# Scheduling space-to-ground optical communication under cloud cover uncertainty

Mateusz Polnik, Ashwin Arulselvan, and Annalisa Riccardi

***Abstract*—Any reliable model for scheduling optical space-to-ground communication must factor in cloud cover conditions due to attenuation of the laser beam by water droplets in the clouds. In this work, we provide two alternative models of uncertainty for cloud cover predictions: a Robust Optimisation model with a polyhedral uncertainty set and a Distributionally Robust Optimisation model with a moment-based ambiguity set. We computationally analyse their performance over a realistic communications system with one satellite and a network of ground stations located in the United Kingdom. The models are solved to schedule satellite operations for six months utilising cloud cover predictions from official weather forecasts.**

**We found that the presented formulations with the treatment of uncertainty outperform in the long term models in which uncertainty is ignored. Both treatments of uncertainty exhibit similar performance. Nonetheless, the novel variant with the polyhedral uncertainty set is considerably faster to solve.**

***Index Terms*—Cloud Cover Uncertainty, Satellite Quantum Key Distribution, Scheduling Satellite Operations, Robust Optimisation, Distributionally Robust Optimisation.**

## I. Introduction

Radio Frequency (RF) communication has been the primary technology for space-to-ground communication over the last decades [1]. Its success is partly because short radio waves belong to the part of the electromagnetic spectrum that is not absorbed by the Earth's atmosphere. Therefore, these wavelengths can pass through clouds unimpeded and not affected by direct sunlight. Nonetheless, the users of the RF communications systems must deal with several inherent restrictions. The maximum throughput of the RF X-band communications system for small satellites saturates at 800 megabits per second [2], which is a bottleneck for transmission of high-resolution data collected during Earth observation missions [1]. Furthermore, the available bandwidth frequencies are limited and subject to licensing. Finally, the size of components necessary to build an RF communications system and their energy utilisation requirements are constraining for the design of small satellites, which are cheaper to assemble and launch [3].

Optical communications systems [4] do not suffer from such limitations. A laser beam allows for several orders of magnitude faster data transfer than radio. Besides, transmission devices are suitable for miniaturisation and require less energy [5]. Similarly, there is no need for bandwidth licensing because a laser beam has a small circumference and can be received in a limited area [5]. Moreover, possible interference with a signal leaves detectable traces in the message. Consequently, the receiver can recognize that information sent from the satellite may have been compromised or eavesdropped. This property is paramount for the Satellite Quantum Key Distribution (SatQKD) application [6], which involves using a spacecraft to deliver cryptographic keys to a network of Optical Ground Stations (OGS).

Despite its potential, the use of optical communications systems has some obstacles. As light is susceptible to diffraction, clouds containing droplets of water attenuate the laser beam, which may render the communication impossible in adverse weather conditions [5]. Unsurprisingly, accounting for weather conditions in planning space-to-ground optical communication results in more reliable schedules with less disruption [7].

The quantum-secured communications system considered here consists of one satellite and a network of OGS. The satellite is equipped with a quantum source (a laser), a quantum random number generator and an optical assembly for precise pointing. We assume that the satellite operates in downlink configuration, as the transmitter, and in a trusted-node architecture, free from malicious attacks. The satellite communicates to the OGS sequentially to transmit the private key. Then it broadcasts the XOR hash of keys pair over a public channel. The OGS can use the XOR hash and their private keys to decrypt the public keys. Uplink configurations, where the satellite is the receiver, and an untrusted node architecture, where entangled photon pairs are sent simultaneously to a pair of OGS, have also been studied in SatQKD but they are out of the scope of this paper. The SatQKD communications system has been chosen as an applicative case study for the methodology we proposed here because it is an emerging application and it has been studied as an optimisation problem for the first time by us in [8]. Nevertheless, the only assumptions in the following study are that the communications system has a single satellite operating in downlink configuration and adversely affected by cloud coverage.

### A. Contributions

We listed noteworthy contributions of this paper below.

1) We developed a novel approach to build a polytope uncertainty set for the Robust Optimisation framework, which combines a box uncertainty set with a vector autoregressive time series model. The modelling structure preserves both geospatial and temporal correlation of uncertain parameters which occur naturally in weather conditions. We used the framework to solve the SatQKD

M. Polnik and A. Riccardi are with the Department of Mechanical & Aerospace Engineering, University of Strathclyde, Glasgow.

A. Arulselvan is with Strathclyde Business School, Glasgow.

Manuscript received April 8, 2021, 2020.

problem with uncertain cloud cover and compared the results to Distributionally Robust Optimisation with a second-order conic ambiguity set. The proposed approach provides similar performance but has a significantly shorter computation time. Despite the targeted application, one could adapt the solution framework to other variants of scheduling satellite communication, e.g., the maximisation of the data volume downloaded, by changing the formulation of the objective function.

2) We provide several examples on how to exploit the information available in the weather forecast to build schedules more resilient to disruptions caused by cloud cover. To the best of our knowledge, it is the first work in the context of satellite communication to use official weather forecasts to model uncertainty and real weather observations to evaluate the performance of schedules.

### B. Paper Structure

The paper is organised as follows. Section II presents the literature review. Section III introduces the Satellite Quantum Key Distribution optimisation problem and illustrates alternative formulations of cloud cover uncertainty. Section IV is devoted to the analysis and discussion of computational results. The last section concludes the paper. All proofs were moved to the Appendix.

## II. Literature Review

In the literature review, we cover modelling techniques to represent uncertain parameters in mathematical programs and the formulations used for scheduling satellite operations.

The techniques used to handle uncertainty in mathematical programming differ in the way uncertainties are modelled and treated in the problem. If a decision-maker assumes to know exactly the probability distribution of the uncertain parameters, then Stochastic Optimisation (SO) can be applied to model and solve the problem [9]. Sample Average Approximation (SAA) [10], [11] and the Conditional Value at Risk (CVaR) [12], [13] can be used to approximate the evaluation of objective and constraints. However, the complexity of these methods depends on the number of scenarios considered. Furthermore, SO is not suitable when little is known about the probability distribution. A more conservative approach that assumes no a priori knowledge on the probability distribution of the uncertain parameters is Robust Optimisation (RO) [14], [15] focusing on the worst-case scenario. The domain of the uncertain parameter in RO is declared using an uncertainty set, and the satisfaction of constraints which involve uncertain parameters is required for all possible values they may assume. Such systems of constraints known as a Robust Counterpart are semi-infinite if the uncertainty set is continuous. Hence the initial RO model is reformulated using duality into an equivalent formulation with a finite number of constraints [16] [17].

The seminal paper of Delange and Ye [18] rekindled the interest in the Distributionally Robust Optimisation (DRO). In this framework, the decision-maker models a family of probability distribution functions formulating constraints on the properties the true but unknown probability distribution exhibits. Such a description is called an ambiguity set. An essential building block of the ambiguity set definition is the support set of the probability distribution, which corresponds to the uncertainty set in the RO framework. The ambiguity set proposed by [18] constraints the mean of the probability distribution to lie within an ellipsoid from the nominal value and bounds from above the covariance matrix using a Positive-Semidefinite (PSD) constraint. The researchers have shown that the worst-case expectation over such an ambiguity set can be computed in polynomial time under reasonable assumptions, i.e., the convexity of the expression inside the expectation [18].

Subsequent researchers made efforts to devise unified, generic and practically solvable DRO frameworks with moment-based ambiguity sets. For instance, the approach of [19] targets the Second Order Conic (SOC) representable ambiguity sets for Linear Programming (LP). Their ambiguity sets include a declaration of the support set, an arbitrary number of equality constraints with an expectation of an affine expression, and constraints with an expectation of a SOC representable function bounded from above. The SOC representable ambiguity sets can also describe higher-order moments by the introduction of additional uncertain variables, which is known as lifting. Overall, an LP with a SOC ambiguity set has an equivalent representation as a classical RO model using the duality theory for moment problems [20]. Consequently, the ambiguity sets with upper bounds on the second marginal or partial moments have SOC reformulations and can be solved using commercial MIP solvers. A similar, albeit more generic framework supporting PSD constraints (i.e., to bound a covariance matrix from above) was proposed by [21].

