Spatial-Multiplexing for Photon-Counting MIMO-FSO Communication Systems

Chadi Abou-Rjeily, *Senior Member IEEE*

***Abstract*—In this paper, we consider the problem of Multiple-Input-Multiple-Output (MIMO) Free-Space Optical (FSO) communications under the Poisson photon-counting detection model. Aiming for high bit rate objectives, we consider the spatial-multiplexing (SMux) solution with M-ary pulse position modulation (PPM) where we propose appropriate optimal and suboptimal decoders and evaluate their complexities. Such novel decoder designs are needed since the widely spread Gaussian noise based MIMO decoders are not suitable for the Poisson model. We also carry out an asymptotic performance analysis that guides a candidate constellation confinement where transmissions are limited to some selected information vectors of the multi-dimensional SMux constellation in an attempt for compromising the multiplexing gains for error-rate improvements. The analyzed SMux solutions with both the unconfined and confined constellations transmit at higher data rates compared to the existing single-aperture systems, MIMO systems with repetition coding and MIMO systems based on spatial modulation.**

***Index Terms*—Free-Space Optics, FSO, Multiple-Input-Multiple-Output, MIMO, maximum-likelihood, spatial multiplexing, Poisson noise, suboptimal decoders, constellation confinement, performance analysis.**

I. Introduction

Multiple-Input-Multiple-Output (MIMO) techniques take advantage from the underlying spatial degree of freedom in wireless communication systems for the sake of achieving enhanced capacities in a bandwidth-efficient and power-efficient manner. MIMO solutions were extensively considered not only in the context of Radio Frequency (RF) systems but also for Free-Space Optical (FSO) systems [1]–[14]. The resemblance between MIMO-RF and MIMO-FSO systems resides in the variability of the path gains even though the sources of this randomness are different where they originate from multi-path propagation (fading) in the former case and from atmospheric turbulence (scintillation) in the latter case. The differentiation between the analysis of MIMO-RF and MIMO-FSO systems arises from the type of detection where intensity-modulation with direct-detection (IM/DD) is commonly used with FSO systems. This has a direct impact on the transmitted modulation schemes where non-negative real-valued signal constellations, such as On-Off Keying (OOK) or Pulse Position Modulation (PPM), are often used. The second main difference resides in the nature of noise where the Gaussian noise model is often adopted for RF systems while FSO systems are described by the more general Poisson photon-counting detection model where the number of photons generated by the optical information-carrying signal and by the background radiation is modeled by a Poisson point process [8], [9], [15], [16]. It is worth noting that the signal-independent additive white Gaussian noise (AWGN) model is a common approximation that is often adopted for the analysis of FSO systems. This approximation is valid only when the shot noise caused by background radiation is dominant with respect to the other noise components such as thermal noise and dark currents [10]–[12].

Numerous MIMO-FSO IM/DD solutions were investigated in the literature including space-time coding (STC) [1], [2], repetition coding (RC) [3]–[9], spatial multiplexing (SMux) [10]–[12] and optical spatial modulation (OSM) [13], [14]. Real-valued STCs were proposed in [1] and [2] for OOK and PPM, respectively, where the AWGN model was considered. On the other hand, RC constitutes the most widely investigated MIMO-FSO scheme capable of achieving spatial diversity gains in a simplified manner where the same information symbol is repeated from all transmit apertures [3]–[9]. In [3], it has been proven that RC outperforms STC with OOK; similarly, it has been proven in [4] that RC outperforms parallel-relaying cooperative systems with binary-PPM. Both [3] and [4] considered the Gaussian noise model. The AWGN model was also adopted in [5]–[7] where, in [5], the outage probability of RC-PPM was analyzed over lognormal, exponential and gamma-gamma turbulence-induced channels; in [6], expressions for the ergodic capacity with RC-OOK were derived over gamma-gamma channels while [7] targeted the bit error rate (BER) analysis of RC-OOK systems with equal gain combining (EGC) and maximum ratio combining (MRC) over gamma-gamma channels. Unlike the RC AWGN-based studies [3]–[7], references [8] and [9] considered RC with the Poisson photon-counting detection model for PPM and OOK modulations, respectively. While the BER was analyzed in [8] over lognormal and exponential channels, multiple-symbol detection was tackled in [9].

While the MIMO-FSO RC scheme is fully diverse with an appealing decoding complexity that is practically the same as that of Single-Input-Single-Output (SISO) systems, this scheme is limited by its incapability of achieving any multiplexing gains where the MIMO-RC schemes transmit at the same data rate as SISO systems. This motivated the investigation of MIMO-FSO SMux solutions where independent data streams are transmitted from the P transmit apertures resulting in a P-fold increase in the bit rate [10]–[12]. The diversity-multiplexing tradeoff over lognormal channels was investigated in [10], the performance of MIMO-FSO SMux systems under lognormal fading and pointing errors was

The author is with the Department of Electrical and Computer Engineering of the Lebanese American University (LAU), PO box 36 Byblos 961, Lebanon. (e-mail: chadi.abourjeily@lau.edu.lb).

evaluated in [11] while different MIMO-FSO IM/DD schemes were compared in [12] with OOK. In this context, it is worth noting that all of the existing MIMO-FSO SMux systems were analyzed exclusively under the Gaussian noise model rendering the input-output baseband relations and signal-to-noise ratio (SNR) expressions very similar to their RF counterparts. While the highly spectral-efficient SMux solutions suffer from the decoding complexity where joint detection needs to be carried out on P data streams, OSM solutions constitute simpler alternatives to SMux where only one transmit aperture is activated in a time-slot thus eradicating the inter-channel interference at the receiver [13], [14]. While both the SMux and OSM schemes do not profit from the transmit diversity (unlike STC and RC), OSM solutions suffer from reduced bit rates compared to SMux. In fact, for SMux the number of bits per channel use (pcu) is multiplied by P with respect to SISO systems while for OSM only $\log_2(P)$ additional bits pcu can be communicated. As the SMux solutions, the MIMO-FSO OSM systems were considered exclusively in the context of AWGN noise.

The problem of sequence detection with photon-counting receivers was addressed in [9], [15], [16] with the following central differences with the detectors proposed in this work. References [9], [15], [16] considered the transmission of sequences of L OOK symbols $[s_1, \ldots, s_L]$ where the transmitted symbols do not interfere with other. This holds for the SISO systems in [15], [16] as well as the MIMO-RC system in [9]. Therefore, the noise statistics in the l-th bit duration depend only on bit s_l independently from the values taken by the other bits $s_{l'}$ for $l' \neq l$. Consequently, the L Poisson random variables observed in the L bit durations are independent; moreover, they are identically distributed for the same bit values. On the other hand, for MIMO-SMux systems, the P transmitted M-PPM symbols $[s_1, \ldots, s_P]$ interfere with each other over the same symbol duration. Therefore, not only the parameter of the Poisson random variable observed in the m-th PPM slot will depend on the values taken by all symbols $s_1, \ldots, s_P$, but also this parameter will vary from one symbol duration to another rendering the detection problem completely different. Finally, unlike this work where we assume that the CSI is available at the receiver, [9], [15], [16] operate in the absence of CSI at the receiver. For these references, sequence detection rather than symbol-by-symbol detection was considered in order to explicitly account for the unknown values of the channel irradiances. In this context, operating in the presence of CSI at the receiver not only results in better performance but it is also judged to be not very problematic since the pilot-symbols overhead is negligible given the large coherence times of the FSO channels.

This work targets the design and analysis of MIMO-FSO systems in the case where the channel state information (CSI) is available at the receiver but not at the transmitter in a way that is analogous to [10]–[14]. In particular, we consider the problem of SMux with the more general Poisson photon-counting detection where this adopted model clearly distinguishes the current work from the previous works on MIMO-FSO SMux systems that all considered the AWGN model [10]–[12]. The Poisson model affects not only the performance analysis but it profoundly alters the design of the decoders where the existing MIMO-RF decoders that were designed to separate the MIMO data streams in the case of Gaussian noise will fail in guaranteeing successful detection with Poisson noise. The existing optimal Maximum-Likelihood (ML) decoders [17], linear minimum mean square error (MMSE) decoders and V-Blast decoders [18] are all tailored to systems corrupted by AWGN not to systems based on photon-counting. Even the MRC scheme that maximizes the SNR with Gaussian noise [19] is not suitable under Poisson statistics. Based on what preceded, the first set of contributions of this work can be summarized as follows:

- Proposing an optimal MIMO-FSO ML decoder under Poisson statistics.
- Proposing two simplified suboptimal MIMO-FSO decoders under Poisson statistics. One of these decoders is capable of achieving optimal detection in the MISO case.
- Evaluating the complexities of the above decoders.

Building this work around spatial multiplexing stems from the need to meet the primary interest of increasing the data rate of FSO systems through MIMO techniques. While, under specific bandwidth requirements, the bit rate of RF systems can be easily increased by increasing the cardinality of the transmitted constellation (often QAM or PSK) without affecting the spectral efficiency, this issue is more subtle with FSO systems. In fact, FSO IM/DD systems are often associated with OOK or PPM. Therefore, for a given baud rate and transceiver bandwidth, the use of OOK fixes the achievable bit rate and all one-dimensional PAM signal expansions will severely deteriorate the performance. On the other hand, the FSO IM/DD bit rate can be increased by increasing the cardinality of the PPM signal set. Despite the fact that this increase reduces the error rate, it results in a decrease in the spectral efficiency. Consequently, moderate sizes of the PPM constellations are considered in practice. As a conclusion, since the use of OOK fixes the bit rate while the use of M-PPM with large values of M limits the spectral efficiency, the SMux solution constitutes a viable option for delivering higher rates with FSO communications.

While the SMux solution meets the high data rate requirements, this comes at the expense of reduced performance levels [12]. This motivates the second direction of research in this work which corresponds to proposing a variant of SMux that compromises the bit rate to BER. As such, the second set of contributions is as follows:

- Carrying out an asymptotic performance analysis of MIMO-FSO SMux systems. This results in the classification of the pairwise error probabilities into a number of categories.
- Proposing an adequate constellation confinement based on the above classification. The cardinality of this constellation exceeds the cardinality of RC (M symbols with M-PPM) in [3]–[7] and that of space shift keying (P symbols) in [13], [14] for all values of M and P. This cardinality also exceeds that of the OSM system based on joint position and antenna modulation in [13] (MP

symbols) for all values of P when $M > 2$.

II. System Model

Consider a $P \times Q$ MIMO system where the transmitter and receiver are equipped with P and Q apertures, respectively. The MIMO-FSO system under consideration is based on IM/DD with M-ary PPM. In this case, the symbol duration T_s is divided into M slots and a light signal is sent in only one of these slots. We denote by $s_p \in \{1, \ldots, M\}$ the position of the pulse transmitted by the p-th aperture for $p = 1, \ldots, P$. The analyzed system revolves around spatial multiplexing where the transmitted vector is denoted by $\mathbf{s} = [s_1, \ldots, s_P] \in \{1, \ldots, M\}^P$.

The channel irradiance between the p-th transmit aperture and q-th receive aperture will be denoted by $I_{q,p}$ for $p = 1, \ldots, P$ and $q = 1, \ldots, Q$. The PQ FSO channels are assumed to be independent and identically distributed according to the channel model proposed in [20]. In this case, the channel irradiance can be expressed as $I = I_l I_a I_p$ taking into account the combined effects of path loss (I_l), atmospheric turbulence-induced scintillation (I_a) and misalignment-induced fading caused by pointing errors (I_p). The path loss is given by $I_l = e^{-\sigma d}$ where σ stands for the attenuation coefficient while d stands for the distance between the transmitter and receiver. In this work, we adopt the gamma-gamma model to characterize the atmospheric scintillation where the probability density function (PDF) of I_a is given by:

$$f_{I_a}(I_a) = \frac{2(\varphi_1\varphi_2)^{(\varphi_1+\varphi_2)/2}}{\Gamma(\varphi_1)\Gamma(\varphi_2)} I_a^{\frac{\varphi_1+\varphi_2}{2}-1} K_{\varphi_1-\varphi_2}\left(2\sqrt{\varphi_1\varphi_2 I_a}\right) \; ; \; I_a \geq 0, \quad (1)$$

where $\Gamma(.)$ is the Gamma function and $K_n(.)$ is the modified Bessel function of the second kind of order n. The distance-dependent parameters of the gamma-gamma distribution are given by $\varphi_1 = \left[\exp\left(0.49\sigma_R^2/(1+1.11\sigma_R^{12/5})^{7/6}\right) - 1\right]^{-1}$ and $\varphi_2 = \left[\exp\left(0.51\sigma_R^2/(1+0.69\sigma_R^{12/5})^{5/6}\right) - 1\right]^{-1}$ where $\sigma_R^2 = 1.23 C_n^2 k^{7/6} d^{11/6}$ is the Rytov variance, $k = \frac{2\pi}{\lambda}$ is the wave number and C_n^2 denotes the refractive index structure parameter.

The PDF of the nonzero boresight pointing error was derived in [20]:

$$f_{I_p}(I_p) = \frac{\xi^2 \exp\left(-\frac{s^2}{2\sigma_s^2}\right)}{A_0^{\xi^2}} I_p^{\xi^2-1} I_0\left(\frac{s}{\sigma_s^2}\sqrt{\frac{\omega_{z_{eq}}^2 \ln\frac{A_0}{I_p}}{2}}\right) \; ; \; 0 \leq I_p \leq A_0, \quad (2)$$

where $I_0(.)$ is the modified Bessel function of the first kind of order zero. In (2), $A_0 = \text{erf}^2(\nu)$, $\omega_{z_{eq}}^2 = \omega_z^2 \frac{\sqrt{\pi}\text{erf}(\nu)}{2\nu\exp(-\nu^2)}$ and $\xi = \frac{\omega_{z_{eq}}}{2\sigma_s}$ where $\nu = \sqrt{\frac{\pi}{2}}\frac{a}{\omega_z}$. In these relations, erf(.) is the error function while a, ω_z, σ_s and s stand for the receiver radius, beam waist, jitter standard deviation and boresight displacement, respectively.

