A Robust Optimization Model for Managing Elective Admission in Hospital

Daizhuo Chen*, Fanwen Meng†, James Ang‡, Singfat Chu§, Melvyn Sim¶, Palvannan Kannapiran‖

August 29, 2011

Abstract

The admission of emergency inpatients in a hospital is unscheduled, urgent and takes priority over elective patients, who are usually scheduled several days in advance. Hospital beds are a critical resource and the management of elective admissions by enforcing quotas could reduce incidents of shortfall. We propose a distributionally robust optimization approach for managing elective admissions to determine these quotas. Based on an adjustable family of distributions, we propose two robust models, one with fixed *budget of variation* and the other with optimized budget of variation subject to an expected bed shortfall constraint. We solve the robust optimization model by deriving a second order conic problem (SOCP) equivalent of the model. The proposed model is tested in simulations based on real hospital admission data and we report favorable results for adopting the robust optimization models.

Keywords: Robust optimization, health care, hospital operations

*Singapore-MIT Alliance, National University of Singapore. Email: daizhuo@nus.edu.sg.

†Centre for Maritime Studies, National University of Singapore. Email: cmsmf@nus.edu.sg.

‡NUS Business School, National University of Singapore. Email: bizangsk@nus.edu.sg.

§NUS Business School, National University of Singapore. Email: bizchucl@nus.edu.sg.

¶NUS Business School and Risk Management Institute, National University of Singapore. Email: dscsimm@nus.edu.sg.

‖Health Services & Outcomes Research, National Healthcare Group, Singapore. Email: palvannan_kannapiran@nhg.com.sg.

1 Introduction

Beds are a critical resource in hospital operations. Overcrowding of Accident and Emergency (A & E) is often due to availability (or rather the shortage) of hospital beds (Wardrope and Driscoll (2003)); so is cancellation of elective surgeries (Robb et al. (2004)). However, bed resources are expensive inasmuch as the hospitals need highly trained personnel to man these beds. Work has been done in the area of the acquisition and utilization of bed resources (e.g., Harper and Shahani (2002), Kao et al. (1981), Cochran and Roche (2007) and Teow and Tan (2008)). Harper and Shahani (2002) acknowledged the complexity of the internal dynamics of a hospital (especially bed management), and used a simulation model for patient flows and bed matching over time.

Typically, Day-of-Week (DoW) patterns of a hospital exhibit a wide range of variations. Emergency admissions are beyond the control of the hospital, while elective admissions are scheduled by the hospital. Nevertheless, often the relative variation is largest in elective admissions, and larger in discharges than emergency admissions (Proudlove et al. (2007)). On days with high bed occupancy, long wait time is encountered. On days with low bed occupancy, beds are under-utilized. We have tightness of usage on one hand, and looseness on the other. It is not the desired state.

Elective surgeries account for the majority of elective admissions, though medical electives (non-surgical cases) do make up for some of these admissions. Elective surgeries are procedures planned in advance and can be divided into day surgery (DS), same day surgery admission (SDA) and inpatient admission (IP). DS cases do not "consume" beds, while SDA cases require beds to accommodate patient day after surgery. IP cases require beds one day before the surgery.

In general, hospitals will admit all emergency cases. As such, in a tight bed situation, the tradeoff is to reduce the number of beds designated for elective admissions. But a more prudent and sensible approach would be to make adjustments on a dynamic basis. What this entails is that when emergency cases are fewer, then more beds could be assigned to elective cases, and vice versa. This leads to an optimal control policy, which is to maximize bed utilization on a daily basis by controlling the number of elective admissions. This requires a more prudent scheduling of operating theatre sessions. However, a higher level of complexity in planning ensues because of the high degree of uncertainty involved in bed availability and its effect on admission rates.

Various models for managing patient admissions have been proposed in the literature. Esogbue and Singh (1976) developed a method for determining optimal distribution of beds in a ward using cut-off level via shortage and holding costs. They assumed Poisson patient arrival distribution and negative exponential distribution for length of stay. Kao and Tung (1981) proposed an approach for periodically reallocating beds to services to minimize the expected overflows, using queueing models to approximate the population dynamics. In fact, queueing theory and stochastic simulation are the main methodological tools in studies of bed allocation and bed capacity (Vassilacopoulos (1985), Gorunescui et al. (2004), Cochran and Roche (2007), Lamiri et al. (2008)). The underlying rationale for researchers relying on these methodological tools is the uncertain nature of the hospital unit vis-à-vis the number of patients as a result of random arrivals and random lengths of stay. A thorough review on OR applications in healthcare services in the United Kingdom can be found in Proudlove et al. (2007).

The admission of emergency inpatients is unscheduled and they are usually warded within hours. In contrast, admission of elective patients is less pressing and they can be warded on the day of admission or even several weeks later. The flexibility vis-à-vis elective patients allows hospitals to manage the flow of elective patients in a way as to "smooth out" the daily bed occupancy, a modus operandus known as "elective smoothing". This will ensure that on days with spikes in emergency cases, the admission rate for elective patients can be reduced. The converse applies. Some hospitals in Singapore have already incorporated this mechanism into their decision support systems and it has led to improvements when elective patient flow is high (Teow et al. (2007)). In these hospitals, the admission quotas for elective patients are obtained by solving a deterministic linear optimization problem without taking into account of the variability of patient arrivals and stay durations. While this achieves smoothing in expectation, it is conceivable that the efficacy would diminish when variability is high.

Due to the difficulties of obtaining true probability distributions and solving stochastic optimization problems, it is common in real world deployment of optimization technology to ignore uncertainty. A fine level of analysis would be required to obtain the distributions of patient arrivals and departure profile as a function of admission quotas, which may not necessarily lead to a computational tractable optimization problem. In recent years, robust optimization offers an attractive alternative for addressing uncertainty in optimization modeling without having to specify exact probability distributions. In many interesting cases, the approach leads to computationally tractable optimization problems; see for instance Ben-Tal and Nemirovski (1998), Bertsimas and Sim (2004), El Ghaoui et al. (1998). In classical robust optimization, uncertainty is represented by an adjustable *uncertainty set*, which is usually a simple geometric convex set such as a l-norm ball intersected with the *support set*, the minimal convex set that contains the uncertainty. The modeler requires to articulate her *ambiguity attitude*[1] by specifying the *budget of uncer-*

[1]We distinguish between risk and ambiguity. Risk deals with uncertainty with known distributions

tainty parameter, which relates to the size of the uncertainty set that she would like to immune against.

While there are several proposed uncertainty sets and heuristics for specifying budgets of uncertainty, these approaches may not naturally characterize the uncertainty relating to patient movements within the hospital. In this paper, we adopt the distributionally robust optimization approach for managing elective admission in hospital, where uncertainty is characterized by the support set and a restricted family of probability distributions; see for instance Chen et al. (2007), Chen et al. (2010), Delage and Ye (2010) and Goh and Sim (2010a, 2010b). Similar to uncertainty set in classical robust optimization, the proposed family of distributions is adjustable via a so called *budget of variation* parameter, which is the bound on the coefficient of variation of the uncertainty parameters. The family of distributions is enlarged by increasing the budget of variation, which leads to greater uncertainty in the patient movements. We also propose an approach to optimize the budget of variation while ensuring that the worst-case expected maximum bed requirement over the planning horizon falls below the bed capacity of the hospital. Interestingly, this could be achieved by solving a small collections of computationally tractable optimization problems. We also study the performance of this approach in case study using real data.

The rest of this paper is organized as follows. In Section 2, we establish a distributionally robust optimization model for managing elective admission in hospital with incomplete information of uncertainties. We then investigate deterministic tractable formulation to this model by deriving a second order conic problem in Section 3. Numerical experiments using real data are carried out in Section 4. Section 5 concludes the paper.

Notation: We denote a random variable with a tilde sign, such as $\tilde{z}$. Matrices and vectors are represented as upper and lower case boldface characters respectively. If $\boldsymbol{x}$ is a vector, we use the notation x_i to denote the ith component of the vector. We represent uncertainty by a state-space Ω and a set (σ-algebra) $\mathcal{F}$ of events. We use the notation $\tilde{x} \geq \tilde{y}$ to denote state-wise dominance over all attributes, i.e, $\tilde{x}(\omega) \geq \tilde{y}(\omega)$ for all $\omega \in \Omega$. We use $\mathbb{P}$ to denote a probability measure on Ω and $\mathbb{E}_{\mathbb{P}}(\tilde{x})$, $\sigma_{\mathbb{P}}(\tilde{x})$ and $\mathrm{cv}_{\mathbb{P}}(\tilde{x})$ denote respectively the expectation, standard deviation and coefficient of variation of $\tilde{x}$ under $\mathbb{P}$.

2 Model formulation

We consider a planning horizon of T days indexed by $t = 0, 1, \ldots, T-1$. Let η_t be the decision variable representing the elective inpatients quota for the tth day within the

while ambiguity does not.

planning horizon. For simplicity of model presentation, we assume that all inpatients are of the same type. We can easily refine the model to consider quotas for different types of inpatients that may be characterized by gender, discipline, and so forth. At the beginning of day $t = 0$ (say at 8 am when clinics open), the quotas $\boldsymbol{\eta} = (\eta_0, \ldots, \eta_{T-1})'$ will be determined and integrated within hospital decision support system for assignment of admissions during the operating hours of the elective clinics. As we proceed to the next day, the process is repeated and a new set of quotas will be computed using the latest information on admission status. We let X be the feasible space of admissible quotas, which is determined strategically by the hospital adminstration. For instance, X could impose a lower bound on the total number of quotas available during the planning horizon. In the rolling horizon implementation, it is also imperative to ensure that the new set of quotas is able to accommodate previously assigned elective admissions. For example if 15 elective admissions have already been assigned on day $t = 6$, we would impose a constraint $\eta_6 \geq 15$. Such linear constraints could easily be encapsulated in X.

We next describe the dynamics of patient flow. Let L be the maximum duration of stay for any patient. Note that by definition, inpatients are patients who are warded for at least one day. To account for the total number of inpatients on the tth day, we need to keep track of the admission status up to $L - 1$ days before the planning horizon. Let $\mathcal{T}^+ = \{0, \ldots, T-1\}$, $\mathcal{T}^{--} = \{-L+1, \ldots, -1\}$ and $\mathcal{T} = \mathcal{T}^{--} \cup \mathcal{T}^+$. We denote $\tilde{p}_{t,l}$ and $\tilde{a}_{t,l}$ to be respectively the number of emergency and elective inpatients arriving on the tth day, $t \in \mathcal{T}$ and would be warded for at least l days, $l \in \{1, \ldots, L\}$. For instance, $\tilde{p}_{1,1}$ refers to the total number of emergency inpatients on day $t = 1$ and its value is uncertain. If $\tilde{d}$ of these patients are discharged on day $t = 2$, then $\tilde{p}_{1,2} = \tilde{p}_{1,1} - \tilde{d}$. Likewise $\tilde{a}_{-1,2}$ refers to the number of elective inpatients that arrive on the previous day ($t = -1$) and would be warded for at least 2 days. At the beginning of day $t = 0$, doctors may not have reviewed the cases for discharge. Hence, the parameter $\tilde{a}_{-1,2}$ is generally uncertain. For our purpose, we need to account for the number of inpatients during the planning horizon, i.e., on the days in $\mathcal{T}^+$. For inpatients arriving on day $t \in \mathcal{T}^{--}$, only the inpatients with the length of stay of at least l days, $l \geq 1 - t$, may remain warded in the hospital during the planning horizon. On the other hand, for patients arriving on day $t \in \mathcal{T}^+$, only the information associated with inpatients with length of stay of least l days, $l \leq \min\{L, T - t\}$ will be needed to compute the quotas. Hence, for notational convenience, we define $\mathcal{L}_t = \{\max\{1, 1-t\}, \max\{1, 1-t\} + 1, \ldots, \min\{L, T-t\}\}$, $t \in \mathcal{T}$, which can be viewed as an *active index set*.

