Abstract—In this paper, we consider the problem of remotely controlling an unstable first-order noiseless linear plant over a discrete symmetric memoryless channel using anytime source-channel coding. We propose a transmission scheme, exploiting unequal error protection (UEP) based on Luby transform (LT) codes under sequential belief propagation (BP) decoding. Sufficient conditions on mean square stability of the linearly controlled closed-loop system are derived for the anytime UEP-LT codes.

I. INTRODUCTION

Networked control systems (NCS) are distributed systems comprised of sensors, actuators (plants) and controllers communicating over a shared medium. Such systems consider interactions between stochastic control theory and information theory (see e.g. [1]–[4] and references therein). From information theory, we know that transmission reliability over noisy rate-limited communication networks can be ensured by coding an arbitrary long sequence without taking delay constraints into account [5]. On the other hand, in many control applications the system performance is typically very sensitive to delay in action. Therefore, it is particularly interesting to apply transmission schemes for which the reliability is improved with time. The problem of remotely controlling a dynamic source over a noisy channel is an emerging application of NCS that has attracted much attention in recent years. Such problems have been addressed over various communication channels, see for example [6]–[10]. Most of the works in this area consider conventional joint source-channel coding (JSCE), and do not involve channel coding due to encoding/decoding delay.

In contrast to the JSCE approach, Sahai in [11] proposed an information theoretical design methodology regarding channel coding for delay constrained systems where prototypes established by Shannon do not suffice. Indeed, in [11], the authors have shown the equivalence between the stabilization of a linear plant over a noisy closed-loop feedback channel and sequential communication over a noisy channel with a noiseless feedback. In [12], anytime error exponents have been determined based on a time-sharing anytime channel code over a discrete memoryless channel (DMC) with perfect channel feedback. In [13], distortion convergence degree and rate have been derived for certain anytime coding schemes over a binary erasure channel (BEC) and assuming no channel feedback. The authors in [14], have shown asymptotic sufficient conditions on the error exponent required for stabilizing vector-valued processes over BEC using anytime tree codes. Sufficient stability conditions using anytime codes have been derived over compound channels in [15].

In this paper, we propose to exploit anytime rateless codes for the purpose of communication and control. The rateless code is a class of low-complexity random sparse channel codes in which the encoder generates an unlimited number of symbols such that the decoder can recover the source symbols from a sufficiently large subset of channel outputs. In real-time applications, it is also important to allocate coding redundancy over streaming information bits according to their importance level, which is called unequal error protection (UEP). This property is also critical for applications such as controlling unstable plants using the anytime design since not only delay is of great importance, but also some parts of information bits need to be treated differently. In the literature, rateless codes have been shown to have advantages exploiting UEP principles, see for instance [16], [17].

The main contribution of this paper is the design and analysis of anytime UEP-rateless code under a sequential belief propagation (BP) decoder over a binary symmetric channel (BSC) in a closed-loop feedback system. To the best of our knowledge, anytime rateless code with the objective of stabilization in feedback control has not been considered before. We will employ Luby transform (LT) [18] code as a realization of our anytime UEP-rateless code. The motivation behind exploiting the UEP-LT coding scheme is the low complexity of the encoder and the decoder. For the purpose of analysis, an upper-bound on the end-to-end distortion is derived in order to determine the degree and the rate of convergence to zero as time increases. These results immediately give sufficient conditions for stability of a linear first-order plant over a BSC in mean square sense [2], [6], [9], [11], [19]. For performance evaluation, we compare our codes with anytime UEP-repetition codes. Furthermore, the sufficient conditions for stability will be compared with the result in [11] using the Gallager random coding.

The rest of the paper is organized as follows. In Section II,
As will be clear later, the binary expansion fits well into $X \subseteq \{0, 1\}^a$ and is drawn according to a known distribution at time $t \in \mathbb{N}$. The source samples are conveyed over a BSC with bit cross-over probability $\epsilon$. The inputs and outputs of the channel are denoted respectively by $y(t) \in \{0, 1\}$ and $z(t) \in \{0, 1\}$. The conditional probability of the channel is time-invariant, i.e., $\Pr(y(t)|z(t)) = \Pr(y|z)$.

