Fig. 3.56 show the BER/OSNR results for a 28 GBaud system. DFB lasers with 1 MHz linewidth are assumed. With such a small linewidth, the block phase estimator performs almost as good as the other three phase estimators which track the phase drift.

Fig. 3.57 shows simulated BER/SNR curves for the same phase estimators and a 10 GBaud system with DFB lasers that have 1 MHz linewidth and a residual frequency mismatch of 20 MHz. While the block phase estimator is not able to cope with high phase noise, both SML phase approximation versions have results similar to the weighted averaging method. They fulfill the high phase noise requirements of 10 GBaud transmission systems with coarsely controlled DFB lasers. The angle-based phase estimation algorithm is a good replacement for the optimized moving average phase estimator. It has been used for the first realtime QPSK transmission [64], also with polarization multiplex and control [85].

In addition to the polarization transformation at the carrier frequency, polarization mode dispersion (PMD) and chromatic dispersion (CD) need also to be compensated. A surprising lead in this field has been taken by Nortel [86]. Another set of application-specific integrated circuits and its application in coherent QPSK transmission was presented in [87].

CD can as well be pre-compensated electronically in the transmitter that is equipped with an I&Q modulator [88]. In that case not even a coherent receiver is needed. Rather, for intensity modulation a standard direct-detection ASK receiver is sufficient.

CD and PMD can be electronically equalized with finite impulse response (FIR) filters in the time domain. Beside this, convolution in the time domain can be replaced by Fast Fourier Transform (FFT), multiplication in the frequency domain and inverse FFT (IFFT). Various equalizer configurations become thereby possible [89].

One possibility is it to perform the IFFT at the transmit end and the FFT at the receive end. In this way the symbols at the IFFT input can be understood as the amplitudes of orthogonal, narrowly spaced carriers. In order to maintain orthogonality in the presence of dispersion the temporal duration of each transmitted IFFT frame is cyclically extended by usually a few percent. This scheme, called orthogonal frequency division multiplex (OFDM), is well known in the wireless world but is also a serious candidate for high-performance optical communication [89–97].

3.3.6 Digital Coherent QAM Receiver

Polarization-multiplexed QPSK is that among the modulation formats which ideally work with no more than 18 photons/bit at a BER = 10^{-9} with the highest information rate, 4 bit/symbol. QPSK itself is a special case among the square quadrature amplitude modulation (QAM) schemes. We will extend and generalize the foregoing from QPSK to M-ary QAM schemes. M-QAM, in particular 16-QAM, is attractive for shorter transmission lengths where ultimate OSNR performance is not needed. Fig. 3.58 shows various QAM constellations with $q = 4$-fold angular symmetry, commonly referred to as square QAM [98, 99]. 4-QAM is the same as QPSK.
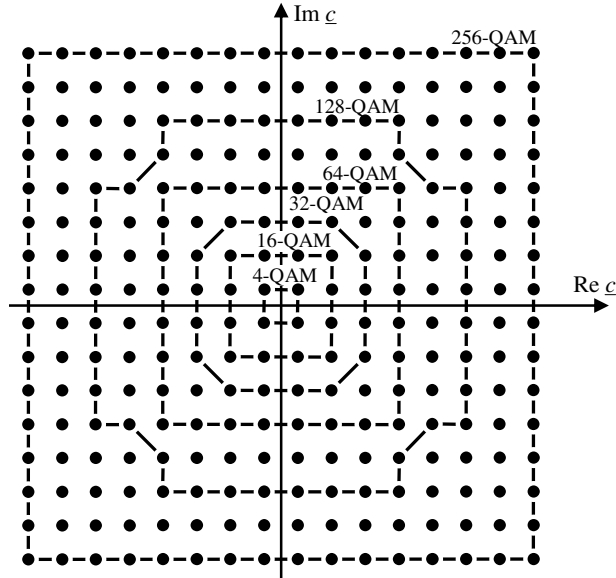
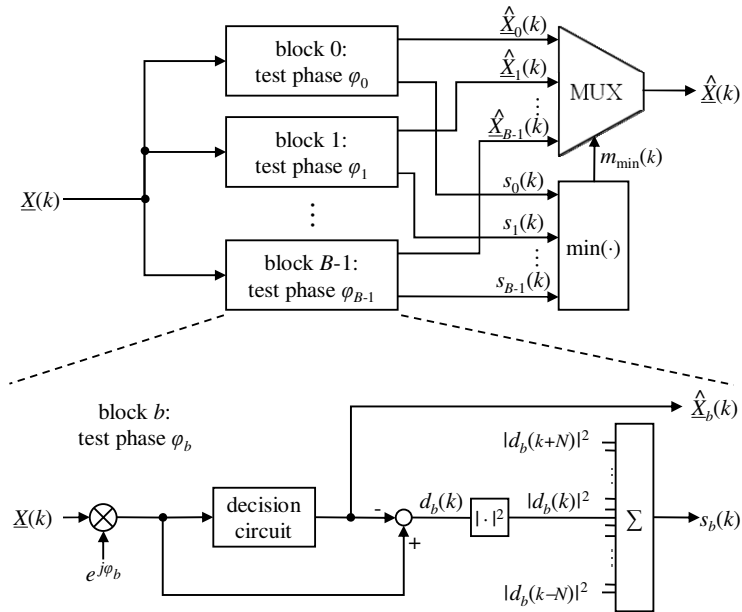**Fig. 3.58** Square QAM constellation diagrams. © 2009 IEEE.

Fig. 3.59 Feedforward carrier recovery using B test phase values φ_b. © 2009 IEEE.

For true square M-QAM signals, the inphase and quadrature symbols (or optical field levels) can be written as

$$c = 2m + 1 - \sqrt{M} , \qquad m \in \left\{ 0, 1, ..., \sqrt{M} - 1 \right\}. \tag{3.214}$$

The mean squared field is $\left\langle c^2 \right\rangle = \dfrac{M-1}{3}$. We divide by $\sqrt{2 \left\langle c^2 \right\rangle}$ and write a complex QAM signal with unity mean squared magnitude,

$$\underline{c} = \sqrt{\frac{3}{2(M-1)}} \left(\left(2m_1 + 1 - \sqrt{M} \right) + j \left(2m_2 + 1 - \sqrt{M} \right) \right),$$

$$m_1, m_2 \in \left\{ 0, 1, ..., \sqrt{M} - 1 \right\}, \qquad\qquad \left\langle |\underline{c}|^2 \right\rangle = 1 . \tag{3.215}$$

For QPSK = 4-QAM this is identical with the $\underline{c}_p$ defined in Table 3.6. The modulation variables in the two quadratures are m_1, m_2.

