Energy-Efficient Implementation of Carrier Phase Recovery for Higher-Order Modulation Formats

Energy-Efficient Implementation of Carrier Phase Recovery for Higher-Order Modulation Formats

Erik Börjeson, *Student Member, IEEE* and Per Larsson-Edefors, *Senior Member, IEEE*

***Abstract*—We introduce circuit implementations of one- and two-stage carrier phase recovery (CPR) for 256QAM coherent optical receivers. We describe in detail the optimizations of algorithms, such as modified Viterbi-Viterbi (mVV), blind phase search (BPS), and principal component-based phase estimation (PCPE), that are required to develop energy-efficient CPR circuits and show how design parameter settings and limited fixed-point resolution affect the SNR penalty. 30-GBaud CPR circuit netlists synthesized in a 22-nm CMOS process technology allow us to study trade-offs between energy per bit and SNR penalty. We show that it is possible to reach an energy dissipation of around 1 pJ/bit at an SNR penalty of 0.6 dB for two-stage PCPE+BPS and mVV+BPS implementations, and that PCPE+BPS is the preferred choice thanks to its smaller area.**

I. Introduction

Coherent technology has been a key to increasing the capacity of long-haul fiber-optic communication. Thanks to its high spectral efficiency and high receiver sensitivity, it is also an interesting alternative for shorter fiber systems. However, since complex digital signal processing (DSP) is required in coherent schemes, DSP power dissipation has been an issue. Low DSP power dissipation is necessary for a proliferation of coherent technology into shorter, more cost-sensitive systems.

For shorter fibers, chromatic dispersion and polarization-mode dispersion become negligible. This allows us to, first, do away with the chromatic dispersion compensation unit and, second, shorten the adaptive equalizer; two power-hungry units of long-haul coherent receivers. There are, however, other receiver units whose design complexity does not scale down as the fiber reach is reduced: The analog-digital converter (ADC) and the carrier phase recovery (CPR) unit. As far as ADC is concerned, the power dissipation depends on resolution and sampling rate. Four 8-bit 70-GSa/s ADCs [1] in a 32-nm CMOS process technology were estimated to dissipate 3.5 pJ/bit for 60-GBaud 16QAM [2]. Note that CMOS technology scaling improves energy efficiency of this type of ADC architecture; a newer generation high-speed ADC in 14-nm FinFET technology [3] showed a 35% improvement in energy per bit. As far as CPR goes, we recently introduced a blind phase search (BPS) implementation in a 22-nm CMOS process technology that dissipates 1.1 pJ/bit for 32-GBaud 16QAM [4]. While the power dissipation for this BPS-based CPR circuit was shown to represent a relatively small portion of the total DSP power dissipation of a 16QAM coherent receiver [2], the energy per bit of a single-stage BPS-based CPR implementation for 32-GBaud 256QAM was 3.1 pJ/bit [4] which is almost three times higher than for 16QAM.

Clearly, if pilot symbols are used, the energy per bit for 256QAM CPR can be improved significantly; down to 0.34 pJ/bit for a 32-GBaud 256QAM CPR circuit [5]. However, since pilot symbols reduce spectral efficiency, it is interesting to explore if non-data-aided CPR implementations for higher-order modulation formats can be optimized, to achieve substantially lower energy per bit. Thus, this paper investigates how to implement energy-efficient blind CPR circuits for higher-order modulation formats, especially focusing on how to efficiently combine several algorithms in a multi-stage CPR implementation. We limit our study to two-stage CPR implementations. This is because adding more stages would significantly increase the power dissipation, so the addition must significantly reduce the penalty. It has, however, previously been shown that using more than two stages only results in small phase estimate improvements [6].

II. CPR Algorithms

There is a wide range of blind CPR algorithms available. This section briefly reviews the algorithms that we have chosen to implement in our VLSI circuit evaluations.

