Distributed Successive Refinement of Multiview Images Using Broadcast Advantage

Zichong Chen, *Student Member, IEEE*, Guillermo Barrenetxea, and Martin Vetterli, *Fellow, IEEE*

Abstract—In environmental monitoring applications, having multiple cameras focus on common scenery increases robustness of the system. To save energy based on user demand, successive refinement image coding is important, as it allows us to progressively request better image quality. By exploiting the broadcast nature and correlation between multiview images, we investigate a two-camera setup and propose a novel two-encoder successive refinement scheme which imitates a ping-pong game. For the bivariate Gaussian case, we prove that this scheme is successively refinable on the theoretical rate-distortion limit of distributed coding (*Wagner surface*) under arbitrary settings. For stereo-view images, we develop a practical successive refinement coding algorithm using the same idea. The simulation results show that this scheme operates close to the distributed coding bound.

Index Terms—Energy-constrained, error resilient, hazard monitoring, source channel coding, video coding.

I. INTRODUCTION

THANKS to the availability of low-power image sensor chips, such as CMOS cameras, wireless sensor networks are extending their functionalities to image monitoring. However, for visual applications to become ubiquitous, we need to tackle the problem of processing and transmitting large amounts of data, under the severe energy constraints of wireless sensor networks. In the context of environmental monitoring, robustness of the system is also an important concern as the camera itself is experiencing unpredictable weather conditions and consequently can be affected by surrounding major events. A multi-camera wireless network can provide a solution to these two main concerns.

1) Robustness: By placing multiple cameras in different locations observing a common scene of interest, we generate correlated multi-view images. In case of camera failure, as long as there is still one surviving, we will not lose the visual access to the scene of interest.

2) Energy saving: Multiple cameras can reduce per-camera energy consumption by sharing the sampling and communication task among all cameras [1].

Figure 1(a) shows the typical setup of a wireless camera monitoring network (in this paper we focus on the two-camera case): two cameras are deployed to monitor an overlapped scenery, and communicate independently with the base station (BS) to deliver the captured stereo-view images. Direct communication between two cameras is not present[1], but each camera overhears the information transmitted by the other one.

With this setup, an interesting problem arises: if two cameras need to transmit the stereo-view images X and Y to the BS, can we separate the transmission into several phases, so that the BS can recover the image X and Y with progressively increasing quality (see Figure 2). This is particularly useful in an energy-limited communication scenario, because we can decide at the receiver whether a high resolution image is really needed after the low resolution version is displayed. Such an idea is called successive refinement and the single source case has been studied extensively in both theory and practical schemes (see Section II-B). Several theoretical variants of the multiple source case have been recently investigated (see Section II-A), but there is hardly any practical successive refinement schemes for multi-view images.

In this paper, we develop our prior work [2] and propose a novel two-encoder successive refinement scheme that imitates the ping-pong game and takes advantage of the broadcast nature of wireless communications. The theoretical analysis for the bivariate Gaussian case shows that this scheme has no loss in coding efficiency as compared to the conventional distributed coding limit, for arbitrary successive refinement settings (any number of stages and rate combinations). We then apply this scheme to real image data and propose a practical stereo-view image coding algorithm capable of successive refinement. The simulation results show that this scheme operates close to the distributed coding bound.

The paper is organized as follows: In Section II, we review the background and related work on successive refinement and distributed coding. In Section III, we describe the broadcast nature and inter-view correlation model as the main assumptions in this paper. In Section IV, we propose a new ping-pong

[1]If there is direct communication between the cameras, then we do not have a distributed scenario anymore.
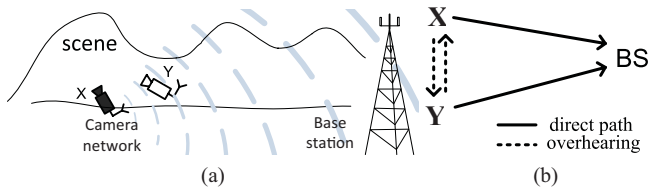
Fig. 1. Two cameras X and Y are deployed to monitor an overlapping scenery, and communicate independently with the BS. When the active camera X is transmitting an image to the BS, the other camera can overhear the same information. (a) Illustration of the setup. (b) Abstract model.
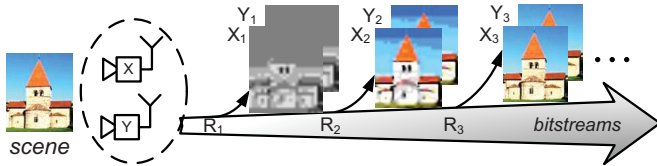

Fig. 2. Successive refinement of stereo-view images. Bitstreams are transmitted in several stages and the BS can recover X and Y with increasing quality.

like two-encoder successive refinement scheme and present the theoretical analysis for the Gaussian quadratic case. This leads to Theorem 1, which proves the successive refinability of this scheme on the distributed rate-distortion limit. Section V applies this idea to stereo-view images and presents a practical stereo-view image coding algorithm capable of successive refinement. The simulation results then show the good performance of the scheme, close to the distributed coding bound.

II. BACKGROUND AND RELATED WORK

A. Distributed Source Coding

In a typical multi-camera network, two cameras transmit images independently to the base station (BS), which can be well modeled as a *two-encoder distributed source coding (DSC2)* setup. When broadcast is employed, a passive communication link exists as well. Then partially separated encoders provides a better model (see Figure 3).

To deliver images under the severe energy constraints of wireless networks, multi-terminal source coding has an important role as it can cut the rate to the theoretical lower limit. The general multi-terminal source coding problem [3] has been posed more than thirty years ago. The rate region for the distributed lossless source coding problem has been solved by Slepian and Wolf [4]. However, the general lossy case is not fully determined yet. Wyner and Ziv [5] solved a special case when one of two sources is entirely known at the decoder. Recently, Wagner et al. [6] gave the rate-distortion region for the two-encoder quadratic Gaussian case. To give an illustration of Wagner's four-dimensional rate-distortion region (R_X, R_Y, D_X, D_Y), we pick the sum-rate $R_X + R_Y$ as a measure of the rate budget and show the coding limits in a three-dimensional space $(R_X + R_Y, D_X, D_Y)$. The surface (*Wagner Surface*) as defined by Eq. (3) of [6] represents the rate distortion limits that any distributed coding scheme should operate on or above.
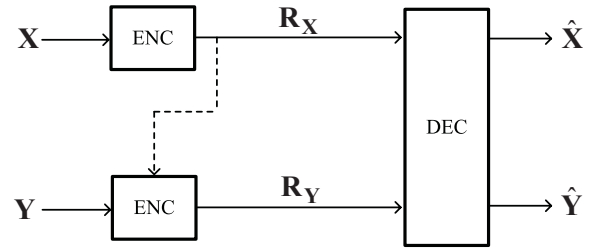

Fig. 3. Distributed source coding of two partial separated encoders. Overhearing provides a passive communication link (dashed line).

Practical distributed single-view video coding schemes [7], [8], and distributed multi-view video coding schemes [9] have been developed in recent years. These schemes require no inter-frame/inter-view communication, but are also very restricted to strong correlation between frames/views. For instance, if the video frame rate decreases or there are occlusion between different views, the coding efficiency drops very quickly.

Distributed source coding using the broadcast advantage is a particular form of the source coding problem with partially separated encoders. The rate-distortion region for two-encoder case was first addressed in [10], where coding theorems are determined for two cases: 1) one source is reproduced perfectly at the receiver; and 2) one source is perfectly revealed to the other source. [11] and [12] further develop this idea in a lossless Slepian-Wolf setup when an encoder can observe the codeword from the other encoder. It is proved that the admissible rate region is not enlarged. However, the general rate-distortion region for a lossy setup as in Figure 3 is still unknown today.

B. Successive Refinement Coding

Comparing with conventional coding methods, successive refinement splits the single codeword into multiple pieces and make it possible to gradually send/reconstruct source(s) with increasing quality. From the rate-distortion perspective, the rate-distortion $R(D)$ of the given source(s) characterizes the performance limit of successive refinement coding: As depicted in Figure 4 for a single source, (R_1, D_1) and $(R_1 + R_2, D_2)$ are the R-D pairs achieved at the 1st and 2nd stage respectively. Generally, for most source types, successive refinement operates above the $R(D)$ curve. If (R_1, D_1) and $(R_1 + R_2, D_2)$ both operate exactly on the $R(D)$ curve of a source, for any (R_1, R_2), then we say that successive refinement of this source is optimal (*successive refinability*).

