Benefits of optical transceivers employing intentionally nonuniform quantization for advanced modulation formats

Stefanos Dris and André Richter
VPIphotonics GmbH, Carnotstr. 6, 10587, Berlin, Germany

ABSTRACT

A significant amount of R&D effort has been expended recently in finding Shannon capacity-approaching modulation schemes for optical communications. Probabilistic shaping (PS) of QAM constellations has emerged as a particularly attractive solution, allowing fine-grain adjustment of bit-loading, which can be traded off for transmission reach; this approach is ideal for realizing flexible, bandwidth-variable transceivers. PS-QAM, as well as other techniques such as digital subcarrier multiplexing (DSCM), pose significant challenges for the design of transceivers. In particular, the resolution of the digital-to-analog and analog-to-digital converters (DACs/ADCs) becomes critical, if the full benefits of advanced formats are to be obtained.

We present results of our investigation on applying intentionally nonuniform quantization in optical transceivers, as a means of relaxing DAC resolution requirements. By matching the quantizer's transfer function to the distribution of the signal amplitudes, quantization noise can be minimized. This novel approach can lower component cost and power consumption, potentially bringing advanced modulation formats to short-haul/metro links. Moreover, transceivers in the less cost-sensitive long-haul market segment can also profit from increased performance, due to higher signal-to-quantization noise ratio (SQNR). We show how to derive the nonuniform levels for any given modulation format, and quantify by means of extensive simulations the performance gain of the overall coherent system.

Keywords: quantization, DAC, coherent, optical communications, QAM, APSK, probabilistic shaping, digital subcarrier multiplexing

1. INTRODUCTION

The advent of high-order modulation formats for coherent optical communication systems, as well as the inclusion of advanced digital signal processing (DSP) functionality in the transmitter, results in digital baseband signals that are inherently multi-level. Converting these into analog electrical signals with sufficient fidelity therefore places stringent requirements on the resolution of the digital-to-analog converters (DACs). In modern optical transceivers, the DACs employ uniform quantization transfer functions, which are optimal only if the information source signals being converted are continuous, and have uniform amplitude probability distributions. This is clearly not the case with commonly used digital modulation formats; if one is to make efficient use of the DACs' resolution, the quantization transfer function should be tailored to the statistics of the signal being processed. So it is intuitive to argue that nonuniform DACs outperform their uniform counterparts, and can therefore reduce the digital resolution requirements for optical communication systems, yielding equal (or better) performance and reduced power consumption.

The concept of nonuniform quantization has been known and applied for several decades, with research on finding the optimum levels dating back as far as the beginning of the digital information age (see e.g.[1,2]). An example of practical application is the ITU-T G.711 PCM standard, which defines two algorithms (μ-law and A-law) for so-called companding[3]: The input signal is transformed by a nonlinear transfer function (a compressor), followed by uniform quantization. The inverse nonlinearity is then applied by an expander, to produce the output. The scheme was specifically designed for voice signals (telephony).

Recently, several relevant simulation and experimental studies have appeared in the optical communications literature. Since high-speed nonuniform DACs for optical communications have not been realized yet, the experimental demonstrations so far have been achieved by selecting a subset of the levels of a uniform DAC. In[4] the concept was applied in a direct-detection system employing discrete multi-tone (DMT), where the performance of the Gaussian-shaped signals was enhanced using a 3-bit nonuniform DAC. The nonuniform DAC has also been proposed for geometrically- and probabilistically-shaped (PS) circular QAM[5], the idea being that such PS signals are Gaussian-like, and also benefit from the concept. A segmented-electrode IQ Mach-Zehnder modulator (IQ-MZM) concept was presented in[6]. This is capable of multi-level format creation, using only binary driving electronics, in what is termed an 'optical DAC' (oDAC). The device's transfer function is programmable with fine granularity, and simulation studies of 4-bit nonlinear DAC operation showed its suitability for generation of high-order PS square M-QAM formats, with power consumption that is far less than can be achieved by an IQ-MZM driven by conventional electronic DACs and amplifiers. The device is of particular interest, as it represents a potentially practical realization of nonuniform quantization for optical communications.

