Maximum Likelihood Laplacian Correlation Channel Estimation in Layered Wyner-Ziv Coding

Nikos Deligiannis, *Member, IEEE,* Adrian Munteanu, *Member, IEEE,* Shuang Wang, *Member, IEEE,* Samuel Cheng, *Member, IEEE,* and Peter Schelkens, *Member, IEEE*

***Abstract*—The distributed source coding theory teaches that efficient compression can be achieved by exploiting the correlation between encoded sources at the decoder only. A critical issue in this coding paradigm is how to accurately estimate the correlation statistics at the decoder. This paper proposes a novel correlation channel estimation method designed for generic layered Wyner-Ziv coding. Driven by real-life applications, i.e., video coding, the correlation noise is assumed to be memoryless zero-mean Laplacian with a block-stationary behavior. Unlike existing methods, the proposed technique derives a novel maximum likelihood estimate of the noise scaling parameter per stationarity block. Adhering to layered coding, the derived estimate is successively refined as more stages are decoded. The Crámer-Rao lower bound (CRLB) for our estimator is derived and the mean squared error performance of the proposed successive refinement algorithm is evaluated through simulations. Apart from achieving an accuracy close to the derived CRLB, the proposed algorithm is shown to yield a Wyner-Ziv rate-distortion performance approaching the ideal but unrealistic scenario where the decoder knows *a priori* the correlation statistics. Moreover, the proposed technique has been integrated into our latest Wyner-Ziv video codec for wireless capsule endoscopy. Experimentation reveals that the proposed method delivers improved compression performance compared to state-of-the-art techniques, while simultaneously reducing the overall decoding complexity.**

***Index Terms*—Layered Wyner-Ziv (WZ) coding, correlation channel estimation (CCE), maximum likelihood estimate (MLE), Crámer-Rao lower bound (CRLB), wireless capsule endoscopy.**

I. Introduction

DISTRIBUTED source coding (DSC) refers to the separate encoding and joint decoding of correlated sources. Slepian and Wolf [1] derived the theoretically achievable rate region for the distributed coding of two discrete sources in a lossless compression scenario. Wyner and Ziv [2] established the lower rate bound under a distortion constraint when a source is encoded independently but decoded using a correlated signal as side information. In successively refined Wyner-Ziv (WZ) coding [3], [4], a source is encoded in multiple stages and decoded with different distortion and possibly different side information at each stage.

Expressing the statistical dependency between the distributed sources in the form of a virtual channel, one may design Slepian-Wolf (SW) coding based on binning realized by channel codes [5]–[7]. Alternatively, distributed arithmetic codes [8], [9] can also implement SW coding. Practical WZ constructions include nested lattice or trellis codes [10] and SW coded nested [11] or non-nested [6] lattice quantization. In [4], successively refined WZ coding was constructed by nested scalar quantization followed by SW coding of bit-planes with low-density parity-check (LDPC) codes.

DSC finds applications in several domains, including wireless sensor networks [12], biometrics, and tampering localization [13]. Distributed video coding (DVC) [6], [12], [14], in particular, targets lightweight multimedia applications, such as wireless visual sensors and wireless capsule endoscopy [15]. In this context, layered WZ coding [4] provides scalability, a key trait for transmission over heterogeneous wireless networks with varying resources and bandwidth.

A common assumption in DSC schemes, e.g., [4], [10], [11], [16], is that the joint decoding of the sources is carried out with a priori knowledge of the correlation statistics. However, in real-world signals, for instance, obtained by sensors that measure temperature, humidity, or capture sound or visual data, the correlation statistics vary both spatially and temporally. The varying correlation statistics cannot be directly measured since the sources are independently encoded. Hence, accurate correlation channel estimation (CCE) is critical to achieve high joint decoding performance.

A. Related Work

Various works have addressed CCE for SW coding of binary sources, where the correlation is modeled as a binary symmetric channel (BSC). To estimate the crossover probability of the BSC, residual redundancies in the log-likelihood ratios were exploited in [17], while expectation-maximization (EM) algorithms were proposed in [18], [19]. In [20], a maximum likelihood estimate (MLE) was obtained using the source and the side information syndromes, which served as an initialization for the EM algorithm. Nevertheless, in [17]–[20] the correlation noise was assumed to be stationary within a

This work is supported by the FWO Flanders (projects G.0391.07 and G.0047.12), by the iMinds Project Little Sister and by NSF (CCF 1117886).

N. Deligiannis was with the Department of Electronics and Informatics, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium, and with iMinds, Gaston Crommenlaan 8 (b102), 9050 Ghent, Belgium. He is now with the Electronic and Electrical Engineering Department, University College London, Roberts Building, Torrington Place, London, WC1E 7JE, UK. (email: n.deligiannis@ucl.ac.uk).

A. Munteanu, and P. Schelkens are with the Department of Electronics and Informatics, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium, and with iMinds, Gaston Crommenlaan 8 (b102), 9050 Ghent, Belgium (email: acmuntea@etro.vub.ac.be; pschelke@etro.vub.ac.be).

S. Wang is with the Division of Biomedical Informatics, University of California, San Diego, San Diego, CA (email: shw070@ucsd.edu).

S. Cheng is with the School of Electrical and Computer Engineering, University of Oklahoma, Tulsa, OK, 74135 (email: samuel.cheng@ou.edu).

codeword. A CCE approach that can track the varying BSC crossover parameter in a codeword by incorporating particle filters in an LDPC factor graph, was proposed in [21]. Particle filtering was subsequently substituted by an expectation-propagation algorithm, resulting in comparable accuracy and reduced computational overhead [22].

Concerning WZ coding, several CCE methods have been proposed in the context of DVC. The majority of the schemes assume an additive noise channel, where the correlation noise is an independent memoryless zero-mean Laplacian distributed random variable [6], [23]. Several stationarity levels of the correlation noise have been considered. In transform-domain DVC, for instance, the correlation noise can be estimated for every frequency band per frame or even per transform coefficient [23]. To improve the CCE in the transform domain, the use of a pixel-domain CCE and the spatial correlation of the noise signal is proposed in [24]. A similar concept was exploited in the TRACE algorithm [25], where refinement of the noise variance upon decoding of each DCT band was carried out. Exploiting cross-band and inter-bit-plane dependencies [26] for noise frame residue refinement was put forward in [27]. Alternatively, block-based classification of the Laplace scaling parameter using offline training was performed in [28]. An adaptive CCE method based on clustering of DCT blocks, updated at the bit-plane level, was presented in [29]. Nevertheless, these CCE techniques for DVC are application-specific; many of which are even explicitly tailored to particular codec components; e.g., the methods in [23], [24], [27]–[29] are designed for motion-compensated interpolation based side information creation.

An application-generic approach to CCE for WZ coding, the method in [30] exploits particle filtering based belief propagation, enabling adaptation to fine noise stationarity levels. However, it does not support layered coding since all the bit-planes are decoded on a joint factor graph. In addition, the increased code length and the use of particle filtering put a high computational strain on the decoder. Joint bit-plane decoding was also used in [31] as to exploit inter-bit-plane correlation in the WZ decoding and in the CCE refinement.

Yet, only few CCE solutions can address application-generic layered WZ coding. The authors of [32] extended the method in [20] to count for the inter-bit-plane dependencies expressed using a Laplace correlation noise model. In prior work [33], we proposed an efficient bit-plane-per-bit-plane successively refined CCE for the side-information-dependent (SID) correlation channel. In [32], [33], however, the correlation noise was assumed stationary per DCT band of each frame.

B. Contributions

Targeting at efficient generic layered WZ coding, this paper proposes a novel approach to progressively refined maximum likelihood (ML) estimation of the Laplacian correlation channel. The proposed technique features key novelties with respect to the state-of-the-art.

First, the proposed method considers the correlation noise statistics to vary per block (a.k.a. chunk) of consecutive samples in a codeword and provides accurate CCE even when small noise stationarity blocks are assumed. This opposes the methods in [32], [33], where the correlation noise was considered stationary within a codeword.

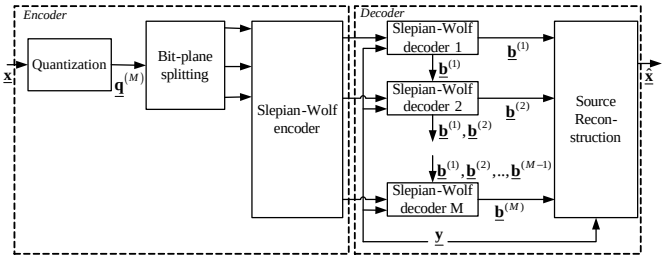


Fig. 1. The layered Wyner-Ziv coding scheme.

Second, at each WZ decoding stage, we derive a novel estimate of the correlation noise scaling parameter per stationarity block by exploiting the side information and the decoding information from the previous decoded stages. Conversely to our previous work [33], it is proven that the derived estimate constitutes the MLE of the correlation noise parameter. We also derive the Crámer-Rao lower bound (CRLB) on the accuracy of our MLE per decoding stage. We show that, as the number of decoded stages increases, the derived CRLB approaches the oracle-based CRLB, that is, the ideal bound attained when knowing the correlation noise samples.

Third, unlike methods solely designed for DVC, e.g., [23]–[25], [27], [28], the proposed approach is application-generic. When applied to DVC, particularly, the proposed technique is neither confined to a specific side information generation method (unlike [23], [24], [27]) nor does it impose a fixed decoding order of the DCT bands like the approach in [25].

Simulations on synthetic data show that, per decoding stage, the mean squared error (MSE) of the proposed MLE is vastly decreased compared to our previous state-of-the-art estimator from [33] and approaches the derived CRLB. Furthermore, the WZ rate-distortion (RD) performance obtained by using the proposed method is close to the one obtained by assuming perfect knowledge of the noise statistics. When implemented in our state-of-the-art DVC system for wireless capsule endoscopy, the proposed method yields systematic compression gains over the TRACE technique in [25] and our prior SID CCE method in [33] (i.e., rate savings up to 7.13% and 4.22%, respectively). The results also show that the proposed algorithm causes a reduction of the decoding complexity.

Concerning the remainder of the paper, Section II elaborates on layered WZ coding and states the CCE problem. The proposed CCE algorithm is presented in Section III, while bounds on the estimation accuracy are derived in Section IV. Section V expands on the application of the proposed CCE in wireless capsule endoscopy, and Section VI reports our experimental results. Finally, Section VII concludes the work.

II. Problem Setting

A. Layered Wyner-Ziv Coding Architecture

We consider layered WZ coding of a source X given correlated side information Y available at the decoder. Contrary

to [4], in the considered code design (see Fig. 1), source clustering is realized by scalar quantization without nesting. The design is motivated by constructions [34] that employ classical quantization and exclusively assign the binning operation to a highly-dimensional channel code. Such code constructions are commonly used in DVC, e.g., [6], [15], [25], [32].

1) Encoder: Consider a quantization function $Q_M(X)$, $Q_M : \mathcal{X} \rightarrow \{1, 2, ..., 2^M\}$, that is a mapping from the source to the quantization alphabet. Given the partition Q_M, the encoder quantizes the incoming source samples n-tuple $\mathbf{x} \triangleq x_1, x_2, ..., x_n$, $x_i \in \mathcal{X}$, to a quantization index n-tuple $\mathbf{q}^{(M)} \triangleq q_1^{(M)}, q_2^{(M)}, ..., q_n^{(M)}$, $q_i^{(M)} \in \{1, 2, ..., 2^M\}$. The quantization index n-tuple $\mathbf{q}^{(M)}$ is then divided into a number of M bit-planes of length n, i.e., $\underline{\mathbf{b}}^{(1)}$, $\underline{\mathbf{b}}^{(2)}$, ..., $\underline{\mathbf{b}}^{(M)}$, where $\mathbf{b}^{(m)} \triangleq b_1^{(m)}, b_2^{(m)}, ..., b_n^{(m)}$ denotes the m-th bit-plane and $b_i^{(m)} = \{0, 1\}$ represents the $i-$th bit in the m-th bit-plane.

