IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier xx.xxxx/ACCESS.xxxx.DOI

Far-End Crosstalk Mitigation for Future Wireline Networks Beyond G.mgfast: A Survey and An Outlook

YANGYISHI ZHANG, RONG ZHANG, JIANKANG ZHANG, TONG BAI, ANAS F. AL RAWI, MARC MOONEN, and LAJOS HANZO.

The financial support of the EPSRC projects EP/N004558/1, EP/PO34284/1, COALESCE, as well as of the European Research Council's Advanced Fellow Grant QuantCom and the Royal Society Grant IF170002 and of the Global Research Challenge Grant is gratefully acknowledged. The DSL channel measurements of Fig. 4 are kindly provided by British Telecom (BT).

ABSTRACT With the escalating demand for high speed, high reliability, low latency, low cost and ubiquitous connectivity, the telecommunications industry is entering a new era where the ultimate optimality of the current wireline-wireless access network has to be achieved. Regarding the current wireline network paradigm, dominated by the copper-based digital subscriber lines (DSL) technology, multi-Gigabit data rate is the ambitious design objective at the customer end for the forthcoming ITU-T G.mgfast standard. In order to prepare for the new challenges in the era of total network convergence, both the wireline and the wireless community must be able to think beyond their respective conventions and learn from each other if necessary. Overall, the current DSL-based wireline network architecture is prone to the mutual interference resulting in far-end crosstalk (FEXT). The newly expanded 424/848 MHz spectrum of the ambitious G.mgfast project introduces far higher FEXT than that over the current 212/30 MHz G.fast/VDSL2 band. Additionally, the coexistence of multiple standards will also cause 'alien' FEXT. In these cases, using the plain zero forcing precoding (ZFP) will no longer attain a sufficiently high performance. However, as shown in the field of wireless communications, using lattice reduction as a signal space remapping technique significantly improves the performance of traditional multi-user detectors (MUD) and of the respective multi-user precoders (MUP). These promising techniques have largely remained unexploited in commercial wireless communications, due to their complexity in the face of the rapidly fluctuating wireless channels.
In this survey, we present an overview of the state-of-the-art in wireline access network and an outlook for recent technological advances in the holistic 'wireline+wireless' access network in the context of network convergence, focusing on the dominant challenge of FEXT mitigation in future DSL networks. Against this background, we investigate both the family of linear precoding and of the Tomlinson-Harashima precoding (THP) schemes conceived for classic DSL. Furthermore, we present a tutorial on the family of lattice reduction aided MUPs, as well as quantifying their expected performances in realistic DSL scenarios. As a by-product, we also demonstrate the duality between MUP and MUD, in the hope that the fifty years' history of MUD could be used to accelerate the development of efficient near-optimal MUPs for future DSL. Under the recommended operating conditions of the 212 MHz profile of G.fast, our benchmark comparisons indicate that the lattice reduction aided techniques are very powerful compared to conventional schemes. In particular, their good performances do not rely on optimized spectrum balancing, and they are also shown to be relatively robust against channel state information (CSI) estimation error, under the assumption of perfectly time-invariant DSL channels.

INDEX TERMS wireline access network, network convergence, fibre-to-the-X, G.fast, G.mgfast, far-end crosstalk, multi-user precoding, zero forcing, Tomlinson-Harashima, lattice reduction, vector perturbation, sphere encoder, dynamic spectrum management.

LIST OF ACRONYMS

ACM	Adaptive Coding and Modulation
AI	Artificial Intelligence
ARQ	Automatic Repeat reQuest
ATP	Aggregate Transmit Power
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
CAPEX	Capital Expenditure
CIR	Carrier-to-Interference Ratio
CO	Central Office
CP	Cyclic Prefix
CPE	Custom Premise Equipment
CR	Cyclic Redundancy
CRB	Cramer-Rao Bound
CSI	Channel State Information
CT	Communication Technology
CVP	Closest Vector Problem
CVPP	Closest Vector Problem with Preprocessing
DFE	Decision Feedback Equalization
DMT	Discrete Multitone
DP	Distribution Point
DPC	Dirty Paper Coding
DSB	Dynamic Spectrum Balancing
DSL	Digital Subscriber Lines
DSM	Dynamic Spectrum Management
DSP	Digital Signal Processing
EC	Echo Cancellation
EZF	Extended Zanatta-Filho
FD	Full Duplexing
FDE	Frequency Domain Equalizer
FEXT	Far-End Crosstalk
FSD	Fixed-complexity Sphere Decoder
FTTC	Fibre to the Cabinet
FTTdp	Fibre to the Distribution Point
FTTH	Fibre to the Home
FTTX	Fibre to the X
GFDM	Generalized Frequency Domain Multiplexing
HKZ	Hermite-Korkine-Zolotareff
IDFT	Inverse Discrete Fourier Transform
ISDN	Integrated Service Digital Network
ICI	Inter-Carrier Interference
IFP	Integer Forcing Precoding
IN	Impulsive Noise
ISI	Inter-Symbol Interference
ISP	Internet Service Provider
IT	Information Technology
ITU-T	International Telecommunication Union Telecommunication Standardization Sector
IWF	Iterative Water-Filling
LDPC	Low Density Parity Check
LLL	Lenstra-Lenstra-Lovász
LLR	Log Likelihood Ratio
LLU	Local Loop Unbundling
LMS	Least Mean Square
LR-ZFP	Lattice-Reduction-aided Zero-Forcing Precoding
LR-THP	Lattice-Reduction-aided Tomlinson-Harashima Precoding
LRMUP	Lattice-Reduction-aided Multi-User Precoding
MAP	Maximum A Posteriori
MIMO	Multiple-Input-Multiple-Output
MLD	Maximum Likelihood Detection
MMSE	Minimum Mean Square Error
MUD	Multi-User Detection
MUP	Multi-User Precoding
NEXT	Near-End Crosstalk
NFV	Network Function Virtualization
OFDM	Orthogonal Frequency Division Multiplexing
ONU	Optical Network Unit
OSB	Optimal Spectrum Balancing
PAM	Pulse Amplitude Modulation
PAPR	Peak-to-Average Power Ratio
POTS	Plain Old Telephone Service
PSD	Power Spectral Density
QAM	Quadrature Amplitude Modulation
QoE	Quality of Experience
QoS	Quality of Service
RCoF	Reverse Compute-and-Forward
RFI	Radio Frequency Interference
RS	Reed-Solomon
SDN	Software Defined Networking
SER	Symbol Error Rate
SIC	Successive Interference Cancellation
SINR	Signal-to-Interference-and-Noise Ratio
SIVP	Shortest Independent Vector Problem
SLNR	Signal-to-Leakage-and-Noise Ratio
SNR	Signal-to-Noise Ratio
SQRD	Sorted QR Decomposition
SRA	Seamless Rate Adaptation
SSB	Static Spectrum Balancing
SVP	Shortest Vector Problem
TDD	Time-Division Duplexing
TDSL	Terabit DSL
THP	Tomlinson-Harashima Precoding
TxPSD	Transmit Power Spectral Density
URLLC	Ultra Reliable Low Latency Communications
V-BLAST	Vertical Bell-Lab Space Time
VCE	Vectoring Control Entity
VF	Vectoring Feedback
VMR	Vectoring Mapping Region
VP	Vector Perturbation
WZC	Wyner-Ziv Coding
ZFP	Zero-Forcing Precoding

I. INTRODUCTION

A. THE FIXED-MOBILE CONVERGENCE

With the arrival of the next generation cellular wireless standard and the emerging of the Internet of Things, the evolution of the communications network has reached a critical point, where the requirement for high speed, high reliability, low latency, low cost and ubiquitous connectivity requires seamless universal convergence of the currently fragmented network infrastructure. The universal convergence entails a multi-dimensional [1] overhaul, spanning the entire existing communications ecosystem that aims for bridging the fragmented sections of conventional networks, such as copper and fibre in fixed wireline broadband [2], as well as the cloud and the edge in wireless cellular broadband [3]. Furthermore, the convergence is expected to ultimately lead to a seamless end-to-end system from wireline to wireless [4]–[7]. On the other hand, converging the network infrastructure demands a corresponding convergence in the solution domain as well. Based on information theory, a seamless integration of information technology (IT) and communication technology (CT) is anticipated to be a future-proof solution for the telecommunications community [8]. The convergence of IT-CT may be facilitated by machine learning techniques [9]–[11] and software defined networking (SDN) [12]–[15] for an intelligent and flexible next-generation network architecture.

The universal convergence of communication networks means that the performance bottleneck between each network-segment has to be eliminated for approaching the holistically optimal network performance. These requirements impose challenges on the future generations of both the wireline and the wireless broadband networks. In particular, the wireline network is typically tasked with high-rate ultra-reliable communication that covers significantly longer range than a wireless cellular network. The investigations conducted in [16] [2] have explicitly demonstrated that the next generation wireline access network, providing distributed gateways and backhaul for mobile devices will be responsible for the majority of future access network traffic. Therefore, optimizing the performance of the next generation wireline network is extremely important both for wired broadband access itself, and also for other indoor wireless networks such as Wi-Fi and visible light communication systems.

B. OVERVIEW OF THE WIRELINE BROADBAND

The family of wireline access network implementations based on the digital subscriber lines (DSL) technology has been dominating the global fixed broadband market [22] since the début of the original integrated service digital network (ISDN) in the 1980s. As a flexible technology, DSL is constantly evolving to meet the escalating demand for high-speed ultra-reliable communications. Initially deployed over the established plain old telephone service (POTS) network to exploit vacant baseband spectrum for Internet services, the wireline broadband network has developed through multiple generations distinguished by their performance metrics and architecture. According to the International Telecommunication Union (ITU), there are four main DSL generations (Fig. 1) which we can classify into two eras:

1) Legacy DSL Era

Prior to the introduction of the optical network unit (ONU), the first two universally defined generations of the DSL standards, HDSL (ITU-T G.991.x series) and ADSL (ITU-T G.992.x series), entirely relied on copper-based local telephony loops from the central office (CO) all the way to the end users' customer premise equipment (CPE). Compatibility with the POTS was not achieved until the standardization of discrete multi-tone (DMT) modulation in the ADSL series. DMT initialized the trend of spectrum expansion for all future DSL generations in order to attain ever-increasing data rates. However, due to the high propagation loss of the local loops, the legacy DSL architecture can only support a total data rate of 25 Mbps, covering a maximum radius of 1.5 km from the CO [23], which is considerably lower than the average requirement of 50-70 Mbps for typical ultra-high-definition TV streaming (using ITU-T H.265 encoding [24]). As a result, the legacy DSL technologies experienced a considerable decline in their global market share, except for Africa [22].

2) Fibre to the X (FTTx) Era

The development of optical fibre technology has revolutionized the conventional wireline access networks. Utilization of the intermediate data relay sites known as the ONUs, which connect to the CO via fibre, significantly boosts both the data rate and reach of wireline broadband beyond the limit of the legacy DSL. The current wireline broadband ecosystem can generally be depicted as a coexistence of multiple fibre-copper hybrid deployments known as fibre-to-the-X (FTTx) (Fig. 2).

In general, due to the superior reliability of optical fibre over conventional copper loops for long-distance communications, the ONU is gradually moving closer to the CPEs, resulting in a shift from the fibre-to-the-cabinet (FTTC) based VDSL (ITU-T G.993.x series) to the fibre-to-the-distribution-point (FTTdp) based G.fast (mainly ITU-T G.9701 [25]). In VDSL2, a total data rate of up to 200 Mbps with 2.5 km coverage has been achieved, whereas G.fast supports a total data rate of up to 2 Gbps for 100m loops and 200 Mbps for a maximum of 500m coverage.

In preparation for the multi-Gigabit requirement of the next generation access network, the general performance of wireline access networks should be fibre-like, which has led to massive fibre-to-the-home (FTTH) roll-out in developing countries such as China [22]. However, as investigated in [26], FTTH incurs considerably more capital expenditure (CAPEX) than the 'FTTdp + G.fast' hybrid model, even though the latter also requires an extensive deployment of distribution points (DP) due to the reduced coverage of G.fast. The rate versus cost trade-off depicted in Fig. 2 characterizes the average case over different geotypes [26], where in rural and suburban areas a more significant cost reduction can be achieved by using FTTdp. Therefore, FTTH deployment is highly situational and it is in general not an economical solution. Furthermore, the advent of powerful communication processors [27] and network function virtualization (NFV) [28] creates the possibility of reaching fibre-like performances using the existing FTTx architecture. The proposition of Cioffi in [29] suggested that new transmission modes in the millimetre-wave band accompanied

TABLE 1: A comparison of surveys on the development of DSL wireline network

Research Area	[17]	[18]	[19]	[20]	[21]	Our Work
Network Architecture	Legacy xDSL	Legacy xDSL	G.fast	G.mgfast	G.mgfast	Future
Channel Modelling	✓		✓			✓
Cost Assessment				✓		✓
Signalling Protocol			✓		✓	✓
Vectored Transmission	✓		✓		✓	✓
Spectrum Balancing		✓	✓			✓

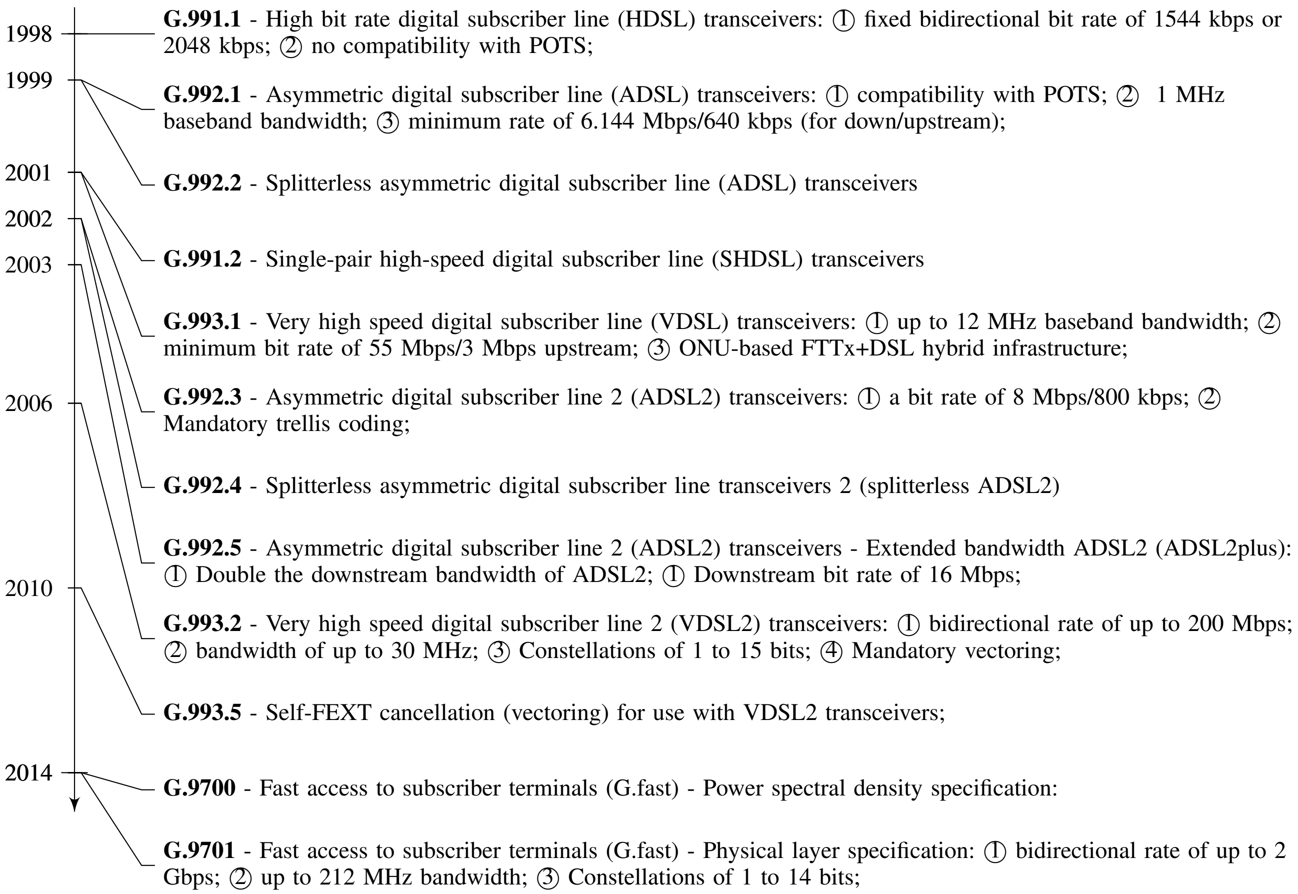

FIGURE 1: Timeline of the ITU-T DSL standards by date of initial publication. Based on the performance and architecture difference, there are four generations consisting of the HDSL (G.991 series), ADSL (G.992 series), VDSL (G.993 series) and G.fast (G.970x series).

by enhanced digital signal processing (DSP) techniques can potentially achieve Terabit performance target (i.e. TDSL) using the DSL framework.

C. OUTLINE

In this survey, we aim to provide an overview of candidate solutions for the future physical layer architecture of the next generation DSL, i.e. G.mgfast, and beyond. Additionally, it is widely recognized that the performance of the current G.fast broadband network is predominantly impaired by crosstalk, which is going to affect G.mgfast even worse due to the channel's frequency response characteristics. In this regard, we will also offer a survey of the state-of-the-art algorithms with respect to the mandatory downstream vectoring in legacy and current DSL standards.

In the spirit of encouraging further research into combating the hostile crosstalk-infested channel environment and therefore providing high performance Internet services in the converged access network era, we firstly review the multi-user precoding (MUP) vs. the multi-user detection (MUD) duality, in the hope that the fifty years' worth of literature of multi-user detection [32] will accelerate the development of powerful and efficient vectoring techniques. Secondly, following the unique stability characteristics of the DSL channels, we offer a tutorial on the family of powerful lattice reduction aided precoding techniques (in wireless

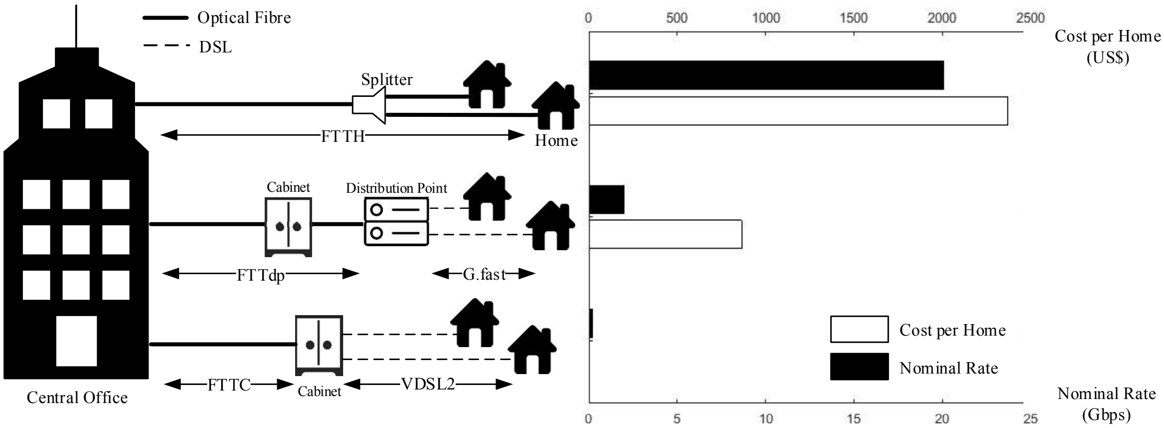

FIGURE 2: Nominal aggregate 'upstream+downstream' data rate and deployment cost for each FTTx-type hybrid wireline broadband architecture, showing the coexistence of FTTH (top), 'FTTdp + G.fast' (middle), 'FTTC + VDSL2' (bottom). An universally established VDSL2 environment, known as 'brownfield' [26] is assumed. The nominal rates are from their respective ITU-T standard [25], [30], [31].

systems), which specifically exploit the quasi-static nature of the channels. Finally, we present a survey of dynamic spectrum balancing (DSB), which is an important historical approach to crosstalk mitigation. Moreover, we conclude this survey by providing comparative simulation results for the most relevant benchmark algorithms, in addition to the practical lessons we learned about crosstalk mitigation in DSL systems. The outline of this survey is as follows.

I **Introduction**
- I-A The Fixed-Mobile Convergence
- I-B Overview of the Wireline Broadband
- I-C Outline
- I-D Novel Contributions

II **The New DSL**
- II-A Channel Characteristics
- II-B System Architecture

III **Multi-User Precoding in DSL**
- III-A The MUP-MUD Duality
- III-B Linear Precoding
- III-C Tomlinson-Harashima Precoding

IV **Lattice Reduction aided MUP**
- IV-A Lattices in Telecommunications
- IV-B Approximate Lattice Precoding
- IV-C Integer Forcing Precoding
- IV-D Vector Perturbation

V **Spectrum Balancing for Vectored Transceivers**
- V-A System Model
- V-B Spectrum Balancing Algorithms

VI **Benchmark Comparisons**
- VI-A Level-3 MUP Performance
- VI-B Multi-Level DSM Performance

VII **Practical Lessons**
- VII-A Complexity vs. Performance
- VII-B The Near-Far Problem
- VII-C Implementation Issues

VIII **Conclusions**

D. NOVEL CONTRIBUTIONS

Due to the particular nature of this treatise, we summarise our tutorial contributions to the DSL literature in Tab. 1. On the other hand, our contributions to the lattice-based communications literature, as well as to the practical application of LR in wireline networks are respectively presented in Tab. 2 and as follow:

- A novel duality between MUP and MUD techniques from a lattice theoretical perspective is conceived for evaluating the performance of emerging MUP techniques based on the rich literature of MUD;
- With the aid of our novel duality principle, the potential of the family of LRMUPs in future commercial DSL networks is characterized in terms of their performances and practical constraints;

II. THE NEW DSL

A. CHANNEL CHARACTERISTICS

Despite the fact that copper-based communication systems have had a long history within the telecommunications industry thanks to the popularization of POTS and DSL, the understanding of the twisted-wire channel is still limited. Due to the renewed interest in boosting the performance of

TABLE 2: A comparison of recent research on lattice-based communications.

Research Area	[33], [34]	[35]	[36]	[37]	[38], [39]	[40], [41]	[42]	Our Work
Lattice Reduction Algorithm	✓							
LR-aided Classic Precoding			✓				✓	✓
Constellation Shaping				✓	✓			
Sphere Encoder		✓			✓	✓		✓
Adaptive Modulation						✓	✓	✓
Power Allocation		✓			✓		✓	✓

DSL, alternative signal propagation modes over twisted copper are being investigated by both the research community [29], [43]–[45] and the standardization body [46], such as the phantom mode and the waveguide mode. Although the quasi-static nature of a DSL channel is widely assumed in the wireline communications literature, the modelling of its transfer characteristics is far from trivial, especially with the discovery of new operational modes. This section is dedicated to a comparative overview of the transfer characteristics of the copper channel for cutting-edge DSL.

1) Direct Channel

The premise of conventional twisted-copper communication methods is the two-port network model relying on the classic transmission line theory. In the simplest case, a pair of twisted copper wires, i.e. a copper pair, carrying differential voltages constitutes a direct channel[1]. Given the length of the copper pair d, as well as the propagation constant $\gamma(f)$ of the two-port network as a function of the differential voltage's frequency f, then the transfer function $H(f;\, d)$ of this single-pair channel can be formulated as:

$$H(f;\, d) = e^{-d\gamma(f)}, \tag{1}$$

if the pair is perfectly terminated. In general, $H(f;\, d)$ is a gradually decaying function of both f and d^2. Many existing DSL channel models proposed for the VDSL/VDSL2 band are inapplicable both to G.fast and to the next generation G.mgfast band [47], mainly because they do not consider the substantial change of the twisted pair's self-coupling characteristics at high frequencies, where the signal wavelength and the copper pair's twist length become comparable. In practice, these (average-case) channel models are not suitable candidates for the network operator. Because they do not exactly match the actual channel transfer function of particular copper pairs, they cannot be used to assess the critical worst-case performance and they cannot be mitigated by time-diversity[3]. For this reason, channel measurements are also extensively used by the DSL research community in order to acquire realistic performance estimates.

[1] In this survey, we make the realistic assumption that all connections other than those directly associated with CPEs consist of optical fibre as seen in Fig. 2, therefore the multi-segment DSL case discussed in [17] can be skipped.

[2] The length of a copper wire is constant after being placed, therefore d is usually considered a 'parameter' of the transfer function.

[3] Time-diversity is a common phenomenon in wireless communications where the severity of the worst-case performance is naturally mitigated due to the time dependence of wireless channels.

2) Crosstalk Channel

In areas that are close to the DP, the copper pairs connecting CPEs from different houses are bundled together as a large multi-pair DSL binder. However, because the POTS network was originally designed for carrying 3.4 kHz voice signals, the high-frequency signals of DSL may cause considerable electromagnetic leakage and unintended coupling between different pairs within a DSL binder. Due to the binder sheath which reduces the alien interferences from outside sources, the unintended coupling inside a binder becomes the dominant source of multi-user interference and constitutes the crosstalk channel. In particular, the forward interference travelling in the same direction as the direct channel signal is termed the far-end crosstalk (FEXT), whereas the returning interference travelling in the opposite direction of the direct channel signal is termed the near-end crosstalk (NEXT) (Fig. 3). We note that there are in fact many other external sources of interference in multi-pair DSL, most of which are however neither measurable nor static. Therefore, they do not constitute interfering channels and are more widely classified as noise, with the rare exception of alien crosstalk both from other DSL binders [48]–[50] and from the co-existing power line communication systems [51] [52].

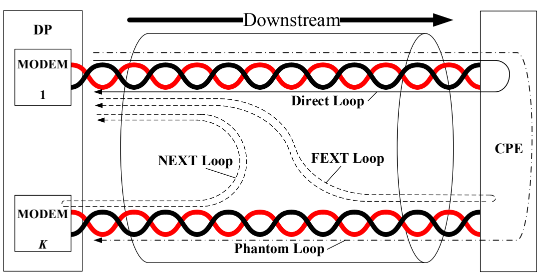

FIGURE 3: Virtual local loop illustration of DSL channels operating in transmission line mode. Both the forward path and the return path are shown for each virtual loop. Signals transmitted by modem 1 are received at the end of each virtual loop's forward path. The CPE must be a single entity or a set of coordinated entities in order to utilize the phantom loop.

The mathematical modelling of crosstalk channels is considerably different from that of direct channels. In general, the coupling characteristics between adjacent copper pairs is not universally deterministic, subject to the exact geometry of the DSL binder's interior, as well as to the dielectric behaviour of the binder's sheath. Hence, the frequency response of crosstalk channels is commonly modelled by stochastic functions [53] [54], which can then be used to generate a particular crosstalk channel realization. Regardless, general practical bounds were established using the '1% worst-case' power spectral density (PSD) model for FEXT and NEXT[4], formulated respectively as [55]:

$$|H_F(f;\, d)|^2 = |H(f;\, d)|^2 (\frac{K-1}{49})^{0.6} (9 \times 10^{-20}) d \cdot f^2 \,, \tag{2}$$

$$|H_N(f;\, d)|^2 = (\frac{K-1}{49})^{0.6} \cdot 10^{-13} f^{1.5} \,, \tag{3}$$

where it is shown that the PSD of crosstalk channels increases with the number of encapsulated copper pairs K. For FEXT, the interfering signal must travel the full length of the DSL binder, whereas NEXT is a form of returning interference, as seen in Fig. 3. In general, the gain of the NEXT channel diminishes with the distance from the transmitter. Eq. (3) represents the integral of this diminishing returning interference over the full length d of the DSL binder, and thus the effect of d is eliminated in the model. On the other hand, the channel gain of FEXT over the direct channel $|H_F(f;\, d)|^2/|H(f;\, d)|^2$, known as the equal level FEXT [56], is shown to increase with the frequency in Eq. (2), implying the deterioration of multi-user interference at high frequencies (Fig. 4). We refer the readers to further references for more in-depth modelling of crosstalk channels [56]–[58]. The excessive level of FEXT at high frequencies has led to some of the most pronounced research challenges in the wireline communications community.

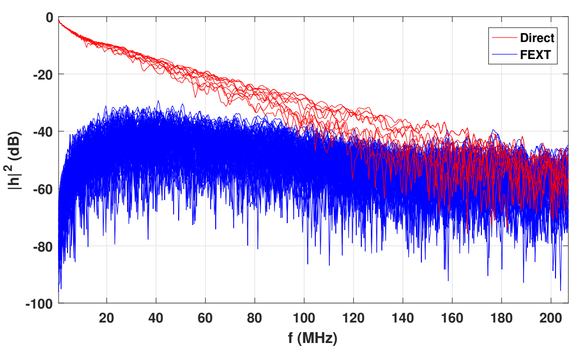

FIGURE 4: Frequency response of the direct link channels and FEXT link channels for the forthcoming 212 MHz G.fast profile of a 100-meter cable containing 10 twisted pairs.

3) Phantom Channel

The phantom and common mode signalling [44], [59], [60], which have not been considered by the standardization bodies until recently [46], relies on the well-established multi-pair bonding technology [61]–[63]. Akin to the multi-antenna transceivers of wireless communications, corporate buildings may have access to a whole binder of copper pairs for a set of coordinated CPEs or a single large CPE such as a server. It is plausible that multi-pair bonding can be used to glean multiplexing gain from the additional spatial dimensions. Hence, the main benefit of multi-pair bonding is that of obtaining additional channels beyond the direct channels and the crosstalk channels.

For a binder of K copper pairs, there are in theory K common-mode direct channels in addition to the K differential-mode direct channels, if each of the $2K$ wires is virtually paired with a different wire from another adjacent copper pair [43]. However, the proposed common mode operation [43] requires heavy modifications to the DSL binders, which makes it impractical. Instead, a better approach to exploiting the 'hidden' circuits of a multi-pair binder is the proverbial 'phantom-mode' transmission.

For a two-pair binder, there exists a third, 'phantom' channel which consists of the two existing copper pairs as its 'wires'. The phantom-mode signal equals to the voltage difference between the means of the two differential voltages carried by the two physical copper pairs [44]. For a K-pair binder, we can utilize up to $K-1$ phantom channels by pairing the physical channels. Therefore, phantom channels can be modelled as conventional transmission lines similar to the physical direct channels, and the multi-pair bonding technology may also benefit the hybrid group of channels for gleaning more multiplexing gain [64]. On the other hand, phantom channels are orthogonal to the physical channels and therefore the phantom mode does not interfere with the usual differential mode. The channel used by phantom mode is depicted in Fig. 3.

4) Surface Plasmon Polariton Channel

The state-of-the-art DSL technologies relying on channel models of the previous types have limited performance even under the G.mgfast specification, compared to the achievable rates of FTTH systems, as seen in Fig. 2. In essence, the superiority of optical fibre comes from its role as a waveguide rather than a conventional transmission line in terms of the propagation of electromagnetic waves. Fortunately, it was recently shown that metal wires may be used for signalling in the surface plasmon polariton mode at the Terahertz (THz) frequency band [65] [66]. The pioneer study of Cioffi et al. in [29] has attempted to utilize this particular waveguide-like signalling mode within the multitude of existing copper wires in DSL binders in order to realize the ultra fast TDSL and therefore align the copper access network performance with that of the fibre access networks.

Due to the complete change of signal propagation mode, the transfer characteristics associated with this waveguide channel are dramatically different from that associated with the transmission line channel, even though their transfer functions may depend on the same set of variables (1) [29].

[4]The crosstalk channel PSD predicted by this model is guaranteed to be lower than 99% of the realizations.