A phase-noise tolerant, parallelizable feedforward carrier recovery for QAM signals, subsequently abbreviated as FCR-QAM, has been presented by Pfau [98, 99]. The setup is depicted in Fig. 3.59. The k-th input signal $\underline{X}(k)$ of the coherent receiver is sampled at the symbol rate, and perfect clock recovery and equalization are assumed. To recover the carrier phase the received signal $\underline{X}(k)$ is rotated by B test carrier phase angles φ_b with

$$\varphi_b = \frac{b}{B} \cdot \frac{2\pi}{q} , \qquad b \in \left\{ 0, 1, ..., B - 1 \right\}. \tag{3.216}$$

Then all rotated symbols are fed into a decision circuit and the squared distance

$$\left| d_b(k) \right|^2 = \left| \underline{X}(k) e^{j\varphi_b} - \left\lfloor \underline{X}(k) e^{j\varphi_b} \right\rfloor_D \right|^2 = \left| \underline{X}(k) e^{j\varphi_b} - \hat{X}_b(k) \right|^2 \tag{3.217}$$

to the closest constellation point is calculated in the complex plane, taking advantage of the rotational symmetry with angles $2\pi / q$. Here $\lfloor \cdot \rfloor_D$ means the rounding to the nearest of the symbols (3.215) as executed in the decision circuit. In order to remove noise, the distances of $2N + 1$ consecutive test symbols rotated by the same carrier phase angle φ_b are summed up,

$$s_b(k) = \sum_{n=-N}^{N} \left| d_b(k-n) \right|^2 . \tag{3.218}$$

The optimum value of N depends on the laser linewidth times symbol rate product. $N = 6 ... 10$ is a fairly good choice. After filtering the optimum phase angle is determined by searching the minimum sum of distance values. As the decoding was already executed in (3.217), the decoded output symbol $\underline{\hat{X}}(k)$ can be selected

from the $\underline{\hat{X}}_b(k)$ by a switch controlled by the index $m_{\min}(k)$ of the minimum distance sum. Unless an error has occurred $\underline{\hat{X}}(k)$ is equal to the k-th transmitted symbol $\underline{c}(k)$ (3.215), but rotated by any of the q angles $2\pi \cdot (1,2,...,q-1)/q$.

In the case $q = 4$ with 4-fold ambiguity of the recovered phase, the first two bits which determine the quadrant of the complex plane should be differentially Gray-encoded. The differential encoding and decoding process is the same as for QPSK, see chapter 3.3.5. Decoding is described by

$$n_o(k) = (n_r(k) - n_r(k-1) + n_j(k))\mod q$$
$$n_o(k), n_r(k), n_j(k) \in \{0, 1, 2, \ldots, q-1\} \tag{3.219}$$

where $n_o(k)$ is the differentially decoded quadrant number, $n_r(k)$ is the received quadrant number and $n_j(k)$ is the jump number. The only required modification of the decoding process compared to QPSK is that quadrant jumps are detected according to

$$n_j(k) = \begin{cases} 1 & B/2 < m_{\min}(k) - m_{\min}(k-1) \\ 0 & -B/2 \le m_{\min}(k) - m_{\min}(k-1) \le B/2 \\ q-1 & m_{\min}(k) - m_{\min}(k-1) < -B/2 \end{cases} \tag{3.220}$$

For all other bits that determine the symbol within a quadrant of the complex plane normal Gray-coding is sufficient and no differential encoding/decoding is required. Fig. 3.60 illustrates the bit to symbol assignment including differential encoding/decoding exemplarily for square 16-QAM.

This FCR-QAM algorithm can also be applied to arbitrary QAM constellations. A q-fold rotational symmetry is already foreseen in (3.216), (3.219). If there is no symmetry then $q = 1$ holds. $\lceil \log_2 q \rceil$ bits must be differentially encoded/decoded, where $\lceil u \rceil$ is the smallest integer larger than or equal to u.

With polarization division multiplex, one may use two separate carrier recoveries. But due to the intrinsically low phase noise tolerance of QAM schemes it is advisable to implement a common carrier recovery for both polarizations. As a consequence, (3.218) is replaced by

$$s_b(k) = \sum_{p=1}^{2} \sum_{n=-N}^{N} |d_{p,b}(k-n)|^2 \tag{3.221}$$

where the added index p is the index of the polarization. Because twice as much data is available to determine the carrier phase angle, N can be halved, thereby increasing the phase noise tolerance by roughly a factor of 2.

Fig. 3.60 16-QAM bit to symbol assignment: The dashed ellipses mark imperfect Gray coding of four symbol pairs due to differential quadrant encoding. © 2009 IEEE.
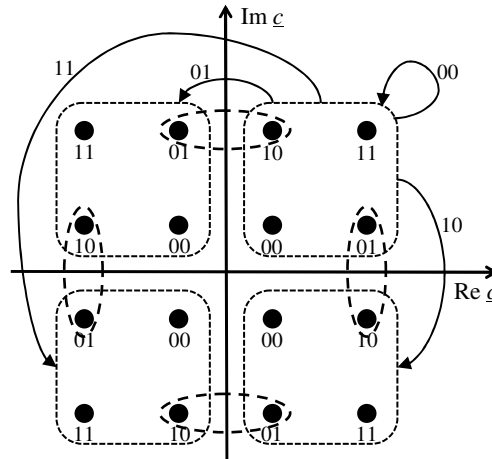
The rotation of a symbol in the complex plane normally requires a complex multiplication, consisting of four real-valued multiplications with subsequent summation. This would lead to a large number of multiplications to be executed, while achieving a sufficient resolution B for the carrier phase values φ_b. The hardware effort would therefore become prohibitive. Applying the CORDIC (coordinate rotation digital computer) algorithm [100, 101] can dramatically reduce the necessary hardware effort to calculate the B rotated test symbols. This algorithm can compute vector rotations simply by summation and shift operations. As for the calculation of the B rotated copies of the input vector intermediate results can be reused for different rotation angles, only $\sum_{b=1}^{\log_2 B} 2^{b+1}$ shift-and-add operations are required to generate the B test symbols. For example, to generate $B = 32$ rotated copies of Z_k the CORDIC algorithm requires only 124 shift-and-add operations instead of 124 real-valued multiplications and 62 summations.

