Soft-Decision Multiple-Symbol Differential Sphere Detection and Decision-Feedback Differential Detection for Differential QAM Dispensing with Channel Estimation in the Face of Rapidly Fading Channels

Chao Xu, *Member, IEEE*, Li Wang, *Member, IEEE*, Soon Xin Ng, *Senior Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—Turbo detection performed by exchanging extrinsic information between the soft-decision QAM detector and the channel decoder is beneficial for the sake of exploring the bit dependency imposed both by modulation and by channel coding. However, when the soft-decision coherent QAM detectors are provided with imperfect channel estimates in rapidly fading channels, they tend to produce potentially unreliable LLRs that deviate from the true probabilities, which degrades the turbo detection performance. Against this background, in this paper, we propose a range of new soft-decision multiple-symbol differential sphere detection (MSDSD) and decision-feedback differential detection (DFDD) solutions for differential QAM (DQAM), which dispense with channel estimation in the face of rapidly fading channels. Our proposed design aims for solving the two inherent problems in soft-decision DQAM detection design, which have also been the most substantial obstacle in the way of offering a solution for turbo detected MSDSD aided differential MIMO schemes using QAM: 1) how to facilitate the soft-decision detection of the DQAM's amplitudes, which—in contrast to the DPSK phases—do not form a unitary matrix, and 2) how to separate and streamline the DQAM's soft-decision amplitude and phase detectors. Our simulation results demonstrate that our proposed MSDSD aided DQAM solution is capable of substantially outperforming its MSDSD aided DPSK counterpart in coded systems *without imposing a higher complexity*. Moreover, our proposed DFDD aided DQAM solution is shown to outperform the conventional solutions in literature. Our discussions on the important subject of coherent versus noncoherent schemes suggest that compared to coherent square QAM relying on realistic imperfect channel estimation, MSDSD aided DQAM may be deemed as a better candidate for turbo detection assisted coded systems operating at high Doppler frequencies.**

***Index Terms*—Absolute-amplitude phase shift keying, decision-feedback differential detection, differential-amplitude phase shift keying, differential QAM, multiple-symbol differential sphere detection, soft-input, soft-output, turbo detection.**

Manuscript received September 14, 2015; revised January 11, 2016; accepted March 8, 2016. Date of publication March 14, 2016; date of current version June 7, 2016. This work was supported by the European Research Council 10.13039/501100000781 Advanced Fellow Grant. The associate editor coordinating the review of this paper and approving it for publication was L. Song.

C. Xu, S. X. Ng, and L. Hanzo are with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: cx1g08@ecs.soton.ac.uk; sxn@ecs.soton.ac.uk; lh@ecs.soton.ac.uk).

L. Wang is with Huawei Technology Sweden R&D Competence Center, Stockholm, Sweden (e-mail: leo.li.wang@huawei.com).

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

Digital Object Identifier 10.1109/TWC.2016.2541665

I. Introduction

HIGH-COMPLEXITY joint channel- and data-estimation is capable of approaching the performance of perfect channel estimation in slowly fading channels [1]–[6], but both its complexity and pilot-overhead escalate in high-Doppler scenarios [7]. Moreover, in pursuit of an increased data-rate, the high-order 64QAM and 256QAM schemes have been included in the ITU-R IMT Advanced 4G standards [8] and in IEEE 802.11ac [9]. As the modulation order increases, the dependency between the modulated bits is also increased. Therefore, it is beneficial to exchange extrinsic information between a soft-decision QAM detector and a channel decoder in turbo detection assisted coded systems [10]–[14], so that the capacity limits of the coded QAM systems may be closely approached. However, when the soft-decision coherent demodulators are provided with imperfect Channel State Information (CSI) in the face of rapidly fading channels, they tend to produce potentially unreliable LLRs that deviate from the true probabilities [14]. This degrades the performance of turbo detection. By contrast, the noncoherent detection of Differential QAM (DQAM) mitigates the above-mentioned problems, where the CSI does not have to be estimated.

More explicitly, as an instantiation of Star QAM, the differential encoding principle of DPSK was first applied to DQAM's phase only, while DQAM's data-carrying amplitude was directly transmitted in [15]. Without differential encoding on amplitude, this transmission mechanism may be termed as absolute-amplitude DQAM, which may include Absolute-amplitude Differential Phase Shift Keying (ADPSK) [15], Twisted ADPSK (TADPSK) [16] and TADPSK associated with Joint Mapping (TADPSK^{JM}) [17], where TADPSK introduces a ring-amplitude-dependent phase rotation in order to increase the Star QAM constellation distances, while TADPSK^{JM} jointly maps its bits to DQAM's phase index and ring-amplitude index in order to increase the correlation between the two terms. Moreover, as a popular alternative, Differential

Amplitude Phase Shift Keying (DAPSK) [18] applies differential encoding both to the phase and to the ring-amplitude, which constitutes the family of differential-amplitude DQAM schemes that may also include the constellation variants of Twisted DAPSK (TDAPSK) and TDAPSK associated with Joint Mapping ($\text{TDAPSK}^{\text{JM}}$). In this paper, the notational form of M-DQAM(M_A, M_P) is used for all the DQAM schemes, where M, M_A and M_P refer to the number of modulation levels, ring-amplitudes and phases, respectively. They have the relationship of $M = M_A M_P$.

In the absence of CSI, the DQAM's Conventional Differential Detection (CDD) suffers from a performance erosion compared to its coherent counterparts [18], [19]. In order to improve the CDD's performance, Multiple-Symbol Differential Detection (MSDD) was conceived both for DPSK in [20]–[22] and for DQAM in [22]. More explicitly, the MSDD extends the CDD's observation window width from $N_w = 2$ to $N_w \geq 2$, where a total of $(N_w - 1)$ data-carrying symbols are jointly detected. Consequently, the MSDD complexity may grow exponentially with N_w. In order to circumvent this problem, Decision-Feedback Differential Detection (DFDD) was conceived for DPSK in [23], [24] and for DQAM in [17], [25]–[27], where the decision feedbacks concerning a total of $(N_w - 2)$ data-carrying symbols are obtained from the previous detection windows, so that only a single symbol has to be detected. However, the DFDD's error propagation problem results in a performance loss. In order to retain the optimum MSDD performance, the concept of Multiple-Symbol Differential Sphere Detection (MSDSD) was proposed for DPSK in [28]–[30], where the Sphere Decoder (SD) was invoked for MSDD.

Inspired by the technical breakthrough of Turbo Code (TC) [31], [32], soft-decision DQAM detection has also been developed throughout the last two decades. Explicitly, Trellis decoded DQAM using the Viterbi algorithm was proposed in [33]. Then, MSDD was invoked for DQAM in the context of multilevel coding in [34]. Moreover, a low-complexity soft-decision CDD was conceived for DAPSK in Rayleigh fading channels in [35], where the ring-amplitude and phase are separately detected. However, no iteration was invoked between the channel decoder and the DQAM detector in these contributions. In [36], the CDD aided DAPSK was employed for turbo detection, where the ring-amplitude and phase are jointly detected. This soft-decision CDD conceived for DAPSK was further streamlined in [37], where the authors also discovered that completely separately detecting the DAPSK's ring-amplitude and phase may impose a performance loss, which is more substantial in coded systems. Furthermore, in [38], a new MSDD/MSDSD arrangement was proposed for soft-decision DAPSK detection, which may be termed as MSDD using Iterative Amplitude/Phase processing (MSDD-IAP). Explicitly, the MSDD-IAP of [38] invokes MSDD and MSDSD for detecting the DAPSK's ring-amplitudes and phases, and then the two detectors may iteratively exchange their decisions in order to achieve the near-optimum MSDD performance. Against this background, at the time of writing, the soft-decision SD has not been invoked for the DQAM's ring-amplitude detection. Furthermore, more soft-decision MSDSD arrangements have to be conceived, because the MSDD-IAP of [38] cannot be directly applied to all the aforementioned DQAM constellations.

In recent years, the differential Multiple-Input Multiple-Output (MIMO) schemes have attracted a lot of attention. More explicitly, Differential Space-Time Modulation (DSTM) design based on group codes was proposed in [39]–[41]. Moreover, Differential Space-Time Block Codes (DSTBCs) were developed based on orthogonal code design principles in [42], [43]. The DSTBCs were further developed for using QAM in [44]–[46] and for employing both non-orthogonal as well as non-unitary designs in [47]. Moreover, in pursuit of a higher rate, the concept of Differential Linear Dispersion Code (DLDC) was proposed in [48], which disperses a total of N_Q real-valued symbols with the aid of N_Q Hermitian dispersion matrices, and then the Cayley transform is used for converting the resultant Hermitian matrix into a unitary matrix. Inspired by the DLDC and the recently developed Spatial Modulation (SM) concept [49], Differential Space-Time Shift Keying (DSTSK) was proposed in [50], where a single one out of a total of N_Q unitary dispersion matrices is activated in order to disperse a single complex-valued PSK/QAM symbol [50]–[52]. Furthermore, the concept of Differential SM (DSM) was proposed in [53], which may be viewed as a special case of DSTSK [54], where only a single RF chain is employed at the DSM transmitter. The DSM was further developed for using QAM in [55].

Apart from these CDD aided differential MIMO schemes operating in slowly fading channels, the hard-decision MSDD aided DSTBC was developed in [56] for slowly fading channels, while the hard-decision MSDD aided DSTM was derived in [57] for rapidly fading channels. Moreover, the hard-decision DFDD aided DSTM and the hard-decision MSDSD aided DSTM/DSTBC using PSK operating in rapidly fading channels were proposed in [57] and [58], respectively.

At the time of writing, the employment of soft-decision MSDSD/DFDD aided differential MIMO using QAM would remain infeasible without solving the inherent soft-decision DQAM's amplitude detection problems. Let us consider the hard-decision MSDSD aided DSTBC using QAM in [51] as an example. In uncoded systems the amplitude of the first DSTBC matrix transmitted in a MSDSD window, which does not carry source information is estimated based on the decision-feedback obtained from the previous MSDSD window in [51]. However, in coded systems we observe that the potential error propagation problem may severely erode the LLR reliability of the soft-decision DQAM detection, which degrades the turbo detection performance. Moreover, without reducing the complexity of soft-decision DQAM detection, its extension to differential MIMO schemes can hardly be affordable in turbo detection assisted coded systems.

Against this background, in this paper, we opt to develop a comprehensive solution for the soft-decision DQAM detection in rapidly fading Single-Input Multiple-Output (SIMO) channels, which aims for solving the following two major problems: (1) *how to facilitate the soft-decision detection of the DQAM's amplitudes, which – in contrast to the DPSK phases – do not form a unitary matrix;* (2) *how to separate and streamline the DQAM's soft-decision amplitude and phase detectors.* The soft-decision MSDSD arrangements proposed in this paper are

TABLE I
THE SOFT-DECISION MSDSD ARRANGEMENTS PROPOSED FOR DQAM IN THIS PAPER

	Original set	HDD set	SDD set
Type 1 (Generic Algorithm)	MSDSD	HDD-MSDSD	SDD-MSDSD
Type 2 (RC Algorithm)	RC MSDSD	RC HDD-MSDSD	RC SDD-MSDSD
Type 3 (RC & IAP Algorithm)	RC MSDSD-IAP	RC HDD-MSDSD-IAP	RC SDD-MSDSD-IAP

TABLE II
MODULATION OF THE DQAM'S DATA-CARRYING SYMBOLS

(a) Absolute-Amplitude DQAM ($s_n = \frac{1}{|s_{n-1}|} x_{n-1} s_{n-1}$)

ADPSK	$x^m = \gamma^a \omega^p = \frac{\alpha^{\check{a}}}{\sqrt{\beta}} \exp(j\frac{2\pi}{M_P}\check{p})$
TADPSK	$x^m = \gamma^a \omega^p \psi^a = \frac{\alpha^{\check{a}}}{\sqrt{\beta}} \exp(j\frac{2\pi}{M_P}\check{p}) \exp(j\frac{2\pi}{M}\check{a})$
TADPSK$^{\text{JM}}$	$x^m = \frac{\alpha^{(\check{m} \bmod M_A)}}{\sqrt{\beta}} \exp(j\frac{2\pi}{M}\check{m})$

(b) Differential-Amplitude DQAM ($s_n = x_{n-1} s_{n-1}$)

DAPSK	$x^m = \gamma^a \omega^p = \frac{\alpha^{[(\check{a}+\mu_{n-1}) \bmod M_A]}}{\alpha^{\mu_{n-1}}} \exp(j\frac{2\pi}{M_P}\check{p})$
TDAPSK	$x^m = \gamma^a \omega^p \psi^a = \frac{\alpha^{[(\check{a}+\mu_{n-1}) \bmod M_A]}}{\alpha^{\mu_{n-1}}} \exp(j\frac{2\pi}{M_P}\check{p}) \exp(j\frac{2\pi}{M}\check{a})$
TDAPSK$^{\text{JM}}$	$x^m = \frac{\alpha^{\{[(\check{m} \bmod M_A)+\mu_{n-1}] \bmod M_A\}}}{\alpha^{\mu_{n-1}}} \exp(j\frac{2\pi}{M}\check{m})$

summarized in Table I. More explicitly, the novel contributions of this paper are as follows:

1) First of all, we propose to invoke the soft-decision SD for both ring-amplitude and phase detection, which has not been presented in the open literature. Depending on the specific treatment of the first DQAM symbol's amplitude Γ_1 of each MSDSD window, we propose the original set, the Hard-Decision-Directed (HDD) set and the Soft-Decision-Directed (SDD) set of MSDSD arrangements of Table I, where Γ_1 is either detected as an unknown variable or alternatively, it is estimated based on hard-decision feedback or soft-decision feedback from the previous MSDSD window, respectively. We will demonstrate that both the original set and the SDD set are suitable for the differential-amplitude DQAM schemes of DAPSK, TDAPSK and TDAPSK$^{\text{JM}}$, while the HDD set is a better choice for the absolute-amplitude schemes of ADPSK, TADPSK and TADPSK$^{\text{JM}}$.
2) Secondly, depending on the DQAM mapping technique, we propose three types of MSDSD arrangements using three new MSDSD algorithms, as seen in Table I. Explicitly, the first type of generic MSDSD arrangements in Table I may be invoked by the joint-mapping based DQAM schemes of TADPSK$^{\text{JM}}$ and TDAPSK$^{\text{JM}}$. The second type of Reduced-Complexity (RC) MSDSD arrangements in Table I may be employed by the twisted DQAM of TADPSK and TDAPSK. Moreover, the third type of RC MSDSD arrangements using IAP in Table I, which separately and iteratively detect the DQAM's ring-amplitudes and phases may be employed by ADPSK and DAPSK. Specifically, we will demonstrate that the RC HDD-MSDSD-IAP aided ADPSK is capable of outperforming its MSDSD aided DPSK counterpart [30] *without imposing a higher complexity*.
3) Furthermore, the important subject of coherent versus noncoherent detection is discussed. Our simulation results suggest that compared to the coherent Square QAM relying on the classic Pilot Symbol Assisted Modulation (PSAM) [7], the DQAM schemes employing MSDSD may be deemed as a better candidate for turbo detection aided coded systems operating at high Doppler frequencies.
4) Moreover, we further propose improved soft-decision DFDD solutions conceived for DQAM, which are equivalent to the MSDD/MSDSD operating in decision-feedback mode. We will demonstrate that the proposed DFDD solutions outperform the conventional prediction-based DFDD solutions of [17], [25]–[27].

The rest of this paper is organized as follows. The DQAM constellations and their notations are introduced in Sec. II. The soft-decision MSDD is configured for DQAM in Sec. III, where the concepts of HDD and SDD are introduced. The three new soft-decision MSDSD algorithms are proposed in Sec. IV, and the improved soft-decision DFDD solutions are developed in Sec. V. Our simulation results are provided in Sec. VI, while our conclusions are offered in Sec. VII.

II. DQAM CONSTELLATIONS

The mapping of the DQAM data-carrying symbols x and transmitted symbols s is summarized in Table II, where the notations γ, ω and ψ represent the ring-amplitude, phase and ring-amplitude-dependent phase rotation of x, respectively. Hence we have $x = \gamma\omega\psi$. Similarly, the notations Γ, Ω and Ψ represent the ring-amplitude, phase and ring-amplitude-dependent phase rotation of s, respectively, where we have $s = \Gamma\Omega\Psi$. The modulation index $m = \text{bin2dec}(b_1 \cdots b_{\text{BPS}})$, ring-amplitude index $a = \text{bin2dec}(b_{\text{BPS}_P+1} \cdots b_{\text{BPS}})$ and phase index $p = \text{bin2dec}(b_1 \cdots b_{\text{BPS}_P})$ are Gray coded indices $\check{m}$, $\check{a}$ and $\check{p}$, respectively. Moreover, for Star QAM, the notations α and $\beta = \frac{\sum_{\mu=0}^{M_A-1} \alpha^{2\mu}}{M_A}$ respectively represent the ring ratio and constellation normalization factor. The advantageous choices in Rayleigh fading channels are $\alpha = 2.0$ for twin-ring Star QAM [59] [60] and $\alpha = 1.4$ for quadruple-ring Star QAM [33], respectively.

As seen in Table II, the ADPSK scheme [15], [16] invokes the absolute-amplitude differential encoding process as:

$$s_n = \frac{1}{|s_{n-1}|} x_{n-1} s_{n-1}, \tag{1}$$

which starts from $s_1 = \frac{1}{\sqrt{\beta}}$. More explicitly, the ADPSK's data-carrying symbols x_{n-1} in (1) are modulated as Star QAM symbols as seen in Table II, and then thanks to the normalization of

$\frac{1}{|s_{n-1}|}$ in (1), the transmitted symbols always have the absolute-amplitude of $\Gamma_n = |s_n| = |x_{n-1}| = \gamma_{n-1}$.

