

Real-Time Communication Systems based on Taut Strings

Eric Setterqvist and Robert Forchheimer

The self-archived postprint version of this journal article is available at Linköping University Institutional Repository (DiVA):

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-148263>

N.B.: When citing this work, cite the original publication.

Setterqvist, E., Forchheimer, R., (2018), Real-Time Communication Systems based on Taut Strings, *Journal of Communications and Networks*, 20(2), 207-218.

<https://doi.org/10.1109/JCN.2018.000027>

Original publication available at:

<https://doi.org/10.1109/JCN.2018.000027>

Copyright: Institute of Electrical and Electronics Engineers (IEEE)

<http://www.ieee.org/index.html>

©2018 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.



Real-Time Communication Systems based on Taut Strings

Eric Setterqvist and Robert Forchheimer

Abstract—We consider buffered real-time communication over channels with time-dependent capacities which are known in advance. The real-time constraint is imposed in terms of limited transmission time between sender and receiver. For a network consisting of a single channel it is shown that there is a coding rate strategy, geometrically characterized as a taut string, which minimizes the average distortion with respect to all convex distortion-rate functions. Utilizing the taut string characterization further, an algorithm that computes the optimal coding rate strategy is provided. We then consider more general networks with several connected channels in parallel or series with intermediate buffers. It is shown that also for these networks there is a coding rate strategy, geometrically characterized as a taut string, which minimizes the average distortion with respect to all convex distortion-rate functions. The optimal offline strategy provides a benchmark for the evaluation of different coding rate strategies. Further, it guides us in the construction of a simple but rather efficient strategy for channels in the online setting which alternates between a good and a bad state.

Index Terms—Buffers, Distortion, Real-time systems, Source coding, Taut string.

I. INTRODUCTION

TODAY, streaming of video and other media have become quite popular applications on the Internet. Further, the Internet is used for broadcasting of live events and demanding interactive services such as video conferences and online gaming.

On the Internet and other IP based computer networks the available transmission capacity typically varies significantly over time. This implies variable distortion, measured by distortion-rate functions, throughout the sending period. One way of smoothing the variation in transmission capacity is to use buffers to temporarily store data. However, small playback delay and other real-time constraints impose bounds on the transmission time between sender and receiver, i.e. data can only be stored for a prescribed time in the buffers. Given constraints on network capacity and transmission time, a problem then emerges concerning how to control buffers and coding rate in order to minimize the average distortion.

Although today's real-time services typically work well there is a price for this in terms of over provisioning of network resources. With emerging services like streaming of 4K video one can expect that the demand and need of network

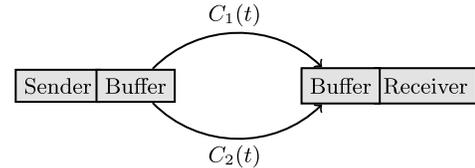


Fig. 1. Communication system with two parallel channels with capacities $C_1(t)$ and $C_2(t)$ respectively.



Fig. 2. Communication system with two serial channels with capacities $C_1(t)$ and $C_2(t)$ respectively.

resources will continue to grow. Coding rate strategies which efficiently utilize the network resources can therefore moderate the need of new investments in network infrastructure.

In this paper we consider communication systems consisting of one sender and one receiver which are linked through a network of channels connected in parallel or in series. The sender has an output buffer, all intermediate nodes between sender and receiver act as combined input and output buffers while the receiver has an input buffer. Further, the sender contains an encoder which through lossy compression makes an approximate representation of the source to fit within the network capacity while the receiver contains a decoder for approximating/reconstructing the source. The channel capacities of the network are time-dependent and are assumed to be known. Note that we allow for each channel to have its own unique time-dependent capacity function. See Figures 1 and 2 for examples of the two types of networks under study.

With deterministic channels, we are studying the so called offline setting. The assumption of knowing the future channel capacities might be considered restrictive. However, there are applications where this assumption seems to be applicable. One example is Internet communication on trains where the train operator provides WiFi to its passengers through train aggregation routers. These routers are connected to multiple mobile operators in order to maximize performance and reliability. For an overview of communication technologies for high speed trains see e.g. [1] and [2]. In [3], communication characteristics of train aggregation routers are analyzed on a large data set from Swedish railways. Substantial delays occur during sections of many journeys making real-time applications challenging. The collected data set suggests that the geographic variability of the aggregate transmission capacity along the routes are approximately stationary in time. Also,

E. Setterqvist is with the Department of Mathematics, Linköping University, SE-581 83 Linköping, Sweden. Present address: Computational Science Center, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria (e-mail: eric.setterqvist@univie.ac.at)

R. Forchheimer is with the Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden (e-mail: robert.forchheimer@liu.se).

the number of passengers, and thus the available transmission capacity for each of them for different sections of the train route can be approximated in advance from the ticket sales. With such advance estimates of the transmission capacity available, offline coding rate strategies becomes applicable.

The main contribution of this paper is the following: for the parallel and serially connected networks with given available capacity there exists a unique coding rate strategy which minimizes the average distortion for *all* convex distortion-rate functions. One appealing feature of the optimal coding rate strategy is that it can be geometrically interpreted as a taut string. Besides giving a characterization, this interpretation provides a framework for an algorithm for computing the optimal coding rate strategy. In the online setting, where the channel capacities are not given in advance, the optimal coding rate strategy gives a benchmark to compare heuristic coding rate strategies with. Further, the properties of the optimal coding rate strategy guides us to the construction of a simple but rather efficient coding rate strategy for a channel which randomly alternates between a good and bad state.

The paper is organized as follows. First, related work are discussed in Section II. In Section III, the general coding rate problem is introduced. Taut strings are introduced in Section IV and it is shown that the coding rate problem, under certain conditions, can be viewed as a taut string problem. Thereafter, in Section V, an algorithm for computing the optimal coding rate strategy is given. Analyzing different types of networks, Section VI considers the coding rate problem for broadcasting of live events while Section VII considers the corresponding problem in the case of stored media. In Section VIII we provide a heuristic strategy, the *mean strategy*, for a channel which randomly alternates between a good and bad state. Numerical evaluations are done which compare the performance of the optimal coding rate strategy, the mean strategy and a greedy strategy in this setting. Finally, conclusions are found in Section IX.

II. RELATED WORK

Buffered data transmission have been considered in various contexts. The practical motivation for our work comes from [4] which considered buffer management strategies, based on rate control through source coding, for a real-time communication system consisting of one sender and one receiver linked by a channel with known time-dependent capacity. In a broader context, rate-distortion optimized strategies for real-time transmission of video have attracted considerable interest and we mention as examples of investigations [5]-[12]. When exploring this area of research, we have not found works with focus on rigorous establishment of optimal coding rate strategies which is the theme of our paper.

In [13], the authors consider transmission of a stored pre-encoded variable-bit-rate video from a server to a client. Given a client buffer of fixed size, an algorithm that achieves the greatest possible reduction in rate variability (optimal smoothing) when transmitting the video is presented. It is shown in [13] that the transmission schedule computed by the algorithm is optimal with respect to all convex rate functions.

The results of [13] was extended to a network of serially connected channels in [14]. In comparison, we study strategies for minimizing the average distortion by adaptively encoding the media to cope with time-varying network conditions. Moreover, the techniques used in our paper are different than the approach of [13], which is based on the theory of majorization, and allow for extension to some classes of time-varying distortion-rate functions.

Using techniques based on calculus of variations, energy-efficient transmission schedules for a single channel with QoS-constraints are investigated in [15]. In this setting, it is shown that there exists a transmission schedule which minimizes all convex power-rate functions. Also, time-varying power-rate functions are considered in [15].

We next mention works in the area of energy harvesting communications system which are related to our investigations. In [16], the authors study the problem of minimizing the transmission time for a given amount of data in a system with finite energy storage capacity and causality constraints on data and energy arrivals. For the offline situation, the resulting problem may be viewed as a taut string problem. The authors outline a principle, named directional water-filling algorithm, for computing the optimal offline transmission strategy. In [17], structural properties of optimal offline strategies are investigated and explicit algorithms for computing such strategies are provided. Besides considering a different application, the investigations in [16] and [17] do not address the invariance of the optimal offline transmission strategy with respect to convex functions and do not consider intermediate nodes between the transmitter and the receiver. Further, the algorithms of [17] are iterative and hence differs from the recursive algorithm given in this paper.

