TCP Performance Over Satellite-Based Hybrid FSO/RF Vehicular Networks: Modeling and Analysis

THANG K. NGUYEN\textsuperscript{1,}, CHUYEN T. NGUYEN\textsuperscript{1,}, HOANG D. LE\textsuperscript{2,} (Graduate Student Member, IEEE), AND ANH T. PHAM\textsuperscript{2,} (Senior Member, IEEE)

\textsuperscript{1}School of Electronics and Telecommunications, Hanoi University of Science and Technology, Hanoi 100000, Vietnam
\textsuperscript{2}Computer Communications Laboratory, The University of Aizu, Aizuwakamatsu 965-8580, Japan

Corresponding author: Chuyen T. Nguyen (chuyen.nguyenthanh@hust.edu.vn)

This work was supported by the Hanoi University of Science and Technology (HUST) under Project T2020-SAHEP-011.

ABSTRACT Recent years have witnessed a growing interest in Internet access from space assisted by low earth orbit (LEO) satellite networks. In the domain of the last-mile access for the Internet of Vehicles (IoV), hybrid free-space optical (FSO)/radio-frequency (RF) communication has recently attracted worldwide research efforts. While the transmission control protocol (TCP) is the most widely deployed transport protocol on the Internet, its performance in the error-prone environment of LEO satellite-assisted hybrid FSO/RF vehicular networks is not well understood. This paper develops a comprehensive analytical model based on the cross-layer approach for TCP performance, considering the FSO and RF satellite fading channels, modeled by the Gamma-Gamma and Nakagami-\(m\) distributions, respectively. The error-control solutions, including the Reed-Solomon (RS) code and Selective repeat automatic repeat request (SR-ARQ), are also employed. Numerical results quantitatively demonstrate the impact of transmission errors at last-mile links and different parameters/settings of error-control solutions on the TCP performance. The paper also supports the selection of proper TCP variants for the considered networks.

INDEX TERMS Satellite networks, hybrid FSO/RF last-mile, internet of vehicles, transmission control protocols, error-control solutions.

I. INTRODUCTION

Thanks to the rapid technological development, several well-known companies have been working for years on plans to provide high-speed Internet access anywhere on Earth, such as SpaceX’s Starlink, Amazon’s Project Kuiper, the UK’s OneWeb, and ViaSat. The typical approach is to use low Earth orbit (LEO) satellites in relative proximity to Earth’s surface. The shorter delay in data transfer of LEO satellites than GEO satellites is one of the key reasons why the planned mega-constellations want to use them. Nevertheless, the challenge is that LEO satellites continuously orbit the Earth. As a result, they are only ever accessible for a short period from any one point on Earth. Therefore, a world-spanning network of thousands of such LEO satellites is supposed to enable quick data connections and the transfer of large quantities of data. These are substantial opening steps for the “Internet from Space” era with satellite internet access provision. With the Internet from Space, the Internet of Vehicles (IoV), which is one of the most rapid development applications of the Internet of Things (IoT) with massive access, has become possible. Satellite assisted-IoV will soon emerge as the scenario for the fifth and sixth generation networks (5G), (6G) in the near future.

In most existing satellite-assisted Internet networks, radio frequency (RF) bands are used for last-mile access. However, because of limited data rates, RF may not be effective for the high-speed satellite network. Free-space optical (FSO) communications, with higher data rates, lower power, and higher directivity due to ultra-narrow beams [1], is therefore being considered in this network. The challenge is that it is highly susceptible to atmospheric turbulence and heavily
dependent on weather effects as cloud [2]. On the other hand, traditional RF communication is more robust and less affected by atmospheric disturbances or weather conditions. Therefore, to take advantage of the high data rates of FSO and the reliability of RF communication technologies, the hybrid FSO/RF emerges as a promising candidate for last-mile access of satellite-assisted Internet networks.

In the transport layer, the transmission control protocol (TCP) is the most popular protocol for many Internet applications that require reliable transmissions [3]. However, it is challenging for the TCP, especially the standard variant using Additive Increase and Multiplicative Decrease (AIMD) algorithm, to maintain the performance over the Long Fat Network (LFN) of satellite-based hybrid FSO/RF communications. First, TCP tends to perform poorly in wireless environments because it misinterprets the losses due to bad channel conditions as an indication of network congestion. Second, the conventional TCP variants are practically ineffective in the LFN due to the slow linear growth function of the congestion window.

A. LITERATURE REVIEW AND MOTIVATIONS

The existing studies on the TCP performance over satellite networks mainly focus on RF-based last-mile access [4]–[11]. Particularly, Zhu et al. studied the performance of TCP enhancements for a hybrid terrestrial–satellite network in [5]. In [7], the authors evaluated the quality of video streaming applications using TCP over satellite links. The interactions of TCP at the transport layer and random access schemes at the Media Access Control (MAC) layer for machine-to-machine services over the satellite links were investigated in [8]. In [9], the performance of multi-path TCP was analyzed through a satellite network, which contains several WiFi access points connected to a server together with a convoy of drones for providing Internet access. In addition, a new technique called Path-Based Network Coding (PBNC) is proposed in [10] to maximize the TCP goodput for Terrestrial-Satellite Mobile Systems. And most recently, the performance evaluation of TCP NewReno over satellite-based Digital Video Broadcasting-Return Channel via Satellite (DVB-RCS2) links was investigated in [11].

There are also several works on the TCP performance under the impact of FSO-based last-mile access [12]–[15]. Particularly, in [12], [13], the authors analyzed the throughput performance of TCP-FSO over satellite-to-fixed ground station channel. In these works, the test-bed system with an optical emulator is used to conduct the measurements and validate the analytical results. In [14], the authors demonstrated the impact of atmospheric turbulence on the TCP throughput and confirmed the effectiveness of TCP Cubic in FSO-based satellite communication systems. And most recently, in [15], the TCP performance with Incremental Redundancy Hybrid ARQ (IR-HARQ) protocol was analytically investigated in FSO based LEO satellite-to-vehicles networks. Also, in this work, the authors showed that in the low error-rate conditions, TCP Cubic outperforms TCP Hybla but, when the transmission errors happen more frequently, TCP Hybla can maintain a better performance than TCP Cubic.

While the conventional TCP variants using AIMD algorithms can still maintain a good performance in satellite-based RF networks [11], it is challenging for these TCP variants to achieve high throughput in the enormous bandwidth of satellite-based FSO networks. For instance, the results in [14] illustrated that TCP variants using the AIMD algorithm, e.g., Reno, are not effective in the satellite FSO-based LFN. As a result, it is crucial to select the proper TCP variants for different network scenarios. At the time of writing, we realize that there are no works that address the TCP performance issue in satellite-based hybrid FSO/RF networks, as summarized in Table 1. For such networks, it is essential to answer two critical questions, as follows.

Q1. How about the TCP performance under the impact of fading channels of hybrid FSO/RF last-mile access?
Q2. Which TCP variants should be used for the scenario of satellite-assisted hybrid FSO/RF vehicular networks?

B. CONTRIBUTIONS

This paper aims to address the questions mentioned above, in which we investigate the performance of TCP variants in the scenario of satellite-assisted hybrid FSO/RF vehicular networks. The primary contributions of this paper can be summarized as follows.

1) We, for the first time, address the TCP performance over satellite-assisted hybrid FSO/RF vehicular networks. To this purpose, we introduce a novel cross-layer framework incorporating the considered physical Hybrid FSO/RF channels together with Forward Error Correction (FEC) code, link-layer Automatic Repeat Request (ARQ), and TCP variants at the transport layer.

2) Unlike [14]–[16], where only congestion avoidance phase was considered, we instead develop a complete analytical model for TCP performance evaluation taking into account other phases of TCP operation. Several well-known TCP variants, including TCP NewReno, TCP Hybla, High-speed TCP (HSTCP), and TCP Cubic, are investigated.

3) We provide the insightful results that allow selecting the suitable TCP variants for our considered networks. Furthermore, Monte-Carlo simulations are also performed to confirm the validity of analytical results.

It is noteworthy that an essential novelty of the work is the complete model that we set up by jointly considering the impact of both physical and link layers for different

| TABLE 1. Research status of TCP over satellite networks. |
|---------------------------------------------------------|
| Last-mile access | TCP over satellite networks |
|------------------|----------------------------|
| Radio frequency (RF) | Well investigated [4]–[11] |
| Free space optics (FSO) | Moderately investigated [12]–[15] |
| Hybrid FSO/RF | Not available |
TCP variants. These issues include the hybrid FSO/RF satellite channels, the presence of error-control methods of both physical and link layers, i.e., Reed-Solomon (RS) code and Selective Repeat ARQ (SR-ARQ). Each of them separately has been well studied. Still, to our best knowledge, no prior work puts all of these strongly interacting pieces together to understand the performance of TCP variants in satellite Internet access.