By contrast, the DAPSK scheme [18], [37] invokes the differential encoding process in the same way as the classic DPSK as:

$$s_n = x_{n-1} s_{n-1}, \tag{2}$$

which also starts from $s_1 = \frac{1}{\sqrt{\beta}}$. More explicitly, the DAPSK's transmitted symbols s_n in (2) are encoded to be Star QAM symbols as $s_n = \Gamma_n \Omega_n = \frac{\alpha^{\mu_n}}{\sqrt{\beta}} \exp(j \frac{2\pi}{M_P} q_n)$, where the transmitted symbol's ring-amplitude and phase indices are given by $[\mu_n = (\check{a} + \mu_{n-1}) \bmod M_A]$ and $[q_n = (\check{p} + q_{n-1}) \bmod M_P]$, respectively. As a result, the modulation of the DAPSK's data-carrying symbol x_{n-1} in (2) is determined both by the data-carrying modulation index m as well as by the previous transmitted ring-amplitude $\Gamma_{n-1} = \frac{\alpha^{\mu_{n-1}}}{\sqrt{\beta}}$, as seen in Table II.

Based on the classic ADPSK and DAPSK, four DQAM variants in literature are also considered in Table II. More explicitly, the TADPSK and TDAPSK schemes impose a ring-amplitude-dependent phase rotation to the ADPSK and DAPSK schemes, respectively [16], [17]. Moreover, the TADPSK$^{\text{JM}}$ and TDAPSK$^{\text{JM}}$ schemes jointly modulate the ring-amplitude and phase for the TADPSK and TDAPSK schemes, respectively, as suggested in [16], [17]. We will demonstrate later that a higher correlation between the ring-amplitude and phase may improve the iteration gain on the EXIT charts, which may result in a performance advantage in specific coded systems.

III. Multiple-Symbol Differential Detection

First of all, the signal received by the N_R antennas may be modelled as $\mathbf{Y}_n = s_n \mathbf{H}_n + \mathbf{V}_n$, where the N_R-element row-vectors $\mathbf{Y}_n$, $\mathbf{H}_n$ and $\mathbf{V}_n$ model the received signal, the Rayleigh fading and AWGN, respectively. Then the N_w observations may be modelled by the MSDD as:

$$\mathbf{Y} = \mathbf{SH} + \mathbf{V} = \mathbf{APOH} + \mathbf{V}, \tag{3}$$

where $\mathbf{Y} = [\mathbf{Y}_{N_w}^T, \mathbf{Y}_{N_w-1}^T, \ldots, \mathbf{Y}_1^T]^T$, $\mathbf{H} = [\mathbf{H}_{N_w}^T, \mathbf{H}_{N_w-1}^T, \ldots, \mathbf{H}_1^T]^T$ and $\mathbf{V} = [\mathbf{V}_{N_w}^T, \mathbf{V}_{N_w-1}^T, \ldots, \mathbf{V}_1^T]^T$ are of size $(N_w \times N_R)$. Moreover, $\mathbf{S} = \text{diag}\{[s_{N_w}, s_{N_w-1}, \ldots, s_1]\}$, $\mathbf{A} = \text{diag}\{[\Gamma_{N_w}, \Gamma_{N_w-1}, \ldots, \Gamma_1]\}$, $\mathbf{P} = \text{diag}\{[\Omega_{N_w}, \Omega_{N_w-1}, \ldots, \Omega_1]\}$ and $\mathbf{O} = \text{diag}\{[\Psi_{N_w}, \Psi_{N_w-1}, \ldots, \Psi_1]\}$ in (3) are all of size $(N_w \times N_w)$. We note that $\mathbf{O}$ is an identity matrix for the ADPSK and DAPSK schemes. The MSDD aims for detecting the $(N_w - 1)$ data-carrying symbols $\{x_v\}_{v=1}^{N_w-1}$, rather than the N_w transmitted symbols $\{s_v\}_{v=1}^{N_w}$. Therefore, the reference symbol $s_1 = \Gamma_1 \Omega_1 \Psi_1$ should be separated from the transmitted symbols seen in (3) following it, which leads to:

$$\mathbf{Y} = \bar{\mathbf{S}}\bar{\mathbf{H}} + \mathbf{V} = \bar{\mathbf{A}}\bar{\mathbf{P}}\bar{\mathbf{O}}\bar{\mathbf{H}} + \mathbf{V}, \tag{4}$$

where the v^{th} diagonal element in $\bar{\mathbf{P}}$ is given by $\bar{\Omega}_v = \Omega_v \Omega_1^*$, which leads to $\bar{\Omega}_1 = 1$ and $\bar{\Omega}_v = \omega_{v-1}\bar{\Omega}_{v-1} = \prod_{t=1}^{v-1} \omega_t$ for $v > 1$. Similarly, the v^{th} diagonal element in $\bar{\mathbf{O}}$ is given by $\bar{\Psi}_v = \Psi_v \Psi_1^*$, which leads to $\bar{\Psi}_1 = 1$ and $\bar{\Psi}_v = \psi_{v-1}\bar{\Psi}_{v-1} = \prod_{t=1}^{v-1} \psi_t$ for $v > 1$. As a result, the v^{th} row in $\bar{\mathbf{H}}$ is given by $\bar{\mathbf{H}}_v = \Omega_1 \Psi_1 \mathbf{H}_v$, where the constant phase $\Omega_1 \Psi_1$ does not change the correlations between fading samples, i.e. $\text{E}\{\mathbf{H}_v^H \mathbf{H}_{v'}\} = \text{E}\{\bar{\mathbf{H}}_v^H \bar{\mathbf{H}}_{v'}\}$. However, unlike Ω_1 and Ψ_1, the value of Γ_1 does affect the MSDD decision. Therefore, when $\bar{\mathbf{A}}$ in (3) is detected by the MSDD, Γ_1 is considered as a known term, which is either obtained based on previous MSDD decisions or detected separately as an unknown variable. As a result, there are $M_A^{(N_w-1)}$ combinations for $\bar{\mathbf{A}}$ in (3). Specifically, for the absolute-amplitude ADPSK/TADPSK/TADPSK$^{\text{JM}}$ using (1), the v^{th} diagonal element in $\bar{\mathbf{A}}$ is given by $\Gamma_v = \gamma_{v-1}$. By contrast, for the differential-amplitude DAPSK/TDAPSK/TDAPSK$^{\text{JM}}$ using (2), we have $\Gamma_v = \gamma_{v-1}\Gamma_{v-1} = \left(\prod_{t=1}^{v-1} \gamma_t\right) \Gamma_1$.

Based on (4), the MSDD may invoke the optimum Log-MAP [14], [61] as:

$$\begin{aligned} L_p(b_k) &= \ln \frac{\sum_{\forall \Gamma_1 \in \left\{\frac{\alpha^{\mu_1}}{\sqrt{\beta}}\right\}_{\mu_1=0}^{M_A-1}} \sum_{\forall \bar{\mathbf{S}} \in \bar{\mathbf{S}}_{b_k=1}} p(\mathbf{Y}|\Gamma_1, \bar{\mathbf{S}}) p(\Gamma_1) p(\bar{\mathbf{S}})}{\sum_{\forall \Gamma_1 \in \left\{\frac{\alpha^{\mu_1}}{\sqrt{\beta}}\right\}_{\mu_1=0}^{M_A-1}} \sum_{\forall \bar{\mathbf{S}} \in \bar{\mathbf{S}}_{b_k=0}} p(\mathbf{Y}|\Gamma_1, \bar{\mathbf{S}}) p(\Gamma_1) p(\bar{\mathbf{S}})} \\ &= L_a(b_k) + L_e(b_k), \end{aligned} \tag{5}$$

where $L_p(b_k)$, $L_e(b_k)$ and $L_a(b_k)$ represent the *a posteriori* LLR and the extrinsic LLR produced by the MSDD as well as the *a priori* LLR gleaned from a channel decoder, respectively, while the subsets $\bar{\mathbf{S}}_{b_k=1}$ and $\bar{\mathbf{S}}_{b_k=0}$ refer to the MSDD combination sets associated with $\bar{\mathbf{S}} = \bar{\mathbf{A}}\bar{\mathbf{P}}\bar{\mathbf{O}}$ of (4), with the specific bit b_k being fixed to 1 and 0, respectively. Furthermore, the probability of receiving $\mathbf{Y}$ given Γ_1 and $\bar{\mathbf{S}}$ is formulated as [21], [22], [28]:

$$p(\mathbf{Y}|\Gamma_1, \bar{\mathbf{S}}) = \frac{\exp\left\{-\text{rvec}(\mathbf{Y}) \mathbf{R}_{YY}^{-1} [\text{rvec}(\mathbf{Y})]^H\right\}}{\pi^{N_R N_w} \det(\mathbf{R}_{YY})}, \tag{6}$$

where the operation rvec($\mathbf{Y}$) forms a $N_R N_w$-element row-vector by taking the rows of $\mathbf{Y}$ one-by-one. As a result, the correlation matrix seen in (6) is given by $\mathbf{R}_{YY} = \text{E}\{[\text{rvec}(\mathbf{Y})]^H \text{rvec}(\mathbf{Y})\} = (\bar{\mathbf{O}}^H \bar{\mathbf{P}}^H \mathbf{C} \bar{\mathbf{P}} \bar{\mathbf{O}}) \otimes \mathbf{I}_{N_R}$, where both $\bar{\mathbf{P}}$ and $\bar{\mathbf{O}}$ are unitary matrices, while the operation $\otimes$ represents the Kronecker product. Moreover, the $(N_w \times N_w)$-element channel correlation matrix $\mathbf{C}$ is given by $\mathbf{C} = \bar{\mathbf{A}}^H \mathbf{R}_{hh} \bar{\mathbf{A}} + \mathbf{R}_{vv}$, where the fading correlation matrix $\mathbf{R}_{hh} = \text{Toeplitz}([\, \rho_0 \;\; \rho_1 \;\; \cdots \;\; \rho_{N_w-1} \,])$ and the AWGN correlation matrix $\mathbf{R}_{vv} = N_0 \mathbf{I}_{\mathbf{N_w}}$ are the same as in the case of DPSK using $N_R = 1$ in [28], [29]. However, since $\bar{\mathbf{A}}$ is not a unitary matrix, it cannot be separated from $\mathbf{C}$ for DQAM detection.

The low-complexity Max-Log-MAP algorithm may be invoked by the MSDD for the sake of simplifying the Log-MAP of (5), and it is given by [61]:

$$\begin{aligned} L_p(b_k) = &\max_{\forall \Gamma_1 \in \left\{\frac{\alpha^{\mu_1}}{\sqrt{\beta}}\right\}_{\mu_1=0}^{M_A-1}} \quad \max_{\forall \bar{\mathbf{S}} \in \bar{\mathbf{S}}_{b_k=1}} d(\Gamma_1, \bar{\mathbf{S}}) \\ &- \max_{\forall \Gamma_1 \in \left\{\frac{\alpha^{\mu_1}}{\sqrt{\beta}}\right\}_{\mu_1=0}^{M_A-1}} \quad \max_{\forall \bar{\mathbf{S}} \in \bar{\mathbf{S}}_{b_k=0}} d(\Gamma_1, \bar{\mathbf{S}}). \end{aligned} \tag{7}$$

Based on (6), the probability metric seen in (7) may be expressed as:

$$\begin{aligned} d(\Gamma_1, \bar{\mathbf{S}}) = &-\text{tr}\left(\mathbf{Y}^H \bar{\mathbf{P}} \bar{\mathbf{O}} \mathbf{C}^{-1} \bar{\mathbf{O}}^H \bar{\mathbf{P}}^H \mathbf{Y}\right) - N_R \ln[\det(\mathbf{C})] \\ &+ \sum_{\bar{k}=1}^{(N_w-1)\text{BPS}} \widetilde{b}_{\bar{k}} L_a(b_{\bar{k}}), \end{aligned} \tag{8}$$

where the determinant in (6) is given by $\det(\mathbf{R}_{YY}) = \det(\mathbf{C})^{N_R}$, while $\{\widetilde{b}_{\bar{k}}\}_{\bar{k}=1}^{(N_w-1)\mathrm{BPS}}$ denotes the bit mapping corresponding to the MSDD candidate $\bar{\mathbf{S}}$.

Moreover, instead of assuming that Γ_1 is an equiprobable variable in (7), soft-decision feedback on $\{p(\Gamma_1)\}_{\forall\Gamma_1}$ may be obtained from the previous MSDD window, which leads us to a SDD-MSDD. In more detail, $d(\Gamma_1, \bar{\mathbf{S}})$ of (8) may be modified to take $\{p(\Gamma_1)\}_{\forall\Gamma_1}$ into account as:

$$d(\Gamma_1, \bar{\mathbf{S}}) = -\mathrm{tr}\left(\mathbf{Y}^H\bar{\mathbf{P}}\bar{\mathbf{O}}\mathbf{C}^{-1}\bar{\mathbf{O}}^H\bar{\mathbf{P}}^H\mathbf{Y}\right) - N_R\ln[\det(\mathbf{C})] + \sum\nolimits_{\bar{k}=1}^{(N_w-1)\mathrm{BPS}}\widetilde{b}_{\bar{k}}L_a(b_{\bar{k}}) + \ln[p(\Gamma_1)]. \quad (9)$$

The probabilities may be updated according to $\{\ln\left[p(\Gamma_{N_w})\right] = \max_{\forall\bar{\mathbf{A}}\in\langle\forall\Gamma_1,\Gamma_{N_w}\rangle} d(\Gamma_1, \bar{\mathbf{S}})\}$, where the ring-amplitudes Γ_1 and Γ_{N_w} transmitted both at the start and end of the MSDD window may represent trellis states $\langle\Gamma_1, \Gamma_{N_w}\rangle$, while the data-carrying ring-amplitudes $\bar{\mathbf{A}}$ govern the state transition. As a result, all initial states $\forall\Gamma_1$ and all transitions $\forall\bar{\mathbf{A}}$ that lead to the specific trellis termination state of Γ_{N_w} have to be taken into account for evaluating $\ln\left[p(\Gamma_{N_w})\right]$. Then the newly updated $\{\ln p(\Gamma_{N_w})\}_{\forall\Gamma_{N_w}}$ may be passed on to the next MSDD window as $\{\ln p(\Gamma_1)\}_{\forall\Gamma_1}$.

If the hard-decision on $\hat{\Gamma}_1$ is fed back from the previous MSDD window, then a HDD-MSDD invoking the Max-Log-MAP may be simply formulated as:

$$L_p(b_k) = \max_{\forall\bar{\mathbf{S}}\in\bar{\mathbf{S}}_{b_k=1}} d(\hat{\Gamma}_1, \bar{\mathbf{S}}) - \max_{\forall\bar{\mathbf{S}}\in\bar{\mathbf{S}}_{b_k=0}} d(\hat{\Gamma}_1, \bar{\mathbf{S}}), \quad (10)$$

where the hard-decision concerning $\hat{\Gamma}_{N_w}$ may be passed on to the next MSDD window as $\hat{\Gamma}_1$.

IV. Multiple-Symbol Differential Sphere Detection

A. Generic MSDSD Algorithm

Similar to MSDSD aided DPSK of [29], [30], SD may be invoked for MSDD aided DQAM using the Max-Log-MAP of (7), where the maximization is converted to minimization as:

$$L_p(b_k) = \min_{\forall\Gamma_1\in\left\{\frac{\alpha^{\mu_1}}{\sqrt{\beta}}\right\}_{\mu_1=0}^{M_A-1}} \min_{\forall\bar{\mathbf{S}}\in\bar{\mathbf{S}}_{b_k=1}} d - \min_{\forall\Gamma_1\in\left\{\frac{\alpha^{\mu_1}}{\sqrt{\beta}}\right\}_{\mu_1=0}^{M_A-1}} \min_{\forall\bar{\mathbf{S}}\in\bar{\mathbf{S}}_{b_k=0}} d, \quad (11)$$

while the probability metrics should be guaranteed to have positive values as:

$$d = \sum\nolimits_{v=1}^{N_w}\left\|\sum\nolimits_{t=1}^{v}\widetilde{l}_{v-t+1,1}\bar{\Psi}_t^*\bar{\Omega}_t^*\mathbf{Y}_t\right\|^2 + N_R\ln\left(\Gamma_1^2\rho_0 + N_0\right) + N_R\sum\nolimits_{v=2}^{N_w}\ln\left[\left(\Gamma_v^2\rho_0 + N_0\right) - \widetilde{\mathbf{e}}_v^T\widetilde{\mathbf{C}}_{v-1}^{-1}\widetilde{\mathbf{e}}_v\right] - \sum\nolimits_{v=2}^{N_w}\xi_{v-1} - \sum\nolimits_{v=2}^{N_w}\sum\nolimits_{\bar{k}_v=1}^{\mathrm{BPS}}\left[\widetilde{b}_{\bar{k}_v}L_a\left(b_{\bar{k}_v}\right) - \overline{C}_{a,\bar{k}_v}\right]. \quad (12)$$