Considerable efforts have been devoted to the study of buffer management strategies for network switches supporting Quality of Service (QoS). We mention in particular the pioneering work of [18]. For an overview of this area we refer to the survey [19]. In the QoS setting, each packet has an intrinsic value given by a positive real number. The goal of the network switch is to maximize the total value of the transmitted packets. The maximization of the value of transmitted packets is not influenced by the actual variation of the incoming flow of packets throughout the sending period. In our work on the other hand, the corresponding variation of channel capacities has a profound influence on the optimal coding rate strategy and therefore the management of the buffers. This is the fundamental difference between our work and [18].

We also note that the notion of taut strings have appeared in applications in statistics, see [20], in image processing, see [21, Chapter 4.4], and in stochastic processes, see [22].

III. MODEL SETTING AND FORMULATION OF THE CODING RATE PROBLEM

It is assumed that time is discrete with uniformly distributed time samples $\{t_1, t_2, \dots\}$ according to

$$t_{j+1} - t_j = \Delta, \quad j = 1, 2, \dots$$

where $\Delta > 0$. The time resolution is assumed to be sufficiently high for capturing the variations of the channels and the source. Besides time, all quantities considered in the paper are assumed to be continuous.

We assume the cut-through model where data can arrive and be transmitted from a node in the network at the same time instance. It should be observed though that the channels may have a time-uniform transmission delay, this has no effect on our study on minimization of the distortion.

The buffers are assumed to work according to the FIFO principle. Hence, data in the buffers are sent through the attached channels in the same order as they have entered the buffers.

Let $R(t_i)$ denote the amount of encoded data for representing the source at sample t_i . So, $R(t_i)$ can be interpreted as the source coding rate at t_i . The total distortion of the representation of a source with n time samples is given by

$$\sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R(t_i)}{\lambda(t_i)}\right) \quad (1)$$

where φ is a real-valued convex function on the real line and $\lambda(t_i) \in (0, \infty)$ is a weight. When dividing (1) with $t_n - t_1$, we obtain the average distortion. Note that the class of functions φ includes the classical distortion-rate functions, i.e. continuous monotonically decreasing convex functions, with one example being $\varphi(x) = 4^{-x}$. The weights $\lambda = (\lambda(t_i))_{i=1}^n$ gives possibility to take into account some time-varying distortion-rate functions.¹

We impose a maximal transmission time $m\Delta$, for some $m \in \{1, 2, \dots\}$, between sender and receiver in order to have a bound on the playback delay. That is, data can be stored for a prescribed time $m\Delta$ in the buffers. Given a source with samples at t_1, \dots, t_n and transmission delay $m\Delta$, we distinguish between coding period, playback period and sending period. These periods are visualized in Figure 3.

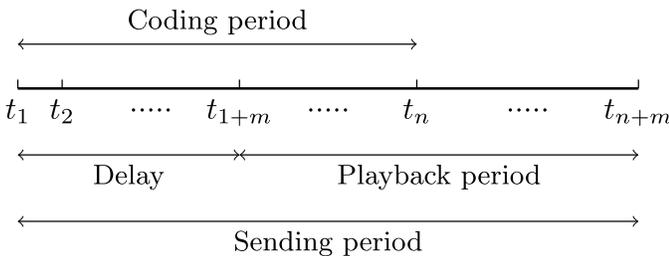


Fig. 3. The different time periods.

Let $U = (U(t_i))_{i=1}^n \in \mathbb{R}^n$ denote the cumulative transmission and storage capacity of the communication system. That is, $U(t_i)$ is the cumulative capacity at t_i taking into account the full capacities of the channels and the full storage capacities of the buffers under the delay constraint. The corresponding coding rate strategy $R_U = (R_U(t_i))_{i=1}^n$ where $R_U(t_1) = U(t_1)$ and $R_U(t_i) = U(t_i) - U(t_{i-1})$, $i = 2, \dots, n$, can be thought of as the greedy coding rate strategy which at each time instance stores as much data as possible in the buffers without thinking

¹For instance, the distortion-rate function for Gaussian sources with time-varying bandwidth.

of need at future time instances. Let $R_L = (R_L(t_i))_{i=1}^n$ denote the coding rate strategy which on the contrary utilize the buffers as late as possible during the coding period in order to guarantee that $\sum_{i=1}^n R_L(t_i) = U(t_n)$, i.e. both strategies R_U and R_L will result in the same total amount of encoded data. By $L = (L(t_i))_{i=1}^n \in \mathbb{R}^n$, we denote the cumulative amount of encoded data resulting from the strategy R_L . The cumulative curves U and L correspond to the notion of arrival curve respectively minimal departure curve in the theory of network calculus, see [23].

In the case of a single channel between the sender and the receiver we have $R_L(t_i) = C(t_i)$, $i = 1, \dots, n - 1$. That is, at t_i for $i = 1, \dots, n - 1$ the quantity $L(t_i)$ corresponds to the cumulative data transmission between the sender and the receiver and the difference $U(t_i) - L(t_i)$ is the amount of buffered data at the sender. However, $R_L(t_n) = \sum_{j=n}^{n+m} C(t_j)$ in order to utilize the allowed delay and take into account the channel capacities at t_{n+1}, \dots, t_{n+m} .

If there are several channels connected in series between the sender and the receiver the situation becomes more complicated. Then there is a possibility that the capacity bottleneck alters between different channels during the sending period. This means that the strategy R_L might need to save data in the buffers also during the time instances t_1, \dots, t_{n-1} in order to fully utilize the transmission capacity of the system. We will consider the derivation of L and U in detail for specific classes of communication systems in Sections VI and VII.

With L and U introduced, the set $\Omega_{L,U}$ is defined:

Definition 1 (The set $\Omega_{L,U}$) The set $\Omega_{L,U} \subset \mathbb{R}^n$ contains all vectors $R = (R(t_i))_{i=1}^n \in \mathbb{R}^n$ which fulfills

$$R(t_j) \geq 0 \quad (2)$$

and

$$L(t_j) \leq \sum_{i=1}^j R(t_i) \leq U(t_j) \quad (3)$$

for $j = 1, \dots, n$.

It is clear that any admissible coding rate strategy will satisfy (2) and (3) and therefore is in $\Omega_{L,U}$. We will consider communication systems where the conditions (2) and (3) characterize an admissible coding rate strategy R . For such systems, the set $\Omega_{L,U}$ is precisely the set of all admissible coding rate strategies and we obtain the following coding rate problem:

Problem 1 (The coding rate problem) Given L , U , and λ over a coding period $\{t_1, \dots, t_n\}$ and a convex distortion-rate function φ , find a strategy $R_{\varphi,\lambda} \in \Omega_{L,U}$ which fulfills

$$\sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R_{\varphi,\lambda}(t_i)}{\lambda(t_i)}\right) = \min_{R \in \Omega_{L,U}} \sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R(t_i)}{\lambda(t_i)}\right).$$

IV. THE TAUT STRING PROBLEM AND THE OPTIMAL CODING RATE STRATEGY

In this section we will analyze optimal coding rate strategies when the conditions (2) and (3) characterize an admissible

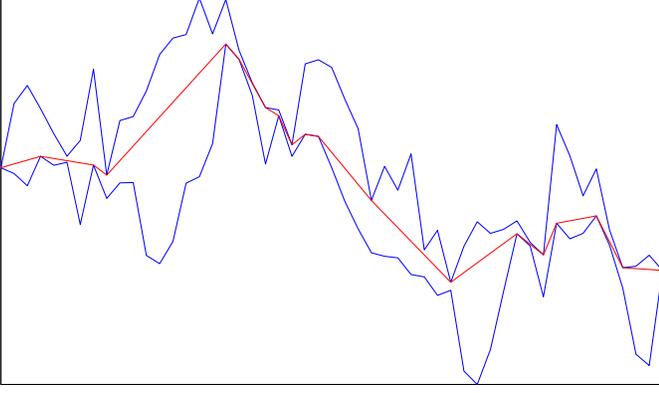


Fig. 4. Illustration of the taut string problem which concerns finding the function with minimal graph length (colored in red) inside a corridor defined by two continuous piecewise linear functions (colored in blue).

coding rate strategy. By establishing a close connection of the coding rate problem and the taut string problem, we will show that there is a unique coding rate strategy which is optimal with respect to every convex function φ .