The decisions in IM/DD photon-counter receivers are based on the numbers of photoelectrons detected in the M PPM slots. The average number of photoelectrons generated by the information-carrying light signal (in the absence of scintillation) in a PPM slot will be denoted by λ_s. Similarly, the average number of noise photoelectrons generated by background radiation (and dark currents) in a PPM slot will be denoted by λ_b. These two quantities are given by:

$$\lambda_s = \eta\frac{P_s T_s/M}{hf} \triangleq \eta\frac{E_s}{hf} \quad ; \quad \lambda_b = \eta\frac{P_b T_s/M}{hf}, \quad (3)$$

where P_s and P_b stand for the incident optical power and the power of background noise, respectively. η is the detector's quantum efficiency, h is Planck's constant and f is the optical center frequency taken to be 1.94×10^{14} Hz (corresponding to a wavelength of 1550 nm). Finally, $E_s = P_s T_s/M$ corresponds to the received optical energy per PPM slot.

Denote by $R_{q,m}$ the random variable corresponding to the number of photoelectrons detected by the q-th receive aperture in the m-th slot. $R_{q,m}$ follows the Poisson distribution with parameter:

$$\text{E}\left[R_{q,m}\right] = \frac{\lambda_s}{P}\sum_{p=1}^{P}\delta_{s_p,m} I_{q,p} + \lambda_b, \quad (4)$$

where E[.] stands for the averaging operator while $\delta_{i,j}$ stands for the Kronecker delta with $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ otherwise. The term $\delta_{s_p,m}$ indicates whether the p-th aperture is transmitting in the m-th PPM slot or not. Finally, the normalization by P in (4) follows from evenly splitting the power among the P transmit apertures in the absence of CSI at the transmitter.

III. Optimal and Suboptimal Decoders

A. ML MIMO Decoder (ML)

1) Decoder Implementation: The maximum likelihood (ML) detection procedure is based on the decision variables $R_{q,m}$ (for $q = 1, \ldots, Q$ and $m = 1, \ldots, M$) as follows:

$$\hat{\mathbf{s}} = \arg\max_{\mathbf{s}\in\{1,\ldots,M\}^P}\left\{\prod_{q=1}^{Q}\prod_{m=1}^{M}\Pr\left(R_{q,m} = r_{q,m}\right)\right\}, \quad (5)$$

where $r_{q,m}$ stands for the actual number of photoelectrons detected in the m-th PPM slot at the q-th receiver aperture. From the Poisson parameters in (4), the ML rule in (5) can be written as:

$$\hat{\mathbf{s}} = \arg\max_{\mathbf{s}\in\{1,\ldots,M\}^P}\left\{\prod_{q=1}^{Q}\prod_{m=1}^{M}\frac{e^{-\lambda_b}\lambda_b^{r_{q,m}}}{r_{q,m}!} e^{-\frac{\lambda_s}{P}\sum_{p=1}^{P}\delta_{s_p,m}I_{q,p}}\left(1 + \frac{\lambda_s}{P\lambda_b}\sum_{p=1}^{P}\delta_{s_p,m}I_{q,p}\right)^{r_{q,m}}\right\}. \quad (6)$$

Removing the term $\frac{e^{-\lambda_b}\lambda_b^{r_{q,m}}}{r_{q,m}!}$ that does not depend on $\mathbf{s}$, equation (6) is equivalent to:

$$\hat{\mathbf{s}} = \arg \max_{\mathbf{s}\in\{1,\ldots,M\}^P} \left\{ e^{-\frac{\lambda_s}{P}\sum_{q=1}^{Q}\sum_{p=1}^{P} I_{q,p}\sum_{m=1}^{M}\delta_{s_p,m}} \prod_{q=1}^{Q}\prod_{m=1}^{M}\left(1+\frac{\lambda_s}{P\lambda_b}\sum_{p=1}^{P}\delta_{s_p,m}I_{q,p}\right)^{r_{q,m}} \right\}, \quad (7)$$

where, since s_p assumes a unique value in $\{1,\ldots,M\}$, then $\sum_{m=1}^{M}\delta_{s_p,m}$ is always equal to 1 implying that the first term in (7) simplifies to the expression $e^{-\frac{\lambda_s}{P}\sum_{q=1}^{Q}\sum_{p=1}^{P} I_{q,p}}$ that does not depend on $\mathbf{s}$ and, hence, can be removed from the ML decision rule. Now, taking the logarithm of the remaining part of (7) results in the following optimal ML decoder:

$$\hat{\mathbf{s}} = \arg \max_{\mathbf{s}\in\{1,\ldots,M\}^P} \left\{ \sum_{q=1}^{Q}\sum_{m=1}^{M} r_{q,m}\log\left(1+\frac{\lambda_s}{P\lambda_b}\sum_{p=1}^{P}\delta_{s_p,m}I_{q,p}\right) \right\} \quad (8)$$

$$\triangleq \arg \max_{\mathbf{s}\in\{1,\ldots,M\}^P} \left\{ \sum_{q=1}^{Q}\chi^{(q)}(\mathbf{s}) \right\}. \quad (9)$$

2) Decoder Complexity: The complexity of the decoder will be evaluated based on the number of multiplications. In other words, we ignore the complexity that is associated with the additions that incur much less processing requirements compared to the multiplication operations.

At a first glance, (8) might suggest that $(MQ)^{M^P}$ multiplications are needed for the implementation of the ML decoder. However, the term $\log\left(1+\frac{\lambda_s}{P\lambda_b}\sum_{p=1}^{P}\delta_{s_p,m}I_{q,p}\right)$ might be zero if there are no apertures transmitting in the m-th PPM slot implying that no multiplications will be needed for the evaluation of the corresponding term $r_{q,m}\log\left(1+\frac{\lambda_s}{P\lambda_b}\sum_{p=1}^{P}\delta_{s_p,m}I_{q,p}\right)$ in (8). Therefore, a more thorough analysis is needed. It can be observed that, for a particular value of q, the set $\left\{\log\left(1+\frac{\lambda_s}{P\lambda_b}\sum_{p=1}^{P}\delta_{s_p,m}I_{q,p}\right)\right\}_{\substack{s\in\{1,\ldots,M\}^P \\ m=1,\ldots,M}}$ comprises exactly 2^P-1 non-zero elements each corresponding to a non-empty subset of $\{1,\ldots,P\}$. Therefore, defining the term $\alpha_{\mathcal{P}}^{(q)}$ as:

$$\alpha_{\mathcal{P}}^{(q)} \triangleq \log\left(1+\frac{\lambda_s}{P\lambda_b}\sum_{p\in\mathcal{P}}\delta_{s_p,m}I_{q,p}\right) \quad \text{for} \quad \mathcal{P}\subset\{1,\ldots,P\}, \quad (10)$$

then $\left\{\log\left(1+\frac{\lambda_s}{P\lambda_b}\sum_{p=1}^{P}\delta_{s_p,m}I_{q,p}\right)\right\}_{\substack{s\in\{1,\ldots,M\}^P \\ m=1,\ldots,M}} = \left\{\alpha_{\mathcal{P}}^{(q)}\right\}_{\mathcal{P}\subset\{1,\ldots,P\}}$ with $\alpha_{\phi}^{(q)} = 0$ (ϕ is the empty set) where both sets in the last equality have a cardinality of 2^P.

The 2^P-1 quantities $\left\{\alpha_{\mathcal{P}}^{(q)}\right\}_{\mathcal{P}\neq\phi}$ do not depend on the transmitted information vector and, hence, they can be predetermined in an initialization phase of the decoding algorithm. In other words, these quantities need to be determined only once for each block-fading duration (rather than being calculated for each tested vector $\mathbf{s}$). Given the very large coherence times of the FSO channels compared to the symbol duration, the complexity arising from the evaluation of these terms can be adequately neglected. In fact, for the outdoor FSO channels, the coherence time is in the order of 1 ms [21]. Given that the data rates of FSO systems range between 100 Mbits/s and 1 Gbits/s, this implies that a block of symbols comprises between 10^5 and 10^6 symbols (with 2-PPM). These numbers are very large when compared to RF systems where the fading blocks extend typically over a number of symbols that ranges between 10^2 and 10^3.

Define $u(\mathbf{s})$ as the number of unique elements of the vector $\mathbf{s}$ indicating the number of slots in which all transmissions from the P transmit apertures take place. We next calculate the number of vectors $\mathbf{s}$ for which $u(\mathbf{s})=n$ where, evidently, $n\leq M$ and $n\leq P$. First, the n distinct slots can be selected from the total number of M slots in $\binom{M}{n}$ different ways. For each one of these selections, denote by l_i the number of apertures transmitting in the i-th unique slot for $i=1,\ldots,n$ where $l_1+\cdots+l_n=P$. The number of ways the P transmit apertures can be divided into n groups containing $l_1,l_2,\ldots,l_n$ elements each is given by $\frac{P!}{l_1!\cdots l_n!}$. Therefore, there are $\binom{M}{n}\sum_{\substack{l_1,\ldots,l_n\in\{1,\ldots,P\} \\ l_1+\cdots+l_n=P}}\frac{P!}{l_1!\cdots l_n!}$ vectors $\mathbf{s}$ satisfying $u(\mathbf{s})=n$. On the other hand, the evaluation of $\sum_{q=1}^{Q}\chi^{(q)}(\mathbf{s})$ in (9) for $u(\mathbf{s})=n$ involves nQ multiplications since the photoelectrons counts in the positions not included in $\mathbf{s}$ will be multiplied by zero. Therefore, the implementation of the ML decoder in (8) necessitates carrying out the following number of multiplications:

$$N_{\text{mul}}^{(\text{ML})} = Q\sum_{n=1}^{\min\{M,P\}} n\times\left[\binom{M}{n}\sum_{\substack{l_1,l_2,\cdots,l_n\in\{1,\ldots,P\} \\ l_1+l_2+\cdots+l_n=P}}\frac{P!}{l_1!l_2!\cdots l_n!}\right]. \quad (11)$$

B. Simplified ML MISO Decoder (Simp-ML)

The optimal ML decoding procedure in (8)-(9) can be further simplified in the case of MISO systems. This will be referred to as the simplified optimal ML (Simp-ML) receiver that results in a significant reduction in the number of multiplications as will be explained later.

1) Decoder Implementation: Setting $Q=1$ in (8)-(9) results in:

$$\hat{\mathbf{s}} = \arg \max_{\mathbf{s}\in\{1,\ldots,M\}^P}\left\{\sum_{m=1}^{M} r_{1,m}\log\left(1+\frac{\lambda_s}{P\lambda_b}\sum_{p=1}^{P}\delta_{s_p,m}I_{1,p}\right)\right\}$$
$$= \arg \max_{\mathbf{s}\in\{1,\ldots,M\}^P}\left\{\chi^{(1)}(\mathbf{s})\right\}. \quad (12)$$

Define $i_1,\ldots,i_M$ as the integers obtained by sorting the decision variables $\{r_{1,m}\}_{m=1}^{M}$ in a decreasing order $r_{1,i_1}\geq r_{1,i_2}\geq\cdots\geq r_{1,i_M}$. The following cases arise.

(*i*): Assume that the tested vector $\mathbf{s}$ satisfies $u(\mathbf{s})=1$ implying that all transmit apertures are transmitting in the same PPM slot. In this case, determining the metric coefficient for the vector $\mathbf{s}=[m'\cdots m']$ results in the value $\chi^{(1)}(\mathbf{s}) = r_{1,m'}\log\left(1+\frac{\lambda_s}{P\lambda_b}\sum_{p=1}^{P}I_{1,p}\right)$. Evidently, this value of $\chi^{(1)}(\mathbf{s})$ is maximized for $m'=i_1$ implying that the vector $\mathbf{s}$ satisfying $u(\mathbf{s})=1$ while having the maximum value

of $\chi^{(1)}(\mathbf{s})$ is $\mathbf{s}=[i_1\cdots i_1]$. In other words, all of the remaining vectors $[i_2\cdots i_2]$, ..., $[i_M\cdots i_M]$ will have a smaller value of $\chi^{(1)}(\mathbf{s})$ and, hence, there is no need to find their corresponding metrics since the ML algorithm searches for the vector $\mathbf{s}$ that maximizes $\chi^{(1)}(\mathbf{s})$. As a conclusion, for $u(\mathbf{s})=1$, the most probable candidate solution is $\mathbf{s}=[i_1\cdots i_1]$ with a weight of $r_{1,i_1}\log\left(1+\frac{\lambda_s}{P\lambda_b}\sum_{p=1}^{P}I_{1,p}\right)=r_{1,i_1}\alpha^{(1)}_{\{1,\ldots,P\}}$ based on the definition in (10).

(*ii*): Assume that $u(\mathbf{s})=2$ implying that the apertures are transmitting in two PPM positions denoted by m' and m''. Denote by $\mathcal{P}_1$ (resp. $\mathcal{P}_2$) the subset of $\{1,\ldots,P\}$ indicating the apertures transmitting in slot m' (resp. m''). In this case, the corresponding weight would be $\chi^{(1)}(\mathbf{s})=r_{1,m'}\alpha^{(1)}_{\mathcal{P}_1}+r_{1,m''}\alpha^{(1)}_{\mathcal{P}_2}$. Evidently, this weight is maximized for $(m',m'')=(i_1,i_2)$ if $\alpha^{(1)}_{\mathcal{P}_1}\geq\alpha^{(1)}_{\mathcal{P}_2}$ and $(m',m'')=(i_2,i_1)$ if $\alpha^{(1)}_{\mathcal{P}_2}\geq\alpha^{(1)}_{\mathcal{P}_1}$. Therefore, the surviving solutions will be obtained as follows. The set $\{1,\ldots,P\}$ needs to be partitioned as $\{1,\ldots,P\}=\mathcal{P}_1\cup\mathcal{P}_2$. Now, for each possible partition, sort the values $\alpha^{(1)}_{\mathcal{P}_1}$ and $\alpha^{(1)}_{\mathcal{P}_2}$ such that $\alpha^{(1)}_{\mathcal{P}_{j_1}}\geq\alpha^{(1)}_{\mathcal{P}_{j_2}}$. The resulting candidate solution can be written as $\mathbf{s}=i_1\sum_{p_1\in\mathcal{P}_{j_1}}\mathbf{e}_{p_1}+i_2\sum_{p_2\in\mathcal{P}_{j_2}}\mathbf{e}_{p_2}$ with a corresponding weight of $\chi^{(1)}(\mathbf{s})=r_{1,i_1}\alpha^{(1)}_{\mathcal{P}_{j_1}}+r_{1,i_2}\alpha^{(1)}_{\mathcal{P}_{j_2}}$ where $\mathbf{e}_p$ stands for the p-th row of the $P\times P$ identity matrix.