In the THz band, the wavelength of electromagnetic waves falls in the millimetre range, which is comparable to the radii of the existing copper pairs. Therefore, THz signals can travel along the surfaces of copper wires as well as in the surrounding free space. For free space electromagnetic waves guided by the copper wires in DSL binders, the other wires of the binder, as well as the metallic sheath if there is one, can be used as reflectors. These natural reflectors can prevent surface waves from leaking when the designated waveguide bends, and hence the total internal reflection effect of optical fibres may theoretically be recreated in TDSL. The physical modelling of these effects are more thoroughly investigated in [67]–[69].

In general, the new operational modes of DSL are not sufficiently investigated at the moment. The significant potential predicted for the somewhat distant future is far from its theoretical completeness and industrial fruition. We will not consider these new types of channels in further details for the remainder of this survey due to its scope. However, the importance of these new modes will gradually become explicit when the system models become mature.

B. SYSTEM ARCHITECTURE

The overall hostile environment of DSL access networks is the combined result of severe frequency-selectivity, intensive self-crosstalk and coloured noise. Additionally, because of the limited computation power of signal processors in the past, the suboptimal performances of legacy and current DSL standards are also largely due to inefficient exploitation of the available spectrum. The ITU-T G series established a family of supplementary technologies that are designed for improving the bandwidth efficiency and quality of service (QoS) of DSL systems, which is highlighted in Fig. 5. In this section, we will present an overview of the state-of-the-art as well as common key technologies under consideration for both the next generation wireless and future wireline access network architecture (Fig. 6). From an industrial and a commercial perspective, the expected technological advances in the forthcoming generation of wireline access networks are investigated in further details in [20] [21].

1) Modulation

From the ADSL era onwards, DSL broadband access became compatible with the POTS. The enabling technology accounting for this feature is the DMT modulation scheme [55], whose passband variant is more commonly known as orthogonal frequency-division multiplexing (OFDM) in wireless and optical communications community. Additionally, the severe frequency-selectivity of DSL ((1)(2)(3) and Fig. 4), which would otherwise cause strong inter-symbol interference (ISI), can be conveniently removed by DMT, resulting in a multitude of frequency bins having small bandwidth and negligible frequency-selectivity, which are known as the tones. Each tone carries a trellis-coded subsymbol from a designated constellation. In G.fast, the constellation is associated with a given order of quadrature amplitude

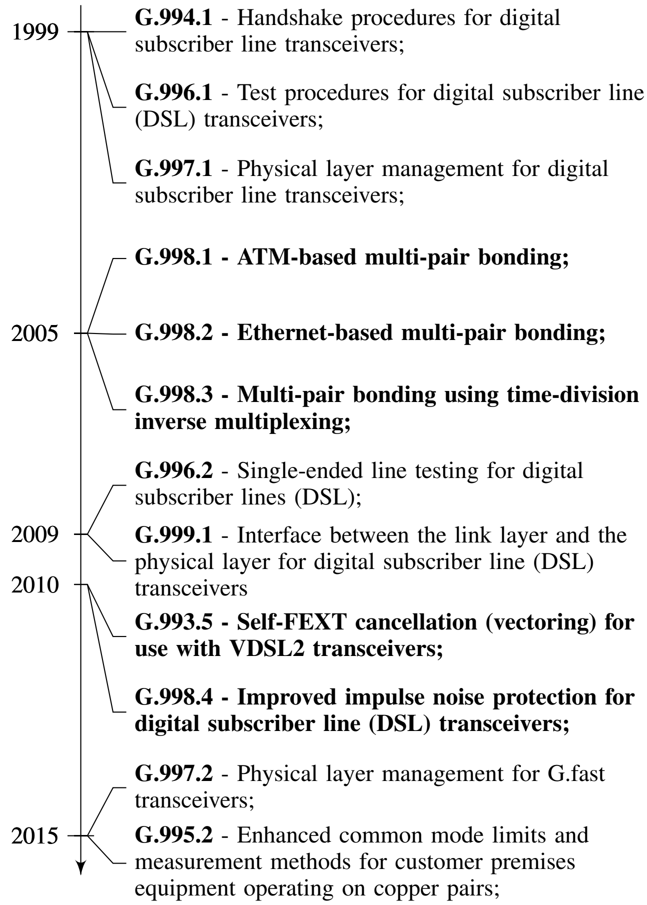

FIGURE 5: Timeline of the ITU-T DSL standards for supplementary technologies by date of initial publication. Bold entries represent technologies which are used to boost bandwidth efficiency.

modulation (QAM). Since DMT operates in the baseband, the absence of carrier-related problems such as frequency and phase offset allows a choice of up to 2^{15}-QAM [70].

In the physical layer architecture characterized by Fig. 7, the downstream transmission of a typical K-pair T-tone G.fast system is shown, assuming that each CPE is connected to the DP via a single twisted pair. The end-to-end system operates in frequency domain, whilst the transmission segment between the front ends of the DP and the CPEs is in time domain. The role of the MUP and the frequency-domain equalizers (FDE) will be discussed in the following sections (bottom of Fig. 7). In order to guarantee tone orthogonality and therefore avoiding inter-carrier interference (ICI), each time-domain subsymbol, obtained via the inverse discrete Fourier transform (IDFT) of the frequency-domain subsymbols, must be transmitted with a sufficiently long cyclic prefix (CP) attached [71]. Upon reception of the time-domain DMT symbol, the CP is removed before the remainder of the DMT symbol is transformed back into frequency domain by DFT. In principle, the duration of CP should be at least identical

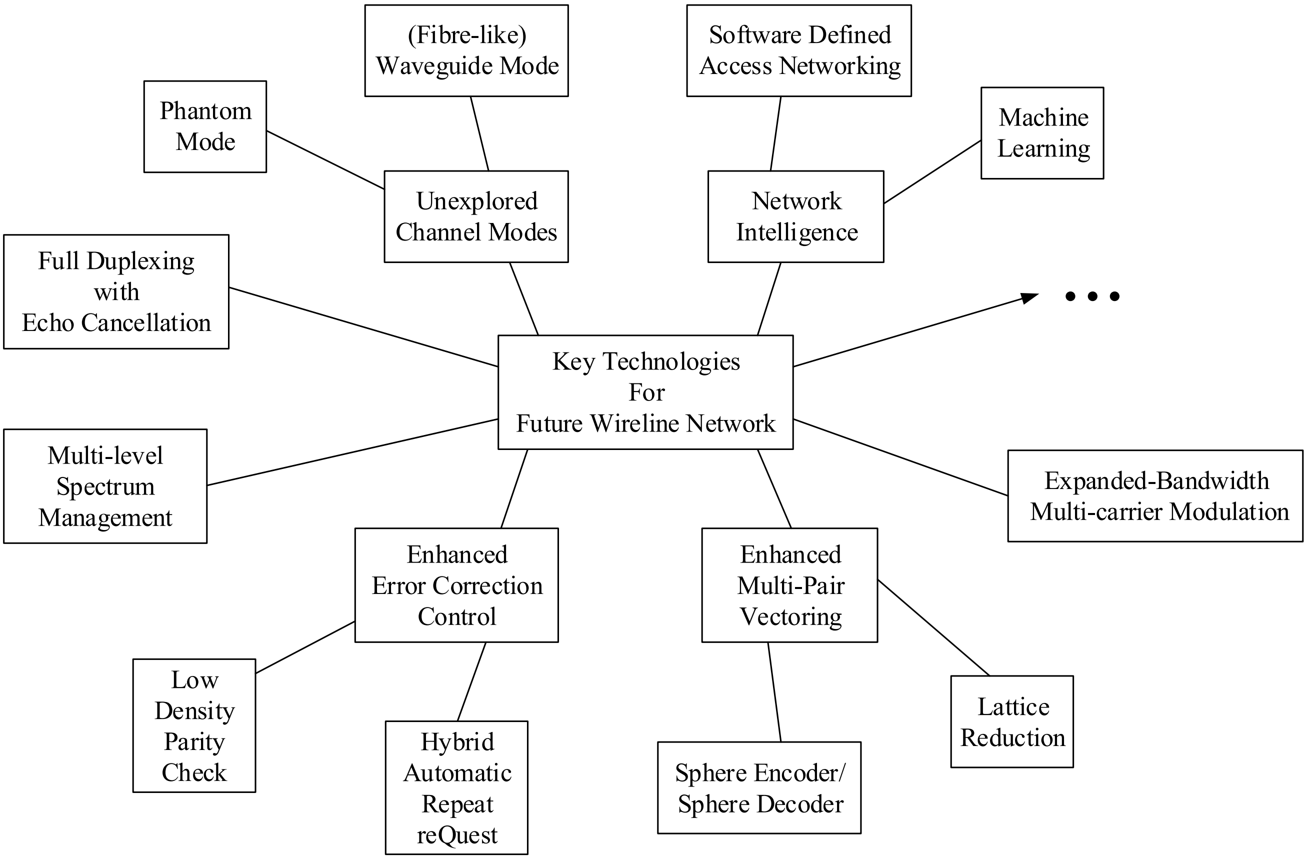

FIGURE 6: Key technologies in consideration for G.mgfast and/or future generation wireline access network.

to the DSL channel's delay spread, which is true in G.fast. Therefore, it is reasonable to assume that the top system of Fig. 7 consists of T independent layers of frequency-domain systems having K users each.

Due to the uncertainty of the channel characteristics at very high frequencies far above the current 212 MHz G.fast ceiling, e.g. for 424 or 848 MHz G.mgfast profiles, as well as to the strict power constraints over these bands[5], the enormous peak-to-average-power ratio (PAPR) of conventional DMT with long CP overhead might not be tolerable for future performance requirements. For the sake of good power efficiency, the generalized frequency-division multiplexing (GFDM) technique [73] may be used as an alternative, due to its reduced CP usage and lower implementation complexity compared to OFDM/DMT yet using a similar structure (as showcased in [74]). Furthermore, zero ICI is still achievable in GFDM if Dirichlet filters are used [75]. Nonetheless, the assumption that each frequency-domain channel is orthogonal holds valid.

[5]The power constraints of DSL are mostly enforced for electromagnetic compatibility with radio broadcast services as seen in [72], therefore we may anticipate similar conditions for the currently undefined bands.

2) Vectoring Control

Besides frequency selectivity, which is mitigated by DMT modulation, another implication of Fig. 4 is the fast decrease of the carrier-to-interference ratio (CIR), which has already caused interference problems in the wideband DSL deployment of VDSL2 and in particular G.fast. Since the modems at the ONU (cabinet or DP) are co-located, and usually coordinated by the same Internet service provider (ISP), they may invoke FEXT removal mechanism referred to as vectoring (or vectored transmission) [76], subject to the availability of channel state information (CSI). Due to the orthogonality of DMT modulated subsymbols, the system diagram at the top of Fig. 7 may be decoupled into T instances of K-user subsystems, which we formulate as:

$$\boldsymbol{y}^t = \boldsymbol{H}^t \boldsymbol{x}^t + \boldsymbol{n}^t \quad (t = 1, 2, ..., T), \tag{4}$$

where the $K \times 1$ vectors $\boldsymbol{y}^t$, $\boldsymbol{x}^t$ and $\boldsymbol{n}^t$ represent the received symbol vector, transmitted symbol vector and the noise vector of tone t. If $\boldsymbol{n}^t$ is 'white', it has the same PSD across all frequency tones. However, in practice there might be other sources of noise which are not white. The $K \times K$ multi-input-multi-output (MIMO) transfer matrix $\boldsymbol{H}^t$ is the

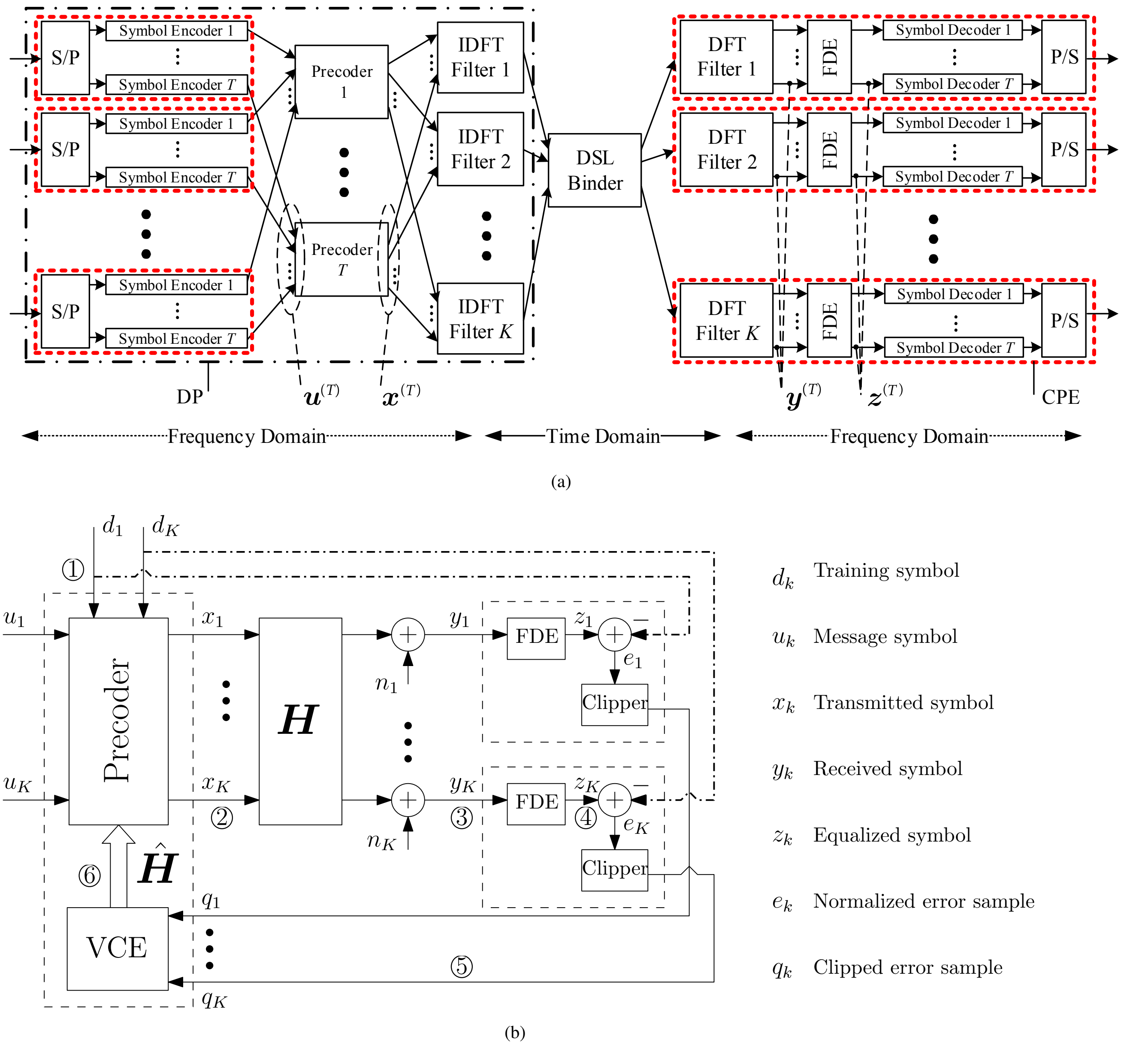

FIGURE 7: (a) Physical layer architecture of multi-user G.fast, depicting (b) the frequency domain single-tone vectoring control loop. The analogue front end is omitted (Copyright [25])

frequency domain multi-user DSL channel of the tth tone. In DSL, the diagonal entry $h^t_{k,k}$ in $\boldsymbol{H}^t$ is the direct channel gain associated with pair k, whilst the off-diagonal elements $h^t_{k,l}$ for $k \neq l$ represent the FEXT channels contaminating pair k. With the emerging of G.mgfast, the DSL network becomes more vulnerable to alien FEXT due to the coexistence of multiple standards, as well as to the typical local loop unbundling (LLU) problem investigated in [77]. The general definition of a vectored transmission group does not deal with alien FEXT. However, the power of SDN/NFV has the promise of flexible traffic management and signal coordination between multiple ISPs [14] [78] in the future, so that vectoring may be expanded to remove 'alien' FEXT (which will then become domestic).

Without loss of generality, we will consider a frequency-domain subsystem whose tone index is neglected (bottom of Fig. 7). For the vectored downstream single-tone system, FEXT is pre-cancelled by the MUP, while in the upstream counterpart the FEXT channel is equalized by the MUD, i.e. a FEXT canceller. In both cases, the ability to remove FEXT relies on the CSI knowledge. In G.fast, the vectoring control entity (VCE in Fig. 7) is responsible for obtaining the prerequisite CSI knowledge mainly for the MUP, whereas MUD configuration is vendor discretionary.

Under the specifications of [25], the operations of downstream vectoring are divided into the initialization stage for

CSI acquisition and the main operational stage. The initialization stage invokes a training-aided channel estimation technique using the vectoring feedback loop of Fig. 7. The full loop consists of the following blocks for the single-tone example:

① The training symbol vector $\boldsymbol{d} = [d_1, d_2, \cdots, d_K]^T$ is fed into the uninitialized precoder. The identity information of $\boldsymbol{d}$ is transmitted as part of the training sequence (termed as the probe sequence in [25]), therefore the training symbol vector is virtually known to the CPEs.

② The transmitted symbol vector $\boldsymbol{x}$ is normally a function of $\boldsymbol{d}$ and the estimated downstream CSI $\hat{\boldsymbol{H}}$. During initialization, the uninitialized precoder forwards $\boldsymbol{d}$, which is directly fed into the channel as the transmitted symbol vector $\boldsymbol{x}$. Hence the transmitted symbol of each customer is uncorrelated.

③ The symbol vector $\boldsymbol{y}$ received at the decentralized CPEs (right side of Fig. 7) is contaminated by the frequency-domain channel $\boldsymbol{H}$ and thus $\boldsymbol{y}$ contains information about the CSI. From each customer's perspective, each received element y_k in (4) can be rewritten as:

$$y_k = h_{k,k} d_k + \sum_{j=1, j \neq k}^{K} h_{k,j} d_j + n_k. \tag{5}$$

Each received symbol contains information both about the channel of the direct link (the first term of (5)), as well as about all FEXT channels coupled to the said link (the second term of (5)).

④ In order to identify the influence of the channel on the training symbol vector, the error between the equalized symbol vector $\boldsymbol{z}$ and the training symbol vector $\boldsymbol{d}$ is recorded as the normalized error sample vector $\boldsymbol{e} = \boldsymbol{z} - \boldsymbol{d}$. For QAM based systems, $\Re(\boldsymbol{e})$ and $\Im(\boldsymbol{e})$ are stored as separate quantities. Since the uninitialized single user equalizer w_k is not invoked, the equalized symbol vector $\boldsymbol{z}$ is identical to $\boldsymbol{y}$.

⑤ Due to the limited bandwidth of the feedback channel, $\boldsymbol{e}$ must be quantized using the quantization format defined in [25]:

$$\boldsymbol{q} = \max\left[-2^{B_{\max}}, \min\left(\lfloor \boldsymbol{e} \cdot 2^{N_{\max}-1} \rfloor, 2^{B_{\max}} - 1\right)\right], \tag{6}$$

where $N_{\max}$ and $B_{\max}$ represent the number of bits that control the quantisation step size and the maximum quantisation range of the clippers, respectively. The clipped error sample vector $\boldsymbol{q}$ is reported back to the VCE on the ONU side (left side of Fig. 7) via the feedback channel.

⑥ The VCE attempts to deduce the original received symbol vector $\boldsymbol{y}$ by reconstructing $\hat{\boldsymbol{y}} = \boldsymbol{d} + \boldsymbol{q}/2^{N_{\max}-1}$. When the VCE receives a sufficient number of clipped error sample vectors, it will be able to produce an estimate $\hat{\boldsymbol{H}}$ of the CSI $\boldsymbol{H}$. For the $K \times K$ channel matrix $\boldsymbol{H}$, at least K clipped error sample vectors (thus containing K^2 clipped error samples) are required, since no unique solution of the K^2 channel coefficients exists, if there are less than K^2 linearly independent equations (in the format of (5)).

The above vectoring feedback scheme is overall sub-optimal in terms of achieving the Cramer-Rao bound (CRB) [79] [80], as a result of the loss from quantization (6) and the feedback channel ⑤. Firstly, the training process may also be triggered at the request of the VCE afterwards in order to update the CSI estimate $\hat{\boldsymbol{H}}$ and subsequently the precoder. Therefore, the quasi-static nature of the channel matrix $\boldsymbol{H}$ may be tackled by regularly updated channel estimation [81] [82]. Secondly, since G.fast currently uses time-division duplexing (TDD) to separate upstream and downstream transmissions, the VCE may exploit the channel's reciprocity [83] to acquire the downstream CSI based on the upstream CSI estimate, the latter of which trivially approaches the CRB [84]. We should note that the in-band full duplexing operation to be introduced in G.mgfast also benefits from this reciprocity.

3) Duplexing

In order to boost the bandwidth efficiency of DSL links beyond the Shannon limit of a single channel use, recent proposals [2] [85] have suggested simultaneous upstream and downstream transmissions over the same bandwidth, i.e. full duplexing (FD), for the forthcoming G.mgfast. Furthermore, FD is also a strong candidate which has been widely studied in the wireless context [86]–[88]. By definition, FD allows doubled channel use within a single DMT symbol duration. However, the capacity gain due to FD is typically less than 100% [89] [90] as a consequence of the resultant strong self-interference, which consists of signal reflection (due to imperfect receiver side impedance matching, sometimes known as the echo), as well as of NEXT in DSL systems. In particular, FD achieves 100% efficiency in both frequency and time domain at the expense of losing orthogonality between the upstream and the downstream data transmissions, resembling a typical non-orthogonal multiple access (NOMA) scenario.

Echo cancellation (EC) [91] and NEXT cancellation [92] are critical techniques for guaranteeing the performance of multi-pair FD DSL systems. Akin to FEXT cancellation, NEXT cancellation is only possible at the DP side, or for either a subgroup of coordinated CPEs or a single CPE connected to multiple twisted pairs (the upstream dual of Fig. 3). On the other hand, since EC is only associated with a single line, it does not require coordinated transmission with other transceivers. However, when the CPEs are not co-located, the CPE side NEXT is shown to be consistently lower than the DP side NEXT [2]. In general, EC/NEXT cancellation can be done in both the time and the frequency domain. The FD operation of a single G.mgfast transceiver unit (CPE or one modem in DP) is depicted in Fig. 8. Overall, cancelling the self-interference in frequency domain is simpler to implement, but the signal reception and transmission must be properly aligned in time. On the other hand, self-interference cancellation in the time domain does not require synchronized signalling, but the complexity of the time

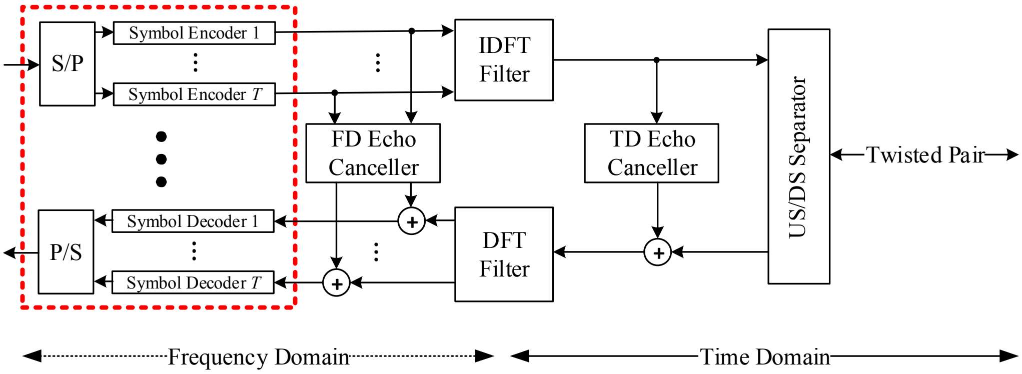

FIGURE 8: Single-pair T-tone full duplexing DSL transceiver model. EC may be done in frequency domain or time domain. The upstream and downstream transmission is separated by a special hybrid circuit. In the multi-pair case [2], the NEXT canceller is included in the EC module.

domain approach is higher. Additionally, self-interference cancellation may also be implemented in time domain at the analogue front end, if the analogue-to-digital converter becomes saturated in the face of strong self-interference. The reader is redirected to [2] for more details on the multi-pair FD design as well as for the performance study of G.mgfast, where a throughput increase of nearly 100% is observed for FD compared to TDD using the same bandwidth. In the case of coexisting G.fast/mgfast deployment, it has also been shown in [93] that a throughput increase of more than 50% is achievable.

4) Error Control

Besides alien FEXT, Impulsive noise (IN) constitutes another category of impairments, whose non-stationary nature cannot be accurately captured by the relatively static CSI. Furthermore, the measurable additive white Gaussian noise (AWGN) floor is typically at -150 dBm/Hz in DSL systems, which is considerably higher than the -174 dBm/Hz AWGN floor characterized for common cellular wireless systems. In general, the noise sources cannot be mitigated by vectoring techniques. Hence, impulsive noise protection is implemented by means of forward error correction coding and retransmission, in addition to adding noise margin in the transceiver design [94]. The bursty nature of IN and its influence on the performance of DSL systems is characterized in [95]. On the other hand, alien impairment in DSL systems such as the radio frequency interference (RFI) caused by other wireless systems in close proximity must be handled by error correction coding as well [96]. Aerial sections of a DSL network, such as drop wires, are susceptible to RF interference, acting just like antennas. Unlike vectoring, error correction coding generally follows the same rules for both upstream and downstream transmissions. The error correction paradigm for the next generation G.mgfast systems consists of the following strategies:

- **Channel Coding**: The performance of the standardized channel coding approach, relying on Reed-Solomon (RS) coding aided trellis coded modulation, is unable to satisfy the increasing demand for high QoS. Therefore, capacity-approaching coding scheme such as low density parity check (LDPC) codes are employed as the successor. Additionally, for IN protection, interleaving will remain an effective approach due to its inherent capability of dispensing bursty errors. Even though interleaving is typically applied in the frequency domain, time domain techniques may be specifically designed to combat the non-uniform spreading of time domain IN in the frequency domain upon demodulation [97].
- **Automatic Repeat reQuest (ARQ)**: When relying on transmitting cyclic redundancy (CR) along with the payload, the channel decoder becomes capable of identifying the incorrectly received DMT symbols and triggering retransmission of the same payload until either the DMT symbol in question is correctly received or the maximum number of retransmissions is reached. ARQ has demonstrated great potential in boosting the throughput of DSL systems [98]. In the future, hybrid ARQ schemes exploiting the availability of past retransmissions are also under consideration. An LDPC coded solution is specifically investigated in [99].

Given the main scope of this survey, the additional benefit of bit-level error correction will not be considered in further details. However, we should note that error control constitutes one of the three most widely recognized signal processing challenges for DSL systems, together with dynamic spectral load balancing and multi-user vectored transmission [100] (Tab. 3). In particular, error control is the base level of the wide-sense dynamic spectrum management (DSM)

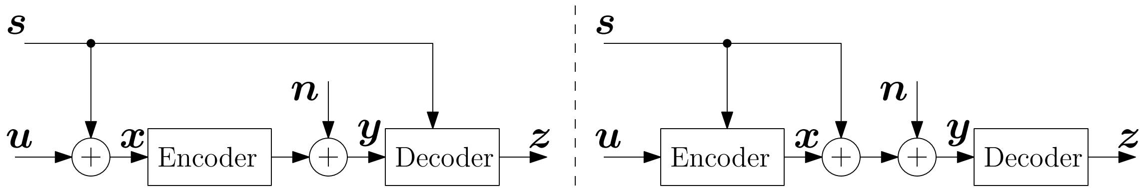

FIGURE 9: A special case of Wyner-Ziv coding (left) versus Dirty Paper coding (right).

paradigm (Tab. 3). The reader is redirected to [101] and the references therein for a more in-depth coverage of (impulsive) noise mitigation and error control in the next generation wireline networks.

TABLE 3: Dynamic Spectrum Management Level Definitions [100]

Level	Functionality
0	N/A
1	Single-pair-based Forward-error-correction Configuration (This section and also [101])
2	Multi-pair-based Spectral Load Balancing and Optimization (cf. Section V)
3	Multi-pair-based Vectored Signal Transmission and Reception (cf. Section III and IV)

5) Network Intelligence

In recent years we have seen a rapid increase in the research of artificial intelligence (AI) and machine learning, which are revolutionizing many areas in communications, such as resource management [102] and routing optimization [103] [104] in SDN-aided system architectures, where the network characteristics can be efficiently learned. Therefore, the machine learning assisted SDN constitutes an intelligent solution for the converged access network architecture [9].

However, harnessing these powerful tools in the existing multi-standard multi-ISP based DSL network paradigm requires further research. As mentioned in Section II-B2, DSL access networks are particularly susceptible to alien crosstalk both from LLU and from other binders in close proximity, in addition to the severe in-binder FEXT and NEXT. In [14], a software-defined, open access network infrastructure was proposed for mitigating the contaminating effects of LLU. Specifically, a management interface, maintained by a third party, is employed for coordinating the traffic associated with different ISPs sharing a single DSL binder. Subsequently, the third party coordinator may invoke vectoring for conveying the transmitted data across the ISPs using the same binder, subject to authentication for preventing violation of privacy.

We should note that, apart from the aforementioned macroscopic benefit of AI for the DSL-aided future access network infrastructure, AI can also be used for tackling the challenges associated with other parts of the system architecture we have discussed. In particular, evolutionary algorithm aided MUDs [105] and deep neural network assisted MUDs [106] have already been conceived for MIMO systems, in addition to the 'auto-encoder' type of end-to-end system design of [107]. However, we should note that the near-optimality of these particular applications has only been characterized for low-dimensional systems.

III. MULTI-USER PRECODING IN DSL

A. THE MUP-MUD DUALITY

It is widely acknowledged that there exists a duality between the uplink and the downlink of cellular wireless networks. In particular, if the channel reciprocity holds between the uplink and the downlink, while the sum of uplink tranmit power constraints equals the downlink transmit power constraint, then the optimal transmission and reception strategies are equivalent in both directions, despite the difference in the capacity regions [108]. In the DSL standards (e.g. [72]), downstream transmission obeys per-pair power constraints similar to its upstream counterpart, thus the duality becomes strict-sense. Nonetheless, the upstream-downstream duality specifically focuses on the duality between the optimal signalling strategies for each direction, which is essentially the duality between MUP and MUD, a different duality of independent research interest.

The MUP-MUD duality originated from a pair of achievable bounds for the coding with side information problem, achieved by Dirty Paper coding (DPC) [109] and Wyner-Ziv coding (WZC)[6] [112], respectively. As portrayed in Fig. 9, DPC achieves optimal encoding/precoding using 'blind' (as in having no knowledge of the interference $\boldsymbol{s}$ in Fig. 9) decoding/detection, while WZC achieves optimal decoding/detection using 'blind' encoding/precoding. Namely, given a Gaussian distributed source $\boldsymbol{u} \sim \mathcal{N}(0, \sigma_u^2)$ and a Gaussian distributed noise source $\boldsymbol{n} \sim \mathcal{N}(0, \sigma_n^2)$, then the Shannon limit of the coded system is $\mathcal{C} = \log_2(1 + \sigma_u^2/\sigma_n^2)$ for both WZC and DPC, regardless of the variance of the interference $\boldsymbol{s}$. The WZC-DPC duality characterizes the holy grail of a multi-user system, i.e. "(non-causally) known interference does not matter even if one side is blind". However, the MUP-MUD duality can be further expanded to sub-optimal schemes in an algorithmic and structural sense. For instance, the widely studied linear schemes, such as the zero

[6]We note that WZC was originally a solution of the distributed source coding problem. However, as mentioned in [110], its lossless counterpart Slepian-Wolf coding [111] has a strong duality with channel coding.

forcing and the minimum mean square error estimator, are shown to constitute MUP-MUD pairs [113] [114].