We now account for the total number of inpatients on the tth day during the planning horizon, $t \in \mathcal{T}^+$. For example, the total number of inpatients on day $t = 0$ can be computed as follows

$$
\begin{array}{ll}
\tilde{a}_{0,1} + \tilde{p}_{0,1} + & \text{(arrivals/adimissions on } t = 0) \\
\tilde{a}_{-1,2} + \tilde{p}_{-1,2} + & \text{(arrivals/adimissions on } t = -1 \text{ and warded for at least 2 days)} \\
\cdots + \tilde{a}_{-L+1,L} + \tilde{p}_{-L+1,L}. & \text{(arrivals/adimissions on } t = -L+1 \text{ and warded for up to } L \text{ days)}
\end{array}
$$

In general, it follows that the total inpatients on day $t \in \mathcal{T}^+$ can be computed as

$$\sum_{(\tau,l)\in\mathcal{U}_t} (\tilde{a}_{\tau,l} + \tilde{p}_{\tau,l}),$$

where the index set $\mathcal{U}_t$ is given by

$$\mathcal{U}_t = \{(\tau, l) \ : \ \tau \in \mathcal{T},\ l \in \mathcal{L}_\tau, l + \tau = t + 1\}.$$

A bed shortfall occurs whenever the total inpatients exceeds the bed capacity, which we denote by c_t, $t \in \mathcal{T}^+$. Note that for generality, we assume that bed capacity, which encompasses the physical beds and manpower availability, is time dependent. Before we could specify an optimization problem, we first need to account for the uncertainty concerning patients arrival and departure.

2.1 Characterizing patient arrivals and departures uncertainty

We describe a nonparametric approach for characterizing the uncertainty on patient arrivals and departures using information obtained from patient movement records. Our aim is to introduce a model of uncertainty without imposing excessive burden on the information requirement, which may otherwise deter practical implementation. Instead of ignoring variability and assuming deterministic parameters taking values at their empirical averages, which is usually done in practice, we assume that the parameters are random variables with known means but their precise distributions are unavailable but belongs to a restrict family of distributions. To avoid being overconservative, we control the "size" of the family of distributions by specifying the budget of variation, μ, which is the upper bound of the coefficients of variations of all the uncertain parameters.

We next show how the uncertain parameters $\tilde{p}_{t,l}$ and $\tilde{a}_{t,l}$ are interrelated, which is the basis for characterizing the support of the uncertainty. Observe that by definition, $\tilde{p}_{t,l}$ and $\tilde{a}_{t,l}$ are nonincreasing in l. For inpatients arriving before $t = 0$, their total admissions are known but their durations of stay may be uncertain. Let p^0_t and a^0_t, $t \in \mathcal{T}^{--}$, be respectively the number of remaining emergency and elective inpatients who have arrived on day t and are still being warded up to the beginning of day 0. The support of the uncertain parameters $\tilde{p}_{t,l}$ and $\tilde{a}_{t,l}$ is given by

$$p_t^0 \geq \tilde{p}_{t,l} \geq \tilde{p}_{t,l'} \geq 0,$$
$$a_t^0 \geq \tilde{a}_{t,l} \geq \tilde{a}_{t,l'} \geq 0,$$

for all $t \in \mathcal{T}^{--}$, $l, l' \in \mathcal{L}_t, l' > l$. Similarly, for inpatients arriving during the planning horizon $t \in \mathcal{T}^+$, the support of the associated uncertain parameters $\tilde{p}_{t,l}$, $\tilde{a}_{t,l}$, is given by

$$p_t^0 \geq \tilde{p}_{t,l} \geq \tilde{p}_{t,l'} \geq 0,$$
$$\eta_t \geq \tilde{a}_{t,l} \geq \tilde{a}_{t,l'} \geq 0,$$

for all $t \in \mathcal{T}^+$, $l, l' \in \mathcal{L}_t, l' > l$. Note that the input parameter, p_t^0 is a prescribed upper bound of $\tilde{p}_{t,l}$.

Instead of assuming a probability distribution, we specify the family of distributions such that for each distribution, $\mathbb{P}$ in the family, the uncertain parameters are random variables with known mean values and their coefficients of variations are bounded below μ. Specifically, for inpatients arriving before $t = 0$, i.e., $t \in \mathcal{T}^{--}$, we assume that

$$\mathbb{E}_{\mathbb{P}}(\tilde{p}_{t,l}) = \bar{p}_{t,l},$$
$$\mathbb{E}_{\mathbb{P}}(\tilde{a}_{t,l}) = \bar{a}_{t,l},$$

for all $l \in \mathcal{L}_t$, where $\bar{p}_{t,l}$ and $\bar{a}_{t,l}$ are respectively the empirical averages of $\tilde{p}_{t,l}$ and $\tilde{a}_{t,l}$. Since these patients are already admitted, in principle, the parameters $\bar{p}_{t,l}, \bar{a}_{t,l}$ may be inferred from the patients' likely duration of stay assessed by their doctors. If such information is unavailable, then one may also use values that are empirically estimated from historical records.

Observe that during the planning horizon, $t \in \mathcal{T}^+$, the parameters characterizing elective admissions, $\tilde{a}_{t,l}$, $l \in \mathcal{L}_t$, depend on the quota η_t. For instance, if the quota for elective admission is fully assigned, which is usually the case, we would have $\tilde{a}_{t,1} = \eta_t$. Likewise, if $\eta_t = 0$, then it is clear that $\tilde{a}_{t,l} = 0$ for all $l \in \mathcal{L}_t$. Unlike the previous case, the uncertain parameters $\tilde{p}_{t,l}$ and $\tilde{a}_{t,l}$, $t \in \mathcal{T}^+$, are associated with inpatients who have yet to arrive at the hospital. Again, from patient movement records we are able to obtain the empirical averages and specify them as follows,

$$\mathbb{E}_{\mathbb{P}}(\tilde{p}_{t,l}) = \bar{p}_{t,l},$$
$$\mathbb{E}_{\mathbb{P}}(\tilde{a}_{t,l}) = \bar{a}_{t,l} = \alpha_{t,l}\eta_t,$$

for all $t \in \mathcal{T}^+$, $l \in \mathcal{L}_t$. The assumption that $\mathbb{E}_{\mathbb{P}}(\tilde{a}_{t,l})$ being proportional to η_t is motivated from the observation that elective admission is almost always fully assigned. Hence, $\alpha_{t,l}$ may be interpreted as the average fraction of the elective inpatients who are warded for at least l days. As it will become clearer, this assumption is also crucial for formulating a tractable optimization problem.

Finally, the coefficients of variation of these parameters are bounded below the budget of variation, μ as follows

$$\begin{aligned}\mathrm{cv}_{\mathbb{P}}(\tilde{p}_{t,l}) &\leq \mu,\\ \mathrm{cv}_{\mathbb{P}}(\tilde{a}_{t,l}) &\leq \mu,\end{aligned}$$

for all $t \in \mathcal{T}$, $l \in \mathcal{L}_t$. Hence, $\mu = 0$, implies that the parameters are almost surely certain and take values at their means. On the other extreme with $\mu = \infty$, then essentially the variabilities of these parameters are not constrained by μ, but could otherwise be limited by the support. We present the family of distributions as a function of the budget of variation, μ as follows

$$\mathbb{F}(\mu) = \left\{ \mathbb{P} \; : \; \begin{array}{ll} \multicolumn{2}{l}{\mathbb{P}\left(\left\{ \begin{array}{ll} p_t^0 \geq \tilde{p}_{t,l} \geq \tilde{p}_{t,l'} \geq 0 & \forall t \in \mathcal{T}^{--}, l, l' \in \mathcal{L}_t, l' > l \\ a_t^0 \geq \tilde{a}_{t,l} \geq \tilde{a}_{t,l'} \geq 0 & \forall t \in \mathcal{T}^{--}, l, l' \in \mathcal{L}_t, l' > l \\ \tilde{p}_{t,l} \geq \tilde{p}_{t,l'} \geq 0 & \forall t \in \mathcal{T}^{+}, l, l' \in \mathcal{L}_t, l' > l \\ \eta_t \geq \tilde{a}_{t,l} \geq \tilde{a}_{t,l'} \geq 0 & \forall t \in \mathcal{T}^{+}, l, l' \in \mathcal{L}_t, l' > l \end{array} \right\}\right) = 1} \\ \mathbb{E}_{\mathbb{P}}(\tilde{p}_{t,l}) = \bar{p}_{t,l} & \forall t \in \mathcal{T}, l \in \mathcal{L}_t \\ \mathbb{E}_{\mathbb{P}}(\tilde{a}_{t,l}) = \bar{a}_{t,l} & \forall t \in \mathcal{T}^{--}, l \in \mathcal{L}_t \\ \mathbb{E}_{\mathbb{P}}(\tilde{a}_{t,l}) = \alpha_{t,l}\eta_t & \forall t \in \mathcal{T}^{+}, l \in \mathcal{L}_t \\ \sigma_{\mathbb{P}}(\tilde{p}_{t,l}) \leq \bar{p}_{t,l}\mu & \forall t \in \mathcal{T}, l \in \mathcal{L}_t \\ \sigma_{\mathbb{P}}(\tilde{a}_{t,l}) \leq \bar{a}_{t,l}\mu & \forall t \in \mathcal{T}^{--}, l \in \mathcal{L}_t \\ \sigma_{\mathbb{P}}(\tilde{a}_{t,l}) \leq \alpha_{t,l}\eta_t\mu & \forall t \in \mathcal{T}^{+}, l \in \mathcal{L}_t \end{array} \right\},$$

where the parameters for characterizing the family of distributions, $\{p_t^0, a_t^0 \; : \; t \in \mathcal{T}^{--}\}$, $\{p_t^0 \; : \; t \in \mathcal{T}^{+}\}$, $\{\bar{p}_{t,l} \; : \; t \in \mathcal{T}, l \in \mathcal{L}_t\}$, $\{\bar{a}_{t,l} \; : \; t \in \mathcal{T}^{--}, l \in \mathcal{L}_t\}$ and $\{\alpha_{t,l} \; : \; t \in \mathcal{T}^{+}, l \in \mathcal{L}_t\}$ are values that could easily be estimated from the patient movement records. Observe that the set $\mathbb{F}(\mu)$ is nondecreasing in μ, i.e.,

$$\mathbb{F}(\mu) \subseteq \mathbb{F}(\mu') \qquad \forall \mu' \geq \mu.$$

2.2 Distributionally robust optimization models

To circumvent the difficulties of obtaining probability distributions and solving the complex stochastic model, the elective smoothing approach ignores uncertainty and solves the following deterministic optimization problem:

$$Z_D = \min_{\boldsymbol{\eta} \in X} \left(\max_{t \in \mathcal{T}^{+}} \left\{ \sum_{(\tau,l) \in \mathcal{U}_t} (\bar{a}_{\tau,l} + \bar{p}_{\tau,l}) - c_t \right\} \right), \tag{2.1}$$

where $\bar{a}_{t,l} = \alpha_{t,l}\eta_t$ for $t \in \mathcal{T}^+$, $l \in \mathcal{L}_t$. Hence, if X is a polyhedron, then this essentially becomes a linear optimization problem. The rationale for elective smoothing is to ensure that the average number of patients adjusted for bed availability are evenly spread across the time periods, which is a reasonable heuristic for reducing incidents of bed shortfalls.

