




# Predicting atmospheric turbulence for secure quantum communications in free space

TAREQ JAOUNI,<sup>1,†</sup>  LUKAS SCARFE,<sup>1,†</sup> FRÉDÉRIC BOUCHARD,<sup>2</sup>  
MARIO KRENN,<sup>3</sup> KHABAT HESHAMI,<sup>1,2</sup>  
FRANCESCO DI COLANDREA,<sup>1,4,\*</sup>  AND EBRAHIM KARIMI<sup>1,2,3</sup> 

<sup>1</sup>Nexus for Quantum Technologies, University of Ottawa, Ottawa ON, K1N 5N6, Canada

<sup>2</sup>National Research Council of Canada, 100 Sussex Drive, Ottawa ON, K1A 0R6, Canada

<sup>3</sup>Max Planck Institute for the Science of Light, Staudtstrasse 2, 91058 Erlangen, Germany

<sup>4</sup>Dipartimento di Fisica, Università di Napoli Federico II, Complesso Universitario di Monte Sant'Angelo, Via Cintia, 80126 Napoli, Italy

<sup>†</sup>The authors contributed equally to this work.

\*francesco.dicolandrea@unina.it

**Abstract:** Atmospheric turbulence is the main barrier to large-scale free-space quantum communication networks. Aberrations distort optical information carriers, thus limiting or preventing the possibility of establishing a secure link between two parties. For this reason, forecasting the turbulence strength within an optical channel is highly desirable, as it allows for knowing the optimal timing to establish a secure link in advance. Here, we train a recurrent neural network, TAROQO, to predict the turbulence strength within a free-space channel. The training is based on weather and turbulence data collected over 9 months for a 5.4 km intra-city free-space link across the City of Ottawa. The implications of accurate predictions from our network are demonstrated in a simulated high-dimensional quantum key distribution protocol based on orbital angular momentum states of light across different turbulence regimes. TAROQO will be crucial in validating a free-space channel to optimally route the key exchange for secure communications in real experimental scenarios.

© 2025 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

## 1. Introduction

The ability to forecast weather conditions is a relevant achievement in modern society, impacting agriculture [1], energy management [2], climate science [3], and public health [4]. Traditional forecasts involve tangible parameters such as temperature, wind speed, and precipitation, while the complex phenomenon of atmospheric turbulence has mainly been restricted to astronomical observations [5], and only recently extended to free-space communications [6,7]. Atmospheric turbulence originates from the rapid variation in time of different weather parameters, which results in optical signals experiencing a continuously varying refractive index when traveling through the atmosphere [8]. This severely affects the reliability of optical experiments realized in “free space”, which has led to the dominance of fibre-based communications networks. Scintillation of low-power light, beam wandering, and wavefront distortion are representative examples of common optical aberrations, globally quantified by the structure parameter  $C_n^2$  [9].

Atmospheric turbulence significantly limits optical communications, both in the classical [10] and quantum [11] regime. Typical effects are high crosstalk and lower power transmission. For these reasons, adaptive-optics correction systems have been proposed and experimentally demonstrated, showing promising results for free-space quantum key distribution (QKD) in turbulent channels [12–14]. These achievements set the baseline for future ground-to-ground [15] and ground-to-satellite [16] configurations. In the broader context of quantum networks, predicting the trend of atmospheric turbulence is key to establishing a secure connection between

two parties, leading to informed decisions about the future usage of available channels within the network.

Based on historical weather data, standard approaches to forecasting atmospheric turbulence relied upon empirical models [17], remote sensing [18] and, more recently, artificial neural networks [19–22]. In particular, Grose and Watson explored the application of different instances of recurrent neural networks (RNNs) to predict  $C_n^2$  values up to a few hours in advance [20]. These architectures are specifically designed to model sequential data, which makes them suitable candidates for capturing the temporal dependencies within the input weather data in the form of a time series. Although excellent levels of accuracy have been reported in this first case study, the final implementation suffered from relevant limitations, most importantly the lack of real night-time data, a training dataset only spanning a single season, and a rigid fixed-time prediction in the future [20].

In this paper, we train TAROQO [23], an RNN processing dataset encompassing 9 months of weather data with a minute-by-minute time resolution to forecast  $C_n^2$  values within a 5.4 km intra-city channel over the City of Ottawa. Our scheme outputs predictions within a flexible number of hours in the future (up to 12 hours) and with a custom time resolution (down to one minute). This computational toolbox provides the preliminary validation step of a turbulent channel for near-term QKD experiments. This analysis is complemented with the numerical simulation of a high-dimensional BB84 QKD protocol employing spatial modes of light under different turbulent regimes. The results of the latter are used to ultimately infer the turbulence threshold under which the actual experiment can be conducted. This enables informed decisions regarding the feasibility of a real key exchange, thereby optimizing the use of time, resources, and costs, by avoiding running experiments under unfavorable conditions.

Our findings are useful not only for QKD applications and, in principle, all free-space experiments utilizing structured light can benefit from this tool, since knowing the turbulence strength in advance would allow for predicting their success rate [24–27]. The same analysis also applies to nearby satellite ground stations, which would allow for determining the optimal channel for maximum key exchange on a given satellite pass.

## 2. Theory

### 2.1. Atmospheric turbulence

Atmospheric turbulence can be quantified in terms of the structure parameter of the refractive index  $C_n^2$ . The  $C_n^2$  is defined as the variance of the refractive index over a given optical path, normalized to the path length [28]:

$$C_n^2 = \frac{\overline{(n(\vec{x}) - n(\vec{x} + \vec{r}))^2}}{r^{\frac{2}{3}}} \approx \frac{\text{var}(n)}{r^{\frac{2}{3}}}, \quad (1)$$

where  $n$ ,  $\vec{x}$ , and  $\vec{r}$  are the refractive index of the atmosphere, the position along the path, and the distance from the sender to the receiver, respectively, with the average taken over all positions within the path.

The effect of atmospheric turbulence on optical beams can be modeled as a random sequence of phase masks dislocated along the path. By randomizing the phase screens in multiple realizations, it is possible to retrieve the average effect of the turbulence on the input beam with good approximation [29]. Zernike modes, which are traditionally employed to model optical aberrations [30], provide a natural choice to generate the phase masks. In particular, the value of  $C_n^2$  can be directly related to the variance of each Zernike mode [31].

To measure the value of  $C_n^2$  in our channel, we employ scintillometry [32]. To implement this, an incoherent light source is placed at one end of the channel, while a large-aperture telescope with a photodiode is installed on the other end. The intensity of the light from the emitter side is

measured continuously for one minute at a time. The statistics of the intensity measurements, specifically the variance and the logarithmic amplitude, are used to extract the  $C_n^2$  [28,33].

## 2.2. Quantum key distribution

Quantum key distribution (establishment) is a technique introduced by Bennett and Brassard with the well-known BB84 protocol [34], which utilizes the principles of quantum mechanics to enable provably secure communications between two parties, by establishing a shared secret key. Security is guaranteed when the error in transmission and measurement of the quantum information is below a given threshold, depending on the dimension of the protocol [35]. Extending QKD to high-dimensional states allows for higher information capacity (typically expressed in “bits per photon”), higher tolerance to errors, and innovative protocols [36].

In a  $d$ -dimensional error-free QKD protocol, the number of informational bits per photon is  $R(d) = \log_2(d)$ . Turbulence-induced errors within a free-space channel result in a diminished key rate. For a high-dimensional BB84 protocol [37]:

$$R(d, e_q) = \log_2(d) - 2h(e_q), \quad (2)$$

where  $h(x) = -x \log_2(x/(d-1)) - (1-x) \log_2(1-x)$  is the Shannon entropy, and  $e_q$  is the average error rate of the states sent and detected. When  $R(d, e_q) < 0$ , security is not guaranteed.

It must be noted that QKD protocols are not used to send messages, but only to establish a shared key through which classical communications can be securely encrypted. These keys can also be generated and stored for later use, even if the quantum channel is incapable of generating new keys [38].

Further reading on the practical implementations of QKD can be found in Ref. [39].