Equitz and Cover [13] gave the necessary and sufficient condition for a single source to be successively refinable: To encode source X with rate-distortion function $R(D)$, a coarse description $\hat{X}_1$ with R-D pair (R_1, D_1) is refined to a finer description $\hat{X}_2$ with a rate of R_2. The distortion of $\hat{X}_2$ is denoted as D_2. The optimality, or $R_1 = R(D_1)$ and $R_1 + R_2 = R(D_2)$ are both achievable, is obtained if and only if we can write $\hat{X}_1 \rightarrow \hat{X}_2 \rightarrow X$ as a Markov chain. A Gaussian source with squared-error distortion is one example that satisfies this Markovian condition.
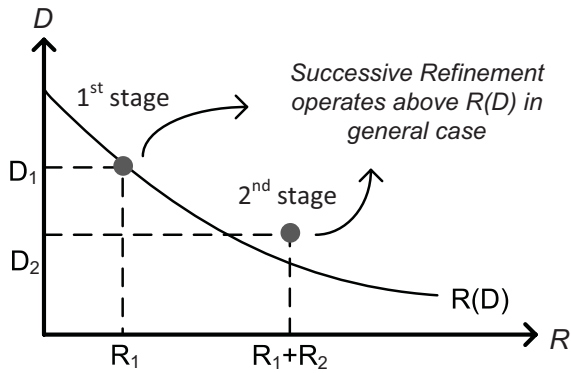
Fig. 4. Rate distortion performance: two-stage successive refinement of a single source.
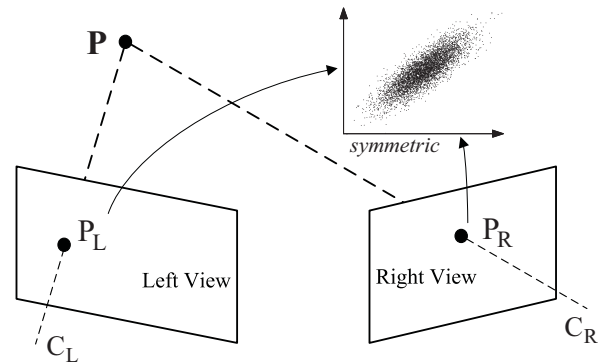

Fig. 5. Intensities of point correspondence (P_L and P_R) of a given point P on stereo-views are statistically linked by a certain distribution with equal marginal distributions.

On the question of successive refinement for multiple sources, [14] proposes a sequential coding of correlated sources for video applications, in which the first source is encoded solely while the subsequent source is encoded based on both sources. This scheme is a weak version of centralized coding as it has access to both sources. However, as it does not fully exploit the joint information due to the first step encoding, the minimum sum-rate is sometimes worse than DSC2. Recently, [15] proposed a successive decoding scheme for the distributed source coding problem (no link between encoders). They prove that successive decoding following a linear fusion achieves the rate-distortion region of DSC2 for the quadratic Gaussian case. However, the final step of fusion actually breaks the successive decoding structure: all results have to be reconstructed after everything is received.

In practical scenarios, images are usually not well modeled by Gaussian sources, thus it is non-trivial to say if a certain image is successively refinable or not. Most of the contemporary single image coders support progressive or scalable coding (synonyms for successive refinement). In particular, the JPEG2000 image coding standard is designed to be scalable in nature [16]: it decomposes the bitstreams into a succession of layers, and each layer contains additional contributions optimized for rate-distortion performance. Therefore, the layer decomposition provides an approximation of successive refinement, as long as the bitstream is truncated at a layer point.

III. MODEL AND ASSUMPTIONS

In this section, we present the image and communication model which leads to a distributed successive refinement coding scheme.

A. Inter-View Correlation Model

The correspondence between two stereo-view images can be completely described by epipolar geometry [17] using the pinhole camera model. Illustrated in Figure 5, a point P in real world is projected onto left and right image planes, and the corresponding projection points are P_L and P_R respectively. The intensities of P_L and P_R depend on the intensity of P (if P lies on a lambertian surface), the light path from P to the image plane, and the optical system of the camera.

Considering the fact that the light path in air is homogeneous and the imaging system of two cameras are almost identical under proper calibration, the intensities of P_L and P_R are expected to be highly correlated. Therefore, any correspondence point pair determined by the epipolar geometry can be statistically modeled by a certain bivariate distribution. Note that this probability distribution should have equal marginal distributions due to the symmetry of the left/right views. If we further assume that such a model is globally *stationary* over the entire image, then the inter-view correlation can be fully described by a single distribution model.

B. Broadcast Advantage

In a typical wireless camera network, inter-camera communication is usually not implemented as it consumes extra energy, but there are potentially existing *passive* links [18]. As wireless systems often use omnidirectional antennas, the cameras within the transmission range of the emitting camera can actually overhear the transmitted messages. To take advantage of the broadcast nature of wireless communication, our proposed scheme is based on several assumptions that are often verified in practice: 1) The scale of the local camera network is small compared with their distance to the BS (see Figure 1a), thus the channel capacity between neighboring cameras is larger than the capacity between the camera and the BS. For simplicity, we assume that the overhearing camera within the transmission range can get the message error-free, as long as this message is received at the BS. 2) Both cameras are synchronized and use time-division multiplexing so that only one camera transmits at a time while the other overhears the broadcast message without interference.

IV. DISTRIBUTED SUCCESSIVE REFINEMENT OF BIVARIATE GAUSSIAN SOURCES USING BROADCAST ADVANTAGE

In this section, we investigate a theoretical framework of distributed successive refinement. Based on the assumptions in Section III, we propose a coding scheme that imitates the ping-pong game, and characterize the corresponding rate-distortion limit when two distributed sources are statistically linked by a jointly bivariate Gaussian model.

A. Setup and Coding Procedure

The broadcast nature of wireless communication provides a free overhearing mechanism that can be exploited to reduce the transmission rate between the cameras and the BS. For instance, in a two-encoder setup as in Figure 3, based on the codeword sent by source X, source Y can be encoded at a lower rate by exploiting the correlation between X and Y.

Inspired by the ping-pong game, we extend such an idea to a new distributed source coding scheme, which has a successive refinement structure. As Figure 6 shows, the broadcast messages act like a ping-pong ball, which is flipped back and forth between the two encoders. We call such a scheme *two-encoder Distributed Successive Approximation Coding using Broadcast Advantage (*DiSAC2*)*, which was initially presented in [2]. We denote the number of encoding stages in DiSAC2 as n (e.g., Figure 6 illustrates 4 stages), and ENCx and ENCy as two separate encoders for sources X and Y respectively.

The sketch of the coding procedure for a n-stage DiSAC2 is as follows:

1) At the first stage, ENCx encodes X without any knowledge of Y. A codeword C_1 is generated and has a rate of R_1. (X_1, Y_1) is reconstructed at the BS after C_1 is received. The corresponding distortion pair is (D_{X_1}, D_{Y_1})[2].

2) At the second stage, ENCy overhears the codeword C_1 while it is being transmitted to the BS, so it only transmits the refinement which fully exploits the joint information between the source Y and C_1. C_2 is the corresponding codeword sent in the second stage, which has a rate of R_2. (X_2, Y_2) is reconstructed at the BS based on (C_1, C_2). The corresponding distortion pair is (D_{X_2}, D_{Y_2}).

3) Similarly, at the stage k, ENCx or ENCy (depends on if k is odd or even) encodes X or Y based on the codewords $(C_1, \cdots, C_{k-1})$. A codeword C_k is generated and has a rate of R_k. (X_k, Y_k) is reconstructed at the BS based on $(C_1, \cdots, C_k)$. The corresponding distortion pair is (D_{X_k}, D_{Y_k}).

If $X = Y$ in Figure 6, it reduces to a successive refinement of a single source. As we know from Section II-B, this is successively refinable on the *{rate, distortion}* curve when the Markovian condition is satisfied (e.g., Gaussian source). Similar results can be investigated for the DiSAC2 scheme. In the following, we specifically discuss the n-stage DiSAC2 with jointly Gaussian sources and quadratic distortion (DiSAC2-Gaussian): Let $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ be a bivariate Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ, and assume $(X, Y) \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$, where

$$\boldsymbol{\mu} = (0, 0), \quad \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \text{ for } |\rho| < 1.$$

B. DiSAC2-Gaussian is Successively Refinable

As we know, a Gaussian source with quadratic distortion is successively refinable on its *{rate, distortion}* curve. In this

[2]All distortions are defined as $\mathbb{E}d(x, \hat{x})$ where $d(\cdot, \cdot)$ is the quadratic error measure.
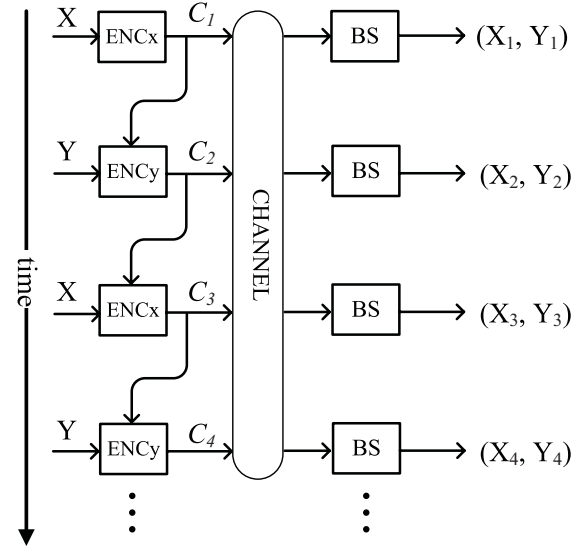

Fig. 6. Setup of a two-encoder distributed successive approximation coding using broadcast advantage (DiSAC2). The rate of codeword C_k sent at the ith stage is R_k. The corresponding reconstruction at the BS has a distortion pair (D_{X_k}, D_{Y_k}). Four stages are depicted.

section, we show that a n-stage DiSAC2 in the quadratic Gaussian case is also successively refinable on the *{sum-rate, distortion pair} surface* characterized by the rate-distortion region of the DSC2 [6].