In this section and the next, we will discuss the existing MUP algorithms in the current DSL communications paradigm, as well as reviewing a family of powerful MUP algorithms based on lattice reduction. Based on the MUP-MUD duality, we hope that our study of the more advanced MUP techniques will inspire further exploration into the subject of downstream vectoring techniques in an algorithmic sense, where the rich literature of MUD [32] will assist us tremendously in selecting algorithms whose duals are applicable to downstream vectoring. Overall, the main difference between MUP and MUD is considered to be their information transfer characteristics, since the latter assumes a Gaussian type conditional probability distribution due to white noise, while the former usually does not rely on nor has access to the noise statistics. Therefore, MUD algorithms that strongly exploit the knowledge of noise statistics such as [115] are generally inapplicable to MUP. Another problematic aspect is that MUP does not benefit from using soft information, which is a consequence of the limited bandwidth of each tone. For this reason, iterative MUD algorithms which rely on log-likelihood ratio (LLR) feedback such as the turbo MUD of [84] is also impractical for MUP design.

A classification of the candidate MUP algorithms that we will cover in this survey is portrayed in Fig. 10. Despite the theoretical optimality of DPC, the processing delay incurred by the precoder also has an impact on the final achievable throughput. Therefore, it is sometimes preferable to use a low-complexity suboptimal precoder to a high-complexity near-optimal one. This design choice is usually justified when the FEXT power is low, as in VDSL2, where the performance gap between the optimal and suboptimal precoding schemes is negligible. Even in the 106 MHz version of G.fast, low-complexity linear precoding is eminently suitable for downstream vectoring. In this section, we will cover the basics of the conventional linear and non-linear precoding algorithms, based on extensions of both the classic zero forcing precoding (ZFP) and the Tomlinson-Harashima precoding (THP) [116] [117], respectively. To avoid confusion, the terminologies ZFP and THP are used to represent the classic member of the respective family (cf. Sec. III-B1 and III-C1), unless a particular variant is specified.

We should note that the widely recognized sum rate upper bound, obtained by using either the maximum ratio combining or its downstream counterpart, the maximum ratio transmission[7], is only capable of characterizing the best-case single-user performance, where both the direct channel and every FEXT channel associated with the designated user are gleaned for obtaining the specific user's maximum diversity gain. In maximum ratio transmission, the FEXT channels coupled into the direct are exploited, whereas in maximum ratio combining, the FEXT channels leaking from the direct are gleaned. In realistic multi-user wireline systems, it is impossible for all users to achieve their respective best-case performance simultaneously using the maximum ratio strategies. Therefore, we will not classify them as 'multi-user' precoding techniques. Instead, we may consider those as beamformers suitable for the single-CPE multi-pair scenario of Fig. 3.

[7]These bounds are also known as the matched filter bounds, or single-user bounds [19].

B. LINEAR PRECODING

1) Plain Zero Forcing

In order to force the FEXT to zero at the CPEs of Fig. 7, an obvious choice of the precoder $\boldsymbol{G}$ would be channel inversion. Given an invertible $K \times K$ DSL channel $\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_K]$, we have:

$$\boldsymbol{x} = \boldsymbol{G}\boldsymbol{u}, \tag{7}$$

where $\boldsymbol{G} = \boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{H}^H)^{-1} = \boldsymbol{H}^{-1}$. As implied by Fig. 4, $\|\boldsymbol{h}_k\|^2 \ \forall k$ is small in the high-frequency band, which subsequently suggest that its inverse $\boldsymbol{G}$ usually has a high power. Since current and next-generation wireline systems will operate under a strict transmit PSD (TxPSD) mask, the operations in (7) generally result in violation of the TxPSD mask and therefore we have to normalize the transmitted symbols. The TxPSD mask imposes a power limit on a per tone per pair basis [72], and hence we generally have to normalize the transmitted symbols with respect to the peak TxPSD per tone over all pairs $k = 1, 2, \cdots, K$. But for analytical tractability, we can gain meaningful insights into the performance of precoding by limiting the sum of TxPSD over all pairs with the aid of a scalar power assignment policy, similar to the total power constraint of an antenna group in wireless communications. This is also known in wireline communications as static spectrum balancing (SSB, as opposed to DSB) since there is no optimization involved. In fact, it is shown in [42] that the two TxPSD characterization methods have a simple geometric relationship in the signal space. Meanwhile, scalar power assignment can also significantly reduce the complexity of spectrum balancing, in which case practical seamless rate adaptation (SRA) techniques [118] may be efficiently invoked. Assuming $\gamma = \|\boldsymbol{H}^{-1}\boldsymbol{u}\|^2$, the combined equivalent channel seen by the message symbol vector $\boldsymbol{u}$ is a normalized identity matrix $\boldsymbol{I}/\sqrt{\gamma}$:

$$\begin{aligned}\boldsymbol{y} &= \frac{\boldsymbol{H}}{\sqrt{\gamma}}(\boldsymbol{H}^{-1}\boldsymbol{u}) + \boldsymbol{n} \\ &= \frac{\boldsymbol{u}}{\sqrt{\gamma}} + \boldsymbol{n},\end{aligned} \tag{8}$$

which results in amplified noise for the constellation demappers, hence decreases the detector's signal-to-noise ratio (SNR). More explicitly, the equalized symbol vector becomes:

$$\boldsymbol{z} = \boldsymbol{u} + \boldsymbol{n}\sqrt{\gamma}. \tag{9}$$

An Example

For the forthcoming G.fast 212 MHz standard, an example of the basic features of downstream vectoring with perfect

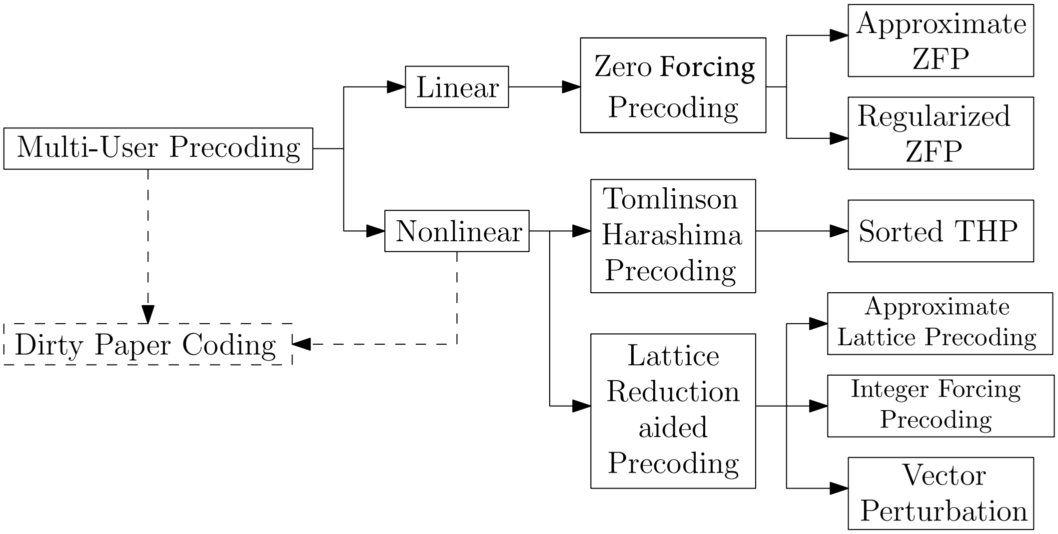

FIGURE 10: Family tree of the MUP techniques. Dirty paper coding characterizes an idealized non-linear scheme which achieves capacity.

DP-side transmit CSI is depicted in Fig. 11 and Fig. 12. In the two-customer test system, each customer is assigned one of the two pairs within the binder and the AWGN has a PSD of -150 dBm/Hz. The message signal transmitted via the first pair (P1) $u_1(f)$ has a flat PSD of -50 dBm/Hz and a constant phase of π over the full 212 MHz bandwidth, whereas the second pair (P2) transmits the blank message $u_2(f) = 0$. In the first scenario, $u_1(f)$ is injected directly into P1 without precoding. In the second scenario, the plain ZFP-based vectoring is applied both to $u_1(f)$ and to $u_2(f)$ for all frequencies, and the resultant signals $x_1(f)$ and $x_2(f)$ are injected into P1 and P2, respectively. A TxPSD mask of -50 dBm/Hz per pair is introduced for all frequencies.

When P2 has no signal injection, the signal received at the output of P2 consists solely of FEXT from P1. Injecting the vectored signal $x_1(f)$ and $x_2(f)$ eliminates said FEXT, since the PSD of the FEXT signal is seen to be reduced to the noise level in Fig. 11, while the phase of the message $u_1(f)$ is recovered in Fig. 12. However, in order to comply with the TxPSD mask, the power of the signal received at the output of P1 is penalized. The SNR penalty shown in Fig. 11 is the direct consequence of the channel's degradation upon increasing the frequency.

Eq. (7) is attractive in terms of its low complexity. Additionally, the near-optimality of the plain ZFP-like algorithms was also widely recognized during the legacy DSL era [119] [120] due to the diagonally dominant structure of $\boldsymbol{H}$ for all tones over the entire bandwidth of up to 30 MHz. However, applying the plain ZFP to the tones beyond 106 MHz will result in excessively amplified noise due to the TxPSD mask. Therefore, it is plausible that the equivalent parallel AWGN channels characterized by (9) is not capacity-approaching. In order to reduce the capacity loss associated with the plain ZFP, a number of linear optimizations may be used. An efficient relative of the plain ZFP is constituted by the diagonalizing precoder of [121], which generates a diagonal (but not identity) channel matrix when combined with $\boldsymbol{H}$. Namely, the following decomposition of $\boldsymbol{H}$ is invoked:

$$\boldsymbol{H} = \boldsymbol{D}\boldsymbol{L}, \tag{10}$$

where $\boldsymbol{D}$ is a diagonal matrix and the diagonal of $\boldsymbol{L}$ is a unit vector. Diagonalizing precoding avoids full channel inversion at the transmitter, which results in a reduced equivalent noise of $\|\boldsymbol{L}^{-1}\boldsymbol{u}\|\boldsymbol{D}^{-1}\boldsymbol{n}$ rather than the noise contribution of $\|\boldsymbol{H}^{-1}\boldsymbol{u}\|\boldsymbol{n}$ of the plain ZFP in (9), given the assumption that CSI knowledge is perfect at both the DP and the CPEs.

Since the processing delay incurred by MUP strongly affects the achievable sum rate, as well as the realization of ultra reliable low latency communications (URLLC) in next generation access networks, it is crucial to maintain minimum MUP complexity. Historical approaches to low complexity MUP in the DSL literature include relaxing the ZF criterion, which led to partial/approximate cancellation schemes such as those of [122] [123]. However, their performances heavily depend on the diagonal dominance of the channel matrices. With the evolution of signal processing hardware and parallel computing, the complexity saving of partial/approximate ZFP becomes negligible, while their performance loss compared to full cancellation is expected to become significantly more pronounced in G.fast/mgfast. Therefore, we will not discuss these schemes further, because their performance is well bounded by the plain ZFP.

2) Regularized Zero Forcing

In the ideal noiseless scenario, the zero forcing criterion is optimal for the QAM constellation demappers at the receiver side. However, in the presence of both noise and FEXT, the plain ZFP is unable to strike an attractive trade-off between noise enhancement and residual FEXT. A common linear improvement of the plain ZFP usually chooses to optimize

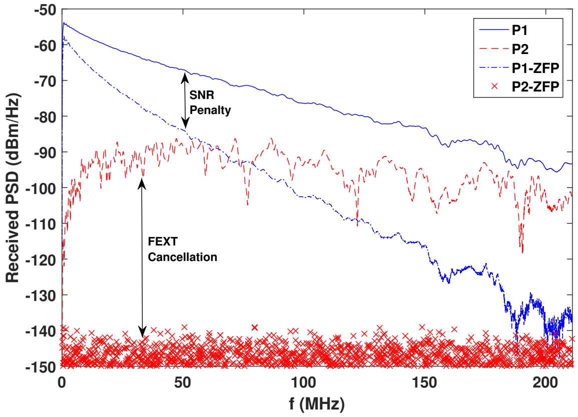

FIGURE 11: Received PSD at the output of a two-pair binder. The intended message $u_1(f)$ carried by pair 1 (P1) is an impulse-like signal with flat PSD of -50 dBm/Hz and a constant phase of π spanning the 212 MHz band while pair 2 (P2) carries blank message $u_2(f) = 0$. The transmitted signal is normalized to the flat TxPSD mask of -50 dBm/Hz per pair and the DP has perfect non-causal downstream channel knowledge. Comparing with the case of plain transmission without vectoring, employing plain ZFP reduces the FEXT signal picked up by P2 to the noise level, while the received PSD of P1 is also penalized as a result of channel inversion.

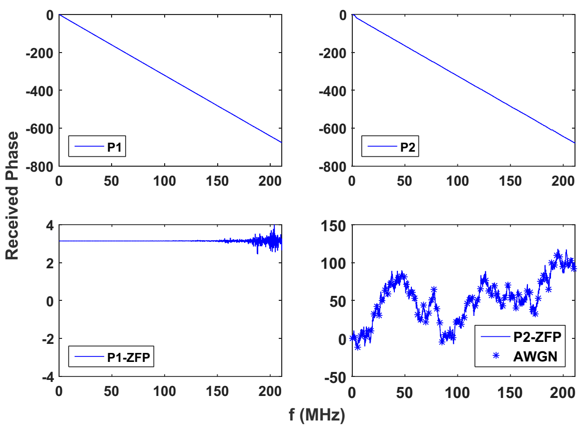

FIGURE 12: Received phase at the output of a two-pair binder. The system configuration of Fig. 11 is used. Without vectoring, the phase response shows the linear phase of the direct channel of P1 as well as the FEXT channel coupling into P2. Using the plain ZFP, the constant phase of $u_1(f)$ is recovered at the output except for high frequency channels where the the signal is heavily distorted by AWGN. Meanwhile, the phase response at the output of P2 is the phase of the observed AWGN sample.

the received signal-to-interference-and-noise ratio (SINR) of $\boldsymbol{z}$ for achieving an improved trade-off, by regularizing the precoding matrix $\boldsymbol{H}^{-1}$. Formally, the regularized version of ZFP, which maximizes the received SINR [124] is formulated as:

$$\boldsymbol{G}_{\text{mSINR}} = \boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{H}^H + \alpha \boldsymbol{I}_K)^{-1}, \tag{11}$$

where the optimal choice of the positive constant α is $\alpha = K\sigma_n^2$ [124]. Coincidentally, (11) is also the closed form solution of the minimum mean square error (MMSE) signal reception criterion in the limit of large binder sizes K, corresponding to minimizing the following cost function:

$$\boldsymbol{G}_{\text{MMSE}} = \arg\min_{\boldsymbol{G}} E\{\|(\boldsymbol{H}\boldsymbol{G} - \boldsymbol{I}_K)\boldsymbol{u} + \boldsymbol{n}\|^2\}. \tag{12}$$

Extending the result of (12), the authors of [125]–[127] recently proposed more flexible precoder designs by using the weighted MMSE criterion, which may be customized to achieve different optimization criteria such as the sum rate, QoS or fairness. The flexible configuration of weighted MMSE makes it a preferable candidate for joint vectoring-spectrum-balancing, i.e. multi-level DSM. Another precoding scheme closely related to the max-SINR precoding scheme optimizes the received signal-to-leakage-and-noise ratio (SLNR) [128] [129]. In contrast to max-SINR precoding, which minimizes the total received interference $|\sum_{j=1,j\neq k}^{K} h_{k,j}x_j|^2$ of (5) from other pairs, max-SLNR precoding minimizes the total leakage $|\sum_{k=1,k\neq j}^{K} h_{k,j}x_j|^2$ coupled into other pairs. As a result of the TDD/FD channel reciprocity in G.fast/mgfast, max-SLNR precoding for the downstream actually constitutes the dual counterpart of max-SINR detection for the upstream [130] and vice versa.

On a historical note, it is worth mentioning that instead of computing the closed form Wiener filtering matrix, the solution of (12) may also be found using classic iterative algorithms such as the least mean square (LMS) method, or alternatively using the simpler sign error feedback scheme proposed in [81] [131]. However, the success and fast convergence of these iterative algorithms heavily rely on the conditioning of the channel matrices. Given the evaded near-orthogonality of the frequency domain channel matrices in G.mgfast, the performance of these iterative algorithms in next-generation wireline access networks has to be further investigated.

3) Remarks

Since the practical choice of MUP algorithm for downstream DSL transmission is vendor-specific [30] [25], the research behind FEXT mitigation for next-generation G.mgfast and for future wireline networks is heavily scenario-dependent and measurement-based. However, according to the collective comments from the industrial experts who have reviewed the preliminary versions of this manuscript, we have found that most of the proposed MUP designs should be classified as regularized ZFP. Unfortunately, due to the paucity of sufficiently accurate modelling of the channel matrices, the theoretical performance with respect to general large-scale multi-pair DSL networks has so far only been qualitatively intimated by the asymptotic analysis of [132].

On the other hand, it was shown in [133] that for wireless MIMO channel matrices, which do not exhibit diagonal dominance, the gain of regularized ZFP over plain ZFP (and its simplified extensions) does not vanish for high SNRs σ_u^2/σ_n^2. Unfortunately, these low complexity, linear precoding schemes are known to be suboptimal in wireless systems [134], where the channel matrices are far from orthogonal. Since the DSL channel matrices also explicitly evaded orthogonality in the high frequency band, the suitability of linear MUP schemes for high performance wireline communications in the 100+ MHz band may become much more dependent upon the optimization of sophisticated joint vectoring-and-spectrum-balancing strategies, as reported by the investigations of [135] [136], and also by the spatially-targeted hybrid MUP-beamforming technique of [137]. These techniques may become even more plausible for low utility rate wireline networks, where the number of active pairs at any given time is sparse compared to the binder size. In particular, it is plausible that the optimal MUP design converges to the linear maximum ratio transmitter, when the network becomes more sparse.

C. TOMLINSON-HARASHIMA PRECODING

1) ZF-THP

As a dual counterpart of both successive interference cancellation (SIC) and of decision feedback equalisation (DFE), the THP relies on triangular factorization of $\boldsymbol{H}$ at the transmitter. This version obeying the ZF reception criterion is sometimes referred to as the ZF-THP. Following the design of [138] [76] seen in Fig. 13, for the $(K \times K)$-element square matrix $\boldsymbol{H}$, we invoke the QR decomposition of its conjugate transpose given by:

$$\boldsymbol{H}^H = \boldsymbol{Q}\boldsymbol{R}, \tag{13}$$

where $\boldsymbol{R}$ is a $(K \times K)$-element upper triangular matrix, and $\boldsymbol{Q}$ is a unitary matrix, i.e. $\boldsymbol{Q}^H\boldsymbol{Q} = \boldsymbol{I}$. Instead of using $\boldsymbol{x} = \boldsymbol{H}^{-1}\boldsymbol{u}$ as we did for the plain ZFP, we define $\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{x}'$ as the final transmitted symbol vector. The channel output $\boldsymbol{y}$ in this case is:

$$\boldsymbol{y} = \boldsymbol{R}^H\boldsymbol{x}' + \boldsymbol{n}. \tag{14}$$

By defining the diagonal matrix $\boldsymbol{W} = \mathrm{diag}[\mathrm{diag}(\boldsymbol{R}^H)] = \mathrm{diag}\{r_{1,1}, r_{2,2}, \cdots, r_{K,K}\}$, we have a naive FEXT pre-canceller in the form of:

$$\boldsymbol{W}\boldsymbol{u} = \boldsymbol{R}^H\boldsymbol{x}'. \tag{15}$$

Eq. (15) is favourable in that it exploits the triangular structure of $\boldsymbol{R}^H$. Since $\boldsymbol{R}^H$ is now a lower triangular matrix, the first message symbol u_1 does not experience FEXT, hence its leakage into all subsequent pairs from the same cable may be determined and subtracted, which now makes the second pair FEXT-free. However, using the naive successive FEXT pre-cancellation will inevitably amplify the unnormalized TxPSD. Therefore, we may recursively apply FEXT subtraction and modulo reduction to construct the pre-equalized symbol vector $\boldsymbol{x}'$ as follows:

$$x'_k = \begin{cases} u_1 & k = 1 \\ \Gamma_\phi\left[u_k - \sum\limits_{m=1}^{k-1} \dfrac{r_{k,m}}{r_{k,k}} x'_m\right] & k = 2, 3, \cdots, K. \end{cases} \tag{16}$$

The corresponding matrix notation is:

$$\boldsymbol{x}' = \Gamma_\phi\left[\boldsymbol{u} + (\boldsymbol{I} - \boldsymbol{W}^{-1}\boldsymbol{R}^H)\boldsymbol{x}'\right], \tag{17}$$

where $\Gamma_\phi[\cdot]$ represents the complex version of a real-valued modulo-ϕ reduction for a complex vector $\boldsymbol{a}$:

$$\Gamma_\phi[\boldsymbol{a}] = \Re(\boldsymbol{a}) - \phi\lfloor\frac{\Re(\boldsymbol{a})}{\phi} + \frac{1}{2}\rfloor + j(\Im(\boldsymbol{a}) - \phi\lfloor\frac{\Im(\boldsymbol{a})}{\phi} + \frac{1}{2}\rfloor). \tag{18}$$

The in-phase part and quadrature-phase part of a QAM symbol drawn from rectangular constellations usually share a common modulo base. If we denote the minimum phaser spacing and maximum amplitude of a rectangular QAM constellation by ξ and c respectively, then we have $\phi = 2c+\xi$. For square M-QAM, the modulo base is simplified to $\phi = \xi\sqrt{M}$. The substitution of (18) in (16) shall guarantee that all elements of $\boldsymbol{x}'$ are located inside the square bounded by $(-\phi/2, \phi/2] + j(-\phi/2, \phi/2]$, therefore strictly confines the average power of $\boldsymbol{x}'$. In fact, the output of the modulo-ϕ operation constitutes a uniformly distributed signal set over the square region. Meanwhile, the phase rotator $\boldsymbol{Q}$ does not cause any power enhancement. We may therefore expect that the final unnormalized TxPSD of $\boldsymbol{x}$ is nearly identical to the average power density of $\boldsymbol{u}$.

By applying element-wise modulo reduction and single user FDE, we can recover $\boldsymbol{z}$ from $\boldsymbol{y}$ in (14) as:

$$\boldsymbol{z} = \boldsymbol{u} + \Gamma_\phi\left[\boldsymbol{W}^{-1}\boldsymbol{n}\right]. \tag{19}$$

Since the equalization matrix $\boldsymbol{W}^{-1}$ represents the inversion of the direct links only, it will not amplify the received noise $\boldsymbol{n}$ nearly as much as the plain ZFP did in (9). On the other hand, $\Gamma_\phi[\cdot]$ transforms the received Gaussian noise into modulo-Gaussian detection noise and thus increases the detection noise. However, the latter usually has little or no impact on the performance of the ZF-THP, especially when the detection SNR is high. It should also be noted that the differential entropy of $\boldsymbol{z}$ is identical to that of $\boldsymbol{u}$, which generally results in information loss at low SNRs. This is known as the modulo loss (cf. Section III-C3).

2) Sorted THP

As shown in Eq. (19), the detection SNR required by the QAM demappers is related to the gain of the (rotated) direct channel $r_{k,k}$. Because of the DFE nature of the THP, the magnitude of $r_{k,k}$ and thus the detection SNR usually degrades with the user's index k during transmit precoding. The standard ZF-THP is therefore subject to the worst-case dominance effect, because the system's overall performance will typically be dominated by the last user having the worst performance. On the other hand, the DFE structure of

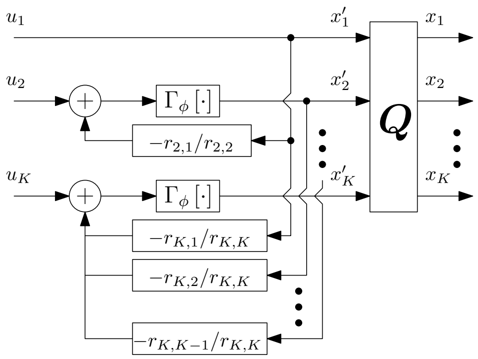

FIGURE 13: A ZF-THP based downstream transmitter. The decentralized FDEs at the CPE side (not shown) are responsible for equalizing the direct channels as well as for remapping the received signal to the Galois field associated with the QAM alphabet.

THP is also subject to error propagation effects, when the realistic imperfect CSI feedback protocol of Fig. 7 is invoked. In particular, any erroneously encoded QAM symbol will result in consecutive errors for all subsequent symbols to be transmitted over the same tone. For these reasons, it is critical to have an optimal ordering of the twisted pairs for ensuring that the performance loss associated with THP's DFE structure is minimized. Given a specific channel matrix, we can optimize the system by initializing the sorted THP to the specific pair associated with the worst direct channel for ensuring that its corresponding detection SNR is maximized. Several propositions have already been suggested using the sorted THP for next generation wireline access networks [139] [140].

Mathematically, a binary permutation matrix $\boldsymbol{E}$ is used for indicating the extra sorting during channel triangularization, so that the minimum detection SNR per vector is maximized [141]. Given the optimal choice $\boldsymbol{E}^*$, we have:

$$\boldsymbol{H}^H\boldsymbol{E}^* = \boldsymbol{QR}, \text{ where } \boldsymbol{E}^* = \arg\max_{\boldsymbol{E}} \left[\min\{r_{1,1}^2, \cdots, r_{K,K}^2\}\right], \tag{20}$$

while for a fixed channel $\boldsymbol{H}$ and any given $\boldsymbol{E}$, the product of all direct channel gain is constant:

$$\det(\boldsymbol{H}^H\boldsymbol{H}) = \det(\boldsymbol{R}\boldsymbol{E}^H\boldsymbol{E}\boldsymbol{R}^H) = \det(\boldsymbol{R})\det(\boldsymbol{R}^H) = \prod_{k=1}^{K} |r_{k,k}|^2. \tag{21}$$

Hence, max-min optimization compensates the worst pair's performance at the expense of good pairs. As a consequence, the largest performance gap between the pairs diminishes. Another sorting scheme conceived in [142] for multi-antenna wireless systems prioritizes the good channels, while turning off those having hostile channel conditions, so that the remaining active links achieve a higher sum-throughput than the entire group did before. However, this technique is only beneficial for point-to-point MIMO systems where fairness is irrelevant.

Under the constraint of (21), (20) represents the process of iteratively minimising $|r_{k,k}|$ of line k for $k = 1, 2, \cdots, K$. A very similar problem was also repeated for the sequential detection of space time codes [143]. The V-BLAST solution was found to be inefficient due to multiple inversions of the channel in the algorithm. Consequently, the authors of [144] proposed the sorted QR decomposition (SQRD) to solve the same problem, which also showed that SQRD solves the same user ordering problem by investing at most 60% of the complexity required by V-BLAST, while only imposing negligible loss on the detection performance.

3) Remarks

As a very popular non-linear precoding technique representing the original vectored transmission proposal [76], the performance of the THP is an extensively studied subject in multi-user communication systems. As for the ZF-THP scheme of Section III-C1, it is widely recognized that there is a gap between the maximum achievable bandwidth efficiency of the ZF-THP using square constellations and the Shannon limit. Without loss of generality, the three dominant sources of information loss of the ZF-THP for the ideal AWGN channel were studied in [145]. We should however note that the following types of loss are inherent to ZF-THP[8] and they are therefore not due to the specific types of communication channels. In fact, the wireline communications industry has already started investigating the fundamental limitations of the benchmark ZF-THP as shown in [146] [147]. For AWGN channels, these limitations are portrayed in Fig. 14 and described as follows:

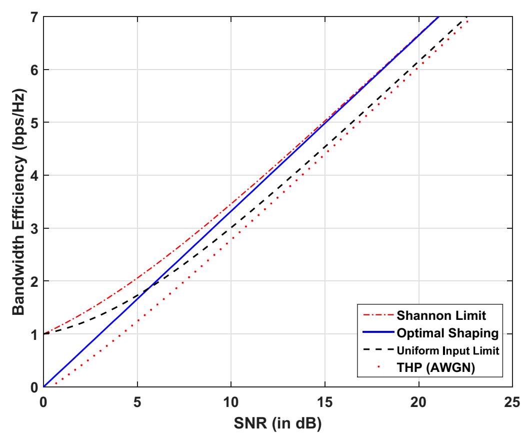

FIGURE 14: The bandwidth efficiency limit of the ZF-THP in interference-free AWGN channel.

- **Modulo Loss**: Due to the modulo operation (18) at the receiver, each QAM symbol to be demapped to (coded) bits is distributed within the square region bounded by

[8] In the following sections, we may notice that they are in fact related to the underlying modulo arithmetic and to the uniform distribution of QAM symbols, and therefore the analysis is applicable to non-linear MUPs in general.

the modulo base ϕ. Therefore, $\boldsymbol{z}$ and $\boldsymbol{u}$ have identical differential entropy. As a consequence, the maximum bandwidth efficiency of the ZF-THP will be upper bounded by $\log_2(\sigma_u^2/\sigma_n^2)$ rather than by the Shannon limit of $\log_2(1+\sigma_u^2/\sigma_n^2)$. A study of this particular phenomenon in a G.fast environment was presented in [147]. Fortunately, the lost term '1' inside the logarithm operator may be regained if we regularize the channel decomposition of (13) in a way similar to (11) to obtain precoding matrices based on the MMSE-THP scheme instead [148]. As characterized by the 'optimal shaping' in Fig. 14, the modulo loss is the most pronounced one in the low to medium SNR regime.

- **Precoding Loss**: Due to the modulo operation at the transmitter and the particular choice of the modulo base, the convex hull of the transmitted symbols occupies a slightly larger volume in the signal space than the specific constellation that they come from. As shown previously in Section III-C1, this precoding loss is negligible (precisely $(M-1)/M$ for large M) and it converges to one, when the constellation order M tends to infinity. Since the precoding loss of the ZF-THP mainly affects low-order constellations, it becomes most prominent in the low-SNR regime.
- **Shaping Loss**: Unlike the modulo loss and the precoding loss, the shaping loss is known to be associated with the square shaped QAM constellations, rather than with the non-linearity of the ZF-THP transmitter and the modulo receivers. Consequently, the shaping loss occurs even in linear precoding schemes such as the plain ZFP. As characterized by the 'Uniform input limit' seen in Fig. 14, the shaping loss grows to a constant of $1/2 \cdot \log_2(2\pi e/12) \approx 0.255$ bps/Hz or equivalently 1.53 dB in the high-SNR regime. Even though the case portrayed in Fig. 14 represents a single AWGN channel, we should note that the same shaping loss also applies to DSL and general wireless channels. However, it has been discovered that for channels with memory (such as fibre optical channels), the shaping loss, representing the gain reduction without optimal shaping can be as high as 1.88 dB [149]. Interestingly, we will see in Section IV that the aforementioned SNR penalty associated with MUP and DSL channels strongly resembles the shaping loss in a geometric sense.

The near-optimal information rate of the ZF-THP is readily recognized and highly appreciated by both the wireline and wireless communications community [76], [150], [151]. However, to what degree the optimality of the ZF-THP in practical multi-pair DSL systems is approached heavily relies on sophisticated DSB [18] strategies, such as the classic adaptive coding and modulation (ACM) protocols of wireless systems. Therefore, we also have to carefully assess the performance vs. complexity trade-off attained by DSB in order to quantify the practical performance of the THP family and those of the other MUPs.

