

# Adaptive Stochastic Control for the Smart Grid

Roger N. Anderson, IEEE member, Albert Boulanger, Warren B. Powell, IEEE member, and Warren Scott

**Abstract**—Approximate Dynamic Programming driven Adaptive Stochastic Control for the Smart Grid holds the promise of providing the autonomous intelligence required to elevate the electric grid to efficiency and self-healing capabilities more comparable to the Internet. To that end, we demonstrate the load and source control necessary to optimize management of distributed generation and storage within the Smart Grid.

**Index Terms**—Smart Grid, Adaptive Stochastic Control, Approximate Dynamic Programming, Control Systems.

## I. INTRODUCTION

**A**UTONOMOUS Control Systems for field operations such as at Electric Utilities and Independent System Operators, and especially for the Smart Grid, are more difficult than those required to control indoor and site-specific systems (e.g. factory assembly lines, petrochemical plants, and nuclear power plants). Below we describe such an Adaptive Stochastic Control (ASC) system for load and source management of real-time Smart Grid operations.

Electric utilities operate in a difficult, outdoor environment that is dominated by stochastic (statistical) variability, primarily driven by the vagaries of the weather and by equipment failures. Within the Smart Grid, advanced dynamic control will be required for simultaneous management of real time pricing, curtailable loads, Electric Vehicle recharging, solar, wind and other distributed generation sources, many forms of energy storage, and microgrid management (Fig. 1).

Computationally, controlling the Smart Grid is a multi-stage, time-variable, stochastic optimization problem. ASC using Approximate Dynamic Programming (ADP) offers the capability of achieving autonomous control using a computational learning system to manage the Smart Grid. Within the complexities of the Smart Grid (Fig. 1), ADP driven ASC is used as a decomposition strategy that breaks the problem of continuous Smart Grid management, with its long

Authors contributed equally:

R. N. Anderson and A. Boulanger are with the Center for Computational Learning Systems, Columbia University, NY, NY 10027. Their work is supported in part by Consolidated Edison of New York, Inc. and the Department of Energy through American Recovery and Reinvestment Act of 2009 contract E-OE0000197 by way of sub-award agreement SA-SG003.

W. B. Powell and W. Scott are with the Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544. Their work is supported in part by the Air Force Office of Scientific Research, grant number FA9550-08-1-0195 and the National Science Foundation, grant CMMI-0856153.

time horizons, into a series of short-term problems that a Mixed-Integer Nonlinear Programming solver can handle with sufficient speed and computational efficiency to make it practical for system-of-systems control.

In this paper, we consider a specific application for distributed electricity dispatch involving a multidimensional control variable (the flow of energy from different sources to serve different loads that we term “load and source control”), where for each time period, the distributed generation is linked to a storage device. We describe an ADP algorithm for solving this ASC problem with hundreds or thousands of variables, and demonstrate that the ADP solution produces results that are extremely close to optimal. We then address the problem of energy storage (e.g. in a large battery) in the presence of a more complex “state of the world” variable. In this problem, the state of the system includes not only the energy stored in the battery, but also variations in wind, load demand and electricity prices. These experiments demonstrate both the potential of ADP for solving high dimensional energy allocation problems but also some potential pitfalls (such as using Approximate Policy Iteration with poorly chosen basis functions). These results are important for the use of ADP for any energy optimization problem.

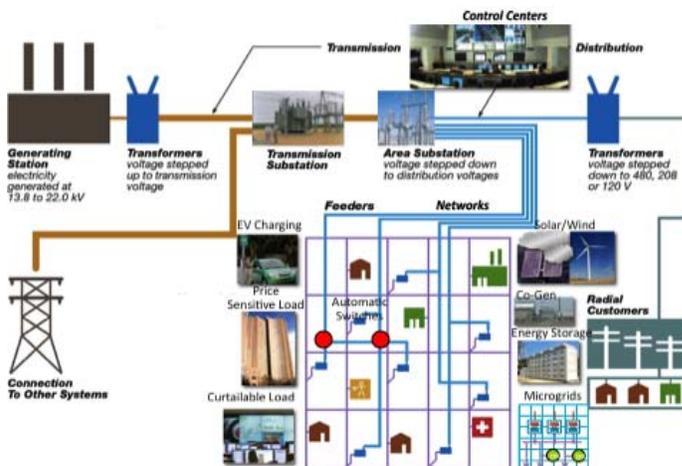


Fig. 1. The Adaptive Stochastic Control of the Smart Grid must simultaneously optimize supply and demand from many new, distributed loads and sources such as: price sensitive and curtable loads, intermittent solar and wind generation, distributed energy storage, EV charging, and microgrids. (Source: Modified after Con Edison drawing).

## II. CONTROL OF THE SMART GRID

Within the Smart Grid, any control technology must automate energy management so that real-time data is converted to information fast enough so that problems are

diagnosed instantly, corrective actions are identified and executed dynamically in the field, and feedback loops provide metrics that verify that the work done is producing the desired effects. Our view of Adaptive Stochastic Control requires the following characteristics:

- Self-healing: automatic repair or removal of potentially faulty equipment from service before it fails, and reconfiguration of the system to reroute supplies of energy to sustain power to all customers,
- Flexible: rapid and safe interconnection of distributed generation and energy storage at any point in the system at any time,
- Predictive: use of statistics, machine learning, adaptive algorithms, and predictive models (for example weather impact projections) to provide the next most likely events so that appropriate actions are taken to reconfigure the system before next worst events can happen,
- Interactive: appropriate information is provided transparently regarding the status of the system in near real time,
- Optimal: both Smart Grid operators and customers act to allow all key participants in the energy system to most efficiently and economically manage contingencies with environmentally sound actions,
- Secure: cyber- and physical-security, so that two-way communications protect all critical assets of the Smart Grid.

### III. MAJOR NEW COMPONENTS OF THE SMART GRID

In order to autonomously control the Smart Grid, it will be necessary to optimally manage new, intelligent equipment at all critical transmission, distribution, and consumption points. It is our view that for this new intelligence to become an effective part of the operations of an integrated Smart Grid system, control technologies must be integrated into an Adaptive Stochastic Control system. The ASC optimizes load and source management within a system-of-systems that provides secure communications, efficient data management, diagnostic analysis, and work management integration [1]. The Smart Grid must operate as an integrated machine that simultaneously controls *at least* the following new technologies that are briefly described below.

