THE ARCHITECTURE OF GREEN ENERGY SYSTEMS*

MICHAEL C. FERRIS† AND ANDY PHILPOTT ‡

Abstract. Energy production throughout the world is transitioning from fossil fuels to renewable sources such as wind power and solar power. This transition has been gradual - over half of the world's electricity is still produced by coal, oil and gas - but must accelerate to meet global emission targets. This paper examines the contributions that mathematical modelling can make to help accelerate this transition. The models we catalog are confined to optimization and equilibrium models, but cover a range of physical scales and time scales. Our focus is on novel model formulations that can help overcome the challenges of the transition by unpicking the complexity inherent in many settings and quantifying the tradeoffs that must be made when developing energy policy.

Key words. green energy transition, renewable electricity, carbon prices

MSC codes. 49-02, 65K10, 90C90, 91B74

1. Introduction.

The world is undergoing a transition from using fossil-fuel energy that emits greenhouse gases (mainly carbon dioxide) to using energy that does not. This transition is a global response to calls to limit global warming that has been caused by the emission of greenhouse gases over the post-industrial era. The current scale and speed of this transition appears insufficient to keep global temperatures below agreed targets. There are many technical, economic, social and political reasons for this slowness that have been canvased in a number of recent reports (see e.g., [1, 2, 10]).

Our purpose in this paper is to examine the contribution that mathematics and mathematical models can make to achieving the goals of the transition and to understanding and overcoming the barriers that are faced in the transition. Those barriers include affordability, sustainability, reliability, industrial competitiveness, and trusted information. The contribution of the paper is primarily to present mathematics; it is not intended to be a survey of existing energy models, of which there are many (see, e.g., [67, 27]). Care is needed since differences between precise mathematical definitions and socially accepted descriptions can lead to confusion and discord. For example, reliability is a complex notion and can include aspects of robustness, resource adequacy and resilience.

In particular we will focus on what we call the *architecture* of energy systems, which consists not only of the physical infrastructure for generating and transporting energy, but also the market and contractual arrangements that give incentives for investing in this infrastructure and that allow for it to be operated in an efficient manner. Our aim is not so much to deliver the correct answer or define an optimal solution, but rather to pose questions that can benefit from a mathematical modelling approach. Many of our approaches incorporate techniques to promote flexibility [20], including multiple types of dispatchable generation, demand response, energy storage and enhanced connectivity. While there are many different mathematical techniques

---

*Submitted to the editors DATE.

Funding: This work was funded by MBIE, and completed at the Institute for Mathematical and Statistical Innovation (IMSI), which is supported by the National Science Foundation (Grant No. DMS-1929348).

†Computer Sciences Department and Wisconsin Institute for Discovery, University of Wisconsin, Madison, WI 53706 (ferris@cs.wisc.edu)

‡Electric Power Optimization Centre, University of Auckland, New Zealand (a.philpott@auckland.ac.nz).

that could impact the transition, we will focus primarily on optimization, including aspects of game theory and links to machine learning (forecasting and prediction) and simulation approaches.

We are interested in the architecture of systems that generate mainly *green* energy, a catch-all term that encompasses renewable energy from sources that are constantly and naturally renewed such as hydroelectric power, wind power and solar power, and energy from other sources with negligible carbon emissions (such as nuclear and geothermal electricity), or net-zero emissions (such as biofuels). Such systems will be an essential part of the transition, along with new technologies that fill gaps in our operating landscape.

Our use of the adjective green in this context might be viewed by some as contentious, as some activities associated with green energy production (such as building hydroelectric dams or mining lithium) can damage the natural environment. As we discuss later in the paper, some of this damage might be justifiable when traded off against the damage avoided by reducing carbon emissions, so it would be unwise to preclude such activities from the mix of green energy we study.

Designing the green energy system of the future is a global problem involving interactions between countries across the world and requiring long term investments, changes of operational procedures, trade-offs and innovations. While internationally coordinated efforts are likely to be the most effective and economical, this is hampered by political discord, disparate goals and perspectives on the severity of the issue, and different ideas on the best course of action to transition into a green energy system. Even within countries, different agents view the risk of inaction, or incorrect actions, in contrasting ways, and will make decisions in their own interests in response to incentives and regulations. At both of these scales (national and individual), desired outcomes of fairness and equity require careful analysis and can be informed by mathematical modelling using metrics that capture the goals.

The challenge then lies mainly in designing appropriate incentives and regulations, so agents with different attitudes to risk align their actions with the objective of global emissions reduction. Our approach in this paper is to look at tools that capture the risk in each agents problems, suggest models and approaches to invest in a portfolio of technologies that may reduce the variability in outcomes and enhance the ability to finance their adoption, whilst quantifying the differences between these agent-driven results and one that might arise with a system-wide perspective.

A green energy system can be viewed along three orthogonal dimensions. We show two of these in Figure 1.

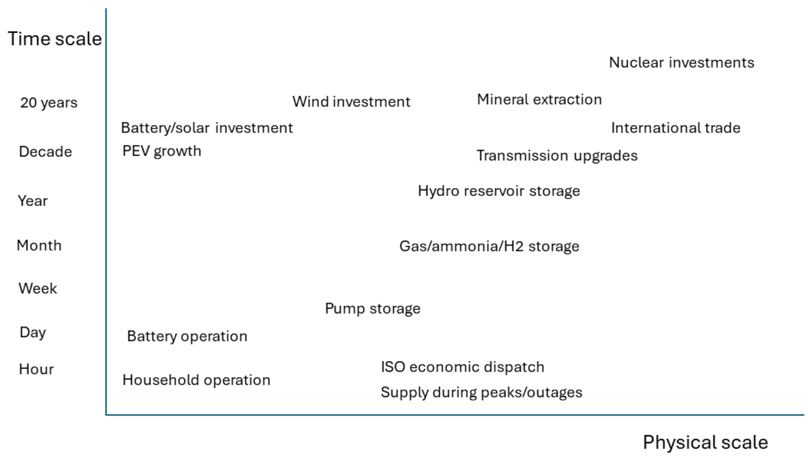

Fig. 1. *The energy transition in two dimensions*

In one dimension one can vary the physical scale of the system. At the smallest scale, one might consider a household with solar panels, a battery and a plug-in electric vehicle. This could form part of a micro grid, which in turn connects to a larger system with industrial electricity supply and demand. High voltage electricity transmission lines link these together into an electricity network, that may be connected to neighbouring networks by transfers along tie lines. The system might transport energy from place to place using other carriers such as hydrogen, ammonia, natural gas, oil, coal or uranium. Transfers of energy are accompanied by financial flows, and derivative instruments that derive their value from these transactions. At the largest (global) scale the energy and financial flows are between different regions and economies; the design and operation of new forms of contracts and financial flows are critical to enabling the transition process.

The overall system is a collection of technologies at different physical scales, connected through a network that might be electrical or some other energy transport. To answer questions about the architecture of this system, or the design and operation of a component, one can consider a particular scale, in which case the interplay with larger (or smaller) scales needs to determine boundary interactions. Such boundary interactions may be physical, financial, regulatory or involve some form of incentives. Interactions occur among sectors of the economy including transport, buildings, industry and agriculture, and reuse and disposal of raw materials necessary for the technologies must also be considered. Furthermore, specific mathematical models may be needed to capture the details of the technologies with some formulations being a more accurate representation of reality, while others are approximations with properties amenable to deeper analysis.

The second important dimension to consider is time, and implicitly the evolution of uncertainties over varying time scales. Energy is produced and consumed continuously, but questions about the architecture of energy systems are posed with different temporal resolution. Also, information flows are often uncertain, and are resolved at a variety of time scales. Predicting new technologies or policy changes, or the increase in electrical demand due to transitions in domestic heating or transport, or the installation and closing of different generation plants can involve complex models and forecasts and these can evolve over time within a physical or computational learning process. Learning can lead to exponentially decreasing costs in new technologies, with rates that are hard to predict a priori. Dealing with uncertainty in forecasts requires models of some sophistication. In the short term, the intermittency of solar and wind power requires backup supply in the form of fast-start generation, load reduction or batteries, so that supply is reliable. On a longer time scale, energy might need to be stored (e.g., in a hydro reservoir) for use in future months when the supply of other sources of energy are lower. The aforementioned issues relate to parametric uncertainties - things we know the form of but are unclear about their actual levels. In contrast, model (or structural) uncertainty arises in problems that involve long-lived capacity choices and need to account for many possible states of the world (e.g., emission constraints, technology changes, political environment) in future decades.

The third important dimension represents social and political or behavioral aspects. These can involve interplay with other (political) institutions, agencies (countries or adversaries) or policies and information. While we discuss models of behavior related to (mathematical) game theory, this paper does not address social/political factors or their evolution. Nonetheless, it is understood that interactions of these types can affect the efficiency of designed systems and how local or national behavior influences the outcomes of a given architecture.

The paper examines a number of policy questions arising in the green energy transition that can be viewed in the above three dimensions. Despite the enthusiasm of advocates for silver bullet solutions to the green-energy transition, the policy questions that arise are complex and do not admit simple intuitive solutions. Our interest in this paper is in formulating these questions in mathematical terms with a view to representing the complexity of the tradeoffs involved. Problem formats that model interactions, and determine what regimes are active at any given time are important in understanding overall structure of solutions, even if specific details are abstracted or approximated.

Our mathematical framework draws on two core methodologies: optimization and game theory. Optimization is a powerful tool for exploring the tradeoffs that are inevitable when comparing competing technologies. For example, it is tempting to remove all fossil-fuelled electricity capacity from a region to make its electricity 100% renewable, but this might be very expensive compared with a system with 1% of fossil-fuelled generation capacity that is used sparingly (see, e.g. [33]). System optimization models make these tradeoffs explicit, and enable decision makers to arrive at optimal combinations of technologies that will meet desired emission goals at least cost. For models involving time and uncertainty, the optimization models become more complicated, and must deal with estimates of probability distributions and attitudes to and representations of risk.

The second methodology guiding our approach is game theory. The transition to green energy emerging in most countries is driven by competing commercial agents, responding to incentives and regulations set by governments. In its simplest form, this setup is known by economists as a *principal-agent* problem [39], in which a leader takes some action and a number of followers respond by optimizing their own objectives in a competitive environment. There are many different versions of this simple game model that arise from varying assumptions on the degree of strategic behavior of agents and the knowledge that each agent has at their disposal. The models can capture features such as barriers to entry, collaboration or contrasting risk attitudes.

In summary, the mathematical study of the architecture of green energy systems involves suites of models encompassing different resolutions in each dimension. The models can be optimized to determine some *social plan* of action that maximizes overall welfare subject to constraints, e.g., on emissions. This gives a gold-standard benchmark for more realistic policies that will attempt to achieve results through incentives (e.g., carbon taxes) and regulations (e.g., renewable energy standards). The extent to which the outcomes of these policies fall short of the gold-standard benchmark can be evaluated by game-theory models.

The paper is laid out as follows. In the next section we classify in mathematical terms the types of optimization and equilibrium models that will be applied to the various settings we study. Section 3 then describes a collection of example problems that can be studied using a selection of models cataloged in Section 2. Section 4 is devoted to a discussion of risk, and how one might devise models that represent the partial equilibrium that emerges when agents have contrasting risk measures. We then make some concluding remarks in Section 5.

2. Mathematical Models.

Models are used in different ways in the sciences. They can be explanatory and promote understanding of unintended consequences through the revelation of counterintuitive linkages between decisions and outcomes. Our colleague, Tom Rutherford, remarked that it is always important to "show me a model in which this assertion makes sense". But models can also be informative in a quantitative sense addressing questions such as: what is the size of a particular effect, or what will this policy or action cost? More recently, the notion of a data driven model (coming from the AI or Machine Learning communities) or the notion of a digital twin (that incorporates a model of some physical process) has become more prominent as AI has risen to the fore. This can be augmented by simulation modelling, defined as the process of creating and analyzing a digital prototype of a physical system to predict its performance in the real world, often as a parametric function of uncertain inputs. Approaches include the Monte Carlo method, agent-based modelling, discrete event simulation, and system dynamics modelling.

While there are many mathematical models that could influence the choice of architecture, we will confine ourselves in this paper to discussing approaches that are based in the field of optimization, and specifically to approaches that utilize constraints to model the underlying physical nature of the problems at hand. It is understood that any such model needs to be populated with data that instantiates these mathematical relationships. Different data will be relevant for models at disparate scales, but we will not cover the acquisition details of this. Nevertheless, we will consider the uncertain nature of these data and suggest models that account for this uncertainty using stochastic optimization approaches. In this section, we briefly outline the main formats that we will use in the sequel.