The rest of the paper is organized as follows. The network scenario is described in Section II. The hybrid FSO/RF switch-over system and fading channel modeling for FSO and RF are presented in Section III. The TCP performance analysis, including the TCP segment loss model, average RTT delay, and the comprehensive loss-based throughput, is shown in Section IV. Numerical results are given in Section V, and finally, Section VI concludes the paper.

![Diagram](image)

**FIGURE 1.** The IoV network scenario with last-mile satellite internet access.

### II. REFERENCE NETWORK SCENARIO

#### A. NETWORK DESCRIPTIONS

As shown in Fig. 1, we consider an end-to-end network scenario where a vehicle downloads data from a remote server (the TCP source) through LEO satellite Internet access. The satellite is connected to a TCP source placed within the wired section of the Internet. The connection between the LEO satellite and the vehicle is assumed to be the potential bottleneck in the network where the hybrid FSO/RF link is used. In this connection, FEC and SR-ARQ are respectively used for the physical layer and link layer-based error control protocol.

A hybrid FSO/RF system including two FSO and RF links is employed at the physical layer. Many degrading factors, including interference, absorption, scattering, refraction, and turbulence caused by the atmosphere and weather conditions (fog, rain), induce harmful effects for the FSO and RF links in different ways [17]. For example, the main deteriorating factors of FSO links are fog and atmospheric turbulence. On the other hand, while RF links are not severely affected by fog and atmospheric turbulence, they are susceptible to heavy rains. The integration of FSO and RF links with a switch-over system, therefore, has emerged as a promising solution for providing reliable and effective transport for critical real-time traffic in an outdoor wireless environment [18]. For both links, the popular Reed-Solomon code (RS) is employed to improve the error performance at the physical layer.

In the link layer, we assume that each TCP segment is divided and encapsulated into $n_f$ smaller packets, then data frames before being transmitted over the last-mile link. Selective Repeat ARQ (SR-ARQ) is applied for both RF and FSO links to ensure reliable data transmission by re-transmitting the incorrect frames with the persistent level of SR ARQ is defined as $X$ (i.e., the maximum number of transmission attempts for a frame).

We consider several different variants of TCP, including Cubic, Hybla, NewReno, and HSTCP. The performance of TCP depends on the congestion window control’s algorithms that are represented as functions of time to deal with loss events. We also assume that the loss events, indicated to the source by three-duplicate ACKs or time-out events, are caused by two main reasons: transmission loss and congestion loss. While errors at the physical layer cause transmission losses, congestion losses are caused by the buffer overflow at routers on the Internet.

#### B. TCP VARIANTS

In this section, we briefly review TCP congestion control algorithms considered in the paper, including NewReno, Hybla, Cubic and HSTCP.

1) TCP NewReno

TCP NewReno, an upgrade version of TCP Reno, is the most widespread congestion control protocol used in most Internet applications. The congestion window size, denoted as $W_{\text{NewReno}}$, is updated after a reception of non-duplicate ACK received or loss indications. This update is based on the AIMD algorithm by which the window size is increased by $\frac{\text{MSS}}{\tau + \text{RTT}}$ bytes for each ACK or one segment for each round-trip time (RTT) with (MSS) is the maximum segment size. It is reduced to half of the current window size for each segment loss detection by three-duplicate ACKs or 1 if time-out happens. The TCP NewReno’s window growth function can be expressed as

$$W_{\text{NewReno}}(x, \tau) = \begin{cases} \frac{2\pi^2}{\tau + \text{RTT}} & \text{SS phase,} \\ \frac{x}{2} & \text{CA phase.} \end{cases} \quad (1)$$

with $x$ is the last window reduction, $\tau$ is the elapsed time since the last lost event, and RTT is the average round-trip time. When the loss event is detected, TCP NewReno reduces its current window size by half.

2) TCP HYBLA

TCP Hybla behaves similarly to TCP NewReno. The difference is in the approach used by TCP Hybla to remedy the impact of high-latency terrestrial or satellite radio links. Hybla removes the effect of extended RTTs by using a
normalized value, given as $\rho = \frac{\text{RTT}}{\text{RTT}_0}$, where RTT$_0$ is the round-trip time of the reference connection to which we aim to equalize our performance. With $\rho$, the congestion window size of TCP Hybla can be expressed as

$$W_{\text{Hybla}}(x, \tau) = \begin{cases} \rho \times 2 \pi^2, & \text{SS phase,} \\ \frac{\rho^2 \tau}{\text{RTT}} + \frac{x}{2}, & \text{CA phase.} \end{cases}$$ (2)

It is noted that as RTT $\leq$ RTT$_0$, then TCP Hybla behaves as the standard TCPs (e.g., TCP NewReno) [19].

3) HIGH-SPEED TCP

High-Speed TCP (HSTCP) is the congestion control mechanism for extensive bandwidth connections. Different from standard TCP, HSTCP lets the congestion window increases by $a(w)$ segments per RTT in the absence of loss events and decreases to $W_{\text{HSTCP}}(1 - b(w))$ segments in response to an RTT with one or more loss events. Given $a(w)$ and $b(w)$, the window growth function of HPTCP can be given as

$$W_{\text{HSTCP}}(x, \tau) = \begin{cases} 2 \pi^2, & \text{SS phase,} \\ a(x) \frac{x}{\text{RTT}} + (1 - b(x))x, & \text{CA phase.} \end{cases}$$ (3)

where $a(w)$, $b(w)$ are the modified increase and decrease parameters of HSTCP as defined in RFC3649 [20] and $\tau$ is the elapsed time since the last loss event.

4) TCP CUBIC

The window size of TCP Cubic is a cubic function of time since the last loss event. Let $W_{\text{Cubic}}(x, \tau)$ denote the window size as a function of the last window reduction $x$ and elapsed time $\tau$, the Cubic’s window growth function can be given as

$$W_{\text{Cubic}}(x, \tau) = \begin{cases} 2 \pi^2, & \text{SS phase,} \\ C(\tau - \sqrt[3]{\beta \frac{x}{C}})^3 + x, & \text{CA phase.} \end{cases}$$ (4)

where $C$ and $\beta$ are the Cubic and the multiplicative decrease factors, respectively [16].

III. HYBRID FSO/RF LAST-MILE MODELING

In this Section, we first review the fading channel models of both FSO and RF links. Then, the design of link switching and channel-state for the considered hybrid FSO/RF last-mile system are presented.

A. FAADING CHANNEL MODELS

1) FSO LINK

The atmospheric turbulence phenomenon causes the scintillation effect resulting in the signal power fluctuations at the vehicle’s detector. As reported in [21], the Gamma-Gamma (GG) model can be used to describe a wide range of turbulence conditions for the down-link from satellite to vehicle, in which the probability density (PDF) of received signal-to-noise ratio (SNR), denoted by $\gamma_f$, is given as [21, (1)]

$$f_{\gamma_f}(\gamma_f) = \frac{(\alpha \beta)^{\alpha + \beta}}{\Gamma(\alpha) \Gamma(\beta) \gamma_f^{\alpha + \beta - 1}} \left(2 \sqrt{\gamma_f \gamma_f} \right).$$ (5)

where $\gamma_f$ is the average value of $\gamma_f$, $K_{\alpha - \beta}(\cdot)$ is the modified Bessel function of the second kind of order $\alpha - \beta$, and $\Gamma(\cdot) := \int_0^\infty t^{\alpha - 1}e^{-t}dt$ is the Gamma function. Here, $\alpha$ and $\beta$ are effective numbers of large-scale and small-scale eddies of the scattering environment, which can be written as [22, (2)]

$$\alpha = \left\{ \exp\left[\frac{0.49 \sigma_R^2}{(1 + 1.11 \sigma_R^{12/5})^{1/6}} \right] - 1 \right\}^{-1},$$

$$\beta = \left\{ \exp\left[\frac{0.51 \sigma_R^2}{(1 + 0.69 \sigma_R^{12/5})^{1/6}} \right] - 1 \right\}^{-1},$$

where $\sigma_R^2$ is the Rytov variance and in case of plane wave propagation, it is written as [21, (2)]

$$\sigma_R^2 = 2.25 k^2 / 6 \sec^{11/6}(\xi) \int_H^{\infty} C_n^2(h)(h - h_v)^{5/6} dh,$$ (8)

where $k = \frac{2\pi}{\lambda}$ is the optical wave number with $\lambda$ is the optical wavelength, $\xi$ is the zenith angle, $H$ is the satellite altitude, $h_v$ is the vehicle height, and $\sec(\cdot)$ is secant function. In addition, $C_n^2(h)$ represents the turbulence strength accordingly to the altitude $h$, and it is determined based on the most widely used Hufnagel-Vally model as [23, (12.1)]

$$C_n^2(h) = 0.00594 \left( \frac{v_{\text{wind}}}{27} \right)^2 \left( 10^{-5} h \right)^{10} \exp \left( -\frac{h}{1000} \right) + 2.7 \times 10^{-16} \exp \left( -\frac{h}{1500} \right) + C_n^2(0) \exp \left( -\frac{h}{100} \right).$$ (9)

where $v_{\text{wind}}$ is the root mean square (RMS) wind speed, and $C_n^2(0)$ is the ground turbulence level varying from $10^{-17} m^{-2/3}$ to $10^{-13} m^{-2/3}$.