In our anytime transmission scheme, the two main functional units, source coding and channel coding, are considered separately. At the transmitter side, we, throughout the paper, employ truncated binary expansion as the source encoder to map $x(t)$ into a sequence of bits $\{b_1, \ldots, b_{j(t)}\}$, where $j(t)$ denotes the number of information bits revealed at time $t$; furthermore, $j(t) \geq j(t-1)$. In general, $j(t)$ depends on source rate, i.e., if the source generates a bit every $1/R$ time units, then $j(t) = \lceil tR \rceil$. As will be clear later, the binary expansion fits well into the anytime framework that provides anytime reliability. The binary expansion is defined by a map $\mathcal{E}^t_e : \mathcal{X} \mapsto \mathbb{B}^{j(t)}$, where $\mathbb{B}^{j(t)}$ takes values from $\{0, 1\}^{j(t)}$. The anytime channel encoder is described by a map $\mathcal{E}^t_c : \mathbb{B}^{j(t)} \mapsto \mathcal{Y}$ which outputs a bit at each time according to the function $E_c(b_{1}^{j(t)}) = y(t)$.

At the remote receiver, the channel decoder, specified by a map $\mathcal{D}^t_c : \mathcal{Z} \mapsto \mathbb{B}^{j(t)}$, is allowed to exploit the information from the current received symbol as well as the previous received symbols for the purpose of estimation, i.e., $\hat{x}_{1}^{j(t)} = \mathcal{D}^t_c(x_{1}^{j(t)})$. Finally, the source decoder can be written as a map $\mathcal{D}^t_s : \mathbb{B}^{j(t)} \mapsto \mathbb{X}$. The source decoder outputs the reconstructed value $\hat{x}_t$ at each time instant $t$. The control unit at any time instant $t$ takes an action based on the current and previous estimates $\hat{x}_{1}^{t}$ which is fed back to the linear plant.

The block diagram of our dynamic closed-loop system using an anytime code is sketched in Figure 1 in which we consider a scalar control-driven discrete-time LTI system as

$$x(t+1) = ax(t) + u(t)$$

where $\{x(t), u(t)\} \subseteq \mathbb{R}$ are state and control processes. The open-loop system is unstable, i.e., $a > 1$, and $x(0)$ is uniformly distributed on $[0, 1]$. We are interested in mean square stabilization which is defined as follows.

**Definition 1**: A system is said to be mean square stable if

$$\lim_{t \to \infty} \mathbb{E}[x^2(t)] = 0,$$

regardless of the initial state $x(0)$.

In order to stabilize the system, the controller takes an action using the estimate of $\hat{x}(0)$ at time instant $t$, denoted by $\hat{x}(0|t)$. More specifically, at $t = 0$, the control action is $u(0) = -a\hat{x}(0|0)$. At $t > 1$, the control signal, which is linear optimal, is given by $u(t) = -a^{t+1}(\hat{x}(0|t) - \hat{x}(0|t-1))$. Thus, the system equation can be formulated as

$$x(t+1) = a^{t+1}(\hat{x}(0|t) - x(0)).$$

Then, the second moment of the state at any time $t$ is given by

$$\mathbb{E}[x^2(t+1)] = a^{2t+1} \mathbb{E}[|x(0) - \hat{x}(0|t)|^2] \triangleq a^{2t+1} \Delta^2(t),$$

where $\Delta(t)$ denotes the end-to-end distortion.

Therefore, if $\lim_{t \to \infty} a^{2t} \Delta^2(t) = 0$ then the given system is mean square stabilized. The goal is to design a channel encoder/decoder pair in order to estimate $x(0)$ precisely out of the received data up to time $t$ so that the $a^{2t}$ in (2) is dominated by $\Delta^2(t)$ as time increases. It should be noted that the end-to-end distortion is due to the source distortion which is caused by the lossy quantizer encoder, and the channel distortion which results from the channel uncertainty. Let us define the source and channel distortions, respectively, as

$$\Delta_s(t) \triangleq \mathbb{E}[|x(t) - \hat{x}(0|t)|^2],$$

$$\Delta_c(t) \triangleq \mathbb{E}[|x(t) - \hat{x}(0|t)|^2],$$

where $x(0|$ is the quantized value of $x(x)$ at time instant $t$. The end-to-end, source and channel distortion measures
are related as
\[ \Delta(t) = \mathbb{E}^{1/2} \| x(0) - x_s(0) | t \| + x_s(0) | t \| - \tilde{x}(0) | t \| \]  
\[ \leq (a) \Delta_s(t) + \Delta_c(t), \]
where \((a)\) follows from the Minkowski inequality. For simplicity of notation, we will use \( x, x_s(t) \) and \( \tilde{x}(t) \) for \( x(0), x_s(0) \) and \( \tilde{x}(0) \), respectively. Therefore, in order to determine an upper-bound on the end-to-end distortion, the source and channel distortions are required as given by the following two lemmas.

**Lemma 1:** The source distortion subjected to the binary expansion at time \( t \) is given by
\[ \Delta_s(t) = \frac{1}{\sqrt{3}} 2^{-K}, \]
where \( K \) is the number of information bits arrived at the anytime channel encoder.

