Statistically Feasible Mixed-Integer Programming for Integrated Energy System Scheduling*

Huilan Xu
the Department of Industrial Engineering and Management, College of Engineering
Peking University
Beijing, China
hlxu@stu.pku.edu.cn

E. Liu
the School of Electrical and Mechanical Engineering
Heze University
Heze, China
gaoyi2313@gmail.com

Pengcheng You
the Department of Industrial Engineering and Management, College of Engineering
Peking University
Beijing, China
pcyou@pku.edu.cn

2024 IEEE 8th Conference on Energy Internet and Energy System Integration (EI2) | 979-8-3315-2352-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/EI264398.2024.10991441

***Abstract*—Integrated energy systems coordinate multiple energy flows to enhance energy efficiency and reduce costs. However, the uncertain and volatile availability of emerging renewable energy generation poses significant challenges to traditional scheduling paradigms of integrated energy systems. To address this issue, we propose a robust sample-driven approach to ensure the statistical feasibility of operational constraints – a notion that is mild conservative and requires no prior knowledge of randomness. We formulate the scheduling problem of an integrated energy system as a statistically feasible mixed-integer program with discrete control characteristics of combined heat and power. To handle the non-linearities from chance constraints, we use shape learning and shape calibration to construct an uncertainty set with a statistical feasibility guarantee. The approach finally employs Benders decomposition to handle the resulting deterministic linear model, offering a rigorous solution to scheduling an integrated energy system under vast uncertainty. We validate our analysis through extensive numerical studies.**

***Index Terms*—Integrated energy system, renewable energy generation, statistical feasibility, Benders decomposition**

I. Introduction

Integrated energy systems coordinate multiple kinds of energy between supply and demand [1]. Electricity, heat, and natural gas can be flexibly converted, improving energy efficiency and reducing costs. Renewable energy generation, such as wind turbines and solar photovoltaics, is gaining popularity in integrated energy systems yet also brings along vast uncertainty and intermittency [2]. This significantly challenges traditional scheduling methods for integrated energy systems, especially prediction-based deterministic ones that sufficed to accommodate mild randomness [3].

There is a large literature on managing integrated energy systems under uncertainty. The main methodologies include robust optimization, stochastic optimization, chance-constrained optimization and distributionally robust optimization [4]–[8]. In particular, [4] constructs a polyhedral uncertainty set with predictions and uses a two-stage robust optimization method to improve computational efficiency. [8] further extends the two-stage model with Wasserstein distributionally robust optimization. [5] develops a multi-scenario stochastic programming model to account for uncertainty and generates scenarios of renewable output based on non-parametric kernel density estimation. Similar approaches are adopted in [6] and [7], yet multiple objectives and chance constraints are respectively taken into account. However, most, if not all, of the work along these lines uses either empirical distributions or data samples to handle uncertainty, ignoring the intrinsic discrepancy compared with underlying true distributions. This may lead to catastrophic consequences when uncertainties jeopardize operational constraints.

Our work complements the literature by adopting an emerging notion of statistical feasibility, first proposed in [9], to establish a novel bi-level chance-constrained model for integrated energy systems and propose a rigorous solution to tackle the resulting scheduling problem. More specifically, in modeling diverse components of an integrated energy system, we introduce a bi-level form of chance constraints to account for the uncertainty in operational constraints. Such an approach provides a guarantee to satisfy operational constraints with high probability based only on historical data samples. On this basis, we formulate the scheduling problem of the integrated energy system as a statistically feasible mixed-integer program. To deal with the nonlinear and discrete nature of the problem, we respectively resort to a constructed uncertainty set and Benders decomposition, which jointly lead to an efficient solution. Numerical studies further validate that our approach maintains a good balance between performance and conservativeness: it achieves a low cost close to a chance-constrained benchmark while controlling the violation rate of operational constraints under any specified requirement.

Our major contributions can be summarized as follows:

- A novel statistically feasible mixed-integer programming formulation is proposed for scheduling an integrated energy system. Without prior knowledge of uncertainty, the formulation still allows the fulfillment of operational constraints with high probability using history data samples, and meanwhile avoids being over-conservative.
- To facilitate efficient computation of a solution to the scheduling problem, we propose to construct a sample-driven uncertainty set such that the statistical feasibility requirement is linearly inner-approximated. Then Bender

This work was supported by NSFC through grants 72431001, 72201007, 723B1001, T2121002, and 72131001.

decomposition is employed to iteratively solve the resulting mixed-integer linear program with rigor.