Theorem 1 (Successive Refinability): Given two jointly Gaussian sources $(X, Y) \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$, the n-stage two-encoder Distributed Successive Approximation Coding using Broadcast Advantage (DiSAC2-Gaussian) with rates $(R_1, R_2, \ldots, R_n)$, and sum-rate distortion pairs

$$\{R_1, (D_{X_1}, D_{Y_1})\}, \ldots, \{R_1 + \ldots + R_n, (D_{X_n}, D_{Y_n})\},$$

achieves the {sum-rate, distortion pair} surface characterized by the rate-distortion region of DSC2 $R_{\text{DSC2}}(D_X, D_Y)$ [6], or equivalently

$$R_{\text{DSC2}}(D_{X_k}, D_{Y_k}) = R_1 + \ldots + R_k, \quad \text{for } k = 1, \ldots, n.$$

Proof: The detailed proof of Theorem 1 is given in the Appendix together with two preparatory lemmas. To provide the reader the methodology we used, we sketch the proof for the simplest two-stage case. In the following, $\mathcal{G}_{\mu, \sigma^2}(X)$ represents a Gaussian function $\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(X-\mu)^2}{2\sigma^2}}$. For brevity, we do not explicitly write out realizations in a probability density. For instance, $p(Y|X) = \mathcal{G}_{\rho X, 1-\rho^2}(Y)$ represents that the conditional probability of Y given X is a Gaussian distribution with a mean ρX and a variance $1 - \rho^2$, where X and Y are treated as the realizations on the right-hand side of the equation. The proof for general n-stage case is based on induction, and employs similar reasoning.

From [6], the minimum sum-rate for DSC2 is

$$R_{\text{DSC2}}(D_X, D_Y) = \frac{1}{2}\log\frac{(1-\rho^2)\left(\sqrt{\frac{4D_X D_Y \rho^2}{(1-\rho^2)^2} + 1} + 1\right)}{2D_X D_Y}.$$

$$(1)$$

For a two-stage DiSAC2 with jointly Gaussian sources and quadratic distortion, we need to prove

$$\begin{cases} R_1 = R_{\text{DSC2}}(D_{X_1}, D_{Y_1}) \\ R_1 + R_2 = R_{\text{DSC2}}(D_{X_2}, D_{Y_2}), \end{cases}$$

so that it achieves the {sum-rate, distortion pair} surface characterized by the rate-distortion region of the DSC2, for any rate pair (R_1, R_2).

Following the coding procedure in Section IV-A, at the first stage, X is encoded to C_1 using the random codebook argument with R-D pair

$$R_1 = \frac{1}{2} \log \frac{1}{D_{X_1}}, \quad D_{X_1} \le 1.$$

Then X_1 is the reconstruction of X after the first stage, and it can be decoded as: $X_1 = C_1$, thus $\mathbb{E}d(X, X_1) = D_{X_1}$. As $p(X) = \mathcal{G}_{0,1}(X)$, the *test channel* in the first stage is $p(X|C_1) = \mathcal{G}_{C_1, D_{X_1}}(X)$. Substituting $X_1 = C_1$,

$$p(X|C_1) = \mathcal{G}_{X_1, D_{X_1}}(X). \tag{2}$$

To decode Y_1 (reconstruction of Y at the first stage), we calculate the conditional probability

$$p(Y|C_1) = \int_{-\infty}^{+\infty} p(YX|C_1) \, dX$$

$$= \int_{-\infty}^{+\infty} p(Y|X) \cdot p(X|C_1) \, dX, \tag{3}$$

where the second equality follows from the fact that C_1 is encoded and decoded from X (a definite function of X), thus $p(Y|C_1 X) = p(Y|X)$. From the joint distribution of (X, Y),

$$p(Y|X) = \mathcal{G}_{\rho X, 1-\rho^2}(Y). \tag{4}$$

Substituting (4) and (2) into (3) leads to

$$p(Y|C_1) = \mathcal{G}_{\mu_1, \sigma_1^2}(Y), \tag{5}$$

where $\mu_1 = \rho X_1$, and $\sigma_1^2 = D_{X_1}\rho^2 - \rho^2 + 1$. Y_1 can be decoded as: $Y_1 = \mu_1$, and $\mathbb{E}d(Y, Y_1) = \sigma_1^2$.

To sum up, the {sum-rate, distortion pair} in the first stage is:

$$\begin{cases} R_1 = \frac{1}{2} \log \frac{1}{D_{X_1}} \\ (D_{X_1}, D_{Y_1}) = \left(D_{X_1}, \sigma_1^2\right). \end{cases} \tag{6}$$

Combining (1) and (6) leads to

$$R_1 = R_{\text{DSC2}}(D_{X_1}, D_{Y_1}). \tag{7}$$

At the second stage, C_1 is known due to the broadcast advantage, thus according to (5), $p(Y - \mu_1) = \mathcal{G}_{0, \sigma_1^2}(Y)$. $Y - \mu_1$ is encoded to C_2 using the random codebook argument, with R-D pair

$$R_2 = \frac{1}{2} \log \frac{\sigma_1^2}{D_{Y_2}}, \quad D_{Y_2} \le \sigma_1^2.$$

Y_2 can be decoded as: $Y_2 = C_2 + \mu_1$, thus

$$\mathbb{E}d(Y, Y_2) = \mathbb{E}d(Y - \mu_1, Y_2 - \mu_1)$$
$$= \mathbb{E}d(Y - \mu_1, C_2)$$
$$= D_{Y_2}.$$

Similar reasoning as for the first stage calculation gives

$$p(X|C_1 C_2) = \mathcal{G}_{\mu_2, \sigma_2^2}(X), \tag{8}$$

where $\mu_2 = (X_1(1 - \rho^2) + D_{X_1} Y_2 \rho / D_{X_1} \rho^2 - \rho^2 + 1)$, and $\sigma_2^2 = \frac{D_{X_1}\left((\rho^2 - 1)^2 - D_{X_1}\rho^2(-D_{Y_2} + \rho^2 - 1)\right)}{(D_{X_1}\rho^2 - \rho^2 + 1)^2}$.

X_2 can be decoded as: $X_2 = \mu_2$, and $\mathbb{E}d(X, X_2) = \sigma_2^2$.

Thus, the {sum-rate, distortion pair} in the second stage is:

$$\begin{cases} R_1 + R_2 = \frac{1}{2} \log \frac{\sigma_1^2}{D_{X_1} D_{Y_2}} \\ (D_{X_2}, D_{Y_2}) = \left(\sigma_2^2, D_{Y_2}\right). \end{cases} \tag{9}$$

Combining (9) and (1), we can verify the following after some calculations:

$$R_1 + R_2 = R_{\text{DSC2}}(D_{X_2}, D_{Y_2}). \tag{10}$$

∎

Observe that the coding procedure of DiSAC2-Gaussian actually suggests a simple algorithm to exploit correlations between refinement and previously sent codewords: a linear prediction of previous reconstructions can be used for predictive coding. We will adapt this idea into the image coding algorithm in Section V. Note that the correlation ρ between X and Y has to be sent to the BS for decoding. In more general cases where the joint statistics between two sources cannot be explicitly expressed, we can directly transmit the linear prediction coefficients of the encoders to the BS for decoding, which just requires a few bytes.

Figure 7 gives a visual illustration of Theorem 1: We choose an initial operating point $\{R_1, (D_{X_1}, D_{Y_1})\}$ at the first stage of DiSAC2-Gaussian, and then send an additional rate R_2 at the second stage. As the dashed curve on the gray surface suggests, $\{R_1 + R_2, (D_{X_2}, D_{Y_2})\}$ with any R_2 travels along a one-dimensional curve on the *Wagner Surface*. Similarly, after we send an additional rate R_3 at the third stage, $\{R_1 + R_2 + R_3, (D_{X_3}, D_{Y_3})\}$ still travels along a one-dimensional curve on the *Wagner Surface*. This means that we can split the overall rate into any number of pieces, send successively, and they will all operate on the *Wagner Surface* at each stage.

