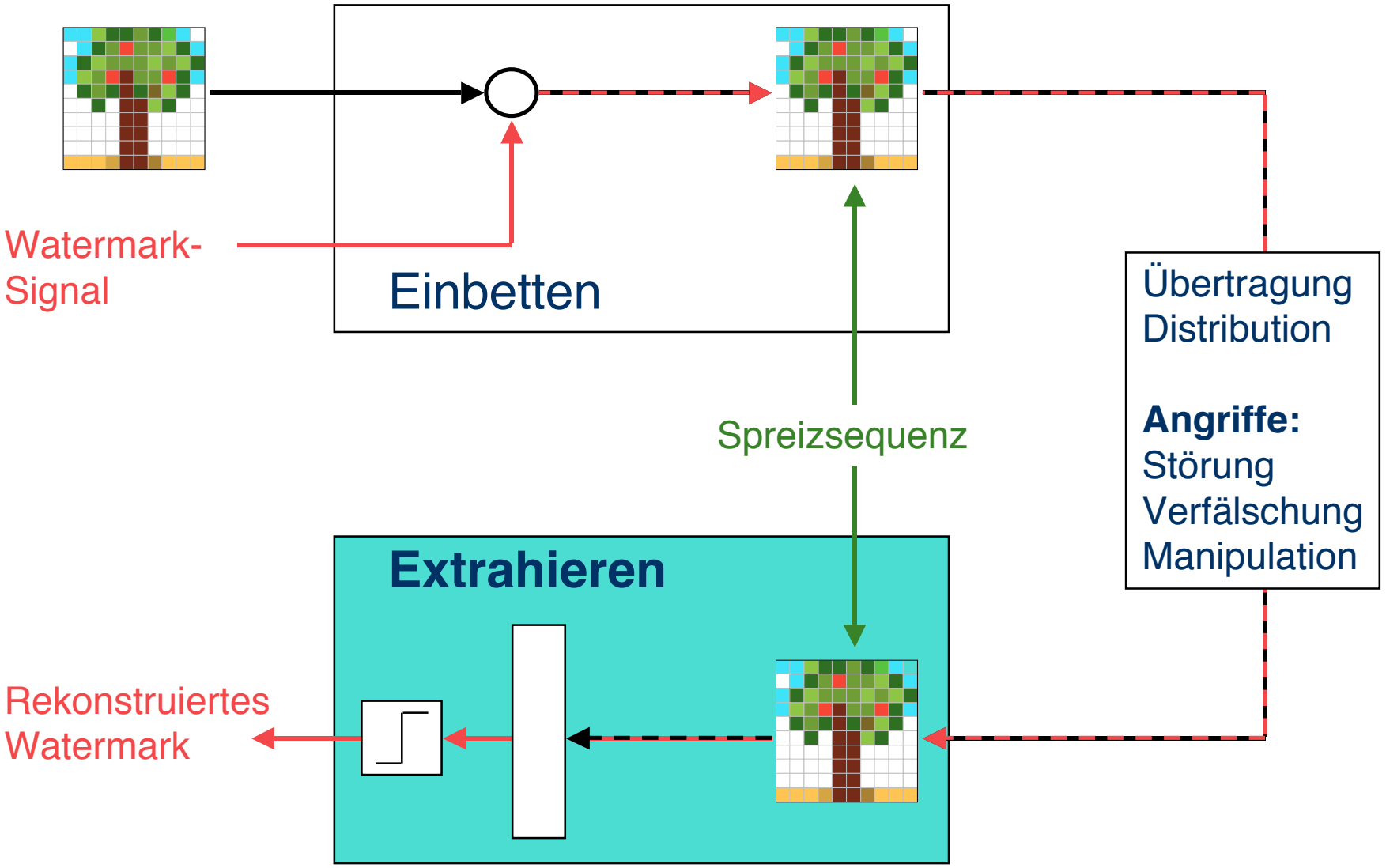
Ein vereinfachtes Beispiel

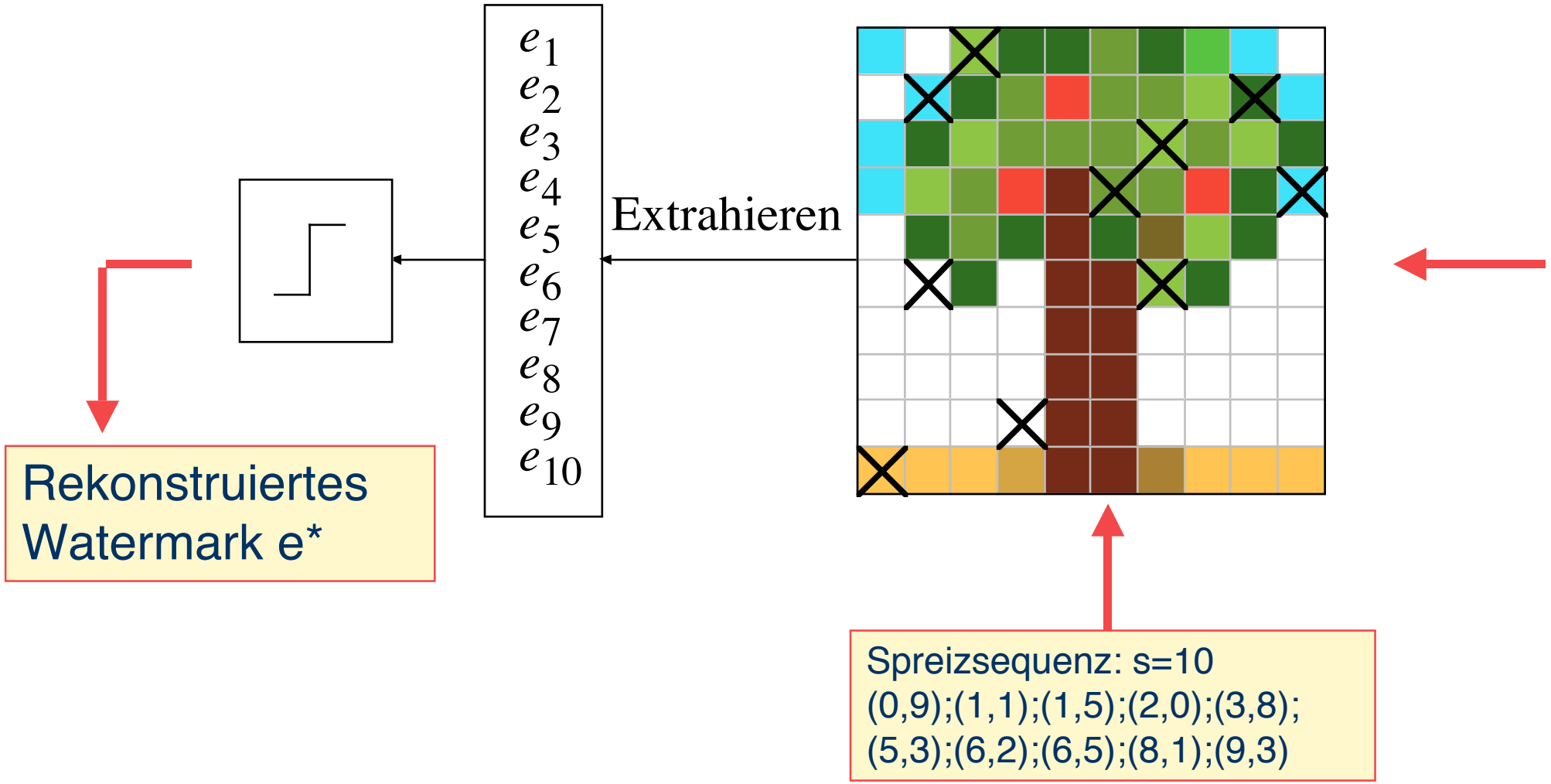


Ein vereinfachtes Beispiel

