

## **Enabling wireless uncompressed HDTV connectivity with a unique video-modem approach: A technical overview**

**Meir Feder, Chief Technology Officer, AMIMON Ltd.**

### **Abstract**

Delivering uncompressed HDTV content wirelessly has become the Holy Grail of wireless video connectivity. Consumers are interested in a wireless connection option for their TVs and projectors in order to make installation easier and more flexible, and to enable simple whole-home connectivity between video sources and displays. Enabling applications such as wireless HDTVs and wireless HD projectors requires delivery of uncompressed HDTV video with video rates as high as 3Gbps (uncompressed 1080p). However, existing and emerging wireless standards, such as 802.11n and UWB, are not capable of delivering such high video rates.

This paper introduces a new video-modem approach for wireless video which bridges this gap. This paper will describe this new approach which uses joint-source channel coding to optimize the wireless modem for video delivery. The paper will explain how this video optimization enables more than a 10X improvement when compared to the traditional data-modem method for wireless video, thereby enabling wireless delivery of very high uncompressed video rates with high reliability and with a range that could cover the whole home.

### **Introduction**

The transformation to high definition content has created consumer demand for high-quality digital video connectivity between high-definition (HD) displays such as high-definition televisions (HDTVs) and HD projectors to other consumer electronic (CE) devices that produce HD sources, such as HD-DVDs, Blu-ray, HD set-top boxes, game consoles and so on. For wire connectivity, high-definition multimedia interface (HDMI™) has emerged in recent years as the enabling digital interface for HD and consumer electronics that can transfer uncompressed HD video and audio at a rate of over 3Gbps required for 1080p video. As a result, HDMI is on the verge of becoming a ubiquitous audio visual interface. Can a similar solution be devised for wireless connectivity?

The success of WiFi and other wireless standards show that consumers love wireless. To fit the slick design of the wall-hanging flat panel displays, consumers wish to get rid of the cumbersome wires and have a flexible, easy-to-use and -install wireless connectivity. However, existing and emerging wireless standards such as 802.11n and UWB are not capable of delivering the high video rates required for high-quality video connectivity. Something else is needed in order to achieve the wireless counterpart of HDMI.

Wireless high-quality video transfer is a tough problem. One aspect is the high video rate requiring a communication channel with enough bandwidth and signal-to-noise ratio (SNR), i.e., enough capacity. But more so, the wireless channel is unstable and unpredictable. Its characteristics change rapidly; due to fading and interference, its SNR and capacity vary considerably. In data transfer, buffers and re-transmissions can compensate for these problems. This is impossible in video (and audio) connectivity, where the transfer must be done in real time with no delay and the high fidelity must remain intact throughout the transfer.

This paper outlines the novel video-modem approach for wireless video transfer, based on the information theoretic principle of Joint Source Channel Coding (JSCC). The video-modem approach overcomes the challenges of the wireless channels for video delivery. AMIMON uses this approach in its Wireless High-Definition Interface (WHDI™) solution which enables true robust uncompressed wireless A/V connectivity with a range that could cover the whole home. The WHDI™ solution uses a multi-input-multi-output (MIMO) 20MHz/40MHz bandwidth channel over the 5GHz band, where MIMO provides an extra bandwidth and diversity boost to the solution. The video-modem JSCC approach can be used over other radios and frequency band as well, enabling a robust 10 times improvement compared to the traditional data-modem method for wireless video, thereby enabling wireless delivery of very high uncompressed video rates.

### **Background from Information Theory**

How to transfer an information source to a destination? More specifically, consider an HDTV audio/video stream over a noisy, fading wireless channel. From the information theory point of view, a communication channel has a capacity, denoted  $C$ , that specifies the maximal rate (in bits per second) that can be sent reliably over the channel.  $C$  depends in general on the specific channel characteristics. Yet a good yardstick providing a simple approximated formula is the famous Shannon capacity of the additive white Gaussian noise (AWGN) channel:

$$C = W \log (1+\text{SNR}) \text{ bits per second}$$

where  $W$  is the channel bandwidth (in Hz) and the logarithm is base 2.