2.1. Optimization models.

Our models will consider decision variables x that live in a finite dimensional space R^n. These variables are constrained to lie in a subset X of R^n and are used to define an objective function f that maps R^n to the real line, and a vector valued function g that is constrained to lie in some convex cone K, resulting in the optimization problem

$$\min_{x \in X} f(x) \text{ s.t. } g(x) \in K. \tag{2.1}$$

Special cases of the data of this problem lead to formats under consideration, namely:

1. if f is a linear function, and g is an affine function, X is polyhedral (possibly R^n) and $K = \{0\}^p \times \mathrm{R}^m_-$, then (2.1) is a linear program (LP)
2. if in addition $X \subset \mathrm{Z}^{n_1} \times \mathrm{R}^{n_2}$ (i.e. some of the variables can only take on discrete values), then (2.1) is a mixed integer program (MIP)
3. K can model both equations and inequalities or a mixture of both (as shown item 1 above)
4. if f and g are convex functions of x then (2.1) is a convex optimization problem
5. if ξ is a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $f(x) = c(x) + \mathbb{E}_{\mathbb{P}}[Q(x, \xi)]$ where $Q(x, \xi)$ is the optimal value of the second-stage problem
$$\min_y q(y, \xi) \text{ s.t. } T(\xi)x + Wy = h(\xi)$$
then (2.1) is a two-stage stochastic programming problem
6. if in addition $X = \{x \geq 0 : Ax = b\}$, $c(x) = c^T x$ and $q(y, \xi) = q(\xi)^T y$, then the problem is a two-stage stochastic linear programming problem.

The formulation of the above two-stage problem assumes that the second-stage data ξ is modelled as a random vector with a known probability distribution. In many applications the expectation $\mathbb{E}$ can be replaced by a more general risk measure ρ.

It is important to note that differences in terminology abound in the literature that deals with uncertainty in future events. In colloquial use the word "scenario" denotes a possible narrative of future events and decisions, so a scenario encapsulates both policy decisions and the outcomes that result from them. This terminology is used widely in the literature studying future decarbonization pathways. On the other hand, in the stochastic programming community, scenarios denote realized sequences of random events that occur independently of any actions or policies. The scenario approach is then a technique for obtaining solutions to robust optimization and chance-constrained optimization problems based on a sample of an underlying random variable. Robust approaches often deal with worst-case outcomes, whereas probabilistic approaches deal with distributions. Acronyms such as LTA (learn-then-act or wait-and-see) compared with ATL (act-then-learn or here-and-now) can lead to misunderstandings of the underlying modelling approach.

Other data modelling issues are important to consider in stochastic programming. Often the notion of representative days are used for quantification of issues related to wind generation. Various papers explore the pros and cons of such approaches, see e.g. [62]. Also the spatial correlations between random variables modelling renewable generation can have distinct impact on the conclusions of a model that is informed by such data, see e.g. [75, 77]. Another correlation issue is the timing of failures after storms. The distribution seems to have delay or relates to the occurrence of multiple events.

It is possible to interpret policies specified by decision makers in different mathematical ways, some of which do not exactly capture the intended notion. Is a stated goal of zero carbon by 2050 to be interpreted as an almost sure constraint, or one that is to be satisfied with high probability [33]? Is the constraint "only once in 5 years" to be interpreted as a chance constraint, or the assertion that if the event occurs then in cannot occur again in the next four years. Models exploring these concepts are further discussed in [31].

The two-stage stochastic programming problem can be extended to a multistage stochastic programming problem, in which decisions are made in many stages $t =$

$1, 2, \ldots, T$ and the random variables define a stochastic process $\xi_t, t = 1, 2, \ldots, T$. After each stage t the values of ξ_t are realized, and adaptive decisions made in the light of this information. Such problems are useful in studying investment problems over long time horizons when new information might require existing capacity to be retired or replaced.

A useful special case of multistage stochastic programming is the discrete-time stochastic optimal control problem. Here the random variables ξ_t at each stage t are assumed to be independent of those at previous and later stages, and the decision variables divide into states x and controls u. This gives constraints:

$$x_{t+1} = g_t(x_t, u_t, \xi_t), \quad u_t \in \mathcal{U}_t, \quad t = 1, 2, \ldots, T-1$$

and objective

$$f(x) = \mathbb{E}[\sum_{t=1}^{T} f_t(x_t, u_t, \xi_t)].$$

In this case the problem has a finite horizon; infinite-horizon versions replace the sum in the objective with a discounted infinite series. Stochastic optimal control problems are amenable to solution by (approximate) dynamic programming [12, 72].

It is important here to be specific about the nature of the uncertainty in the above models. In most stochastic optimization problems, the random variables are assumed to have known distributions that can be estimated from a sample of historical data. A popular approach is to solve a sample average approximation problem using the finite empirical distribution [79]. Convergence of this approach with increasing sample size relies on laws of large numbers and the central limit theorem, which may not hold for heavy-tailed distributions. For stochastic optimization problems involving planning decisions made many years in the future, probabilities (e.g., of a new technology emerging) are impossible to estimate from historical data, and some expert assessment must be made and tested. As identified by Mercure et al [59], risks and opportunities in these settings are more important to identify than net present values based on discounted expected cash flow. A real-options [18] approach has some appeal here though this is difficult to apply in system settings where there are many competing and complementary investment options, and limited hedging instruments. Much of the work in the stochastic programming literature deals with problems with exogenous uncertainty (where the optimization decisions cannot influence the stochastic process). Decision dependent uncertainty problems [64] have endogenous uncertainty, and can model, for example, learning in systems.

Risk-averse stochastic programming problems formulated in scenario trees provide another alternative framework that models upside optionality as well as downside risk. Binary variables in these models can represent timing decisions, e.g. when to build or shut down generating plants, albeit with an increase in computational complexity. It is important to recognize that these models are *look-ahead* optimization models [71], with the goal of specifying a well-hedged first-stage decision. The intention after the first stage decision is implemented, is to re-solve a new model in a rolling-horizon fashion with updated estimates of parameters. How far to look ahead, how to appropriately approximate the future, and how to implement the solutions in practice are all interesting research questions, with answers that can generally only be settled by numerical experiments with context-specific models.

Finally in some settings one might seek a solution that performs well over a set of varying problem data. *Robust optimization* provides a numerically efficient way of doing this by specifying a convex uncertainty set $\mathcal{U}$ that defines the data variations (see [13]). For example, when the constraint data are uncertain we obtain:

$$\min f(x) \text{ s.t. } x \in X(u), \quad u \in \mathcal{U}. \tag{2.2}$$

This notion can be extended to compute a *distributionally robust* solution to a stochastic optimization problem that performs well for every probability distribution lying in a set $\mathcal{P}$. An example formulation would be as follows.

$$\min_{x \in X} \max_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}[f(x, \xi)]. \tag{2.3}$$

2.2. Complementarity models. A complementarity problem is a generalization of the optimality conditions of (2.1). In this setting we seek a variable x such that

$$x \in X, F(x) \in X^*, x^T F(x) = 0$$

where $F : \mathrm{R}^n \mapsto \mathrm{R}^n$, X is now a cone (in many settings the positive orthant in R^n) and X^* is the dual cone $X^* := \{w : z^T w \geq 0, \forall z \in X\}$. The third constraint indicates that x and $F(x)$ form a complementary pair and is often written as $x \perp F(x)$. The complementary slackness conditions of linear programming are a special case of a complementarity problem. While there are many examples of the use of complementarity formulations in engineering and economics (see [32, 35]), one particular modelling use allows the formulation to automatically switch between regimes of operation. For example, in [21] complementarity constraints are used to model automatic tap-changing transformers and other switched electrical devices. Given the following constraints,

$$\begin{aligned}
&v = \bar{v} + v^+ - v^-, \\
&0 \leq (q - q^{min}) \perp v^+ \geq 0, \\
&0 \leq (q^{max} - q) \perp v^- \geq 0,
\end{aligned}$$

it is easy to see that when q is strictly between q^{min} and q^{max} then v is at set point $\bar{v}$, whereas if q is at one of its bounds, then v is allowed to move away from the set point value.

A generalization of the complementarity problem is a variational inequality, where

$$x \in X \text{ and } F(x)^T (z - x) \geq 0, \text{ for all } z \in X.$$

This is sometimes termed a generalized equation, since in the special case of $X = \mathrm{R}^n$ it simplifies to the solution of a square nonlinear system $F(x) = 0$. It is also clear that when X is a cone, this is identical to the (cone) complementarity problem. When X is a convex set (not necessarily a cone), then the optimality conditions of

$$\min_{x \in X} f(x)$$

are in the form of a variational inequality:

$$x \in X \text{ and } \nabla f(x)^T (z - x) \geq 0, \ \forall z \in X,$$

when f is smooth enough. These are necessary and sufficient for optimality under a convexity assumption. For the optimality conditions of (2.1), where the constraints

$g(x) \in K$ have a particular representation, Lagrange multipliers can be introduced and the variational inequality are the so-called KKT-conditions. In this setting, a constraint qualification may be needed to prove equivalence to the optimization. The motivation to call this problem format an equilbrium problem arises from the consideration of the variational form of the Signorini problem [32]. Specialized techniques for solution are given in [53], for example.

A bilevel program is an example of a hierarchical optimization where a parametric version of (2.1), the so-called lower level (follower) problem, is embedded in the constraint set of an upper level (leader) case of (2.1). Formally,

$$\min_{(x,y)\in X} f_U(x,y) \text{ s.t. } g_U(x,y) \in K_U, y \in \mathrm{SOL}_L(x) \tag{2.4}$$

where

$$\mathrm{SOL}_L(x) := \arg\min_{z\in Y} f_L(x,z) \text{ s.t. } g_L(x,z) \in K_L.$$

In other settings, SOL_L might consist of the optimal solutions of several linked optimization problems as in a non-cooperative game. Here the lower level problem $y \in \mathrm{SOL}_L(x)$ can be replaced by a set valued inclusion $(x, y) \in \mathrm{SOL}_L$ that represents a more general parametric equilibrium:

$$\min_{(x,y)\in X} f_U(x,y) \text{ s.t. } g_U(x,y) \in K_U, (x,y) \in \mathrm{SOL}_L \tag{2.5}$$

For example, there may be many followers $f_L(i)$, $i \in I$, where given the leader's policy choice x, the followers' actions are assumed to be chosen to give a *Nash equilibrium*, that is, no unilateral improvement for any follower. The leader seeks a policy that maximizes overall welfare. The mathematical formulation (2.5) of this problem is called a *Mathematical Program with Equilibrium Constraints* or *MPEC*. In fact, Mathematical Program with Equilibrium Constraints can encompass bilevel programs where the lower level parametric optimization problem is replaced by its variational form, thus

$$\min_{(x,y)\in X} f_U(x,y) \text{ s.t. } g_U(x,y) \in K_U, y \in Y,\ \nabla_y f_L(x,y)^T(z-y) \geq 0,\ \forall z \in Y$$

where for notational ease we have simplified the lower level problem to

$$\min_{z\in Y} f_L(x,z). \tag{2.6}$$

Assumptions are needed to guarantee that the variational form is necessary and sufficient for optimality in (2.6).

The principal-agent problem is an instance of the bilevel programming problem. In this case, the leader is the principal (owner) and the agent (manager) is the follower. The agent's actions $y = a$ are chosen to optimize their expected utility $V_A(w, a)$ given that the principal sets a reward $x = w$. The principal optimizes their expected utility $V_P(w, a)$. Note that the agent only accepts the contract if $V_A(w, a) \geq v_0$, so a participation constraint is added to the upper level problem. The bilevel form is thus:

$$\max_{(w,a)\in X} V_P(w,a) \text{ s.t. } V_A(w,a) \geq v_0, a \in \arg\max_{z\in Y} V_A(w,z). \tag{2.7}$$

The last constraint in this model ensures that the chosen action is also the agent's best response. It is of course possible to convert this to an MPEC under assumptions that guarantee the lower level optimization can be replaced by its variational form.

2.3. Machine learning models. Dynamic models and the process of learning (both in the model and in the underlying system being modelled) is a critical part of both strategic and operational analyses. The problem of estimating the cost of new technologies and the evolution of these over time with large scale adoption/market entry is difficult and learning often generates much greater changes than anticipated in models [89]. This may be due to the large scale replication of some technologies and potentially gives advantages to the growth of these provisions over other alternatives (e.g. nuclear, or carbon capture and storage (CCS)) that have high development costs, isolated deployment and lack of learning. The paper [89] argues that principled use of these models, complete with confidence intervals related to uncertainties is an important tool in debates such as the one considered in this paper.

A second setting where learning a model (or a solution map) has enormous potential is the following, as advocated in [16, 87]. Suppose a particular aspect of the problem is captured by a (complicated) model, but formulating that within a larger model or decision process is difficult. Instead of using the explicit submodel, machine learning (or other stochastic modelling techniques) could be used to train a surrogate that is easy to use. That surrogate could be an approximation of the solution map (for example a deep neural net) of the submodel, or some alternative. Often the training of the surrogate model is expensive but can be done offline, possibly using advanced computing techniques to speed the process. However, once the surrogate is built, it is very easy, fast and cheap to use. While deep learning is stressed in the approaches cited above, other techniques such as SDDP [25] are also possible. An important application is to deal with security constraints or reliability, where these surrogates can be used in contingency checking for reserves design. There may be issues about prices when using such approximations are used and questions to address regarding out-of-sample performance. In many cases, the surrogate could be used to populate a lookup table for approximation of the original phenomenon. There is also an extensive literature on reduced order models [11] that have the same kinds of properties that are being advocated here, allowing the use of these models in multiple scale instances for example.

A specific example involves the learning of linear constraints (cuts) using machine learning. The context could be failure modes of an electricity system, where the original model is difficult, for example a non-convex AC optimal power flow (ACOPF) model. On loss of a transmission line, the "Latta formula" [68] is used to compute the flows on each of the resulting paths and the change in temperature induced then generates a time to failure. Currently constraints are added in "by hand" to guard against this possibility. Can we instead use a huge sample (offline learning) to generate a feasibility cut (rule of thumb replacement). Similar suggestions are found in [33] for a machine learning model trained to replicate hydro storage policies that could be used to inform the mix of other generation.