Using a predefined order—e.g., from the most-significant bit-plane (MSB), $m = 1$, to the least significant bit-plane (LSB), $m = M$,—the source bit-planes $\mathbf{b}^{(1)}$, $\mathbf{b}^{(2)}$, ..., $\mathbf{b}^{(M)}$ are sequentially passed to a SW encoder that can ideally achieve the compression rate $H\left(Q_M(X)|Y\right)$. In this work, structured SW binning is implemented by the rate-adaptive LDPC accumulate (LDPCA) codes of [5].

2) Decoder : The decoder exploits the correlation statistics to sequentially decode the source bit-planes $\mathbf{b}^{(1)}$, $\mathbf{b}^{(2)}$, ..., $\mathbf{b}^{(M)}$ with the aid of the side information samples n-tuple $\underline{\mathbf{y}}$. In particular, for every source bit-plane $\mathbf{b}^{(m)}$, $m = \{1, 2, ..., M\}$, the correlation statistics between the source n-tuple $\mathbf{x}$ and the side information n-tuple $\underline{\mathbf{y}}$ are converted to soft-input estimates, namely, log-likelihood ratios (LLRs). These LLRs, which provide *a priori* information about the probability of each bit to be $b_i^{(m)} = 0$ or $b_i^{(m)} = 1$, are passed to the soft channel decoder. Then, an iterative soft-decoding algorithm (e.g., the sum-product algorithm [35]) is executed to decode the source bit-planes with an error probability close to zero.

For optimal layered WZ coding, during the formulation of the LLRs, information given by the side information and the already decoded source bit-planes is taken into account. Specifically, let $b_i^{(m)}$, $m \in \{1, 2, ..., M\}$, $i \in \{1, 2, ..., n\}$, be the bit under consideration, y_i be the corresponding side information sample value and $b_i^{(1)}, b_i^{(2)}, ..., b_i^{(m-1)}$ be the already decoded bits in the previous $m-1$ bit-planes. Then, the input LLR is given by $\log \frac{P\left(b_i^{(m)}=0|y_i, b_i^{(1)}, b_i^{(2)}, ..., b_i^{(m-1)}\right)}{P\left(b_i^{(m)}=1|y_i, b_i^{(1)}, b_i^{(2)}, ..., b_i^{(m-1)}\right)}$, where the numerator and the denominator are calculated by integrating the conditional probability density function (pdf) of the correlation channel, i.e., $f_{X|Y}(x|y)$, over the quantization bin defined by the already decoded bits $b_i^{(1)}, b_i^{(2)}, ..., b_i^{(m-1)}$ and by setting $b_i^{(m)} = 0$ or $b_i^{(m)} = 1$, respectively. Because of this chained approach to construct the LLRs, the total rate conforms to the chain rule of entropies [36]. This means that, *ideally*, there is no performance loss between non-layered and layered WZ coding [4].

Once all the source bit-planes are decoded, they are combined to form the decoded quantization indices of the source. These indices are then used to reconstruct the source with the help of the side information and the correlation statistics. In particular, if the mean squared error (MSE) distortion metric is employed, the optimal reconstruction of a source sample x_i is obtained as the centroid of the random variable X given the corresponding side information sample y_i and the decoded quantization index $q_i^{(M)}$. Namely, $\hat{x}_i = \mathbb{E}\left[x_i|y_i, q_i^{(M)}\right]$, where $\mathbb{E}[\bullet|\bullet]$ is the conditional expectation operator. Closed-form expressions for the reconstruction function can be found in [4] and [37] for a Gaussian and a Laplacian pdf, respectively.

3) Rate Control: Early DSC designs [4], [10], [11], assume the correlation statistics to be stationary and perfectly known *a priori* at the encoder and decoder. Under this assumption, the encoder and decoder agree on an efficient code rate driven by the SW rate. Yet, in practice, the correlation channel exhibits a non-stationary behavior, which neither can be known in advance nor can it be measured in practice. Therefore, a rate-adaptive decoder-driver rate control approach using a feedback channel is a good solution for many applications [6], [15], [32].

B. The Correlation Channel Estimation Problem

We assume that the correlation between the source X and the side information Y is expressed by a memoryless channel, $X = Y + Z$, where the correlation noise Z is zero-mean Laplacian, i.e., $Z \sim Laplace(0, \lambda)$, and independent of the side information, which is arbitrarily distributed. The assumption of independent correlation noise opposes our previous work [33], where the noise was dependent on the side information. Our motivation stems from the fact that most existing DVC systems employ the former model to express the correlation channel [6], [14], [23]–[32]. The correlation channel pdf is then given by

$$f_{X|Y}(x|y) = \frac{\lambda}{2} \exp(-\lambda|x - y|), \tag{1}$$

where $\lambda = \frac{\sqrt{2}}{\sigma_Z}$ is the scaling parameter of the distribution and $\sigma_Z^2 = \mathbb{E}\left\{Z^2\right\}$. We assume that the correlation noise Z is *block-stationary*; namely, λ varies per noise stationarity block (chunk), defined by K consecutive samples in the noise n-tuple $\mathbf{z}$, where $n \bmod K = 0$. To express this formally, we adhere to the following notation. An n-tuple $\mathbf{z}$ is written as $\mathbf{z} \triangleq \mathbf{z}_0\mathbf{z}_1\mathbf{z}_2...\mathbf{z}_{L-1}$, where $\mathbf{z}_\ell$ denotes a block of the K succeeding noise samples $z_{\ell \times K+i}$, where $i = \{1, 2, ..., K\}$ counts the samples in the block, $\ell \in \{0, 1, ..., L-1\}$ indexes the blocks, and $L \triangleq \frac{n}{K}$. The set $\{\lambda_\ell\}_{\ell=0}^{L-1}$ contains the scaling parameters that control the noise distribution over the entire n-tuple $\mathbf{z}$.

Under the common but unrealistic assumption that the correlation noise samples $\mathbf{z}$ are known *a priori* to the decoder, an offline MLE for the scaling parameter per stationarity block $\mathbf{z}_\ell$ is given by

$$\hat{\lambda}_\ell^{\text{off}} = K \left(\sum_{i=1}^{K} |z_{\ell \times K+i}| \right)^{-1}, \ell = 0, 1, .., L-1. \tag{2}$$

We remark that $\hat{\lambda}_\ell^{\text{off}}$ in (2) is a minimum variance unbiased (MVU) estimator [38].

In practical DSC designs, however, the correlation noise cannot be directly measured as the source and the side information are located at the encoder and decoder, respectively.

In the following, we derive a novel online MLE for λ_ℓ per stationary block at the decoder, which can be successively refined as source bit-planes are decoded. We also evaluate the asymptotic efficiency of the proposed MLE using the CRLB.

We remark that a block-stationary noise model with a constant number of samples per stationarity block is assumed for simplicity. The focus of the paper is on the derivation and the evaluation of the performance of the MLE λ_ℓ. Devising mechanisms to adapt to complex noise variations through arbitrary shaped blocks is outside the scope of this paper.

III. Maximum Likelihood Correlation Estimation

A. Introduction to the Basic Concept

In the layered WZ coding scheme, discussed in Section II-A, the source n-tuple $\mathbf{x}$ is quantized with a 2^M-stage quantization partition Q_M and successively encoded into M bit-planes. This progressive coding scheme is equivalent to an *embedded* partition $Q_1 Q_2 ... Q_m ... Q_M$, where the m-th stage partition Q_m has 2^m quantization intervals corresponding to indices $q_i^{(m)} \in \{1, 2, ..., 2^m\}$, formed by the bits $b_i^{(1)}\, b_i^{(2)} ... b_i^{(m)}$. We exploit this progressive coding approach at the decoder to design a successively refined MLE of the *block-stationary* correlation channel parameter λ_ℓ.

The proposed progressive ML CCE scheme is described as follows: Starting from initial correlation channel parameter estimates $\left\{\hat{\lambda}_\ell^{(0)}\right\}_{\ell=0}^{L-1}$, the SW decoder decodes the MSB $\underline{\mathbf{b}}^{(1)}$ of the source samples. This initial estimate may for instance be obtained from a previously decoded data n-tuple, or it can be the expected value of the parameter derived from offline observations.

By decoding $\mathbf{b}^{(1)}$, the decoder obtains access to a coarsely quantized version of the source samples, denoted by the n-tuple $\mathbf{q}^{(1)}$, which corresponds to partition Q_1. The decoder can exploit the available quantized source n-tuple $\mathbf{q}^{(1)}$ and the side information n-tuple $\underline{\mathbf{y}}$ to perform CCE for the next bit-plane. Analogous to the decomposition of the noise samples into stationarity blocks, let us write $\mathbf{q}^{(1)} \triangleq \mathbf{q}_0^{(1)} \mathbf{q}_1^{(1)} \mathbf{q}_2^{(1)} ... \mathbf{q}_{L-1}^{(1)}$ and $\underline{\mathbf{y}} \triangleq \mathbf{y}_0 \mathbf{y}_1 \mathbf{y}_2 ... \mathbf{y}_{L-1}$. Namely, per stationarity block ℓ, the decoder derives an MLE $\hat{\lambda}_\ell^{(1)}$ using the samples $q_{\ell \times K+i}^{(1)}$ and $y_{\ell \times K+i}$, $i \in \{1, 2, ..., K\}$, in the K-tuples $\mathbf{q}_\ell^{(1)}$ and $\mathbf{y}_\ell$, respectively. The proposed derivation of the MLE is detailed in Section III-B. The updated parameters $\left\{\hat{\lambda}_\ell^{(1)}\right\}_{\ell=0}^{L-1}$ are then used to derive the soft-input information to SW decode the next source bit-plane $\mathbf{b}^{(2)}$. The decoder then combines the available $\mathbf{b}^{(1)}$ and $\mathbf{b}^{(2)}$ to form a more accurate version, i.e., $\mathbf{q}^{(2)}$, of the quantized source samples (corresponding to partition Q_2), which is in turn used to obtain a refined MLE $\hat{\lambda}_\ell^{(2)}$ per stationary block ℓ.

In a recursive manner, after SW decoding of the $\underline{\mathbf{b}}^{(m)}$ bit-plane ($1 \leq m \leq M$), the proposed algorithm forms a quantized version $\mathbf{q}^{(m)}$ of the source samples based on the available bit-planes $\mathbf{b}^{(1)} \mathbf{b}^{(2)} ... \mathbf{b}^{(m)}$. Using $\mathbf{q}_\ell^{(m)}$ and $\mathbf{y}_\ell$ the decoder obtains an MLE $\hat{\lambda}_\ell^{(m)}$ for the $\ell-$th stationary block. We highlight that when all bit-planes have been decoded, a.k.a., $m = M$, the proposed MLE algorithm is carried out again so as to obtain the channel parameters $\left\{\hat{\lambda}_\ell^{(M)}\right\}_{\ell=0}^{L-1}$, which are used to perform optimal minimum MSE (MMSE) reconstruction. Since per additional decoded bit-plane $\underline{\mathbf{b}}^{(m)}$ the available approximation $\mathbf{q}^{(m)}$ of the source samples at the decoder becomes more accurate, the obtained MLE $\left\{\hat{\lambda}_\ell^{(m)}\right\}_{\ell=0}^{L-1}$ are becoming more accurate as well, thereby leading to a progressively improved CCE.