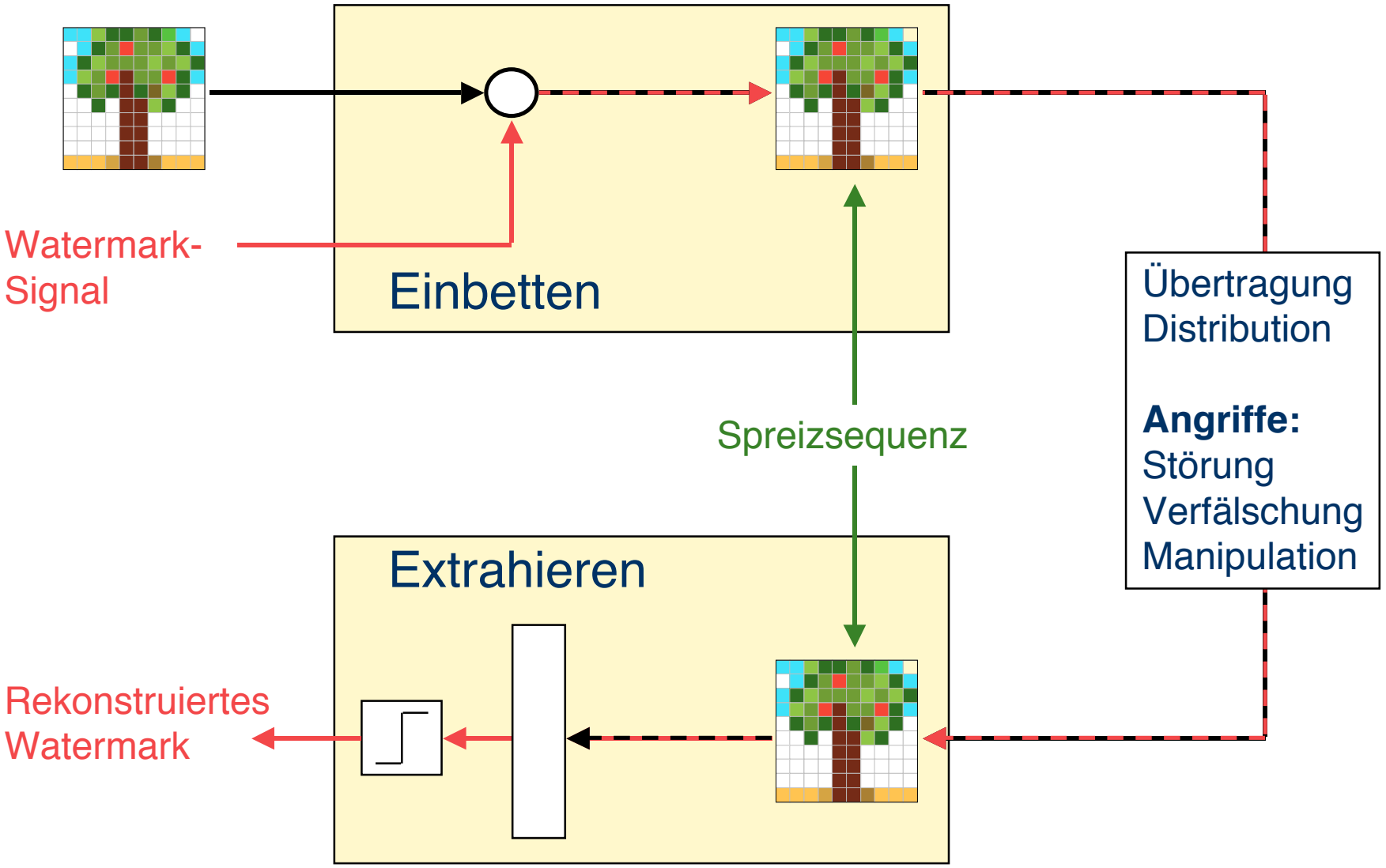


Ein vereinfachtes Beispiel




e_1
e_2
e_3
e_4
e_5
e_6
e_7
e_8
e_9
e_{10}
Extrahieren
Rekonstruiertes Watermark e*
Spreizsequenz: s=10
(0,9);(1,1);(1,5);(2,0);(3,8);
(5,3);(6,2);(6,5);(8,1);(9,3)

Ein vereinfachtes Beispiel



Etwas formaler ...