2.4. Forecasting and prediction models. There is an enormous literature on forecasting that utilizes methodologies such as deep neural nets, statistical learning [48] and data analytics. In this paper we assume such methods are used to generate forecasts that can be used for data provision in our models, but do not describe them further since their black-box nature makes it difficult to interpret results and understand the model constructs generated. Some references can be found in the following survey papers [46, 85].

3. Examples. In this section we look at examples of problems arising in the architecture of green energy systems that can be modelled using the approaches outlined in Section 2. Our catalog of examples is loosely ordered by their scale, from the small to the large. Furthermore, the models are broadly conditioned on looking at issues of flexibility in planning, ensuring the problems determine decisions on technologies and capacities that are informed by operational characteristics of the desired energy system.

3.1. Household electricity planning. The simplest agent engaged in the transition to green energy is the individual person or household. They make decisions on the level and type of energy consumption for heating, refrigeration, cleaning, entertainment, and transport. While there is an interplay between human behaviour and smart control systems, the increasing use of available information and optimized operations (e.g. via model predictive control) has the potential to reduce load and/or use green resources more fully. Households might choose to use a combination of rooftop solar energy, batteries and electric vehicles to meet their needs. If they are exposed to carbon charges and time-varying electricity prices, then they face a capacity planning problem that chooses the capacity of heat pumps, solar panels, battery and car battery, and an operating policy of electricity consumption and battery charging/discharging to meet expected energy needs. This is a two-stage stochastic program in which the first stage defines capacity choices and the second stage is an infinite-horizon stochastic optimal control problem that defines the operating policy.

$$
\begin{aligned}
\min_{z,x,u} \quad & K(z) + V \\
\text{s.t.} \quad & V = \mathbb{E}[\sum_{t=0}^{\infty} \beta^t f_t(x_t, u_t, \xi_t)], \\
& z \in Z, \quad x_t \in \mathcal{X}(z, \xi), \quad u_t \in \mathcal{U}(z, \xi).
\end{aligned}
$$

Note that the constraint set Z can encode many complicated engineering relationships involving the investments z. The state variable x_t represents storage and the control u_t represents charge and discharge of storage as well as electricity purchases and load shedding. The set $\mathcal{U}(z, \xi)$ represents both household demand for electricity and supply of power from investments z. The operating costs $f_t(x_t, u_t, \xi_t)$ are discounted with discount factor β. Details and data for the capital, operating and lost load costs and the demand profile are not specified here, but represent samples for different operational cases. Of course, many households make investment decisions in solar panels and batteries without this sort of analysis as they are typically not exposed to varying electricity price and the household savings from optimal operations are too small to warrant the solution of a complicated optimization model.

While much of the energy management can be carried out "behind the meter", agents might interact directly with the electricity market whenever they have a deficit or excess of power. Choices between purchase or load reduction (turning off appliances) can be price directed. Some companies install solar panel systems with built in controls that promise guaranteed electricity savings over a fixed time horizon, obviating the need for households to optimize individually. Such disaggregated control has some drawbacks as potential system stability problems may ensue if appliances of many agents respond simultaneously to a single price signal without some coordination.

3.2. Aggregators and micro grids. Solar generation falls into two categories, residential (often called roof-top) and utility-scale (often called solar farms). Determining the sizing of these farms is an optimization problem. Is it better to have a

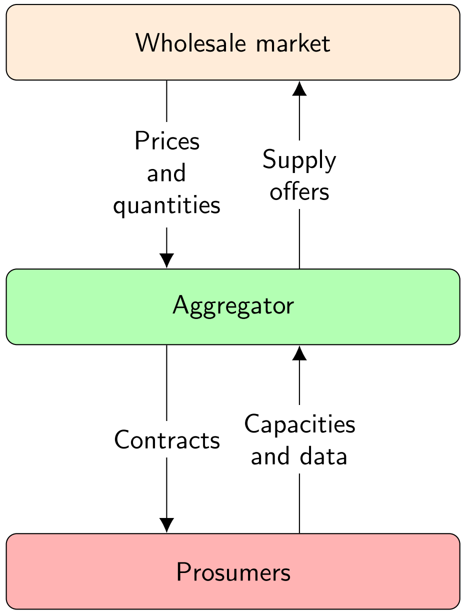

Fig. 2. *Aggregator as intermediary between prosumer and electricity market: based on [36]*

large single facility or a distributed collection of smaller ones? The answer will depend on land availability, and issues relating to the connection of this supply to the electrical grid.

Aggregators combine household demand and solar generation into a single energy source. This allows an aggregator to act as a virtual power plant and provide promises to deliver at least a certain amount of power/energy in a given time frame. These resources are part of a growing collection of distributed energy resources (DER), and their effective management and scheduling are mathematical problems that have received some recent attention [63]. Individual households typically cannot make such strong promises due to variability in the amount they can supply. Aggregation can reduce that variability, a property that is utilized to give diversified investments in the financial industry. Additionally, an aggregator can handle issues such as construction delays (a solar farm takes anywhere from 6 to 12 months to build), local and municipal permitting and approval processes, and ongoing maintenance and operation concerns [15]. The main concerns here are electrical engineering issues (and possible legality) related to distributed injection of supply, such as voltage support and frequency regulation. Questions arise around the regulatory policy (see, e.g., [28]) vis-a-vis the size of the aggregate supplier, and also to whether innovations such as digital transformers can provide alternative technical solutions [60]).

A schematic showing the typical operation of an aggregator system is shown in Figure 2.

Operational models for aggregators can vary. In [47], aggregators are the intermediaries between a collection of prosumers (the combination of a producer and a consumer) and the electricity market, whereas in [66] a different approach is taken where consumers are aggregated in a demand response setting. The aggregator's design problem is to select from a collection of distributed solar energy sources those that in aggregate will generate a certain volume of energy with the smallest variation in output (essentially the Markowitz model [58] in finance). We consider a design where solar energy sources are aggregated and augmented with batteries to smooth short-time fluctuations. If we let Q represent the matrix of covariances in energy output of solar sources, r be the vector of expected energy outputs, and $x = (x_i)$ be a binary variable that includes source i or not, we solve

$$\begin{aligned} \min_{x \in X} \quad & c^T x + \varphi(x^T Q x) \\ \text{s.t.} \quad & r^T x \geq d. \end{aligned}$$

X captures other constraints on x, and the objective adds the cost of solar installation to the cost $\varphi(\cdot)$ of batteries to deal with the overall variation in supply. The constraint then ensures average power output is above a threshold for interactions with the electricity grid.

In the context of distributed green energy systems, one concern is whether it is better to design the system for local use (i.e. use rooftop solar to power residential air conditioners directly behind the meter) and store excess locally in some form for later use (disaggregated storage), or is it better to directly deliver the excess to the electricity market, or have an aggregator manage the (excess) supply? These choices are compounded by supply intermittency when the local user has a deficit of energy and needs to procure it from elsewhere. The choice of storage mechanism is part of the design, and requires understanding the usage pattern - short or long time storage, power or energy requirements. In another section we touch on other aspects of storage or aggregated control related to reliability guarantees of the overall system.

Direct interaction with the market by a prosumer can be modelled as a special case of the aggregator problem. Interactions with the electricity market are governed by standard mechanisms described in Subsection 3.5. The remaining design decisions relate to the pricing of energy flows between the prosumer and the aggregator, and the mechanism to control the prosumer demand. For example, the aggregator can rent the consumer's roof at a fixed price, install its own solar panels, and then control the energy flows as part of a (large) virtual prosumer. An issue for the aggregator is to determine what roof space to rent and at what price (connection charge and per unit cost or payment), a so-called two-part pricing model. These models form a contract between the prosumer and the aggregator and such contracts can take on many forms. A rental contract could pay a fixed amount per month, or might provide retail power to the household at a reduced rate. The latter contract must specify how the price is indexed to the price of energy, and there is a need to understand how long term increases in demand will be treated, a topic that is well-understood by electricity retailers. Four different models of how to integrate distributed energy resources (DER) into electricity markets are given in [36]. They all rely on following a participant two-part pricing model (connection charge and selling price of the aggregator), but differ in the regulations that the aggregator faces.

Net metering is a billing method that allows consumers to use electricity they have generated at a different time that when it was actually generated. It's especially important for non-dispatchable renewable energy sources like wind and solar. An issue is the possible over-compensation for solar energy, and new policies are being generated to address this. Mathematical models could address related issues of fairness and equity.

Aggregation is also possible for plug-in electric vehicles that are currently controlled by their owners. Imagine a world where a fleet is owned and controlled by a corporation and cars are available on demand for a particular trip. This enables the corporation to control charging and vehicle use using a similar model to those outlined above.

3.3. Distribution network architecture. Distribution companies operate the low voltage networks that distribute electricity from the high voltage transmission grid to consumers. These operations are subject to variability from local demand and generation but also from equipment failure. Distribution companies can install special devices and configure the topology of the network to make it resilient to this variability. Dynamic topology control that switches lines in and out of the network also provides flexibility [34, 40, 41, 54]. For example, a mesh design (that provides redundancy in the form of multiple connection paths) can be configured as a radial network, allowing failures to be accurately identified and isolated. Lines (including those that are switched out) can be reinforced to reconnect the distribution service in case of failure (see for example [81]). In addition to these actions, the distribution company can procure flexibility services from battery storage or interruptible load. In a green energy system that has distributed battery capacity, these could be utilized for short term supply during a reconfiguration process. The type and amount of services to be procured depends on their offered cost, the existing flexibility actions available to the distribution company, and the level of reliability they require.

Since the distribution network is largely radial, and flows are simpler to calculate in this setting (similar to TCP/IP in internet applications), there is a tradeoff related to the design of redundant links and their capacities. The idea is to have additional lines built for reliability but switch them in and out to have a tree for flow calculation and control. There is an extensive literature on this, related to tie lines, and back-feeding. It is unclear if there are new issues related specifically to the green energy transition.

3.4. Load forecasting. Estimating load on the electricty system is crucial for many, if not all, models. Load forecasting is often categorized into: 1) Short-term (one hour to one week), 2) Medium-term (week to a year), and 3) Long-term (longer than a year) settings that are appropriate for different use cases. New policy issues, disruptive technologies to facilitate the transition, engineering and economic enhancements that change usage patterns, and efforts to electrify both heating and transport lead to substantive changes in electric demand. In fact, the fast growth in the use of LLM's across society and the world had led to huge increases in the use of computational resources and consequently in energy to power them. Some see this as a principal limitation to the AI revolution. Such perturbations must be included in the load forecasts for them to be at all useful. A recent survey is provided in [65].

A popular approach is to use a neural network approach [6] for the load forecasts. The paper [91] solves an optimal load dispatch model of a grid-connected community microgrid which contains residential power load, photovoltaic arrays, electric vehicles (EV), and energy storage systems (ESS), under three contrasting scheduling scenarios. In the load dispatch model, the residential power load and the photovoltaic power output were obtained from the forecasting results of a neural net model. The total cost of the proposed model consists of transaction costs between the microgrid and the main power grid, depreciation cost of EV and ESS, and treatment cost of pollutant emissions. Simple limit constraints specify interaction with the electrical grid.

A related forecasting problem is that of determining the value of lost load (VOLL), a single parameter that defines the price above which customers would wish to be disconnected rather than consume energy. It is clear that such a price would be different for every consumer, so a point estimate for VOLL is a coarse approximation.

Nevertheless, VOLL is widely used in security calculations and capacity planning models. These might benefit from a more nuanced model of demand reduction at high prices.

3.5. Electricity system operations. The economic dispatch model consists of buses $\mathcal{B}$, lines $\mathcal{L}$ and generators $\mathcal{G} \subset \mathcal{B}$ in an optimization:

$$\min_{(q,\theta,y)\in X} \quad \sum_{i\in\mathcal{G}} c_i(q_i^g) \tag{3.1}$$

$$\text{s.t.} \quad q_i^g - q_i^d = \sum_{j\in\delta^+(i)} y_{ij} - \sum_{j\in\delta^-(i)} y_{ji}, \qquad i \in \mathcal{B} \tag{3.2}$$

$$B_{ij}(\theta_i - \theta_j) = y_{ij}, \qquad (i,j) \in \mathcal{L} \tag{3.3}$$

$$-\bar{y}_{ij} \le y_{ij} \le \bar{y}_{ij}, \qquad (i,j) \in \mathcal{L} \tag{3.4}$$

$$q_i^{min} \le q_i^g \le q_i^{max}, \qquad i \in \mathcal{G} \tag{3.5}$$

where $\delta^+(i) = \{j \in \mathcal{B} : (i,j) \in \mathcal{L}\}$, $\delta^-(i) = \{j \in \mathcal{B} : (j,i) \in \mathcal{L}\}$ specify the network structure, B_{ij}, $q_i^{min,max}$, $\bar{y}_{ij}$ are electrical properties and c_i are production cost functions (most often linear or quadratic), and q_i^d is demand, see for example [84]. Variables determine active generated power q^g, voltage phase angles θ and active power flows y. Extensions of this basic problem can be used to incorporate different load conditions, failures, and maintenance schedules for instance (see for example [49]).