(*iii*): The above procedure can be generalized for any value of $n\in\{1,\ldots,\min\{M,P\}\}$ such that $u(\mathbf{s})=n$.

As such, the flowchart of the proposed Simp-ML decoder can be summarized as follows:

Data: $\{r_{1,m}\}_{m=1}^{M}$ and $\{\alpha^{(1)}_{\mathcal{P}}\}_{\mathcal{P}\subset\{1,\ldots,P\}}$ from (10);
Result: The ML decoded vector $\hat{\mathbf{s}}$ from (12);
initialization: counter=0;
sort $\{r_{1,m}\}_{m=1}^{M}$ in descending order: $r_{1,i_1}\geq\cdots\geq r_{1,i_M}$;
for $n=1,\ldots,\min\{M,P\}$ **do**
  Partition the set $\{1,\ldots,P\}$ into n subsets: $\{1,\ldots,P\}=\mathcal{P}_1\cup\cdots\cup\mathcal{P}_n$;
  for *each candidate partition* **do**
    counter=counter+1;
    Sort the values $\{\alpha^{(1)}_{\mathcal{P}_1},\ldots,\alpha^{(1)}_{\mathcal{P}_n}\}$ in descending order: $\alpha^{(1)}_{\mathcal{P}_{j_1}}\geq\cdots\geq\alpha^{(1)}_{\mathcal{P}_{j_n}}$;
    Evaluate the candidate solution: $\mathbf{s}(counter)=i_1\sum_{p_1\in\mathcal{P}_{j_1}}\mathbf{e}_{p_1}+\cdots+i_n\sum_{p_n\in\mathcal{P}_{j_n}}\mathbf{e}_{p_n}$;
    Evaluate the corresponding weight $w(counter)=\chi^{(1)}(\mathbf{s}(counter))=r_{1,i_1}\alpha^{(1)}_{\mathcal{P}_{j_1}}+\cdots+r_{1,i_n}\alpha^{(1)}_{\mathcal{P}_{j_n}}$;
  end
  $c=\arg\max_{counter}\{w(counter)\}$;
  $\hat{\mathbf{s}}=\mathbf{s}(c)$;
end

Algorithm 1: The Simp-ML MISO Decoder

2) Decoder Complexity: The partitioning of the set $\{1,\ldots,P\}$ into n subsets $\mathcal{P}_1,\ldots,\mathcal{P}_n$ must take into consideration the fact that this partitioning is to be followed by the sorting of $\alpha^{(1)}_{\mathcal{P}_1},\ldots,\alpha^{(1)}_{\mathcal{P}_n}$. Consequently, the order of the subsets in the partition is not important. Therefore, in order to avoid evaluating equivalent partitions, the following rules must be respected.

- *Rule 1*: The cardinalities of the subsets $\mathcal{P}_1,\ldots,\mathcal{P}_n$ must be arranged in increasing order. For example, for $P=3$ and $n=2$, the partitions $\{2\}\cup\{1,3\}$ and $\{1,3\}\cup\{2\}$ are equivalent since in both cases the ML decoder will base its decision on the values $\left\{\min\{\alpha^{(1)}_{\{2\}},\alpha^{(1)}_{\{1,3\}}\},\max\{\alpha^{(1)}_{\{2\}},\alpha^{(1)}_{\{1,3\}}\}\right\}$ that do not change when permuting the subsets.
- *Rule 2*: Even when rule 1 is respected, the cardinalities of some consecutive subsets $\mathcal{P}_i,\mathcal{P}_{i+1},\cdots$ might still be the same. In such cases, the permutations among the subsets having the same cardinalities must not be considered. For example, for $P=5$ and $n=3$, the partitions $\{1\}\cup\{2,3\}\cup\{4,5\}$ and $\{1\}\cup\{4,5\}\cup\{2,3\}$ are equivalent.

Therefore, the partitioning of $\{1,\ldots,P\}$ into n subsets, while respecting the above rules, can be carried out in the following number of ways:

$$\frac{P!}{l_1!l_2!\cdots l_n!}\times\frac{1}{[(\sum_{i=1}^{n}\delta_{l_i,1})!]\cdots[(\sum_{i=1}^{n}\delta_{l_i,P})!]};\ \begin{cases}1\leq l_1\leq l_2\leq\cdots\leq l_n\leq P\\ l_1+l_2+\cdots+l_n=P\end{cases},\quad(13)$$

where $l_i\triangleq|\mathcal{P}_i|$ for $i=1,\ldots,n$. The relation $l_1\leq\cdots\leq l_n$ is introduced to satisfy rule 1. The summation $\sum_{i=1}^{n}\delta_{l_i,p}$ yields the number of elements among $\{l_1,\ldots,l_n\}$ that have the same value of p in $\{1,\ldots,P\}$. The division by $(\sum_{i=1}^{n}\delta_{l_i,p})!$ takes into consideration the fact that the permutations among the $\sum_{i=1}^{n}\delta_{l_i,p}$ corresponding subsets are equivalent satisfying rule 2.

In order to better highlight on the partitioning rules that directly impact the complexity of the Simp-ML decoder, a number of examples will be provided in the case $M\geq P$.

- For $n=1$, $l_1=P$ and $\mathcal{P}_1=\{1,\ldots,P\}$ constitutes the only partitioning option.
- For $n=P$, $(l_1,\ldots,l_P)=(1,\ldots,1)$ and $\{1,\ldots,P\}$ is partitioned as $\{1\}\cup\{2\}\cup\cdots\cup\{P\}$ which, from (13), can be carried out in $\frac{P!}{1!\cdots1!}\times\frac{1}{P!0!\cdots0!}=1$ way.
- For $P=3$ and $n=2$, the only possible solution for (l_1,l_2) (satisfying $1\leq l_1\leq l_2\leq 3$ and $l_1+l_2=3$) is $(l_1,l_2)=(1,2)$. This results in $\frac{3!}{1!2!}\times\frac{1}{1!1!0!}=3$ partitioning options $\{1\}\cup\{2,3\}$, $\{2\}\cup\{1,3\}$ and $\{3\}\cup\{1,2\}$.
- Consider the case $P=4$. For $n=2$, $(l_1,l_2)\in\{(1,3),(2,2)\}$ where the $\frac{4!}{1!3!}\times\frac{1}{1!0!1!0!}=4$ and $\frac{4!}{2!2!}\times\frac{1}{0!2!0!0!}=3$ corresponding partitions are $\{1\}\cup\{2,3,4\}$, $\{2\}\cup\{1,3,4\}$, $\{3\}\cup\{1,2,4\}$ and $\{4\}\cup\{1,2,3\}$ as well as $\{1,2\}\cup\{3,4\}$, $\{1,3\}\cup\{2,4\}$ and $\{1,4\}\cup\{2,3\}$, respectively. In the last case, the partitions $\{2,3\}\cup\{1,4\}$, $\{2,4\}\cup\{1,3\}$ and $\{3,4\}\cup\{1,2\}$ are redundant (the two corresponding sets are flipped) and, hence, must not be considered.

Since the number of transmit apertures P is fixed, then the candidate partitions of $\{1,\ldots,P\}$ can be stored in lookup tables in order to simplify the search process. A key point in the proposed Simp-ML algorithm is that the ordering of $\{r_{1,1},\ldots,r_{1,M}\}$ renders the number of visited candidate information vectors $\mathbf{s}$ independent from M for $M\geq P$ implying that signal constellations with large cardinalities can be advantageously used. Given that the evaluation of $\chi^{(1)}(\mathbf{s})$ requires n multiplications for $u(\mathbf{s})=n$, then from (13), the

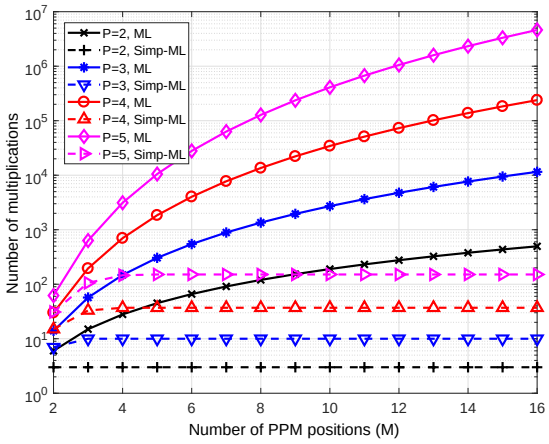

Fig. 1. Number of multiplications needed for the implementation of the ML and Simp-ML algorithms for $Q = 1$.

number of multiplications needed to implement the Simp-ML decoding algorithm takes the following value:

$$N_{\text{mul}}^{(\text{Simp-ML})} = \sum_{n=1}^{\min\{M,P\}} n \times \left[\sum_{\substack{1 \le l_1 \le l_2 \le \cdots \le l_n \le P \\ l_1 + l_2 + \cdots + l_n = P}} \frac{P!}{l_1! l_2! \cdots l_n!} \times \frac{1}{[(\sum_{i=1}^{n} \delta_{l_i,1})!] \cdots [(\sum_{i=1}^{n} \delta_{l_i,P})!]} \right]. \quad (14)$$

Fig. 1 shows the variations of the numbers of multiplications of the ML and Simp-ML algorithms in (11) and (14), respectively, as a function of M for different values of P and for $Q = 1$. It is worth noting that (11) is proportional to the number of receive apertures Q. This figure highlights the significant reduction in the number of multiplications that results from implementing the Simp-ML decoding algorithm. For example, for $M = 8$ and $P = 3$, the simplified MISO detection procedure reduces the number of multiplications from 1352 to 10.

C. The First Simplified Suboptimal MIMO Decoder (Simp-Sub1)

In this section, we propose a simplified suboptimal decoder (denoted by Simp-Sub1) that can be associated with MIMO systems. In the case of MISO systems, the Simp-Sub1 decoder is capable of achieving optimal detection and it simplifies to the Simp-ML decoder presented in Section III-B. The main challenge behind MIMO detection under Poisson statistics resides in the fact that the weighing coefficients $\alpha_{\mathcal{P}}^{(q)}$ in (10) depend jointly on the transmitted symbols and the channel coefficients in a nonlinear manner. This is remarkably different from the MIMO-AWGN systems where the ML rule is given by $\arg\min_{\mathbf{s}} \|\mathbf{y} - \mathbf{H}\mathbf{s}\|^2$ where the channel matrix $\mathbf{H}$ and the symbol vector $\mathbf{s}$ appear separately in this equation ($\mathbf{y}$ is the received vector). In Section III-B, the challenging ML problem, that is equivalent to determining which apertures are pulsed in each of the slots, was simplified by associating ordered decision variables to ordered weighing coefficients related to the tested partition of the set of transmit apertures. While this association is feasible in the case $Q = 1$, it will evidently fail in the case $Q > 1$ since the involved decision variables and weighing coefficients might be ordered in different ways for each of the receive apertures and the ordering might even revolve around different slots for each of these apertures. In other words, the sorting of $\{r_{q,m}\}_{m=1}^{M}$ and of $\{\alpha_{\mathcal{P}}^{(q)}\}_{\mathcal{P} \subset \{1,\ldots,P\}}$ (in algorithm-1) might not be the same for different values of q. Moreover, for $M > P$, the PPM positions of the P largest elements of $\{r_{q,m}\}_{m=1}^{M}$ might vary from one receive aperture to another.

Instead of basing its decision on the QM random variables $R_{q,m}$ for $q = 1, \ldots, Q$ and $m = 1, \ldots, M$, the Simp-Sub1 decoder bases its decision on the summation of the numbers of photoelectrons detected by the different receiver apertures in each of the M PPM slots. In this case, the M resulting decision variables can be sorted in a unique way and algorithm-1 can be readily applied. Defining the random variable R'_m as $R'_m = \sum_{q=1}^{Q} R_{q,m}$, then from (4), R'_m is a Poisson random variable with parameter $\mathrm{E}[R'_m] = \frac{\lambda_s}{P} \sum_{p=1}^{P} \delta_{s_p,m} \left(\sum_{q=1}^{Q} I_{q,p} \right) + Q\lambda_b$. Now, algorithm-1 that was built around the Poisson random variables $\{R_{1,m}\}_{m=1}^{M}$ can be readily applied with the Poisson random variables $\{R'_m\}_{m=1}^{M}$. In this case, the Simp-Sub1 decoder corresponds to applying algorithm-1 while replacing the data inputs $\{r_{1,m}\}_{m=1}^{M}$ and $\{\alpha_{\mathcal{P}}^{(1)}\}_{\mathcal{P} \subset \{1,\ldots,P\}}$ with the inputs $\{r'_m\}_{m=1}^{M}$ and $\{\alpha'_{\mathcal{P}}\}_{\mathcal{P} \subset \{1,\ldots,P\}}$, respectively, where $r'_m = \sum_{q=1}^{Q} r_{q,m}$ and from (10):

$$\alpha'_{\mathcal{P}} = \log \left[1 + \frac{\lambda_s}{QP\lambda_b} \sum_{p \in \mathcal{P}} \delta_{s_p,m} \left(\sum_{q=1}^{Q} I_{q,p} \right) \right] \quad \text{for } \mathcal{P} \subset \{1, \ldots, P\}. \quad (15)$$

The relation $R'_m = \sum_{q=1}^{Q} R_{q,m}$ suggests that the Simp-Sub1 decoder is performing equal gain combining (EGC) prior to applying algorithm-1. In this context, other combining schemes (including MRC) of the form $R'_m = \sum_{q=1}^{Q} a_q R_{q,m}$ (for some arbitrary weights $a_1, \ldots, a_Q$) are not adapted to the noise model and decoding algorithms under consideration since the resulting random variable R'_m will not follow the Poisson distribution. Finally, since the evaluation of the new inputs $\{r'_m\}_{m=1}^{M}$ and $\{\alpha'_{\mathcal{P}}\}_{\mathcal{P} \subset \{1,\ldots,P\}}$ involves only additions, then the number of multiplications required for implementing the Simp-Sub1 algorithm is as provided in (14).

