MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Han-Kobayashi and Dirty-Paper Coding for Superchannel Optical Communications

Koike-Akino, T.; Kojima, K.; Millar, D.S.; Parsons, K.; Kametani, S.; Sugihara, T.; Yoshida, T.; Ishida, K.; Miyata, Y.; Matsumoto, W.; Mizuochi, T.

TR2015-007 January 2015

Abstract

Superchannel transmission is a candidate to realize Tb/s-class high-speed optical communications. In order to achieve higher spectrum efficiency, the channel spacing shall be as narrow as possible. However, densely allocated channels can cause nonnegligible inter-channel interference (ICI) especially when the channel spacing is close to or below the Nyquist bandwidth. In this paper, we consider joint decoding to cancel the ICI in dense superchannel transmission. To further improve the spectrum efficiency, we propose the use of Han-Kobayashi (HK) superposition coding. In addition, for the case when neighboring subchannel transmitters can share data, we introduce dirty-paper coding (DPC) for pre-cancellation of the ICI. We analytically evaluate the potential gains of these methods when ICI is present for sub-Nyquist channel spacing.

IEEE Journals of Lightwave Technology

Han–Kobayashi and Dirty–Paper Coding for Superchannel Optical Communications

Toshiaki Koike-Akino, Keisuke Kojima, David S. Millar, Kieran Parsons, Soichiro Kametani, Takashi Sugihara, Tsuyoshi Yoshida, Kazuyuki Ishida, Yoshikuni Miyata, Wataru Matsumoto, and Takashi Mizuochi

***Abstract*—Superchannel transmission is a candidate to realize Tb/s-class high-speed optical communications. In order to achieve higher spectrum efficiency, the channel spacing shall be as narrow as possible. However, densely allocated channels can cause non-negligible inter-channel interference (ICI) especially when the channel spacing is close to or below the Nyquist bandwidth. In this paper, we consider joint decoding to cancel the ICI in dense superchannel transmission. To further improve the spectrum efficiency, we propose the use of Han–Kobayashi (HK) superposition coding. In addition, for the case when neighboring subchannel transmitters can share data, we introduce dirty–paper coding (DPC) for pre-cancellation of the ICI. We analytically evaluate the potential gains of these methods when ICI is present for sub-Nyquist channel spacing.**

***Index Terms*—Superchannel optical communications, Han–Kobayashi coding, dirty–paper coding, joint decoding**

I. INTRODUCTION

RAPID demand of increasing data rates in optical communications has necessitated high-throughput technologies, such as faster-than Nyquist [1]–[3] and superchannel transmission [4]–[14], where parallel subchannel transmitters multiplex independent data with different wavelengths to increase the total throughput without increasing the processing speed of each subchannel transceiver. The spectrum efficiency can increase as the channel spacing decreases. However, inter-channel interference (ICI) can be a major limiting factor to realize a dense channel allocation. Therefore, the ICI is often desired to be minimized.

In superchannel transmission, the interference from adjacent channels may be caused by both linear crosstalk [12]–[14] and nonlinear crosstalk [10]–[12]. With nonlinear crosstalk, the interference level becomes non-identical in different channels. In order to balance the interference, the launch power of each subchannel transmitter can be adjusted [10]–[12]. However, optimizing the power levels of all channels is cumbersome because changing the power of one channel affects other channels through fiber nonlinearity. Moreover, we have shown in [12] that rate control for all subchannels can be better than power control to increase the spectrum efficiency. It was also confirmed that joint rate and power control offers only a marginal improvement in superchannel transmission. Therefore, we consider uniform power transmission for all subchannels with appropriate rate control.

Although linear crosstalk is a simpler problem than the nonlinear crosstalk, mitigating ICI is still not straightforward because the parallel subchannel receivers normally operate independently. If all subchannel receivers can share signals from different channels, ICI cancellation can be accomplished with multiple-input multiple-output (MIMO) signal processing as studied in [14]. It was shown that the MIMO equalization can realize a super-dense channel spacing of 50% baud rate. However, the subchannel receivers cannot work in parallel for MIMO equalization and require high-speed interconnects to share received signals. Hence, we consider an alternative method in this paper. From the information theoretical aspect, it has been proved in [15], [16] that very strong interference does not limit the channel capacity because interference itself conveys meaningful data and the interference data can be decoded and canceled out for strong interference cases. This theory suggests that suppressing undesired interference may not always be the best solution if the receivers employ some kind of interference cancellation, such as joint decoding [15].

However, joint decoding does not work well for weak interference cases because it is hard to decode the interference signals. For such cases, the so-called Han–Kobayashi (HK) scheme [16]–[19] has been proposed to achieve higher data rates. The HK scheme uses superposition coding of two split data; one data is decoded at a desired receiver and the other data is decoded at all adjacent receivers. By controlling the power split for those two data, we can resolve the drawbacks of both the conventional decoding and the joint decoding.

If adjacent transmitters can share data, the spectrum efficiency can be further improved by introducing dirty–paper coding (DPC) [20]–[22]. Costa proved in [20] that the channel capacity cannot be limited by interference signals no matter how strong it is if the transmitter can exploit the side information of the interference signals. In practice, such a DPC scheme can be realized by a modulo-lattice precoding [21] to cancel the ICI signals in advance at each transmitter. The DPC scheme may be useful for superchannel optical

Manuscript received October 6, 2014; revised December 2, 2014 and January 16, 2015; accepted January 22, 2015.

T. Koike-Akino, K. Kojima, D. S. Millar, and K. Parsons are with Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139, U.S.A.

S. Kametani, T. Sugihara, T. Yoshida, Y. Miyata, and W. Matsumoto are with Information Technology R&D Center, Mitsubishi Electric Corporation, 5-1-1, Ofuna, Kamakura, Kanagawa 247-8501, Japan.

K. Ishida is with Communication Networks Center, Mitsubishi Electric Corporation, 8-1-1, Tsukaguchi, Amagasaki, Hyogo 661-8661, Japan.

T. Mizuochi is with Advanced Technology R&D Center, Mitsubishi Electric Corporation, 8-1-1, Tsukaguchi, Amagasaki, Hyogo 661-8661, Japan.

