Measurement-based VLC channel characterization for I2V communications in a real urban scenario

S. Caputo, L. Mucchi, F. S. Cataliotti, M. Seminara, T. Nawaz and J. Catani

Abstract—Visible light communication (VLC) is nowadays envisaged as a promising technology to enable new classes of services in intelligent transportation systems ranging, e.g., from assisted driving to autonomous vehicles. The assessment of the performance of VLC for automotive applications requires as a basic step a model of the transmission pattern and propagation of the VLC signal when real traffic-lights and road scenarios are involved. In this paper an experimental measurement campaign has been carried out by using a regular traffic-light as source (red light) and a photoreceiver positioned, statically, at different distances and heights along the road. A linear regression technique is used to come up with different propagation models. The proposed models have been compared, in terms of accuracy and complexity, to the conventional Lambertian model to describe the VLC channel in a realistic outdoor scenario. The proposed models provide a significant higher accuracy with comparable complexity.

Index Terms— Visible light communications, experimental measurements, channel modeling, vehicular communications.

I. INTRODUCTION

The ongoing substitution of conventional light sources with light emitting diodes (LEDs) has recently fuelled scientific and industrial activity in visible light communication (VLC) technology [1], [2], [3]. Such an interest stems from the possibility offered by LEDs for fast modulation of light. In the context of automotive and vehicular networks [4], [5], where low latency and reliability are of crucial importance, VLC can offer significant advantages [6], [7]. In addition, VLC is an inherently energy-efficient interconnection since the energy is, in any case, already needed to illuminate the road or for road signaling (e.g., traffic lights) [8]. VLC could therefore be significant for vehicle-to-vehicle (V2V) or infrastructure-to-vehicle (I2V) communications, both issues of decisive importance for road safety, in particular in the context of assisted- or unmanned-driving [9], [10]. A reliable assessment of suitability of VLC technology for V2V or I2V communications [11], [12] requires an accurate characterization of the VLC channel. In literature, the approaches to this problem has been theoretical or empirical. The theoretical approach aims at a mathematical reduction of the problem with approximations, such as Lambertian emission and reflection pattern [8]. Ray tracing is another method of investigation, allowing for a software-based physical simulation [13]. This kind of approaches is usually characterized by a simplified scenario, aimed at reducing the computational complexity. On the other hand, the empirical approach typically involves a measurement campaign followed by the extraction of a mathematical model of the channel as in [14]. However, whilst the optical wireless channel has been characterized in literature for infrared (IR) communication [15], and recently a comparison between IR and VLC appeared [16], in order to enable the use of VLC technology for safety-critical / smart driving applications [17], [18], it is mandatory to test and analyze the VLC-based V2V and I2V communication channel in a realistic outdoor scenario, still lacking, due to intrinsic difficulties mainly represented by the influence of non-ideal external factors, such as ambient light and irregularities of the emission pattern of the LED-based traffic lights or headlights [21], [22]. In this paper we present an outdoor measurement campaign, aimed at the extensive characterization of the VLC channel in a real urban scenario, carried out in the city of Prato (Italy) in collaboration with ILES srl, a private company which develops urban signaling systems. The measurements have been taken in a real urban road, with regulatory traffic light emitting a VLC signal to the receiver located along the same road. The measurement grid has been replicated setting the receiver at three different heights corresponding, respectively, to car headlights, dashboard and internal mirror. Measurements have been performed also in the presence of stray light sources such as sunlight or other cars’ headlights. The data has been used to extract a mathematical model of the VLC signal propagation. In particular, the optical source plus propagation transfer function is modeled. We will show that the Lambertian model is far from being accurate when real traffic-light lamps are considered for VLC transmission. A linear regression technique is used to come up with a more accurate model of the transmission-propagation pattern. The results of the measurement campaign highlights attainable...
and $A$ is the amplitude, $T$ is the duration of the single rectangular pulse, $s_k$ is the transmitted symbol and $N_b$ is the number of transmitted symbol sequences. The symbol sequence is $s = \{1, -1, 1, -1, 1, -1\}$, chosen in order to encompass all of the possible logic transitions appearing in Manchester encoding, where no more than two consecutive symbols can be of the same sign [23] (see Fig. 2).

The received signal, in each point on the grid, is the convolution of the transmitted signal with four impulse responses

$$s_{RX}(t) = s_{TX}(t) * h_{el}^{TX}(t) * h_{op}^{TX}(t) * h_{op}^{RX}(t) * h_{el}^{RX}(t) + n(t) \quad (3)$$

where
- $n(t)$ is the (white Gaussian) thermal noise at the receiver,
- $h_{el}^{TX}(t)$ takes into account the electronic at the transmitter,
- $h_{op}^{TX}(t)$ the effects of the optics at the transmitter,
- $h_{op}^{RX}(t)$ the light propagation,
- $h_{el}^{RX}(t)$ the optics at the receiver, and
- $h_{el}^{RX}(t)$ the electronic conversion at the receiver.

In the frequency domain, Eq. (3) becomes

$$S_{RX}(f) = S_{TX}(f) H_{el}^{TX}(f) H_{op}^{TX}(f) H_{op}^{RX}(f) H_{el}^{RX}(f) + N(f)$$

$$H(f) \quad (4)$$

It is important to note that $h_{op}(t)$ does not actually depend on time as well, since during the measurements both TX and RX were static, and no variations in the environment occurred. In our study, we aim at the characterization of ultimate performance of our system in a real scenario. For this motivation, for each point of the grid, we chose to align the optical axis of the receiver towards the lamp, adjusting its horizontal and vertical alignment angles. The benefits of
such setting are mainly two. First, with our set of focal length and detector size, the image of the lamp falls entirely into the active area of the detector for all points on the grid, so that the reconstructed intensity pattern coincides with the lamp pattern. Secondly, this configuration avoids any interference effect given by reflections of the lamp’s signal from tarmac, as no reflected beam falls into the field of view (FOV) of our receiver. A demonstration of this is given in Appendix B.