The source information content is also measured in bits per second. If the HD video source is simply quantized to say, 8 bits per pixel per color, then 720 p requires about 1.3Gbps, 1080i about 1.5Gbps, and 1080p about 3Gbps. This is probably an overestimate of the information content since by lossless compression this amount of bits can be reduced by, say, a factor of two. Furthermore, if more distortion is allowed, the bit rate needed to represent the video source within that distortion is further reduced.

The information theory quantity  $R(D)$ , the rate-distortion function of the source, defines the minimal bit-rate that is needed to represent the source within a specified distortion,  $D$ . It is common to measure  $D$  by the average square error between the original source and its representation, as compared to the source dynamic range. This measure, called peak signal to noise ratio (PSNR), is for 8-bit representation of the video:

$$\text{PSNR} = 10 \log_{10} (255)^2/D \text{ dB}$$

While  $R(D)$  of a video source is not easily specified and depends on the video content and the correlation between the pixels and the frames, for low  $D$  (i.e., high PSNR), it can be assumed that the fine quality video samples are independent. At this operating point, a bit per samples is required to reduce the error magnitude by a factor of two (or achieve 6dB better PSNR). 720p video has 55.3 Mega pixels per second or  $55.3 \times 3 \approx 166$  M samples per second, 1080i has 62.2 Mega pixels or 186.6 Mega samples per second, and 1080p has 124.4 Mega pixels or 373.2 Mega samples per second. In other words, for high-quality video,

it requires 166, 186.6 and 373.2 Mbps respectively for improving the video quality by 6dB at the high PSNR region.

The best attainable performance in transferring a source over an information channel using any communication scheme is given by Shannon's celebrated "source-channel coding theorem," asserting that this best performance (minimum distortion) of maintains:

$$R(D) = C$$

i.e., the source is distorted by an amount,  $D$ , that cannot be better than  $R^{-1}(C)$ . As channel capacity depends on the channel signal-to-noise ratio (CSNR), to assess the performance, one has to draw a graph of the PSNR as a function of the channel capacity or the CSNR.

Shannon's theorem statement is intuitively understandable.  $C$  measures the maximal bit rate that can be transferred reliably over the channel, while  $R(D)$  is the minimal bit rate required to represent the source up to a distortion,  $D$ . Furthermore, Shannon's theorem suggests a solution given by the separation principle: The optimal performance can be achieved by first compressing the source to  $R$  bits-per-second, inflicting a distortion  $D$  on the source, then (assuming that  $R$  is smaller than the channel capacity,  $C$ ) sending these bits without error over the channel. This is the common approach. However, while this approach can (theoretically) achieve optimal performance, it cannot do so in practice at reasonable complexity and delay for reasons discussed below.

An alternative approach is joint source-channel coding (JSCC), where there is no separation between source coding and channel coding. This approach leads to the video-modem architecture. As described below, this approach fits the varying and unpredictable nature of the wireless channel and can attain (close to) the optimal performance via a robust, cost-effective solution that is adaptive to the varying channel capacity.

### Feasibility assessment for the 5GHz, 20/40Hz MIMO channel

The expression above allows checking the feasibility of any proposed solution to deliver HD video over a channel. Let us then check AMIMON's proposed WHDI™ solution: Is the channel capacity indeed greater than the source rate-distortion function at the desired low distortion? Can it provide a reliable wireless video connectivity that can cover the whole home?

What is the capacity? The WHDI™ solution uses a wireless MIMO channel over 5GHz band, with four transmit antennas and 20-40MHz bandwidth. The wireless MIMO channel capacity, when the channel matrix is  $H$  (known to the receiver) and with Gaussian noise is:

$$C = \log_2 \det(I + \rho H H^H) \quad \text{bits/sec/Hz}$$

where  $\rho$  is the CSNR per transmit antenna,  $H$  is the channel matrix, and the superscript  $^H$  indicates the complex conjugate. If  $n_R$ , the number of antennas at the receiver, is greater or equal  $n_T$ , the number of transmit antennas, and the channel matrix is non-singular, the capacity formula is approximated by:

$$C \approx n_T \log_2 (1 + \rho(H)) \quad \text{bits/sec/Hz}$$

where  $\rho(H)$  is an average effective CSNR that depend on the specific fading channel matrix  $H$ , which in turn depends on the environment and the distance. Thus, the capacity is not constant but random, changing according to the varying effective CSNR. With  $n_T=4$  and 20MHz bandwidth, the capacity can vary from about 800Mbps for 30dB CSNR to about 500Mbps for 18dB CSNR, and it can get down to 250Mbps for  $\sim 10$ dB CSNR, in high fading and large home distance.