In more detail, firstly, the trace term in (8) may be evaluated by $\mathrm{tr}(\mathbf{Y}^H\bar{\mathbf{P}}\bar{\mathbf{O}}\mathbf{C}^{-1}\bar{\mathbf{O}}^H\bar{\mathbf{P}}^H\mathbf{Y}) = \|\mathbf{L}^T\bar{\mathbf{O}}^H\bar{\mathbf{P}}^H\mathbf{Y}\|^2 = \sum_{v=1}^{N_w}\left\|\sum_{t=1}^{v} l_{N_w-t+1,N_w-v+1}\bar{\Psi}_t^*\bar{\Omega}_t^*\mathbf{Y}_t\right\|^2$, where $\{\{l_{N_w-t+1,N_w-v+1}\}_{t=1}^{v}\}_{v=1}^{N_w}$ are elements in the lower triangular matrix $\mathbf{L}$, which is obtained from the decomposition of $\mathbf{C}^{-1} = \mathbf{L}\mathbf{L}^T$. We note that both $\mathbf{C} = \bar{\mathbf{A}}^H\mathbf{R}_{hh}\bar{\mathbf{A}} + \mathbf{R}_{vv}$ and $\mathbf{L}$ remain unknown, until all ring-amplitudes in $\bar{\mathbf{A}}$ are detected. In order to solve this problem by invoking SD, we define the $(v\times v)$-element partial channel correlation matrix $\widetilde{\mathbf{C}}_v$, which may be evaluated with the aid of the SD's previous decisions concerning $\{\Gamma_t\}_{t=1}^{v-1}$ and a single variable Γ_v as:

$$\widetilde{\mathbf{C}}_v = \begin{bmatrix} \Gamma_v^2\rho_0 + N_0 & \Gamma_v\Gamma_{v-1}\rho_1 & \cdots & \Gamma_v\Gamma_1\rho_{v-1} \\ \Gamma_{v-1}\Gamma_v\rho_1 & \Gamma_{v-1}^2\rho_0 + N_0 & \cdots & \Gamma_{v-1}\Gamma_1\rho_{v-2} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_1\Gamma_v\rho_{v-1} & \Gamma_1\Gamma_{v-1}\rho_{v-2} & \cdots & \Gamma_1^2\rho_0 + N_0 \end{bmatrix} = \begin{bmatrix} \Gamma_v^2\rho_0 + N_0 & \widetilde{\mathbf{e}}_v^T \\ \widetilde{\mathbf{e}}_v & \widetilde{\mathbf{C}}_{v-1} \end{bmatrix}, \quad (13)$$

while the $(v-1)$-element column-vector $\widetilde{\mathbf{e}}_v$ in (13) is given by $\widetilde{\mathbf{e}}_v = [\Gamma_v\Gamma_{v-1}\rho_1, \ldots, \Gamma_v\Gamma_1\rho_{v-1}]^T$. It can be readily seen that $\widetilde{\mathbf{C}}_v$ is a submatrix of $\mathbf{C}$, but they become equal, when the SD index is increased to $v = N_w$. As a result, we also have the relationship of $\{l_{N_w-t+1,N_w-v+1} = \widetilde{l}_{v-t+1,1}\}_{t=1}^{v}$, where $\{\widetilde{l}_{v-t+1,1}\}_{t=1}^{v}$ are elements in the $(v\times v)$-element lower triangular matrix $\widetilde{\mathbf{L}}_v$, which is a submatrix of $\mathbf{L}$, and $\widetilde{\mathbf{L}}_v$ may be directly obtained from the submatrix decomposition of $\widetilde{\mathbf{L}}_v\widetilde{\mathbf{L}}_v^T = \widetilde{\mathbf{C}}_v^{-1}$. Secondly, according to the Leibniz formula [62], the determinant of $\widetilde{\mathbf{C}}_v$ in (13) may be evaluated by $\det(\widetilde{\mathbf{C}}_v) = \det(\widetilde{\mathbf{C}}_{v-1})[(\Gamma_v^2\rho_0 + N_0) - \widetilde{\mathbf{e}}_v^T\widetilde{\mathbf{C}}_{v-1}^{-1}\widetilde{\mathbf{e}}_v]$. Therefore, the complete determinant term $N_R\ln[\det(\mathbf{C})]$ in (8) may be evaluated by $N_R\ln[\det(\widetilde{\mathbf{C}}_1)] + N_R\sum_{v=2}^{N_w}\ln[\left(\Gamma_v^2\rho_0 + N_0\right) - \widetilde{\mathbf{e}}_v^T\widetilde{\mathbf{C}}_{v-1}^{-1}\widetilde{\mathbf{e}}_v]$, where the initial term is given by $N_R\ln[\det(\widetilde{\mathbf{C}}_1)] = N_R\ln\left(\Gamma_1^2\rho_0 + N_0\right)$, as seen in (12). Thirdly, the constant $\xi_{v-1} = \min_{\forall\Gamma_1\cdots\forall\Gamma_v} N_R\ln[(\Gamma_v^2\rho_0 + N_0) - \widetilde{\mathbf{e}}_v^T\widetilde{\mathbf{C}}_{v-1}^{-1}\widetilde{\mathbf{e}}_v]$ and the constant $\overline{C}_{a,\bar{k}_v} = \frac{1}{2}\left[|L_a(b_{\bar{k}_v})| + L_a(b_{\bar{k}_v})\right]$ are artificially added in order to maintain a non-negative Euclidean Distance (ED) in (13), as discussed in the context of (17) in [30]. We note that adding the constants of $\sum_{v=2}^{N_w}(\xi_{v-1} - \overline{C}_{a,\bar{k}_v})$ to the MSDD metric of (8) does not impose any performance difference, and the constants $\{\xi_{v-1}\}_{v=2}^{N_w}$ may be obtained by brute-force search in an off-line fashion before performing MSDSD.

As a result, based on (12), the SD's Partial ED (PED) may be defined as:

$$\begin{aligned} d_v &= \left\|l_{N_w,N_w}\mathbf{Y}_1\right\|^2 + N_R\ln\left(\Gamma_1^2\rho_0 + N_0\right) \\ &\quad + \sum\nolimits_{\bar{v}=2}^{v}\left\|\sum\nolimits_{t=1}^{\bar{v}}\widetilde{l}_{\bar{v}-t+1,1}\bar{\Psi}_t^*\bar{\Omega}_t^*\mathbf{Y}_t\right\|^2 + \sum\nolimits_{\bar{v}=2}^{v}\Xi_{\bar{v}} \\ &\quad - \sum\nolimits_{\bar{v}=2}^{v}\sum\nolimits_{\bar{k}_{\bar{v}}=1}^{\mathrm{BPS}}\left[\widetilde{b}_{\bar{k}_{\bar{v}}}L_a(b_{\bar{k}_{\bar{v}}}) - \overline{C}_{a,\bar{k}_{\bar{v}}}\right] \\ &= d_{v-1} + \Delta_{v-1}. \end{aligned} \quad (14)$$

where the PED increment is given by:

$$\begin{aligned} &\Delta_{v-1} \\ &= \left\|\widetilde{l}_{1,1}\bar{\Psi}_{v-1}^*\bar{\Omega}_{v-1}^*\mathbf{Y}_v + \omega_{v-1}\psi_{v-1}\left(\sum\nolimits_{t=1}^{v-1}\widetilde{l}_{v-t+1,1}\bar{\Psi}_t^*\bar{\Omega}_t^*\mathbf{Y}_t\right)\right\|^2 \\ &\quad + \Xi_v - \sum\nolimits_{\bar{k}_v=1}^{\mathrm{BPS}}\left[\widetilde{b}_{\bar{k}_v}L_a(b_{\bar{k}_v}) - \overline{C}_{a,\bar{k}_v}\right], \end{aligned} \quad (15)$$

TABLE III
PSEUDOCODE FOR THE SCHNORR–EUCHNER SEARCH STRATEGY TAILORED FOR SOFT-DECISION MSDSD AIDED DQAM

Subfunction: $[\{\Delta^{\bar{m}}_{v-1}\}_{\bar{m}=0}^{M-1}, \{x^{\bar{m}}_{v-1}\}_{\bar{m}=0}^{M-1}, n_{v-1}] = \textbf{sortDelta}(\{\mathbf{Y}_t\}_{t=1}^{v}, \{\hat{\Gamma}_t\}_{t=1}^{v-1}, \{\hat{\Psi}_t\}_{t=1}^{v-1}, \{\hat{\Omega}_t\}_{t=1}^{v-1}, \{L_a(b_{\bar{k}_v})\}_{\bar{k}_v=1}^{\mathrm{BPS}}, \{\overline{C}_{a,\bar{k}_v}\}_{\bar{k}_v=1}^{\mathrm{BPS}})$

Requirements: $\{\widetilde{l}_{v-t+1,1}\}_{t=1}^{v}$ in (15) are taken from $\widetilde{\mathbf{L}}(\{\hat{\Gamma}_t\}_{t=1}^{v-1}, \Gamma_v) = \widetilde{\mathbf{L}}_{\mathbf{v}}$, which are pre-evaluated and pre-stored. $\Xi(\{\hat{\Gamma}_t\}_{t=1}^{v-1}, \Gamma_v) = \Xi_v$ in (15) are also pre-evaluated and pre-stored.

1:	**for** m=0 **to** $M-1$	//visit all M child nodes.
2:	(ADPSK/TADPSK/TADPSK$^{\mathrm{JM}}$:) $\Gamma_v = \gamma_{v-1}$	//visit $x^m_{v-1} = \gamma_{v-1}\omega_{v-1}\psi_{v-1}$.
	(DAPSK/TDAPSK/TDAPSK$^{\mathrm{JM}}$:) $\Gamma_v = \gamma_{v-1}\hat{\Gamma}_{v-1}$	
3:	$\Delta^m_{v-1} = \|\widetilde{l}_{1,1}\hat{\Psi}^*_{v-1}\hat{\Omega}^*_{v-1}\mathbf{Y}_v + \omega_{v-1}\psi_{v-1}(\sum_{t=1}^{v-1}\widetilde{l}_{v-t+1,1}\hat{\Psi}^*_t\hat{\Omega}^*_t\mathbf{Y}_t)\|^2 + \Xi(\{\hat{\Gamma}_t\}_{t=1}^{v-1}, \Gamma_v) - \sum_{\bar{k}_v=1}^{\mathrm{BPS}}[\widetilde{b}_{\bar{k}_v}L_a(b_{\bar{k}_v}) - \overline{C}_{a,\bar{k}_v}]$	//evaluate PED increments of (15).
4:	**end for**	
5:	$[\{\Delta^{\bar{m}}_{v-1}\}_{\bar{m}=0}^{M-1}, \{x^{\bar{m}}_{v-1}\}_{\bar{m}=0}^{M-1}] = \textbf{sort}(\{\Delta^m_{v-1}\}_{m=0}^{M-1})$	//rank PED increments in increasing order.
6:	$n_{v-1} = 0$	//initialize child node counter.

while the determinant term is defined as $\Xi_v = N_R \ln[(\Gamma_v^2\rho_0 + N_0) - \widetilde{\mathbf{e}}_v^T\widetilde{\mathbf{C}}_{v-1}^{-1}\widetilde{\mathbf{e}}_v] - \xi_{v-1}$.

Based on the PED of (14), the MSDSD algorithm of [29] may be invoked, but its "sortDelta" subfunction should be revised as summarized in Table III, where the subscript $m \in \{0, \ldots, M-1\}$ represents the data-carrying constellation point index which may be directly translated back to binary source bits as $[b_1 \cdots b_{\mathrm{BPS}}] = \mathrm{dec2bin}(m)$. Furthermore, the subscript $\overline{m} \in \{\overline{0}, \ldots, \overline{M-1}\}$ represents the constellation point index ordered according to the increasing values of PED increment Δ_{v-1}. We note that the MSDD model of (3) stores received signal vectors in a reverse order compared to the one seen in [29] in order to detect the phases according to $\Omega_v = \omega_{v-1}\Omega_{v-1}$, instead of detecting them backwards as $\Omega_{v-1} = \omega^*_{v-1}\Omega_v$. As a result, the MSDSD algorithm of [29] may now commence from index $v = 2$ and update the sphere radius at index $v = N_w$. The child node counter n_{v-1} in Table III has the revised range of $0 \le n_{v-1} \le (M-1)$ throughout the SD search, which accords with the range of the constellation point index m. Moreover, similar to the pseudo-code presented in [29], the MSDSD may initialize the PED as $d_1 = 0$ for the sake of simplicity, but the Γ_1-related term $d_1 = \left\|l_{N_w,N_w}\mathbf{Y}_1\right\|^2 + N_R\ln\left(\Gamma_1^2\rho_0 + N_0\right)$ in (14) should be added to the SD's output radius before comparing EDs over Γ_1 as seen in (11). It is also worth noting that a total of $\sum_{v=1}^{N_w} M_A^v$ candidates for $\{\widetilde{\mathbf{L}}_v\}_{v=2}^{N_w}$ and $\{\Xi_v\}_{v=2}^{N_w}$ seen in (15) may be pre-evaluated and pre-stored in an off-line fashion, before performing MSDSD. The memory required for storing $\{\widetilde{\mathbf{L}}_v\}_{v=2}^{N_w}$ and $\{\Xi_v\}_{v=2}^{N_w}$ is small compared to that of MSDD.

Therefore, the optimum ED is given by $d_{MAP} = \min_{\forall\Gamma_1}\left(\min_{\forall\bar{\mathbf{S}}} d\right)$, where $\left(\min_{\forall\bar{\mathbf{S}}} d\right)$ is found by the SD without the MSDD full search, while the corresponding hard-bit decisions $\{b_k^{MAP}\}_{k=1}^{(N_w-1)BPS}$ may also be obtained along with d_{MAP}. Following this, the Max-Log-MAP of (11) may now be completed by:

$$L_p(b_k) = \begin{cases} -d_{MAP} + \bar{d}_{MAP}, & \text{if } \mathrm{b}_k^{\mathrm{MAP}} = 1, \\ -\bar{d}_{MAP} + d_{MAP}, & \text{if } \mathrm{b}_k^{\mathrm{MAP}} = 0, \end{cases} \tag{16}$$

where $\bar{d}_{MAP}$ is obtained by invoking the SD again for evaluating each $L_p(b_k)$, where the specific bit b_k is fixed to be the toggled MAP decision $\bar{b}_k^{MAP}$, i.e. we have $\bar{d}_{MAP} = \min_{\forall\Gamma_1}(\min_{\forall\{\bar{\mathbf{S}}\}_{b_k=\bar{b}_k^{MAP}}} d)$. As a result, the SD is invoked a total of $M_A[1 + (N_w - 1)\mathrm{BPS}]$ times for producing $\{L_p(b_k)\}_{k=1}^{(N_w-1)\mathrm{BPS}}$ of (16). As suggested in [30], the repeated SD calculations may be avoided by labelling the previously visited nodes and storing their PED metrics.

Moreover, both HDD-MSDD and SDD-MSDD of Sec. III may be implemented as HDD-MSDSD and SDD-MSDSD, respectively. For HDD-MSDSD, $\hat{\Gamma}_1$ is known from previous MSDSD window. Therefore, HDD-MSDSD may also produce $L_p(b_k)$ in the same way as MSDSD of (16), except that the comparisons over the different Γ_1 values for obtaining d_{MAP} and $\bar{d}_{MAP}$ in (16) may be eliminated, and there is no need to evaluate the Γ_1-related term $d_1 = \left\|l_{N_w,N_w}\mathbf{Y}_1\right\|^2 + \ln\left(\Gamma_1^2\rho_0 + N_0\right)$. By contrast, for SDD-MSDSD, the Γ_1-related term has to take $p(\Gamma_1)$ into account as $d_1 = \left\|l_{N_w,N_w}\mathbf{Y}_1\right\|^2 + \ln\left(\Gamma_1^2\rho_0 + N_0\right) + \ln[p(\Gamma_1)]$. In order to obtain $\{\ln[p(\Gamma_1)]\}_{\forall\Gamma_1}$, the SD's visited EDs d may be utilized for evaluating $\ln\left[p(\Gamma_{N_w})\right] = \max_{\forall\bar{\mathbf{A}}\in\langle\forall\Gamma_1,\Gamma_{N_w}\rangle} d(\Gamma_1, \bar{\mathbf{S}})$, where we have $\{d(\Gamma_1, \bar{\mathbf{S}}) = -d\}_{\forall\Gamma_1,\forall\bar{\mathbf{A}}}$. Following this, $\{\ln p(\Gamma_{N_w})\}_{\forall\Gamma_{N_w}}$ may become $\{\ln p(\Gamma_1)\}_{\forall\Gamma_1}$ for the next MSDSD window. Apart from the extra term of $\{\ln[p(\Gamma_1)]\}_{\forall\Gamma_1}$, SDD-MSDSD may proceed in the same way as MSDSD of (16).

B. RC MSDSD Algorithm

It was proposed in [30] that by exploring the constellation symmetry provided by the M_PPSK scheme's Gray-coded labelling, the number of constellation points visited by the SD's Schnorr-Euchner search strategy may be significantly reduced for the soft-decision MSDSD aided DPSK. In this section, we further conceive RC design for a range of DQAM constellations, including DAPSK, TDAPSK, ADPSK and TADPSK, which modulate the ring-amplitude and phase separately, so that their M_PPSK phase detection may be performed in the same way as the RC DPSK detection of [30]. First of all, let us rewrite the PED increment of (15) as:

$$\Delta_{v-1} = \widetilde{\Delta}_{v-1} + \overline{\Delta}_{v-1}, \tag{17}$$

where the ring-amplitude-related term $\widetilde{\Delta}_{v-1}$ derived from (15) is given by:

$$\widetilde{\Delta}_{v-1} = \left\|\widetilde{l}_{1,1}\mathbf{Y}_v\right\|^2 + \left\|\sum\nolimits_{t=1}^{v-1}\widetilde{l}_{v-t+1,1}\bar{\Psi}^*_t\bar{\Omega}^*_t\mathbf{Y}_t\right\|^2 + \Xi_v + \overline{C}_{a,v-1} - \frac{L_a(b_1) + L_a(b_2)}{2}, \tag{18}$$

while we have the constant $\overline{C}_{a,v-1} = \sum_{\bar{k}_v=1}^{\text{BPS}} \overline{C}_{a,\bar{k}_v}$. We note that $\widetilde{\Delta}_{v-1}$ of (18) is invariant over the different phase candidates for ω_{v-1} in (15). Furthermore, the M_PPSK-related term $\overline{\Delta}_{v-1}$ seen in (17) is given by:

$$\begin{aligned}\overline{\Delta}_{v-1} &= -2\Re\left[(\omega'_{v-1})^* z'_{v-1}\right] - \sum_{\bar{k}_v=1}^{\text{BPS}} \widetilde{b}_{\bar{k}_v} L_a(b_{\bar{k}_v}) \\ &\quad + \frac{L_a(b_1)+L_a(b_2)}{2} \\ &= -2\Re(\omega'_{v-1})\Re(z'_{v-1}) - 2\Im(\omega'_{v-1})\Im(z'_{v-1}) \\ &\quad - \sum_{\bar{k}_v=1}^{\text{BPS}} \widetilde{b}_{\bar{k}_v} L_a(b_{\bar{k}_v}) + \frac{L_a(b_1)+L_a(b_2)}{2}, \end{aligned} \tag{19}$$

where $\Re(\cdot)$ and $\Im(\cdot)$ take the real part and the imaginary part of a complex number. In (19), we deliberately rotate all the detected M_PPSK constellations (except for the special case of $M_P = 2$) anti-clockwise by (π/M_P) as $\omega'_{v-1} = \omega_{v-1}\exp\left(j\frac{\pi}{M_P}\right)$, so that there are exactly $(M_P/4)$ PSK phases for ω'_{v-1} in each quadrant [13], [30]. Accordingly, the decision variable z'_{v-1} used for detecting ω'_{v-1} in (19) is given by:

$$z'_{v-1} = \mathbf{A}_{v-1}^{\text{SD}} \left(\mathbf{B}_{v-1}^{\text{SD}}\right)^H \exp\left(j\frac{\pi}{M_P}\right), \tag{20}$$

where we have $\mathbf{A}_{v-1}^{\text{SD}} = \widetilde{l}_{1,1}\psi^*_{v-1}\bar{\Psi}^*_{v-1}\bar{\Omega}^*_{v-1}\mathbf{Y}_v$ and $\mathbf{B}_{v-1}^{\text{SD}} = -\sum_{t=1}^{v-1}\widetilde{l}_{v-t+1,1}\bar{\Psi}^*_t\bar{\Omega}^*_t\mathbf{Y}_t$.