For the above considerations, under our configuration we can assume \( h^\text{op}_{\text{RX}} \) to be nearly constant into the whole measurement grid. The impact of different collecting optics and angles on the overall system performances, as well as possible effects of reflections on the optical channel quality will be addressed in future work.

In this paper we focus on the modelling of \( h^\text{op}_{\text{RX}} * h_p(t) \). It is common in literature to model this as Lambertian, but we will demonstrate that the Lambertian model does not provide for an accurate description when the source is a real LED-based traffic-light, equipped with regulatory shaping optical lens, in an urban scenario. We then propose a linear regression method to come up with a mathematical model that better fits the propagation of VLC signals from traffic-light to vehicles.

III. Measurement Campaign in Urban Scenario

The on-field measurements campaign has been carried out in the city of Prato (Italy) in collaboration with ILES srl, a company producing and installing road signaling infrastructures. The company has installed one traffic light in a real urban scenario consisting of a two-lane road with buildings on both sides (see Fig. 3). The traffic light has been positioned on the right side of the rightmost lane, with 0.75 m indentation, at the height of 2.83 m, in accordance to Italian regulations UNI11248 [24] and UNI13201-2 [25]. The measurements have been carried out during one entire day (sunny conditions). The red light of the traffic light has been modulated with the information signal (see Fig. 3). In particular, the red LED lamp has been controlled by our driver circuit (see Sec. III-A1) which, besides controlling and supplying the DC nominal operating current, allows for the insertion of the bit sequence via an external function generator (see Sec. III-A1). The LED driver is the only component that has been replaced in the commercial traffic light.

The photodetector has been positioned on a grid on points in front of the emitting traffic light (Fig. 3). The grid has been centered at the position of the traffic light, with the \( x \)-axis along the road length, the \( y \)-axis along the road width and a vertical \( z \)-axis. The grid points along the \( x \)-axis ranges from 3 to 30 m, with more density in the first 10 m: [3 m, 4 m, 5 m, 7 m, 10 m, 15 m, 20 m, 25 m, 30 m]. Along the \( y \)-axis we included an offset of 0.75 m since the traffic light has this indentation on the sidewalk. The step between measurement points is fixed to 1 m: [0.75 m, 1.75 m, 2.75 m, 3.75 m, 4.75 m, 5.75 m]. This two-dimensional grid is repeated for three different heights (\( z \)-axis) to simulate where the receiver could be placed on the vehicle: car headlights (0.75 m), dashboard (1 m) or rear-view inside mirror (1.35 m). The photodetector is connected to an oscilloscope to display and record the VLC signal coming from the traffic light. The oscilloscope is triggered by a sync output of the function generator whose period, highlighted in Fig 2, corresponds to a full Manchester modulation cycle. The record length is 200 kpts at a sampling rate of 0.5 GHz, encompassing 10 full modulation periods. For each measurement we acquired both a single shot and the average of 4 traces.

A. TX-RX Hardware Design

The VLC TX and RX stages have been designed in order to feature analog bandwidths above 150 kHz and to reject low-frequency high-intensity components coming from sunlight or traffic illumination without saturating the first amplification stage.

1) Transmitter: The schematic block diagram of TX hardware is reported in Fig. 4. TX hardware is composed by the LED light source, and its current driver. Our implementation allows to generate the required current (\( \approx 0.7 \) A) for the traffic light LEDs to provide the nominal luminous flux, as well as to insert a current modulation proportional to an external signal, which in turn allows for insertion of data streams into the optical carrier using any kind of protocol based on light intensity modulation. The modulator section (cyan shaded area of Fig. 4) is placed after a P-I regulation stage (purple shaded area), stabilizing the supply current, which is sensed by a
been chosen as the highest still preserving a bandwidth larger than any coloured filter at receiver. The photoreceiver gain has been limited to about 3x to avoid the need for RX gains without the risk of saturating the amplifier by stray ambient lights. Such configuration allowed for high signal to pass through the first stage of the transimpedance amplification stage. The time constant of the actuation chain in the open-loop response, affect the maximum achievable transmission bandwidth. We have limited the relative modulation amplitude to 30% of the average DC value of 700 mA, in order to avoid a possible overburden of the LED sources due to excessive currents in the positive periods of the modulation pattern of the nominal DC value. A function generator (Tektronix AFG1022) has been used to provide the modulation circuit with the modulation waveform (see Fig. 2). The transmitted waveform has been chosen to match a OOK Manchester encoding, according to PHY I of standard IEEE 802.15.7 for Outdoor VLC [26]. In order to embed all possible bit configurations of the Manchester encoding, the data package is constructed by attaching two square-wave blocks with frequencies of 50 kHz and 100 kHz, respectively, with a global periodicity of 40 $\mu$s, corresponding to the packet duration. The Manchester encoding grants a constant average current (Fig. 2), leading to a constant illumination intensity emitted by the traffic light lamp.