Consider now the source. As discussed above, it is hard to assess the HD video rate-distortion function since the video is composed of many different images – some require a very small rate (like a uniform color image) and some require a higher rate. Yet, as confirmed by the fact that the 8-bit HDTV samples can be losslessly compressed by about a factor of two, it can be assumed that the HD (1080i or 720p) rate distortion value is below  $\sim 700$  Mbps at that 8-bit distortion level. If a higher distortion is allowed, a lower rate is sufficient to represent the source. Following reports on the performance of compression algorithms the HDTV source requires about 100-200 Mbps for representing the HDTV source at better than 40dB PSNR. Note that an assessment for the rate-distortion value of a 1080p source is twice the assessment of the 1080i source.

Summarizing the above, for 720p/1080i, the 20MHz channel used by WHDI™ has a capacity of 250-800 Mbps throughout the home, which is higher than that HDTV rate-distortion value at video quality of 40dB PSNR. However, it emphasizes the importance of being close to the Shannon bound and utilizing the optimal performance. The JSCC approach can do it “on-the-fly,” while the separation approach struggles and, as noted above, requires a reliable transmission of hundreds of Mega bits per second in order to improve the source quality (PSNR) by 6dB.

### The traditional separation approach

Separation of source coding and channel coding is a major principle of traditional video communication systems; Figure 1 shows such system design for wireless video delivery.

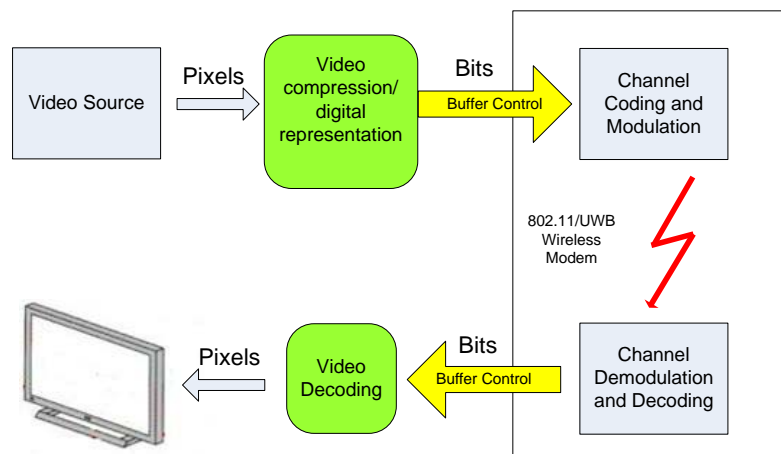


Figure 1: Traditional system design based on source-channel separation

Source-channel separation leads to modular system design allowing independent optimization of source and channel coders. Separation also allows interoperability by providing a common digital interface. So what is wrong with this modular, traditional approach?

It turns out that several major problems occur when the separation principle is used for wireless video transmission. First, separation theorem assumes known channel capacity. But this is never the case in wireless communication where the channel is varying and therefore its capacity is unknown in advance. Thus, Shannon's optimality claim does not hold, and there is a performance loss due to separation. Second, a separation based system is non-robust. The source code (compressed data) is sensitive, requiring that its output is protected by a strong channel code with guaranteed performance (to guarantee very low BER) which is hard to achieve. Finally, separation leads to a highly complex solution with a large delay due to compression and the need for reliable data transfer. In summary, realistic systems based on the separation principle operate far away from the optimal possible performance.

### The Joint Source-Channel Coding (JSCC) approach

To overcome these problems, the JSCC approach is used. Such a system design is depicted in Figure 2.