An alternative approach to building ambiguity sets is to define a ball in the space of probability distribution functions using some distance metric for probability distributions such as the Kullback-Leibler [22] divergence or the Wasserstein metric [23]. The probability distribution in the centre of the sphere is inferred from training samples, and the radius controls conservativeness. To differentiate from the moment-based ambiguity sets, the former approach is named Data-Driven DRO. Theoretical results motivate the use of metric-based ambiguity sets. For the Wasserstein ambiguity set, one can derive probabilistic guarantees that the sphere of the given radius contains the true probability distribution which generated the data sample [23]. Moreover, probabilistic guarantees can be computed for out-of-sample performance. Nonetheless, in general, data-driven approaches to modelling ambiguity sets at the current stage are considered less scalable than moment-based formations [23]. For that reason, we used the latter approach to model ambiguity sets in this work.

The dominating approach to scheduling of space-to-ground communication in the literature is to discretise time and formulate the scheduling problem as a mixed-integer linear program which maximises the volume of data transferred to a fixed network of ground stations [24], [25] or the revenue accumulated for the accomplished tasks [26]. Such formulations have been applied to problems with single [24] or multiple

satellites [25], [26] placed in orbits which remain unchanged over the scheduling horizon. Realistic problem instances can be solved to optimality within seconds due to a tight bound of the LP relaxation [24], [25], i.e., 1-2% at the root of the branch and bound tree [8]. For a thorough overview of the deterministic methods for scheduling satellite operations and other related optimisation problems, we refer an interested reader to our previous paper [8].

Besides scheduling space to ground communication, cloud cover is of primary importance for image acquisition using the Earth Observation Satellite (EOS). Understandably, clouds obscuring the target in the satellite image may require repeating the image acquisition task. Consequently, the resources spent on the unsuccessful attempt are wasted, and the need to redo the task in the future causes disruptions. The first treatment of cloud cover uncertainty in EOS scheduling is due to [27]. The authors formulated the problem as a stochastic program with a rolling horizon [28] that maximised the expected reward for the observations performed and the number of the accomplished tasks. The execution of a given task may fail due to adverse weather conditions modelled assuming the Bernoulli probability distribution. As a result, the expectation in the objective had a closed-form expression. An alternative formulation in which minimum profit is subject to a chance constraint was proposed by [29]. The author also assumed the Bernoulli probability distribution and approximated the chance constraint using SAA. In a recent paper, [30] considered scheduling of multiple EOS with uncertain rewards for the acquisition of an image modelled using the budgeted uncertainty set [31].

Regarding wireless space-to-ground and terrestrial communication, optimisation has also been used for beamforming (BF) [32]. This radio communication technique utilises multiple antennas simultaneously to control the direction and the shape of the beam by applying constructive or destructive interference between signals emitted by individual antennas. As a result, BF can improve the transfer rate and reduce the area in which the signal can be received, causing less interference. The authors [33] considered the optimisation problem of finding a set of beamformer weights for antenna arrays which maximise the aggregate transfer rate subject to domain-specific constraints. They prevented potential eavesdroppers from capturing the signal, enforced the power consumption budget, controlled the interference and noise. The angle between the transmitting antenna and the direction of the beam (a.k.a. the angle of departure) for the terrestrial communication and the channel error for the space-to-ground communication were sources of uncertainty. The authors modelled them using uncertainty sets. Since the resulting optimisation problem was non-convex, the researchers applied a series of transformations (i.e., a successive convex approximation, a Tailor expansion, and sampling the uncertainty set) to substitute the non-convex expressions by their convex equivalents. An alternative variant of the problem with the objective function focused on mitigating the risk of eavesdropping communication has been explored in [34]. In a similar vein, a deterministic variant of the problem for the multicast space-to-ground communication and the terrestrial communication with clustering of end-users to improve the usage of available spectrum was studied by [35].

To conclude, the techniques adopted for modelling uncertain cloud cover in satellite scheduling applications were limited to RO and SO, in which researchers assumed the probability of accomplishing a mission objective in given weather conditions. Alternative modelling techniques, such as DRO have not been studied for scheduling satellites. Finally, the results discussed in the literature do not use official weather forecasts, and their performance is not verified in real weather conditions.

## III. Models

### A. Problem Description

The SatQKD problem has been formulated for the first time in our previous work on deterministic scheduling using historical weather observations [8]. It aims to schedule a sequence of space-to-ground data transfers within some prescribed scheduling horizon. During a data transfer, a satellite generates cryptographic keys, defined as 256 bit of randomness, and sends them to a ground station. The ground stations use the keys to establish secure duplex connections which require each party to contribute one key. After the connection is closed, the keys are removed as their reuse is prohibited for security reasons. The objective is to supply a ground station with keys in proportion to its importance, expressed as a weight, in the communications system. Otherwise, if some ground stations stockpile all keys, the remaining communication nodes will not be able to establish a secure connection. As a result, the network will become effectively disconnected. We assume the keys a ground station received during a given scheduling horizon become available for use in the next scheduling horizon. Therefore, the objective function can be evaluated once at the end of the current scheduling horizon.

Several external conditions which are consequences of physics and technological limitations must be satisfied to perform a successful data transfer:

1) To establish an optical connection, the ground station must be visible from space to the satellite. Specifically, the keys can be transferred when the elevation angle between the satellite and the ground station is at least $15^\circ$. The lower the elevation angle, the longer distance the laser beam needs to travel in the atmosphere. Consequently, we assume that below $15^\circ$, the errors are too high for a successful data transfer to happen.
2) The position of the satellite in the sky at a given time was computed using an approximate analytical model presented in [8]. Therefore, the occurrence and the duration of communication windows is realistic. Furthermore, the connection between the satellite and a ground station can be established when neither the ground station nor the satellite is in sunlit. A satellite passing over the UK mainland territory at night hours between late September and March remains in the Earth shadow.
3) The number of keys a satellite can send to a ground station in the cloud-free line of sight is estimated using the model of the optical communication in the atmospheric channel described in [8] which is parametrised by the

elevation angle between the satellite and the ground station. For ease of exposition, we assume a satellite sends keys at a rate a ground station can receive them, and each ground station has the same communication capabilities.

4) For simplicity, we assume that the satellite can observe one ground station at a time. The observation is defined as a single continuously attempted optical connection between the satellite and a given ground station. The transfer of keys to another ground station can start after some idle time since stopping observations of the previous ground station (i.e., 30 seconds). This period is spent on activities required to establish a new connection, e.g. the recalibration of the optical apparatus.

Before outlining the deterministic formulation of the problem, we introduce the following symbols.

### B. Notations

We are given the set $N := \{1, \ldots, n\}$ representing "regular" ground stations. $\bar{N} := N \cup \{0\}$ is an extended variant of $N$ containing the auxiliary ground station 0, which we use to model the mandatory transition period that the satellite should go through before it switches ground stations for transmission.

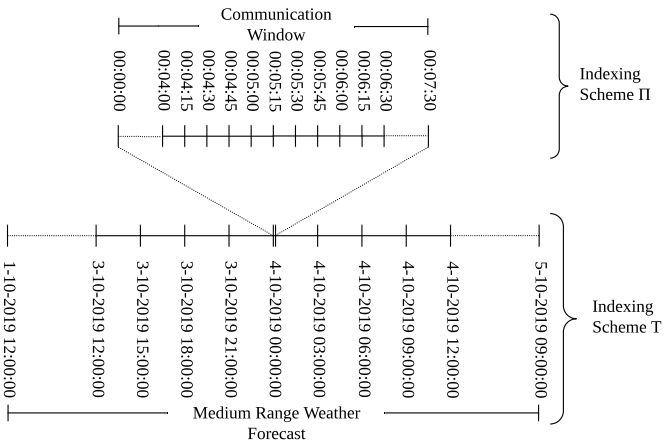


Fig. 1. Indexing schemes employed in the problem definition. The indexing scheme $\Pi$ is discontinuous and covers communication windows between the spacecraft and ground stations. Each night, depending on the geographical location of the ground station, there are either one or two visibility periods observed by the ground stations which are not fully covered by clouds. The indexing scheme $T$ is continuous and covers the entire scheduling horizon. Its resolution depends on the frequency of cloud cover prediction changes in a weather forecast.