- Markiertes Objekt D(x,y) entsteht durch
- pixelweise Addition des
- originalen Objektes N(x,y) mit der
- Sequenz S(x,y)

$$D(x,y) = N(x,y) + S(x,y)$$

- Jedes Informationsbit b_i des Watermarks wird in S(x,y) repräsentiert durch eine sog. Basisfunktion φ_i
- S(x,y) ergibt sich nach:

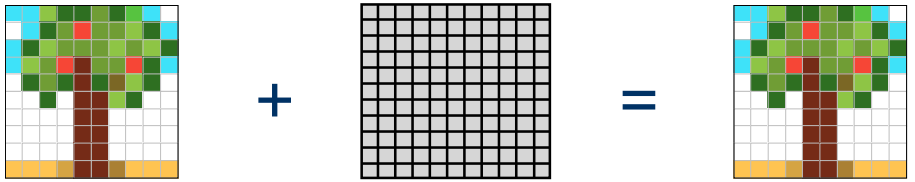
$$S(x,y) = \sum_i b_i \varphi_i(x,y)$$

Etwas formaler ...

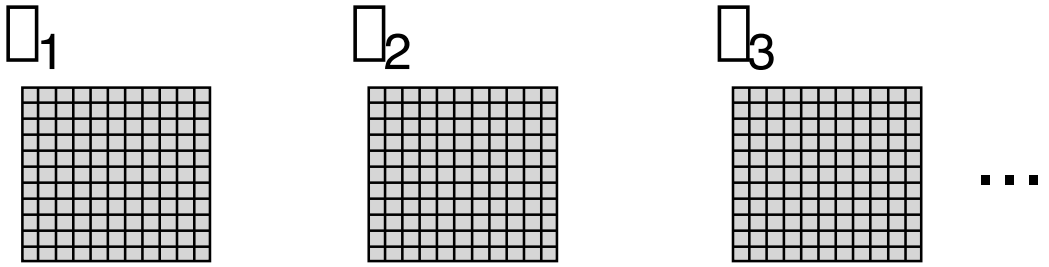
Einbetten:

$$D(x,y) = N(x,y) + S(x,y)$$

$$S(x,y) = \sum_i b_i \, \Psi_i(x,y)$$



- Watermark = $(b_1, b_2, b_3, \ldots, b_i, \ldots)$ (Bitvektor)
- Die Basisfunktionen Ψ_i sollten orthogonal zueinander sein.
- Im einfachsten Fall sind das unabhängig voneinander gebildete Zufallszahlen.
- Die Basisfunktionen Ψ_i:

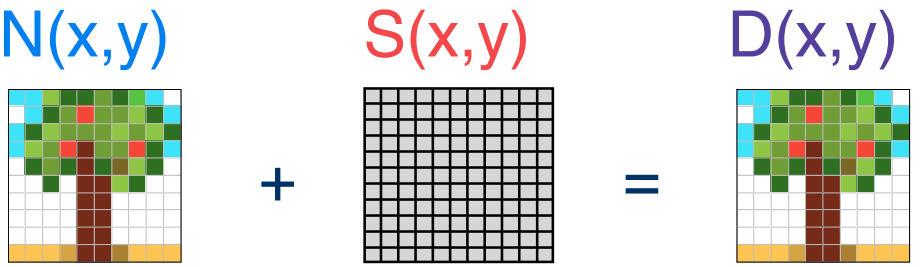


Etwas formaler ...

Einbetten:

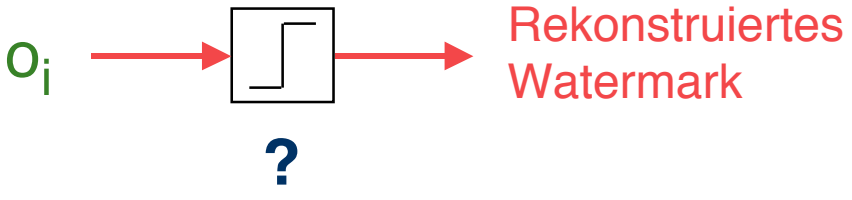
$$D(x,y) = N(x,y) + S(x,y)$$

$$S(x,y) = \sum_i b_i \, \psi_i(x,y)$$



Extrahieren:

$$o_i = \sum_{x,y} D(x,y) \, \psi_i(x,y)$$



Was bleibt?

- Jemand könnte sich eine Kombination aus selbst gewähltem Watermark b und Basisfunktionen „basteln“, so daß er ein Objekt als seines ausgeben könnte, obwohl er es nie markiert hat.

- Notwendigkeit der **Registrierung** des Marks und der Basisfunktionen bzw. der Sequenz S.

§ **Geschäftsmodell** notwendig, wenn man etwas beweisen können will.