Given the PDF of $\gamma_f$ in (5), its cumulative distribution function (CDF) can be computed as [24, (4)]

$$F_{\gamma_f}(\gamma_f) = \frac{1}{\Gamma(\alpha) \Gamma(\beta)} G_{1,3}^{2,1} \left[ \frac{1}{\sqrt[3]{\gamma_f}} \left( \sqrt[3]{1} \right) \right].$$ (10)

where $G_{a,b,d}^{c,d}[\cdot]$ is Meijer’s G-function.

On the other hand, one of the essential parameters to deeply understand the fading channels is the level crossing rate (LCR). It is defined as the average number of times per second that received SNR passes the certain threshold $\gamma_{th}$ (either positive or negative direction). The LCR for GG model defined in [25, (26)] is, however, not available in the closed-form expression. By applying the Gauss-Laguerre quadrature approximation $\int_{-\infty}^{\infty} \exp(-x) g(x)dx \approx \sum_{i=1}^{n} w_i g(x_i)$ in [26, (25.4.45)], the closed-form expression for LCR can be obtained in (11), as shown at the bottom of the next page, where $n$ is the Gauss-Laguerre approximation order, while $w_i$ and $x_i$ are respectively the weight factor and the $i$-th zero of
Laguerre polynomial \( L_n(x) \) found in [26, Table 25.9]. In addition, \( \sigma_e^2 \) is the log-intensity variance found in [1], and \( t_{ci} \) is the FSO fading channel coherence time defined as \( t_{ci} = \frac{v_{wind}}{c} \) [21, (7)].

2) RF LINK
The RF-based satellite fading channel can be well characterized by the Nakagami-\( m \) model thanks to its capability of describing different fading scenarios, from line-of-sight (LOS) to non-line-of-sight (NLOS) [27]. Furthermore, it was verified in [28] that using the Nakagami-\( m \) model simplifies the analysis while maintaining good accuracy. As a result, the PDF of received SNR in RF fading channels, denoted by \( \gamma_r \), can be expressed as [28, Table 3]

\[
f_{\gamma_r}(\gamma_r) = \frac{m^m \gamma_r^{m-1}}{\tilde{\gamma}_r^m \Gamma(m)} \exp \left( -\frac{m\gamma_r}{\tilde{\gamma}_r} \right), \quad \gamma_r \geq 0,
\]

where \( \tilde{\gamma}_r \) is the average value of \( \gamma_r \), and \( m \) represents different fading scenarios (refer to [28] for more details). Given PDF of \( \gamma_r \), its CDF can be determined as

\[
F_{\gamma_r}(\gamma_r) = \frac{1}{\Gamma(m)} \gamma_r^m, \quad \gamma_r \geq 0,
\]

where \( \gamma(m, x) := \int_0^x t^{m-1} e^{-t} dt \) is the lower incomplete Gamma function.

On the other hand, the LCR of Nakagami-\( m \) fading model, denoted as LCR(\( \gamma_{th} \)), is given by [29, (10)]

\[
\text{LCR} (\gamma_{th}) = \sqrt{2\pi} \frac{m \gamma_{th}}{\tilde{\gamma}_r} f_d \left( \frac{m \gamma_{th}}{\tilde{\gamma}_r} \right)^{m-1} \exp \left( -\frac{m\gamma_{th}}{\tilde{\gamma}_r} \right).
\]

where \( f_d \) is the maximum Doppler frequency, which is due to the relative motion between LEO satellite and vehicles. Given the frequency carrier as \( f_c \), the Doppler frequency shift can be written in (15), as shown at the bottom of the page, where \( c \) is the speed of light in vacuum, \( \theta_{\text{max}} \) is the maximum elevation angle, \( w_F \) is the relative angular velocity of the satellite with respect to the vehicle, \( r_E \) is the radius of Earth, and \( r \) is the distance between the satellite and the center of Earth [30, (5)].

In addition, for the analysis of the Doppler effect with respect to time, it is necessary to investigate the satellite visibility window duration. It is defined as the total duration that the LEO satellite can communicate effectively with the vehicle within a small percentage of the satellite orbit period and can be expressed as [30, (11)]

\[
\tau(\theta_{\text{max}}) = \frac{2}{w_F} \arccos \left( \frac{\cos \left( \arccos \left( \frac{r_c}{r} \cos (\theta_{\text{max}}) \right) - \theta_{\text{max}} \right) - \theta_{\text{min}} \right) \right),
\]

where \( \theta_{\text{min}} \) and \( \theta_{\text{max}} \) are the minimum and maximum values of elevation angle in a satellite pass.

An example of Doppler frequency shift according to visibility time is illustrated in Fig. 2. The satellite orbit is \( H = 600 \text{ km}, f_c = 2.4 \text{ GHz} \), the vehicle’s velocity is \( v_c = 50 \text{ km/h} \), the minimum elevation angle is \( \theta_{\text{min}} = 10^\circ \). Also, different maximum elevation angles \( \theta_{\text{max}} \), i.e., 20°, 30°, 60°, 90° are taken into account. As seen, the maximum Doppler frequency shift is ranging from approximately −50 kHz to 50 kHz.

### FIGURE 2. An example of Doppler shift between a LEO satellite and a vehicle.

### B. LINK SWITCHING AND CHANNEL-STATE DESIGN

1) LINK SWITCHING DESIGN
Fig. 3 shows a hybrid FSO/RF with a link switching system including two line-of-sight (LOS) FSO and RF channels that operate at different transmission rates. A feedback link is deployed to convey the channel state information (CSIs) of both FSO or RF links back to the transmitter. The FSO link,
with higher bandwidth, is considered a primary link in this system, while the RF link acts as a backup when the FSO link is down. Also, the feedback link is assumed to be error-free and no latency.

Based on the last CSI of FSO as a default link, the system decides the transmission link at the beginning of a transmission cycle. Initially, the system transmits data by using the FSO link, which has a higher transmission rate. Whenever the environmental condition worsens, which leads to the decrease of FSO channel quality, the system switches to the backup link as RF. When using the RF link, if the quality of the RF link is not good and the FSO’s CSI is not favorable either, the system is supposed to be in the outage state. Whenever the FSO link CSI gets better, the system switches back to the FSO link. Denote $S$ as the link selection for this transmission, $S$ can be expressed as

$$
S = \begin{cases} 
\text{FSO}, & \gamma_{\text{thres}} \leq \gamma < \infty, \\
\text{RF}, & \gamma_{\text{out}} \leq \gamma < \gamma_{\text{thres}}, \\
\text{Outage}, & 0 \leq \gamma < \gamma_{\text{out}},
\end{cases}
$$

(17)

where $\gamma_f$, $\gamma_r$ are the instantaneous received SNRs from FSO and RF links, respectively.

The probability of link switching from FSO to RF therefore can be given as

$$
P_{\text{switch}} = P(\gamma_f < \gamma_{\text{thres}}) = F_{\gamma_f}(\gamma_{\text{thres}}),
$$

(18)

where $F_{\gamma_f}$ is the CDF of FSO channel. Obviously, at high average received SNR of the FSO link, the link switching probability is very low and it is increasing when the quality of the FSO link degrades. Atmospheric turbulence and zenith angle also greatly affect the link switching probability. For example, comparing Fig. 4(a) and Fig. 4(b), to keep the link switching probability under 0.1, the required average received SNR is higher than 7 dB and 14 dB corresponding to the weak and strong turbulence conditions.