To determine the closest constellation point $\hat{X}_b(k)$ the rotated symbols are fed into a decision circuit. The squared distance (3.217) at instant k can be rewritten as

$$\left|d_b(k)\right|^2 = \left(\text{Re}\left(\underline{X}(k)e^{j\varphi_b}\right) - \text{Re}\,\hat{X}_b(k)\right)^2 + \left(\text{Im}\left(\underline{X}(k)e^{j\varphi_b}\right) - \text{Im}\,\underline{\hat{X}}_b(k)\right)^2 \qquad (3.222)$$

and requires to two multiplications and three additions/subtractions. However, the subtraction results are small in magnitude and the required result resolution is moderate. Therefore $\left|d_b(k)\right|^2$ is most efficiently determined by a look-up table or basic logic functions.

Highly parallelized systems allow for a very resourceful implementation of the summation of $2N+1$ consecutive values in (3.218), (3.221). The adders can be arranged in a binary tree structure where intermediate results from different modules are reused in neighbor modules.

The overall FCR-QAM hardware effort is on the order of B times higher than for QPSK. This may be viewed as dramatic since B is in the range of 16 to 64. But taking into account that the more bits/symbol are transmitted and that the hardware effort for electronic polarization control and PMD and CD compensation is also many times higher than that for QPSK carrier recovery it appears that the overall effort is reasonable. Furthermore it is possible to implement the FCR-QAM algorithm in two stages, which reduces the required hardware effort [102].

The FCR-QAM algorithm has been simulated for the constellations 4-QAM (QPSK), 16-QAM, 64-QAM and 256-QAM. These true square constellations are easiest to generate [103, 104] and are optimally immune against additive white Gaussian noise (AWGN) [105]. The transversal filter halfwidth is always set to $N = 9$, and each data point is based on the simulation of 200,000 symbols. The results are compared against the theoretically achievable sensitivity [105]

$$\frac{E_S}{N_0} = \frac{M-1}{3}\left(Q^{-1}\left(\frac{\log_2 M}{2}\left(1-\frac{1}{\sqrt{M}}\right)^{-1}\left(1-\sqrt{1-BER}\right)\right)\right)^2 . \qquad (3.223)$$

E_S/N_0 is the optical signal to noise ratio (OSNR), M is the number of constellation points, BER is the target bit error rate and $Q(x) = \frac{1}{2}\mathrm{erfc}\left(\frac{x}{\sqrt{2}}\right)$ is the Q function. The inverse of (3.223) is

$$BER = 1-\left(1-\frac{2}{\log_2 M}\left(1-\frac{1}{\sqrt{M}}\right)Q\left(\sqrt{\frac{3}{M-1}\frac{E_S}{N_0}}\right)\right)^2 . \qquad (3.224)$$

A crucial quantity for FCR-QAM is the required number B of test phases φ_b. Fig. 3.61a shows the sensitivity penalty at the bit error rate 10^{-3} for 4-QAM and 16-QAM. FCR-QAM and a receiver with ideal carrier recovery were simulated with different resolutions for the carrier phase. Ideal carrier recovery means that the receiver knows the exact carrier phase (which is only realizable in simulation) and therefore the sensitivity penalty is only caused by differential quadrant encoding and quantization effects.

4-QAM attains a minimum penalty of 0.5 dB for the ideal receiver and 0.7 dB for FCR-QAM. The penalty difference of 0.2 dB is thus the implementation-induced penalty. For 16-QAM the minimum penalties decreases (0.4 dB for the ideal receiver, 0.6 dB for FCR-QAM), because only 2 out of 4 transmitted bits are differentially encoded. For all four receivers it can be seen that almost no additional penalty is induced due to the quantization of the carrier phase for $\log_2 B \geq 5$. Therefore in all following simulations for 4-QAM and 16-QAM B is set to 32.

Fig. 3.61b shows the same for 64-QAM and 256-QAM. The minimum penalty for 64-QAM is 0.3 dB with ideal carrier recovery and 0.5 dB with FCR-QAM. For 256-QAM the respective values are 0.35 dB and 0.55 dB. For both constellations

the penalty due to the quantization of the carrier phase is tolerable only if $\log_2 B \geq 6$. The number of test phase values for 64-QAM and 256-QAM is therefore subsequently chosen as $B = 64$.
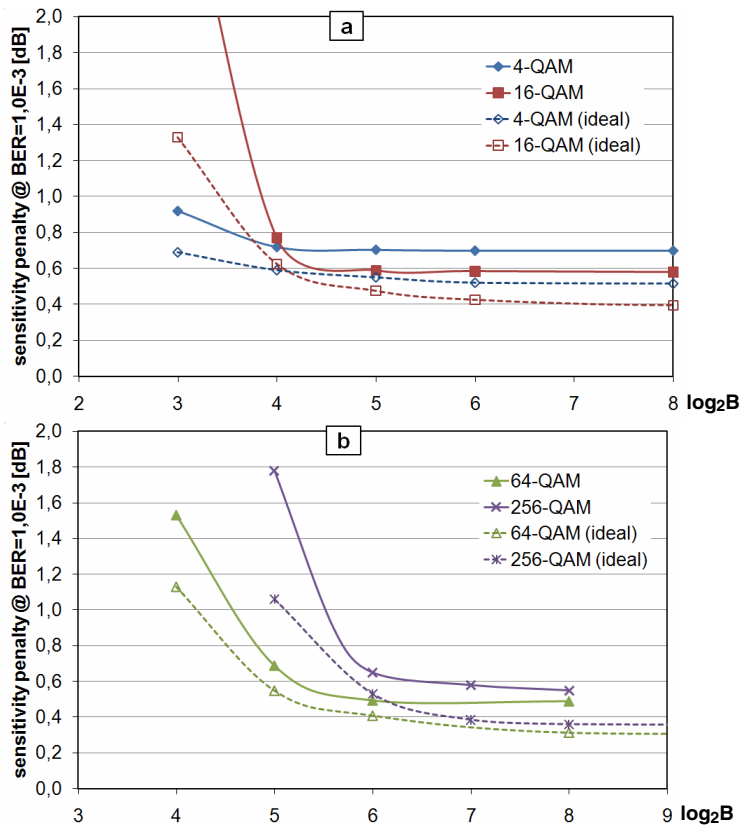

Fig. 3.61 Sensitivity penalty for different numbers of test phase values φ_b for (a) 4- and 16-QAM; (b) 64-QAM and 256-QAM. © 2009 IEEE.