Nevertheless, despite being a tractable linear optimization problem, the model ignores the potential impact of uncertainty and could lead to severe shortfalls in hospital beds whenever bad scenarios arises. A natural extension of the elective smoothing approach to incorporate uncertainty is to minimize the worst-case expected maximum bed excess over the planning horizon as follows

$$Z_R(\mu) = \min_{\boldsymbol{\eta} \in X} \sup_{\mathbb{P} \in \mathbb{F}(\mu)} \mathbb{E}_{\mathbb{P}} \left(\max_{t \in \mathcal{T}^+} \left\{ \sum_{(\tau,l) \in \mathcal{U}_t} (\tilde{a}_{\tau,l} + \tilde{p}_{\tau,l}) - c_t \right\} \right). \tag{2.2}$$

In the absence of uncertainty, i.e., $\mu = 0$, it is clear that model (2.1) is the same as model (2.2), hence $Z_D = Z_R(0)$. As we increase the budget of variation μ, the model takes into consideration more potential variations in the admission process. It is the modeler's choice to set the value of μ, i.e., the budget of variations that will be protected against. We refer to this model the fixed budget model.

Optimized budget of variation model

The main challenge of model (2.2) is how to specify the value of μ that would yield the desired level of performance in controlling bed shortfalls. Intuitively, an underly or overly specified budget of variation, μ may not adequately protect against potential bed shortfalls when the actual uncertainty is realized. In practice, the parameter μ has to be turned accordingly so that it gives the best overall performance on real data.

We note that model (2.2) is only a means to cope with the issue of bed shortfalls. In a well managed hospital, it is imperative that beds capacity should exceed average demands, which implies $Z_D = Z_R(0) \leq 0$. Extending this notion to incorporate uncertainty, if $Z_R(\mu) \leq 0$, for $\mu > 0$, then we are guaranteed a solution that ensures that for all $\mathbb{P} \in \mathbb{F}(\mu)$, the expected maximum bed excess across the time periods is less then zero, i.e.,

$$\mathbb{E}_{\mathbb{P}} \left(\max_{t \in \mathcal{T}^+} \left\{ \sum_{(\tau,l) \in \mathcal{U}_t} (\tilde{a}_{\tau,l} + \tilde{p}_{\tau,l}) - c_t \right\} \right) \leq 0 \qquad \forall \mathbb{P} \in \mathbb{F}(\mu).$$

In light of above discussion, we propose another robust optimization approach, i.e., to find the most reliable solution that would protect against the worst uncertainty that might lead to bed shortfalls. To maximize the level of robustness, we push the boundary of uncertainty by maximizing the budget of variation, μ subject to $Z_R(\mu) \leq 0$ as follows

$$\begin{aligned} \mu^* = \max \quad & \mu \\ \text{s.t.} \quad & Z_R(\mu) \leq 0 \\ & \mu \in [0, \infty). \end{aligned} \tag{2.3}$$

Since the set $\mathbb{F}(\mu)$ is nondecreasing in μ, the function $Z_R(\mu)$ is also nondecreasing in μ. As a result, model (2.3) is feasible if and only if $Z_R(0) \leq 0$ and $Z_R(\infty) \geq 0$. Moreover, it is reasonable to assume that the inequalities are strict so that the bed capacity is sufficient to meet average demands but also not overly excessive. As opposed to the fixed budget model, we refer to this model the optimized budget model.

The optimum solution of model (2.3) can easily be obtained by binary search on μ^* so that $Z_R(\mu^*) = 0$. In the following section, we next show how to solve the subproblem, which is model (2.2).

3 Tractable formulation

In this section, we develop a tractable formulation of model (2.2). Since this model is a minimax optimization problem, our approach is to formulate a minimization problem that is equivalent to the inner maximization problem in model (2.2). As a result, model (2.2) can be turned into a deterministic minimization problem.

We focus on the inner maximization problem in model (2.2), i.e.

$$\begin{aligned} \sup \quad & \mathbb{E}_{\mathbb{P}}\left(\max_{t \in \mathcal{T}^+}\left\{\textstyle\sum_{(\tau,l)\in\mathcal{U}_t}(\tilde{a}_{\tau,l} + \tilde{p}_{\tau,l}) - c_t\right\}\right) \\ \text{s.t.} \quad & \mathbb{E}_{\mathbb{P}}(\tilde{p}_{\tau,l}) = \bar{p}_{\tau,l}, && \forall \tau \in \mathcal{T}, l \in \mathcal{L}_\tau, \\ & \mathbb{E}_{\mathbb{P}}(\tilde{p}^2_{\tau,l}) \leq \bar{p}^2_{\tau,l}(1+\mu^2), && \forall \tau \in \mathcal{T}, l \in \mathcal{L}_\tau, \\ & \mathbb{E}_{\mathbb{P}}(\tilde{a}_{\tau,l}) = \bar{a}_{\tau,l}, && \forall \tau \in \mathcal{T}^{--}, l \in \mathcal{L}_\tau, \\ & \mathbb{E}_{\mathbb{P}}(\tilde{a}^2_{\tau,l}) \leq \bar{a}^2_{\tau,l}(1+\mu^2), && \forall \tau \in \mathcal{T}^{--}, l \in \mathcal{L}_\tau, \\ & \mathbb{E}_{\mathbb{P}}(\tilde{a}_{\tau,l}) = \alpha_{\tau,l}\eta_\tau, && \forall \tau \in \mathcal{T}^{+}, l \in \mathcal{L}_\tau, \\ & \mathbb{E}_{\mathbb{P}}(\tilde{a}^2_{\tau,l}) \leq \alpha^2_{\tau,l}(1+\mu^2)\eta^2_\tau, && \forall \tau \in \mathcal{T}^{+}, l \in \mathcal{L}_\tau, \\ & \mathbb{E}_{\mathbb{P}}[1] = 1, \\ & \mathbb{P}\left\{(\tilde{p}_{\tau,l}, \tilde{a}_{\tau,l})_{\tau\in\mathcal{T}, l\in\mathcal{L}_\tau} \in W_p \times W_a(\boldsymbol{\eta})\right\} = 1, \end{aligned} \tag{3.4}$$

where

$$\begin{aligned} W_p &:= \left\{ (p_{\tau,l})_{(\tau,l)\in\mathcal{I}} \;:\; p^0_\tau \geq p_{\tau,l} \geq p_{\tau,l'} \geq 0,\ \forall\, (\tau,l),(\tau,l') \in \mathcal{I}_{\mathcal{T}^{--}}, l' > l \right\}, \\ W_a(\boldsymbol{\eta}) &:= \left\{ (a_{\tau,l})_{(\tau,l)\in\mathcal{I}} \;:\; \begin{array}{ll} a^0_\tau \geq a_{\tau,l} \geq a_{\tau,l'} \geq 0 & \forall\, (\tau,l),(\tau,l') \in \mathcal{I}_{\mathcal{T}^{--}}, l' > l, \\ \eta_\tau \geq a_{\tau,l} \geq a_{\tau,l'} \geq 0 & \forall\, (\tau,l),(\tau,l') \in \mathcal{I}_{\mathcal{T}^{+}}, l' > l. \end{array} \right\}, \end{aligned}$$

and

$$\mathcal{I}_{\mathcal{T}^+} := \{(\tau,l) : \tau \in \mathcal{T}^+, l \in \mathcal{L}_\tau\}, \quad \mathcal{I}_{\mathcal{T}^{--}} := \{(\tau,l) : \tau \in \mathcal{T}^{--}, l \in \mathcal{L}_\tau\}, \quad \mathcal{I} := \mathcal{I}_{\mathcal{T}^+} \cup \mathcal{I}_{\mathcal{T}^{--}}.$$

By definition, it follows that $\mathcal{I} = \{(\tau,l) : \tau \in \mathcal{T}, l \in \mathcal{L}_\tau\}$. Note also that W_p and $W_a(\boldsymbol{\eta})$ are actually the cross products of a number of sets with respect to parameters $\tau \in \mathcal{T}$. For instance, W_p can be rewritten as

$$W_p = \Pi_{\tau\in\mathcal{T}} W^\tau_p,$$

where

$$W^\tau_p := \left\{ (p_{\tau,l})_{l\in\mathcal{L}_\tau} \;:\; p^0_\tau \geq p_{\tau,l} \geq p_{\tau,l'} \geq 0, \forall\, l, l' \in \mathcal{L}_\tau,\ l' > l \right\}, \quad \tau \in \mathcal{T}.$$

Problem (3.4) is a maximization problem over a probability distribution function, which is generally an intractable optimization problem, see for instance Murty and Kabadi (1987). However, under our model of uncertainty, we will show an equivalent formulation of problem (3.4), namely its dual problem, is a minimization problem with a second order conic programming (SOCP) representation. As a result, problem (3.4) can be readily solved by existing commercialized SOCP solvers, such as CPLEX and MOSEK.

For convenience in description, for each $t \in \mathcal{T}$, let $z^t_{\tau,l}$ denote the indicator function defined by

$$z^t_{\tau,l} = \begin{cases} 1, & \text{if } \tau + l = t + 1, \\ 0, & \text{otherwise}, \end{cases}$$

for any $(\tau,l) \in \mathcal{I}$. Noticing that $\mathcal{U}_t = \{(\tau,l) \in \mathcal{I} : \tau + l = t+1\}$, item $\sum_{(\tau,l)\in\mathcal{U}_t} (\tilde{a}_{\tau,l} + \tilde{p}_{\tau,l}) - c_t$ in the objective of (3.4) can then be expressed as below

$$\sum_{(\tau,l)\in\mathcal{I}} (\tilde{a}_{\tau,l} + \tilde{p}_{\tau,l}) z^t_{\tau,l} - c_t, \ \ \forall\, t \in \mathcal{T}^+.$$

We further define the following sets:

$$\begin{aligned} W^\tau_1 &:= \left\{ (p_{\tau,l})_{l\in\mathcal{L}_\tau} \;:\; p^0_\tau \geq p_{\tau,l} \geq p_{\tau,l'} \geq 0, \forall\, l, l' \in \mathcal{L}_\tau,\ l' > l \right\}, \ \tau \in \mathcal{T}, \\ W^\tau_2 &:= \left\{ (a_{\tau,l})_{l\in\mathcal{L}_\tau} \;:\; a^0_\tau \geq a_{\tau,l} \geq a_{\tau,l'} \geq 0, \forall\, l, l' \in \mathcal{L}_\tau,\ l' > l \right\}, \ \tau \in \mathcal{T}^{--}, \\ W^\tau_3(\boldsymbol{\eta}) &:= \left\{ (a_{\tau,l})_{l\in\mathcal{L}_\tau} \;:\; \eta_\tau \geq a_{\tau,l} \geq a_{\tau,l'} \geq 0, \forall\, l, l' \in \mathcal{L}_\tau,\ l' > l \right\}, \ \tau \in \mathcal{T}^{+}. \end{aligned}$$

By applying duality theory of Isii (1963), we derive an equivalent formulation of problem (3.4) as follows:

Theorem 3.1 *Problem (3.4) has the same objective as the following optimization problem:*

$$\inf_{\rho,(s_{\tau,l},u_{\tau,l},v_{\tau,l},w_{\tau,l})_{(\tau,l)\in\mathcal{I}}} \quad \Bigg\{\rho + \sum_{(\tau,l)\in\mathcal{I}} \bar{p}_{\tau,l}s_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}} \bar{p}^2_{\tau,l}(1+\mu^2)u_{\tau,l} +$$

$$\sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}--}} \bar{a}_{\tau,l}v_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}--}} \bar{a}^2_{\tau,l}(1+\mu^2)w_{\tau,l} +$$

$$\sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}+}} \alpha_{\tau,l}\eta_\tau v_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}+}} \alpha^2_{\tau,l}(1+\mu^2)\eta^2_\tau w_{\tau,l}\Bigg\} \tag{3.5}$$

$$\begin{aligned}
\text{s. t.} \quad & \sum_{\tau\in\mathcal{T}} \pi^t_1(\boldsymbol{s}_\tau - \boldsymbol{z}^t_\tau, \boldsymbol{u}_\tau) + \sum_{\tau\in\mathcal{T}^{--}} \pi^t_2(\boldsymbol{v}_\tau - \boldsymbol{z}^t_\tau, \boldsymbol{w}_\tau) + \\
& \sum_{\tau\in\mathcal{T}^+} \pi^t_3(\boldsymbol{v}_\tau - \boldsymbol{z}^t_\tau, \boldsymbol{w}_\tau) + \rho + c_t \geq 0, \ \forall\, t\in\mathcal{T}^+, \\
& u_{\tau,l}, \ \ w_{\tau,l} \geq 0, \ \forall\, (\tau,l)\in\mathcal{I},
\end{aligned}$$

where for $t\in\mathcal{T}^+$,

$$\begin{aligned}
\pi^t_1(\boldsymbol{s}_\tau - \boldsymbol{z}^t_\tau, \boldsymbol{u}_\tau) &:= \min\left\{\sum_{l\in\mathcal{L}_\tau}\left((s_{\tau,l}-z^t_{\tau,l})p_{\tau,l} + u_{\tau,l}p^2_{\tau,l}\right)\Big|(p_{\tau,l})_{l\in\mathcal{L}_\tau}\in W^\tau_1\right\}, \ \tau\in\mathcal{T}, \\
\pi^t_2(\boldsymbol{v}_\tau - \boldsymbol{z}^t_\tau, \boldsymbol{w}_\tau) &:= \min\left\{\sum_{l\in\mathcal{L}_\tau}\left((v_{\tau,l}-z^t_{\tau,l})a_{\tau,l} + w_{\tau,l}a^2_{\tau,l}\right)\Big|(a_{\tau,l})_{l\in\mathcal{L}_\tau}\in W^\tau_2\right\}, \ \tau\in\mathcal{T}^{--}, \\
\pi^t_3(\boldsymbol{v}_\tau - \boldsymbol{z}^t_\tau, \boldsymbol{w}_\tau) &:= \min\left\{\sum_{l\in\mathcal{L}_\tau}\left((v_{\tau,l}-z^t_{\tau,l})a_{\tau,l} + w_{\tau,l}a^2_{\tau,l}\right)\Big|(a_{\tau,l})_{l\in\mathcal{L}_\tau}\in W^\tau_3(\boldsymbol{\eta})\right\}, \ \tau\in\mathcal{T}^{+},
\end{aligned}$$

and $\rho\in\Re$, $\boldsymbol{s}_\tau = (s_{\tau,l})_{l\in\mathcal{L}_\tau}$, $\boldsymbol{u}_\tau = (u_{\tau,l})_{l\in\mathcal{L}_\tau}$, $\boldsymbol{v}_\tau = (v_{\tau,l})_{l\in\mathcal{L}_\tau}$, $\boldsymbol{w}_\tau = (w_{\tau,l})_{l\in\mathcal{L}_\tau}$, $\boldsymbol{z}^t_\tau = (z^t_{\tau,l})_{l\in\mathcal{L}_\tau}$.

Proof. Using the duality theory of infinite-dimensional linear programming in probability spaces (Anderson and Nash (1987), Isii (1963), Vandenberghe et al. (2007)) the dual of problem (3.4) is

$$\inf_{\rho,(s_{\tau,l},u_{\tau,l},v_{\tau,l},w_{\tau,l})_{(\tau,l)\in\mathcal{I}}} \left\{\rho+\sum_{(\tau,l)\in\mathcal{I}}\bar{p}_{\tau,l}s_{\tau,l}+\sum_{(\tau,l)\in\mathcal{I}}\bar{p}^2_{\tau,l}(1+\mu^2)u_{\tau,l}+\right.$$
$$\sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}--}}\bar{a}_{\tau,l}v_{\tau,l}+\sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}--}}\bar{a}^2_{\tau,l}(1+\mu^2)w_{\tau,l}+$$
$$\left.\sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}+}}\alpha_{\tau,l}\eta_\tau v_{\tau,l}+\sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}+}}\alpha^2_{\tau,l}(1+\mu^2)\eta^2_\tau w_{\tau,l}\right\}$$
$$\text{s.t.}\quad \rho+\sum_{(\tau,l)\in\mathcal{I}}s_{\tau,l}p_{\tau,l}+\sum_{(\tau,l)\in\mathcal{I}}u_{\tau,l}p^2_{\tau,l}+\sum_{(\tau,l)\in\mathcal{I}}v_{\tau,l}a_{\tau,l}+\sum_{(\tau,l)\in\mathcal{I}}w_{\tau,l}a^2_{\tau,l} \tag{3.6}$$
$$\geq \sum_{(\tau,l)\in\mathcal{U}_t}(a_{\tau,l}+p_{\tau,l})-c_t,\ \forall\, t\in\mathcal{T}^+,\ \forall\,(p_{\tau,l},a_{\tau,l})_{(\tau,l)\in\mathcal{I}}\in W_p\times W_a(\boldsymbol{\eta}),$$
$$u_{\tau,l},\ \ w_{\tau,l}\geq 0,\ \forall\,(\tau,l)\in\mathcal{I},$$

where $s_{\tau,l},u_{\tau,l},\ v_{\tau,l},w_{\tau,l},\ l\in\mathcal{L}_\tau$, and ρ are the Lagrange multipliers corresponding to equality/inequality constraints concerning the first and second moments of $\tilde{p}_{\tau,l}$ and $\tilde{a}_{\tau,l}$, $(\tau,l)\in\mathcal{I}$, together with the constraint $\mathbb{E}_{\mathbb{P}}[1]=1$. Evidently, the multipliers, $u_{\tau,l}$ and $w_{\tau,l}$, $(\tau,l)\in\mathcal{I}$, corresponding to the inequality constraints are all nonnegative. This property, as we shall see, is very important in the subsequent analysis.

Note that the system of inequality constraints in problem (3.6) consists of infinitely many constraints. To deal with this complex system, we turn to investigate its reduced reformulation. Using the notation of vector $z^t_{\tau,l}$, it is not hard to see that the inequality system in the dual problem (3.6) can be rewritten as:

$$\sum_{(\tau,l)\in\mathcal{I}}\left((s_{\tau,l}-z^t_{\tau,l})p_{\tau,l}+u_{\tau,l}p^2_{\tau,l}\right)+\sum_{(\tau,l)\in\mathcal{I}}\left((v_{\tau,l}-z^t_{\tau,l})a_{\tau,l}+w_{\tau,l}a^2_{\tau,l}\right)$$
$$\geq\ -\rho-c_t,\ \forall\, t\in\mathcal{T}^+,\ \forall\,(p_{\tau,l},a_{\tau,l})_{(\tau,l)\in\mathcal{I}}\in W_p\times W_a(\boldsymbol{\eta}).$$

Evidently, the right hand side of the above system of infinite inequalities is independent on $(p_{\tau,l},a_{\tau,l})$ while two summation terms concerning $p_{t,l}$ and $a_{t,l}$ on the left hand side are independent. Then, the above system can be significantly simplified as follows:

$$\min\left\{\sum_{(\tau,l)\in\mathcal{I}}(s_{\tau,l}-z^t_{\tau,l})p_{\tau,l}+\sum_{(\tau,l)\in\mathcal{I}}u_{\tau,l}p^2_{\tau,l}\Big|(p_{\tau,l})_{(\tau,l)\in\mathcal{I}}\in W_p\right\}$$
$$+\min\left\{\sum_{(\tau,l)\in\mathcal{I}}(v_{\tau,l}-z^t_{\tau,l})a_{\tau,l}+\sum_{(\tau,l)\in\mathcal{I}}w_{\tau,l}a^2_{\tau,l}\Big|(a_{\tau,l})_{(\tau,l)\in\mathcal{I}}\in W_a(\boldsymbol{\eta})\right\}$$
$$\geq-\rho-c_t,\ \forall\, t\in\mathcal{T}^+.$$

Note that the objectives in the above system are separable in $(p_{\tau,l})_{l\in\mathcal{L}_\tau}$, $(a_{\tau,l})_{l\in\mathcal{L}_\tau}$ for $\tau \in \mathcal{T}$, respectively. Noticing that W_p and $W_a(\boldsymbol{\eta})$ can be written as the cross products of some sets with respect to the parameter $\tau \in \mathcal{T}$, thereby the "min" and "sum" operators on the left hand side are exchangeable. By recalling the definitions of W_1^τ, W_2^τ, and $W_3^\tau(\boldsymbol{\eta})$, we have

$$\sum_{\tau\in\mathcal{T}} \pi_1^t(\boldsymbol{s}_\tau - \boldsymbol{z}_\tau^t, \boldsymbol{u}_\tau) + \sum_{\tau\in\mathcal{T}^{--}} \pi_2^t(\boldsymbol{v}_\tau - \boldsymbol{z}_\tau^t, \boldsymbol{w}_\tau) + \sum_{\tau\in\mathcal{T}^+} \pi_3^t(\boldsymbol{v}_\tau - \boldsymbol{z}_\tau^t, \boldsymbol{w}_\tau) \geq -\rho - c_t, \ \forall\, t \in \mathcal{T}^+.$$

Thus, the desired result follows immediately. This completes the proof. □

Note that the equivalent formulation (3.5) in Theorem 3.1 is a deterministic counterpart of the objective function, $Z_R(\mu)$, of robust optimization model (2.2). To derive a tractable reformulation, in what follows we investigate the underlying minimization problems in the constraints of (3.5), i.e., π_i^t, $i = 1, 2, 3$, $t \in \mathcal{T}^+$. First, for any $\tau \in \mathcal{T}$, define an index set $\mathcal{L}_\tau^+ := \mathcal{L}_\tau \cup \{1 + \min\{L, T - \tau\}\}$. We state the results as follows.

Proposition 3.1 *Given $\gamma \in \Re$. For $t \in \mathcal{T}^+$, the following statements hold true.*

(i) For any $\tau \in \mathcal{T}^{--}$, the system of inequality

$$\pi_1^t(\boldsymbol{s}_\tau - \boldsymbol{z}_\tau^t, \boldsymbol{u}_\tau) \geq \gamma \tag{3.7}$$

is second order cone representable in the sense that there exist $\lambda_{\tau,l}^t \geq 0$, $l \in \mathcal{L}_\tau^+$, such that (3.7) is equivalent to

$$\begin{aligned} &\sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^t + p_\tau^0 \lambda_{\tau,1-\tau}^t + \gamma \leq 0, \\ &4u_{\tau,l} y_{\tau,l}^t \geq \left(s_{\tau,l} - z_{\tau,l}^t + \lambda_{\tau,l}^t - \lambda_{\tau,l+1}^t\right)^2, \ \forall\, l \in \mathcal{L}_\tau, \\ &y_{\tau,l}^t \geq 0, \forall\, l \in \mathcal{L}_\tau. \end{aligned} \tag{3.8}$$

(ii) For any $\tau \in \mathcal{T}^+$, the system of inequality (3.7) is second order cone representable in the sense that there exist $\lambda_{\tau,l}^t \geq 0$, $l \in \mathcal{L}_\tau^+$, such that (3.7) is equivalent to

$$\begin{aligned} &\sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^t + p_\tau^0 \lambda_{\tau,1}^t + \gamma \leq 0, \\ &4u_{\tau,l} y_{\tau,l}^t \geq \left(s_{\tau,l} - z_{\tau,l}^t + \lambda_{\tau,l}^t - \lambda_{\tau,l+1}^t\right)^2, \ \forall\, l \in \mathcal{L}_\tau, \\ &y_{\tau,l}^t \geq 0, \forall\, l \in \mathcal{L}_\tau. \end{aligned} \tag{3.9}$$

(iii) For any $\tau \in \mathcal{T}^{--}$, the system of inequality

$$\pi_2^t(\boldsymbol{v}_\tau - \boldsymbol{z}_\tau^t, \boldsymbol{w}_\tau) \geq \gamma \tag{3.10}$$

is second order cone representable in the sense that there exist $\lambda^t_{\tau,l} \geq 0,\ l \in \mathcal{L}^+_\tau$*, such that*

$$\begin{aligned}
&\sum_{l\in\mathcal{L}_\tau} y^t_{\tau,l} + a^0_\tau \lambda^t_{\tau,1-\tau} + \gamma \leq 0,\\
&4w_{\tau,l} y^t_{\tau,l} \geq \left(v_{\tau,l} - z^t_{\tau,l} + \lambda^t_{\tau,l} - \lambda^t_{\tau,l+1}\right)^2,\ \forall\, l \in \mathcal{L}_\tau,\\
&y^t_{\tau,l} \geq 0, \forall\, l \in \mathcal{L}_\tau.
\end{aligned} \tag{3.11}$$