In the present study we carry out simulations in VPItransmissionMaker Optical Systems in order to quantify the benefit of employing intentionally nonuniform quantization levels in the transfer functions of transmitter DACs. The performance gain that can be achieved with a 5-bit nonuniform DAC versus its uniform counterpart is quantified in terms of the signal-to-quantization noise ratio (SQNR). We consider two modulation formats: Digital subcarrier multiplexing (DSCM) and probabilistically-shaped M-ary amplitude phase shift keying (PS-M-APSK)[7,8]. The former is an interesting use case for nonuniform DACs, since its amplitude distribution tends toward Gaussianity, with increasing number of subcarriers. On the other hand, PS-M-APSK is both probabilistically and geometrically shaped, and the results could be used to draw more general conclusions about the kinds of single-carrier QAM formats that can benefit from the concept. Moreover, PS-M-APSK is a special case of an advanced PS format that offers superior performance (in terms of mutual information) to square M-QAM, while also allowing for Gray mapping. We propose a simple improvement to finding the optimal DAC levels for the PS-M-APSK case, which increases performance and is applicable to QAM constellations in general.

2. OPTIMIZATION OF TRANSMITTER DAC QUANTIZATION LEVELS

2.1 System Description

We consider the coherent optical transmission system shown in Figure 1, comprising a dual-polarization (DP) IQ-MZM, a channel, and a polarization-diversity coherent receiver. For each polarization on the transmitter side, the DSP block creates the required in-phase and quadrature digital signals that are converted to analog electrical signals by the DACs, before being amplified to $2V_\pi$ and fed to the RF inputs of the IQ-MZM. In order to isolate the effect of the DAC resolution on the signal quality, we consider that the components and channel are ideal, and do not introduce any noise or distortions.

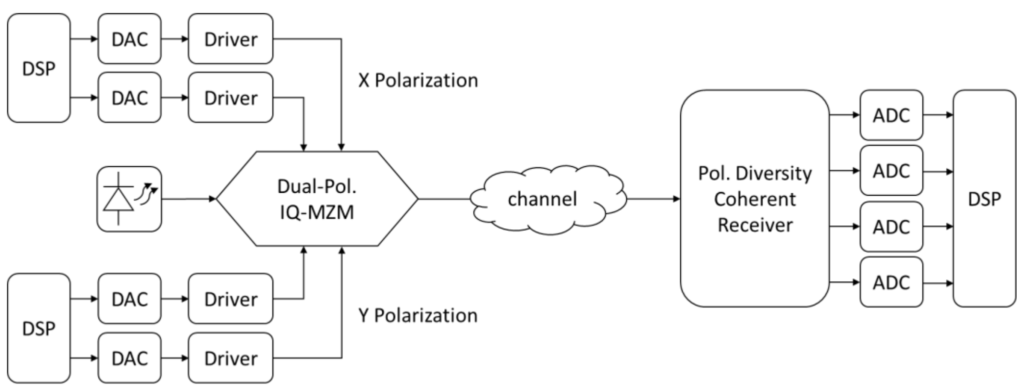

Figure 1. Dual-polarization coherent optical transmission system model

The transmitter DSP is primarily responsible for mapping the information bits to complex symbols, according to the desired modulation format. Optional, but commonly employed functionality also includes (root) raised cosine (RRC) pulse shaping and digital pre-distortion to mitigate transmitter component impairments. The latter may be static, as in the case of the linearization of the MZM transfer function via the arcsin function, or frequency-dependent, where a linear FIR filter is used to pre-equalize the frequency response of the transmitter. In this work we consider only the former, since we want to ignore any bandwidth limitations and distortions.

PS-M-APSK creation and detection

Figure 2 represents an illustration of the DSP functions used in the simulation studies to create and detect PS-M-APSK signals. For the case of PS-64-APSK signals, shaping is performed to yield spectral efficiencies (SEs) of 4.5, 5.0, 5.5 and 6.0 bits/symbol/polarization (the latter corresponding to uniform symbol probabilities, i.e. no shaping). The amplitudes of the rings of the constellations are optimized for each SE[7]. They are shown in Figure 3, with added noise (same OSNR for all cases) such that the differences in symbol probability are evident.