FCR-QAM is advantageous because of its phase noise tolerance. Today's commercial transmission systems usually employ DFB lasers, because they are cost-efficient and have a small footprint. Their linewidth is in the range $100\,\text{kHz} < \Delta f_{\text{DFB}} < 10\,\text{MHz}$. The sum linewidth $\Delta f = 2\Delta f_{\text{DFB}}$ of signal and LO lasers is twice as much. Assuming a symbol rate of 20 Gbaud, linewidth times symbol duration products down to $10^{-5} < \Delta f \cdot T < 10^{-3}$ can be realized. Fig. 3.62 shows the sensitivity penalty of the FCR-QAM algorithm against the $\Delta f \cdot T$ product. Table 3.7 shows the maximum tolerable linewidth times symbol duration product for a sensitivity penalty of 1 dB at a BER of 10^{-3}. In a polarization multiplexed system with common carrier recovery these values can be approximately doubled.
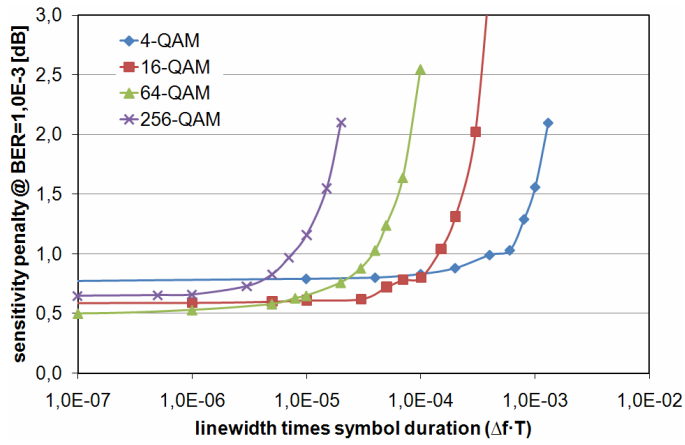
Fig. 3.62 Receiver tolerance against phase noise for different square QAM constellations. © 2009 IEEE.
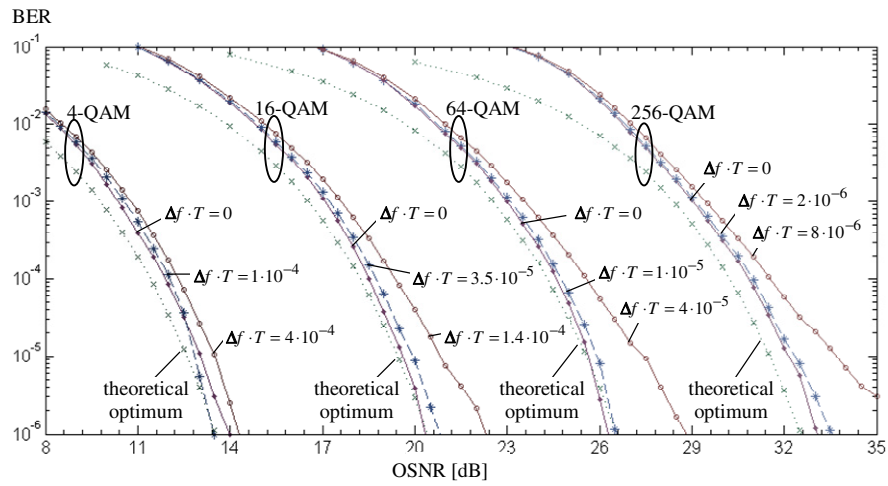

Fig. 3.63 Impact of different linewidth times symbol duration products on the receiver sensitivity of coherent QAM receivers. © 2009 IEEE.

Table 3.7 Linewidth requirements for feedforward carrier recovery with different square QAM constellations. © 2009 IEEE.

Constellation	Max. tolerable $\Delta f \cdot T$ for 1 dB penalty @ BER = 10^{-3}	Max. tolerable Δf_{DFB} for $1/T_S$ = 20 Gbaud
4-QAM	$4.1 \cdot 10^{-4}$	4.1 MHz
16-QAM	$1.4 \cdot 10^{-4}$	1.4 MHz
64-QAM	$4.0 \cdot 10^{-5}$	400 kHz
256-QAM	$8.0 \cdot 10^{-6}$	80 kHz

BER vs. OSNR has additionally been evaluated for selected values of $\Delta f \cdot T$ (Fig. 3.63). For BERs below 10^{-5} the results become inaccurate due to the low number of errors which occurred within the $2 \cdot 10^6$ symbols that were simulated per data point. The theoretical optimum is calculated by (3.224). A $\Delta f \cdot T_S$ value equal to 1/4 of that causing a ~1 dB penalty at BER = 10^{-3} gives excellent results at least down to BER = 10^{-6}.

For high OSNR the contribution of AWGN to phase noise can be considered to be also Gaussian with the variance $\left(2 E_S / N_0\right)^{-1}$ [106]. Therefore the efficiency $e(N)$ of the phase estimator depending on the filter half width N is given by

$$e(N) = \frac{\dfrac{1}{2N+1}\left(2\dfrac{E_S}{N_0}\right)^{-1}}{\left\langle (\psi - \hat{\psi})^2 \right\rangle} \le 1 . \tag{3.225}$$

where the numerator is the Cramer-Rao lower bound [107], and the denominator is the mean squared error of the phase estimator output. Fig. 3.64 shows both the mean squared error together with the theoretical optimum given by the Cramer-Rao lower bound (top row) and the resulting estimator efficiency $e(N)$ (bottom row) for the different QAM constellations.

For higher QAM constellations the efficiency of the phase estimation reduces. This can be related to the fact that for lower OSNR other limiting factors like quantization and phase noise become more dominant, Note that for a 5 bit (4-QAM, 16-QAM) and 6 bit quantization (64-QAM, 256-QAM) of $\hat{\psi}$ the minimum mean squared errors are ~$2 \cdot 10^{-4}$ and ~$5 \cdot 10^{-5}$, respectively.