#### A. AMI's, Demand Response and Curtailable Loads

Many people, especially in the public sector, consider the Smart Grid to be only Advanced Metering Infrastructure (AMI). Such systems provide 2-way consumption control at the customer site, as well as, distributed load management, and customer communications at the utility site [2]. An extension of AMI is the Home Area Network (HAN) that additionally provides demand response functionality such as automated control of refrigerators, air conditioners, thermostats, and home entertainment systems. In addition, many utilities, energy services and aggregator companies provide automated curtailment programs through subscription services. When controlled by ADP algorithms, self-healing capabilities more common to the Internet can potentially be built into automated reconfiguration regimes when

information is passed through such curtailment programs, [3,4].

#### B. Flexible Power Electronics

Other classes of Smart Grid devices that must be optimally managed are power flow routers such as fault current limiters, sectionalizing switches, FACTS devices [5], and Smart Wires [6]. For example, FACTS devices can be used to route power around load congestion. The Smart Grid must manage these internet-like “routers” along with the only present alternative, incentive-based nodal pricing in states with competitive, real-time markets [7].

#### C. Photovoltaics and Solar Heating

Photovoltaics (PV) provide local load relief for the Smart Grid. However, the inherent unpredictability caused by cloud cover variations makes the certainty of fixed quantities of power impossible, unless distributed storage is coupled with the PV systems. That said, entire countries depend upon solar heating for hot water subsystems for all homes, such as in Cyprus. PV and solar heating are fundamentally a curtailment service whereby grid electricity is replaced by PV locally available to a home or business. Also, small amounts of power can be sent back into the grid to relieve load in a local area [8].

#### D. Recharging Electric Vehicles

Electric Vehicles and Plug-in Hybrid Electric Vehicles (grouped as EV's here) present unique problems for Smart Grid control because they are mobile sinks for power in the day and fixed sinks at night [9]. ASC management of EV charging is mostly needed during the day in large urban areas, when large populations of EV's will plug into the grid upon arrival at work, just as the electricity consumption is ramping up towards peak loads and electric transportation systems such as subways are in their morning rush hours. A further homeland security requirement will likely be that each EV must receive at least a partial recharge so all vehicles can make it out of the city in case of an emergency. Thus, load transfer to storage facilities linked to EV charging stations is needed in addition to grid charging to manage such variable demand.

“Green Garages” are beginning to appear in cities like New York. They certify that the power used to charge EV's comes from renewable energy sources, often on roofs. Also, EV's could represent a significant mobile source of emergency power in case of crisis situations such as blackouts. These Vehicle to Grid technologies (V2G) could then provide additional power particularly to nearby homes. Many countries are promoting EV use that will drive market penetration, such as the introduction of laws like “The Electric Drive Vehicle Deployment Act” of 2010<sup>1</sup> in the United States. Such mobile load and source complexities must be managed within the ASC.

<sup>1</sup> c.f.:

[http://markey.house.gov/index.php?option=com\\_content&task=view&id=4006&Itemid=141](http://markey.house.gov/index.php?option=com_content&task=view&id=4006&Itemid=141) .

### E. Microgrids

Microgrids are small scale, largely independent grids that remain connected to the Smart Grid. Within microgrids, distributed generation sources such as PV and wind, along with distributed generators, are linked to distributed storage and EV recharging stations to provide a self-sustaining local grid. They provide local electric distribution for a neighborhood, campus, military base, or manufacturing facility that can be independently “islanded” from the grid in emergencies. Microgrids also include local load and source control using Building Management Systems (BMS) and often power Heating Ventilation and Air Conditioning (HVAC) of large groups of buildings. Microgrids are designed to be able to stand alone from the electric grid (thus the islanding) in times of crisis so that the power in the area can be maintained via local generation. Microgrids can also be sites of significant curtailable load for utilities during critical load relief periods of peak demand [10]. The ASC must be cognizant of financial and market valuations critical to the benefits of having a microgrid in the first place [11], [12].

### F. Energy Storage

A critical addition to the Smart Grid control solution comes from the addition of significant energy storage capability. Intermittent power sources like PV, Solar Thermal, and Wind require some place to store the electricity to fill needs during cloudy and/or windless times. The Electricity Storage Organization tracks the cost of both large and small scale energy storage systems, from Lithium-Ion, Nickel-Cadmium and Lead-Acid batteries, through fly wheels and super-capacitors, to various large scale battery storage devices, and finally to large scale cavern storage of compressed air and

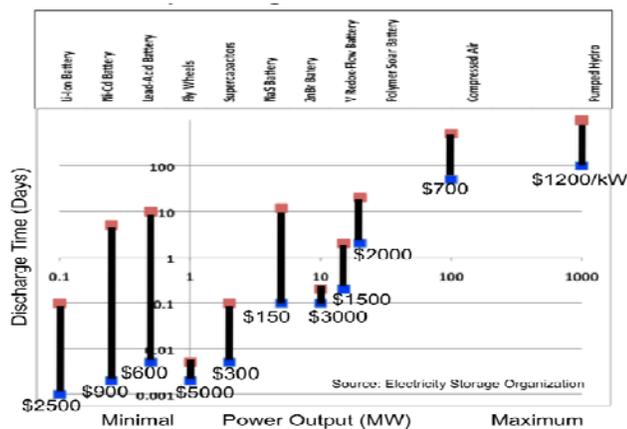


Fig. 2. The relative power output, discharge time, and cost per kWh for various Energy storage devices.

hydroelectric storage that involves pumping water back upstream during nights (Fig. 2).

In addition, other electricity storage devices such as those that melt salt, heat vegetable oils, freeze ice, and use fuel cells have attained widespread but limited, deployment. All these storage technologies are viable, if affordable and controllable: barriers that have not yet been fully conquered. However, their certain entry into alternative energy systems make PV, wind, EV recharging, and microgrids manageable.

### G. Distributed Generation

The Smart Grid also must be able to control small-scale generation owned by customers. Facilities such as combined-heat-and-power (CHP) co-generation and emergency diesel generators will be managed along with PV, EV and microgrid sources and storage facilities in order to preserve adequate power margins at all times.