Locational marginal prices (LMPs), defined by the Lagrange multipliers (dual variables) on (3.2), can be shown to maximize total welfare of producers and consumers in perfectly competitive markets under assumptions of convexity and completeness. Under some additional assumptions this is true in dynamic stochastic settings as well [30]. This feature is becoming important for renewable systems with storage.

Locational marginal prices are less attractive when optimizing systems with large thermal plant having minimum operating levels and fixed costs for switching on and off. In the setting above, we might add a constraint and binary variables x

$$q_i^{min} x \le q_i^g \le q_i^{max} x, x \in \{0,1\}$$

to force a particular generator to operate at 0, or in the range $[q_i^{min}, q_i^{max}]$, $q_i^{min} > 0$. Here the lack of convexity invalidates the classical welfare theorems. In practice most system operators in LMP markets solve mixed integer programming problems to determine what plant should run, and when. Marginal prices from such a dispatch are not always sufficient to pay for generators' costs, and so "make-whole" payments are required to provide incentives for participation in the market. See [5] for a recent detailed discussion of the merits of such centrally dispatched systems in contrast to self-dispatched systems.

Some electricity market system operators (such as New Zealand and Australia) solve (convex) dispatch problems formulated as linear programs. To enable this they require supply curves to represent minimum operating levels and start-up and shut-down costs in the offered "marginal" cost curve. In other words, in a single-period setting, a plant that is currently off might mark up the marginal cost of its offer by an amount that would cover the cost of switching on if it were dispatched. A plant that was currently operating would offer at a discount to ensure that it was not switched off. Such a dispatch model treats these as truthful marginal cost declarations and yields LMPs that reflect these. The welfare theorems of convex markets obviate the need for make-whole payments.

There are two disadvantages with this approach. Unlike conventional marginal costs that can be calculated from fuel costs and heat rates, amortized start-up and shut-down costs are difficult to estimate. For example, should a start-up cost be amortized over a 30 minute period or over the expected period that the unit is on? To avoid a shortfall, suppliers will be conservative, and so the cost of dispatch will generally be higher than one obtained by solving a MIP. This loss in efficiency will be more pronounced when there are many large thermal units that can operate in different combinations. A MIP that accurately models starts and shuts can cut through these to yield a less expensive dispatch.

A second disadvantage comes from the increased difficulty in monitoring the potential strategic behavior of market participants who are now freed from any imposed regulatory constraint to offer at short-run marginal cost. In markets that use MIPs to dispatch generation plant, the start-up and shut-down costs and no-load costs are also much harder to estimate than fuel costs, so there is admittedly a similar incentive for generators to mark these up above their true values without being detected.

As electricity markets include growing amounts of intermittent generation and storage devices, the make-whole payments required to incentivize participation have been increasing (see [43]). While LMPS are currently computed using deterministic models, the dynamic stochastic features of markets with green energy seem to require a different approach to price formation to properly reward flexibility [26]. It is possible that the replacement of coal and gas plant by wind and solar generators will decrease economies of scale and lead to dispatch problems that can be well approximated by convex stochastic optimization problems, reducing the need for make-whole payments.

Stochastic market clearing models have a new set of challenges, even if convexity can be assumed. Even in markets approximated as a two-stage stochastic program with a finite probability distribution the optimal solution cannot be both budget balanced (where the independent system operator does not lose money) and recover each agent's costs (each market participant does not lose money) in every scenario (see [19]). It is possible under some strong assumptions on completeness of the risk market to ensure budget balance and cost recovery in risk-adjusted expectation which at least makes participation individually rational. A deeper philosophical problem with stochastic dispatch is an assumption that agents agree on the underling probability distribution used in the stochastic program. Rather than imposing a distribution, markets are supposed to be a mechanism for eliciting these probability distributions from a range of participants who each "put their money where their mouth is".

Stochastic market clearing models must also be dynamic, treating many trading periods at once, so they are stochastic optimal control problems rather than two-stage problems. Since the realized values of random variables in the future will inevitably differ from those in any model, the optimal control problems need to be updated in a rolling horizon fashion, as these values are discovered. Currently, a number of markets adopt this rolling horizon approach in a deterministic setting where single forecasts are updated. Such look-ahead dispatch models can yield efficient dispatch solutions, but can cause consistency problems in the resulting LMPs [42].

There is a large body of research related to non-convex models for electricity operations and markets. Some of this is related to AC optimal power flow (ACOPF) problems that are represented as non-convex nonlinear optimization models (or convex approximations of these). Such models are typically more challenging to solve compared with the linear models outlined above, and can be less favourable for analyses due to a lack of accompanying duality theory. In addition to this, models of unit commitment are also non-convex, typically modelled using mixed integer programs, leading to concerns about the specification of energy prices. Much research continues into the effective inclusion of these models into practice and the additional value to the consumer that this facilitates, and that will be important while traditional generation (such as gas turbines used as backup) is still in operation. As we move through the green transition, it is clear that the dynamic stochastic models will remain at the cutting edge for operational considerations, but it may be that eventually the non-convex ACOPF and unit-commitment issues will become less critical.

3.6. Emissions trading. Many countries have implemented cap-and-trade markets for greenhouse gas emissions [3, 86]. These differ in their implementation but generally involve a decreasing cap on annual emissions permits that must be surrendered each year by organizations to account for their emissions. The permits are auctioned by governments and traded in a secondary market. Given a price for a permit each emitter in the economy faces an optimization problem that equilibrates the price of permits against the marginal cost of reducing emissions.

In practice, emissions markets are subject to political intervention. Some sectors of the economy (e.g. farmers whose animals emit biogenic methane) are made exempt (at least temporarily) from surrendering permits. The reason is that the carbon charge imposes a cost that they cannot avoid in the short term by technological means. Extra costs might make them uncompetitive in international markets. This is unsustainable in the long run, as biogenic emissions must be reduced. Indeed many countries are beginning to add emission tariffs to imported goods, which effectively imposes the costs on farmers that were not imposed by emissions charges in their own country [61].

A second political intervention comes from the effect of emission charges on energy prices, notably gasoline and electricity. These price increases affect poor households disproportionately (as they spend a higher proportion of income on energy than wealthy households). Moreover poor households have limited access to cheap capital, so replacing legacy technologies such as gasoline cars and gas water heating is expensive. This results in strong advocacy for energy subsidies or for more substantial income redistribution through taxation policy to enable poor households to reduce emissions.

Ideally a global cap-and-trade market would result in a world carbon price that would reduce emissions in the most efficient way. A number of authors (see e.g. [51] have pointed to potential deficiencies in such a market. Lack of effective verification of permits can cause "carbon leakage" to less compliant countries and weakening in permit prices as experienced for about ten years after 2008. There are also potential market failures. Consider a least-cost optimal solution for the world to reach a desired emission target that requires a poor country to face a large fixed cost to be able to reduce emissions (say by building a large hydroelectric dam). A global emissions price might be insufficient to incentivize this. A subsidy from the rest of the world will enable this solution to be realized.

A related issue is how to measure the carbon intensity in consumed products or services. This is important for many companies who are aiming to demonstrate their engagement in green energy practices. If a carbon price is incorporated into the dispatch model, one could run simulations with and without that price, and observe dispatch differences for example on a heat map. Allocation of these differences to particular consumers requires new models. Another idea would be to fix the original dispatch and run calculations through that solution to find instantaneous carbon intensities, and then integrate through time to get overall intensities. It is unclear how effective such calculations could be given all the other uncertainties in the process.

There is an analogy here with uplift payments in optimal dispatch, where the marginal energy price is insufficient to produce the socially optimal outcome. Uplift payments incentivize participation of all generating plant in the optimal dispatch solution.

3.7. The role of storage, peaking and load shedding. The most popular forms of green electricity are generated by the wind and the sun. These sources are both intermittent and uncertain. Intermittency (the fact that the sun does not shine at night) and the (random) variability (due to cloud cover or other effects) can be treated separately [90]. In some areas solar insolation is reasonably predictable but is not available at night time. If the solar power exceeds demand during the day and is not exported then some form of energy storage might be desirable to use the power generated during the day in the evening and night time. This storage is intended to be cycled on a daily basis, and will save its operators money by reducing night-time power consumption that must otherwise be bought off the grid [82]. Batteries are typically used to perform this function if the discounted electricity cost saved over the battery life covers its capital cost. Batteries also can be used to transfer energy between time periods for other variable sources of energy such as wind power [50]. Specific mathematical models of batteries for use in storage models can be found in [73], for example.

Like any inventory, battery storage also plays a role when supply and demand are unpredictable [22]. Energy storage then provides a hedge against future uncertainty. The optimal sizing, location and operation of batteries under these circumstances requires a stochastic optimization model that represents the short-term uncertainty in supply, e.g., when predicted wind does not eventuate [94].

An alternative approach installs fast-start peaking generators to deal with uncertain and intermittent renewable energy supply. These typically are open-cycle natural gas turbines, but they could be configured to run on biofuel or green methane produced from carbon capture and hydrogen. The optimal sizing, location and operation of such peaking plant also requires a stochastic optimization model. Instead of installing peaking capacity, the system might arrange for (industrial) consumers to shed load in response to price. This *demand response* essentially performs the same function as a peaking plant. Estimating demand response for different customer types requires some estimate of their marginal value of electricity, which is much harder to determine compared with a price of natural gas. Another alternative is to use a battery to provide the peaking functionality [23]. The related concept of *load balancers* is considered in the internet reliability literature. Load balancing techniques make internet operations more failsafe than electricity grid operations. It may be possible to leverage approaches such as those labelled “canary” to increase the reliability of the green energy grid.

Storage can also operate over a longer time scale (see [78]). For example in some regions where energy supply is seasonal, hydroelectric reservoirs are used to transfer water from melting snow or wet season rainfall to dry seasons of the year. The water in these systems stores energy. In contrast to short-term battery storage that can be used to overcome a limitation on electricity *capacity*, reservoir storage is a response to seasonal *energy* limitations.

Electricity markets have traditionally not needed to consider moving energy over time and so their design is not necessarily ideal when the infiltration of storage devices and other methods to move load across time becomes significant. The use of stochastic programming models has drawbacks that include (a) the need to agree on the scenarios used for uncertain parameters in the model, and (b) the mispricing of the option value of energy storage and the value of increasing current dispatch to meet future ramping constraints, and (c) the need for uplift payments that compensate participants for the fact that the system operator forecast the future incorrectly. In a recent paper [69], we have proposed a new class of economic dispatch models that attempt to overcome these drawbacks, based on agent decision rules (ADRs). Forecasting future outcomes or scenarios passes from the system operator to market participants who make state-dependent offers of energy using these decision rules. When generators and battery owners face future uncertainty they take positions that risk losses. Some of these losses result from being dispatched in advance of a realized random price under which they would have preferred to be dispatched differently. In our dispatch model, we propose that the generators should factor this possibility into their ADRs. The ADRs will involve a short-run marginal cost and a state-dependent future cost function that enables the system operator to dispatch them in a single period optimization. ADR's can accommodate existing supply function bids, but also enable modelling of flexible load shifting, demand response and reserve bids.

3.8. Transmission. Electricity transmission architecture is a key component of the transition to green energy. Historically, transmission of electricity has been driven by economies of scale in generation. Electricity generation from large-scale coal and nuclear plant needs transmission to make it available to consumers that can be located many miles from generator locations. The cost of transmission lines has historically been low compared with the costs of proliferating small plants for local electricity generation. Even as these costs fall, transmission remains important since renewable sources of energy (e.g. offshore wind) are not always located where demand is.

In most electricity markets, transmission is separated from energy production, and is owned and operated by an independent regulated monopoly. Designing transmission systems to achieve desirable social outcomes is nevertheless a challenging optimization problem. Examples of models that study this are [57] in a deterministic setting, [88] in a setting with random wind and transmission switching, and [74] and [93] in a principal-agent setting.

For switching problems, the economic dispatch problem can be updated to replace constraints (3.3) and (3.4) by

$$
\begin{aligned}
&B_{ij}(\theta_i - \theta_j) - M_{ij}(1 - x_{ij}) \le y_{ij} \le B_{ij}(\theta_i - \theta_j) + M_{ij}(1 - x_{ij}) \\
&- \bar{y}_{ij} x_{ij} \le y_{ij} \le \bar{y}_{ij} x_{ij},
\end{aligned}
$$

for $(i, j) \in \mathcal{L}$, where M_{ij} represent so-called big-M constants that facilitate the switching on and off of a given line ij, and binary variables x represent switching decisions.

Reconfiguration and initial design share many similar features, particularly if a given set of choices is specified a-priori. In this case, investment costs could be added to the objective:

$$
\sum_{i \in G} c_i(q_i^g) + \sum_{ij} b_{ij} x_{ij}.
$$

3.9. Energy/resource tradeoffs. Land is finite, and using it for energy generation such as in solar farms, or more generally for climate renewal as in reforestation, precludes agricultural production or other uses. Similarly, biofuel production (corn for ethanol instead of feed) and dam building for new hydro generation uses land for energy while reducing its availability for other uses. In this context equilibrium models are relevant, allowing a price to determine efficient allocation of scarce resources to a variety uses. Certainly, the tradeoff does not need to be limited to energy and land, but could involve other finite resources, or other environmental concerns.

