Error Exponents of Typical Random Trellis Codes

Neri Merhav

The Andrew & Erna Viterbi Faculty of Electrical Engineering

Technion—Israel Institute of Technology

Haifa 3200004, Israel

ITW 2019, Visby, Gotland, Sweden, August 25–28, 2019

Typical Random Codes

Traditional random coding error exponents are defined as

$$E_{\mathsf{r}}(R) = \lim_{n\to\infty}\left[-\frac{\ln \mathbf{E} P_{\mathsf{e}}(\mathcal{C}_n)}{n}\right].$$

We define typical–code error exponents as

$$E_{\mathsf{typ}}(R) = \lim_{n\to\infty}\left[-\frac{\mathbf{E} \ln P_{\mathsf{e}}(\mathcal{C}_n)}{n}\right].$$

- By Jensen's inequality, $E_{\mathsf{typ}}(R) \geq E_{\mathsf{r}}(R)$.
- $E_{\mathsf{r}}(R)$ – dominated by bad codes; $E_{\mathsf{typ}}(R)$ – dominated by typical codes.

Let $\mathcal{G}_E = \{\mathcal{C}_n : \quad P_{\mathsf{e}}(\mathcal{C}_n) \doteq e^{-nE}\}$.

$$\overline{P_{\mathsf{e}}(\mathcal{C}_n)} \doteq \sum_E P(\mathcal{G}_E)\cdot e^{-nE} \doteq P(\mathcal{G}_E^*)\cdot e^{-nE^*}.$$

Otoh, $E_{\mathsf{typ}}(R) = \sum_E P(\mathcal{G}_E)\cdot E = E_0$, where $P[\mathcal{G}_{E_0}] \to 1$.

Motivation

- $E_{\text{typ}}(R)$ is never worse than $E_{\text{r}}(R)$.
- Code selected once and for all: no LLN to support $\mathbf{E}P_{\text{e}}(\mathcal{C}_n)$.
- Once selected, w.h.p. $P_{\text{e}}(\mathcal{C}_n) \sim e^{-nE_0}$, forever.
- Theoretical framework for random–like codes (Battail, 1995).
- Analogy: physics of disordered sys. – quenched vs. annealed average.

Q: With all these motivations, why wasn't it explored much more before?

A: Not so easy to analyze (also in physics) ....

Related Work

- Barg & Forney ('02): i.i.d. random coding, BSC:

  At low rates: $E_{\text{typ}}(R) = E_{\text{ex}}(2R) + R$.

- Nazari ('11); Nazari, Anastasopoulos & Pradhan ('14):

  upper and lower bounds for the α–decoder.

- Stat. phys. literature: Kabashima ('08), Mora & Riviore ('06), ...:

  LDPC codes - replica analysis and cavity method.

- Battail ('95):

  random–like codes.

- Merhav ('18): fixed–composition rand. coding, DMC, gen. likelihood dec.

  Exact error exponent of the typical random code (TRC).

Contributions

Deriving the error exponent of the typical random trellis code for:

- A general DMC (not just MBIOS).
- A general rational coding rate (not only $1/n$).
- A general random selection of a time–varying code.