The outage event happens when the system is operating at RF link and $\gamma < \gamma_{\text{out}}$; outage probability thus can be determined as

$$
P_{\text{outage}} = P(\gamma_f < \gamma_{\text{thres}})P(\gamma_r < \gamma_{\text{out}}) = F_{\gamma_f}(\gamma_{\text{thres}})F_{\gamma_r}(\gamma_{\text{out}}),
$$

(19)

where $F_{\gamma_f}$, $F_{\gamma_r}$ are the CDF of FSO and RF channel, respectively.

Fig. 5 depicts the outage probability as a function of both $\gamma_f$ and $\gamma_r$ with different turbulence conditions and Nakagami factors $m$. As evident, the atmospheric turbulence has a significant impact on the outage probability. For example, to keep the outage probability under $10^{-2}$, the required SNRs are $\{\gamma_f, \gamma_r > 15\text{dB}\}$ and $\{\gamma_f, \gamma_r > 25\text{dB}\}$ corresponding to the weak and strong turbulence conditions.

2) CHANNEL-STATE MODELING

To effectively facilitate the frame transmissions, we now design a channel-state model, in which a range of SNRs
defines a state. The duration of the channel state is selected long enough for the transmission of multiple frames can be completed within a state, and it should be equal and shorter than the channel fading coherence time.\(^1\) Both FSO and RF channels are organized into non-overlapping consecutive states. The channel is said to be in-state \(k\)-th if the received SNR falls into the interval of \([\gamma_k, \gamma_{k+1})\). In addition, the duration of channel-state \(k\)-th, denoted by \(\bar{\tau}_k\), depends on the statistical characteristic of fading channels, which can be obtained as \([31, (27)]\)

\[
\bar{\tau}_k = \frac{P(\gamma_k \leq \gamma \leq \gamma_{k+1})}{LCR(\gamma_k) + LCR(\gamma_{k+1})}, \quad (20)
\]

where \(LCR(\gamma_k)\) is the LCR given in (11) for FSO and (14) for RF. Additionally, \(P_k\) is the probability of the channel-state \(k\)-th, which is determined as

**FSO model:**

\[
P_k = \int_{\gamma_k}^{\gamma_{k+1}} f(\gamma) d\gamma = F(\gamma_{k+1}) - F(\gamma_k), \quad (21)
\]

**RF model:**

\[
P_k = \int_0^{\gamma_{\text{thres}}} f(\gamma) d\gamma \int_{\gamma_k}^{\gamma_{k+1}} f(\gamma) d\gamma = F(\gamma_{\text{thres}})[F(\gamma_{k+1}) - F(\gamma_k)], \quad (22)
\]

where \(f(\gamma)\) and \(F(\gamma)\) are the channel PDF and CDF, respectively.

We now determine the SNR threshold levels \(\gamma_k\) for the channel-state model, including both FSO and RF. Given the fading channel coherence time of \(\tau_c\) (FSO) and \(\tau_c\) (RF), the algorithm for searching these thresholds of FSO and RF models is summarized as in the Algorithm below. After the searching process, all \(M + 1\) and \(N + 1\) threshold levels corresponding to \(M\) and \(N\) channel-states of FSO and RF models, respectively, can be determined. Here, we set \(\gamma_{\text{thres}} = \gamma_{i_{\text{int}}}\) and \(\gamma_{\text{outs}} = \gamma_{i_{\text{ext}}}\).

Given \([\gamma_k, \gamma_{k+1})\), the average bit error rate (BER) at the channel-state \(k\)-th can be computed as

**FSO model:**

\[
\text{BER}_k = \frac{1}{P_k} \int_{\gamma_k}^{\gamma_{k+1}} \text{BER}(\gamma_k, \gamma_{k+1}) d\gamma = \frac{0.1}{P_k} \Gamma(\alpha, \beta)B(\alpha, \beta, 1 - \alpha + \beta)(K - L), \quad (23)
\]

where \(B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x + y)}\) is the Beta function, \(K\) and \(L\) can be expressed as \([31, (14-16)]\)

\[
K = \sum_{i=0}^{\infty} \frac{a_i(\alpha, \beta)}{P_i} \frac{g_i}{z_i}, \quad \frac{g_i}{z_i} = \frac{1}{1 + \frac{g_i}{z_i}}
\]

Algorithm 1 SNR Thresholds Searching

**Require:** SNR thresholds for channel-state model

**Input:**

- **FSO model:** frame duration \(T_f\) and \(c_f\) (frames/state)
- **RF model:** frame duration \(T_f\) and \(c_f\) (frames/state)

**Output:** \(\{\gamma_m\}_{m=1}^M\) for FSO and \(\{\gamma_n\}_{n=1}^N\) for RF models

1. **Initialization** \(\gamma_1 = 0, \gamma_1 = 0\) and \(m = n = 2\)
2. while \(m < M + 1\) do
3. Search \(\gamma_m\) that satisfies \(\gamma_m = c_f T_f \leq t_f\)
4. \(m = m + 1\)
5. end while
6. while \(n < N + 1\) do
7. Search \(\gamma_n\) that satisfies \(\gamma_n = c_f T_f \leq t_c\)
8. \(n = n + 1\)
9. end while
10. \(\gamma_{(M+1)} = \gamma_{(N+1)} = +\infty\)

\[
\times \left[ \Gamma \left( \frac{i + \beta}{2}, g_z \gamma_k \right) - \Gamma \left( \frac{i + \beta}{2}, g_z \gamma_{k+1} \right) \right], \quad (24)
\]

\[
L = \sum_{i=0}^{\infty} \frac{a_i(\beta, \alpha)}{P_i} \frac{g_i}{z_i}\]

\[
\times \left[ \Gamma \left( \frac{i + \alpha}{2}, g_z \gamma_k \right) - \Gamma \left( \frac{i + \alpha}{2}, g_z \gamma_{k+1} \right) \right], \quad (25)
\]

with \(g_z = \frac{3}{2(Z - 1)}\) and \(a_i(x, y) = \frac{(xy)^{i+y}}{\Gamma(i + x + y + 1)!}\)

**RF model:**

\[
\text{BER}_k = \frac{1}{P_k} \int_{\gamma_k}^{\gamma_{k+1}} \text{BER}(\gamma_k, \gamma_{k+1}) d\gamma = \frac{0.2}{P_k} \frac{m}{(m, b_n \gamma_{k+1})} - \frac{\text{BER}(\gamma)}{b_n \gamma_{k+1}}, \quad (27)
\]

where \(b_n = \frac{m}{P_k} + \frac{3}{2(Z - 1)}\) and \(\text{BER}(\gamma)\) is the instantaneous BER, which is well approximated as \([32, (28)]\)

\[
\text{BER}(\gamma) = 0.2 \exp \left[ -\frac{3\gamma}{2(Z - 1)} \right], \quad (28)
\]

where \(Z\) is the signal constellation size and \(Z = 2\) for the BPSK modulation.

**IV. TCP PERFORMANCE ANALYSIS**

This section focuses on the throughput analysis of TCP variants, including TCP NewReno, Hybla, HSTCP, and Cubic with SR-ARQ and RS code over the hybrid FSO/RF fading channel. The TCP throughput depends on the behavior of the congestion window and the average round-trip delay (RTT), which will be described as follows.
**A. TCP SEGMENT LOSS**

The TCP segment loss is assumed to be caused by two main factors: transmission losses at the last-mile link and congestion losses on the Internet.

1) TRANSMISSION LOSS

The transmission loss is caused by the errors on the last-mile link. Given $\text{BER}_{f}$, the RS coded frame error rate for each channel of FSO and RF links can be expressed as [33, (16)]

**FSO link:**

$$\text{FER}_{f_{m}} = 1 - \sum_{j=0}^{n_{0}} \left( \frac{L_{f}}{j} \right) (\text{BER}_{f_{m}})^{j} \left[ 1 - (\text{BER}_{f_{m}}) \right]^{L_{f}-j},$$  

(29)

**RF link:**

$$\text{FER}_{r_{n}} = 1 - \sum_{j=0}^{n_{0}} \left( \frac{L_{r}}{j} \right) (\text{BER}_{r_{n}})^{j} \left[ 1 - (\text{BER}_{r_{n}}) \right]^{L_{r}-j},$$  

(30)

where $L_{f}$ is the frame size, and $n_{0}$ is the maximum number of errors that can be corrected.