(iv) For any $\tau \in \mathcal{T}^+$*, the system of inequality*

$$\pi^t_3(\boldsymbol{v}_\tau - \boldsymbol{z}^t_\tau, \boldsymbol{w}_\tau) \geq \gamma \tag{3.12}$$

is second order cone representable in the sense that there exist $\lambda^t_{\tau,l} \geq 0,\ l \in \mathcal{L}^+_\tau$*, such that*

$$\begin{aligned}
&\sum_{l\in\mathcal{L}_\tau} y^t_{\tau,l} + \eta_\tau \lambda^t_{\tau,1} + \gamma \leq 0,\\
&4w_{\tau,l} y^t_{\tau,l} \geq \left(v_{\tau,l} - z^t_{\tau,l} + \lambda^t_{\tau,l} - \lambda^t_{\tau,l+1}\right)^2,\ \forall\, l \in \mathcal{L}_\tau,\\
&y^t_{\tau,l} \geq 0, \forall\, l \in \mathcal{L}_\tau.
\end{aligned} \tag{3.13}$$

Proof. (i). By definition, problem $\pi^t_1(\boldsymbol{s}_\tau - \boldsymbol{z}^t_\tau, \boldsymbol{u}_\tau)$ can be written as

$$\begin{aligned}
\pi^t_1(\boldsymbol{s}_\tau - \boldsymbol{z}^t_\tau, \boldsymbol{u}_\tau) = \min \quad & \sum_{l\in\mathcal{L}_\tau} \left(u_{\tau,l} p^2_{\tau,l} + (s_{\tau,l} - z^t_{\tau,l}) p_{\tau,l}\right)\\
\text{s.t.} \quad & p^0_\tau \geq p_{\tau,l} \geq p_{\tau,l'} \geq 0,\ \ \forall\, l, l' \in \mathcal{L}_\tau,\ l' > l.
\end{aligned} \tag{3.14}$$

Note that the above problem is a quadratic programming in which the coefficients concerning the second degree are nonnegative as $u_{\tau,l} \geq 0$ for $l \in \mathcal{L}_\tau$ by Theorem 3.1. To solve this problem, we consider its dual as given below:

$$\max_{\boldsymbol{\lambda}_\tau \geq 0} \zeta(\boldsymbol{\lambda}^t_\tau), \tag{3.15}$$

where $\zeta(\boldsymbol{\lambda}^t_\tau)$ is the associated Lagrange dual function, $\boldsymbol{\lambda}^t_\tau \in \Re^{|\mathcal{L}_\tau|+1}$ denotes the vector of the corresponding Lagrange multipliers, and for any given set S, $|S|$ denotes the cardinality of S.

Let $\boldsymbol{p}_\tau = (p_{\tau,l})_{l\in\mathcal{L}_\tau}$. For convenience in description and without loss of generality, we assume the indices of the entries in vector $\boldsymbol{\lambda}^t_\tau$ are consistent with those of $\boldsymbol{p}_\tau$, i.e., $\boldsymbol{\lambda}^t_\tau = (\lambda^t_{\tau,l})_{l\in\mathcal{L}^+_\tau}$. Applying some basic operations, it gives the Lagrange dual function as follows:

$$\zeta(\boldsymbol{\lambda}^t_\tau) := \min_{(p_{\tau,l})_{l\in\mathcal{L}_\tau}} \left\{ \sum_{l\in\mathcal{L}_\tau} \left(u_{\tau,l} p^2_{\tau,l} + (s_{\tau,l} - z^t_{\tau,l} + \lambda^t_{\tau,l} - \lambda^t_{\tau,l+1}) p_{\tau,l}\right) - p^0_\tau \lambda^t_{\tau,1-\tau} \right\}.$$

Note that the Slater's condition holds true since the interior of the feasible region of problem (3.14) is nonempty. By the strong duality theorem, the system of inequality (3.7) can then be written as what follows. There exist $\lambda^t_{\tau,l} \geq 0,\ l \in \mathcal{L}^+_\tau$, such that

$$\min_{(p_{\tau,l})_{l\in\mathcal{L}_\tau}} \left\{ \sum_{l\in\mathcal{L}_\tau} \left(u_{\tau,l} p^2_{\tau,l} + (s_{\tau,l} - z^t_{\tau,l} + \lambda^t_{\tau,l} - \lambda^t_{\tau,l+1}) p_{\tau,l} \right) \right\} - p^0_\tau \lambda^t_{\tau,1-\tau} \geq \gamma,$$

which, by virtue of the separability of the above minimization problem in $p_{\tau,l}$, can be further reformulated as

$$\sum_{l\in\mathcal{L}_\tau} \left\{ \min_{p_{\tau,l}} \left(u_{\tau,l} p^2_{\tau,l} + (s_{\tau,l} - z^t_{\tau,l} + \lambda^t_{\tau,l} - \lambda^t_{\tau,l+1}) p_{\tau,l} \right) \right\} - p^0_\tau \lambda^t_{\tau,1-\tau} \geq \gamma. \tag{3.16}$$

To investigate the quadratic programming problems on the left hand side of (3.16), we consider the following two cases: (a) $u_{\tau,l} > 0$ for all $l \in \mathcal{L}_\tau$; (b) $u_{\tau,l} = 0$ for some $l \in \mathcal{L}_\tau$, respectively.

For case (a), solving the optimality condition of each minimization problem involved, i.e., $2u_{\tau,l} p_{\tau,l} + s_{\tau,l} - z^t_{\tau,l} + \lambda^t_{\tau,l} - \lambda^t_{\tau,l+1} = 0,\ l \in \mathcal{L}_\tau$, we immediately derive the optimal solution and the corresponding optimal value, which are denoted by $p^*_{\tau,l}$ and $f^*_{\tau,l}$ as follows:

$$\begin{aligned} p^*_{\tau,l} &= \frac{1}{2u_{\tau,l}} \left(z^t_{\tau,l} - s_{\tau,l} + \lambda^t_{\tau,l+1} - \lambda^t_{\tau,l} \right),\ \ l \in \mathcal{L}_\tau, \\ f^*_{\tau,l} &= -\frac{1}{4u_{\tau,l}} \left(s_{\tau,l} - z^t_{\tau,l} + \lambda^t_{\tau,l} - \lambda^t_{\tau,l+1} \right)^2,\ \ l \in \mathcal{L}_\tau. \end{aligned}$$

Substituting the optimal value $f^*_{\tau,l}$ to the inequality (3.16), it yields that

$$\sum_{l\in\mathcal{L}_\tau} \frac{1}{4u_{\tau,l}} \left(s_{\tau,l} - z^t_{\tau,l} + \lambda^t_{\tau,l} - \lambda^t_{\tau,l+1} \right)^2 + p^0_\tau \lambda^t_{\tau,1-\tau} \leq -\gamma. \tag{3.17}$$

To derive a second order cone representation, we introduce the additional variables $y^t_{\tau,l},\ l \in \mathcal{L}_\tau,\ t \in \mathcal{T}^+$ such that

$$\frac{1}{4u_{\tau,l}} (s_{\tau,l} - z^t_{\tau,l} + \lambda^t_{\tau,l} - \lambda^t_{\tau,l+1})^2 \leq y^t_{\tau,l},\ l \in \mathcal{L}_\tau.$$

Thereby, system (3.17) is equivalent to

$$\begin{aligned} &\sum_{l\in\mathcal{L}_\tau} y^t_{\tau,l} + p^0_\tau \lambda^t_{\tau,1-\tau} \leq -\gamma, \\ &4u_{\tau,l} y^t_{\tau,l} \geq \left(s_{\tau,l} - z^t_{\tau,l} + \lambda^t_{\tau,l} - \lambda^t_{\tau,l+1} \right)^2,\ \forall\, l \in \mathcal{L}_\tau, \\ &y^t_{\tau,l} \geq 0, \forall\, l \in \mathcal{L}_\tau, \end{aligned} \tag{3.18}$$

which is a second order cone representation as desired.

For case (b), the analysis is similar to case (a), but becomes much simple, as the underlying problem reduces to a linear programming in this case. Noticing that $\pi_1^t(\boldsymbol{s}_\tau - \boldsymbol{z}_\tau^t, \boldsymbol{u}_\tau)$ is lower bounded by a constant γ, we then have $s_{\tau,l} - z_{\tau,l}^t + \lambda_{\tau,l}^t - \lambda_{\tau,l+1}^t = 0$ and $p_{\tau,l}^* = 0$. Thereby, system (3.18) is valid as well.

$(ii) - (vi)$. The arguments for these cases are similar to case (i). For brevity, here we omit the details. This completes the proof. □

Using Theorem 3.1 and Proposition 3.1, we are ready to derive the following result concerning the tractability of robust optimization model (2.2), which is a main result of this paper.

Theorem 3.2 *Robust optimization model (2.2) is equivalent to the following second order cone programming problem*

$$
\begin{aligned}
\inf_{\substack{\rho, \boldsymbol{\eta}, (\boldsymbol{s}_\tau, \boldsymbol{u}_\tau, \boldsymbol{v}_\tau, \boldsymbol{w}_\tau)_{\tau\in\mathcal{T}},\\ (\boldsymbol{\lambda}_\tau^p, \boldsymbol{\lambda}_\tau^a, \boldsymbol{y}_\tau^p, \boldsymbol{y}_\tau^a)_{\tau\in\mathcal{T}}}} \quad & \Bigg\{\rho + \sum_{(\tau,l)\in\mathcal{I}} \bar{p}_{\tau,l} s_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}} \bar{p}_{\tau,l}^2 (1+\mu^2) u_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}^{--}}} \bar{a}_{\tau,l} v_{\tau,l} + \\
& \sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}^{--}}} \bar{a}_{\tau,l}^2 (1+\mu^2) w_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}^{+}}} \alpha_{\tau,l} v_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}^{+}}} \alpha_{\tau,l}^2 (1+\mu^2) w_{\tau,l} \Bigg\} \\
\text{s. t.} \quad & \sum_{\tau\in\mathcal{T}} \sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,p} + \sum_{\tau\in\mathcal{T}^{--}} p_\tau^0 \lambda_{\tau,1-\tau}^{t,p} + \sum_{\tau\in\mathcal{T}^{+}} p_\tau^0 \lambda_{\tau,1}^{t,p} + + \sum_{\tau\in\mathcal{T}} \sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,a} \\
& + \sum_{\tau\in\mathcal{T}^{+}} \lambda_{\tau,1}^{t,a} + \sum_{\tau\in\mathcal{T}^{--}} a_\tau^0 \lambda_{\tau,1-\tau}^{t,a} \le \rho + c_t, \ \forall\, t\in\mathcal{T}^+, \qquad (3.19) \\
& 4u_{\tau,l} y_{\tau,l}^{t,p} \ge \left(s_{\tau,l} - z_{\tau,l}^t + \lambda_{\tau,l}^{t,p} - \lambda_{\tau,l+1}^{t,p}\right)^2, \ \forall\, l\in\mathcal{L}_\tau,\ \forall\,\tau\in\mathcal{T},\ \forall\, t\in\mathcal{T}^+, \\
& 4w_{\tau,l} y_{\tau,l}^{t,a} \ge \left(v_{\tau,l} - z_{\tau,l}^t + \lambda_{\tau,l}^{t,a} - \lambda_{\tau,l+1}^{t,a}\right)^2, \ \forall\, l\in\mathcal{L}_\tau,\ \forall\,\tau\in\mathcal{T}^{--},\ \forall\, t\in\mathcal{T}^+, \\
& 4w_{\tau,l} y_{\tau,l}^{t,a} \ge \left(v_{\tau,l} - \eta_\tau z_{\tau,l}^t + \lambda_{\tau,l}^{t,a} - \lambda_{\tau,l+1}^{t,a}\right)^2, \ \forall\, l\in\mathcal{L}_\tau,\ \forall\,\tau\in\mathcal{T}^{+},\ \forall\, t\in\mathcal{T}^+, \\
& u_{\tau,l} \ge 0,\ w_{\tau,l} \ge 0,\ \forall\, l\in\mathcal{L}_\tau,\ \forall\,\tau\in\mathcal{T}, \\
& \lambda_{\tau,l}^{t,p} \ge 0,\ \lambda_{\tau,l}^{t,a} \ge 0,\ \forall\, l\in\mathcal{L}_\tau^+,\ \forall\,\tau\in\mathcal{T},\ \forall\, t\in\mathcal{T}^+, \\
& y_{\tau,l}^{t,p} \ge 0,\ y_{\tau,l}^{t,a} \ge 0,\ \forall\, l\in\mathcal{L}_\tau,\ \forall\,\tau\in\mathcal{T},\ \forall\, t\in\mathcal{T}^+, \\
& \boldsymbol{\eta} \in X.
\end{aligned}
$$