V. DISTRIBUTED SUCCESSIVE REFINEMENT OF STEREO-VIEW IMAGES USING BROADCAST ADVANTAGE

Under the stationary assumption of Section III-A, we extend the DiSAC2 scheme for bivariate Gaussian source (DiSAC2-*Gaussian*) to a practical image coding scheme, namely DiSAC2 for stereo-view images (DiSAC2-*Stereo*). With the help of broadcast, we combine layer decomposition and predictive coding to achieve successive refinement coding of stereo-view images. Then, simulation results are given by applying the algorithm to stereo-view image datasets.
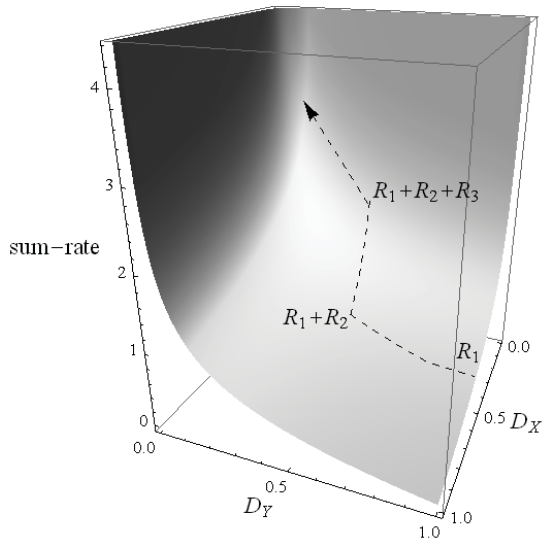
Fig. 7. DiSAC2-Gaussian is successfully refinable. The *{sum-rate, distortion pair}* travels on the *Wagner surface*. Three stages are illustrated, further stages follow a similar pattern.
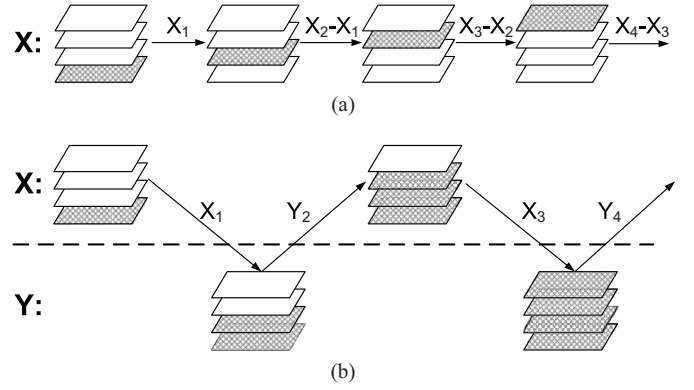

Fig. 8. Successive refinement using layer decomposition (four-stage setup). (a) Single image case: an incremental layer is transmitted at each stage. (b) Two correlated images case: incremental layer with lower layers is transmitted together at each stage. Shadow: Transmitted layers.

A. Algorithm

Like in the Gaussian case, we need the correlation between two sources to perform the distributed successive refinement. Image registration technique is the conventional tool to establish point-wise correspondence between two images. At the initial stage, one camera first transmit a coarse version of its image to the BS; the other camera overhears the same image and applies view registration to the stereo-view images[3]. The estimated registration parameters are then transmitted to the BS (a few bytes) and therefore are known at both cameras due to overhearing. After the raw stereo-view images are properly registered and aligned with each other, we investigate the successive refinement coding of the two aligned images X and Y (same notations as in Section IV). In our algorithm, only the overlapping parts of left and right views are considered in the coding process. Non-overlapping parts are encoded independently using conventional methods because they are not correlated.

Layer decomposition is the conventional method to achieve single image successive refinement: Layered bitstreams are transmitted in an incremental manner from low level layer to high level layer (see Figure 8a). By extending the idea of layer decomposition to an interleaved setup, we get an intuitive approach for the successive refinement of X and Y: As shown in Figure 8b, two images are encoded in ping-pong fashion as in the DiSAC2 scheme, and a new *incremental* layer is transmitted at each new stage. It is worth mentioning that, unlike the single image successive refinement as in Figure 8a, the complete k lowest layers are transmitted at kth stage (see Figure 8b). This is because the lower layers between the left view and the right view can be slight different due to registration error. The images transmitted at each stage are reconstructed at the BS, and we denote them

by $X_1, Y_2, X_3, Y_4, \ldots$, to be consistent with the notations in Section IV.

However, as X and Y are highly correlated, the lower layers of $X_1, Y_2, X_3, Y_4, \ldots$, are also correlated and therefore direct transmission is inefficient. Based on the coding procedure of DiSAC2-Gaussian, we propose to use linear prediction of previous reconstructions to eliminate the redundant information.

For a n-stage DiSAC2-Stereo scheme, at a even stage $2k \leq n$, we need to encode Y of the right-view camera and transmit it to the BS[4]. Previous reconstructions $(X_1, Y_2, \ldots, Y_{2k-2}, X_{2k-1})$ are fully known to the right-view camera due to the previous broadcasts. Thus we seek to maximize the quality of reconstruction:

$$\max_{\alpha_i,\, i \in [1, 2k-1]} \text{PSNR}\,(Y_{2k}, Y), \qquad (11)$$

where Y_{2k} is the reconstruction of Y at the BS, and α_i are $2k - 1$ variables for predictive coding.

The predictive coding procedure is as follows:

1) Y is decomposed into n layers with increasing quality. By omitting the highest $n - 2k$ layers, we obtain a coarse version of Y with the lowest $2k$ layers, which is denoted as *Temp*.

2) Compute the residual using linear prediction of previous reconstructions:

$$C_{2k} = Temp - (\alpha_1 \cdot X_1 + \alpha_2 \cdot Y_2 + \cdots + \alpha_{2k-1} \cdot X_{2k-1}),$$

3) Encode the residual C_{2k} with a desired rate R_{2k}. $\widehat{C}_{2k}$ is the corresponding decoded reconstruction, and Y_{2k} is obtained by:

$$Y_{2k} = \widehat{C}_{2k} + (\alpha_1 \cdot X_1 + \alpha_2 \cdot Y_2 + \cdots + \alpha_{2k-1} \cdot X_{2k-1}).$$

Based on the layer decomposition and the predictive coding above, we depict the entire algorithm of a n-stage DiSAC2-Stereo scheme in Algorithm 1.

B. Simulations

Using the algorithms described in Section V-A, we show the simulation results of a three-stage DiSAC2-*Stereo*.

[3]See [19] for details on the image registration algorithm, which performs robust under unequal image qualities between two images.

[4]For an odd stage, a similar procedure is used.

Algorithm 1 Algorithm of a *n*-Stage *DiSAC2-Stereo* Scheme

1: **for** stage $i = 1 \rightarrow n$ **do**
2:    **if** $i == 1$ (the first stage) **then**
3:       The left camera directly sends the base layer of its image to the BS.
4:       Meanwhile, the right camera overhears the left view and applies view registration algorithm.
5:       The right camera transmits the estimated registration parameters to the BS.
6:       The left camera overhears the registration parameters.
7:       Using the registration parameters, two aligned images X and Y for the left camera and the right camera are generated.
8:    **else**
9:       Two cameras take turns to send refinements.
10:      E.g., at a even stage $i = 2k$, we solve (11) using predictive coding procedure.
11:      The optimal prediction coefficients and the codewords are transmitted to the BS.
12:      The other camera overhears the same message and obtains a reconstruction for the coding of the next stage.
13:    **end if**
14: **end for**

To investigate the algorithm performance in a comprehensive way, we choose two experimental stereo-view datasets as shown in Figure 9: *Park* is captured by a rotating camera which is suitable for image registration. In contrast, *Church* is captured by two cameras far apart and has complex depth structure, which potentially brings registration error to the aligned image pair.

For all the simulations in this section, DiSAC2-*Stereo* is evaluated with rate triplet (R_1, R_2, R_3) in the range $\{R_1 \in (0.01, 0.05), R_2 \in (0.01, 0.15), R_3 \in (0.01, 0.1)\}$, among which 245 uniformly distributed grid samples are chosen to reduce the computation burden.

1) Prediction Coefficient Searching: According to (11), for a three-stage DiSAC2-*Stereo*, there are one coefficient α_1, and two coefficient α_1', α_2' needed to be optimized for the 2nd and 3rd stage coding respectively:

$$\begin{cases} \max_{\alpha_1} \ \mathrm{PSNR}\,(Y_2, Y) \\ \max_{\alpha_1', \alpha_2'} \ \mathrm{PSNR}\,(X_3, X). \end{cases}$$

Such optimizations can be done with extensive grid searching. To speed up the searching process, we use the following strategies:

1) The second optimization problem is simplified to a 1D searching with constraint $\alpha_1' + \alpha_2' = 1$.
2) First use linear searching with a very sparse grid to find a good initial point.
3) Starting from the chosen point, we obtain the optimal coefficient by using a typical sequential quadratic programming (SQP) solver[5].

---

[5]See fmincon function in MATLAB.