Color versions of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier xx.xxx/JLT.2015.xxx

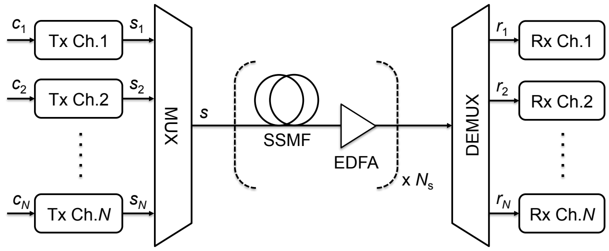


Fig. 1. Superchannel optical fiber communications systems.

communications, in which adjacent subchannel transmitters share data for ICI pre-cancellation.

In this paper, we theoretically investigate the potential benefits obtained by the HK and DPC schemes for superchannel optical transmission. In [13], we have reported a preliminary analysis for the simplest case of a two-channel system in back-to-back configuration. In the presence of strong ICI with dense channel spacing, a significant gain up to 2.5-times higher spectrum efficiency has been observed. To the best of authors' knowledge, our study has been the first attempt applying HK or DPC to optical communications in the literature. This paper extends our preliminary work. The major contributions of this paper are two-fold: i) we introduce a scalable way to extend HK and DPC to more than two channels, ii) we theoretically analyze the benefit of HK and DPC for more realistic optical links in the presence of fiber nonlinearity.

II. Superchannel Transmission Systems

A. Superchannel Signal Description

Fig. 1 depicts a schematic of superchannel optical transmission for N channels. The n-th subchannel transmitter for $n \in \{1, 2, \ldots, N\}$ sends a band-limited signal s_n at a carrier frequency of f_n. The equivalent complex signal at a time of t can be expressed as

$$s_n(t) = \sum_k g_{\mathrm{tx}}(t - kT_{\mathrm{s}}) c_n(k) \mathrm{e}^{\jmath 2\pi f_n t}, \tag{1}$$

where $g_{\mathrm{tx}}(t)$ is an impulse response of a transmitter filter, T_{s} is the symbol duration, $c_n(k)$ is a constellation symbol transmitted at the k-th symbol instance, and $\jmath = \sqrt{-1}$ is the imaginary unit. The multiplexer sums up all signals as

$$s(t) = \sum_{n=1}^{N} s_n(t). \tag{2}$$

The transmitted signal $s(t)$ propagates through N_{s} spans of standard single-mode fiber (SSMF) with Erbium-doped fiber amplifier (EDFA). We assume that EDFA compensates for the attenuation in each fiber span, and adds amplified spontaneous emission (ASE) noise. After optical demultiplexer, the n-th subchannel receiver obtains

$$r_n(t) = g_{\mathrm{rx}}(t) \star \big(h(t) \star s(t) + z(t)\big) \mathrm{e}^{-\jmath 2\pi f_n t}, \tag{3}$$

where $g_{\mathrm{rx}}(t)$ is an impulse response of a receiver filter, an operator $\star$ denotes a convolution, $h(t)$ is an impulse response of fiber due to linear chromatic dispersion, and $z(t)$ is an effective additive noise, which includes ASE noise due to EDFAs and nonlinear interference (NLI) due to Kerr fiber nonlinearity. The ASE noise is modeled as circular-symmetric white Gaussian distribution [23]. We also use the Gaussian noise (GN) model [24] for NLI. The received signal is rewritten as

TABLE I
Main Parameters for Superchannel Fiber Communications

Number of channels	$N = 3$
Symbol rate	$B = 32$Gbaud
Normalized channel spacing	$\delta f = 0.50, 0.95$
TRx filter	Root-raised-cosine (RRC)
Roll-off factor	$\alpha = 0.01$
Launch power	-2.0dBm per channel
Fiber type	Standard single-mode fiber (SSMF)
Fiber distance	80km per span
Fiber loss	0.20dB/km
Chromatic dispersion	16.5ps/nm/km
Nonlinearity coefficient	1.3/W/km
Amplifier type	Erbium-doped fiber amplifier (EDFA)
Noise figure	4dB
Number of spans	$N_{\mathrm{s}} = 10$

$$r_n(t) = \sum_{m=1}^{N} \sum_k h_{n,m}(t - kT_{\mathrm{s}}) c_m(k) + z_n(t), \tag{4}$$

where $h_{n,m}(t)$ is an impulse response of the effective total filter from the m-th subchannel transmitter to the n-th subchannel receiver and $z_n(t)$ is a band-limited additive noise. The total impulse response $h_{n,m}(t)$ is a convolution of $g_{\mathrm{tx}}(t)$, $h(t)$, and $g_{\mathrm{rx}}(t)$ with a carrier frequency offset of $f_n - f_m$. The term of $h_{n,n}(t)$ corresponds to the desired signal for the n-th subchannel receiver, whereas the other responses, i.e., $h_{n,m}(t)$ for $m \neq n$, contribute to undesired ICI.

For simplicity, we consider a uniform channel spacing over all N channels. Letting $B = 1/T_{\mathrm{s}}$ be the baud rate, the channel spacing normalized by the baud rate is defined as

$$\delta f = \frac{f_{n+1} - f_n}{B}. \tag{5}$$

We assume root-raised-cosine (RRC) filters with a roll-off factor of α for both the transmitter filter $g_{\mathrm{tx}}(t)$ and the receiver filter $g_{\mathrm{rx}}(t)$. In fact, the RRC filter may not be the best choice when the receivers employ equalization for joint decoding. We leave the filter optimization as future work. Table I lists several parameters under consideration. We use $N = 3$ channels, $B = 32$Gbaud, $\delta f = 0.50$ or 0.95 channel spacing, $\alpha = 0.01$ roll-off factor, and $N_{\mathrm{s}} = 10$ spans unless otherwise stated. We mainly consider a sub-Nyquist channel spacing, i.e., $\delta f < 1$, to analyze the impact of ICI.