**Proof:** The proof is given in Appendix A.

**Lemma 2:** The channel distortion is related to the bit error probability as
\[ \Delta_c(t) = \sum_{k=1}^{K} 2^{-2k} P_e_k(t), \]
where \( P_e_k(t) \) is the bit error probability of the \( k \)th information bit at time instant \( t \).

**Proof:** The proof is given in Appendix B.

The end-to-end distortion can be also characterized by the rate and degree of convergence to zero which quantify the goodness of coding schemes. Particularly, if for some polynomial \( C(t) \), the end-to-end distortion decays at least as fast as
\[ \Delta(t) \leq C(t)2^{-\zeta t}, \]
then \( 0 < \zeta \leq 1 \), and \( \zeta > 0 \) specify the degree and the rate of convergence to zero, respectively. The constant \( \zeta \) is also known as the distortion exponent and can be determined as
\[ \zeta = \lim_{t \to \infty} \left[ \frac{-1}{t} \log \Delta(t) \right]. \]

To be able to stabilize the closed-loop feedback system according to Definition 1, it is desired that \( \Delta(t) \) decays exponentially, i.e., \( \tau = 1 \). In the next sections, we assess the performance of two anytime coding schemes based on UEP-repetition [13], [20] and UEP-LT [21], [22] codes with respect to the degree and rate of convergence over the BSC.

**III. ANYTIME UEP-REPLICATION CODING**

As a part of the channel code, an anytime repetition strategy is used, which is a special block repetition code. More specifically, how frequently an information bit is coded is determined by its importance level. For example, the first bit \( b_1 \) is repeated most often since it is the most significant bit. If the bit stream \( b_1^{(t)} \) is generated up to time \( t \), the channel encoder transmits a bit \( y(t) \) according to the sequence \( \{ b_1; b_1; b_2; b_1; b_2; b_3; \ldots \} \) over the channel, where semicolons indicate time instants at which the repetition begins, and commas denote channel uses. Let us assume that at time instant \( t \), \( K \) bits have arrived at the anytime channel encoder, where \( K \leq j(t) \). Then, the decoder receives a bit \( z(t) \), the noisy version of \( y(t) \), and employs majority logic decoding (MLD) so as to decide \( b_i(t) = 0 \) (or, 1) \( \forall 1 \leq i \leq K \) if the number of received zeros (or, ones) is more than the number of ones (or, zeros) in the sequence \( x(t) \). The decoder should also be able to decode if the number of received zeros or ones is equal which is performed by deciding \( b_i(t) = 1 \) or \( b_i(t) = 0 \) using a Bernoulli trial. The following theorem provides the degree and rate of convergence of the end-to-end distortion using this scheme.

**Theorem 1:** Given the BSC with bit cross-over probability \( \epsilon \), and using the anytime UEP-repetition code under MLD, there exists a polynomial \( C(t) \) such that the end-to-end distortion \( \Delta(t) \) of the source-channel coding system of Figure 1 is upper-bounded as
\[ \Delta(t) \leq C(t)2^{-\zeta t}, \]
where
\[ \zeta = \left\{ \begin{array}{ll} \sqrt{2} \frac{1}{\sqrt{2}} (1 + \log \sqrt{\epsilon(1-\epsilon)}) & , \epsilon < \epsilon' \\ \frac{1}{\sqrt{2}} (1 + \log \sqrt{\epsilon(1-\epsilon)}) & , \epsilon \geq \epsilon' \end{array} \right. \]

Furthermore, \( \epsilon \geq \epsilon' \) is the region where \( 1 + \frac{1}{2} \log \left( \frac{4(1-\epsilon)}{(1-\epsilon')^2} \right) + \frac{1}{\eta} \log \frac{1}{\eta} \geq 0 \), and \( \eta \) is a continuous variable in the interval \( \frac{1}{\epsilon} \leq \eta \leq 1 \).

**Proof:** The proof is given in Appendix C.

The anytime repetition coding strategy, despite its low encoding/decoding computational complexity, cannot provide exponential convergence in time. In the next section, we develop an anytime transmission scheme exploiting anytime rateless codes achieve exponential convergence.

**IV. ANYTIME UEP-LT CHANNEL CODING**

In this section, we first describe the encoding and decoding procedure for anytime UEP-LT channel codes. Then, theoretical analysis of the proposed scheme will be carried out.