IV. LATTICE REDUCTION AIDED MUP

A. LATTICES IN TELECOMMUNICATIONS

The concept of lattices is among the most fundamental and influential analytical tools in information theory. Lattice based methodologies are usually among the optimal candidates in a wide range of IT and CT related areas such as quantum-attack-resistant cryptography [152], capacity-achieving channel coding [37] and more relevantly, optimal MUP/MUD design [145]. We should note that even though lattice coding and the family of lattice reduction aided MUPs (LRMUP) to be investigated in this section both exploit the geometric goodness of lattices, their approaches are rather distinct. In the case of lattice coding, we have to construct a lattice codebook, which gives us the desired properties of good channel codes. By contrast, LRMUPs usually exploit the existing lattice structure spawned by the multi-user channel. Popularized by the celebrated Lenstra-Lenstra-Lovász [153] algorithm, the recent developments [154] [155] in lattice reduction algorithms and LRMUPs demonstrate that they have significant practical value and potential in the telecommunication industry.

1) Multi-user System as a Lattice

In the complex-valued K-dimensional Euclidean space $\mathbb{C}^K$, the $(K \times K)$-element generator matrix $\boldsymbol{G} = [\boldsymbol{g}_1, \boldsymbol{g}_2, \cdots, \boldsymbol{g}_K]$ spawns a lattice $\mathcal{L}(\boldsymbol{G})$ whose column vectors $\boldsymbol{g}_k$ represent the basis. By definition, lattices are periodic arrangements of discrete points. As a consequence, there is an infinite number of legitimate basis for any given lattice $\mathcal{L}(\boldsymbol{G})$, where $K > 1$. Therefore, the points of the lattice are formulated as:

$$\mathcal{L}(\boldsymbol{G}) = \{\boldsymbol{G}\boldsymbol{l} : \boldsymbol{l} \in \mathbb{G}^K\}, \tag{22}$$

where $\mathbb{G}$ denotes the set of all complex-valued integers and $\boldsymbol{G}\boldsymbol{l}$ is the standard matrix-vector multiplication. Using the above definition (22), it may be readily seen that the K-pair T-tone system of (4) is closely related to the union of linear spans of K-dimensional lattices $\mathcal{L}(\boldsymbol{H}^t)$ from T independent signal spaces, i.e. $\bigcup_{t=1}^{T} \text{span}(\mathcal{L}(\boldsymbol{H}^t))$ or simply $\bigcup_{t=1}^{T} \text{span}(\boldsymbol{H}^t)$. Meanwhile, since the majority of number theory problems originate from the real-valued domain, the complex-valued lattices are usually decoupled into real-valued ones. In particular, we may decouple one of the T signal spaces in $\mathbb{C}^K$ into $\mathbb{R}^{2K}$ using the following transformation:

$$\begin{bmatrix} \Re(\boldsymbol{y}) \\ \Im(\boldsymbol{y}) \end{bmatrix} = \begin{bmatrix} \Re(\boldsymbol{H}) & -\Im(\boldsymbol{H}) \\ \Im(\boldsymbol{H}) & \Re(\boldsymbol{H}) \end{bmatrix} \begin{bmatrix} \Re(\boldsymbol{x}) \\ \Im(\boldsymbol{x}) \end{bmatrix} + \begin{bmatrix} \Re(\boldsymbol{n}) \\ \Im(\boldsymbol{n}) \end{bmatrix}. \tag{23}$$

As mentioned in the definition above, a lattice of at least two dimensions has an infinite amount of basis. However, the fundamental parallelotopes constructed by two legitimate basis $\boldsymbol{G}_1 \neq \boldsymbol{G}_2$ of the same lattice have the same volume given by:

$$vol[\mathcal{L}(\boldsymbol{G})] = \sqrt{\det(\boldsymbol{G}_1^H \boldsymbol{G}_1)} = \sqrt{\det(\boldsymbol{G}_2^H \boldsymbol{G}_2)}, \tag{24}$$

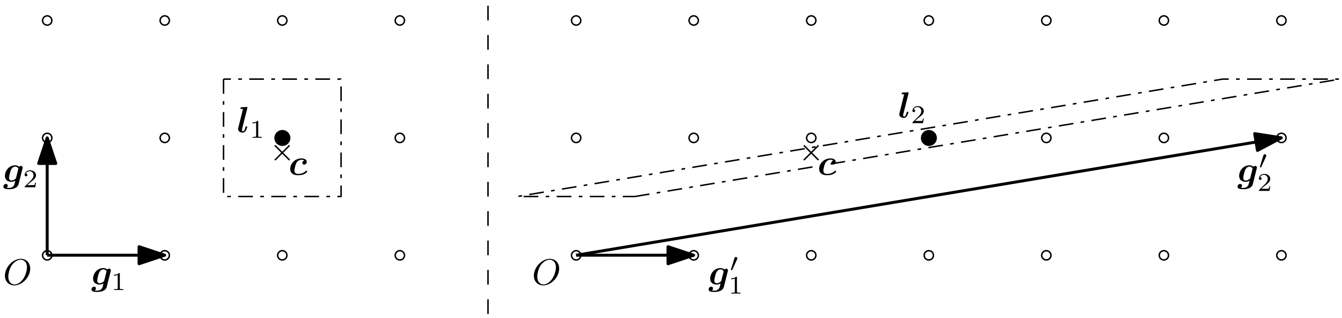

FIGURE 15: The effect of lattice reduction on quantizing a complex number. For the optimal basis on the left, the quantization region associated with any given complex-valued integer is the set of complex numbers that have the shortest Euclidean distances to said integer then to any other complex-valued integer. The minimum distance criterion is not satisfied by the basis on the right, therefore the quantization result is erroneous. This criterion can only be satisfied by orthogonal basis, and by the Voronoi cell (cf. Section IV-A2) of lattices that do not have an orthogonal basis.

where it may be observed that the pair of generator matrices are related by the unimodular transformation matrix $\boldsymbol{Z}$:

$$\boldsymbol{G}_1\boldsymbol{Z} = \boldsymbol{G}_2, \text{ where } \det(\boldsymbol{Z}) = \pm 1. \tag{25}$$

Since the desirable diagonal-dominance and quasi-orthogonality of low-frequency DMT channels is no longer achieved over the majority of the wide G.fast/mgfast spectrum, increasing the grade of orthogonality for these badly shaped channels emerges as a valuable performance-boosting strategy. A widely accepted measure of basis orthogonality for a lattice $\mathcal{L}(\boldsymbol{G})$ having the generator matrix $\boldsymbol{G}$ is termed as the orthogonality defect formulated as:

$$\delta(\boldsymbol{G}) = \frac{\prod_{k=1}^{K} \|\boldsymbol{g}_k\|}{vol[\mathcal{L}(\boldsymbol{G})]}, \tag{26}$$

where we can explicitly see that the most orthogonal basis is simultaneously the basis with the shortest vectors for any given lattice. A set of perfectly orthogonal basis vectors $\boldsymbol{G}^*$ satisfies $\delta(\boldsymbol{G}^*) = 1$. The primary task of LR is therefore to find the shortest basis of a given lattice (Minkowski's criterion [156]), or one of the shorter ones (LLL criterion [153]), using a known but long basis. We will not discuss in detail the lattice reduction algorithms themselves, but motivated readers might like to consult the survey and tutorial in [155]. A more specific survey on the LLL reduction algorithm is also available in [154].

An Example

Let us consider the simple problem of quantizing a complex number $\boldsymbol{c} = (2, 0.9j)$ to the nearest complex-valued integer. The correct answer $\boldsymbol{l}_1 = (2, j)$ is obtained using the conventional basis of $\mathbb{C}$, i.e. $\boldsymbol{g}_1 = (1, 0j), \boldsymbol{g}_2 = (0, j)$. This is because $\boldsymbol{c}$ falls inside the quantization region of $\boldsymbol{l}_1$ (left of Fig. 15). However, as seen on the right side of Fig. 15, if a bad/long basis such as $\boldsymbol{g}_1' = (1, 0j), \boldsymbol{g}_2' = (6, j)$ is used, then $\boldsymbol{c}$ falls in the quantization region of $\boldsymbol{l}_2 = (3, j)$. Therefore, the quantizer will erroneously consider $\boldsymbol{l}_2$ as the nearest integer neighbour of $\boldsymbol{c}$. As shown in [145], the quantization problem we have considered here represents the foundation of a wide range of non-linear MUP algorithms.

As an illustration in Fig. 16, we depict the quality of the basis associated with the inverse channel matrices, as well as that of the LLL-reduced basis, for a 10 pair DSL binder over the 212 MHz bandwidth. Observe from Fig. 16 that the orthogonality improvement offered by LLL reduction is quite significant, starting from around 90 MHz. As proven in the pioneering contribution of Babai [157] and empirically shown later by our simulations in Section VI, the LLL-reduced basis determines the lower bound of the worst-case performance of heuristic algorithms (e.g. rounding-off [157]) associated with MUPs (e.g. the plain ZFP). By contrast, for an arbitrary basis the worst-case performance remains unbounded. Moreover, since lattice reduction is invoked at the initialization stage as preprocessing, its complexity overhead does not detrimentally affect the performance of DSL systems.

2) The Essence of Vectoring and the Duality

Previously, we have shown that both the optimization criterion as well as the design of MUP and MUD may be viewed as each other's dual pairs based on the WZC-DPC duality. In the following, we will see that the algorithmic core of MUP and MUD may also be interpreted as each other's dual pairs, which is a by-product of using lattice reduction strategies for multi-user systems.

In terms of the MUD design for upstream DSL, the widely recognized optimal decision criterion is the maximum likelihood (ML) criterion, or the maximum a posteriori (MAP) criterion if the data symbol at the source $\boldsymbol{u}$ does not follow the uniform distribution. Formally, the ML detector (MLD) optimizes the following non-convex cost function:

$$\boldsymbol{z}_{opt} = \arg\min_{\boldsymbol{z}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{z}\|^2, \ \boldsymbol{z} \in \mathcal{M}, \tag{27}$$

where $\mathcal{M}$ represents a square QAM constellation. Since $\mathcal{M}$ represents a scaled version of (a subset of) the integer

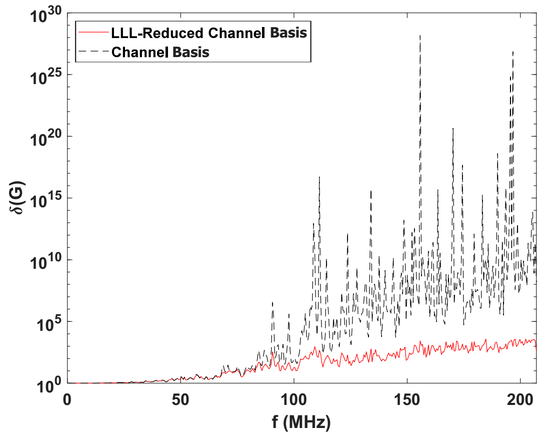

FIGURE 16: Orthogonality defect $\delta(\boldsymbol{G})$ associated with the signal space spawned by the inverse DMT channel matrices $\boldsymbol{G}$, as well as by the LLL-reduced ones, corresponding to a 10-pair binder in $\mathbb{C}^{10}$.

lattice $\mathbb{Z}^{2K}$, it is readily seen that the main task of (27) is to synthesize a point in $\mathcal{L}(\boldsymbol{H})$ that is the closest to another given point $\boldsymbol{y}$ in $\mathbb{R}^{2K}$. This is one of the most influential and fundamental problems in the development of lattice theory which is termed as the closest vector problem (CVP) [158] [159] that captures the essence of MUD. By extension, the other MUD algorithms may be conceived as heuristic approaches to an approximate solution of (27). Nonetheless, each MUD algorithm uniquely defines a tessellation pattern, and more importantly, a decision region surrounding each point of $\mathcal{L}(\boldsymbol{H})$. As shown in [160], the shape of the decision region directly determines the optimality of the associated MUD. It is well-known that the decision region associated with MLD is the basis-invariant Voronoi cell of $\mathcal{L}(\boldsymbol{H})$ [158]. The Voronoi cell $\mathcal{V}(\boldsymbol{H}, \hat{\boldsymbol{y}}_i)$ centred at $\hat{\boldsymbol{y}}_i \in \mathcal{L}(\boldsymbol{H})$ is defined as:

$$\mathcal{V}(\boldsymbol{H}, \hat{\boldsymbol{y}}_i) = \{\boldsymbol{y} \in \mathbb{R}^{2K} : \ \|\boldsymbol{y} - \hat{\boldsymbol{y}}_i\| \leq \|\boldsymbol{y} - \hat{\boldsymbol{y}}_j\| \quad \forall i \neq j\}. \tag{28}$$

By comparison, any other form of the decision region is suboptimal, because they violate the minimum Euclidean distance criterion (cf. Fig. 15). From a similar perspective, we will now characterize the optimal criterion for MUP.

With respect to the MUP design for downstream DSL, it is known from (9)(19) and Fig. 4 that the SNR penalty is the dominant problem over the wide bandwidth of G.mgfast (also in [161] [42]) compared to the losses discussed in Sec. III-C3. More specifically, a transmit-power-minimizing MUP is capable of achieving the optimal diversity order, regardless of the fundamental limits of non-linear MUPs. Therefore, it is reasonable to consider the transmit signal power as the cost function for MUP optimization. For example, we may use the simple formulation of [162] by introducing a perturbation vector defined as:

$$\boldsymbol{l}_{opt} = \arg\min_{\boldsymbol{l}} \|\boldsymbol{G}(\boldsymbol{u} + \boldsymbol{l})\|^2, \tag{29}$$

where $\boldsymbol{G}$ is the inverse channel matrix. Under the principle of DPC, the choice of the perturbation vector $\boldsymbol{l}$ should be limited to the set of scaled integers $\phi\mathbb{Z}^{2K}$ so that the modulo receivers $\Gamma_\phi[\cdot]$ can reconstruct the transmitted signal without knowing the exact value of $\boldsymbol{l}$. The same principle applies to the THP. For analytical simplicity, uniformly distributed continuous constellations are widely considered in non-linear MUP analysis [42], [148], [163]. Let $\mathcal{U}$ be a continuous square set formulated as:

$$\mathcal{U} = \{u : \ -\phi/2 \leq \Re(u) < \phi/2, -\phi/2 \leq \Im(u) < \phi/2 \ \}. \tag{30}$$

We now immediately notice the strong similarity between (27) and (29). More explicitly, in (29), we seek to synthesize a lattice point $\boldsymbol{Gl}$ in the lattice $\mathcal{L}(\boldsymbol{G})$ that is the closest one to the given point $\boldsymbol{Gu}$. It was also shown in [163] that the output of the MUP $\boldsymbol{x} = \boldsymbol{G}(\boldsymbol{u} + \boldsymbol{l})$ is uniformly distributed over $\mathcal{V}(\boldsymbol{G}, \boldsymbol{0})$. This MUP design is known as the vector perturbation (VP) [162] precoding.

Based on the concept of the decision region of MUDs, the authors of [42] proposed the dual concept of the vectoring mapping region (VMR). Because $\mathcal{L}(\boldsymbol{G})$ and $\mathcal{L}(\boldsymbol{H})$ constitute a pair of dual lattices [155], the VMR and the MUD decision regions are closely related. Without loss of generality, the VMRs of all the MUPs investigated in this survey are portrayed in Fig. 18, where the lattice is spawned by the inverse of a two-user system employing pulse amplitude modulation (PAM) in $\mathbb{R}^2$. We note that this limitation is a result of the fact that we cannot graphically demonstrate the VMRs of a two-user QAM system, because it is in $\mathbb{R}^4$.

B. APPROXIMATE LATTICE PRECODING

In [160] and [42], the concept of approximate lattice decoding and approximate lattice encoding were proposed independently for MUD and MUP. We shall adopt this notion and observe the influence of LR on the plain ZFP and the ZF-THP approaches, instead of using the potentially ambiguous terminology of lattice reduction aide broadcast precoding [164], the latter of which essentially constitutes extensions to the vector perturbation scheme we will cover later. In essence, approximate lattice precoding heuristically solves the closest vector problem with preprocessing (CVPP), the solution of which can be further used as preprocessing for other schemes, such as the approximate message passing algorithm of [165]. However, the strict-sense CVPP in lattice theory typically exploits the geometry of the Voronoi cell and utilizes exponential-sized memory/buffer for very high dimensional lattices [166], whereas LRMUP algorithms simply acquire a good basis to work with for low to medium dimensional lattices.

The nature of lattice reduction, given a generator matrix $\boldsymbol{G}$, may be described with the aid of the following decomposition:

$$\boldsymbol{G} = \boldsymbol{Q}\boldsymbol{R}\boldsymbol{Z}^{-1}, \tag{31}$$

where $\boldsymbol{Q}$ and $\boldsymbol{R}$ are unitary and upper triangular matrices, respectively. The unimodular matrix $\boldsymbol{Z}$ in (31) is carefully

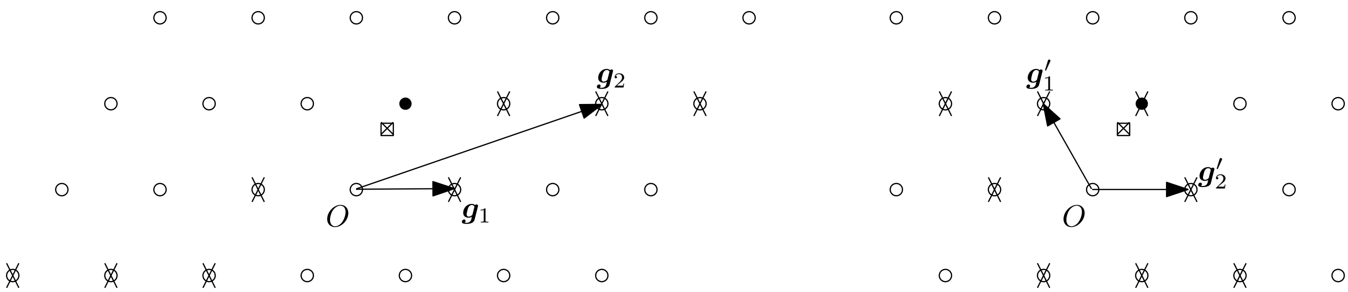

FIGURE 17: The effect of lattice reduction on the complexity of enumerating the closest lattice point problem $\arg\min_{\boldsymbol{l}} \|\boldsymbol{G}(\boldsymbol{u}+\boldsymbol{l})\|^2$. The crossed square represents the given point $-\boldsymbol{G}\boldsymbol{u}$ and the solid dot represents the closest lattice point to $-\boldsymbol{G}\boldsymbol{u}$. The lattice point O is given as the first heuristic solution during the sphere encoder's initialization, while the eight crossed circles on each side are the minimum integer combinations of the respective basis vectors. Therefore these are the first batch of lattice points to be enumerated. Compared with the case of the reduced basis on the right side where the final solution can be found within the first eight enumerations, finding the closest lattice point over the long basis on the left cannot be completed within the first eight enumerations. Therefore lattice reduction decreases the complexity of enumerating the closest lattice point.

constructed for ensuring that $\boldsymbol{G}_{\text{red}} = \boldsymbol{Q}\boldsymbol{R}$ satisfies one of the lattice reduction criteria. Eq. (20) closely resembles the format of (31), which indicates that the SQRD may be interpreted in a wide sense as a type of primitive lattice reduction. Lattice reduction is generally a hard problem, regardless of the reduction criteria used. There are several independently proposed reduction criteria, with the most extensively studied ones being the Minkowski criterion [156], the Hermite-Korkine-Zolotareff (HKZ) [167] criterion and the suboptimal Lenstra-Lenstra-Lovász (LLL) [153] criterion. The LLL reduction is the weakest criterion of the three, but it is still more favourable in practice because it is the first, polynomial-complexity, algorithm that produces reasonably short (i.e. whose length is strictly upper bounded) basis vectors. By constrast, both the Minkowski reduction and the HKZ reduction are considered as 'optimal' in the sense of obtaining the shortest basis, they rely on solving multiple iterations of the NP-hard shortest vector problems (SVP) [168]. Even though lattice reduction aided schemes are usually implemented in the real-valued domain, complex-domain algorithms also do exist. By comparison, it is empirically proven by simulations in both [169] and [34] that the complex-valued Minkowski and LLL algorithms are indeed equivalent to the real-valued versions.

1) LR-ZFP

In order to mitigate the noise enhancement of the plain ZFP, we may construct an LLL-reduced inverse channel $\boldsymbol{F} = \boldsymbol{H}^{-1}\boldsymbol{Z}$ as the precoder, in which case we have to shift the K-dimensional point $\boldsymbol{u}$ to a different position due to the unimodular transformation. Denoting the unimodular transformed message symbol vector as $\tilde{\boldsymbol{u}}$, we have:

$$\boldsymbol{x} = \frac{1}{\sqrt{\gamma}}\boldsymbol{F}\tilde{\boldsymbol{u}} \quad \text{where} \quad \tilde{\boldsymbol{u}} = \Gamma_\phi\left[\boldsymbol{Z}^{-1}\boldsymbol{u}\right]. \tag{32}$$

Since $\tilde{\boldsymbol{u}}$ is uniformly distributed over the basis parallelotope constructed by the reduced basis $\boldsymbol{F}$, we may observe that the

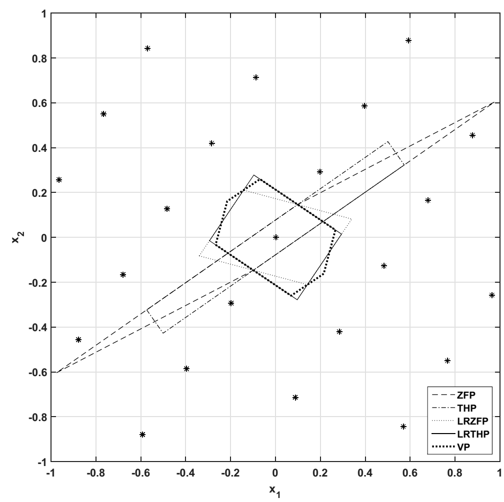

FIGURE 18: Geometric boundaries indicating the VMRs associated with the MUPs for a two-pair PAM based system. The basis of the lattice in $\mathbb{R}^2$ is spawned by the (*LLL-reducible*) channel inverse matrix, and its lattice points are shown. The joint input to the MUPs is the uniformly distributed input $\mathcal{U}$ of (30), which approximates the union of two independent PAM constellations. The *second moment* of each VMR indicates the average encoded signal power, while the *second moment per dimension* indicates the average encoded signal power per user. These properties are graphically characterized by the degree of 'isotropy' of the VMRs.

VMR of the LR-ZFP is the parallelogram region portrayed in Fig. 18. The corresponding equalized symbol vector is:

$$\boldsymbol{z} = \boldsymbol{u} + \Gamma_\phi\left[\boldsymbol{n}\sqrt{\gamma}\right], \tag{33}$$

where $\gamma = \|\boldsymbol{F}\tilde{\boldsymbol{u}}\|^2$. In this case, the expected performance improvement mainly accrues from the assumption that the inverse of any high-frequency DMT channel never constitutes a shorter basis than an LLL-reduced one. Otherwise, (33) would exhibit an even higher detection noise than (9) due to the modulo operation. The transmitter structure of the LR-ZFP scheme is given in Fig. 19. In comparison to the ZF-THP based transmitter of Fig. 13, the unimodular matrix filter $\boldsymbol{Z}^{-1}$ of the LR-ZFP based transmitter of Fig. 19 substitutes the decision feedback loop of the ZF-THP based transmitter of Fig. 13. The remaining components are structured similarly in both cases, except for the extra direct channel equalizers $(1/r_{k,k})$ required for a ZF-THP based receiver as formulated in (19). Since matrix filters can be implemented efficiently with the advent of parallel computing, whereas the decision feedback loop cannot, the LR-ZFP based transmitter would incur a lower processing delay than the ZF-THP based transmitter.

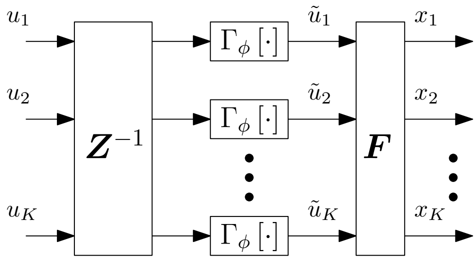

FIGURE 19: A LR-ZFP based downstream transmitter.

2) LR-THP

Since the sorted THP may be considered to constitute a low-level LR-THP, we may also employ a triangular decomposition, while replacing the sorting procedure by the more powerful LLL reduction, as formulated in:

$$\boldsymbol{H}^H = \boldsymbol{S}\boldsymbol{B}\boldsymbol{Z}^{-1}, \tag{34}$$

where $\boldsymbol{S}$ has orthogonal columns and $\boldsymbol{B}$ is of an upper triangular structure, with $\text{diag}(\boldsymbol{B}) = \text{diag}(\boldsymbol{I})$. As seen in [170], the optimal ordering of (20) and the implicit ordering associated with (34) may be stacked. However, Chang. et al. [171] proved that a combination of LLL reduction and SQRD would be similar to using LLL reduction alone, since the sorting operation within the LLL algorithm results in similar ordering to that of SQRD. Therefore, in our case we will dispense with any additional sorting.

Denoting the unimodular transformed message symbol vector as $\tilde{\boldsymbol{u}} = \Gamma_\phi\left[\boldsymbol{Z}^H\boldsymbol{u}\right]$, we have the precoded symbol vector $\boldsymbol{x}$ given by:

$$\boldsymbol{x} = \frac{1}{\sqrt{\gamma}}\boldsymbol{S}^{-H}\boldsymbol{x}', \text{ where } \boldsymbol{x}' = \Gamma_\phi\left[\tilde{\boldsymbol{u}} + (\boldsymbol{I} - \boldsymbol{B}^H)\boldsymbol{x}'\right]. \tag{35}$$

We may now see that the transmitted symbol vector is distributed over the LR-THP's VMR in Fig. 18, which is associated with the parallelotope/orthotope $\boldsymbol{S}^{-H}$. Then (35) results in the exact same form of the equalized symbol vector $\boldsymbol{z}$ as we obtained in (33), except that in the case of the LR-THP, $\boldsymbol{S}^{-H}$ has a more isotropic shape and thus it results in an even lower power amplification than the precoding matrix $\boldsymbol{F}$ does in the LR-ZFP. The transmitter structure of the LR-THP scheme is given in Fig. 20, which may be considered as a hybrid of the ZF-THP based transmitter of Fig. 13 and the LR-ZFP transmitter of Fig. 19. In fact, the main difference between the LR-THP and the LR-ZFP is akin to the difference between the ZF-THP and the plain ZFP. Since the LR-ZFP scheme of Fig. 19 may be viewed as a plain ZFP scheme (starting from $\tilde{u}_k$ in Fig. 19) implemented within a unimodular mapping process, the LR-THP scheme may also be conceptually considered as a conventional ZF-THP scheme[9] (starting from $\tilde{u}_k$ in Fig. 20) implemented within a unimodular mapping process.

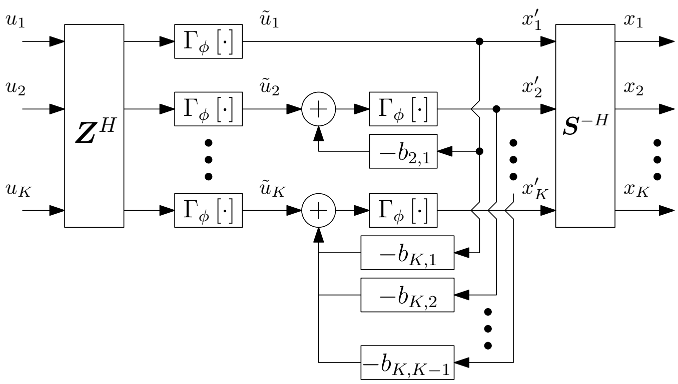

FIGURE 20: A LR-THP based downstream transmitter, $b_{i,j}$ is the element on the ith row and jth column of $\boldsymbol{B}^H$.

C. INTEGER FORCING PRECODING

As an emerging lattice reduction aided multi-user technique, the integer forcing precoding (IFP) [36] solution provides a novel yet familiar perspective of approximate lattice precoding. Conceptually, IFP is inspired by the reverse compute-and-forward (RCoF) protocol [175] used in wireless systems, which is also based on the parallel modulo Gaussian channel principle of all non-linear MUPs. The IFP scheme proposed in [36] may in fact be considered as a generalization of the aforementioned approximate lattice precoding schemes. By comparing the schematic of the LR-ZFP in Fig. 19 to that of the LR-THP in Fig. 20, it is plausible that both of the approximate lattice precoding schemes have three common building blocks at the transmitter. In IFP, these are generalized as follows:

- **Integer-valued filter**. In conventional approximate lattice precoding, this is exactly the unimodular transformation matrix, which is the solution of the SVP associated with the act of lattice reduction itself. However, instead of using a strictly invertible integer-valued matrix, IFP relaxes the criterion and only requires the integer-valued matrix to be invertible over Galois fields. The

[9] This is somewhat different from the typical ZF-THP transceiver structure because the CPE-side decentralized FDEs $1/r_{k,k}$ presented in (19) can be 'transferred' to the DP side if the decomposition of (13) produces a triangular matrix having unit diagonal.

TABLE 4: Milestones in the Development of Lattice Reduction Algorithms and LR-aided MIMO Techniques.

Year	Authors	Contribution
1873-1891 [156] [167]	Korkine and Zolotareff and Minkowski	Proposed the performance criterion for the *optimal* Minkowski-reduced and KZ-reduced lattice basis, respectively.
1982 [153]	Lenstra *et al.*	Proposed the suboptimal LLL criterion and the first 'polynomial-complexity' lattice reduction algorithm, namely the LLL algorithm.
1986 [157]	Babai	Analyzed the geometric properties of the LLL-reduced basis and proposed the *rounding-off* and the *nearest-plane* approximate solutions for the closest lattice point problem.
2004 [164]	Windpassinger *et al.*	Proposed a nonlinear precoding scheme based on the LLL-reduction and Babai's approximations, which significantly outperforms the conventional ZF-THP.
2007-2008 [172] [173]	Ling and Jaldén *et al.*	Independently proposed upper bounds for the average complexity of the LLL reduction algorithm, i.e. $\mathcal{O}(K^3 \log K)$ and $\mathcal{O}(K^2 \log K)$.
2008-2009 [174] [169]	Gan and Ma *et al.*	Proposed complex-valued LLL reduction algorithms directly applicable to MIMO systems using QAM.
2012 [33]	Zhang *et al.*	Proposed realistic algorithms for finding both Minkowski- and KZ-reduced lattice basis.
2013 [171]	Chang *et al.*	Investigated the effect of the LLL algorithm on the accuracy of Babai's approximation, as well as on the complexity of the sphere decoder.
2012-2017 [36], [175], [176]	Hong and Zhan *et al.*	Developed an *integer-forcing*-aided MIMO framework based on both lattice reduction and compute-and-forward technique. Instead of solving for the exact codeword, the IF technique finds the integer combination of the codewords which operates under an integer-valued effective CSI matrix.
2017 [34]	Ding *et al.*	Proposed a complex-valued lattice reduction algorithm for finding the optimal Minkowski-reduced basis, as well as a relaxed version of the algorithm.

underlying lattice problem is the shortest independent vectors problem (SIVP) [168], whose solution does not necessarily constitutes a basis, but is obtainable via lattice reduction strategies [177].

- **Lattice coding**. The lattice mapping of IFP is not necessarily carried out over either the multi-user channel lattice or its dual lattice. By contrast, the codebook is constructed using a pair of nested lattices [178] [179] with the standard 'construction A' approach of [180]. Recent research in lattice coding shows that we may achieve the same goal using only a single 'good' lattice [37]. Nonetheless, the coding procedure itself is ultimately a CVP solver.
- **Linear beamformer**. For approximate lattice precoding, the front-end linear beamformer is obtained via factorizing the reduced lattice basis. IFP allows for more relaxed choices of this front-end, subject to the specific power constraint of the system. In particular, if the identity matrix is used, then the IFP scheme becomes equivalent to the standard RCoF [36].