II. Problem Formulation

This section introduces the models of diverse components of an integrated energy system and then formulates a statistically feasible scheduling problem under the uncertainty of renewable energy generation. The framework of the inte-

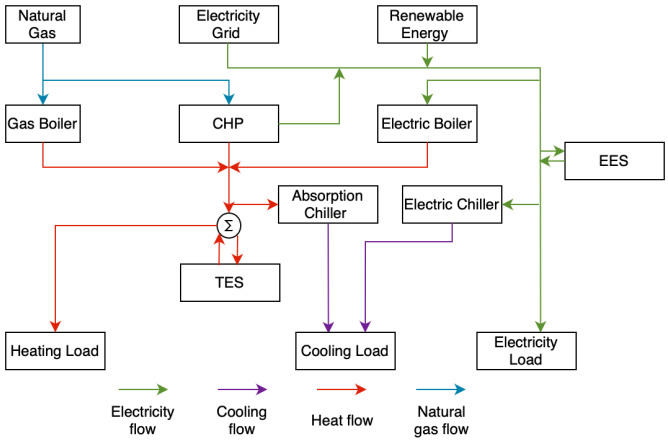

Fig. 1. Energy flow of an integrated energy system.

grated energy system is shown in Fig. 1. Consider a discrete-time horizon $t \in \{0,1,2,...T\}$, where electricity and natural gas from the supply side go through complicated energy conversion to meet the demand of heat, cooling, and power consumption.

1) Natural Gas and Electricity Supply: The integrated energy system purchases natural gas to fuel Combined Heat and Power (CHP) and a gas boiler:

$$V^{\text{grid}}(t) = V^{\text{CHP}}(t) + V^{\text{GB}}(t). \tag{1}$$

Here $V^{\text{grid}}(t)$ is the purchased natural gas. $V^{\text{CHP}}(t)$ and $V^{\text{GB}}(t)$ are the natural gas consumption of the CHP and the gas boiler, respectively.

Electricity comes from four sources. The first is the intermittent renewable energy generation $P^{\text{R}}(t)$ captured by

$$P^{\text{R}}(t) = \hat{P}^{\text{R}}(t) + \xi(t),$$

where $\hat{P}^{\text{R}}(t)$ denotes a given prediction of renewable energy generation while $\xi(t)$ is a random variable representing the corresponding prediction error. The second is the purchased electricity from the grid, denoted by $P^{\text{grid}}(t)$. The third and the last are the electricity outputs from the CHP and electrical energy storage (EES), denoted by $P^{\text{CHP}}(t)$ and $P^{\text{EES}}(t)$, both of which will be modeled shortly. Note that in the case of negative $P^{\text{EES}}(t)$, the EES will consume electricity to charge, thus becoming a load. In the meantime, there are an electric boiler and an electric chiller that also consume $P^{\text{EB}}(t)$ and $P^{\text{EC}}(t)$ amount of electricity, respectively. Then the remaining electricity is all used to meet given inelastic power demand $P^{\text{load}}(t)$.

Sufficient supply of electricity is required to satisfy all the demand. However, the uncertainty of renewable energy generation makes it a challenging task: On the one hand, we want to avoid conservative decisions to over-purchase electricity with unnecessary costs. On the other hand, we have to abide by operational constraints (at least with high probability). We consider a setting where only a dataset $\mathscr{D}$ of history renewable generation is available. To exploit such a dataset, we use the notion of statistical feasibility to construct a bi-level chance constraint for supply meeting demand as follows [9]:

$$\mathbb{P}_{\mathscr{D}}\Big(\mathbb{P}_{\xi(t)}\Big(P^{\text{CHP}}(t) + P^{\text{grid}}(t) + P^{\text{EES}}(t) - P^{\text{EC}}(t) - P^{\text{EB}}(t) + \xi(t) + \hat{P}^{\text{R}}(t) \geq P^{\text{load}}(t)\Big) \geq 1-\delta\Big) \geq 1-\varepsilon, \tag{2}$$

where δ and ε are predetermined small constants. The inner of (2) is a conventional chance constraint, while the outer ensures that the inner chance constraint is satisfied with high probability (w.r.t actual distribution) given the history data (empirical distribution).

2) Heating System: Heat could be provided by the CHP, the gas boiler, the electric boiler and (potentially) thermal energy storage (TES), and is used to supply given inelastic heating load, an absorption chiller and (potentially) TES.

The CHP could convert natural gas into both electricity and heat [10]. The CHP is specified by a tuple of parameters. $P^{\text{CHP, min}}$ and $P^{\text{CHP, max}}$ are the lower and upper bounds of electricity generation, respectively. Let $\eta_{1\,\text{e}}^{\text{CHP}}(kWh/m^3)$ and $\eta_{2\text{e}}^{\text{CHP}}(kWh)$ be the CHP electrical efficiency curve components, and $\eta_{1\,\text{h}}^{\text{CHP}}(kWh/m^3)$ and $\eta_{2\,\text{h}}^{\text{CHP}}(kWh)$ be the CHP thermal efficiency curve components. Given such a specification, the CHP outputs electricity $P^{\text{CHP}}(t)$ and heat $Q^{\text{CHP}}(t)$ according to

$$P^{\text{CHP}}(t) = V^{\text{CHP}}(t)\eta_{1\,\text{e}}^{\text{CHP}} + \eta_{2\text{e}}^{\text{CHP}} I^{\text{CHP}}(t), \tag{3}$$

$$I^{\text{CHP}}(t)P^{\text{CHP,min}} \leq P^{\text{CHP}}(t) \leq I^{\text{CHP}}(t)P^{\text{CHP,max}}, \tag{4}$$

$$Q_h^{\text{CHP}}(t) = V^{\text{CHP}}(t)\eta_{1\,\text{h}}^{\text{CHP}} + \eta_{2\,\text{h}}^{\text{CHP}} I^{\text{CHP}}(t), \tag{5}$$

$$I^{\text{CHP}}(t) \in \{0,1\}, \tag{6}$$

where $I^{\text{CHP}}(t)$ is a binary variable denoting the CHP on/off status.