Basically, the encoding scheme we employ in this paper is an adaptation of the expanding window fountain (EWF) codes [17] fitted into the anytime transmission scheme. Let the information bits (input bits) and the encoded bits (output bits) be allocated to some overlapping information windows and disjoint encoding windows, respectively. Particularly, at time instant \( t \), the current information window, denoted by index \( j \), contains a portion of streaming information bits with the fixed length \( K_j \leq \lfloor R | t \rfloor \). Associated with the \( j \)th information window, there exists a degree distribution \( \Omega_j(x) \) according to which the bits in the corresponding window are uniformly chosen to be XOR'ed and generate an encoded bit \( y(t) \) that is located in the \( j \)th encoding window with the total length \( T_j \). It should be noted that this scheme provides \( j \) levels of importance in which each contains \( K_m = K_{m-1} - 1 \leq m \leq j \) bits. The encoding procedure is shown in Figure 2.

Now, let \( \Omega_j(x) = \sum_{d=j}^{\infty} \Omega_d x^d \) denote the output bit degree distribution corresponding to the \( j \)th encoding window,
where $\Omega_d$ represents the probability of choosing $d_j$ number of bits from the $j^{th}$ information window, and $D_j$ shows the maximum order of the degree distribution. We define $\beta_j = \Omega_d(1)$ as the average degree of an output bit in the $j^{th}$ encoding window. Similar node degree distributions can be specified for the streaming input bits with the difference that they depend on time as well as node degree. For this purpose, we first assume that the anytime algorithm occurs up to the $j$ level, that is, $\Lambda_j$ represents the probability that a bit in the $m^{th}$ level is of degree $n_m$ at time $t$. The degree distribution $\Lambda_m(t, x)$ can be obtained from the degree distributions of the input bits from the $m^{th}$ information window up to the $j^{th}$ information window induced only by the edges connected to their corresponding encoding windows which is denoted by $I_m(t, x) = \sum_{n_m} \Lambda_m(t)x^{n_m}$ denote the input bit degree distribution of the $m^{th}$ level $(1 \leq m \leq j)$ level of importance, where $\Lambda_m(t)$ represents the probability that a bit in the $m^{th}$ level is of degree $n_m$ at time $t$. The degree distribution $\Lambda_m(t, x)$ can be obtained from the degree distributions of the input bits from the $m^{th}$ information window up to the $j^{th}$ information window induced only by the edges connected to the $m^{th}$ encoding window. It can be verified that $I_m(t)$ is drawn according to the binomial distribution (refer to [23], [24] for details), i.e.,

$$I_m(t) \sim \left\{ \begin{array}{ll}
\binom{T_i}{n_i} p_i^{n_i}(1 - p_i)^{T_i - n_i} & m \leq i \leq j - 1 \\
\binom{t_i}{n_j} p_j^{n_j}(1 - p_j)^{i - n_j} & i = j
\end{array} \right. \quad \text{(11)}$$

where $p_i = \frac{\alpha_i}{\alpha_i + \beta_i}$ and $1 \leq n_i \leq T_i$ for $m \leq i \leq j - 1$. Moreover, $p_j = \frac{\alpha_j}{\alpha_j + \beta_j}$, and $1 \leq n_j \leq t_i$, where $i = t - \sum_{k=1}^{j-1} T_k$. The probability generating function of $I_m(t)$ is given by

$$I_i(t, x) = \left\{ \begin{array}{ll}
[(1 - p_i) + p_i x]^{T_i} & m \leq i \leq j - 1 \\
[(1 - p_j) + p_j x]^i & i = j
\end{array} \right. \quad \text{(12)}$$

Factoring $I_i(t, x)$ in the form of $(1 - p_i)^{T_i} \left(1 + \frac{p_i x}{1 - p_i}\right)^i$, and applying the inequality $(1 \pm x)^k \leq e^{\pm kx}$, $\forall x \in [0, 1]$, and $k \in \mathbb{R}$, to the both terms, we obtain

$$I_i(t, x) \leq e^{\frac{\alpha_i(T_i - m - 1)}{\sigma_i - 1}} \quad m \leq i \leq j - 1$$

$$I_j(t, x) \leq e^{\frac{\alpha_j(T_j - m - 1)}{\sigma_j - 1}} \quad i = j \quad \text{(13)}$$

In order to proceed to find $\Lambda_m(t, x)$, it is well-known that the generating function of a sum of independent (not necessarily identical) random variables is equal to product of their generating functions. Hence

$$\Lambda_m(t, x) = \prod_{i=m}^{j} I_i(t, x)$$

$$\leq e^{\sum_{i=m}^{j} \alpha_i(T_i - m - 1)} e^{\sum_{i=m}^{j-1} \alpha_i(T_j - m - 1)} \quad \text{(14)}$$