The (red) LED emitter (Lux Potentia OJ200-R07, 1A 12V) is composed of a series of 3 high-power LEDs. In our scheme, the original power supply has been bypassed by our MOSFET-based current driver/modulator, while preserving the original case of the LED series, so that the global features of the traffic light illumination pattern are unaltered. We also checked that different colour lamps do not have a significantly different frequency response to modulation, as the LED substrates have similar parasitic capacitance, so the bandwidth of the system is the same and the temporal dependence of Eq. 1 is unaltered. A red-coloured Fresnel lens and a precision resistor, from direct current (DC) to $\sim 1.5$ kHz, given by a proper adjustment of the servo resistor-capacitor (RC) constant (see Fig. 4).

In such configuration, any modulation above the PI servo cut frequency $f_{TX} = \frac{1}{2\pi RC} \approx 1.5$ kHz will not be compensated by the $P-I$ loop, and will be added as a current modulation through the MOSFET transistor. The open loop bandwidth of the op-amp chain can virtually exceed several MHz, whereas, on the other hand, the large parasitic capacitance of the large-area LED module embedded in the traffic light lamp, as well as the presence of non linearity of the actuation chain in the open-loop response, affect the maximum achievable transmission bandwidth. We have limited the relative modulation amplitude to 30% of the average DC value of 700 mA, in order to avoid a possible overburden of the LED sources due to excessive currents in the positive periods of the modulation pattern of the nominal DC value. A function generator (Tektronix AFG1022) has been used to provide the modulation circuit with the modulation waveform (see Fig. 2). The transmitted waveform has been chosen to match a OOK Manchester encoding, according to PHY I of standard IEEE 802.15.7 for Outdoor VLC [26]. In order to embed all possible bit configurations of the Manchester encoding, the data package is constructed by attaching two square-wave blocks with frequencies of 50 kHz and 100 kHz, respectively, with a global periodicity of 40 $\mu$s, corresponding to the packet duration. The Manchester encoding grants a constant average signal (Fig. 2), leading to a constant illumination intensity emitted by the traffic light lamp.

The (red) LED emitter (Lux Potentia OJ200-R07, 1A 12V) is composed of a series of 3 high-power LEDs. In our scheme, the original power supply has been bypassed by our MOSFET-based current driver/modulator, while preserving the original case of the LED series, so that the global features of the traffic light illumination pattern are unaltered. We also checked that different colour lamps do not have a significantly different frequency response to modulation, as the LED substrates have similar parasitic capacitance, so the bandwidth of the system is the same and the temporal dependence of Eq. 1 is unaltered. A red-coloured Fresnel lens and a precision resistor, from direct current (DC) to $\sim 1.5$ kHz, given by a proper adjustment of the servo resistor-capacitor (RC) constant (see Fig. 4).

In such configuration, any modulation above the PI servo cut frequency $f_{TX} = \frac{1}{2\pi RC} \approx 1.5$ kHz will not be compensated by the $P-I$ loop, and will be added as a current modulation through the MOSFET transistor. The open loop bandwidth of the op-amp chain can virtually exceed several MHz, whereas, on the other hand, the large parasitic capacitance of the large-area LED module embedded in the traffic light lamp, as well as the presence of non linearity of the actuation chain in the open-loop response, affect the maximum achievable transmission bandwidth. We have limited the relative modulation amplitude to 30% of the average DC value of 700 mA, in order to avoid a possible overburden of the LED sources due to excessive currents in the positive periods of the modulation pattern of the nominal DC value. A function generator (Tektronix AFG1022) has been used to provide the modulation circuit with the modulation waveform (see Fig. 2). The transmitted waveform has been chosen to match a OOK Manchester encoding, according to PHY I of standard IEEE 802.15.7 for Outdoor VLC [26]. In order to embed all possible bit configurations of the Manchester encoding, the data package is constructed by attaching two square-wave blocks with frequencies of 50 kHz and 100 kHz, respectively, with a global periodicity of 40 $\mu$s, corresponding to the packet duration. The Manchester encoding grants a constant average signal (Fig. 2), leading to a constant illumination intensity emitted by the traffic light lamp.

2) Receiver: The receiver hardware is implemented by modifying a Thorlabs PDA36-A active photodiode with a physical AC decoupling of the photodiode chip from the first transimpedance amplification stage. The time constant $f_{RX} \approx 5$ kHz of such decoupling is realized through a multielement RLC network, with parameters chosen in such a way to filter out all of the unwanted low-frequency light variations (headlights, sunlight ecc.), still allowing for the modulation signal to pass through the first stage of the receiving electronics. Such configuration allowed for high RX gains without the risk of saturating the amplifier by a large light background, and practically avoids the need for any coloured filter at receiver. The photoreceiver gain has been chosen as the highest still preserving a bandwidth $> 150$ kHz. The concentrator used in this RX setup (rightmost panel) is the shortest focal, low-cost, plastic lens we had available, i.e. a 25 mm diameter, $F = 30$ mm focal length uncoated aspheric singlet. The choice of short-focal, aberration-corrected optical concentrators is crucial in order to keep a good trade-off between reasonably high acceptance angle and global optical gain when the solution to increase the input optics diameter beyond 1” is unaffordable. The data collection is accomplished by recording both the transmitted and received waveforms through a 70 MHz, 1Gs/s digital oscilloscope (Tektronix TBS2074).

IV. PROPAGATION MODEL

The experimental (raw) data, recorded in each point of the 3D grid, has been processed to reduce the high-frequency noise effects and estimate the received amplitude in the frequency domain. The processing steps on the raw data can be summarized as follows:

- Data Binning;
- Fast Fourier Transform, normalized to a reference value chosen as maximum signal obtained in the “reference point” (see Fig. 3);
- Reconstruction of the amplitude of the received signal in each position $(x, y, z)$ of the grid.