For simple modulation formats which encode data on the phase only, the M^{th}-power (also known as Viterbi-Viterbi) CPR algorithm can be used [7]. By taking the M^{th} power of the input symbols for M-PSK, the phase modulation can be removed and the phase noise estimated, once the additive white Gaussian noise (AWGN) has been reduced by averaging. The relatively low complexity of this method promises to deliver an energy-efficient circuit implementation, however, the move to higher-order formats, such as QAM, requires modifications to account for the non-constant amplitude of the symbols. One suggested method is to use QPSK partitioning of the symbols, where only Class-1 symbols having modulation angles of $\pi/4 + n\pi/2$, for $n = 0...3$, are used for phase estimation [8]. This works well for 16QAM, but for higher-order QAM the modulus of many Class-1 symbols are similar to other constellation points, reducing the number of usable candidate symbols. The result is that only a small portion of the received symbols are being used for phase estimation, making tracking of fast phase changes infeasible due to the long averaging window needed to reduce AWGN [6].

A modified Viterbi-Viterbi (mVV) CPR algorithm for 64QAM was presented in [6], where the usable Class-1 symbols are combined with outermost edge symbols, increasing

This work was financially supported by the Knut and Alice Wallenberg Foundation.

E. Börjeson and P. Larsson-Edefors are with the Department of Computer Science and Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden (e-mail: erikbor@chalmers.se, perla@chalmers.se).

the number of symbols used for estimation. This method shows a larger laser-linewidth tolerance than standard QPSK partitioning and can be extended to 256QAM. However, the required length of the averaging window is still relative large, resulting in poor performance for larger linewidths. For this case, a two-stage CPR approach, where mVV is used as a coarse-grain CPR in conjunction with a fine-grain constellation transformation (CT) algorithm, has been suggested [6].

In the original CT algorithm [9], a 16QAM constellation is reduced to a QPSK constellation by analyzing the I and Q parts of the received symbols. After transformation, a conventional M^{th}-power estimator is used. Bilal et al. [6] extended CT to work also for 64QAM by using a stepwise transformation of the constellation, first to 16QAM and then to QPSK, and an equivalent method can be used for 256QAM. But for CT to work, only a minor amount of phase noise can be accepted in the symbols, limiting its use to a fine-grain CPR stage.

Blind phase search (BPS) for fiber-optic systems was originally suggested by Pfau et al. [10]. Here, the input symbols are rotated by test phases and the symbol closest to a constellation point is selected as the output. The key to keeping BPS circuit complexity low is to minimize the number of test phases [11]. However, higher-order modulation formats require more BPS test phases, so the energy efficiency of BPS degrades: A single-stage 256QAM BPS circuit requires three times more energy per bit than a single-stage 16QAM BPS circuit [4].

A recently published CPR algorithm, called principal component-based phase estimation (PCPE) [12], uses principal component analysis to estimate the phase noise. In general, the PCPE algorithm requires less complex hardware than BPS does. In addition, PCPE scales much better than BPS to higher-order modulation formats. However, PCPE suffers from residual phase noise, which is especially pronounced at high signal-to-noise ratios (SNRs). The problem with residual phase noise can be mitigated by cascading PCPE with BPS, using a limited number of test phases [12].

Other two-stage algorithms have been suggested, such as two-stage BPS [13] and mVV followed by a maximum likelihood estimator (MLE) [6]. The former is not included in this work, since BPS+BPS cannot reconcile low energy dissipation and low SNR penalty for 256QAM. The latter could be a potential candidate with a performance similar to that of mVV+CT. However, the implementation of MLE-based CPR can be assumed to be more complex than that of CT.

III. System Model

Our CPR circuits were defined in a hardware description language (HDL) and the evaluation was performed using MATLAB-HDL co-simulation, where bit-true HDL code is embedded inside the MATLAB environment: The data and channel impairments were modelled in MATLAB and fed to a simulator that performs logic simulation of the HDL unit. The results were then imported back into MATLAB for demodulation and evaluation.

Fig. 1 shows our system model, which includes AWGN and phase noise, modelled as a Wiener process. All other impairments, e.g., chromatic and polarization-mode dispersion are assumed to be fully compensated by other DSP units. We also neglect the effect of non-linear impairments.

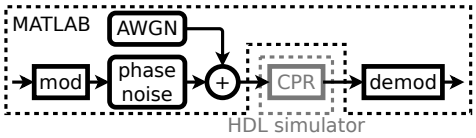


Fig. 1: System model used for our simulations, with the MATLAB parts shown in black and the HDL part in gray.