B. Maximum Likelihood Estimation Formulations

We now concentrate on the proposed derivation of the MLE $\lambda_\ell^{(m)}$ from the samples in the K-tuples $\mathbf{q}_\ell^{(m)}$ and $\mathbf{y}_\ell$. Since we focus on the samples in a stationarity block and in order to maintain the simplicity of the notations, without loss of generality, we hereby drop the subscript ℓ that indexes the stationarity block. Consider the following definition.

Definition 1 (m-th Stage Estimator): Let $\mathbf{q}^{(m)} \triangleq q_1^{(m)}, q_2^{(m)}, ..., q_K^{(m)}$, $\mathbf{q}^{(m)} \in \{1, 2, ..., 2^m\}^K$ be a K-tuple of quantization indices of the WZ source samples, formed by m SW decoded bit-planes, $1 \leq m \leq M$. Also, let $\mathbf{y} \triangleq y_1, y_2, ..., y_K$ be a K-tuple of side information samples with $\mathbf{y} \in \mathcal{Y}^K$. An m-th stage estimator $\hat{\lambda}^{(m)}$ of the Laplace correlation channel in (1) is defined as a function $\varphi : \{1, 2, ..., 2^m\}^K \times \mathcal{Y}^K \rightarrow \Lambda$, where $\Lambda = (0, +\infty)$ is the parameter set for $\hat{\lambda}^{(m)}$ and K is the samples' size.

We consider the m-th stage *conditional likelihood function* $\mathcal{L}(\lambda) = p(\mathbf{q}^{(m)}|\mathbf{y}; \lambda)$, where $p(\mathbf{q}^{(m)}|\mathbf{y}; \lambda)$ is the conditional probability mass function (pmf) of the quantization indices K-tuple given the side information samples' K-tuple, parameterized by the parameter λ of the correlation channel pdf. Since the correlation channel is assumed memoryless, we have

$$\mathcal{L}(\lambda) = p(\mathbf{q}^{(m)}|\mathbf{y}; \lambda) = \prod_{i=1}^{K} p(q_i^{(m)}|y_i; \lambda), \tag{3}$$

where $p(q_i^{(m)}|y_i; \lambda)$, $\forall i \in \{1, 2, ..., K\}$, is the conditional pmf for the quantization index $q_i^{(m)}$ of a source sample x_i given its corresponding side information sample y_i. The conditional pmf $p(q_i^{(m)}|y_i; \lambda)$ is defined by the integration of the correlation channel pdf $f_{X|Y}(x|y)$ over the quantization bin defined by $q_i^{(m)}$, that is,

$$\begin{aligned} p(q_i^{(m)}|y_i; \lambda) &= \int_{u_i^{(m)}}^{v_i^{(m)}} f_{X|Y}(x|y_i) dx \\ &= \int_{u_i^{(m)}}^{v_i^{(m)}} \frac{\lambda}{2} \exp(-\lambda |x - y_i|) dx, \end{aligned} \tag{4}$$

where $u_i^{(m)}$ and $v_i^{(m)}$ are respectively the lower and upper bound of the quantization bin indexed by $q_i^{(m)}$ in the quantization partition Q_m. Depending on the relative position of y_i with respect to the bounds $[u_i^{(m)}, v_i^{(m)})$ of the bin indexed by

$q_i^{(m)}$, the probability $p(q_i^{(m)}|y_i;\lambda)$ can be computed as

$$p(q_i^{(m)}|y_i;\lambda) = \begin{cases} \frac{1}{2}e^{-\lambda(u_i^{(m)}-y_i)} - \frac{1}{2}e^{-\lambda(v_i^{(m)}-y_i)}, & y_i < u_i^{(m)} \\ \frac{1}{2}e^{\lambda(v_i^{(m)}-y_i)} - \frac{1}{2}e^{\lambda(u_i^{(m)}-y_i)}, & y_i \geq v_i^{(m)} \\ 1 - \frac{1}{2}e^{\lambda(u_i^{(m)}-y_i)} - \frac{1}{2}e^{-\lambda(v_i^{(m)}-y_i)}, & \\ \qquad\qquad u_i^{(m)} \leq y_i < v_i^{(m)} \end{cases} \quad (5)$$

We wish to find an m-th stage MLE $\hat{\lambda}^{(m)}$, such that

$$\hat{\lambda}^{(m)} = \arg\max_{\lambda} \mathcal{L}(\lambda). \quad (6)$$

To ease the calculations, we derive $\hat{\lambda}^{(m)}$ by equivalently maximizing a monotone transformation of (3), that is

$$\ln \mathcal{L}(\lambda) = \ln p(\mathbf{q}^{(m)}|\mathbf{y};\lambda) = \sum_{i=1}^{K} \ln p(q_i^{(m)}|y_i;\lambda). \quad (7)$$

Then, the m-th stage MLE $\hat{\lambda}^{(m)}$ that achieves the maximum in (6), can be found as the solution of the following m-th stage *conditional log-likelihood equation*

$$\frac{\partial}{\partial\lambda} \ln \mathcal{L}(\lambda) = 0 \Rightarrow \sum_{i=1}^{K} \frac{\partial}{\partial\lambda} \ln p(q_i^{(m)}|y_i;\lambda) = 0, \quad (8)$$

where the partial derivative $\frac{\partial}{\partial\lambda} \ln p(q_i^{(m)}|y_i;\lambda)$ is derived in Appendix A for each of the three forms of (5), corresponding to the cases $y_i < u_i^{(m)}$, $y_i \geq v_i^{(m)}$ and $u_i^{(m)} \leq y_i < v_i^{(m)}$. Concerning the solution of (8), the following holds.

Theorem 1 (Existence and Uniqueness): There is a unique finite m-th stage MLE $\hat{\lambda}^{(m)}$ of the Laplace correlation channel pdf $f_{X|Y}(x|y)$ for K samples, under the condition that there exists at least one side information sample y_i in $\mathbf{y}$ that is not between or on the bounds of the quantization bin defined by its corresponding m-th stage quantization index $q_i^{(m)}$.

Proof: The proof of Theorem 1 is given in Appendix B preceded by a preparatory lemma. ■

Corollary 1: If every side information sample y_i in $\mathbf{y}$ is located between or on the bounds of the quantization bin defined by its corresponding m-th stage quantization index $q_i^{(m)}$, namely, if $u_i^{(m)} \leq y_i \leq v_i^{(m)}$, $\forall i \in \{1,2,...,K\}$, then the m-th stage MLE is given by $\hat{\lambda}^{(m)} = +\infty$.

Proof: Intuitively, when $u_i^{(m)} \leq y_i \leq v_i^{(m)}$ $\forall y_i$ in $\mathbf{y}$, then the likelihood is maximized at $\sigma_Z = 0$ or $\lambda = +\infty$. Formally, from the proof of Theorem 1, if $u_i^{(m)} \leq y_i \leq v_i^{(m)}, \forall i \in \{1,2,...,K\}$ then $\lim_{\lambda\to+\infty} \frac{\partial}{\partial\lambda} \ln p(\mathbf{q}^{(m)}|\mathbf{y};\lambda) = 0$. Combining this result with Lemma 1 in Appendix B ends the proof. ■

The following two propositions describe the conditions under which an analytical derivation of the finite MLE, defined in Theorem 1, is feasible.

Proposition 1 (Analytical Solutions): The MLE can be found analytically if the following sufficient conditions are satisfied: (i) the quantizer has uniform intervals and (ii) $y_i \notin \left[u_i^{(m)}, v_i^{(m)}\right), \forall i = \{1,2,...,K\}$.

Proof: The proof is sketched in Appendix C. ■

Proposition 2 (High Rate Assumptions): Under high rate assumptions, the MLE is approximated by $\hat{\lambda}^{(m)} = K\left(\sum_{i=1}^{K} |\xi_i^{(m)} - y_i|\right)^{-1}$, where $\xi_i^{(m)} = \frac{u_i^{(m)}+v_i^{(m)}}{2}$.

Proof: Let $\Delta_i^{(m)} = v_i^{(m)} - u_i^{(m)}$denote the quantization interval at the m-th decoding stage. When $\Delta_i^{(m)} \to 0$ the pdf $f_{X|Y}(x|y)$ is assumed to be constant in the quantization interval and thus, the pmf in (4) can be approximated as $p(q_i^{(m)}|y_i;\lambda) \simeq \frac{\lambda}{2}e^{-\lambda|\xi_i^{(m)}-y_i|}\Delta_i^{(m)}$. Inserting the latter in (8) and solving for λ ends the proof. ■

Remark 1: In the general case, the conditional log-likelihood equation $\frac{\partial}{\partial\lambda} \ln \mathcal{L}(\lambda) = 0$ is an exponential equation on λ and cannot be solved analytically. Hence, we hinge on an numerical solution. A grid search for the $\hat{\lambda}^{(m)}$ that attains the maximum in (6) is not computationally feasible, since the parameter set for λ is continuous and not confined to a finite interval, i.e., $\Lambda = (0, +\infty)$. To derive the MLE, we implement a numerical root finding algorithm that combines bracketing, bisection, secant, and inverse quadratic interpolation methods [39]. After a maximum number of iterations J_{MAX}, if the algorithm cannot find the root of equation (8), with a given accuracy ϵ_{acc}, then we set $\hat{\lambda}^{(m)} = \hat{\lambda}^{(m-1)}$. Namely, the correlation channel parameter estimate from the previous decoding level is retained. The same holds for the case that falls into Collorary 1. It is noteworthy that as more bit-planes are decoded the quantization intervals in the m-th stage partition Q_m in the decoder become finer and in turn, the probability of having a case corresponding to Collorary 1 drops. This effect is in accordance with the progressively refined nature of the proposed estimator.

The proposed ML CCE algorithm, as applied for the $\ell-$th stationarity block, is summarized in Algorithm 1.

Algorithm 1 ML Estimation at Stationary Block Level

```
function MLE(q_ℓ^(m), y_ℓ)
    if u_{ℓ×K+i}^(m) ≤ y_{ℓ×K+i} ≤ v_{ℓ×K+i}^(m), ∀i ∈ {1,2,...,K}
    then
        Set λ̂_ℓ^(m) = λ̂_ℓ^(m−1)
    else
        if y_{ℓ×K+i} ∉ [u_{ℓ×K+i}^(m), v_{ℓ×K+i}^(m)) ∧
            Δ_{ℓ×K+i}^(m) = Δ^(m), ∀i ∈ {1,2,...,K}
        then
            λ̂_ℓ^(m) is analytically found by (18) or (19).
        else
            Find λ̂_ℓ^(m) by numerically solving the equation
            ∂/∂λ ln L(λ) = 0 defined in (8).
        end if
    end if
end function
```

Remark 2: For simplicity, the above presentation of the ML estimation formulations consider identical side information samples $\underline{\mathbf{y}}$ at every refinement level $m \in \{1,2,...,M\}$. Yet, since the proposed ML CCE algorithm is applied per refinement stage m, it can be directly generalized to the case where the side information samples $\underline{\mathbf{y}}^{(m)}$ are refined per level.

Remark 3: The proposed progressively refined ML CCE algorithm is applicable irrespective of the side information's distribution and the quantization function used at the layered WZ encoder. The MLE is derived based on the side information samples and the bounds $[u_i^{(m)}, v_i^{(m)})$ of the quantization bin corresponding to each quantized source sample $q_i^{(m)}$.

We conjecture that the proposed technique can be extended to estimate the standard deviation of a zero-mean Gaussian correlation noise as well. The proof and validation of this conjecture is left as topic for further research.