Dataset name	Stage	Mean (step)	Standard deviation
Park	2nd	2.5	0.7
	3rd	2.5	0.8
Church	2nd	3.1	0.8
	3rd	2.6	0.9

By comparing the optimal coefficients obtained by grid searching, our method is shown to be efficient and accurate in practice. Table I shows that the SQP solver usually converges within 3 steps.

2) Successive Refinement: Similar to Figure 7 for DiSAC2-*Gaussian*, we verify now how close DiSAC2-*Stereo* performs to the performance limit of distributed stereo-view coding. To the best of our knowledge, currently there is no distributed stereo-view image coder available. As an alternative, we use conventional prediction-based centralized image coding to approximate the distributed coding limit: By enumerating rate combinations of two encoders, we can obtain the distortion pairs achieved using centralized image coding. The green surfaces in Figure 10 show the operating points over all possible combinations, which represent the performance limits of centralized coding.

The two curves in Figure 10a or Figure 10b represent the operating points of a three-stage DiSAC2-*Stereo* at the 2nd stage and the 3rd stage, respectively. It can be seen that this scheme does perform quite close to the limit for the two datasets. To give a better illustration of coding losses, we measure the PSNR loss of DiSAC2-*Stereo* with respect to the performance limit of centralized coding, both using the same rate triplet (R_1, R_2, R_3). With the simulation results of 245 rate triplets as mentioned in the experimental setup, Figure 11 shows the distribution (PDF) of coding losses for the 2nd stage and the 3rd stage, respectively. We can see that the losses are mostly distributed within the range of 3dB: The maximum likelihood values of PSNR losses are 0.6dB/0.8dB at the 2nd/3rd stages for the *Park*, and 1.2dB/1.4dB at the 2nd/3rd stages for the *Church*. There are several reasons to explain these losses:

1) It is well known that a distributed setup has coding loss with respect to a centralized setup [20]. As a result, the centralized coding that we used provides an upper bound for the performance of distributed coding, and thus the actual coding loss of DiSAC2-*Stereo* will be smaller.
2) The successive refinability does not generally hold for any kind of sources. The quadratic Gaussian condition of Theorem 1 does not necessarily model natural images very well.
3) The error of stereo-view registration can degrade the correlation model of Section III-A. Particularly, we can see from Figure 11 that *Park* performs better than *Church* because it is taken by a fixed rotating camera and is simpler for registration.
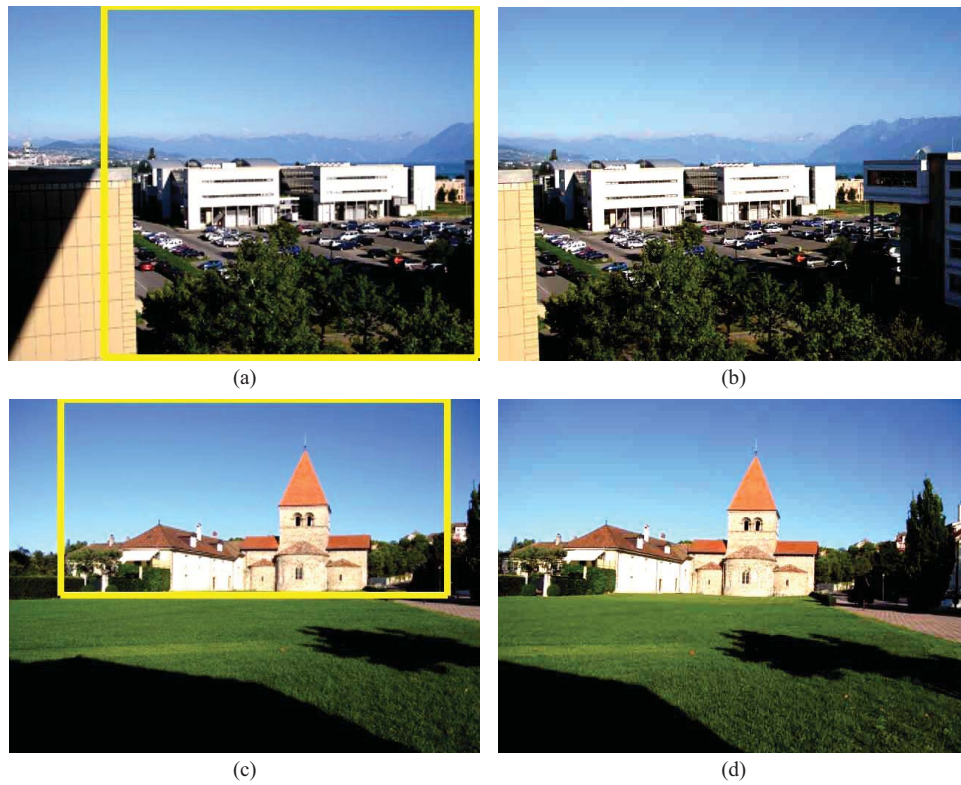
Fig. 9. The stereo-view datasets *Park* and *Church*. The yellow box in the left view shows the overlapping part of left and right views detected by the registration algorithm. (a) Left view of *Park*. (b) Right view of *Park*. (c) Left view of *Church*. (d) Right view of *Church*.
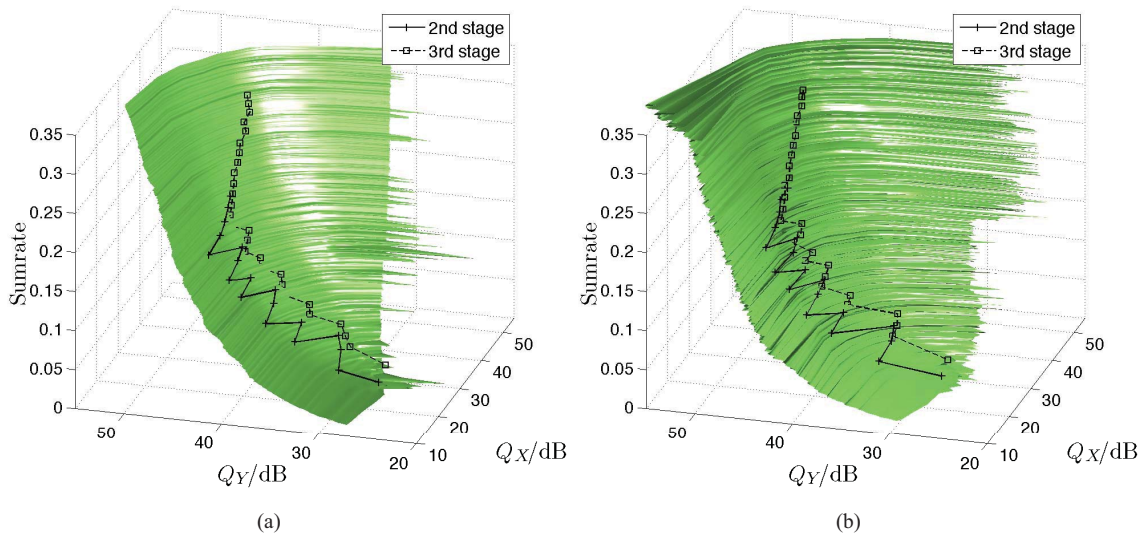

Fig. 10. Plots for (a) *Park* and (b) *Church*. Surface plotted in $(Q_X, Q_Y, \text{Sumrate})$ domain shows the performance limit of distributed stereo-view coding approximated by centralized coding. Q_X, Q_Y denote the reconstruction quality of stereo images X, Y, in terms of PSNR. The curves with markers show the operating points of a three-stage DiSAC2-*Stereo* at the second stage and the third stage, respectively. For clarity, not all samples are plotted in each curve.