Figure 1 illustrates the time indexing schemes employed to formulate the problem. We are given a scheduling horizon of $T$, and we discretise it using two schemes. The same symbol is used to represent the set and the time interval it discretises. However, the notation will be clear from the context. In the coarse index scheme, we have a fully ordered set of time intervals $T := \{1, \ldots, T\}$, with a discretisation interval $t \in T$ corresponding to a few hours in the time horizon. It models the weather forecast update frequency assumed by the decision-maker, e.g., 3 hours. For each coarse-grained time interval $t \in T$, we have an ordered set for the fine-grained indexing scheme, $\Pi_t := \{|\Pi_{t-1}| + 1, |\Pi_{t-1}| + 2, \ldots, |\Pi_t|\}$, and $\Pi_0$ is an empty set. The time intervals in these sets have a resolution in the range of seconds, e.g., 30 seconds. It is the period after which we update the position of the satellite in the sky, which affects the throughput of the satellite connection. Later in Section IV-F, we discuss the issues surrounding the adjustment of this resolution. We now have $\Pi = \bigcup_{i=1}^{T} \Pi_i$ and we index this set with $\pi$. Another way of looking at the schemes is that $T$ provides a partition of $\Pi$. The coarse-grained indexing scheme covers the entire scheduling horizon, whereas the fine-grained indexing scheme maps the periods when communication with at least one ground station is possible. For some ground station $n \in N$, the notation $\Pi^n$ restricts the index to only those time intervals in which communication with the ground station $n$ is possible, i.e, $\Pi^n := \{\pi \in \Pi : k_{n\pi} > 0\})$, where $k_{n\pi}$ denotes the set of keys that can be transmitted in the time interval $\pi$ to the ground station $n$. Furthermore, for some coarse-grained time interval $t \in T$, the notation $\Pi_t^n$ additionally discards the time intervals $\pi \in \Pi^n$ which are not contained in $t$, i.e., $\Pi_t^n := \Pi^n \cap \Pi_t$.

The two-dimensional vector $\mathbf{x} \in \mathbb{B}^{|\bar{N}| \times |\Pi|}$ contains binary decision variables that indicate whether a given ground station is observed by the satellite at a given period. The vector $\mathbf{c} \in [0,1]^{|N| \times |T|}$ contains parameters representing uncertain cloud cover. We consider several formulations which adopt different approaches to model that uncertainty. The vector of constants $\mathbf{k} \in \mathbb{R}_{\geq 0}^{|N| \times |\Pi|}$ stores the number of keys which can be delivered to a given ground station over a given period in the cloud-free line of sight [8]. If clouds are present, we assume that the number of keys possible to deliver decreases linearly as cloud cover increases, e.g., for some $t \in T$ and $\pi \in \Pi_t$ such that the effective number of keys delivered to the ground station $n \in N$ during the period $\pi$ equals $(1 - c_{nt})k_{n\pi}$. This approximation scheme is simplistic but commonly assumed in the optimisation of the optical communications systems [1] and the analysis of solar irradiance [36]. The constant $\omega_n$ represents the weight assigned to the ground station $n$. Finally, the constant $\beta_n$ indicates the number of keys a ground station stores in its buffer at the beginning of the scheduling horizon.

### C. Feasible Schedule

A valid schedule satisfies the following constraints.

$$\sum_{n \in \bar{N}} x_{n\pi} \leq 1 \qquad \forall \pi \in \Pi \tag{1}$$

$$x_{n\pi} \leq x_{n\pi-1} + x_{0\pi-1} \qquad \forall \pi \in \Pi \backslash \{1\}, \quad \forall n \in N \tag{2}$$

$$\mathbf{x} \in \mathbb{B}^{|\bar{N}| \times |\Pi|} \tag{3}$$

Constraint (1) ensures in each period the satellite either sends keys to at most one ground station or remains in the transition phase preceding a transfer of keys to the next ground station. Constraint (2) states that if the satellite sends keys to some ground station at a given time interval, then the same ground station received keys in preceding time interval or the satellite was going through the transition phase. Finally, Constraint (3) enforces the binary restrictions. Overall,

Constraints (1 - 3) are the building blocks employed in the rest of the models we discuss. Therefore, for conciseness, we declare them using the notation below.

$$\mathcal{X} := \{\mathbf{x} \in \mathbb{B}^{|\bar{N}|\times|\Pi|} : \mathbf{x} \text{ satisfies (1) and (2)}\}$$

The following expression, denoted as $\Lambda(\mathbf{x}, \mathbf{c})$, is another component present in all models either as the objective or a constraint:

$$\Lambda(\mathbf{x}, \mathbf{c}) = \min_{n\in N} \{\lambda_n(\mathbf{x}, \mathbf{c})\}$$

where $\lambda_n(\mathbf{x}, \mathbf{c}) := \frac{1}{\omega_n}\left(\beta_n + \sum_{t\in T}\sum_{\pi\in\Pi_t^n}(1 - c_{nt}) \cdot k_{n\pi}x_{n\pi}\right)$. For a given feasible schedule $\mathbf{x}$ and cloud cover information $\mathbf{c}$, the expression returns a non-negative real number, which we call a scaling factor and denote using $\lambda := \Lambda(\mathbf{x}, \mathbf{c})$. Note that the scaling factor is proportionate to the least number of keys received by any ground station. From the definition of $\Lambda$, it follows that for each ground station $n \in N$, the expression $\omega_n\lambda$ bounds from below the total number of keys the ground station $n$ would store at the end of the scheduling horizon if no keys were consumed. Intuitively, an increase in $\lambda$ indicates a growth in the number of keys across all ground stations in proportion to their weight. Consequently, the maximisation of $\Lambda(\mathbf{x}, \mathbf{c})$ as the objective for scheduling transfers of keys avoids extreme situations in which one ground station stockpiles the majority of keys that cannot be used because other ground stations do not have a sufficient number of keys to communicate.

### *D. Deterministic Model*

The deterministic model of the SatQKD presented below was first proposed in [8].

$$\max_{\mathbf{x}\in\mathcal{X}} \Lambda(\mathbf{x}, \mathbf{c}') \tag{DSF}$$

The vector $\mathbf{c}'$ stores cloud cover predictions. This simplistic model treats the predictions as deterministic values, which are uncertain parameters in real-world.

The DSF problem always has a feasible solution. For instance, a schedule with no data transfers is a valid solution. The objective value for such a schedule, $\min_{n\in N}\{\beta_n/\omega_n\}$, gives us a lower bound. One can compute a trivial upper bound by relaxing the constraint $\mathbf{x} \in \mathcal{X}$. It corresponds to a scenario in which the satellites can transfer keys to all ground stations simultaneously and with no setup time. Consequently, the objective value is bounded from above by the expression $\min_{n\in N} \lambda_n(\mathbf{1}, \mathbf{c}')$, where $\mathbf{1}$ is a vector with all its entries as 1.

The complexity of the SatQKD problem is stated in Theorem 1.

**Theorem 1.** *The deterministic SatQKD problem is strongly NP-hard.*

Proof in the Appendix A.

As a consequence from the theorem, no Fully Polynomial Time Approximation Scheme is available for the SatQKD problem unless P = NP. Since looking for a dedicated optimisation algorithm is futile, we will solve the formulation directly using a MIP solver.

The performance of the deterministic model strongly relies on accurate cloud cover predictions. Unfortunately, our experience indicates that prediction errors for the UK grow significantly with the length of the forecasting horizon. For that reason, we also consider a variant of the problem with the rolling horizon in which we update the weather forecast at noon every day and solve the SatQKD problem for the remaining period. We refer to the model using the DRH tag. The formulation benefits from the opportunity to assess past decisions based on weather conditions observed in real-world and to reconsider future decisions using more recent and thus often more accurate weather forecast. Even when the forecasting models provide accurate predictions, the deterministic scheduling decisions made based on point estimates fair poorly in comparison with the schedules that account for uncertainty as we will see in Section IV.

There are specific applications for which deterministic models would prove to be sufficient. Given historical weather observations, such a model could be used to evaluate a design of the future communications system [8]. For instance, to decide where the ground stations should be established, how many of them should be operational, finally, which configuration of the satellite orbit should be selected. Nonetheless, when operating a communications system in real-world, cloud cover predictions are provided in the form of a weather forecast, and their accuracy decreases significantly over time. It motivates our efforts to develop models which treat cloud cover as an uncertain parameter.

### *E. Robust Optimisation Model*

To include information about changes of cloud cover conditions over time and space, we assume cloud cover follows a Vector AutoRegression (VAR) process [37], which is an example of a multivariate time series model, wherein the uncertainty is expressed as a linear combination of past observations. The VAR process has been a popular choice to model various meteorological phenomena, such as temperature and precipitation [38].

The VAR process of the order (lag) $h$ is defined below.

$$\mathbf{c}_t = \mathbf{A}_1\mathbf{c}_{t-1} + \mathbf{A}_2\mathbf{c}_{t-2} + \cdots + \mathbf{A}_h\mathbf{c}_{t-h} + \mathbf{a}_0 + \epsilon$$

The vector $\mathbf{c}_t \in \mathbb{R}^{|N|}$, for some $t \in T$, contains the time series variables corresponding the cloud cover for all ground stations at time $t$. The time-invariant matrices $\mathbf{A}_1, \ldots, \mathbf{A}_h \in \mathbb{R}^{|N|\times|N|}$ and the vector $\mathbf{a}_0 \in \mathbb{R}^{|N|}$ are the parameters of the model. Finally, $\epsilon$ is a white noise process. Taken together, according to the VAR(h) model, the present state is a sum of the linear combination of the past $h$ states and white noise.