IFP bridges the lattice coding approach and the lattice reduction approach, the two main applications of lattice theory in telecommunications. In particular, IFP has the potential of borrowing good practices from lattice coding for improving the MUP performance beyond that of the conventional approximate lattice precoding. As a consequence of the more flexible lattice coding design, the VMR of IFP is not generally deterministic and it may also overlap with the VMR associated with another LRMUP. More interestingly, the distributed IFP-aided receivers can 'virtually' cooperate, which is contrary to the common belief, namely that the distributed downstream receivers cannot cooperate. Specifically, cooperative signal processing is typically associated with (multi-user) matrix filtering. In IFP, the receivers have to acquire the columns of the integer-valued matrix during initialization, which virtually enables matrix filtering. This particular aspect may lead to other cooperative distributed receiver designs relying on quantized information (such as CSI) gleaned from the downstream transmitter.

D. VECTOR PERTURBATION

Even though Babai had shown in [157] that the performance gap between LLL based approximate lattice decoding and MLD is reasonably small, attaining the optimal performance is still relevant, if we want to reach the maximum bandwidth efficiency in the forthcoming G.mgfast and to be fully prepared for the era of the converged access network in the near-future. Reflecting on the full-diversity MUP criterion of (29), the algorithm that finds its exact solution is known as the vector perturbation [162] scheme. The most widely recognized solver of the underlying exact CVP problem is more generally known as the sphere encoder[10]. We note that the cost function of (29) can still be optimized heuristically, which results in the MUPs of [164] [42]. However, the exact solution is valuable in that it allows full exploitation of the multi-dimensional signal space spawned by the multi-pair channels, because the performance gap between the suboptimal MUPs and the optimal MUP can be enormous

[10]For MLD this is typically called the sphere decoder. As seen in the difference between (27) and (29), the exact CVP solver operates in a finite lattice for MLD while the VP scheme can potentially work in infinite lattice. To avoid ambiguity, the latter is denoted as sphere encoder in this survey.

for DSL binders enclosing a large number of copper pairs[11]. In particular, with the popularization of SDN and NFV, future wireline networks may be able to consolidate different signal spaces resulting from LLU for all vectored groups simultaneously, in order to regain the loss imposed by alien FEXT.

The VP scheme of Fig. 22 constitutes an extension of the plain ZFP scheme, where the perturbation process (from u_k to x'_k) is introduced in addition to the plain ZFP encoding matrix $\boldsymbol{G}$. The operations of the VP based transmitter may be formulated as:

$$\boldsymbol{x} = \frac{1}{\sqrt{\gamma}}\boldsymbol{G}(\boldsymbol{u}+\boldsymbol{l}), \quad \boldsymbol{G} = \boldsymbol{H}^{-1}. \tag{36}$$

The exact solution of (29) is found by the enumeration-based sphere encoder. The original depth-first type sphere decoder was conceived in [181], which was then further improved both in [182] and [183]. The breadth-first type [184]–[186] and the more recent best-first type [187]–[189] enumeration algorithms were conceived as more hardware-friendly versions, since they can be readily pipelined to have a fixed throughput. All versions of the enumeration algorithms were claimed to approach the MLD performance, relying on a simple fact of the CVP problem, namely that if a lattice point is the closest one to another point in $\mathbb{R}^{2K}$ within a hypersphere region centred at the given point, then said lattice point is also the closest to the given point within the entire signal space. However, choosing the optimal radius of the hypersphere is a difficult problem in all cases. If the radius is too large, then there are too many lattice points to be enumerated in the hypersphere, whilst if the radius is too small, then the CVP search will fail and we are forced to extend the radius to restart the enumeration. Additionally, in the breadth- and best-first versions, tuning the 'breadth parameter' K is also a difficult problem, if both the MLD-like performance and the minimum buffer size are desired. In this section, we will characterize the depth-first version of [182] in Alg. 1 as the most established candidate. For a more in-depth analysis of the sphere encoder, the reader is redirected to the contribution of [190] and the references therein.

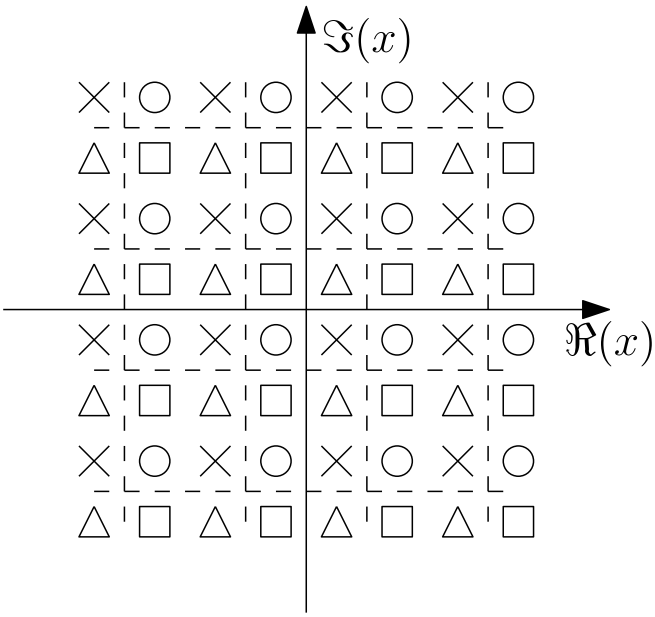

FIGURE 21: The effect of the additive perturbation. The four possible states of the 4QAM are depicted using one of the four symbols. Given that the original 4QAM constellation constitutes a square centred at the origin, introducing the perturbation vector results in an infinitely expanded QAM constellation. After perturbation, each message symbol from the original 4QAM is remapped onto one of the locations denoted by the same symbol on the infinite QAM constellation. From the perspective of the modulo receiver, all locations denoted by the same symbol on the infinite QAM constellation are indistinguishable.

[11] In fact, as a known result in lattice theory, both the exact and the near-optimal solution (whose loss is upper bounded by a constant) of the CVP are obtained in exponential complexity order. Conversely, solutions found by typical heuristic algorithms usually have exponential loss.

1) LR-aided Sphere Encoder

The sphere encoder consists of the lattice reduction preprocessing of Section IV-B followed by the search-tree-based enumeration. In the example of Fig. 17, the lattice is spawned in the 2D plane $\mathbb{R}^2$ having the generator matrix $\boldsymbol{G} = [\boldsymbol{g}_1, \boldsymbol{g}_2]$. In order to determine the closest lattice point (the black dot) to the given point $\bar{\boldsymbol{u}} = -\boldsymbol{G}\boldsymbol{u}$, the sphere encoder commences by transforming the long basis of $\boldsymbol{G}$ into a short one $\boldsymbol{G}' = [\boldsymbol{g}'_1, \boldsymbol{g}'_2]$ using lattice reduction. The main purpose of such a transformation is to accelerate the subsequent enumeration process as proven in [171].

Adopting the $2K$-dimensional real-valued system model for the remainder of this section and exploiting the lattice reduction notion of $\boldsymbol{G} = \boldsymbol{Q}\boldsymbol{R}\boldsymbol{Z}^{-1}$ in (31) for the inverse channel $\boldsymbol{G} = \boldsymbol{H}^{-1}$, we can map the sphere encoding problem of (29) from a general matrix $\boldsymbol{G}$ onto a triangular one $\boldsymbol{R}$ by rotation as follows:

$$\boldsymbol{l}'_{opt} = \arg\min_{\boldsymbol{l}'} \|\boldsymbol{v} - \boldsymbol{R}\boldsymbol{l}'\|^2, \text{ where } \boldsymbol{v} = \boldsymbol{Q}^H\bar{\boldsymbol{u}}/\phi, \ \boldsymbol{l}' = \boldsymbol{Z}^{-1}\boldsymbol{l}/\phi. \tag{37}$$

The constant ϕ is removed so that $\boldsymbol{l}'$ becomes an integer-valued vector in $\mathbb{Z}^{2K}$. For the initialization of the sphere encoder, an approximate solution to (37) can be easily obtained by solving the respective least squares problem, where we solve $\boldsymbol{l}'$ for a real-valued vector in $\mathbb{R}^{2K}$ first, and then we round off the entries of the real-valued vector to their respective nearest integer. Denoting the Euclidean distance from $\boldsymbol{v}$ to this approximate solution as $\sqrt{\beta}$, we may enumerate the lattice points within the $2K$-dimensional hypersphere of radius $\sqrt{\beta}$ centred at $\boldsymbol{v}$. Let $r_{k,j}$ be the entry at the kth row and jth column of the upper-triangular matrix $\boldsymbol{R}$ of (37).

Then we may define the following DFE structure as in [191]:

$$c_k = \begin{cases} \dfrac{v_k}{r_{k,k}} & k = 2K, \\ \dfrac{v_k}{r_{k,k}} - \displaystyle\sum_{j=k+1}^{2K} \dfrac{r_{k,j}}{r_{k,k}} l'_j & k = 2K-1, \cdots, 1. \end{cases} \tag{38}$$

Similar to (16), Eq. (38) has a bottom-up decision feedback structure. Therefore, the sphere constraint may be invoked in each iteration of (38), corresponding to the subspaces having lower than $2K$ dimensions. A necessary condition for a lattice point to fall inside the hypersphere centred at $\boldsymbol{v}$, i.e. $\|\boldsymbol{v} - \boldsymbol{R}\boldsymbol{l}'\|^2 < \beta$, is formulated as [191]:

$$r_{k,k}^2 (l'_k - c_k)^2 < \beta - \sum_{j=k+1}^{2K} r_{j,j}^2 (l'_j - c_j)^2, \quad k = 2K, 2K-1, \cdots, 1. \tag{39}$$

Eq. (39) represents a tree structure from the root node of $k = 2K$ to the leaf node of $k = 1$, in which we may compute the partial Euclidean distance of $\sum_{j=k+1}^{2K} r_{j,j}^2 (l'_j - c_j)^2$ from the enumerated lattice point to $\boldsymbol{v}$ in the $(2K - k + 1)$-dimensional subspace. The sphere encoder progresses towards the leaf node by one step whenever (39) is satisfied, hence the algorithm is depth-first. Otherwise, it backs up towards the root node by one level to evaluate a different lattice point. The evaluation at each node is one-dimensional along the integer sequence $\mathbb{Z}$ in the order of $l'_k, l'_k - 1, l'_k + 1, l'_k - 2, l'_k + 2, \ldots$, starting from the integer $l'_k = \lceil c_k \rfloor$, where $\lceil \cdot \rfloor$ is the rounding-off operator. Whenever the leaf node of $k = 1$ is reached, the radius $\sqrt{\beta}$ of the hypersphere boundary is updated and the corresponding state represents a new candidate lattice point. The final solution of (37) is the last lattice point obtained before the sphere encoder terminates its operation.

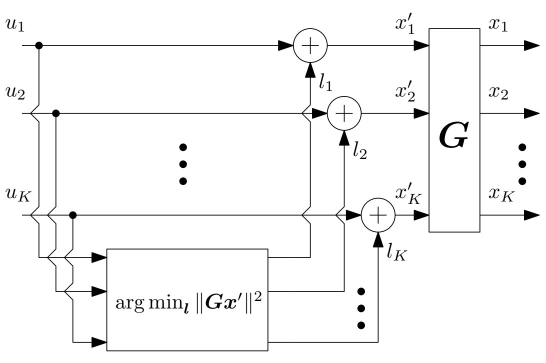

FIGURE 22: A VP based downstream transmitter. The additive vector $\boldsymbol{l}$ is classically solved either by the sphere encoder [162], or by Babai's approximations [164] [42].

2) Remarks

Albeit VP is capable of achieving the same diversity as the optimal MLD while also functioning as a simple attachment to the classic plain ZFP, it exhibits a high PAPR and a potentially prohibitive sphere encoding complexity. In terms of PAPR, the spliced constellation of Fig. 21 suggests

Algorithm 1: Depth-First Sphere Encoder

Input: Lattice Generator $\boldsymbol{G}$, a given point $\bar{\boldsymbol{u}} \in \mathbb{R}^{2K}$, the sphere radius $\sqrt{\beta} = \infty$;
Output: Integer vector $\boldsymbol{l} \in \mathbb{Z}^{2K}$, s.t. $\bar{\boldsymbol{u}} \in \mathcal{V}(\boldsymbol{G}, \boldsymbol{G}\boldsymbol{l})$;
$[\boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{Z}] = \text{Reduction}(\boldsymbol{G})$;
$\boldsymbol{v} \leftarrow \boldsymbol{Q}^H \bar{\boldsymbol{u}} / \phi$;
$c_{2K} \leftarrow v_{2K} / r_{2K,2K}$;
$l'_{2K} \leftarrow \lceil c_{2K} \rfloor$;
$k \leftarrow k - 1$;
while $true$ **do**
  if $\sum_{j=k+1}^{2K} r_{j,j}^2 (l'_j - c_j)^2 < \beta$ **then**
    if $k > 1$ **then**
      $c_k \leftarrow (v_k - \sum_{j=k+1}^{2K} r_{k,j} l'_j) / r_{k,k}$;
      $l'_k \leftarrow \lceil c_k \rfloor$;
      $k \leftarrow k - 1$;
    else
      //A valid intermediate $\boldsymbol{l}'$ is found
      $\beta \leftarrow \|\boldsymbol{v} - \boldsymbol{R}\boldsymbol{l}'\|^2$;
      $k \leftarrow 2$;
      Choose the next value for l'_2 in order;
  else
    $k \leftarrow k + 1$;
    if $k > 2K$ **then**
      Terminate;
    Choose the next value for l'_k in order;
$\boldsymbol{l} \leftarrow \phi \boldsymbol{Z} \boldsymbol{l}'$;

that the dynamic range required by the front-end filter $\boldsymbol{G}$ is significantly increased, especially since the constellation expansion depends on the worst-case channel quality of all tones. Therefore we in general need more expensive circuitry for the DP of G.mgfast using the VP-based approach. As a design alternative of sphere encoding, Zhang et al. [161] proposed the expanded constellation mapping scheme that solves the closest point problem of a finite lattice, which is more closely related to the sphere decoding based implementation of MLDs. The sphere decoding algorithm naturally works on a finite lattice, where the candidate lattice points are confined in a pre-determined region. However, sphere decoding usually has a higher complexity than sphere encoding because of the extra overhead of the former required for remapping the invalid solutions, which are found outside the boundaries of the finite lattice [192]. Meanwhile, the family of powerful lattice reduction techniques such as LLL are generally not favoured in sphere decoding, since it would be difficult to keep track of the boundary of the finite lattice. Instead, less sophisticated transformations such as the SQRD should be used, which however increases the complexity of the enumeration process. Regarding other characteristics of the existing G.fast system with respect to its potential practical deployment in VP, its backwards compatibility with linear precoding based receivers was addressed in [193], while a particularly robust vectoring feedback error mitigation arrangement was proposed in [194].

As a well-investigated subject, the (depth-first) sphere decoder has both an average-case and a worst-case complexity that grows exponentially with the system's dimensionality [195], which constitutes a substantial disadvantage compared to the deterministic polymonial (quadratic) complexity order of the conventional linear and nonlinear vectoring schemes, as well as compared to the IFP (depending on the choice of the lattice codebook) and to the family of approximate lattice precoding schemes, when ignoring the complexity overhead of initialization. This had led to renewed efforts invested in reducing the complexity of the conventional sphere decoder/encoder [196]–[198], as well as to the optimization of the lattice reduction preprocessing [199]. A rather different practical technique of reducing the complexity to a manageable level is constituted by parallel computing. The branches of the search tree in the sphere decoder cannot be enumerated concurrently in the conventional sense, since the sphere constraint is constantly updated during run-time. As a consequence, each parallel enumeration thread must be able to communicate with other threads in order to achieve a beneficial efficiency boost. The parallel enumeration algorithm of [200] achieves an efficiency improvement proportional to the number of parallel threads. Theoretically, the currently provable best-case algorithm for finding the exact solution of CVP has a complexity order of $\mathcal{O}(2^{\mathcal{O}(2K)})$ relying on a buffer size of $\mathcal{O}(2^{\mathcal{O}(2K)})$ [166] [201], given that the solver has a priori knowledge of the Voronoi cell. However, the particular solver invoked in this case is very impractical for high throughput transmission compared to the sphere encoder, and the performance difference is negligible for medium-sized systems.

Since the VP scheme constitutes a nonlinear expansion of the plain ZFP, the concept of regularization, as mentioned in Sec. III-B2, may also be used for enhancing the performance of VP. The MMSE criterion based VP scheme has been proposed relying on either a continuous-and-discrete hybrid perturbation [202] [203] strategy, or a regularized channel inversion strategy [162] [35]. As seen in Section III-C3, this is to overcome the same modulo loss caused by the non-linear receivers. Additionally, as studied in [39], VP is a flexible scheme that, when expanded over multiple symbol durations, can fully regain the 'shaping loss' via the construction of nested lattice constellations. The expansion of VP over time allows for a more efficient exploitation of the non-causal CSI knowledge (towards the original concept of DPC), whose gain was also empirically characterized in [204] [205].

V. SPECTRUM BALANCING FOR VECTORED TRANSCEIVERS

DSM is a prominent research subject in the history of wireline communications. Based on the historic definition of [100] (Table 3), past research of DSM was mainly concentrating on independent bit-loading based level-2, i.e. DSB, as shown in the landmark contributions of [18], [213]–[216]. However, with the introduction of ITU-T G.993.5, MUP-based level-3 DSM became mandatory. As a consequence, a joint spectrum management strategy must be utilized, whose objective is to simultaneously optimize both the power and constellation assignment for maintaining the same bit error rate (BER) for all tones and all CPEs, while cancelling all known interferences as well as maximizing the sum rate of the vectored DSL binder. We note that depending on the applications, the BER requirement may rarely be unequal within a vectored group [217]. In general, DSB strategies relying on the proverbial 'water-filling' criterion are commonly employed to avoid the 'worst-case dominance' of plain vectoring without DSB.

In our context, the total transmit power is partitioned across two dimensions, namely the frequency and the signal space. On the other hand, the power constraints are valid for each individual pair of a binder, imposed on both a PSD and a total power basis. Despite the knowledge of the theoretical rate region for the majority of the MUP schemes determined under the usual sum-power constraint or per-antenna power contraint in wireless systems, e.g. [126] [218], achieving the optimal multi-level DSM remains a critical research challenge in the developing of wireline access networks in the literature [135], [219]–[221]. This is particularly true for the hostile crosstalk-intensive environment, operating under the radical power constraints of both G.fast and the forthcoming G.mgfast. In this section, we will conduct an empirical case study on the performance of multi-level DSM, i.e. joint DSB-vectoring, employing both the conventional MUPs and the LRMUPs.

A. SYSTEM MODEL

Given the DMT modulated multi-user DSL system of (4), we have to find a set of $T \times K$ appropriate M-QAM constellations, which results in the maximum sum rate, whilst meeting the BER target, which is formulated as:

$$\max \sum_{t=1}^{T} \sum_{k=1}^{K} b_k^t, \text{ where } b_k^t = \log_2(M_k^t) \text{ and } b_k^t \leq b_{\max}, \tag{40}$$

subject to the bit cap $b_{\max}$ and to the per-pair TxPSD mask $\mathbb{P}^t$ as well as to the per-pair ATP budget $\mathbb{A}$ [72]:

$$E\{|x_k^t|^2\} \leq \mathbb{P}^t \qquad \forall t, k, \tag{41}$$

$$\sum_{t=1}^{T} E\{|x_k^t|^2\} \leq \mathbb{A} \qquad \forall k. \tag{42}$$

For vectored DSL systems, the effective channel between the equalized symbol vector $\boldsymbol{z}^t$ and the message symbol vector $\boldsymbol{u}^t$ constitutes a diagonal matrix. Hence the average constellation energy $E\{|u_k^t|^2\}$ of each message symbol's alphabet, i.e. the power allocated to each message symbol, is determined only by the detection SNR requirement of the equalized symbol z_k^t. On the other hand, the choice of

TABLE 5: Milestones in the Development of Vector Perturbation Transmit Precoding.

Year	Authors	Contribution
2005 [124] [162]	Peel and Hochwald *et al.*	First proposed the concept of vector perturbation precoding by extending the plain ZFP (i.e. channel inversion), as a practical method of approaching the capacity of a multi-user/multi-antenna channel. VP may be conceptually regarded as the counterpart to MLD at the transmitter side.
2006-2007 [202] [203]	Kim and Chun and Chua *et al.*	Independently investigated the problem of achieving MMSE criterion based VP by using a real-valued perturbation vector in addition to the original integer vector.
2008-2009 [206] [163]	Ryan *et al.*	Analyzed the performance of VP from a lattice-theoretical perspective and proposed a tight lower bound for the average power of VP-encoded signal.
2008 [207]	Chae *et al.*	Proposed the block diagonalized version of VP as a counterpart to the linear block diagonalization precoding.
2009-2010 [208] [209]	Han *et al.*	Analyzed the modulo loss incurred by VP and proposed an improved input-dependent search algorithm, which also reduces the encoder's complexity.
2010-2011 [210] [211]	Yao *et al.*	First proposed the MBER criterion based VP scheme for a joint transceiver design that achieves improved BER performance over other VP schemes at no complexity sacrifice.
2010 [193]	Vetter and Sun	Addressed the backward compatibility issue of VP precoding where the receivers of a subset of users cannot perform necessary actions such as modulo operation.
2011-2015 [38] [39]	Avner *et al.*	Analyzed the theoretical performance of the VP scheme and its space-time extension. The proposed sum-rate lower bounded of VP asymptotically achieves the sum-capacity at high SNR, regardless of the system's dimensionality. Additionally, the VP scheme was shown to compensate the typical 'shaping loss', which prominently exists in the conventional precoding schemes such as the ZF-THP, with an optimized choice of the encoding lattice.
2012-2013 [212] [35]	Mazrouei-Sebdani and KrzymieÅĎ	Proposed an optimization technique for the MMSE-VP scheme employed under *per-antenna-group* power constraints. This version outperforms the conventional VP at a complexity increase.
2014 [194]	Masouros *et al.*	Proposed a constructive VP scheme that enhances the robustness of the conventional VP scheme. In the face of limited CSI feedback, constructive VP removes the typical 'error floor' at the expense of a slight diversity loss.
2015-2016 [40] [41]	Li and Masouros	First proposed a multi-alphabet joint VP scheme specifically for systems employing adaptive modulations, which fully achieves the optimality of the conventional VP scheme.
2017 [161]	Zhang *et al.*	First proposed the expanded constellation mapping (ECM) scheme based on a power-restricted version of VP for use with G.fast systems. The ECM scheme exhibits an overall performance analogous to that of the conventional VP.

constellation is restricted by the detection SNR in the form of the standard capacity expression:

$$b = \log_2 M = \log_2(1 + \frac{\eta}{\sigma}), \tag{43}$$

where b is the bandwidth efficiency, i.e. the number of bits per message symbol, and η is the detection SNR. For a given square (i.e. even-bit) QAM constellation, its SNR gap σ [222] with respect to the Shannon limit of Gaussian channel is defined in terms of the corresponding symbol error rate (SER) target:

$$\sigma = \frac{1}{3}\left[\sqrt{2}\,\mathrm{erfc}^{-1}\left(\frac{SER}{2}\right)\right]^2, \tag{44}$$

where $\mathrm{erfc}^{-1}(\cdot)$ denotes the inverse (Gaussian) complementary error function. The SER target can be trivially converted to the BER target which we aim for, based on the bit mapping scheme of the constellation. As demonstrated in Fig. 23, each QAM scheme has a specific operating point with respect to the given SER target, which may be calculated individually by the exact SER expression of the QAM constellation relying on either odd or even number of bits per symbol [223] [108].

Given our fixed BER target, (43) represents the unique mapping between the minimum SNR requirement η and the bandwidth efficiency b. However, due to the bit cap $b_{\max}$ and the discrete nature of bit loading, the solution to the combinatorial optimization problem of (40) is obtained using look-up table based methods rather than using the standard water-filling type of power allocation algorithms. For a given $b_{\max}$, we can construct a look-up table for mapping the power allocated to the corresponding choice of constellation with respect to a fixed BER target and the PSD of the AWGN. The general configuration of the spectrum balancing policy within the multi-level DSM operation will be discussed in Section V-A1, while the TxPSD of the MUPs under the designated power assignment policy is characterized in Section V-A2. Finally, the bit loading algorithms will be presented in Section V-B.

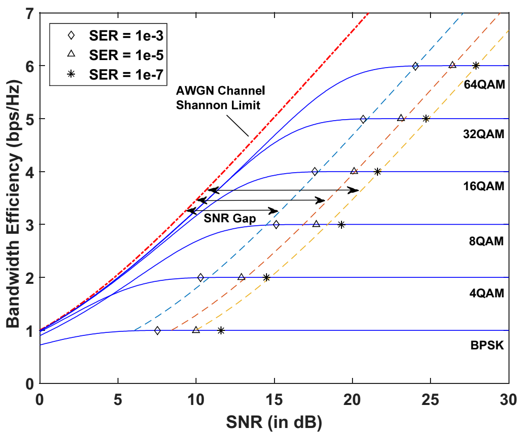

FIGURE 23: The SNR gap towards the Shannon limit of AWGN Channel. For large and even-bit QAM constellations, the SNR gap approximation locates the minimum required SNR for a given symbol error rate target. The suboptimal odd-bit QAM constellations, e.g. BPSK and 8QAM, demonstrate (slightly) wider gaps towards the Shannon limit.

1) Inner Spectrum Balancing Policy

Let the real-valued non-negative diagonal matrix $\boldsymbol{P}^t = \text{diag}\{P_1^t, P_2^t, \cdots, P_K^t\}$ determine the power assigned to tone t. Then the corresponding amplitude scaling matrix is $\boldsymbol{A}^t = \text{diag}\{\sqrt{P_1^t}, \sqrt{P_2^t}, \cdots, \sqrt{P_K^t}\}$. Since diagonal matrices are generally not commutative in multiplication, we cannot apply $\boldsymbol{P}^t$ arbitrarily within the DP. For the plain ZFP and the ZF-THP, the equalized symbol vectors are formulated as:

$$\boldsymbol{z}^t = \boldsymbol{A}^t\boldsymbol{u}^t + \boldsymbol{n}^t \text{ for plain ZFP,} \tag{45}$$

$$\boldsymbol{z}^t = \boldsymbol{A}^t\boldsymbol{u}^t + \Gamma_{\phi^t\boldsymbol{A}^t}\left[(\boldsymbol{W}^t)^{-1}\boldsymbol{n}^t\right] \text{ for ZF-THP,} \tag{46}$$

where $\boldsymbol{W}^t$ follows the same definition of (15) for tone t. Since we have $E\{|u_k^t|^2\} = 1$, the detection SNR experienced by the QAM demapper can be expressed as:

$$\text{SNR}_{k,\text{ZFP}}^t = \frac{P_k^t}{E\{|n_k^t|^2\}}, \tag{47}$$

$$\text{SNR}_{k,\text{THP}}^t = \frac{P_k^t|r_{k,k}^t|^2}{E\{|\bar{n}_k^t|^2\}}, \tag{48}$$

where $\bar{n}_k^t$ represents the AWGN term which corresponds to n_k^t enhanced by the modulo operator (18). This is exactly the 'modulo loss' presented in Section III-C3. Near the typical operating point at the BER target, of say 10^{-7}, the detection SNR is sufficiently high for ensuring that the noise-enhancement effect of the modulo receiver becomes negligible. Therefore we may assume that $\bar{n}_k^t$ follows the Gaussian distribution, and consequently, the SNR gap of (44) is applicable both to the ZFP based linear vectoring, as well as to all other modulo receiver based vectoring schemes. However, for low modulation orders, the modulo loss is not negligible and the exact value may be determined via simulation, as done in [224].

Generally, using a common modulo base ϕ is mandatory, if we apply the family of lattice reduction aided precoding schemes to a vector of symbols drawn from different constellations. The enumeration process of VP discussed in Section IV-D also requires a common modulo base for all message symbols [40]. Therefore, the power allocation conceived for lattice reduction aided precoding has to be applied after the modulo operations. Thus, with the aid of Eq. (32), (35) and (36), we can define the power allocation schemes for the LRMUPs as follows:

$$\boldsymbol{x}^t = \boldsymbol{F}^t\boldsymbol{A}^t\Gamma_{\phi^t}\left[(\boldsymbol{Z}^t)^{-1}\boldsymbol{u}^t\right] \text{ for LR-ZFP,} \tag{49}$$

$$\boldsymbol{x}^t = (\boldsymbol{S}^t)^{-H}\boldsymbol{A}^t(\boldsymbol{x}^t)' \text{ for LR-THP,} \tag{50}$$

$$\boldsymbol{x}^t = \boldsymbol{G}^t\boldsymbol{A}^t(\boldsymbol{u}^t + \boldsymbol{l}^t) \text{ for VP,} \tag{51}$$

where

$$(\boldsymbol{x}^t)' = \Gamma_{\phi^t}\left[\Gamma_{\phi^t}\left[(\boldsymbol{Z}^t)^H\boldsymbol{u}^t\right] + (\boldsymbol{I} - (\boldsymbol{B}^t)^H)(\boldsymbol{x}^t)'\right].$$

However, if $\boldsymbol{A}^t$ is a non-scalar matrix, then the operations in (49) and (50) are no longer capable of fully pre-compensating for the FEXT signal in the context of decentralized receivers. Therefore we have to use a scalar matrix $\boldsymbol{A}^t$ for these approximate lattice precoding schemes, which corresponds to SSB.

On the other hand, for the VP based vectoring scheme, the power assignment policy of (51) affects the choice of the perturbation symbol vector $\boldsymbol{l}^t$. If the scaling matrix $\boldsymbol{A}^t$ of (51) is non-scalar, then the optimization of (29) may no longer necessarily produce the optimal perturbation vector with respect to the power-controlled transmitter of (51). The main reason is that the two lattices $\mathcal{L}(\boldsymbol{G}^t\boldsymbol{A}^t)$ and $\mathcal{L}(\boldsymbol{G}^t)$ are not normally isomorphic. For the sake of analytical tractability, we shall restrict $\boldsymbol{A}^t$ to be a scalar matrix for the VP based vectoring of (51), in which case $\mathcal{L}(\boldsymbol{G}^t\boldsymbol{A}^t)$ and $\mathcal{L}(\boldsymbol{G}^t)$ are isomorphic lattices.