As mentioned in the introduction, many forms of green energy may involve some use of finite resources. Batteries involve the extraction of rare-earth materials, and deforestation occurs in the extraction of copper. How can our models capture these effects? Do we need to consider more complex life cycle models accounting for all inputs, for example. Or is a pricing mechanism an effective way to encourage capital investment in alternatives?

More generally, energy generation and consumption is part of a broader economic landscape where energy and the products and services it enables are transferred between different sectors of the economy. The effect of a change in the energy architecture will be felt in all sectors and requires a model of the whole economy to evaluate. Integrated Assessment Models (IAMs) of which there are many (see [67, 14]) aim to model these intersectoral energy flows in a system optimization framework. Alternative approaches use computable general equilbrium models of the economy (see, e.g., [92, 14]). Such models have already looked at the viability for example of biofuels, and could easily be adapted to other alternative technologies.

3.10. Engineering and operational models. The growth in green energy that is required to reach net zero might be less than projected if one focuses on the services that use energy rather than the energy that they currently consume. According to [29], “Energy services are those functions performed using energy which are means to obtain or facilitate desired end services or states.” Consumer services might be redesigned so that they consume much less energy. Car and bike sharing are such services, others relate to batteries. Operations research models can determine prices, location of equipment and operational procedures to make these more efficient. Improvements can accrue from locating the service near to the source of energy generation or demand, with conversion from a storage entity occurring prior to use. Note that these models could benefit from an analogy to content delivery/distribution networks (CDN’s) in the internet delivery literature and practice.

Typical infrastructure planning models are also applicable to problems involving the type, quantity and siting of EV charging stations [9, 8]. Such models combine a node-based approach with a flow-based approach to model the needs of EVs to recharge on intermediary stops on long-haul travels, and uses a bilevel approach. Application to other infrastructure needs for the energy transition could build upon these and similar models. In general, it is possible to convert any form of energy into another target form, having different properties from the source form. Only 40% of the energy used in the United States is currently supplied by electricity. The majority of the remaining 60% of energy is supplied by directly combusting fossil fuels like gasoline to power cars or by burning natural gas for heat and cooking.

3.10.1. Conversion for Storage. As mentioned above, electricity can be converted to a chemical form in a battery for example that allows for energy to be stored over short time periods, or water can be pumped uphill creating potential energy for later conversion using gravity and turbines. Such conversions are lossy, in that some energy is expended and lost in the conversion process. Electricity is expensive to store since it incurs these losses both in conversion and possibly over time due to leakage.

Storage also requires capital and this adds to the expense. Batteries have high conversion efficiencies but have a high capital cost per MWh stored. A principal use of batteries is therefore to transfer electrical energy over short time periods, allowing repeated use of the battery over time to arbitrage prices so as to recover capital costs from high utilization. The timing of charge/discharge can be determined effectively using stochastic control models.

For longer time frames of storage, batteries are not as effective since they are used less frequently and so cannot recover their capital costs. In this setting, there may be conversions of the electrical energy that are less efficient from an energy conversion perspective, but allow the energy to be moved across time to where it is much more valuable. These conversions may even be relatively inexpensive from a capital perspective, as they might only use excess capacity of existing/deployed technologies (such as ammonia generation or hydrogen to methane conversion). More generally, conversions could be done locally, converting generated energy into a form suitable for local storage and later use at that location or for more effective transport (e.g. methane is more easily transported in pipes with lower losses than hydrogen). Optimization again can be used to determine what conversions to do, where to do them, and at what scale.

3.10.2. Portfolio of Storage. System optimization models can shed light on these conversions and which ones are effective in a given portfolio. We illustrate this with a toy example. Consider a set K of different storage types (say ammonia, green methane, hydrogen, pumped storage, and battery), with variables for the amount of energy stored $s_{kt}(\omega)$ in storage type k in a scenario ω at time $t = 1 \ldots, T$, and the related charging $q^+_{kt}(\omega)$ and discharging $q^-_{kt}(\omega)$ profiles. Integer variables x_k determine how many units of k are installed. The overall cost of operation is given by

$$\sum_k c_k x_k + \mathbb{E}_\omega \left(\sum_t \gamma_k (q^+_{kt}(\omega) + q^-_{kt}(\omega)) + p_t(\omega)(q^+_{kt}(\omega) - q^-_{kt}(\omega)) \right)$$

where c_k is the per period capital charge for storage k, γ_k represents the cost due to cycling the battery and $p_t(\omega)$ is the price paid for energy at t. The system dynamics are modelled by

$$s_{k(t+1)}(\omega) = s_{kt}(\omega) + e_k q^+_{kt}(\omega) - q^-_{kt}(\omega)$$

where e_k is the charging efficiency, and composition of the portfolio of storage is determined using

$$s_{kt}(\omega) \leq \mathcal{S}_k x_k$$

with $\mathcal{S}_k$ being the size of a unit of the storage k. Residual demand $r_t(\omega)$ is related to storage via

$$r_t(\omega) = \sum_k q^-_{kt}(\omega) - q^+_{kt}(\omega)$$

This can be augmented with spill on the left hand side (that is penalized in the definition of cost perhaps) and the addition of a peaking plant supply on the right if desired. The key to such models is in the data $(K, T, c_k, e_k, \mathcal{S}_k, r_t(\omega))$: we specify T as the number of hours in a year, and generate the demand $d_t(\omega)$ uniformly at random

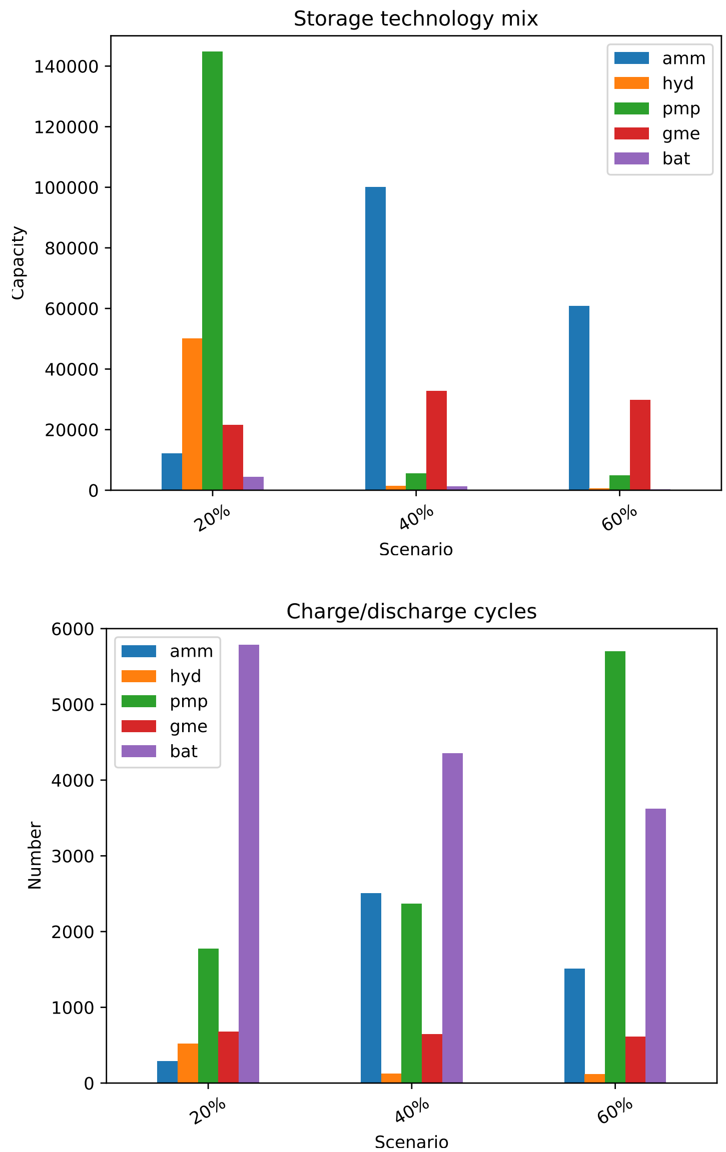

Fig. 3. *Storage portfolio and charging frequency under different generation design scenarios.*

(using an upper bound on the random sample in each time step generated by a seasonal underlying curve supplemented by daily deviations to capture the day/night cycles). Supply is specified so it provides an overbuild factor $1 + \eta$ more than the demand from generators, and residual demand is the difference of demand and supply. Other data are taken from estimates in the literature.

Figure 3 shows optimal installed capacity and the number of charge/discharge events for three different levels ($\eta = 0.2, 0.4, 0.6$) of renewable overbuild, in a free disposal regime without peaking plants. Installed battery capacity has high capital costs so the storage capacity chosen is small. It is used primarily to deal with demand

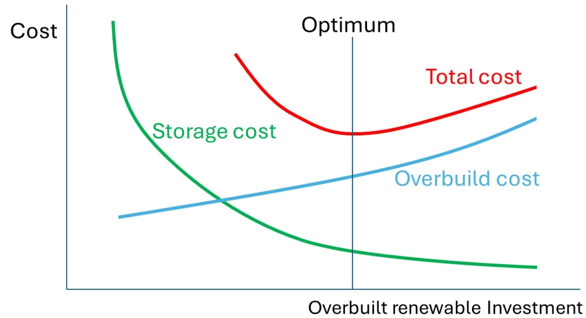

FIG. 4. *Optimizing renewable overbuild and storage*

peaks, so the frequency of its usage is large as shown in the lower panel of Figure 3. At low levels of excess renewable energy supply, the portfolio of storage investment is biased strongly towards the more efficient storage technologies (batteries and pump storage) to use the excess energy most effectively to avoid shortages. As the levels of renewable oversupply increase, ammonia and green methane become more attractive: the energy wasted by these less efficient storage technologies is less costly if there is a large surplus of energy and is outweighed by the lower capital cost of these technologies. Fewer batteries are built as oversupply increases, since this reduces peaking requirements that are increasingly handled by (less efficient) pump storage.

This simple model shows that a single choice of storage technology will not be optimal: we require a mix of storage technologies depending on the level of renewable overbuild. Of course the total costs of storage decrease as the amount of overbuilt renewable capacity increases, so there will be an optimal setting where the marginal cost of this equals the marginal decrease in storage cost. This is shown schematically in Figure 4. With an appropriate representation of the transmission network, the model can also be extended to determine the location of energy storage as well as its technology and size.

3.10.3. Conversion for Transport. Electricity is what we call a secondary energy source. It is created by converting primary sources of energy like fossil fuels, wind and solar energy, into electricity. It is a particularly useful form of energy because it can be quickly and efficiently transported over long distances and is readily usable in a multitude of settings (lighting, heat, mechanics, transport, etc). Electricity is also referred to as an energy carrier, which means it can be reconverted to other forms of energy such as mechanical energy or heat.

Transmission of electricity over long distances incurs losses through dissipated heat. (These losses are reduced by increasing the voltage and decreasing the electrical current.) The capital cost of the transmission infrastructure and the cost of energy losses can be compared with alternative forms of energy transport.

Since around only 40% of global carbon dioxide (CO2) emissions originate from power generation which can be decarbonized via electrification, the other 60% coming mostly from industry, buildings and transport have to be decarbonized using alternative means, one of which is green hydrogen. Green hydrogen is a versatile energy carrier that can be used directly or in the form of its derivatives like methanol or ammonia. Generating green hydrogen needs water (some regions are considering desalination to provide this) and uses various forms of electrolysis. One could imagine converting electricity to hydrogen gas at a large generation plant, transporting the hydrogen to a city, and then storing it and converting it back to electricity through combustion or fuel cells when it is needed. This enables the energy to be available at peak times. Note, however, that each conversion incurs a loss of energy and hydrogen is very expensive to transport (being light but requiring heavy pressure vessels, or susceptible to leaks from conventional gas pipes).

An alternative model transports electricity to the city and makes hydrogen locally. Electrolysers to make hydrogen can be made cheaply at very small scale, and require only electricity and fresh water as fuel. This means that electricity rather than hydrogen is transported, and hydrogen can be made and stored locally where the demand occurs. Such a model requires a transmission grid to be dimensioned to meet extra demand, but avoids the much higher costs of hydrogen transport. The model in Subsection 3.10.2 can be easily extended to address these issues.

Demand for energy can change due to changes in behavior of users. There are concerns about the electrification of urban transport expressed for example in [17]. While a very high gasoline tax would yield some interesting developments, it is unclear how elastic the demand is, and whether such policies would lead to more working from home, more use of public transport and electric vehicles. For another example, air transportation is very energy intensive and currently not very green. Transition strategies are focused on sustainable aviation fuel (SAF), liquid hydrogen and electric power, both pure and hybrid [38]. The aggregation of transport by sea or pipeline instead of airlines or trucking could reduce emissions substantially, perhaps at the cost of longer transport times. Passenger travel via sea instead of by air might also involve much longer times, but at a smaller energy cost per person. Models could shed light on the underlying properties that are being utilized here - is the key simply economies of scale? Tradeoffs based on behavior change are not limited to the energy sector but will impact other sectors such as tourism and industrial productivity.