IV. Asymptotic Efficiency Bounds

The m-th stage MLE, which results from solving (8), is biased since (8) is a non-linear equation on λ. To study the asymptotic efficiency of the proposed MLE, we derive the m-th stage CRLB$^{(m)}$ for the problem. We observe that the domain in which the m-th stage conditional likelihood function in (3) is nonzero does not depend on λ. Hence, the regularity condition $\mathbb{E}\left[\frac{\partial}{\partial\lambda}\ln p\left(\mathbf{q}^{(m)}|\mathbf{y};\lambda\right)|Y\right]=0$ is satisfied [38]. As a consequence, the CRLB$^{(m)}$ is given by[1]

$$\text{CRLB}^{(m)} = -\mathbb{E}\left[\frac{\partial^2}{\partial\lambda^2}\ln p(\mathbf{q}^{(m)}|\mathbf{y};\lambda)|Y\right]^{-1}, \qquad (9)$$

where $\mathbb{E}[\bullet|Y]$ is the conditional expectation with respect to the conditional pmf $p(\mathbf{q}^{(m)}|\mathbf{y};\lambda)$. The average curvature of the conditional log-likelihood function can be computed as

$$\begin{aligned}
&\mathbb{E}\left[\frac{\partial^2}{\partial\lambda^2}\ln p(\mathbf{q}^{(m)}|\mathbf{y};\lambda)|Y\right]\\
&=\sum_{\mathbf{q}^{(m)}}\frac{\partial^2}{\partial\lambda^2}\ln p(\mathbf{q}^{(m)}|\mathbf{y};\lambda)\,p(\mathbf{q}^{(m)}|\mathbf{y};\lambda)\\
&=\sum_{q_1^{(m)}=1}^{2^m}\ldots\sum_{q_K^{(m)}=1}^{2^m}\sum_{i=1}^{K}\frac{\partial^2}{\partial\lambda^2}\ln p\left(q_i^{(m)}|y_i;\lambda\right)\prod_{i=1}^{K}p\left(q_i^{(m)}|y_i;\lambda\right)\\
&=\sum_{i=1}^{K}\sum_{q_i^{(m)}=1}^{2^m}\frac{\partial^2}{\partial\lambda^2}\ln p\left(q_i^{(m)}|y_i;\lambda\right)\,p\left(q_i^{(m)}|y_i;\lambda\right), \qquad (10)
\end{aligned}$$

where the forms of $p\left(q_i^{(m)}|y_i;\lambda\right)$ and $\frac{\partial^2}{\partial\lambda^2}\ln p\left(q_i^{(m)}|y_i;\lambda\right)$ are given by (5) and (12a)-(12c), respectively. Lemma 1 in Appendix B proves that $\frac{\partial^2}{\partial\lambda^2}\ln p\left(\mathbf{q}^{(m)}|\mathbf{y};\lambda\right)<0$, thereby verifying that CRLB$^{(m)}>0$. For the analytical forms of the MLE, obtained under the conditions described in Proposition 1, the m-th stage CRLB$^{(m)}$'s are given in Appendix C.

It is worth noticing that, as the number of decoded WZ stages increases, the CRLB$^{(m)}$ asymptotically (i.e., $\Delta_i^{(m)}\rightarrow 0$) approaches the oracle (or offline) CRLB. The latter, which is defined by $\frac{\lambda^2}{K}$, refers to the unrealistic case where the decoder has access to the noise samples. This convergence to the oracle CRLB can be readily verified by replacing the approximated high rate pmf (given in the proof of Proposition 2) into (10).

[1]For the sake of simplicity, we again drop the subscript ℓ that indexes the stationarity block size.

V. Application to Wyner–Ziv Coding of Endoscopic Video

The practical merit of the proposed MLE is shown using our endoscopic DVC (EDVC) [15] system. EDVC delivers state-of-the-art DVC performance, outperforming several relevant up-to-date codecs, including [14], [33], [40], [41].

A. Codec Architecture Overview

In EDVC [15], the key frames are intra-coded with JPEG or H.264/AVC Intra, whereas the WZ frames are encoded in a hash and a WZ stage. The hash is a reduced resolution version of each WZ frame coded at a low quality with JPEG or H.264/AVC Intra. Each WZ frame undergoes the discrete cosine transform (DCT) and then the coefficients of each band β are encoded following the codec in Section II-A. Namely, the quantized coefficients per band $\mathbf{q}^{(M)}(\beta)$ are turned into bit-planes $\mathbf{b}^{(M)}(\beta)$ that are passed to an LDPCA encoder [5].

In the decoder, the decoded key frames are used together with the decoded and up-scaled hash to produce a motion-compensated prediction of the WZ frame (see [15] for details). Transforming the prediction frame yields side information. In [15], correlation estimation was performed using our SID CCE algorithm from [33] that assumes band-level noise stationarity per frame. In this work, however, we implement our novel ML CCE, which is applied for each stationarity block per band of each frame. The obtained statistics are used to decode the bit-planes of the WZ bands and perform reconstruction. Inverse integer DCT is finally carried out to return to the pixel domain.

B. Proposed ML Correlation Estimation Technique in EDVC

The correlation between the DCT coefficients of a WZ frame and its side information is considered Laplacian with a scaling parameter varying per group of neighboring coefficients in each band per frame. For consistency, following the notations from Section II, the size of a band of a frame is denoted by $n=V\times H$, where V, H represent the number of DCT blocks in the vertical and horizontal direction of a transformed frame, respectively. Coefficients of the same band, say β, that belong to $K=K_1\times K_2$ neighboring DCT blocks are clustered into correlation noise stationarity blocks, where K_1, K_2 are divisors of V, H, respectively. By definition, there are $L=\frac{n}{K_1\times K_2}$ stationarity blocks of coefficients per band β of a frame; namely, for a side information frame band we can write $\underline{\mathbf{y}}(\beta)\triangleq\mathbf{y}_0(\beta)\mathbf{y}_1(\beta)\mathbf{y}_2(\beta)...\mathbf{y}_{L-1}(\beta)$. The $K=K_1\times K_2$ coefficients in each block $\mathbf{y}_\ell(\beta)$ share the same correlation noise parameter $\lambda_\ell(\beta)$. The set $\{\lambda_\ell(\beta)\}_{\ell=0}^{L-1}$ refers to the scaling parameters of the noise over the band β of a frame.

The proposed ML CCE technique is executed per band β of each WZ frame. The scaling parameters $\left\{\hat{\lambda}_\ell^{(0)}(\beta)\right\}_{\ell=0}^{L-1}$ needed to initiate the algorithm are copied from the corresponding band of the previously decoded WZ frame[2]. By decoding each

[2]For the first WZ frame in the sequence a predefined set of scaling parameters is used to decode the first bit-plane per band. Specifically, based on offline measurements, one parameter per band is determined as the average value from all the frames in several test sequences.

bit-plane $\mathbf{b}^{(m)}(\beta)$ in the band, the decoder learns a recursively updated quantized version $\mathbf{q}^{(m)}(\beta)$ of the WZ coefficients. Based on the K coefficients in each $\mathbf{q}_\ell^{(m)}(\beta)$ and $\mathbf{y}_\ell(\beta)$, an updated MLE $\hat{\lambda}_\ell^{(m)}(\beta)$ is derived, as detailed in Section III-B.

The stationarity block size K, which is regulated by the parameters K_1 and K_2, affects the accuracy of the algorithm. Low K_1 and K_2 values (correspondingly, low K values) provide adjustment to the regional variation of the channel statistics but, on the other hand, challenge accurate statistical inference due to narrow statistical support. In particular, a low value of K increases the probability of the special case described in Corollary 1, i.e., $\hat{\lambda}_\ell^{(m)}(\beta) = +\infty$. For this reason, the algorithm also estimates a single $\hat{\lambda}^{(m)}(\beta)$ per band of each frame, using all the band samples in $\underline{\mathbf{y}}(\beta)$ and $\underline{\mathbf{q}}^{(m)}(\beta)$. This estimate is used in case an MLE for a stationarity block could not be derived.

VI. Experimental Results

A. Evaluation on Synthetic Data

We implement the layered WZ coding system detailed in Section II-A. By definition (see Section II-B), the source is formed as $X = Y + Z$, where the side information Y is considered either Gaussian or Laplacian with a zero mean and $\sigma_Y = 50$. The correlation noise samples are drawn from an independent zero-mean Laplacian distribution. Per stationarity block $\ell = 0, 1, ..., L-1$ in a codeword, the noise scaling parameter $\lambda_\ell = \frac{\sqrt{2}}{\sigma_{Z_\ell}}$ is randomly selected from the interval $\lambda_\ell \in [0.2020, 2.8284]$ (corresponding to $\sigma_{Z_\ell} \in [0.5, 6]$) with uniform probability. The source sample (codeword) length is set to $n = 1584$ or 6336. Source quantization is realized by a uniform scalar quantizer with 2^8 levels, corresponding to $M = 8$ bit-planes. To SW decode the first bit-plane of the encoded source n-tuple and thus, to initiate the proposed MLE algorithm, a fixed scaling parameter value, i.e., $\left\{\hat{\lambda}_\ell^{(0)}\right\}_{\ell=0}^{L-1} = 1.5152$, is used. This initial value constitutes the mean of the interval from which λ_ℓ is drawn. Concerning the configuration of the numerical root-finding algorithm, we set the initial value to 0.01, the maximum number of iterations to $J_{MAX} = 1000$, and the accuracy to $\epsilon_{acc} = 10^{-10}$. Average results over 400 independent trials are presented.

We begin by evaluating the MSE performance of the proposed MLE per decoded stage (represented by the number m of SW decoded bit-planes) with respect to the derived m-th stage CRLB$^{(m)}$ and the oracle CRLB. The MSE performance obtained with the initial estimator (i.e., using the fixed parameter value $\left\{\hat{\lambda}_\ell^{(0)}\right\}_{\ell=0}^{L-1} = 1.5152$ to decode all the bit-planes) and our previous SII estimator[3] from [33] are also added in the comparison. Four experimental setups are considered, corresponding to a Gaussian or Laplacian side information[4] and to a source length of $n = 1584$ or 6336. In this experimental setup, we set $n = K$ (i.e., $L = 1$). The results in Fig. 2 indicate that the proposed MLE systematically yields a significantly lower MSE performance in comparison to the naive initial estimator as well as our previous successively refined SII estimator from [33]. The vast MSE decrease compared to our prior estimator is due to the fact the proposed estimate is the MLE of the noise scaling parameter. In fact, the MSE of our MLE closely follows the CRLB$^{(m)}$ per decoding stage m, irrespective of the distribution of the side information or the source length. Hence, the proposed MLE is asymptotically efficient given the decoded information at stage m. It is important to observe that, as m increases, the CRLB$^{(m)}$ and the MSE of the proposed MLE progressively converge to the oracle CRLB. This behavior corroborates the successively refined nature of the proposed estimator.

Subsequently, we compare the RD performance of layered WZ coding using the proposed MLE algorithm against the performance obtained when assuming perfect knowledge of the noise scaling parameters at the decoder (offline/oracle estimation). The RD performance when the initial fixed scaling parameter is used to decode all the bit-planes is also assessed. In this experimental setup, the source length is set to 1584 samples, the side information is Laplacian and various stationarity block sizes, that is, $K = 198, 132, 99,$ or 44 [5] are considered. Fig. 3 depicts the distortion—expressed in terms of the signal-to-noise ratio SNR $= 10\log(\sigma_X^2/(\frac{1}{n}\sum(x_i - \hat{x}_i)^2))$—versus the rate required for the compression of bit-planes $m = 2$ to 8. The results show that, in case of large stationarity block sizes [see Figs. 3(a)-(b)], the proposed MLE achieves an RD performance almost identical to offline estimation irrespective of the RD point. In case of smaller stationarity block sizes [see Fig. 3(c)-(d)], a performance loss compared to offline estimation is observed due to the reduced statistical support for accurate derivation of the MLE. Compared to decoding with the fixed initial parameter, the proposed MLE systematically yields higher RD performance. In effect, the obtained improvements are mounting with the rate because the proposed MLE is successively refined as more information is decoded.