The gas boiler also converts natural gas into heat. It is specified by its thermal efficiency $\eta_{\text{h}}^{\text{GB}}$, the low-level heat value of natural gas L_{gas} $(9.78kWh/m^3)$, and the lower/upper bounds $V^{\text{GB,min}}/V^{\text{GB,max}}$ of natural gas consumption. The amount of heat converted, denoted as $Q_{\text{h}}^{\text{GB}}(t)$, needs to satisfy

$$V^{\text{GB}}(t) = \frac{Q_{\text{h}}^{\text{GB}}(t)}{\eta_{\text{h}}^{\text{GB}} \times L_{\text{gas}}}, \tag{7}$$

$$V^{\text{GB,min}} \leq V^{\text{GB}}(t) \leq V^{\text{GB,max}}. \tag{8}$$

The electric boiler converts electricity into heat. Its parameters include the thermal efficiency $\eta_{\text{h}}^{\text{EB}}$ and the lower/upper

bounds $P^{\mathrm{GB,min}}/P^{\mathrm{GB,max}}$ of electricity consumption. Then the heat $Q_{\mathrm{h}}^{\mathrm{EB}}(t)$ is generated according to

$$Q_{\mathrm{h}}^{\mathrm{EB}}(t) = \eta_{\mathrm{h}}^{\mathrm{EB}} P^{\mathrm{EB}}(t), \tag{9}$$

$$P^{\mathrm{EB,min}} \le P^{\mathrm{EB}}(t) \le P^{\mathrm{EB,max}}. \tag{10}$$

All the heat generated has to sufficiently satisfy the aggregate demand from the heating load $Q_{\mathrm{h}}^{\mathrm{load}}(t)$ and the absorption chiller $Q_{\mathrm{h}}^{\mathrm{AC}}(t)$, while the TES discharges or charges thermal energy $Q^{\mathrm{TES}}(t)$ (positive for discharge and negative for charge) as needed. This is captured by

$$Q_{\mathrm{h}}^{\mathrm{EB}}(t) + Q_{\mathrm{h}}^{\mathrm{CHP}}(t) - Q_{\mathrm{h}}^{\mathrm{AC}}(t) + Q_{\mathrm{h}}^{\mathrm{GB}}(t) + Q^{\mathrm{TES}}(t) \ge Q_{\mathrm{h}}^{\mathrm{load}}(t). \tag{11}$$

3) Cooling System: The absorption chiller and the electric chiller are both cooling devices but respectively use heat and electricity. The absorption chiller operates according to

$$Q_{\mathrm{c}}^{\mathrm{AC}}(t) = \eta_{\mathrm{c}}^{\mathrm{AC}} Q_{\mathrm{h}}^{\mathrm{AC}}(t), \tag{12}$$

$$Q_{\mathrm{h}}^{\mathrm{AC,min}} \le Q_{\mathrm{h}}^{\mathrm{AC}}(t) \le Q_{\mathrm{h}}^{\mathrm{AC,max}}, \tag{13}$$

where $Q_{\mathrm{c}}^{\mathrm{AC}}(t)$ denotes the output cooling power, $\eta_{\mathrm{c}}^{\mathrm{AC}}$ denotes the cooling efficiency, and $Q_{\mathrm{h}}^{\mathrm{AC,min}}/Q_{\mathrm{h}}^{\mathrm{AC,max}}$ denote the lower/upper bounds of heat consumption.

The electric chiller operates according to

$$Q_{\mathrm{c}}^{\mathrm{EC}}(t) = \eta_{\mathrm{c}}^{\mathrm{EC}} P^{\mathrm{EC}}(t), \tag{14}$$

$$P^{\mathrm{EC,min}} \le P^{\mathrm{EC}}(t) \le P^{\mathrm{EC,max}}, \tag{15}$$

where $Q_{\mathrm{c}}^{\mathrm{EC}}(t)$ denotes the output cooling power, $\eta_{\mathrm{c}}^{\mathrm{EC}}$ denotes the cooling efficiency, and $P_{\mathrm{h}}^{\mathrm{EC,min}}/P_{\mathrm{h}}^{\mathrm{EC,max}}$ denote the lower/upper bounds of electricity consumption.