Regarding the decoding procedure, BP is recognized as one of the most efficient algorithms in decoding graphical codes approaching Shannon capacity. It is an iterative procedure such that at each iteration input and output bits exchange messages, containing log-likelihood ratio (LLR). Let $B^{(l)} \in \mathbb{R}, Y^{(l)} \in \mathbb{R}$ denote the messages passed from an input (output) bit $b(y)$ to an input (output) bit $b = 1$ or 0 at the $l^{th}$ iteration of the BP. The updating rules of the messages at each iteration are as follows

$$\text{tanh}(Y^{(l)}/2) = \text{tanh}(q(t)/2) \prod_{adj(y)} \text{tanh}(B^{(l)}/2)$$

$$B^{(l+1)} = \sum_{adj(b)} Y^{(l)}$$

where $q(t) \triangleq \ln[\Pr(g(t) = 0 | z(t))/\Pr(g(t) = 1 | z(t))]$ denotes channel LLR at time $t$. In the case of BSC, $q(t) = \ln \left( \frac{1}{2} \right) (1 - z(t))$. Moreover, $adj(b)$ ($adj(y)$) represents all the output (input) bits adjacent to the input (output) bit $b$ ($y$) excluding the edge from $y$ to $b$. Now, the messages are available at the decoder, therefore, the previous edges connecting the input to the output bits are known at time $t$. It is worth pointing out that the messages $Y^{(l)}$ and $B^{(l)}$ are symmetric random variables as shown in [25], i.e., if $f$ is a probability density function of a messages $X$, then $f$ satisfies $f(-x) = e^{-x}f(x)$. Furthermore, to simplify the analysis of the codes using the BP, one can approximate the probability density of messages passed at each iteration by simple functions instead of tracking the true densities. Commonly in practice, e.g., [24], [25], $B^{(l)}$ and $Y^{(l)}$ in (15) are assumed symmetric Gaussian random variables which is a realistic assumption according to central limit theorem. Thus, the variance of variables is twice the mean, by which we only need to determine the mean. We will use these facts to analyze the UEP-LT codes.

The random behavior of the LT codes is characterized by its degree distribution, and, indeed, an unoptimized design
of the degree distribution may lead to a poor performance of the process. Interested readers are referred to [22] for the details of degree distribution design using anytime UEP-LT codes. Herein, we briefly outline the design which is focused on minimizing the number of encoding symbols under error probability constrains. More specifically, the design aims at minimizing delay under the constraint that the mean of the input-to-output bit messages increases at each iteration. As a result, the mean of the output-to-input bit messages increases at each iteration, and converges to the channel LLR, i.e., $q(t)$, which is known as the stability condition [24].

The following theorem gives the degree and rate of convergence for the proposed anytime UEP-LT codes.

**Theorem 2:** Given the BSC with bit cross-over probability $\epsilon$, and using the anytime UEP-LT codes when the number of BP iterations tends to infinity, there exists a polynomial function $C(t)$ for which the end-to-end distortion ($\Delta(t)$) of the source-channel coding system in Figure 1 is upper-bounded as

$$\Delta(t) \leq C(t)2^{-\xi t},$$

where

$$\xi = \frac{\beta_j}{2K_j \ln 2} \left(\frac{1 - \epsilon}{4\epsilon(1 - \beta_j/K_j)} + 1\right).$$

Furthermore, $K_j$, $\alpha_j$, and $\beta_j$ are respectively the size of $j^{th}$ information window, average degree of an input bit in the $j^{th}$ information window and average degree of an output bit in the $j^{th}$ encoding window.

**Proof:** The proof is given in Appendix D.

**Remark 2:** Since the anytime UEP-LT encoder is based on XOR’ing a portion of $[Rt]$ number of bits, the number of encoding operations grows at most like $O(t)$. The decoding complexity based on optimized degree distribution and the sequential BP grows linearly in time at each iteration, i.e., of the order $O(t)$. The decoding complexity of other relevant coding schemes, e.g., anytime binary random convolution codes, grows exponentially in time using maximum likelihood sequential decoding such as the Viterbi algorithm. Other sequential decoding of anytime convolutional codes, such as Fano algorithm, although having the decoding complexity grows like $O(t)$ in time, the scheme needs to operate at rates below computational cut-off rate for successful decoding.

Following Definition 1 and (2), sufficient stability conditions using the anytime schemes are derived in Corollary 1.