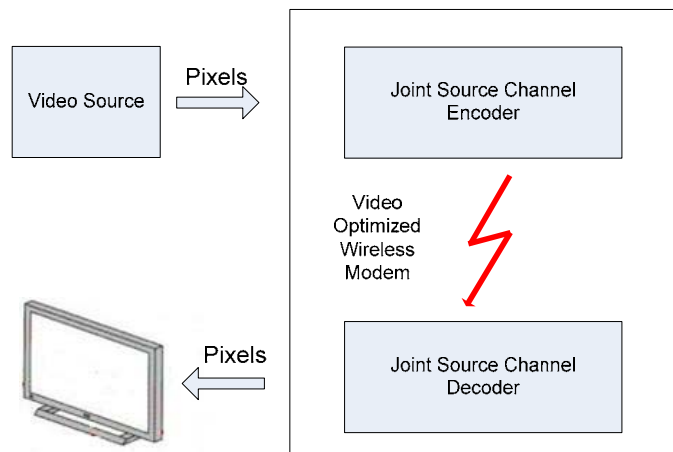


Figure 2: System design based on joint source-channel coding

The JSCC is composed of three elements:

1. Video processing and representation to prioritize the video components according to their importance. Some examples of prioritization include:
  - a. The most significant bits (MSBs) of a pixel are more important than the least significant bits (LSBs).
  - b. Lower spatial frequencies are more important than higher frequencies.
  - c. The luminance component is more important than the chrominance components.
2. Unequal error protection (UEP) to encode the most significant bits of the important components by better (lower rate) error correcting code, than the least significant bits of the less important components.
3. Combination of modulation and UEP to generate the proper constellation in the channel signal space. For example:
  - a. Use coarser constellation for the important components and finer constellation for the less important components.
  - b. Use "noisier" frequency bands for the less important components.

- c. Use space-time coding for better protection of the important components.

This process translates the video pixels, with a very low latency of several image lines, to modulated symbols, e.g., orthogonal frequency division modulation (OFDM) symbols. A block diagram of a joint source-channel coding is depicted in Figure 3.



Figure 3: JSCC

### Advantages of JSCC over traditional approach

The approach, based on joint source-channel coding, has several profound advantages over the traditional approach. First, since the video components and their bit representation are not equally important, JSCC uses unequal error protection (UEP) while traditional systems provide equal protection to all bits. Thus, in traditional systems the most significant bits of the important components are not protected enough, while the least significant bits of the less important video components are protected too much (and consequently waste channel resources). This situation is depicted in Figure 4.

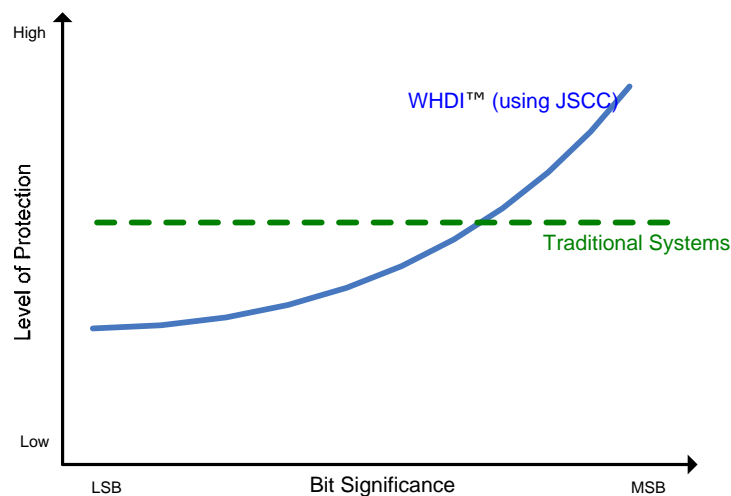


Figure 4: Unequal error protection vs. equal protection

The JSCC approach enables a better utilization of the available channel capacity, even when it is varying. Traditional systems should work at a rate that is below the worst-case channel capacity, since otherwise the video communication will not be reliable. By allocating whatever available capacity to send information that is less sensitive, JSCC utilizes capacity almost to its fullness. It should be noted that some traditional systems can “average out” short periods when the capacity decreases abruptly, but this requires large buffers leading to high complexity and delay. Other traditional systems adapt to the varying channel characteristics by using feedback. But feedback is cumbersome and requires complex return channel. Furthermore, since the wireless channel can change quickly, by the time

the transmitter gets the feedback information, the channel has been changed. The comparison in channel utilization is described in Figure 5.