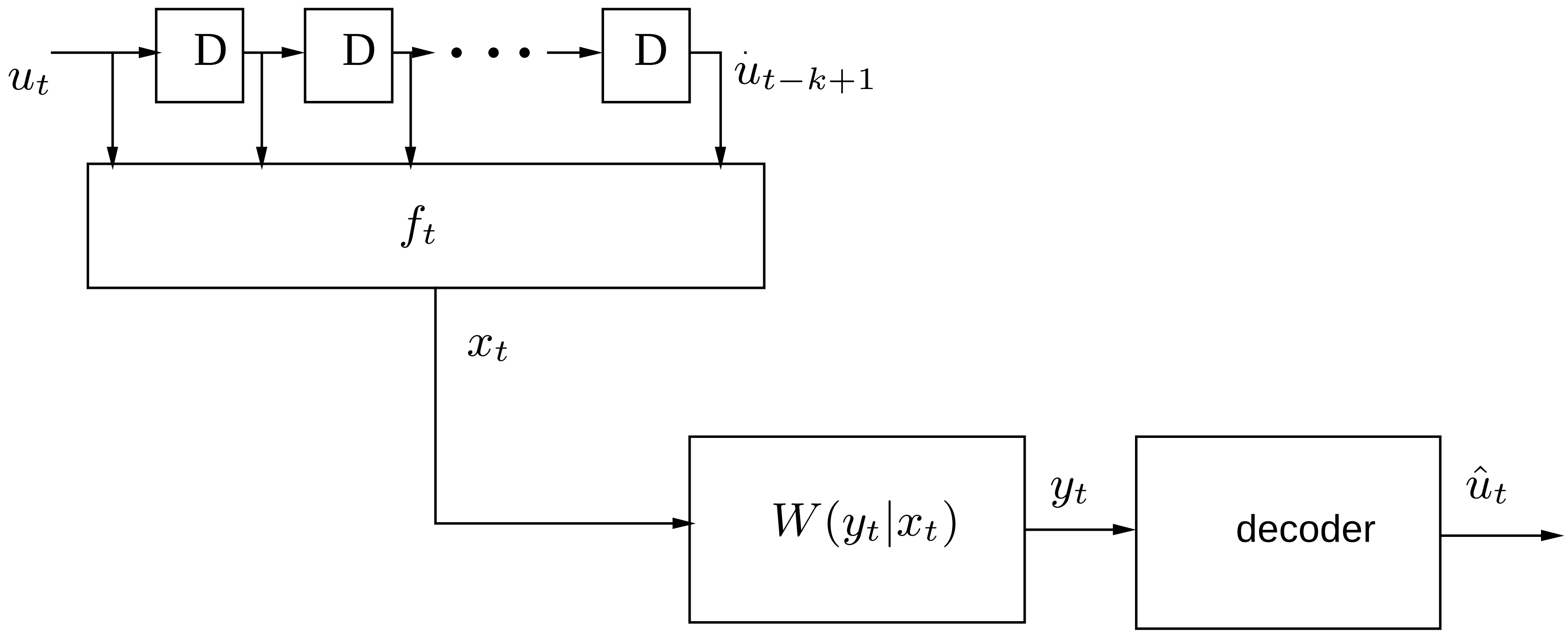
The analysis method (based on the MoT) provides insights on:

- Structure of the typical random trellis code.
- Dominant error events.

Additional extensions:

- Channels with finite input memory (ISI).
- Mismatched decoding metric.

Problem Setting



- $\{\boldsymbol{u}_t\}$ – m–vectors of purely random bits.
- $f_t : \{0,1\}^{mk} \to \mathcal{X}^n$ randomly selected according to Q^n.
- Coding rate: $R = m/n$; constraint length: $K = mk$.
- For convolutional codes, $\{f_t\}$ are linear.
- $W =$ DMC.
- Asymptotic regime: $k \to \infty$ while m and n are held fixed.

Background

Traditional performance metric: $E(R,Q) = \liminf_{K\to\infty}[-\log \mathbf{E}P_{\mathsf{e}}]/K$.

$$E(R,Q) = \begin{cases} R_0(Q)/R & R < R_0(Q) \\ E_0(\rho,Q)/R & R > R_0(Q) \end{cases}$$

where ρ satifies: $\rho R = E_0(R,Q)$.
For $R > R_0(Q)$: $\exists$ matching converse.
For $R < R_0(Q)$: improvement by an expurgated bound,

$$R_0(Q)/R \to E_{\mathsf{cex}}(R,Q) \stackrel{\triangle}{=} E_{\mathsf{x}}(\rho,Q)/R$$

with $\rho R = E_{\mathsf{x}}(\rho,Q)$.
In [Viterbi & Odenwalder, 1969]: for at least half of the convolutional codes

$$P_{\mathsf{e}} \leq \left(\frac{2L}{1-2^{-\epsilon/\rho R}}\right)^{\rho} \cdot \exp\{-KE_{\mathsf{cex}}(R,Q)\}.$$

If $2L$ is replaced by $100L$, the bound applies to 99% of the codes.