Fig. 3.64 also shows that the selected filter halfwidth N = 9 is always close to the optimum filter half width for a minimum mean squared error. In principle, receiver performance could be improved by optimizing N for each parameter set {OSNR, $\Delta f \cdot T_S$, M}.
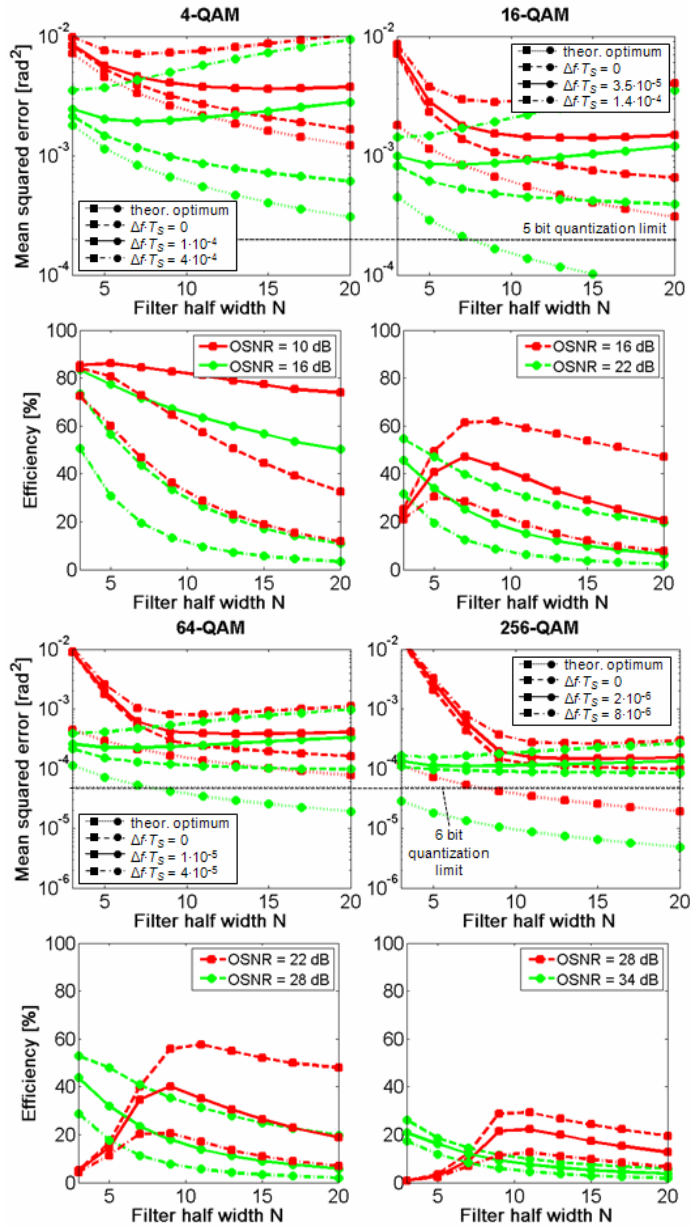
Fig. 3.64 Phase estimator mean squared error and efficiency $e(N)$ vs. filter half width N for different square QAM constellations. © 2009 IEEE.

Fig. 3.65 shows the effect of the analog-to-digital converter (ADC) resolution on receiver sensitivity. The necessary ADC resolution increases approximately by 1 bit if the number of constellation points is multiplied by 4. Table 3.8 summarizes the ADC requirements for a 100 Gb/s polarization multiplexed transmission system. Since commercial systems will also contain PMD and CD compensation, which necessitates oversampling, the values for $T_S/2$ sampling are also given. Tremendous progress is currently being made, and ADC resolution and speed targets are being met [108, 109].
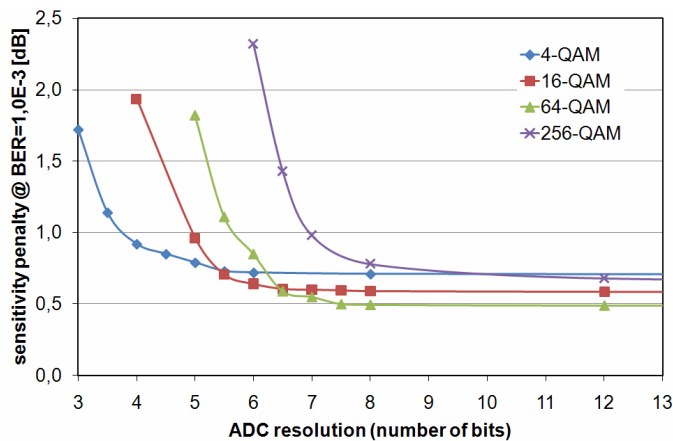

Fig. 3.65 Receiver sensitivity penalty vs. analog-to-digital converter resolution for different square QAM constellations. © 2009 IEEE.

Table 3.8 Analog-to-digital converter requirements for polarization-multiplexed 100 Gb/s transmission. © 2009 IEEE.

Constellation	ADC bandwidth	ADC sampling rate ($T_S/2$ sampling)	ADC effective number of bits
4-QAM	25 GHz	50 Gs/s	> 3.8
16-QAM	12.5 GHz	25 Gs/s	> 4.9
64-QAM	8.33 GHz	16.7 Gs/s	> 5.7
256-QAM	6.25 GHz	12.5 Gs/s	> 7.0

Fig. 3.66 shows the receiver sensitivity penalty against different resolutions of Re d_b and Im d_b. For all considered constellations a resolution ≥ 4 bits is sufficient. As for $|d_b|^2$ the penalty for a resolution ≥ 5 bits is tolerable (Fig. 3.67), and the square operations in (3.222) can be realized with a small look-up table (4 bit input, 4 bit output). The reason for the similar requirements in all QAM constellations is that the distance to the closest constellation point is independent of the number of constellation points. So, the hardware effort of FCR-QAM increases only moderately with the QAM order, which in turn determines overall spectral efficiency and hence the permissible cost frame.

To demonstrate the feasibility of optical high-order QAM transmission Figs. 3.68−3.71 depict the constellation diagrams at the receiver before and after carrier recovery for the different modulation formats and different optical signal-to-noise-ratios (OSNR). All constellation diagrams can be recovered using FCR-QAM for the different QAM modulation schemes and their respective linewidth times symbol rate products.
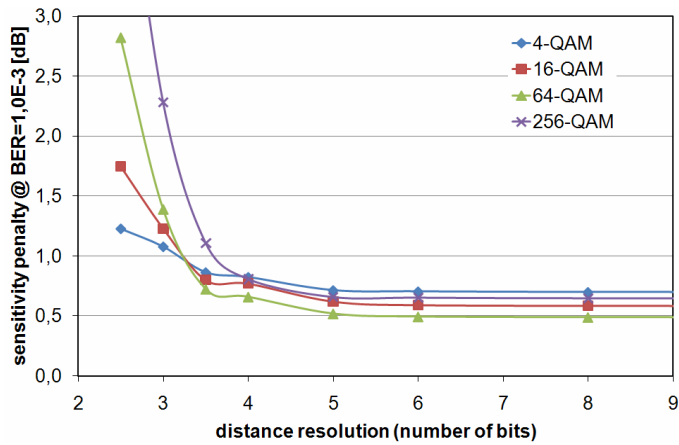