4. Risk.

In the classical finance literature, risk is identified with variance. In some settings this makes it beneficial to reduce variance through aggregation. As in the model of Subsection 3.2, a collection of wind turbines with uncorrelated variable wind generation can be aggregated to give a more predictable supply, which presents advantages to economic dispatch models. Similarly the capital asset pricing model translates variance in returns into a discount rate that can be used to assess the risk of uncertain cash flows, so reducing variance with no change in expected reward is deemed to be beneficial.

However, as noted by [59] the energy transition presents decision makers with risks (downside variance) and opportunities (upside variance). Ideally, optimization models should be able to take advantage of opportunities while minimizing risks. In contrast with models that minimize variance, risk-averse stochastic programming models using *coherent* risk measures [79] provide a principled approach for doing this.

Risk in settings with many agents requires careful handling. Each agent type is exposed to a unique set of risk factors that arise from their technology choices, climate, fuel source, exchange rates, and regulatory intervention. Some of these risks can be reduced through hedge contracts signed with counterparties who see reward opportunities in the risks faced by others. We give some examples of these transactions.

4.1. Short-term risk instruments.

A popular form of hedge contract is called a *contract for differences* (CFD). Arranged at some strike price f, this is a financial agreement to pay a counterparty $p - f$ where p is the observed price of electricity. So if party A intends to sell Q MWh to counterparty B at some future time, then Q CFDs arranged at f will hedge the unknown future price and conduct the transaction at known price f.

Weather derivatives are also a mechanism for reducing risk. Consider distributed solar, and demand from air-conditioning. In the event of a very sunny day, the air conditioners need more energy to run and the price would rise, but solar farms are producing more. A weather derivative in which the solar farm guarantees the air conditioner a certain amount of energy whenever the temperature (or insolation) is above a certain level will reduce the risk of losses of both parties.

For a second example of weather-based derivatives consider a geothermal generator. This has high capital costs and very low operating costs, so it make sense to run as a base-load plant. In the middle of the day when solar power is at a maximum, it might make sense for the electricity system to control geothermal output to avoid spilling energy. A solar farm might arrange a derivative contract with a geothermal plant that pays out when the sun shines, but imposes a cap on geothermal output at this time [44].

Can hedge contracts remove all risk? In an uncertain environment an *Arrow-Debreu security* is a derivative contract that pays \$1 to the holder if a particular future state of the world occurs. If these exist for every possible future state then in principle an agent can insure against any conceivable loss (at some ex-ante cost) by purchasing an appropriate Arrow-Debreu security off a counterparty.

This highly idealized situation would never occur in practice but it is a useful model to study risk and contracts. A relatively recently developed theory (see [76, 70, 30]) shows that if markets for energy are perfectly competitive and convex, and all agents are endowed with coherent risk measures, and the market for Arrow-Debreu securities is complete, then agents will trade their risk using these securities until no more risk can be hedged. The remaining risk is then treated by each agent as if they were using the risk measure of the least risk-averse agent. For example if some agents such as speculators were actually risk-neutral then a complete market for Arrow-Debreu securities will result in every agent optimizing the expectation of their costs and benefits (i.e., acting as neutral to risk). This theory enables one to establish useful welfare theorems that demonstrate that the markets deliver socially optimal outcomes.

In practice, risk markets are incomplete, so the welfare theorems do not hold. Computational studies show that removing some risk using CFDs and other instruments can improve welfare outcomes in incomplete markets. It is also possible to find counterexamples where adding instruments makes welfare worse [4]. Furthermore the computation of equilibria in incomplete settings is difficult as these might fail to exist or not be unique [37]. This is an active area of research in scientific computation (see, e.g. [52, 45]).

4.2. Long-term risk. The transition from a largely fossil-fueled energy system to a renewable system is expected to take decades. Although we can develop sophisticated planning models to guide the decisions made, these decisions will in many cases be made by commercial organizations in pursuit of profits, but also facing many uncertainties. Investment in energy production and infrastructure development is financed largely by borrowing, and the cost of this finance depends on the risk of the investment, and so organizations making investment decisions need to understand the risk of the investment as well as its (uncertain) reward.

Capacity investments must make non-negative risk-adjusted returns to be justified. In the risk-averse stochastic programming setting this amounts to a non-negative net present value with stochastic discount rates. In a complete market for risk, the trade of Arrow-Debreu securities leads companies to share the same stochastic discount rates. This allows the optimal capacity decisions for companies to be determined by a social planner who maximizes social NPV with the same discounting.

In practice, as in the short-term setting, risk markets are not complete, so a social planning solution might not match a risked equilibrium. The latter, however, can often be computed as the solution to a complementarity problem. As an example, consider the following equilibrium problem formulated in [19] where each generator chooses generating capacities and generation levels and retailers of energy choose amounts to buy[1]. Each agent a solves the problem:

$$
\begin{aligned}
P(a): \min_{(\boldsymbol{x}^a,\boldsymbol{z}^a,\boldsymbol{q}^a)\geq 0.} & \rho^a(Z^a) \\
\text{s.t. } Z^a(\omega) = & \sum_{k\in\mathcal{K}} \mathrm{K}_k \cdot z_k^a \\
& + \sum_{t\in\mathcal{T}}\sum_{k\in\mathcal{K}} (\mathrm{c}_{kt}(\omega) - \pi_t(\omega)) \cdot x_{kt}^a(\omega) \\
& + \sum_{t\in\mathcal{T}} (\pi_t(\omega) - \mathrm{r}) \cdot (\mathrm{d}_t^a(\omega) - q_t^a(\omega)) \\
& + \sum_{t\in\mathcal{T}} \mathrm{v} \cdot q_t^a(\omega) && \forall \omega\in\Omega, && (4.1)\\
x_{kt}^a(\omega) \leq & \ \mathrm{m}_{kt}(\omega)\cdot z_k^a && \forall k\in\mathcal{K},\omega\in\Omega,t\in\mathcal{T}, && (4.2)\\
\sum_{t\in\mathcal{T}} x_{kt}^a(\omega) \leq & \ \mathrm{n}_k(\omega)\cdot z_k^a && \forall k\in\mathcal{K},\omega\in\Omega && (4.3)\\
q_t^a(\omega) \leq & \ \mathrm{d}_t^a(\omega) && \forall \omega\in\Omega, t\in\mathcal{T}. && (4.4)
\end{aligned}
$$

The objective for each agent, a, is to minimize their own risk-adjusted disbenefit $\rho^a(Z^a)$. Here ρ^a is a coherent risk measure and $Z^a(\omega)$ is the net cost from investing and operating their fleet of generation in scenario ω as defined by (4.1). The constraints contain terms for both generators and retailers and so some will not be present for each type of agent. The generator a produces $x_{kt}^a(\omega)$ from plant type k and the retailer buys power at wholesale price $\pi_t(\omega)$ and sells it at fixed price $\boldsymbol{r}$. In the first line of (4.1), we have the physical capacity investment cost, $\sum_{k\in\mathcal{K}} \mathrm{K}_k \cdot z_k^a$, where the sum is over investment technologies. In the second line of (4.1), we have the component of the disbenefit from generation, $(\mathbf{c}-\boldsymbol{\pi})\boldsymbol{x}$, with $\mathbf{c}$ giving the marginal cost of generation, $\boldsymbol{\pi}$ the spot market price, and $\boldsymbol{x}$ the output of generation.

In the third term, we define the disbenefit from meeting demand. The per unit cost of meeting demand is given by $\boldsymbol{\pi} - \mathbf{r}$ with the agent having to purchase the electricity directly from the spot market at $\boldsymbol{\pi}$ and given $\mathbf{r}$ by the consumer. The demand met by the retail component of the agent is given by $\mathbf{d} - \boldsymbol{q}$. The exogenous demand of each consumer is given by $\mathbf{d}$, and $\boldsymbol{q}$ is how much the retail company decides to curtail. The overall profit is given by $(\boldsymbol{\pi} - \mathbf{r})(\mathbf{d} - \boldsymbol{q})$.

In the final term, we define the penalty the retail agent must pay for unmet

[1]In [19] there is also an ISO agent that dispatches power through a transmission network. We assume a single node model for simplicity.

demand, $\boldsymbol{q}$. The penalty is the value of lost load, $\mathbf{v}$, which is much higher than typically observed spot market prices. This penalty is added to the lost revenue from not meeting all of the consumer demand for electricity generation.

In equations (4.2) through (4.4), we define the physical constraints on generation and curtailment. Equation (4.2) limits the power output $\boldsymbol{x}$ of each plant, depending on the capacity investment $\boldsymbol{z}$ and some multiplicative adjustment, $\mathbf{m}$, that depends on the scenario and load block. Equation (4.3) limits the energy output of a generation plant. Finally, (4.4) limits consumption to be at most the level of demand.

To form a complementarity problem, the KKT conditions from problem $\mathrm{P}(a)$ for each agent a are added to the following market clearing conditions:

$$
\begin{aligned}
&0 \leq \sum_{a\in\mathcal{A},k\in\mathcal{K}} x^a_{kt}(\omega) + \sum_{a\in\mathcal{A}} q^a_t(\omega) - \sum_{a\in\mathcal{A}} \mathrm{d}^a_t(\omega) \perp \pi_t(\omega) \geq 0, && \forall \omega \in \Omega, t \in \mathcal{T},\\
&0 \leq \sum_{a\in\mathcal{A}} q^a_t(\omega) \perp \mathbf{r} + \mathbf{v} - \pi_t(\omega) \geq 0, && \forall \omega \in \Omega, t \in \mathcal{T}.
\end{aligned}
$$

These complementarity conditions ensure that supply meets demand at a competitive price. We have free disposal of power within our model, allowing supply to exceed demand at each node. However, when this occurs, the spot market price for electricity at this node will be 0. And when some positive amount of load is shed then the price hits its maximum value $\mathbf{r}+\mathbf{v}$. As mentioned above, the incompleteness of the market for trading risk complicates the existence, uniqueness and computation of equilibrium in these models, but in many practical instances equilibria exist and can be computed (see [55] and [4]).

As alluded to by [59], long-term investment decisions should maximize opportunity while controlling risk. Stochastic programming models that represent such real options are multistage, since opportunities are revealed over time as random variables are realized. Multistage risk-averse optimization has many variations depending on the form of conditional risk measure used. We mention two.

Given an adapted set of actions at each node of a scenario tree, an *end-of-horizon* risk measure sums the payoffs at each node along a path from root to leaf to give a scenario payoff. The risk of the set of actions is then evaluated using a coherent risk measure applied to this distribution of scenario payoffs. This is the predominant risk measure used in software for solving multistage models of capacity expansion under uncertainty (see, e.g., [24]).

Given an adapted set of actions at each node of a scenario tree, a *nested* risk measure computes the risk-adjusted payoff at the parent of each leaf node, using the payoffs at this node and its children. This risked "value-to-go" function is then used to evaluate the risk-adjusted payoff of the set of decisions at the grandparent of each leaf in a recursive pattern. This recursive definition ensures that the dynamic risk measure is time-consistent.

Dynamic risked equilibrium (see [30]) of many agents can be viewed as an open-loop problem or a closed-loop problem. In the former setting, agents choose every action in every state of the world on day 1, assuming other agents have fixed theirs. The response of an agent is then computed in response to this knowledge. Such an equilibrium is not subgame perfect. In a closed-loop equilibrium, an equilibrium is computed for every state of the world at the final time. The payoffs in this equilibrium then inform actions at the penultimate time, and the solution is computed recursively. As shown in [30], these two solution concepts yield the same result in perfectly competitive convex markets with complete risk markets. In imperfect or incomplete markets they are not the same. Developing computational methods for these problems is an active area of research (see [80]).

Why are these models important? Much effort has been devoted to developing integrated assessment models (IAMs) for understanding the transition to green energy. These models are (often deterministic) social planning models with high levels of physical fidelity, but treating the future as predictable scenarios. Including uncertainty and risk aversion in these models makes them more realistic, but the results need to be reconciled with commercial investment decisions of competing agents. Welfare theorems give some justification for using risk-averse IAMs as gold-standard benchmarks for the dynamic risked equilibria in incomplete markets that we believe are closer representations to what will actually occur.

4.3. Architecture for resilience. Unexpected outages (that can arise from operator mistakes, major storms or environmental disturbances, or even deliberate sabotage by adversarial actors) are a general concern in electrical energy systems. However, the more distributed nature of green energy systems may allow some enhancements, whereby cascading failures can be avoided by isolating subnetworks of the overall grid. Since more batteries or other storage devices are installed (to provide transfer of energy over time), those same resources could be made available (along with existing distributed generation) to facilitate balancing while isolated. This is a novel use of additional functionality installed in the system to improve overall resilience.

In any disaggregated system, the need arises for additional information to facilitate better overall control and stability. There is a large existing literature in the energy domain related to information, privacy and mechanism design (for markets, auctions, etc). The underlying question regarding the much finer scales of disaggregation that might come about in a green energy system brings up questions as to whether these existing mechanisms are sufficient in these new operating environments, or what changes and enhancements are needed. An alternative approach is outlined in [77].

4.4. Capacity markets. The transition to green energy will be costly. According to the International Energy Agency over 60% of the world's electricity in 2021 was generated from fossil fuels. Given that total electricity generation will increase from electrification of transport and industrial processes, the scale of the investment in green electricity capacity is immense.