B. Nonlinear Interference (NLI)

To evaluate performance in realistic fiber communications systems, we consider NLI in addition to ICI and ASE noise. We use the GN model [24] to calculate the power spectrum density of NLI induced by Kerr fiber nonlinearity. The GN model has been studied in several ways [25]–[27] and widely used for analysis of nonlinear fiber communications [12], [28]. Let $G_{\mathrm{NLI}}(f)$ be NLI power spectrum density at a frequency of f, before passing through receiver filter. The NLI spectrum, $G_{\mathrm{NLI}}(f)$, calculated by the GN model is shown in Fig. 2

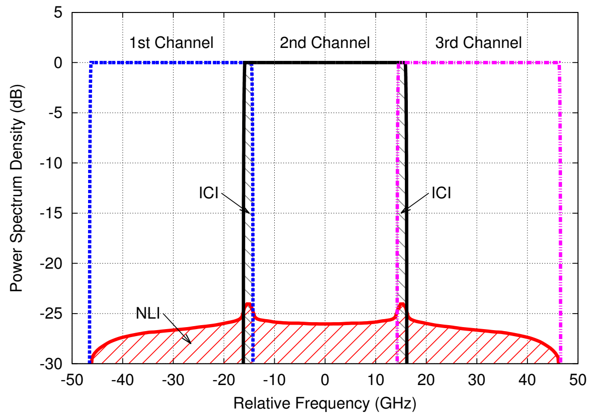


Fig. 2. Power spectrum density of nonlinear interference (NLI) and inter-channel interference (ICI) for $\delta f = 0.95$ and $\alpha = 0.01$ over $N_\mathrm{s} = 10$ spans of SSMF at launch power of -2.0dBm.

for parameters given in Table I. As shown in this figure, the NLI is usually larger at the center subchannel and smaller at the two outermost subchannels [11], [12]. For sub-Nyquist channel spacing, the NLI spectrum has some peaks around the overlapping band because of the impact of ICI.

Using numerical calculation of $G_\mathrm{NLI}(f)$ based on the GN model, we compute an effective noise variance of $z_n(t)$ for each channel. More specifically, the noise variance for the n-th subchannel receiver becomes

$$\sigma_n^2 = \int \left| G_\mathrm{rx}(f - f_n) G_\mathrm{NLI}(f) \right|^2 \mathrm{d}f + \sigma_\mathrm{ASE}^2, \tag{6}$$

where $G_\mathrm{rx}(f)$ denotes the transfer function of the receiver filter and σ_ASE^2 is the variance of the ASE noise passed through the receiver filter. The ASE noise power σ_ASE^2 due to EDFA is a function of the number of spans, the baud rate, the noise figure, the fiber loss to compensate for, and photon energy [23]. Although the NLI and ASE passed through the receiver filter are not really white, we assume whitened Gaussian noise with an equivalent noise variance of σ_n^2 at the n-th subchannel receiver for simplicity of analysis.

C. Inter-Channel Interference (ICI)

Provided that the transmitting constellation symbol $c_n(k)$ has unit energy on average, the desired signal power at the n-th subchannel receiver is given as $\int |H_{n,n}(f)|^2 \mathrm{d}f$, while the ICI power from the m-th subchannel (for $m \neq n$) is expressed as $\int |H_{n,m}(f)|^2 \mathrm{d}f$, which is an accumulated power of an overlapping ICI band as shown in Fig. 2. Here, $H_{n,m}(f)$ is the Fourier transform of $h_{n,m}(t)$, i.e., the transfer function of a linear effect through overall the fiber communications systems from the m-th subchannel transmitter to the n-th subchannel receiver. The signal-to-noise ratio (SNR) and signal-to-interference ratio (SIR) can be, respectively, calculated as

$$\rho_n = \frac{\int |H_{n,n}(f)|^2 \mathrm{d}f}{\sigma_n^2}, \tag{7}$$

$$\frac{1}{\beta} = \frac{\int |H_{n,n}(f)|^2 \mathrm{d}f}{\int |H_{n,n+1}(f)|^2 \mathrm{d}f}, \tag{8}$$

where we denote β as an inverse SIR. The transmission performance can be generally characterized by signal-to-interference plus noise ratio (SINR), i.e., $\rho_n/(1 + 2\beta\rho_n)$ for the center subchannel ($n = 2$), wherein 2β comes from the ICI of the first subchannel and that of the last subchannel. Note that we have $\beta \leq 1$ because the ICI cannot exceed the signal power for superchannel transmission.

In reality, the ICI signal has memory as discussed in [13], where we evaluated 3-tap and 19-tap equalizers to decode the ICI signal. To shorten the ICI memory, we have used a relatively large roll-off factor [13]. However, in this paper, we consider more realistic optical communications systems with chromatic dispersion, which may have longer memory size required for equalization. For this case, the memory size issue for ICI equalization can be less important in comparison to chromatic dispersion compensation. Consequently, we consider a matched filter bound (MFB) [3], [29], assuming infinite-tap equalization or frequency-domain equalization. It has been shown that the MFB performance can be achieved by practical turbo equalization [2], [30], [31]. Since the filter memory issue for equalization is not our main focus in this paper, we use near-zero roll-off factor of $\alpha = 0.01$, which is applicable to modern optical communications [9].

III. Han–Kobayashi and Dirty–Paper Coding

Here, we introduce the HK scheme and DPC to cope with the ICI. We first describe a conventional scheme and joint decoding, both of which are special cases of the HK scheme. Those four different methods are illustrated in Fig. 3.