The coefficients of the model for the given order are computed in the fitting procedure. The implementation [39] we have employed, estimates the coefficients using the ordinary least squares method, and the lag is set according to one of the supported decision criteria to create a parsimonious model, i.e., the minimisation of the Bayesian Information Criterion (BIC) [40].

We fitted model parameters and inferred the lag using BIC minimisation on a training dataset of historical cloud cover observations in 2018. Overall, the tuning procedure suggested setting the lag to one, which agrees with the results presented by other researchers. For instance, cloud cover and sunshine can be predicted using Markov processes [36], in which the next state depends exclusively on the present state and the transition matrix. Similarly, [38] used a VAR model of order one to simulate temperature and precipitation. Consequently, hereafter, we assume the order one, this simplifies the reformulation, i.e., we stripped subscripts from the vector $a$ and the matrix $A$.

Let us define the uncertainty set $\mathcal{V}$, which combines the VAR model with interval uncertainty.

$$\mathcal{V} = \left\{ \mathbf{c} \in \mathbb{R}^{|N|\times|T|} \;\middle|\; \begin{array}{ll} \underline{\mathbf{c}} \le \mathbf{c} \le \bar{\mathbf{c}} & \\ \mathbf{c}_t - \mathbf{e} \le \mathbf{a}_0 + A\mathbf{c}_{t-1} & \forall t \in T\backslash\{1\} \\ \mathbf{a}_0 + A\mathbf{c}_{t-1} \le \mathbf{c}_t + \mathbf{e} & \forall t \in T\backslash\{1\} \end{array} \right\}$$

In the description of the uncertainty set $\mathcal{V}$, we replaced $\epsilon$, which denotes the white noise process in the VAR definition, by the vector of residuals $\mathbf{e} \in \mathbb{R}^{|N|}_{\ge 0}$. The vector of residuals performs a similar role in controlling the degree of conservativeness as the parameter $\Gamma$ in the budgeted uncertainty set [41], [31]. In the latter uncertainty set, the uncertain parameters can assume values whose sum of absolute deviations from nominal values does not exceed $\Gamma$ regardless of the correlations between the parameters. On the other hand, in our model, the values of uncertain parameters must fit into the VAR model. Therefore, preserving spatial and temporal correlations between random variables, which is desirable in modelling weather. In this study, we estimated the vector of residuals by analysing how precisely the VAR model trained using historical weather observations explains the changes in cloud cover predictions. See Section IV-C for more details on our approach.

The SatQKD problem with the $\mathcal{V}$ uncertainty set is defined below.

$$\max_{\mathbf{x}\in\mathcal{X}} \min_{\mathbf{c}\in\mathcal{V}} \; \Lambda(\mathbf{x}, \mathbf{c}) \tag{BTS}$$

Let us make a couple of remarks here. (i) A chance-constrained program would be a direct extension of DSF to incorporate uncertainty:

$$\max t : \text{Prob}_{\mathbf{c}\sim P}\{\Lambda(\mathbf{x}, \mathbf{c}) \ge t\} \ge 1 - \epsilon \tag{4}$$

This requires knowledge of the distribution, $P$, of the uncertain $\mathbf{c}$. Even if $P$ is known and simple, it is numerically tedious to evaluate these constraints. Our RO model, BTS, provides an analytical approximation to (4). We do not discuss other models that approximates (4), such as CVaR ([12], [13]), that were significantly outperformed by BTS in our experiments. (ii) Any statistical model could be used to construct the uncertainty set $\mathcal{V}$, but our choice of the VAR model allows a polyhedral representation. This permits the problem to be reformulated as a mixed-integer linear program. We formalise this in the following Proposition.

**Proposition 1.** *The SatQKD problem with the uncertainty set $\mathcal{V}$ has the following reformulation.*

$$\begin{aligned}
& & \max \Theta & & \\
\textit{s.t.:} \quad & \omega_n\Theta - \beta_n - \sum_{t\in T}\sum_{\pi\in\Pi^n_t} k_{n\pi}x_{n\pi} & & & \\
& \quad + \sum_{t\in T} \bar{c}_{nt}\bar{z}_{nt} - \sum_{t\in T} \underline{c}_{nt}\underline{z}_{nt} & & & \\
& \quad - \sum_{t\in T\backslash\{1\}} (a-e)\bar{v}_{nt} & & & \\
& \quad + \sum_{t\in T\backslash\{1\}} (a+e)\underline{v}_{nt} \le 0 & & & \forall n \in N \\
& \bar{\mathbf{z}}_t - \underline{\mathbf{z}}_t + A^\top\bar{\mathbf{v}}_{t+1} - A^\top\underline{\mathbf{v}}_{t+1} & & & \\
& -\bar{\mathbf{v}}_t - \underline{\mathbf{v}}_t - \sum_{\pi\in\Pi^n_t} k_{n\pi}x_{n\pi} = 0 & & \forall n \in N & \forall t \in T\backslash\{1\} \\
& & \mathbf{x} \in \mathcal{X} & & \\
& & \bar{\mathbf{v}}, \underline{\mathbf{v}}, \bar{\mathbf{z}}, \underline{\mathbf{z}} \in \mathbb{R}_{\ge 0}{}^{|N|\times|T|} & &
\end{aligned}$$

*where $\bar{\mathbf{v}}, \underline{\mathbf{v}}, \bar{\mathbf{z}}, \underline{\mathbf{z}}$ are the dual multipliers of the inequalities in $\mathcal{V}$ and $\Theta$ is the auxiliary variable representing the objective function in the dual reformulation, with $\Theta \le \lambda_n(\mathbf{x}, \mathbf{c})$ for all $\mathbf{c} \in \mathcal{V}$ and some ground station $n \in N$.*

Proof in the Appendix B.

### F. Distributionally Robust Optimisation Model

The model we have studied so far makes no assumptions about the probability distribution. Nonetheless, some of its properties as mean or variance could be reliably estimated given a data sample of sufficient size. Here, we use the DRO framework proposed by [19] in which ambiguity sets are SOC representable. The ambiguity sets we define are based on the marginal moments, and we are not incorporating cross-moment estimations. This tends to be a conservative model, but we are providing it for benchmarking purposes and including cross-moment could be future research work. We use the lifted ambiguity sets, which contain the auxiliary variables $\mathbf{u} \in \mathbb{R}^{N\times T}$. For $n \in N, t \in T$, and a function defined as $g_{nt} := (c_{nt} - m_{nt})^2$, where $m_{nt}$ is the mean value of the random parameter $c_{nt}$, its epigraph is given by:

$$\text{epi } g_{nt} = \{(c_{nt}, u_{nt}) \in \mathbb{R}\times\mathbb{R} \;\mid g_{nt}(c_{nt}) \le u_{nt}\}$$

Let us formulate the support set $\mathcal{W}$ and the ambiguity set $\mathbb{F}$.

$$\mathcal{W} = \left\{ (\mathbf{c}, \mathbf{u}) \in \mathbb{R}^{N\times T} \times \mathbb{R}^{N\times T} \;\middle|\; \begin{array}{l} \underline{\mathbf{c}} \le \mathbf{c} \le \bar{\mathbf{c}} \\ (c_{nt} - m_{nt})^2 \le u_{nt} \\ \quad \forall n \in N \;\forall t \in T \end{array} \right\}$$

$$\mathbb{F} = \left\{ \mathbb{P} \in \mathcal{P}_0(\mathbb{R}^{N\times T} \times \mathbb{R}^{N\times T}) \;\middle|\; \begin{array}{l} (\tilde{\mathbf{c}}, \tilde{\mathbf{u}}) \sim \mathbb{P} \\ \mathbb{E}_{\mathbb{P}}(\tilde{\mathbf{c}}) = \mathbf{m} \\ \mathbb{E}_{\mathbb{P}}(\tilde{\mathbf{u}}) \le \mathbf{d} \\ \mathbb{P}((\tilde{\mathbf{c}}, \tilde{\mathbf{u}}) \in \mathcal{W}) = 1 \end{array} \right\}$$

The vector $\mathbf{d} \in \mathbb{R}_{\geq 0}{}^{|N|\times|T|}$ stores upper bounds of the variance, which will be estimated from data. We relax the equality restriction on the second moment as this permits the SOC reformulation.