IV. CPR Implementation

Our 256QAM CPR circuits target a symbol rate of 30 GBaud, which calls for parallelization of the design in $P = 32$ lanes at a clock rate of 937.5 MHz. These values correspond to a data rate of 400 Gbit/s, if we assume transmission over two polarizations and a 20% overhead for forward error correction (FEC). Block diagrams of the proposed mVV+CT and PCPE+BPS implementations are shown in Fig. 2 and Fig. 3, with the parallel parts marked in gray. Moving as many operations outside of the parallel portion of the design as possible is an important way of reducing circuit power dissipation. The word length of all internal signals are calculated from the input symbol word length, W, and kept as low as possible without severely impacting the SNR penalty. The remainder of this section will describe the most important features of our developed implementations.

a) mVV: A block diagram of the mVV implementation is shown in the left part of Fig. 2. The first unit calculates the magnitude of the input symbols by using an unrolled coordinate rotation digital computer (CORDIC) [14] circuit. The input is then partitioned using the magnitudes to determine if an input symbol should be included in the phase estimation or not. The I and Q values of symbols that should be ignored are set to zero. Calculation of the 4^{th} power is done using two complex squarers with rounding, since our integrated circuit design tools recognize these constructs and can optimize them efficiently. To reduce implementation complexity, averaging is implemented as a multi-input addition of the parallel symbols; a method that is used in all of our described implementations. After averaging, the result is shifted as much a possible to the left without changing the sign of the data. This shift ensures maximum utilization of the word length, despite varying number of symbols included in the average, and provides an acceptable input to the angle operation, also implemented using CORDIC. After unwrapping, the estimated phase is sent to the compensation unit.

The optimal setting for the length of the averaging window, L, is dependent on the linewidth symbol-duration product $\Delta f T_s$ and a larger laser linewidth results in a smaller optimal L, as shown in Fig. 4a. In this figure, the bit-error rate (BER) is plotted as a function of L at an SNR of 17 dB, chosen to be close to the soft-decision FEC threshold. Since the average length, $L > P$, the parallelism can be abandoned for the angle and unwrap operations and multiple clock cycles can be used for signal propagation in these units. Clock gating can also be used to minimize unnecessary logic signal switching and reduce power dissipation [15].

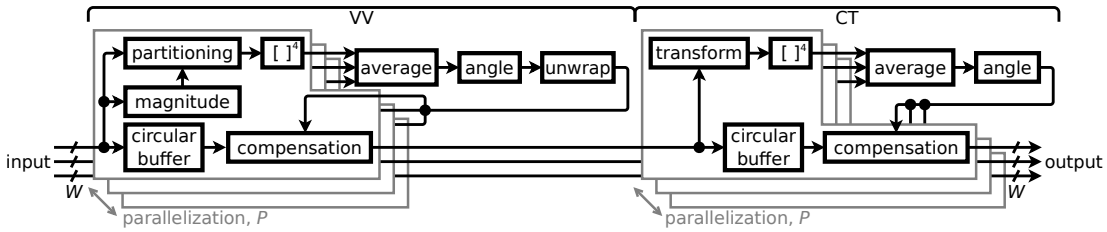


Fig. 2: Block diagram of the mVV+CT CPR implementation, where W is the input word length and P is the number of parallel lanes.

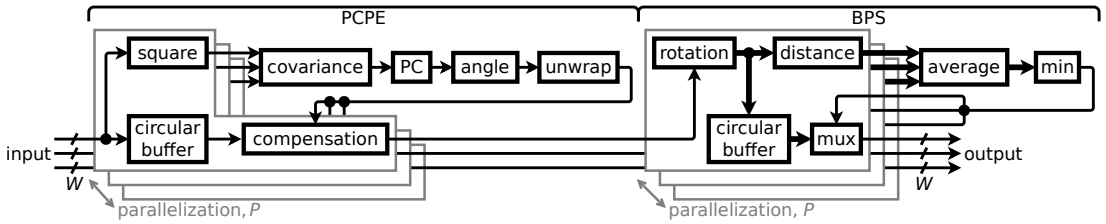
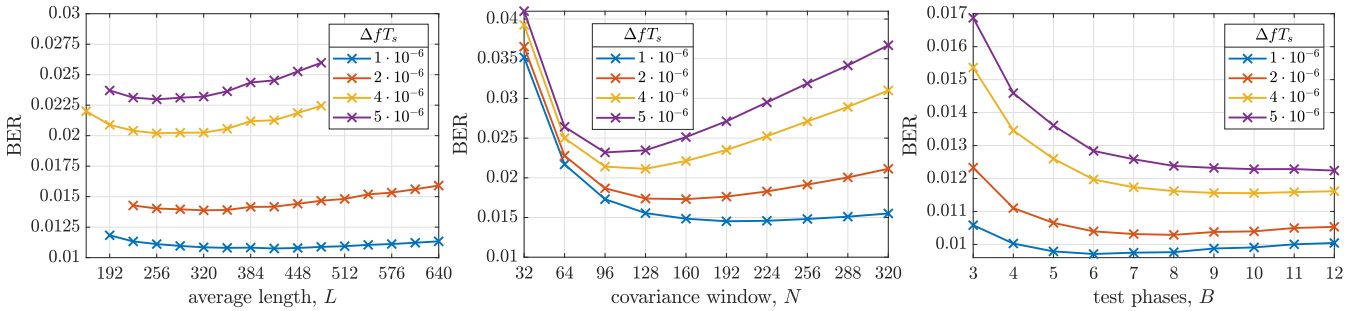