A. Conventional Decoding

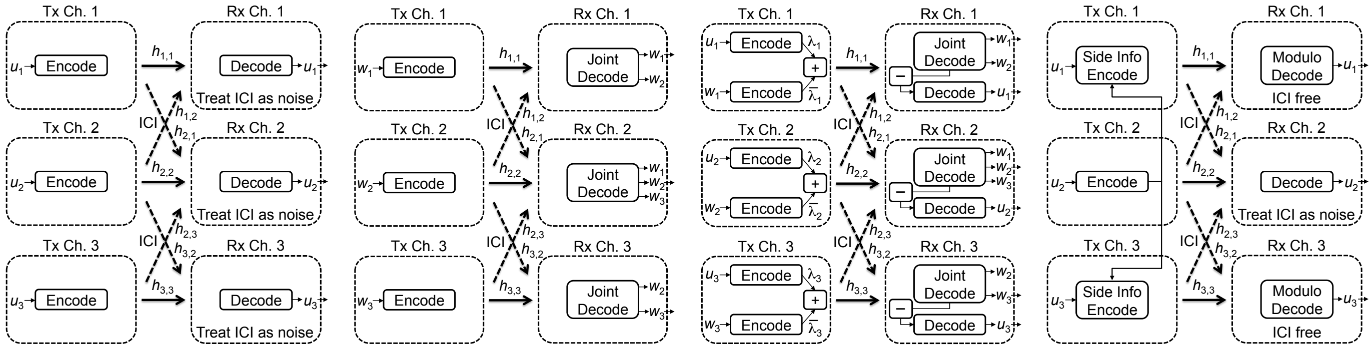
Fig. 3(a) depicts a schematic of conventional decoding scheme. In the conventional scheme, each subchannel transmitter sends independent data in parallel and the corresponding subchannel receiver decodes only the desired data. For this case, the ICI components in $h_{n,m}(t)$ for $n \neq m$ are treated as additional noise. Using the SNR ρ_n from (7) and the inverse SIR β from (8), the achievable sum rate for N-channel transmission is expressed as follows:

$$R_\mathrm{conv} = C\Big(\frac{\rho_1}{1 + \beta\rho_1}\Big) + \sum_{n=2}^{N-1} C\Big(\frac{\rho_n}{1 + 2\beta\rho_n}\Big) + C\Big(\frac{\rho_N}{1 + \beta\rho_N}\Big), \tag{9}$$

where $C(\rho) = \log_2(1+\rho)$ denotes a Shannon limit for additive white Gaussian noise at an SNR of ρ. In (9), the SINR replaces the SNR to take ICI and NLI into account. The first and last terms are for the two outermost subchannels, which may have lower ICI crosstalk than others. Note that the achievable rate degrades rapidly with an increased ICI power β, which appears in the denominator of the SINR. Hence, it is preferred to minimize the ICI power in general. This situation is different for the following methods.

B. Joint Decoding

Fig. 3(b) illustrates joint decoding scheme. For joint decoding, the desired signal and the ICI signals are jointly decoded



(a) Conventional Decoding (b) Joint Decoding (c) Han–Kobayashi Coding (d) Dirty–Paper Coding

Fig. 3. Different coding/decoding schemes for superchannel transmission (u_n is private data for individual decoding and w_n is public data for joint decoding).

at each subchannel receiver. Note that this joint decoding does not need to share information between neighboring subchannel receivers, unlike MIMO equalization [14]. With joint decoding, each receiver can achieve interference-free multiple-access channel (MAC) capacity [15].

For the three-channel case, the achievable sum rate with joint decoding is expressed as

$$R_{\mathrm{joint3}} = \min\Big\{C\big((1+2\beta)\rho_2\big), C\big(2\beta\rho_2\big) + C\big(\rho_2\big), \\ C\big((1+\beta)\rho_2\big) + C\big(\rho_1\big), 2C\big(\rho_1\big) + C\big(\rho_2\big)\Big\}. \quad (10)$$

The first term of R_{joint3} is based on the total SNR at the center subchannel, the second term is based on the sum of two ICI signal powers at the center subchannel, the third term is based on the sum of the desired signal and one of the ICI signals at the center subchannel, and the last term is based on individual signal power. Here, we assumed $\rho_2 \leq \rho_1 = \rho_3$. Although the expression of the achievable rate becomes complicated for more than 3 channels, we can analyze any number of channels in a straightforward manner.

As shown in (10), the increased interference power β can improve the achievable rate. It suggests that minimizing the ICI power may not be the best strategy. Note that the rate can be severely constrained when the interference power is very weak [16]. To solve this problem, we may adaptively select the decoding strategy from the joint decoding to the conventional decoding if the ICI signal is too weak to decode. We consider such an adaptive joint decoding for analysis.

C. Han–Kobayashi (HK) Coding

One of the best known strategies to solve the drawbacks of both the conventional decoding and the joint decoding for weak interference channels is the so-called HK scheme [16]–[18], which uses superposition coding and partial joint decoding. The HK scheme is depicted in Fig. 3(c). For each transmitter, the coded sequence c_n is a superposition of two codewords of source data u_n and w_n. One is for private data, which are decoded only at the intended receiver as in conventional decoding. The other is used for public data, which are decoded at all adjacent receivers as in joint decoding. Two codewords (for private u_n and public data w_n) are superimposed with a power splitting factor of λ_n and $\bar{\lambda}_n = 1 - \lambda_n$. At each subchannel receiver, all public data from adjacent subchannel transmitters (e.g., w_{n-1}, w_n, and w_{n+1}) are first decoded jointly. The decoded public data are used to cancel out for another decoder to decode private data u_n. Note that the unintended public data, i.e., w_k for $k \neq n$ at the n-th subchannel receiver, are discarded in the end after joint decoding and ICI cancellation.

The benefit of the HK scheme comes from the power splitting. As a special case, setting $\lambda_n = 1$ for all channels reduces to the conventional decoding strategy, whereas setting $\lambda_n = 0$ for all channels becomes the joint decoding strategy. By controlling the power splitting, we can mitigate the ICI impact for the private data while achieving the joint decoding gain for public data. The HK scheme was originally proposed for a two-channel case. There is no obvious way to extend the HK scheme for more than two channels in the literature. Even for the two-channel case, optimizing the power splitting factors is not straightforward, as addressed in [17], [18]. In [18], it was shown that identical power splitting, i.e., $\lambda_1 = \lambda_2$, is not always optimal even for symmetric interference channels.