Based on the channel state modelling of hybrid FSO/RF channel, the average frame error rate over last-mile link can be given as [34, (28)]

$$\text{FER} = \sum_{m=1}^{M} R_{m} P_{m} \text{FER}_{m} + \sum_{n=1}^{N} R_{n} P_{n} \text{FER}_{n},$$  

(31)

where $\text{FER}_{m}$, $\text{FER}_{n}$, $R_{m}$, $P_{m}$, $R_{n}$, $P_{n}$ are the average frame error rate, steady-state probability of FSO link $m$-th state and RF link $n$-th state while $R_{t}$, $R_{c}$ are the transmission rate of each link, respectively.

Considering the operation of SR-ARQ, when a frame loss is detected, the sender keeps re-transmitting this frame; if this frame does not get through the networks after $X$ re-transmissions (i.e., there is no ACK), the corresponding TCP segment is assumed to be lost. Given a $\text{FER}$, the TCP transmission loss probability over the last-mile link can be obtained as

$$P_{t} = 1 - \left[ 1 - \text{FER}^{X+1} \right]^{n_{f}},$$  

(32)

where $X$ is the persistent level of SR-ARQ, and $n_{f}$ is the number of frames per segment.

2) CONGESTION LOSS

The congestion loss is caused by buffer overflow at routers on the Internet section. For the sake of simplicity, it is assumed that congestion happens when TCP window size attains the maximum window size $W_{\text{max}}$, which is approximately the sum of the buffer size $B$ of the Internet section and the average Bandwidth-Delay Product (BDP) of last-mile link so it can be given as [35, (2)]

$$W_{\text{max}} = B + \frac{\text{BDP}}{L_{\text{TCP}}} = B + \frac{\bar{R} \times \text{RTT}}{L_{\text{TCP}}},$$  

(33)

where $L_{\text{TCP}}$ is the TCP segment size while $\bar{R}$, $\text{RTT}$ are the average transmission rate and TCP round-trip time of the last-mile link that will be discussed later.

In addition, since the Slow Start (SS) period is very concise in comparison with the Congestion Avoidance (CA), the congestion loss probability can be well-approximated as [36, (2)] based on the process of a transmission cycle between two consecutive loss events

$$P_{c} = \frac{1}{N_{c}} \approx \begin{cases} \frac{8}{3 W_{\text{max}}^{2}} \rho^{2}, & \text{TCP NewReno}, \\ \frac{8}{3 W_{\text{max}}^{2}} b(w)^{2} W_{\text{max}}^{2}, & \text{TCP Hybla}, \\ \frac{b(w)^{2}}{C}, & \text{HTCP}, \\ \frac{4 \text{RTT}}{\beta W_{\text{max}}^{4}}, & \text{TCP Cubic}, \end{cases}$$  

(34)

where $N_{c}$ is the total number of segments that are transmitted in a cycle (as in Fig. 6), $W_{\text{max}}$ is the maximum congestion window size, $C$ and $\beta$ are the Cubic and multiplicative decrease parameters, $W = W_{\text{max}} \left[ 1 - b(W_{\text{max}}) \right]$ and $a(w), b(w)$ are the modified increase, decrease parameters of HSTCP, respectively, while $\rho$ is the normalized Hybla factor.

3) UNIFIED LOSS MODEL

This section derives the unified loss probability (ULP) under the impact of transmission and congestion losses. Figure 6 shows the TCP congestion window evolution when only congestion loss is considered (left) and when both transmission and congestion losses (unified loss) are considered, respectively.

Denote $N_{\text{max}}$ as the maximum number of segments that can be sent when a congestion loss occurs in the unified loss model. Based on the similar approach as [36], by approximating $N_{\text{max}}$ to a constant value that is equal to $N_{c}$, we can obtain the expression of the ULP ($\bar{P}_{\text{loss}}$) as

$$\bar{P}_{\text{loss}} = 1 - \frac{P_{t}}{1 - \exp \frac{P_{c}}{N_{c}}},$$  

(35)

where $\bar{N}$ is the average number of segments that can be sent during a cycle between two consecutive losses in the unified loss model while $P_{t}, P_{c}$ are the transmission loss and...
congestion loss probabilities, respectively. Eq. (35) can be validated as

\[
\lim_{P_t \to 0} P_{\text{loss}}' = \lim_{P_t \to 0} \frac{P_t}{1 - \exp \left( \frac{P_t}{P_c} \right)} = P_c,
\]
(36)

\[
\lim_{P_c \to 0} P_{\text{loss}}' = \lim_{P_c \to 0} \frac{P_t}{1 - \exp \left( \frac{P_t}{P_c} \right)} = P_t.
\]
(37)

**FIGURE 7.** Actual model and equivalent model.

As all losses are equally treated in the TCP, for the sake of simplicity, we can re-design an equivalent model with congestion losses only, as in Fig. 7. It is seen that, while the actual network model, which is characterized by its Unified Loss \(P_{\text{loss}}\), including both transmission loss \(P_t\) and congestion loss \(P_c\), the equivalent model can be described by congestion event only, which is represented as \(\bar{P}_c\). To equalize the two models, we have

\[
\bar{P}_{\text{loss}} = \bar{P}_c.
\]
(38)

Based on the analysis above, (38) can be given as

\[
\bar{P}_{\text{loss}} = \frac{1}{N} = \bar{P}_c = \frac{1}{N_c},
\]
(39)

where \(N, N_c\) is the average number of segments that can be sent during a cycle between two consecutive losses in the actual and equivalent model, respectively. By substituting (35) into (39), \(\bar{P}_c\) can be expressed as

\[
\bar{P}_c = \frac{1}{N_c} = \frac{P_t}{1 - \exp \left( \frac{P_t}{P_c} \right)} = \frac{P_t}{1 - \exp \left( \frac{P_t}{P_c} \right)}.
\]
(40)

where \(P_t, P_c\) are the transmission and congestion loss probability of the actual network model. With \(N_c\) is a function of \(W_{\text{max}}\) and \(N_c\) is a function of \(\bar{W}_{\text{max}}\) that are defined in (33), the formation of \(\bar{W}_{\text{max}}\) can be given as

\[
\bar{W}_{\text{max}} \approx \left\{ \begin{array}{ll}
\frac{2}{3} \sqrt{8 \frac{1 - \exp \left( -\frac{P_t W_{\text{max}}}{2} \right)}{P_t}}, & \text{TCP NewReno,} \\
\frac{2}{3} \sqrt{8 \frac{1 - \exp \left( -\frac{P_t W_{\text{max}}}{2} \right)}{P_t}}, & \text{TCP Hybla,} \\
\frac{2}{3} \sqrt{8 \frac{1 - \exp \left( -\frac{P_t W_{\text{max}}}{2} \right)}{P_t}}, & \text{HTCP,} \\
\frac{4}{3} \sqrt{1 - \exp \left( -\frac{P_t W_{\text{max}}}{2} \right)} P_t, & \text{TCP Cubic,} \\
\end{array} \right.
\]
(41)

where \(G = \frac{4 - \beta}{4RTT} \sqrt{\frac{\beta W_{\text{max}}^2}{C}}\) is the Cubic constant parameter while \(K = \left[ \frac{P_t W_{\text{max}}}{2} \right]^2 \) is the approximated parameter of HSTCP and \(\rho\) is the normalized Hybla factor.

**B. AVERAGE TCP ROUND-TRIP DELAY**

The average end-to-end-trip delay time for a TCP segment, denoted as \(\bar{RTT}\) can be approximated as

\[
\bar{RTT} = T_{\text{OD}} + T_{\text{LD}} = T_{\text{OD}} + \left( \frac{L_f}{\bar{R}} + 2T_{\text{prop}} \right)(\bar{X} + 1) \gamma_t, \tag{42}
\]

where \(T_{\text{OD}}\) is a fixed delay caused by other parts of the network, which is deterministic. \(T_{\text{LD}}\) is the average delay in the last-mile link that depends on the data frame re-transmissions occurred at the link-layer with \(T_{\text{prop}}\) is the propagation delay from satellite to vehicle, \(L_f\) is the frame size and \(\bar{R}\) is the average transmission rate of the last-mile link, which is given as [34, (31)]

\[
\bar{R} = \sum_{m=1}^{M} R_t P_{in} + \sum_{n=1}^{N} R_t P_{in}, \tag{43}
\]

Here, \(R_t, M, R_t, N\) are the transmission rate, number of channel-states of FSO and RF link, respectively. \(L_f/\bar{R}\) is therefore the average transmission time for one frame. In addition, \(\bar{X}\) is the average number of retransmissions for one frame and it can be obtained as [37, (22)]

\[
\bar{X} = \frac{\sum_{x=0}^{X} FE^{x}(1 - FE)}{1 - FE^{X+1}} = \frac{FE - (X + 1)FE^{X+1}}{1 - FE^{X+1}}, \tag{44}
\]

where \(X\) is the persistent level of SR-ARQ and FE is the frame loss probability of the last-mile link.