Fig. 3: Block diagram of the PCPE+BPS CPR implementation, where W is the input word length and P is the number of parallel lanes. The thick arrows represent buses carrying data for multiple rotations.



(a) (b) (c)

Fig. 4: BER at an SNR of 17 dB and $W = 10$ as a function of (a) the length of the mVV averaging window in mVV+CT, (b) the covariance window for PCPE, and (c) the number of test phases used in PCPE+BPS with optimal selection of N.

b) CT: The CT implementation shown in Fig. 2 is very similar to that of mVV, described above. The partitioning is replaced by a transformation operation and the unwrapping is removed, since only small phase fluctuations are expected. We implement the transformation in a single unit, i.e., 256QAM is directly converted to QPSK, instead of using multiple, successive transformations as suggested in [6]. A short CT averaging window is needed to track fast phase changes; the optimal setting was found to be 32 for all CT implementations.

c) PCPE: In the PCPE circuit, the parallel input symbols are first squared prior to the calculation of the covariance done over a window of N consecutive symbols. The optimal value of N, resulting in the minimal SNR penalty, is dependent on the amount of phase noise present in the received symbol stream. A larger $\Delta f T_s$ results in a smaller optimal N, as shown in Fig. 4b. Typically $N > P$ and this makes it possible to abandon the parallelism after this point. The resulting covariance matrix is a 2×2 symmetric matrix; thus only three values need to be calculated and these are stored in a register which is reset each N/P clock cycle. Since we are only interested in the relative differences between the values of the covariance matrix, we find the largest of the three values and count the number of leading zeros (z) in this value. All values of the covariance matrix are left shifted with $z - 1$ and we remove as many of the least significant bits as we can, without increasing the SNR penalty. The same idea is used after calculation of the principal component (PC), to reduce the number of bits propagated to the next operation. The phase angle is calculated using CORDIC and unwrapped before being fed to the compensation unit. Since $N > P$, all operations following the covariance calculation will have N/P clock cycles to complete, enabling clock gating to reduce power dissipation.

d) BPS: The selection of BPS test phases is controlled by two parameters: The number of test phases B needed and, when BPS is used as a second CPR stage, the angle these test phases will span, S. The optimal choice of B depends on the phase noise and SNR properties of the channel; Fig. 4c shows the BER as a function of B for a BPS used as a second stage preceded by PCPE. For $\Delta f T_s = 10^{-5}$, it is enough to choose $B = 4$, but for larger linewidths the optimal number of test phases increases. This number is, however, much smaller than the >28 test phases needed for a single-stage 256QAM BPS implementation with reasonable SNR penalty [4]. The optimal selection of S varies with the combination of $\Delta f T_s$ and B, but its value has no effect on the power dissipation. For low

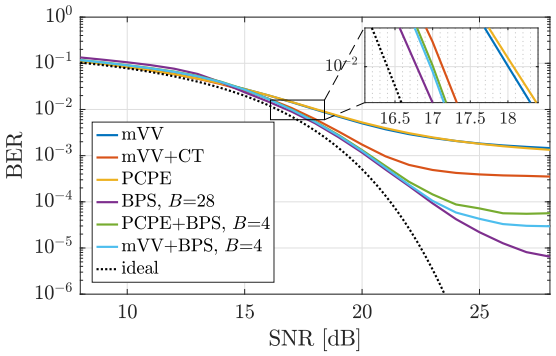


Fig. 5: BER as a function of SNR for $\Delta f T_s = 10^{-6}$ and a word length of $W = 10$.