Proof. First, we rewrite the system of inequality constraints of problem (3.5) as

$$
\begin{aligned}
& \sum_{\tau\in\mathcal{T}} \pi_1^t(\boldsymbol{s}_\tau - \boldsymbol{z}_\tau^t, \boldsymbol{u}_\tau) + \sum_{\tau\in\mathcal{T}^{--}} \pi_2^t(\boldsymbol{v}_\tau - \boldsymbol{z}_\tau^t, \boldsymbol{w}_\tau) + \sum_{\tau\in\mathcal{T}^{+}} \pi_3^t(\boldsymbol{v}_\tau - \boldsymbol{z}_\tau^t, \boldsymbol{w}_\tau) \qquad (3.20) \\
& \ge -\rho - c_t,\ \forall\, t\in\mathcal{T}^+.
\end{aligned}
$$

Then according to Proposition 3.1 and applying some necessary operations, for each $t \in \mathcal{T}^+$, there exist some Lagrange multipliers $\lambda_{\tau,l}^{t,p}$ and $\lambda_{\tau,l}^{t,a}$, $l \in \mathcal{L}_\tau^+$, $\tau \in \mathcal{T}$, such that (3.20) is equivalent to

$$\begin{aligned}
&\sum_{\tau\in\mathcal{T}^+}\left(\sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,p} + p_\tau^0\lambda_{\tau,1}^{t,p}\right) + \sum_{\tau\in\mathcal{T}^{--}}\left(\sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,p} + p_\tau^0\lambda_{\tau,1-\tau}^{t,p}\right) + \sum_{\tau\in\mathcal{T}^+}\left(\sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,a} + \eta_\tau\lambda_{\tau,1}^{t,a}\right)\\
&+\sum_{\tau\in\mathcal{T}^{--}}\left(\sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,a} + a_\tau^0\lambda_{\tau,1-\tau}^{t,a}\right) \le \rho + c_t,\ \forall\, t\in\mathcal{T}^+,\\
&4u_{\tau,l}y_{\tau,l}^{t,p} \ge \left(s_{\tau,l} - z_{\tau,l}^t + \lambda_{\tau,l}^{t,p} - \lambda_{\tau,l+1}^{t,p}\right)^2,\ \forall\, l\in\mathcal{L}_\tau,\ \forall\,\tau\in\mathcal{T},\ \forall\, t\in\mathcal{T}^+,\\
&4w_{\tau,l}y_{\tau,l}^{t,a} \ge \left(v_{\tau,l} - z_{\tau,l}^t + \lambda_{\tau,l}^{t,a} - \lambda_{\tau,l+1}^{t,a}\right)^2,\ \forall\, l\in\mathcal{L}_\tau,\ \forall\,\tau\in\mathcal{T},\ \forall\, t\in\mathcal{T}^+,\\
&y_{\tau,l}^{t,p} \ge 0,\ \ \forall\, l\ \in\mathcal{L}_\tau,\ \forall\,\tau\in\mathcal{T},\ \forall\, t\in\mathcal{T}^+,\\
&y_{\tau,l}^{t,a} \ge 0,\ \ \forall\, l\in\mathcal{L}_\tau,\ \forall\,\tau\in\mathcal{T},\ \forall\, t\in\mathcal{T}^+.
\end{aligned}$$

On the other hand, according to Theorem 3.1, model (2.2) is actually a "min-min" two-stage problem Thus, model (2.2) is equivalent to the following problem:

$$\begin{aligned}
\inf_{\substack{\rho,\boldsymbol{\eta},(\boldsymbol{s}_\tau,\boldsymbol{u}_\tau,\boldsymbol{v}_\tau,\boldsymbol{w}_\tau)_{\tau\in\mathcal{T}},\\(\boldsymbol{\lambda}_\tau^p,\boldsymbol{\lambda}_\tau^a,\boldsymbol{y}_\tau^p,\boldsymbol{y}_\tau^a)_{\tau\in\mathcal{T}}}}\quad & \Bigg\{\rho + \sum_{(\tau,l)\in\mathcal{I}} \bar{p}_{\tau,l}s_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}} \bar{p}_{\tau,l}^2(1+\mu^2)u_{\tau,l}\\
&+\sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}^{--}}} \bar{a}_{\tau,l}v_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}^{--}}} \bar{a}_{\tau,l}^2(1+\mu^2)w_{\tau,l}\\
&+\sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}^{+}}} \alpha_{\tau,l}\eta_\tau v_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}^{+}}} \alpha_{\tau,l}^2(1+\mu^2)\eta_\tau^2 w_{\tau,l}\Bigg\}\\
\text{s. t.}\quad & \sum_{\tau\in\mathcal{T}^+}\left(\sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,p} + p_\tau^0\lambda_{\tau,1}^{t,p}\right) + \sum_{\tau\in\mathcal{T}^{--}}\left(\sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,p} + p_\tau^0\lambda_{\tau,1-\tau}^{t,p}\right)\\
&+\sum_{\tau\in\mathcal{T}^+}\left(\sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,a} + \eta_\tau\lambda_{\tau,1}^{t,a}\right) + \sum_{\tau\in\mathcal{T}^{--}}\left(\sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,a} + a_\tau^0\lambda_{\tau,1-\tau}^{t,a}\right)\\
&\le \rho + c_t,\ \forall\, t\in\mathcal{T}^+, \qquad\qquad (3.21)\\
&4u_{\tau,l}y_{\tau,l}^{t,p} \ge \left(s_{\tau,l} - z_{\tau,l}^t + \lambda_{\tau,l}^{t,p} - \lambda_{\tau,l+1}^{t,p}\right)^2,\ \forall\, l\in\mathcal{L}_\tau,\ \forall\,\tau\in\mathcal{T},\ \forall\, t\in\mathcal{T}^+,\\
&4w_{\tau,l}y_{\tau,l}^{t,a} \ge \left(v_{\tau,l} - z_{\tau,l}^t + \lambda_{\tau,l}^{t,a} - \lambda_{\tau,l+1}^{t,a}\right)^2,\ \forall\, l\in\mathcal{L}_\tau,\ \forall\,\tau\in\mathcal{T},\ \forall\, t\in\mathcal{T}^+,\\
&\lambda_{\tau,l}^{t,p} \ge 0,\ \lambda_{\tau,l}^{t,a} \ge 0,\ \forall\, l\in\mathcal{L}_\tau^+,\ \forall\,\tau\in\mathcal{T},\ \forall\, t\in\mathcal{T}^+,\\
&u_{\tau,l} \ge 0,\ w_{\tau,l} \ge 0,\ \forall\, l\in\mathcal{L}_\tau,\ \forall\,\tau\in\mathcal{T},\\
&y_{\tau,l}^{t,p} \ge 0,\ y_{\tau,l}^{t,a} \ge 0,\ \forall\, l\in\mathcal{L}_\tau,\ \forall\,\tau\in\mathcal{T},\ \forall\, t\in\mathcal{T}^+,\\
&\boldsymbol{\eta}\in X.
\end{aligned}$$

Note that the objective function of problem (3.21) is nonlinear (actually it is of degree three) with respect to the decision variables. Fortunately, we turn to linearize the objective function using the transformation of the underlying variables. Specifically, we make the following transformations on variables under consideration:

$$\bar{v}_{\tau,l} := \eta_\tau v_{\tau,l},\ \bar{w}_{\tau,l} := \eta_\tau^2 w_{\tau,l}, \forall\, l \in \mathcal{L}_\tau, \tau \in \mathcal{T}^+,$$
$$\bar{\lambda}_{\tau,l}^{t,a} := \eta_\tau \lambda_{\tau,l}^{t,a},\ \forall\, l \in \mathcal{L}_\tau^+, \tau \in \mathcal{T}^+,\ t \in \mathcal{T}^+.$$

Then, the objective function of (3.21) can be converted as

$$\rho + \sum_{(\tau,l)\in\mathcal{I}} \bar{p}_{\tau,l} s_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}} \bar{p}_{\tau,l}^2 (1+\mu^2) u_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}--}} \bar{a}_{\tau,l} v_{\tau,l}$$
$$+ \sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}--}} \bar{a}_{\tau,l}^2 (1+\mu^2) w_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}+}} \alpha_{\tau,l} \bar{v}_{\tau,l} + \sum_{(\tau,l)\in\mathcal{I}_{\mathcal{T}+}} \alpha_{\tau,l}^2 (1+\mu^2) \bar{w}_{\tau,l},$$

which, clearly, is a linear function. Note also that according to the previous arguments, we have

$$\bar{w}_{\tau,l} \geq 0,\ \forall\, l \in \mathcal{L}_\tau,\ \forall\, \tau \in \mathcal{T}^+;\ \bar{\lambda}_{\tau,l}^{t,a} \geq 0,\ \forall\, l \in \mathcal{L}_\tau^+,\ \forall\, \tau \in \mathcal{T}^+,\ \forall\, t \in \mathcal{T}^+.$$

Next, we consider the terms associated with the underlying variables $v_{\tau,l},\ w_{\tau,l}$ and $\lambda_{\tau,l}^{t,a}$ involved in the constraints of (3.21), namely,

$$\sum_{\tau\in\mathcal{T}} \sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,p} + \sum_{\tau\in\mathcal{T}^+} p_\tau^0 \lambda_{\tau,1}^{t,p} + \sum_{\tau\in\mathcal{T}^{--}} p_\tau^0 \lambda_{\tau,1-\tau}^{t,p} + \sum_{\tau\in\mathcal{T}} \sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,a} + \sum_{\tau\in\mathcal{T}^+} \eta_\tau \lambda_{\tau,1}^{t,a} + \sum_{\tau\in\mathcal{T}^{--}} a_\tau^0 \lambda_{\tau,1-\tau}^{t,a}$$
$$\leq \rho + c_t,\ \forall\, t \in \mathcal{T}^+,$$

and

$$4 w_{\tau,l} y_{\tau,l}^{t,a} \geq \left(v_{\tau,l} - z_{\tau,l}^t + \lambda_{\tau,l}^{t,a} - \lambda_{\tau,l+1}^{t,a} \right)^2,\ \forall\, l \in \mathcal{L}_\tau,\ \forall\, \tau \in \mathcal{T}^+, \forall\, t \in \mathcal{T}^+.$$

Accordingly, using the introduced transformations, the above two systems can be rewritten as

$$\sum_{\tau\in\mathcal{T}} \sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,p} + \sum_{\tau\in\mathcal{T}^+} p_\tau^0 \lambda_{\tau,1}^{t,p} + \sum_{\tau\in\mathcal{T}^{--}} p_\tau^0 \lambda_{\tau,1-\tau}^{t,p} + \sum_{\tau\in\mathcal{T}} \sum_{l\in\mathcal{L}_\tau} y_{\tau,l}^{t,a} + \sum_{\tau\in\mathcal{T}^+} \bar{\lambda}_{\tau,1}^{t,a} + \sum_{\tau\in\mathcal{T}^{--}} a_\tau^0 \lambda_{\tau,1-\tau}^{t,a}$$
$$\leq \rho + c_t,\ \forall\, t \in \mathcal{T}^+,$$

and

$$4 \bar{w}_{\tau,l} y_{\tau,l}^{t,a} \geq \left(\bar{v}_{\tau,l} - \eta_\tau z_{\tau,l}^t + \bar{\lambda}_{\tau,l}^{t,a} - \bar{\lambda}_{\tau,l+1}^{t,a} \right)^2,\ \forall\, l \in \mathcal{L}_\tau,\ \forall\, \tau \in \mathcal{T}^+,\ \forall\, t \in \mathcal{T}^+,$$

respectively.