Let us introduce the taut string problem. Given an interval $[a, b] \subset \mathbb{R}$ and $n + 1$ points $a = x_0 < x_1 < \dots < x_n = b$, consider two continuous functions F and G on $[a, b]$ which are linear on the intervals $[x_{i-1}, x_i]$, $i = 1, \dots, n$, with $F \leq G$, $F(a) = G(a)$ and $F(b) = G(b)$. Given such functions F and G , we now introduce the set $\Gamma_{F,G}$:

Definition 2 Let $\Gamma_{F,G}$ denote the set of all continuous functions H which are linear on the intervals $[x_{i-1}, x_i]$, $i = 1, \dots, n$, and satisfy $F \leq H \leq G$.

The taut string problem can now be formulated as

Problem 2 Find the function $H_* \in \Gamma_{F,G}$ that satisfies

$$\int_a^b \sqrt{1 + (H'_*(x))^2} dx = \min_{H \in \Gamma_{F,G}} \int_a^b \sqrt{1 + (H'(x))^2} dx.$$

Note that the derivative of $H \in \Gamma_{F,G}$ is defined everywhere except at x_i , $i = 0, 1, \dots, n$. We will interpret the derivative of H as a piecewise constant function on the intervals (x_{i-1}, x_i) , $i = 1, \dots, n$. Denote the set of the derivatives of the functions in $\Gamma_{F,G}$ with $\Gamma'_{F,G}$. The existence and uniqueness of H_* follow from the facts that $\Gamma'_{F,G}$ is a closed and convex subset of piecewise constant functions, $F(a) = G(a)$ and $\varphi(x) = \sqrt{1 + x^2}$ is a strictly convex function on \mathbb{R} .

The minimization with respect to graph length provides a geometrical characterization of H_* , i.e. the graph of H_* can be viewed as a taut string between $(a, F(a))$ and $(b, F(b))$ inside the corridor with lower boundary given by $F(x)$ and upper boundary given by $G(x)$, $a \leq x \leq b$. For an illustration, see Figure 4.

A notable feature of the taut string H_* is that it satisfies

$$\int_a^b \varphi(H'_*(x)) dx = \min_{H \in \Gamma_{F,G}} \int_a^b \varphi(H'(x)) dx \quad (4)$$

for all convex functions $\varphi : \mathbb{R} \rightarrow \mathbb{R}$. For a proof see e.g. Theorem 5.2 and Remark 5.2 in [24]. So H_* minimizes, besides graph length, also energy (choose $\varphi(x) = x^2$) and

total variation (choose $\varphi(x) = |x|$) among the functions in $\Gamma_{F,G}$.

We now link the taut string problem to the coding rate problem under study. Let λ , L and U be given over a coding period $\{t_1, \dots, t_n\}$. We associate a certain set $\Gamma_{F_{L,\lambda}, G_{U,\lambda}}$ to λ , L and U .

Definition 3 (The set $\Gamma_{F_{L,\lambda}, G_{U,\lambda}}$) Given λ , L and U over a coding period $\{t_1, \dots, t_n\}$, put $x_0 = 0$ and $x_j = x_{j-1} + \lambda(t_j)$, $j = 1, \dots, n$. Construct associated continuous piecewise linear functions $F_{L,\lambda}$ and $G_{U,\lambda}$ on the intervals $[x_{j-1}, x_j]$, $j = 1, \dots, n$, according to

$$\begin{aligned} F_{L,\lambda}(x_0) &= 0, \\ F_{L,\lambda}(x_j) &= L(t_j), \quad j = 1, \dots, n, \end{aligned} \quad (5)$$

and

$$\begin{aligned} G_{U,\lambda}(x_0) &= 0, \\ G_{U,\lambda}(x_j) &= U(t_j), \quad j = 1, \dots, n. \end{aligned} \quad (6)$$

Let $\Gamma_{F_{L,\lambda}, G_{U,\lambda}}$ denote the set of all continuous piecewise linear functions on the intervals $[x_{j-1}, x_j]$, $j = 1, \dots, n$, which are in between $F_{L,\lambda}$ and $G_{U,\lambda}$.

We have now introduced the necessary notation in order to formulate our main result:

Theorem 1 Let λ , L and U be given over a coding period $\{t_1, \dots, t_n\}$. Suppose the set $\Omega_{L,U}$ characterize the set of all admissible coding rate strategies. Then there is a unique admissible coding rate strategy $R_\lambda \in \Omega_{L,U}$ which satisfies

$$\sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R_\lambda(t_i)}{\lambda(t_i)}\right) = \min_{R \in \Omega_{L,U}} \sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R(t_i)}{\lambda(t_i)}\right)$$

for all convex functions $\varphi : \mathbb{R} \rightarrow \mathbb{R}$. Moreover, R_λ is given by

$$R_\lambda(t_i) = H_*(x_i) - H_*(x_{i-1}), \quad i = 1, \dots, n,$$

where H_* is the taut string in $\Gamma_{F_{L,\lambda}, G_{U,\lambda}}$.

Proof. Recalling (4) and the definition of $\Gamma_{F_{L,\lambda}, G_{U,\lambda}}$, we conclude that the taut string $H_* \in \Gamma_{F_{L,\lambda}, G_{U,\lambda}}$ satisfies

$$\begin{aligned} \int_{x_0}^{x_n} \varphi(H'_*(x)) dx &= \\ \sum_{i=1}^n (x_i - x_{i-1}) \varphi\left(\frac{H_*(x_i) - H_*(x_{i-1})}{x_i - x_{i-1}}\right) &= \\ \sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{H_*(x_i) - H_*(x_{i-1})}{\lambda(t_i)}\right) &\leq \\ \sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{H(x_i) - H(x_{i-1})}{\lambda(t_i)}\right) &= \\ \sum_{i=1}^n (x_i - x_{i-1}) \varphi\left(\frac{H(x_i) - H(x_{i-1})}{x_i - x_{i-1}}\right) &= \\ \int_{x_0}^{x_n} \varphi(H'(x)) dx & \end{aligned} \quad (7)$$

for all $H \in \Gamma_{F_{L,\lambda}, G_{U,\lambda}}$ and all convex functions $\varphi : \mathbb{R} \rightarrow \mathbb{R}$.

Let $R_\lambda = (R_\lambda(t_j))_{j=1}^n \in \mathbb{R}^n$ be defined by

$$R_\lambda(t_j) = H_*(x_j) - H_*(x_{j-1}), \quad j = 1, \dots, n.$$

As $H_* \in \Gamma_{F_{L,\lambda}, G_{U,\lambda}}$, it is clear that R_λ satisfies (3). Since both $F_{L,\lambda}$ and $G_{U,\lambda}$ are increasing functions, it follows that the taut string H_* also is an increasing function since it minimizes graph length in $\Gamma_{F_{L,\lambda}, G_{U,\lambda}}$. This means that the element R_λ also must satisfy (2). So, $R_\lambda \in \Omega_{L,U}$ and is therefore an admissible coding rate strategy by assumption.