2^{16} symbols are generated by the mapper, after which they are upsampled and pulse-shaped at twice the symbol rate by an ideal RRC filter with roll-off = 0.2. The amplitude clipping stage is used only with a uniform quantized DAC; it can be tuned to maximize the SQNR, and thus enable a fair comparison with the nonuniform DAC. In this work, the ac-coupled signal is clipped symmetrically about the 0 level, and the amount of clipping is quantified as a percentage of the full signal range. The digital signal is finally pre-distorted by the inverse static MZM transfer function, and sent to the DAC.

On the receiver side, the signal is sampled at twice the symbol rate by an ideal ADC (with practically 'infinite' resolution), and passed through the matched RRC filter, before being downsampled to 1 sample/symbol. The error vector magnitude (EVM) of the resulting complex 2D signal is then estimated, which, in the absence of any other impairments, is related to the SQNR by $SQNR = 1/EVM^2$.

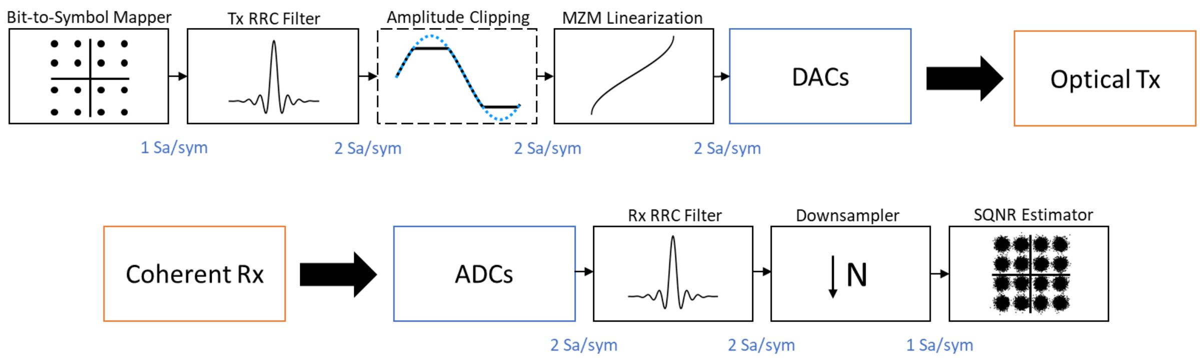

Figure 2. Transmitter and receiver DSP employed in the simulation studies for the PS-M-APSK signals. For simplicity, only one polarization is illustrated. The signal sampling rate is shown at each DSP block.

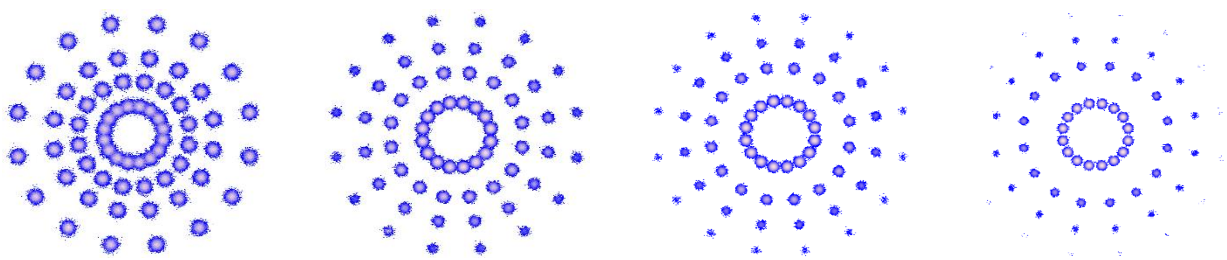

Figure 3. The geometrically- and probabilistically-shaped 64-APSK constellations considered in the simulations, each with different spectral efficiency