For notational convenience, we rewrite the variables $\bar{v}_{\tau,l}$, $\bar{w}_{\tau,l}$, $\bar{\lambda}^{t,a}_{\tau,l}$ by $v_{\tau,l}$, $w_{\tau,l}$, $\lambda^{t,a}_{\tau,l}$, respectively. Then problem (3.19) follows straightforwardly. This completes the proof. □

According to Theorem 3.2, we obtain a second order cone programming reformulation of the original robust model (2.2). It is well known that second order cone programming is tractable and powerful in computation, which has become to more and more popular in numerical optimization in recent years, especially for solving large scale problems arising from operations management in industry.

4 Empirical studies

In this section, we study the performance of our robust optimization models using real data from a hospital in Singapore. Our data set consists of daily admission and length of stay of both emergency and elective patients throughout the year of 2008. For data sensitivity considerations, we scaled the original data in a proportionate manner, and all following discussions are based on the adjusted data. Table 1 presents basic statistics concerning the admissions and length of stay of both emergency patients and elective patients. Emergency patients, averaging about 119 daily, account for about 82% of daily admissions at National University Hospital (NUH) of Singapore. Their mean length of stay (LOS) at 3.57 days exceeds that of elective patients by about 1 day. Out of the mean daily 119 emergency admissions, about 41 (24, 16, 10, 8, 6, 4, ...) stay for 1 (2, 3, 4, 5, 6, 7, ...) days. However, both groups have about similar relative volatility in their LOS, as indicated by the coefficient of variation statistics.

Patient Type	Emergency	Elective
Daily Admissions: Mean	119.10	25.98
Daily Admissions: Std Dev	16.15	13.12
Daily Admissions: CV	0.14	0.51
LOS: Mean	3.57	3.21
LOS: Std Dev	2.97	2.70
LOS: CV	0.83	0.84

Table 1: Statistics of Emergency Patients and Elective Patients

Figure 1 shows the autocorrelation plot of daily emergency admissions at NUH across the year of 2008. Our investigation of seasonality in daily emergency admissions reveals volatility across the days rather than across the months. The patterns of elective admissions more or less mirror those appearing in the graphics below for emergency admissions.

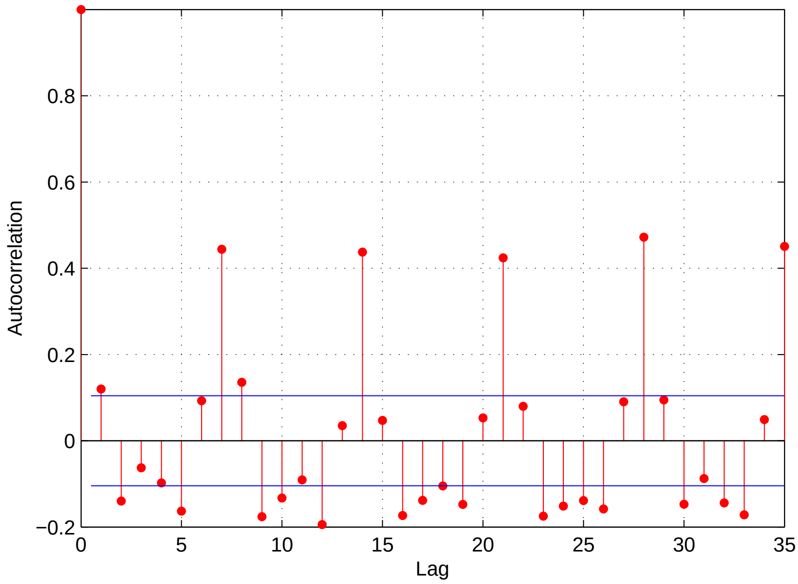

Figure 1: Autocorrelation of Daily Emergency Admissions

Figure 2 shows the average daily emergency admission pattern within a week. There is an obvious weekly pattern. On average, we observe less emergency admissions during the weekend. Within a week, we see the greatest number of emergency admissions on Monday.

4.1 Methodology

We separate the one-year time duration of our data set into two periods. The first period, consisting of $T_0 = 199$ days, is the learning period. We use the data in this period to estimate the means of admission variables. The second period, consisting of the rest 167 days, is the simulation period. We simulate the admission process using the deterministic model and the robust model, respectively.

A rolling horizon approach is adopted in our simulation. Specifically, we will solve the elective admission problem once a week, taking into consideration the next $T = 21$ days (referred to as the planning horizon). The resultant optimal elective admission quota will cover 21 days, however, only the quota for the first 7 days will be implemented in the simulation. The quota for the rest 14 days will *not* be used. If we do not include

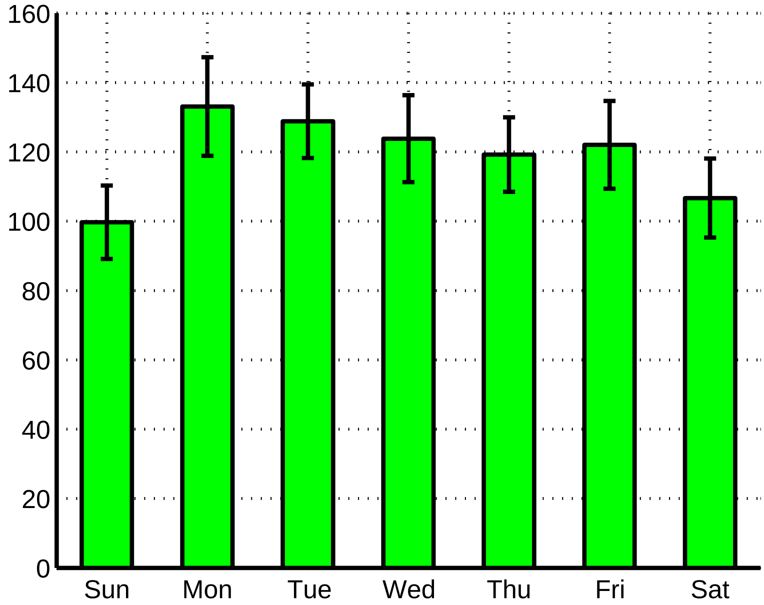

Figure 2: Average Daily Emergency Admissions by Weekday (Error Bars Indicate Standard Deviations)

extra days (extra in the sense that the quota of these days will not be implemented) in the planning horizon, the elective admission model tends to assign overly large admission quota to the last day(s) of the planning horizon, which will lead to high bed occupancy immediately following the planning period, which is not properly accounted for by the model. Such effect is called the boundary effect. Including extra days in the planning horizon allows us to mitigate the boundary effect.

We impose daily upper bounds and lower bounds for the quota sizes. In addition, there is a weekly minimum quota to be assigned for every week. For simplicity, we assume that these bounds are the same across days and weeks. As such, the feasible region X for the quota has the following form:

$$X = \left\{ \boldsymbol{\eta} \in \Re^T \ : \ \kappa^u \geq \eta_\tau \geq \kappa^b, \forall \tau \in \mathcal{T}^+; \sum_{\tau=7(n-1)}^{7n-1} \eta_\tau \geq \kappa^w, \forall n = 1, \ldots, T/7 \right\},$$

where κ^u, κ^b, κ^w are nonnegative scalars given in advance. The constraints as set in X are mainly based on the following considerations. To improve the service level for emergency inpatients, the hospital should try to avoid the over-utilization of the limited sources such as, doctors, nurses, medical equipments. This means that the hospital should not reserve or assign bed slots to emergency patients as many as possible without considering the demands from elective patients during the planning period. Then, the hospital requires that the weekly number of bed slots allocated to elective patients during the planning horizon $\mathcal{T}^+$ is no less than the given threshold κ^w.

Once we solve our model and obtain the elective admission quota, we will simulate the patient admission process for the following week. We will use the real emergency admission and LOS data because it is reasonable to assume that such data are independent of the elective patient admission quota. However, on day t, the real elective admission data $\{a_{t,l}\}_{l=1}^L$ is based on the real elective admission quota (namely, $a_{t,1}$). Hence such data must be changed to reflect the new quota imposed by our model. This is done in the following way. For any day t, let η_t^* be the optimal quota of the day determined by our model. We scale the number of elective patients admitted on day t with their LOS's ranging from 1 up to L days, each by a factor of $\eta_t/a_{t,1}$ (again ignoring the integrality constraint). Essentially, we preserve the proportion of patients with different LOS's, but scale their numbers so that the imposed admission quota is met.

In order to run the simulation, we need to estimate the means of admission variables $(\tilde{p}_{t,l}, \tilde{a}_{t,l})_{(t,l)\in\mathcal{I}_{\mathcal{T}^+}}$. Ideally these estimations should be achieved using the input of doctors. Due to the absence of such input in our simulation, however, we obtain such estimation purely via historical admission and LOS data in the learning period. In particular, we estimate the following parameters from the historical data.

- $\gamma_l^{p,\text{DOW}}$: the average number of emergency patients admitted on a particular day of week (DOW) and will be warded for at least l days.
- $\pi_{l_1,l_2}^{p,\text{DOW}}$($\pi_{l_1,l_2}^{a,\text{DOW}}$): the probability that an emergency (or elective) patient admitted on a particular day of week will be warded for at least l_2 days given that this patient has already stayed for l_1 days.

For $t \in \mathcal{T}^{--}$, the values of $\tilde{p}_{t,-t}$ and $\tilde{a}_{t,-t}$ are already known by the time of planning. Thus we have the following estimation

$$\mathbb{E}_{\mathbb{P}}[\tilde{p}_{t,l}] = \tilde{p}_{t,-t}\pi_{-t,l}^{p,\text{DOW}},$$

$$\mathbb{E}_{\mathbb{P}}[\tilde{a}_{t,l}] = \tilde{a}_{t,-t}\pi_{-t,l}^{a,\text{DOW}}.$$

For $t \in \mathcal{T}^+$, we make the following estimation

$$\mathbb{E}_{\mathbb{P}}[\tilde{p}_{t,l}] = \gamma_l^{p,\text{DOW}},$$

$$\mathbb{E}_{\mathbb{P}}[\tilde{a}_{t,l}] = \eta_t\pi_{1,l}^{a,\text{DOW}}.$$

4.2 Simulation results

Under the above setting, our model can be solved efficiently. For $L = 14$ and $T = 21$, the SOCP equivalent of model (2.2) has 60,424 variables. Solving this model takes less than 10 seconds on a 12-core 2.4GHz Mac Pro computer using either the CPLEX solver or the MOSEK solver.

We are interested in comparing the performance of the robust model against that of the deterministic model in simulations. We consider different scenarios by using different bed capacities and different starting dates of the simulation period. The constraints on elective admission quota are set as follows. The daily upper bounds for quota is set to be 50 and the daily lower bounds for the quota is set to be 5. The weekly lower bounds for quota is set to be 175. We assume that the bed capacity is fixed throughout the simulation period.

To determine the value of μ in the fixed budget model (2.2), we note that the empirical coefficient of variation for daily emergency and elective patient admission is respectively 0.14 and 0.51 (see Table 1). For our simulation, we set μ to be 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6 respectively. We also note that the deterministic model (2.1) corresponds to the fixed budget model with $\mu = 0$.