As a result of the discussion above, the scaling matrices $\boldsymbol{A}^t$ in (49), (50) and (51) shall all be scalar matrices, implying that the LRMUPs invoke SSB for level-2 DSM[12]. The equalized symbol vector has the same common form for the LR-ZFP, the LR-THP and VP, expressed as:

$$\boldsymbol{z}^t = \boldsymbol{A}^t\boldsymbol{u}^t + \Gamma_{\phi^t\boldsymbol{A}^t}\left[\boldsymbol{n}^t\right], \tag{52}$$

which results in the detection SNR formulated as:

$$\text{SNR}_{k,\text{LRMUP}}^t = \frac{P_k^t}{E\{|\bar{n}_k^t|^2\}}. \tag{53}$$

2) TxPSD Characterization

Since both (41) and (42) are defined on a per-pair basis, we have to evaluate TxPSD for each active pair, given the power assignment policy and the precoder. Based on [224] and the

[12]This is not necessarily the case for all LRMUPs since DSB is possible in VP-type of MUP such as in the scheme of [42]

general assumption where $E\{|u_k^t|^2\} = 1$, in the plain ZFP and the ZF-THP, $E\{|x_k^t|^2\}$ is related to $\boldsymbol{P}^t$ by:

$$E\{|x_k^t|^2\}|_{\text{ZFP}} = \sum_{j=1}^{K} |g_{k,j}^t|^2 P_j^t \qquad \text{with } g_{k,j}^t \text{ drawn from } \boldsymbol{G}^t, \tag{54}$$

$$E\{|x_k^t|^2\}|_{\text{THP}} = \sum_{j=1}^{K} |Q_{k,j}^t|^2 \rho_j^t P_j^t \qquad \text{with } Q_{k,j}^t \text{ drawn from } \boldsymbol{Q}^t. \tag{55}$$

The modulo operator found in the ZF-THP based transmitter in Fig. 13 causes a slight increase of the average constellation energy characterized by ρ_j^t, whose value is uniquely determined by the constellation b_j^t. We have seen in Section III-C3 that this is the 'precoding loss'. For square M_j^t-QAM constellations, $\rho_j^t = M/(M-1)$ converges quickly to one upon increasing $b_j^t = \log_2 M_j^t$ and can be safely ignored for large b_j^t values. This is exactly the precoding loss previously defined in Section III-C3. The output of both the LR-ZFP-based and of the LR-THP-based transmitters can be characterized in a format similar to (55) based on the power assignment policy of (49) and (50). The pair of approximate lattice precoding schemes follows the same encoding steps constituted by the modulo operations, power allocation and linear filtering, as implied by Eq. (49) and (50). Since the result of modulo operation exhibits the same PSD in the approximate lattice precoding schemes as well as in the ZF-THP, we can characterize the TxPSD of the LR-ZFP and the LR-THP as follows:

$$E\{|x_k^t|^2\}|_{\text{LRZFP}} = \sum_{j=1}^{K} |f_{k,j}^t|^2 \rho^t P^t \text{ with } f_{k,j}^t \text{ drawn from } \boldsymbol{F}^t, \tag{56}$$

$$E\{|x_k^t|^2\}|_{\text{LRTHP}} = \sum_{j=1}^{K} |s_{k,j}^t|^2 \rho^t P^t \text{ with } s_{k,j}^t \text{ drawn from } (\boldsymbol{S}^t)^{-H}. \tag{57}$$

Because the choice of constellation is shared by all pairs on the same tone for the LR-ZFP and the LR-THP, we emphasize this fact by neglecting the pair index j for the terms ρ^t and P^t in Eq. (56) and (57).

In terms of the transmitter's power transfer characteristics between the input $\boldsymbol{u}^t$ and the output $\boldsymbol{x}^t$, Eq. (54), (55), (56) and (57) exhibit a similar format. As stated in Section III-C, the output of the modulo operator $\Gamma_\phi[\cdot]$ of (18) is distributed over the square-shaped region characterized by the set $\mathcal{U}$ of (30). The second moment of $\mathcal{U}$ represents the average energy of a large QAM constellation having the modulo base ϕ. Hence, for the ZF-THP, the LR-ZFP and the LR-THP based vectoring schemes, $\boldsymbol{x}^t$ may be considered as a vector of $\mathcal{U}$-symbols passing through the scaling matrix $\boldsymbol{A}^t$ and an inverse-channel-related matrix thereafter. For the plain ZFP based linear vectoring, modelling the MUP's input as the set $\mathcal{U}$ will penalize the admissible power allocation policy by ρ. Consequently, $\boldsymbol{x}^t$ is distributed over some origin-centred parallelotope in the Euclidean space $\mathbb{C}^K$ or $\mathbb{R}^{2K}$. For each vectoring scheme (except for VP), the parallelotope associated with the specific inverse-channel-related matrix represents the VMR (Fig. 18).

By contrasting (51) to both (49) and (50), it may be readily seen that the VP-based transmitter does not share the aforementioned power transfer characteristics of the other modulo encoders. In fact, $\boldsymbol{x}^t$ is distributed over the origin-centred Voronoi cell of the lattice $\mathcal{L}(\boldsymbol{G}^t)$. Ryan et al. had shown in [163] that for lattices spawned by the inverse of wireless channels, the geometric properties of the resultant Voronoi cell are very similar to those of a hypersphere. In particular, it is demonstrated in [163] that the second moment of a Voronoi cell is closely lower bounded by that of the hypersphere having the same dimension and volume. The second moment of the Voronoi cell of $\mathcal{L}(\boldsymbol{G}^t)$ represents the average total TxPSD of the VP-based transmitter having the input alphabet of $\mathcal{U}/\phi$ and no amplitude scaling, which was formulated in [163] as follows:

$$E\{\|\boldsymbol{G}^t(\boldsymbol{u}^t + \boldsymbol{l}^t)\|^2\} \geq \frac{K\Gamma(K+1)^{1/K}}{(K+1)\pi} \det[(\boldsymbol{H}^t)^H \boldsymbol{H}^t]^{-1/K}. \tag{58}$$

Given the channel characteristics of commercial DSL systems (e.g. Fig. 4), we may observe that the associated channel matrices are quasi-orthogonal (thus diagonally-dominant) over the low frequencies, but this quasi-orthogonality is eroded for the channel matrices above the bandwidth of the G.fast generation in operation at the time of writing. Therefore, the corresponding Voronoi cells are orthotope-like for the low-frequency multi-pair channels and hypersphere-like for the high-frequency ones. For the channel matrices in the high-frequency range, we extend the empirical result of [163] hypothetically by assuming that the mapping region of VP is a hypersphere[13], in which case the equality holds in (58). Since the hypersphere is perfectly isotropic, each VP-encoded symbol x_k^t of the same tone t will have an identical share of the total TxPSD $E\{\|\boldsymbol{x}^t\|^2\}|_{\text{VP}}$. Given that $\boldsymbol{A}^t$ is a scalar matrix in (51), the per-pair TxPSD of the VP encoded symbol vector may be characterized by:

$$E\{|x_k^t|^2\}|_{\text{VP}} = \frac{E\{\|\boldsymbol{x}^t\|^2\}|_{\text{VP}}}{K} P^t. \tag{59}$$

If we disregard the modest effect of the modulo-encoder-related power penalty ρ^t and invoke the identity matrix $\boldsymbol{A}^t = \boldsymbol{I}$ for scaling, then we can compare each vectoring scheme's mapping behaviour graphically using the mapping regions of Fig. 18. For all vectoring schemes except for the VP based one, the output of the symbol encoder is distributed over some parallelogram-shaped region. The shape of the parallelogram depends on the front-end filter matrix (e.g. $\boldsymbol{F}$ of Fig. 19). For the VP based vectoring scheme, the mapping region

[13]Since the hypersphere does not tessellate the Euclidean space $\mathbb{R}^{2K}$, a Voronoi cell cannot be an exact hypersphere. However, using this approximation will allow us to investigate the absolute upper bound, albeit unachievable.

constitutes a hexagon-shaped Voronoi cell in this example. The Voronoi cell of a higher dimensional lattice becomes a general convex polytope, whose exact geometry may be computed using the algorithms of [225] or [226]. However, these algorithms have prohibitively high complexity, hence they are not suitable for vectoring.

B. SPECTRUM BALANCING ALGORITHMS

1) Historical Note

Conventional spectrum balancing strategies are conceived for isolated level-2 DSM in legacy DSL systems. These strategies are largely based on approximations of the convex optimization approaches such as water-filling. In fact, [227] and later [228] established general duality principles between non-convex optimization problems in multi-carrier systems and their convex counterparts. More broadly, DSB approaches can be classified into two main categories:

- **Centralized Algorithms**: Relying on complete CSI knowledge, the DP can employ centralized DSB strategies to achieve optimal spectrum balancing (OSB [229]). Due to the high complexity of OSB in the face of high-dimensional systems, subgroup-based [230] and iteration-based [231] [232] OSB were proposed. The family of centralized algorithms will constitute competitive candidates for the next generation, because they can achieve optimal or near-optimal sum rate. However, they have the main disadvantage of being complicated to reconfigure, if the spectral load has to be adjusted due to unexpected IN strikes.
- **Distributed Algorithms**: Mainly used in legacy systems dispensing with vectoring prior to VDSL2, distributed DSB algorithms are typically outperformed by centralized ones, because each transmitting modem of the distributed regime can only optimize itself. The best known distributed algorithm is the iterative water-filling (IWF) scheme of [233] [213] and by extension the selective IWF scheme of [214]. Distributed algorithms tend to have lower complexity (e.g. the distributed DSB scheme of [234]) than centralized ones, but the complexity of SSB is still the lowest. Additionally, autonomous algorithms [215] [216] relying on a hybrid of centralized and iterative distributed approaches to OSB were shown to have comparable performance to the centralized algorithms in legacy DSL systems.

With respect to vectored DSL systems such G.mgfast and/or G.fast, the DSB paradigm becomes slightly more complicated. In essence, multi-level DSM employs both complex-valued (coordinated QAM signalling, i.e. vectoring) and real-valued (coordinated gain control, i.e. DSB) spectrum management strategies for achieving the optimal sum-rate of a multi-pair channel. From a holistic perspective, the optimal multi-level DSM scheme should ideally aim for jointly optimizing all operational layers defined in Table 3, and for all known interferences [235]. The general multi-level DSM paradigm and its algorithms may be considered as follows depending on the transmission link direction:

- **Upstream**. In the upstream, a multi-pair DSL channel is effectively reduced to a diagonal interference channel, i.e. K independent single-pair channels with only background noise and no crosstalk, when either the linear ZF or the ZF-DFE MUD is invoked. In this case, DSB reduces to a trivial, water-filling-like power allocation. When (weighted-)MMSE crosstalk cancellation is used, the MUD itself becomes coupled with the inner spectrum balancing policy. In this case, a joint optimization is necessary.
- **Downstream**. In accordance with our discussion in Sec. V-A and unlike the case with the upstream, the MUPs can decouple a downstream multi-pair channel to independent single-pair channels, but the inner spectrum balancing policy retains the cross-correlation among all of the copper pairs (cf. Sec. V-A2). In this case we may choose to fix the MUP's configuration and subsequently optimize the inner spectrum balancing policy (for the ZF criterion, cf. Sec. V-A1). Alternatively, we may also jointly optimize the MUP and the spectrum balancing policy (for the MMSE criterion).

For the general system model presented in Sec. V-A, each transmitted QAM symbol exhibits an average TxPSD $E\{|x_k^t|^2\}$ that depends on the power allocation P_k^t for all $k = 1, 2, ..., K$ of the same tone t. Therefore, we will mainly consider centralized DSB algorithms. Furthermore, due both to the non-convexity of VMR computation (which is required for per pair TxPSD and per pair ATP characterization) and to the paucity of literature for joint LRMUP-DSB, we shall use the greedy heuristic bit loading algorithm based on [236] [224] as an extension of the provably optimal single-pair case of [237] for fairly assessing the performances of the vectoring schemes. At the time of writing, we have not found successful application of the duality principle of [227] [228] to overcome the non-convexity associated with lattice reduction. As shown in [238], using the result of the convex optimization as an initial solution is capable of improving the efficiency of the subsequent heuristic algorithm. However, the application is limited to one-dimensional optimization with respect to a single-pair multitone scenario, and extending the approach to 2D remains an open problem. Motivated readers are encouraged to consult [18] for an in-depth survey of the DSB algorithms conceived for DSL transceivers.

2) Greedy Search aided Algorithm

Greedy search based techniques applied to global optimization problems in general do not necessarily lead to the optimal solution. However, for single-pair multi-tone systems, a greedy search based bit loading algorithm has been proven to be optimal. In this section we will use the extended approach of [224] for the multi-pair multi-tone case. Let $F(b_k^t)$ denote the power required for meeting a given BER target, when transmitting at b bits/symbol on tone t of pair k. The Extended Zanatta-Filho algorithm of [224] consists of a pair of consecutive bit-removal phases obeying (41) and (42), respectively. For the first phase, EZF seeks to comply with the TxPSD mask on each tone. The message symbols of all tones are assigned the maximum number of bits $b_{\max}$ and the corresponding power $P_k^t = F(b_{\max})\ \forall t, k$. When using precoding, each transmitted symbol exhibits an average TxPSD of $E\{|x_k^t|^2\}$ that can be computed, given the knowledge of the power assignment policy $\boldsymbol{P}^t$ and the precoder, which we will formulate in Section V-A2. On a given tone t, if the TxPSD constraint is not fully satisfied, EZF will find the specific pair, where the highest TxPSD occurs, which is formulated as $k_{\max} = \arg\max_k(E\{|x_k^t|^2\})$. For the pair $k_{\max}$, the particular pair k^* where removing one bit from $b_{k^*}^t$ would have caused the largest reduction of $E\{|x_{k_{\max}}^t|^2\}$ is selected, and one bit is subtracted from $b_{k^*}^t$. The power assignment policy for that tone t is then updated accordingly and the new maximum TxPSD based on the updated power assignment policy will be handled in a similar way, until the maximum TxPSD of tone t finally obeys the mask. The first phase of EZF iteratively continues to make such reductions, until $E\{|x_k^t|^2\} \leq \mathbb{P}^t$ is achieved $\forall t, k$.

The second half of EZF will seek to comply with the ATP requirement, as detailed in Alg. 3. Starting with the bit loading results gleaned from the first phase, Alg. 3 will then find the specific line $\bar{k}_{\max}$ associated with the largest ATP $\bar{k}_{\max} = \arg\max_k \sum_{t=1}^{T} E\{|x_k^t|^2\}$. It will then seek to find b_{k*}^{t*}, where removing one bit would have caused the largest reduction of $\sum_{t=1}^{T} E\{|x_{\bar{k}_{\max}}^t|^2\}$. These steps are repeated until the ATP constraint is finally satisfied on all lines.

Algorithm 2: TxPSD-Limited Bit Removal

Initialization: $b_k^t = b_{\max} \quad \forall t, k$. $\boldsymbol{P}^t = F(b_{\max})\boldsymbol{I}_K$;
for *all tones* $t = 1, ..., T$ **do**
  while $\max_k(E\{|x_k^t|^2\}) > \mathbb{P}^t$ **do**
    Find the pair $k_{\max} = \arg\max_k E\{|x_k^t|^2\}$;
    for *all candidate pairs* $k = 1, ..., K$ **do**
      Compute $[\Delta E\{|x_{k_{\max}}^t|^2\}]|_k = E\{|x_{k_{\max}}^t|^2\}|_{P_k^t = F(b_k^t)} - E\{|x_{k_{\max}}^t|^2\}|_{P_k^t = F(b_k^t - 1)}$;
    Find the pair $k^* = \arg\max_k [\Delta E\{|x_{k_{\max}}^t|^2\}]|_k$;
    $b_{k^*}^t \leftarrow b_{k^*}^t - 1$;
    Determine the new $\boldsymbol{P}^t$ and $E\{|x_k^t|^2\}$ for all pairs;

VI. BENCHMARK COMPARISONS

In this section, we present comparative simulation results for the benchmark MUP algorithms present in Fig. 10. Their performance will be characterized in terms of the SER and the sum rate. It is worth noting that the results presented in this section only characterize the performances under the particular set of channel measurements portrayed in Fig. 4. However, for channel measurements taken with other DSL binders of the same type and physical parameters, the performance fluctuations should be minimal.

Algorithm 3: ATP-Limited Bit Removal

initialization: b_k^t and $\boldsymbol{P}^t$ from Part 1 $\forall t, k$;
while $\max_k(\sum_{t=1}^{T} E\{|x_k^t|^2\}) > \mathbb{A}$ **do**
  Find the pair $k_{\max} = \arg\max_k \sum_{t=1}^{T} E\{|x_k^t|^2\}$;
  for *all tones* $t = 1, ..., T$ *and pairs* $k = 1, ..., K$ **do**
    Compute $[\Delta E\{|x_{k_{\max}}^t|^2\}]|_{(t,k)} = E\{|x_{k_{\max}}^t|^2\}|_{P_k^t = F(b_k^t)} - E\{|x_{k_{\max}}^t|^2\}|_{P_k^t = F(b_k^t - 1)}$;
  Find $(t^*, k^*) = \arg\max_{t,k} [\Delta E\{|x_{k_{\max}}^t|^2\}]|_{(t,k)}$;
  $b_{k^*}^{t^*} \leftarrow b_{k^*}^{t^*} - 1$;
  Determine the new $\boldsymbol{P}^{t^*}$ and $\sum_{t=1}^{T} E\{|x_k^t|^2\}$ for all pairs;

A. LEVEL-3 MUP PERFORMANCE

In order to compare the performance of each MUP for transmission over the DSL binder having frequency domain channels characterized by Fig. 4, we simulate the average SER of the multi-pair system having SSB for level-2 DSM, versus the average ATP per pair, using the system configuration of Table 6. A power constraint is invoked by normalizing the TxPSD to the peak value of the elements of $\boldsymbol{x}^t$ for ensuring that the constraint is satisfied for all pairs and for each transmission. A quantitative discussion of optimized joint-level DSM will be presented in the next section regarding the sum rate achieved.

TABLE 6: Default Vectoring Configurations

Parameter	Value
Constellation	16QAM
Modulation	DMT
Channel Coding	N/A
Lower Spectral Bound	517.5 kHz
Upper Spectral Bound	212 MHz
Tone Spacing	517.5 kHz
Number of Pairs	10
AWGN	-150 dBm/Hz
Binder Length	100 m
$B_{\max}$	12
$N_{\max}$	12

Fig. 24 demonstrates the SER performance of the vectoring schemes over the expanded 212 MHz G.fast channel profile assuming that the DP as the downstream transmitter has perfect CSI knowledge. The performance of classic linear precoding schemes and the THP schemes, as well as of the LRMUPs is compared. It is clearly seen in Fig. 24 that the best SER achieved by the conventional precoding schemes (THP) is approximatley 8 times higher than the worst-case performance of LRMUP (LR-ZFP) at the recommended operating point [72] of 4 dBm per-pair ATP. At the relaxed 8 dBm per-pair ATP operating point, the SER of conventional precoding is ten times higher than that of the LRMUP. The gain of the linear MMSE precoding over the plain ZFP becomes most prominent at low to medium SNRs, which exceeds that of the LR-ZFP at the ATP of -8 dBm or lower. The sorted THP scheme is seen to be the most advantageous one at medium SNRs, outperforming the ZF-THP. However, the gain of these (SSB-based) linearly improved schemes is insignificant compared to that of lattice reduction.

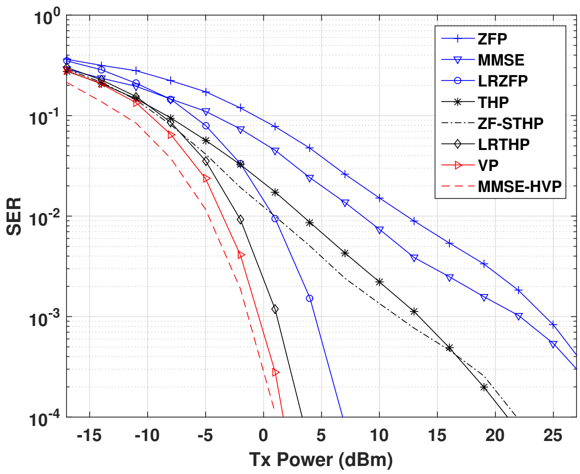

FIGURE 24: Average SER for transmission over the 212 MHz G.fast channel profile versus the average ATP per pair. The message symbol alphabet is 16QAM and a scalar power assignment policy is enforced. The (linear) MMSE scheme is based on the regularized ZFP of [124].

Fig. 25 compares the robustness of the conventional precoding schemes and their lattice reduction aided counterparts in the face of CSI estimation errors, when the vectoring mechanism of Fig. 7 is invoked under the assumption that the forward signalling channel is perfectly time-invariant. The approximate lattice precoding schemes significantly outperform their counterparts operating without lattice reduction. Additionally, it is also apparent that the LR-THP and VP both exhibit high robustness against imperfect CSI estimation, whilst all the other precoding schemes suffer from a substantial SNR loss, as well as from a high SER floor above 10^{-3}. However, the LR-ZFP achieves the same SER performance as the ZF-THP at the 4 dBm operating point, even if the DP has access to perfect transmit CSI knowledge in the case of the ZF-THP. We should note that the CSI estimation error is due to a variety of other sources in practical systems in addition to the quantization error characterized by Fig. 26, including amongst others the influence of impedance mismatching, when quantifying the multi-pair DSL binder's transfer functions relying on the classic transmission line theory. Moreover, the realistic imperfect vectoring feedback channel further aggravates the effects of CSI estimation error.

It has been shown in [136] that more advanced linear MUP

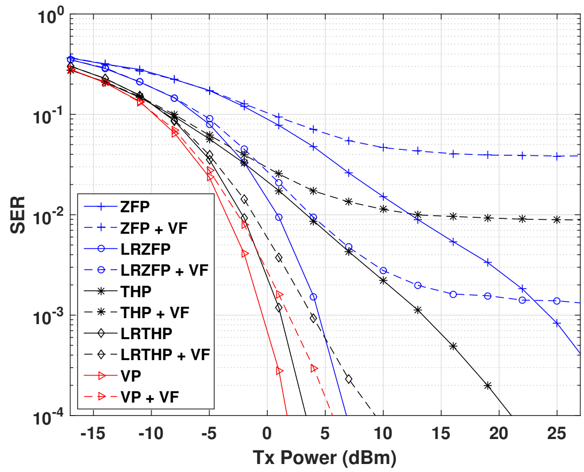

FIGURE 25: Average SER for transmission over the 212 MHz G.fast channel profile with respect to the robustness against imperfect CSI knowledge. The case where the ONU has access to perfect non-causal knowledge of the downstream CSI is compared against the case, where the DP acquires the downstream CSI with the aid of the vectoring feedback loop of Fig. 7.

schemes, such as the one proposed and investigated in [135], may become capable of outperforming the ZF-THP at certain operating points associated with moderate degrees of CSI estimation error, subject to an optimized level-2 DSB policy. This is not observed for the operating point defined in this section based on the standard operations described in both Section II-B2 and [25], due both to the potential difference in the measured channel and to the lack of optimized multi-level DSM in this section. Additionally, we also recommend further investigations of the practical operating point regarding the tolerable degree of CSI imperfection in G.mgfast systems. On the other hand, the seemingly surprising result showing the superiority of linear MUP over the classic ZF-THP reported in [136] was considered to be due to the fact that the ZF-THP as a greedy scheme (i.e. first user gets the best performance) is susceptible to instability, therefore it is more sensitive to imperfect CSI in the face of ill-conditioned multi-pair channel. This observation is consistent with our comparisons and it is explained in the beginning of Section III-C2. In particular, this has led to the conception of stability-improvement schemes such as the sorted THP and the LR-THP. Furthermore, the authors of [239] demonstrated that LR-aided MMSE-SIC, which is the MUD counterpart of the MMSE variant of the LR-THP, may be configured for improving the robustness of its conventional counterpart dispensing with LR.

B. MULTI-LEVEL DSM PERFORMANCE

Perfect DP-side CSI knowledge is assumed for the performance characterization of multi-level (joint level-2 and level-3) spectrum management. In Fig. 26, the throughput per pair is shown for each MUP. The greedy bit loading technique of Alg. 2 and Alg. 3 is invoked under the bit cap of $b_{\max} = 14$ and 4 dBm ATP limit per pair, under the standard TxPSD mask defined in [72]. The SSB policy employed by the LRMUPs constrains their degree of freedom. Thus their performance is compromised as a result of the associated worst-case dominance. However, using the LR-ZFP under the SSB policy will still increase the binder's total sum rate by 6% over that of the plain ZFP using greedy bit loading. Under the general assumption that the DSL channel is quasi-static, the long-term complexity of the LR-ZFP will be identical to that of the plain ZFP, since the additional complexity of initialization can be ignored.

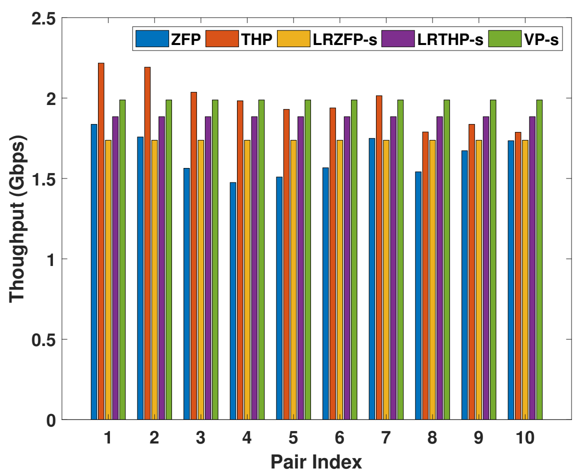

FIGURE 26: Throughput per pair for the vectoring schemes using the 100-meter 10-pair cable characterized by Fig. 3. The TxPSD mask of [72] is invoked and the ATP limit is 4 dBm per pair. The simulation uses an AWGN floor of $N_0 = -150$ dBm/Hz and the bit cap is $b_{\max} = 14$ bits per pair per tone. The power policies for the plain ZFP and the ZF-THP are optimized with the EZF bit loading algorithm of [224], while SSB policies (-s) are employed by the LR-ZFP, the LR-THP and VP.

Fig. 27 quantifies the average bit-loading over all pairs per tone for each MUP. As indicated by the channel quality degradation characterized in Fig. 4 and Fig. 16, the number of supported bits drops at high frequencies for all MUPs. We note that the sum rate of the idealized VP, where the VMR is a perfect hypersphere, is slightly higher than that of the DSB-aided ZF-THP, even though the former does not rely on DSB policy. Even though the hypersphere VMR and subsequently the ideal performance of VP is not achievable in reality, it may be practically achievable using the optimal lattice coding strategy for IFP.

Additionally, it has been identified in Section III-C3 that the only loss of the optimized ZF-THP is the 0.255 bps/Hz 'shaping loss' at high SNRs. However, the influence of the 'modulo loss' becomes significant particularly for the low SNR range as depicted in Fig. 14 and also reported in [224]. At low SNRs, the precoding loss is also considerably higher due to the lower admissible constellation size whose effect has been investigated in Section V-A. Under these considerations, the sum rate achieved by the LR-THP relying on SSB is 4.5% lower than that of the ZF-THP. However, it was discovered in [42] that an alternative MUP whose VMR overlaps with that of the LR-THP achieves the same near-optimality as the optimized ZF-THP. More interestingly, as portrayed in Tab. 7, if all MUPs employ the SSB policy, then the achievable sum rate of the LR-ZFP becomes marginally higher than that of the ZF-THP, despite the fact that the former has lower run-time complexity. However, we should note that this observation heavily relies on the goodness of the (reduced-)lattice basis in the multi-dimensional signal space, which is in practice dependent on the channel's profile.

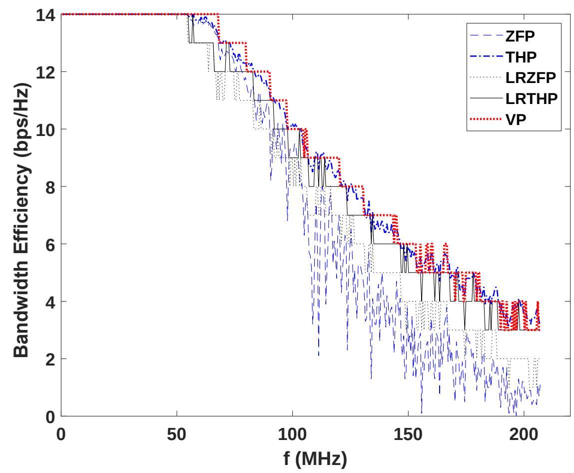

FIGURE 27: Average bit loading per tone over all pairs. The simulation configuration from Fig. 26 is used. The spectral load of the plain ZFP and the ZF-THP are optimized with the greedy bit loading algorithm of Alg. 2 and 3, while SSB policies (-s) are employed by the LR-ZFP, the LR-THP and VP.

Finally, the influence of SDN-aided cross-ISP vectoring is demonstrated in Tab. 7, given that the minimum-complexity SSB policy is used. Assuming that the two vectored groups of five subscribers each from two individual ISPs treat each other as alien FEXT in the case of LLU, we may observe a sum rate boost of upto 86% from employing cross-ISP vectoring. Moreover, both approximate lattice precoding schemes attain a higher net gain in sum rate than their standard counterparts operating without lattice reduction.

VII. PRACTICAL LESSONS

A. COMPLEXITY VS. PERFORMANCE

Shannon's channel capacity quantifies the maximum mutual information associated with a single channel use. When the sum rate associated with multiple DMT symbol durations is considered, the effective channel capacity must be evaluated

TABLE 7: Sum Rate Performance Comparison (Gbps) for 10-pair 100-metre DSL binder

	ZFP	THP	LR-ZFP	LR-THP	VP
LLU-SSB	8.7	9.2	9.7	10.5	N/A
SSB	14.5	17.2	17.4	18.8	19.9
DSB	16.4	19.7	17.4	18.8	19.9

with respect to the processing delay of the MUP algorithms as well. Firstly, the initialization overhead associated with the LLL lattice reduction algorithm and other channel matrix factorization operations such the QR decomposition does not affect the processing delay of the MUP during run-time, and the overhead itself may be deemed affordable on average (i.e. polynomial[14]). Secondly, their operations may be expedited by invoking parallel algorithms, such as the parallel sphere encoder of [200]. This is also conceptually the approach taken both by the K-best [186] and by the fixed complexity sphere decoding (FSD) [241] algorithm, as well as by the parallel THP algorithms of [242] [243]. Additionally, parallel computing can also be used for pipelining the lattice reduction algorithm [244] [245]. However, we should note that these reduced-complexity variants are generally suboptimal compared to their original sequential counterparts, because the former typically ignores a sizeable part of the solution set that has a low probability of containing the global optimum. Nonetheless, even though for low-dimensional systems the performances of low-complexity algorithms match those of their original counterparts sufficiently well, the trade-off must be reinvestigated for large-scale systems.

B. THE NEAR-FAR PROBLEM

Since the telecommunications industry has developed according to a demand-driven model, the design philosophy of access networks is gradually shifting from the network-centric paradigm to user-centric [246] [87], where the main focus becomes quality of experience (QoE) rather than the conventional QoS [247]. From a user-centric perspective, the access network must be capable of providing satisfactory data rate and high reliability for each user, subject to their individual QoE requirement. However, due to the different wire-length of each user in the wireline access network, providing good QoE for all users is a challenging task in the face of the near-far problem.

The detrimental effect of the near-far problem in DSL networks mainly results in the violation of the user fairness. Recall from Fig. 2 that each customer premise and hence the CPE is generally located at a different distance from the G.fast DP. Given the propagation characteristics we discussed in Section II-A, the signals of the users who are far away from the DP are often overwhelmed by those of the users that are closer. The overall performance associated with mixed binder length is typically much worse than that of fixed-length binder due to the worst-case dominance effect. Furthermore, as demonstrated by the results of [137], the performance gap between different MUPs is sometimes also affected by the mixed binder length. The typical solution to the near-far problem in wireless communications is to eliminate the interference with the aid of SIC. However, this solution cannot be readily applied in DSL networks because unlike the mobile terminals in wireless networks, the CPEs cannot be relocated to improve user fairness.

It was shown in [42] that lattice reduction based preprocessing can be used for improving the fairness guarantee of the THP-like approach. In particular, this may be interpreted as the equalization of the channel matrix eigenvalues, representing the balanced CIR of each user. However, the drawback of the LR-aided approach is that the users having good CIRs must sacrifice their performance for the sake of fairness. From a user-centric perspective, multiple lattice basis associated with the channel matrix may be stored so that the precoding matrix may be adjusted based on the predicted QoE requirement.