### H. Storm Management

The key exogenous variable is weather, and its corollary is accurate storm forecasting. Weather is the principal forcing function driving the uncertainties that must be optimized by all Smart Grid control systems. New methods linking these erratic sources to storage are required if we are to treat these renewable sources and sinks as dispatchable loads [13], [14]. Experiments with ADP control of such distributed load and source combinations will be presented in Sections V and VI below.

### I. Massive Solar Thermal and Wind Generation Facilities

Solar thermal power generation facilities have been very successful in linking large arrays of mirrors that focus the sun’s energy into a storage medium, usually a salt that is melted or a vegetable oil that is heated. The heat storage medium can be used to power steam generators to produce electricity for many hours after sunset. This combination has allowed the design of very large solar thermal power plants. Similarly, national visions of a hydrocarbon free future of energy independence have led to the installation of gigantic fields of wind turbine power generation in several countries across the globe. It is theoretically possible to control the input from many such plants distributed across large deserts and seas so that as much electricity could be generated from this source as from nuclear and hydro electric power plants. For example, Arizona has begun construction of the first 280 MW of an intended 4300 MW solar thermal plant south of Phoenix.

A successful ASC for the Smart Grid must combine predictive capabilities for cloud cover and strong but erratic winds in areas of solar and wind generation plants, such as those in Arizona, West Texas and the North Sea, with large energy storage facilities. Compressed Air Energy Storage (CAES) facilities in underground caverns or emptied natural gas reservoirs are realistic examples. Swider [15] has demonstrated the economic market modeling needed to justify the combined investment of wind generators with CAES. Payback is minimized only if the laying of regional transmission lines needed to get the power to market is part of the up-front investment. This was a hard lesson learned in West Texas where, as much as half of the 2000+ MW of wind power is dormant at any given time because of transmission limitations [16], [17].

### J. Nanotechnologies

Above all, controllers for the Smart Grid must have the capacity to adapt to new technologies not yet invented, or in long-term development, such as nuclear fusion, or more likely, nanotechnologies. Smalley [18] has presented examples of future nanotechnologies that will likely be important

distributed energy sources and storage media within the next 10 years, including:

- Nanophotovoltaics that may drop PV costs by 100 fold or more,
- Nanophoto catalysts that reduce CO<sub>2</sub> emissions during the formation of methanol,
- Nanothermochemical catalysts that directly convert light and water to hydrogen to work efficiently at temperatures lower than 900 degrees C,
- Nanofuel cells that drop the cost by 10-100x and provide low temperature starting capacity that is reversible,
- Nanobatteries and super-capacitors, that along with low friction nanoflywheels will improve efficiency by 10-100x for transportation and distributed generation applications,
- Nanoelectronics that produce nanocomputers, and nanosensors for better SCADA systems
- Nanolighting to replace incandescent, fluorescent and LED,
- Nanopaints for the exterior of buildings that generate electricity, and ultimately,
- Quantum wires (QW) that might rewire the transmission grid and enable continental, and even worldwide electricity transport by replacing copper and aluminum transmission wires.

Perhaps the most promising of these nanotechnologies for the Smart Grid are Quantum Wires that will have the electrical conductivity of copper at one-sixth the weight, but a strength beyond Kevlar. QW can be spun into polypropylene-like “rope” and used for transmission lines of the future. This “Fullerene tube” rope will form a super-material of extreme strength, lightness, high temperature resistance, and unidirectional thermal conductivity (electrons just fit into each tube, and so have only one place to go), but they also “magically” quantum-jump from one tube to the next [19], [20].

#### IV. ADAPTIVE STOCHASTIC CONTROL

We propose that the key to the successful implementation of the Smart Grid is to create the ASC management system for

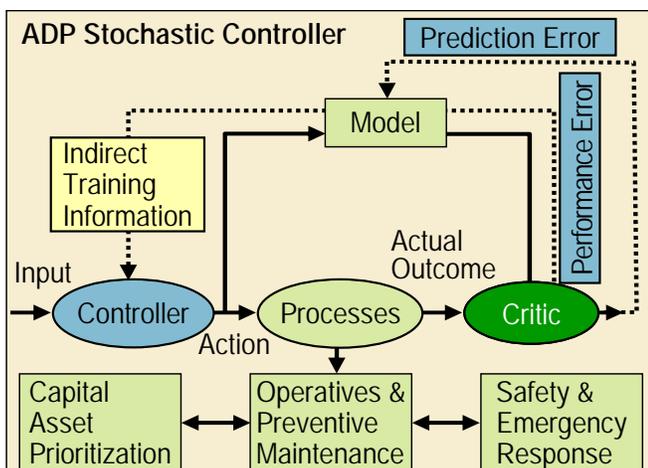


Fig. 3. The feedback loops for an Adaptive Stochastic Controller for the Smart Grid optimally interprets incoming data from many new distributed sources and simultaneously manages asset prioritization, operational actions, maintenance tasks, and emergency responses.

control of the electric grid. The ASC will be able to optimize amongst all combinations of loads and sources above, and in even more unimaginable combinations, and at all points along the Smart Grid. A tall task indeed

In order to make this vision a reality, the ASC must receive, interpret, and act on all manner of new data coming from SCADA sources throughout the grid (Fig. 3). It will send commands to manage contingencies and optimize power flow, initiate preventive maintenance, control switching, minimize loads and optimize capital investment, all the while dealing with erratic solar and wind generation and distributed storage, equipment failures and weather variations.

Utilities now use complex, computationally driven, command and control systems like the ASC only in nuclear power plant management. However, these systems are particularly focused on preventive maintenance and identification of out-of-normal operational performance. They are very good at identifying the “next worst” condition that can happen to the plant at any given time, but they are not so good at determining the “next most likely” condition to occur within the facility. The ASC for the Smart Grid must do both.

In Operations Research, control of such systems presents an extremely complex multi-stage, time-variant, stochastic optimization problem. An ASC requires the use of algorithms that perform complex mathematics using model simulations of the future in near real-time. That is, ADP solvers are needed that are more familiar to the military, petrochemical and transportation industries.