In our simulation, we consider different scenarios by imposing different bed capacities (540, 550, 560 and 570) and using different simulation starting dates (199, 219, 239, 259 and 279). For each scenario, we measure the performance of various models by the total bed shortage in the simulation period (in the unit of bed·day).

In Table 2, we present the comparison of results of different models when the bed capacity is 540. The total bed shortage is shown for different scenarios found by varying the starting date of simulation period. In addition, we summarize the performance of a model by summing up the total bed shortage resulting from this model for different starting dates. This is shown in the last row of the table. We present the results for other capacities in Table 3, 4 and 5.

By examining the computational results, we note that as we increase the budget of variation from 0 to 0.6, the performance level of the fixed μ model initially improves, but then deteriorates. The optimal level of performance is achieved when μ is around 0.2. This implies that for fixed μ models, there exists an optimal level of budget of variation.

In addition, we note that our optimized budget of variation model consistently performs better than the deterministic model (i.e., the model with $\mu = 0$), irrespective of our choice of bed capacity. This suggests that our optimized budget model is superior to the deterministic model.

In our simulation results, despite the model with fixed budget at $\mu = 0.2$ achieves better performance compared to the optimized budget model, we note that it is unclear how we can properly choose the value of μ prior to the simulation. The advantage of the optimized budget model is that it doesn't need us to find the best value of μ, and it performs well without much tinkering.

	Total bed shortage (capacity = 540)							
T_0	Opt budget	$\mu = 0$	$\mu = 0.1$	$\mu = 0.2$	$\mu = 0.3$	$\mu = 0.4$	$\mu = 0.5$	$\mu = 0.6$
199	196	206	219	211	206	209	240	240
219	133	135	177	107	109	130	95	173
239	161	161	143	146	186	184	196	182
259	39	43	46	50	90	97	93	86
279	33	80	37	38	63	81	149	178
Sum	561	626	624	553	653	702	772	858

Table 2: Comparison of Models: Varying Starting Date (Bed Capacity is 540).

	Total bed shortage (capacity = 550)							
T_0	Opt budget	$\mu = 0$	$\mu = 0.1$	$\mu = 0.2$	$\mu = 0.3$	$\mu = 0.4$	$\mu = 0.5$	$\mu = 0.6$
199	99	108	82	101	77	83	116	129
219	64	62	98	41	25	32	16	66
239	76	74	71	73	83	86	103	97
259	13	12	8	7	33	34	32	18
279	7	4	8	10	23	28	64	81
Sum	259	261	267	232	241	263	331	391

Table 3: Comparison of Models: Varying Starting Date (Bed Capacity is 550).

5 Conclusions

In this study, we present a new robust approach to manage elective admissions in hospital. Our model also contributes to the methodology of robust optimization. In formulating our optimization model, instead of using the worst-case performance as the objective, we propose to maximize the level of uncertainty such that the worst-case performance meets a pre-specified target. In our problem, this method proves to provide fairly good performance without tinkering with the model parameters. We showed how to solve our

	Total bed shortage (capacity = 560)							
T_0	Opt budget	$\mu = 0$	$\mu = 0.1$	$\mu = 0.2$	$\mu = 0.3$	$\mu = 0.4$	$\mu = 0.5$	$\mu = 0.6$
199	42	52	25	46	31	31	55	67
219	25	24	45	6	0	1	0	9
239	36	36	33	32	34	36	52	47
259	0	2	0	0	5	4	4	2
279	0	0	0	0	2	6	4	14
Sum	103	115	103	83	71	79	116	139

Table 4: Comparison of Models: Varying Starting Date (Bed Capacity is 560).

	Total bed shortage (capacity = 570)							
T_0	Opt budget	$\mu = 0$	$\mu = 0.1$	$\mu = 0.2$	$\mu = 0.3$	$\mu = 0.4$	$\mu = 0.5$	$\mu = 0.6$
199	12	17	0	6	10	11	18	26
219	7	10	17	0	0	0	0	0
239	15	13	16	19	11	11	24	24
259	0	0	0	0	0	0	0	0
279	0	0	0	0	0	0	0	1
Sum	33	40	33	25	21	23	42	51

Table 5: Comparison of Models: Varying Starting Date (Bed Capacity is 570).

model via transformation into a second order conic problem. We also performed empirical studies based on real data obtained from hospital. Finally, we tested the proposed model in a simulation study based on real hospital admission data. Simulation results suggest that the robust optimization approach has better performance compared to the deterministic approach.

Acknowledgement

The authors would like to thank Joe Sim (CEO, National University Hospital, Singapore) for the connections and support of this project and Kiok-Liang Teow (National Healthcare Group, Singapore) for his feedbacks and suggestions for this paper.

References

[1] Anderson, E. J., and P. Nash. 1987. *Linear Programming in Infinite-dimensional Spaces: Theory and Applications.* Wiley, Singapore.

[2] Balintfy, J. L. 1960. A stochastic model for the analysis and prediction of admissions and discharges in hospitals. *Management Science* **2** 288 – 299.

[3] Ben-Tal, A. and A. Nemirovski. 1998. Robust convex optimization. *Mathematics of Operations Research*, **23** 769– 805.

[4] Bertsimas, D. and M. Sim. 2004. The price of robustness. *Operations Research*, **52**(1) 35–53.

[5] Cardoen, B., E. Demeulemeester., and J. Beliën. 2010. Operating room planning and scheduling: A literature review. *European Journal of Operational Research* **201**(3) 921 – 932.

[6] Chen, W., M. Sim, J. Sun, and C.-P. Teo. 2010. From CVaR to uncertainty set: implications in joint chance constrained optimization. *Operations Research* **58**(2) 470 – 485.

[7] Chen, X., M. Sim, P. Sun, and J. Zhang. 2008. A linear-decision based approximation approach to stochastic programming. *Operations Research* **56** 344 – 357.

[8] Chen, X., M. Sim, and P. Sun 2007. A robust optimization perspective on stochastic programming. *Operations Research* **55** 1058 – 107.

[9] Cochran, J. K., and K. Roche. 2007. A queuing-based decision support methodology to estimate hospital inpatient bed demand. *Journal of the Operational Research Society* advance online publication, September.

[10] Delage E., and Y. Ye. 2010. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research* **58**(3) 596 – 612.

[11] Denton, B., and D. Gupta. 2003. A sequential bounding approach for optimal appointment scheduling. *IIE Transactions* **35** 1003 – 1016.

[12] Denton, B. T., A. J. Miller, H. J. Balasubramanian, and T. R. Huschka. 2010. Optimal allocation of surgery blocks to operating rooms under uncertainty. *Operations Research* **58** 802 – 816.

[13] El Ghaoui, L., F. Oustry, and H. Lebret. 1998. Robust solutions to uncertain semidenite programs. *SIAM Journal of Optimization*, **9** 33 – 52.

[14] Esogbue, A. O., and A. J. Singh. 1976. A stochastic model for an optimal priority bed distribution problem in a hospital ward. *Operations Research* **24** 884 – 898.

[15] Gerhak, Y., D. Gupta, and M. Henig. 1996. Reservation planning for elective surgery under uncertain demand for emergency surgery. *Management Science* **42** 321 – 334.

[16] Ghaoui L. El, M. Oks, and F. Oustry. 2003. Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research* **51** 543 – 556.

[17] Gilboa, I., and D. Schmeidler. 1989. Maximin expected utility theory with non-unique prior. *Journal of Mathematical Economics* **18** 141 – 153.

[18] Goh, J., and M. Sim. 2010a. Distributionally robust optimization and its tractable approximations. *Operations Research* **58** 902 – 917.

[19] Goh, J. and M. Sim. 2010b. Robust optimization made easy with ROME. accepted in *Operations Research.*

[20] Gorunescui, F., S. I. McClean, and P. H. Millard. 2004. Using a queueing model to help plan bed allocation in a department of Geriatric medicine. *Health Care Management Science* **5** 307 – 312.

[21] Guerriero, F., and R. Guido. 2011. Operational research in the management of the operating theatre: a survey. *Health Care Management Science* **14** 89 – 114.

[22] Harrison, G. W., A. Shafer1, and M. Mackay. 2005. Modelling variability in hospital bed occupancy. *Health Care Management Science* **8** 325 – 334.

[23] Harper, P. R., and A. K. Shahani. 2002. Modelling for the planning and management of bed capacities in hospitals. *Journal of the Operational Research Society* **53** 11 – 18.

[24] Isii, K. 1963. On sharpness of Tchebysheff-type inequalities. *Annals of the Institue of Statistical Mathematics* **14** 185 – 197.

[25] Jebali, A., A. B. Hadjalouane, and P. Ladet. 2005. Operating rooms scheduling. *International Journal of Production Economics* **99** 52 – 62.

[26] Kao, E. P. C., and G. G. Tung. 1981. Bed allocation in a public health care delivery system. *Management Science* **27** 507 – 521.

[27] Kim, S-C, I. Horowitzb, K. K. Young, and T. A. Buckleye. 2000. Flexible bed allocation and performance in the intensive care unit. *Journal of Operations Management* **18** 427 – 443.

[28] Kolesar, P. 1970. A Markovian model for hospital admission scheduling. *Management Science* **16** 384–396.

[29] Kokangul, A. 2008. A combination of deterministic and stochastic approaches to optimize bed capacity in a hospital unit source. *Computer Methods and Programs in Biomedicine* **90** 56 – 65.

[30] Lamiri, M., X. Xie, A. Dolgui, and F. Grimaud. 2008. A stochastic model for operating room planning with elective and emergency demand for surgery. *European Journal of Operational Research* **185** 1026 – 1037.

[31] Lian, J., K. Distefano, S. Shields, C. Heinichen, M. Giampietri, and L. Wang. 2010. Clinical appointment process: improvement through schedule defragmentation. *IEEE Engineering in Medicine and Biology Magazine* March/April 127 – 134.

[32] Min, D., and Y. Yih. 2010. Scheduling elective surgery under uncertainty and downstream capacity constraints. *European Journal of Operational Research* **206** 642 – 652.

[33] Murty, K. G., and S. N. Kabadi. 1987. Some NP-Complete Problems in Quadratic and Nonlinear Programming, *Mathematical Programming* **39** 117 – 129.

[34] Proudlove, N. C., S. Black, and A. Fletcher. 2007. OR and the challenge to improve the NHS: modelling for insight and improvement in in-patients flow. *Journal of the Operational Research Society* **58** 145—158.

[35] Robb, W. B., M. J. O'Sullivan, A. E. Brannigan, and D. J. Bouchier-Hayes. 2004. Are elective surgical operations cancelled due to increasing medical admissions. *Irish Journal of Medical Sciences* **173** 129 –132.

[36] Rohleder, T., and K. Klassen. 2002. Rolling horizon appointment scheduling: a simulation study. *Health Care Management Science* **5** 201 – 209.

[37] Teow, K. L., and W. S. Tan. 2008. Allocation of hospital beds in an existing hospital. *Journal of Operations and Logistics* **2** 2–2.

[38] Teow, K.L., Heng, BH, Chong Y, Sim J. 2007. Using Mathematical Programming to Optimize Bed Occupancy by Smoothing Elective against Emergency Admissions. *NHG Annual Scientific Congress.*

[39] Vandenberghe, L., S. Boyd, and K. Comanor. 2007. Generalized Chebyshev bounds via semidefinite programming. *SIAM Review* **49** 52–64.

[40] Vassilacopoulos, G. 1985. A simulation model for bed allocation to hospital inpatient departments. *Simulation* **45** 233 – 241.

[41] Wardrope, J., and P. Driscoll.2003. Turbulent times. *Emergency Medicine Journal* **20** 116 – 116.

[42] Zemach, R. 1970. A model of health service utilization and resource allocation. *Operations Research* **18** 1071 – 1086.