Finally, the aggregate cooling power needs to meet the cooling load $Q_{\mathrm{c}}^{\mathrm{load}}(t)$:

$$Q_{\mathrm{c}}^{\mathrm{AC}}(t) + Q_{\mathrm{c}}^{\mathrm{EC}}(t) \ge Q_{\mathrm{c}}^{\mathrm{load}}(t). \tag{16}$$

4) Storage System: The EES and the TES are typical storage subject to the following system dynamics and operational constraints:

$$-P^{\mathrm{EES,max}} \le P^{\mathrm{EES}}(t) \le P^{\mathrm{EES,max}}, \tag{17}$$

$$SOC^{\mathrm{EES}}(t) = SOC^{\mathrm{EES}}(t-1) - \eta_e^{\mathrm{EES}} P^{\mathrm{EES}}(t), \tag{18}$$

$$SOC^{\mathrm{EES,min}} \le SOC^{\mathrm{EES}}(t) \le SOC^{\mathrm{EES,max}}, \tag{19}$$

$$-Q^{\mathrm{TES,max}} \le Q^{\mathrm{TES}}(t) \le Q^{\mathrm{TES,max}}, \tag{20}$$

$$SOC^{\mathrm{TES}}(t) = SOC^{\mathrm{TES}}(t-1) - \eta_{\mathrm{h}}^{\mathrm{TES}} Q^{\mathrm{TES}}(t), \tag{21}$$

$$SOC^{\mathrm{TES,min}} \le SOC^{\mathrm{TES}}(t) \le SOC^{\mathrm{TES,max}}. \tag{22}$$

Here $P^{\mathrm{EES,max}}/Q^{\mathrm{TES,max}}$ are the power/heat rating. $\eta_{\mathrm{e}}^{\mathrm{EES}}$ and $\eta_{\mathrm{h}}^{\mathrm{TES}}$ are the dissipation rates. $SOC^{\mathrm{EES,min}}/SOC^{\mathrm{EES,max}}$ and $SOC^{\mathrm{TES,min}}/SOC^{\mathrm{TES,max}}$ are respectively the EES and TES lower/upper bounds on their state-of-charge. $SOC^{\mathrm{EES}}(t)/SOC^{\mathrm{TES}}(t)$ represents the energy/heat stored in the EES/TES. The energy injection and withdrawal limits are given in (17) and (20). The EES and TES system dynamics are described in (18) and (21), respectively, with the states-of-charge bounded by (19) and (22).

Given the characteristics of all the components above, the goal of the integrated energy system is to minimize the total costs of purchasing electricity and natural gas from external sources. Suppose $c_e(t)$ and $c_g(t)$ are respectively the prices of electricity and natural gas power at time t. Renewable energy generation is assumed free of charge. Let $\mathbf{w}$ be a shorthand for all the variables above. Then we formulate the statistically feasible scheduling problem as:

$$\min_{\mathbf{w}} \quad \sum_{t=1}^{T} \Big(c_e(t) P^{\mathrm{grid}}(t) + c_g(t) V^{\mathrm{grid}}(t) \Big) \tag{23}$$

$$\text{s.t.} \quad (1)-(22)$$

This problem is challenging in two ways: 1) the bi-level chance constraint (2) is nonlinear and 2) the CHP involves discrete decision variables.

III. Robust Sample-Driven Approach with Benders Decomposition

In this section, we transfer the bi-level chance constraint into a linear deterministic constraint by constructing an uncertainty set $\mathscr{U}$ using the given dataset. Then we iteratively solve the deterministic mixed-integer linear program with Benders decomposition.

A. Uncertainty Set Construction

To solve the statistically feasible scheduling problem, we first consider an approximation of the inner constraint of (2):

$$P^{\mathrm{CHP}}(t) + P^{\mathrm{grid}}(t) + P^{\mathrm{EES}}(t) - P^{\mathrm{EC}}(t) - P^{\mathrm{EB}}(t) + \xi(t) + \hat{P}^{\mathrm{R}}(t) \ge P^{\mathrm{load}}(t), \forall \xi(t) \in \mathscr{U}, \tag{24}$$

where $\mathscr{U}$ is an uncertainty set. Obviously, $\xi(t) \in \mathscr{U}$ implies:

$$P^{\mathrm{CHP}}(t) + P^{\mathrm{grid}}(t) + P^{\mathrm{EES}}(t) - P^{\mathrm{EC}}(t) - P^{\mathrm{EB}}(t) + \xi(t) + \hat{P}^{\mathrm{R}}(t) \ge P^{\mathrm{load}}(t). \tag{25}$$

We choose $\mathscr{U}$ that covers a $1-\delta$ content of $\xi(t)$, which means that $\mathbb{P}(\xi(t) \in \mathscr{U}) \ge 1-\delta$, then

$$\mathbb{P}\Big(P^{\mathrm{CHP}}(t) + P^{\mathrm{grid}}(t) + P^{\mathrm{EES}}(t) - P^{\mathrm{EC}}(t) - P^{\mathrm{EB}}(t) + \xi(t) + \hat{P}^{\mathrm{R}}(t) \ge P^{\mathrm{load}}(t) \Big) \ge \mathbb{P}(\xi(t) \in \mathscr{U}) \ge 1-\delta. \tag{26}$$

Assume that the given data set $\mathscr{D} = \{\xi(1), ..., \xi(n)\}$, $\xi(t)$ is i.i.d. sampled from the actual probability distribution P. If we can construct $\mathscr{U} = \mathscr{U}(\mathscr{D})$ to be a $1-\delta$ content set for P with a confidence level $1-\varepsilon$,

$$\mathbb{P}_{\mathscr{D}}(\mathbb{P}(\xi(t) \in \mathscr{U}) \ge 1-\delta) \ge 1-\varepsilon, \tag{27}$$

then according to [9, Lemma 2], any solution that satisfies (24) with the particular $\mathscr{U}$ subject to (27) will also be feasible for the inner constraint of (2) with probability $1-\varepsilon$, i.e., (2) is met.