Digital SCM creation and detection

Generation of digital SCM signals involves the extra step of single-sideband (SSB) modulation of each QAM-modulated RF subcarrier, as shown in Figure 4. We generated 2, 4 and 8 subcarriers, each carrying either QPSK or 16-QAM symbols. The roll-off of the RRC filters are set to 0.05, with just enough spacing between the carriers to avoid inter-carrier interference (ICI). The sampling rate is set to (2 × number of subcarriers × symbol rate per subcarrier). It follows that the samples/symbol is then a function of the number of subcarriers (4 Sa/symbol for the 2-subcarrier case, and up to 16 Sa/symbol for the 8-subcarrier case).

Again, signal clipping is only employed for the uniform quantizer. The receiver side (omitted for simplicity) comprises of sampling by an ideal ADC at the same sampling rate as the transmitter DAC, RF subcarrier downconversion, RRC matched filtering, downsampling and finally EVM/SQNR estimation.

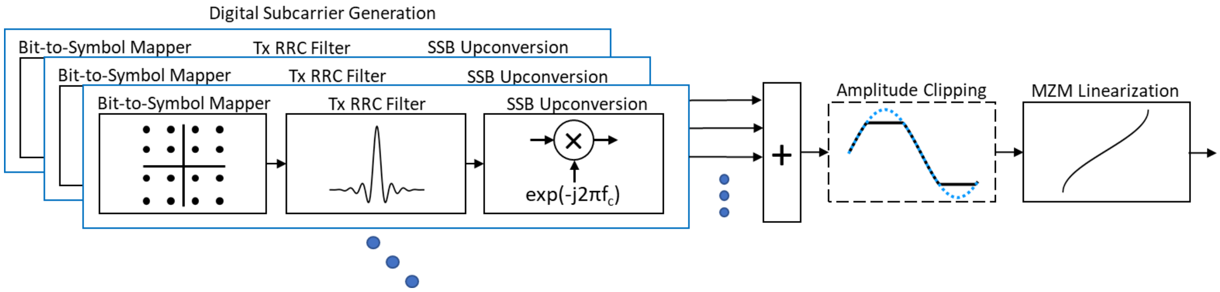

Figure 4. Transmitter DSP applied for generating digital SCM signals

2.2 Simulation Results

Digital SCM

Figure 5 shows the probability distributions of the in-phase amplitudes, for all simulated DSCM cases, created using the transmitter DSP of Figure 4. In general, the distribution of a DSCM signal tends to become more Gaussian with increasing modulation order and number of carriers. In our case, the distribution tails are slightly stretched, due to the MZM linearization function of the DSP. These digital signals are used as training signals to determine the optimum 5-bit nonuniform DAC transfer functions, using Lloyd's algorithm[1,2], which is a recursive method that minimizes the mean square error for a given training set. As can be seen from the transfer functions obtained (Figure 6), this procedure converges to a result that naturally clips the signal. Moreover, the DAC transfer functions are fairly similar, suggesting that a single transfer function can be used for all signals. To test this, we also derive a 'universal' nonuniform DAC transfer function trained on a combination of all signals.

The SQNR results after the receiver DSP are compared against the performance achieved with a 5-bit uniform DAC, for which the amount of clipping is varied to maximize the SQNR for every simulated signal (Figure 7). The nonuniform DACs achieve ~3.5 dB better performance w.r.t. the uniform for 2-carrier DSCM, while the difference is even higher with more carriers (ΔSQNR ~4 dB with 8-carrier DSCM). Also in Figure 7 (left), the performance of the 16-QAM DSCM using the universal transfer function is plotted. As can be seen, negligible penalty is observed compared to the tailored transfer functions.