Fig. 3.66 Receiver sensitivity penalty vs. internal resolution for the distances Re d_m and Im d_m for different square QAM constellations. © 2009 IEEE.
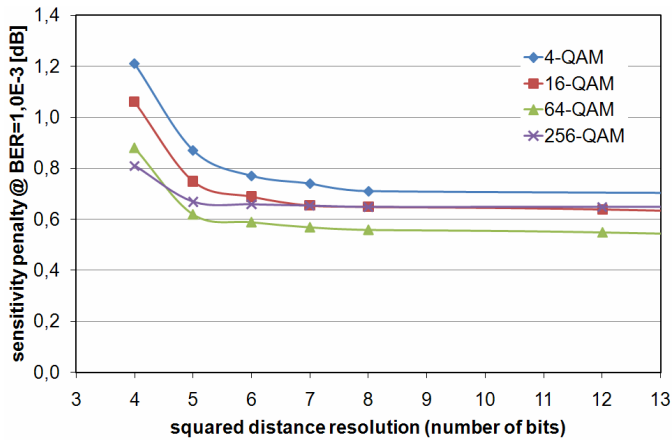

Fig. 3.67 Receiver sensitivity penalty vs. internal resolution for the squared distance $|d_b|^2$ for different square QAM constellations. © 2009 IEEE.
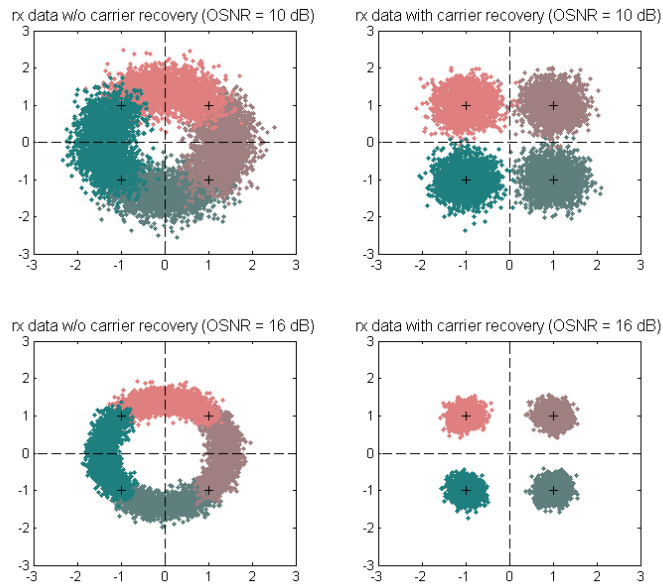
Fig. 3.68 4-QAM constellation diagrams before and after carrier recovery for $\Delta f \cdot T_S = 4.1 \cdot 10^{-4}$. © 2009 IEEE.
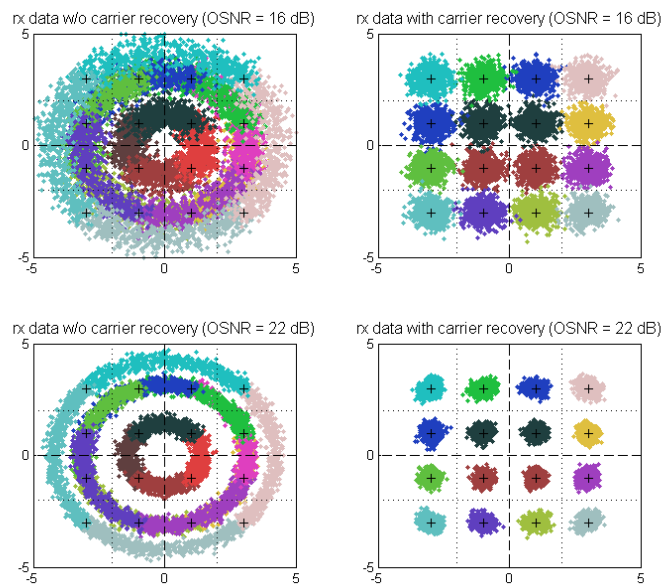

Fig. 3.69 16-QAM constellation diagrams before and after carrier recovery for $\Delta f \cdot T_S = 1.4 \cdot 10^{-4}$. © 2009 IEEE.
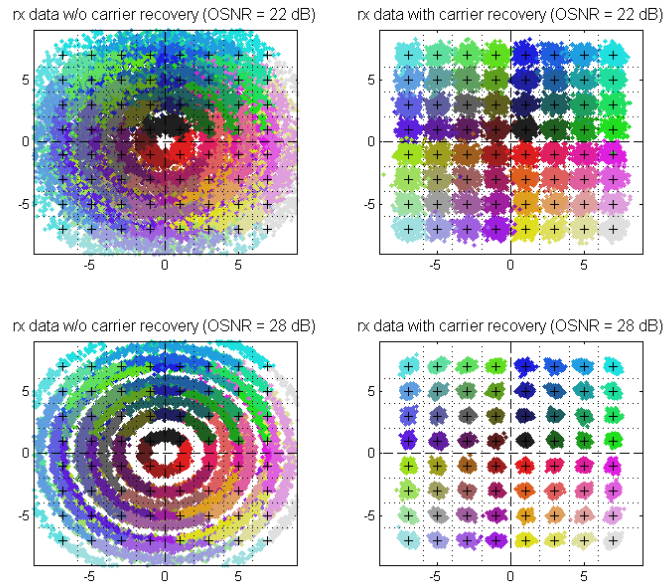
rx data w/o carrier recovery (OSNR = 22 dB)   rx data with carrier recovery (OSNR = 22 dB)

rx data w/o carrier recovery (OSNR = 28 dB)   rx data with carrier recovery (OSNR = 28 dB)

Fig. 3.70 64-QAM constellation diagrams before and after carrier recovery for $\Delta f \cdot T_S = 4.0 \cdot 10^{-5}$. © 2009 IEEE.

rx data w/o carrier recovery (OSNR = 28 dB)   rx data with carrier recovery (OSNR = 28 dB)

rx data w/o carrier recovery (OSNR = 34 dB)   rx data with carrier recovery (OSNR = 34 dB)
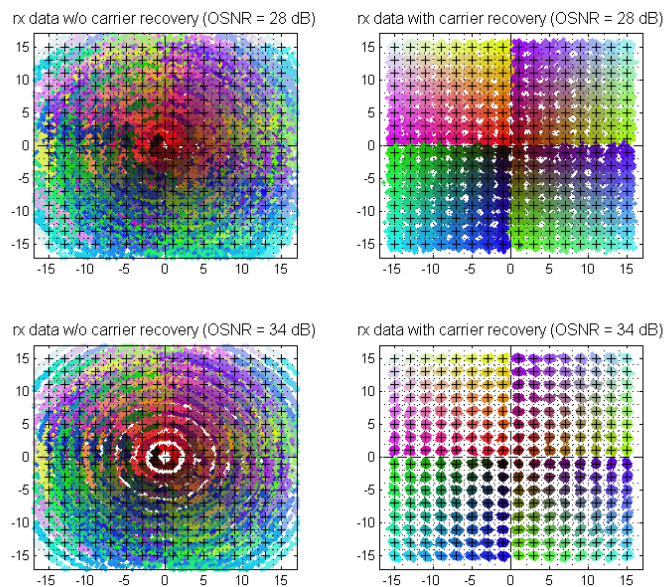
Fig. 3.71 256-QAM constellation diagrams before and after carrier recovery for $\Delta f \cdot T_S = 8.0 \cdot 10^{-6}$. © 2009 IEEE.