In order to extend to more than two channels, we propose a scalable way to control power splitting for all N channels, in which two factors λ_{odd} and λ_{even} are used for odd-channel splitting λ_{2i-1} and even-channel splitting λ_{2i}, respectively. To make it even simpler, we consider five strategies with only one parameter λ as follows:

1) symmetric: $\lambda_{\mathrm{odd}} = \lambda_{\mathrm{even}} = \lambda$,
2) even-private asymmetric: $\lambda_{\mathrm{even}} = 1$, $\lambda_{\mathrm{odd}} = \lambda$,
3) even-public asymmetric: $\lambda_{\mathrm{even}} = 0$, $\lambda_{\mathrm{odd}} = \lambda$,
4) odd-private asymmetric: $\lambda_{\mathrm{odd}} = 1$, $\lambda_{\mathrm{even}} = \lambda$, and
5) odd-public asymmetric: $\lambda_{\mathrm{odd}} = 0$, $\lambda_{\mathrm{even}} = \lambda$.

We will later compare all the five strategies. For three-channel transmission with two splitting parameters λ_{even} and λ_{odd}, the achievable rate of HK scheme can be expressed as

$$R_{\mathrm{HK3}} = 2C\Big(\frac{\lambda_{\mathrm{odd}}\rho_1}{1+\beta\lambda_{\mathrm{even}}\rho_1}\Big) + C\Big(\frac{\lambda_{\mathrm{even}}\rho_2}{1+2\beta\lambda_{\mathrm{odd}}\rho_2}\Big) + \\ \min\Big\{C\big((\bar{\lambda}_{\mathrm{even}} + 2\beta\bar{\lambda}_{\mathrm{odd}})\rho_2'\big), C\big(2\beta\bar{\lambda}_{\mathrm{odd}}\rho_2'\big) + C\big(\bar{\lambda}_{\mathrm{even}}\rho_2'\big), \\ C_{12} + C\big(\bar{\lambda}_{\mathrm{odd}}\rho_1'\big), 2C\big(\bar{\lambda}_{\mathrm{odd}}\rho_1'\big) + C\big(\bar{\lambda}_{\mathrm{even}}\rho_2'\big)\Big\}, \quad (11)$$

where

$$C_{12} = \min\Big\{C\big((\bar{\lambda}_{\mathrm{even}} + \beta\bar{\lambda}_{\mathrm{odd}})\rho_2'\big), C\big((\bar{\lambda}_{\mathrm{odd}} + \beta\bar{\lambda}_{\mathrm{even}})\rho_1'\big)\Big\},$$

$$\rho_1' = \frac{\rho_1}{1+(\lambda_{\text{odd}}+\beta\lambda_{\text{even}})\rho_1},$$
$$\rho_2' = \frac{\rho_2}{1+(\lambda_{\text{even}}+2\beta\lambda_{\text{odd}})\rho_2}.$$

The first two terms in (11) are based on conventional decoding as in (9) for private data, and the reminder is based on joint decoding as in (10) for public data. Parameters ρ_n' correspond to effective SINRs before ICI cancellation of the public data.

D. Dirty–Paper Coding (DPC)

When subchannel transmitters know the message of adjacent subchannel transmitters, we can further improve the spectrum efficiency by introducing the DPC scheme [20], in which the ICI is pre-canceled at the transmitter, similar to Tomlinson-Harashima precoding. Fig. 3(d) shows the schematic of DPC scheme. The subchannel transmitter exploits encoded data from adjacent subchannel transmitters as side information to cancel the ICI by using modulo-lattice encoding [21]. The corresponding receiver can achieve an ICI-free signal with modulo-lattice decoding. Even with the DPC method, we cannot cancel the ICI at all the subchannel receivers because of causality for side information encoding. In this paper, we propose a comb-like DPC strategy, in which every even-channel transmitter sends encoded data to neighboring odd-channel transmitters to cancel the ICI with DPC. With this DPC method, the achievable sum rate becomes

$$R_{\text{DPC}} = \begin{cases} \displaystyle\sum_{i=1}^{N/2} C(\rho_{2i-1}) + \sum_{i=1}^{N/2-1} C\Big(\frac{\rho_{2i}}{1+2\beta\rho_{2i}}\Big) + C\Big(\frac{\rho_N}{1+\beta\rho_N}\Big), \\ \displaystyle\sum_{i=1}^{(N+1)/2} C(\rho_{2i-1}) + \sum_{i=1}^{(N-1)/2} C\Big(\frac{\rho_{2i}}{1+2\beta\rho_{2i}}\Big), \end{cases} \tag{12}$$

respectively, for the cases when N is even and odd. Note that a practical DPC based on repeat-accumulate codes and modulo-lattice precoding is reported by Erez and ten Brink [21], in which the optimized code closely approaches the theoretical DPC bound. In [22], it is also shown that practical joint source-channel coding achieves the DPC bound.

IV. Performance Analysis

We evaluate the performance of the HK and DPC schemes for parameters listed in Table I. To obtain the spectrum efficiency, the achievable sum rates (R_{HK}, R_{conv}, R_{joint}, and R_{DPC}) are normalized by a total bandwidth consumption of $B(1+\alpha+(N-1)\delta f)$.

First, we show the performance curves as a function of a channel spacing δf in Fig. 4. Here, we use an optimized launch power for each data point. For low- or no-ICI regimes above quasi-Nyquist spacing for $\delta f > 1-\alpha = 0.99$, there is no gain of the HK scheme compared to the conventional scheme. However, the HK and DPC schemes show significant tolerance against stronger ICI when the channel spacing narrows down, and can achieve $2\sim 4$ times higher spectrum efficiency than the conventional decoding for $\delta f < 0.85$. More remarkably, the HK scheme can realize 50% channel spacing with a high spectrum efficiency, which is comparable to the no-ICI performance at $\delta f = 1{+}\alpha = 1.01$. This suggests that the HK scheme

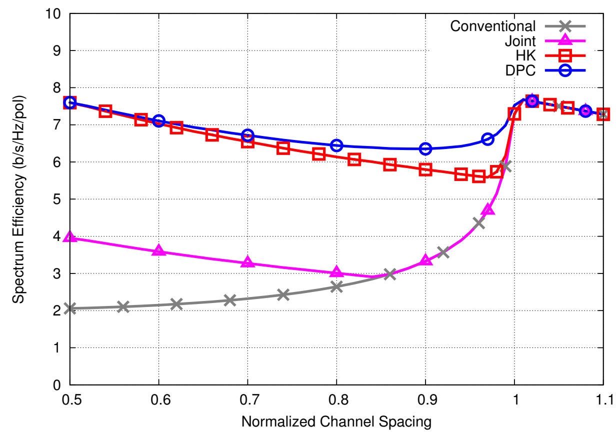


Fig. 4. Spectrum efficiency vs. channel spacing δf.