**C. TCP THROUGHPUT ANALYSIS**

1) LOSS-BASED TCP MODELLING

The TCP throughput performance can be evaluated by analyzing the TCP window variation between two adjacent losses. Specifically, the window behavior of TCP can be modeled as a Markov chain. By deriving the stationary distribution of the chain, the average throughput can then be obtained. In particular, the range of TCP window size \((0, W_{\text{max}}]\) is first separated into \((N + 1)\) intervals. The zero interval \(0\)-th is \((0, SST)\) (SST means Slow-Start threshold) denotes the Slow-Start phase, and other \(N\) intervals \(I\)-th to \(N\)-th in the Congestion Avoidance phase are equally divided in the range of \((SST, W_{\text{max}}]\), where \(W_{\text{max}}\) is the maximum TCP window size. Each interval represents to a Markov state.

The TCP window is said to be in the \(m\)-th state if its value belongs to the corresponding interval. In that case, the window size is considered the midpoint of the interval, denoted as \(z_m\). When TCP is in \(0\)-th state the TCP window size is \(\frac{SST}{2}\). If TCP is in state \(m\)-th of Congestion Avoidance
Phase, its window interval is \( (m - 1) \frac{W_{\text{max}} - \text{SST}}{N}, m \frac{W_{\text{max}} - \text{SST}}{N}, m \in \{1, 2, 3, \ldots, N\} \) and the TCP window size is equal to \( (m - 0.5) \frac{W_{\text{max}} - \text{SST}}{N}, m \in \{1, 2, 3, \ldots, N\} \).

We observe the Markov chain at the time instant when a window reduction occurs, which is indicated by Retransmission Time-out (RTO) or Triple Duplicate ACKs (TD ACKs). The transition between \( m \)-th and \( n \)-th state \((n \neq 0)\) corresponds to the state transition between two consecutive loss events, as illustrated in Fig. 8. When a window reduction event happens, the window size reduces to \( \beta \) times of its size just before the event \( (z_m \rightarrow \beta z_m) \). There are two possible cases: if the value of reduction window size is below the Slow-Start threshold \((\beta z_m < \text{SST})\), TCP window grows from \( \beta z_m \) to SST by Slow Start growth function, after that with Congestion Avoidance growth function, it grows from SST to \( z_n \) without encountering another loss event (as in Fig. 8(a)). Otherwise, if \((\beta z_m \geq \text{SST})\), the window size grows from \( \beta z_m \) to \( z_n \), according to Congestion Avoidance growth function (see Fig. 8).

\[ P_{m,n} = \begin{cases} 1 - P_{\text{TO}(z_m)} & \text{TCP NewReno,} \\ \frac{2}{W_{\text{max}} \text{RTT}} & \text{TCP Hybla,} \\ \frac{2 \rho}{W_{\text{max}} \text{RTT}} & \text{HTCP,} \\ \frac{2 \rho}{W_{\text{max}} \text{RTT}} & \text{TCP Cubic,} \end{cases} \]

where \( T_{\text{loss}} \) is the average time between two consecutive losses. In the equivalent model, it is the duration when the TCP window size increases from \( \beta W_{\text{max}} \) to \( W_{\text{max}} \) (as in Fig. 7). \( W_{\text{max}} \) is the maximum window size of equivalent model, while RTT is the average RTT of last-mile link, defined in (42). \( C, \beta \) are the Cubic and multiplicative decrease parameters, \( \rho \) is the normalized Hybla factor and \( a(w), b(w) \) are the modified increase and decrease parameters of HSTCP with \( W = \left[1 - b(W_{\text{max}})\right] W_{\text{max}} \).

From (45), (46), (48), with \((N + 1)\) is the total number of states, the complete transition probability matrix of the Markov chain \( P = \left[P_{m,n}\right]_{(N+1) \times (N+1)} \) can be obtained. Denote \( \pi = [\pi_0, \pi_1, \ldots, \pi_N] \) as the matrix of steady-state probabilities, where \( \pi_m \) is the probability of the \( m \)-th state. Following the Markov chain theory, we can derive \( \pi_m \) by solving

\[ \pi = \pi \times P, \quad \sum_{m=0}^{N} \pi_m = 1. \]  

2) TCP THROUGHPUT

Given \( Z(t) \) is the total amount of data transmitted during time \( t \), by applying the renewal reward theorem, the average TCP throughput can be expressed as

\[ \bar{\eta} = \lim_{t \to \infty} \frac{Z(t)}{t} = \frac{\sum_{m=0}^{N} \sum_{n=0}^{N} \pi_m P_{m,n} A_{m,n}}{\sum_{m=0}^{N} \sum_{n=0}^{N} \pi_m P_{m,n} T_{m,n}}, \]
where $P_{m,n}$, $T_{m,n}$ and $A_{m,n}$ are the transition probability, average time duration and total transmission data during the transition between $m$-th and $n$-th state, respectively.

As shown in Fig. 8 and the analysis above, the average time duration between $m$-th and $n$-th state, $T_{m,n}$, can be given as

$$T_{m,n} = \begin{cases} D_{CA}(z_m, z_n), & \beta z_m \geq \text{SST}, \\ D_{SS}(z_m, \text{SST}), & \beta z_m < \text{SST}, \end{cases}$$

where RTO is the re-transmission time-out, $D_{SS}(x, y)$ is the time duration when the window size grows from $x$ to $y$ in the Slow-Start phase. For different TCP variants: TCP NewReno, TCP Hybla, HSTCP, and TCP Cubic, according to (1), (2), (3), and (4), it can be obtained as

$$D_{SS}(x, y) = \frac{\text{RTT}}{\rho} \ln \frac{y}{x}, \quad \text{for TCP Hybla},$$

$$D_{SS}(x, y) = \frac{\text{RTT}}{\rho} \ln \frac{y}{x}, \quad \text{otherwise},$$

where $\text{RTT}$ is the average RTT of the last-mile link, and $\rho$ is the normalized Hybla factor.

$D_{CA}(x, y)$ is the time duration when the window size grows from $x$ to $y$ in the Congestion Avoidance phase, and it is given as

$$D_{CA}(x, y) = \begin{cases} \text{RTT}(y - x), & \text{TCP NewReno}, \\ \text{RTT}(y - x), & \text{TCP Hybla}, \\ \frac{\text{RTT}(y - x)}{\alpha^2(1)}, & \text{HSTCP}, \\ \frac{\sqrt{\frac{y - x}{\rho}}}{\sqrt{\frac{1}{\beta}}} + \frac{\sqrt{\frac{\beta y}{\rho(1 - \beta)}}}{\sqrt{\frac{1}{\beta}}}, & \text{TCP Cubic}, \end{cases}$$

where $C$, $\beta$ are the Cubic and multiplicative decrease parameters, $\alpha(w)$ is the modified increase parameter of HSTCP, and $\rho$ is the normalized Hybla factor.

Let $D_1 = D_{CA}(z_m, z_n)$, $D_2 = D_{SS}(\beta z_m, \text{SST})$ and $D_3 = D_{CA}(\text{SST}, z_n)$, the total transmission data during the transition between $m$-th and $n$-th state, $A_{m,n}$, can be expressed as

$$A_{m,n} = \int_0^{T_{m,n}} \frac{W(z_m, T_{m,n})}{\text{RTT}} \, d\tau$$

where $W(x, \tau)$ is the TCP window function, including $W_{SS}$ and $W_{CA}$ in the Slow-Start and Congestion Avoidance phases, respectively.