**Corollary 1:** According to Theorem 1, the anytime UEP-repetition scheme over the BSC cannot stabilize the system (1) if the open-loop constant $a \geq 1$ since the degree of convergence cannot be beaten by $1/2$. On the other hand, based on Theorem 2, the sufficient condition for stabilization of a scalar linear plant using the anytime UEP-LT codes over the BSC is given by

$$\log(a) \leq \frac{\beta_j}{2K_j \ln 2} \left(\frac{1 - \epsilon}{4\epsilon(1 - \beta_j/K_j)} + 1\right).$$

The sufficiency results in [11, theorem 5.2] show that the LTI system (1) over a general DMC is mean square stable if $E_r(\log a) > 2\log(a)$, where $E_r(\cdot)$ is the Gallager random coding error exponent calculated in base two. Using the above result and the Gallager error exponent for the BSC (see [26, page 146] for details), the closed-loop feedback system (1) is stable if

$$\log(a) < \frac{1}{3} \left[1 - 2\log(\sqrt{\epsilon} + \sqrt{1 - \epsilon})\right].$$

In the next section, we will compare the two sufficient conditions for the stability of the closed-loop feedback system.

**V. NUMERICAL RESULTS**

In this section, we quantify the performance of the anytime transmission schemes in terms of end-to-end distortion, degree and rate of convergence.

We have shown the performance of the anytime UEP-LT codes with five and ten levels of importance, which are denoted respectively by 5-UEP-LT and 10-UEP-LT. The lengths of information and encoding windows for 5-UEP-LT set to $K = [12, 30, 50, 75, 120]$ and $T = [50, 150, 250, 300, 350]$, respectively, and $K = [10, 12, 20, 30, 40, 50, 60, 75, 95, 120]$ and $T = [50, 55, 60, 70, 80, 100, 120, 140, 160, 180]$ for 10-UEP-LT codes, correspondingly. Furthermore, the degree distributions are optimized for the BSC with bit cross-over probability $\epsilon = 0.11$ (for which the channel capacity is approximately 0.5) using the approach in [21] which yields the average degree of output node $\beta = [0.6, 1.26, 1.53, 2.09, 2.84]$ and $\beta = [0.68, 0.78, 1.12, 1.47, 1.62, 1.79, 2.00, 2.24, 2.82, 3.56]$ for the 5-UEP-LT and 10-UEP-LT codes, respectively.

First, in Figure 3, we depict a performance comparison of the anytime UEP-repetition and UEP-LT codes using 5000 Monte-Carlo simulations. The number of BP iterations is also set to 15 rounds. The degree of convergence plays a crucial role which can be seen from the gap between the two coding schemes. Furthermore, the analytic distortion exponent obtained by (10) for the anytime UEP-repetition codes.
scheme is equal to $\zeta = 0.4782$, whereas that of the anytime 5-UEP-LT codes is $\zeta = 0.0661$ and the 10-UEP-LT codes is $\zeta = 0.0524$ which is given by (17). The latter can also be verified by Figure 4 where distortion exponents determined using simulation and analysis of the 5-UEP-LT and 10-UEP-LT codes are reported as well as the analytic distortion exponent using Gallager random coding derived in (19). The results of Figure 4 also show the sufficient conditions for mean square stabilization in feedback control over a BSC using the proposed anytime UEP-LT codes with the configurations explained and the Gallager random coding. Although the stability region using the proposed scheme is a subset of that of the Gallager random coding in this setup, it can be improved by increasing the number of protection levels. It should be noted that the stability bound using the Gallager random coding might not be achieved in practice since the decoding complexity grows unboundedly.

VI. CONCLUSION

We have used a modular approach in order to design two main components of a networked control system; an encoder and a decoder based on practical anytime causal transmission strategies. The components are designed such that they have low complexity, and provide transmission reliability and system stability. In our studies, we have focused on linear plants and binary symmetric communication channels in order to derive degree and rate of convergence for end-to-end distortion using UEP-repetition and UEP-LT codes. Moreover, the sufficient condition on mean square stability of a closed-loop feedback system has been derived using the proposed anytime UEP-LT codes.

The results of this paper can be easily extended to other classes of binary memoryless symmetric channels such as binary input additive white Gaussian (BIAWGN) channel. Exploiting channel feedback in order to progressively refine the performance can be a subject of future work. Furthermore, trade-offs between the number of protection levels and stability regions remain to be explored.