Each step is detailed in the following.

We have observed that the frequency components of the received signal over 500 kHz show a negligible amplitude, thus we applied a data binning, equivalent to a downsampling procedure, to reduce the frequency observation interval hence filtering out unnecessary high-frequency components, laying above our electronic system bandwidth. The binning procedure takes a cluster of consecutive (time) samples and replaces the cluster with the average value of the samples.

The binning value can be defined as

$$N_{\text{bin}} = \frac{f_c}{2f_{\text{max}}} \quad (5)$$
where $f_c$ is the initial sampling frequency of the received signal and $f_{\text{max}}$ is the highest frequency of the desired observation window. In our experiment, $N_{\text{bin}} = 250$, since the sampling frequency of the oscilloscope is 250 MHz.

The ratio between the transfer function\(^1\) (TF) of the whole system calculated in each single point of the grid and the TF calculated in the reference position can be used to define the propagation model of the VLC signal from the traffic light to the vehicle

$$\Delta H_i = \frac{H_i(f)}{H_{\text{ref}}(f)}$$

with

$$H_i(f) = \frac{S_{RX}(f)}{S_{TX}(f)}; \quad H_{\text{ref}}(f) = \frac{S_{RX_{\text{ref}}}(f)}{S_{TX_{\text{ref}}}(f)}$$

where $S_{TX}(f)$ and $S_{RX}(f)$ is the Fast Fourier Transform (FFT) of the transmitted and received signal, respectively, at location $i$, $H_i(f)$ represents the TF between the traffic light and the $i$th point on the grid and $H_{\text{ref}}(f)$ stands for the TF between the traffic light and the reference point. All measurements have been triggered on the same signal, and thus the FFT of transmitted signal will be the same for each point $S_{TX}(f) = S_{TX_{\text{ref}}}(f)$. Thus, Eq. (6) can be then rewritten as

$$\Delta H_i = \frac{S_{RX}(f)}{S_{RX_{\text{ref}}}(f)}$$

Eq. (8) we can deduce that $\Delta H_i$ depends only on the spectrum of the received signal at the $i$-th location and at the reference point\(^2\). In addition, we can safely assume that $\Delta H_i$ does not depend on the carrier and modulation frequencies, as the large spectral width of the optical carrier emitted by the LED source makes the contribution of absorption lines of air absolutely irrelevant for any channel loss, whilst the large difference between modulation and carrier frequencies avoids any frequency-dependent interference effect in line-of-sight tests.

Fig. 5 shows the map of the amplitudes $\Delta H_i$ measured in each point of the grid at different height [0.75, 1, 1.35] m. Amplitudes are reported in logarithmic (dB) scale. We have used the point on the grid that shows the maximum amplitude (0 dB) as reference (“reference point” in Fig. 5). The intensity is lower when the height is higher. In fact, at a distance $x = 20$ m the intensity clearly decreases moving from $h = 0.75$ m ($I \approx -15$ dB) to $h = 1.35$ m ($I \approx -20$ dB). This difference is then reflected in the accuracy map in Fig. 8.

It is worth to point out that the ratio in Eq. (8) makes us to neglect the effects of the electronic components, $H_{TX}(f)$ and $H_{RX}(f)$, of the $i$-th received signal. We also note that spatial-dependent contributions of collecting optics can be neglected at the receiver, and hence the TF of the optical part at the receiver can be considered as constant for every point of the grid. This is due to the fact that the focused image of lamp is fully contained into the active area of the photodiode for all points in the measurement grid (see Sec. II). Due to the

\[^1\]The ratio $\frac{S_{RX}(f)}{S_{TX}(f)}$ is usually defined as the transfer function of a generic system (black box) that has $s_{TX}(t)$ as input and $s_{RX}(t)$ as output.

\[^2\]The effect of the thermal noise at the receiver has been neglected.

operations described above, $\Delta H_i$ in Eq. (8) does not depend on the frequency.

A. Lambertian Model

In the Lambertian model [8], the path loss is given by

$$I = \frac{I_0 \cos^\nu(\phi)}{d^2}$$

where $I_0$ is the intensity in the axis of irradiance, $\phi$ is the irradiance angle, $\nu$ is a parameter for taking into account the influence of an optical lens in the LED emission, and $d$ is a distance between the transmitter and the receiver. The angle $\phi$ is referred to the optical axis of traffic light lens, identified as the line of maximum irradiance of the source (in our case, this is found to coincide with a line connecting the center of the LED lamp and the point (14, 0, 0) m, i.e. located at 14 m on the ground in front of the traffic light). Assuming rotational symmetry around the optical axis, the angle $\phi$ is illustrated in Fig. 6.

B. Proposed models for Intensity

To mathematically characterize the propagation of the VLC signal, we aim to model the amplitude $\Delta H_i$ of the received signal over the grid, which in turn is quantified by its intensity.

Fig. 5: Amplitude map (dB) of the VLC signal $\Delta H_i$ measured over the grid for three different heights [0.75, 1, 1.35] m corresponding to car headlights, can desk and car internal mirror, respectively. The map reports the measurement data (intensity) over the field test grid shown in Fig. 3.
The intensity of the received signal depends on the position of the receiver in spherical coordinates

\[ I(\alpha, \beta, d) \]

where \( \alpha \) is the elevation, \( \beta \) is the azimuth and \( d \) is the distance between the transmitter and the receiver. Angles \( \alpha \) and \( \beta \) are shown in Fig. 6.