For different TCP variants, $A_{m,n}$ is determined as

If $\beta z_m \geq \text{SST}$:

$$A_{m,n}^{(1)} = \begin{cases} \frac{1}{\text{RTT}} \int_0^{D_1} W_{CA}(z_m, D_1) \, d\tau, & \text{TCP NewReno}, \\ \frac{1}{\text{RTT}} \int_0^{D_1} W_{SS}(D_2) \, d\tau, & \text{TCP Hybla}, \\ \frac{1}{\text{RTT}} \int_0^{D_1} W_{SS}(D_2) \, d\tau, & \text{HSTCP}, \\ \frac{1}{\text{RTT}} \int_0^{D_1} W_{SS}(D_2) \, d\tau, & \text{TCP Cubic}, \end{cases}$$

If $\beta z_m < \text{SST}$:

$$A_{m,n}^{(2)} = \begin{cases} \frac{1}{\text{RTT}} \left( \frac{2D_2}{\ln 2} \right) + A_{m,n}^{(1)} |D_2|, & \text{TCP NewReno}, \\ \frac{1}{\text{RTT}} \left( \frac{2D_2}{\ln 2} \right) + A_{m,n}^{(1)} |D_2|, & \text{TCP Hybla}, \\ \frac{1}{\text{RTT}} \left( \frac{2D_2}{\ln 2} \right) + A_{m,n}^{(1)} |D_2|, & \text{HSTCP}, \\ \frac{1}{\text{RTT}} \left( \frac{2D_2}{\ln 2} \right) + A_{m,n}^{(1)} |D_2|, & \text{TCP Cubic}, \end{cases}$$

where $L = \frac{\beta \alpha(y)}{\rho C}$ with $C$, $\beta$ are the Cubic and multiplicative decrease parameters, $\alpha(w)$, $\beta(w)$ are the modified increase and decrease parameters of HSTCP, respectively, and $\rho$ is the normalized Hybla factor.

If Time-out event occurs: $A_{m,n} = 0$.

V. NUMERICAL RESULTS AND DISCUSSIONS

In this Section, we present numerical results and comparatively discuss the performance of TCP variants with different parameters/settings in the scenario of satellite-assisted hybrid FSO/RF vehicular networks. Monte Carlo simulations are also performed to validate the analytical model.

The parameters used in the analysis, unless otherwise noted, are as follows. For the LEO satellite: satellite altitude $H = 600$ km, zenith angle $\xi = 55^\circ$, RF frequency carrier $f_c = 2.4$ GHz, optical wavelength $\lambda = 1550$ nm, FSO bit rate $R_t = 1$ Gbps, RF bit rate $R_e = 150$ Mbps. Regarding the vehicular: car’s height $h_c = 1.5$ m, optical aperture diameter $D_t = 10$ cm, car’s speed $v_c = 50$ km/h. As for the ARQ and FEC: SR-ARQ’s persistent level $X = 2$, RS code RS(255, 223) with $n = 16$ and data frame size $L_f = 223$ bytes. Other parameters used in this paper: Nakagami factor $m = 1.5$, average received SNR of the RF link $\gamma_r = 15$ dB, rms wind speed $v_{\text{wind}} = 21$ m/s, $c_f = 2000$ frames/slot for FSO, $c_t = 2$.
frames/slot for RF, \( n_f = 5 \) frames/segment. In addition, the parameters related to TCP variants are given in Table 2.

| Parameters               | Values                      |
|--------------------------|----------------------------|
| TCP segment size         | \( L_{TCP} = 1155 \) bytes |
| TCP Cubic factor         | \( C_f = 4 \) Mbps         |
| Multiplicative decrease factor | \( \beta = 0.2 \)         |
| Hybla RTT reference      | RTT_{H} = 25 ms            |
| HSTCP high-window        | HS_f = 83000 segments      |
| HSTCP low-window         | HS_t = 38 segments         |
| HSTCP High-P             | HS_P = 10^{-7}            |
| HSTCP High decrease      | HS_D = 0.1                |
| Slow-Start threshold     | SST = 1 Mb                |
| Retransmission Timeout   | RTO = 150 ms              |
| Delay of other parts     | T_{OD} = 30 ms            |
| Buffer size              | \( B = 1000 \) segments   |

Firstly, we focus on analyzing the effectiveness of the link-layer error-control design, including SR-ARQ and error correction, for the improvement of TCP throughput performance in this network. Fig. 10 depicts the TCP Cubic performance comparison of different SR-ARQ persistent levels on the condition with or without the RS code. Obviously, a significant throughput improvement is observed when SR-ARQ is employed. For instance, in Fig. 10(b), when \( \gamma_f = 20 \) dB with the persistent level \( X = 2 \), TCP Cubic throughput can achieve 850 Mbps, while when \( X = 1 \) the achieved throughput is only 300 Mbps. When SR-ARQ is not employed, throughput drops to almost zero. This emphasizes the importance of the SR-ARQ protocol by mitigating the impact of frame loss and improve the TCP performance. On the other hand, while the higher persistent level of SR-ARQ may provide higher throughput, it may increase the overall delay in the network, which is sensitive to LEO satellite-to-vehicle communications design. Therefore, the selection of a reasonable persistent level \( X \) is necessary. In addition, the RS code also holds an important role in improving the effectiveness of this network. Fig. 10(a) and Fig. 10(b) confirm the effect of RS on the throughput performance. Given \( \gamma_f = 15 \) and \( X = 2 \), with RS code, TCP throughput can achieve 400 Mbps while it achieves under 100 Mbps in the uncoded system.

Next, we investigate the impact of segment size on the TCP performance over the last-mile access so that optimal segment sizes can be determined to maximize the system throughput at a certain atmospheric turbulence condition. Fig. 11 depicts the relationship between segment size and the throughput of several TCP variants under the different turbulence conditions, weak, moderate, and strong, respectively. As seen in Fig. 11(a), when the turbulence is weak, which is equivalent to low transmission loss probability in this network, larger segment sizes do not lead to a significant increase in the segment loss probability. Better throughput performance, therefore, can be obtained. For example, in Fig. 11(a), when the segment size is 500 bytes, TCP Cubic only achieves 800 Mbps, while when it is 1500 bytes, the throughput is increased to 900 Mbps. Nevertheless, when the turbulence gets stronger, increasing segment size does not mean better throughput performance. In Fig. 11(b), throughput improvement can only be seen in TCP Cubic, HSTCP when the segment size increases from 500 to 1000 bytes. At the same time, the throughput performance degradation is severe in TCP Hybla and TCP NewReno. With TCP Hybla, when the segment size is 500 bytes, the throughput is 500 Mbps. However, TCP Hybla can only reach less than 100 Mbps when the segment size is 1500 bytes.

The throughput performance deterioration of these TCP variants becomes more severe in the case turbulence condition is strong, as shown in Fig. 11(c). This phenomenon is logical as when the turbulence becomes moderate and strong, it increases the transmission loss probability. Therefore, with
a larger segment size, the segment loss probability tends to increase. Based on this analysis for this figure, optimal segment sizes can be determined to achieve the maximum throughput in a particular turbulence condition. For example, the segment size should be set to 1000 bytes for TCP Cubic and HSTCP when the turbulence strength is moderate.

Another essential issue of this network is the impact of the zenith angle on the TCP throughput performance. Fig. 12 illustrates the relationship between the throughput of TCP Cubic and zenith angle with different satellite altitudes and turbulence conditions. As is evident, the throughput gets higher when the zenith angle is low because, in the LEO satellite-to-vehicle optical link, the turbulence strength is weaker in the case of small zenith angles. Obviously, the increase of satellite altitude causes a longer transmission duration, and it exceeds the channel coherence time, leading to the degradation of the throughput performance.

We now comparatively discuss the performance of TCP variants with different parameters/settings in the scenario of satellite-assisted hybrid FSO/RF vehicular networks. Fig. 13 shows the comparison between TCP variants in terms of average throughput for a range of average received SNR of the FSO link. First, it is seen that, overall, TCP Cubic achieves the best throughput performance, while NewReno has the worst performance compared to HSTCP and TCP Hybla. The performance difference can be explained by the way that TCP protocols react when the loss events happen. For instance, different from HSTPC, Hybla, and NewReno, the window growth function of TCP Cubic is defined in real-time, independent from RTT, which can be extended in LEO satellite links. In addition, because Cubic increases its window size slowly when it is close to the previous saturation point (i.e., the last loss event), it makes Cubic’s performance less efficient when the losses frequently occur (i.e., low SNR). With HSTCP, its performance is better because the modified congestion control mechanism is flexible with increase and decrease parameters.