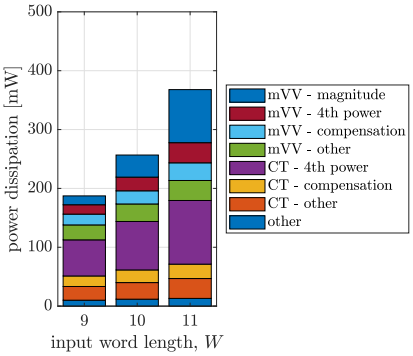


Fig. 6: Power dissipation distribution for mVV+CT for different settings of W.

values of $\Delta f T_s$, the optimal value of S will be very small and, due to the limited fixed-point resolution, the SNR penalty will actually increase with increasing values of B, as can be seen for the curves representing a $\Delta f T_s$ of 10^{-6} and $2 \cdot 10^{-6}$ in Fig. 4c. There is no need for additional phase unwrapping when BPS is used as a fine-grain second stage, due to the small amount of phase noise left after the first stage.

e) Compensation: In all implementations except when BPS is used as a second stage, the input symbols are delayed in a circular buffer, which can be clock gated efficiently, to match the pipelining delay of the estimation circuits. The actual phase compensation is then performed using complex multipliers. In the second-stage BPS, the rotated values are already available and all rotated input values are stored in a buffer. The estimated phase is used to select the value resulting in the smallest average distance to a constellation point.

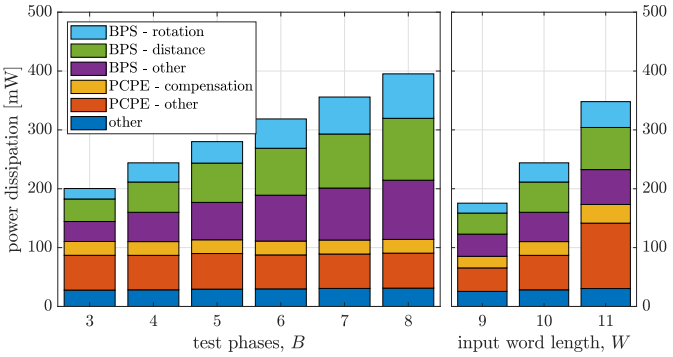


Fig. 7: Power dissipation for the major units of the PCPE+BPS implementation for (a) different number of test phases B and (b) different input word lengths W.

V. Results

The BER as a function of SNR for our circuit implementations is shown in Fig. 5, for $\Delta f T_s = 10^{-6}$ and a word length $W = 10$ bits. The effect of the residual phase noise left after single-stage mVV and PCPE can be seen at higher SNRs, where the BER levels out. This behavior is even more pronounced at larger $\Delta f T_s$ (not shown in the figure), especially for mVV. At a BER of 10^{-2}, the two-stage mVV+BPS and PCPE+BPS implementations have similar BER performance, but the latter shows worse performance at higher SNRs. This is most probably due to the larger residual phase noise left after PCPE, which affects the performance of the second-stage BPS but is too small to influence the BER results of the single-stage PCPE. The mVV+CT combination shows an additional penalty, compared to the other two-stage implementations, of just over 0.1 dB. Compared to a single-stage 256QAM BPS implementation with $B = 28$ and interpolation [4], the performance of the two-stage CPR implementations is slightly worse, but as we will show this is accompanied by a significantly reduced power dissipation.

Using an industrial-grade application-specific integrated circuit (ASIC) design flow [15], we synthesized the two-stage CPR circuit implementations using Cadence Genus to a cell library in a 22-nm CMOS process technology at slow process corners, using a 0.72-V 125-°C characterization, to avoid overly optimistic timing results. The synthesized netlists were simulated using MATLAB-generated input vectors and the resulting switching statistics were back-annotated into Genus for the power analysis, which assumed typical process corners, at a 0.8-V 85-°C characterization, to be close to a normal usage case. Note that the power dissipation results in this section are presented for CPR on a single polarization.