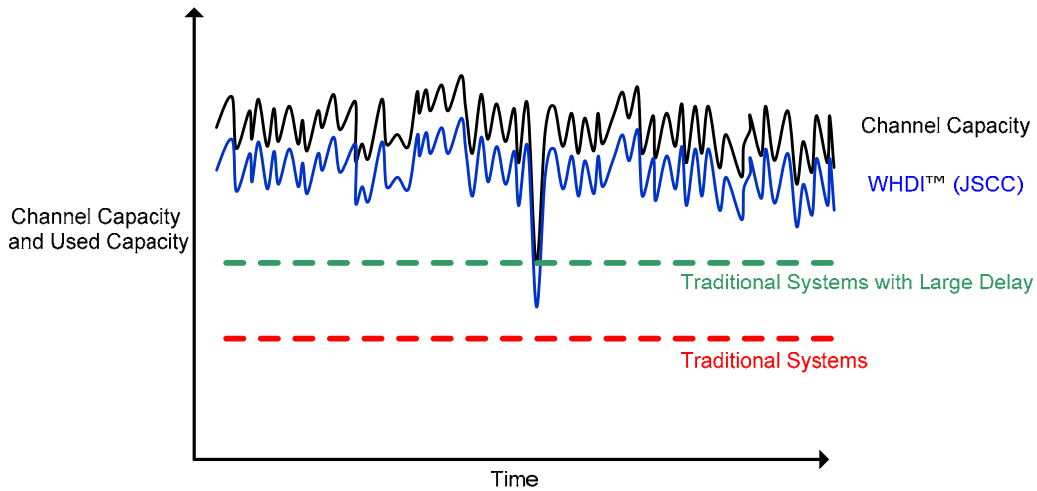


Figure 5: Utilization of channel capacity

The main advantage of JSCC over traditional systems lies in its ability to gracefully adapt to the varying channel capacity and the channel SNR. Traditional systems suffer from a threshold effect – They must guarantee a minimal SNR, otherwise the entire communication fails. To lower the threshold, traditional systems reduce the rate, e.g. by using deeper compression. But then the quality degrades. In any case, traditional systems always have a “quality ceiling,” where the picture quality cannot improve, even if the channel becomes better, over the pre-defined quality associated with the worst-case design. The JSCC approach does not have a threshold SNR, and the picture quality improves when channel conditions improve. The video quality of traditional and JSCC in varying channel conditions is depicted in Figure 6.

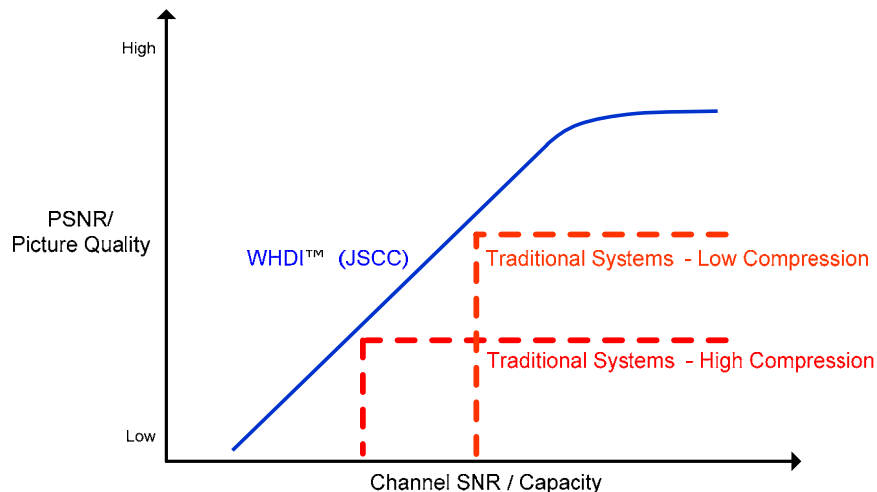


Figure 6: Graceful adaptation vs. threshold/ceiling effects

The CSNR is mostly between 20-30dB for the video-modem used by WHDI™, allowing a PSNR of over 40-45dB for almost perfect reconstruction. But even if the CSNR occasionally drops, the resulting PSNR is still good, while the traditional system will then completely break down.