## 3. Methodology

### 3.1. Data collection and preparation

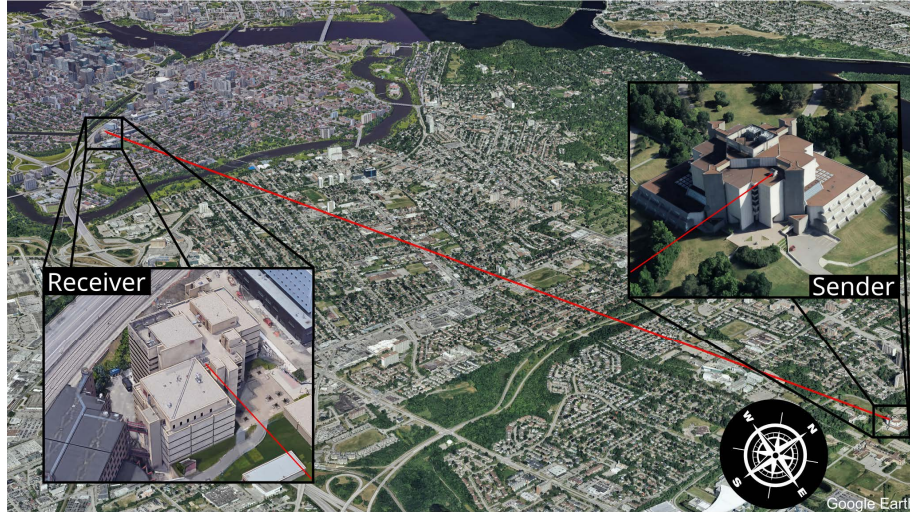
We employ the Scintech BLS450 large aperture scintillometer to record the value of  $C_n^2$  over our channel [40]. The system has been active 24/7 measuring the turbulence of the channel for 9 months. An aerial view of the monitored channel is provided in Fig. 1. The  $C_n^2$  values have been recorded once per minute. Fog and heavy snowfall can cause system outages, corresponding to missing data points in our training set. To remove spurious information deriving from high-frequency noise, a moving average is performed for each data point over a one-hour time window.

Other weather parameters, such as temperature, solar radiation, and humidity, have been obtained from Environment and Climate Change Canada [41]. The weather station at which the parameters are measured lies 5 km south-west from the receiver. For these parameters, no additional processing or filtering was applied. The averaged  $C_n^2$  values are time correlated with the other meteorological data, and finally separated into batches for training the RNN.

The input layer of the network consists of 12 hours of minute-by-minute weather data, with  $n$  hours of future  $C_n^2$  values as the target output. The dataset is divided into a training part where the network learns, a validation dataset to provide feedback on the network’s training, and a test dataset to evaluate the network’s performance on unseen data. All input and output features are normalized according to:

$$X \rightarrow \frac{X - \min(X_T)}{\max(X_T) - \min(X_T)}, \quad (3)$$

where  $\min(X_T)$  and  $\max(X_T)$  denote, respectively, the minimum and maximum of the feature  $X$  over the training dataset.



**Fig. 1. Channel path.** The path over which the turbulence data was collected from July 2023 to March 2024. The receiver is on the University of Ottawa campus, while the sender is on the Canadian National Research Council 5.4 km East-North-East. This configuration was chosen to minimize the background levels caused by the sunlight.

### 3.2. Recurrent neural network

We train a gated recurrent unit (GRU) neural network, named TAROQQO, to forecast the evolution of the  $C_n^2$  values over time. GRUs are a simplified derivative of the long short-term memory (LSTM) recurrent units, which capture long-term features of the data. Figure 2 illustrates our neural network architecture, which takes 12 hours of prior combined meteorological and scintillometer data as input and outputs the  $C_n^2$  forecast over 6 hours. Following the permutation feature importance method detailed in Appendix A, temperature ( $^{\circ}\text{C}$ ), solar radiation ( $\text{kJ}/\text{m}^2$ ), relative humidity (%),  $\log_{10} C_n^2$ , and the UTC time (s) are used as input features. The time feature  $t$  is made periodic over one day with the following map:

$$\begin{aligned} t_x &= \cos\left(\frac{2\pi t}{T}\right), \\ t_y &= \sin\left(\frac{2\pi t}{T}\right), \end{aligned} \quad (4)$$

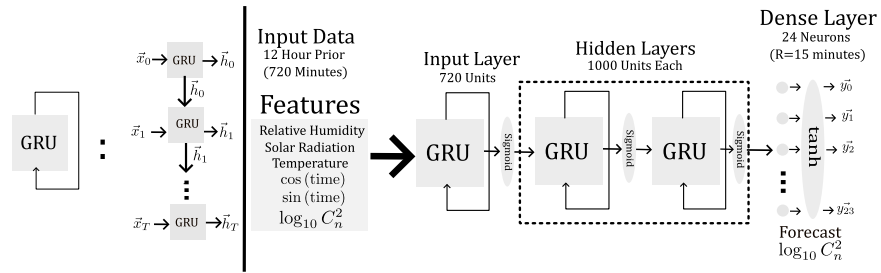
where  $T = 86400$  s. In particular, this transformation induces a smooth transition from the end of one day to the beginning of another, which also reflects how our physical quantities change.

Our network outputs the turbulence forecast through a fully connected dense layer that takes as input the final hidden state of the incident GRU layer. The number of output neurons is given by  $N_{\text{out}} = H/R$ , where  $H$  is the desired number of hours in the future (6 hours) and  $R$  is the desired output time resolution (15 minutes).

The cost function used for training is the mean squared error (MSE):

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N |\vec{y}_{i,\text{pred}} - \vec{y}_{i,\text{true}}|^2, \quad (5)$$

where  $\vec{y}_{\text{pred}}$  and  $\vec{y}_{\text{true}}$  denote the predicted and true output  $C_n^2$  forecast, and  $N$  is the size of the evaluated dataset. The complete list of hyperparameters used for training is reported in Appendix B.



**Fig. 2. Neural network architecture.** The network processes 12 hours of input data. The data is fed through a series of gated recurrent unit (GRU) layers, allowing the network to learn long-term trends through time in the input series. The hidden state variable  $\vec{h}$  is standardized using nonlinear activation layers at each step. The network outputs the forecast in a single shot through the final fully connected dense layer.

## 4. Results

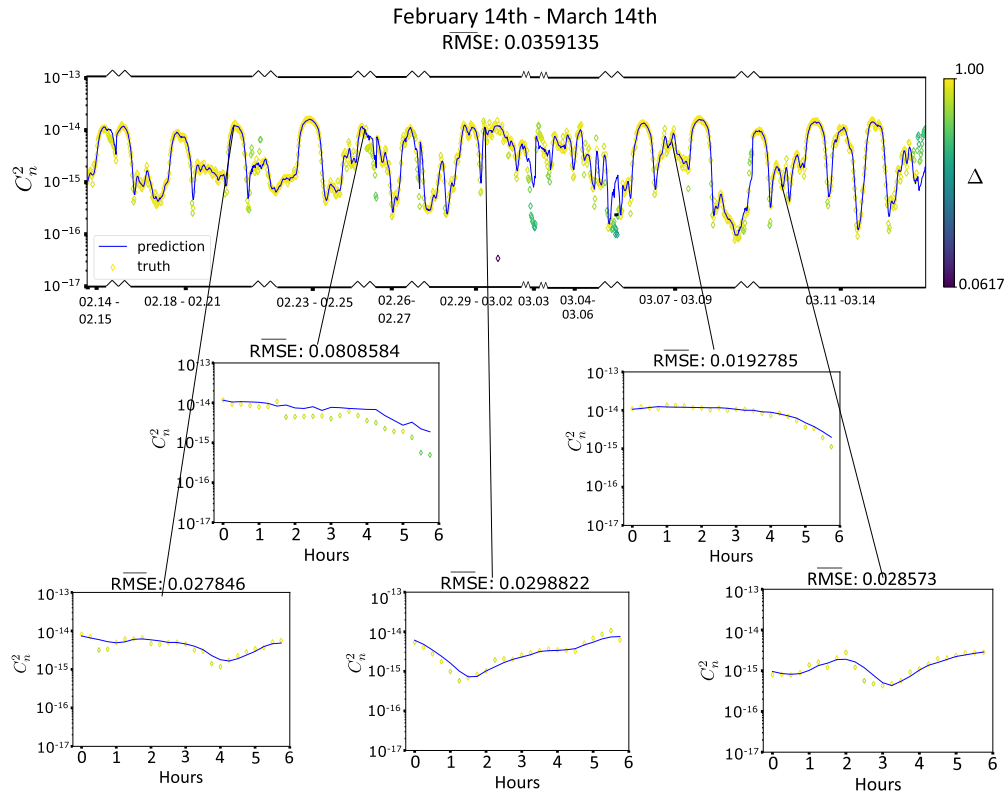
### 4.1. $C_n^2$ forecast

The network is validated over two test periods, one spanning October 2023 and the other from February 14th to March 14th, 2024. To do so, we partition the dataset in two different ways, each having one of the aforementioned periods as the test dataset, then train the network on each partition (see Appendix B). In particular, while the February-March period exhibits smooth oscillations of atmospheric turbulence, the dataset from October 2023 has been specifically chosen as it features a significantly variable evolution of the  $C_n^2$  compared to other months, which makes it a challenging validation. The performance of the network is evaluated as the average root mean squared error, defined as  $\overline{\text{RMSE}} = \frac{1}{N} \sqrt{\text{MSE}}$ . This error is computed using the predicted and actual  $\log_{10} C_n^2$  values, normalized according to Eq. (3). In the following, we report the results for 6 hours in advance with a 15-minute time resolution. However, we have observed adequate results for up to 12 hours in advance, even with a time resolution down to 1 minute.