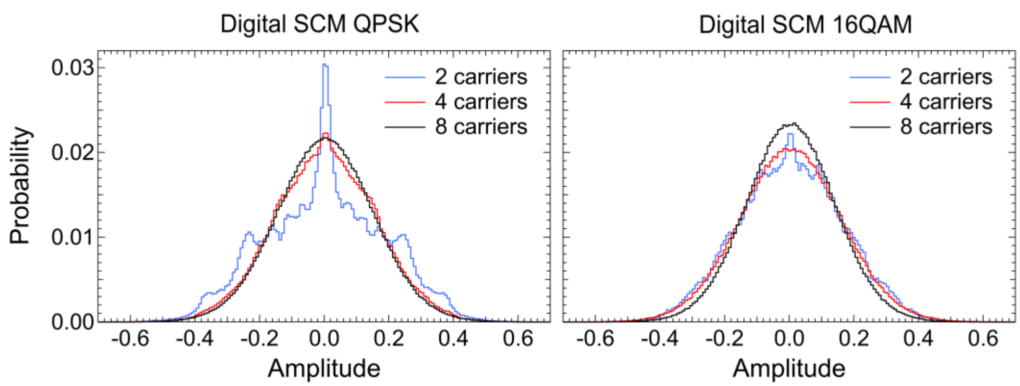

Figure 5. In-phase amplitude histograms of the digital SCM signals for QPSK (left) and 16-QAM (right), for 2, 4 and 8 carriers, after being processed with the transmitter-side DSP of Figure 4

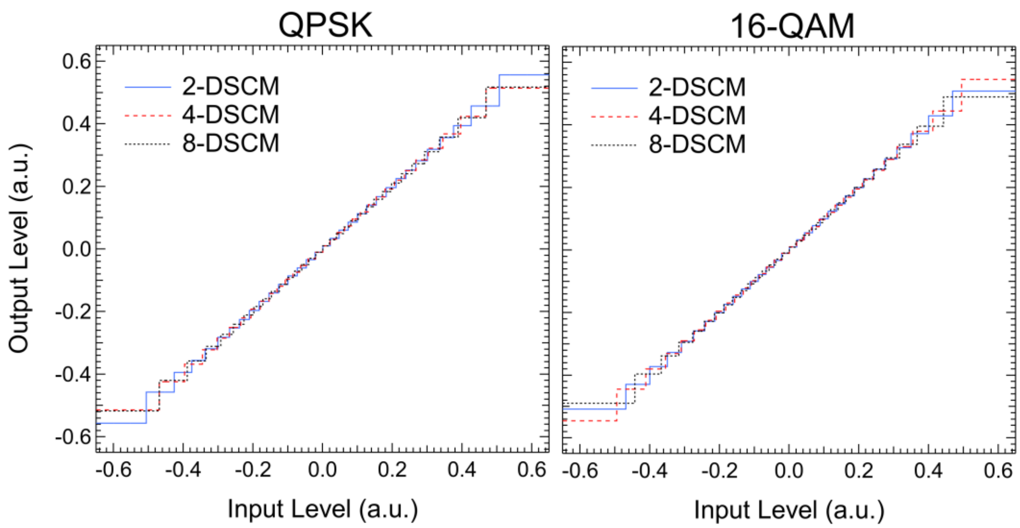

Figure 6. Optimal nonuniform DAC transfer functions for the DSCM signals.

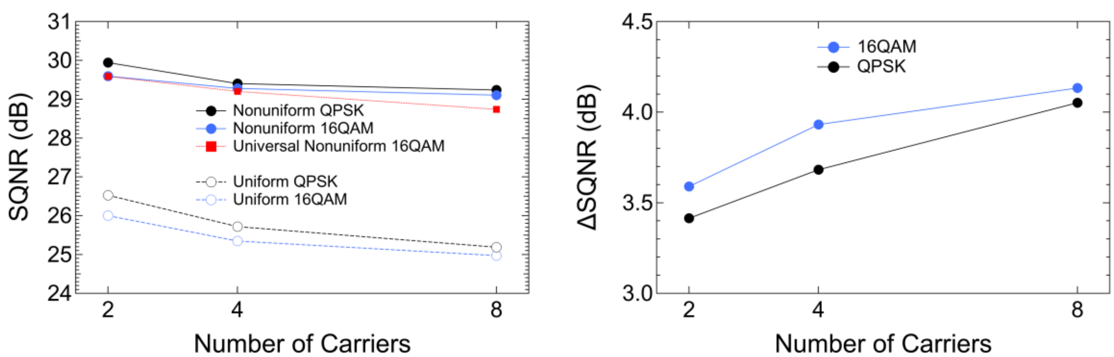

Figure 7. SQNRs obtained with the 5-bit uniform and nonuniform DACs for the DSCM systems (left), and corresponding SQNR gain achieved with the nonuniform DAC, w.r.t the uniform one (right)