Let the transmitted symbol **c** and its recovered replica **r** (ideally equal to **c**, except for the time delay) have unit powers in each polarization ($\left|r_p\right|^2 = \left|c_p\right|^2 = 1$, p = 1, 2). The decision-directed electronic polarization control algorithm (3.200), (3.201) works also for polarization-multiplexed QAM.

Matters are more complicated for the constant-modulus polarization control algorithm (3.200), (3.204), which is non-data-aided. If it is used as is then the control gain g must be chosen so low that the deviations of $\left|\underline{X}_{1,2}\right|^2$ from 1 are averaged out effectively. This slows the control down considerably. In this context a CMA for 8-QAM signals [110] and a decision-directed CMA for QAM [111] have been published.

For adaptation of the CMA for QAM we modify (3.204) to become

$$\mathbf{T} = \begin{bmatrix} \Delta P_{1,\min} & 0 \\ 0 & \Delta P_{2,\min} \end{bmatrix} \mathbf{X}\mathbf{R}^+, \tag{3.226}$$

$$\Delta P_{p,\min} = \left[\Delta P_{p,h_p}\right]_{\min_{\forall h_p}\left|\Delta P_{p,h_p}\right|} \qquad \Delta P_{p,h_p} = \hat{P}_{p,h_p} - \left|\underline{X}_p\right|^2. \tag{3.227}$$

$\Delta P_{p,h_p}$ is the power difference between the observed signal powers $\left|\underline{X}_p\right|^2$ in both polarizations p = 1, 2 and all expected values $\hat{P}_{p,h_p}$ (h_p = 1, 2, ... , H) of signal powers in case of zero polarization crosstalk. $\Delta P_{p,\min}$ is that value $\Delta P_{p,h_p}$ which has the smallest magnitude, hence most likely power difference. Once the algorithm has converged all power differences $\Delta P_{p,\min}$ would be zero in the absence of noise. The number H of distortion-free signal powers $\hat{P}_{p,h_p}$ equals the number of circles around the origin that are needed to touch all symbols in the chosen QAM constellation. The value is H = 1, 3, 6, 9 for 4-, 16-, 64- and 256-QAM, respectively.

The extension (3.205), (3.206) of the CMA for differential phase compensation is even more difficult to adapt for QAM. Similar as for FCR-QAM, a number of test phases

$$0 \leq \varphi_c < 2\pi/q \tag{3.228}$$

is subtracted from the observed phase difference $\arg \underline{X}_1 - \arg \underline{X}_2$. Which test phases are needed depends on the signal powers $\hat{P}_{p,h_p}$ that have been estimated in the minimization process (3.227). Depending on the estimated powers $\hat{P}_{1,h_1}$,

$\hat{P}_{2,h_2}$ there are different numbers of test phases. Then one selects, for usage in (3.205),

$$\Delta\zeta = \left[\Delta\zeta_c\right]_{\min_{\forall\,\varphi_c}|\Delta\zeta_c|} \qquad \Delta\zeta_c = \left(\left(\arg\underline{X}_1 - \arg\underline{X}_2 - \varphi_c + \varphi_f + \frac{\pi}{4}\right)\bmod\frac{\pi}{2}\right) - \frac{\pi}{4}$$

(3.229)

i.e., that value $\Delta\zeta$ among the various phase differences $\Delta\zeta_c$ found for the applicable set of test phases φ_c which has the lowest magnitude $\left|\Delta\zeta_c\right|$. For the time being let $\varphi_f = 0$. In order to improve accuracy in the presence of noise the phase difference can be weighted according to the available or estimated powers, for example using

$$\Delta\zeta = \left[\Delta\zeta_c\right]_{\min_{\forall\,\varphi_c}|\Delta\zeta_c|} \hat{P}_{1,h_1}\,\hat{P}_{2,h_2}\,.$$

(3.230)

Consider 16-QAM as an example. We assume normalization with respect to the mean power. Then the expected powers in the individual polarization channels and the needed test phases are

$$\hat{P}_{p,h_p} = \begin{cases} 0.2 & \text{for } h_p = 1 \\ 1 & \text{for } h_p = 2 \\ 1.8 & \text{for } h_p = 3 \end{cases}, \qquad \varphi_c = \begin{cases} 0 & \text{for } h_1,h_2 \in \{1,3\} \\ 0,\pm 0.64 & \text{for } h_1 = h_2 = 2 \\ \pm 0.46 & \text{otherwise} \end{cases}. \qquad (3.231)$$

While the CMA with differential phase compensation can, in this configuration, track a once-acquired optimum, initial false locking is all the same possible. Similar to the *p*-fold phase ambiguity there exist different locking points, but only one of them yields a product **MJ** proportional to the unity matrix, while in the other cases there is a static phase shift between the two polarization channels. In order to quit a false optimum one can proceed similar as in FCR-QAM. One calculates $\Delta\zeta$ according to (3.229) not only for $\varphi_f = 0$ but for D equidistant phase offsets

$$\varphi_f = \frac{f}{F}\cdot\frac{2\pi}{q}, \qquad f \in \{0,1,...,F-1\}.$$

(3.232)

A good choice is $F = 8\sqrt{M}$; hence F may be equal to B specified in (3.216). The squares of the various $\Delta\zeta = \Delta\zeta(f,k)$ are added up over K subsequent symbols,

$$W_f = \sum_{k=1}^{K}(\Delta\zeta(f,k))^2\,.$$

(3.233)

One determines that integer f which corresponds to the smallest W_f. This indicates that a better optimum is available if one introduces a differential phase shift φ_f between the polarizations. So, after K symbols one sets

$$\mathbf{M} := \begin{bmatrix} e^{j\varphi_f/2} & 0 \\ 0 & e^{-j\varphi_f/2} \end{bmatrix} \mathbf{M} . \qquad (3.234)$$

Thereafter the summation process (3.233) may start anew. But this is usually not needed because a single application of (3.234) usually yields the a differential phase very close to the optimum, which is subsequently improved and tracked as described above. As a consequence, the full set (3.229) obtained for all f and the summation (3.233) may be executed less frequently than the tracking calculation for $\varphi_f = 0$. This reduces the required hardware effort.