Figure 3 plots the prediction of  $C_n^2$  values over the February-March period. The plots are obtained by cascading 6-hour predictions sequentially in time. The colorscale encodes the accuracy on the final prediction:  $\Delta(\text{RMSE}) = 1 - \text{RMSE}$ , with higher values of  $\Delta$  indicating better performances. A typical day features  $C_n^2$  peaks at  $10^{-14}$  during the afternoon, dropping in the evening to values approaching  $10^{-16}$ , followed by a local peak overnight around  $10^{-14}$ . The insets show representative examples of individual 6-hour predictions, generated from portions of the test set equally spaced in time. For the February-March test set, the network performs very well in matching the actual  $C_n^2$ , with an average RMSE within the order of  $10^{-2}$  ( $\Delta \geq 0.90$ ). This corresponds to an average relative error of about 9% on the actual  $C_n^2$  values. It also shows a certain degree of robustness to minor deviations from the typical periodic trend. The predictions for the October test set are reported in Appendix C.

### 4.2. Numerical QKD experiment

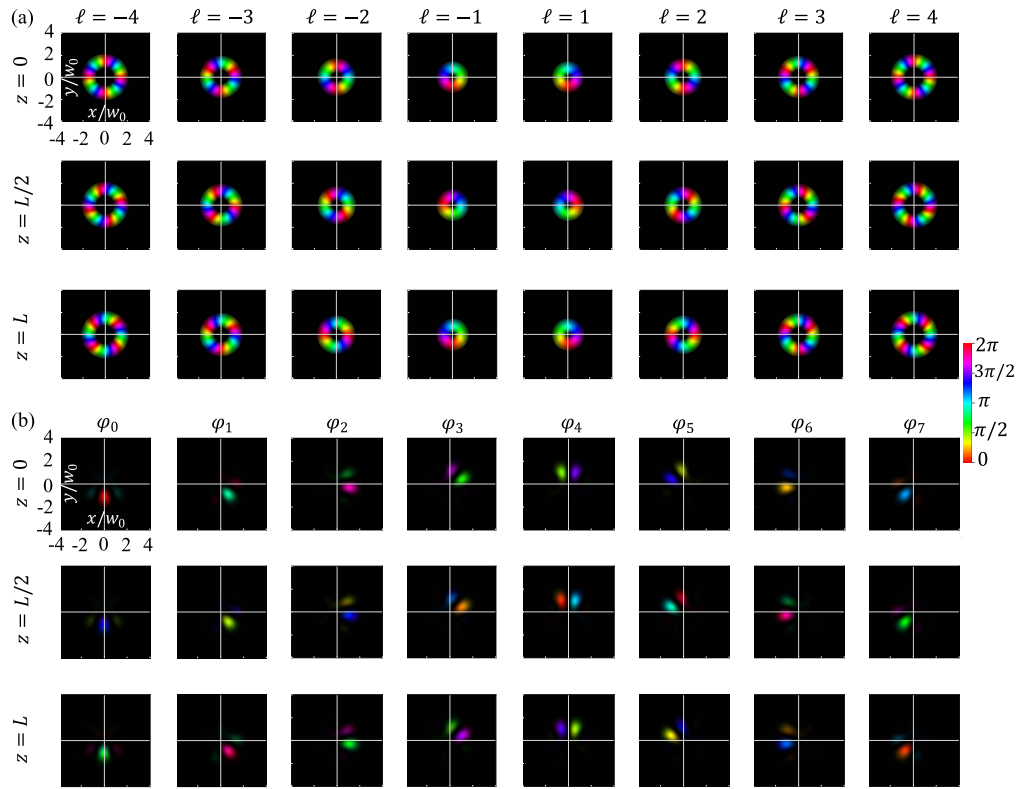
We numerically investigate the performance of an 8-dimensional BB84 protocol, by simulating a realistic key exchange across the channel with increasing turbulence strengths. This analysis is conducted to determine the tolerable levels of turbulence within our channel, beyond which a QKD experiment would be ineffective and no longer feasible. In particular, weak ( $C_n^2 = 10^{-16}$ ), moderate ( $C_n^2 = 10^{-15}$ ), and strong ( $C_n^2 = 10^{-14}$ ) turbulence are investigated. The high-dimensional protocol exploits Laguerre-Gaussian (LG) spatial modes of light, carrying a discrete amount of orbital angular momentum (OAM)  $\ell\hbar$  [42], encoding a qudit of information. The Fourier-conjugate



**Fig. 3. Predictions on one test set.** Network predictions for the month from mid-February to mid-March 2024. We have our trained model forecast 6 hours in the future, then cascade individual predictions to cover one month. Jagged lines indicate missing data points from the scintillometer system. A one-hour moving average is applied to the training set to remove high-frequency noise. The scatter points refer to the raw data (before the moving average is performed). The insets show examples of individual 6-hour predictions, equally spaced throughout the month. Good performances are observed on average, with larger deviations occurring in correspondence with strong feature variations.

basis (ANGLE modes) is used as a mutually unbiased basis (MUB). The explicit expression of these modes in the position representation is provided in Appendix D.