PS-64-APSK

For the PS-64-APSK signals, the clipping ratio is optimized when using the uniform DACs, just as in the DSCM simulations. The procedure to find the optimal nonlinear DAC transfer function, however, is modified. Given that we operate the transmitter DSP at 2 Sa/symbol, 50% of the samples are placed at the optimal symbol sampling time instant, while the others are placed at the transitions between symbols. The histograms of the in-phase components of these two sampling time instants are plotted separately in Figure 8.

Training the nonuniform DAC turned out to be challenging. In[6] we noted that training it on only the symbol instants, yielded better results than training the nonuniform DAC on the entire signal (i.e. including the transitions). However, in that case the transmitter pulse shaping filter had a raised cosine response (instead of root raised cosine), meaning there was no ISI at the symbol instants out of the transmitter DACs. Attempting the same training on the signals studied here yields suboptimal results, and in some cases may result in excessive clipping, introducing severe nonlinear distortion. While the latter may not manifest itself as a degradation in the SQNR calculation, it is undesirable since it cannot be mitigated by the linear equalizer in the digital coherent receiver. On the other hand, training on the entire signal including the transitions (as in[5]) results in slightly suboptimal SQNR, since the symbol sampling points are clearly more important.

In the present work we propose a different approach that strikes a trade-off: We train the nonuniform DAC levels on a signal containing the samples at all of the symbols, but only a fraction of the samples at the transitions. We vary this fraction from 0 to 100%, and choose the point at which the SQNR is maximized. For the tested formats, we find that the

SQNR is maximized when the fraction of transition samples used is between 40-60%. The highest SQNR improvement of 0.6 dB w.r.t. using all the transition samples is observed for the 5.5 bits/symbol case. Conversely, when a signal contains transitions that swing above the maximum symbol amplitude, ignoring these will cause severe clipping; thus, for the 6.0 bits/symbol case, only ~0.06 dB SQNR gain is achieved using our training optimization.