## Robustness and complexity

Robustness is a highly important feature for the wireless video connectivity solution. Following the above, JSCC is much more robust to the varying characteristics of the wireless channel than the traditional approach, and so it is much less sensitive to obstacles and distance variations. Figure 7(a) shows a typical graph representing the probability that in a given instance, the channel supports a bit rate that is greater than  $B$ . To ensure an operating system, the traditional approach must work at a point where the attainable bit rate is achieved with 99.99 percent (otherwise, there will always be re-transmissions). As can be seen from this typical graph, this rate can be, say, 5 times less than the average capacity and more than 10 times the maximal capacity [see Figure 7(b)]. Even at that work point, there is always the chance (say, 0.01 percent) that the channel will not even support the required bit rate, and so the system based on the separation principle will collapse and result in large errors, especially if the data had to be compressed beforehand to fit the available bit rate at the work point. On the other hand, the JSCC approach allows use of the instantaneous capacity and so it enjoys a much higher capacity, which can be translated to better video quality [as can be seen in Figure 7(c)], yet it does not have to worry about the chance that the capacity and available bit rate drops below some specific value.

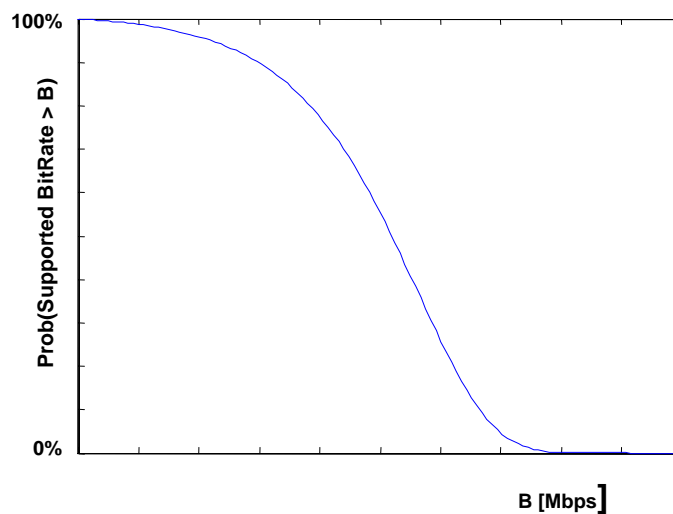


Figure 7(a): *The Wireless Fading Channel – Probabilistic behavior*