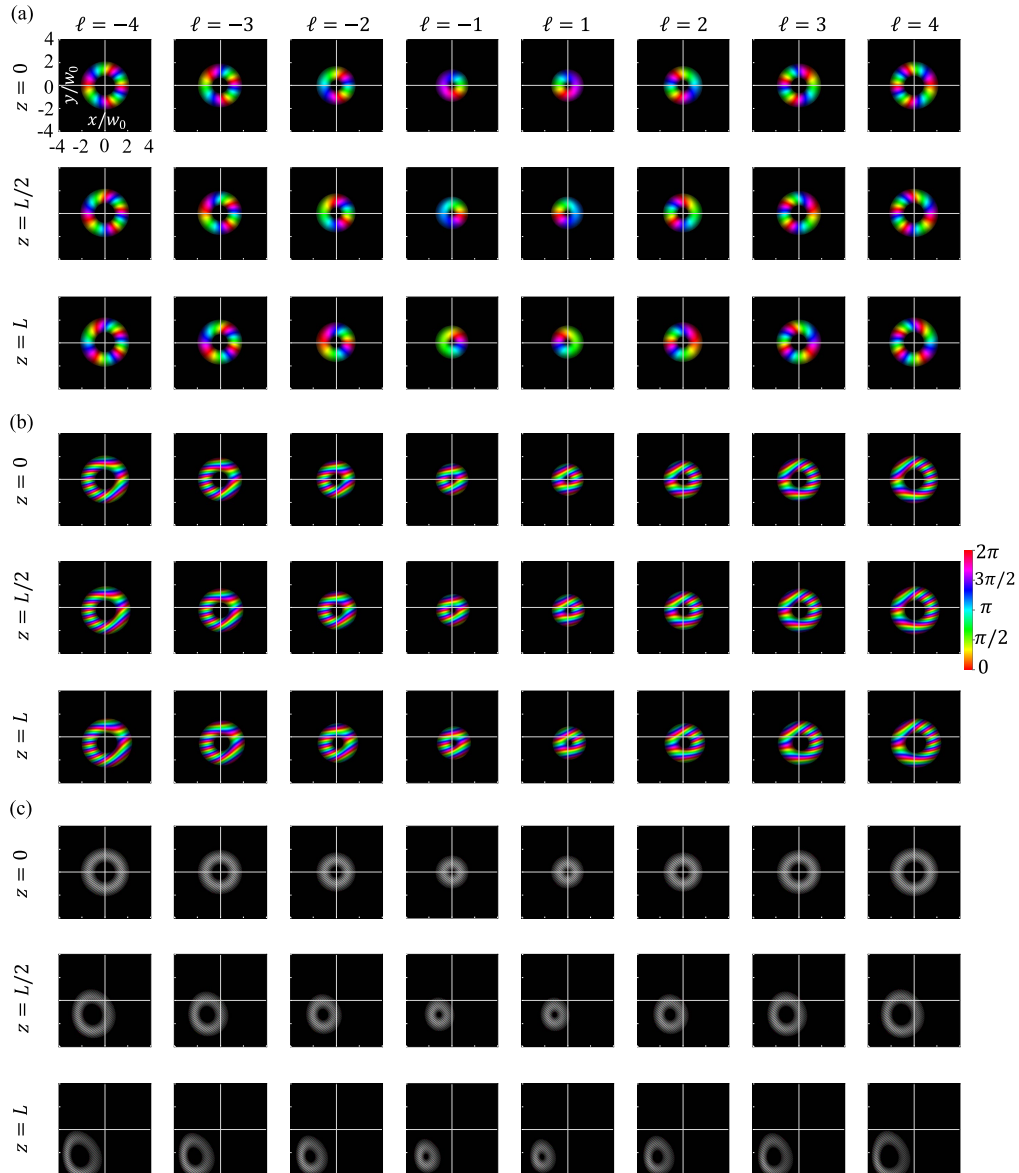
The parameters used in our simulated experiments are:  $w_0 = 8$  cm,  $\lambda = 810$  nm,  $L = 5.4$  km,  $D = 30$  cm, where  $w_0$  is the beam waist at  $z = 0$ ,  $\lambda$  is the operating wavelength,  $L$  is the channel length, and  $D$  is the receiver aperture. Figure 4(a) and (b) show the OAM and ANGLE modes in the sender plane ( $z = 0$ ), halfway across the channel ( $z = L/2$ ), and at the receiver end ( $z = L$ ), when turbulence is not present, with the opacity and the hue colorscale encoding the amplitude and the phase of the field, respectively. The  $\ell = 0$  mode is removed from the OAM basis, due to non-negligible crosstalks with adjacent modes [14]. In our simulations, the effect of turbulence is modeled as a single phase mask located at  $z = 0$ , which is generated from a random combination of Zernike modes, where the contribution of each mode is extracted within the corresponding variance associated with the simulated  $C_n^2$  [31]. However, a more general approach would require including multiple phase objects across the beam propagation, as discussed in Sec. 2.1. For each level of turbulence, we run 100 numerical experiments, from which the average crosstalk matrices are extracted. Further details on the numerical simulations are provided in Appendix E.



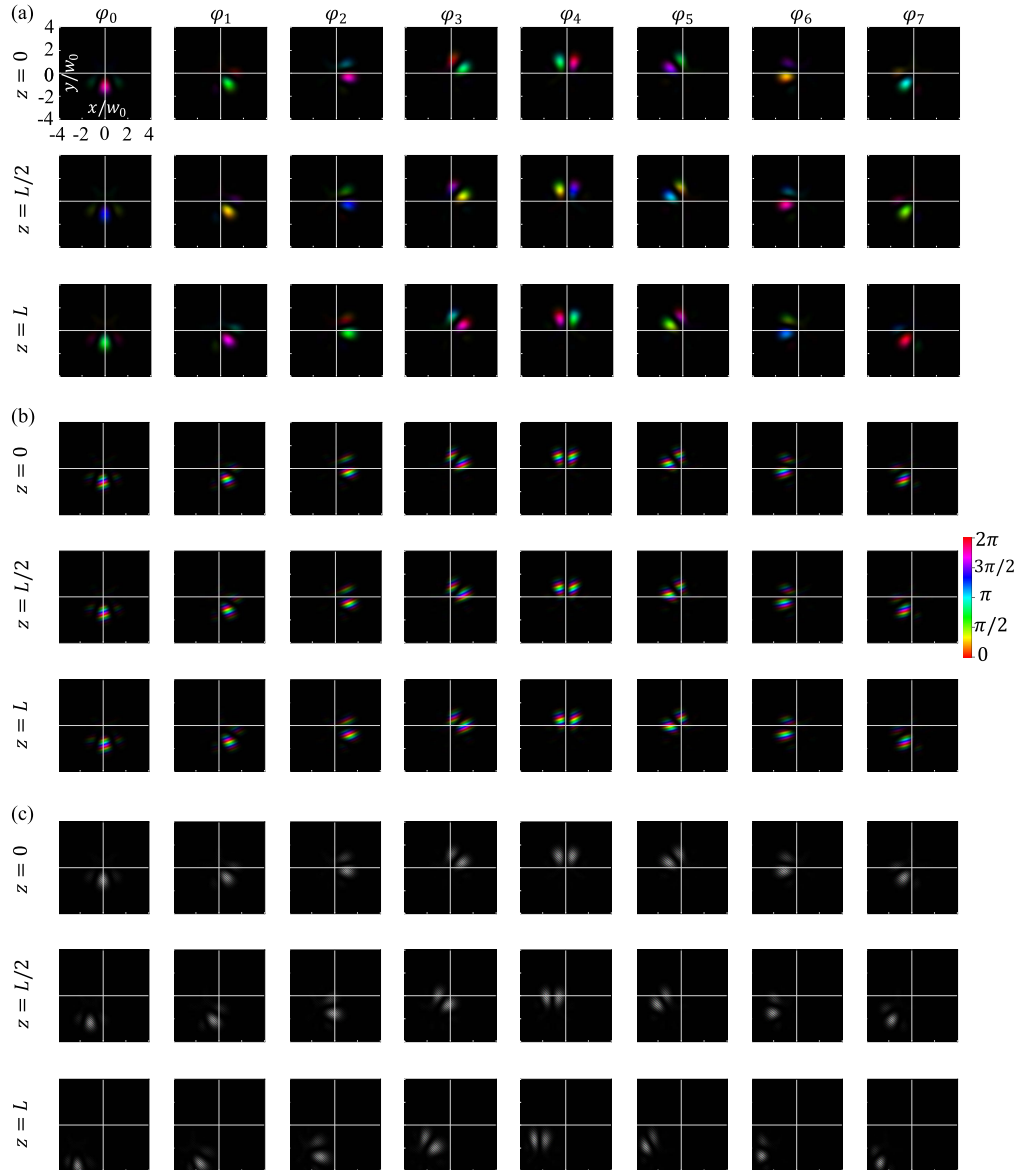
**Fig. 4. 8-dimensional QKD.** (a) 8-dimensional OAM space, spanning modes from  $\ell = -4$  to  $\ell = 4$ , excluding  $\ell = 0$ . (b) ANGLE modes  $\{\varphi_0, \dots, \varphi_7\}$  provide a possible MUB for the high-dimensional protocol. The input field at  $z = 0$ , and the resulting field upon propagation at  $z = L/2$  and  $z = L$ , with  $L = 5.4$  km being the length of the free-space link, are provided for each mode. Propagation is simulated without turbulence. Opacity and hue colorscale encode the amplitude and the phase of the field, respectively.

Figure 5 provides a visualization of the effect of different turbulence levels on the input OAM beams, from weak (a) to moderate (b) to strong (c) turbulence. As a representative example, we plot the realization corresponding to the maximum error within the same turbulence strength, i.e., the one associated with the larger beam distortion. It is worth noticing that the statistical contribution of each Zernike mode decreases with the mode index [31]. Accordingly, low-rank aberrations such as tip and tilt typically dominate the beam dynamics, resulting in an overall decentering effect. This is also evident in Fig. 6, showing the effect of the same aberrations on ANGLE modes.