The problem SatQKD with the ambiguity set $\mathbb{F}$ defined on the support set $\mathcal{W}$ is formulated below.

$$\max_{\mathbf{x}\in\mathcal{X}} \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}(\Lambda(\mathbf{x}, \tilde{\mathbf{c}})) \tag{MSTD}$$

Let us transform the formulation into a classical RO problem by applying Theorem 1 of Bertsimas, Sim and Zhang [19]. We consider the MSTD as a two-stage optimisation problem in which decisions $\mathbf{x} \in \mathcal{X}$ are made first, and the scaling factor $\lambda$ is computed after uncertain parameters $\tilde{\mathbf{c}}$ are observed.

The problem $\Lambda(\mathbf{x}, \mathbf{c})$ is always feasible, regardless of input $\mathbf{x}$ and $\mathbf{c}$. Therefore, we conclude the problem $\Lambda(\mathbf{x}, \mathbf{c})$ has a complete recourse as required by [19]. By applying the aforementioned theorem we reformulate the MSTD problem for the given assignment of decision variables $\mathbf{x}$ as follows.

$$\begin{aligned} R(\mathbf{x}) = \max & f + \mathbf{g}^\top \mathbf{m} - \mathbf{h}^\top \mathbf{d} \\ \text{s.t.:} \quad & f + \mathbf{g}^\top \mathbf{c} - \mathbf{h}^\top \mathbf{u} \leq \Lambda(\mathbf{x}, \mathbf{c}) \quad \forall (\mathbf{c}, \mathbf{u}) \in \mathcal{W} \\ & f \in \mathbb{R},\ \mathbf{g} \in \mathbb{R}^{|N|\times|T|},\ \mathbf{h} \in \mathbb{R}_{\geq 0}{}^{|N|\times|T|} \end{aligned}$$

The vectors of dual multipliers $\mathbf{g}$ and $\mathbf{h}$ replaced the expectation operators from the ambiguity set $\mathbb{F}$. Now, let us embed the definition of the second stage problem $\Lambda(\mathbf{x}, \mathbf{c})$ into the formulation below.

$$\begin{aligned} R(\mathbf{x}) = \max & f + \mathbf{g}^\top \mathbf{m} - \mathbf{h}^\top \mathbf{d} && (5) \\ \text{s.t.:} \quad & f + \mathbf{g}^\top \mathbf{c} - \mathbf{h}^\top \mathbf{u} \\ & \qquad \leq \lambda_n(\mathbf{x}, \mathbf{c}) \quad \forall (\mathbf{c}, \mathbf{u}) \in \mathcal{W}, \quad \forall n \in N && (6) \\ & f \in \mathbb{R}, \quad \mathbf{g} \in \mathbb{R}^{|N|\times|T|}, \quad \mathbf{h} \in \mathbb{R}_{\geq 0}{}^{|N|\times|T|} \end{aligned}$$

The last model is a classical RO problem. However, let us remark that Constraint (6) exists in $|N|$ copies, i.e., one instance of the constraint for a ground station. Each such a constraint contains the summation term of uncertain cloud cover across all ground stations. Therefore, the uncertainty is not constraint-wise [16]. Instead of transforming the formulation into constraint-wise uncertainty, we solve the MSTD formulation using the adversarial method [42]. It is an iterative approach in which the uncertainty set $\mathcal{W}$ is replaced by a finite number of samples from that set. New samples are obtained from the uncertainty set on-demand only if the solution computed in the last iteration violates some constraints for those samples.

Formally, let $\hat{\mathcal{W}}$ denote the set of samples from $\mathcal{W}$. The adversarial method iteratively solves the following mixed-integer linear program to obtain a solution to the MSTD problem.

$$\begin{aligned} & \max f + \mathbf{g}^\top \mathbf{m} - \mathbf{h}^\top \mathbf{d} && (7) \\ \text{s.t.:} \quad & f + \mathbf{g}^\top \bar{\mathbf{c}} - \mathbf{h}^\top \bar{\mathbf{u}} \\ & \qquad \leq \lambda_n(\mathbf{x}, \bar{\mathbf{c}}) \quad \forall (\bar{\mathbf{c}}, \bar{\mathbf{u}}) \in \hat{\mathcal{W}}, \quad \forall n \in N && (8) \\ & \mathbf{x} \in \mathcal{X} \\ & f \in \mathbb{R}, \quad \mathbf{g} \in \mathbb{R}^{|N|\times|T|}, \quad \mathbf{h} \in \mathbb{R}_{\geq 0}{}^{|N|\times|T|} \end{aligned}$$

Then, given the assignment of the decision variables $\bar{\mathbf{x}}, \bar{f}, \bar{\mathbf{g}}, \bar{\mathbf{h}}$, for each ground station $n \in N$, we attempt to find a sample for which the current solution violates Constraint (8) the most by solving the subproblem below, which is a SOC program.

$$\max_{(\mathbf{c},\mathbf{u})\in\mathcal{W}} \bar{f} + \bar{\mathbf{g}}^\top \mathbf{c} - \bar{\mathbf{h}}^\top \mathbf{u} - \lambda_n(\bar{\mathbf{x}}, \mathbf{c}) \tag{9}$$

If the objective of the optimal solution to the subproblem is non-positive for every ground station, then Constraint (8) is satisfied for all $(\mathbf{c}, \mathbf{u}) \in \mathcal{W}$. Therefore, the solution computed by the adversarial method is feasible. Otherwise, we found some samples $(\bar{\mathbf{c}}, \bar{\mathbf{u}})$ for which the previous solution is invalid. Consequently, we extend the set $\hat{\mathcal{W}}$ by adding the new samples and commence the next iteration.

## IV. Results and Discussion

Let us present computational results obtained by solving the formulations described in the previous section. As we shortly illustrate, the advantages of the model with the rolling horizon are significant. For that reason, the variant with the rolling horizon was implemented for both formulations which handle uncertainty.

### A. Implementation

The models were developed in C++ using the Gurobi MIP solver. Data processing scripts that generate problem instances were implemented in Python. The coefficients of the VAR model were computed using the statsmodels library [39]. All software we implemented to obtain results discussed in this paper is available open-source [43].

The formulations were solved using a workstation with the AMD Ryzen 7 processor and 32 GB of RAM. The discretisation step of communication windows was 30 seconds. The stopping criterion was reaching the optimality gap limit of $10^{-2}$ or exceeding the computational time limit of 4 hours.

### B. Data

Problem instances were built using official weather forecasts and observations of actual weather conditions published by the OpenWeatherMap web service [44]. Our data set comprises 34 problem instances. They are available online for testing and benchmarking purposes [45]. Each instance requires solving the SatQKD problem for the scheduling horizon of five consecutive days, which is the span of medium-range weather forecasts. Typically cloud cover predictions are not available in weather forecasts for more extended periods. This fact seems to be the result of the maximum period for which cloud cover is predictable in the future with reasonable accuracy.

Figure 2 contrasts an example cloud cover forecast for Glasgow and cloud cover observed in real-world. In the illustrated example, cloud cover was predicted reliably only for a short time in the future, e.g., the standard deviation of cloud cover prediction error in Glasgow in the next 120 hours is over 40%.

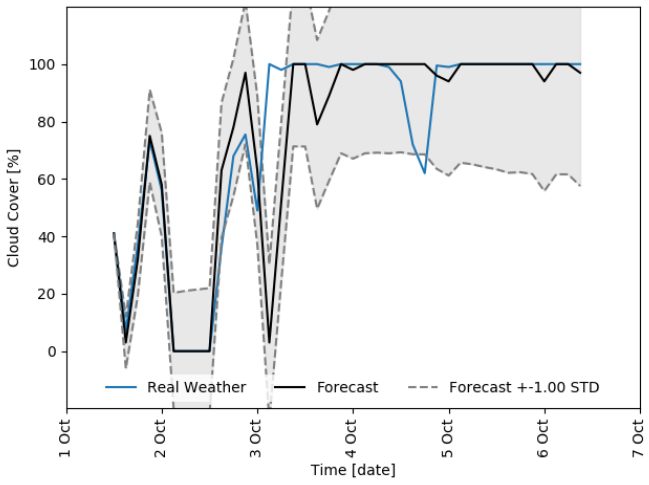


Fig. 2. Medium-range cloud cover forecast for Glasgow. The forecast was downloaded on the 1st of October 2019 at noon. The filled area between the dashed lines denotes the distance of one standard deviation of the prediction error. Standard deviation was estimated using a sample of cloud cover prediction errors calculated as the marginal difference between the cloud cover forecast and the conditions observed later in real-world between June 2019 and March 2020.