This raises several important questions. What incentive structures are needed to ensure that the right mix of capacity is built? Is the dynamic risked equilibrium that emerges from commercial decisions enough to give the capacity increases that we need? Finally, will this equilibrium be achieved in time to avert a climate catastrophe? As an aside, the US is experiencing large interconnection queues [56] for the approval of new wind and solar projects. Models to explain this process are developed in [7].

The first question is an area of active research. As mentioned in Subsection 3.5 locational marginal prices (LMPs) are not always sufficient to incentivize optimal participant behavior. In perfectly competitive, convex energy-only markets LMPs provide economic rents that support optimal levels of investment at the margin determined by a *screening-curve* analysis [83] as depicted in Figure 5.

The screening curve shows the annual total cost per MW capacity plotted against the number of annual operating hours. The total cost is a combination of fixed and variable cost based on the number of production hours in a year. A minimum cost for each capacity factor can be found by combining the screening curve with the *load duration curve* (LDC), here approximated by 10 load blocks with piecewise constant demand. The projection produces the least-cost capacity combination that can serve

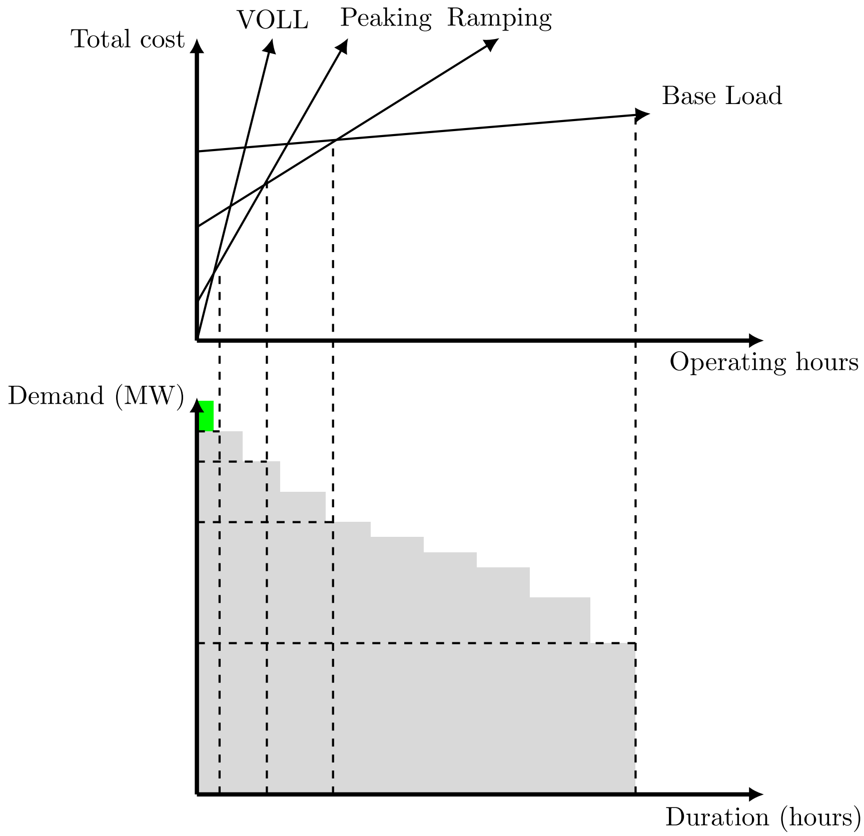

FIG. 5. *The screening curve: how capacity is traditionally planned in electricity systems.*

the load profile. For example, to supply the part of the LDC that has higher capacity factor (*i.e.*, running most of the year), base load is the least cost option. As the number of operating hours decreases, the plants that are less expensive to build but more costly to run begin to become more economical. For a small number of hours near the tip of the duration curve, high variable cost peakers are the most economical, while some load is shed at VOLL.

This picture is complicated by intermittent generation sources that are not dispatchable, and by risk aversion that affects the equilibrium as discussed in the previous section. And even in the simple deterministic case, energy prices might need to be very high on occasions to sustain the peaking investment needed to make the system avoid shedding load. For example if load shedding is acceptable in at most four or five hours per year, then prices need to become very high to pay for the annual capital cost of a peaking plant that runs only during these periods. The uncertainty of receiving these cash flows every year makes such an investment too risky.

Contracts between energy suppliers can resolve some of the risks faced by generators in deciding capacity investments. For example, a hydroelectric generator could arrange a two-way option contract with a coal plant to keep the coal plant available for periods of low reservoir inflows. The hydroelectric generator buys a call option off the coal plant, and the coal plant buys a put option (at a lower strike price) from the hydro generator. These contracts (that can be arranged to have the same price) enable the coal plant to receive revenue even when wholesale prices are below its marginal cost of generation in return for some loss of revenue in peak periods.

Capacity markets that arrange additional payments for committed generation capacity ahead of time are a popular mechanism intended to overcome these problems. Opinions differ on the effectiveness of these mechanisms in comparison with energy-only markets, and studying their design and operation is an active area of research.

In dealing with the transition to green energy, capacity markets serve to answer the second question as they can procure the desired capacity of different energy technologies at auction. So governments can decide to increase this as needed to meet demand growth. It is not clear whether the same outcome might be achieved at lower cost with an energy-only solution.

Large-scale expensive new generation capacity expansion requires electricity demand to pay for it. One source of this is private, behind-the-meter specialized applications such as data centers or large industrial concerns. One might expect to see generation capacity expansion occur behind the meter backed by such demand.

The final question of timing is important. A green-energy risked equilibrium must be viewed over a long time scale and achieve a green energy system in time to avert a climate catastrophe. Dynamic equilibrium models might give some confidence that commercial investment will deliver in time, but betting the planet's future on this might be too risky for policy makers. As evidence of climate change becomes more obvious, generational shifts in voter preferences might lead to more direct government intervention in planning and implementing the transition. In this case, relying on competitive electricity markets to achieve the transition might be viewed by governments as too much of a risk.

5. Conclusions. In this paper we have outlined some of the questions arising in the transition to green energy, and presented some mathematical approaches to address them. The models we discuss are formulations of optimization problems and related complementarity problems, in settings with a variety of physical scales, and dealing with different time scales. The costs of the physical and institutional architecture required to bring about the transition will be substantial and will involve risk. Mathematical models will be essential in understanding the complex tradeoffs that have to be made in planning and incentivizing the transition to enable it to occur at a low cost and in time to avoid global temperatures rising to unacceptable levels.

Acknowledgments. We would like to acknowledge the contributions of Laura Diaz Anadon, Dennice Gayme, Eddie Anderson and Michel De Lara to this document.

REFERENCES

[1] *Green protectionism will slow the energy transition*, Economist, (December 9, 2023).
[2] *The renewables business faces a make-or-break moment*, Economist, (December 9, 2023).
[3] *How carbon prices are taking over the world*, Economist, (October 1, 2023).
[4] I. Abada and A. Ehrenmann, *When market incompleteness is preferable to market power: Insights from power markets*, Available at SSRN 4402888, (2023).
[5] V. Ahlqvist, P. Holmberg, and T. Tangerås, *A survey comparing centralized and decentralized electricity markets*, Energy Strategy Reviews, 40 (2022), p. 100812.
[6] K. Amarasinghe, D. L. Marino, and M. Manic, *Deep neural networks for energy load forecasting*, in 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE), IEEE, June 2017, https://doi.org/10.1109/isie.2017.8001465.
[7] E. J. Anderson, *Interconnection models*, 2024. Presentation at ISMP 2024, Montreal, Canada.
[8] M. F. Anjos, I. Bouras, L. Brotcorne, A. G. Weldeyesus, C. Alasseur, and R. Zorgati, *Integrated Location, Sizing, and Pricing for EV Charging Stations*, Springer Nature

Switzerland, 2024, pp. 431–448, https://doi.org/10.1007/978-3-031-57603-4_18.

[9] M. F. Anjos, B. Gendron, and M. Joyce-Moniz, *Increasing electric vehicle adoption through the optimal deployme nt of fast-charging stations for local and long-distance travel*, European Journal of Operational Research, 285 (2020), pp. 263–278, https://doi.org/10.1016/j.ejor.2020.01.055.

[10] P. Barbrook-Johnson, J.-F. Mercure, S. o. Sharpe, C. Peñasco, C. Hepburn, L. D. Anadon, J. D. Farmer, and T. M. Lenton, *Economic modelling fit for the demands of energy decision makers*, Nature Energy, (2024), https://doi.org/10.1038/s41560-024-01452-7.

[11] P. Benner, S. Gugercin, and K. Willcox, *A survey of projection-based model reduction methods for parametric dynamical systems*, SIAM Review, 57 (2015), pp. 483–531, https://doi.org/10.1137/130932715.

[12] D. Bertsekas and S. Shreve, *Stochastic optimal control: the discrete-time case*, vol. 5, Athena Scientific, 1996.

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

[14] C. Böhringer and T. F. Rutherford, *Integrating bottom-up into top-down: A mixed complementarity appr oach*, SSRN Electronic Journal, (2005), https://doi.org/10.2139/ssrn.770725.

[15] S. Burger, J. P. Chaves-Ávila, C. Batlle, and I. J. Pérez-Arriaga, *A review of the value of aggregators in electricity systems*, Renewable and Sustainable Energy Reviews, 77 (2017), pp. 395–405, https://doi.org/10.1016/j.rser.2017.04.014.

[16] M. Chatzos, T. W. K. Mak, and P. V. Hentenryck, *Spatial network decomposition for fast and scalable ac-opf learni ng*, IEEE Transactions on Power Systems, 37 (2022), pp. 2601–2612, https://doi.org/10.1109/tpwrs.2021.3124726.

[17] S. Cloete, *The 10 big problems with simply replacing fossil cars with electric*, Dec. 2021, https://energypost.eu/the-10-big-problems-with-simply-replacing-fossil-cars-with-electric/.

[18] T. Copeland and V. Antikarov, *Real Options: A Practitioner's Guide*, Texere, 2001, https://books.google.co.nz/books?id=fnhPAAAAMAAJ.

[19] R. Cory-Wright, A. Philpott, and G. Zakeri, *Payment mechanisms for electricity markets with uncertain supply*, Operations Research Letters, 46 (2018), pp. 116–121.

[20] R. de Neufville and S. Scholtes, *Flexibility in Engineering Design*, The MIT Press, 2011, https://doi.org/10.7551/mitpress/8292.001.0001.

[21] T. De Rubira and A. Wigington, *Extending complementarity-based approach for handling voltage band regulation in power flow*, in 2016 Power Systems Computation Conference (PSCC), IEEE, 2016, pp. 1–6.

[22] P. Denholm, E. Ela, B. Kirby, and M. Milligan, *The Role of Energy Storage with Renewable Electricity Generation*, Office of Scientific and Technical Information (OSTI), Jan. 2010, pp. 1–61, https://doi.org/10.2172/972169.

[23] P. Denholm, J. Nunemaker, P. Gagnon, and W. Cole, *The potential for battery energy storage to provide peaking capacity in the United States*, Renewable Energy, 151 (2020), pp. 1269–1277, https://doi.org/10.1016/j.renene.2019.11.117.

[24] A. Downward, R. Baucke, and A. Philpott, *JuDGE.jl: a Julia package for optimizing capacity expansion*, Optimization Online, (2020), https://optimization-online.org/2020/11/8086/.

[25] O. Dowson and L. Kapelevich, *SDDP.jl: a Julia package for stochastic dual dynamic programming*, INFORMS Journal on Computing, 33 (2021), pp. 27–33.

[26] B. Eldridge, B. Knueven, and J. Mays, *Rethinking the price formation problem–part 1: Participant incentives under uncertainty*, IEEE Transactions on Energy Markets, Policy and Regulation, (2023).

[27] A. Fattahi, J. Sijm, and A. Faaij, *A systemic approach to analyze integrated energy system modeling tools: A review of national models*, Renewable and Sustainable Energy Reviews, 133 (2020), p. 110195.

[28] Federal Energy Regulatory Commission, *Order no. 2222: Participation of distributed energy resource aggregations in markets operated by regional transmission organizations and independent system operators*, 2020, https://www.ferc.gov/sites/default/files/2020-09/E-1_0.pdf.

[29] M. J. Fell, *Energy services: A conceptual review*, Energy Research & Social Science, 27 (2017), pp. 129–140, https://doi.org/10.1016/j.erss.2017.02.010.

[30] M. Ferris and A. Philpott, *Dynamic risked equilibrium*, Operations Research, 70 (2022), pp. 1933–1952.

[31] M. Ferris and A. Philpott, *On chance constraints in practice*, In preparation, (2024).

[32] M. C. Ferris and J. S. Pang, *Engineering and Economic Applications of Comple-*

mentarity Problems, SIAM Review, 39 (1997), pp. 669–713, https://doi.org/10.1137/S0036144595285963.

[33] M. C. Ferris and A. Philpott, *Renewable electricity capacity planning with uncertainty at multiple scales*, Computational Management Science, 20 (2023), https://doi.org/10.1007/s10287-023-00472-0. 10.21203/rs.3.rs-2207695/v1.

[34] E. B. Fisher, R. P. O'Neill, and M. C. Ferris, *Optimal Transmission Switching*, IEEE Transactions on Power Systems, 23 (2008), pp. 1346–1355, https://doi.org/10.1109/TPWRS.2008.922256, http://www.cs.wisc.edu/~ferris/papers/transmission.pdf.