The crosstalk integrals between the modes are plotted in Fig. 7. In the case of weak turbulence (a-d), the channel is capable of supporting a high-dimensional BB84 protocol with an information capacity greater than a single bit of information per photon. Interestingly, for moderate (b-e) and strong (c-f) turbulence, the ANGLE basis proves more robust than OAM. This is ascribed to the concentration of optical power in a smaller region, associated with a reduced effective beam waist. More rigorously, the security of a certain basis across a channel is quantified by the quantum dit error rate (QDER), which is the percentage of incorrect measurements in the shared key after the security verification. In 8 dimensions, a QDER greater than 24.7% results in the impossibility of guaranteeing secure communications as the key rate (given in Eq. (2)) will be negative. This value is related to the amount of information that can be transferred per

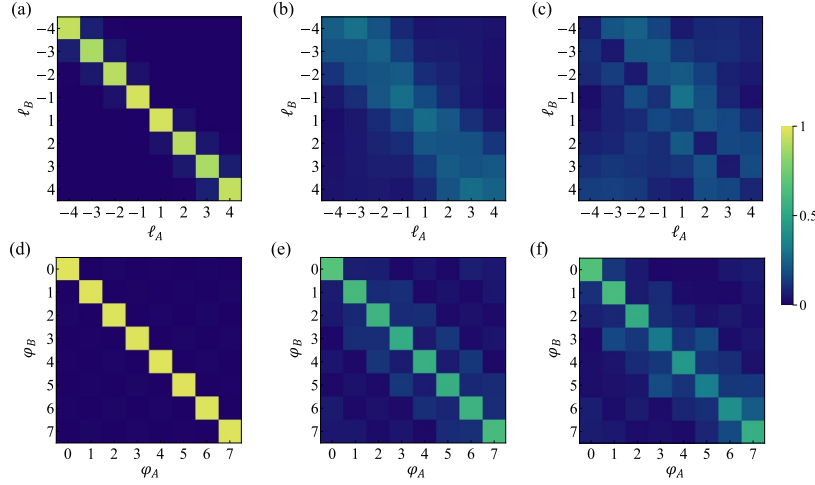


**Fig. 5. Aberrated OAM modes.** Aberrations affect the OAM beam propagation over the channel. Different regimes are considered: (a) weak ( $C_n^2 = 10^{-16}$ ), (b) moderate ( $C_n^2 = 10^{-15}$ ), and (c) strong turbulence ( $C_n^2 = 10^{-14}$ ). The panels refer to the realization yielding the least secure communication, i.e., the one minimizing the diagonal overlap integrals.



**Fig. 6. Aberrated ANGLE modes.** Aberrations affect the ANGLE beam propagation over the channel. Different regimes are considered: (a) weak ( $C_n^2 = 10^{-16}$ ), (b) moderate ( $C_n^2 = 10^{-15}$ ), and (c) strong turbulence ( $C_n^2 = 10^{-14}$ ). The panels refer to the realization yielding the least secure communication, i.e., the one minimizing the diagonal overlap integrals.

photon sent through the channel. The average QDERs extracted from the numerical simulation are provided in Table 1, where the information capacity is expressed in “bits per photon” (b/p). An ideal 8-dimensional channel would support 3 b/p of information capacity. The simulations confirm that secure communications can only be established in the weak-turbulence regime ( $C_n^2 \approx 10^{-16}$ ).



**Fig. 7. Crosstalk matrix.** Average crosstalk integrals between different OAM and ANGLE modes across a turbulent channel. The average is computed over 100 realizations of (a-d) weak ( $C_n^2 = 10^{-16}$ ), (b-e) moderate ( $C_n^2 = 10^{-15}$ ), and (c-f) strong turbulence ( $C_n^2 = 10^{-14}$ ). Each row of individual panels refers to the normalized average crosstalk over 100 independent realizations.

**Table 1. Average QDER for the OAM and ANGLE (ANG) bases, along with the overall error rate and subsequent sifted key rate per photon, extracted from 100 numerical realizations of different turbulence levels over the channel. Security is guaranteed only in the low-turbulence regime. Negative information capacity is reported as 0.**

$C_n^2$	OAM QDER	ANG QDER	$e_q$	$R(e_q)$ (b/p)
$10^{-16}$	8.18%	2.33%	5.25%	2.11
$2 \times 10^{-16}$	17.80%	5.40%	11.60%	1.31
$5 \times 10^{-16}$	55.07%	21.77%	38.42%	0
$10^{-15}$	77.00%	41.47%	59.23%	0
$10^{-14}$	92.07%	51.96%	72.02%	0

## 5. Conclusion

We employed a recurrent neural network to forecast future turbulence conditions within an optical channel. The accuracy of the predictions over significantly long periods demonstrates that our surrogate model has achieved a robust learning of the temporal variation of  $C_n^2$  values correlated to relevant weather parameters. Moreover, we have shown how the predictions from TAROQO can be used to foresee the error rate of a QKD experiment, by simulating a high-dimensional protocol with the real parameters of our free-space link. This result could also apply to QKD ground-to-satellite systems to optimize key exchange rates. While numerical simulations have

only been performed for well-known spatial modes of light, future studies could address the performance of other encoding schemes for QKD under turbulence, such as vector beams and the time-frequency domain.

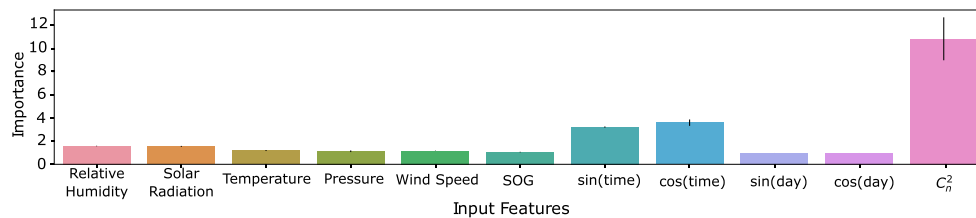
Scintillometer data acquisition will continue over the region of Ottawa, continually expanding the current dataset for enhanced training. Additionally, improved performance could be achieved by exploring alternative neural network architectures, such as xLSTMs [43], which offer enhanced capabilities for modeling long-term dependencies within the data. Promising directions also include physics-informed neural networks, which are particularly well-suited for learning the underlying differential relationships within the data [44–46], and autoencoder architectures, similar to those used in machine translation [47,48]. To compensate for limited data, it will also be interesting to explore data augmentation techniques, particularly those involving generative adversarial networks [49].

## Appendix A. Permutation feature importance

The importance of each feature is assessed through the permutation feature importance (PFI) technique [50]. Given a model trained with every possible input feature, we corrupt the dataset by permuting one of the features. The importance of each feature is quantified by evaluating the model on the corrupted dataset and comparing its performance with the original model:

$$I = \frac{\epsilon_{\text{perm}}}{\epsilon_{\text{orig}}}, \quad (6)$$

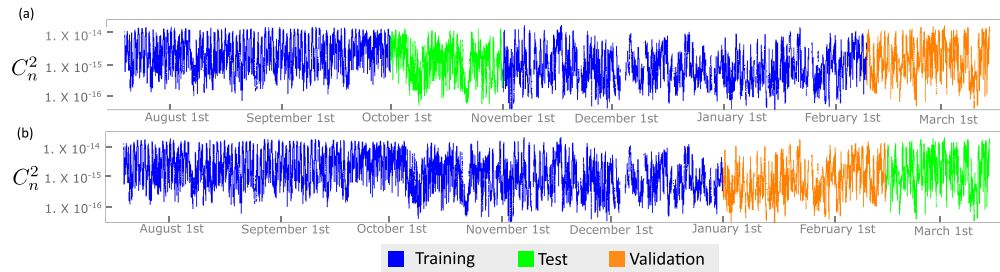
where  $I$  is the feature importance, and  $\epsilon_{\text{perm}}$  and  $\epsilon_{\text{orig}}$  are the RMSE of the corrupted and original model, respectively. In addition to the input features indicated in Fig. 2, we also considered pressure (hPa), snow on ground (SOG, cm), wind speed (m/s), and the day according to the Julian calendar. We then execute PFI on the first 1000 examples of the dataset. For each feature, we repeat the permutation three times to minimize the effects induced by the variability of the permutation. The results are shown in Fig. 8. Prior values of  $C_n^2$  appear to be the most important feature, followed by time, relative humidity, and solar radiation. This is in congruence with the analysis carried out in Ref. [20]. It must be noted that this kind of analysis tends to penalize the importance of correlated features, such as pressure and temperature [51]. To address this issue, we permute pressure, temperature, humidity, and wind speed simultaneously, effectively considering their importance together [52]. Indeed, increased values for importance were reported.