We consider solving the SatQKD problem for the period between late September 2019 and March 2020. Our previous study indicates that the communication windows are not affected by the illumination of the spacecraft by the Sun during that time frame [8]. Therefore, the weather is the sole factor affecting the performance of the communications system. In this study, our primary focus is to analyse the performance of the models in real-world. We decided to solve each problem instance independently rather than passing the number of keys delivered during the preceding scheduling horizons. Therefore, unsatisfactory performance of a given model for one problem instance does not influence the solution's performance for another scheduling horizon.

### C. Parameter Tuning

Robust and distributionally robust models contain additional parameters which control the conservativeness of the formulation. We aimed to find the configurations that maximise the sum of the scaling factors for the set of problem instances recalculated using real weather observations. For hyperparameter tuning, we used the subset of 26 problem instances picked in chronological order.

In particular, we tested different sizes of the uncertainty set/support set measured in standard deviations from the nominal value, i.e., $\tilde{u} \in [u - r \cdot \sigma, u + r \cdot \sigma]$, where $u$ is the nominal value and $r$ is the multiplicative factor of standard deviation $\sigma$. We tested 80 settings of the radius $r$ starting from 0.05 until 4.0 with the step of 0.05. Ultimately, the parameter $r$ was set to $0.45$ for the support set $\mathcal{W}$, and it was lowered to $0.2$ for the uncertainty set $\mathcal{V}$. Furthermore, we obtained the best performance by assigning the vector of residuals values that are large enough to comprise 99% of cloud cover prediction changes in weather forecasts. In case the uncertainty set $\mathcal{V}$ is empty because no realisation of uncertainty follows the VAR time series model, we seek the lowest coefficient to multiply the residual terms that makes the uncertainty set non-empty.

### D. Computational Results

For every formulation illustrated in the paper, we built optimisation models using an initial weather forecast or a sequence of the forecasts updated in the rolling horizon framework. For the final evaluation of solutions, we recomputed the scaling factor $\lambda$ for the weather conditions observed in real-world at the time the schedule would be executed. Ultimately, the analysis of how well the schedules performed in real-world should be relevant to the decision-makers which we will demonstrate on an example.

Table I displays the absolute scaling factor for every model and a problem instance. Besides the performance of a schedule, we provide supplementary information about Mean Cloud Cover (MCC) obtained from the initial weather forecast and Mean Absolute Percentage Error (MAPE) between the cloud cover forecast and real cloud cover observed during communication windows. These auxiliary metrics provide some additional insight into the problem instance. For example, if a weather forecast indicates overcast, which blocks the opportunity for communication, using more advanced models is not expected to make an impact. Similar intuitive reasoning concerns inaccurate weather forecasts. If they contain significant prediction errors, models which rely on them are going to make suboptimal decisions.

The BTS model accumulated the largest total scaling factor for all problem instances. Its closest competitor, who for the majority of cases recorded a similar performance, was the MSTD model. On the other hand, the deterministic model with a single weather forecast (DSF) performed the worst. Such an outcome was predictable because all other models implemented a rolling horizon and had the opportunity to modify future decisions using more recent weather forecasts. In particular, the deterministic model with a rolling horizon (DRH) performed noticeably better.

### E. Computational Time

Table II displays the average and worst-case computational time for solving the benchmark problems using models studied in this paper.

Each problem instance was solved using the deterministic and RO model in at most 2 seconds. The models solved using the adversarial method required considerably more time.

### F. Model Scalability

We restrict the analysis of the scalability to the RO model BTS, which was the best formulation in our study. Table III presents the descriptive statistics of computational time for different settings of the discretisation step.

TABLE I
ABSOLUTE PERFORMANCE OF THE SOLUTIONS TO THE SATQKD PROBLEM OBSERVED IN REAL-WORLD.

| | MCC | MAPE | DSF | DRH | BTS | MSTD |
|---|---|---|---|---|---|---|
| 1 | 65 | 23 | 5,689 | 4,941 | 5,368 | 5,642 |
| 2 | 74 | 16 | 4,109 | 1,344 | 2,218 | 2,456 |
| 3 | 74 | 14 | 10,259 | 9,074 | 9,271 | 9,262 |
| 4 | 63 | 22 | 4,999 | 14,485 | 12,226 | 14,486 |
| 5 | 59 | 21 | 16,450 | 11,403 | 12,985 | 12,331 |
| 6 | 73 | 18 | 4,096 | 7,285 | 6,965 | 7,006 |
| 7 | 56 | 21 | 666 | 703 | 696 | 696 |
| 8 | 82 | 25 | 1,325 | 6,457 | 7,994 | 7,994 |
| 9 | 66 | 28 | 327 | 786 | 660 | 660 |
| 10 | 64 | 25 | 327 | 4,825 | 4,825 | 3,923 |
| 11 | 72 | 22 | 2,156 | 6,938 | 6,930 | 6,938 |
| 12 | 93 | 8 | 552 | 1,002 | 861 | 634 |
| 13 | 61 | 22 | 9,389 | 13,961 | 15,596 | 15,401 |
| 14 | 43 | 11 | 19,318 | 15,858 | 15,922 | 17,970 |
| 15 | 73 | 17 | 6,161 | 12,959 | 13,654 | 13,776 |
| 16 | 68 | 17 | 660 | 1,153 | 1,663 | 1,559 |
| 17 | 79 | 12 | 758 | 6,475 | 6,593 | 7,416 |
| 18 | 70 | 20 | 4,199 | 9,127 | 6,276 | 2,775 |
| 19 | 86 | 21 | 646 | 712 | 525 | 711 |
| 20 | 57 | 32 | 4,094 | 9,796 | 9,790 | 10,291 |
| 21 | 56 | 32 | 5,014 | 1,625 | 2,757 | 2,955 |
| 22 | 80 | 13 | 327 | 4,166 | 4,166 | 4,166 |
| 23 | 50 | 27 | 5,572 | 11,368 | 11,364 | 11,319 |
| 24 | 68 | 18 | 7,083 | 8,221 | 11,121 | 6,994 |
| 25 | 76 | 15 | 504 | 3,119 | 4,123 | 3,568 |
| 26 | 77 | 12 | 666 | 4,708 | 4,502 | 4,628 |
| 27 | 50 | 14 | 12,995 | 12,558 | 12,580 | 13,135 |
| 28 | 60 | 26 | 14,062 | 7,358 | 6,726 | 6,988 |
| 29 | 80 | 20 | 3,096 | 9,431 | 9,022 | 10,175 |
| 30 | 79 | 11 | 2,510 | 3,788 | 552 | 5,075 |
| 31 | 77 | 18 | 372 | 4,000 | 3,941 | 4,000 |
| 32 | 58 | 18 | 16,860 | 9,893 | 16,957 | 10,444 |
| 33 | 68 | 9 | 9,914 | 8,385 | 10,319 | 10,319 |
| 34 | 76 | 18 | 4,586 | 9,722 | 8,964 | 8,209 |
| Total | - | - | 179,741 | 237,627 | 248,112 | 239,525 |

TABLE II
AVERAGE AND WORST-CASE COMPUTATIONAL TIME FOR SOLVING THE BENCHMARK PROBLEMS.

| Method | Mean [s] | Max [s] |
|---|---|---|
| DSF | 0.10 | 0.30 |
| DRH | 0.30 | 0.50 |
| BTS | 1.30 | 2.00 |
| MSTD | 25,525.40 | 50,271.30 |

TABLE III
AVERAGE AND WORST-CASE COMPUTATIONAL TIME TO PROVING OPTIMALITY FOR BTS MODEL DEPENDING ON TIME DISCRETISATION OF COMMUNICATION WINDOWS.

| Scheme [s] | Mean [s] | Max [s] |
|---|---|---|
| 30 | 1.00 | 1.70 |
| 15 | 1.90 | 3.30 |
| 10 | 2.70 | 7.20 |
| 5 | 7.40 | 21.70 |
| 1 | 98.40 | 428.20 |

Intuitively, the computational time grows as the resolution of the discretisation scheme for communication windows increases. This is due to the number of constraints in (2) and variables in (3) which grow with the resolution of the fine-grained indexing scheme, $\Pi$, we introduced in Section III-B. Ultimately, the model remains practically solvable to optimality even if communication windows are split into segments of one second.