APPENDIX

A. proof of Lemma 1

The proof is based on the definition of the source distortion (3), i.e.,

$$\Delta_s^2(t) \triangleq \mathbb{E} \left[ (x_s(t) - \hat{x}(t))^2 \right] = \mathbb{E} \left[ \left( \sum_{k=1}^{K} b_k 2^{-k} - \sum_{k=1}^{K} b_k 2^{-l_k} \right)^2 \right]$$

$$= \mathbb{E} \left[ \sum_{k=1}^{K} \sum_{l=K+1}^{\infty} b_k b_l 2^{-(l_k + l - t)} \right] = \sum_{k=1}^{K} \sum_{l=1}^{\infty} 2^{-(l_k + l - t)} \mathbb{E}[b_k b_l],$$

(20)

where straightforward mathematical manipulations are used as well as the linear property of the expectation operator. Further, since $b_k$’s are iid distributed as Bernoulli$(1/2)$, $\mathbb{E}[b_k b_l]$ can be obtained as

$$\mathbb{E}[b_k b_l] = \frac{1}{4} + \frac{1}{4} \delta(k - l),$$

(21)

where $\delta(k - l)$ is the delta Dirac function. Plugging (21) into (20), and applying the sum of geometric series, the result is obtained.

B. proof of Lemma 2

The proof follows from the definition of the channel distortion (3), i.e.,

$$\Delta_c^2(t) \triangleq \mathbb{E} \left[ (x_c(t) - \hat{x}(t))^2 \right] = \mathbb{E} \left[ \left( \sum_{k=1}^{K} b_k 2^{-k} - \sum_{k=1}^{K} \hat{b}_k 2^{-k} \right)^2 \right]$$

$$= \sum_{k=1}^{K} 2^{-2k} \mathbb{E}[|b_k - \hat{b}_k(t)|^2]$$

$$= \sum_{k=1}^{K} 2^{-2k} \mathbb{P}(\hat{b}_k(t) = 1|b_k = 0) = \sum_{k=1}^{K} 2^{-2k} P_{e_k}(t),$$

(22)

where straightforward mathematical manipulations are used as well as the linear property of the expectation operator, and the last equality follows by defining $P_{e_k}(t) \triangleq \mathbb{P}(\hat{b}_k(t) = 1|b_k = 0)$.

C. Proof of Theorem 1

According to the repetition strategy, $j(t)$ bits are generated at time $t$ in which $K$ bits arrive at the channel encoder where $t = \frac{K(K+1)}{2}$. In order to analyze the channel distortion, the probability of error for the bits that are repeated odd number of times, $P_{e_{odd}}$, is considered individually from those which are repeated even number of times, $P_{e_{even}}$, where $n$ is the number of repetitions of the bits. Therefore, if we assume that at time $t$, $K$ is odd, then the probability of error using
The second term in $P_{e_n}^{even}$ is due to the errors in which the number of zeros is equal to the number of ones.

Using Lemma 2, the channel MSE can be expressed as

$$\Delta_i^2(t) = \sum_{k=1}^{K-1} 2^{-2k} P_{e_n}^{odd}_{k-K-1} + \sum_{k=even}^{K-1} 2^{-2k} P_{e_n}^{even}_{k-K-1},$$

(23)

Therefore, $\Delta_i(t)$ is upper-bounded as

$$\Delta_i(t) \leq \frac{1}{\sqrt{3}} 2^{-K} + 2 \sum_{k=odd}^{K} 2^{-2k} \sqrt{P_{e_n}^{odd}_{k-K-1}}$$

(25a)

$$\leq \frac{1}{\sqrt{3}} 2^{-K} + 2 \sum_{k=odd}^{K} 2^{-2k} \sqrt{(K+1)} D(K+1,K+1) \|\epsilon\|$$

(25b)

$$\leq \frac{1}{\sqrt{3}} 2^{-K} - \min \left\{ \frac{1}{\sqrt{2}} \sum_{k=odd}^{K} 2^{-2k} \sqrt{(K+1)} D(K+1,K+1) \|\epsilon\| \right\}$$

(25c)

$$\leq C_1 2^{-\sqrt{2}t/2} + C_2(t) 2^{-\sqrt{2}t/2} \min \left\{ \frac{1}{\sqrt{2}} \sum_{k=odd}^{K} 2^{-2k} \sqrt{(K+1)} D(K+1,K+1) \|\epsilon\| \right\}$$

(25d)

where inequality (25a) follows from the Minkowski inequality and induction, i.e., $P_{e_n}^{even} \leq P_{e_n}^{odd}$. The inequality (25b) follows by defining the binary relative entropy function $D(x||y) = x \log \frac{x}{y} + (1-x) \log \frac{1-x}{1-y}$ and the corresponding inequality in [26, pp. 530-1]. The equality (25c) is obtained by defining $\eta = \frac{1}{\sqrt{2}} \sum_{k=odd}^{K} 2^{-2k} \sqrt{(K+1)} D(K+1,K+1) \|\epsilon\|$, and changing the variable from $k$ to $\eta$. The inequality (25d) is acquired by defining the appropriate constant $C_1$ and polynomial $C_2(t)$ as well as the trivial inequality $\frac{\sqrt{2}}{\sqrt{2}t/2} \leq \sqrt{2}t/2$.