For a generic DQAM scheme, we may consider the M constellation points as $M/4$ groups of symmetric QPSK-like constellation points that are associated with the same magnitudes but different polarities. Considering the 16-ADPSK(2,8) scheme as an example, the $M = 16$ constellation points are assigned to $M/4 = 4$ groups, as seen in Fig. 1. As a result, the four candidates for $\overline{\Delta}_{v-1}$ of (19) corresponding to the four QPSK-like constellation points in each group may always be expressed in the following form:

$$\overline{\Delta}_{v-1} = \pm t_{v-1}^{Re_g} \pm t_{v-1}^{Im_g} - \sum_{\bar{k}_v=3}^{\text{BPS}} \widetilde{b}_{\bar{k}_v} L_a(b_{\bar{k}_v}), \tag{21}$$

where the real and imaginary parts of z'_{v-1} are associated with $L_a(b_2)$ and $L_a(b_1)$, respectively:

$$\begin{aligned} t_{v-1}^{Re_g} &= \overline{A}^{\overline{g}}\Re(z'_{v-1}) - \frac{L_a(b_2)}{2}, \\ t_{v-1}^{Im_g} &= \overline{B}^{\overline{g}}\Im(z'_{v-1}) - \frac{L_a(b_1)}{2}. \end{aligned} \tag{22}$$

The coordinates of the rotated M_PPSK constellation points for ω'_{v-1}, which are located in the first quadrant may be denoted by $\{(A^{\overline{g}}, B^{\overline{g}})\}_{\overline{g}=0}^{M_P/4-1}$, and we have $\overline{A}^{\overline{g}} = 2A^{\overline{g}}$ as well as $\overline{B}^{\overline{g}} = 2B^{\overline{g}}$ in (22). The relationship between the group index $g = \text{bin2dec}(\widetilde{b}_3 \cdots \widetilde{b}_{\text{BPS}})$, the PSK phase group index $\overline{g} = \text{bin2dec}(\widetilde{b}_3 \cdots \widetilde{b}_{\text{BPS}_P})$ and the ring-amplitude index $a = \text{bin2dec}(\widetilde{b}_{\text{BPS}_P+1} \cdots \widetilde{b}_{\text{BPS}})$ is given by $\{\{g = \overline{g}\cdot M_A + a\}_{\overline{g}=0}^{M_P/4-1}\}_{a=0}^{M_A-1}$.

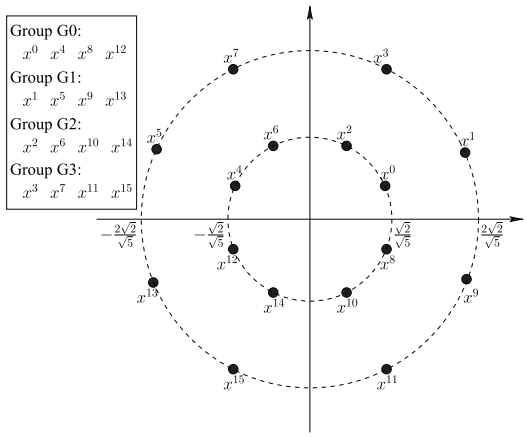


Fig. 1. Constellation diagram for detecting 16-ADPSK(2,8)'s data-carrying symbols, where the 8PSK phases are rotated anti-clockwise by $\pi/8$ for the sake of reduced-complexity detection design.

Considering the 16-ADPSK(2,8) example of Fig. 1, the four candidates for $\overline{\Delta}_{v-1}$ of (19) for Group G2 may be expressed as seen in (23), shown at the bottom of the page, which may all be expressed in the form of (21) as $\overline{\Delta}_{v-1}^{2,6,10,14} = \pm t_{v-1}^{Re_2} \pm t_{v-1}^{Im_2} - L_a(b_3)$, where we have $t_{v-1}^{Re_2} = 2\sin(\frac{\pi}{8})\Re(z'_{v-1}) - \frac{L_a(b_2)}{2}$ and $t_{v-1}^{Im_2} = 2\cos(\frac{\pi}{8})\Im(z'_{v-1}) - \frac{L_a(b_1)}{2}$.

It may be observed that the only difference between the four candidates of $\overline{\Delta}_{v-1}$ in (21) is the polarity of $t_{v-1}^{Re_g}$ and $t_{v-1}^{Im_g}$, This allows us to directly obtain the local minimum metric of Group Gg ($g \in \{0 \ldots, M/4-1\}$) by simply evaluating:

$$\overline{\Delta}_{v-1}^{g} = -\left|t_{v-1}^{Re_g}\right| - \left|t_{v-1}^{Im_g}\right| - \sum_{\bar{k}_v=3}^{\text{BPS}} \widetilde{b}_{\bar{k}_v} L_a(b_{\bar{k}_v}). \tag{24}$$

As a result, instead of evaluating and comparing four metrics according to the four constellation points in each group, only the constellation point in the first quadrant is visited by the one-step evaluation of (24). Moreover, after obtaining the $M/4$ local minima of each group according to (24), the global minimum PED increment candidate for Δ_{v-1} of (17) may be simply given by:

$$\Delta_{v-1} = \min_{\forall g\in\{0\ldots,M/4-1\}} \overline{\Delta}_{v-1}^{g} + \widetilde{\Delta}_{v-1}^{a}, \tag{25}$$

$$\begin{aligned}
\overline{\Delta}_{v-1}^{2} &= -2\sin\left(\frac{\pi}{8}\right)\Re\left(z'_{v-1}\right) - 2\cos\left(\frac{\pi}{8}\right)\Im(z'_{v-1}) - L_a(b_3) + \frac{L_a(b_1)+L_a(b_2)}{2} = -t_{v-1}^{Re_2} - t_{v-1}^{Im_2} - L_a(b_3), \\
\overline{\Delta}_{v-1}^{6} &= 2\sin\left(\frac{\pi}{8}\right)\Re(z'_{v-1}) - 2\cos\left(\frac{\pi}{8}\right)\Im(z'_{v-1}) - L_a(b_2) - L_a(b_3) + \frac{L_a(b_1)+L_a(b_2)}{2} = t_{v-1}^{Re_2} - t_{v-1}^{Im_2} - L_a(b_3), \\
\overline{\Delta}_{v-1}^{10} &= -2\sin\left(\frac{\pi}{8}\right)\Re(z'_{v-1}) + 2\cos\left(\frac{\pi}{8}\right)\Im(z'_{v-1}) - L_a(b_1) - L_a(b_3) + \frac{L_a(b_1)+L_a(b_2)}{2} = -t_{v-1}^{Re_2} + t_{v-1}^{Im_2} - L_a(b_3), \\
\overline{\Delta}_{v-1}^{14} &= 2\sin\left(\frac{\pi}{8}\right)\Re(z'_{v-1}) + 2\cos\left(\frac{\pi}{8}\right)\Im(z'_{v-1}) - L_a(b_1) - L_a(b_2) - L_a(b_3) + \frac{L_a(b_1)+L_a(b_2)}{2} = t_{v-1}^{Re_2} + t_{v-1}^{Im_2} - L_a(b_3),
\end{aligned} \tag{23}$$

TABLE IV
PSEUDO-CODE FOR THE SCHNORR–EUCHNER SEARCH STRATEGY TAILORED FOR SOFT-DECISION RC MSDSD AIDED DQAM (PART I)

Subfunction: $[\{|t_{v-1}^{Re_g}|\}_{g=0}^{M/4-1}, \{|t_{v-1}^{Im_g}|\}_{g=0}^{M/4-1}, \{\widetilde{\Delta}_{v-1}^{a}\}_{a=0}^{M_A-1}, \{\mathrm{CW}_{v-1}^{g}\}_{g=0}^{M/4-1}, \{\mathrm{CW}m_{v-1}^{g}\}_{g=0}^{M/4-1}, \{n_{v-1}^{g}\}_{g=0}^{M/4-1}, \Delta_{v-1}, m_{v-1}, n_{v-1}] = \mathbf{findBest}(\{\mathbf{Y}_t\}_{t=1}^{v}, \{\hat{\Gamma}_t\}_{t=1}^{v-1}, \{\hat{\Psi}_t\}_{t=1}^{v-1}, \{\hat{\Omega}_t\}_{t=1}^{v-1}, \{L_a(b_{k_v})\}_{k_v=1}^{\mathrm{BPS}}, \overline{C}_{a,v-1})$

Requirements: We define the *a priori* knowledge of group index as $\{P^g = \sum_{k_v=3}^{\mathrm{BPS}} \widetilde{b}_{k_v} L_a(b_{k_v})\}_{g=0}^{M/4-1}$.

1: **for** $a = 0$ **to** $M_A - 1$
2: (ADPSK/TADPSK/TADPSK$^{\mathrm{JM}}$:) $\Gamma_v = \gamma_{v-1}^{a}$ //fix the specific Γ_v for $\widetilde{\mathbf{L}}_{\mathbf{v}}$ and Ξ_v
(DAPSK/TDAPSK/TDAPSK$^{\mathrm{JM}}$:) $\Gamma_v = \gamma_{v-1}^{a}\hat{\Gamma}_{v-1}$
3: $\mathbf{A}_{v-1}^{\mathrm{SD}} = \widetilde{l}_{1,1}\psi_{v-1}^{*}\bar{\Psi}_{v-1}^{*}\bar{\Omega}_{v-1}^{*}\mathbf{Y}_v$ //evaluate $\mathbf{A}_{v-1}^{\mathrm{SD}}$ according to (20)
4: $\mathbf{B}_{v-1}^{\mathrm{SD}} = -\sum_{t=1}^{v-1}\widetilde{l}_{v-t+1,1}\bar{\Psi}_t^{*}\bar{\Omega}_t^{*}\mathbf{Y}_t$ //evaluate $\mathbf{B}_{v-1}^{\mathrm{SD}}$ according to (20)
5: $z'_{v-1} = \mathbf{A}_{v-1}^{\mathrm{SD}}(\mathbf{B}_{v-1}^{\mathrm{SD}})^H \exp\left(j\frac{\pi}{M_P}\right)$ //evaluate decision varible of (20)
6: $\widetilde{\Delta}_{v-1}^{a} = \|\mathbf{A}_{v-1}^{\mathrm{SD}}\|^2 + \|\mathbf{B}_{v-1}^{\mathrm{SD}}\|^2 + \Xi_v + \overline{C}_{a,v-1} - \frac{L_a(b_1)+L_a(b_2)}{2}$ //evaluate ring-amplitude-related term of (18)
7: **for** $\overline{g} = 0$ **to** $M_P/4 - 1$
8: $g = \overline{g}M_A + a$ //update group index
9: $|t_{v-1}^{Re_g}| = |\overline{A}^{\overline{g}}\Re(z'_{v-1}) - \frac{L_a(b_2)}{2}|$ //associate $\Re(z'_{v-1})$ to $L_a(b_2)$ as defined in (22)
10: $|t_{v-1}^{Im_g}| = |\overline{B}^{\overline{g}}\Im(z'_{v-1}) - \frac{L_a(b_1)}{2}|$ //associate $\Im(z'_{v-1})$ to $L_a(b_1)$ as defined in (22)
11: $\mathrm{CW}_{v-1}^{g} = -|t_{v-1}^{Re_g}| - |t_{v-1}^{Im_g}| - P^g + \widetilde{\Delta}_{v-1}^{a}$ //update local minimum of each group
12: $b_1 = (t_{v-1}^{Im_g} < 0)$ //update local optimum child node index
13: $b_2 = (t_{v-1}^{Re_g} < 0)$
14: $\mathrm{CW}m_{v-1}^{g} = \mathbf{bin2dec}(b_1 b_2 b_3 \cdots b_{\mathrm{BPS}})$ //recall that we have $[b_3 \cdots b_{\mathrm{BPS}}] = \mathbf{dec2bin}(g)$
15: $n_{v-1}^{g} = 0$ //initialize child node counter for each group
16: **end for**
17: **end for**
18: $[\Delta_{v-1}, \hat{g}] = \mathbf{min}(\{\mathrm{CW}_{v-1}^{g}\}_{g=0}^{M/4-1})$ //update global minimum
19: $m_{v-1} = \mathrm{CW}m_{v-1}^{\hat{g}}$ //initialize global optimum child node index
20: $n_{v-1} = 0$ //update global child node counter

which is obtained by evaluating and comparing a reduced number of $M/4$ metrics according to the $M/4$ constellation points in the first quadrant.

Similar to [30], a Comparion Window (CW) may be introduced in order to invoke the Schnorr-Euchner search strategy, where the DQAM constellation points are visited in a zigzag fashion. More explicitly, when the SD visits a specific index v for the first time, the CW stores the local minima $\{\overline{\Delta}_{v-1}^{g} + \widetilde{\Delta}_{v-1}^{a}\}_{g=0}^{M/4-1}$ from all groups and produces the global minimum according to (25), which is represented by the "findBest" subfunction of Table IV. When the SD re-visits a specific index v, the CW has to update a new local child node for the group which is the one that produced the previous global child node. A group may be marked 'completed' when all its four QPSK-like child nodes have been tested using a zigzag pattern. By contrast, each incomplete group may provide a local child node candidate in the CW, and the CW may once again produce the global child node according to (25). This strategy is represented by the "findNext" subfunction of Table V. In summary, the soft-decision RC MSDSD algorithm designed for DPSK in Table I of [30] may be invoked for DQAM, but the subfunctions should be replaced by Tables IV and V of this paper.

Owing to the fact that only $M/4$ candidates are evaluated and compared by the CW, up to 75% of the child nodes are avoided by our RC MSDSD design, which is verified by the examples portrayed in Fig. 2. It can be seen in Fig. 2 that with the same number of SD steps, the soft-decision RC MSDSD aided 16-DAPSK(2,8) using the Schnorr-Euchner algorithm of Tables IV and V visits a substantially reduced number of constellation points compared to the soft-decision MSDSD aided 16-TDAPSK$^{\mathrm{JM}}$(2,8) invoking Table III.

Moreover, as discussed in [30], the RC Schnorr-Euchner algorithm of Tables IV and V avoids to invoke the sorting algorithm as seen in line 5 of Table III, where the average number of comparisons required by the classic sorting algorithms (e.g. Bubble sort, Timsort, Library sort, etc.) [63], [64] is as high as $O(M \log M)$.

C. RC MSDSD-IAP Algorithm

In order to further reduce the MSDSD complexity in coded DQAM systems, we propose the RC MSDSD-IAP algorithm as follows:

Step 1: An initial estimate of the phase matrix $\bar{\mathbf{P}}$ of (4) may be obtained by the CDD as introduced in [37]. Explicitly, the data-carrying phases are given by $\hat{\omega}_{v-1} = \exp(j\frac{2\pi}{M_P}\check{p}_{v-1})$, where we have $\check{p}_{v-1} = \lfloor \frac{M_P}{2\pi}\angle(\mathbf{Y}_v\mathbf{Y}_{v-1}^H)\rceil$ for $v \in \{2, \ldots, N_w\}$. Then the transmitted phases in $\hat{\bar{\mathbf{P}}}$ may be obtained by $\hat{\bar{\Omega}}_v = \hat{\omega}_{v-1}\hat{\bar{\Omega}}_{v-1}$ commencing from $\hat{\bar{\Omega}}_1 = 1$.