In the utility industry, only the Independent System Operators use such complex control algorithms, and then only for economic dispatch of power. For the Smart Grid, the electricity industry will have to successfully adapt these advanced ADP control algorithms for the distributed distribution of electricity or the system will risk catastrophic failure. For example, future distribution control rooms will be required to manage the *margin* between local loads and multi-owner sources. Margin is currently managed only at the transmission level. We predict that economic benefits will be substantial when distribution control centers also manage margin. Significant economic gains have been measured after transition to similar autonomous, adaptive system-of-systems control in many other industries [1, 21].

Momoh [22] offers an excellent summary of the currently-used and next-generation of control techniques including ADP, and describes how they might be used by the utility industry for optimal Smart Grid control. Werbos [23] further describes the intelligence that must be mathematically managed using computational learning system theory. Chuang and McGranaghan [24] further develop requirements for such intelligent controllers for the Smart Grid to include simple distributed generation and storage devices and the interfaces needed to connect to electricity market participation.

Building upon those successes, the next step is autonomous, Adaptive Stochastic Control that can simultaneously coordinate distributed generation and storage, utility operations and customer responses to stochastically varying system and market conditions.

Such a dynamic, stochastic system is described by five basic components:

1. The State Variables – and their three core components:

- a. The physical state – This would capture the amount of energy in a battery, the status of a diesel generator (on/off), or other physical dimensions of the system.
  - b. The information state – This includes current and historical demand, energy availability from wind/solar, etc., and electricity prices.
  - c. The belief state – For systems in which we are uncertain about the distribution of quantities such as demand, the reliability of the network, or best prices, we estimate via belief. Thusly, probability distributions make up our belief in the state of the system.
2. The Decisions (actions/controls) – These include whether to charge/discharge the battery, draw power from or pump it into the grid, or use backup generation, etc.
  3. The Exogenous Information – This captures all the dimensions of uncertainty such as possible changes in demand, price of electricity, and/or supply of energy (e.g. from clouds obstructing the sun). This would also include any network contingencies and emergency failures the system is experiencing.
  4. The Transition Function – Given the state, decisions and exogenous information, the transition function determines the state at the next point in time. This is a set of equations that describe how the system is likely to evolve over time.
  5. The Objective Function – This is the metric(s) that governs how we make those decisions and evaluate the performance of policies the controller designs.

An important component of the requirement for success of ASC involves specifying all five of these core elements of the problem. In addition, we also have to specify the control structure. For example, are we controlling a single-agent system whereby utilities are managing their own electrical grids, or a multi-agent system, whereby individual building operators on the customer-end and the Independent System Operator on the transmission-end are also participating in the management of local power distribution systems? All the above must be accommodated within the ASC algorithms.

#### A. Policies

Decisions are made using a policy  $X^\pi(S) \rightarrow x$  that maps the information in state  $S$  to a decision  $x$ . For our problem, it is useful to define the state variable using

$$S_t = (R_t, \rho_t, K_t)$$

which is the state variable, capturing energy resources  $R_t$ , exogenous information  $\rho_t$ , and the belief (or knowledge) state  $K_t$ .  $R_t = (R_{ta})_{a \in \mathcal{A}}$  is the *resource state vector*, where  $R_{ta}$  is the quantity of resources with attribute vector  $a \in \mathcal{A}$ .  $R_t$  describes the status of dispatchable power generation, the amount of energy in storage, the state of maintainable parts, the locations of mobile storage, generation, and curtailable load.

For example, if  $a$  refers to a particular diesel generator, then we might have  $R_{ta} = 1$  to indicate that the generator is turned

on. If  $a$  refers to a type of generator (of which there may be many), then  $R_{ta}$  might refer to the total kilowatts of capacity that are available to be used. For a battery,  $R_t$  would be the kilowatt-hours of energy in storage. For a mobile generator, we might let  $a$  be the location of the generator, and we use  $R_{ta} = 1$  to indicate that a generator is at location  $a$ .

The problem of choosing the right type of policy, and then the sub-problem of choosing the best parameters within a class of policies, is written as:

$$\sup_{\pi \in \Pi} V^\pi = \mathbb{E} \left\{ \sum_{t=0}^T \gamma^t C(S_t, X^\pi(S_t)) \right\} \quad (1)$$

$V^\pi$  is known variously as the value of a policy or the cost-to-go function (sometimes denoted as  $J^\pi$ ). Here,  $C(S_t, X^\pi(S_t))$  can be a cost function if we are minimizing, or a contribution function if we are maximizing. This may include the cost of generating electricity, purchasing fuel, losses due to energy conversion, and/or the cost of demand response. It may also include the cost of repair, and penalties for curtailing loads from buildings.

We can tune a policy to minimize costs while also maintaining a level of risk, for example, of being short of water in a reservoir. This is known as a root finding problem in stochastic search, for which the classic Robbins-Monro stochastic approximation procedure was designed. We simulate a policy (e.g. by fixing  $\theta$  in the policy  $X^\pi(S_t | \theta)$  above), and after each sample path we observe if we ran out of water or not. We then adjust  $\theta$  up or down to solve the constraint  $\mathbb{P}(\mathcal{E} | \theta) - q = 0$  where  $\mathcal{E}$  is the event that we run out of water, and  $q$  is the desired probability.

The focus of the ASC is to design a robust policy  $\pi$  that most optimally controls the components of the system, including whether to charge/discharge a battery/storage unit, when to run a distributed generator, and how much energy to draw from or add to the grid. Furthermore, this has to be done for every customer in every network and circuit in the utility's service area. Therefore, an adequate model of the system is required. Policies come in four broad classes:

- 1) Myopic Policies – Those policies are short term and are by definition unable to see into the future. Myopic policies minimize the next-period cost without regard to the impact of decisions for future states. Myopic functions lack an explicit forecast of events or costs in the future (for that, see the next three categories), but the tuning of selected parameters can produce policies that result in good behaviors over time, so these policies alone can often produce good results during *routine operational states*.
- 2) Look-ahead Policies – Also classified under names such as model predictive control and rolling horizon procedures, these policies involve optimizing over some time horizon using a forecast of the possible variability of exogenous events such as weather, demand and prices. Look-ahead policies can be broadly divided into three categories:

1. Deterministic forecasts = Optimization over a time horizon using point estimates of what might happen in the future,
2. Stochastic forecasts = Optimization over a time horizon using an approximation such as a sample realization of random outcomes that might happen within the range of the horizon,

Look-ahead policies with a stochastic forecast are typically hard to solve, while deterministic forecasts can produce decisions that are vulnerable to variations from the forecast.