We also compare our scheme with an independent intra-coding scheme where inter-view correlation is not exploited: If the broadcast is not utilized, each cameras just apply conventional single image successive refinement coding to its own image. In contrast, DiSAC2-*Stereo* takes advantage of the broadcast nature to exploit interview correlation. Figure 12 shows the distribution (PDF) of coding gains (in

terms of PSNR) of DiSAC2-*Stereo* with respect to the independent intra-coding scheme, at the 2nd stage and the 3rd stage, respectively. We can see that the coding gains are distributed over a wide range of 5dB, especially at the 3rd stages. Again, *Park* performs better than *Church* because it has a smaller registration error. These results show the benefit of exploiting inter-view correlations by using broadcast advantage.
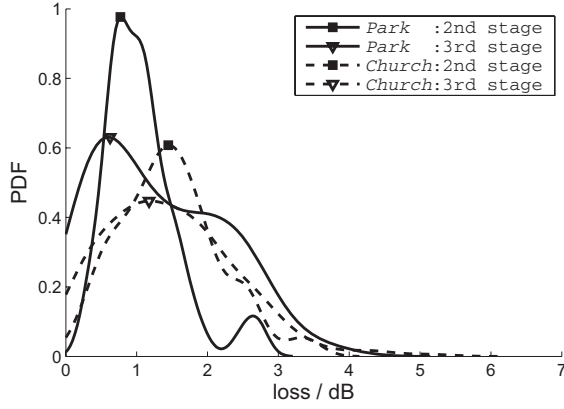
Fig. 11. Probability density (PDF) of PSNR losses of DiSAC2-*Stereo* compared to the performance limit of centralized coding, at the second stage and the third stage for two datasets, respectively. The distribution is obtained by measuring 245 different rate triplet combinations. The density curves are smoothed by normal kernel and normalized. The marker on each curve represents the maximum likelihood value of PSNR loss.
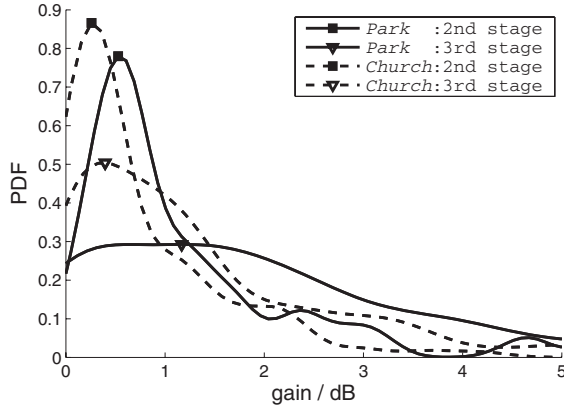

Fig. 12. Probability density (PDF) of coding gains of DiSAC2-*Stereo* as compared to independent intra-coding where the broadcast is not utilized, at the second stage and the third stage for two datasets, respectively. The distribution is obtained by measuring 245 different rate triplet combinations. The density curves are smoothed by normal kernel and normalized. The marker on each curve represents the maximum likelihood value of coding gains.

VI. CONCLUSION

We introduced a new successive refinement scheme for two encoders (DiSAC2), which takes advantage of the free broadcast nature present in wireless networks. The coding procedure imitates the ping-pong game, and has a successive refinement structure. For the quadratic Gaussian case, we prove that the DiSAC2-Gaussian is successively refinable on the *{sum-rate, distortion pair}* surface, which is characterized by the rate-distortion region of DSC2.

For environmental monitoring purpose, using broadcast advantage, we apply the same idea to encode stereo-view images. Based on the layer decomposition and linear prediction method, DiSAC2-Stereo is proposed to achieve successive refinement of registered stereo-view images. Simulation results of a three-stage DiSAC2-Stereo show that this scheme operates close to the distributed coding bound.

Last but not least, although we restrict the discussions to two cameras case in this paper, the same idea can be readily

extended to more cameras. One straightforward extension is to create a "multiple two-camera" system so that any part of the overall scenery is covered by at least one pair of cameras. In this case, to avoid the interferences among multiple cameras, we can use the adaptive synchronization algorithm proposed in [1] to properly coordinate the cameras.

APPENDIX
PROOF OF SUCCESSIVE REFINABILITY
OF n-STAGE DISAC2-GAUSSIAN

Lemma 1: For a n-stage DiSAC2 with Gaussian sources and quadratic distortion[6], the conditional probability at ith stage is Gaussian distributed: $p(\bar{Z}|C_1 \cdots C_i) = \mathcal{G}_{\mu_i, \sigma_i^2}(\bar{Z})$, where

$$Z = \begin{cases} X & i \text{ is odd} \\ Y & i \text{ is even,} \end{cases} \quad \text{and } \bar{Z} \text{ vice versa.} \quad (12)$$

μ_i and σ_i^2 can be calculated in a recursive way from $G_i(\bar{Z})$ in the following:

$$\begin{cases} G_i(\bar{Z}) = \int\limits_{-\infty}^{+\infty} H_i(X, Y) \, dZ \\ H_i(X, Y) = \dfrac{H_{i-1}(X, Y)}{G_{i-1}(Z)} \cdot \mathcal{G}_{Z_i, D_{Z_i}}(Z), \end{cases} \quad (13)$$

where $H_i(X, Y)$ denotes $p(XY|C_1 \cdots C_i)$, $G_i(\bar{Z})$ denotes $p(\bar{Z}|C_1 \cdots C_i)$, and the initial condition is

$$H_0(X, Y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{X^2 + Y^2 - 2\rho XY}{2(1-\rho^2)}}. \quad (14)$$

Furthermore, $H_i(X, Y)$ can be expressed as $K_i \cdot e^{-h_i(X,Y)}$, where $h_i(X, Y)$ is a polynomial of (X, Y) of degree 2, and K_i is a coefficient that keep $H_i(X, Y)$ normalized. The exponent $h_i(X, Y)$ can be calculated recursively by:

$$h_i(X, Y) = h_{i-1}(X, Y) - \frac{(Z - \mu_{i-1})^2}{2\sigma_{i-1}^2} + \frac{(Z - Z_i)^2}{2D_{Z_i}}. \quad (15)$$

Proof: At the initial stage $i = 0$, it is immediate to verify that $H_0(X, Y)$ in (14) is consistent with $p(XY)$ and $h_0(X, Y)$ is a polynomial of (X, Y) of degree 2. From (13), $G_0(X) = \int\limits_{-\infty}^{+\infty} H_0(X, Y) \, dY = \frac{1}{2\pi} e^{-\frac{X^2}{2}}$, which is also consistent with $p(X)$. Thus, Lemma 1 holds when $i = 0$.

By induction, we first assume that Lemma 1 holds when $i = 2k$, $k \geq 0$, from which we know that

$$\begin{cases} G_{2k}(X) = p(X|C_1 \cdots C_{2k}) = \mathcal{G}_{\mu_{2k}, \sigma_{2k}^2}(X) \\ H_{2k}(X, Y) = p(XY|C_1 \cdots C_{2k}) = K_{2k} \cdot e^{-h_{2k}(X,Y)} \quad (16) \\ h_{2k}(X, Y) \text{is a polynomial of } (X, Y) \text{ with the degree of 2.} \end{cases}$$

When $i = 2k + 1$, the coding procedure is as follows: $(C_1, \ldots, C_{2k})$ are already known due to the broadcast advantage. Thus according to (16), $p(X - \mu_{2k}) = \mathcal{G}_{0, \sigma_{2k}^2}(X)$. $X - \mu_{2k}$

---

[6]Notations inherited from Section IV.

is then encoded to C_{2k+1} using random codebook argument, with R-D pair

$$R_{2k+1} = \frac{1}{2} \log \frac{\sigma_{2k}^2}{D_{X_{2k+1}}}, \quad D_{X_{2k+1}} \le \sigma_{2k}^2.$$

X_{2k+1} can be decoded as: $X_{2k+1} = C_{2k+1} + \mu_{2k}$.

As $p(X - \mu_{2k}|C_1 \cdots C_{2k}) = \mathcal{G}_{0,\sigma_{2k}^2}(X)$, the *test channel* in this stage is

$$p(X - \mu_{2k}|C_1 \cdots C_{2k}C_{2k+1}) = \mathcal{G}_{C_{2k+1},D_{X_{2k+1}}}(X).$$

Thus

$$\begin{aligned} p(X|C_1 \cdots C_{2k+1}) &= \mathcal{G}_{C_{2k+1}+\mu_{2k},D_{X_{2k+1}}}(X) \\ &= \mathcal{G}_{X_{2k+1},D_{X_{2k+1}}}(X). \end{aligned} \quad (17)$$

To decode Y_{2k+1}, the conditional probability

$$\begin{aligned} &p(Y|C_1 \cdots C_{2k+1}) \\ &= \int_{-\infty}^{+\infty} \frac{p(XY|C_1\cdots C_{2k})}{p(X|C_1\cdots C_{2k})} \cdot p(X|C_1 \cdots C_{2k+1}) \, dX, \end{aligned}$$

which follows from a similar argument as (3) and the fact that C_{2k+1} is encoded and decoded from $(X, C_1, \cdots, C_{2k})$. To calculate the integral, $p(XY|C_1 \cdots C_{2k})$ and $p(X|C_1 \cdots C_{2k})$ are already known from (16) and $p(X|C_1 \cdots C_{2k+1}) = \mathcal{G}_{X_{2k+1},D_{X_{2k+1}}}(X)$ due to (17). Therefore,

$$\begin{cases} H_{2k+1}(X,Y)7 = p(XY|C_1 \cdots C_{2k+1}) \\ \qquad = \dfrac{H_{2k}(X,Y)}{G_{2k}(X)} \cdot \mathcal{G}_{X_{2k+1},D_{X_{2k+1}}}(X) \\ G_{2k+1}(Y) = p(Y|C_1 \cdots C_{2k+1}) = \displaystyle\int_{-\infty}^{+\infty} H_{2k+1}(X,Y) \, dX. \end{cases} \quad (18)$$