Objectives

Studying the typical ensemble performance,

$$\mathcal{E}_{\text{trtc}}(R,Q) = \liminf_{K\to\infty} \frac{-\mathbf{E}\log P_{\text{e}}}{K}$$

as well as $\mathcal{E}_{\text{trcc}}(R,Q)$ defined similarly for convolutional codes.

- Deriving both “Csiszár–style” and “Gallager–style” expressions.
- Comparing to the random coding exponent and expurgated exponent.
- Comparing to typical random block codes of the same complexity.

Main Result

For $R < R_0(Q)$:

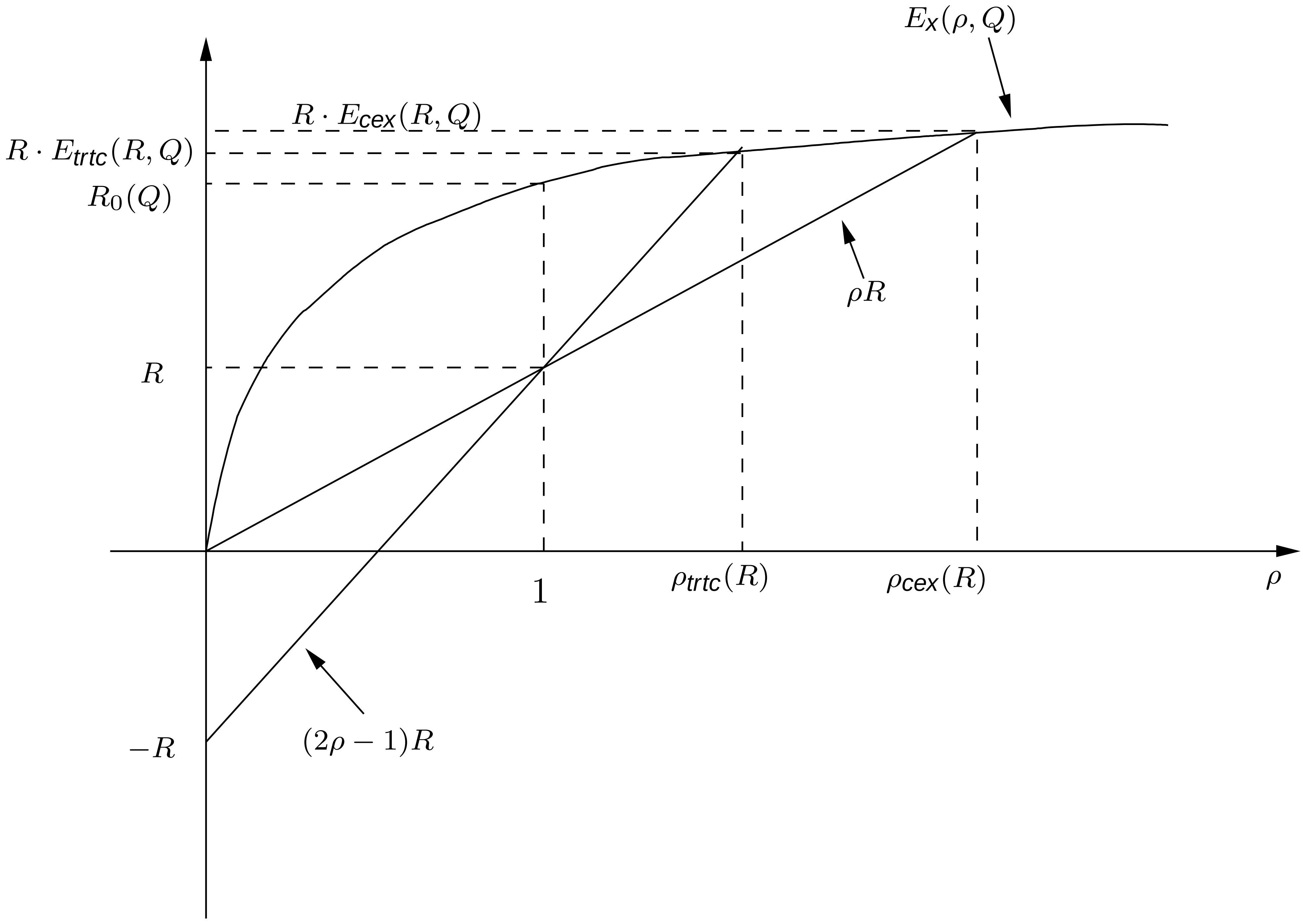
$$\mathcal{E}_{\mathsf{trtc}}(R,Q) \geq E_{\mathsf{trtc}}(R,Q) \triangleq \frac{E_{\mathsf{X}}(\rho,Q)}{R},$$