Fig. 6 shows the distribution of the power dissipation for mVV+CT, using three different setting of W and optimal settings of L for an SNR of 17 dB and $\Delta f T_s = 10^{-6}$. The magnitude operation accounts for one third of the power dissipation of mVV, while the more complex 4^{th}-power calculation draws much less power; note though that the area of the latter is almost four times larger. This mismatch is due to a large part of the symbols being set to zero in the partitioning, which greatly reduces the switching activity of the multipliers in the 4^{th}-power operation. In the CT implementation, all symbols are used in the 4^{th}-power operation and this operation accounts for more than half of the total CT power dissipation. The power dissipation of the mVV magnitude operation is significantly higher for the larger word lengths, as the number of computational stages in the unrolled CORDIC has to be increased. The compensation unit is similar for mVV and CT, showing comparable power results.

There are two main design parameters that affect the PCPE+BPS power dissipation: The number of test phases, B, and the input word length, W. Increasing the number of

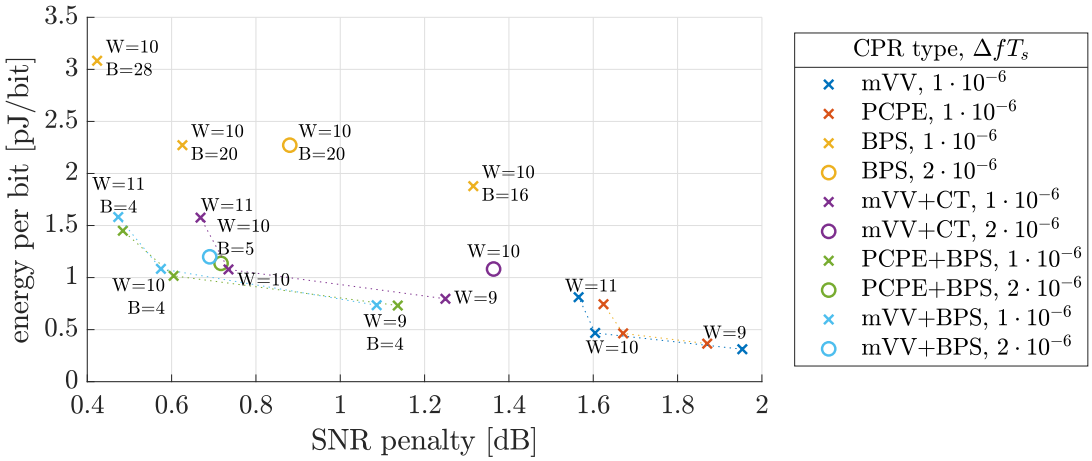


Fig. 8: Energy dissipated per bit as a function of the SNR penalty for CPR implementations with varying parameter settings.

test phases affects the power dissipation of the BPS unit, in particular the rotation and distance components, as shown in Fig. 7a for a design optimized for $\Delta f T_s = 10^{-6}$ and an SNR of 17 dB, with $N = 192$ and $W = 10$. For each extra test phase used, the total power dissipation increases with approximately 39 mW, corresponding to 0.16 pJ/bit. Fig. 7b shows how the power dissipation is affected by the choice of word length. Due to the large number of multipliers inside the square and covariance components of the PCPE unit and the rotation component of the BPS unit, these circuit portions are more affected than others when increasing W from 10 to 11 bits. Since the same DSP components are used for mVV+BPS as for the other two-stage implementations, the power-dissipation distribution for mVV+BPS can be extracted from Fig. 6 and Fig. 7.

The trade-off between energy efficiency and SNR penalty is one of the most interesting factors when comparing DSP implementations. Thus, we synthesized a number of different CPR designs with comparable parameter settings. Fig. 8 shows the results in terms of energy dissipation per bit versus penalty for two different $\Delta f T_s$ settings. The dotted lines in this graph connects implementations with all parameters held constant, except the word length W. The lowest penalty is reached for a single-stage BPS implementation with $W = 10$ and $B = 28$; we have previously shown [4] that increasing these parameters further does not significantly reduce the penalty. Reducing the number of BPS test phases decreases the power dissipation, but comes at a cost of a higher SNR penalty, which rapidly increases for lower B as hinted in Fig. 8.