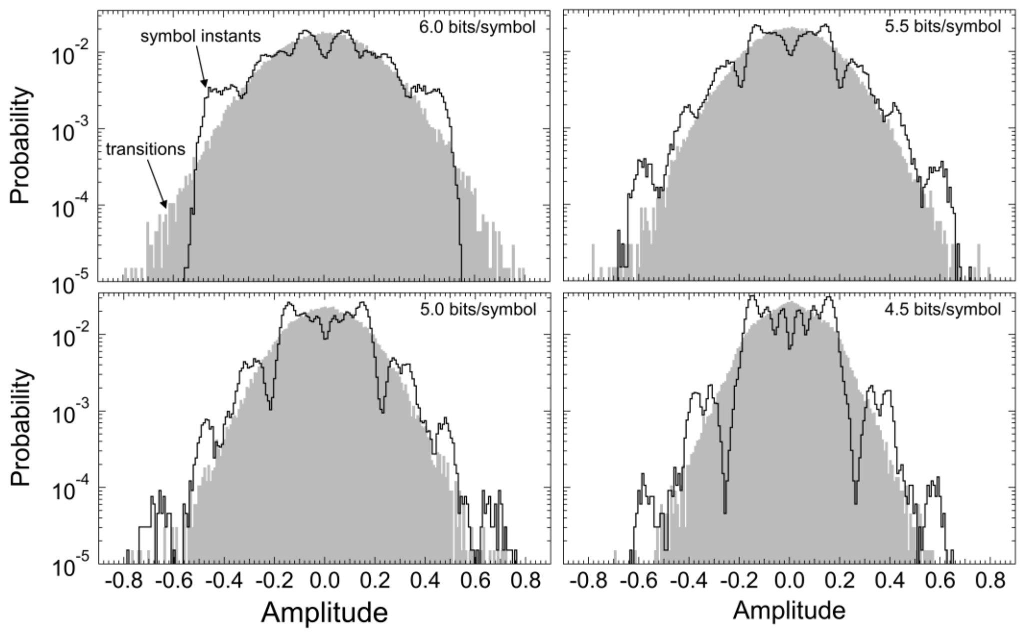

Figure 8. In-phase amplitude histograms at the symbol instants (solid lines) and the transitions (grey shaded area) for PS-64-APSK with varying amount of probabilistic shaping, after being processed with the transmitter-side DSP of Figure 2

The optimal nonuniform DAC transfer functions obtained are plotted in Figure 9. It is immediately obvious that they are vastly different, depending on the amount of probabilistic shaping applied. In contrast to the DSCM signals, it is therefore not possible to derive a single DAC transfer function that will work well with all PS-64-APSK signals. In terms of performance, the nonuniform DAC achieves a gain of ~3.1 dB with respect to the unshaped 64-APSK. The gain increases with the application of more probabilistic shaping, reaching ~5.1 dB for the 4.5 bits/symbol PS-64-APSK.

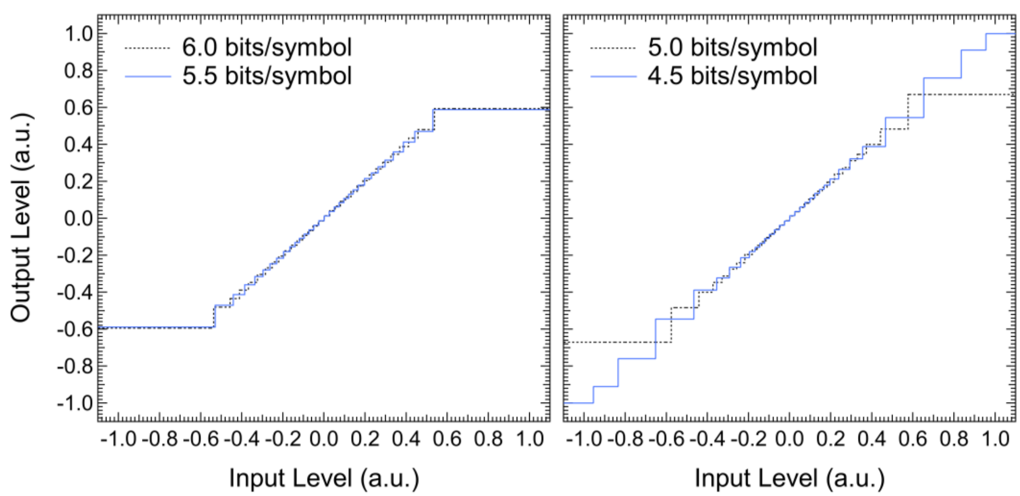

Figure 9. Optimal nonuniform DAC transfer functions for the PS-64-APSK signals

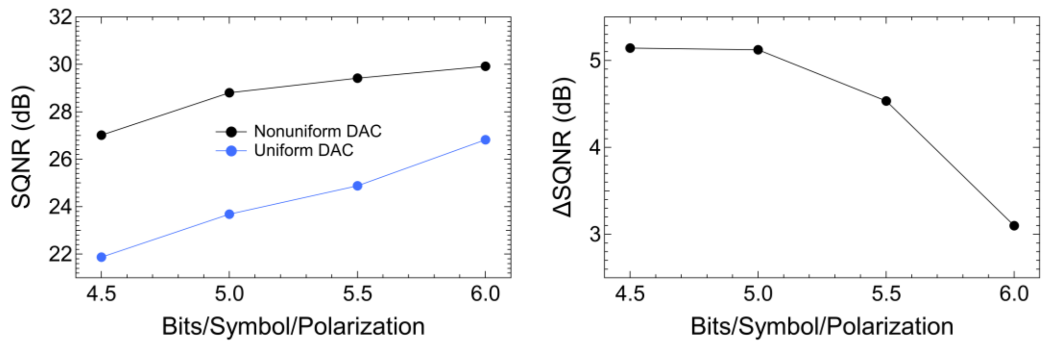

Figure 10. SQNRs obtained with the 5-bit uniform and nonuniform DACs for PS-64-APSK (left), and corresponding SQNR gain achieved with the nonuniform DAC, w.r.t to the uniform one (right)