Step 2: Upon obtaining $\hat{\bar{\mathbf{P}}}$, the transmitted ring-amplitudes Γ_1 and $\bar{\mathbf{A}}$ of (4) may be estimated by the Multiple-Symbol Differential Amplitude Sphere Detection (MSDASD) as:

$$\left\{\hat{\Gamma}_1, \hat{\bar{\mathbf{A}}}\right\} = \arg \min_{\forall \Gamma_1} \min_{\forall \bar{\mathbf{A}}} d^{MSDASD}. \tag{26}$$

where the MSDASD's ED of d^{MSDASD} is given by the MSDSD's ED of (12) associated with the fixed phases $\{\hat{\bar{\Omega}}_v\}_{v=1}^{N_w}$ as:

TABLE V
PSEUDO-CODE FOR THE SCHNORR–EUCHNER SEARCH STRATEGY TAILORED FOR SOFT-DECISION RC MSDSD AIDED DQAM (PART II)

Subfunction: $[\{\mathrm{CW}_{v-1}^{g}\}_{g=0}^{M/4-1}, \{\mathrm{CW}m_{v-1}^{g}\}_{g=0}^{M/4-1}, \{n_{v-1}^{g}\}_{g=0}^{M/4-1}, \{\mathrm{cd}_{v-1}^{g}\}_{g=0}^{M/4-1}, \Delta_{v-1}, m_{v-1}, n_{v-1}]$
$= \textbf{findNext}(\{|t_{v-1}^{Re_g}|\}_{g=0}^{M/4-1}, \{|t_{v-1}^{Im_g}|\}_{g=0}^{M/4-1}, \{\widetilde{\Delta}_{v-1}^{a}\}_{a=0}^{M_A-1}, \{\mathrm{CW}_{v-1}^{g}\}_{g=0}^{M/4-1}, \{\mathrm{CW}m_{v-1}^{g}\}_{g=0}^{M/4-1},$
$\{n_{v-1}^{g}\}_{g=0}^{M/4-1}, \{\mathrm{cd}_{v-1}^{g}\}_{g=0}^{M/4-1}, \Delta_{v-1}, m_{v-1}, n_{v-1})$

1: $b_1 \cdots b_{\mathrm{BPS}} = \textbf{dec2bin}(m_{v-1})$ //obtain the previously tested child node
2: $\hat{a} = \textbf{bin2dec}(b_{\mathrm{BPS}_P+1} \cdots b_{\mathrm{BPS}})$ //previously tested group's amplitude index
3: $\hat{g} = \textbf{bin2dec}(b_3 \cdots b_{\mathrm{BPS}})$ //update previously tested group's index
4: $n_{v-1}^{\hat{g}}++$ //update child node counter
5: **switch** $n_{v-1}^{\hat{g}}$
6: **case** 1: $\mathrm{cd}_{v-1}^{\hat{g}} = \textbf{sign}(|t_{v-1}^{Re_{\hat{g}}}| - |t_{v-1}^{Im_{\hat{g}}}|)$ //update the condition of group $\hat{g}$
7: **if** $\mathrm{cd}_{v-1}^{\hat{g}} == 1$ //the case of $|t_{v-1}^{Re_{\hat{g}}}| > |t_{v-1}^{Im_{\hat{g}}}|$
8: $\mathrm{CW}_{v-1}^{\hat{g}} = -|t_{v-1}^{Re_{\hat{g}}}| + |t_{v-1}^{Im_{\hat{g}}}| - P^{\hat{g}} + \widetilde{\Delta}_{v-1}^{\hat{a}}$ //alter the imaginary part of the local minimum
9: $\mathrm{CW}m_{v-1}^{\hat{g}} = \textbf{bin2dec}(\bar{b}_1 b_2 b_3 \cdots b_{\mathrm{BPS}})$ //alter b_1 in the mapping of the local optimum
10: **else**
11: $\mathrm{CW}_{v-1}^{\hat{g}} = |t_{v-1}^{Re_{\hat{g}}}| - |t_{v-1}^{Im_{\hat{g}}}| - P^{\hat{g}} + \widetilde{\Delta}_{v-1}^{\hat{a}}$ //alter the real part of the local minimum
12: $\mathrm{CW}m_{v-1}^{\hat{g}} = \textbf{bin2dec}(b_1 \bar{b}_2 b_3 \cdots b_{\mathrm{BPS}})$ //alter b_2 in the mapping of the local optimum
13: **end if**
14: **break**
15: **case** 2: $\mathrm{CW}_{v-1}^{\hat{g}} = -\overline{\Delta}_{v-1} - 2P^{\hat{g}} + 2\widetilde{\Delta}_{v-1}^{\hat{a}}$ //alter decision made in Case 1
16: $\mathrm{CW}m_{v-1}(\hat{g}) = \textbf{bin2dec}(\bar{b}_1 \bar{b}_2 b_3 \cdots b_{\mathrm{BPS}})$
17: **break**
18: **case** 3: $\mathrm{CW}_{v-1}^{\hat{g}} = |t_{v-1}^{Re_{\hat{g}}}| + |t_{v-1}^{Im_{\hat{g}}}| - P^{\hat{g}} + \widetilde{\Delta}_{v-1}^{\hat{a}}$ //alter the local optimum child node
19: **if** $\mathrm{cd}_{v-1}^{\hat{g}} == 1$ $\mathrm{CW}m_{v-1}^{\hat{g}} = \textbf{bin2dec}(\bar{b}_1 b_2 b_3 \cdots b_{\mathrm{BPS}})$ //alter b_1 in the mapping decision maded in Case 2
20: **else** $\mathrm{CW}m_{v-1}^{\hat{g}} = \textbf{bin2dec}(b_1 \bar{b}_2 b_3 \cdots b_{\mathrm{BPS}})$ //alter b_2 in the mapping decision maded in Case 2
21: **break**
22: **end switch**
23: $\Delta_{v-1} = \inf$ //initialize global minimum
24: **for** $g = 0$ **to** $M/4 - 1$
25: **if** $\mathrm{CW}_{v-1}^{g} < \Delta_{v-1}$ **and** $n_{v-1}^{g} < 4$ //compare local minima from 'incompleted' groups
26: $\Delta_{v-1} = \mathrm{CW}_{v-1}^{g}$ //update global minimum
27: $m_{v-1} = \mathrm{CW}m_{v-1}^{g}$ //update global child node
28: **end if**
29: **end for**
30: $n_{v-1}++$ //update global child node counter

$$d^{MSDASD} = \sum_{v=1}^{N_w} \left\| \sum_{t=1}^{v} \widetilde{l}_{v-t+1,1} \bar{\Psi}_t^* \hat{\bar{\Omega}}_t^* \mathbf{Y}_t \right\|^2 + N_R \ln\left(\Gamma_1^2 \rho_0 + N_0\right) + \sum_{v=2}^{N_w} \Xi_v - \sum_{v=2}^{N_w} \sum_{\bar{k}_v=\mathrm{BPS}_P+1}^{\mathrm{BPS}} \left[\widetilde{b}_{\bar{k}_v} L_a(b_{\bar{k}_v}) - \overline{C}_{a,\bar{k}_v}\right]. \quad (27)$$

We note that the phase-related term of $\{-\sum_{v=2}^{N_w}\sum_{\bar{k}_v=1}^{\mathrm{BPS}_P}[\widetilde{b}_{\bar{k}_v} L_a(b_{\bar{k}_v}) - \overline{C}_{a,\bar{k}_v}]\}$ in (12) is omitted from (27). Explicitly, the MSDASD of (26) may obtain $(\min_{\forall \bar{\mathbf{A}}} d^{MSDASD})$ with the aid of the SD algorithm introduced in Sec. IV-A. Since all phases $\{\hat{\bar{\Omega}}_v\}_{v=1}^{N_w}$ are known for the MSDASD, there are only a total of M_A candidates for the PED increment seen in line 3 of Table III, which may be expressed as $\Delta_{v-1}^{a} = \left\| \sum_{t=1}^{v} \widetilde{l}_{v-t+1,1} \bar{\Psi}_t^* \hat{\bar{\Omega}}_t^* \mathbf{Y}_t \right\|^2 + \Xi_v - \sum_{\bar{k}_v=\mathrm{BPS}_P+1}^{\mathrm{BPS}} [\widetilde{b}_{\bar{k}_v} L_a(b_{\bar{k}_v}) - \overline{C}_{a,\bar{k}_v}]$.

Step 3: After estimating the ring-amplitudes $\hat{\Gamma}_1$, $\hat{\mathbf{A}}$ and hence also obtaining $\hat{\bar{\mathbf{O}}}$ and $\hat{\mathbf{C}} = \hat{\bar{\mathbf{A}}}^H \mathbf{R}_{hh} \hat{\bar{\mathbf{A}}} + \mathbf{R}_{vv}$, the estimate of the M_PPSK candidates may be improved by the Multiple-Symbol Differential Phase Sphere Detection (MSDPSD) as:

$$\hat{\bar{\mathbf{P}}} = \arg \quad \min_{\forall \bar{\mathbf{P}}} \quad d^{MSDPSD}, \quad (28)$$

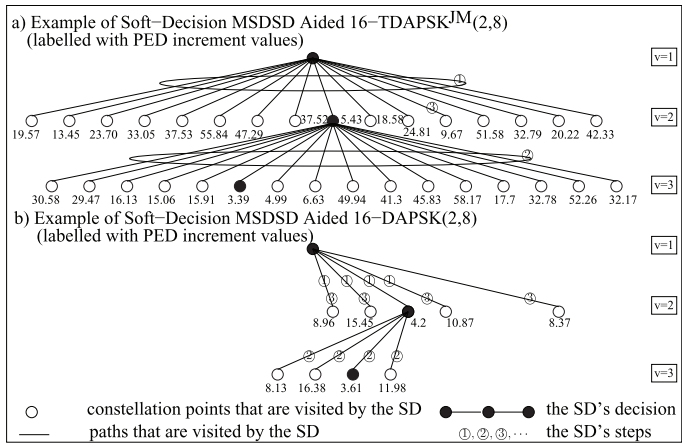


Fig. 2. Example of soft-decision MSDSD aided 16-TDAPSK$^{\mathrm{JM}}$(2,8) invoking the Schnorr-Euchner search strategy of Table III in comparison to soft-decision RC MSDSD aided 16-DAPSK(2,8) invoking Tables IV and V, which are recorded at SNR = 9 dB, $f_d = 0.03$ and $I_A = 0.3$, $N_R = 2$ and $N_w = 3$.

where the MSDPSD's ED of d^{MSDPSD} is given by the MSDSD's ED of (12) associated with the fixed $\hat{\Gamma}_1$, $\hat{\mathbf{A}}$ and $\{\hat{\Psi}_v\}_{v=1}^{N_w}$ as:

$$d^{MSDPSD} = \sum_{v=1}^{N_w} \left\| \sum_{t=1}^{v} l_{N_w-t+1,N_w-v+1} \hat{\bar{\Psi}}_t^* \bar{\Omega}_t^* \mathbf{Y}_t \right\|^2 - \sum_{v=2}^{N_w} \sum_{\bar{k}_v=1}^{\mathrm{BPS}_P} \left[\widetilde{b}_{\bar{k}} L_a(b_{\bar{k}}) - \overline{C}_{a,\bar{k}_v} \right]. \tag{29}$$

In more detail, $\{\{l_{N_w-t+1,N_w-v+1}\}_{t=1}^{v}\}_{v=1}^{N_w}$ are elements in the lower triangular matrix $\mathbf{L}$, which is decomposed from $\mathbf{L}\mathbf{L}^H = \hat{\mathbf{C}}^{-1}$. Moreover, the ring-amplitude-related term of $\{N_R \ln\left(\Gamma_1^2 \rho_0 + N_0\right) + \sum_{v=2}^{N_w} \Xi_v - \sum_{v=2}^{N_w} \sum_{\bar{k}_v=\mathrm{BPS}_P+1}^{\mathrm{BPS}} [\widetilde{b}_{\bar{k}_v} L_a(b_{\bar{k}_v}) - \overline{C}_{a,\bar{k}_v}]\}$ in the ED expression (12) is omitted from (29). It may be observed that the MSDPSD's ED (29) is in the same form as (17) in [30], where the vector in (17) of [30] is now given by $\{\{\mathbf{U}_{t,v} = l_{N_w-t+1,N_w-v+1} \hat{\bar{\Psi}}_t^* \mathbf{Y}_t\}_{t=1}^{v}\}_{v=1}^{N_w}$ according to (29), while the phase variables $\{\bar{s}_t\}_{t=1}^{v}$ in (17) of [30] correspond to $\{\bar{\Omega}_t\}_{t=1}^{v}$ in (29). Therefore, the RC MSDSD aided DPSK of [30] may be directly invoked for the MSDPSD of (28).

Step 4: In order to achieve a near-optimum MSDD performance, Step 2 and Step 3 may be repeated IR_{AP} times. Finally, the ring-amplitude-related soft-bit decisions may be made by the MSDASD as:

$$L_p(b_k) = \begin{cases} -d_{MAP}^{MSDASD} + \bar{d}_{MAP}^{MSDASD}, & \text{if } \mathrm{b_k^{MAP}} = 1, \\ -\bar{d}_{MAP}^{MSDASD} + d_{MAP}^{MSDASD}, & \text{if } \mathrm{b_k^{MAP}} = 0, \end{cases} \tag{30}$$

where $d_{MAP}^{MSDASD} = \min_{\forall \Gamma_1} \min_{\forall \tilde{\mathbf{A}}} d^{MSDASD}$ is obtained by the MSDASD of (26) in Step 2, while the corresponding hard-bit decisions $\{\{b_k^{MAP}\}_{k=(v-2)\mathrm{BPS}+\mathrm{BPS}_P+1}^{(v-1)\mathrm{BPS}}\}_{v=2}^{N_w}$ are also recorded along with d_{MAP}^{MSDASD}. Then $\bar{d}_{MAP}^{MSDASD} = \min_{\forall \Gamma_1} \min_{\forall \{\tilde{\mathbf{A}}\}_{b_k=\bar{b}_k^{MAP}}} d_{N_w}^{MSDASD}$ in (30) may be obtained by invoking the MSDASD algorithm again for each $L_p(b_k)$, where b_k is fixed to be the flipped MAP decision $\bar{b}_k^{MAP}$. Furthermore, the phase-related soft-bit decisions may be produced by the MSDPSD as:

$$L_p(b_k) = \begin{cases} -d_{MAP}^{MSDPSD} + \bar{d}_{MAP}^{MSDPSD}, & \text{if } \mathrm{b_k^{MAP}} = 1, \\ -\bar{d}_{MAP}^{MSDPSD} + d_{MAP}^{MSDPSD}, & \text{if } \mathrm{b_k^{MAP}} = 0, \end{cases} \tag{31}$$

where $d_{MAP}^{MSDPSD} = \min_{\forall \tilde{\mathbf{P}}} d^{MSDPSD}$ and the corresponding $\{\{b_k^{MAP}\}_{k=(v-2)\mathrm{BPS}+1}^{(v-2)\mathrm{BPS}+\mathrm{BPS}_P}\}_{v=2}^{N_w}$ may be obtained by (28) in Step 3, while $\bar{d}_{MAP}^{MSDPSD} = \min_{\forall \{\tilde{\mathbf{P}}\}_{b_k=\bar{b}_k^{MAP}}} d^{MSDPSD}$ is obtained by invoking the MSDPSD algorithm again for soft-bit decisions $\{\{L_p(b_k)\}_{k=(v-2)\mathrm{BPS}+1}^{(v-2)\mathrm{BPS}+\mathrm{BPS}_P}\}_{v=2}^{N_w}$, when the specific bit b_k is fixed to be the flipped MAP decision $\bar{b}_k^{MAP}$.

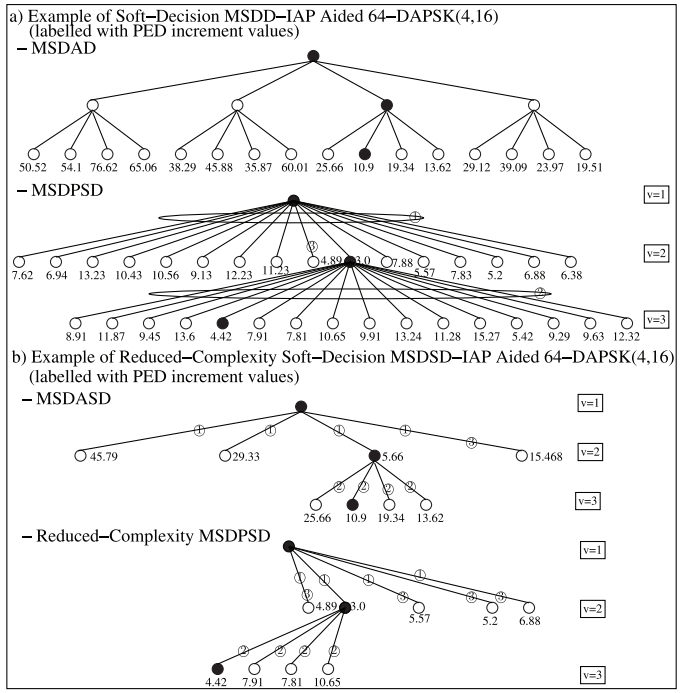


Fig. 3. Example of soft-decision MSDD-IAP of [38] and the proposed soft-decision RC MSDSD-IAP conceived for 64-DAPSK(4,16), which are recorded at SNR = 9 dB, $f_d = 0.03$ and $I_A = 0.3$, $N_R = 2$ and $N_w = 3$.

Fig. 3 exemplifies the comparison between the conventional MSDD-IAP of [38] and the RC MSDSD-IAP conceived for 64-DAPSK(4,16). It is evidenced by Fig. 3 that with the aid of sphere decoding, the MSDASD effectively reduces the number of ring-amplitude candidates visited by the conventional Multiple-Symbol Differential Amplitude Detection (MSDAD), while with the aid of the RC M_PPSK phase detection of [30], the RC MSDPSD also successfully reduces the number of M_PPSK candidates visited by the conventional MSDPSD.

V. Decision-Feedback Differential Detection

In order to conceive the DFDD that is equivalent to MSDD/MSDSD operating in decision-feedback mode, the MSDD of (7) may be revised for DFDD as:

$$L_p(b_k) = \max_{\forall \{\gamma_{N_w-1}, \omega_{N_w-1}\}_{b_k=1}} d(\gamma_{N_w-1}, \omega_{N_w-1}) - \max_{\forall \{\gamma_{N_w-1}, \omega_{N_w-1}\}_{b_k=0}} d(\gamma_{N_w-1}, \omega_{N_w-1}), \tag{32}$$

where the DFDD's probability metric is given by toggling the polarity of the MSDSD's PED increment Δ_{v-1} of (15) associated with the index $v = N_w$ as:

$$d(\gamma_{N_w-1}, \omega_{N_w-1}) = -\| l_{1,1} \hat{\Psi}_{N_w-1}^* \hat{\Omega}_{N_w-1}^* \mathbf{Y}_{N_w} + \omega_{N_w-1} \psi_{N_w-1} \left(\sum_{t=1}^{N_w-1} l_{N_w-t+1,1} \hat{\Psi}_t^* \hat{\Omega}_t^* \mathbf{Y}_t \right) \|^2 - \widetilde{\Xi}_{N_w} + \sum_{\bar{k}=1}^{\mathrm{BPS}} \widetilde{b}_{\bar{k}} L_a(b_{\bar{k}}). \tag{33}$$

The variable ring-amplitude γ_{N_w-1} determines $\{l_{N_w-t+1,1}\}_{t=1}^{N_w}$, ψ_{N_w-1} and $\widetilde{\Xi}_{N_w}$ in (33), where we have

$\widetilde{\Xi}_{N_w} = \ln\left[(\Gamma_{N_w}^2 + N_0) - \widetilde{\mathbf{e}}_{N_w}^T \widetilde{\mathbf{C}}_{N_w-1}^{-1} \widetilde{\mathbf{e}}_{N_w}\right]$. The constant ξ_{N_w} in Ξ_{N_w} as well as the constant $\ln \overline{C}_A^{N_w-1}$ seen in (15) may be ignored by the DFDD.