For a given element $R \in \Omega_{L,U}$, we can construct a function $H \in \Gamma_{F_{L,\lambda}, G_{U,\lambda}}$ where $H(x_0) = 0$ and

$$H(x_j) = \sum_{i=1}^j R(t_i), \quad j = 1, \dots, n.$$

Therefore, we obtain from the estimates of (7) that

$$\sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R_\lambda(t_i)}{\lambda(t_i)}\right) = \min_{R \in \Omega_{L,U}} \sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R(t_i)}{\lambda(t_i)}\right) \quad (8)$$

holds for all convex functions $\varphi : \mathbb{R} \rightarrow \mathbb{R}$. As H_* is the unique function with shortest graph length in $\Gamma_{F_{L,\lambda}, G_{U,\lambda}}$, it follows that R_λ is the unique strategy in $\Omega_{L,U}$ which fulfills (8). \square

Remark 1 If $x_j - x_{j-1} = \lambda(t_j) = c$ for $j = 1, \dots, n$, i.e. the distortion-rate function is uniform throughout the coding period, then R_λ can be interpreted as the derivative of H_* . More precisely, the coefficients of the piecewise constant function H_*' multiplied with the factor c equal the elements of the vector $R_\lambda \in \mathbb{R}^n$.

V. AN ALGORITHM FOR THE CODING RATE PROBLEM

Given λ, L and U over a coding period $\{t_1, \dots, t_n\}$, we now consider an algorithm for computing the optimal coding rate strategy R_λ when the conditions

$$R(t_j) \geq 0,$$

and

$$L(t_j) \leq \sum_{i=1}^j R(t_i) \leq U(t_j),$$

for $j = 1, \dots, n$ characterize an admissible coding rate strategy R . In order to do this, we first introduce Algorithm 1 for computing the taut string $H_* \in \Gamma_{F,G}$ given continuous piecewise linear functions F and G on a general collection of intervals $[x_{i-1}, x_i]$, $i = k+1, \dots, m$, with $F \leq G$, $F(x_k) = G(x_k)$ and $F(x_m) = G(x_m)$. We have:

Lemma 1 *The taut string algorithm computes the taut string $H_* \in \Gamma_{F,G}$ in a finite number of steps.*

Proof. For the proof of Lemma 1 we refer to the Appendix. \square

The taut string algorithm is recursive and divides, until termination, the original problem into smaller and smaller subproblems in accordance with a binary tree structure. One might view the algorithm as a binary tree sort. As such, the algorithm has worst-case complexity $O((m-k)^2)$ and average-complexity $O((m-k) \log(m-k))$ when considering a taut string problem over the intervals $[x_{i-1}, x_i]$, $i = k+1, \dots, m$.

Algorithm 1 Taut string algorithm

```

function TAUT( $k, m, F, G$ )
   $H_*(x_k) = F(x_k)$ 
   $H_*(x_m) = F(x_m)$ 
  for  $i = k$  to  $m$  do
     $V(x_i) = \frac{x_m - x_i}{x_m - x_k} H_*(x_k) + \frac{x_i - x_k}{x_m - x_k} H_*(x_m)$ 
  end for
   $C = \max_{i \in \{k+1, \dots, m-1\}} \left( \min_{x \in [F(x_i), G(x_i)]} |x - V(x_i)| \right)$ 
  if  $C = 0$  then
    for  $i = k+1$  to  $m-1$  do
       $H_*(x_i) = V(x_i)$ 
    end for
  else
    Find an index  $j \in \{k+1, \dots, m-1\}$  such that
     $\min_{x \in [F(x_j), G(x_j)]} |x - V(x_j)| = C$ 
    if  $V(x_j) < F(x_j)$  then
       $H_*(x_j) = F(x_j)$ 
    else
       $H_*(x_j) = G(x_j)$ 
       $F(x_j) = G(x_j)$ 
    end if
    if  $j - k > 1$  then
      return TAUT( $k, j, F, G$ )
    end if
    if  $m - j > 1$  then
      return TAUT( $j, m, F, G$ )
    end if
  end if
end function

```

With the taut string algorithm given, we now turn to the problem of computing the optimal coding rate strategy. Construct the continuous piecewise linear functions $F_{L,\lambda}$ and $G_{U,\lambda}$ according to Definition 3. For computing the optimal coding rate strategy, Algorithm 2 is now introduced.

Algorithm 2 The coding rate algorithm

```

Run TAUT( $0, n, F_{L,\lambda}, G_{U,\lambda}$ )
for  $i = 1$  to  $n$  do
   $R_\lambda(t_i) = H_*(x_i) - H_*(x_{i-1})$ 
end for

```

We have the following result:

Theorem 2 *The coding rate algorithm constructs the optimal coding rate strategy R_λ .*

Proof. From Lemma 1 we know that TAUT($0, n, F_{L,\lambda}, G_{U,\lambda}$) computes the taut string $H_* \in \Gamma_{F_{L,\lambda}, G_{U,\lambda}}$. With the taut string H_* at our disposal, we can compute $R_\lambda(t_i) = H_*(x_i) - H_*(x_{i-1})$, $i = 1, \dots, n$, as was established in Theorem 1. \square

VI. LIVE BROADCASTING

In this section we consider live media as the source material. It will be shown that there exists a unique optimal coding rate

strategy, with respect to all convex distortion-rate functions, for networks of channels in parallel and series between sender and receiver. We assume a coding period $\{t_1, \dots, t_n\}$ and a playback delay of $m\Delta$ units of time which results in a sending period $\{t_1, \dots, t_{n+m}\}$.

In contrast to stored media considered in the next section, for live media we do not have possibility to encode data of the source sample $R(t_j)$ in advance at t_i , $i < j$. Considering a sending period with given channel capacities, the greater flexibility of coding of stored media means that its optimal coding rate strategy will have lower or at most equal distortion compared with the optimal coding rate strategy for live media. For a comparison of the situations it might be instructive to compare Figures 6 and 9 where the corridors defined by L and U are illustrated for the same basic communication system transmitting live respectively stored media.

A. The basic communication system

Consider the basic communication system consisting of a sender, with an output buffer, and a receiver, with an input buffer, which are linked by a communication channel with known capacity $C(t)$. The lower and upper bounds L and U

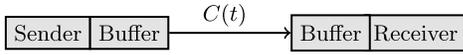


Fig. 5. The basic communication system.

of the cumulative amount of encoded source data are given by

$$U(t_j) = \sum_{i=1}^{j+m} C(t_i), j = 1, \dots, n,$$

and

$$L(t_j) = \begin{cases} \sum_{i=1}^j C(t_i), j = 1, \dots, n-1, \\ \sum_{i=1}^{n+m} C(t_i), j = n. \end{cases}$$

In addition to the sent data $\sum_{i=1}^j C(t_i)$, i.e. data that have reached the receiver at t_j , $U(t_j)$ store $\sum_{i=j+1}^{j+m} C(t_i)$ amount of data in the sender buffer corresponding to the sum of channel capacities for the next m time instances (recall the delay $m\Delta$). Considering L , we should for any t_i , $i = 1, \dots, n-1$, have used the cumulative channel capacity for transmitting encoded source data to the receiver. At t_n , the end of the coding period, we should have used all available channel capacity during the sending period.

In Figure 6, we illustrate an example of a corridor given by L and U . The dotted piecewise linear graph, i.e. the shortest path between $(0,0)$ and $(t_6, U(t_6))$, corresponds to the cumulative optimal coding rate strategy.

Given L and U , recall that the set $\Omega_{L,U}$ is the set of all strategies $R = (R(t_j))_{j=1}^n$ which satisfies

$$R(t_j) \geq 0$$

and

$$L(t_j) \leq \sum_{i=1}^j R(t_i) \leq U(t_j)$$

for $j = 1, \dots, n$. For the basic communication system, any $R \in \Omega_{L,U}$ is an admissible coding rate strategy as we

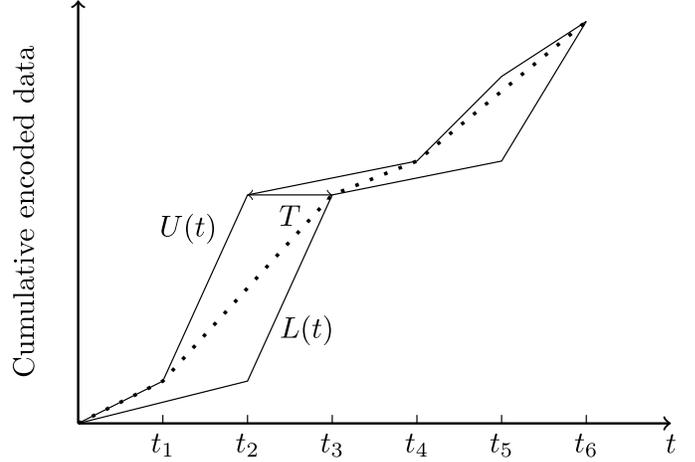


Fig. 6. Illustration of corridor given by L and U for coding of live media with coding period $\{t_1, \dots, t_6\}$ and delay $T = \Delta$ for a basic communication system. The dotted piecewise linear graph corresponds to the cumulative optimal coding rate strategy.

can redistribute buffer storage between different time samples within the constraints given by channel capacities and delay.