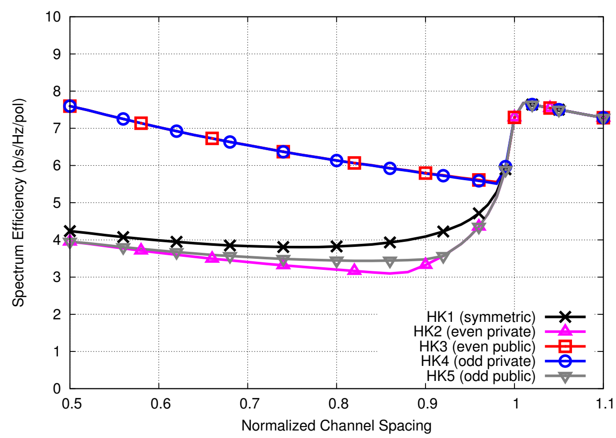


Fig. 5. Comparison between five strategies for HK scheme.

is a good candidate to realize such high-density superchannel transmission because the HK scheme does not require MIMO signal processing or high-speed interconnects between parallel subchannel receivers, unlike in [14]. For such a high-density channel spacing, the conventional decoding cannot send high-rate data especially at the center subchannel receiver because SINR is less than 0dB. Although the achievable spectrum efficiency for super-dense 50% channel spacing even with the HK scheme cannot outperform that for the zero-ICI case in quasi-Nyquist channel spacing, the performance curve against the channel spacing becomes much more robust around a normalized channel spacing of $\delta f = 0.50$, compared to quasi-Nyquist channel spacing. Note that joint decoding becomes useless for weak ICI regimes because the joint decoding requires high power for high-rate ICI signal to decode, as discussed in Section III-B.

In Fig. 5, we next compare five different strategies proposed for the HK scheme in Section III-C. It is seen that the first strategy with symmetric power split ($\lambda_{\text{odd}} = \lambda_{\text{even}}$) does not provide higher spectrum efficiency than the third strategy ($\lambda_{\text{even}} = 0$) and the fourth strategy ($\lambda_{\text{odd}} = 1$). The fifth strategy ($\lambda_{\text{odd}} = 0$) and the second strategy ($\lambda_{\text{even}} = 1$) are even worse than the symmetric power split. For all the channel spacing, the fourth strategy offers the best spectrum efficiency. This is preferable in terms of complexity because the HK

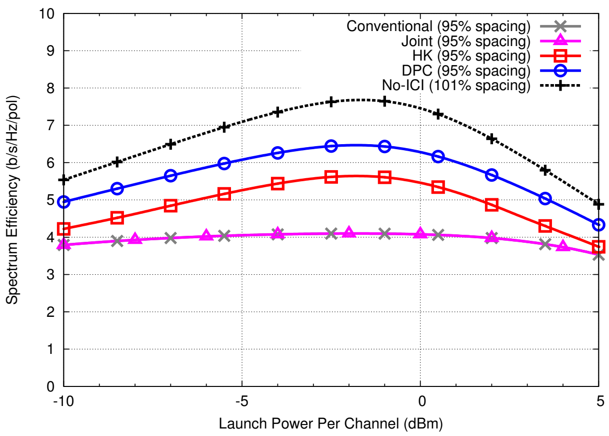


Fig. 6. Spectrum efficiency vs. launch power ($\delta f = 0.95$).

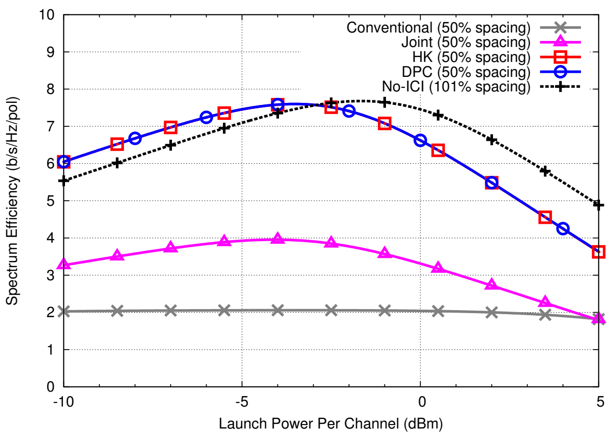


Fig. 7. Spectrum efficiency vs. launch power ($\delta f = 0.50$).

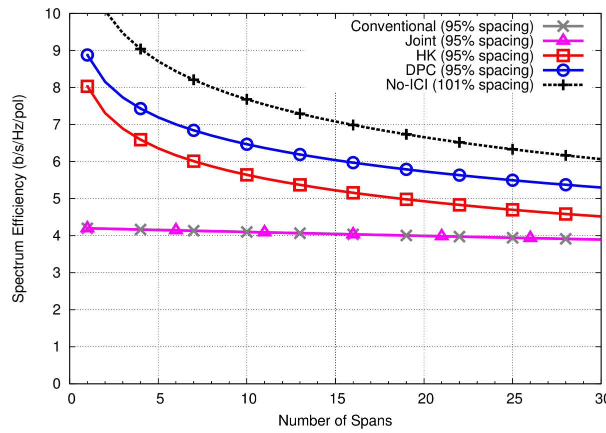


Fig. 8. Spectrum efficiency vs. span length N_s ($\delta f = 0.95$).