Finally, the performance of TCP variants is analyzed under the impact of turbulence conditions on the LEO satellite-to-vehicle channel at the last-mile link. Fig. 14 depicts the average TCP throughput versus atmospheric turbulence strength, represented as ground turbulence levels $C_{n}^2(0)$. It is clear that the strong turbulence significantly deteriorates the TCP performance, especially for the Hybla and NewReno. For example, TCP Cubic can achieve the maximum throughput for the turbulence range from $C_{n}^2(0) = 5 \times 10^{-15} \text{m}^{-2/3}$ to $C_{n}^2(0) = 7 \times 10^{-14} \text{m}^{-2/3}$ while with TCP Hybla, it is only from $C_{n}^2(0) = 5 \times 10^{-15} \text{m}^{-2/3}$ to $C_{n}^2(0) = 10^{-14} \text{m}^{-2/3}$. In another example, when $C_{n}^2(0) = 3 \times 10^{-13} \text{m}^{-2/3}$ as strong turbulence condition, TCP Cubic can achieve 400 Mbps, whereas TCP Hybla achieve only 50 Mbps. This throughput...
degradation by the effect of atmospheric turbulence is caused by the fact that when the turbulence strength increases, transmission losses happen more and more and reduces the TCP sending rate, which leads to performance degradation.

VI. CONCLUSION
We presented a cross-layer analysis for TCP performance over LEO satellite-assisted hybrid FSO/RF vehicular networks. A truncated SR-ARQ and RS code were employed to improve the link and physical layer performance, respectively. Using the proposed channel model and error-control method, a comprehensive end-to-end throughput model for several TCP variants, including NewReno, Hybla, Cubic, and HSTCP under the impact of both congestion losses in the Internet section and LEO satellite-based Hybrid FSO/RF transmission errors at the last-mile link, was developed. We quantitatively illustrated the impact of transmission errors at last-mile links, comparatively discussed throughput performance of TCP variants, and the selection of proper TCP for our networks. The results also revealed that Cubic is most suitable for the considered networks among the four TCP variants. The HSTCP and Hybla could be used; however, NewReno was inefficient in the high-speed satellite networks, especially in strong turbulence conditions. Finally, the Monte Carlo simulations were also performed to validate the correctness of the analysis model.

REFERENCES
[1] H. Kaushal and G. Kaddoum, “Optical communication in space: Challenges and mitigation techniques,” IEEE Commun. Surveys Tuts., vol. 19, no. 1, pp. 57–96, 1st Quart., 2017.
[2] H. Le, T. Nguyen, and A. Pham, “Cloud attenuation statistical model for satellite-based FSO communications,” IEEE Antennas Wireless Propag. Lett., vol. 20, no. 5, pp. 643–647, May 2021.
[3] J. F. Kurose and K. W. Ross, Computer Networking: A Top-Down Approach, 6th ed. London, U.K.: Pearson, 2012.
[4] T. De Cola and M. Marchese, “Performance analysis of data transfer protocols over space communications,” IEEE Trans. Aerosp. Electron. Syst., vol. 41, no. 4, pp. 1200–1223, Oct. 2005.
[5] J. Zhu, S. Roy, and J. H. Kim, “Performance modelling of TCP enhancements in terrestrial-satellite hybrid networks,” IEEJ/ACM Trans. Netw., vol. 14, no. 4, pp. 753–766, Aug. 2006.
[6] Y. Sun, Z. Ji, and H. Wang, “TFRC-satellite: A TFRC variant with a loss differentiation algorithm for satellite networks,” IEEE Trans. Aerosp. Electron. Syst., vol. 49, no. 2, pp. 716–725, Apr. 2013.
[7] X. Li and J. Wang, “Elastically reliable video transport protocol over lossy satellite links,” IEEE J. Sel. Areas Commun., vol. 36, no. 5, pp. 1097–1108, May 2018.
[8] M. Bacco, T. De Cola, G. Giambene, and A. Gotta, “TCP-based M2M traffic via random-access satellite links: Throughput estimation,” IEEE Trans. Aerosp. Electron. Syst., vol. 55, no. 2, pp. 846–863, Apr. 2019.
[9] S. R. Pokhrel, J. Jin, and H. L. Vu, “Mobility-aware multipath communication for unmanned aerial surveillance systems,” IEEE Trans. Veh. Technol., vol. 68, no. 6, pp. 6088–6098, Jun. 2019.
[10] G. Giambene, D. K. Luong, T. De Cola, V. A. Le, and M. Muhammad, “Analysis of a packet-level block coding approach for terrestrial-satellite mobile systems,” IEEE Trans. Veh. Technol., vol. 68, no. 9, pp. 8117–8132, Aug. 2019.
[11] J. Liu, Z. Han, and W. Li, “Performance analysis of TCP new Reno over satellite DVB-RCS2 random access links,” IEEE Trans. Wireless Commun., vol. 19, no. 1, pp. 435–446, Oct. 2020.
[12] Y. Hasegawa, “A transmission control protocol for free-space optical communications,” in Proc. GLOBECOM IEEE Global Commun. Conf., Dec. 2017, pp. 1–7.
[37] V. V. Mai, T. C. Thang, and A. T. Pham, “Performance of TCP over free-space optical atmospheric turbulence channels,” IEEE/OSA J. Opt. Commun. Netw., vol. 5, no. 11, pp. 1168–1177, Dec. 2013.

[38] S. Hassayoun, P. Maille, and D. Ros, “On the impact of random losses on TCP performance in coded wireless mesh networks,” in Proc. IEEE INFOCOM, Mar. 2010, pp. 1–9.

[39] W. Bao, V. W. S. Wong, and V. C. M. Leung, “A model for steady state throughput of TCP CUBIC,” in Proc. IEEE Global Telecommun. Conf. (GLOBECOM), Dec. 2010, pp. 1–6.

[40] S. R. Pokhrel, H. L. Vu, and A. L. Cricenti, “Adaptive admission control for IoT applications in home WiFi networks,” IEEE Trans. Mobile Comput., vol. 19, no. 12, pp. 2731–2742, Dec. 2020.

THANG K. NGUYEN is currently pursuing the B.E. degree in electronics and telecommunications with Hanoi University of Science and Technology (HUST), Vietnam. He is also working as a Research Assistant with the Communications Theory and Applications Research Group (CTARG), HUST. His research interests include optical wireless communications, hybrid FSO/RF for satellite networks, and transmission control protocols.

CHUYEN T. NGUYEN received the B.E. degree in electronics and telecommunications from Hanoi University of Science and Technology (HUST), Vietnam, in 2006, the M.S. degree in communications engineering from the National Tsing Hua University, Taiwan, in 2008, and the Ph.D. degree in informatics from Kyoto University, Japan, in 2013. From September to November 2014, he was a Visiting Researcher with The University of Aizu, Japan. He is currently an Associate Professor with the School of Electronics and Telecommunications, HUST. His current research interests include in areas of MAC protocol design and reliable transmission in wireless/optical networks. He received the Fellow Award from Hitachi Global Foundation, in August 2016; the First Best Paper Award in 2019 IEEE ICT; and the Best Paper Awards in 2018 KICS/IEEE ICTC and 2019 IEEE ICC.

HOANG D. LE (Graduate Student Member, IEEE) received the B.E. degree in electronics and communications from Hanoi University of Science and Technology (HUST), Vietnam, in 2016, and the M.E. degree in computer science and engineering from The University of Aizu (UoA), Japan, in 2018. He is currently pursuing the Ph.D. degree with the Computer Communications Laboratory, UoA. His current research interests include optical wireless communications, hybrid FSO/RF for satellite/UEV networks, cross-layer design, and error-control protocols. He was a recipient of a number of awards, including the 2018 KICS/IEEE ICTC Best Paper Award, the 2019 IEEE VTS Tokyo Chapter Young Researcher Encouragement Award, and the 2019 IEEE ComSoc Sendai Chapter Student Excellent Research Award.

ANH T. PHAM (Senior Member, IEEE) received the B.E. and M.E. degrees in electronics engineering from Hanoi University of Technology, Vietnam, in 1997 and 2000, respectively, and the Ph.D. degree in information and mathematical sciences from Saitama University, Japan, in 2005. From 1998 to 2002, he was with NTT Corporation, Vietnam. Since 2005, he has been a Faculty Member with The University of Aizu, where he is currently a Professor and the Head of the Computer Communications Laboratory, Division of Computer Engineering. He has authored/coauthored over 160 peer-reviewed articles on these topics. His research interests include the broad areas of communication theory and networking with a particular emphasis on modeling, design, and performance evaluation of wired/wireless communication systems and networks. He is a member of IEICE and OSA.