Hence, it follows that for a given polynomial $C(t)$, the end-to-end distortion is upper-bounded as

$$\Delta(t) \leq C(t) 2^{-\zeta t/2},$$

(26)

where the distortion exponent $\zeta$ is obtained by interpreting the behavior of $\frac{1}{\sqrt{3}} 2^{-K} - \min \left\{ \frac{1}{\sqrt{2}} \sum_{k=odd}^{K} 2^{-2k} \sqrt{(K+1)} D(K+1,K+1) \|\epsilon\| \right\}$ in (25d). More specifically, introducing the threshold transition probability $\epsilon'$, the derivative of $\frac{1}{\sqrt{3}} 2^{-K} - \min \left\{ \frac{1}{\sqrt{2}} \sum_{k=odd}^{K} 2^{-2k} \sqrt{(K+1)} D(K+1,K+1) \|\epsilon\| \right\}$ with respect to $\eta$ is positive if $\epsilon \geq \epsilon'$. Therefore, it is monotonically increasing, and the minimum value is attained at $\eta = \frac{1}{\sqrt{2}}$, which is negative; thus the channel distortion exponent becomes dominant. On the contrary, for $\epsilon < \epsilon'$, the minimum value does not depend on $K$ (equivalently $t$); hence, the channel and source distortion exponents are the same, then $\zeta = \sqrt{2}$. This completes the proof of Theorem 1.

**D. Proof of Theorem 2**

We consider the anytime transmission happens at the interval $T_j$, then by following Lemma 2, the channel distortion can be upper-bounded as

$$\Delta_{e_j}^{(i)}(t) \leq \sum_{i=1}^{j} \sum_{k=1}^{K_i} 2^{-k} \sqrt{P_{j,i}^{(i)}}(t),$$

(27)

where $P_{e_j}^{(i)}(t)$ denotes bit error probability of the input bits in the $j$th level of importance at iteration $l$ of the BP. Assuming the Gaussian distribution for the output-to-input node messages and all-zero sequence transmitted over the channel, $P_{e_j}^{(i)}(t)$ can be obtained by (28), where $Y_{j,i}^{(i)}$ is the message sent from output bits in the $j$th encoding window to the input bits of the $i$th level at iteration $l$ of the BP. Moreover, $I_{nj}$ is the probability that an input bit (in the $i$th information window) connected only to the $j$th encoding window is of degree $nj$ regardless of the other edges. Thus,

$$P_{e_j}^{(i)}(t) = \frac{1}{2} \prod_{m=1}^{j} \sum_{m,i} I_{nj} \exp \left\{ -m \epsilon \left[ Y_{j,i}^{(i)}(t) \right] / 4 \right\}$$

(29a)

$$= \frac{1}{2} \prod_{m=1}^{j} \sum_{m,i} \exp \left\{ -m \epsilon \left[ Y_{j,i}^{(i)}(t) \right] / 4 \right\}$$

(29b)

where (29a) is obtained by applying the Chernoff bound for the Q-function, i.e., $Q(x) \leq \frac{1}{2} \exp(-x^2/2)$ for $x \geq 0$, to (28). Furthermore, by following (14), (29b) can be derived, where $\epsilon_i$ and $\beta_i$ (1 $\leq i \leq j$) are defined in Section IV.

As the BP iterates, $\left[ Y_{j,i}^{(i)} \right] (i \leq m \leq j)$ converges to the channel LLR, i.e., $q = \ln \left( \frac{\epsilon}{\epsilon'} \right)$, because of the stability condition provided by the optimal design of degree distributions [21], [24].

By considering the error probability in (29b), and combining it with the channel distortion (27) and the source distortion (5), there exists polynomials $C(t)$ and $C(t)$ for which the end-to-end distortion is given by

$$\Delta(t) \leq \frac{1}{\sqrt{3}} 2^{-K_j} + C(t) 2^{-\zeta t/2} \left( \frac{1}{\sqrt{2}} \sum_{k=odd}^{K} 2^{-2k} \sqrt{(K+1)} D(K+1,K+1) \|\epsilon\| \right)$$

(30)

where the last inequality follows from the fact that $K_j$ is chosen in such a way that $\Delta(t)$ is always decaying, i.e., the source distortion is smaller than the channel distortion at each encoding interval. Thus, with the right choice of $K_j$,
the distortion decays exponentially, and the rate of convergence becomes
\[
\frac{\beta}{2K_j \ln 2} \left( \frac{1-e^{-\frac{t}{4\beta K_j N_j}}} {4(1-\beta K_j N_j)} + 1 \right).\]
This completes the proof of Theorem 2.

REFERENCES