B. Evaluation on Endoscopic Video Data

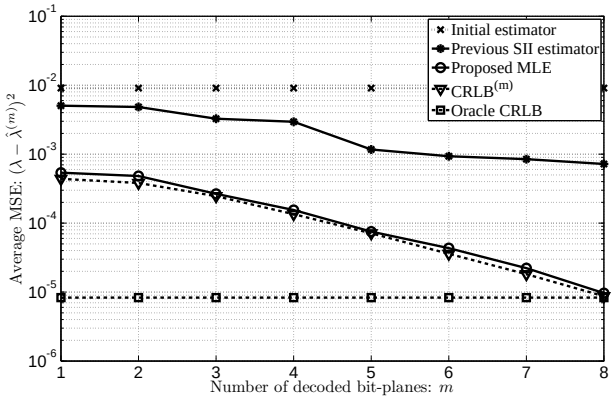
To assess the impact of the proposed algorithm on the performance of our state-of-the-art EDVC [15] system, we carry out experiments on three conventional endoscopic and five capsule[6] endoscopic test video sequences obtained from clinical examinations performed at the Gastroenterology Clinic of the Universitair Ziekenhuis Brussels. The capsule endoscopic sequences, which visualize diverse areas of the gastrointestinal track of 2 patients, contain 150 frames each, acquired at a

[3]The SID estimator in [33] is designed for the SID correlation channel, where the distribution of the correlation noise $Z = X - Y$ is zero-mean Laplacian with a scaling parameter $\lambda(y)$ that varies depending on the realization y of the side information. For practicality [33], a finite number of scaling parameters $\lambda(\tilde{y}_i)$ are estimated by assigning side information samples to finite indices $\tilde{y}_i$ via quantization. In this specific experiment, the quantizer is uniform with 2^7 levels. According to [33, Lemma 2] one may derive the scaling parameter estimate for the SII channel from the estimates derived for the SID channel as $\hat{\lambda}_{SII} = \left[\sum_{\tilde{y}_i} p_{\tilde{Y}}(\tilde{y}_i)\frac{1}{\hat{\lambda}^2(\tilde{y}_i)}\right]^{-1/2}$, where $p_{\tilde{Y}}(\tilde{y}_i)$ is the pmf of the quantized side information.

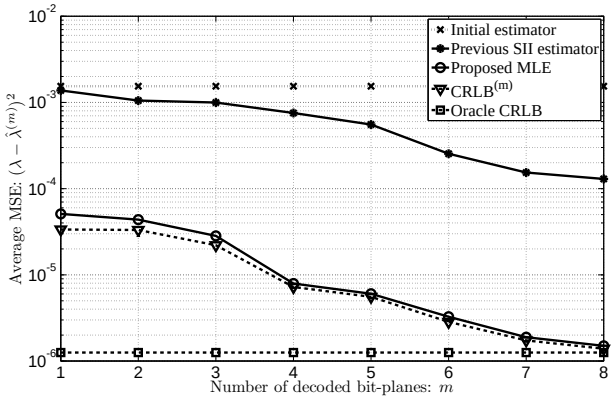
[4]As mentioned in *Remark 3*, the algorithm is applied unaltered irrespective of the distribution of the side information.

[5]The scaling parameter of the correlation noise is constant for the K samples within each stationarity block.

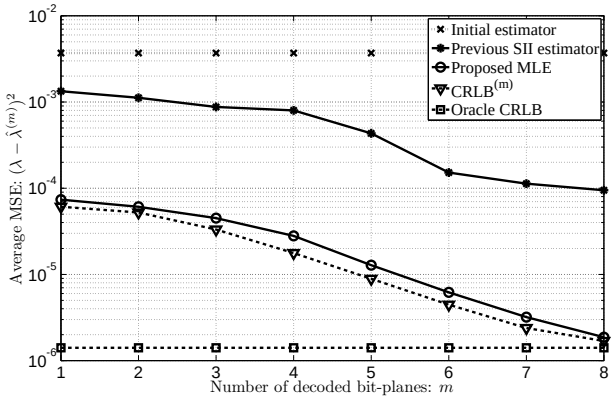
[6]The capsule endoscopic video data were obtained with the PillCam SB2 [42] from Given Imaging.



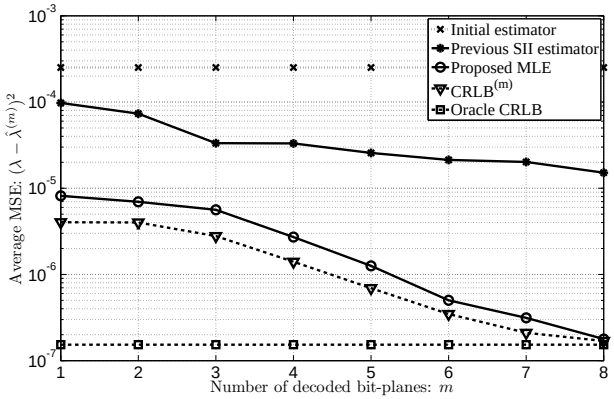
(a) Y Laplace, $n = 1584$.



(b) Y Gaussian, $n = 1584$.



(c) Y Laplacian, $n = 6336$.



(d) Y Gaussian, $n = 6336$.

Fig. 2. The average MSE of the proposed MLE, our previous SII estimator from [33], and the initial estimator $\hat{\lambda} = 1.5152$ with respect to the m-th stage $\text{CRLB}^{(m)}$ and the oracle CRLB, plotted versus the number of decoded bit-planes.

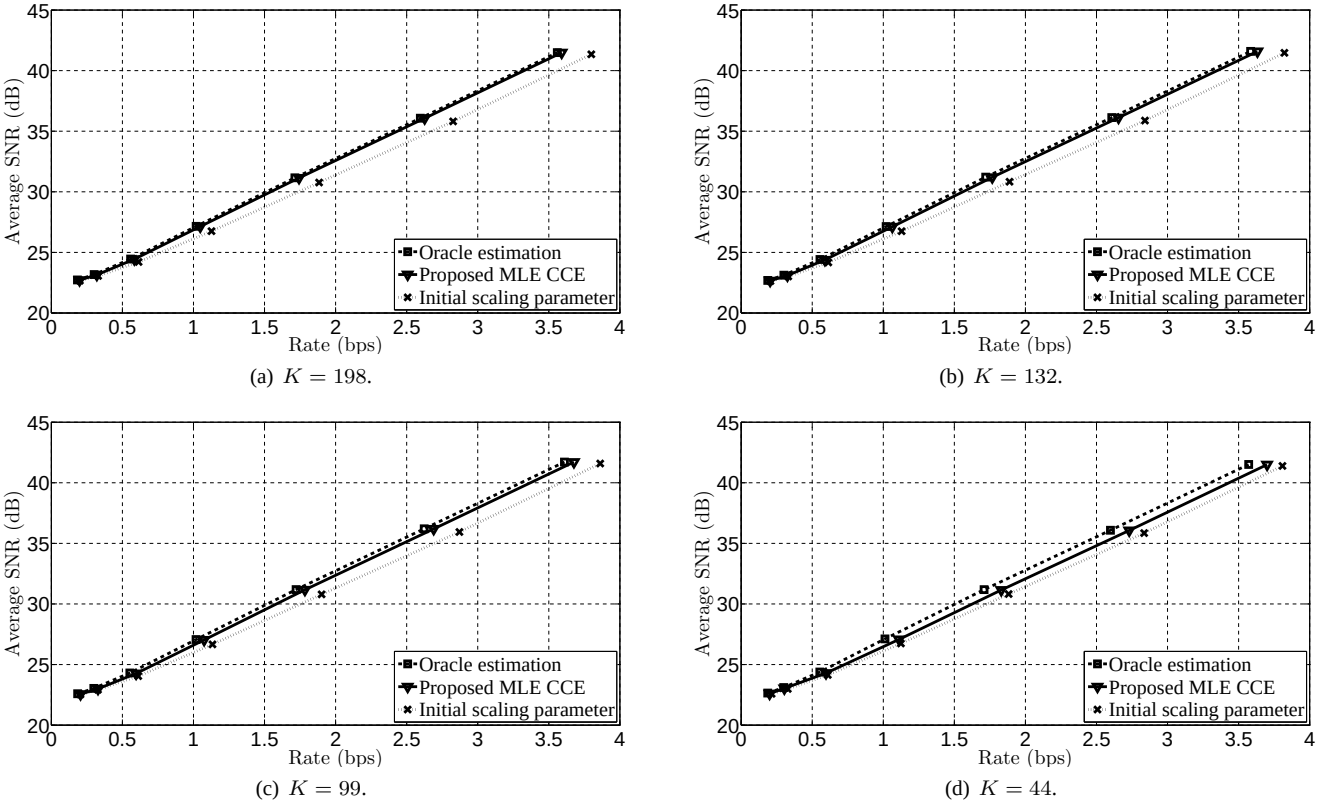
TABLE I
BJØNTEGAARD RATE DELTAS, ΔR(%), ON THE EDVC PERFORMANCE WHEN USING VARIOUS STATIONARITY BLOCK SIZES VERSUS THE PERFORMANCE OBTAINED ASSUMING NOISE STATIONARITY PER BAND PER FRAME.

K	1024	256	64	16	4
$(K_1 \times K_2)$	(32×32)	(16×16)	(8×8)	(4×4)	(2×2)
Capsule Video 1	-0.18	-0.51	-0.99	-1.25	-0.99
Capsule Video 2	-0.03	-0.32	-0.60	-0.88	-0.76
Capsule Video 3	-0.18	-0.41	-0.64	-0.81	-0.48
Capsule Video 4	-0.11	-0.39	-0.59	-0.68	-0.18
Capsule Video 5	-0.20	-0.44	-0.62	-0.85	-0.39
K	384	64	24	16	4
$(K_1 \times K_2)$	(24×16)	(8×8)	(6×4)	(4×4)	(2×2)
Endoscopy Video 1	-0.43	-0.95	-1.14	-1.21	-0.86
Endoscopy Video 2	-0.45	-0.73	-0.85	-0.94	-0.63
Endoscopy Video 3	-0.35	-0.77	-0.92	-1.13	-1.06

rate of 2Hz with a frame resolution of 256×256 pixels. Contrary, each endoscopic sequence contains 100 frames with a resolution of 480×320 pixels at a frame rate of 30Hz. The EDVC system was configured with a GOP of size two and all three Y, U and V components of the sequences were encoded, as detailed in [15].