3. CONCLUSION

We have demonstrated via simulations of coherent optical communication systems that quantization with a nonuniform transfer function offers superior performance than with a uniform one at the same resolution. We used Lloyd's algorithm to match the DAC levels to the signal sources' amplitude distributions. The latter are dependent on both, the modulation format being used and the transmitter DSP (such as pulse-shaping and pre-distortion). The SQNR gain for digital subcarrier modulation ranged between 3.5-4 dB with 5 bits of resolution, while for PS-64-APSK it was even higher, reaching as much as 5 dB of improvement when more probabilistic shaping was applied.

If we are to leverage this scheme in optical communication systems, the challenge is for component designers to come up with nonuniform quantizers that are realizable in terms performance and power consumption. A potential barrier is that of the requirement for programmability: When dealing with generalized QAM formats (including the PS variety), we have shown that there is no one-size-fits-all transfer function, and this may complicate practical implementation. On the other hand, signals such as DSCM and DMT which tend toward Gaussianity, can take advantage of a single transfer function regardless of the number of carriers or bit loading. A static transfer function nonuniform DAC optimized for these scenarios may therefore be easier to implement.

REFERENCES

[1] S. P. Lloyd, "Least squares quantization in PCM," IEEE Trans. Inform. Theory (Special Issue on Quantization), vol. IT-28, pp. 129–137, Mar. 1982.

[2] J. Max, "Quantizing for Minimum Distortion," IRE Transactions on Information Theory, Vol. IT-6, pp. 7–12, March, 1960.

[3] N. Judell and L. Scharf, "A simple derivation of Lloyd's classical result for the optimum scalar quantizer," IEEE Transactions on Information Theory, 32 (2), 326–328 (1986).

[4] J. Peng, L. Han, Q. Zhu, C. Qiu, Y. Zhang, C. Tremblay, Y. Su, "SQNR Improvement Enabled by Nonuniform DAC Output Levels for IM-DD OFDM Systems," IEEE Photonics Journal, 9 (2), April 2017.

[5] S. Almonacil, F. Jardel, P. Jennevé, and Patricia Layec, "Nonuniform DAC Design for Combined Geometrically and Probabilistically Shaped Circular QAM," in Proc. Advanced Photonics Congress (SPPCom) 2018, Zurich, Switzerland, paper SpTu4G.5.

[6] A. Aimone, S. Dris, J. Sundermeyer, J. H. Choi, C. Zerna, A. Richter, K. Velthaus, "Programmable transfer function optical-DAC using an InP segmented Mach-Zehnder modulator," in Proc. 20th International Conference on Transparent Optical Networks (ICTON 2018), paper Tu.B5.1.

[7] Zaishuang Liu, Qiuliang Xie, Kewu Peng, and Zhixing Yang, "APSK Constellation with Gray Mapping," IEEE Communication Letters, 15 (12), pp 1271-1273, December, 2011.

[8] S. Zhang, F. Yaman, Y.-K. Huang, J. D. Downie, D. Zou, W. A. Wood, A. Zakharian, R. Khrapko, S. Mishra, V. Nazarov, J. Hurley, I. B. Djordjevic, E. Mateo, Y. Inada, "Capacity-Approaching Transmission over 6375 km at Spectral Efficiency of 8.3 bit/s/Hz," in Proc. 2016 Optical Fiber Communications Conference and Exhibition (OFC), 22-24 March, Anaheim, CA, USA (2016).