To address the conservativeness, we carefully construct $\mathscr{U}$ based on the following two criteria: 1) the volume of $\mathscr{U}$ should be as small as possible and 2) $\mathbb{P}(\xi(t) \in \mathscr{U})$ is not only greater but close to $1-\delta$ and $\mathbb{P}_{\mathscr{D}}(\mathbb{P}(\xi(t) \in \mathscr{U}) \ge 1-\delta)$ is not only greater but close to $1-\varepsilon$.

We shall construct an uncertainty set $\mathscr{U}$ that is a $1-\delta$ content prediction error set with a confidence $1-\varepsilon$. The smaller

the volume of $\mathscr{U}$, the less conservative the approximation to the chance constraint, but it may not satisfy the constraint (2). We want to construct the smallest uncertainty set $\mathscr{U}$ such that $\mathscr{U}$ covers $1-\delta$ portion of the historical samples with confidence $1-\varepsilon$. This leads us to suggest a two-phase strategy. We first partition $\mathscr{D} = \{\xi(1), ..., \xi(n)\}$ into two parts $\mathscr{D}_1$ and $\mathscr{D}_2$, with size of n_1 and n_2, respectively.

*1) **Shape learning**:* We set the shape $\mathscr{U} = \left\{ \frac{(\xi(t)-\mu)^2}{\sigma} \leq \rho, t \in \{1,2,\cdots,T\} \right\}$. μ and σ are respectively the sample mean and sample variance of $\mathscr{D}_1$.

*2) **Shape calibration**:* ρ is the size of $\mathscr{U}$. We use $\mathscr{D}_2$ to calibrate ρ to satisfy the statistical feasibility and get the smallest size. We set a transformation map $F(\xi(t)) = \frac{(\xi(t)-\mu)^2}{\sigma}$. Then we transform all $\xi_t \in \mathscr{D}_2$ by the transformation map $F(\cdot)$, and sort $F(\xi(t))$ to $\{F(\xi(t))_1 \leq ... \leq F(\xi(t))_{n_2}\}$. We shall find the sample index i^*, which satisfies:

$$i^* = \min \left\{ r : \sum_{k=0}^{r-1} \binom{n_2}{k} (1-\delta)^k (\delta)^{n_2-k} \geq 1-\varepsilon \right\}, \qquad (28)$$

where $1-(1-\delta)^{n_2} \geq 1-\varepsilon$. The value of ρ we seek equals to $F(\xi(t))_{i^*}$. Then we get the shape $\mathscr{U}$. We can demonstrate the statistical guarantee following the same routine for [9, Th. 1 and Lemma 3].

We can furtherly turn (24) into a simpler and deterministic constraint by using the shape parameters of the constructed uncertainty set. Mathematically, $P^{\text{CHP}}(t) + P^{\text{grid}}(t) + P^{\text{EES}}(t) - P^{\text{EC}}(t) - P^{\text{EB}}(t) + \xi(t) + \hat{P}^{\text{R}}(t) \geq P^{\text{load}}(t), \forall \xi(t) \in \mathscr{U}$, with $\mathscr{U}$ characterized by $\{\frac{(\xi(t)-\mu)^2}{\sigma} \leq \rho, t \in \{1,2,\cdots,T\}\}$, can be transformed into the equivalent linear constraint:

$$P^{\text{load}}(t) - (P^{\text{CHP}}(t) + P^{\text{grid}}(t) + P^{\text{EES}}(t) - P^{\text{EC}}(t) - P^{\text{EB}}(t) + \hat{P}^{\text{R}}(t)) \leq \mu - \sqrt{\rho\sigma}. \qquad (29)$$

Then the statistically feasible scheduling problem (23) is transformed into a deterministic mixed-integer linear program:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \sum_{t=1}^{T} \left(c_e(t) P^{\text{grid}}(t) + c_g(t) V^{\text{grid}}(t) \right) \qquad (\mathbf{P}) \\ \text{s.t.} \quad & (1), (3)-(22), (29) \end{aligned}$$

B. Benders Decomposition

Benders decomposition is an efficient method to solve **P**. We decompose **P** into a master problem and a slave problem. The master problem involves only integer variables and serves an approximation for the original problem that always provides a lower-bound objective value. Meanwhile, the slave problem involves only continuous variables and serves as an inner problem of the original problem with fixed integer variables.