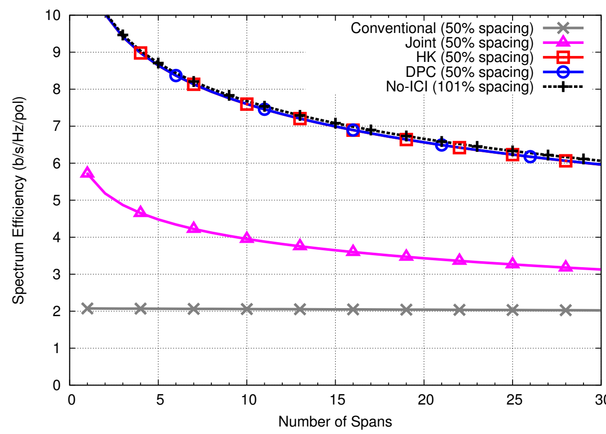


Fig. 9. Spectrum efficiency vs. span length N_s ($\delta f = 0.50$).

scheme with superposition coding is required at only the center subchannel transmitter, while the two-side subchannel transmitters can operate conventionally.

We then consider, in detail, two cases: i) the worst case of channel spacing at $\delta f = 0.95$, around which the HK scheme has the lowest spectrum efficiency as shown in Fig. 4, and ii) the best case of channel spacing at $\delta f = 0.50$, in which the HK scheme offers the highest gain from the conventional decoding. Both cases are sub-Nyquist spacing, while the first case has only 5% margin from Nyquist spacing. As a reference, we also present the performance of no-ICI case for quasi-Nyquist spacing of $\delta f = 1.01$.

In Figs. 6 and 7, we plot the performance as a function of launch power for $\delta f = 0.95$ and 0.50, respectively. The spectrum efficiency can be maximized at a launch power of -2dBm and -4dBm, respectively, for $\delta f = 0.95$ and 0.50. Note that high-density channel spacing creates strong ICI, leading to larger NLI noise because the NLI power is a cubic function of signal power [24]. It is shown that the conventional decoding has almost flat performance, which is because ICI dominates ASE and NLI. The HK and DPC schemes can compensate for such an ICI-limiting performance. For 50% spacing, the HK scheme achieves DPC performance, and its peak spectrum efficiency is comparable to the no-ICI case. Note that the HK scheme for $\delta f = 0.50$ can be better than no-ICI case for $\delta f = 1.01$ in linear regimes for a launch power below -4dBm because of the denser channel allocation.

We then show the spectrum efficiency versus the number of spans in Figs. 8 and 9 for $\delta f = 0.95$ and 0.50, respectively. Here, the launch power is optimized for all data points. Due to ICI, the conventional decoding has limited performance over fiber spans. The HK and DPC schemes significantly outperform the conventional decoding and approach no-ICI performance especially for highly dense channel spacings.

The HK and DPC schemes can be applied to any number of channels. We show the performance as a function of the number of channels N in Fig. 10 for a channel spacing of $\delta f = 0.50$. It is noted that high spectrum efficiencies close to no-ICI performance can be maintained by the HK and DPC schemes, in particular for the cases when N is an odd number. In contrast, joint decoding performs relatively well for $N = 2$ and 4, whereas a considerable loss can be seen for $N = 3$.

V. Conclusions

Through theoretical analyses, we have shown a significant benefit of HK scheme to tolerate against ICI in superchannel optical communications. Compared to conventional decoding, the HK scheme achieves $2 \sim 4$ times higher spectrum efficiency for a dense channel allocation. The HK scheme can be one of candidates to realize very high-density superchannels

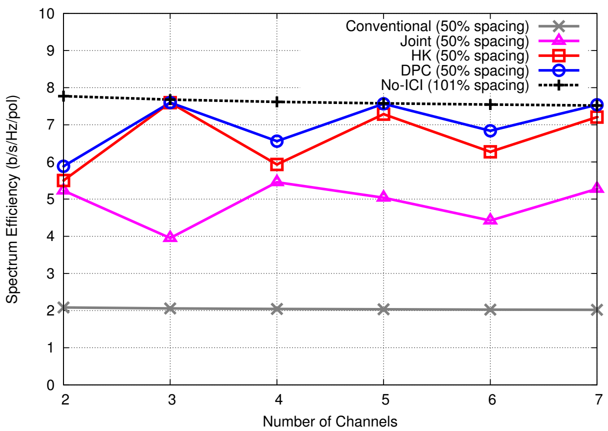


Fig. 10. Spectrum efficiency vs. number of channels N ($\delta f = 0.50$).

such as 50% channel spacing, while it does not require any MIMO equalization or high-speed interconnects between receivers. We have also evaluated DPC, showing a slightly better performance than the HK scheme. Although the HK and DPC schemes require additional complexity in order to maintain spectrum efficiency comparable to the zero-ICI case, it is expected that the ICI management techniques will enable increased robustness to practical hardware imperfections such as laser frequency drift and mistuning.

We have done some theoretical analyses in fiber communications systems with ASE and NLI based on the GN model, assuming NLI is additive Gaussian noise. If we can exploit NLI with nonlinear equalization, the HK scheme may be more useful because the total interference power can be stronger to decode. The generalization to handle nonlinear crosstalk remains as a future work. The demonstration of the HK and DPC methods using a practical modulation, coding, decoding, and equalization in a real fiber plant experiment will be also an important work in the future.

Acknowledgment

This work was in part supported by the Lambda Reach project of the National Institute of Information and Communications Technology (NICT), as part of a program of the Ministry of Internal Affairs and Communications (MIC) of Japan.