From (16) and (18),

$$H_{2k+1}(X,Y) = \frac{K_{2k}\sigma_{2k}^2}{D_{X_{2k+1}}} \cdot e^{-\left(h_{2k}(X,Y)-\frac{(X-\mu_{2k})^2}{2\sigma_{2k}^2}+\frac{(X-X_{2k+1})^2}{2D_{X_{2k+1}}}\right)}.$$

Thus, $H_{2k+1}(X,Y)$ can be expressed as $K_{2k+1} \cdot e^{-h_{2k+1}(X,Y)}$, where

$$h_{2k+1}(X,Y) = h_{2k}(X,Y) - \frac{(X-\mu_{2k})^2}{2\sigma_{2k}^2} + \frac{(X-X_{2k+1})^2}{2D_{X_{2k+1}}}. \quad (19)$$

Finally, from (18) and the fact that $h_{2k+1}(X,Y)$ is a polynomial of (X, Y) with the degree of 2, it is immediate to verify that $p(Y|C_1 \cdots C_{2k+1}) = \int_{-\infty}^{+\infty} K_{2k+1} \cdot e^{-h_{2k+1}(X,Y)} \, dX$ is Gaussian distributed, and can be denoted as

$$\mathcal{G}_{\mu_{2k+1},\sigma_{2k+1}^2}(Y). \quad (20)$$

In conclusion, from (18), (19), and (20), it is proven that Lemma 1 holds when $i = 2k + 1$. Similar arguments can be used to prove the induction from stage $2k+1$ to stage $2k+2$. Therefore, Lemma 1 is proved.   ■

Given the recursive calculation process for σ_i^2 as Lemma 1, we get the recursive expression of σ_i^2 as follows.

Lemma 2: For a n-stage DiSAC2 with Gaussian sources and quadratic distortion, σ_i^2 at each stage $(i \ge 2)$ can be expressed as:

$$\frac{1}{2\sigma_i^2} = a_{i-1} - \frac{\rho^2}{\frac{\rho^2}{a_{i-1}} + \frac{2(1-\rho^2)^2}{D_{Z_i}}}, \quad (21)$$

where $a_i > 0$ and can be computed recursively by

$$a_i = \frac{\rho^2}{4a_{i-1}(1-\rho^2)^2} + \frac{1}{2D_{Z_i}}. \quad (22)$$

The notation for Z is the same as (12).

The initial condition is given by

$$\begin{cases} a_1 = \dfrac{1}{2(1-\rho^2)} - \dfrac{1}{2} + \dfrac{1}{2D_{X_1}} \\ \dfrac{1}{2\sigma_1^2} = \dfrac{1}{2(1-\rho^2)} - \dfrac{\rho^2}{4a_1(1-\rho^2)^2}. \end{cases} \quad (23)$$

Proof: From Lemma 1, we know that $h_i(X, Y)$ is a polynomial of (X, Y) of degree 2. Thus, we rewrite

$$h_1(X, Y) = a_1(X - b_1 Y - c_1)^2 + p_1(Y - q_1)^2 + s_1, \quad (24)$$

where $a_1, b_1, c_1, p_1, q_1, s_1$ are real-valued coefficients.

From (13):

$$\begin{aligned} \mathcal{G}_{\mu_1,\sigma_1^2}(Y) &= G_1(Y) \\ &= \int_{-\infty}^{+\infty} K_1 \cdot e^{-(a_1(X-b_1 Y-c_1)^2+p_1(Y-q_1)^2+s_1)} \, dX \\ &= K_1' \cdot e^{-p_1(Y-q_1)^2}. \end{aligned} \quad (25)$$

By checking the definition of $\mathcal{G}_{\mu_1,\sigma_1^2}(Y)$, $p_1 = \frac{1}{2\sigma_1^2}, q_1 = \mu_1$, K_1' is the normalized coefficient, and $a_1 > 0$ (otherwise (25) not integrable).

From (14), (15) and $\mu_0 = 0, \sigma_0^2 = 1$, we know that

$$h_1(X, Y) = \frac{X^2 + Y^2 - 2\rho XY}{2(1-\rho^2)} - \frac{X^2}{2} + \frac{(X-X_1)^2}{2D_{X_1}}. \quad (26)$$

By comparing the coefficients of terms X^2, XY, Y^2 in (24) and (26),

$$\begin{cases} a_1 = \dfrac{1}{2(1-\rho^2)} - \dfrac{1}{2} + \dfrac{1}{2D_{X_1}} \\ a_1 b_1 = \dfrac{\rho}{2(1-\rho^2)} \\ \dfrac{1}{2\sigma_1^2} = p_1 = \dfrac{1}{2(1-\rho^2)} - \dfrac{\rho^2}{4a_1(1-\rho^2)^2}, \end{cases}$$

which proves the initial condition (23). Using induction, we first assume that Lemma 2 holds when $i = 2k - 1$, $k \ge 1$, from which we know that

$$\begin{cases} h_{2k-1}(X, Y) = a_{2k-1}(X - b_{2k-1}Y - c_{2k-1})^2 \\ \qquad\qquad + p_{2k-1}(Y - q_{2k-1})^2 + s_{2k-1} \\ a_{2k-1} b_{2k-1} = \dfrac{\rho}{2(1-\rho^2)} \\ p_{2k-1} = \dfrac{1}{2\sigma_{2k-1}^2}. \end{cases} \quad (27)$$

When $i = 2k$, since $h_{2k}(X, Y)$ is a polynomial of (X, Y) of degree 2, we rewrite

$$h_{2k}(X, Y) = a_{2k}(Y - b_{2k}X - c_{2k})^2 + p_{2k}(X - q_{2k})^2 + s_{2k}, \tag{28}$$

where $a_{2k}, b_{2k}, c_{2k}, p_{2k}, q_{2k}, s_{2k}$ are real-valued coefficients. Using similar induction as before, from (13) (15) (27),

$$\begin{cases} a_{2k} = a_{2k-1}b_{2k-1}^2 + p_{2k-1} - \dfrac{1}{2\sigma_{2k-1}^2} + \dfrac{1}{2D_{Y_{2k}}} \\[2mm] a_{2k}b_{2k} = \dfrac{\rho}{2(1 - \rho^2)} \\[2mm] \dfrac{1}{2\sigma_{2k}^2} = p_{2k} = a_{2k-1} - a_{2k}b_{2k}^2, \end{cases}$$

which can be further reduced by comparing with (27):

$$\begin{cases} a_{2k} = \dfrac{\rho^2}{4a_{2k-1}(1 - \rho^2)^2} + \dfrac{1}{2D_{Y_{2k}}} \\[2mm] a_{2k}b_{2k} = \dfrac{\rho}{2(1 - \rho^2)} \\[2mm] \dfrac{1}{2\sigma_{2k}^2} = a_{2k-1} - \dfrac{\rho^2}{\frac{\rho^2}{a_{2k-1}} + \frac{2(1-\rho^2)^2}{D_{Y_{2k}}}}. \end{cases} \tag{29}$$

Thus, (29) proves that Lemma 2 holds when $i = 2k$. Similar arguments can be used to prove the induction from $i = 2k$ to $2k + 1$. Therefore, Lemma 2 is proved. ∎

Proof of Theorem 1: Due to the ping-pong structure of the coding procedure, there are some differences between even and odd stages. Without loss of generality, we prove the induction from $2k$ stage DiSAC2 to $2k + 1$ stage DiSAC2. Similar arguments can be used to verify the induction from $2k+1$ stage DiSAC2 to $2k + 2$ stage DiSAC2 (omitted here for brevity).

By induction, we first assume that a $2k$ stage DiSAC2-Gaussian is successively refinable ($k \geq 1$), that is:

$$R_{\text{DSC2}}(D_{X_{2k}}, D_{Y_{2k}}) = R_1 + \cdots + R_{2k}, \tag{30}$$

where $R_{\text{DSC2}}(D_X, D_Y)$ is the rate-distortion region of the DSC2 [6] as defined by (1).