Fig. 6: [Left side] Angle \( \phi \) of the Lambertian model (9). [Right side] Angles \( \alpha \) and \( \beta \) of the proposed models (10), (11). The horizontal axis is parallel to the road axis.

The first model proposed here (10) is derived by modifying the well-known Lambertian propagation model [8]. Here the optical intensity is taken as:

\[ I_1(\alpha, \beta, d) = \frac{\beta}{d^2}, \]  

(10)

scaling as inversely proportional to the square of the distance. Differently from Lambert’s law, the numerator has been generalized as a function of \( \alpha \) and \( \beta \).

The Lambert’s cosine propagation law assumes that the source of light is a homogeneous diffuser. Whilst remaining a valid approximation in the analysis of emission of many direct sources as LEDs, this hypothesis can be highly inaccurate in case of shaped beam patterns, i.e. in presence of lenses (as in traffic lights). Thus, the second propagation model proposed here only preserves a global \( 1/d^2 \) dependence of radiated optical intensity:

\[ I_2(\alpha, \beta, d) = \frac{\beta}{d^2}. \]  

(11)

The numerator is instead a generic polynomial function of \( \alpha \) and \( \beta \).

A more general approach can consider the optical intensity as a polynomial function of \( \alpha, \beta, d \) (with a given order) without an explicit \( 1/d^2 \) weight, i.e.,

\[ I_3 = f(\alpha, \beta, d) \]  

(12)

Depending on the specific needs and applications, one can choose a general or a more specific model which in general will feature very different convergence performances as a function of the number of parameters employed in the fitting procedure.

The multiple generalized linear regression (MGLR) method has been used to find the best fitting parameters for the three models \( I_1, I_2 \) and \( I_3 \). Multiple linear regression is a generalization of simple linear regression to the case of more than one independent variable, and a special case of general linear models, restricted to one dependent variable.

The generalized linear model (GLM) is typically used to model an irregular emission pattern [27], [28], [29] as it is a flexible generalization of ordinary linear regression that allows for response variables that have error distribution models other than a normal distribution. The GLM generalizes linear regression by allowing the linear model to be related to the response variable via a link function and by allowing the magnitude of the variance of each measurement to be a function of its predicted value. The output of the MGLR method is a set of polynomial coefficients of the function that fits the intensity.

In our procedure, the inputs of the MGLR method are the following parameters: the model under evaluation \( (I_1, I_2, I_3) \), the maximum order of the polynomial function for each variable \((\alpha, \beta, d)\), the statistical distribution of the error \((\text{Normal, Poisson, Gamma})\)

Let \( y_1, \cdots, y_n \) denote \( n \) independent observations on a response. We treat \( y_i \) as a realization of a random variable \( Y_i \).

In the general linear model we assume that \( Y_i \) has a normal distribution with mean \( \mu_i \) and variance \( \sigma \). We further assume that the expected value \( \mu_i \) is a linear function of \( p \) predictors that take values \( x_i = (x_{i1}, x_{i2}, \cdots, x_{ip}) \) for the \( i \)-th case, so that

\[ E[Y_i] = \mu_i = x_i^T \beta \]  

(13)

where \( \beta \) is a vector of unknown parameters. Once found \( \beta \) from (13) we can write

\[ \mu_i = b_0 + b_1x_{i1} + b_2x_{i2} + \cdots + b_px_{ip} \]  

(14)

Different (from Normal) statistical distribution of the error can be used by taking into account the generalized linear model. In a generalized linear model, the outcome \( Y_i \) of the dependent variables is assumed to be generated from a particular distribution in the exponential family (Normal, Poisson and Gamma). The vector of the mean \( \mu_i \) of the distribution depends on the independent variables \( x_i \) through

\[ g(\mu_i) = g^{-1}(x_i^T \beta) \]  

(15)

where \( g(\cdot) \) is the link function.

To find the highest polynomial order for each of the variables \((\alpha, \beta, d)\) to be inserted into the MGLR model still avoiding the overfitting problem, a k-fold method is used [31], [32], [33]. The k-fold has been applied to every model \((I_1, I_2, I_3)\) for each one of the error distribution (Normal, Poisson, Gamma). Thus, nine models have been evaluated. The polynomial order of \( \alpha \) and \( d \) ranges from 1 to 9, while the order of \( \beta \) can assume only even values, from 2 to 8, because of azimuthal-symmetric nature of the emitted pattern.

Incidentally, by assuming that Fresnel lenses provide the same beam shape for the three LED lamps (red, orange and...
green), we remark that our model allows to directly retrieve intensity maps for all of them, as the only difference among the three is given by the different height with respect to ground and by the relative position of the emission wavelength peak with respect to the responsivity curve of Silicon photodiode. This can be simply taken into account with a different attenuation factor (depending on the specific colour) for all of the three models given by Eqs. (10), (11) and (12).

V. Results

The percentage error for each of the proposed models \( I_1(\alpha, \beta, d), I_2(\alpha, \beta, d), I_3(\alpha, \beta, d) \) as a function of the number of polynomial terms with a Normal (Fig. 7a), Gamma (Fig. 7b) and Poisson (Fig. 7c) distribution in the MGLR method. Each point corresponds to a specific parameter configuration. For example, one red square in Fig. 7a represents the error produced by the model \( I_3(\alpha, \beta, d) \) with a specific number of polynomial terms for \( \alpha, \beta \) and \( d \) coming out from the application of the k-fold and MGLR procedures. The error is calculated as the RMSE between the original measured data and the corresponding values of the polynomial fitting function over the whole grid. Solid lines are a guide to the eye connecting the best parameter configuration minimizing the error for a specific number of terms. Given a specific error threshold, it is possible to reach it with the lowest number of polynomial terms, after which the minimization procedure looses efficiency.