A more hardware-efficient way to avoid false locking is the following: After the decision circuits a framing information is detected which indicates whether data is being received correctly in both polarization channels. If not, then (3.234) is executed with a suitably chosen φ_f, or with all values given in (3.232), until data recovery is correct. This doesn't slow down normal control, since it occurs only at initial signal acquisition.

CMA-QAM and DPC-CMA-QAM were verified in short simulations without noise, for 4-, 16-, 64- and 256-QAM. Subsequently noise has been added, and standard CMA, CMA-QAM and DPC-CMA-QAM have been compared for 16-QAM [76]. Fig. 3.72 shows the sensitivity degradation at BER = 10^{-3} as a function of control gain g, for a sum linewidth times symbol duration product of $\Delta f \cdot T = 2 \cdot 10^{-4}$. $1/g$ is proportional to the small-signal polarization control time constant, so a large g is advantageous. The total penalty reaches 2 dB for the standard CMA at $g \approx 2^{-6.5}$ while the CMA-QAM can control polarization about 3 times faster, with $g = 2^{-5}$. For both, the two carrier recoveries processed 19 symbols in parallel. The DPC-CMA-QAM with one joint carrier recovery for both polarizations processes only 9 temporal samples. This means that phase noise is better tolerated, and sensitivity is therefore increased by ~0.5 dB. Fig. 3.73 shows the same simulations, but at a doubled $\Delta f \cdot T = 4 \cdot 10^{-4}$. The sensitivity advantage of the common carrier recovery in the case of DPC-CMA-QAM is even more pronounced.

All QAM polarization control algorithms are compared in Fig. 3.74, again as a function of polarization control gain g [76]. Performance is worst for the original decision-directed algorithm (ODDA) [63] with $\mathbf{T} = (\mathbf{1} - \mathbf{Q})\mathbf{M}$ employed in (3.200) and $\mathbf{Q}$ given by (3.199). Penalty traces for CMA-QAM and DPC-CMA-QAM are fairly identical, due to $\Delta f \cdot T = 0$. The modified decision-directed algorithm (MDDA) (3.200), (3.201) performs best. At 0.5 dB penalty the MDDA is ~15 times faster than the standard CMA and ~4 times faster than the DPC-CMA-QAM.
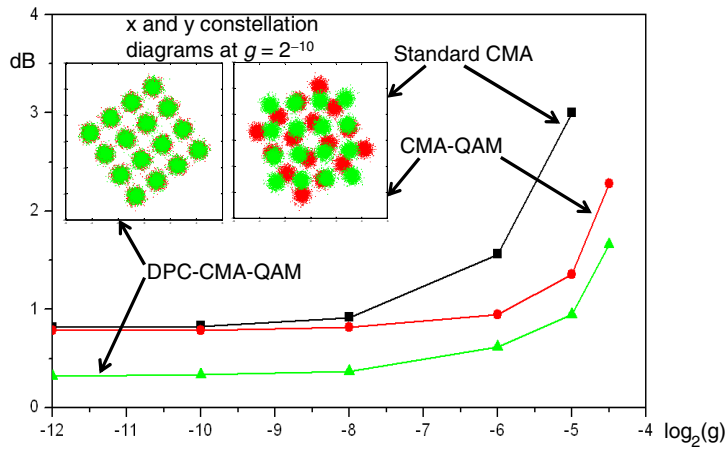
Fig. 3.72 Sensitivity penalty at BER = 10^{-3} of various CMAs applied to polarization-multiplexed 16-QAM signals vs. control gain g, for $\Delta f \cdot T_S = 2 \cdot 10^{-4}$, **MJ = 1**. Each data point corresponds to 200,000 symbols. Insets: Constellation diagrams, with phase difference between polarization channels compensated in the case of DPC-CMA-QAM. © 2010 IEEE.
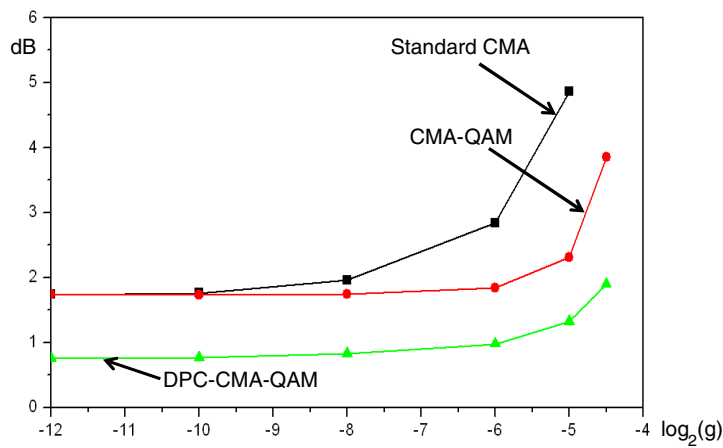

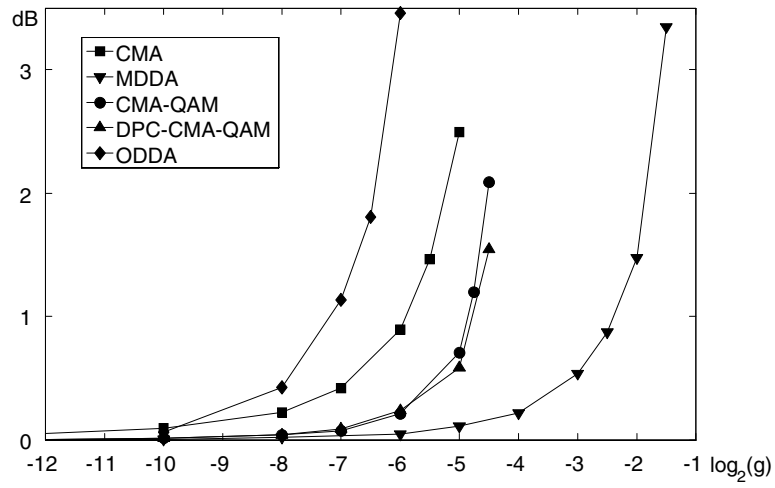**Fig. 3.73** As above, but for $\Delta f \cdot T_S = 4 \cdot 10^{-4}$. © 2010 IEEE.

Fig. 3.74 Sensitivity penalty at BER = 10^{-3} of various 16-QAM polarization control algorithms vs. control gain g, for $\Delta f \cdot T_S = 0$, **MJ = 1**. The MDDA outperforms all other algorithms. © 2010 IEEE.