Furthermore, the first transmitted ring-amplitude Γ_1 of each DFDD window should still be treated as a separate variable for the differential-amplitude DQAM schemes. More explicitly, for the differential-amplitude DQAM using (2), any erroneous decisions concerning Γ_1 and $\{\gamma_t\}_{t=1}^{v-1}$ may degrade the decision reliablility concerning $\Gamma_v = \left(\prod_{t=1}^{v-1} \gamma_t\right)\Gamma_1$. By contrast, according to (1), we always have $\Gamma_v = \gamma_{v-1}$ for the absolute-amplitude DQAM schemes, which do not have the error propagation problem. Therefore, for the differential-amplitude DQAM of DAPSK, TDAPSK and TDAPSK$^{\mathrm{JM}}$, the DFDD using the Max-Log-MAP of (33) may be revised as (34), shown at the bottom of the page, where the probability metric is given by the MSDSD's ED of (12) associated with decision feedback based on $\{\hat{x}_v = \hat{\gamma}_v \hat{\omega}_v \hat{\psi}_v\}_{v=1}^{N_w-2}$ and hence also on $\{\hat{s}_v = \hat{\Gamma}_v \hat{\Omega}_v \hat{\Psi}_v\}_{v=2}^{N_w-1}$ as (35), shown at the bottom of the page, where the constant of $\sum_{v=2}^{N_w-1}[\sum_{\bar{k}=(v-2)\mathrm{BPS}+1}^{(v-1)\mathrm{BPS}} \widetilde{b}_{\bar{k}} L_a(b_{\bar{k}}) - \ln \overline{C}_A^{v-1}] - \ln \overline{C}_A^{N_w-1}$ in the MSDSD's ED of Eq. (12) is ignored. In this way, the potentially erroneous decision concerning Γ_{N_w} made during the current DFDD window will not degrade the following DFDD windows.

We note that the conventional DFDD in [17], [26], [27] ignored the problem of having a ring-amplitude-dependent channel correlation matrix $\mathbf{C} = \bar{\mathbf{A}}^H \mathbf{R}_{hh} \bar{\mathbf{A}} + \mathbf{R}_{vv}$. More explicitly, the DFDD probability metric for absolute-amplitude DQAM is given by [17]:

$$d(x_{N_w-1}) = -\frac{\left\| \mathbf{Y}_{N_w} - \frac{x_{N_w-1}\hat{s}_{N_w-1}}{\hat{\Gamma}_{N_w-1}} \left[\sum_{t=1}^{N_w} \overline{w}_t \mathbf{Y}_t / (\hat{s}_t)\right] \right\|^2}{1 + N_0 - \mathbf{e}_{N_w}^T \overline{\mathbf{w}}} + \sum_{\bar{k}=1}^{\mathrm{BPS}} \widetilde{b}_{\bar{k}} L_a(b_{\bar{k}}), \tag{36}$$

while that for differential DQAM is formulated as:

$$d(x_{N_w-1}) = -\frac{\left\| \mathbf{Y}_{N_w} - x_{N_w-1}\hat{s}_{N_w-1} \left[\sum_{t=1}^{N_w} \overline{w}_t \mathbf{Y}_t / (\hat{s}_t)\right] \right\|^2}{1 + N_0 - \mathbf{e}_{N_w}^T \overline{\mathbf{w}}} + \sum_{\bar{k}=1}^{\mathrm{BPS}} \widetilde{b}_{\bar{k}} L_a(b_{\bar{k}}), \tag{37}$$

where the prediction-based filter taps are $\overline{\mathbf{w}} = [\overline{w}_{N_w-1}, \ldots, \overline{w}_1]^T = \{\mathrm{Toeplitz}([\rho_0, \ldots, \rho_{N_w-2}]) + N_0 \mathbf{I}_{\mathbf{N_w-1}}\}^{-1} [\rho_1, \ldots, \rho_{N_w-1}]$, which are directly given by the filter taps of classic DPSK detection [24]. Consequently, a performance loss is inevitable, when they are compared to the proposed DFDD using (33) and (35).

VI. Performance Results

In this section, the DQAM detection results are examined in RSC, TC as well as IRCC and URC coded systems, where the simulation parameters are the same as those summarized in Table V of [30]. We note that the MSDSD window-width is set to be $N_w = 4$ in this paper. We will demonstrate that MSDSD associated with $N_w = 4$ is a compelling choice for DQAM in terms of its performance advantage over its coherent QAM counterpart, when relying on realistic imperfect CSI estimation in rapidly fading channels.

A. Suitable MSDSD Arrangements for Different DQAM Constellations

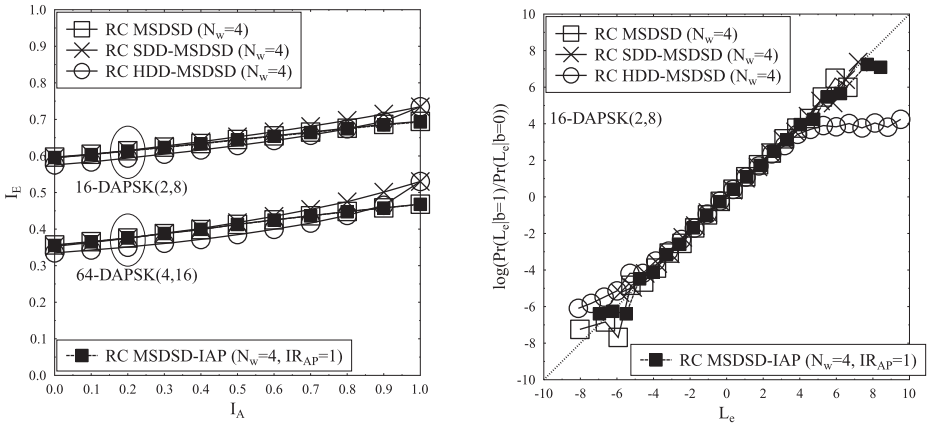
First of all, the EXIT charts of DAPSK employing RC MSDSD, RC SDD-MSDSD and RC HDD-MSDSD using the RC MSDSD algorithm of Sec. IV-B are portayed in Fig. 4(a), which demonstrate that SDD-MSDSD exhibits a better performance compared to both MSDSD and HDD-MSDSD. However, considering that SDD-MSDSD has to produce a soft decision feedback, which only provides a small improvement over MSDSD in Fig. 4(a), we opted for invoking the IAP algorithm of Sec. IV-C for RC MSDSD in form of RC MSDSD-IAP associated with $IR_{AP} = 1$, which does not impose any noticeable performance difference for DAPSK, as evidenced by Fig. 4(a).

As a further insightful aspect, the accuracy of the extrinsic LLRs produced by the MSDSD algorithms seen in Fig. 4(a) are tested as portrayed in Fig. 4(b), where the two PDFs $\{p(L_e|b)\}_{b=\{0,1\}}$ are obtained by estimating the histograms of L_e, with the source bits being $b = \{0, 1\}$. If the LLR definition of $L_e = \ln \frac{p(L_e|b=1)}{p(L_e|b=0)}$ is statistically accurate, then the LLRs accuracy test may result in a diagonal line. However, the LLRs produced by HDD-MSDSD aided DAPSK seen in Fig. 4(b) exhibit a poor integrity. This is because that according to the

$$L_p(b_k) = \max_{\forall \Gamma_1} \max_{\forall \{\gamma_{N_w-1}, \omega_{N_w-1}\}_{b_k=1}} d(\Gamma_1, \gamma_{N_w-1}, \omega_{N_w-1}) - \max_{\forall \Gamma_1} \max_{\forall \{\gamma_{N_w-1}, \omega_{N_w-1}\}_{b_k=0}} d(\Gamma_1, \gamma_{N_w-1}, \omega_{N_w-1}), \tag{34}$$

$$d(\Gamma_1, \gamma_{N_w-1}, \omega_{N_w-1}) = -\left\| l_{N_w,N_w} \mathbf{Y}_1 \right\|^2 - \ln\left(\Gamma_1^2 \rho_0 + N_0\right) - \sum_{v=2}^{N_w-1} \left\| \sum_{t=1}^{v} l_{N_w-t+1,N_w-v+1} \hat{\Psi}_t^* \hat{\Omega}_t^* \mathbf{Y}_t \right\|^2 - \left\| l_{1,1} \hat{\Psi}_{N_w-1}^* \hat{\Omega}_{N_w-1}^* \mathbf{Y}_{N_w} + \omega_{N_w-1} \psi_{N_w-1} \left(\sum_{t=1}^{N_w-1} l_{N_w-t+1,1} \hat{\Psi}_t^* \hat{\Omega}_t^* \mathbf{Y}_t \right) \right\|^2 - \sum_{v=2}^{N_w} \widetilde{\Xi}_v + \sum_{\bar{k}=1}^{\mathrm{BPS}} \widetilde{b}_{\bar{k}} L_a(b_{\bar{k}}), \tag{35}$$

$$C_{DCMC}^{MSDD}(SNR) = \frac{\sum_{a=0}^{M_A-1} \sum_{i=0}^{M^{(N_w-1)}-1} E\left\{ \log_2 \left[\frac{M^{(N_w-1)} \sum_{\tilde{a}=0}^{M_A-1} p(\mathbf{Y}|\bar{\mathbf{S}}^i, \Gamma^{\tilde{a}})}{\sum_{\tilde{a}=0}^{M_A-1} \sum_{\tilde{i}=0}^{M^{(N_w-1)}-1} p(\mathbf{Y}|\bar{\mathbf{S}}^{\tilde{i}}, \Gamma^{\tilde{a}})} \right] | \bar{\mathbf{S}} = \bar{\mathbf{S}}^i, \Gamma_1 = \Gamma^a \right\}}{(N_w - 1) M_A M^{(N_w-1)}}, \tag{38}$$



(a) EXIT charts (b) LLR accuracy test

Fig. 4. EXIT charts and LLR accuracy test of DAPSK employing RC MSDSD, RC SDD-MSDSD, RC HDD-MSDSD and RC MSDSD-IAP recorded at SNR = 9 dB and $f_d = 0.03$, where we have $N_R = 2$.

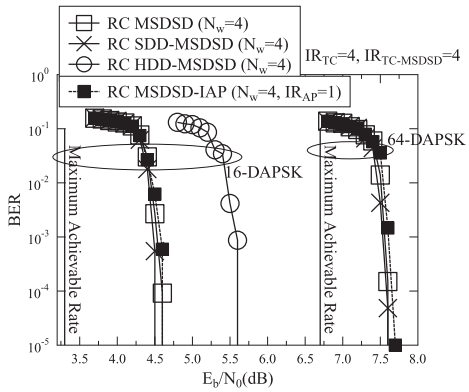


Fig. 5. BER performance comparison of RC MSDSD, RC SDD-MSDSD, RC HDD-MSDSD and RC MSDSD-IAP, when they are employed for DAPSK detection in TC coded system, where we have $f_d = 0.03$ and $N_R = 2$.

DAPSK's differential encoding of (2), the potentially erroneous hard-decision feedback concerning Γ_1 and $\{\gamma_t\}_{t=1}^{v-1}$ may degrade the accuracy of the following decisions on $\Gamma_v = \left(\prod_{t=1}^{v-1} \gamma_t\right) \Gamma_1$.

The EXIT charts and LLRs accuracy analysis of Fig. 4 are confirmed by the BER performance of Fig. 5, which demonstrates that RC MSDSD, RC SDD-MSDSD and RC MSDSD-IAP perform similarly, but RC HDD-MSDSD performs much worse than its counterparts, where the BER curve of TC coded 64-DAPSK(4,16) employing RC HDD-MSDSD cannot even be portrayed within our standard E_b/N_0 range. Therefore, we conclude that RC MSDSD-IAP may be suggested for DAPSK.

Moreover, we note that the Discrete-input Continuous-output Memoryless Channel (DCMC) capacity [14] of the MSDD aided DQAM systems is given by (38), shown at the bottom of the previous page, where $p(\mathbf{Y}|\bar{\mathbf{S}}^i, \Gamma^a)$ is given by (6), while we have $\{p(\Gamma^{\widetilde{a}}) = p(\Gamma^{\bar{a}}) = \frac{1}{M_A}\}_{\forall \widetilde{a}, \forall \bar{a}}$. For the case of employing HDD-MSDD, the perfect decision feedback leads to $\{p(\Gamma^{\widetilde{a}}) = p(\Gamma^{\bar{a}}) = 1\}_{\widetilde{a}=\bar{a}=a}$ and $\{p(\Gamma^{\widetilde{a}}) = p(\Gamma^{\bar{a}}) = 0\}_{\forall \widetilde{a} \neq a, \forall \bar{a} \neq a}$. The resultant DCMC capacity $C_{DCMC}^{HDD-MSDD}$ may be revised from C_{DCMC}^{MSDD} of (38), where both $\sum_{\widetilde{a}=0}^{M_A-1}$ and $\sum_{\bar{a}=0}^{M_A-1}$ are omitted, while both $p(\mathbf{Y}|\bar{\mathbf{S}}^i, \Gamma^{\widetilde{a}})$ and $p(\mathbf{Y}|\bar{\mathbf{S}}^{\bar{i}}, \Gamma^{\bar{a}})$ may be replaced by $p(\mathbf{Y}|\bar{\mathbf{S}}^i, \Gamma^a)$. According to the definitions of extrinsic information $I_E = I(b; L_e)$ and DCMC capacity $C_{DCMC} = I(\{\bar{\mathbf{S}}, \Gamma_1\}; \mathbf{Y})$, the area property of the EXIT charts [65] may be expressed as $\int_0^1 I_E(SNR, I_A) dI_A \approx \frac{C^{DCMC}(SNR)}{\text{BPS}}$, which implies that the area under the EXIT curve is approximately equal to the normalized DCMC capacity. The maximum achievable rates seen in Fig. 5 are obtained based on this property, which indicate the SNRs required for the half-rate channel coded DQAM systems to achieve their full capacity of 0.5BPS. It can be seen in Fig. 5 that the turbo detection performed by exchanging extrinsic information between the TC decoder and the MSDSD is capable of effectively exploiting the iteration gain of the MSDSD portrayed by Fig. 4(a), which results in a near-capacity performance.

TABLE VI
SUMMARY OF SUITABLE MSDSD ARRANGEMENTS FOR DIFFERENT DQAM CONSTELLATIONS

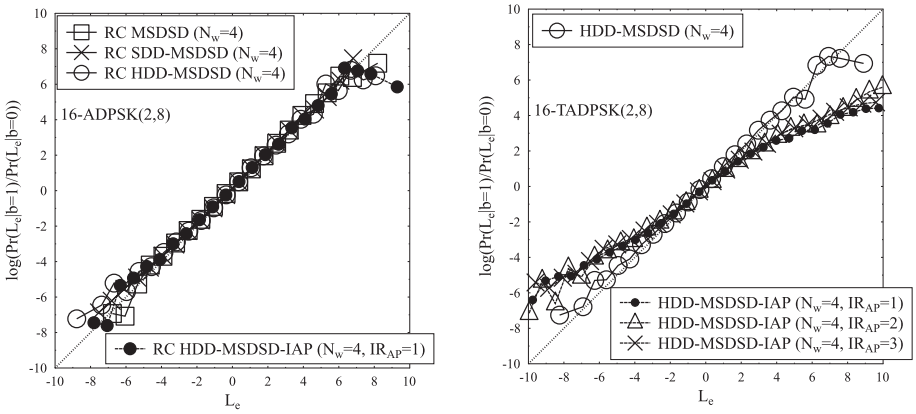
(a) Absolute-Amplitude DQAM

ADPSK	RC HDD-MSDSD-IAP
TADPSK	RC HDD-MSDSD
TADPSK$^{\text{JM}}$	HDD-MSDSD

(b) Differential-Amplitude DQAM

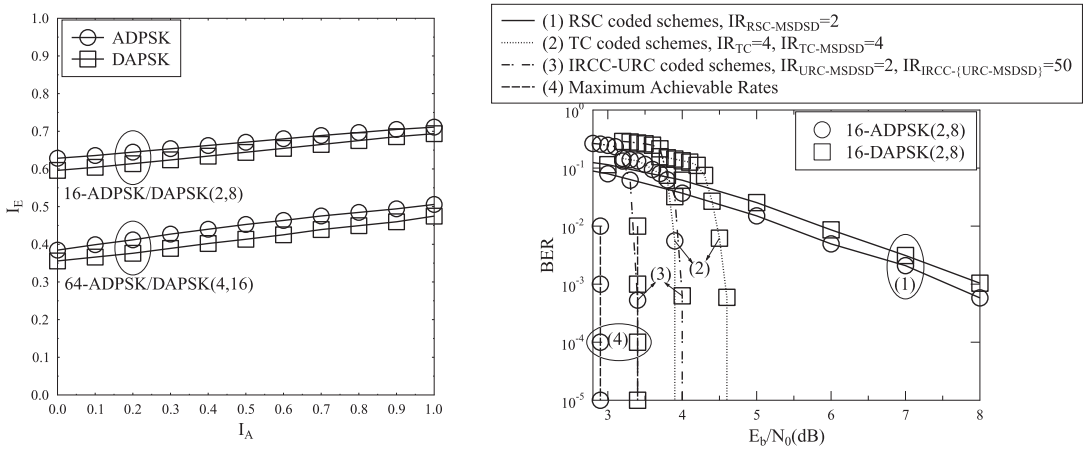
DAPSK	RC MSDSD-IAP
TDAPSK	RC MSDSD
TDAPSK$^{\text{JM}}$	MSDSD

The same analysis relying on EXIT charts, on the LLR accuracy test and on the BER performance may be carried out for all DQAM constellations. Due to the journal's space limit, we directly offer our conclusions in Table VI. More explicitly, the absolute-amplitude DQAM schemes of ADPSK, TADPSK and TADPSK$^{\text{JM}}$ may employ the HDD family of MSDSD arrangements. For example, the LLRs produced by RC HDD-MSDSD aided ADPSK seen in Fig. 6(a) exhibit a good integrity. This is because according to (1), the absolute-amplitude DQAM schemes have $\Gamma_v = \gamma_{v-1}$, which do not cause error propagation. Furthermore, owing to the amplitude-phase correlation, the twisted DQAM schemes of TADPSK and TDAPSK cannot employ the IAP algorithm of Sec. IV-C, which is exemplified by Fig. 6(b) for the case of RC HDD-MSDSD-IAP aided TADPSK. Moreover, the joint-mapping DQAM schemes of



(a) ADPSK (b) TADPSK

Fig. 6. LLR accuracy test of ADPSK and TADPSK employing different MSDSD arrangements recorded at SNR = 9 dB, $f_d = 0.03$ and $N_R = 2$.



(a) EXIT charts (b) BER performance

Fig. 7. EXIT charts and BER comparison between ADPSK employing RC HDD-MSDSD-IAP ($N_w = 4$, $IR_{AP} = 1$) and DAPSK employing RC MSDSD-IAP ($N_w = 4$, $IR_{AP} = 1$), where we have $f_d = 0.03$ and $N_R = 2$.