- 3) Policy Function Approximations – These are functions that return an action given a state, without solving any form of optimization problem. These come in different flavors, including:

1. Rule-based lookup tables (“if” in this state, “then” take that action).
2. Parameterized rules (“if the electricity price is over some number  $\rho^U$ , then draw energy from the battery; if it is below  $\rho^L$ , then store energy in the battery”). Another form of parameterization arises when we have to combine the cost of electricity against the risk that we will have to ask customers to curtail usage.
3. Statistical functions – If  $x$  is the amount of energy to draw from the grid, then let

$$x = \theta_0 + \theta_1 \phi_1(S) + \theta_2 \phi_2(S)$$

where  $\phi_j(s)$  are predefined basis functions.

- 4) Policies based on Value Function Approximations – Optimal policies characterized by the Hamilton-Jacobi-Bellman (**HJB**) equation:

$$V_t(S_t) = \max_{x_t} (C(S_t, x_t) + E\{V_{t+1}(S^M(S_t, x_t, W_{t+1})) | S_t\})$$

Solving the HJB equation for Smart Grid problems incurs not one but **three** ‘Curses of Dimensionality’:

- 1) *the state variable*,
- 2) *the exogenous information*  $W_{t+1}$ ,  
(wind, solar, prices, demands), and
- 3) *the vector of actions*  $x_t$ .

We overcome these ‘curses’ using several devices: a) approximate the value function around the post-decision state to eliminate the expectation, b) replace the value function with a computationally tractable approximation, and c) solve the resulting deterministic maximization problem using a commercial solver.

These policies can be combined when, for example, we can use a short-term forecast of weather to optimize over, say, an eighty-four hour future time horizon (model predictive control) and then use value function approximations (given by  $\bar{V}_{t+1}(S_{t+1})$ ) to capture the impact of being in a state at the end of the eighty-four hours. These, in turn, can be further combined with tunable policies (policy function approximations), to take advantage of simple rules for determining when batteries should be charged or discharged.

## B. ADP and the Post-Decision State

Adaptive Stochastic Control for the Smart Grid involves the design of robust policies that work well over many sample realizations. Given the diversity of problems, choosing a controller requires finding a policy structure that works well given the specific structure of the physical problem. Particular concerns are the complexity of the state variable and the dimensionality of the control variable. For example, model predictive control is well suited to complex states and multidimensional actions, but the resulting models can be computationally large and time-consuming to solve, especially if uncertainties are handled explicitly. Run times can grow from a few seconds if we are willing to optimize over a short time horizon to many hours with longer horizons.

Policy function approximations require the design of specific functions that return an action given a state. The most flexible uses a lookup table (“if in this state, then take this action”) but this strategy is limited to extremely small state spaces. More common are parametric models (“store energy if the energy from wind exceeds some value”), which usually requires tuning one or more parameters. Parametric models of policy functions require the ability to recognize the structure of a policy. When this is possible, parametric models work very well.

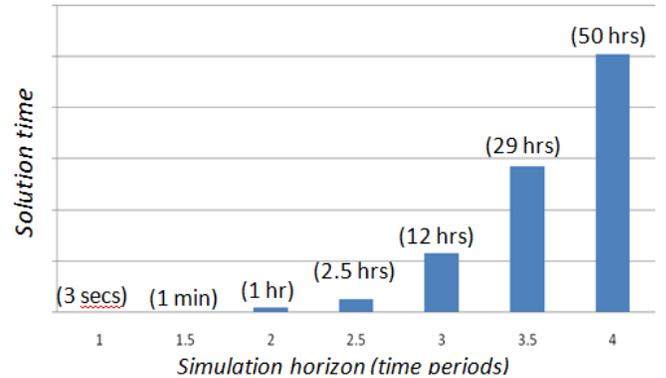


Fig. 4. Computation times can grow dramatically when optimizing over longer planning horizons.

When the structure of a policy is not obvious, it is necessary to turn to policies based on value function approximations. The theoretical foundation of this strategy is based on solving the HJB equation. Using different strategies for sampling the value of being in a state, these sampled values can be used to produce statistical estimates of the value of being in a state (sometimes referred to as the “critic” as in Fig. 3). These sampled estimates are typically then used to fit a parametric or nonparametric statistical model.

Policies based on value function approximations produce a decomposition that reduces Smart Grid problems with long horizons into a series of smaller problems that can be particularly easy to solve. Some applications, such as the load and source controller, might require solving integer programs. Using modern solvers such as Cplex, such problems can be exceptionally easy to solve over short horizons, but become exponentially harder as horizons grow (as might happen when using model predictive control). This is illustrated in Fig. 4.

Implicit in the use of policies based on value function approximations is that we can solve the optimization problem:

$$x_t = \arg \max_{x \in \mathcal{X}} (C(S_t, x) + \mathbb{E} \bar{V}_{t+1}(S_{t+1}))$$

When  $x$  is a vector, solving the maximization problem is problematic, since we also typically cannot compute the expectation exactly. In fact, there is an entire area of operations research known as stochastic search that focuses on solving problems of the general form:

$$\max_x \mathbb{E} F(x, W)$$

where  $W$  is a random variable.

In the Smart Grid, we solve the problem of “nested” expectation by approximating the value function around the *post-decision state*:

$$S_t^x = S^{M,x}(S_t, x_t),$$

which is computed deterministically from the current state  $S_t$  and action  $x_t$ . This transforms our decision problem to

$$x_t = \arg \max_{x \in \mathcal{X}} (C(S_t, x) + \bar{V}_t^x(S_t^x)).$$

Note that we are now solving a deterministic problem without an imbedded expectation. This makes it possible to bring into play commercial solvers that can handle vector-valued decisions. The solver is then used to handle multiple networks of storage devices, curtailable loads, and distributed generators interacting with a distribution grid.