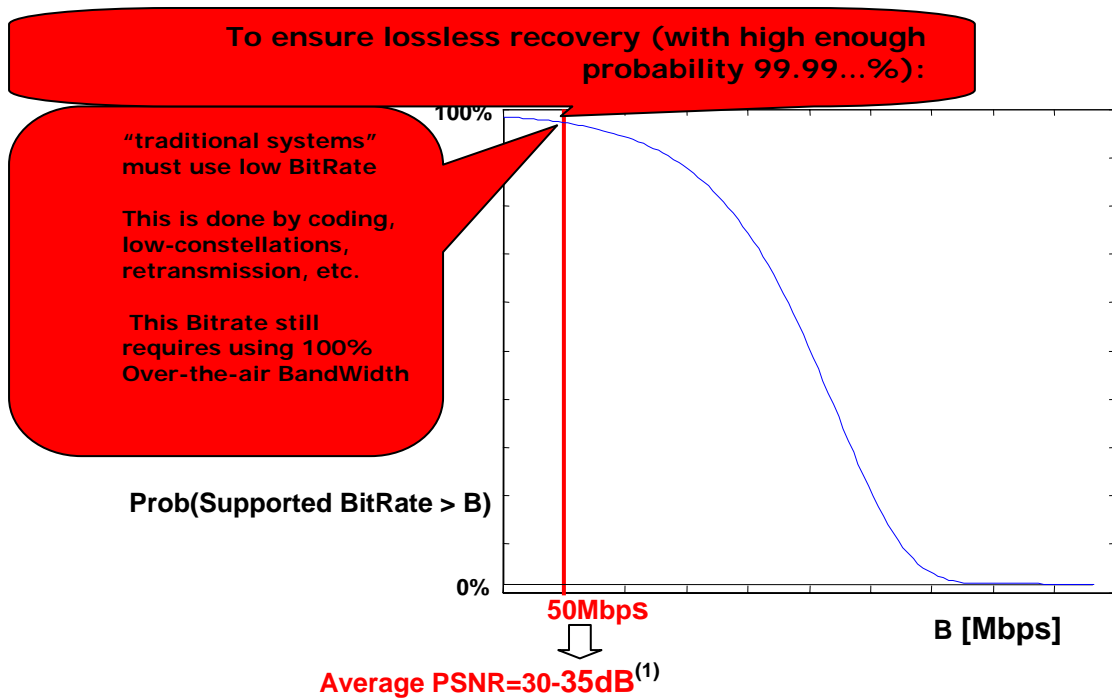


Figure 7(b): Work point of traditional systems

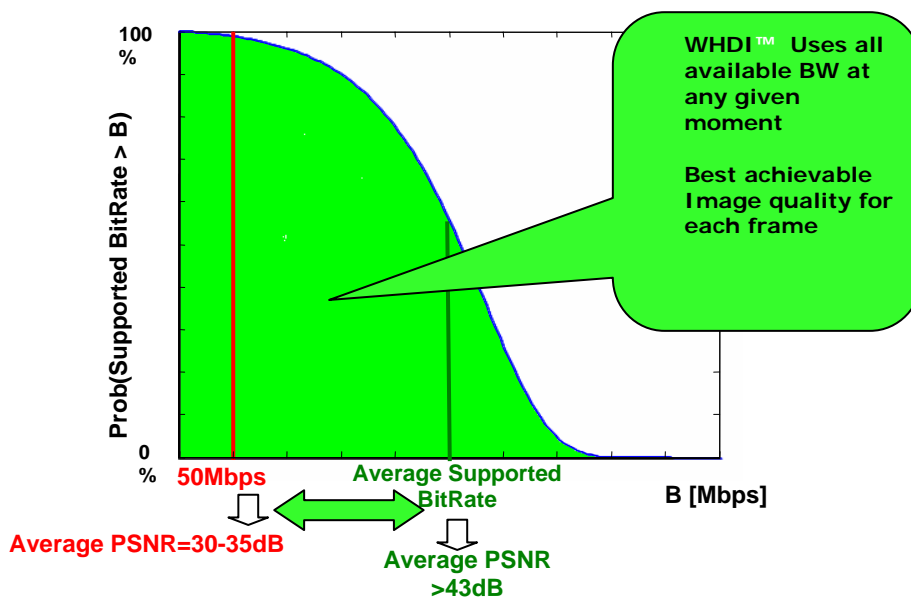


Figure 7(c): JSCC performance – PSNR is better by 8-13dB than traditional systems over the same channel

The JSCC performance and robustness are attained with much less computation and memory complexity and much simpler system architecture than the traditional systems. JSCC does not require complex HDTV compression. It does not require large buffers for compression and for modem re-transmissions. Its latency is very small (less than 1 ms). It can work essentially one way, with no feedback, without compromising the communication link and the video quality. This is an important feature since it allows a natural point-to-multipoint system architecture. Its built-in robustness and adaptability allow the video modem based on JSCC to achieve a 10 times improvement over traditional systems!

**Wrap-up**

The following table provides a comparison of AMIMON’s JSCC vs. traditional systems based on separation, where both systems are operating over the same wireless channel. The table indicates the advantages of JSCC.

	AMIMON JSCC used in WHDI™	“Traditional Systems”
<b>Quality</b>	Average 43dB PSNR	Average 30-35dB PSNR
<b>Robustness</b>	Superb	BER may cause undecodable frames
<b>Recompression</b>	None	Needed: May cause artifact enhancement
<b>System Cost and Complexity</b>	Simple: Virtually no latency	DSP HW on both ends (\$\$) Possible delay and memory (\$\$) More complex system (\$\$)
<b>Video-Audio Sync</b>	Less than 1msec latency	Possible audio-video sync issue