TADPSK$^{\text{JM}}$ and TDAPSK$^{\text{JM}}$ can only employ the generic MSDSD algorithm of Sec. IV-A, which jointly detect the ring-amplitude and phase.

B. Comparison Between DQAM Constellations

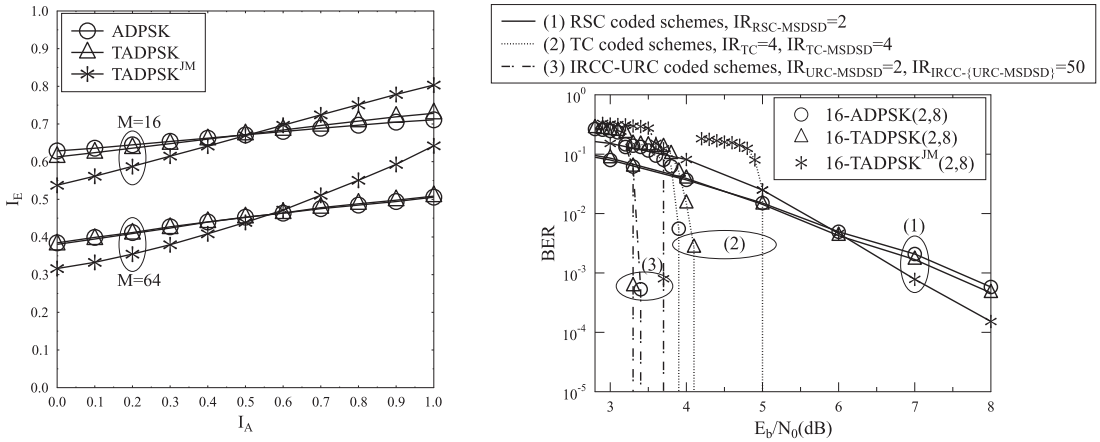
First of all, Fig. 7 portrays comparison between ADPSK and DAPSK. More explicitly, the EXIT charts of Fig. 7(a) indicates that ADPSK outperforms DAPSK, which confirms the capacity results shown in [16]. Furthermore, the BER performance comparison of Fig. 7(b) explicitly demonstrates that 16-ADPSK(2,8) outperforms its 16-DAPSK(2,8) counterpart in all three of the RSC, TC and RSC-URC coded systems.

Fig. 8 further portrays our performance comparison between the classic ADPSK and its twisted counterparts of TADPSK and TADPSK$^{\text{JM}}$. More explicitly, the EXIT charts of Fig. 8(a) demonstrate that the ADPSK, TADPSK and TADPSK$^{\text{JM}}$ exhibit a similar achievable rate, which is reflected by the area under the EXIT curves [65]. Nonetheless, TADPSK exhibits a slightly higher iteration gain than ADPSK, while TADPSK$^{\text{JM}}$ achieves the highest iteration gain, as demonstrated by Fig. 8(a). As a result, it is demonstrated by the BER performance of Fig. 8(b) that 16-TADPSK$^{\text{JM}}$(2,8) outperforms its counterparts in RSC coded system, while 16-ADPSK(2,8) performs the best in TC coded system. This is because that the steep EXIT curve of TADPSK$^{\text{JM}}$ matches better the EXIT curve shape of RSC, while the less steep EXIT curve of ADPSK matches better to the horizontal EXIT curve of TC [14]. Furthermore, Fig. 8(b) also shows that 16-TADPSK(2,8) may outperform its counterparts in IRCC and URC coded near-capacity system. Explicitly, the number of iterations between the URC and MSDSD is given by $IR_{URC-MSDSD} = 2$, which may not be sufficient for reaping the full benefit of the high iteration gain of TDAPSK$^{\text{JM}}$, but unfortunately a higher $IR_{URC-MSDSD}$ may not be affordable. As a result, 16-TADPSK$^{\text{JM}}$(2,8) performs the worst in IRCC and URC coded system, as evidenced by Fig. 8(b).

In conclusion, ADPSK is a better choice than DAPSK in channel coded systems, according to Fig. 7. Moreover, the ADPSK's twisted modulated counterparts of TADPSK and TADPSK$^{\text{JM}}$ may exhibit their advantages in different coded systems, as suggested by Fig. 8.

C. Comparison Between DQAM and DPSK

The performance comparison of Fig. 9 demonstrates that both ADPSK and DAPSK substantially outperform their DPSK



(a) EXIT charts (b) BER performance

Fig. 8. EXIT charts and BER comparison between ADPSK employing RC HDD-MSDSD-IAP ($N_w = 4$, $IR_{AP} = 1$), TADPSK employing RC HDD-MSDSD ($N_w = 4$) and TADPSK$^{\mathrm{JM}}$ employing HDD-MSDSD ($N_w = 4$), where we have $f_d = 0.03$ and $N_R = 2$.

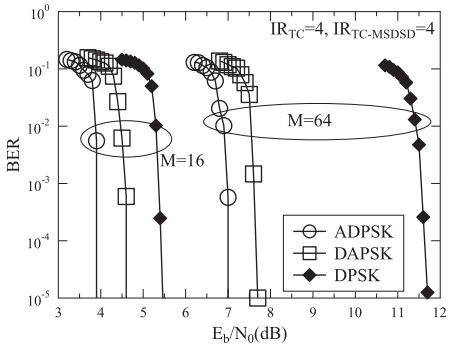


Fig. 9. BER comparison between ADPSK employing RC HDD-MSDSD-IAP ($N_w = 4$, $IR_{AP} = 1$), DAPSK employing RC MSDSD-IAP ($N_w = 4$, $IR_{AP} = 1$) and DPSK employing RC MSDSD ($N_w = 4$) in TC coded systems, where we have $f_d = 0.03$ and $N_R = 2$.

counterparts. Explicitly, Fig. 9 shows that the RC HDD-MSDSD-IAP ($N_w = 4$, $IR_{AP} = 1$) aided 16-ADPSK(2,8) outperforms RC MSDSD ($N_w = 4$) aided 16-DPSK of [30] by about **1.6 dB** in TC systems, where we have $N_R = 2$ and $f_d = 0.03$. This performance difference becomes a more substantial **4.7 dB** for $M = 64$, as seen in Fig. 9.

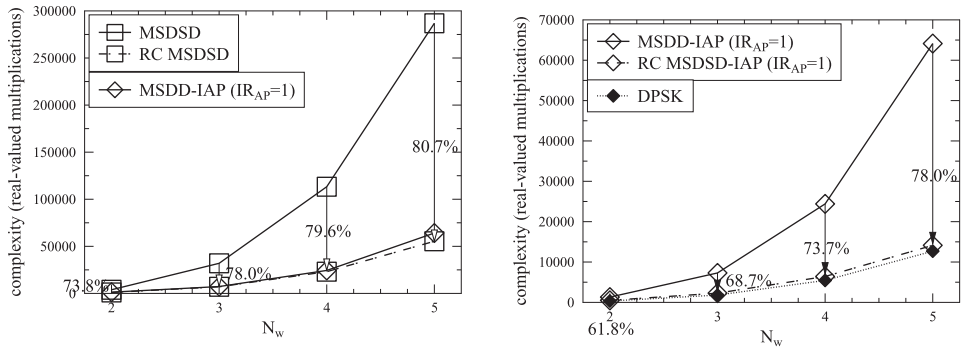
Moreover, the complexity of our MSDSD algorithms is further quantified in terms of the total number of real-valued multiplications in Fig. 10. Explicitly, Fig. 10(a) demonstrates that the RC MSDSD algorithm of Sec. IV-B may offer a substantial $73.8\% \sim 80.7\%$ complexity reduction compared to the generic MSDSD algorithm of Sec. IV-A. As a result, the MSDSD complexity becomes as low as that of the MSDD-IAP of [38], as seen in Fig. 10(a). Furthermore, it is demonstrated by Fig. 10(b) that the RC MSDSD-IAP of Sec. IV-C achieves an additional $61.8\% \sim 78.0\%$ complexity reduction compared to the conventional MSDD-IAP of [38]. Consequently, the complexity of RC MSDSD-IAP aided DAPSK becomes comparable to that of the RC MSDSD aided DPSK of [30], as evidenced by Fig. 10(b).

Fig. 11 provides our complexity comparison of DQAM and DPSK schemes. Firstly, it is demonstrated by Fig. 11 that among the absolute-amplitude DQAM schemes of ADPSK, TADPSK and TADPSK$^{\mathrm{JM}}$, the HDD-MSDSD aided TADPSK$^{\mathrm{JM}}$ of Sec. IV-A exhibits the highest complexity, which is followed by the RC HDD-MSDSD aided TADPSK of Sec. IV-B and then by the RC HDD-MSDSD-IAP aided ADPSK of Sec. IV-C. Secondly, Fig. 11 also confirms that RC MSDSD-IAP aided DAPSK exhibits a higher complexity than RC HDD-MSDSD-IAP aided ADPSK, which avoided the comparison over Γ_1. Last but not least, Fig. 11(a) shows that the complexity of RC HDD-MSDSD-IAP ($IR_{AP} = 1$) aided 16-ADPSK(2,8) is similar to that of RC MSDSD aided 16-DPSK of [30], while Fig. 11(b) demonstrates that the complexity of RC HDD-MSDSD-IAP ($IR_{AP} = 1$) aided 64-ADPSK(4,16) is even lower than that of RC MSDSD aided 64-DPSK. This is because the complexity order of the Schnorr-Euchner search strategy of the RC MSDSD aided DPSK of [30] is given by $O(M/4)$, where only the $M/4$ constellation points in the first quadrant are visited. Meanwhile, the complexity order for the case of RC HDD-MSDSD-IAP ($IR_{AP} = 1$) aided ADPSK is given by $O(M_A + M_P/4)$ according to Sec. IV-C. As a result, we have $O(M_A + M_P/4) = O(M/4)$ for $M = 16$ and $O(M_A + M_P/4) < O(M/4)$ for $M = 64$, which verifies our complexity results of Fig. 11.

D. Comparison Between Coherent and Noncoherent Schemes

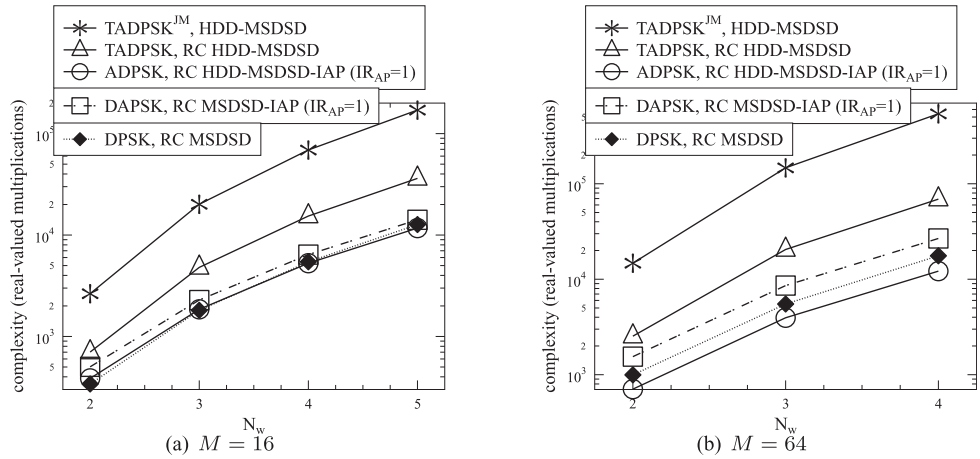
In this section, the important subject of coherent versus noncoherent schemes is discussed. For the noncoherent ADPSK scheme, RC HDD-Subset MSDSD-IAP is employed, where the consecutive MSDSD windows are overlapped by $N_{OL} = 3$ observations and the $(N_{OL} - 1 = 2)$ symbols detected at the edges may be discarded, as suggested by [58]. For the coherent Square QAM scheme, PSAM [7] associated with pilot spacing of N_{PS} and observation window width of N_{OW} (number of filter taps) is invoked for channel estimation.

Fig. 12 demonstrates that when f_d is increased from 0.001 to 0.03, a severe deviation from the LLR definition emerges for the extrinsic LLRs produced by the PSAM aided coherent 16QAM detector. This is because the coherent detectors assume having



(a) MSDSD and RC MSDSD (b) MSDD-IAP and RC MSDSD-IAP

Fig. 10. Complexity (real-valued multiplications) comparison between MSDSD of Sec. IV-A, RC MSDSD of Sec. IV-B, MSDD-IAP of [38] and RC MSDSD-IAP of Sec. IV-C employed by 16-DAPSK(2,8). Complexity of RC MSDSD aided 16-DPSK of [30] is also portrayed. The results are recorded at SNR $= 0$ dB and $I_A = 1$, where we have $f_d = 0.03$ and $N_R = 2$.



(a) $M = 16$ (b) $M = 64$

Fig. 11. Complexity (real-valued multiplications) comparison between DQAM schemes employing their suitable MSDSD arrangements. Complexity of RC MSDSD aided DPSK of [30] is also portrayed. The results are recorded at SNR $= 0$ dB and $I_A = 1$, where we have $f_d = 0.03$ and $N_R = 2$.

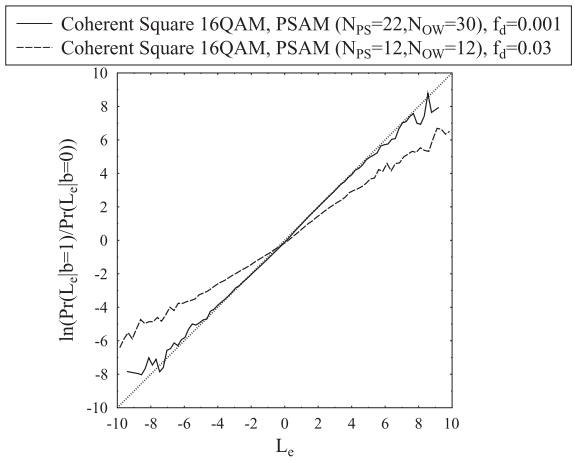


Fig. 12. LLR accuracy test of PSAM aided Square 16QAM recorded at $E_b/N_0 = 2$ dB and $I_A = 0$, where we have $N_R = 1$. When we have $f_d = 0.03$, PSAM's N_{PS} is reduced to 12 in order to sample the channel more frequently, while N_{OW} is also reduced to 12, due to the weak temporal correlation.

perfect knowledge of the CSI, which is especially unrealistic, when the fading channel fluctuates rapidly.

Finally, Fig. 13(a) demonstrates that when we have $f_d = 0.001$, PSAM aided coherent 16QAM outperforms RC HDD-Subset MSDSD-IAP assisted 16-ADPSK(2,8) in the RSC coded system, TC coded system as well as in the IRCC and URC coded system. However, when we have $f_d = 0.03$, 16-ADPSK(2,8) exhibits an impressive performance advantage over coherently detected Square 16QAM, which is 0.5 dB and 0.9 dB in the context of our TC and IRCC-URC coded systems, respectively, as evidenced by Fig. 13(b).

E. Performance Results for DFDD

The BER performance results of the proposed DFDD solutions and the conventional DFDD solutions [17], [26], [27] are compared in Fig. 14 in the context of TC coded ADPSK and DAPSK systems. The DFDD window-width is set to $N_w = 3$, because any further performance improvement for $N_w > 3$ was shown to become negligible by our results not shown in this paper due to the strict space limit. It can be seen in Fig. 14 that the MSDSD solutions generally outperform their DFDD counterparts. Nonetheless, Fig. 14 also demonstrates that the proposed DFDD solutions substantially improve the performance of the conventional DFDD solutions by 1.4 dB and 1.3 dB in TC coded ADPSK and DAPSK systems, respectively.

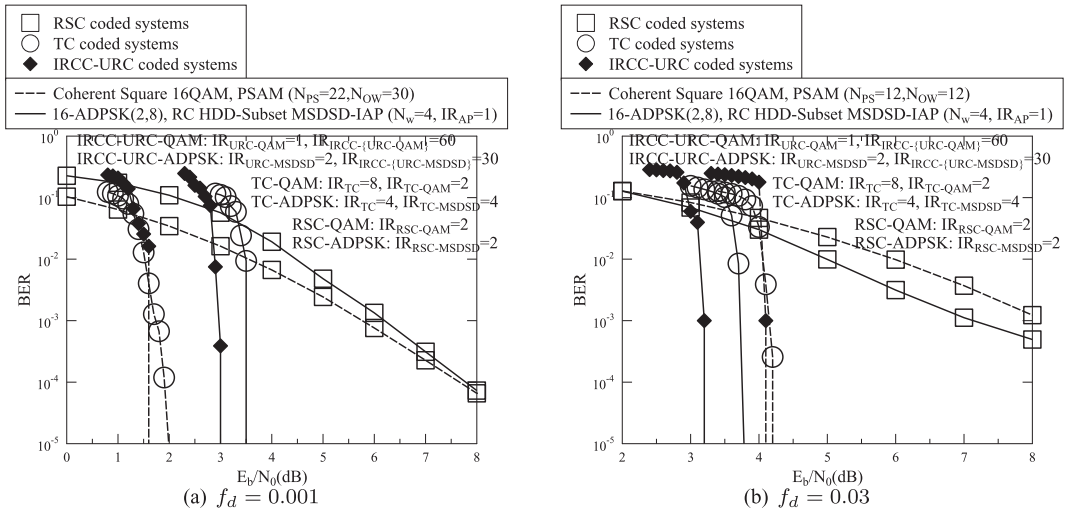


Fig. 13. BER performance comparison between RSC/TC/IRCC-URC coded PSAM aided Square 16QAM and its 16-ADPSK(2,8) counterpart employing RC HDD-Subset MSDSD-IAP ($N_w = 4$, $IR_{AP} = 1$), where we have $N_R = 2$.

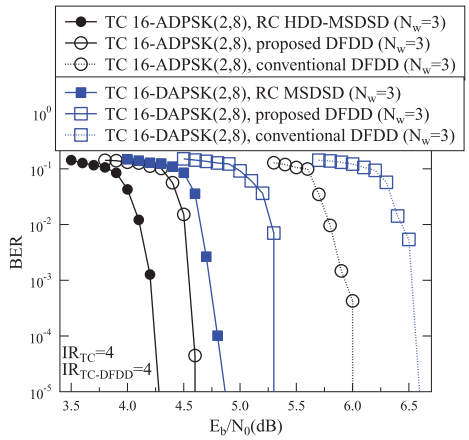


Fig. 14. BER performance results of TC coded DFDD ($N_w = 3$) aided ADPSK and DAPSK, where we have $N_R = 2$ and $f_d = 0.03$. The proposed DFDD solutions use the probability metrics of (33) and (35), while the conventional DFDD solutions [17], [26], [27] use (36) and (37).