### C. Designing policies

When using policy function approximations, we face the challenge of finding a function that maps a state to an action. For example, we might write the policy  $X^\pi(S_t)$  using a simple regression model as

$$X^\pi(S_t) = \theta_0 + \theta_1 \phi_1(S_t) + \theta_2 \phi_2(S_t). \quad (1)$$

Alternatively, the policy may be based on a value function approximation

$$X^\pi(S_t) = \arg \max_{x \in \mathcal{X}} (C(S_t, x) + \bar{V}_t^x(S_t^x)). \quad (2)$$

Whether we are approximating the policy itself using a policy function approximation as in (1), or a policy based on a value function approximation (2), we face the problem of approximating a function. Three fundamental strategies for approximating functions are i) lookup tables (an action for each state, or a value for each state), ii) parametric models (as in (1)), and iii) nonparametric models.

One of the most popular strategies is to use a parametric model for the value of being in each state, where we would write

$$\bar{V}(S) = \sum_{f \in \mathcal{F}} \theta_f \phi_f(S)$$

With this strategy, we face the challenge of first identifying the basis functions  $\phi_f(S)$ , and then tuning the parameters  $\theta$ . This approach is popular and can work quite well, but it introduces the undesirable “art” of identifying the basis functions.

A powerful and flexible strategy is to use nonparametric statistical representations to approximate either the policy or value functions. Then one must evaluate competing statistical strategies, such as

- Kernel regression
- Support Vector regression
- Neural networks
- Dirichlet process mixtures.

For example, Hannah et al. [25], [26], have developed an algorithm called DP-GLM, which uses Dirichlet Process mixtures of Generalized Linear Models. This method offers some powerful features:

- a) It can handle high-dimensional covariates (state variables),
- b) The covariates can be discrete, continuous or categorical,
- c) It is asymptotically unbiased.

The DP-GLM algorithm has recently been implemented in Java where it has been tuned to handle incremental updates, which is important for approximate dynamic programming. DP-GLM is a Bayesian strategy, but additional research is needed to handle the specification of priors.

These nonparametric methods offer the ability to handle complex functions. Also, they can be used in situations where data is becoming available in the future which we are not aware of now. This might be a common situation as we move into a future Smart Grid that we do not fully envisage right now. However, considerable empirical research is needed to ensure that an approximation strategy is robust.

Approximation strategies work best when they take advantage of the structure of the problem. For example, we may have a good idea what a battery charge/discharge policy should look like. A major component of the Smart Grid deployment at Con Edison involves the storage of energy, the management of dispatchable power such as distributed generators and storage units. These decisions all act on the resource vector  $R_t$ . Fortunately, resource allocation problems exhibit concavity (when we are maximizing), which is a particularly useful property both for approximating a value function, and for optimizing vector-valued decisions.

Godfrey and Powell [27], Topaloglu and Powell [28] and Powell [29] (chapters 11 and 12) show how this property can be exploited to solve large-scale stochastic resource allocation problems. This strategy was recently adapted to develop SMART, a stochastic, multi-scale energy policy model that can handle high-dimensional energy dispatch and storage over a large network and hundreds of thousands of time periods [29].

#### D. Policy search

There are two fundamental strategies for searching for policies:

1) Direct policy search for policy function approximations – Since we cannot compute the expectation exactly, we have to depend on Monte Carlo sampling (which might be online or offline). This draws on the broader field of stochastic search [31]. Policies can be optimized using methods such as sequential kriging. Frazier et al. [32], [33], and Scott et al. [34] develop the idea of using the knowledge gradient (see Section VI) for stochastic search, and this has proven to be quite effective for policy optimization. The knowledge gradient chooses measurements that maximize the expected value of a measurement. Direct policy search can be highly effective when the structure of a policy is fairly apparent. For example, deciding when to charge and discharge a battery may be a simple function of time of day, prices and energy availability from wind or solar. Generally, policy search is performed when the behavior of a policy is governed by a relatively small number of tunable parameters.

2) Bellman residual minimization for value function approximations – This is the most widely used strategy for optimizing policies, and encompasses a variety of algorithmic approaches that include approximate value iteration (including temporal difference learning) and approximate policy iteration (Bertsekas and Tsitsiklis [35], Bertsekas [36], Powell [37], and the references cited therein).

Bellman residual minimization uses classical statistical methods to observe the value of being in a state, and then uses this to develop an approximation of the value function as a function of the state. There are two broad strategies for performing this approximation: approximate value iteration (also known as TD(0)), where the value of being in a state depends directly on the current value function approximation (a form of statistical bootstrapping), and Approximate Policy Iteration (API), that requires simulating a policy into the future to approximate the value of being in a state, and then uses this to update the value function approximation.

Approximate value iteration is faster and easier to implement, but it has been shown to be unstable [38], [39]. However, it is very effective for problems that can be classified as resource allocation problems, which is true for Smart Grid decisions such as how much energy should be held in a storage device, how many distributed generators should be used at a point in time, and how many mobile storage devices should be moved to a particular location.

#### E. Convergence results

There are surprisingly few provably convergent algorithms in approximate dynamic programming [38], and none for general applications with continuous state variables. Fortunately, we have much stronger results when we can exploit the concavity that arises from resource allocation problems (Fig. 5). Ma and Powell [39] review the literature on convergence proofs and present a provably convergent algorithm for a parametric representation, but the proof makes

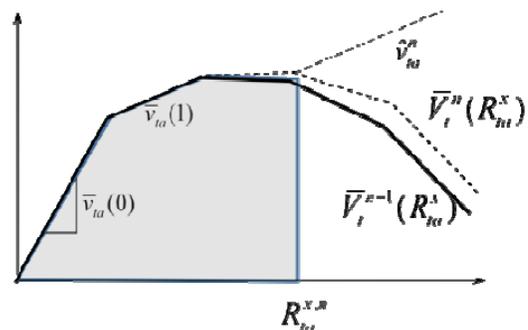


Fig. 5. Piecewise linear value function approximation for energy storage, showing stochastic update while maintaining concavity.

the very strong (and critical) assumption that the true value function can be perfectly represented in the space of value functions represented by the particular parametric representation.

Ormonet and Sen [42] present a convergence proof using kernel regression, but their algorithm assumes a finite action space (in practice, it has to be small) and kernel regression will not scale to more than a few dimensions in the state space. Ma and Powell [39] present a further convergence proof for an algorithm that assumes continuous and multidimensional states and actions using kernel regression, but it does not resolve the issue of exploration, which remains a difficult algorithmic issue when using nonparametric representations.