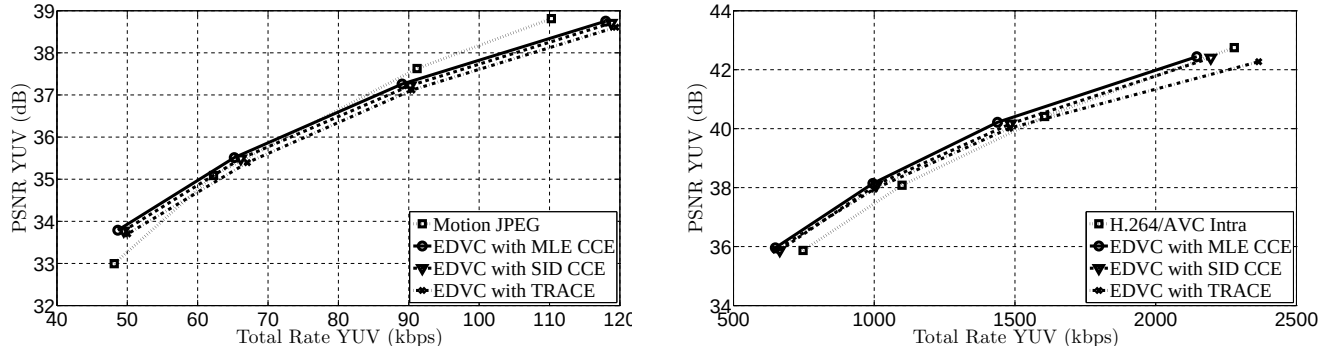
Initially, we study the influence of the stationarity block size on the CCE accuracy and in turn, on the RD performance. The codec's configuration where a Laplace scaling parameter $\hat{\lambda}^{(m)}(\beta)$ is estimated per band of each frame (i.e., $n = K = 64 \times 64 = 4096$ for capsule endoscopic and $n = K = 120 \times 80 = 9600$ for endoscopic sequences) is set as benchmark. Table I depicts the Bjøntegaard Deltas [43] (BD) on the performance of EDVC obtained with different stationarity block sizes, $K_1 \times K_2$, compared to the benchmark configuration. The results corroborate the trade-off between capturing the spatial variations of the noise and maintaining adequate support for the accurate derivation of the MLE. Overall, the best performance is achieved when regulating the stationarity block size to $K = K_1 \times K_2 = 16$ coefficients.

The performance of the EDVC system with the proposed MLE method is thereafter compared with the performances obtained when using alternative up-to-date CCE techniques. The results in Fig. 4 and Table II show that the proposed MLE (using $K = 16$) systematically outperforms the TRACE [25] method and our previous SID CCE [33] technique, with respective BD rate savings of up to 7.13% and 4.22%. TRACE has also been used as a benchmark to evaluate other CCE methods, including [27], [29], [33]. Unlike TRACE and our SID CCE method, the proposed MLE is asymptotically efficient per decoding stage. In addition, contrary to TRACE, the proposed method performs a bit-plane-per-bit-plane refinement of the CCE. This is a feature of other existing CCE methods as well, including [26], [27], [30], [31] and our SID CCE



(a) $K = 198$. (b) $K = 132$. (c) $K = 99$. (d) $K = 44$.

Fig. 3. RD performance of layered WZ coding with the proposed MLE. The side information is Laplacian, while various noise stationarity levels are tested.



(a) (b)

Fig. 4. Compression performance of EDVC using different online CCE methods versus Motion JPEG or H.264/AVC Intra: (a) Capsule Video 1; (b) Endoscopy Video 1. The rate is expressed in kilo-bits-per-second (kbps), while the distortion is quantified using the peak-signal-to-noise-ratio $\left(\text{PSNR}_{\text{YUV}} = \frac{4\times\text{PSNR}_{\text{Y}}+\text{PSNR}_{\text{U}}+\text{PSNR}_{\text{V}}}{6}\right)$. The key and the hash frames of EDVC are encoded with (a) JPEG, or (b) H.264/AVC Intra.

technique. Compared with our SID CCE, however, the proposed MLE can adapt better to small noise stationarity levels, thereby improving the accuracy of the correlation estimation.

The results in Fig. 4(a) show that EDVC, equipped with the proposed MLE technique, yields comparable performance to Motion JPEG despite the highly-complex motion conditions encountered in capsule endoscopic video content. These conditions are due to the erratic motion of the capsule in the gastrointestinal tract and the low acquisition frame rates. Yet, when temporally fluent endoscopic video content is encoded [see Fig. 4(b)], EDVC introduces a notable BD rate gain of 8.80% over H.264/AVC Intra (Main Profile with CABAC).

C. Complexity Assessment

The complexity of the proposed MLE technique is comparatively evaluated using execution time measurements obtained with the software implementation of our EDVC system[7]. This complexity examination methodology is commonly followed in the literature [25], [30], [33]. Table III depicts the execution time per WZ frame for the proposed MLE (configured with $K = 4096$ or $K = 16$) and our previous SID CCE [33] method. The total decoding time and the number of feedback

[7]Experiments were executed on an Intel® Core™ i7 CPU running at 2.20GHz with 16GB of RAM. Our EDVC was written in C++, compiled with Visual Studio 2008 and running in release mode under Windows 7.

TABLE II
BJØNTEGAARD DELTAS ON THE PERFORMANCE OF EDVC WHEN USING THE PROPOSED MLE METHOD VERSUS UP-TO-DATE CCE TECHNIQUES.

	vs. TRACE [25]		vs. SID CCE [33]	
	ΔR(%)	ΔPSNR(dB)	ΔR(%)	ΔPSNR(dB)
Capsule Video 1	-4.36	0.25	-1.82	0.11
Capsule Video 2	-5.15	0.33	-1.48	0.10
Capsule Video 3	-5.70	0.32	-1.89	0.11
Capsule Video 4	-5.04	0.29	-2.96	0.18
Capsule Video 5	-7.13	0.40	-4.22	0.23
Endoscopy Video 1	-4.67	0.24	-3.11	0.17
Endoscopy Video 2	-6.16	0.31	-3.50	0.21
Endoscopy Video 3	-5.05	0.25	-3.87	0.23

TABLE III
DECODING EXECUTION TIME (SEC) AND NUM. OF FEEDBACK REQUESTS PER WZ-FRAME OF THE EDVC SYSTEM CONFIGURED WITH DIFFERENT CCE TECHNIQUES; (A) CAPSULE VIDEO 4, (B) CAPSULE VIDEO 5.

		SID CCE [33]			MLE ($K=4096$)			MLE ($K=16$)		
		CCE time	Dec. time	#FB reqs.	CCE time	Dec. time	#FB reqs.	CCE time	Dec. time	#FB reqs.
(a)	RD pt-1	0.03	27.40	66.5	0.09	25.86	54.60	0.11	25.29	54.37
	RD pt-2	0.23	32.05	96.62	0.46	30.25	86.83	0.59	29.39	85.03
	RD pt-3	0.42	50.43	148.17	0.64	46.72	139.52	0.90	44.03	136.06
	RD pt-4	0.76	62.23	193.15	0.87	57.43	185.00	1.19	52.78	180.75
(b)	RD pt-1	0.03	28.53	68.62	0.08	26.89	58.09	0.10	26.56	56.25
	RD pt-2	0.18	33.74	98.52	0.22	31.40	93.37	0.37	30.77	91.38
	RD pt-3	0.34	52.36	158.08	0.38	48.58	147.71	0.63	46.27	145.50
	RD pt-4	0.72	63.80	204.08	0.79	58.93	196.15	1.10	55.02	193.25

requests per WZ frame of EDVC, configured with each of the assessed CCE techniques, is also provided.

The results indicate that the correlation estimation generally has a minor impact on the total decoding computational demands compared to other components of the decoder, particularly, the side information generation and the LDPCA decoding. Contrasting the complexities of the different CCE methods, it can be observed that, when the noise scaling parameter is estimated at a DCT band level (i.e., $K = 4096$), the proposed MLE induces a complexity comparable to that of our previous SID technique in [33]. Lowering the value of K increases the complexity of the proposed method, as the MLE algorithm is run per stationarity block of each band.

Nevertheless, incorporating the proposed MLE (especially using $K = 16$) reduces significantly the overall decoding complexity of EDVC compared to employing our prior SID CCE method. In effect, as shown in Table III, by improving the CCE accuracy over our prior SID technique, the proposed MLE method induces fewer feedback channel requests and LDPCA decoding iterations. Hence, the feedback channel rate and the associated latency are also reduced.

VII. CONCLUSION

A novel CCE technique tailored to generic layered WZ coding has been proposed in this paper. Unlike existing methods, e.g., [30], [32], [33], the proposed approach derives a novel MLE of the scaling parameter of the Laplacian correlation noise. Adhering to quality scalable WZ coding, the derived estimator is progressively refined per decoding stage. The MSE of the proposed estimator is proven to approximate the derived m-th stage CRLB$^{(m)}$, converging to the oracle CRLB at high decoding stages. It is also shown that the proposed MLE delivers an RD performance close to offline estimation. When incorporated into our EDVC system, the proposed method systematically yields superior RD performance over state-of-the-art techniques. Specifically, when endoscopic video content is encoded, the proposed method brings respective BD [43] rate savings of up to 7.13% and 4.22% over TRACE [25] and our previous SID CCE [33] technique. Finally, as verified by experimentation, the proposed algorithm is computationally efficient. In fact, on account of the improved accuracy in CCE, the LDPCA decoder converges faster, leading to reduced decoding delays that benefit applications like wireless capsule endoscopy.

APPENDIX A
CALCULATION OF $\frac{\partial}{\partial\lambda}\ln p(q_i^{(m)}|y_i;\lambda)$ AND $\frac{\partial^2}{\partial\lambda^2}\ln p(q_i^{(m)}|y_i;\lambda)$

Applying the natural logarithm to the three cases of the pmf in (5) and then taking the first partial derivative with respect to λ yields the three forms of $\frac{\partial}{\partial\lambda}\ln p(q_i^{(m)}|y_i;\lambda)$ shown in (11a), (11b) and (11c), where $\Delta_i^{(m)}$ is the quantization interval of the $i-$th quantization cell, $\theta_i^{(m)} = y_i - u_i^{(m)}$, and $\eta_i^{(m)} = v_i^{(m)} - y_i$. Next, applying the partial derivative with respect to λ on (11a), (11b) and (11c) gives the three forms of $\frac{\partial^2}{\partial\lambda^2}\ln p(q_i^{(m)}|y_i;\lambda)$ expressed by (12a), (12b) and (12c).

APPENDIX B
EXISTENCE AND UNIQUENESS OF THE MLE

Lemma 1: The m-th stage conditional log-likelihood function $\ln\mathcal{L}(\lambda)$ for a Laplace correlation channel is concave at $\lambda \in (0, +\infty)$.