References

[1] E. Mazo, “Faster-than-Nyquist signaling,” *Bell System Technical Journal*, vol. 54, no. 8, pp. 1450–1462, Oct. 1975.
[2] G. Colavolpe, T. Foggi, A. Modenini, and A. Piemontese, “Faster-than-Nyquist and beyond: how to improve spectral efficiency by accepting interference,” *Optics Express*, vol. 19, no. 27, pp. 26600-26609, Dec. 2011.
[3] F. Rusek and J. B. Anderson, “Constrained capacities for faster-than-Nyquist signaling,” *IEEE Trans. Information Theory*, vol. 55, no. 2, pp. 764–775, Feb. 2009.
[4] B.-E. Olsson, A. Kristiansson, and A. Alping, “RF-assisted optical dual-carrier 112 Gbit/s polarization-multiplexed 16-QAM transmitter,” *OFC*, OM.K5, Mar. 2010.
[5] V. Sleiffer, D. van den Borne, V. Veljanovski, M. Kuschnerov, M. Hirano, Y. Yamamoto, T. Sasaki, S. L. Jansen, and H. de Waardt, “Transmission of 448-Gb/s dual-carrier POLMUX-16QAM over 1230 km with 5 flexi-grid ROADM passes.” *OFC*, OW4C-3, Mar. 2012.
[6] H.-S. Chung, “Dual-carrier DQPSK based 112 Gb/s signal transmission over 480 km of SMF link carrying 10 Gb/s NRZ channels,” *OFC*, JW2A-3, Mar. 2012.
[7] J. Li, M. Karlsson, and P. A. Andrekson, “1.94 Tb/s (11×176Gb/s) DP-16QAM superchannel transmission over 640km EDFA-only SSMF and two 280GHz WSSs,” *ECOC*, Th.2.C.1, Sept. 2012.
[8] J. H. Ke, Y. Gao, and J. C. Cartledge, “400 Gbit/s single-carrier and 1 Tbit/s three-carrier superchannel signals using dual polarization 16-QAM with look-up table correction and optical pulse shaping,” *Optics Express*, vol. 22, no. 71, pp. 71–84, Jan. 2014.
[9] J. Renaudier, R. R. Muller, L. Schmalen, P. Tran, P. Brindel, and G. Charlet, “1-Tb/s PDM-32QAM superchannel transmission at 6.7-b/s/Hz over SSMF and 150-GHz-Grid ROADMs,” *ECOC*, Tu.3.3.4, Sept. 2014.
[10] X. Liu, S. Chandrasekhar, P. J. Winzer, T. Lotz, J. Carlson, J. Yang, G. Cheren, and S. Zederbaum, “1.5-Tb/s guard-banded superchannel transmission over 56×100-km (5600-km) ULAF using 30-Gbaud pilot-free OFDM-16QAM signals with 5.75-b/s/Hz net spectral efficiency,” *ECOC*, Th.3.C.5, Sept. 2012.
[11] O. Vassilieva, T. Yamauchi, S. Oda, I. Kim, T. Hoshida, Y. Aoki, J. Rasmussen, and M. Sekiya, “Systematic analysis of intra-superchannel nonlinear crosstalk in flexible grid networks,” *ECOC*, Mo.4.3.6, Sept. 2014.
[12] K. Kojima, D. Millar, T. Koike-Akino, K. Parsons, S. Kametani, and T. Sugihara, “Maximizing transmission capacity of superchannels using rate-adaptive FEC,” *ECOC*, P.3.23, Sept. 2014.
[13] T. Koike-Akino, K. Kojima, D. S. Millar, K. Parsons, S. Kametani, T. Sugihara, T. Yoshida, K. Ishida, Y. Miyata, W. Matsumoto, and T. Mizuochi, “Interference management with Han-Kobayashi coding: Dual-carrier coherent optical communications,” *ECOC*, Mo.3.5.1, Sept. 2014.
[14] F. Hamaoka, K. Saito, T. Matsuda, and A. Naka, “Super high density multi-carrier transmission system by MIMO processing,” *ECOC*, Mo.3.5.4, Sept. 2014.
[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd Edition*, Wiley, July 2006.
[16] T. Han and K. Kobayashi, “A new achievable rate region for the interference channel,” *IEEE Trans. Information Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.
[17] R. Etkin, D. Tse, and H. Wang, “Gaussian interference channel capacity to within one bit,” *IEEE Trans. Information Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.
[18] O. Mehanna, J. Marcos, and N. Jindal, “On achievable rates of the two-user symmetric Gaussian interference channel,” *Allerton Conf. on Communication, Control, and Computing*, IL, Sept. 2010.
[19] Y. Zhao, C. W. Tan, A. S. Avestimehr, S. N. Diggavi, and G. J. Pottie, ”On the maximum achievable sum-rate with successive decoding in interference channels,” *IEEE Trans. Information Theory*, vol. 58, no. 6, pp. 3798–3820, June 2012.
[20] M. Costa, “Writing on dirty paper,” *IEEE Trans. Information Theory*, vol. 29, pp. 439–441, May 1983.
[21] U. Erez and S. ten Brink. “A close-to-capacity dirty paper coding scheme,” *IEEE Trans. Information Theory*, vol. 51, no. 10, pp. 3417–3432, Sept. 2005.
[22] Y. Sun, Y. Yang, A. D. Liveris, V. Stankovic, and Z. Xiong, “Near-capacity dirty-paper code design: A source-channel coding approach,” *IEEE Trans. Information Theory*, vol. 55, no. 7, 3013–3031, July 2009.
[23] G. P. Agrawal, *Fiber-Optic Communication Systems, 4th Ed.*, Wiley, Oct. 2010.
[24] P. Poggiolini, “The GN model of non-linear propagation in uncompensated coherent optical systems,” *IEEE Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3857–3879, Dec. 2012.
[25] P. Serena and A. Bononi, “An alternative approach to the Gaussian noise model and its system implications,” *IEEE Journal of Lightwave Technology*, vol. 31, no. 22, pp. 3489-3499, Oct. 2013.
[26] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, “EGN model of non-linear fiber propagation,” *Optics Express*, vol. 22, no. 13, pp. 16335–16362, June 2014.
[27] P. Serena, A. Bononi, and N. Rossi, “The impact of the modulation dependent nonlinear interference missed by the Gaussian noise model,” *ECOC*, Mo.4.3.1, Sept. 2014.
[28] E. Agrell, A. Alvarado, G. Durisi, and M. Karlsson, “Capacity of a nonlinear optical channel with finite memory,” arXiv preprint arXiv:1403.3339, Mar. 2014.
[29] S. U. H. Qureshi, “Adaptive equalization,” *Proc. IEEE*, vol. 73, no. 9, pp. 1349–1387, Sept. 1985.
[30] A. M. Chan and G. W. Wornell, “Approaching the matched-filter bound using iterated-decision equalization with frequency-interleaved encoding,” *GLOBECOM*, vol. 1, pp. 297–301, Dec. 2002.
[31] M. Tüchler, R. Koetter, and A. C. Singer, “Turbo equalization: principles and new results,” *IEEE Trans. Communications*, vol. 50, no. 5, pp. 754–767, May 2002.