As it can be seen in Fig. 7a, the error decreases faster for model \( I_2(\alpha, \beta, d) \). The other two models can get more accurate, at the expenses of introducing a high number of polynomial terms. Hence depending on the particular needs one can either choose a more “efficient” or a more “accurate” model for the VLC system.

If a Gamma or Poisson distribution is used in the MGLR method, the error decreases faster than using the Normal distribution. The \( I_2 \) model with Poisson distribution reaches the lowest error, but it requires a high number of polynomial terms (80) for the fitting function. If a low number of polynomial terms is desired, the model \( I_1 \) with Gamma or Poisson distribution reaches a lower error. A low number of polynomial terms is useful for extending the model to an area larger than the one where the measurements are taken. The k-fold method is used to avoid the overfitting problem in the area of the measurements.

Only two models have been compared to define a propagation model which is valid for distances out of measurement’s range. The first one is \( I_1 \) with Gamma distribution, while the second is \( I_2 \) with Poisson distribution. The first one has been selected because the error decreases more rapidly, the second one has been selected because it reaches the lowest error.

In case of \( I_1(\alpha, \beta, d) \) with Gamma distribution, the vector \( y \) of the intensity measured over the grid is

\[
y = (\Delta H_1 d_1^2, \Delta H_2 d_2^2, \ldots, \Delta H_n d_n^2)
\]  

(16)

where \( n \) is the number of points over the grid of the measurements, \( \Delta H_i \) is the intensity (in dB) in the \( i \)-th point of the grid and \( d_i \) is the distance of the \( i \)-th point of the grid from the source of the VLC signal (the traffic light).

The predictor matrix is composed by

\[
X = [x_1 \ x_2 \ \cdots \ x_n]^T
\]

(17)

with

\[
x_i = [1 \ \alpha_i \ \beta_i \ \alpha_i \otimes \beta_i]^T
\]

(18)

where \( \otimes \) stands for the element-by-element product operator, \( \alpha_i = (\alpha_i^1, \alpha_i^2, \ldots, \alpha_i^p) \) is the vector of the elevation angles and \( \beta_i = (\beta_i^1, \beta_i^2, \ldots, \beta_i^{2q}) \) is the vector of the azimuth (squared) angles. The parameters \( p \) and \( 2q \) represent the maximum polynomial order for the elevation parameter \( \alpha \) and the azimuth parameter \( \beta \), respectively.
In case of \( I_2(\alpha, \beta, d) \) with Poisson distribution, the vector \( y \) of the intensity measured over the grid is the same (16), while
\[
x_i = [1 \ \alpha_i \ \beta_i \ d_i \ \alpha_i \otimes \beta_i \ \otimes d_i]
\]
(19)
where \( d_i = (d_{i1}, d_{i2}, \ldots, d_{i\tau}) \). The parameter \( \tau \) is the maximum polynomial order for the distance parameter \( d \).

The objective is to find the coefficients \( b = (b_1, \cdots, b_m)^T \) (with \( m = pq + 1 \)) so that
\[
E[y]^{-1} = Xb
\]
in case of model \( I_1(\alpha, \beta, d) \) with Gamma distribution, and
\[
\log(E[y]) = Xb
\]
in case of model \( I_2(\alpha, \beta, d) \) with Poisson distribution (with \( m = pq\tau + 1 \)).

Table I and II show the different solutions for (20) and (21), respectively, increasing the order of the polynomial function. The error associated to each model compared with the measurements is reported as average percentage error
\[
\epsilon_{perc} = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - I_{1i}|}{y_i} \times 100
\]
(22)
as well as root mean squared error (RMSE)
\[
\epsilon_{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - I_{1i})^2}{n}}
\]
(23)
with \( i = 1, \cdots, n \) and \( l = 1, 2, 3 \).

### TABLE I: Number of terms, orders of the variables and error for the solution (20) of model \( I_1(\alpha, \beta, d) \) with Gamma distribution. Only values of \( \alpha, \beta \) that minimize the error are reported (see the curve min \( I_1 \) in Fig. 7b).

| No. of terms | Order \( \alpha \) | Order \( \beta \) | Error (%) | RMSE |
|--------------|----------------|----------------|------------|------|
| 30           | 6              | 8              | 3.7307     | 0.014001 |
| 25           | 6              | 8              | 4.3831     | 0.017496 |
| 22           | 6              | 6              | 4.5449     | 0.018839 |
| 14           | 6              | 6              | 6.5972     | 0.02563 |
| 12           | 4              | 4              | 8.3442     | 0.030691 |
| 9            | 2              | 6              | 16.002     | 0.061841 |
| 6            | 2              | 4              | 20.056     | 0.084346 |
| 5            | 2              | 2              | 30.182     | 0.10461 |

### TABLE II: Number of terms, orders of the variables and error for the solution (21) of model \( I_2(\alpha, \beta, d) \) with Poisson distribution. Only values of \( \alpha, \beta \) that minimize the error are reported (see the curve min \( I_2 \) in Fig. 7c).

| No. terms | Order \( \alpha \) | Order \( \beta \) | Order \( d \) | Error (%) | RMSE |
|-----------|----------------|----------------|-------------|------------|------|
| 90        | 8              | 8              | 2           | 0.8221     | 0.0016 |
| 86        | 7              | 8              | 2           | 1.9923     | 0.0049 |
| 60        | 6              | 4              | 4           | 2.5728     | 0.006811 |
| 48        | 5              | 6              | 3           | 3.0709     | 0.010701 |
| 40        | 6              | 6              | 1           | 3.3717     | 0.010796 |
| 40        | 6              | 2              | 1           | 5.9458     | 0.014649 |
| 20        | 5              | 2              | 1           | 8.5654     | 0.029661 |
| 12        | 3              | 2              | 1           | 31.776     | 0.056186 |

The proposed models have been compared to the conventional Lambertian model [1], widely used in literature. Fig. 8 shows the accuracy of the Lambertian model as well as of the proposed new models, over the experimental measurements grid. It is important to note that the accuracy has been calculated only over the points where we had the value corresponding to a real measurement. We could not measure the intensity on the other side of the traffic-light since there was no lane available, as in a typical urban road.