From Theorem 1 now follows that there is a unique coding rate strategy $R_\lambda \in \Omega_{L,U}$ which satisfies

$$\sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R_\lambda(t_i)}{\lambda(t_i)}\right) = \min_{R \in \Omega_{L,U}} \sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R(t_i)}{\lambda(t_i)}\right)$$

for all convex functions $\varphi: \mathbb{R} \rightarrow \mathbb{R}$. Further, we can find R_λ by the coding rate algorithm of Section V.

B. Generalization of the basic communication system: channels in parallel

Consider a system of N parallel communication channels between the sender and the receiver. See Figure 7 for an illustration when $N = 2$. The N channels are assumed to have

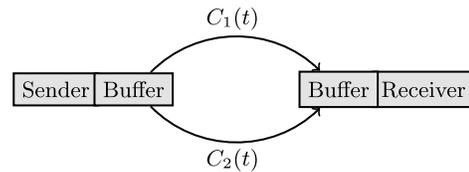


Fig. 7. Communication system with two parallel channels.

known and independent capacities $C_1(t_i), \dots, C_N(t_i)$ throughout the sending period $\{t_1, \dots, t_{n+m}\}$. This system can be reduced to a basic communication system of Subsection VI-A with a single channel of capacity $C(t_i) = C_1(t_i) + \dots + C_N(t_i)$ together with an attached output buffer. So, for the parallel system there exists a unique coding rate strategy which minimizes the total distortion for all convex distortion-rate functions and is computable by the coding rate algorithm.

C. Generalization of the basic communication system: channels in series

We now turn to a communication system with N serially connected channels. The channels have known capacities

$C_1(t_i), \dots, C_N(t_i)$ during the sending period $\{t_1, \dots, t_{n+m}\}$. See Figure 8 for an illustration of a system with two serially connected channels.



Fig. 8. Communication system with two serial channels.

Let us now derive the lower and upper bounds L and U of the cumulative transmission capacity of the system. The situation is more complicated than for the basic communication system because the capacity bottleneck can alter between different channels during the sending period. Therefore, generic expressions of L and U cannot be obtained.

We start by determining the upper bound U . Let $b_i^U(t_j)$ denote the number of bits saved in buffer i , $i = 1, \dots, N$, at t_j , $j = 1, \dots, n$, by the coding rate strategy R_U corresponding to U . The buffers are numbered $1, 2, \dots$ starting from the sender and moving through the network towards the receiver. A procedure for determining U is given by Algorithm 3.

Remark 2 After the coding period, i.e. at time instances t_{n+1}, \dots, t_{n+m} redistribution of old stored data as in Algorithm 3 is done for bringing all data generated by R_U to the receiver.

Besides its greedy nature, the derivation of U follows two basic principles at each time instance t_i during the coding period. First, old stored data in the buffers, starting with buffer N and moving backwards, is sent as close as possible to the receiver. Then new data, encoded at t_i , is sent as close as possible to the receiver. So, the whole network between sender and receiver can be considered to work according to a FIFO principle.

We only consider strategies following these principles which will exclude some admissible coding rate strategies from further consideration. However, optimal coding rate strategies are not excluded by these principles which is now explained. First, prioritizing sending of previously stored data to the receiver release corresponding storage space in the buffers for data of $R(t_i)$. Further, with the distribution of previously stored data fixed, we should send the new data $R(t_i)$ as close to the receiver as possible in view of alternating capacity bottleneck between the channels at subsequent time instances.

Let $b_i^L(t_j)$ denote the number of bits saved in buffer i , $i = 1, \dots, N$, at t_j , $j = 1, \dots, n$, by the coding rate strategy R_L corresponding to L . From U , we can obtain L working according to Algorithm 4.

Remark 3 Redistribution of old stored data is done at t_{n+1}, \dots, t_{n+m} , as in Algorithm 3, in order to bring all data generated by R_L to the receiver.

Let us explain the main principles behind the construction of L . We start at t_n by assigning $L(t_n) = U(t_n)$. The buffer storage of L at t_n must equal the corresponding buffer storage of U , otherwise $L(t_n) \neq U(t_n)$. For $k = n, \dots, 2$, we subtract $\min_{1 \leq j \leq N} C_j(t_k)$ from $L(t_k)$ when constructing $L(t_{k-1})$. The quantity $\min_{1 \leq j \leq N} C_j(t_k)$ represents the amount of data which can be transmitted at t_k only using channel capacities. At $k = n$, we in addition subtract $b_1^L(t_n)$ as this quantity

Algorithm 3 A procedure for determining U

```

Initialization:
for  $i = 1$  to  $N$  do
   $b_i^U(t_1) = 0$ 
end for
 $U(t_1) = 0$ 
for  $k = 1$  to  $n$  do
  Redistribution of old stored data:
  for  $j = N$  to  $1$  do
    for  $i = N$  to  $j$  do
       $x = \min \left\{ b_j^U(t_k), \min_{j \leq l \leq i} C_l(t_k) \right\}$ 
      for  $l = j$  to  $i$  do
         $C_l(t_k) = C_l(t_k) - x$ 
      end for
       $b_j^U(t_k) = b_j^U(t_k) - x$ 
      if  $i < N$  then
         $b_i^U(t_k) = b_i^U(t_k) + x$ 
      end if
    end for
  end for
  Sending of new data to receiver buffer:
   $x = \min_{1 \leq i \leq N} C_i(t_k)$ 
  for  $i = 1$  to  $N$  do
     $C_i(t_k) = C_i(t_k) - x$ 
  end for
   $U(t_k) = U(t_k) + x$ 
  Sending of new data to intermediate buffers:
  if  $k < n$  then
    for  $j = N$  to  $1$  do
      if  $j > 1$  then
         $x_1 = \min_{1 \leq l \leq j-1} C_l(t_k)$ 
         $x_2 = \sum_{i=k+1}^{k+m} \min_{j \leq l \leq N} C_l(t_i) - \sum_{i=j}^N b_i^U(t_k)$ 
         $x = \min \{x_1, x_2\}$ 
        for  $i = 1$  to  $j$  do
           $C_i(t_k) = C_i(t_k) - x$ 
        end for
         $U(t_k) = U(t_k) + x$ 
         $b_j^U(t_k) = b_j^U(t_k) + x$ 
      else
         $x = \sum_{i=k+1}^{k+m} \min_{1 \leq l \leq N} C_l(t_i) - \sum_{i=1}^N b_i^U(t_k)$ 
         $U(t_k) = U(t_k) + x$ 
         $b_1^U(t_k) = b_1^U(t_k) + x$ 
      end if
    end for
  end if
  Initialization for  $t_{k+1}$ :
   $U(t_{k+1}) = U(t_k)$ 
  for  $i = 1$  to  $N$  do
     $b_i^U(t_{k+1}) = b_i^U(t_k)$ 
  end for
end if
end for
  
```

can be completely represented by data encoded at t_n . For $k = n - 1, \dots, 1$, $b_1^L(t_k) = 0$ as the strategy R_L does not