D. The Second Simplified Suboptimal MIMO Decoder (Simp-Sub2)

The second simplified suboptimal decoder (Simp-Sub2) is based on limiting the ML search to M_r positions out of the total M positions where $M_r < M$. In other words, the candidate vectors tested by the ML algorithm (in Section III-A) will be limited to a subset of cardinality M_r^P carved from the set $\{1, \ldots, M\}^P$ of cardinality M^P. Since the P transmitted light pulses can occupy P positions at most, then M_r must also be chosen to satisfy $M_r \ge P$ since the choice $M_r < P$ will inevitably result in an error (even for large values of E_s) when the transmitted pulses occupy between $M_r + 1$

and P positions. Therefore, M_r must satisfy $P \leq M_r < M$. While the Simp-Sub2 decoder limits the ML search to M_r positions with $P \leq M_r < M$ in the case $P < M$, the Simp-Sub1 decoder limits this search to $\min\{M, P\} = P$ positions in this case. In other words, both suboptimal algorithms apply the concept of limiting the search to a subset of positions with the Simp-Sub1 decoder examining a smaller number of positions. Further limiting the number of positions to $M_r' < P$ with Simp-Sub1 will result in error floors where errors will always occur when the P transmit apertures transmit in more than M_r' positions.

Since the presence of a light signal in a certain PPM slot increases the number of detected photoelectrons (on average), then the M_r slots with maximum photoelectron counts should be selected. Given that the ordering of $\{r_{q,m}\}_{m=1}^{M}$ varies from one receive aperture to another, then the sorting must be based on a set of M decision variables that comprises the contributions from all receive apertures. The decision variables $\{r'_m = \sum_{q=1}^{Q} r_{q,m}\}_{m=1}^{M}$ constitute a feasible option for selecting the M_r slots. In other words, ordering $\{r'_m\}_{m=1}^{M}$ in descending order as $r'_{i_1} \geq r'_{i_2} \geq \cdots \geq r'_{i_M}$, then the candidate PPM slots would be $i_1, \ldots, i_{M_r}$ implying that the simplified ML search will be limited to the set $\{i_1, \ldots, i_{M_r}\}^P$:

$$\hat{\mathbf{s}} = \arg \max_{\mathbf{s} \in \{i_1, \ldots, i_{M_r}\}^P} \left\{ \sum_{q=1}^{Q} \chi^{(q)}(\mathbf{s}) \right\}. \tag{16}$$

Given that the selection of the M_r candidate positions does not involve any multiplications (since it is based on sorting), then the number of multiplications required for the implementation of the Simp-Sub2 decoder can be obtained by replacing M with M_r in (11).

Comparing the complexities of the two suboptimal MIMO decoders, it can be observed from (11) and (14) that the complexity of Simp-Sub2 increases linearly with Q while the complexity of Simp-Sub1 is independent of Q (because of EGC). Denoting the numbers of multiplications in (11) and (14) by $N_{\text{mul}}^{(\text{ML})}(Q, P, M)$ and $N_{\text{mul}}^{(\text{Simp-ML})}(P, M)$, respectively, then:

$$\begin{aligned} &N_{\text{mul}}^{(\text{Simp-ML})}(P, M) \leq N_{\text{mul}}^{(\text{Simp-ML})}(P, \infty) \leq N_{\text{mul}}^{(\text{ML})}(1, P, P) \\ &\leq N_{\text{mul}}^{(\text{ML})}(Q, P, P) \leq N_{\text{mul}}^{(\text{ML})}(Q, P, M_r) \leq N_{\text{mul}}^{(\text{ML})}(Q, P, M), \end{aligned} \tag{17}$$

where the last two inequalities follow since the complexity of the ML decoder increases with the number of tested positions (where $P \leq M_r < M$) while the third inequality follows since this complexity increases with Q. Similarly, the first inequality follows from $N_{\text{mul}}^{(\text{Simp-ML})}(P, M)$ being an increasing function of M. Finally, the second key inequality follows from Fig. 1. As a conclusion, given that the complexities of the Simp-Sub1 and Simp-Sub2 decoders are given by $N_{\text{mul}}^{(\text{Simp-ML})}(P, M)$ and $N_{\text{mul}}^{(\text{ML})}(Q, P, M_r)$, respectively, then the Simp-Sub2 decoder requires a larger number of multiplications for all values of M_r. This additional complexity will be associated with improved near-optimal performance levels as will be highlighted later.

While the decoders' complexities have been evaluated in terms of the number of multiplications, sorting procedures need to be implemented by these decoders. In particular, the Simp-ML and Simp-Sub1 decoders require carrying out two types of sorting; namely, the sorting of $\{\alpha'_{\mathcal{P}}\}_{\mathcal{P} \subset \{1,\ldots,P\}}$ and $\{r'_m\}_{m=1}^{M}$ where these quantities simplify to $\{\alpha_{\mathcal{P}}^{(1)}\}_{\mathcal{P} \subset \{1,\ldots,P\}}$ and $\{r_{1,m}\}_{m=1}^{M}$ for the Simp-ML decoder. On the other hand, the Simp-Sub2 decoder necessitates the sorting of the M decision variables $\{r'_m\}_{m=1}^{M}$. The weights $\alpha'_{\mathcal{P}}$ do not depend on the values of the transmitted symbols and, consequently, the sorting of these elements can be carried out at the beginning of each fading block that extends over thousands of symbol durations in the case of FSO systems that have large coherence times. Therefore, the involved computational complexity can be ignored. Consequently, the main additional complexity arises from the sorting of $\{r'_m\}_{m=1}^{M}$ that needs to be carried out on a symbol-by-symbol basis with Simp-Sub1 and Simp-Sub2.

However, the sorting algorithms are based on comparisons and do not involve any complicated processing-demanding multiplication operations. In this context, many powerful sorting algorithms have been proposed in the literature and any of these algorithms can be readily used to perform the required sorting with a marginal impact on the decoders' complexities. While the sorting algorithms have their challenges, typical algorithms have a good running time complexity and space complexity of $O(n \log(n))$ and $O(1)$, respectively, where n is the size of the list to be sorted. For the applications considered in this paper, the required sorting of M elements is judged to be unchallenging given that the number of positions M does not assume excessively large values. In fact, the sorting algorithms, that are designed to sort thousands of data entries, can easily handle the sorting of M elements where the values of M do not exceed 16 in practice. With such small number of elements to be sorted, the running time and memory usage are usually not critical. This is especially true since the number of elements to be sorted, and hence the number of required addition-based comparisons, is much smaller than the number of multiplications in (11) and (14). For example, for $P = 5$ and $M = 4$, $N_{\text{mul}}^{(\text{ML})} = 3124$ and $N_{\text{mul}}^{(\text{Simp-ML})} = 146$ that are both large compared to 4.

Finally, it is worth highlighting that the sorting requirements can be further relaxed where effectively only the P largest elements among $\{r'_m\}_{m=1}^{M}$ need to be sorted with Simp-Sub1 for $P < M$ while the M_r largest elements among $\{r'_m\}_{m=1}^{M}$ need to be sorted with Simp-Sub2.

The proposed algorithms possess appealing parallelism capabilities where the search can be evenly split among N_{proc} processors thus reducing the processing time and latency by a factor of N_{proc}. This can be readily realized by partitioning the sets $\{1, \ldots, M\}^P$ and $\{i_1, \ldots, i_{M_r}\}^P$ in (9) and (16) into N_{proc} subsets each spanned by one of the available processors. The latency constraints further justify the need for suboptimal decoders with low processing requirements. There requirements are particularly met by the Simp-Sub1 decoder that entails a very small number of multiplications as highlighted in Fig. 1.

IV. Performance Analysis and Constellation Confinement

In this section, we derive the symbol error probability (SEP) of the considered SMux scheme. The conditional SEP (conditioned on the channel irradiances) can be expressed as:

$$P_e = \frac{1}{M^P} \sum_{\mathbf{s}\in\{1,\ldots,M\}^P} \sum_{\substack{\mathbf{s}'\in\{1,\ldots,M\}^P \\ \mathbf{s}'\neq\mathbf{s}}} \Pr(\mathbf{s}\rightarrow\mathbf{s}'), \qquad (18)$$

where $\Pr(\mathbf{s}\rightarrow\mathbf{s}')$ is the pairwise error probability of transmitting the information vector $\mathbf{s}$ and deciding in favor of $\mathbf{s}'\neq\mathbf{s}$. The SEP calculations will be based on the ML decision rule given in (8) where we assume that ties are always broken in favor of the erroneous symbols with the consequence that the derived SEP expression corresponds to an upper-bound.

A. Mathematical Preliminaries

The probabilities in this section can be written under the general form $\Pr\left(\sum_i a_i X_i \geq 0\right)$ for some constants $a_1, a_2, \ldots$ where $X_1, X_2, \ldots$ are Poisson random variables (r.v.s) with parameters $\lambda_1, \lambda_2, \ldots$. This probability will be upper-bounded using the Chernoff bound $\Pr(X\geq b)\leq e^{-bt}M_X(t)$ for every $t>0$ where $M_X(t)=\mathrm{E}[e^{tX}]$ stands for the moment generating function of the r.v. X. Consequently, $\Pr\left(\sum_i a_i X_i\geq 0\right)\leq M_{\sum_i a_i X_i}(t)=\prod_i M_{X_i}(a_i t)$ resulting in:

$$\Pr\left(\sum_i a_i X_i \geq 0\right) \leq \exp\left(\sum_i \lambda_i (e^{a_i t}-1)\right) \;;\; \forall\, t>0, \qquad (19)$$

since $M_X(t) = \exp\left(\lambda\left(e^t-1\right)\right)$ for a Poisson r.v. with parameter λ.

While the expression in (19) will be used for bounding the SEP, a consequent simpler expression will be used for the sake of confining the transmitted constellation based on an asymptotic analysis. This target can be achieved by first observing that, based on (4), the parameters of the concerned r.v.s X_i can be written under the general form $\lambda_i=\mathrm{E}[X_i]=k_i\lambda_s+\lambda_b$ where the constant k_i can be either zero or positive. For asymptotically large values of λ_s, the r.v.s with $k_i=0$ (whose parameters are equal to λ_b) will generate a number of photoelectrons that is considerably smaller than that generated by the Poisson r.v.s for which $k_i\neq 0$. Therefore, $\sum_i a_i X_i$ can be approximated by $\sum_{i\ |\ k_i\neq 0} a_i X_i \triangleq \sum_j a_j X_j$ for $\lambda_s\gg 1$.

Now, the parameters λ_i of the Poisson r.v.s having $k_i\neq 0$ are very large implying that they can be approximated by the Gaussian distribution. Using the Gaussian approximation, the r.v. $\sum_j a_j X_j$ can be approximated with a Gaussian r.v. with mean $\mu=\sum_j a_j\lambda_j$ and variance $\sigma^2=\sum_j a_j^2\lambda_j$. Consequently, for large values of λ_s, (19) can be further approximated by $\frac{1}{2}\mathrm{erfc}\left(-\frac{\mu}{\sqrt{2}\sigma}\right)$ where $\mathrm{erfc}(x)=\frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}\mathrm{d}t$ is the complementary error function. Replacing μ and σ by their values, (19) can be further approximated by the following asymptotic expression:

$$\Pr\left(\sum_i a_i X_i \geq 0\right) \approx \frac{1}{2}\mathrm{erfc}\left(-\frac{\sum_{i\ |\ \lambda_i\neq\lambda_b}\left(a_i(\lambda_i-\lambda_b)\right)}{\sqrt{2\sum_{i\ |\ \lambda_i\neq\lambda_b}\left(a_i^2(\lambda_i-\lambda_b)\right)}}\right), \qquad (20)$$

where the last approximation follows since $\lambda_i=k_i\lambda_s+\lambda_b$ tends to $k_i\lambda_s$ (i.e. $\lambda_i-\lambda_b$) for $\lambda_s\gg 1$.

B. $P=2$ Transmit Apertures

In order to offer more insights on the calculation procedures, we first consider the special case of $P=2$. We denote the transmitted vector by $\mathbf{s}=(s_1,s_2)$ and the output of the ML decoder by $\mathbf{s}'=(s_1',s_2')$. The two following cases arise depending on whether $s_1=s_2$ or $s_1\neq s_2$.

1) Case I: $s_1=s_2\triangleq s$ implying from (4) that $R_{q,s}$ is a Poisson r.v. with parameter $(I_{q,1}+I_{q,2})\frac{\lambda_s}{2}+\lambda_b$ while $R_{q,m}$ has a parameter λ_b for $m'\neq s$. The following cases arise[1].

Case 1.1: In this case $s_1'=s_2'\triangleq s'\neq s$. Based on the ML decision rule in (8), the error probability can be written as $\Pr\left(\sum_{q=1}^Q \alpha_{\{1,2\}}^{(q)} R_{q,s'}\geq\sum_{q=1}^Q\alpha_{\{1,2\}}^{(q)}R_{q,s}\right)$ where the constant $\alpha_{\mathcal{P}}^{(q)}$ is defined in (10). Based on (19), this probability can be bounded as:

$$\mathfrak{p}_{1,1}\leq\exp\left(\sum_{q=1}^Q\left[\left((I_{q,1}+I_{q,2})\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{-\alpha_{\{1,2\}}^{(q)}t}-1\right)+\lambda_b\left(e^{\alpha_{\{1,2\}}^{(q)}t}-1\right)\right]\right). \qquad (21)$$

Following from (20):

$$\mathfrak{p}_{1,1}\approx\frac{1}{2}\mathrm{erfc}\left(\frac{\sum_{q=1}^Q\alpha_{\{1,2\}}^{(q)}(I_{q,1}+I_{q,2})\frac{\lambda_s}{2}}{\sqrt{2\sum_{q=1}^Q[\alpha_{\{1,2\}}^{(q)}]^2(I_{q,1}+I_{q,2})\frac{\lambda_s}{2}}}\right)\mapsto\mathrm{erfc}\left(\sqrt{\max_q\{I_{q,1}+I_{q,2}\}\lambda_s}\right), \qquad (22)$$

where $f(\lambda_s)\mapsto f'(\lambda_s)$ means that the function $f(\lambda_s)$ behaves asymptotically like the function $f'(\lambda_s)$ as $\lambda_s\rightarrow\infty$. The asymptotic value in (22) follows since $\alpha_{\mathcal{P}}^{(q)}\mapsto\log(\lambda_s)$ which is much smaller than λ_s as $\lambda_s\rightarrow\infty$.

Finally, $\mathfrak{p}_{1,1}$ appears $M(M-1)$ times in the SEP expression in (18) corresponding to the number of ways of selecting the two distinct positions s and s' from the total M positions.

Case 1.2: In this case $s_1'\neq s_2'$ which results in the three following possibilities.