VII. Conclusions

In this paper, we developed a comprehensive solution for the soft-decision DQAM detection in rapidly fading channels. First of all, we proposed the original set, the HDD set and the SDD set of MSDD/MSDSD solutions, which respectively correspond to having no decision feedback, hard-decision feedback and soft-decision feedback on the first DQAM symbol's ring-amplitude in each MSDD/MSDSD window. In this way, the potential error propagation problem in soft-decision DQAM's ring-amplitude detection may be avoided. Secondly, we proposed to invoke soft-decision SD for both DQAM's ring-amplitude and phase detection, which has not been seen in open literature. Thirdly, by exploring the symmetry provided by Gray-labelled DQAM constellations, we proposed a RC MSDSD algorithm, which visit a reduced number of constellation points without imposing any performance loss. Fourthly, we further proposed a RC MSDSD-IAP algorithm, which separately and iteratively detecting the DQAM's ring-amplitudes and phases by two soft-decision SDs. Fifthly, we developed the improved soft-decision DFDD solutions based on the MSDD/MSDSD operating in decision-feedback mode, which substantially outperfom the existing prediction-based DFDD solutions that do not take into account the DQAM ring-amplitudes' effect on the channel's correlation matrix.

With the aid of EXIT charts and LLR accuracy test, the most suitable soft-decision MSDSD arrangements of different DQAM constellations were suggested. Our simulation results demonstrate that the absolute-amplitude ADPSK schemes outperform their differential-amplitude DAPSK counterparts in coded systems. Furthermore, among the absolute-amplitude DQAM schemes, ADPSK, TADPSK and TADPSK^{JM} have a better BER performance in TC coded, IRCC-and-URC coded as well as RSC coded systems, respectively. Thanks to the proposed reduced-complexity design, the RC HDD-MSDSD-IAP aided ADPSK is capable of outperforming its MSDSD aided DPSK counterparts in coded systems *without imposing a higher complexity*. Moreover, our simulation results also verify that compared to coherent Square QAM relying on realistic imperfect channel estimation, MSDSD aided DQAM is a more suitable candidate for turbo detection assisted coded systems operating in rapidly fading channels.

References

[1] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[2] S. Adireddy, T. Lang, and H. Viswanathan, "Optimal placement of training for frequency-selective block-fading channels," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2338–2353, Aug. 2002.

[3] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.

[4] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[5] M. Arti, R. Mallik, and R. Schober, "Joint channel estimation and decoding of space-time block codes in AF MIMO relay networks," in *Proc. Int. Conf. Signal Process. Commun. (SPCOM)*, Jul. 2012, pp. 1–5.

[6] M. Arti, "Channel estimation and detection in hybrid satellite-terrestrial communication systems," *IEEE Trans. Veh. Technol.*, 2015, to be published [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=7156170

[7] J. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 686–693, Nov. 1991.

[8] ITU-R. (2009). *Coexistence Between IMT-2000 PCDMA-DS and IMT-2000 OFDMA TDD WMAN in the 2 500-2 690 MHz Band Operating in Adjacent Bands in the Same Area* [Online]. Available: http://www.itu.int/pub/R-REP-M.2146-2009

[9] Qualcomm. (2012). *IEEE802.11ac: The Next Evolution of Wi-Fi Standards* [Online]. Available: http://www.qualcomm.com/media/documents/files/ieee802-11ac-the-next-evolution-of-wi-fi.pdf

[10] J. Hagenauer, "The turbo principle: Tutorial introduction and state of the art," in *Proc. Int. Symp. Turbo Codes Related Topics*, 1997, pp. 1–11.

[11] M. El-Hajjar and L. Hanzo, "EXIT charts for system design and analysis," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 127–153, Feb. 2014.

[12] S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative demapping and decoding for multilevel modulation," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM'98)*, 1998, vol. 1, pp. 579–584.

[13] C. Xu, D. Liang, S. Sugiura, S. X. Ng, and L. Hanzo, "Reduced-complexity approx-log-MAP and max-log-MAP soft PSK/QAM detection algorithms," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1415–1425, Apr. 2013.

[14] L. Hanzo, O. Alamri, M. El-Hajjar, and N. Wu, *Near-Capacity Multi-Functional MIMO Systems: Sphere-Packing, Iterative Detection and Cooperation*. Hoboken, NJ, USA: Wiley, May 2009.

[15] W. Weber, "Differential encoding for multiple amplitude and phase shift keying systems," *IEEE Trans. Commun.*, vol. COM-26, no. 3, pp. 385–391, Mar. 1978.

[16] R. Fischer, L. Lampe, and S. Calabr, "Differential encoding strategies for transmission over fading channels," *Int. J. Electron. Commun.*, vol. 54, no. 1, pp. 59–67, 2000.

[17] L. Lampe and R. Schober, "Low-complexity iterative demodulation for noncoherent coded transmission over Ricean-fading channels," *IEEE Trans. Veh. Technol.*, vol. 50, no. 6, pp. 1481–1496, Nov. 2001.

[18] W. Webb, L. Hanzo, and R. Steele, "Bandwidth efficient QAM schemes for Rayleigh fading channels," *IEE Proc. I, Commun. Speech Vis.*, vol. 138, no. 3, pp. 169–175, Jun. 1991.

[19] F. Adachi and M. Sawahashi, "Performance analysis of various 16 level modulation schemes under Rayleigh fading," *Electron. Lett.*, vol. 28, no. 17, pp. 1579–1581, Aug. 1992.

[20] D. Divsalar and M. K. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 300–308, Mar. 1990.

[21] P. Ho and D. Fung, "Error performance of multiple-symbol differential detection of PSK signals transmitted over correlated Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 40, no. 10, pp. 1566–1569, Oct. 1992.

[22] D. Divsalar and M. Simon, "Maximum-likelihood differential detection of uncoded and trellis coded amplitude phase modulation over AWGN and fading channels-metrics and performance," *IEEE Trans. Commun.*, vol. 42, no. 1, pp. 76–89, Jan. 1994.

[23] H. Leib and S. Pasupathy, "The phase of a vector perturbed by Gaussian noise and differentially coherent receivers," *IEEE Trans. Inf. Theory*, vol. 34, no. 6, pp. 1491–1501, Nov. 1988.

[24] R. Schober, W. Gerstacker, and J. Huber, "Decision-feedback differential detection of MDPSK for flat Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1025–1035, Jul. 1999.

[25] R. Schober, W. Gerstacker, and J. Huber, "Decision-feedback differential detection scheme for 16-DAPSK," *Electron. Lett.*, vol. 34, no. 19, pp. 1812–1813, Sep. 1998.

[26] W. Gerstacker, R. Schober, and J. Huber, "Decision-feedback differential detection for 16DAPSK transmitted over Rician fading channels," in *Proc. IEEE 50th Veh. Technol. Conf. (VTC'99-Fall)*, Amsterdam, The Netherlands, Sep. 1999, vol. 5, pp. 2515–2519.

[27] R. Schober, W. H. Gerstacker, and J. B. Huber, "Decision-feedback differential detection based on linear prediction for 16DAPSK signals transmitted over flat Ricean fading channels," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1339–1342, Aug. 2001.

[28] L. Lampe, R. Schober, V. Pauli, and C. Windpassinger, "Multiple-symbol differential sphere decoding," *IEEE Trans. Commun.*, vol. 53, no. 12, pp. 1981–1985, Dec. 2005.

[29] V. Pauli, L. Lampe, and R. Schober, ""Turbo DPSK" using soft multiple-symbol differential sphere decoding," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1385–1398, Apr. 2006.

[30] C. Xu, X. Zuo, S. X. Ng, R. G. Maunder, and L. Hanzo, "Reduced-complexity soft-decision multiple-symbol differential sphere detection," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3275–3289, Sep. 2015.

[31] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *Proc. IEEE Int. Conf. Commun. (ICC'93)*, Geneva, Switzerland, May 1993, vol. 2, pp. 1064–1070.

[32] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, Oct. 1996.

[33] T. May, H. Rohling, and V. Engels, "Performance analysis of Viterbi decoding for 64-DAPSK and 64-QAM modulated OFDM signals," *IEEE Trans. Commun.*, vol. 46, no. 2, pp. 182–190, Feb. 1998.

[34] R. Fischer, L. Lampe, and S. Muller-Weinfurtner, "Coded modulation for noncoherent reception with application to OFDM," *IEEE Trans. Veh. Technol.*, vol. 50, no. 4, pp. 910–919, Jul. 2001.

[35] K. Ishibashi, H. Ochiai, and R. Kohno, "Low-complexity bit-interleaved coded DAPSK for Rayleigh-fading channels," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 9, pp. 1728–1738, Sep. 2005.

[36] D. Liang, S. X. Ng, and L. Hanzo, "Soft-decision star-QAM aided BICM-ID," *IEEE Signal Process. Lett.*, vol. 18, no. 3, pp. 169–172, Jan. 2011.

[37] C. Xu, D. Liang, S. X. Ng, and L. Hanzo, "Reduced-complexity noncoherent soft-decision-aided DAPSK dispensing with channel estimation," *IEEE Trans. Veh. Technol.*, vol. 62, no. 6, pp. 2633–2643, Feb. 2013.

[38] L. Wang and L. Hanzo, "Low-complexity near-optimum multiple-symbol differential detection of DAPSK based on iterative amplitude/phase processing," *IEEE Trans. Veh. Technol.*, vol. 61, no. 2, pp. 894–900, Jan. 2012.

[39] B. Hughes, "Differential space-time modulation," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2567–2578, Nov. 2000.

[40] B. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2041–2052, Dec. 2000.

[41] A. Shokrollahi, B. Hassibi, B. Hochwald, and W. Sweldens, "Representation theory for high-rate multiple-antenna code design," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2335–2367, Sep. 2001.

[42] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 7, pp. 1169–1174, Jul. 2000.

[43] H. Jafarkhani and V. Tarokh, "Multiple transmit antenna differential detection from generalized orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2626–2631, Sep. 2001.

[44] X. G. Xia, "Differentially en/decoded orthogonal space-time block codes with APSK signals," *IEEE Commun. Lett.*, vol. 6, no. 4, pp. 150–152, Apr. 2002.

[45] C. S. Hwang, S. H. Nam, J. Chung, and V. Tarokh, "Differential space time block codes using nonconstant modulus constellations," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2955–2964, Nov. 2003.

[46] M. Bhatnagar, A. Hjorungnes, and L. Song, "Precoded differential orthogonal space-time modulation over correlated Ricean MIMO channels," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 2, pp. 124–134, Apr. 2008.

[47] M. Bhatnagar, A. Hjorungnes, and L. Song, "Differential coding for non-orthogonal space-time block codes with non-unitary constellations over arbitrarily correlated Rayleigh channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 3985–3995, Aug. 2009.

[48] B. Hassibi and B. Hochwald, "Cayley differential unitary space-time codes," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1485–1503, Jun. 2002.

[49] R. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.

[50] S. Sugiura, S. Chen, and L. Hanzo, "Coherent and differential space-time shift keying: A dispersion matrix approach," *IEEE Trans. Commun.*, vol. 58, no. 11, pp. 3219–3230, Nov. 2010.

[51] C. Xu, L. Wang, S. Ng, and L. Hanzo, "Multiple-symbol differential sphere detection aided differential space-time block codes using QAM constellations," *IEEE Signal Process. Lett.*, vol. 18, no. 9, pp. 497–500, Sep. 2011.

[52] S. Sugiura, C. Xu, S. X. Ng, and L. Hanzo, "Reduced-complexity coherent versus non-coherent QAM-aided space-time shift keying," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3090–3101, Nov. 2011.

[53] Y. Bian, X. Cheng, M. Wen, L. Yang, H. Poor, and B. Jiao, "Differential spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 64, no. 7, pp. 3262–3268, Jul. 2015.

[54] N. Ishikawa and S. Sugiura, "Unified differential spatial modulation," *IEEE Wireless Commun. Lett.*, vol. 3, no. 4, pp. 337–340, Aug. 2014.

[55] P. Martin, "Differential spatial modulation for APSK in time-varying fading channels," *IEEE Commun. Lett.*, vol. 19, no. 7, pp. 1261–1264, Jul. 2015.
[56] M. Bhatnagar and A. Hjorungnes, "Decoding of differential OSTBC with non-unitary constellations using multiple received data blocks," in *Proc. IEEE Int. Conf. Commun. (ICC'10)*, May 2010, pp. 1–5.
[57] R. Schober and L. Lampe, "Noncoherent receivers for differential space-time modulation," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 768–777, May 2002.
[58] V. Pauli and L. Lampe, "Tree-search multiple-symbol differential decoding for unitary space-time modulation," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1567–1576, Aug. 2007.
[59] X. Dong, N. Beaulieu, and P. Wittke, "Error probabilities of two-dimensional M-ary signaling in fading," *IEEE Trans. Commun.*, vol. 47, no. 3, pp. 352–355, Mar. 1999.
[60] L. Lampe and R. Fischer, "Comparison and optimization of differentially encoded transmission on fading channels," in *Proc. Int. Symp. Power Line Commun. Appl. (ISPLC'99)*, 1999, pp. 107–113.
[61] W. Koch and A. Baier, "Optimum and sub-optimum detection of coded data disturbed by time-varying intersymbol interference," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM'90)*, San Diego, CA, USA, Dec. 1990, vol. 3, pp. 1679–1684.
[62] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas*, 2nd ed. Princeton, NJ, USA: Princeton Univ. Press, 2009.
[63] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. Cambridge, MA, USA: MIT Press, 2001.
[64] M. A. Bender, M. Farach-Colton, and M. Mosteiro, "Insertion sort is O(n log n)," *Theory Comput. Syst.*, vol. 39, no. 3, pp. 391–397, 2006.
[65] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: Model and erasure channel properties," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2657–2673, Nov. 2004.

Chao Xu (S'09–M'14) received the B.Eng. degree from Beijing University of Posts and Telecommunications, Beijing, China, and the B.Sc.(Eng.) (with first class Hons.) from Queen Mary University of London, London, U.K. (through a Sino-UK joint degree program), in 2008, both in telecommunications engineering with management, the M.Sc. degree (with distinction) in radio frequency communication systems, and the Ph.D. degree in wireless communications from the University of Southampton, Southampton, U.K., in 2009 and 2015, respectively. He is currently a Postdoctoral Researcher working with Southampton Wireless Group, University of Southampton. His research interests include reduced-complexity MIMO design, noncoherent detection, extrinsic-information-transfer-chart-aided turbo detection, and co-operative communications. He was awarded the Best M.Sc. Student in Broadband and Mobile Communication Networks by the IEEE Communications Society (United Kingdom and Republic of Ireland Chapter) in 2009. He was the recipient of the 2012 Chinese Government Award for Outstanding Self-Financed Student Abroad.

Li Wang (S'09–M'10) was born in Chengdu, China, in 1982. He received the Ph.D. degree from the University of Southampton, Southampton, U.K., in 2010. From 2010 to 2012, he conducted research as a Senior Research Fellow with the School of Electronics and Computer Science, University of Southampton. During his academic period, he was involved in a number of projects, such as those from Mobile VCE and Indian-UK Advanced Technology Centre (IU-ATC). In March 2012, he joined the R&D Center of Huawei Technologies in Stockholm, Sweden, and is currently working as a Principle Engineer in both radio transmission technology and radio resource management areas. He has authored 36 research papers in IEEE/IET journals and conferences, and also coauthored one JohnWiley/IEEE Press book. His research interests include wireless communications, including PHY layer modeling, link adaptation, cross-layer system design, multicarrier transmission, MIMO techniques, CoMP, channel coding, multiuser detection, noncoherent transmission techniques, advanced iterative receiver design, and adaptive filter. He is currently conducting pioneering cross-discipline researches to build next-generation communication systems with artificial intelligence. He was the recipient of Huawei Individual Contribution Award in 2015.

Soon Xin Ng (S'99–M'03–SM'08) received the B.Eng. degree (first class) in electronic engineering and the Ph.D. degree in telecommunications from the University of Southampton, Southampton, U.K., in 1999 and 2002, respectively. From 2003 to 2006, he was a Postdoctoral Research Fellow working on collaborative European research projects known as SCOUT, NEWCOM, and PHOENIX. Since August 2006, he has been a Member of Academic Staff with the School of Electronics and Computer Science, University of Southampton. He was involved in the OPTIMIX and CONCERTO European projects as well as the IU-ATC and UC4G projects. He is currently an Associate Professor of telecommunications with the University of Southampton. His research interests include adaptive coded modulation, coded modulation, channel coding, space-time coding, joint source and channel coding, iterative detection, OFDM, MIMO, co-operative communications, distributed coding, quantum error correction codes, and joint wireless-and-optical-fibre communications. He is currently working on an EPSRC project on "Cooperative Classical and Quantum Communications Systems." He has authored over 190 papers and coauthored two John Wiley/IEEE Press books in this field. He is a Chartered Engineer and a Fellow of the Higher Education Academy in the U.K.

Lajos Hanzo (M'91–SM'92–F'04) received the degree in electronics in 1976 and the Doctorate degree in 1983. In 2009, he was awarded an honorary doctorate by the Technical University of Budapest, Budapest, Hungary, while in 2015 by the University of Edinburgh, Edinburgh, U.K. During his 40-year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He is an FREng, an FIET, and a Fellow of EURASIP. He has 24,000 citations. He has successfully supervised about 100 Ph.D. students, coauthored 20 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10,000 pages, published 1500+ research entries at the IEEE Xplore, acted both as the TPC and the General Chair of IEEE conferences, presented keynote lectures and has been awarded a number of distinctions. Currently, he is directing a 60-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) U.K., the European Research Council's Advanced Fellow Grant, and the Royal Society's Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. His research is funded by the European Research Council's Senior Research Fellow Grant. He is also a Governor of the IEEE VTS. From 2008 to 2012, he was the Editor-in-Chief of the *IEEE Press* and also a Chaired Professor with Tsinghua University, Beijing, China.