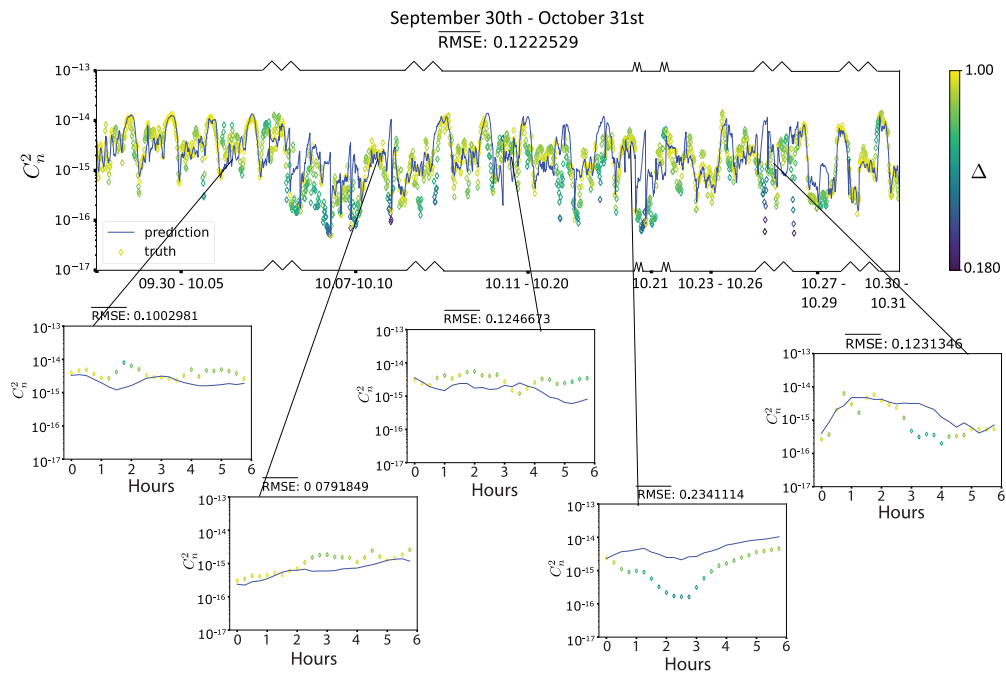
**Fig. 8. Importance of input features.** For each input feature, the mean and standard deviation (error bars) of the importance is computed over three repeated permutations.

## Appendix B. Training hyperparameters

The network is trained and validated using data spanning from July 2023 to March 2024. We list the model hyperparameters in Table 2. One month is removed from the original dataset and used for testing (cf. Sec. 4.1). Figure 9 illustrates the partitions of the complete dataset into training, validation, and test datasets, for the two case studies reported in Fig. 3 and Fig. 10.



**Fig. 9. Partitions of the complete dataset.** (a) The month of October is selected as test set, and an 85:15 train:validation split is applied to the remaining dataset. (b) A 75:15:10 train:validation:test split is applied to the complete dataset, which corresponds to a test set spanning from mid-February to mid-March (cf. Sec. 4.1).



**Fig. 10. Predictions over the month of October.** Strong fluctuations of the input features lead to larger deviations of the final predictions from the measured turbulence, with an average relative error of about 20% on the actual  $C_n^2$  values.

Table 2. Hyperparameters.

Hours of Prior Data	12 Hours
Forecast Length	6 Hours
Input Time Resolution	1 Minute
Output Time Resolution	15 Minutes
Batch Size	32
Train-Validation Split ( <i>with October</i> )	85:15
Train-Validation-Test Split ( <i>no October</i> )	75:15:10
Initial Learning Rate	$10^{-4}$
Number of Epochs at Convergence ( <i>with October</i> )	149
Number of Epochs at Convergence ( <i>no October</i> )	262
Patience	15 Epochs
Reduction Factor	0.1
Training Error at Convergence ( <i>with October</i> )	$1.2 \times 10^{-4}$
Training Error at Convergence ( <i>no October</i> )	$7.2 \times 10^{-5}$
Validation Error at Convergence ( <i>with October</i> )	$1.6 \times 10^{-4}$
Validation Error at Convergence ( <i>no October</i> )	$1.3 \times 10^{-4}$

The training process is realized using the Tensorflow library [53], and model optimization is handled using the Adam optimizer [54]. Training is carried out on the *Narval* supercluster using a NVidia A100SXM4 GPU. An adaptive training strategy is used, whereby the learning rate is reduced by a factor of 0.1 if the validation loss does not decrease significantly within 15 epochs (this number is referred to as *patience*). The number of hours of prior data is kept fixed at 12. Altogether, training is completed in approximately 5 hours.

In addition to the study reported in Appendix A, we have also investigated the performance of different architectures. For instance, with our existing GRU-RNN structure, we have explored longer inputs of 18, 24, and 30 hours, and increased the number of units within each hidden layer to up to 1640 units. However, no significant reduction in the validation loss was observed. We also considered a fully connected neural network, with 2 hidden layers of 720 neurons separated by ReLU activation layers. Here, the input layer admits a flattened dataset where the first 6 entries represent the input features of the first time step. However, the validation error of our flagship model at convergence is three orders of magnitude smaller than that of this architecture.

### Appendix C. Predictions over the month of October

TAROQQO's predictions for the month of October are shown in Fig. 10. Unlike the smoother variations in the February-March dataset (see Section 4.1), the more pronounced variations observed in October prove to be a greater challenge for the network, with an average RMSE within the order of  $10^{-1}$  ( $\Delta < 0.90$ ). The average relative error computed over the month is approximately 20% on the actual  $C_n^2$  values.

### Appendix D. OAM and ANGLE modes of light

The 8-dimensional QKD protocol explored in Sec. 4.2 leverages the MUBs provided by a set of LG modes and the corresponding Fourier-transformed basis. LG modes are labeled by two integers  $\ell$  and  $p$ , representing the azimuthal (OAM) and radial index, respectively. Our protocol only relies on the OAM content of the beam, hence we set  $p = 0$  for all the LG modes. In the

position basis, their expression reads

$$\langle x, y; z | \ell \rangle = N_\ell \frac{w_0}{w(z)} \left( \frac{r\sqrt{2}}{w(z)} \right)^{|\ell|} e^{-r^2/w^2(z)} L_{|\ell|}^0 \left( \frac{2r^2}{w^2(z)} \right) e^{-ik\frac{r^2}{2R(z)}} e^{-i\ell\phi} e^{i\psi(z)}, \quad (7)$$

where  $N_\ell$  is a normalization constant,  $r = \sqrt{x^2 + y^2}$  and  $\phi = \arctan(y/x)$  are the radial and azimuthal coordinates, respectively, and  $L_\ell^p$  are the generalized Laguerre polynomials. The beam waist  $w(z)$ , the radius of curvature  $R(z)$  and the Gouy phase  $\psi(z)$  can be determined from [55]:

$$w(z) = w_0 \sqrt{1 + \frac{z^2}{z_R^2}}; \quad (8a)$$

$$R(z) = z \left( 1 + \frac{z_R^2}{z^2} \right); \quad (8b)$$

$$\psi(z) = (|\ell| + 1) \arctan \frac{z}{z_R}, \quad (8c)$$

where  $z_R = \pi w_0^2/\lambda$  is the Rayleigh range.

The conjugate basis, referred to as the ANGLE basis, is obtained as

$$|\varphi_j\rangle = \frac{1}{\sqrt{d}} \sum_{\ell} |\ell\rangle e^{2\pi i j g(\ell)/d}, \quad (9)$$

where  $d = 8$ , the summations runs from  $\ell = -d/2$  to  $\ell = d/2$  excluding the  $|\ell = 0\rangle$  mode, and  $g(\ell) = d/2 + (\ell - 1)\Theta(\ell) + \ell\Theta(-\ell)$ , with  $\Theta$  the Heaviside step function.