– *Case 1.2.1*: $s_1'=s$ implying, from (8), that the corresponding error probability is

[1]The cases will be named as "Case $n.n'$" where n and n' are the numbers of distinct PPM slots in $\mathbf{s}$ and $\mathbf{s}'$, respectively.

$\Pr\left(\sum_{q=1}^{Q}\left[\alpha_{\{1\}}^{(q)}R_{q,s}+\alpha_{\{2\}}^{(q)}R_{q,s_2'}\right]\geq\sum_{q=1}^{Q}\alpha_{\{1,2\}}^{(q)}R_{q,s}\right)$. Based on (19), this probability can be bounded as:

$$\mathfrak{p}_{1,2,1}\leq\exp\left(\sum_{q=1}^{Q}\left[\left((I_{q,1}+I_{q,2})\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{(\alpha_{\{1\}}^{(q)}-\alpha_{\{1,2\}}^{(q)})t}-1\right)+\lambda_b\left(e^{\alpha_{\{2\}}^{(q)}t}-1\right)\right]\right). \quad (23)$$

– *Case 1.2.2*: $s_2'=s$ implying that the error probability can be obtained by interchanging the subscripts 1 and 2 in (23):

$$\mathfrak{p}_{1,2,2}\leq\exp\left(\sum_{q=1}^{Q}\left[\left((I_{q,1}+I_{q,2})\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{(\alpha_{\{2\}}^{(q)}-\alpha_{\{1,2\}}^{(q)})t}-1\right)+\lambda_b\left(e^{\alpha_{\{1\}}^{(q)}t}-1\right)\right]\right). \quad (24)$$

– *Case 1.2.3*: $s_1'\neq s$ and $s_2'\neq s$. In this case, the error probability can be written as $\Pr\left(\sum_{q=1}^{Q}\left[\alpha_{\{1\}}^{(q)}R_{q,s_1'}+\alpha_{\{2\}}^{(q)}R_{q,s_2'}\right]\geq\sum_{q=1}^{Q}\alpha_{\{1,2\}}^{(q)}R_{q,s}\right)$ resulting in:

$$\mathfrak{p}_{1,2,3}\leq\exp\left(\sum_{q=1}^{Q}\left[\left((I_{q,1}+I_{q,2})\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{-\alpha_{\{1,2\}}^{(q)}t}-1\right)+\lambda_b\left(e^{\alpha_{\{1\}}^{(q)}t}-1\right)+\lambda_b\left(e^{\alpha_{\{2\}}^{(q)}t}-1\right)\right]\right). \quad (25)$$

The asymptotic expressions of $\mathfrak{p}_{1,2,1}$, $\mathfrak{p}_{1,2,2}$ and $\mathfrak{p}_{1,2,3}$ can be obtained by replacing $\alpha_{\{1,2\}}^{(q)}$ in (22) with $\alpha_{\{1,2\}}^{(q)}-\alpha_{\{1\}}^{(q)}$, $\alpha_{\{1,2\}}^{(q)}-\alpha_{\{2\}}^{(q)}$ and $\alpha_{\{1,2\}}^{(q)}$, respectively. Consequently, the asymptotic variations of these three probabilities with λ_s is captured by the expression given in (22).

On the other hand, the expression in (23) appears $M(M-1)$ times in P_e which corresponds to the number of ways of selecting the symbols s and s_1'. The same holds for (24). Finally, the probability in (25) appears $M(M-1)(M-2)$ times which corresponds to the number of ways of selecting the three distinct positions s, s_1' and s_2' from the total of M positions.

2) Case II: In this case, $\mathbf{s}$ occupies two distinct positions $s_2\neq s_1$. For this scenario, R_{q,s_1} and R_{q,s_2} follow the Poisson distributions with parameters $I_{q,1}\frac{\lambda_s}{2}+\lambda_b$ and $I_{q,2}\frac{\lambda_s}{2}+\lambda_b$, respectively, while the parameters of the remaining $M-2$ r.v.s will be equal to λ_b. The two following cases arise depending on whether $\mathbf{s}'$ extends over one or two PPM positions.

Case 2.1: In this case $s_2'=s_1'\triangleq s'$ resulting in one of the following situations:

– *Case 2.1.1*: $s'=s_1$. From (8), the error probability can be derived from $\Pr\left(\sum_{q=1}^{Q}\alpha_{\{1,2\}}^{(q)}R_{q,s_1}\geq\sum_{q=1}^{Q}\left[\alpha_{\{1\}}^{(q)}R_{q,s_1}+\alpha_{\{2\}}^{(q)}R_{q,s_2}\right]\right)$ which, from (19), results in:

$$\mathfrak{p}_{2,1,1}\leq\exp\left(\sum_{q=1}^{Q}\left[\left(I_{q,1}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{(\alpha_{\{1,2\}}^{(q)}-\alpha_{\{1\}}^{(q)})t}-1\right)+\left(I_{q,2}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{-\alpha_{\{2\}}^{(q)}t}-1\right)\right]\right). \quad (26)$$

From (20):

$$\mathfrak{p}_{2,1,1}\approx\frac{1}{2}\mathrm{erfc}\left(\frac{\sum_{q=1}^{Q}\left[\left(\alpha_{\{1\}}^{(q)}-\alpha_{\{1,2\}}^{(q)}\right)I_{q,1}\frac{\lambda_s}{2}+\alpha_{\{2\}}^{(q)}I_{q,2}\frac{\lambda_s}{2}\right]}{\sqrt{2\sum_{q=1}^{Q}\left[\left(\alpha_{\{1\}}^{(q)}-\alpha_{\{1,2\}}^{(q)}\right)^2I_{q,1}\frac{\lambda_s}{2}+\left(\alpha_{\{2\}}^{(q)}\right)^2I_{q,2}\frac{\lambda_s}{2}\right]}}\right). \quad (27)$$

Since $\alpha_{\{1\}}^{(q)}-\alpha_{\{1,2\}}^{(q)}\mapsto k$ (constant) while $\alpha_{\{2\}}^{(q)}\mapsto\log(\lambda_s)$, then $\mathfrak{p}_{2,1,1}\mapsto\mathrm{erfc}\left(\sqrt{\max_q\{I_{q,2}\}\lambda_s}\right)$.

– *Case 2.1.2*: $s'=s_2$. The corresponding error probability can be obtained by interchanging the subscripts 1 and 2 in (26):

$$\mathfrak{p}_{2,1,2}\leq\exp\left(\sum_{q=1}^{Q}\left[\left(I_{q,2}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{(\alpha_{\{1,2\}}^{(q)}-\alpha_{\{2\}}^{(q)})t}-1\right)+\left(I_{q,1}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{-\alpha_{\{1\}}^{(q)}t}-1\right)\right]\right), \quad (28)$$

with $\mathfrak{p}_{2,1,2}\mapsto\mathrm{erfc}\left(\sqrt{\max_q\{I_{q,1}\}\lambda_s}\right)$.

– *Case 2.1.3*: $s'\neq s_1$ and $s'\neq s_2$. In this case, the error probability is given by: $\Pr\left(\sum_{q=1}^{Q}\alpha_{\{1,2\}}^{(q)}R_{q,s'}\geq\sum_{q=1}^{Q}\left[\alpha_{\{1\}}^{(q)}R_{q,s_1}+\alpha_{\{2\}}^{(q)}R_{q,s_2}\right]\right)$ which, from (19), results in:

$$\mathfrak{p}_{2,1,3}\leq\exp\left(\sum_{q=1}^{Q}\left[\left(I_{q,1}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{-\alpha_{\{1\}}^{(q)}t}-1\right)+\left(I_{q,2}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{-\alpha_{\{2\}}^{(q)}t}-1\right)+\lambda_b\left(e^{\alpha_{\{1,2\}}^{(q)}t}-1\right)\right]\right)\mapsto\mathrm{erfc}\left(\sqrt{\max_q\{\max\{I_{q,1},I_{q,2}\}\}\lambda_s}\right). \quad (29)$$

While the symbols s_1 and s_2 can be selected in $M(M-1)$ ways, s' can be selected in 1, 1 and $M-2$ ways according to cases 2.1.1, 2.1.2 and 2.1.3, respectively. Therefore, in (18), there are $M(M-1)$, $M(M-1)$ and $M(M-1)(M-2)$ vectors $\mathbf{s}$ and $\mathbf{s}'$ whose pairwise error probabilities are equal to $\mathfrak{p}_{2,1,1}$, $\mathfrak{p}_{2,1,2}$ and $\mathfrak{p}_{2,1,3}$, respectively.

Case 2.2: In this case $s_2'\neq s_1'$ resulting in one of the six following scenarios:

– *Case 2.2.1*: $s_1'=s_1$ and $s_2'\neq s_2$. In this case, the error probability can be calculated from $\Pr\left(\sum_{q=1}^{Q}\left[\alpha_{\{1\}}^{(q)}R_{q,s_1}+\alpha_{\{2\}}^{(q)}R_{q,s_2'}\right]\geq\sum_{q=1}^{Q}\left[\alpha_{\{1\}}^{(q)}R_{q,s_1}+\alpha_{\{2\}}^{(q)}R_{q,s_2}\right]\right)$ that simplifies to $\Pr\left(\sum_{q=1}^{Q}\left[\alpha_{\{2\}}^{(q)}\left(-R_{q,s_2}+R_{q,s_2'}\right)\right]\geq 0\right)$ that results in the following bound following from (19):

$$\mathfrak{p}_{2,2,1}\leq\exp\left(\sum_{q=1}^{Q}\left[\left(I_{q,2}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{-\alpha_{\{2\}}^{(q)}t}-1\right)+\lambda_b\left(e^{\alpha_{\{2\}}^{(q)}t}-1\right)\right]\right)\mapsto\mathrm{erfc}\left(\sqrt{\max_q\{I_{q,2}\}\lambda_s}\right). \quad (30)$$

– *Case 2.2.2*: $s_1' \neq s_1$ and $s_2' = s_2$. The error probability can be obtained by replacing the subscript 2 with 1 in (30) resulting in:

$$\mathfrak{p}_{2,2,2} \leq \exp\left(\sum_{q=1}^{Q}\left[\left(I_{q,1}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{-\alpha_{\{1\}}^{(q)}t}-1\right)+\lambda_b\left(e^{\alpha_{\{1\}}^{(q)}t}-1\right)\right]\right) \mapsto \text{erfc}\left(\sqrt{\max_q\{I_{q,1}\}\lambda_s}\right). \quad (31)$$

– *Case 2.2.3*: $s_1' = s_2$ and $s_2' \neq s_1$. Following the same calculation procedures as in the previous cases, the error probability can be bounded as:

$$\mathfrak{p}_{2,2,3} \leq \exp\left(\sum_{q=1}^{Q}\left[\left(I_{q,1}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{-\alpha_{\{1\}}^{(q)}t}-1\right)+\left(I_{q,2}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{(\alpha_{\{1\}}^{(q)}-\alpha_{\{2\}}^{(q)})t}-1\right)+\lambda_b\left(e^{\alpha_{\{2\}}^{(q)}t}-1\right)\right]\right), \quad (32)$$

with $\mathfrak{p}_{2,2,3} \mapsto \text{erfc}\left(\sqrt{\max_q\{I_{q,1}\}\lambda_s}\right)$ following a reasoning similar to the one provided in (27).

– *Case 2.2.4*: $s_1' \neq s_2$ and $s_2' = s_1$. By interchanging the roles of the two transmit apertures, the error probability follows from (32):

$$\mathfrak{p}_{2,2,4} \leq \exp\left(\sum_{q=1}^{Q}\left[\left(I_{q,1}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{(\alpha_{\{2\}}^{(q)}-\alpha_{\{1\}}^{(q)})t}-1\right)+\left(I_{q,2}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{-\alpha_{\{2\}}^{(q)}t}-1\right)+\lambda_b\left(e^{\alpha_{\{1\}}^{(q)}t}-1\right)\right]\right) \mapsto \text{erfc}\left(\sqrt{\max_q\{I_{q,2}\}\lambda_s}\right). \quad (33)$$

Each one of the probabilities in (30), (31), (32) and (33) appears $M(M-1)(M-2)$ times in (18) which corresponds to the number of ways of selecting three distinct M-PPM symbols (s_1, s_2 and one of the symbols s_1' or s_2').

– *Case 2.2.5*: $s_1' \neq s_1$, $s_1' \neq s_2$, $s_2' \neq s_1$ and $s_2' \neq s_2$ resulting in:

$$\mathfrak{p}_{2,2,5} \leq \exp\left(\sum_{q=1}^{Q}\left[\left(I_{q,1}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{-\alpha_{\{1\}}^{(q)}t}-1\right)+\left(I_{q,2}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{-\alpha_{\{2\}}^{(q)}t}-1\right)+\lambda_b\left(e^{\alpha_{\{1\}}^{(q)}t}-1\right)+\lambda_b\left(e^{\alpha_{\{2\}}^{(q)}t}-1\right)\right]\right) \mapsto \text{erfc}\left(\sqrt{\max_q\{\max\{I_{q,1},I_{q,2}\}\}\lambda_s}\right). \quad (34)$$

The probability $\mathfrak{p}_{2,2,5}$ appears $M(M-1)(M-2)(M-3)$ times in (18) which corresponds to the number of ways of selecting four distinct symbols.

– *Case 2.2.6*: $s_1' = s_2$ and $s_2' = s_1$. In this case, the error probability follows from $\Pr\left(\sum_{q=1}^{Q}\left[\alpha_{\{1\}}^{(q)}R_{q,s_2}+\alpha_{\{2\}}^{(q)}R_{q,s_1}\right] \geq \sum_{q=1}^{Q}\left[\alpha_{\{1\}}^{(q)}R_{q,s_1}+\alpha_{\{2\}}^{(q)}R_{q,s_2}\right]\right)$ that can be written as: $\Pr\left(\sum_{q=1}^{Q}\left[(\alpha_{\{2\}}^{(q)}-\alpha_{\{1\}}^{(q)})(R_{q,s_1}-R_{q,s_2})\right] \geq 0\right)$ implying that:

$$\mathfrak{p}_{2,2,6} \leq \exp\left(\sum_{q=1}^{Q}\left[\left(I_{q,1}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{(\alpha_{\{2\}}^{(q)}-\alpha_{\{1\}}^{(q)})t}-1\right)+\left(I_{q,2}\frac{\lambda_s}{2}+\lambda_b\right)\left(e^{(\alpha_{\{1\}}^{(q)}-\alpha_{\{2\}}^{(q)})t}-1\right)\right]\right), \quad (35)$$

with $\mathfrak{p}_{2,2,6} \mapsto \text{erfc}\left(\max_q\left\{\frac{|I_{q,1}-I_{q,2}|}{\sqrt{I_{q,1}+I_{q,2}}}\right\}\sqrt{\lambda_s}\right)$ where this terms appears $M(M-1)$ times in (18).