Algorithm 4 A procedure for determining L

```

 $L(t_n) = U(t_n)$ 
for  $i = 1$  to  $N$  do
   $b_i^L(t_n) = b_i^U(t_n)$ 
end for
for  $k = n$  to  $2$  do
  if  $k = n$  then
     $L(t_{k-1}) = L(t_k) - \min_{1 \leq j \leq N} C_j(t_k) - b_1^L(t_k)$ 
  else
     $L(t_{k-1}) = L(t_k) - \min_{1 \leq j \leq N} C_j(t_k)$ 
  end if
   $b_1^L(t_{k-1}) = 0$ 
  for  $i = 1$  to  $N$  do
     $C_i(t_k) = C_i(t_k) - \min_{1 \leq j \leq N} C_j(t_k)$ 
  end for
  if  $N > 1$  then
    for  $j = N$  to  $2$  do
      if  $\min_{j \leq i \leq N} C_i(t_k) > 0$  then
         $x = \min \left\{ \min_{j \leq i \leq N} C_i(t_k), b_j^U(t_k) \right\}$ 
        for  $i = j$  to  $N$  do
           $C_i(t_k) = C_i(t_k) - x$ 
        end for
         $b_j^L(t_{k-1}) = b_j^L(t_k) + x$ 
         $L(t_{k-1}) = L(t_{k-1}) + x$ 
      else if  $\min_{1 \leq i \leq j-1} C_i(t_k) > 0$  then
         $x = \min \left\{ \min_{1 \leq i \leq j-1} C_i(t_k), b_j^L(t_k) \right\}$ 
        for  $i = 1$  to  $j-1$  do
           $C_i(t_k) = C_i(t_k) - x$ 
        end for
         $b_j^L(t_{k-1}) = b_j^L(t_k) - x$ 
         $L(t_{k-1}) = L(t_{k-1}) - x$ 
      end if
    end for
  end if
end for
end if
end for

```

store data unless necessary. We then subtract $\min_{1 \leq j \leq N} C_j(t_k)$ from the channel capacities in order to calculate the remaining capacity at t_k . Next, it might be necessary to store data in the intermediate buffers in order to take advantage of alternating capacity bottleneck of the channels during the sending period. Therefore, we consider buffer j , $j = N, \dots, 2$, and see if it is possible at t_k to (i) send stored data from buffer j to the receiver or (ii) transmit new data to buffer j from the sender. These two possibilities are mutually exclusive. If either of them apply, the corresponding channel capacities are updated together with $b_j^L(t_{k-1})$ and $L(t_{k-1})$. Note that by construction is $b_j^U(t_k)$ the maximal possible amount of data in buffer j at t_k .

The derivation of L takes into account the alternating capacity bottleneck at different channels throughout the sending period and therefore the effect of intermediate buffers. Hence, the output buffer at the receiver can handle all strategies which

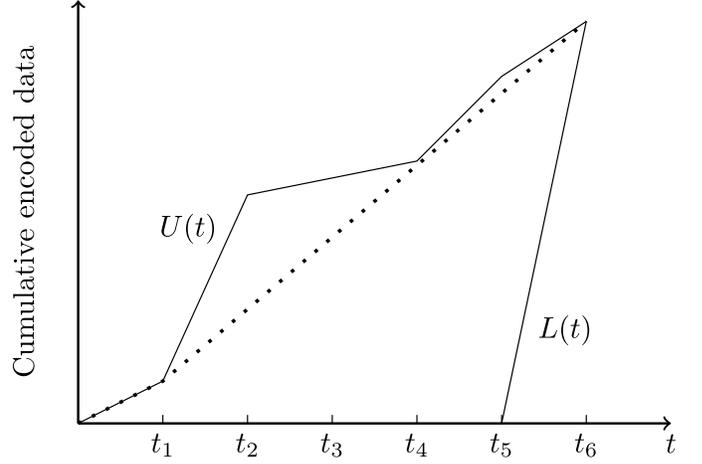


Fig. 9. Illustration of corridor given by L and U for coding of stored media with coding period $\{t_1, \dots, t_6\}$ and delay $T = \Delta$ for a basic communication system. The dotted piecewise linear graph corresponds to the cumulative optimal coding rate strategy.

cumulatively are in between L and U . So, as in the case of the basic communication system, the set $\Omega_{L,U}$ characterize all admissible coding rate strategies. From Theorem 1 then follows that the resulting coding rate problem admits a solution $R_\lambda \in \Omega_{L,U}$ which satisfies

$$\sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R_\lambda(t_i)}{\lambda(t_i)}\right) = \min_{R \in \Omega_{L,U}} \sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R(t_i)}{\lambda(t_i)}\right)$$

for all convex functions $\varphi : \mathbb{R} \rightarrow \mathbb{R}$. The strategy R_λ can be computed by the coding rate algorithm of Section V.

VII. STORED MEDIA

We now consider streaming of stored media for the type of networks discussed in Section VI. As before, we assume a coding period $\{t_1, \dots, t_n\}$ and a playback delay of $m\Delta$ units of time which results in a sending period $\{t_1, \dots, t_{n+m}\}$. The difference compared to Section VI is that we assume that the source material is stored in full quality at the sender. So, the entire source material is available for coding from the beginning of the sending period. Hence, in the case of stored media it is important to note that the encoded data at t_i do not need to correspond to the source sample $R(t_i)$ but can be part of any sample $R(t_j)$, $i \leq j \leq n$.

A. The basic communication system

For the basic communication system, the upper and lower bounds of the cumulative amount of encoded data are given by

$$U(t_j) = \sum_{i=1}^{j+m} C(t_i), j = 1, \dots, n,$$

and

$$L(t_j) = \begin{cases} 0, & j = 1, \dots, n-1, \\ \sum_{i=1}^{n+m} C(t_i), & j = n. \end{cases}$$

The upper bound U is the same as for live broadcasting but the lower bound L is different due to the greater flexibility

of coding the source material. Compare further Figures 6 and 9 where the corridors defined by L and U are illustrated for a basic communication system transmitting live respectively stored media.

With the increased flexibility of transmitting stored media it follows that every $R = (R(t_i))_{i=1}^n \in \mathbb{R}^n$ satisfying

$$R(t_j) \geq 0$$

and

$$L(t_j) \leq \sum_{i=1}^j R(t_i) \leq U(t_j)$$

for $j = 1, \dots, n$ can be realized by the basic communication system. To see this, note that at each time instance t_j during the coding period we encode an amount of data $R_U(t_j)$ and distribute this amount of data freely between the source samples at t_j, \dots, t_n . So, the set $\Omega_{L,U}$ characterizes all admissible coding rate strategies. Theorem 1 therefore implies that the resulting coding rate problem admits a solution R_λ which satisfies

$$\sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R_\lambda(t_i)}{\lambda(t_i)}\right) = \min_{R \in \Omega_{L,U}} \sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R(t_i)}{\lambda(t_i)}\right)$$

for all convex functions $\varphi : \mathbb{R} \rightarrow \mathbb{R}$. As in the case of live broadcasting, we can find R_λ by the coding rate algorithm of Section V.

B. Generalization of the basic communication system: channels in parallel

The parallel system can be analyzed as a basic communication system of Subsection VII-A with a single channel of capacity $C(t_i) = C_1(t_i) + \dots + C_N(t_i)$, $i = 1, \dots, n + m$, together with an attached output buffer.

C. Generalization of the basic communication system: channels in series

U is identical to live broadcasting while L is different and given by

$$L(t_j) = \begin{cases} 0, & j = 1, \dots, n - m, \\ U(t_n), & j = n. \end{cases}$$

The difference in the construction of L occurs as we have access to the entire source file from the beginning of the sending period and therefore can encode content in advance.

All strategies $R = (R(t_j))_{j=1}^n$ fulfilling

$$R(t_j) \geq 0$$

and

$$L(t_j) \leq \sum_{i=1}^j R(t_i) \leq U(t_j)$$

for $j = 1, \dots, n$ are admissible. This follows since at each time instance t_j , $j = 1, \dots, n$, we can encode an amount of data given by $R_U(t_j)$ but distribute this capacity freely between the source samples at t_j, \dots, t_n .

The admissible coding rate strategies are therefore characterized by the set $\Omega_{L,U}$. Theorem 1 then implies that the resulting coding rate problem admits a solution R_λ which satisfies

$$\sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R_\lambda(t_i)}{\lambda(t_i)}\right) = \min_{R \in \Omega_{L,U}} \sum_{i=1}^n \lambda(t_i) \varphi\left(\frac{R(t_i)}{\lambda(t_i)}\right)$$

for all convex functions $\varphi : \mathbb{R} \rightarrow \mathbb{R}$. Finally, the strategy R_λ can be computed by the coding rate algorithm of Section V.