According to the coding procedure (see the sketch of proof in Section IV-B), the rate and the distortion pairs at stage $2k$ and $2k + 1$ are

$$\begin{cases} R_{2k} = \dfrac{\sigma_{2k-1}^2}{D_{Y_{2k}}} \\[2mm] (D_{X_{2k}}, D_{Y_{2k}}) = \left(\sigma_{2k}^2, D_{Y_{2k}}\right) \\[2mm] R_{2k+1} = \dfrac{\sigma_{2k}^2}{D_{X_{2k+1}}} \\[2mm] (D_{X_{2k+1}}, D_{Y_{2k+1}}) = \left(D_{X_{2k+1}}, \sigma_{2k+1}^2\right), \end{cases} \tag{31}$$

where σ_i^2 is as defined in Lemma 1. Plugging (31) into (30), it holds at stage $2k$ that

$$R_{\text{DSC2}}(\sigma_{2k}^2, D_{Y_{2k}}) = \frac{1}{2} \log \frac{\sigma_1^2 \cdots \sigma_{2k-1}^2}{D_{X_1} D_{Y_2} \cdots D_{Y_{2k}}}. \tag{32}$$

To prove the successive refinability at stage $2k + 1$, we need to verify

$$R_{\text{DSC2}}(D_{X_{2k+1}}, \sigma_{2k+1}^2) = R_1 + \cdots + R_{2k+1}$$

$$= \frac{1}{2} \log \frac{\sigma_1^2 \cdots \sigma_{2k-1}^2 \sigma_{2k}^2}{D_{X_1} D_{Y_2} \cdots D_{Y_{2k}} D_{X_{2k+1}}}$$

$$= \frac{1}{2} \log \frac{(1 - \rho^2) \left(\sqrt{\frac{4\sigma_{2k}^2 D_{Y_{2k}} \rho^2}{(1-\rho^2)^2} + 1} + 1\right) \sigma_{2k}^2}{2\sigma_{2k}^2 D_{Y_{2k}} D_{X_{2k+1}}}, \tag{33}$$

where the third equality follows from (1) and (32). By expanding the left side of (33) using (1), (33) can be further reduced to

$$\frac{\sqrt{\frac{4D_{X_{2k+1}} \sigma_{2k+1}^2 \rho^2}{(1-\rho^2)^2} + 1} + 1}{\sigma_{2k+1}^2} = \frac{\sqrt{\frac{4\sigma_{2k}^2 D_{Y_{2k}} \rho^2}{(1-\rho^2)^2} + 1} + 1}{D_{Y_{2k}}}. \tag{34}$$

From Lemma 2, we know that

$$\begin{cases} \sigma_{2k}^2 = 1 \Big/ 2 \left(a_{2k-1} - \dfrac{\rho^2}{\frac{\rho^2}{a_{2k-1}} + \frac{2(1-\rho^2)^2}{D_{Y_{2k}}}}\right) \\[4mm] \sigma_{2k+1}^2 = 1 \Big/ 2 \left(a_{2k} - \dfrac{\rho^2}{\frac{\rho^2}{a_{2k}} + \frac{2(1-\rho^2)^2}{D_{X_{2k+1}}}}\right) \\[4mm] a_{2k} = \dfrac{\rho^2}{4a_{2k-1}(1 - \rho^2)^2} + \dfrac{1}{2D_{Y_{2k}}}. \end{cases} \tag{35}$$

Clearly, it holds that $|\rho| < 1$, $D_{Y_{2k}} > 0$, $D_{X_{2k+1}} > 0$, and $a_{2k-1} > 0$ according to Lemma 2. Under these conditions, by plugging (35) into both sides of (34), we can verify the equality after some extensive calculations[7]. Therefore, it proves that a $2k + 1$ stage DiSAC2-Gaussian is successively refinable.

In Section IV-B, we already proved the initial conditions (7) and (10). Thus, using the induction above, we can conclude that any n-stage DiSAC2-Gaussian is successively refinable. ∎
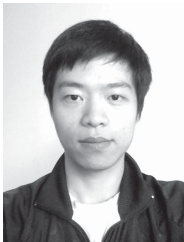
ACKNOWLEDGMENT

REFERENCES

[1] Z. Chen, G. Barrenetxea, and M. Vetterli, "Share risk and energy: Sampling and communication strategies for multi-camera wireless monitoring networks," in *Proc. 31st Annu. IEEE Int. Conf. Comput. Commun.*, Mar. 2012, pp. 1862–1870.

[2] Z. Chen, G. Barrenetxea, and M. Vetterli, "Distributed successive approximation coding using broadcast advantage: The two-encoder case," in *Proc. 48th Annu. Allerton Conf. Commun., Control Comput.*, Sep.–Oct. 2010, pp. 1110–1116.

[3] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications* (CISM Courses and Lectures), vol. 229, G. Longo, Ed. New York: Springer-Verlag, 1978, pp. 171–231.

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

---

[7]One can verify this using computer algebra softwares like Mathematica.

[5] A. 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.

[6] A. B. Wagner, S. Tavildar, and P. Viswanath, "Rate region of the quadratic Gaussian two-encoder source-coding problem," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1938–1961, May 2008.

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

[8] R. Puri, A. Majumdar, and K. Ramchandran, "Prism: A video coding paradigm with motion estimation at the decoder," *IEEE Trans. Image Process.*, vol. 16, no. 10, pp. 2436–2448, Oct. 2007.

[9] C. Yeo and K. Ramchandran, "Robust distributed multiview video compression for wireless camera networks," *IEEE Trans. Image Process.*, vol. 19, no. 4, pp. 995–1008, Apr. 2010.

[10] A. Kaspi and T. Berger, "Rate-distortion for correlated sources with partially separated encoders," *IEEE Trans. Inf. Theory*, vol. 28, no. 6, pp. 828–840, Nov. 1982.

[11] J. Korner, "Successive encoding of correlated sources," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 390–395, May 1983.

[12] Y. Oohama, "Universal coding for correlated sources with linked encoders," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 837–847, May 1996.

[13] W. Equitz and T. Cover, "Successive refinement of information," *IEEE Trans. Inf. Theory*, vol. 37, no. 2, pp. 269–275, Mar. 1991.

[14] H. Viswanathan and T. Berger, "Sequential coding of correlated sources," *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 236–246, Jan. 2000.

[15] H. Behroozi and M. R. Soleymani, "Successively structured Gaussian two-terminal source coding," *Wirel. Personal Commun.*, vol. 48, no. 4, pp. 485–510, 2009.

[16] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 still image coding system: An overview," *IEEE Trans. Consumer Electron.*, vol. 46, no. 4, pp. 1103–1127, Nov. 2000.

[17] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. 2003.

[18] T. Cui, L. Chen, and T. Ho, "Distributed optimization in wireless networks using broadcast advantage," in *Proc. 46th IEEE Conf. Decision Control*, Dec. 2007, pp. 5839–5844.

[19] Z. Chen, "Sensorcam cameras and image registration," EPFL, Lausanne, Switzerland, Tech. Rep. 170457, 2011.

[20] R. Zamir, "The rate loss in the Wyner-Ziv problem," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 2073–2084, Nov. 1996.

Guillermo Barrenetxea received the Master's degree from the Public University of Navarra, Navarra, Spain, and the Ph.D. degree in communication systems from École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

He is a Senior Researcher at EPFL, where he is responsible for the Wireless Sensor Network Research Group, Audiovisual Communication Laboratory. He has published about 25 scientific papers on a variety of topics within sensor networks. He is a co-founder and CTO of the startup company Sensorscope that commercializes environmental sensing technology.

Martin Vetterli (S'86–M86–SM'90–F'95) received the Dipl.El.Ing. degree from ETH Zurich, Zurich, Switzerland, the M.S. degree from Stanford University, Stanford, CA, in 1981, 1982, and 1986, respectively, and the Doctorat és Sciences degree from 'Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, in 1986.

He was an Associate Professor in electrical engineering with Columbia University, New York, NY, and was a Full Professor in electrical engineering and computer sciences with the University of California, Berkeley, before joining the Communication Systems Division of EPFL. He held several positions at EPFL, including the Chair of Communication Systems and the Founding Director of the National Competence Center in Research on Mobile Information and Communication Systems. From 2004 to 2011, he was the Vice President of EPFL, and since March 2011, he has been the Dean of the School of Computer and Communication Sciences, EPFL. He will lead the Swiss National Science Foundation in January 2013. He is involved in research on electrical engineering, computer sciences, and applied mathematics. His research interests include wavelet theory and applications, image and video compressions, self-organized communications systems and sensor networks, and fast algorithms. He has authored over 150 journals papers. He is the co-author of three textbooks, with J. Kovacevic, *Wavelets and Subband Coding* (Prentice-Hall, 1995), with P. Prandoni, *Signal Processing for Communications* (CRC Press, 2008), and with J. Kovacevic and V. Goyal, the forthcoming book *Fourier and Wavelet Signal Processing* (2012).

Mr. Vetterli was a recipient of several awards, including the Best Paper Awards from EURASIP in 1984 and the IEEE Signal Processing Society in 1991, 1996, and 2006, the Swiss National Latsis Prize in 1996, the SPIE Presidential Award in 1999, the IEEE Signal Processing Technical Achievement Award in 2001, and the IEEE Signal Processing Society Award in 2010. He is a fellow of ACM and EURASIP, was a member of the Swiss Council on Science and Technology from 2000 to 2004, and is an ISI highly cited researcher in engineering.

Zichong Chen (S'10) received the B.Sc. degree (Hons.) in electrical engineering from Tsinghua University, Beijing, China, in 2008. He is currently pursuing the Ph.D. degree in communication systems with the École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

He researched real-time 3-D scene reconstruction and human face modeling and rendering, as a Research Intern at Huawei, North America Headquarters (R&D), Santa Clara, CA. He is currently a Research Assistant with the Audiovisual Communications Laboratory, EPFL. His current research interests include distributed smart camera networks, image/video processing, mobile sensing, and statistical learning.