The accuracy is calculated as the difference between the measured intensity in a point of the grid and the corresponding intensity of the fitting polynomial function of the model.

As it can be seen the conventional model shows a significant higher error compared to the models proposed in this paper. The average RMSE of the conventional Lambertian is 0.097, while the one of the two proposed models (with, e.g., 12 parameters) is 0.030 and 0.029 (see Table I and II).

Fig. 9 shows the intensity of the propagation models \( I_1(\alpha, \beta, d) \) model with Gamma distribution and \( I_2(\alpha, \beta, d) \)
The models can be reliably used to evaluate the performance of VLC technology for automotive applications. In addition, the results show that the Lambertian model is less accurate in describing the VLC transmission-propagation from traffic-light to vehicles when real lamps are considered. The proposed models have been compared in terms of complexity and accuracy, together with the conventional Lambertian model. The data have been then used to mathematically model the transmission pattern and the propagation channel. Three models have been proposed and compared in terms of complexity and accuracy, with the conventional Lambertian model. The results show that the Lambertian model is less accurate in describing the VLC transmission-propagation from traffic-light to vehicles when real lamps are considered. The proposed models provide for a more accurate description of the transmission channel, and can be extended to distances wider than those considered during the measurements campaign. In particular, the Lambertian model provides an average error of 25%, while the two proposed polynomial models obtain 8% with 12 coefficients. Anyway, the regression process to calculate the coefficients of the two proposed polynomial models has a computational cost higher than the Lambertian model.

The models can be reliably used to evaluate the performance of VLC technology for automotive applications. In addition,
the theoretical error probability of the VLC signal has been derived by using the experimental data. The results highlight that an uncoded error probability of $10^{-3}$ is achievable at 30 m distance, showing that our implementation can be safely used
for data service in I2V applications. Furthermore, our work can be a valuable starting point to assess the impact of more generic weather scenarios on I2V VLC links, such as those in presence of fog and/or rain.

ACKNOWLEDGMENT

This work has been performed by the Visible Light Communications Research laboratory (VisiCoRe, https://visicore.dinfo.unifi.it/), a partnership of University of Florence, CNR-INO and LENS.

Authors would like to thank the company ILES srl (http://www.ilessrl.com) for supplying the traffic-light and for the important support given during the measurements campaign.

REFERENCES

[1] T. Komine and M. Nakagawa, “Fundamental analysis for visible-light communication system using LED lights,” IEEE Transactions on Consumer Electronics, vol. 50, no. 1, pp. 100–107, 2004.

[2] S. Allattani, “Review of LiFi technology and its future applications,” Journal of Optical Communications, 2018.

[3] N. Chi, H. Haas, M. Kavehrad, T. D. C. Little, and X. Huang, “Visible light communications: demand factors, benefits and opportunities,” IEEE Wireless Communications, vol. 22, no. 2, pp. 5–7, 2015.

[4] F. Martinez, C. Toh, J. C. Cano, C. Calafate, and P. Manzoni, “Emergency services in future intelligent transportation systems based on vehicular communication networks,” IEEE Intelligent Transportation Systems Magazine, vol. 2, pp. 6–20, 2010.

[5] G. Karagiannis, O. Altintas, E. Ekici, G. Heijenk, B. Jarupan, K. Lin, and T. Weil, “Vehicular networking: A survey and tutorial on requirements, architectures, challenges, standards and solutions,” IEEE communications surveys and tutorials, vol. accepted f, no. 4, pp. 584–616, 2011.

[6] S. Wu, H. Wang, and C. Youn, “Visible light communications for 5G wireless networking systems: from fixed to mobile communications,” IEEE Network, vol. 26, no. 6, pp. 41–45, 2014.

[7] N. Chi, Led-based Visible Light Communications. Springer-verlag, Berlin AN, 2018.

[8] M. Akanegawa, Y. Tanaka, and M. Nakagawa, “Basic study on traffic information system using LED traffic lights,” IEEE Intelligent Transportation Systems, vol. 2, no. 4, pp. 197–203, 2001.

[9] A. Correa, A. Hamid, and E. Sparks, “Li-Fi based smart traffic network,” in IEEE Transportation Electrification Conference and Expo (ITEC), Long Beach, CA, USA, 2018, pp. 217–219.

[10] M. Boban, A. Kousaridas, K. Manolakis, J. Eichinger, and W. Xu, “Connected roads of the future: Use cases, requirements, and design considerations for vehicle-to-everything communications,” IEEE Vehicular Technology Magazine, vol. 13, no. 3, pp. 110–123, 2018.

[11] F. Hu, Vehicle-to-vehicle and Vehicle-to-infrastructure Communications: A Technical Approach. CRC Press, 2018.