A similar scalability analysis of a deterministic model for scheduling downloads using multiple satellites to a network of ground stations was conducted by [25]. The researcher has demonstrated computational time to grow linearly with the number of satellites for problem instances with randomly generated communication windows. Considering the influence of the discretisation scheme on computational time, he reported a non-linear growth which agrees with our observations.

## V. CONCLUSIONS

We studied techniques for modelling uncertainty in cloud cover predictions with the application to scheduling satellite quantum key distribution. We considered models both with and without a rolling horizon, a Robust Optimisation model with the polyhedral uncertainty sets and a Distributionally Robust Model with a moment-based ambiguity set.

Each model solved a suite of 34 problem instances built using official medium-range weather forecasts which cover the period between September 2019 and March 2020. The performance of the computed solutions was evaluated using real weather conditions observed at the time when the schedule would have been executed. Arguably, our solution methodology with access to official weather forecasts and performance evaluation in real-world closely imitates the environment in which the satellite operations are being scheduled.

We observe that medium-range cloud cover forecasts tend to lose accuracy over time considerably. Consequently, one could achieve substantial improvements solely by accepting the weather forecasts updates and recomputing the future part of the schedule using more precise information. For the problem instances considered, we recorded the best performance for the Robust Optimisation model with a polyhedral uncertainty set built from a vector autoregressive model, which describes the evolution of the uncertain cloud cover by capturing the spatial and temporal dependence in data. The model was trained using historical weather observations and effectively solved to optimality using an off-the-shelf Mixed-Integer Linear Programming solver.

Future research directions could focus on the formulation of the uncertain scheduling problem itself, i.e. by applying data-driven distributionally robust ambiguity sets, or extending the configuration of the communications system to support a constellation of the satellites.

## APPENDIX A
## COMPUTATIONAL COMPLEXITY OF THE SATQKD PROBLEM

**Theorem 2.** *The deterministic SatQKD problem is strongly NP-hard.*

*Proof.* We will show the proof through a reduction from 3-Partition, where given a set of $3n$ integers $S := \{a_1, \ldots, a_{3n}\}$ that add up to $nB$ such that each integer $a_i$ is strictly between $\frac{B}{4}$ and $\frac{B}{2}$ and to decide whether $S$ can be partitioned into $n$ triplets $S_1, S_2, \ldots, S_n$, such that sum of integers in each subset $S_i$ add up to exactly $B$ is NP-complete [46].

In our reduction, the SatQKD instance has $n$ ground stations. We will let $\beta_i = 0$ and $w_i = B$ for each ground station $i$. We assume that there are $3n$ time intervals, each corresponding to an integer in the set $S$, during which we can transfer keys to the $n$ ground stations, and we can send $a_i$ keys from time interval $i$ to every ground station. The ordering of the intervals does not matter. Hence, we can take some arbitrary ordering. We assume that there is no cloud cover during the entire time horizon. Finally, we assume that the switching time is 0. We have a solution to 3-Partition problem if and only if the solution to the instance of SatQKD is at least 1.

Let us suppose there is a solution to 3-Partition problem, which means there is a partition of $S$ into triplets $S_1, \ldots, S_n$, where each subset $S_i$ add up to exactly $B$. For each subset, we will assign a ground station, and we will assign each integer present in that subset as the time interval of communication to the corresponding ground station. Since each integer in a subset adds to $B$, the solution sends exactly $B$ keys to every ground station resulting in the objective value of 1.

First, notice that for the objective value of SatQKD to be at least 1, every ground station has to receive at least $B$ keys. Since the time integers add to $nB$, and there are $n$ ground stations, they have to receive exactly $B$ keys each for the objective to be at least 1. Moreover, since $a_i$ are strictly between $\frac{B}{4}$ and $\frac{B}{2}$, exactly three intervals need to be assigned to each ground station if each were to receive at least $B$ keys.

Finally, suppose there is no solution to the 3-Partition problem, then for every triplet partition of $S$ at least one of the 3-intervals will add to a value strictly less than $B$. So we have that the SatQKD instance will have an objective value strictly less than 1. □

## APPENDIX B
## REFORMULATION OF THE SATQKD PROBLEM WITH THE UNCERTAINTY SET $\mathcal{V}$

*Proof.* The problem BTS can be reformulated as:

$$\begin{aligned} \max \quad & \Theta \\ \text{s.t.:} \quad & \Theta \leq \lambda_n(\mathbf{x}, \mathbf{c}) \qquad \forall \mathbf{c} \in \mathcal{V} \quad \forall n \in N \\ & \mathbf{x} \in \mathcal{X} \end{aligned}$$

Notice that we can reformulate the constraint $\Theta \leq \lambda_{n'}(\mathbf{x}, \mathbf{c})$ for all $\mathbf{c} \in \mathcal{V}$ some ground station $n' \in N$ as follows:

$$\begin{aligned} \max_{\mathbf{c} \in \mathcal{V}} \left\{ \sum_{t \in T} \sum_{\pi \in \Pi_{n't}} c_{n't} k_{n'\pi} x_{n'\pi} \right\} \leq -\omega_{n'} \Theta \\ + \beta_{n'} + \sum_{t \in T} \sum_{\pi \in \Pi_{n't}} k_{n'\pi} x_{n'\pi} \end{aligned} \tag{10}$$

The dual of the optimisation problem in the left hand side of (10) can be written as:

$$\begin{aligned} \min \quad & \bar{\mathbf{c}}^\top \bar{\mathbf{z}} - \underline{\mathbf{c}}^\top \underline{\mathbf{z}} \\ & - \sum_{t \in T \setminus \{1\}} (a - e) \bar{v}_{n't} \\ & + \sum_{t \in T \setminus \{1\}} (a + e) \underline{v}_{n't} \\ \text{s.t.:} \quad & \bar{\mathbf{z}}_t - \underline{\mathbf{z}}_t + A^\top \bar{\mathbf{v}}_{t+1} - A^\top \underline{\mathbf{v}}_{t+1} \\ & - \bar{\mathbf{v}}_t - \underline{\mathbf{v}}_t - \sum_{\pi \in \Pi_{n't}} k_{n'\pi} x_{n'\pi} \quad = 0 \quad \forall t \in T \setminus \{1\} \\ & \bar{\mathbf{v}}, \underline{\mathbf{v}}, \bar{\mathbf{z}}, \underline{\mathbf{z}} \in \mathbb{R}_{\geq 0}^{|N| \times |T|} \end{aligned}$$

where $\bar{\mathbf{v}}, \underline{\mathbf{v}}, \bar{\mathbf{z}}, \underline{\mathbf{z}}$ are the dual multipliers of the inequalities in $\mathcal{V}$. Plugging this dualised version of (10) back in the BTS would give us the result. □

## ACKNOWLEDGEMENTS

We are grateful to Luca Mazzarella for estimating the number of quantum keys which can be transferred from space-to-ground in the cloud-free line of sight. We also thank Daniel KL Oi for insightful suggestions in defining the deterministic variant of the SatQKD problem and proposing a representative example of the communications system.