#### V. ASC FOR DISTRIBUTED GENERATION DISPATCH IN THE PRESENCE OF STORAGE

The control of distributed energy generation and storage resources for real time Load and Source Control (LSC) has been mostly limited to date to controlling pumped hydroelectric power in a reservoir. However, recent Smart Grid demonstration projects are providing new opportunities to show the value of resource allocation using distributed generation and storage for better control of the electric grid. Candidate high value applications are:

- “Instantaneous” storage
- Ramp-rate-limited distributed generation
- Cycling power supplies for load arbitrage
- Regulation control support
- Voltage and frequency stabilization
- Power quality management
- Reserve power management
- Reliability
- Security
- Load shifting
- Customer energy management
- Stability and optimization of intermittent, renewable power (e.g., wind, PV).

In order to take advantage of this flexibility, the ASC controller uses ADP to derive and execute load-balancing policies based on stochastic inputs of prices, cloud cover estimations, and distributed generation/storage availability.

In this section, we outline an algorithm that can be used to solve the problem of electric power dispatch in the presence of a single storage device. Posed as a maximization problem, we

exploit the property of concavity of the value function. At the same time, we assume that the state variable consists purely of a resource vector without other “state of the world” variables, which can dramatically complicate the problem. This issue is revisited in Section VI that follows.

#### A. Approximate Dynamic Programming for Resource Allocation

An important dimension of Smart Grid management involves making decisions that can be described as resource management or resource allocation: how much energy to store in a battery, whether a diesel generator should be turned on, and whether a mobile storage device (and/or generator) should be moved to a congested location. A general model that captures the state of all available resources uses the vector  $R_t = (R_{ta})_{a \in \mathcal{A}}$  where  $R_{ta}$  is the number of resources with attribute vector  $a$ . We then let  $x_t = (x_{tad})_{a \in \mathcal{A}, d \in \mathcal{D}}$ , where  $x_{tad}$  is the number of resources we act on with a decision of type  $d \in \mathcal{D}$ . A decision  $d$  can be  $(-1, 0, +1)$  to discharge, hold, or recharge a battery, and it can be  $(0, 1)$  to turn a distributed generator off or on, or to repair a component, or it can be a location to which we are sending a mobile storage device for load pocket relief. Note that setting a price signal is not a resource allocation decision which impacts flow conservation (represented on the right hand side of constraints). Prices impact coefficients in the objective function, introducing different challenges for approximating the value function.

Now assume that we represent our state variable as  $S_t = (R_t, \rho_t)$  where  $\rho_t$  is a vector capturing all parameters other than those captured by  $R_t$  (prices, demand, solar input, wind). We then exploit the property (true for many, but not all problems), that the value-function  $V_t(R_t, \rho_t)$  is often concave in  $R_t$ . Below, we summarize how the strategy works, along with the current state of convergence theory and recent experimental applications for energy management.

#### B. Value function approximations for resource allocation

The concavity property (assuming a maximization framework) suggests a powerful approximation strategy: we approximate the value function around the post-decision resource vector  $R_t^x = R^M(R_t, x_t)$  as a separable, piecewise linear function, using:

$$\bar{V}_t(R_t^x, \rho_t) = \sum_{a \in \mathcal{A}} \bar{V}_{ta}(R_{ta}^x)$$

where  $\bar{V}_{ta}(R_{ta}^x)$  is piecewise-linear and concave. It is very easy to estimate piecewise linear, concave functions [27], [30], by iteratively stepping forward through time, and updating value functions as we go. Imagine that we are in iteration  $n$  at time  $t$ , and let

$$\tilde{V}_t^n(R_t^n, \rho_t^n) = \max_{x_t \in \mathcal{X}_t^n} (C(S_t^n, x_t) + \bar{V}_t^{n-1}(R^M(R_t^n, x_t)))$$

where  $\tilde{V}_t^n(R_t^n)$  is a placeholder function and  $\mathcal{X}_t^n$  is the feasible region at time  $t$ , iteration  $n$ , capturing constraints such as flow conservation (e.g. we can only use energy we have stored). Solving the maximization problem can be done using a commercial solver such as *Cplex*. Now let  $\hat{v}_{ta}^n$  be the marginal value of an incremental change in  $R_{ta}^n$ . We might obtain this as a dual variable of the flow conservation constraint for  $R_{ta}^n$ , or using a numerical derivative:

$$\hat{v}_{ta}^n = \tilde{V}_t^{n-1}(R_t^n + e_{ta}, \rho_t^n) - \tilde{V}_t^{n-1}(R_t^n, \rho_t^n),$$

where  $e_{ta}$  is a vector of 0's with a 1 in the element corresponding to attribute  $a$ . We then use  $\hat{v}_{ta}^n$  to update our piecewise linear value function approximation.

We begin by first smoothing  $\hat{v}_{ta}^n$  with the slope of  $\tilde{V}_{ta}^n(R_{ta}^n)$  corresponding to  $R_{ta}^n$ . It is important to maintain concavity during the update. There are several methods to do this. For example, the successive, projective approximation routine (SPAR) first updates the value function, possibly producing a piecewise linear approximation that is no longer concave, and then projects this function back onto the space of concave approximations [30]. The idea is illustrated in Fig. 5.

#### C. Convergence theory

Exploiting the property of concavity has allowed us to obtain convergence results that are not possible with more general dynamic programs. Powell et al. [30] show that a pure exploitation algorithm (that is, the state that we visit is determined by our action) produces provably optimal solutions

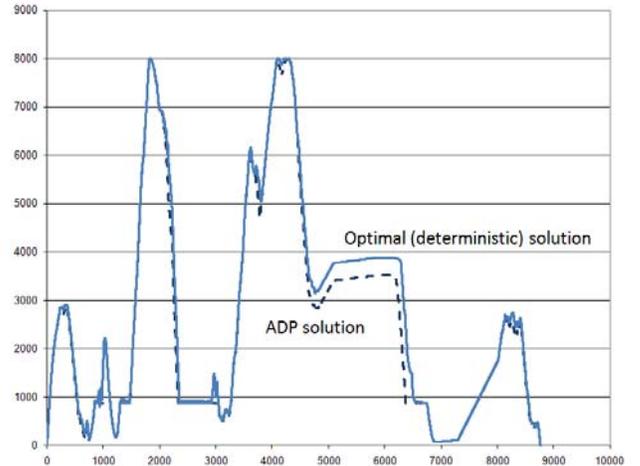


Fig. 6. Comparison of amount of energy held in storage over an entire year for a deterministic problem using a commercial solver (solid line) and approximate dynamic programming (dashed line).

for multidimensional two-stage problems. Pure exploitation algorithms are easy to implement for high-dimensional applications, since we only have to solve the maximization problem given a value function approximation, and then simulate this forward in time. Nascimento and Powell [43],

building on the theory in Powell and Nascimento [44], prove convergence for a *multistage* energy storage problem, if there is only one storage facility. This work produces an algorithm that can handle high-dimensional control vectors, such as those that govern the allocation of energy resources around a network.