[12] M. Boban, K. Manolakis, M. Ibrahim, S. Buzzi, and W. Xu, “Design aspects for 5G V2X physical layer,” in IEEE Conference on Standards for Communications and Networking (CSCN), 2016, pp. 1–7.

[13] J. Lee, S. Y. Kwon, S. Y. Jung, and Y. H. Kwon, “Simulation modeling of visible light communication channel for automotive applications,” in 15th International IEEE Conference on Intelligent Transportation Systems, Anchorage, Alaska, USA, 2012, pp. 1–7.

[14] K. Cui, G. Chen, Z. Xu, and R. D. Roberts, “Experimental characterization of traffic light to vehicle VLC link performance,” in 2011 IEEE GLOBECOM Workshops (GC Wkshps), 2011, pp. 808–812.

[15] J. R. Barry, J. M. Kahn, W. J. Krause, E. A. Lee, and D. G. Messer-chmit, “Simulation of multipath impulse response for indoor wireless optical channels,” IEEE Journal on Selected Areas in Communications, vol. 11, no. 3, pp. 367–379, 1993.

[16] K. Lee, H. Park, and J. R. Barry, “Indoor channel characteristics for visible light communications,” IEEE Communications Letters, vol. 15, no. 2, pp. 217–219, 2011.

[17] I. Yaqoob, I. A. T. Hashem, Y. Mehmood, A. Gani, S. Mokhtar, and S. Guizani, “Enabling communication technologies for smart cities,” IEEE Communications Magazine, vol. 55, no. 1, pp. 112–120, 2017.
The distance $\rho_{\text{min}}$ depends on the specific constellation that has been transmitted. In our experiments, an antipodal (BPSK) constellation was used, thus the $P_e$ can be written as

$$P_e = Q\left(\frac{PR}{\sigma_N^2}\right) = Q\left(\sqrt{\text{SNR}}\right)$$  \hspace{1cm} (28)$$

where $PR$ is the received power and $\sigma_N^2$ is the noise variance. To estimate the noise variance the following procedure has been carried out. Let us first remind that the received signal vector at location $i$ is

$$s_i = [s_{i1} \cdots s_{iN_b}]$$  \hspace{1cm} (29)$$

where $s_{ij}$ is the vector coming from the sampling of the waveform in (2). The received vector at the reference location ($i = \text{ref}$) is averaged over the repetition periods

$$\bar{s}_{\text{ref}} = \frac{1}{N_b} \sum_{j=1}^{N_b} s_{\text{ref}j}$$  \hspace{1cm} (30)$$

Now, the estimated noise vector at location $i$

$$w_i = [w_{i1} \cdots w_{iN_b}]$$  \hspace{1cm} (31)$$

is calculated as the difference between the received signal vector at location $i$ and the averaged vector at reference location multiplied by $\Delta H_i$ (6)

$$w_{ij} = s_{ij} - \bar{s}_{\text{ref}} \Delta H_i \quad j = 1, \cdots, N_b$$  \hspace{1cm} (32)$$

The noise vector in (31) is then used to estimate the noise variance $\sigma_{w_i}^2$ at location $i$. An example of histogram of the noise vector $w_i$ is reported in Fig. 10, where the grid point $i$ is $(x = 30, y = 1.75, z = 0.75)$ m. Fig. 10 shows the occurrences (y-axis) of the noise samples (x-axis) as in (32).

The signal-to-noise ratio (SNR) can thus be calculated as

$$\text{SNR}_i = \frac{E[\bar{s}_{\text{ref}}^2] \Delta H_i^2}{\sigma_{w_i}^2}$$  \hspace{1cm} (33)$$

This result can be used to calculate the error probability (28) in every point of the measurement grid. The map of the error probability is reported in Fig. 11 for the height 0.75 m. A probability of error of $10^{-3}$ (uncoded) can be provided even at 30 m away from the traffic light. The map shows the error probability over the perimeter where the experimental measurements have been carried on. The intensity models proposed in this paper can be anyway used to calculate the performance of a VLC system in different perimeters.

B. FOV and reflections form tarmac

Here we show that in our configuration reflections of lamp’s light from tarmac are never collected by receiver as they fall out of the Field of View (FOV) of the receiver anywhere on the measurements grid. Without loss of generality we restrict our analysis to the case of receiver placed along the optical axis of the traffic light lens ($\beta = 0$). With reference to Fig. 12-a, the FOV $\theta$ of a receiver is given by $\theta/2 = \arctan(l/2F)$, where $l$ and $F$ are the detector width and lens focal length, respectively. As shown in Fig. 12-b, one can see that reflections from tarmac (red line) cannot geometrically enter the detector FOV when $\theta/2 \leq \gamma + \delta$. After easy trigonometric considerations, the above expression reads as:

$$\arctan\left(\frac{l}{2F}\right) \leq \arctan\left(\frac{Z + S}{D}\right) + \arctan\left(\frac{Z - S}{D}\right),$$  \hspace{1cm} (34)$$

where the parameters are described in Fig. 12. The above expression, given our parameters set ($F = 30$ mm, $l = 6$ mm) is always verified in our measurement grid ($D < 35$ m) and for all of the three receiver heights, and turns false only at very large distances ($D \gtrsim 50$ m, see Fig. 12-c), where the detector’s optical axis progressively leans towards a horizontal plane as the distance $D$ grows, hence letting to reflected beams the possibility to enter the FOV cone.
Fig. 12: a) Field of View for a given detector of size $l$ and focal length $F$; b) reflections from tarmac (red line) in relation to parameters used in the text; c) the merit function $\theta/2 - \gamma - \delta$ gets positive when reflections can geometrically enter the detector’s FOV.