C. IMPLEMENTATION ISSUES

1) System Imperfections

Approaching the multi-pair multi-tone channel capacity characterized by DPC requires instantaneous and non-causal knowledge of the communication environment, including both the channel matrices and the highly-coloured noise. Practical DSL transceiver units suffer from the following problems:

- Imperfect MUP Design. Firstly, the imperfect transmit CSI obtained via the standard vectoring feedback loop of Fig. 7b causes violation of the ZF signal reception criterion and therefore it results in residual crosstalk. Additionally, the classic THP-aided non-linear vectoring [76] considered by the DSL community is susceptible to instability in the face of imperfect transmit CSI. Hence the THP-based multi-level DSM was outperformed by linear multi-level DSM in the case of [136]. On the other hand, the performance of LR techniques is dependent upon the numerical precision of the LR algorithm, even though our comparisons have shown that perfect LR may improve the robustness of non-linear MUPs. In particular, the numerical stability of certain versions of the LLL algorithm was studied in [191]. Characterizing the realistic performance of non-linear MUPs, including that of the optimal sphere-encoder, requires further research.
- Imperfect Noise Estimation. Secondly, the sum rate achieved by vectored transmission is strongly influenced by the noise level and distribution. The majority of the existing multi-level DSM research assumes the noise to be white, which consists of the typical AWGN plus a noise margin reserved for the worst-case noise bursts. Naturally, a conservative design philosophy will lead to suboptimal performance. Specifically, the capacity

[14]However, the worst-case complexity of the LLL algorithm has been shown to be infinite in [240].

of the multi-channel system is underestimated due to the overestimated average noise power. Since all residual interference may be considered as noise, e.g. alien crosstalk, RFI and IN, it is challenging to construct an accurate model of the exact noise statistics. As we will discuss in the following section, bursty noise constrains the potential adoption of multi-level DSM because of the associated retrain cost. Consequently, learning the noise statistics may improve the overall performance of DSL wireline networks. In particular, hybrid ARQ-based IN protection protocols may become adaptive to the noise environment, and the latency associated with retransmissions may be reduced. In contrast to the typical multi-level DSM which involves only level 2 and level 3, a holistic design that combines all three levels, with the addition of level 1 error control, may result in significantly better performance than what has been achieved with conventional multi-level DSM.

2) Retrain Cost

So far we have assumed that the vectoring control protocol of Section II-B2 only has to be invoked once at the initialization stage to train the VCE, which will then continue to operate for multiple DMT symbol durations. This is in general not a strong assumption concerning the quasi-static nature of DSL channels. However, there exist other factors, which can substantially change the frequency response of a particular DSL binder, such as physically moving or bending the binder at some midpoint. Since DSL binders are normally placed overhead as drop wires or buried underground, the probability of these events are slim. Therefore, the cost associated with retraining the VCE for updating the CSI knowledge does not generally constitute a performance bottleneck.

However, for the 424 MHz G.mgfast profile, we should note that the number of tones is over 8,000 [2]. In this case, the total initialization overhead associated with LLL reduction may potentially exceed the acceptable processing delay for initialization. However, as implied by Fig. 16, LLL reduction is only required for frequencies above 90 MHz. Furthermore, the average complexity of the LLL algorithm and that of the QR decomposition are both of a polynomial order, the latter of which is the mandatory preprocessing for the THP. Therefore, the practicality of the LRMUP is comparable to that of the THP as non-linear MUP candidates.

On the other hand, due to the DSL's susceptibility to the stochastic IN and RFI, the DSB policy that only specifies the AWGN PSD and a static noise margin has to be frequently updated in practice. The SRA protocol is an existing solution, which is capable of providing a real-time DSB policy update without requiring VCE retraining. However, the complexity associated with the optimization of DSB may become a performance bottleneck in the face of IN and RFI. For this reason, the low complexity of the SSB policy is favourable. As shown in Section VI-B, the idealized VP relying on the SSB policy has a similar performance to that of the THP relying on optimized DSB. Therefore, using an optimally-tuned low-complexity sphere encoder aided VP may potentially become the capacity-achieving solution for vectored DSL in the future.

3) Compatibility

We will consider both the backward compatibility with current DSL standards such as G.fast and VDSL2, as well as the forward compatibility with future standards following our vision for the wireline access network. In general, the compatibility problem occurs as a consequence of the multi-standard operation of DSL. As investigated in [147], using the THP in a system mixed with 20% linear receivers does not significantly downgrade the performance compared to the ideal THP transceiver structure of Fig. 13. Since the modulo receivers are commonly used by both the THP and the LRMUPs, the negative impact of legacy linear receivers on the performance of large-scale deployment of non-linear MUP is modest. For the new hardware requirement at the DP side, we have justified that the non-linear optimization block of VP can be incorporated as a simple attachment into the widely deployed ZFP-like architecture. Therefore, the operational expenditure associated with VP should be moderate compared to the alternative non-linear MUP architectures. This is due to the fact that the VP based transmitter is fully compatible with the existing linear MUP, hence its linear front end and the non-linear optimizer can be maintained or replaced independently.

To overcome the bandwidth efficiency limit of the state-of-the-art DSL deployment in preparation for the next generation access network paradigm, the architecture of the current wireline access network has to be fundamentally refined. In this case, the forward compatibility issue results in a two-fold CAPEX trade-off. Firstly, based on the investigation of [26], fibre placement should be prioritized in areas where the CAPEX associated with FTTdp and that of FTTH is comparable. This route requires the corresponding deployment of fibre-based CPEs. Secondly, if FTTH is significantly more expensive than FTTdp, then the unexplored signalling modes of DSL binders should be employed. In general, utilizing the TDSL transmission mode (Section II-A) hidden within the existing DSL binders requires modifications of the critical components of the state-of-the-art CPE hardware, such as adding THz antennas and RF down converters. However, the phantom mode signalling is at the moment a well-established technology, even though it is not widely exploited in the industry yet. For these reasons, the CAPEX incurred by CPE modifications and last mile fibre placement should be carefully assessed and compared.

VIII. CONCLUSIONS

For the forthcoming metallic wireline broadband access network standard G.mgfast, and the converged wireline-wireless network paradigm beyond 5G, the ultimate optimality of both the wireless and the wireline access networks must be achieved. This vision imposes challenges on the wireline communications community, because wireline access networks are used for the ultra high speed and URLLC as the backhaul of the next generation wireless access network.

In this survey, we presented an overview of the state-of-the-art DSL technologies, as well as of the emerging solutions for future wireline network architectures. More specifically, we investigated the dominant challenge of FEXT precancellation in DSL wireline access networks. For the enhanced vectoring approach, we found that lattice reduction significantly improves the performance of conventional MUPs at a modest extra complexity during run-time, under the general assumption that DSL channels are perfectly time-invariant. Furthermore, our performance assessment of the MUPs indicates that the gain achieved by LRMUP does not necessarily rely on complex DSB strategies. For particular channel conditions, the performance of a low-complexity MUP having lattice reduction may be better than a higher complexity one having no lattice reduction. This phenomenon is observed in our performance assessment for the LR-ZFP and the ZF-THP. However, as we have mentioned, the optimal multi-user algorithm is always a potential solver for the NP-hard exact CVP problem. Finding the most efficient exact CVP solver, which should preferably have an average case sub-exponential (or lower) complexity order, is still an open problem in active research. Solving this will be crucial for large-scale (e.g. 100-pair binder) implementation of the phantom mode DSL.

Finally, under the general trend of network unification, we found that (SDN-aided) cross-ISP vectoring allows more efficient exploitation of the multi-dimensional signal space and it nearly doubles the sum rate in a two-ISP LLU scenario. Furthermore, SDN and NFV also support low-cost realization of fibre-level performance over copper, thanks to the recent discovery of 'hidden' signalling modes using the existing telephony-based DSL binders. However, employing these promising new signalling modes requires considerable amount of physical modelling in future research, and the associated CPE modification cost should be carefully compared against that of FTTH for different geographical regions.

ACKNOWLEDGEMENT

We would like to thank Dr. Jochen Maes from Nokia Bell Labs for spotting multiple technical inaccuracies with respect to the state-of-the-art DSL technologies, and Prof. Lie-Liang Yang for his comments about the fundamental principles of multi-user communications.

REFERENCES

[1] M. Ruffini, "Multidimensional Convergence in Future 5G Networks," Journal of Lightwave Technology, vol. 35, no. 3, pp. 535–549, Feb 2017.

[2] R. Strobel, "Physical Layer Framework for the Multi-Gigabit Copper-Fiber Access Networks," Fiber and Integrated Optics, vol. 37, no. 6, pp. 291–313, 2018.

[3] F. van Lingen, M. Yannuzzi, A. Jain, R. Irons-Mclean, O. Lluch, D. Carrera, J. L. Perez, A. Gutierrez, D. Montero, J. Marti, R. Maso, and a. J. P. Rodriguez, "The Unavoidable Convergence of NFV, 5G, and Fog: A Model-Driven Approach to Bridge Cloud and Edge," IEEE Communications Magazine, vol. 55, no. 8, pp. 28–35, Aug 2017.

[4] A. Tzanakaki, M. Anastasopoulos, I. Berberana, D. Syrivelis, P. Flegkas, T. Korakis, D. C. Mur, I. Demirkol, J. Gutierrez, E. Grass, Q. Wei, E. Pateromichelakis, N. Vucic, A. Fehske, M. Grieger, M. Eiselt, J. Bartelt, G. Fettweis, G. Lyberopoulos, E. Theodoropoulou, and D. Simeonidou, "Wireless-Optical Network Convergence: Enabling the 5G Architecture to Support Operational and End-User Services," IEEE Communications Magazine, vol. 55, no. 10, pp. 184–192, Oct 2017.

[5] P. T. Dat, A. Kanno, N. Yamamoto, and T. Kawanishi, "Seamless Convergence of Fiber and Wireless Systems for 5G and Beyond Networks," Journal of Lightwave Technology, pp. 1–1, 2018.

[6] S. Gosselin, A. Pizzinat, X. Grall, D. Breuer, E. Bogenfeld, S. Krauss, J. A. T. Gijon, A. Hamidian, N. Fonseca, and B. Skubic, "Fixed and mobile convergence: which role for optical networks?" IEEE/OSA Journal of Optical Communications and Networking, vol. 7, no. 11, pp. 1075–1083, November 2015.

[7] K. Kanonakis, I. Tomkos, H. Krimmel, F. Schaich, C. Lange, E. Weis, J. Leuthold, M. Winter, S. Romero, P. Kourtessis, M. Milosavljevic, I. N. Cano, and O. Prat, "An OFDMA-based optical access network architecture exhibiting ultra-high capacity and wireline-wireless convergence," IEEE Communications Magazine, vol. 50, no. 8, pp. 71–78, August 2012.

[8] Z. Li, Telecommunication 4.0: Reinvention of the Communication Network. Springer Singapore, 2017.

[9] J. Xie, F. R. Yu, T. Huang, R. Xie, J. Liu, and Y. Liu, "A Survey of Machine Learning Techniques Applied to Software Defined Networking (SDN): Research Issues and Challenges," IEEE Communications Surveys Tutorials, pp. 1–1, 2018.

[10] Q. Mao, F. Hu, and Q. Hao, "Deep Learning for Intelligent Wireless Networks: A Comprehensive Survey," IEEE Communications Surveys Tutorials, vol. 20, no. 4, pp. 2595–2621, Fourthquarter 2018.

[11] F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An Overview on Application of Machine Learning Techniques in Optical Networks," IEEE Communications Surveys Tutorials, pp. 1–1, 2018.

[12] S. Bemby, H. Lu, K. H. Zadeh, H. Bannazadeh, and A. Leon-Garcia, "ViNO: SDN overlay to allow seamless migration across heterogeneous infrastructure," in 2015 IFIP/IEEE International Symposium on Integrated Network Management (IM), May 2015, pp. 782–785.

[13] K. Kerpez and G. Ginis, "Software-Defined Access Network (SDAN)," in 2014 48th Annual Conference on Information Sciences and Systems (CISS), March 2014, pp. 1–6.

[14] K. J. Kerpez, J. M. Cioffi, G. Ginis, M. Goldburg, S. Galli, and P. Silverman, "Software-defined access networks," IEEE Communications Magazine, vol. 52, no. 9, pp. 152–159, Sep. 2014.

[15] R. Zheng, W. Yang, and J. Zhou, "Future access architecture: Software-defined accesss networking," in 2014 IEEE 11th Consumer Communications and Networking Conference (CCNC), Jan 2014, pp. 881–886.

[16] Cisco, "Global Mobile Data Traffic Forecast Update," Feb 2017.

[17] C. Leung, S. Huberman, K. Ho-Van, and T. Le-Ngoc, "Vectored DSL: Potential, Implementation Issues and Challenges," IEEE Commun. Surveys Tuts., vol. 15, no. 4, pp. 1907–1923, Fourth 2013.

[18] S. Huberman, C. Leung, and T. Le-Ngoc, "Dynamic Spectrum Management (DSM) Algorithms for Multi-User xDSL," IEEE Commun. Surveys Tuts., vol. 14, no. 1, pp. 109–130, First 2012.

[19] S. M. Zafaruddin, I. Bergel, and A. Leshem, "Signal Processing for Gigabit-Rate Wireline Communications: An Overview of the State of the Art and Research Challenges," IEEE Signal Processing Magazine, vol. 34, no. 5, pp. 141–164, Sep. 2017.

[20] O. Lamparter, L. Fang, J. Bischoff, M. Reitmann, R. Schwendener, T. Zasowski, and X. Zhang, "Multi-Gigabit over Copper Access Networks: Architectural Evolution and Techno-Economic Analysis," IEEE Communications Magazine, vol. 57, no. 8, pp. 22–27, August 2019.

[21] V. Oksman, R. Strobel, T. Starr, J. Maes, W. Coomans, M. Kuipers, E. B. Tovim, and D. Wei, "MGFAST: A New Generation of Copper Broadband Access," IEEE Communications Magazine, vol. 57, no. 8, pp. 14–21, August 2019.

[22] ITU-BCSD, "The State of Broadband 2018: Broadband Catalyzing Sustainable Development," Septmber 2018.

[23] ITU-T, "G.992.5 : Asymmetric digital subscriber line 2 (ADSL2) transceivers - Extended bandwidth ADSL2 (ADSL2plus)."

[24] ——, "H.265 : High efficiency video coding."

[25] ——, "G.9701 : Fast Access to Subscriber Terminals (G.fast) - Physical Layer Specification."

[26] J. R. Schneir and Y. Xiong, "Cost Assessment of FTTdp Networks with G.fast," IEEE Commun. Mag., vol. 54, no. 8, pp. 144–152, August 2016.

[27] L. Dake, C. Zhaoyun, and W. Wei, "Trends of communication processors," China Communications, vol. 13, no. 1, pp. 1–16, Jan 2016.
[28] M. T. Raza, S. Lu, M. Gerla, and X. Li, "Refactoring Network Functions Modules to Reduce Latencies and Improve Fault Tolerance in NFV," IEEE Journal on Selected Areas in Communications, vol. 36, no. 10, pp. 2275–2287, Oct 2018.
[29] J. M. Cioffi, K. J. Kerpez, C. S. Hwang, and I. Kanellakopoulos, "Terabit DSLs," IEEE Communications Magazine, vol. 56, no. 11, pp. 152–159, November 2018.
[30] ITU-T, "G.993.5 : Self-FEXT Cancellation (Vectoring) for Use with VDSL2 Transceivers."
[31] ——, "G.9807.1 : 10-Gigabit-capable symmetric passive optical network (XGS-PON)."
[32] S. Yang and L. Hanzo, "Fifty Years of MIMO Detection: The Road to Large-Scale MIMOs," IEEE Commun. Surveys Tuts., vol. 17, no. 4, pp. 1941–1988, Fourthquarter 2015.
[33] W. Zhang, S. Qiao, and Y. Wei, "HKZ and Minkowski Reduction Algorithms for Lattice-Reduction-Aided MIMO Detection," IEEE Trans. Signal Process, vol. 60, no. 11, pp. 5963–5976, Nov 2012.
[34] L. Ding, Y. Wang, and J. Zhang, "Complex Minkowski Reduction and a Relaxation for Near-Optimal MIMO Linear Equalization," IEEE Wireless Commun. Lett., vol. 6, no. 1, pp. 38–41, Feb 2017.
[35] M. Mazrouei-Sebdani and W. A. KrzymieÅĎ, "On MMSE Vector-Perturbation Precoding for MIMO Broadcast Channels With Per-Antenna-Group Power Constraints," IEEE Trans. Signal Process., vol. 61, no. 15, pp. 3745–3751, Aug. 2013.
[36] D. Silva, G. Pivaro, G. Fraidenraich, and B. Aazhang, "On Integer-Forcing Precoding for the Gaussian MIMO Broadcast Channel," IEEE Trans. Wireless Commun., vol. 16, no. 7, pp. 4476–4488, July 2017.
[37] C. Ling and J. Belfiore, "Achieving AWGN Channel Capacity With Lattice Gaussian Coding," IEEE Transactions on Information Theory, vol. 60, no. 10, pp. 5918–5929, Oct 2014.
[38] Y. Avner, B. M. Zaidel, and S. Shamai, "On vector perturbation precoding for the MIMO Gaussian broadcast channel," in Proc. IEEE Int Symp. Information Theory, Jul. 2011, pp. 1723–1727.
[39] Y. Avner, B. M. Zaidel, and S. S. Shitz, "On Vector Perturbation Precoding for the MIMO Gaussian Broadcast Channel," IEEE Trans. Inf. Theory, vol. 61, no. 11, pp. 5999–6027, Nov. 2015.
[40] A. Li and C. Masouros, "A Constellation Scaling Approach to Vector Perturbation for Adaptive Modulation in MU-MIMO," IEEE Wireless Commun. Lett., vol. 4, no. 3, pp. 289–292, June 2015.
[41] ——, "A Two-Stage Vector Perturbation Scheme for Adaptive Modulation in Downlink MU-MIMO," IEEE Trans. Veh. Technol., vol. 65, no. 9, pp. 7785–7791, Sep. 2016.
[42] Y. Zhang, R. Zhang, A. F. A. Rawi, and L. Hanzo, "Approximate Perturbation Aided Lattice Encoding (APPLE) for G.fast and Beyond," IEEE Access, vol. 6, pp. 53 438–53 451, 2018.
[43] S. Jagannathan, V. Pourahmad, K. Seong, J. M. Cioffi, M. Ouzzif, and R. Tarafi, "Common-mode data transmission using the binder sheath in digital subscriber lines," IEEE Transactions on Communications, vol. 57, no. 3, pp. 831–840, March 2009.
[44] W. Foubert, C. Neus, L. V. Biesen, and Y. Rolain, "Exploiting the Phantom-Mode Signal in DSL Applications," IEEE Transactions on Instrumentation and Measurement, vol. 61, no. 4, pp. 896–902, April 2012.
[45] L. F. Alloin, P. Biyani, L. Pierrugues, S. Zafaruddin, and C. V. Kumar, "DM-CM diversity receiver for a wireline communication system," May 26 2015, uS Patent 9,042,498.
[46] ITU-T, "G.995.2 : Enhanced common mode limits and measurement methods for customer premises equipment operating on copper pairs."
[47] D. Acatauassu, S. HÃűst, C. Lu, M. Berg, A. Klautau, and P. O. BÃűrjesson, "Simple and Causal Copper Cable Model Suitable for G.fast Frequencies," IEEE Trans. Commun., vol. 62, no. 11, pp. 4040–4051, Nov 2014.
[48] D. Gomes, E. Medeiros, A. Klautau, and E. Pelaes, "Mitigation of Alien Crosstalk for Downstream DSL Impaired by Multiple Interferers," IEEE Communications Letters, vol. 21, no. 11, pp. 2380–2383, Nov 2017.
[49] A. A. Amayreh, J. L. Masson, M. HÃl'lard, and M. Ouzzif, "Alien crosstalk elimination in digital subscriber line systems," IET Communications, vol. 8, no. 10, pp. 1714–1723, July 2014.
[50] G. Ginis and C.-N. Peng, "Alien Crosstalk Cancellation for Multipair Digital Subscriber Line Systems," EURASIP Journal on Advances in Signal Processing, vol. 2006, no. 1, p. 016828, Feb 2006.
[51] K. M. Ali, G. G. Messier, and S. W. Lai, "DSL and PLC Co-Existence: An Interference Cancellation Approach," IEEE Transactions on Communications, vol. 62, no. 9, pp. 3336–3350, Sep. 2014.
[52] S. Galli, K. J. Kerpez, H. Mariotte, and F. Moulin, "PLC-to-DSL Interference: Statistical Model and Impact on VDSL2, Vectoring, and G.Fast," IEEE Journal on Selected Areas in Communications, vol. 34, no. 7, pp. 1992–2005, July 2016.
[53] M. Sorbara, P. Duvaut, F. Shmulyian, and S. Singh, "Construction of a DSL-MIMO channel model for evaluation of FEXT cancellation systems in VDSL2," in Sarnoff Symposium 2007. Nassau Inn, Princeton, NJ: IEEE, Apr 2007, pp. 1 – 6.
[54] J. Maes, M. Guenach, and M. Peeters, "Statistical MIMO Channel Model for Gain Quantification of DSL Crosstalk Mitigation Techniques," in 2009 IEEE International Conference on Communications, June 2009, pp. 1–5.
[55] T. Starr, J. Cioffi, and P. Silverman, Understanding Digital Subscriber Line Technology, ser. Prentice Hall communications engineering and emerging technologies series. Prentice Hall PTR, 1999, no. v. 1.
[56] R. F. M. van den Brink, "Modeling the Dual-Slope Behavior of in-Quad EL-FEXT in Twisted Pair Quad Cables," IEEE Trans. Commun., vol. 65, no. 5, pp. 2153–2163, May 2017.
[57] O. Ogundapo, A. Duffy, and C. Nche, "Parameter for near end crosstalk prediction in twisted pair cables," in 2016 IEEE International Symposium on Electromagnetic Compatibility (EMC), July 2016, pp. 485–490.
[58] P. Lafata, "Realistic Modeling of Far-End Crosstalk in Metallic Cables," IEEE Communications Letters, vol. 17, no. 3, pp. 435–438, March 2013.
[59] V. Belevitch, "Theory of Proximity Effect in Multiwire Cables. 2." PHILIPS RESEARCH REPORTS, vol. 32, no. 2, pp. 96–117, 1977.
[60] E. Oswald, "Application of A Phantom Circuit for xDSL Communications," in Proc. 13th International OFDM-Workshop, Aug 2008.
[61] ITU-T, "G.998.1: ATM-Based Multi-Pair Bonding."
[62] ——, "G.998.2: Ethernet-Based Multi-Pair Bonding."
[63] ——, "G.998.3: Multi-Pair Bonding Using Time-Division Inverse Multiplexing."
[64] W. Coomans, R. B. Moraes, K. Hooghe, A. Duque, J. Galaro, M. Timmers, A. J. van Wijngaarden, M. Guenach, and J. Maes, "XG-FAST: Towards 10 Gb/s copper access," in 2014 IEEE Globecom Workshops (GC Wkshps), Dec 2014, pp. 630–635.
[65] K. Wang and D. M. Mittleman, "Metal wires for terahertz wave guiding," Nature, vol. 432, no. 7015, p. 376, 2004.
[66] D. M. Mittleman, "Frontiers in terahertz sources and plasmonics," Nature Photonics, vol. 7, no. 9, p. 666, 2013.
[67] T.-I. Jeon, J. Zhang, and D. Grischkowsky, "THz Sommerfeld wave propagation on a single metal wire," Applied Physics Letters, vol. 86, no. 16, p. 161904, 2005.
[68] J. C. Wiltse, "Surface-Wave Propagation on a Single Metal Wire or Rod at Millimeter-Wave and Terahertz Frequencies," in 2006 IEEE MTT-S International Microwave Symposium Digest, June 2006, pp. 970–973.
[69] A. Markov, S. Gorgutsa, H. Qu, and M. Skorobogatiy, "Practical metal-wire THz waveguides," arXiv preprint arXiv:1206.2984, 2012.
[70] ITU-T, "G.993.2 : Very high speed digital subscriber line transceivers 2 (VDSL2)."
[71] W. Shieh, "OFDM for Flexible High-Speed Optical Networks," Journal of Lightwave Technology, vol. 29, no. 10, pp. 1560–1577, May 2011.
[72] ITU-T, "G.9700 : Fast Access to Subscriber Terminals (G.fast) - Power Spectral Density Specification."
[73] G. Fettweis, M. Krondorf, and S. Bittner, "GFDM - Generalized Frequency Division Multiplexing," in VTC Spring 2009 - IEEE 69th Vehicular Technology Conference, April 2009, pp. 1–4.
[74] B. Farhang-Boroujeny and H. Moradi, "Derivation of GFDM based on OFDM principles," in 2015 IEEE International Conference on Communications (ICC), June 2015, pp. 2680–2685.
[75] C.-L. Tai, B. Su, and C. Jia, "Frequency-Domain Decoupling for MIMO-GFDM Spatial Multiplexing," arXiv preprint arXiv:1803.06448, 2018.
[76] G. Ginis and J. M. Cioffi, "Vectored Transmission for Digital Subscriber Line Systems," IEEE J. Sel. Areas Commun., vol. 20, no. 5, pp. 1085 – 1104, June 2002.
[77] D. Hincapie, G. Maierbacher, and M. Leibiger, "Rate and Reach Gains of Vectored DSL in the Current Access Network," in Broadband Coverage in Germany. 9th ITG Symposium. Proceedings, April 2015, pp. 1–6.
[78] J. Rückert, R. Bifulco, M. Rizwan-Ul-Haq, H. Kolbe, and D. Hausheer, "Flexible traffic management in broadband access networks using Software Defined Networking," in 2014 IEEE Network Operations and Management Symposium (NOMS), May 2014, pp. 1–8.

[79] H. Cramér, Mathematical methods of statistics (PMS-9). Princeton university press, 1946, vol. 9.
[80] C. Rao, "Information and accuracy attanaible in the estimation of statistical parameters," Bull. Calcutta Math. Soc., vol. 37, pp. 81–91, 1945.
[81] J. Louveaux and A. . V. D. Veen, "Adaptive DSL Crosstalk Precancellation Design Using Low-Rate Feedback From End Users," IEEE Signal Processing Letters, vol. 13, no. 11, pp. 665–668, Nov 2006.
[82] I. Binyamini and I. Bergel, "Adaptive Precoder Using Sign Error Feedback for FEXT Cancellation in Multichannel Downstream VDSL," IEEE Transactions on Signal Processing, vol. 61, no. 9, pp. 2383–2393, May 2013.
[83] Z. Shi, Q. Wang, J. Jin, D. Jiang, and G. Liu, "Achievability of the channel reciprocity and its benefit in TDD system," in 2010 5th International ICST Conference on Communications and Networking in China, Aug 2010, pp. 1–4.
[84] J. Zhang, S. Chen, R. Zhang, A. F. A. Rawi, and L. Hanzo, "Differential Evolution Algorithm Aided Turbo Channel Estimation and Multi-User Detection for G.Fast Systems in the Presence of FEXT," IEEE Access, vol. 6, pp. 33 111–33 128, 2018.
[85] H. Mariotte, "Overview of ITU-T SG15 Q4 xDSL and G.(mg)fast," May 2017.
[86] D. Kim, H. Lee, and D. Hong, "A Survey of In-Band Full-Duplex Transmission: From the Perspective of PHY and MAC Layers," IEEE Communications Surveys Tutorials, vol. 17, no. 4, pp. 2017–2046, Fourthquarter 2015.
[87] V. Wong, R. Schober, D. Ng, and L. Wang, Key Technologies for 5G Wireless Systems. Cambridge University Press, 2017.
[88] S. K. Sharma, T. E. Bogale, L. B. Le, S. Chatzinotas, X. Wang, and B. Ottersten, "Dynamic Spectrum Sharing in 5G Wireless Networks With Full-Duplex Technology: Recent Advances and Research Challenges," IEEE Communications Surveys Tutorials, vol. 20, no. 1, pp. 674–707, Firstquarter 2018.
[89] X. Wang, H. Huang, and T. Hwang, "On the Capacity Gain from Full Duplex Communications in a Large Scale Wireless Network," IEEE Transactions on Mobile Computing, vol. 15, no. 9, pp. 2290–2303, Sep. 2016.
[90] X. Xie and X. Zhang, "Does full-duplex double the capacity of wireless networks?" in IEEE INFOCOM 2014 - IEEE Conference on Computer Communications, April 2014, pp. 253–261.
[91] N. Ehtiati and B. Champagne, "A General Framework for Mixed-Domain Echo Cancellation in Discrete Multitone Systems," IEEE Transactions on Communications, vol. 61, no. 2, pp. 769–780, February 2013.
[92] R. Nongpiur, D. J. Shpak, and A. Antoniou, "An Analysis of a Near-End Crosstalk Cancelation System That Uses Adaptive Filters," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 55, no. 10, pp. 3306–3316, Nov 2008.
[93] P. Tsiaflakis, Y. Lefevre, W. Coomans, and J. Maes, "Friendly Full Duplex: A Multi-User Full Duplex Method for MGfast in Coexistence with G.fast," in 2018 IEEE Global Communications Conference (GLOBECOM), Dec 2018, pp. 1–6.
[94] ITU-T, "G.998.4 : Improved Impulse Noise Protection for Digital Subscriber Line (DSL) Transceivers."
[95] T. Bai, H. Zhang, R. Zhang, L. L. Yang, A. F. A. Rawi, J. Zhang, and L. Hanzo, "Discrete Multi-Tone Digital Subscriber Loop Performance in the Face of Impulsive Noise," IEEE Access, vol. 5, pp. 10 478–10 495, 2017.
[96] F. Sjoberg, R. Nilsson, P. O. Borjesson, P. Odling, B. Wiese, and J. A. C. Bingham, "Digital RFI suppression in DMT-based VDSL systems," IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 51, no. 11, pp. 2300–2312, Nov 2004.
[97] M. Mirahmadi, A. Al-Dweik, and A. Shami, "BER Reduction of OFDM Based Broadband Communication Systems over Multipath Channels with Impulsive Noise," IEEE Transactions on Communications, vol. 61, no. 11, pp. 4602–4615, November 2013.
[98] J. Neckebroek, M. Moeneclaey, M. Guenach, M. Timmers, and J. Maes, "Comparison of error-control schemes for high-rate communication over short DSL loops affected by impulsive noise," in 2013 IEEE International Conference on Communications (ICC), June 2013, pp. 4014–4019.
[99] T. Bai, C. Xu, R. Zhang, A. F. Al Rawi, and L. Hanzo, "Performance of HARQ-Assisted OFDM Systems Contaminated by Impulsive Noise: Finite-Length LDPC Code Analysis," IEEE Access, vol. 7, pp. 14 112–14 123, 2019.
[100] J. Cioffi, M. Brady, V. Pourahmad, S. Jagannathan, W. Lee, Y. Kim, C. Chen, K. Seong, D. Yu, M. Ouzzif, H. Mariotte, R. Tarafi, G. Ginis, B. Lee, T. Chung, and P. Silverman, "Vectored DSLs with DSM: The road to Ubiquitous Gigabit DSLs," in Proc. World Telecommun. Conf., May 2006.
[101] T. Bai, H. Zhang, J. Zhang, C. Xu, A. F. A. Rawi, and L. Hanzo, "Impulsive Noise Mitigation in Digital Subscriber Lines: the State-of-the-Art and Research Opportunities," submitted to IEEE Communications Magazine, 2019.
[102] U. Challita, L. Dong, and W. Saad, "Proactive Resource Management for LTE in Unlicensed Spectrum: A Deep Learning Perspective," IEEE Transactions on Wireless Communications, vol. 17, no. 7, pp. 4674–4689, July 2018.
[103] S. Sendra, A. Rego, J. Lloret, J. M. Jimenez, and O. Romero, "Including artificial intelligence in a routing protocol using Software Defined Networks," in 2017 IEEE International Conference on Communications Workshops (ICC Workshops), May 2017, pp. 670–674.
[104] R. Alvizu, S. Troia, G. Maier, and A. Pattavina, "Matheuristic with machine-learning-based prediction for software-defined mobile metro-core networks," IEEE/OSA Journal of Optical Communications and Networking, vol. 9, no. 9, pp. D19–D30, Sep. 2017.
[105] J. Zhang, S. Chen, X. Mu, and L. Hanzo, "Evolutionary-Algorithm-Assisted Joint Channel Estimation and Turbo Multiuser Detection/Decoding for OFDM/SDMA," IEEE Transactions on Vehicular Technology, vol. 63, no. 3, pp. 1204–1222, March 2014.
[106] N. Samuel, T. Diskin, and A. Wiesel, "Deep MIMO detection," in 2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), July 2017, pp. 1–5.
[107] T. OâĂŹShea and J. Hoydis, "An Introduction to Deep Learning for the Physical Layer," IEEE Transactions on Cognitive Communications and Networking, vol. 3, no. 4, pp. 563–575, Dec 2017.
[108] Goldsmith, A., Wireless Communications. Cambridge University Press, 2005.
[109] M. Costa, "Writing on Dirty Paper," IEEE Trans. Inf. Theory, vol. 29, no. 3, pp. 439 – 441, May 1983.
[110] A. Wyner, "Recent results in the Shannon theory," IEEE Transactions on Information Theory, vol. 20, no. 1, pp. 2–10, January 1974.
[111] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," IEEE Transactions on Information Theory, vol. 19, no. 4, pp. 471–480, July 1973.
[112] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," IEEE Transactions on Information Theory, vol. 22, no. 1, pp. 1–10, January 1976.
[113] L. Yang, "Multiuser Transmission Via Multiuser Detection: Altruistic-Optimization and Egocentric-Optimization," in 2007 IEEE 65th Vehicular Technology Conference - VTC2007-Spring, April 2007, pp. 1921–1925.
[114] ——, Multicarrier Communications. Wiley, 2009.
[115] L. Wang, L. Xu, S. Chen, and L. Hanzo, "Three-Stage Irregular Convolutional Coded Iterative Center-Shifting K-Best Sphere Detection for Soft-Decision SDMAâĂŞOFDM," IEEE Transactions on Vehicular Technology, vol. 58, no. 4, pp. 2103–2109, May 2009.
[116] M. Tomlinson, "New Automatic Equaliser Employing Modulo Arithmetic," Electronics Letters, vol. 7, no. 5, pp. 138 – 139, Mar 1971.
[117] H. Harashima and H. Miyakawa, "Matched-Transmission Technique for Channels With Intersymbol Interference," IEEE Trans. Commun., vol. 20, no. 4, pp. 774 – 780, Aug 1972.
[118] V. Oksman and B. Heise, "Method for Seamless Bit Rate Adaptation for Multicarrier DSL," Apr. 14 2009, US Patent 7,519,124.
[119] R. Cendrillon, M. Moonen, T. Bostoen, and G. Ginis, "The Linear Zero-Forcing Crosstalk Canceler is Near-Optimal in DSL Channels," in Global Telecommunications Conference, vol. 4. IEEE, Nov 2004, pp. 2334 – 2338.
[120] R. Cendrillon, G. Ginis, E. V. den Bogaert, and M. Moonen, "A Near-Optimal Linear Crosstalk Precoder for Downstream VDSL," IEEE Transactions on Communications, vol. 55, no. 5, pp. 860–863, May 2007.
[121] R. Cendrillon, M. Moonen, J. Verlinden, T. Bostoen, and G. Ginis, "Improved Linear Crosstalk Precompensation for DSL," in 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 4, May 2004, pp. iv–1053–6 vol.4.
[122] R. Cendrillon, G. Ginis, M. Moonen, and K. Van Acker, "Partial crosstalk precompensation in downstream VDSL," Signal Processing, vol. 84, no. 11, pp. 2005–2019, 2004.
[123] A. Leshem and L. Youming, "A Low Complexity Linear Precoding Technique for Next Generation VDSL Downstream Transmission Over