Of the 10-bit single-stage implementations, the PCPE implementation is the most energy efficient, with less than 0.5 pJ/bit for $W = 10$ at $\Delta f T_s = 10^{-6}$. Incrementing W to 11 bits results only in a minor reduction of a significant penalty around 1.6 dB, at the cost of a 60% increase in power dissipation. Similar results can be observed for mVV. The SNR penalty for single-stage PCPE and mVV implementations quickly becomes large for larger linewidths (not shown in the figure), and is over 2.5 dB for $W = 10$ and $\Delta f T_s = 2 \cdot 10^{-6}$. The high SNR penalty is due to the residual phase noise left after mVV and PCPE, causing problems especially for modulation formats higher than 64QAM.

For $\Delta f T_s = 10^{-6}$, a two-stage mVV+BPS or PCPE+BPS implementation with $W = 11$ and $B = 4$ results in almost the same BER performance as that of BPS with the lowest penalty, but only at 55% of its power dissipation. Reducing the word length to 10 bits results in an energy dissipation for mVV+BPS and PCPE+BPS of 1.08 pJ/bit and 1.02 pJ/bit, respectively. Compared to a single-stage implementation using the same word length, the SNR penalty reduction from adding a second BPS stage is more than 1 dB. This reduction is even larger for larger values of $\Delta f T_s$, albeit at the cost of adding more test phases in the BPS unit, which results in a higher power dissipation. This cost is, however, still lower than using a single-stage BPS CPR unit. The SNR penalty for the mVV+CT is higher than for the other two-stage algorithms for all tested values of W, and the difference is increasing for larger $\Delta f T_s$. The penalty difference is due to CT performing worse than BPS as a second stage and is the reason why a two-stage PCPE+CT combination was not explored further.

VI. Conclusion

We have presented circuit implementations of a range of one- and two-stage carrier phase recovery (CPR) methods for 256QAM transmission. Based on our HDL implementations, we synthesized the CPR designs to a 22-nm CMOS process technology, allowing us to show how different optimizations and parameter settings affect performance and power dissipation. These results were then used to illustrate the important design trade-off between energy per bit and SNR penalty, and to compare two-stage CPR approaches to single-stage approaches, involving combinations of blind phase search (BPS), constellation transformation (CT), principal component-based phase estimation (PCPE), and Viterbi-Viterbi (VV) methods.

We have shown how the selection of averaging window size for the modified VV (mVV) algorithm affects the BER, and that this parameter has an optimal setting that varies with the laser linewidth. The mVV partitioning method, in which many symbols are set to zero, has a beneficial effect on the power dissipation of the following 4^{th}-power operation, reinforcing the importance of reducing logic switching activity to reduce

energy per bit [15]. The two most important design parameters for the PCPE+BPS CPR implementation, i.e., the covariance window length and the number of test phases, can be selected to minimize the SNR penalty compared to an ideal case. The window length has a clear optimum, but this choice does not significantly impact power dissipation. The SNR penalty is reduced by increasing the number of test phases, up to a point where a further increase does not affect the phase estimation result. However, the power dissipation increases linearly with the number of test phases. Increasing the input word length to more than 10 bits does not result in a reduction of the SNR penalty for any of our implementations, but would only increase power dissipation.

Parameters selected in a good trade-off between SNR penalty and power dissipation for a 256QAM 30-GBaud PCPE+BPS implementation results in an SNR penalty of 0.6 dB with an energy dissipation of 1.02 pJ/bit, which is less than half the energy per bit of a single-stage BPS implementation with the same SNR penalty. A mVV or PCPE single-stage implementation can never reach these low penalty values due to the residual phase noise. In addition, we have shown a >1-dB penalty improvement when adding a fine-grain BPS stage after either mVV or PCPE. For larger $\Delta f T_s$ values, this penalty improvement increases even more. The two-stage mVV+BPS and PCPE+BPS show similar performance, but the latter would be preferred due to its smaller area. The mVV+CT implementation shows energy efficiencies close to the other two-stage CPR methods, albeit at a higher SNR penalty. When $\Delta f T_s$ is increased, the performance of mVV+CT deteriorates quickly. This behavior is less pronounced when using BPS as a second stage, since we have the possibility to finely regulate the performance of this stage by incrementing the test phase count, suggesting that BPS is a more versatile choice for a second CPR stage.