Mathematically, **P** takes the following form:

$$\begin{aligned} \min_{x,y} \quad & Z(x,y) = f^T y + c^T x && (30a) \\ \text{s.t.} \quad & Ax \leq b && (30b) \\ & Dx + Ey \leq d && (30c) \\ & y \in \mathbb{S} && (30d) \end{aligned}$$

Here $x \in \mathbb{R}^p$ stands for the continuous variables in **P**, while $y \in \mathbb{Z}^q$ is $I^{CHP}(t)$ for $t \in \{0,1,2,...T\}$ in **P**. y is constrained in a feasible set $\mathbb{S} \subset \{0,1\}^q$. $c \in \mathbb{R}^p$ and $f \in \mathbb{R}^q$ respectively denote the cost vectors for x and y, where f is a zero vector. The matrix $A \in \mathbb{R}^{u \times p}$ and the vector $b \in \mathbb{R}^u$ specify the affine inequality constraints on x. Similarly, $D \in \mathbb{R}^{v \times p}$, $E \in \mathbb{R}^{v \times q}$ and $d \in \mathbb{R}^v$ specify the affine inequality constraints that couple x and y.

For notational convenience, we define

$$\Omega := \{(x,y) : (30b), (30c), (30d)\}$$

to be the set of feasible region of the MILP (30), which is assumed to be non-empty. Then we denote its optimal cost as:

$$Z^* = \min_{(x,y) \in \Omega} Z(x,y) \ ,$$

which is assumed to be finite and achieved at an optimal (not necessarily unique) solution.

The master problem is designed to take the following form:

Master Problem

$$\begin{aligned} \min_{y, \hat{Z}} \quad & \hat{Z} && (31a) \\ \text{s.t.} \quad & \text{Feasible Cuts} && (31b) \\ & \text{Infeasible Cuts} && (31c) \\ & y \in \mathbb{S} && (31d) \end{aligned}$$

where $\hat{Z} \in \mathbb{R}$ is a scalar variable that serves as a proxy for the original cost function $Z(x,y)$. It is connected with y via the feasible cuts, which are a series of lower-bound constraints for $\hat{Z}$. The infeasible cuts are constraints on y to exclude those choices of values that will render the slave problem infeasible, as we will see shortly.

The slave problem takes care of all the components related to the continuous variables in the MILP (30), treating the integer variable y as given, i.e., $y = y(k)$ based on the latest output of the master problem. In particular, the slave problem at the k-th iteration takes the following form:

Slave Problem

$$\begin{aligned} \min_{x} \quad & c^T x && (32a) \\ \text{s.t.} \quad & Ax \leq b && (32b) \\ & Dx \leq d - Ey(k) && (32c) \end{aligned}$$

where the constant $f^T y(k)$ in the objective function and the feasible set $\mathbb{S}$ of y are irrelevant and removed, given the fixed value of $y(k)$.

We consider the first possible case for the optimal solution to the slave problem(32). The dual slave problem is given as:

$$\begin{aligned} \max_{\lambda, \mu} \quad & -b^T \lambda + (Ey(k) - d)^T \mu && (33a) \\ \text{s.t.} \quad & c + A^T \lambda + D^T \mu = 0 && (33b) \\ & \lambda \geq 0, \ \mu \geq 0 && (33c) \end{aligned}$$

where $\lambda \in \mathbb{R}^u$ and $\mu \in \mathbb{R}^v$ are dual variables for (32b) and (32c), respectively.

The optimal value of the dual slave problem (33) achieved at $(\lambda(k),\mu(k))$ coincides with the optimal cost of the slave problem (32) and provides an upper bound of Z^*, i.e.,

$$\bar{Z}(k) := f^T y(k) - b^T \lambda(k) + (Ey(k) - d)^T \mu(k) \geq Z^* \ . \quad (34)$$

Each time we obtain $(\lambda(k),\mu(k))$ from the dual slave problem (33), we add a feasible cut as follows in the master problem (31):

Feasible Cut

$$\hat{Z} \geq f^T y - b^T \lambda(k) + (Ey - d)^T \mu(k) \ . \quad (35)$$

Now we consider the second possible case for the optimal solution to the slave problem (32), i.e., (32c) is infeasible with the current choice of $y(k)$. To this end, an infeasible cut is developed:

Infeasible Cut

$$0 \geq -b^T \lambda(k) + (Ey - d)^T \mu(k) \ . \quad (36)$$

Feasible cut and infeasible cut provide feedback for the master problem by adding cutting-plane constraints, or simply "cuts", that drive the master problem solution closer to an optimal solution to the mixed-integer linear program. Benders decomposition keeps solving the two problems in an alternating fashion until it converges to optimum. More details about the whole solution are provided in the technical report [11].

IV. Numerical Studies

This section presents numerical studies designed to validate the efficiency of our approach. Electricity and natural gas prices come from [12]. The loads of electricity, heating, and cooling, as well as the actual and predicted amount of renewable energy generation are all sourced from [13]. The parameters of the integrated energy system are adapted from [4].

We compare our robust sample-driven approach **P** with four benchmarks:

– Offline Optimum (OO) in hindsight: an offline version of the scheduling problem (23) with actual prediction error realizations. This is set as the optimal benchmark.