## REFERENCES

[1] M. Capelle, M.-J. Huguet, N. Jozefowiez, and X. Olive, "Optimizing ground station networks for free space optical communications: Maximizing the data transfer," Networks, vol. 73, no. 2, pp. 234–253, 2019.
[2] S. Grau, "Contributions to the advance of the integration density of cubesats," Doctoral Thesis, Technische Universität Berlin, Berlin, 2019.
[3] A. Shao, J. R. Wertz, and E. A. Koltz, "Quantifying the cost reduction potential for earth observation satellites," in Proceedings of the 12th Reinventing Space Conference, S. Hatton, Ed. Cham: Springer International Publishing, 2017, pp. 199–210.
[4] A. K. Majumdar, Fundamentals of Free-Space Optical Communications Systems, Optical Channels, Characterization, and Network/Access Technology. Elsevier, 2019, ch. 4, pp. 55 – 116.
[5] H. Kaushal and G. Kaddoum, "Optical communication in space: Challenges and mitigation techniques," IEEE Communications Surveys Tutorials, vol. 19, no. 1, pp. 57–96, Firstquarter 2017.
[6] S.-K. Liao, W.-Q. Cai, W.-Y. Liu, L. Zhang, Y. Li, J.-G. Ren, J. Yin, Q. Shen, Y. Cao, Z.-P. Li, F.-Z. Li, X.-W. Chen, L.-H. Sun, J.-J. Jia, J.-C. Wu, X.-J. Jiang, J.-F. Wang, Y.-M. Huang, Q. Wang, Y.-L. Zhou, L. Deng, T. Xi, L. Ma, T. Hu, Q. Zhang, Y.-A. Chen, N.-L. Liu, X.-B. Wang, Z.-C. Zhu, C.-Y. Lu, R. Shu, C.-Z. Peng, J.-Y. Wang, and J.-W. Pan, "Satellite-to-ground quantum key distribution," Nature, vol. 549, no. 7670, pp. 43–47, 2017.
[7] R. J. Alliss and B. Felton, "The mitigation of cloud impacts on free-space optical communications," in Atmospheric Propagation IX, L. M. W. Thomas and E. J. Spillar, Eds., vol. 8380, International Society for Optics and Photonics. SPIE, 2012, pp. 227 – 238.
[8] M. Polnik, L. Mazzarella, M. Di Carlo, D. K. Oi, A. Riccardi, and A. Arulselvan, "Scheduling of Space to Ground Quantum Key Distribution," EPJ Quantum Technology, 2020.
[9] A. Ruszczyński and A. Shapiro, Stochastic Programming Models. SIAM - Society for Industrial and Applied Mathematics, 2009, ch. 1, pp. 1–25.
[10] A. J. Kleywegt, A. Shapiro, and T. Homem-de Mello, "The sample average approximation method for stochastic discrete optimization," SIAM J. on Optimization, vol. 12, no. 2, p. 479–502, Feb. 2002.
[11] B. Verweij, S. Ahmed, A. J. Kleywegt, G. Nemhauser, and A. Shapiro, "The sample average approximation method applied to stochastic routing problems: A computational study," Computational Optimization and Applications, vol. 24, no. 2, pp. 289–333, Feb 2003.

[12] A. Nemirovski and A. Shapiro, "Convex approximations of chance constrained programs," SIAM Journal on Optimization, vol. 17, no. 4, pp. 969–996, 2007. [Online]. Available: https://doi.org/10.1137/050622328

[13] S. Sarykalin, G. Serraino, and S. Uryasev, Value-at-Risk vs. Conditional Value-at-Risk in Risk Management and Optimization. INFORMS, 2014, ch. Chapter 13, pp. 270–294.

[14] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, Robust Optimization, ser. Princeton Series in Applied Mathematics. Princeton University Press, October 2009.

[15] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," SIAM Review, vol. 53, no. 3, pp. 464–501, 2011.

[16] B. L. Gorissen, İhsan Yanıkoğlu, and D. den Hertog, "A practical guide to robust optimization," Omega, vol. 53, pp. 124 – 137, 2015.

[17] A. Ben-Tal and A. Nemirovski, "Robust solutions of uncertain linear programs," Operations Research Letters, vol. 25, no. 1, pp. 1 – 13, 1999.

[18] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," Operations Research, vol. 58, no. 3, pp. 595–612, 2010.

[19] D. Bertsimas, M. Sim, and M. Zhang, "Adaptive distributionally robust optimization," Management Science, vol. 65, no. 2, pp. 604–618, 2019.

[20] K. Isii, "On sharpness of tchebycheff-type inequalities," Annals of the Institute of Statistical Mathematics, vol. 14, no. 1, pp. 185–197, Dec 1962.

[21] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally robust convex optimization," Operations Research, vol. 62, no. 6, pp. 1358–1376, 2014.

[22] Z. Hu and L. Hong, "Kullback-leibler divergence constrained distributionally robust optimization," Available at Optimisation Online, 11 2012, last accessed 17/01/2020. [Online]. Available: http://www.optimization-online.org/DB_HTML/2012/11/3677.html

[23] P. Mohajerin Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations," Mathematical Programming, vol. 171, no. 1, pp. 115–166, Sep 2018.

[24] S. Spangelo, J. Cutler, K. Gilson, and A. Cohn, "Optimization-based scheduling for the single-satellite, multi-ground station communication problem," Computers & Operations Research, vol. 57, pp. 1 – 16, 2015.

[25] J. Castaing, "Scheduling downloads for multi-satellite, multi-ground station missions," in Proceedings of the Small Satellite Conference, Technical Session VIII, SSC14-VIII-4, Logan, United States, 2014, pp. 1–12. [Online]. Available: https://digitalcommons.usu.edu/smallsat/2014/FJRStudentComp/4/

[26] F. Marinelli, S. Nocella, F. Rossi, and S. Smriglio, "A lagrangian heuristic for satellite range scheduling with resource constraints," Computers and Operations Research, vol. 38, no. 11, pp. 1572 – 1583, 2011.

[27] D. Liao and Y. Yang, "Imaging order scheduling of an earth observation satellite," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 37, no. 5, pp. 794–802, Sep. 2007.

[28] S. Sethi and G. Sorger, "A theory of rolling horizon decision making," Annals of Operations Research, vol. 29, no. 1, pp. 387–415, Dec 1991.

[29] J. Wang, E. Demeulemeester, and D. Qiu, "A pure proactive scheduling algorithm for multiple earth observation satellites under uncertainties of clouds," Computers & Operations Research, vol. 74, pp. 1 – 13, 2016.

[30] X. Wang, G. Song, R. Leus, and C. Han, "Robust earth observation satellite scheduling with uncertainty of cloud coverage," IEEE Transactions on Aerospace and Electronic Systems, pp. 1–1, 2019.

[31] D. Bertsimas and M. Sim, "Robust discrete optimization and network flows," Mathematical Programming, vol. 98, pp. 49–71, 09 2003.

[32] J. Litva and T. K. Lo, Digital Beamforming in Wireless Communications, 1st ed. USA: Artech House, Inc., 1996.

[33] Z. Lin, M. Lin, W. Zhu, J. Wang, and J. Cheng, "Robust secure beamforming for wireless powered cognitive satellite-terrestrial networks," IEEE Transactions on Cognitive Communications and Networking, pp. 1–1, 2020.

[34] Z. Lin, M. Lin, B. Champagne, W. P. Zhu, and N. Al-Dhahir, "Secure and energy efficient transmission for rsma-based cognitive satellite-terrestrial networks," IEEE Wireless Communications Letters, pp. 1–1, 2020.

[35] Z. Lin, M. Lin, J. Wang, T. de Cola, and J. Wang, "Joint beamforming and power allocation for satellite-terrestrial integrated networks with non-orthogonal multiple access," IEEE Journal of Selected Topics in Signal Processing, vol. 13, no. 3, pp. 657–670, 2019.

[36] H. Morf, "Sunshine and cloud cover prediction based on markov processes," Solar Energy, vol. 110, pp. 615 – 626, 2014.

[37] E. Zivot and J. Wang, Vector Autoregressive Models for Multivariate Time Series. New York, NY: Springer New York, 2006, ch. 11, pp. 385–429.

[38] N. J. Sparks, S. R. Hardwick, M. Schmid, and R. Toumi, "Image: a multivariate multi-site stochastic weather generator for european weather and climate," Stochastic Environmental Research and Risk Assessment, vol. 32, no. 3, pp. 771–784, Mar 2018.

[39] W. McKinney, J. Perktold, and S. Seabold, "Time series analysis in python with statsmodels," in Proceedings of the 10th Python in Science Conference (SciPy 2011), 2011, pp. 96–102. [Online]. Available: http://conference.scipy.org/proceedings/scipy2011/pdfs/statsmodels.pdf

[40] H. Bozdogan, "Model selection and akaike's information criterion (aic): The general theory and its analytical extensions," Psychometrika, vol. 52, pp. 345–370, 02 1987.

[41] D. Bertsimas and M. Sim, "The price of robustness," Operations Research, vol. 52, no. 1, pp. 35–53, 2004.

[42] D. Bertsimas, I. Dunning, and M. Lubin, "Reformulation versus cutting-planes for robust optimization," Computational Management Science, vol. 13, no. 2, pp. 195–217, 2016.

[43] M. Polnik. (2020) Scheduling space-to-ground optical communication under cloud cover uncertainty - source code repository. Last accessed 20/01/2020. [Online]. Available: https://github.com/pmateusz/uncertain-quake

[44] Open Weather Map. (2019) Service homepage. Last accessed 25/1/2019. [Online]. Available: https://openweathermap.org

[45] M. Polnik, A. Arulselvan, and A. Riccardi. Scheduling space-to-ground optical communication under cloud cover uncertainty - data set. Last accessed 23/04/2020. [Online]. Available: https://1drv.ms/u/s!AobWzOHtlT38hNAtrEC3bU5zqYqEOg?e=NcVOSf

[46] M. R. Garey and D. S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness (Series of Books in the Mathematical Sciences). W. H. Freeman, 1979.