## Appendix E. Details on numerical simulations

In our simulations, aberrations are modeled as a single phase screen  $\xi(x, y)$  located at the sender plane  $z = 0$ :

$$\tilde{E}(x, y; z = 0) = E(x, y; z = 0) e^{i\xi(x, y)}, \quad (10)$$

where  $\tilde{E}$  and  $E$  are the aberrated and input fields, respectively. The phase screen is generated as a superposition of Zernike modes [30]:

$$\xi(\rho, \phi) = \sum_{n, m} c_{n, m} Z_{n, m}(\rho, \phi), \quad (11)$$

where  $\rho$  is the normalized radial distance in the transverse plane:  $\rho = \sqrt{x^2 + y^2}/\text{Max}(\sqrt{x^2 + y^2})$ , where the maximum is taken across the spatial window considered for the simulations (cf. Figs. 4–6). The Zernike mode of order  $(n, m)$ , with  $n \geq |m| > 0$ , is defined as

$$Z_{n, m}(\rho, \phi) = R_{n, |m|}(\rho) \cdot (\Theta(m) \cos m\phi + (1 - \Theta(-m)) \sin m\phi), \quad (12)$$

where  $R_{n, |m|}$  is the radial Zernike polynomial

$$R_{n, |m|}(\rho) = \sum_{k=0}^{\frac{n-|m|}{2}} \frac{(-1)^k (n-k)!}{k! \left(\frac{n+|m|}{2} - k\right)! \left(\frac{n-|m|}{2} - k\right)!} \rho^{n-2k} \quad (13)$$

if  $n - |m|$  is even and 0 if  $n - |m|$  is odd. The weights of individual Zernike modes in Eq. (11) are extracted from a zero-mean normal distribution depending on the turbulence conditions. Indeed,

the variance  $\sigma_{n,m}^2$  of the coefficient  $c_{n,m}$  depends on the turbulence strength via the relation [31]

$$\sigma_{n,m}^2 = \gamma_{n,m} \left( \frac{D}{r_0} \right)^{5/3}, \quad (14)$$

where  $\gamma_{n,m}$  is a mode-dependent coefficient,  $D$  is the receiver aperture, and  $r_0$  is the Fried parameter. The latter can be determined from the  $C_n^2$  [28]:

$$r_0 = 1.68 \left( C_n^2 L k^2 \right)^{-3/5}, \quad (15)$$

where  $L$  is the channel length and  $k = 2\pi/\lambda$  is the light wavevector. For completeness, the list of coefficients  $\gamma_{n,m}$  for the Zernike modes employed in our simulations, as well as the resulting variances under the turbulence regimes that have been considered, are provided in Table 3.

**Table 3. The coefficients of proportionality  $\gamma_{n,m}$  and the variances of the Zernike modes in the experimental setup considered in the paper (see Eq. (14)) are provided for different turbulence regimes.**

$n$	$m$	$\gamma_{n,m}$	$\sigma_{n,m}^2 (C_n^2 = 10^{-16})$	$\sigma_{n,m}^2 (C_n^2 = 10^{-15})$	$\sigma_{n,m}^2 (C_n^2 = 10^{-14})$
1	1	0.582	1.0708	10.708	107.08
1	-1	0.134	0.2466	2.466	24.66
2	0	0.111	0.2042	2.042	20.42
2	-2	0.088	0.1619	1.619	16.19
2	2	0.0648	0.1192	1.192	11.92
3	-1	0.0587	0.1080	1.080	10.80
3	1	0.0525	0.0966	0.966	9.66
3	-3	0.0463	0.0852	0.852	8.52
3	3	0.0401	0.0738	0.738	7.38
4	0	0.0377	0.0694	0.694	6.94
4	2	0.0352	0.0648	0.648	6.48
4	-2	0.0328	0.0604	0.604	6.04
4	4	0.0304	0.0559	0.559	5.59
4	-4	0.0279	0.0513	0.513	5.13

The propagation of the OAM and ANGLE modes up to a distance  $z$  without turbulence can be directly obtained from the analytical expressions of Eq. (7) and Eq. (9), respectively (see Fig. 4). When turbulence is present, there is no analytical expression for the propagation, and Fresnel diffraction of aberrated modes must be numerically simulated (see Fig. 5 and Fig. 6):

$$\tilde{E}(x, y; z) = \frac{e^{ikz}}{i\lambda z} e^{i\frac{\pi}{\lambda z}(x^2+y^2)} \mathcal{F}(x, y; z), \quad (16)$$

where

$$\mathcal{F}(x, y; z) = \iint dx' dy' \tilde{E}(x', y'; 0) e^{i\frac{\pi}{\lambda z}(x'^2+y'^2)} e^{\frac{2\pi i}{\lambda z}(xx'+yy')}. \quad (17)$$

Finally, each element  $(i, j)$  of the crosstalk matrices plotted in Fig. 7 is obtained as the normalized mean of the overlap integrals between the  $i$ -th and the  $j$ -th modes, where the mean is performed over 100 random realizations of weak, moderate, and strong turbulence, and the normalization pertains each row independently. The overlap integral between a sent aberrated

mode  $\tilde{E}_j$  and a detected (non-aberrated) mode  $E_i$  is given by

$$I_{i,j} = \iint dx dy E_i^*(x, y; z) \tilde{E}_j(x, y; z), \quad (18)$$

where \* denotes complex conjugation of the field. Since optical propagation acts as a unitary transformation on the beam, the overlap integrals can be computed at any distance along the propagation. For convenience, the overlap is evaluated at  $z = 0$ .

**Funding.** Canada Research Chairs; Ministero dell'Università e della Ricerca (PE0000023-NQSTI); Canada First Research Excellence Fund; National Research Council Canada (HTSN Challenge Program); Alliance Consortia Quantum Grant (QUINT, ARAQNE).

**Acknowledgment.** The authors would like to thank Alicia Sit for valuable discussions and for help purchasing the scintillometer system. The authors would also like to thank Nazanin Dehghan, Alessio D'Errico, and Ashlin Jacob for their help in setting up the scintillometer system. Finally, the authors would like to thank Environment and Climate Change Canada for providing us with the meteorological data. In particular, we appreciate the help of John Richard of Applied Climatology Services. This work was supported by Canada Research Chairs; Canada First Research Excellence Fund (CFREF); National Research Council of Canada High-Throughput and Secure Networks (HTSN) Challenge Program; the Qeyssat User INvestigation Team (QUINT) and the Alliance for Research and Applications of Quantum Network Entanglement (ARAQNE) Alliance Consortia Quantum grants. FDC further acknowledges support from the PNRR MUR project PE0000023-NQSTI.

**Disclosures.** The authors declare no conflicts of interest.

**Author Contributions.** FDC, LS, and EK conceived the idea. LS collected and prepared the scintillometer data. TJ developed the neural network architecture. FDC performed numerical simulations of QKD under turbulence. TJ, LS, and FDC wrote the first version of the manuscript. FB, MK, KH, and EK supervised the project.

**Data availability.** The source code and data underlying the results presented in this paper are available in Ref. [56].