– Prediction-based Optimization (PO): a reformulation of the scheduling problem (23) using expected prediction errors.

– Chance-constrained Optimization (CC): a reformulation of the scheduling problem (23) using the classic chance-constraint that assumes Gaussian distributed prediction errors.

– Robust Optimization (RO): a reformulation of the scheduling problem (23) using the classic RO method based on history prediction data.

We will use the total purchase costs and the violation rate of the constraint (2) (averaged over 1000 experiment runs) as the main metrics.

Fig. 2 illustrates how the total costs of different methods change with varying sample sizes when $\delta = 0.05$ and $\varepsilon = 0.05$. The optimal cost generated from OO 30507(\$). The performance trends of the three methods are different. The total cost of **P** is lower than that of RO but higher than and closer to the total cost of CC as the sample size increases. The total cost of CC does not change much with the increasing sample size, compared to the total cost of RO and **P**. Fig. 3 illustrates how the violation rates of different methods change with the increasing sample size. The violation rate of CC is much greater than 0.05. RO maintains a rate of 0 but incurs a higher cost than **P** and CC. **P** exploits the flexibility of the bi-level chance constraint, with a violation rate below 0.05 and decreasing costs as the sample size grows.

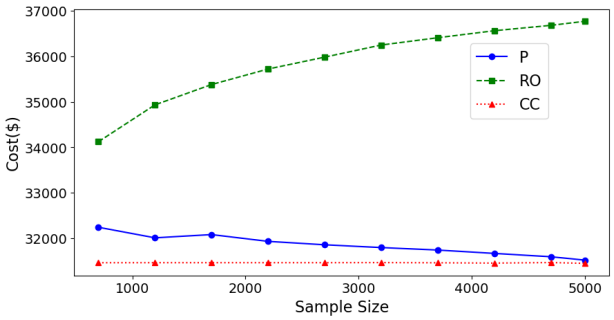

Fig. 2. Impact of sample size on costs.

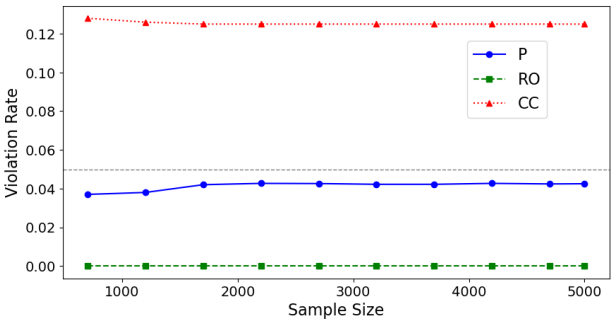

Fig. 3. Impact of sample size on constraint violation rate.

Table I shows the performance of **P** and benchmarks when $n_1 = 3640$ and $n_2 = 1560$. The relative cost increment is the ratio of the difference between the actual cost and the optimal cost to the optimal cost. δ' means the ratio of the number of violated constraints to the length of T. PO has a negative relative cost increment, indicating a lower cost compared to the optimal cost, but the violation rate significantly exceeds 0.05. **P** achieves a low cost close to CC and δ' is below and closest to the desired 0.05.

According to Fig. 4, the grid is the primary source of electricity, supplemented by renewable energy, the EES, and the CHP. The EES discharges when electricity prices are high and charges when prices are low. Fig. 5 shows that most natural gas is used for electricity generation by the CHP because converting natural gas to electricity through the CHP is more cost effective than purchasing the equivalent amount of electricity. Fig. 6 shows that the gas boiler is the primary source of heat because it is more efficient in converting natural gas to heat compared to the CHP. In Fig. 7, the cooling supply mainly comes from the absorption chiller. The system tends to convert heat rather than electricity to cooling due to the low natural gas price.

TABLE I
COMPARISON ACROSS DIFFERENT METHODS

Method	Relative Cost Increment	Constraint Violation Rate
PO	−0.721%	10%
P	3.334%	4.25%
CC	3.101%	12.5%
RO	20.253%	0

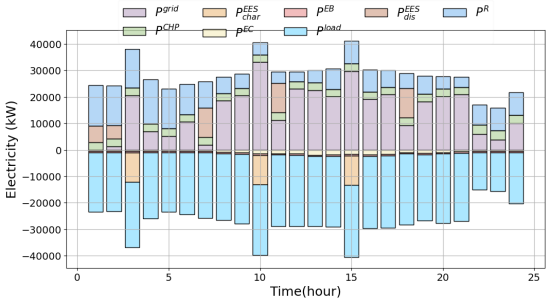

Fig. 4. Electricity flow.

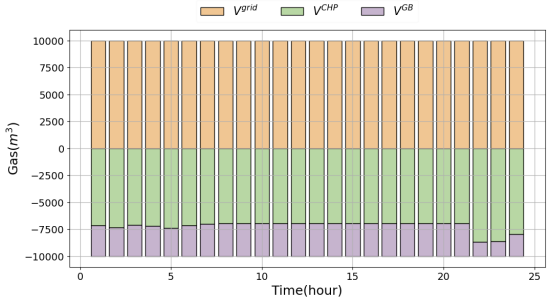

Fig. 5. Natural gas flow.