where ρ satisfies

$$\frac{E_{\mathsf{X}}(\rho,Q)}{2\rho-1} = R.$$

Also,

$$\mathcal{E}_{\mathsf{trcc}}(R,Q) \geq E_{\mathsf{cex}}(R,Q).$$


$E_x(\rho, Q)$
$R \cdot E_{cex}(R, Q)$
$R \cdot E_{trtc}(R, Q)$
$R_0(Q)$
ρR
R
1
$\rho_{trtc}(R)$
$\rho_{cex}(R)$
ρ
$(2\rho - 1)R$
$-R$

Characterizing the Typical Codes

The probability of error

$$P_{\mathsf{e}}(\mathcal{C}) \leq \sum_{\ell \geq 1} 2^{-m\ell} \sum_{\{P_{XX'}\}} N_\ell(P_{XX'}) \exp\{-n(k+\ell)\Delta(P_{XX'})\}.$$

For typical codes, all $\{N_\ell(P_{XX'})\}$ that have a small expectation, vanish simultaneously w.h.p.

This amounts to the condition

$$2\ell R < (k+\ell)D(P_{XX'}\|Q \times Q).$$

Joint types that are "too far" from $Q \times Q$ are not populated.

For populated joint types,

$$N_\ell(P_{XX'}) \sim \exp\{m[2\ell - (k+\ell)D((P_{XX'}\|Q \times Q)/R]\}.$$

An Alternative Expression

$$E_{\mathsf{trtc}}(R,Q) = \inf_{\hat{R}<R} \inf_{\{P_{XX'}:\ D(P_{XX'}\|Q\times Q)<2\hat{R}\}} \frac{\mathbf{E}_P d_{\mathsf{B}}(X,X') + \hat{R}}{R-\hat{R}}.$$

Dominant error events:

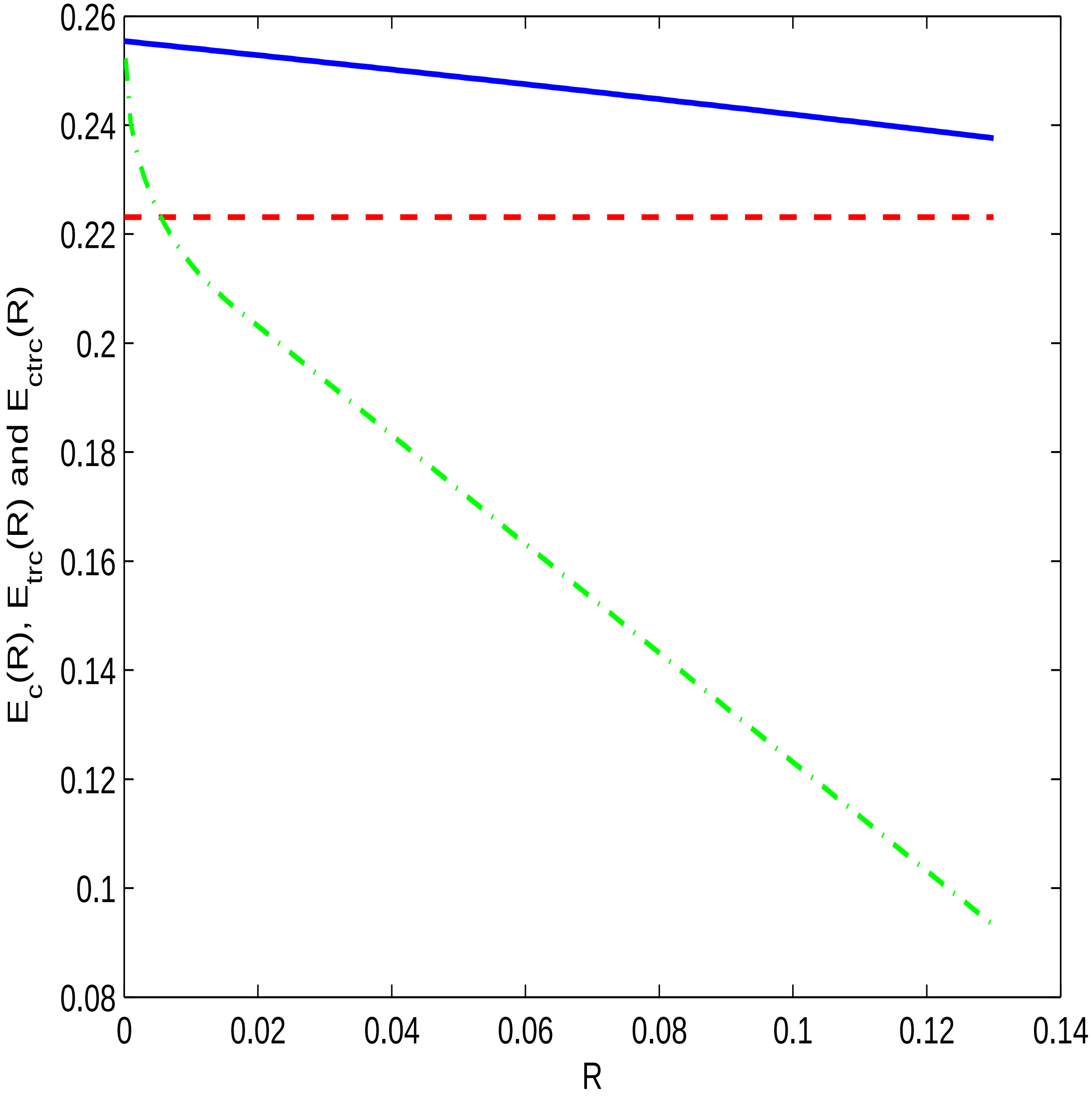
A sub–exponential number of paths with joint type,

$$P_{XX'}(x,x') = \frac{Q(x)Q(x')\exp\{-d_{\mathsf{B}}(x,x')/\rho\}}{Z}$$

and critical length of

$$k+\ell = \frac{kR}{2R - D(P_{XX'}\|Q\times Q)}.$$

A Numerical Example: BSC with $p = 0.1$



Two Words Regarding an Extension

The paper contains also an extension to a channel with mismatch and input memory (ISI),

$$W(\boldsymbol{y}|\boldsymbol{x}) = \prod_t W(y_t|x_t, x_{t-1}).$$

$E_{\mathsf{X}}(\rho, Q)$ is now replaced by $-\rho \log \lambda$, where λ is the Perron-Frobenius eigenvalue of a certain matrix that depends on the Q and on the channel single–letterly (details – in the paper).