REFERENCES

[1] L. Kull, T. Toifl, M. Schmatz, P. A. Francese, C. Menolfi, M. Braendli, M. Kossel, T. Morf, T. M. Andersen, and Y. Leblebici, "A 90GS/s 8b 667mW 64x interleaved SAR ADC in 32nm digital SOI CMOS," in *IEEE Int. Solid-State Circuits Conf.*, Feb. 2014, pp. 378–379.

[2] C. Fougstedt, O. Gustafsson, C. Bae, E. Börjeson, and P. Larsson-Edefors, "ASIC design exploration for DSP and FEC of 400-Gbit/s coherent data-center interconnect receivers," in *Opt. Fiber Commun. Conf. (OFC)*, Mar. 2020, p. Th2A.38.

[3] L. Kull, D. Luu, C. Menolfi, M. Braendli, P. A. Francese, T. Morf, M. Kossel, A. Cevrero, I. Ozkaya, and T. Toifl, "A 24-to-72GS/s 8b time-interleaved SAR ADC with 2.0-to-3.3pJ/conversion and >30dB SNDR at Nyquist in 14nm CMOS FinFET," in *IEEE Int. Solid-State Circuits Conf.*, 2018, pp. 358–360.

[4] E. Börjeson, C. Fougstedt, and P. Larsson-Edefors, "VLSI implementations of carrier phase recovery algorithms for M-QAM fiber-optic systems," *IEEE J. Lightw. Technol.*, vol. 38, no. 14, pp. 3616–3623, 2020.

[5] E. Börjeson, C. Fougstedt, and P. Larsson-Edefors, "ASIC design exploration of phase recovery algorithms for M-QAM fiber-optic systems," in *Opt. Fiber Commun. Conf. (OFC)*, Mar. 2019, p. W3H.7.

[6] S. M. Bilal, C. R. S. Fludger, V. Curri, and G. Bosco, "Multistage carrier phase estimation algorithms for phase noise mitigation in 64-quadrature amplitude modulation optical systems," *IEEE J. Lightw. Technol.*, vol. 32, no. 17, pp. 2973–2980, Sept. 2014.

[7] A. Viterbi and A. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Trans. Inf. Theory*, vol. 29, no. 4, pp. 543–551, 1983.

[8] M. Seimetz, "Laser linewidth limitations for optical systems with high-order modulation employing feed forward digital carrier phase estimation," in *Opt. Fiber Commun. Conf. (OFC)*, Feb. 2008, p. OTuM2.

[9] J. H. Ke, K. P. Zhong, Y. Gao, J. C. Cartledge, A. S. Karar, and M. A. Rezania, "Linewidth-tolerant and low-complexity two-stage carrier phase estimation for dual-polarization 16-QAM coherent optical fiber communications," *IEEE J. Lightw. Technol.*, vol. 30, no. 24, pp. 3987–3992, 2012.

[10] T. Pfau, S. Hoffmann, and R. Noe, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," *IEEE J. Lightw. Technol.*, vol. 27, no. 8, pp. 989–999, Apr. 2009.

[11] H. Sun, K. Wu, S. Thomson, and Y. Wu, "Novel 16QAM carrier recovery based on blind phase search," in *Eur. Conf. Opt. Commun. (ECOC)*, Sept. 2014, p. Tu.1.3.4.

[12] J. C. M. Diniz, Q. Fan, S. M. Ranzini, F. N. Khan, F. D. Ros, D. Zibar, and A. P. T. Lau, "Low-complexity carrier phase recovery based on principal component analysis for square-QAM modulation formats," *Opt. Express*, vol. 27, no. 11, pp. 15 617–15 626, May 2019.

[13] Q. Zhuge, C. Chen, and D. V. Plant, "Low computation complexity two-stage feedforward carrier recovery algorithm for M-QAM," in *Opt. Fiber Commun. Conf. (OFC)*, Mar. 2011, p. OMJ5.

[14] J. E. Volder, "The CORDIC trigonometric computing technique," *IRE Trans. Electronic Computers*, vol. EC-8, no. 3, pp. 330–334, 1959.

[15] P. Larsson-Edefors and E. Börjeson, "Power-efficient ASIC implementation of DSP algorithms for coherent optical communication," in *IEEE Photonics Society Summer Topicals Meeting Series (SUM)*, July 2020, p. MA1.1.