Combining the above results, the conditional SEP in (18) can be upper-bounded as follows:

$$P_e \leq \frac{1}{M^2}M(M-1)\left[\mathfrak{p}_{1,1}+\mathfrak{p}_{1,2,1}+\mathfrak{p}_{1,2,2}+\mathfrak{p}_{2,1,1}+\mathfrak{p}_{2,1,2}+\mathfrak{p}_{2,2,6}+(M-2)\left[\mathfrak{p}_{1,2,3}+\mathfrak{p}_{2,1,3}+\mathfrak{p}_{2,2,1}+\mathfrak{p}_{2,2,2}+\mathfrak{p}_{2,2,3}+\mathfrak{p}_{2,2,4}+(M-3)\mathfrak{p}_{2,2,5}\right]\right]. \quad (36)$$

Finally, it is worth noting that while the inequalities in (21), (23)-(25), (26), (28), (29) and (30)-(35) hold for all positive values of t, the value of t can be further optimized to achieve the tightest possible bound. Independent optimizations need to be carried out for each one of these equations rather than optimizing (36) directly. This approach not only results in simpler optimization procedures, but also it minimizes the upper-bound since each one of the 13 constituent terms in (36) will be optimized separately. The value of t that minimizes an expression that takes the general form in (19) can be obtained by solving the equation $\sum_i a_i\lambda_i e^{a_i t} = 0$. While it is hard to solve this equation analytically (with more than two summands), it can be easily solved numerically using any of the widely available mathematical softwares.

3) Constellation Confinement: The interest of the presented upper-bounding technique resides in the fact that the derived expressions are tractable and, hence, they can be further exploited for the sake of confining the transmitted constellation as a means of enhancing the error performance. We start by defining the following constants that depend on the channel irradiances:

$$\beta_1 = \max_q\{I_{q,1}\} \;\;;\;\; \beta_2 = \max_q\{I_{q,2}\}$$
$$\beta_{1\oplus 2} = \max_q\{I_{q,1}+I_{q,2}\} \;\;;\;\; \beta_{1\otimes 2} = \max_q\{\max\{I_{q,1},I_{q,2}\}\}$$
$$\beta_{1\ominus 2} = \max_q\left\{\frac{|I_{q,1}-I_{q,2}|}{\sqrt{I_{q,1}+I_{q,2}}}\right\}, \quad (37)$$

where β_1, β_2 and $\beta_{1\ominus 2}$ are all smaller than $\beta_{1\otimes 2}$ that, in turn, is smaller than $\beta_{1\oplus 2}$.

We define $c(\mathbf{s},\mathbf{s}')$ as the number of elements that are common to the vectors $\mathbf{s}$ and $\mathbf{s}'$. Based on this definition and on the notations in (37), the probabilities that were derived in

the previous subsection can be categorized as follows:

$$c(\mathbf{s},\mathbf{s}')=0:\ \{\mathfrak{p}_{1,1},\mathfrak{p}_{1,2,3}\}\mapsto \text{erfc}(\sqrt{\beta_{1\oplus 2}\lambda_s})\ ;\ \{\mathfrak{p}_{2,1,3},\mathfrak{p}_{2,2,5}\}\mapsto \text{erfc}(\sqrt{\beta_{1\otimes 2}\lambda_s}) \quad (38)$$

$$c(\mathbf{s},\mathbf{s}')=1:\ \{\mathfrak{p}_{1,2,1},\mathfrak{p}_{1,2,2}\}\mapsto \text{erfc}(\sqrt{\beta_{1\oplus 2}\lambda_s})\ ;\ \{\mathfrak{p}_{2,1,2},\mathfrak{p}_{2,2,2},\mathfrak{p}_{2,2,3}\}\mapsto \text{erfc}(\sqrt{\beta_1\lambda_s})\ ;\ \{\mathfrak{p}_{2,1,1},\mathfrak{p}_{2,2,1},\mathfrak{p}_{2,2,4}\}\mapsto \text{erfc}(\sqrt{\beta_2\lambda_s}) \quad (39)$$

$$c(\mathbf{s},\mathbf{s}')=2:\ \mathfrak{p}_{2,2,6}\mapsto \text{erfc}(\sqrt{\beta_{1\ominus 2}\lambda_s}). \quad (40)$$

Regarding (39), it is worth noting that whenever the term $\mathfrak{p}_{1,2,1}$ appears in $\Pr(\mathbf{s}\to\mathbf{s}')$, then the probability $\mathfrak{p}_{2,1,1}$ will appear in $\Pr(\mathbf{s}'\to\mathbf{s})$. Similarly, the terms $\mathfrak{p}_{1,2,2}$ and $\mathfrak{p}_{2,1,2}$ will always appear together when the roles of $\mathbf{s}$ and $\mathbf{s}'$ are interchanged. On the other hand, the channel state information is not available to the transmitter; hence, the transmitter can not adapt the transmitted constellation to the variables β_1 and β_2. Consequently, whenever the constellation contains vectors satisfying $c(\mathbf{s},\mathbf{s}')=1$, the corresponding probability $\Pr(\mathbf{s}\to\mathbf{s}')+\Pr(\mathbf{s}'\to\mathbf{s})$ will behave asymptotically as $\text{erfc}(\sqrt{\min\{\beta_1,\beta_2\}\lambda_s})$. Regarding (38), the terms $\mathfrak{p}_{1,1}$ and $\mathfrak{p}_{2,2,5}$ can appear alone (depending on the specific constellation confinement) unlike the probabilities $\mathfrak{p}_{1,2,3}$ and $\mathfrak{p}_{2,1,3}$ that always appear together. Therefore, for $c(\mathbf{s},\mathbf{s}')=0$, the corresponding error probability can behave asymptotically as either $\text{erfc}(\sqrt{\beta_{1\oplus 2}\lambda_s})$ or $\text{erfc}(\sqrt{\beta_{1\otimes 2}\lambda_s})$.

Based on the above analysis, the subsets of $\{1,\ldots,M\}^P$ can be classified in terms of the number of admissible common elements between the different information vectors. The following notation will be used:

$$\mathcal{C}_{\mathcal{U}}\triangleq\{\mathbf{s}\mid\{c(\mathbf{s},\mathbf{s}')\ \ \forall\ \ \mathbf{s}'\neq\mathbf{s}\}=\mathcal{U}\}, \quad (41)$$

resulting in different possible constellation confinements with the following options (i): $\mathcal{C}_{\{0\}}$ where $P_e\mapsto\text{erfc}(\sqrt{\beta_{1\oplus 2}\lambda_s})$ or $P_e\mapsto\text{erfc}(\sqrt{\beta_{1\otimes 2}\lambda_s})$, (ii): $\mathcal{C}_{\{1\}}$ or $\mathcal{C}_{\{0,1\}}$ where $P_e\mapsto\text{erfc}(\sqrt{\min\{\beta_1,\beta_2\}\lambda_s})$, (iii): $\mathcal{C}_{\{2\}}$ or $\mathcal{C}_{\{0,2\}}$ where $P_e\mapsto\text{erfc}(\sqrt{\beta_{1\ominus 2}\lambda_s})$, (iv): $\mathcal{C}_{\{1,2\}}$ or $\mathcal{C}_{\{0,1,2\}}$ where $P_e\mapsto\text{erfc}(\sqrt{\min\{\beta_1,\beta_2,\beta_{1\ominus 2}\}\lambda_s})$. Evidently, the options $\mathcal{C}_{\{1\}}$, $\mathcal{C}_{\{2\}}$ and $\mathcal{C}_{\{1,2\}}$ are to be excluded since the alternatives $\mathcal{C}_{\{0,1\}}$, $\mathcal{C}_{\{0,2\}}$ and $\mathcal{C}_{\{0,1,2\}}$, respectively, result in larger numbers of elements while achieving the same asymptotic performance.

– The set $\mathcal{C}_{\{0,1,2\}}$ with maximum cardinality is the set $\{1,\ldots,M\}^P$ corresponding to the SMux case with no constellation confinement.

– Sets of the form $\mathcal{C}_{\{0\}}$ with maximum cardinality can be constructed as follows. Select an even integer $M'\leq M$. Divide the M' positions among unique $M'/2$ pairs and allocate each to a vector of the form (s_1,s_2) with $s_1<s_2$ and allocate the remaining $M-M'$ positions among vectors of the form (s_1,s_1). For example, for $M=3$, the sets $\{(1,1),(2,2),(3,3)\}$ ($M'=0$) and $\{(1,2),(3,3)\}$, $\{(1,3),(2,2)\}$, $\{(2,3),(1,1)\}$ ($M'=2$) are valid options. The cardinality in this case is $M-M'+M'/2=M-M'/2$ which is maximized for $M'=0$. This corresponds to the RC solution [3]–[9] where all apertures transmit in the same PPM position: $\mathbf{s}=(s,s)$ with $s\in\{1,\ldots,M\}$. While this scheme results in the smallest possible value of $P_e\mapsto\text{erfc}(\sqrt{\beta_{1\oplus 2}\lambda_s})$ (since $\beta_{1\oplus 2}$ is the largest among the constants defined in (37)), it suffers from the smallest cardinality that is equal to M (which is the same as in single-aperture systems).

– Sets of the form $\mathcal{C}_{\{0,2\}}$ can be constructed as follows. Select an even integer $M'\leq M$. Divide the M' positions among unique $M'/2$ pairs and allocate each to two vectors of the form (s_1,s_2) and (s_2,s_1) with $s_2\neq s_1$ and allocate the remaining $M-M'$ positions among vectors of the form (s_1,s_1). This results in the maximum cardinality of M implying that RC is better since it achieves the same cardinality with a smaller error $\text{erfc}(\sqrt{\beta_{1\oplus 2}\lambda_s})\leq\text{erfc}(\sqrt{\beta_{1\ominus 2}\lambda_s})$.

– Sets of the form $\mathcal{C}_{\{0,1\}}$ correspond to the constellation confinement that we propose. This can be achieved by avoiding having vectors of the form (s_1,s_2) and (s_2,s_1) together in the signal set (for $s_2\neq s_1$). The maximum cardinality can be obtained by adding $\binom{M}{1}$ (number of ways of selecting vectors of the form (s_1,s_1)) and $\binom{M}{2}$ (number of ways of selecting two distinct positions from the M positions). In other words, the proposed confined constellation in the case of $P=2$ apertures is given by $\mathcal{S}=\{(s_1,s_2)\ \mid\ s_1\leq s_2\}$ having a cardinality of $\frac{M(M+1)}{2}$ with an error probability that scales as $P_e\mapsto\text{erfc}(\sqrt{\min\{\beta_1,\beta_2\}\lambda_s})$.

Comparing the RC, SMux-confined and SMux-unconfined solutions, the constellation cardinalities increase from M to $\frac{M(M+1)}{2}$ to M^2 while the error probabilities increase from $\text{erfc}(\sqrt{\beta_{1\oplus 2}\lambda_s})$ to $\text{erfc}(\sqrt{\min\{\beta_1,\beta_2\}\lambda_s})$ to $\text{erfc}(\sqrt{\min\{\beta_1,\beta_2,\beta_{1\ominus 2}\}\lambda_s})$, respectively. The conditional SEP for the Simp-Sub1 decoder can be obtained from the conditional SEP of the ML decoder by (i): replacing $I_{q,p}$ by $\sum_{q=1}^{Q}I_{q,p}$ for $p=1,2$ in (37), (ii): removing the maximization over q in (37) and (iii): replacing λ_s by λ_s/Q to account for the noise accumulation resulting from the EGC implemented by Simp-Sub1. Performing the above modifications results in the error probabilities of $\text{erfc}(\sqrt{\beta^+_{1\oplus 2}\lambda_s/Q})$, $\text{erfc}(\sqrt{\min\{\beta^+_1,\beta^+_2\}\lambda_s/Q})$ and $\text{erfc}(\sqrt{\min\{\beta^+_1,\beta^+_2,\beta^+_{1\ominus 2}\}\lambda_s/Q})$ for RC, SMux-confined and SMux-unconfined, respectively. In these equations, $\beta^+_p=\sum_{q=1}^{Q}I_{q,p}$ for $p=1,2$, $\beta^+_{1\oplus 2}=\sum_{q=1}^{Q}(I_{q,1}+I_{q,2})$ and $\beta^+_{1\ominus 2}=\frac{|\sum_{q=1}^{Q}(I_{q,1}-I_{q,2})|}{\sqrt{\sum_{q=1}^{Q}(I_{q,1}+I_{q,2})}}$ following from (37). Since β_1, β_2, $\beta_{1\oplus 2}$ and $\beta_{1\ominus 2}$ all increase with Q, then the ML decoder profits from the underlying receive diversity with the RC, SMux-confined and SMux-unconfined schemes. On the other hand, only β^+_1, β^+_2 and $\beta^+_{1\oplus 2}$ are increasing functions of Q while $\beta^+_{1\ominus 2}$ might decrease with Q. Consequently, the EGC-based Simp-Sub1 decoder keeps the same receive diversity advantage as the ML decoder only with the RC and SMux-confined schemes.

C. Any Number of Transmit Apertures

Denote by $\mathbf{I}$ the $Q\times P$ channel matrix whose (q,p)-th element is equal to $I_{q,p}$. For the vector $\mathbf{s}=(s_1,\ldots,s_P)$, define the $Q\times M$ matrix $\mathbf{S}$ as $\mathbf{S}=\mathbf{I}\times[\mathbf{e}_{s_1};\cdots;\mathbf{e}_{s_P}]$ where $\mathbf{e}_m$ stands for the m-th row of the $M\times M$ identity matrix. Similarly, let $\mathbf{S}'=\mathbf{I}\times[\mathbf{e}_{s'_1};\cdots;\mathbf{e}_{s'_P}]$ for $\mathbf{s}'=(s'_1,\ldots,s'_P)$. Based on this notation, the $Q\times M$ matrix $\mathbf{A}=\log\left(1+\frac{\lambda_s}{P\lambda_b}\mathbf{S}\right)$ contains the weighing coefficients associated with the vector $\mathbf{s}$ (i.e. the

weights defined in (10)). Similarly, $\mathbf{A}' = \log\left(1+\frac{\lambda_s}{P\lambda_b}\mathbf{S}'\right)$ contains the weighing coefficients of $\mathbf{s}'$. Finally, define the $Q \times M$ matrix $\mathbf{\Lambda}$ as the matrix containing the parameters of the Poisson random variables in (4) with $\mathbf{\Lambda}_{q,m} = \mathrm{E}[R_{q,m}]$ where $\mathbf{X}_{i,j}$ stands for the (i,j)-th element of the matrix $\mathbf{X}$.