Proof: Taking the second partial derivative of $\ln\mathcal{L}(\lambda)$, defined by (7), with respect to λ, yields

$$\frac{\partial^2}{\partial\lambda^2}\ln\mathcal{L}(\lambda) = \frac{\partial^2}{\partial\lambda^2}\ln\prod_{i=1}^{K} p(q_i^{(m)}|y_i;\lambda) = \sum_{i=1}^{K}\frac{\partial^2}{\partial\lambda^2}\ln p(q_i^{(m)}|y_i;\lambda), \quad (13)$$

where each term $\frac{\partial^2}{\partial\lambda^2}\ln p(q_i^{(m)}|y_i;\lambda)$ is given by (11a), (11b) or (11c) depending on whether $y_i < u_i^{(m)}$, $y_i \geq v_i^{(m)}$ or $u_i^{(m)} \leq y_i < v_i^{(m)}$, respectively. Since $\lambda \in (0, +\infty)$ and $\Delta_i^{(m)} > 0$, the summation terms in (13) that may take the form of (12a) or (12b) are evidently strictly negative. These terms correspond to $(q_i^{(m)}, y_i)$ pairs for which $y_i < u_i^{(m)}$ or $y_i \geq v_i^{(m)}$. We now turn to the terms that may follow the form of (12c). Since for these terms $u_i^{(m)} \leq y_i < v_i^{(m)}$, then by definition $\theta_i^{(m)} \geq 0$ and $\eta_i^{(m)} > 0$, meaning that $0 \leq e^{-\lambda\theta_i^{(m)}} \leq 1$ and $0 \leq e^{-\lambda\eta_i^{(m)}} < 1$. As a consequence, for the denominator of the first fraction in (12c) it holds that $2 - e^{-\lambda\theta_i^{(m)}} - e^{-\lambda\eta_i^{(m)}} > 0$. Namely, the terms in (13) that may follow (12c) are also strictly negative. As a consequence, $\frac{\partial^2}{\partial\lambda^2}\ln p(\mathbf{q}^{(m)}|\mathbf{y};\lambda) < 0$. Therefore, according to the second derivative test, $\ln p(\mathbf{q}^{(m)}|\mathbf{y};\lambda)$ is concave at $\lambda \in (0, +\infty)$. ■

Proof of Theorem 1: From Lemma 1, the function $\frac{\partial}{\partial\lambda}\ln p(\mathbf{q}^{(m)}|\mathbf{y};\lambda)$ is strictly monotonically decreasing at $\lambda\in(0,+\infty)$. Furthermore, the one-sided limit of the function at zero is

$$\lim_{\lambda\to 0^+}\frac{\partial}{\partial\lambda}\ln p(\mathbf{q}^{(m)}|\mathbf{y};\lambda) = \sum_{i=1}^{K}\lim_{\lambda\to 0^+}\frac{\partial}{\partial\lambda}\ln p(q_i^{(m)}|y_i;\lambda) = +\infty, \quad (14)$$

since $\lim_{\lambda\to 0^+}\frac{\partial}{\partial\lambda}\ln p(q_i^{(m)}|y_i;\lambda) = +\infty$ for all possible forms of $\frac{\partial}{\partial\lambda}\ln p(q_i^{(m)}|y_i;\lambda)$, given by (11a), (11b) or (11c).

Concerning the limit at infinity, we have: If $y_i < u_i^{(m)}$, then $\lim_{\lambda\to+\infty}\frac{\partial}{\partial\lambda}\ln p(q_i^{(m)}|y_i;\lambda) = \theta_i^{(m)} < 0$. Alternatively, if $y_i > v_i^{(m)}$, then $\lim_{\lambda\to+\infty}\frac{\partial}{\partial\lambda}\ln p(q_i^{(m)}|y_i;\lambda) = \eta_i^{(m)} < 0$. If $u_i^{(m)} \le y_i \le v_i^{(m)}$, then $\lim_{\lambda\to+\infty}\frac{\partial}{\partial\lambda}\ln p(q_i^{(m)}|y_i;\lambda) = 0$. Hence, if there exists at least one side information sample y_i in $\mathbf{y}$ for which $y_i < u_i^{(m)}$ or $y_i > v_i^{(m)}$, then

$$\lim_{\lambda\to+\infty}\frac{\partial}{\partial\lambda}\ln p(\mathbf{q}^{(m)}|\mathbf{y};\lambda) = \sum_{i=1}^{K}\lim_{\lambda\to+\infty}\frac{\partial}{\partial\lambda}\ln p(q_i^{(m)}|y_i;\lambda) < 0, \quad (15)$$

Combining (14) and (15) with the fact that $\frac{\partial}{\partial\lambda}\ln p(\mathbf{q}^{(m)}|\mathbf{y};\lambda)$ is continuous and strictly monotonically decreasing proves that the equation $\frac{\partial}{\partial\lambda}\ln p(\mathbf{q}^{(m)}|\mathbf{y};\lambda) = 0$ has a unique and finite solution at $\lambda\in(0,+\infty)$. ■

APPENDIX C
ANALYTICAL SOLUTIONS OF THE MLE

Proof of Proposition 1: Provided that $y_i \notin \left[u_i^{(m)}, v_i^{(m)}\right), \forall i = \{1,2,...,K\}$, one can divide the pairs $(q_i^{(m)}, y_i)$, $i\in\{1,2,...,K\}$, in two disjoint sets S_A and S_B, corresponding to the cases $y_i < u_i^{(m)}$ and $y_i \ge v_i^{(m)}$, respectively. Under this condition, (8) can be developed as

$$\sum_{i\in S_A}\left[y_i - v_i^{(m)} + \frac{\Delta_i^{(m)}}{1-e^{-\lambda\Delta_i^{(m)}}}\right] + \sum_{i\in S_B}\left[v_i^{(m)} - y_i + \frac{\Delta_i^{(m)}}{e^{\lambda\Delta_i^{(m)}}-1}\right] = 0, \quad (16)$$

Given that the quantizer Q_m has uniform intervals, namely, $\Delta_i^{(m)} = \Delta^{(m)}$, $\forall i\in\{1,2,..,K\}$, then equation (16) becomes

$$\frac{|S_A|\Delta^{(m)}}{1-e^{-\lambda\Delta^{(m)}}} + \frac{|S_B|\Delta^{(m)}}{e^{\lambda\Delta^{(m)}}-1} = \sum_{i\in S_A}\left[v_i^{(m)} - y_i\right] + \sum_{i\in S_B}\left[y_i - v_i^{(m)}\right], \quad (17)$$

where $|S_A|$, $|S_B|$ denote the cardinalities of the sets S_A and S_B. It holds that: $|S_A| + |S_B| = K$. Next, solving (17) for λ gives

$$\hat{\lambda}^{(m)} = \frac{1}{\Delta^{(m)}} \times \ln\frac{\sum_{i\in S_A}\left[v_i^{(m)} - y_i\right] + \sum_{i\in S_B}\left[y_i - v_i^{(m)}\right] + |S_B|\Delta^{(m)}}{\sum_{i\in S_A}\left[v_i^{(m)} - y_i\right] + \sum_{i\in S_B}\left[y_i - v_i^{(m)}\right] - |S_A|\Delta^{(m)}}. \quad (18)$$

$$\frac{\partial}{\partial\lambda}\ln p(q_i^{(m)}|y_i;\lambda) = \begin{cases} y_i - v_i^{(m)} + \dfrac{\Delta_i^{(m)}}{1-e^{-\lambda\Delta_i^{(m)}}}, & y_i < u_i^{(m)}, \quad (11a)\\ v_i^{(m)} - y_i + \dfrac{\Delta_i^{(m)}}{e^{\lambda\Delta_i^{(m)}}-1}, & y_i \ge v_i^{(m)}, \quad (11b)\\ \dfrac{\theta_i^{(m)}e^{-\lambda\theta_i^{(m)}} + \eta_i^{(m)}e^{-\lambda\eta_i^{(m)}}}{2-e^{-\lambda\theta_i^{(m)}} - e^{-\lambda\eta_i^{(m)}}}, & u_i^{(m)} \le y_i^{(m)} < v_i^{(m)}. \quad (11c)\end{cases}$$

$$\frac{\partial^2}{\partial\lambda^2}\ln p(q_i^{(m)}|y_i;\lambda) = \begin{cases} -\dfrac{\left[\Delta_i^{(m)}\right]^2 e^{-\lambda\Delta_i^{(m)}}}{\left[1-e^{-\lambda\Delta_i^{(m)}}\right]^2}, & y_i < u_i^{(m)}, \quad (12a)\\ -\dfrac{\left[\Delta_i^{(m)}\right]^2 e^{\lambda\Delta_i^{(m)}}}{\left[e^{\lambda\Delta_i^{(m)}}-1\right]^2}, & y_i \ge v_i^{(m)}, \quad (12b)\\ -\dfrac{\left[\theta_i^{(m)}\right]^2 e^{-\lambda\theta_i^{(m)}} + \left[\eta_i^{(m)}\right]^2 e^{-\lambda\eta_i^{(m)}}}{2-e^{-\lambda\theta_i^{(m)}} - e^{-\lambda\eta_i^{(m)}}} - \left[\dfrac{\theta_i^{(m)}e^{-\lambda\theta_i^{(m)}} + \eta_i^{(m)}e^{-\lambda\eta_i^{(m)}}}{2-e^{-\lambda\theta_i^{(m)}} - e^{-\lambda\eta_i^{(m)}}}\right]^2, & \\ & u_i^{(m)} \le y_i^{(m)} < v_i^{(m)}. \quad (12c)\end{cases}$$

Alternatively, if $|S_B| = 0$ or $|S_A| = 0$, namely, if $y_i < u_i^{(m)}$ or $y_i \geq v_i^{(m)}$, $\forall i \in \{1, 2, ..., K\}$ then

$$\hat{\lambda}^{(m)} = \frac{1}{\Delta^{(m)}} \ln \left[1 + \frac{K\Delta^{(m)}}{\sum_{i=1}^{K} \left[y_i - v_i^{(m)} \right]} \right]^{\mp 1}, \qquad (19)$$

where the $-$ and $+$ on the exponent corresponds to the first and second case, respectively. ∎

CRLB$^{(m)}$ for the Analytical Forms: Provided that the conditions listed in Proposition 1 hold, then replacing the second derivative forms of (12a) and (12b) in (10) gives

$$\text{CRLB}^{(m)} = \left[\frac{|S_A|\Delta^{(m)} e^{-\lambda\Delta^{(m)}}}{\left(1 - e^{-\lambda\Delta^{(m)}}\right)^2} + \frac{|S_B|\Delta^{(m)} e^{\lambda\Delta^{(m)}}}{\left(e^{\lambda\Delta^{(m)}} - 1\right)^2} \right]^{-1}. \qquad (20)$$

Setting $|S_A| = K$ and $|S_B| = 0$ or $|S_B| = K$ and $|S_A| = 0$ in (20) yields the respective CRLB$^{(m)}$ for the case $y_i < u_i^{(m)}$ or $y_i \geq v_i^{(m)}$, $\forall i \in \{1, 2, ..., K\}$. ∎

ACKNOWLEDGMENT

The paper is dedicated to the memory of Prof. Theodore L. Deliyannis, a respected colleague and beloved uncle and mentor.

REFERENCES

[1] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, Apr. 1973.

[2] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[3] Y. Steinberg and N. Merhav, "On successive refinement for the Wyner-Ziv problem," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1636–1654, Aug. 2004.

[4] S. Cheng and Z. Xiong, "Successive refinement for the Wyner-Ziv problem and layered code design," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3269–3281, Aug. 2005.

[5] D. Varodayan, A. Aaron, and B. Girod, "Rate-adaptive codes for distributed source coding," *Signal Process.*, vol. 86, no. 11, pp. 3123–3130, Nov. 2006.

[6] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proc. IEEE*, vol. 93, no. 1, pp. 71–83, Jan. 2005.

[7] V. Stankovic, A. D. Liveris, Z. Xiong, and C. N. Georghiades, "On code design for the Slepian-Wolf problem and lossless multiterminal networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1495–1507, Jun. 2006.

[8] M. Grangetto, E. Magli, and G. Olmo, "Distributed arithmetic coding for the Slepian-Wolf problem," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2245–2257, Jun. 2009.

[9] S. Malinowski, X. Artigas, C. Guillemot, and L. Torres, "Distributed coding using punctured quasi-arithmetic codes for memory and memoryless sources," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4154–4158, Oct. 2009.

[10] S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 626–643, Mar. 2003.

[11] Z. Liu, S. Cheng, A. D. Liveris, and Z. Xiong, "Slepian-Wolf coded nested lattice quantization for Wyner-Ziv coding: High-rate performance analysis and code design," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4358–4379, Oct. 2006.

[12] R. Puri, A. Majumdar, P. Ishwar, and K. Ramchandran, "Distributed video coding in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 23, no. 4, pp. 94–106, Jul. 2006.