[35] S. A. Gabriel, A. J. Conejo, B. F. Fuller, J. David an d Hobbs, and C. Ruiz, *Complementarity Modeling in Energy Markets*, Springer New York, 2013, https://doi.org/10.1007/978-1-4419-6123-5.

[36] Z. Gao, K. Alshehri, and J. R. Birge, *Aggregating distributed energy resources: Efficiency and market power*, Manufacturing and Service Operations Management, (2024), https://doi.org/10.1287/msom.2021.0539.

[37] H. Gérard, V. Leclère, and A. Philpott, *On risk averse competitive equilibrium*, Operations Research Letters, 46 (2018), pp. 19–26.

[38] S. Gössling and C. Lyle, *Transition policies for climatically sustainable aviation*, Transport Reviews, 41 (2021), pp. 643–658, https://doi.org/10.1080/01441647.2021.1938284.

[39] S. J. Grossman and O. D. Hart, *An analysis of the principal-agent problem*, Econometrica, 51 (1983), p. 7, https://doi.org/10.2307/1912246.

[40] J. Han and A. Papavasiliou, *The impacts of transmission topology control on the European electricity network*, IEEE Transactions on Power Systems, 31 (2016), pp. 496–507, https://doi.org/10.1109/tpwrs.2015.2408439.

[41] K. W. Hedman, M. C. Ferris, R. P. O'Neill, E. B. Fisher, and S. S. Oren, *Co-optimization of Generation Unit Commitment and Transmission Switching with N-1 Reliability*, IEEE Transactions on Power Systems, 25 (2010), pp. 1052–1063, https://doi.org/10.1109/TPWRS.2009.2037232, http://www.cs.wisc.edu/~ferris/papers/Hedman{_}GenUC.pdf.

[42] W. Hogan, *Electricity market design: Multi-interval pricing models*, https://scholar. harvard.edu/files/whogan/files/hogan_hepg_multi_period_, 62220 (2020).

[43] C. Hohl, C. Prete, and M. Radhakrishnan, A.and Webster, *Intraday markets, wind integration and uplift payments in a regional us power system*, Energy Policy, 175 (2023), p. 113503.

[44] H. Hoschle, H. L. Cadre, Y. Smeers, A. Papavasiliou, and R. Belmans, *An ADMM-based method for computing risk-averse equilibrium in capacity markets*, IEEE Transactions on Power Systems, (2018), https://doi.org/10.1109/tpwrs.2018.2807738.

[45] O. Huber and M. C. Ferris, *Reformulations for convex composite functions and their nested compositions*, tech. report, University of Wisconsin, 2024.

[46] S. Iniyan, L. Suganthi, and A. A. Samuel, *Energy models for commercial energy prediction and substitution o f renewable energy sources*, Energy Policy, 34 (2006), pp. 2640–2653, https://doi.org/10.1016/j.enpol.2004.11.017.

[47] J. Iria and F. Soares, *An energy-as-a-service business model for aggregators of prosumers*, Applied Energy, 347 (2023), p. 121487, https://doi.org/10.1016/j.apenergy.2023.121487.

[48] G. James, D. Witten, T. Hastie, and R. T. hirani, *An Introduction to Statistical Learning*, Springer New York, 2013, https://doi.org/10.1007/978-1-4614-7138-7.

[49] S. Jebaraj and S. Iniyan, *A review of energy models*, Renewable and Sustainable Energy Reviews, 10 (2006), pp. 281–311, https://doi.org/10.1016/j.rser.2004.09.004.

[50] W. Jeon, A. Lamadrid, and T. Mount, *The economic value of distributed storage at different locations on an electric grid*, The Energy Journal, 40 (2019).

[51] J. Johnston, *Problems of equity and efficiency in the design of international greenhouse gas cap-and-trade schemes*, Harv. Envtl. L. Rev., 33 (2009), p. 405.

[52] Y. Kim and M. C. Ferris, *Solving equilibrium problems using extended mathematical programming*, Mathematical Programming Computation, 11 (2019), pp. 457–501, https://doi.org/10.1007/s12532-019-00156-4.

[53] Y. Kim, O. Huber, and M. C. Ferris, *A structure-preserving pivotal method for affine variational inequalities*, Mathematical Programming Computation, 168 (2017), pp. 93–121, https://doi.org/10.1007/s10107-017-1124-9.

[54] B. Kocuk, H. Jeon, S. S. Dey, J. Linderoth, J. Luedtke, and A. Sun, *A cycle-based formulation and valid inequalities for DC power transmission problems with switching*, Operations Research, 64 (2016), pp. 928–938.

[55] C. Kok, A. Philpott, and G. Zakeri, *Value of transmission capacity in electricity markets with risk averse agents*, Tech. Report www.epoc.org.nz/papers/TransmissionPaperOperationsResearch.pdf, EPOC Work-

ing Paper, 2018.

[56] L. B. N. Laboratory, *Grid connection backlog grows by 30dominated by requests for solar, wind, and energy storage*, 2024, https://emp.lbl.gov/news/grid-connection-backlog-grows-30-2023-dominated-requests-solar-wind-and-energy-storage.

[57] C. Li, A. Conejo, P. Liu, B. Omell, J. Siirola, and I. Grossmann, *Mixed-integer linear programming models and algorithms for generation and transmission expansion planning of power systems*, European Journal of Operational Research, 297 (2022), pp. 1071–1082.

[58] H. Markowitz, *Portfolio selection*, The Journal of Finance, 7 (1952), pp. 77–91.

[59] J.-F. Mercure, S. Sharpe, J. Vinuales, M. Ives, M. Grubb, A. Lam, P. Drummond, H. Pollitt, F. Knobloch, and F. J. Nijsse, *Risk-opportunity analysis for transformative policy design and appraisal*, Global Environmental Change, 70 (2021), p. 102359.

[60] P. Moutis and O. Alizadeh-Mousavi, *Digital twin of distribution power transformer for real-time monitoring of medium voltage from low voltage measurements*, IEEE Transactions on Power Delivery, 36 (2021), pp. 1952–1963, https://doi.org/10.1109/tpwrd.2020.3017355.

[61] F. Murphy, A. Pierru, and Y. Smeers, *Measuring the effects of price controls using mixed complementarity models*, European Journal of Operational Research, 275 (2019), pp. 666–676, https://doi.org/10.1016/j.ejor.2018.11.051.

[62] P. Nahmmacher, E. Schmid, L. Hirth, and B. Knopf, *Carpe diem: A novel approach to select representative days for long-term power system modeling*, Energy, 112 (2016), pp. 430–442, https://doi.org/10.1016/j.energy.2016.06.081.

[63] N. Nazir and M. Almassalkhi, *Grid-aware aggregation and realtime disaggregation of distributed energy resources in radial networks*, IEEE Transactions on Power Systems, 37 (2022), pp. 1706–1717, https://doi.org/10.1109/tpwrs.2021.3121215.

[64] O. Nohadani and K. Sharma, *Optimization under decision-dependent uncertainty*, SIAM Journal on Optimization, 28 (2018), pp. 1773–1795, https://doi.org/10.1137/17m1110560.

[65] I. K. Nti, M. Teimeh, O. Nyarko-Boateng, and A. F. Adekoya, *Electricity load forecasting: a systematic review*, Journal of Electrical Systems and Information Technology, 7 (2020), https://doi.org/10.1186/s43067-020-00021-8.

[66] Ö. Okur, P. Heijnen, and Z. Lukszo, *Aggregator's business models in residential and service sectors: A review of operational and financial aspects*, Renewable and Sustainable Energy Reviews, 139 (2021), p. 110702, https://doi.org/10.1016/j.rser.2020.110702.

[67] A. Pfenninger, S.and Hawkes and J. Keirstead, *Energy systems modeling for twenty-first century energy challenges*, Renewable and Sustainable Energy Reviews, 33 (2014), pp. 74–86.

[68] A. Philpott and G. Everett, *On load shedding and transmission grid security*, tech. report, downloadable from https://www.epoc.org.nz/papers/SecurityPaper.pdf, 2004.

[69] A. Philpott, M. Ferris, and J. Mays, *Electricity dispatch and pricing under uncertainty*, In preparation, (2024).

[70] A. B. Philpott, M. C. Ferris, and R. J. B. Wets, *Equilibrium, uncertainty and risk in hydro-thermal electricity systems*, Mathematical Programming B, 157 (2016), pp. 483–513, https://doi.org/10.1007/s10107-015-0972-4.

[71] W. Powell, *A unified framework for stochastic optimization*, European Journal of Operational Research, 275 (2019), pp. 795–821.

[72] W. B. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensiona lity*, Wiley, Aug. 2011, https://doi.org/10.1002/9781118029176.

[73] D. Pozo, *Linear battery models for power systems analysis*, Electric Power Systems Research, 212 (2022), p. 108565, https://doi.org/10.1016/j.epsr.2022.108565.

[74] D. Pozo, E. Sauma, and J. Contreras, *A three-level static MILP model for generation and transmission expansion planning*, IEEE Transactions on Power systems, 28 (2012), pp. 202–210.

[75] H. Rahimian, G. Bayraksan, and T. Homem-de Mello, *Identifying effective scenarios in distributionally robust stochastic programs with total variation distance*, Mathematical Programming, 173 (2018), pp. 393–430, https://doi.org/10.1007/s10107-017-1224-6.

[76] D. Ralph and Y. Smeers, *Risk trading and endogenous probabilities in investment equilibria*, SIAM Journal on Optimization, 25 (2015), pp. 2589–2611.

[77] R. Rossmann, M. Anitescu, J. Bessac, M. Ferris, M. Krock, J. Luedtke, and L. Roald, *A framework for balancing power grid resiliency and efficiency with bi-objective stochastic integer optimization*, tech. report, University of Wisconsin, 2024.

[78] N. Sepulveda, J. Jenkins, A. Edington, D. Mallapragada, and R. Lester, *The design space for long-duration energy storage in decarbonized power systems*, Nature Energy, 6 (2021), pp. 506–516.

[79] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on stochastic programming: mod-*

eling and theory, SIAM, 2021.

[80] J. Shen, *Theory and Computation of the Multistage Stochastic Equilibria with Risk-Averse Players*, PhD thesis, University of Wisconsin –Madison, 2023.

[81] K. Singh, A. Philpott, and K. Wood, *Column generation for design of survivable networks*, tech. report, Electric Power Optimization Centre, 2006, https://www.epoc.org.nz/papers/SinghPhilpottWood.pdf.

[82] R. Sioshansi, P. Denholm, T. Jenkin, and J. Weiss, *Estimating the value of electricity storage in PJM: Arbitrage and some welfare effects*, Energy Economics, 31 (2009), pp. 269–277.

[83] S. Stoft, *Power System Economics*, IEEE Press, Piscataway, 2002.

[84] B. Stott, J. Jardim, and O. Alsac, *DC power flow revisited*, IEEE Transactions on Power Systems, 24 (2009), pp. 1290–1300, https://doi.org/10.1109/tpwrs.2009.2021235, http://ieeexplore.ieee.org/xpls/abs{_}all.jsp?arnumber=4956966.

[85] S.-B. Tsai, Y. Xue, J. Zhang, Q. Chen, Y. Liu, J. Zhou, and W. Dong, *Models for forecasting growth trends in renewable energy*, Renewable and Sustainable Energy Reviews, 77 (2017), pp. 1169–1178, https://doi.org/10.1016/j.rser.2016.06.001.

[86] W. Tushar, T. K. Saha, C. Yuen, and D. S. a nd H. Vincent Poor, *Peer-to-peer trading in electricity networks: An overview*, IEEE Transactions on Smart Grid, 11 (2020), pp. 3185–3200, https://doi.org/10.1109/tsg.2020.2969657.

[87] A. Velloso and P. Van Hentenryck, *Combining deep learning and optimization for security-constrained optimal power flow*, 2020, https://doi.org/10.48550/ARXIV.2007.07002.

[88] J. Villumsen, G. Brønmo, and A. Philpott, *Line capacity expansion and transmission switching in power systems with large-scale wind power*, IEEE Transactions on Power Systems, 28 (2012), pp. 731–739.

[89] R. Way, M. C. Ives, P. Mealy, and J. D. Farmer, *Empirically grounded technology forecasts and the energy transition*, Joule, 6 (2022), pp. 2057–2082, https://doi.org/10.1016/j.joule.2022.08.009.

[90] P. Weber and M. Woerman, *Intermittency or uncertainty? Impacts of renewable energy in electricity markets*, SSRN Electronic Journal, (2022), https://doi.org/10.2139/ssrn.4212066.

[91] L. Wen, K. Zhou, S. Yang, and X. Lu, *Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting*, Energy, 171 (2019), pp. 1053–1065, https://doi.org/10.1016/j.energy.2019.01.075.

[92] N. Winchester and D. White, *The climate policy analysis (C-PLAN) model, version 1.0*, Energy Economics, 108 (2022), p. 105896.

[93] S. Wogrin, D. Tejada-Arango, A. Downward, and A. Philpott, *Welfare-maximizing transmission capacity expansion under uncertainty*, Philosophical Transactions of the Royal Society A, 379 (2021), p. 20190436.

[94] X. Xi, R. Sioshansi, and V. Marano, *A stochastic dynamic programming model for co-optimization of distributed energy storage*, Energy Systems, 5 (2014), pp. 475–505.