Based on the above notations, the pairwise error probability $\Pr(\mathbf{s} \rightarrow \mathbf{s}')$ can be calculated from $\Pr\left[\sum_{q,m}\mathbf{A}'_{q,m}R_{q,m} \geq \sum_{q,m}\mathbf{A}_{q,m}R_{q,m}\right]$ which, from (19), can be bounded as:

$$\Pr(\mathbf{s}\rightarrow\mathbf{s}') \leq \exp\left(\sum_{q=1}^{Q}\sum_{m=1}^{M}\mathbf{\Lambda}_{q,m}\left(e^{\mathbf{\Pi}_{q,m}t}-1\right)\right), \quad (42)$$

where $\mathbf{\Pi} \triangleq \mathbf{A}' - \mathbf{A}$.

An analysis similar to that presented in subsection IV-B shows that the dominant terms among the different pairwise error probabilities in (42) behave asymptotically as either $\mathrm{erfc}(\sqrt{\beta_p\lambda_s})$ or $\mathrm{erfc}(\sqrt{\beta_{p'\ominus p}\lambda_s})$ where extending the notation in (37): $\beta_p = \max_q\{I_{q,p}\}$ for $p = 1,\ldots,P$ and $\beta_{p'\ominus p} = \max_q\left\{\frac{|I_{q,p'}-I_{q,p}|}{\sqrt{I_{q,p'}+I_{q,p}}}\right\}$ for $p' \neq p$. In fact, the first term will appear whenever the components of $\mathbf{s}$ and $\mathbf{s}'$ satisfy $s'_p \neq s_p$ while $s'_{p_0} = s_{p_0}$ for $p_0 \neq p$ (so that the corresponding terms $\mathbf{\Pi}_{q,m}$ in (42) will be equal to zero). On the other hand, probabilities that can be written under the second form will appear when $(s_p, s_{p'}) = (s'_{p'}, s'_p)$ while $s'_{p_0} = s_{p_0}$ for $p_0 \neq p, p'$. On the other hand, the remaining probability terms will behave as $\mathrm{erfc}(\sqrt{k\lambda_s})$ where the constant k takes the general form $k = \max_q\{\min\{\beta_{p_1}, \beta_{p_2}, \cdots, \beta_{p'_n\ominus p_n}, \beta_{p'_{n+1}\ominus p_{n+1}}, \cdots\}\}$ or $k = \max_q\{\min\{\sum_{p\in\mathcal{P}} I_{q,p}, \max_{p\in\mathcal{P}'}\{I_{q,p}\}\}\}$ with $|\mathcal{P}| > 1$ and $|\mathcal{P}'| > 1$. In fact, the first expression of k will follow whenever a pair (s_{p_0}, s'_{p_0}) of initially equal components (to yield β_p for $p_0 \neq p$) now have different components and whenever two initially equal pairs $(s_{p_0}, s_{p'_0}) = (s'_{p_0}, s'_{p'_0})$ (to yield $\beta_{p'\ominus p}$ for $p_0, p'_0 \neq p, p'$) satisfy the new relation $(s_{p_0}, s_{p'_0}) = (s'_{p'_0}, s'_{p_0})$. Moreover, the second value of k will follow whenever the apertures in $\mathcal{P}$ or $\mathcal{P}'$ transmit in the same PPM slot.

Therefore, it follows that the conditional SEP of SMux with no constellation confinement behaves asymptotically as:

$$P_e \mapsto \mathrm{erfc}\left(\sqrt{\min\left\{\min_{p=1,\ldots,P}\{\beta_p\}, \min_{\substack{p,p'=1,\ldots,P\\ p'\neq p}}\{\beta_{p'\ominus p}\}\right\}\lambda_s}\right). \quad (43)$$

The objective of the proposed constellation confinement is to decrease the value of the above probability to $\mathrm{erfc}\left(\sqrt{\min_p\{\beta_p\}\lambda_s}\right)$ by eliminating the information-vector pairs that will make the term $\mathrm{erfc}(\sqrt{\beta_{p'\ominus p}\lambda_s})$ appear in the SEP expression. Based on the above analysis, this term will emerge whenever $\mathbf{s}$ and $\mathbf{s}'$ have $P-2$ elements that are the same while the remaining two elements are flipped. Consequently, the elimination of the terms having the form $\mathrm{erfc}(\sqrt{\beta_{p'\ominus p}\lambda_s})$ can be realized in a simple way by sorting the components of the information vectors in ascending (or descending) order. In this case, the flipping of any two components will result in a vector that falls outside the confined constellation.

As a conclusion, the confined constellation will be designed as follows:

$$\mathbf{S}_{\text{conf}} = \{\mathbf{s} = (s_1,\ldots,s_P) \mid 1 \leq s_1 \leq s_2 \leq \cdots \leq s_P \leq M\}, \quad (44)$$

resulting in the following cardinality:

$$|\mathbf{S}_{\text{conf}}| = \sum_{n=1}^{\min\{M,P\}}\binom{M}{n}\binom{P-1}{n-1}, \quad (45)$$

where n stands for the number of non-empty PPM slots with $\binom{M}{n}$ capturing the number of ways in which these slots can be selected from the total of M PPM slots. Once the nonempty slots are determined, the information vector can be directly deduced in a unique way by finding the number of apertures transmitting in each of these n slots since an aperture with a higher index (compared to another aperture) can not transmit in a prior PPM position following from the symbol sorting in (44). For example, for $P = 5$ and $n = 2$ with 3 (resp. 2) apertures transmitting in the first (resp. second) slot, then apertures 1, 2 and 3 must transmit in the first slot while apertures 4 and 5 must transmit in the second slot resulting in $\mathbf{s} = [s_1, s_1, s_1, s_2, s_2]$ where s_1 and s_2 stand for the indices of the nonempty slots (with $s_1 < s_2$). This justifies the term $\binom{P-1}{n-1}$ appearing in (45) where this number stands for the number of ways of writing P as the sum of n non-zero integers (i.e. number of compositions of P into exactly n parts). Finally, it can be easily proven that $|\mathbf{S}_{\text{conf}}| > M$ implying an increased cardinality compared to RC.

It is worth noting that the confinement targeted the elimination of the terms $\mathrm{erfc}(\sqrt{\beta_{p'\ominus p}\lambda_s})$ rather than the terms $\mathrm{erfc}\left(\sqrt{\beta_p\lambda_s}\right)$ for the following main reason. In the second case, the inclusion of the vector $\mathbf{s} = (s_1,\ldots,s_P)$ in the confined set will incur the exclusion of at least $P(M-1)$ information vectors of the form $(s', s_2, \ldots, s_P)$, $(s_1, s', s_3, \ldots, s_P)$, $\ldots$ $(s_1, \ldots, s_{P-1}, s')$ where s' can take $M-1$ values (that are different from the replaced value in $\mathbf{s}$). On the other hand, in the first case, this will incur the exclusion of $\binom{P}{2}$ vectors of the form $(s_2, s_1, s_3, \ldots, s_P)$, $(s_3, s_2, s_1, \ldots, s_P)$, $(s_4, s_2, s_3, s_1, s_5, \ldots, s_P)$, $\ldots$. Consequently, the first option that we adopt has a higher potential of increasing the cardinality of the confined constellation. This observation is more clearly reflected in the case of $P = 2$ where, from (36) and (39)-(40), each one of the probabilities $\mathrm{erfc}\left(\sqrt{\beta_1\lambda_s}\right)$ and $\mathrm{erfc}\left(\sqrt{\beta_2\lambda_s}\right)$ appears $M(M-1)(2M-3)$ times while the probability $\mathrm{erfc}(\sqrt{\beta_{1\ominus 2}\lambda_s})$ appears $M(M-1)$ times.

Finally, from (45), since evaluating the ML metric for an information vector extending over n non-empty slots (i.e. with n unique elements) is equal to n, then the number of multiplications required by the ML algorithm when associated with the confined constellation is:

$$N_{\text{mul}}^{(\text{ML})} = Q\sum_{n=1}^{\min\{M,P\}} n \times \left[\binom{M}{n}\binom{P-1}{n-1}\right], \quad (46)$$

which can also be obtained by replacing $\frac{P!}{l_1!l_2!\cdots l_n!}$ with 1 in (11) (the corresponding summation in (11) will simplify to the number of compositions of P into n summands given by $\binom{P-1}{n-1}$).

Regarding the Simp-ML decoder, algorithm-1 can be readily applied with the sole modification that the partitioning $\{1,\ldots,P\} = \mathcal{P}_1 \cup \cdots \cup \mathcal{P}_n$ must take into consideration the fact that the elements of $\mathbf{s}$ are sorted in ascending order: $\forall\ (x,y) \in \mathcal{P}_i \times \mathcal{P}_j$, $x < y$ for $i < j$. The number of required multiplications with the confined constellation is:

$$N_{\text{mul}}^{\text{(Simp-ML)}} = \sum_{n=1}^{\min\{M,P\}} n \times \binom{P-1}{n-1}, \tag{47}$$

following from (45) since, for a particular value of n, the knowledge of the composition of P into n summands yields one candidate solution to be tested (with n required multiplications). Moreover, the n slots with maximum EGC counts are selected in a unique way.

While an explicit expression for the diversity order is difficult to obtain in the case of nonzero boresight [20], the diversity order is given by $\zeta \triangleq \min\{\xi^2, \varphi_2\}$ for a single-aperture FSO link with zero boresight. In this scenario, it can be easily proven that the error probabilities in the cases of RC and SMux-confined with the ML decoder scale asymptotically as $\lambda_s^{-PQ\zeta}$ and $\lambda_s^{-Q\zeta}$, respectively. However, for SMux-unconfined, it is very challenging to derive the diversity order since it is hard to study the random quantity $\beta_{p'\ominus p} = \max_q \left\{ \frac{|I_{q,p'} - I_{q,p}|}{\sqrt{I_{q,p'} + I_{q,p}}} \right\}$.

V. Numerical Results

In this section, we present some numerical results for the sake of comparing the different decoders and showing the impact of constellation confinement. The presented figures show the variation of the SEP as a function of the signal energy E_s per information bit for different values of the background-noise power P_b. The former quantity is equal to $\frac{E_s}{P\log_2(M)}$ for SMux with unconfined constellation and $\frac{E_s}{|\mathbf{S}_{\text{conf}}|}$ for SMux with confined constellation where $|\mathbf{S}_{\text{conf}}|$ is given in (45). The quantum efficiency, attenuation constant and refractive index structure constant are set to $\eta = 0.5$, $\sigma = 0.44$ dB/km and $C_n^2 = 1.7\times10^{-14}$ m$^{-2/3}$, respectively. We also set $\sigma_s/a = 3$ while the impact of pointing errors will be captured by varying the values of ω_z/a and s/a.

It is worth highlighting that the comparison between the different transmission schemes will be carried out for the same value of M for the following reason. While increasing M increases the number of bits per symbol, it increases the symbol duration T_s as well. In fact, for a given transmitter/receiver bandwidth, the symbol duration is fixed to $T_s = M\delta$ where the pulse duration δ (that is equal to the duration of one PPM slot) fixes the optical bandwidth. Consequently, the bit rates of the SMux and RC schemes are $\frac{P\log_2(M)}{M\delta}$ and $\frac{\log_2(M)}{M\delta}$, respectively, where both of these quantities decrease with M as in the case of single-aperture systems. In this context, the bit rate of the M-PPM SMux scheme is P times larger than that of the M-PPM RC scheme. Finally, it is worth highlighting that while M-PPM SMux systems and M^P-PPM RC systems transmit the same number of bits per symbol ($P\log_2(M)$ bits), the symbol duration of the corresponding RC scheme is M^{P-1} times larger than that of the SMux scheme thus incurring the reduction of the bit rate by a factor of M^{P-1}. This highlights on the interest of the considered SMux scheme (whether with or without the constellation confinement) where the bit rate can be increased without altering the number of PPM positions.

Fig. 2 shows the performance of 2×1 and 2×2 systems with 4-PPM for $d = 3$ km, $\omega_z/a = 25$ and $s/a = 3$. The level of background noise is determined from $P_bT_s/M = -185$ dBJ and the unconfined constellation is considered. For the 2×1 system, results show that the Simp-ML decoder is achieving exactly the same performance level as the ML decoder thus stressing on the optimality of the former simplified decoder for MISO systems. Among the suboptimal decoders, decoder Simp-Sub1 manifests the worst performance that still outperforms 2×1 systems by about 2.5 dB while decoder Simp-Sub2 is capable of achieving advantageous performance levels that are very close to those achieved by the optimal ML decoder. In particular, for $M_r = 3$, the SEPs of the Simp-Sub2 and ML decoders are practically the same for all values of E_s while, for $M_r = 2$, the Simp-Sub2 decoder results in limited performance losses for the values of $\frac{E_s}{P\log_2(M)}$ exceeding -163 dBJ. In terms of complexity, the Simp-Sub1 decoder is the most appealing with 3 multiplications per information vector while the implementation of the ML decoder is the most demanding with 56 multiplications per information vector. For the Simp-Sub2 decoder, the number of multiplications increases from 12 to 30 as M_r increases from 2 to 3. The proposed optimal and suboptimal decoders can also be applied in the scenario where the CSI is available at the transmitter resulting in the implementation of a power allocation (PA) strategy. In this case, the power normalisation factor $1/P$ in the weighing coefficients in (10) needs to be replaced by the parameter ϕ_p that stands for the fraction of the power allocated to the p-th transmit aperture with $\sum_{p=1}^{P}\phi_p = 1$. Evidently, this extension does not entail any additional complexity given that it only affects the block terms. In such scenarios, the backup RF link might be useful for providing feedback from the receiver to the transmitter. While the resulting performance gains are highly dependent on the adopted PA strategy, developing the optimal PA scheme falls beyond the scope of this work. Nevertheless, in order to provide some insights on the realized performance gains, we show the performance with the simple PA strategy $\phi_p = \frac{\sum_{q=1}^{Q} I_{q,p}}{\sum_{p'=1}^{P}\sum_{q=1}^{Q} I_{q,p'}}$ that allocates more power to the stronger links in a linear fashion. Results in Fig. 2 show that a performance gain in the order of 4 dB can be achieved by this simple PA with the Simp-Sub1 decoder thus motivating the importance of examining the PA strategies in future works.