VIII. ONLINE SETTING AND NUMERICAL EVALUATIONS

The aim of this section is to provide some insight into the online setting. In particular, we numerically evaluate the distortion of the offline strategy and two heuristic online strategies for randomly generated channel capacities. The evaluations are restricted to the basic communication system with a single channel between the sender and the receiver. To thoroughly investigate the online setting with accompanying numerical evaluations is a comprehensive research topic in itself and falls outside the scope of the present paper.

It is assumed that the channel randomly changes between two states during the sending period $\{t_1, \dots, t_{n+m}\}$ and that the states of the channel at different time instances are independent of each other. C_i denotes the channel capacity of state i , $i = 1, 2$, and p denotes the probability of state 1. This model of the channel can easily be extended to the case of N different states, $N \in \mathbb{N}$.

One of the heuristic strategies is the greedy strategy R_U , corresponding to U , which at each time instance during the coding period codes as much data as possible without considering the need at future time instances. The second heuristic strategy, denoted R_M , tries at each time instance during the coding period to code according to the expected capacity

$$(1 + m/n)(pC_1 + (1 - p)C_2) \quad (9)$$

of the channel (taking into account the maximal delay of m time samples). If $\sum_{i=1}^j R_M(t_i) > U(t_j)$ we update $R_M(t_j)$ according to

$$R_M(t_j) = U(t_j) - \sum_{i=1}^{j-1} R_M(t_i)$$

where $\sum_{i=1}^0 R_M(t_i) := 0$. Analogously, if $\sum_{i=1}^j R_M(t_i) < L(t_j)$ we set

$$R_M(t_j) = L(t_j) - \sum_{i=1}^{j-1} R_M(t_i).$$

By construction, we then have

$$L(t_j) \leq \sum_{i=1}^j R_M(t_i) \leq U(t_j), j = 1, \dots, n.$$

Algorithm 5 provides the mean strategy formulated in pseudocode.

Algorithm 5 The mean strategy

```

 $\sum_{i=1}^0 R_M(t_i) := 0$ 
for  $j = 1$  to  $n$  do
   $R_M(t_j) = (1 + m/n)(pC_1 + (1 - p)C_2)$ 
  if  $\sum_{i=1}^j R_M(t_i) > U(t_j)$  then
     $R_M(t_j) = U(t_j) - \sum_{i=1}^{j-1} R_M(t_i)$ 
  else if  $\sum_{i=1}^j R_M(t_i) < L(t_j)$  then
     $R_M(t_j) = L(t_j) - \sum_{i=1}^{j-1} R_M(t_i)$ 
  end if
end for

```

Numerical evaluations of the different strategies were performed with the following parameters: channel capacities $C_1 = 10$ bits and $C_2 = 1$ bits, $n = 100000$ time samples, delay $m = 100$ time samples and uniform weight $\lambda(t_i) = 1$, $i = 1, \dots, n$. The distortion-rate function $\varphi(R) = 4^{-R}$ was used and we tested three different values of p : 0.8, 0.5 and 0.2. For each value of p we ran 5000 numerical evaluations. The averaged distortion of the different strategies from the numerical evaluations are found in Table I.

TABLE I
NUMERICAL RESULTS

| $p = 0.8$ | |
|--------------------------------------|------------|
| Strategy | Distortion |
| Optimal offline strategy R_λ | 1.15 |
| Mean strategy R_M | 16.82 |
| Greedy strategy R_U | 5000.25 |
| $p = 0.5$ | |
| Strategy | Distortion |
| Optimal offline strategy R_λ | 48.6 |
| Mean strategy R_M | 157.2 |
| Greedy strategy R_U | 12500.9 |
| $p = 0.2$ | |
| Strategy | Distortion |
| Optimal offline strategy R_λ | 2062.2 |
| Mean strategy R_M | 2343.2 |
| Greedy strategy R_U | 20000.1 |

Clearly, the greedy strategy R_U has by far the worst performance in all three cases. This is not surprising as the greedy strategy codes the source material with $R_U(t_1) = \sum_{i=1}^{m+1} C(t_i)$ bits at t_1 and with $R_U(t_i) = C(t_{i+m})$ bits for all other time instances t_i , $i = 2, \dots, n$, during the coding period. This typically results in a considerably uneven distribution of channel capacity used for $R_U(t_1)$ and the remaining source samples $R_U(t_i)$, $i = 2, \dots, n$. The performance of the mean strategy R_M , which is not of much more complexity to implement than the greedy strategy, gets close in performance to the optimal offline strategy R_λ when p decreases. The quotient between the distortion of R_λ and R_M is 1.13 when $p = 0.2$ but increases to more than 14 when $p = 0.8$.

Let us explain the reason to consider the mean strategy and its connection with the optimal offline strategy. Note that for each admissible strategy $R \in \Omega_{L,U}$ we have $\sum_{i=1}^n R(t_i) = \sum_{j=1}^{n+m} C(t_j)$ as we otherwise would not use all channel capacity during the sending period. So, the average coding rate is $\sum_{j=1}^{n+m} C(t_j)/n$ for all admissible strategies. Note further that the strategy R_λ can be viewed as the minimal deviation

from the average strategy as it minimizes all central moments of order $r \geq 1$ in $\Omega_{L,U}$:

$$\begin{aligned} & \sum_{i=1}^n |R_\lambda(t_i) - \sum_{j=1}^{n+m} C(t_j)/n|^r = \\ & \min_{R \in \Omega} \sum_{i=1}^n |R(t_i) - \sum_{j=1}^{n+m} C(t_j)/n|^r. \end{aligned} \quad (10)$$

This follows from Theorem 1 by considering the specific convex functions

$$\varphi_r(x) = |x - \sum_{j=1}^{n+m} C(t_j)/n|^r, r \geq 1.$$

In the online setting, the property (10) then guided us to consider R_M which at each time instance tries to stay as close as possible to the expected coding capacity given by (9).

IX. CONCLUSION

We have considered live broadcasting and streaming of stored media through networks of time-varying channels connected in parallel or series. Within the constraints of network capacity and prescribed transmission delay, we investigate how to encode the source optimally with respect to average distortion. Using the framework of taut strings, it has been shown that there exists a coding rate strategy which minimizes the average distortion with respect to all convex distortion-rate functions. The geometrical interpretation in terms of a taut string provides the basis for an algorithm for computing the optimal coding rate strategy.

The optimal coding rate strategy guided us in the construction of a simple but reasonably efficient strategy addressing the online situation when the channel randomly alternates between a good and a bad state. It also provides a benchmark for evaluation of other coding rate strategies.

A natural extension of the work in this paper is to consider traffic of multiple pairs of senders and receivers within the network. We plan to address such an extension in future work.

APPENDIX: PROOF OF LEMMA 1

Proof. We now prove that the taut string algorithm actually constructs the taut string $H_* \in \Gamma_{F,G}$.

Assume first that

$$\max_{i \in \{1, \dots, n-1\}} \left(\min_{x \in [F(x_i), G(x_i)]} |x - V(x_i)| \right) = 0.$$

Then the algorithm assigns

$$H_*(x_i) = V(x_i) := \frac{x_m - x_i}{x_m - x_k} H_*(x_k) + \frac{x_i - x_k}{x_m - x_k} H_*(x_m)$$

for $i = k+1, \dots, m-1$. By construction, $H_* \in \Gamma_{F,G}$. Further, the graph of H_* corresponds to the straight line between $(x_k, F(x_k))$ and $(x_m, F(x_m))$ which clearly minimizes the graph length between these points. Hence, the constructed H_* is the taut string.