## References

1. J.-L. Kong, X.-M. Fan, X.-B. Jin, *et al.*, "Bmae-net: A data-driven weather prediction network for smart agriculture," *Agronomy* **13**(3), 625 (2023).
2. D. Lazos, A. B. Sproul, and M. Kay, "Optimisation of energy management in commercial buildings with weather forecasting inputs: A review," *Renew. Sustain. Energy Rev.* **39**, 587–603 (2014).
3. E. Kendon, N. Roberts, H. Fowler, *et al.*, "Heavier summer downpours with climate change revealed by weather forecast resolution model," *Nat. Clim. Change* **4**(7), 570–576 (2014).
4. B. Merz, C. Kuhlicke, M. Kunz, *et al.*, "Impact forecasting to support emergency management of natural hazards," *Rev. Geophys.* **58**(4), e2020RG000704 (2020).
5. F. Roddier, "The effects of atmospheric turbulence in optical astronomy," in *Prog. Optics*, vol. 19 (Elsevier, 1981), pp. 281–376.
6. X. Zhu and J. Kahn, "Free-space optical communication through atmospheric turbulence channels," *IEEE Trans. Commun.* **50**(8), 1293–1300 (2002).
7. N. Leonhard, G. Sorelli, V. N. Shatokhin, *et al.*, "Protecting the entanglement of twisted photons by adaptive optics," *Phys. Rev. A* **97**(1), 012321 (2018).
8. J. C. Wyngaard, "Atmospheric turbulence," *Annu. Rev. Fluid Mech.* **24**(1), 205–234 (1992).
9. A. N. Kolmogorov, "The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers," *Proc. R. Soc. Lond. A* **434**(1890), 9–13 (1991).
10. J. A. Anguita, M. A. Neifeld, and B. V. Vasic, "Turbulence-induced channel crosstalk in an orbital angular momentum-multiplexed free-space optical link," *Appl. Opt.* **47**(13), 2414–2429 (2008).
11. G. A. Tyler and R. W. Boyd, "Influence of atmospheric turbulence on the propagation of quantum states of light carrying orbital angular momentum," *Opt. Lett.* **34**(2), 142–144 (2009).
12. C. J. Pugh, J.-F. Lavigne, J.-P. Bourgoin, *et al.*, "Adaptive optics benefit for quantum key distribution uplink from ground to a satellite," *Adv. Opt. Technol.* **9**(5), 263–273 (2020).
13. J. Zhao, Y. Zhou, B. Braverman, *et al.*, "Performance of real-time adaptive optics compensation in a turbulent channel with high-dimensional spatial-mode encoding," *Opt. Express* **28**(10), 15376–15391 (2020).
14. L. Scarfe, F. Hufnagel, M. F. Ferrer-Garcia, *et al.*, "Fast adaptive optics for high-dimensional quantum communications in turbulent channels," *arXiv* (2023).
15. R. N. Lanning, M. A. Harris, D. W. Oesch, *et al.*, "Quantum communication over atmospheric channels: A framework for optimizing wavelength and filtering," *Phys. Rev. Appl.* **16**(4), 044027 (2021).
16. M. T. Gruneisen, M. L. Eickhoff, S. C. Newey, *et al.*, "Adaptive-optics-enabled quantum communication: A technique for daytime space-to-earth links," *Phys. Rev. Appl.* **16**(1), 014067 (2021).
17. F. S. Binkowski, "A simple semi-empirical theory for turbulence in the atmospheric surface layer," *Atmos. Environ.* **13**(2), 247–253 (1979).

18. A. Rafalimanana, C. Giordano, A. Ziad, *et al.*, "Prediction of atmospheric turbulence by means of wrf model for optical communications," in *International Conference on Space Optics—ICSO 2020*, vol. 11852 (SPIE, 2021), pp. 1856–1866.
19. Y. Wang and S. Basu, "Estimation of optical turbulence in the atmospheric surface layer from routine meteorological observations: an artificial neural network approach," in *Laser Communication and Propagation through the Atmosphere and Oceans III*, vol. 9224 (SPIE, 2014), pp. 300–307.
20. M. G. Grose and E. A. Watson, "Forecasting atmospheric turbulence conditions from prior environmental parameters using artificial neural networks," *Appl. Opt.* **62**(13), 3370–3379 (2023).
21. K. Bi, L. Xie, H. Zhang, *et al.*, "Accurate medium-range global weather forecasting with 3d neural networks," *Nature* **619**(7970), 533–538 (2023).
22. C. Bodnar, W. P. Bruinsma, A. Lucic, *et al.*, "Aurora: A foundation model of the atmosphere," *arXiv* (2024).
23. In Italian, Tarocco means tarot cards, which are used in fortune telling.
24. M. Krenn, J. Handsteiner, M. Fink, *et al.*, "Twisted light transmission over 143 km," *Proc. Natl. Acad. Sci.* **113**(48), 13648–13653 (2016).
25. M. Krenn, R. Fickler, M. Fink, *et al.*, "Communication with spatially modulated light through turbulent air across vienna," *New J. Phys.* **16**(11), 113028 (2014).
26. M. P. J. Lavery, C. Peuntinger, K. Günthner, *et al.*, "Free-space propagation of high-dimensional structured optical fields in an urban environment," *Sci. Adv.* **3**(10), e1700552 (2017).
27. A. Sit, F. Bouchard, R. Fickler, *et al.*, "High-dimensional intracity quantum cryptography with structured photons," *Optica* **4**(9), 1006–1010 (2017).
28. V. I. Tatarski, *Wave propagation in a turbulent medium* (Courier Dover Publications, 2016).
29. L. Burger, I. Litvin, and A. Forbes, "Simulating atmospheric turbulence using a phase-only spatial light modulator : research article," *S. Afr. J. Sci.* **104**(3), 129–134 (2008).
30. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge University, 1999), 7th ed.
31. R. J. Noll, "Zernike polynomials and atmospheric turbulence\*," *J. Opt. Soc. Am.* **66**(3), 207–211 (1976).
32. H. C. Ward, "Scintillometry in urban and complex environments: a review," *Meas. Sci. Technol.* **28**(6), 064005 (2017).
33. P. H. Deitz and N. J. Wright, "Saturation of scintillation magnitude in near-earth optical propagation," *J. Opt. Soc. Am.* **59**(5), 527–535 (1969).
34. C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," *Theor. Comput. Sci.* **560**, 7–11 (2014).
35. N. J. Cerf, M. Bourennane, A. Karlsson, *et al.*, "Security of quantum key distribution using  $d$ -level systems," *Phys. Rev. Lett.* **88**(12), 127902 (2002).
36. H. Bechmann-Pasquinucci and W. Tittel, "Quantum cryptography using larger alphabets," *Phys. Rev. A* **61**(6), 062308 (2000).
37. L. Sheridan and V. Scarani, "Security proof for quantum key distribution using qudit systems," *Phys. Rev. A* **82**(3), 030301 (2010).
38. P. Sharma, A. Agrawal, V. Bhatia, *et al.*, "Quantum key distribution secured optical networks: A survey," *IEEE Open J. Commun. Soc.* **2**, 2049–2083 (2021).
39. O. Amer, V. Garg, and W. O. Krawec, "An introduction to practical quantum key distribution," *IEEE Trans. Aerosp. Electron. Syst.* **36**(3), 30–55 (2021).
40. Scintech, "Advanced atmospheric sensing," <https://www.scintec.com/>.
41. Government of Canada. [https://climate.weather.gc.ca/index\\_e.html](https://climate.weather.gc.ca/index_e.html).
42. L. Allen, M. W. Beijersbergen, R. J. C. Spreeuw, *et al.*, "Orbital angular momentum of light and the transformation of laguerre-gaussian laser modes," *Phys. Rev. A* **45**(11), 8185–8189 (1992).
43. M. Beck, K. Pöppel, M. Spanring, *et al.*, "xIstm: Extended long short-term memory," *arXiv* (2024).
44. R. T. Chen, Y. Rubanova, J. Bettencourt, *et al.*, "Neural ordinary differential equations," *Adv. Neural Inf. Process. Syst.* **31**, 1 (2018).
45. Z. Li, N. Kovachki, K. Aizzadenesheli, *et al.*, "Fourier neural operator for parametric partial differential equations," *arXiv* (2020).
46. J. Pathak, S. Subramanian, P. Harrington, *et al.*, "Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators," *arXiv* (2022).
47. C. E. A. Mohamed Akram Zaytar, "Sequence to sequence weather forecasting with long short-term memory recurrent neural networks," *Int. J. Comput. Appl.* **143**, 7–11 (2016).
48. I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, (MIT Press, Cambridge, MA, USA, 2014), NIPS'14, p. 3104–3112.
49. B. K. Iwana and S. Uchida, "An empirical survey of data augmentation for time series classification with neural networks," *PLoS One* **16**(7), e0254841 (2021).
50. A. Altmann, L. Tolosi, O. Sander, *et al.*, "Permutation importance: a corrected feature importance measure," *Bioinformatics* **26**(10), 1340–1347 (2010).

51. T. Altmann and G. Casalicchio, *Limitations on Interpretable Machine Learning Methods* (2019). [https://slds-lmu.github.io/iml\\_methods\\_limitations/pfi-correlated.html](https://slds-lmu.github.io/iml_methods_limitations/pfi-correlated.html).
52. T. Parr, K. Turgutlu, C. Csiszar, *et al.*, “Beware default random forest importances,” (2018). <https://explained.ai/rf-importance/index.html>.
53. M. Abadi, A. Agarwal, and P. Barham, “TensorFlow: Large-scale machine learning on heterogeneous systems,” (2015). Software available from [tensorflow.org](https://tensorflow.org).
54. D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv* (2014).
55. M. Padgett and L. Allen, “The poynting vector in laguerre-gaussian laser modes,” *Opt. Commun.* **121**(1-3), 36–40 (1995).
56. T. Jaouni, L. Scarfe, F. Bouchard, *et al.*, <https://github.com/TareqJ1000/TAROQO>.