The simulation setup of Fig. 2 is reproduced in Fig. 3 for 4×1 and 4×4 systems with 8-PPM. The numbers of multiplications required by the Simp-ML, Simp-Sub1, Simp-Sub2 with $M_r = 4$, Simp-Sub2 with $M_r = 5$ and ML decoders are 37, 37, 2800, 7380 and 54240, respectively. This highlights on the advantageous simplicity of the MIMO Simp-Sub1 decoder that, evidently, comes at the expense of performance losses. This renders this type of decoding suitable for low-complexity MIMO systems. On the other hand, for systems possessing higher processing capabilities, the Simp-Sub2 decoder with $M_r = 4$ constitutes an appealing option

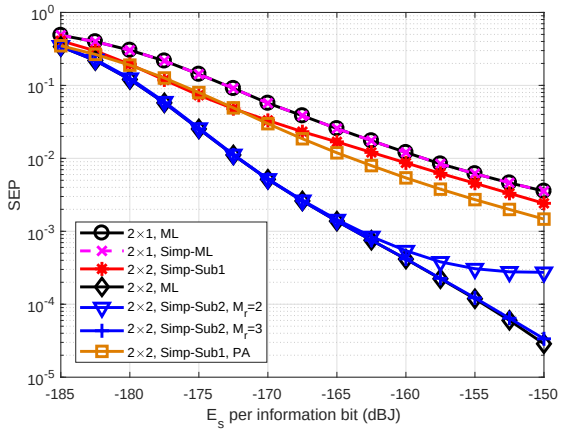

Fig. 2. Performance with 2 transmit apertures and 4-PPM for $d = 3$ km, $\omega_z/a = 25$, $s/a = 3$ and $P_bT_s/M = -185$ dBJ.

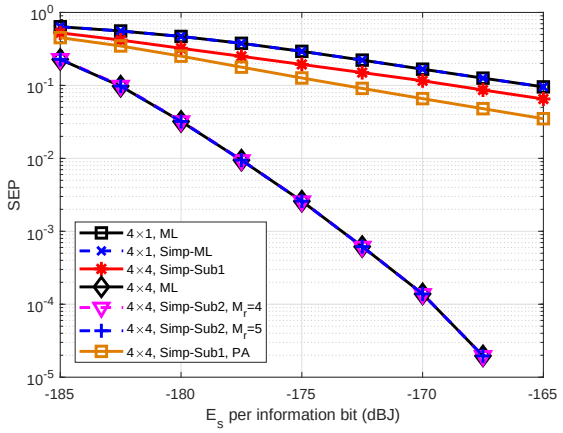

Fig. 3. Performance with 4 transmit apertures and 8-PPM for $d = 3$ km, $\omega_z/a = 25$, $s/a = 3$ and $P_bT_s/M = -185$ dBJ.

where the achieved performance levels are extremely close to those achieved by the optimal ML decoder while performing 19.37 times less multiplications. Compared to the 2×2 systems considered in Fig. 2, the performance gap between the Simp-Sub1 decoder and the ML decoder increases in the case of 4×4 systems. This is justified by the sharp reduction in the number of multiplications required by the Simp-Sub1 decoder when (P, Q, M) increases from $(2, 2, 4)$ to $(4, 4, 8)$. In fact, while the Simp-Sub1 decoder requires 18.6 times less multiplications compared to the ML decoder with 2×2 systems, this ratio increases to 1466 with 4×4 systems. Fig. 3 shows that implementing the PA strategy results in a gain of 4.2 dB with the Simp-Sub1 decoder.

Fig. 4 compares the performance of the confined and unconfined constellations for 3×3 MIMO systems with 8-PPM for $P_bT_s/M = -180$ dBJ with $d = 4$ km in the absence of pointing errors ($\omega_z/a \to \infty$ and $s/a = 0$). Results in Fig. 4 highlight on the enhanced performance levels that can be achieved by confining the transmitted constellation according to (44). Evidently, this SEP enhancement is achieved through compromising the data rate where the cardinalities of the confined and unconfined constellations are equal to 120 and 512, respectively. In this case, the performance gains are higher for larger values of E_s where the constellation confinement results in a performance gain of 5.7 dB at a SEP of 10^{-4} with the ML decoding. Fig. 4 also highlights the important observation that the gap between the Simp-Sub1 and ML decoders is significantly reduced when confining the transmitted constellation. While the performance gap is particularly pronounced with the unconfined constellation with values exceeding 15 dB at 10^{-2}, the Simp-Sub1 decoder is capable of achieving appealing performance levels that are comparable to those achieved by the ML decoder with a much reduced complexity in the case where the constellation confinement is applied. In fact, the Simp-Sub1 decoder performs only 1.7 dB worse than the optimal ML decoder at a SEP of 10^{-4} when the confinement is applied. The confined constellation with the Simp-Sub1 decoder even outperforms the unconfined constellation with the ML decoder where the performance gains are in the order of 4 dB. The results also highlight on the accuracy of the derived approximate SEP expressions in predicting the performance for average-to-large values of E_s. In fact, the approximate SEP expressions vary with E_s in a manner that is comparable with the variations of the exact SEPs. Fig. 4 also shows the performance of RC that achieves the best SEP performance [3]. However, this performance superiority comes at the expense of a sharp drop in the data rate where the transmitted constellation has a cardinality of 8 (3 bits) while the cardinalities of SMux with the confined and unconfined constellations are 120 (6.9 bits) and 512 (9 bits), respectively. Finally, applying SMux with the proposed constellation confinement constitutes a worthwhile compromise between the two extremes of conventional SMux (highest rate and worst SEP) and RC (lowest rate and best SEP).

Fig. 5 shows the performance of 5×5 MIMO systems with 4-PPM, $d = 3$ km, $P_bT_s/M = -185$ dBJ, $\omega_z/a = 12$ and $s/a = 0$. The obtained results re-emphasize on the findings drawn from Fig. 4 whether in terms of the enhanced performance levels that result from confining the transmitted constellation or in terms of the improved decoding potential of the Simp-Sub1 decoder when associated with the confined constellation. It is worth noting that the SMux solutions transmit at the rates of 5.8 bits per channel use (pcu) and 10 bits pcu with and without applying the constellation confinement, respectively. These rates are much higher than those achieved by the SISO systems (and MIMO RC systems) that transmit at the rate of 2 bits pcu. As in Fig. 4, the variations of the SEP curves are accurately predicted by the presented approximate expressions.

VI. Conclusion

We investigated MIMO-FSO IM/DD communication systems with photon-counting receivers. Our analysis revolved around SMux and a proposed novel scheme with constellation confinement that can be perceived as a compromise between SMux (best rate) and RC (best performance). We proposed an optimal decoder as well as two suboptimal decoders for

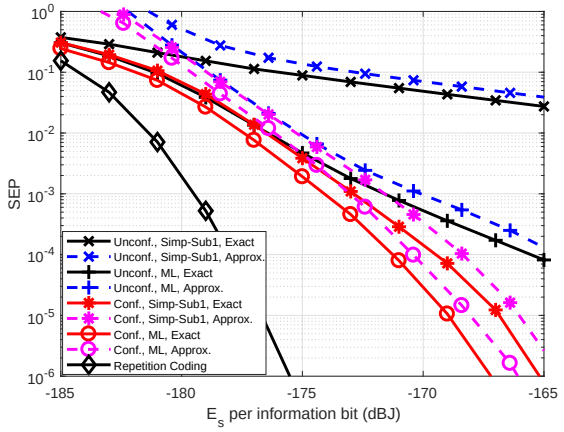

Fig. 4. Performance of 3×3 MIMO systems with 8-PPM for a link distance of 4 km with $P_b T_s/M = -180$ dBJ, $\omega_z/a \to \infty$ and $s/a = 0$. The solid and dashed lines correspond to the exact and approximate SEPs, respectively.

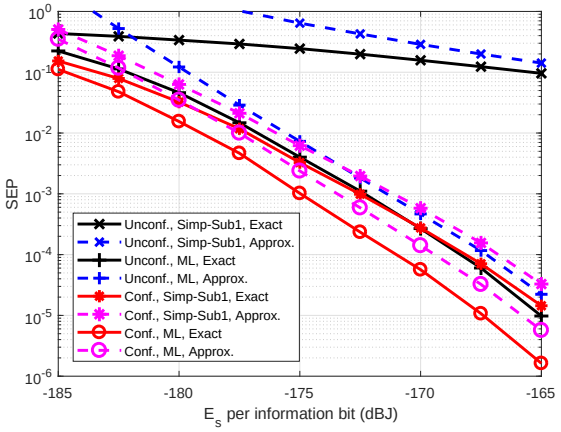

Fig. 5. Performance of 5×5 MIMO systems with 4-PPM for a link distance of 3 km with $P_b T_s/M = -185$ dBJ, $\omega_z/a = 12$ and $s/a = 0$. The solid and dashed lines correspond to the exact and approximate SEPs, respectively.

accomplishing the challenging objective of separating the interfering channels. While the Simp-Sub1 decoder is the simplest, it suffers from the most pronounced performance degradations with a dramatic reduction in such degradations when this decoder is associated with the confined constellation. The Simp-Sub2 decoder has higher complexity but exhibits near-optimal performance.

REFERENCES

[1] M. K. Simon and V. A. Vilnrotter, "Alamouti-type space-time coding for free-space optical communication with direct detection," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 35–39, Jan. 2005.

[2] C. Abou-Rjeily and W. Fawaz, "Space-time codes for MIMO ultra-wideband communications and MIMO free-space optical communications with PPM," *IEEE J. Select. Areas Commun.*, vol. 26, no. 6, pp. 938–947, Aug. 2008.

[3] M. Safari and M. Uysal, "Do we really need OSTBCs for free-space optical communication with direct detection?" *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4445 – 4448, Nov. 2008.

[4] C. Abou-Rjeily, "Performance analysis of FSO communications with diversity methods: Add more relays or more apertures?" *IEEE J. Select. Areas Commun.*, vol. 33, no. 9, pp. 1890 – 1902, Sep. 2015.

[5] N. Letzepis and A. G. I. Fabregas, "Outage probability of the Gaussian MIMO free-space optical channel with PPM," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3682–3690, Dec. 2009.

[6] J. Zhang, L. Dai, Y. Han, Y. Zhang, and Z. Wang, "On the ergodic capacity of MIMO free-space optical systems over turbulence channels," *IEEE J. Select. Areas Commun.*, vol. 33, no. 9, pp. 1925–1934, Sep. 2015.

[7] E. Bayaki, R. Schober, and R. K. Mallik, "Performance analysis of MIMO free-space optical systems in gamma-gamma fading," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3415–3424, Nov 2009.

[8] S. G. Wilson, M. Brandt-Pearce, Q. Cao, and J. H. Leveque, "Free-space optical MIMO transmission with Q-ary PPM," *IEEE Trans. Commun.*, vol. 53, no. 8, pp. 1402–1412, Aug. 2005.

[9] M. L. Riediger, R. Schober, and L. Lampe, "Multiple-symbol detection for photon-counting MIMO free-space optical communications," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5369–5379, Dec. 2008.

[10] H. Nouri, F. Touati, and M. Uysal, "Diversity-multiplexing tradeoff for log-normal fading channels," *IEEE Trans. Commun.*, vol. 64, no. 7, pp. 3119–3129, July 2016.

[11] M. T. Dabiri, M. J. Saber, and S. M. S. Sadough, "On the performance of multiplexing FSO MIMO links in log-normal fading with pointing errors," *IEEE J. Opt. Commun. Netw.*, vol. 9, no. 11, pp. 974–983, Nov. 2017.

[12] G. Yang, M.-A. Khalighi, T. Virieux, S. Bourennane, and Z. Ghassemlooy, "Contrasting space-time schemes for MIMO FSO systems with non-coherent modulation," in *IEEE Int. Workshop on Optical Wireless Commun.*, Oct 2012, pp. 1–3.

[13] T. Özbilgin and M. Koca, "Optical spatial modulation over atmospheric turbulence channels," *J. Lightwave Technol.*, vol. 33, no. 11, pp. 2313–2323, June 2015.

[14] A. Jaiswal, M. R. Bhatnagar, and V. K. Jain, "Performance evaluation of space shift keying in free-space optical communication," *IEEE J. Opt. Commun. Netw.*, vol. 9, no. 2, pp. 149–160, Feb. 2017.

[15] T. Song and P.-Y. Kam, "Robust data detection for the photon-counting free-space optical system with implicit CSI acquisition and background radiation compensation," *J. Lightwave Technol.*, vol. 34, no. 4, pp. 1120–1132, Feb. 2016.

[16] N. D. Chatzidiamantis, G. K. Karagiannidis, and M. Uysal, "Generalized maximum-likelihood sequence detection for photon-counting free space optical systems," *IEEE Trans. Commun.*, vol. 58, no. 12, pp. 3381–3385, Dec. 2010.

[17] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inform. Theory*, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.

[18] 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 *Int. Symp. on Signal, Systems and Electronics*, Sep. 1998, pp. 295–300.

[19] D. G. Brennan, "Linear diversity combining techniques," in *Proc. of the IEEE*, vol. 91, no. 2, Feb. 2003, pp. 331–356.

[20] F. Yang, J. Cheng, and T. A. Tsiftsis, "Free-space optical communication with nonzero boresight pointing errors," *IEEE Trans. Commun.*, vol. 62, no. 2, pp. 713–725, Feb. 2014.

[21] N. Letzepis, K. D. Nguyen, A. G. i Fàbregas, and W. G. Cowley, "Outage analysis of the hybrid free-space optical and radio-frequency channel," *IEEE J. Select. Areas Commun.*, vol. 27, no. 9, pp. 1709–1719, Dec. 2009.