[13] M. Tagliasacchi, G. Valenzise, and S. Tubaro, "Hash-based identification of sparse image tampering," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2491–2504, Nov. 2009.

[14] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov, and M. Quaret, "The DISCOVER codec: Architecture, techniques and evaluation," in *Picture Coding Symp. (PCS)*, Nov. 2007.

[15] N. Deligiannis, F. Verbist, A. Iossifides, J. Slowack, R. Van de Walle, P. Schelkens, and A. Munteanu, "Wyner-Ziv video coding for wireless lightweight multimedia applications," *EURASIP J. Wireless Communications and Networking*, no. 106, 2012.

[16] A. Saxena and K. Rose, "On scalable distributed coding of correlated sources," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2875–2883, Jul. 2010.

[17] Y. Fang, "Crossover probability estimation using mean-intrinsic-LLR of LDPC syndrome," *IEEE Commun. Lett.*, vol. 13, no. 9, pp. 679–681, Sep. 2009.

[18] J. Garcia-Frias and Y. Zhao, "Compression of correlated binary sources using Turbo codes," *IEEE Commun. Lett.*, vol. 5, no. 10, pp. 417–419, Oct. 2001.

[19] A. Zia, J. Reilly, and S. Shirani, "Distributed parameter estimation with side information: A factor graph approach," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2007, pp. 2556–2560.

[20] V. Toto-Zarasoa, A. Roumy, and C. Guillemot, "Maximum likelihood BSC parameter estimation for the Slepian-Wolf problem," *IEEE Commun. Lett.*, vol. 15, no. 2, pp. 232–234, Feb. 2011.

[21] L. Cui, S. Wang, S. Cheng, and M. B. Yeary, "Adaptive binary Slepian-Wolf decoding using particle based belief propagation," *IEEE Trans. Commun.*, vol. 59, no. 9, pp. 2337–2342, Sep. 2011.

[22] L. Cui, S. Wang, and S. Cheng, "Adaptive Slepian-Wolf decoding based on expectation propagation," *IEEE Commun. Lett.*, vol. 16, no. 2, pp. 252–255, Feb. 2012.

[23] C. Brites and F. Pereira, "Correlation noise modeling for efficient pixel and transform domain Wyner-Ziv video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 9, pp. 1177–1190, Sep. 2008.

[24] J. Škorupa, J. Slowack, S. Mys, N. Deligiannis, J. D. Cock, P. Lambert, A. Munteanu, and R. V. de Walle, "Exploiting quantization and spatial correlation in virtual-noise modeling for distributed video coding," *Sig. Proc.: Image Comm.*, vol. 25, no. 9, pp. 674–686, Oct. 2010.

[25] X. Fan, O. C. Au, and N. M. Cheung, "Transform-domain adaptive correlation estimation (TRACE) for Wyner-Ziv video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 11, pp. 1423–1436, Nov. 2010.

[26] X. Huang and S. Forchhammer, "Transform domain Wyner-Ziv video coding with refinement of noise residue and side information," in *Proc. of SPIE*, vol. 7744, 2010, pp. 774 418–1.

[27] ——, "Cross-band noise model refinement for transform domain Wyner-Ziv video coding," *Sig. Proc.: Image Comm.*, vol. 27, no. 1, pp. 16–30, Jan. 2012.

[28] G. R. Esmaili and P. C. Cosman, "Wyner-Ziv video coding with classified correlation noise estimation and key frame coding mode selection," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2463–2474, Sep. 2011.

[29] H. Luong, L. Rakêt, X. Huang, and S. Forchhammer, "Side information and noise learning for distributed video coding using optical flow and clustering," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4782–4796, Dec. 2012.

[30] S. Wang, L. Cui, L. Stankovic, V. Stankovic, and S. Cheng, "Adaptive correlation estimation with particle filtering for distributed video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 5, pp. 649–658, May 2012.

[31] H. Van Luong, X. Huang, and S. Forchhammer, "Parallel iterative decoding of transform domain Wyner-Ziv video using cross bitplane correlation," in *IEEE Int. Conf. Image Process. (ICIP)*, 2011, pp. 2633–2636.

[32] V. Toto-Zarasoa, A. Roumy, and C. Guillemot, "Source modeling for distributed video coding," *IEEE Trans. Circuits Syst. Video Techn.*, vol. 22, no. 2, pp. 174–187, Feb. 2012.

[33] N. Deligiannis, J. Barbarien, M. Jacobs, A. Munteanu, A. Skodras, and P. Schelkens, "Side-information-dependent correlation channel estimation in hash-based distributed video coding," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1934–1949, Apr. 2012.

[34] D. Rebollo-Monedero, S. Rane, A. Aaron, and B. Girod, "High-rate quantization and transform coding with side information at the decoder," *Signal Process., Special Issue on Distributed Source Coding*, vol. 86, no. 11, pp. 3160–3179, Nov. 2006.

[35] S. Chung, T. Richardson, and R. Urbanke, "Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 657–670, Feb. 2001.

[36] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, ser. Wiley Series in Telecommunications. New York, USA: Wiley, 1991.
[37] D. Kubasov, J. Nayak, and C. Guillemot, "Optimal reconstruction in Wyner-Ziv video coding with multiple side information," in *IEEE Workshop Multimedia Signal Process. (MMSP)*, 2007, pp. 251–254.
[38] S. M. Kay, *Fundamentals of statistical signal processing: Estimation theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.
[39] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C*. Cambridge: Cambridge University Press, 1992.
[40] J. Ascenso, C. Brites, and F. Pereira, "A flexible side information generation framework for distributed video coding," *Multimedia Tools and Applications*, vol. 48, no. 3, pp. 381–409, 2010.
[41] N. Deligiannis, M. Jacobs, J. Barbarien, F. Verbist, J. Skorupa, R. Van De Walle, A. Skodras, P. Schelkens, and A. Munteanu, "Joint dc coefficient band decoding and motion estimation in wyner-ziv video coding," in *Int. Conf. Digital Signal Process. (DSP)*, 2011, pp. 1–6.
[42] C. Mc Caffrey, O. Chevalerias, C. O'Mathuna, and K. Twomey, "Swallowable-capsule technology," *IEEE Pervasive Computing*, vol. 7, no. 1, pp. 23–29, Jan.-Mar. 2008.
[43] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves," ITU-T Video Coding Experts Group (VCEG), Tech. Rep. Document VCEG-M33, Apr. 2001.

Nikos Deligiannis (S'08, M'10) received the Diploma degree in electrical and computer engineering from University of Patras, Greece, in 2006, and the Ph.D. degree in engineering (awarded with the highest distinction and congratulations of the jury members) from Vrije Universiteit Brussel (VUB), Brussels, Belgium, in 2012.

From June 2012 to September 2013, he was a post-doctoral researcher with the Department of Electronics and Informatics (ETRO), VUB and a senior research engineer with the 4Media group of the iMinds research institute, Ghent, Belgium. In October 2013, he joined the Department of Electronic and Electrical Engineering at University College London, where he is currently working as research associate. His research interests include multiterminal communications, distributed source coding, wireless sensor networks, image/video coding and location positioning algorithms.

Dr. Deligiannis was the co-recipient of the 2011 ACM/IEEE International Conference on Distributed Smart Cameras Best Paper Award and the recipient of the 2013 Scientific Prize IBM–FWO Belgium for his PhD work.

Adrian Munteanu (M'07) received the M.Sc. degree in electronics from Politehnica University of Bucharest, Bucharest, Romania, in 1994; the M.Sc. degree in biomedical engineering from University of Patras, Patras, Greece, in 1996; and the Ph.D. degree in applied sciences (awarded with the highest distinction and congratulations of the jury members) from Vrije Universiteit Brussel (VUB), Brussel, Belgium, in 2003.

From 2004 to 2010, he was a postdoctoral fellow with the Fund for Scientific Research, Flanders, Belgium. Since 2007, he has been a professor with VUB. He is also a Research Leader of the 4Media group with the iMinds research institute, Ghent, Belgium. He has published more than 200 articles in journals and conference proceedings, book chapters, patent applications and contributions to standards. His research interests include scalable image and video coding, distributed source coding, scalable coding of 3-D graphics, 3-D video coding, error-resilient coding, multiresolution image and video analysis, video segmentation and indexing, multimedia transmission over networks, and statistical modeling.

He was the recipient of the 2004 BARCO-FWO prize for his Ph.D. work. Dr. Munteanu is currently Associate Editor of IEEE Transactions on Multimedia.

Shuang Wang (S'08, M'12) received the B.S. degree in applied physics and the M.S. degree in biomedical engineering from the Dalian University of Technology, China, and the Ph.D. degree in electrical and computer engineering from the University of Oklahoma, Tulsa, OK, in 2012.

He worked with Soft Imaging LLC, Houston, TX, in the areas of digital watermarking for Biomedical Image and GPU-based next-generation parallel genome sequencing algorithm during the summers of 2008 and 2009. He was working as a graduate assistant and a Ph.D. candidate under Dr. Samuel Cheng in OU-Tulsa campus. Currently, he is a Postdoctoral researcher at Division of Biomedical Informatics (DBMI), University of California, San Diego (UCSD). His research interests include information theory, image/signal processing, machine learning, and privacy-preserving algorithms.

He has one patent submission in the area of channel coding. He was awarded the Robert Hughes Centennial Fellowship in the Fall 2009.

Samuel Cheng received the B.S. degree in Electrical and Electronic Engineering from the University of Hong Kong, and the M.Phil. degree in Physics and the M.S. degree in Electrical Engineering from Hong Kong University of Science and Technology and the University of Hawaii, Honolulu, respectively. He received the Ph.D. degree in Electrical Engineering from Texas A&M University in 2004.

He worked in Microsoft Asia, China, and Panasonic Technologies Company, New Jersey, during the summers of 2000 and 2001. In 2004, he joined Advanced Digital Imaging Research, a research company based near Houston, Texas, as a Research Engineer. In 2006, he joined the School of Electrical and Computer Engineering at the University of Oklahoma and is currently an associate professor. His research interests include distributed source coding and image/signal processing

Peter Schelkens (M'97) received the degree in electronic engineering in VLSI-design from the Industrile Hogeschool Antwerpen-Mechelen (IHAM), Campus Mechelen. He then received the M.Sc. degree in applied physics, a biomedical engineering degree (medical physics), and the Ph.D. degree in applied sciences from the Vrije Universiteit Brussel (VUB).

He currently holds a professorship with the Department of Electronics and Informatics (ETRO), VUB. He is a member of the scientific staff of the Interdisciplinary Institute for Broadband Technology, Belgium. Additionally, since 1995, he has also been affiliated to the Interuniversity Microelectronics Institute, Belgium, as scientific collaborator. His research interests are situated in the field of multidimensional signal processing encompassing the representation, communication, security, and rendering of these signals while especially focusing on cross-disciplinary research. He is a coeditor of the books The JPEG 2000 Suite (New York:Wiley, 2009) and Optical and Digital Image Processing (New York: Wiley, 2011). He has published more than 200 papers in journals and conference proceedings, and holds several patents, as well as contributed to several standardization processes.

Dr. Schelkens is the Belgian head of delegation for the ISO/IEC JPEG standardization committee, editor/chair of part 10 of JPEG 2000: Extensions for Three-Dimensional Data and PR Chair of the JPEG committee. Since 2010, he has been a member of the Board of Councilors of the Interuniversity Microelectronics Institute. He is a member of SPIE and ACM, and is currently the Belgian EURASIP Liaison Officer. In 2011, he acted as General co-Chair of IEEE International Conference on Image Processing (ICIP) and the Workshop on Quality of Multimedia Experience (QoMEX).