Copper," IEEE Transactions on Signal Processing, vol. 55, no. 11, pp. 5527–5534, Nov 2007.
[124] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multiantenna Multiuser Communication-Part I: Channel Inversion and Regularization," IEEE Trans. Commun., vol. 53, no. 1, pp. 195 – 202, Jan 2005.
[125] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," IEEE Transactions on Communications, vol. 49, no. 12, pp. 2198–2206, Dec 2001.
[126] S. S. Christensen, R. Agarwal, E. D. Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," IEEE Transactions on Wireless Communications, vol. 7, no. 12, pp. 4792–4799, December 2008.
[127] Q. Shi, M. Razaviyayn, Z. Luo, and C. He, "An Iteratively Weighted MMSE Approach to Distributed Sum-Utility Maximization for a MIMO Interfering Broadcast Channel," IEEE Transactions on Signal Processing, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.
[128] M. Sadek, A. Tarighat, and A. H. Sayed, "A Leakage-Based Precoding Scheme for Downlink Multi-User MIMO Channels," IEEE Trans. Wireless Commun., vol. 6, no. 5, pp. 1711 – 1721, May 2007.
[129] P. A. Lopes and J. A. Gerald, "Leakage-based precoding algorithms for multiple streams per terminal MU-MIMO systems," Digital Signal Processing, vol. 75, pp. 38–44, 2018.
[130] P. Patcharamaneepakorn, S. Armour, and A. Doufexi, "On the Equivalence Between SLNR and MMSE Precoding Schemes with Single-Antenna Receivers," IEEE Commun. Lett., vol. 16, no. 7, pp. 1034 – 1037, July 2012.
[131] J. Louveaux and A. van der Veen, "Adaptive Precoding for Downstream Crosstalk Precancelation in DSL Systems Using Sign-Error Feedback," IEEE Transactions on Signal Processing, vol. 58, no. 6, pp. 3173–3179, June 2010.
[132] I. Bergel, S. Zafaruddin, and A. Leshem, "Large Matrix Asymptotic Analysis of ZF and MMSE Crosstalk Cancelers for Wireline Channels," arXiv preprint arXiv:1803.06019, 2018.
[133] Y. Jiang, M. K. Varanasi, and J. Li, "Performance Analysis of ZF and MMSE Equalizers for MIMO Systems: An In-Depth Study of the High SNR Regime," IEEE Transactions on Information Theory, vol. 57, no. 4, pp. 2008–2026, April 2011.
[134] K. R. Kumar, G. Caire, and A. L. Moustakas, "Asymptotic Performance of Linear Receivers in MIMO Fading Channels," IEEE Transactions on Information Theory, vol. 55, no. 10, pp. 4398–4418, Oct 2009.
[135] W. Lanneer, P. Tsiaflakis, J. Maes, and M. Moonen, "Linear and Nonlinear Precoding Based Dynamic Spectrum Management for Downstream Vectored G.fast Transmission," IEEE Trans. Commun., vol. 65, no. 3, pp. 1247–1259, March 2017.
[136] J. Maes, C. Nuzman, and P. Tsiaflakis, "Sensitivity of nonlinear precoding to imperfect channel state information in G.fast," in 2016 24th European Signal Processing Conference (EUSIPCO), Aug 2016, pp. 290–294.
[137] L. Stigant, R. Zhang, L. D. Humphrey, T. P. Linney, T. Morsman, and A. F. Al Rawi, "Spatially Adaptive Linear Precoding for Systems Beyond G.Fast," in 2018 IEEE Global Communications Conference (GLOBECOM), Dec 2018, pp. 1–6.
[138] G. Ginis and J. M. Cioffi, "A Multi-User Precoding Scheme Achieving Crosstalk Cancellation with Application to DSL Systems," in 34th Asilomar Conference on Signals, Systems and Computers, vol. 2. Pacific Grove, CA, USA: IEEE, Oct 2000, pp. 1627 – 1631.
[139] M. Hekrdla, A. Matera, W. Wang, D. Wei, and U. Spagnolini, "Ordered Tomlinson-Harashima Precoding in G.fast Downstream," in 2015 IEEE Global Communications Conference (GLOBECOM), Dec 2015, pp. 1–6.
[140] K. Zu, F. C. B. F. Müller, C. Lu, P. Eriksson, and A. Klautau, "Rate Balancing Based TomlinsonâĂŞHarashima Precoding for G.Fast Systems," IEEE Communications Letters, vol. 20, no. 8, pp. 1519–1522, Aug 2016.
[141] R. Habendorf and G. Fettweis, "On Ordering Optimization for MIMO Systems with Decentralized Receivers," in International Symposium on Intelligent Signal Processing and Communications Systems, vol. 4. Melbourne: IEEE, May 2006, pp. 1844 – 1848.
[142] S. Kinjo and S. Ohno, "A Tomlinson-Harashima Precoding for Correlated MIMO Channels," in International Symposium on Intelligent Signal Processing and Communications Systems. Naha: IEEE, Nov 2013, pp. 257 – 262.
[143] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An Architecture for Realizing Very High Data Rates Over the Rich-Scattering Wireless Channel," in URSI International Symposium on Signals, Systems, and Electronics. Pisa: IEEE, Oct 1998, pp. 295 – 300.
[144] D. Wubben, J. Rinas, R. Bohnke, V. Kuhn, and K. D. Kammeyer, "Efficient Algorithm for Detecting Layered Space-Time Codes," in International ITG Conference on Source and Channel Coding, Berlin, Jan 2002, pp. 399 – 405.
[145] M. L. Honig, Capacity-Approaching Multiuser Communications over Multiple Input/Multiple Output Broadcast Channels. Wiley-IEEE Press, 2009, pp. 384–.
[146] J. Maes, "Making G.fast Faster," in 2016 G.fast Summit, May 2016.
[147] R. Strobel, A. Barthelme, and W. Utschick, "Implementation aspects of nonlinear precoding for G.fast âĂŤ coding and legacy receivers," in 2017 25th European Signal Processing Conference (EUSIPCO), Aug 2017, pp. 111–115.
[148] R. D. Wesel and J. M. Cioffi, "Achievable rates for Tomlinson-Harashima precoding," IEEE Transactions on Information Theory, vol. 44, no. 2, pp. 824–831, March 1998.
[149] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "On shaping gain in the nonlinear fiber-optic channel," in 2014 IEEE International Symposium on Information Theory. IEEE, 2014, pp. 2794–2798.
[150] A. Barthelme, R. Strobel, M. Joham, and W. Utschick, "Weighted MMSE Tomlinson-Harashima Precoding for G.fast," in 2016 IEEE Global Communications Conference (GLOBECOM), Dec 2016, pp. 1–6.
[151] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," IEEE Transactions on Wireless Communications, vol. 3, no. 4, pp. 1305–1316, July 2004.
[152] J. W. Bos, C. Costello, M. Naehrig, and D. Stebila, "Post-Quantum Key Exchange for the TLS Protocol from the Ring Learning with Errors Problem," in 2015 IEEE Symposium on Security and Privacy, May 2015, pp. 553–570.
[153] A. Lenstra, H. Lenstra, and L. Lovász, "Factoring Polynomials with Rational Coefficients," Math. Ann., vol. 261, pp. 515–534, 1982.
[154] P. Nguyen and B. Vallée, The LLL Algorithm: Survey and Applications, ser. Information Security and Cryptography. Springer Berlin Heidelberg, 2009.
[155] D. Wubben, D. Seethaler, J. Jalden, and G. Matz, "Lattice Reduction," IEEE Signal Processing Magazine, vol. 28, no. 3, pp. 70–91, May 2011.
[156] H. Minkowski, "Ueber die positiven quadratischen Formen und über kettenbruchähnliche Algorithmen." Journal fÃijr die reine und angewandte Mathematik, vol. 107, pp. 278–297, 1891.
[157] L. Babai, "On Lovász' Lattice Reduction and the Nearest Lattice Point Problem," Combinatorica, vol. 6, no. 1, pp. 1–13, 1986.
[158] J. Conway and N. Sloane, Sphere Packings, Lattices, and Groups, ser. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1993.
[159] G. Hanrot, X. Pujol, and D. Stehlé, "Algorithms for the shortest and closest lattice vector problems," in International Conference on Coding and Cryptology. Springer, 2011, pp. 159–190.
[160] C. Ling, "Towards Characterizing the Performance of Approximate Lattice Decoding in MIMO Communications," in 6th International ITG-Conference on Source and Channel Coding, April 2006, pp. 1–6.
[161] R. Zhang, A. F. A. Rawi, L. D. Humphrey, and L. Hanzo, "Expanded Constellation Mapping for Enhanced Far-End-Cross-Talk Cancellation in G.fast," IEEE Commun. Lett., vol. 21, no. 1, pp. 56–59, Jan 2017.
[162] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multiantenna Multiuser Communication-Part II: Perturbation," IEEE Trans. Commun., vol. 53, no. 3, pp. 537 – 544, Mar 2005.
[163] D. J. Ryan, I. B. Collings, I. V. L. Clarkson, and R. W. Heath, "Performance of Vector Perturbation Multiuser MIMO Systems with Limited Feedback," IEEE Trans. Commun., vol. 57, no. 9, pp. 2633 – 2644, Oct 2009.
[164] C. Windpassinger, R. F. H. Fischer, and J. B. Huber, "Lattice-Reduction-Aided Broadcast Precoding," IEEE Trans. Commun., vol. 52, no. 12, pp. 2057–2060, Dec 2004.
[165] S. Lyu and C. Ling, "Hybrid Vector Perturbation Precoding: The Blessing of Approximate Message Passing," IEEE Transactions on Signal Processing, vol. 67, no. 1, pp. 178–193, Jan 2019.
[166] D. Micciancio and P. Voulgaris, "A deterministic single exponential time algorithm for most lattice problems based on Voronoi cell computations," SIAM Journal on Computing, vol. 42, no. 3, pp. 1364–1391, 2013.
[167] A. Korkine and G. Zolotareff, "Sur les formes quadratiques," Math. Ann., vol. 6, no. 3, pp. 366–389, 1873.

[168] D. Micciancio, "Efficient reductions among lattice problems," in Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics, 2008, pp. 84–93.
[169] Y. H. Gan, C. Ling, and W. H. Mow, "Complex Lattice Reduction Algorithm for Low-Complexity Full-Diversity MIMO Detection," IEEE Trans. Signal Process, vol. 57, no. 7, pp. 2701–2710, July 2009.
[170] C. Stierstorfer and R. F. Fischer, "Lattice-Reduction-aided TomlinsonâĂŞHarashima Precoding for Point-to-Multipoint Transmission," AEU - International Journal of Electronics and Communications, vol. 60, no. 4, pp. 328 – 330, 2006.
[171] X. W. Chang, J. Wen, and X. Xie, "Effects of the LLL Reduction on the Success Probability of the Babai Point and on the Complexity of Sphere Decoding," IEEE Trans. Inf. Theory, vol. 59, no. 8, pp. 4915–4926, Aug 2013.
[172] C. Ling and N. Howgrave-Graham, "Effective LLL Reduction for Lattice Decoding," in 2007 IEEE International Symposium on Information Theory, June 2007, pp. 196–200.
[173] J. Jaldén, D. Seethaler, and G. Matz, "Worst- and Average-Case Complexity of LLL Lattice Reduction in MIMO Wireless Systems," in 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, March 2008, pp. 2685–2688.
[174] X. Ma and W. Zhang, "Performance Analysis for MIMO Systems with Lattice-Reduction Aided Linear Equalization," IEEE Trans. Commun., vol. 56, no. 2, pp. 309–318, February 2008.
[175] S. Hong and G. Caire, "Reverse Compute and Forward: A Low-Complexity Architecture for Downlink Distributed Antenna Systems," in 2012 IEEE International Symposium on Information Theory Proceedings, July 2012, pp. 1147–1151.
[176] J. Zhan, B. Nazer, U. Erez, and M. Gastpar, "Integer-Forcing Linear Receivers," IEEE Trans. Inf. Theory, vol. 60, no. 12, pp. 7661–7685, Dec 2014.
[177] S. Stern and R. F. H. Fischer, "Optimal Factorization in Lattice-Reduction-Aided and Integer-Forcing Linear Equalization," in SCC 2017; 11th International ITG Conference on Systems, Communications and Coding, Feb 2017, pp. 1–6.
[178] R. Zamir, Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation, and Multiuser Information Theory. Cambridge University Press, 2014.
[179] U. Erez and R. Zamir, "Achieving 1/2 log (1+SNR) on the AWGN channel with lattice encoding and decoding," IEEE Transactions on Information Theory, vol. 50, no. 10, pp. 2293–2314, Oct 2004.
[180] W. Kositwattanarerk, S. S. Ong, and F. Oggier, "Construction A of Lattices Over Number Fields and Block Fading (Wiretap) Coding," IEEE Transactions on Information Theory, vol. 61, no. 5, pp. 2273–2282, May 2015.
[181] U. Fincke and M. Pohst, "Improved Methods for Calculating Vectors of Short Length in a Lattice, Including a Complexity Analysis," Mathematics of Computation, vol. 44, no. 170, pp. 463–471, 1985.
[182] C. P. Schnorr and M. Euchner, "Lattice Basis Reduction: Improved Practical Algorithms and Solving Subset Sum Problems," Mathematical Programming, vol. 66, no. 1, pp. 181–199, 1994.
[183] M. O. Damen, H. E. Gamal, and G. Caire, "On Maximum-Likelihood Detection and the Search for the Closest Lattice Point," IEEE Trans. Inf. Theory, vol. 49, no. 10, pp. 2389–2402, Oct 2003.
[184] Z. Xie, C. K. Rushforth, R. T. Short, and T. K. Moon, "Joint Signal Detection and Parameter Estimation in Multiuser Communications," IEEE Trans. Commun., vol. 41, no. 8, pp. 1208–1216, Aug 1993.
[185] K. wai Wong, C. ying Tsui, R. S. . Cheng, and W. ho Mow, "A VLSI Architecture of a K-Best Lattice Decoding Algorithm for MIMO Channels," in 2002 IEEE International Symposium on Circuits and Systems. Proceedings, vol. 3, May 2002, pp. III–III.
[186] Z. Guo and P. Nilsson, "Algorithm and Implementation of the K-Best Sphere Decoding for MIMO Detection," IEEE J. Sel. Areas Commun., vol. 24, no. 3, pp. 491–503, March 2006.
[187] T. Fukatani, R. Matsumoto, and T. Uyematsu, "Two Methods for Decreasing the Computational Complexity of the MIMO ML Decoder," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol. 87, no. 10, pp. 2571–2576, 2004.
[188] K. Lee and J. Chun, "ML Symbol Detection Based on the Shortest Path Algorithm for MIMO Systems," IEEE Trans. Signal Process., vol. 55, no. 11, pp. 5477–5484, Nov 2007.
[189] R. Y. Chang and W. Chung, "Best-First Tree Search with Probabilistic Node Ordering for MIMO Detection: Generalization and Performance-Complexity Tradeoff," IEEE Trans. Wireless Commun., vol. 11, no. 2, pp. 780–789, February 2012.
[190] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest Point Search in Lattices," IEEE Trans. Inf. Theory, vol. 48, no. 8, pp. 2201 – 2214, Aug 2002.
[191] X. Xie, X.-W. Chang, and M. A. Borno, "Partial LLL Reduction," in IEEE GlobalComm, 2011.
[192] C. Studer, D. Seethaler, and H. Bölcskei, "Finite lattice-size effects in MIMO detection," in 2008 42nd Asilomar Conference on Signals, Systems and Computers, Oct 2008, pp. 2032–2037.
[193] H. Vetter and Y. Sun, "Legacy user support in multiuser MIMO with vector perturbation precoding," in Proc. Indoor and Mobile Radio Communications 21st Annual IEEE Int. Symp. Personal, Sep. 2010, pp. 887–892.
[194] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Vector Perturbation Based on Symbol Scaling for Limited Feedback MISO Downlinks," IEEE Trans. Signal Process., vol. 62, no. 3, pp. 562–571, Feb. 2014.
[195] J. Jalden and B. Ottersten, "On the Complexity of Sphere Decoding in Digital Communications," IEEE Trans. Signal Process., vol. 53, no. 4, pp. 1474–1484, April 2005.
[196] M. Mohaisen, B. Hui, K. Chang, S. Ji, and J. Joung, "Fixed-complexity Vector Perturbation with Block Diagonalization for MU-MIMO Systems," in 2009 IEEE 9th Malaysia International Conference on Communications (MICC), Dec 2009, pp. 238–243.
[197] C. Zheng, X. Chu, J. McAllister, and R. Woods, "Real-Valued Fixed-Complexity Sphere Decoder for High Dimensional QAM-MIMO Systems," IEEE Trans. Signal Process., vol. 59, no. 9, pp. 4493–4499, Sept 2011.
[198] X. Chen, G. He, and J. Ma, "VLSI Implementation of a High-Throughput Iterative Fixed-Complexity Sphere Decoder," IEEE Trans. Circuits Syst. II, Exp. Briefs, vol. 60, no. 5, pp. 272–276, May 2013.
[199] M. Guenach, "Novel Lattice Reduction Algorithms: Precoder Reduction and Vector Perturbation Tradeoffs," IEEE Commun. Lett., p. 1, 2018.
[200] Ö. Dagdelen and M. Schneider, Parallel Enumeration of Shortest Lattice Vectors. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 211–222.
[201] A. Becker, N. Gama, and A. Joux, "A sieve algorithm based on overlattices," LMS Journal of Computation and Mathematics, vol. 17, no. A, pp. 49–70, 2014.
[202] W. S. Chua, C. Yuen, and F. Chin, "A Continuous Vector-Perturbation for Multi-Antenna Multi-User Communication," in VTC2007-Spring. Dublin: IEEE, Apr 2007, pp. 1806 – 1810.
[203] E. Y. Kim and J. Chun, "Optimum Vector Perturbation Minimizing Total MSE in Multiuser MIMO Downlink," in Proc. IEEE Int. Conf. Communications, vol. 9, Jun. 2006, pp. 4242–4247.
[204] D. J. Ryan, "Space-Time Vector Perturbation Precoding," in 2010 IEEE International Conference on Communications, May 2010, pp. 1–6.
[205] D. A. Karpuk, A. Barreal, O. W. Gnilke, and C. Hollanti, "Nested Lattice Codes for Vector Perturbation Systems," arXiv preprint arXiv:1604.07048, 2016.
[206] D. J. Ryan, I. B. Collings, I. V. L. Clarkson, and R. W. H. Jr., "A Lattice-Theoretic Analysis of Vector Perturbation for Multi-User MIMO Systems," in Proc. IEEE Int. Conf. Communications, May 2008, pp. 3340–3344.
[207] C. B. Chae, S. Shim, and R. W. Heath, "Block Diagonalized Vector Perturbation for Multiuser MIMO Systems," IEEE Trans. Wireless Commun., vol. 7, no. 11, pp. 4051–4057, Nov. 2008.
[208] H. S. Han, S. H. Park, and I. Lee, "Improved Vector Perturbation with Modulo Loss Reduction for Multiuser Downlink Systems," in Proc. IEEE Int. Conf. Communications, Jun. 2009, pp. 1–5.
[209] H. S. Han, S. H. Park, S. Lee, and I. Lee, "Modulo Loss Reduction for Vector Perturbation Systems," IEEE Trans. Commun., vol. 58, no. 12, pp. 3392–3396, Dec. 2010.
[210] W. Yao, S. Chen, and L. Hanzo, "A Transceiver Design Based on Uniform Channel Decomposition and MBER Vector Perturbation," IEEE Trans. Veh. Technol., vol. 59, no. 6, pp. 3153–3159, July 2010.
[211] ——, "Generalized MBER-Based Vector Precoding Design for Multiuser Transmission," IEEE Trans. Veh. Technol., vol. 60, no. 2, pp. 739–745, Feb 2011.
[212] M. Mazrouei-Sebdani and W. A. KrzymieÅĎ, "On MMSE vector-perturbation precoding for MIMO broadcast channels with per-antenna-group power constraints," in Proc. IEEE Int. Conf. Communications (ICC), Jun. 2012, pp. 4841–4845.

[213] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," IEEE Journal on Selected Areas in Communications, vol. 20, no. 5, pp. 1105–1115, June 2002.
[214] Y. Xu, T. Le-Ngoc, and S. Panigrahi, "Selective iterative waterfilling for digital subscriber lines," EURASIP J. Adv. Signal Process, no. 59068, 2007.
[215] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous Spectrum Balancing for Digital Subscriber Lines," IEEE Transactions on Signal Processing, vol. 55, no. 8, pp. 4241–4257, Aug 2007.
[216] P. Tsiaflakis, M. Diehl, and M. Moonen, "Distributed Spectrum Management Algorithms for Multiuser DSL Networks," IEEE Transactions on Signal Processing, vol. 56, no. 10, pp. 4825–4843, Oct 2008.
[217] J. Verdyck and M. Moonen, "Dynamic Spectrum Management in Digital Subscriber Line Networks With Unequal Error Protection Requirements," IEEE Access, vol. 5, pp. 18 107–18 120, 2017.
[218] W. Yu and T. Lan, "Transmitter Optimization for the Multi-Antenna Downlink With Per-Antenna Power Constraints," IEEE Transactions on Signal Processing, vol. 55, no. 6, pp. 2646–2660, June 2007.
[219] A. R. Forouzan, M. Moonen, J. Maes, and M. Guenach, "Joint Level 2 and 3 Dynamic Spectrum Management for Downstream DSL," IEEE Trans. Commun., vol. 60, no. 10, pp. 3111–3122, October 2012.
[220] P. Tsiaflakis, J. Vangorp, J. Verlinden, and M. Moonen, "Multiple Access Channel Optimal Spectrum Balancing for Upstream DSL Transmission," IEEE Communications Letters, vol. 11, no. 4, pp. 398–300, April 2007.
[221] V. Le Nir, M. Moonen, J. Verlinden, and M. Guenach, "Optimal Power Allocation for Downstream xDSL With Per-Modem Total Power Constraints: Broadcast Channel Optimal Spectrum Balancing (BC-OSB)," IEEE Transactions on Signal Processing, vol. 57, no. 2, pp. 690–697, Feb 2009.
[222] J. M. Ciofli, "A Multicarrier Primer," ANSI T1E1, vol. 4, pp. 91–157, 1991.
[223] X. chun Zhang, H. Yu, and G. Wei, "Exact Symbol Error Probability of Cross-QAM in AWGN and Fading Channels," EURASIP Journal on Wireless Communications and Networking, vol. 2010, no. 1, p. 917954, Nov 2010.
[224] J. Neckebroek, M. Moeneclaey, W. Coomans, and M. Guenach, "Novel Bitloading Algorithms for Coded G.fast DSL Transmission with Linear and Nonlinear Precoding," in IEEE International Conference on Communications. London: IEEE, June 2015, pp. 945 – 951.
[225] E. Viterbo and E. Biglieri, "Computing the Voronoi Cell of a Lattice: the Diamond-Cutting Algorithm," IEEE Trans. Inf. Theory, vol. 42, no. 1, pp. 161–171, Jan 1996.
[226] M. Dutour Sikirić, A. Schürmann, and F. Vallentin, "Complexity and algorithms for computing Voronoi cells of lattices," Math. comp., vol. 78, no. 267, pp. 1713–1731, 2009.
[227] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," IEEE Transactions on Communications, vol. 54, no. 7, pp. 1310–1322, July 2006.
[228] P. Tsiaflakis, I. Necoara, J. A. K. Suykens, and M. Moonen, "Improved Dual Decomposition Based Optimization for DSL Dynamic Spectrum Management," IEEE Transactions on Signal Processing, vol. 58, no. 4, pp. 2230–2245, April 2010.
[229] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal Multiuser Spectrum Balancing for Digital Subscriber Lines," IEEE Trans. Commun., vol. 54, no. 5, pp. 922–933, May 2006.
[230] S. Wei, L. Youming, and Y. Miaoliang, "Low-complexity Grouping Spectrum Management in Multi-user DSL Networks," in 2009 WRI International Conference on Communications and Mobile Computing, vol. 1, Jan 2009, pp. 381–385.
[231] R. Cendrillon and M. Moonen, "Iterative spectrum balancing for digital subscriber lines," in IEEE International Conference on Communications, 2005. ICC 2005. 2005, vol. 3, May 2005, pp. 1937–1941.
[232] R. Lui and W. Yu, "Low-complexity near-optimal spectrum balancing for digital subscriber lines," in IEEE International Conference on Communications, 2005. ICC 2005. 2005, vol. 3, May 2005, pp. 1947–1951.
[233] W. Yu, W. Rhee, S. Boyd, and J. M. Ciofli, "Iterative water-filling for Gaussian vector multiple access channels," in Proceedings. 2001 IEEE International Symposium on Information Theory (IEEE Cat. No.01CH37252), June 2001, pp. 322–.
[234] J. Papandriopoulos and J. S. Evans, "Low-Complexity Distributed Algorithms for Spectrum Balancing in Multi-User DSL Networks," in 2006 IEEE International Conference on Communications, vol. 7, June 2006, pp. 3270–3275.
[235] W. Lanneer, J. Verdyck, P. Tsiaflakis, J. Maes, and M. Moonen, "Vectoring-based dynamic spectrum management for G.fast multi-user full-duplex transmission," in 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Oct 2017, pp. 1–5.
[236] D. Z. Filho, R. R. Lopes, R. Ferrari, R. Suyama, and B. Dortschy, "Bit Loading for Precoded DSL Systems," in 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07, vol. 3, April 2007, pp. III–353–III–356.
[237] J. Campello, "Practical Bit Loading for DMT," in 1999 IEEE International Conference on Communications (Cat. No. 99CH36311), vol. 2, 1999, pp. 801–805 vol.2.
[238] T. N. Vo, K. Amis, T. Chonavel, and P. Siohan, "A Computationally Efficient Discrete Bit-Loading Algorithm for OFDM Systems Subject to Spectral-Compatibility Limits," IEEE Trans. Commun., vol. 63, no. 6, pp. 2261 – 2272, June 2015.
[239] J. Park and J. Chun, "Improved Lattice Reduction-Aided MIMO Successive Interference Cancellation Under Imperfect Channel Estimation," IEEE Transactions on Signal Processing, vol. 60, no. 6, pp. 3346–3351, June 2012.
[240] J. Jalden, D. Seethaler, and G. Matz, "Worst- and average-case complexity of LLL lattice reduction in MIMO wireless systems," in 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, March 2008, pp. 2685–2688.
[241] L. G. Barbero and J. S. Thompson, "Fixing the Complexity of the Sphere Decoder for MIMO Detection," IEEE Transactions on Wireless Communications, vol. 7, no. 6, pp. 2131–2142, June 2008.
[242] Y. Gu and K. K. Parhi, "High-Speed Architecture Design of TomlinsonâĂŞHarashima Precoders," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 54, no. 9, pp. 1929–1937, Sep. 2007.
[243] L. Zhang, Y. Cai, R. C. de Lamare, and M. Zhao, "Robust Multibranch TomlinsonâĂŞHarashima Precoding Design in Amplify-and-Forward MIMO Relay Systems," IEEE Transactions on Communications, vol. 62, no. 10, pp. 3476–3490, Oct 2014.
[244] Y. Luo and S. Qiao, "A parallel LLL algorithm." in C3S2E. Citeseer, 2011, pp. 93–101.
[245] C. Ling, W. H. Mow, and N. Howgrave-Graham, "Reduced and Fixed-Complexity Variants of the LLL Algorithm for Communications," IEEE Transactions on Communications, vol. 61, no. 3, pp. 1040–1050, March 2013.
[246] C. I, S. Han, Z. Xu, S. Wang, Q. Sun, and Y. Chen, "New Paradigm of 5G Wireless Internet," IEEE Journal on Selected Areas in Communications, vol. 34, no. 3, pp. 474–482, March 2016.
[247] P. K. Agyapong, M. Iwamura, D. Staehle, W. Kiess, and A. Benjebbour, "Design considerations for a 5G network architecture," IEEE Communications Magazine, vol. 52, no. 11, pp. 65–75, Nov 2014.