We now consider the remaining case, i.e. when there is an index $j \in \{k+1, \dots, m-1\}$ such that

$$\begin{aligned} & \min_{x \in [F(x_j), G(x_j)]} |x - V(x_j)| = \\ & = \max_{i \in \{k+1, \dots, m-1\}} \left(\min_{x \in [F(x_i), G(x_i)]} |x - V(x_i)| \right) > 0. \end{aligned} \quad (11)$$

Assume without loss of generality that $F(x_j) > V(x_j)$. From (11) follows

$$V(x_i) + F(x_j) - V(x_j) \geq F(x_i)$$

for $i = k+1, \dots, m-1$. As before, let H_* denote the taut string in Γ . We introduce the trimmed taut string \widehat{H}_* where $\widehat{H}_*(x_k) = H_*(x_k)$, $\widehat{H}_*(x_m) = H_*(x_m)$ and

$$\widehat{H}_*(x_i) = \min \{H_*(x_i), V(x_i) + F(x_j) - V(x_j)\}$$

for $i = k+1, \dots, m-1$. By construction, $\widehat{H}_*(x_i) \in [F(x_i), G(x_i)]$ for $j = k, \dots, m$. So, $\widehat{H}_* \in \Gamma_{F,G}$.

Consider the specific convex function $\varphi(x) = \sqrt{1+x^2}$, i.e. the graph length function. The taut string H_* and the trimmed taut string \widehat{H}_* differ on intervals $[x_l, x_{l+n}]$, $n \geq 2$, where

$$H_*(x_l) \leq V(x_l) + F(x_j) - V(x_j),$$

$$H_*(x_{l+n}) \leq V(x_{l+n}) + F(x_j) - V(x_j)$$

and

$$H_*(x_i) > V(x_i) + F(x_j) - V(x_j)$$

for $i = l+1, \dots, l+n-1$. It is clear that \widehat{H}_* has shorter graph length than H_* on these intervals. For remaining segments, \widehat{H}_* and H_* coincide by construction. Since the taut string H_* uniquely minimizes graph length in $\Gamma_{F,G}$ we must have $\widehat{H}_* = H_*$ and in particular

$$H_*(x_j) = \widehat{H}_*(x_j) \leq V(x_j) + F(x_j) - V(x_j) = F(x_j).$$

As $H_*(x_j) \geq F(x_j)$ this gives $H_*(x_j) = F(x_j)$.

If there are remaining nodes x_i , $i = k+1, \dots, m-1$, where the taut string H_* has not been fixed, the procedure of the algorithm is repeated for the subintervals $[x_k, \dots, x_j]$ and $[x_j, \dots, x_m]$. By continuing in this recursive manner, the taut string will be constructed for the entire interval $[x_k, x_m]$ after a finite number of steps. \square

ACKNOWLEDGMENT

The authors thank Professor Natan Kruglyak for discussions that have been instrumental for the results of this paper. Eric Setterqvist acknowledges a co-operation research grant from the Research School in Interdisciplinary Mathematics at Linköping University. Robert Forchheimer acknowledges the support by the Swedish Research Council framework grant No. 2014-6230.

REFERENCES

- [1] M. K. Müller, M. Tarantetz and M. Rupp, "Providing current and future cellular services to high speed trains", *IEEE Commun. Mag.*, vol. 53, no. 10, pp. 96-101, Oct. 2015.
- [2] É. Masson, M. Berbineau and S. Lefebvre, "Broadband Internet access on board high speed trains, A technological survey", in *Communication Technologies for Vehicles. Nets4Cars/Nets4Trains/Nets4Aircraft 2015, Sousse, Tunisia, May 6-8, 2015*, M. Kassab, M. Berbineau, A. Vinel, M. Jonsson, F. Garcia, J. Soler, Eds. Cham: Springer, 2015. pp. 165-176.
- [3] J. Garcia, S. Alfredsson and A. Brunstrom, "Examining cellular access systems on trains: Measurements and change detection", in *2017 Network Traffic Measurement and Analysis Conference (TMA), Dublin, Ireland, June 21-23, 2017*. IEEE, <https://doi.org/10.23919/tma.2017.8002916>
- [4] A. Muhammad, P. Johansson and R. Forchheimer, "Performance evaluation of rate-control strategies for real-time systems over a rate-varying channel", *Journal of Communications*, vol. 5, no. 5, pp. 390-399, May 2010.
- [5] J. R. Corbera and S. Lei, "Rate control in DCT video coding for low delay communications", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 172-185, Feb. 1999.
- [6] C.-Y. Hsu, A. Ortega and M. Khansari, "Rate control for robust video transmission over burst-error wireless channels", *IEEE J. Sel. Areas Commun.*, vol. 17, no. 5, pp. 756-773, May 1999.
- [7] H. Song and C.-C. J. Kuo, "Rate control for low-bit-rate video via variable-encoding frame rates", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 4, pp. 512-521, Apr. 2001.
- [8] H. Wang and A. Ortega, "Rate-distortion optimized scheduling for redundant video representations", *IEEE Trans. Image Process.*, vol. 18, no. 2, pp. 225-240, Feb. 2009.
- [9] L. Jiménez-Rodríguez, F. Auli-Llinàs and M. W. Marcellin, "FAST rate allocation for JPEG2000 video transmission over time-varying channels", *IEEE Trans. Multimedia*, vol. 15, no. 1, pp. 15-26, Jan. 2013.
- [10] Q. Chen and D. Wu, "Delay-rate-distortion model for real-time video communication", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 8, pp. 1376-1394, Aug. 2015.
- [11] Y. Shen, K. Mehrotra, P. C. Cosman, L. B. Milstein and X. Wang, "Delay allocation between source buffering and interleaving for wireless video", *EURASIP J. Wirel. Commun. Netw.*, vol. 2016, no. 1, Dec. 2016.
- [12] M. Kazemi, R. Iqbal and S. Shirmohammadi, "Redundancy allocation based on the weighted mismatch-rate slope for multiple description video coding", *IEEE Trans. Multimedia*, vol. 19, no. 1, pp. 54-66, Jan. 2017.
- [13] J. D. Salehi, Z.-L. Zhang, J. Kurose and D. Towsley, "Supporting stored video: reducing rate variability and end-to-end resource requirements through optimal smoothing", *IEEE/ACM Trans. Netw.*, vol. 6, no. 4, pp. 397-410, Aug. 1998.
- [14] J. Rexford and D. Towsley, "Smoothing variable-bit-rate video in an internet network", *IEEE/ACM Trans. Netw.*, vol. 7, no. 2, pp. 202-215, April 1999.
- [15] M. A. Zafer and E. Modiano, "A calculus approach to energy-efficient transmission with quality-of-service constraints", *IEEE/ACM Trans. Netw.*, vol. 17, no. 3, pp. 898-911, June 2009.
- [16] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: optimal policies", *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1732-1743, Sep. 2011.
- [17] J. Yang and S. Ulukus, "Optimal packet scheduling in an energy harvesting system", *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 220-230, Jan. 2012.
- [18] A. Kesselman, Z. Lotker, Y. Mansour, B. Patt-Shamir, B. Schieber and M. Sviridenko, "Buffer overflow management in QoS switches", *SIAM J. Comput.*, vol. 33, no. 3, pp. 563-583, 2004.
- [19] M. H. Goldwasser, "A survey of buffer management policies for packet switches", *ACM SIGACT News*, vol. 41, no. 1, pp. 100-128, March 2010.
- [20] E. Mammen and S. van de Geer, "Locally adaptive regression splines", *Ann. Statist.*, vol. 25, no. 1, pp. 387-413, 1997.
- [21] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier and F. Lenzen, *Variational Methods in Imaging* (Series Applied Mathematical Sciences 167). New York: Springer, 2009.
- [22] M. Lifshits and E. Setterqvist, "Energy of taut strings accompanying Wiener process", *Stoch. Proc. Appl.*, vol. 125, no. 2, pp. 401-427, 2015.
- [23] J.-Y. Le Boudec and P. Thiran, *Network Calculus: A Theory of Deterministic Queuing Systems for the Internet* (Lecture Notes in Computer Science, vol. 2050). Berlin: Springer-Verlag, 2001.
- [24] N. Kruglyak and E. Setterqvist, "Invariant K -minimal sets in the discrete and continuous settings", *J. Fourier Anal. Appl.*, vol. 23, no. 3, pp. 572-611, 2017.