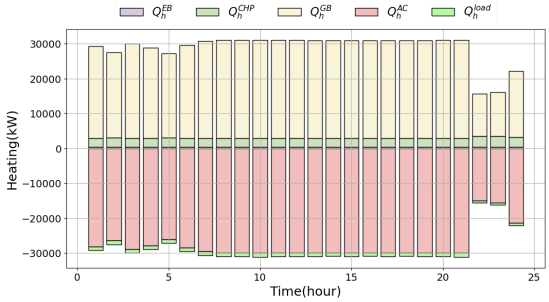

Fig. 6. Heat flow.

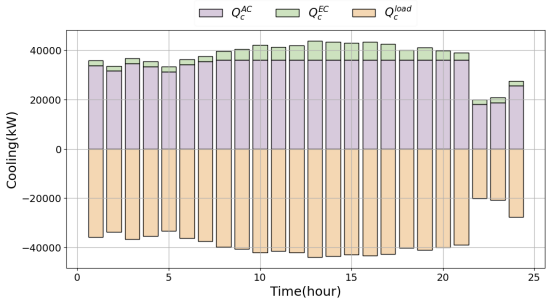

Fig. 7. Cooling flow.

V. CONCLUSION

In this paper, we introduce a novel statistically feasible mixed-integer programming formulation for scheduling an integrated energy system. To address the problem's nonlinear and discrete characteristics, we utilize a constructed uncertainty set and Benders decomposition. Our approach effectively manages the uncertainty of renewable energy generation using historical data, ensuring high probability adherence to the constraints while avoiding over-conservatism. Extensive numerical studies validate that the approach achieves a low cost and a violation rate close to the desired 0.05, compared to other benchmarks. Future work could explore the extension of this method to more uncertainties, further enhancing its applicability.

REFERENCES

[1] E. A. M. Ceseña and P. Mancarella, "Energy systems integration in smart districts: Robust optimisation of multi-energy flows in integrated electricity, heat and gas networks," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 1122–1131, 2018.

[2] C. Zhang, Y. Xu, Z. Li, and Z. Y. Dong, "Robustly coordinated operation of a multi-energy microgrid with flexible electric and thermal loads," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2765–2775, 2018.

[3] C. Wang, Y. Wang, Z. Ding, T. Zheng, J. Hu, and K. Zhang, "A transformer-based method of multienergy load forecasting in integrated energy system," *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 2703–2714, 2022.

[4] M. Wu, J. Xu, L. Zeng, C. Li, Y. Liu, Y. Yi, M. Wen, and Z. Jiang, "Two-stage robust optimization model for park integrated energy system based on dynamic programming," *Applied Energy*, vol. 308, 118249, 2022.

[5] S. Lin, C. Liu, Y. Shen, F. Li, D. Li, and Y. Fu, "Stochastic planning of integrated energy system via Frank-Copula function and scenario reduction," *IEEE Transactions on Smart Grid*, vol. 13, no. 1, pp. 202–212, 2021.

[6] Y. Cao, Y. Mu, H. Jia, X. Yu, K. Hou, and H. Wang, "A multi-objective stochastic optimization approach for planning a multi-energy microgrid considering unscheduled islanded operation," *IEEE Transactions on Sustainable Energy*, vol. 15, no. 2, pp. 1300–1314, 2023.

[7] H. Zhang, J. Wang, X. Zhao, J. Yang, and Z. A. B. sinnah, "Modeling a hydrogen-based sustainable multi-carrier energy system using a multi-objective optimization considering embedded joint chance constraints," *Energy*, vol. 278, 127643, 2023.

[8] J. Li, Z. Xu, H. Liu, C. Wang, L. Wang, and C. Gu, "A Wasserstein distributionally robust planning model for renewable sources and energy storage systems under multiple uncertainties," *IEEE Transactions on Sustainable Energy*, vol. 14, no. 3, pp. 1346–1356, 2022.

[9] L. J. Hong, Z. Huang, and H. Lam, "Learning-based robust optimization: Procedures and statistical guarantees," *Management Science*, vol. 67, no. 6, pp. 3447–3467, 2021.

[10] E. A. M. Ceseña, T. Capuder, and P. Mancarella, "Flexible distributed multienergy generation system expansion planning under uncertainty," *IEEE Transactions on Smart grid*, vol. 7, no. 1, pp. 348–357, 2015.

[11] H. Xu, E. Liu, and P. You, "Statistically feasible mixed-integer programming for integrated energy system scheduling (technical report)." https://pengcheng-you.github.io/desires-lab/papers/XLY_TR.pdf, 2024.

[12] X. Jin, Y. Mu, H. Jia, J. Wu, X. Xu, and X. Yu, "Optimal day-ahead scheduling of integrated urban energy systems," *Applied Energy*, vol. 180, pp. 1–13, 2016.

[13] ASU, "Campus metabolism." https://cm.asu.edu, 2024. Accessed: 2024-08-04.