It is critical to remember that these algorithms depend on pure exploitation. Convergence proofs for general ADP algorithms require some form of explicit exploration strategy to force the algorithm to visit all states. Such requirements are virtually impossible to enforce in practice for high-dimensional applications. Also, these algorithms use approximate value iteration, where an estimate of the value at time  $t$  depends on the value function approximation. Such algorithms are particularly easy to implement, and they scale well for high-dimensional applications.

#### D. Experimental work

There is a growing body of empirical research demonstrating that separable, piecewise linear value functions work for industrial-scale resource allocation problems. For example, Topaloglu and Powell [28] compare their performance for stochastic resource allocation problems.

We applied the approximate dynamic programming problem to an electricity dispatch problem. We modeled energy from wind, along with nuclear, coal, and natural gas generation from the grid. Decisions were made in hourly increments, where each decision problem was linked by a single energy storage device.

A major challenge that arises when using value function approximations is determining the quality of the resulting policy. For this problem class, an interesting benchmark is to fit the value functions for a deterministic problem, and compare the resulting solution to the optimal solution for the deterministic problem, obtained by using a commercial solver. This strategy is limited by the size of the deterministic problem that the solver can handle. For our application, we applied the problem to modeling energy storage in hourly increments over an entire year (8,760 time periods).

The results of this benchmark test are shown in Fig. 6 which shows the solution using approximate dynamic programming (dashed line) with those from a deterministic energy storage problem obtained using a commercial solver (solid line), which gives us the optimal solution. The results are comparable

## VI. ADP FOR A BATTERY STORAGE PROBLEM WITH A GENERAL STATE VARIABLE

We next considered an energy storage problem where we had to model the states of different processes including wind, load demand and prices, as well as the energy in battery storage. The problem, depicted in Fig. 7, includes unlimited energy from the grid (but at a price), free energy from wind or solar (but where the quantity is limited), energy from a storage node in the form of a battery, and a building with random demand.

We used this problem setting to compare two algorithmic strategies, both of which are based on approximating the value function using basis functions. The first is classical

Approximate Policy Iteration (API), which uses Bellman error minimization to estimate the value of being in a state. We then use API to determine a policy based on value function approximations. In the second, we use direct policy search to estimate the regression parameters of the value function approximation.

#### A. The model

Our energy flow model includes five decisions:

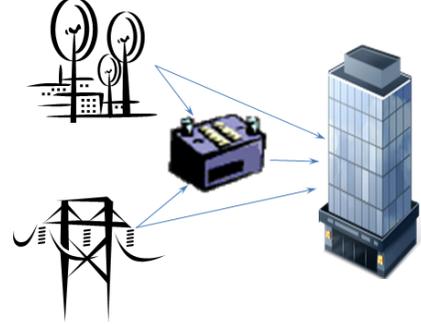


Fig 7. Energy storage network with energy from grid and wind, battery storage and a battery load.

$$x_t = (x_{t,GD}, x_{t,GB}, x_{t,WD}, x_{t,WB}, x_{t,BD}) .$$

These are, respectively, the flow of megawatts from Grid to Demand (GD to the building), Grid to Battery (GB), Wind to Demand (WD), Wind to Battery (WB), and Battery to Demand (BD). Energy from the grid or wind that is first stored in the battery has a conversion loss of  $1 - \rho$ . Power from the grid is unlimited, but at a price that depends on the commitment made the day before for that time period. If we need power that exceeds the commitment, then this has to be purchased on the more expensive real-time spot market. Power requested below the commitment is at a price fixed in the day-ahead market. We let  $E_t$  be the energy available from the wind at time  $t$ , and we let  $D_t$  be the energy required by the building at time  $t$ . The energy flows are governed by

$$x_{t,WB} + x_{t,WD} \leq E_t,$$

$$x_{t,WD} + x_{t,BD} + x_{t,GD} = D_t.$$

The energy storage equation is given by

$$R_{t+1,B} = R_{t,B} + \rho x_{t,GB} + \rho x_{t,WB} - x_{t,BD} + \hat{R}_{t+1,B}.$$

In addition, the wind  $E_t$ , demand  $D_t$ , and real-time prices  $\hat{p}_{t,G}^n$  from the grid evolve according to

$$E_{t+1} = E_t + \hat{E}_{t+1},$$

$$D_{t+1,B} = D_{t,B} + \hat{D}_{t+1,B},$$

$$p_{t+1,G}^n = p_{t,G}^n + \hat{p}_{t+1,G}^n.$$

$\hat{R}_{t+1,B}$  is a random variable used to capture exogenous changes in energy storage. The state of our system is given by

$$S_t = (R_t, E_t, D_t, p_{t,G}^n).$$

Of these, only  $R_t$  is directly affected by our decision vector  $x_t$ . The remainder evolves as a result of the exogenous random variables  $W_t = (\hat{E}_t, \hat{D}_t, \hat{p}_t^n, \hat{R}_t)$ .

### B. Policy optimization

We now face the challenge of designing a policy to determine  $x_t$ . We consider a policy with the structure

$$X^\pi(S_t) = \arg \max_x (C(S_t, x) + \bar{V}(S^{M,x}(S_t, x)))$$

where the value function approximation  $\bar{V}(S^{M,x}(S_t, x))$  is approximated using

$$\bar{V}(S_t^x | \theta) = \sum_{f \in \mathcal{F}} \theta_f \phi_f(S_t^x).$$

Here,  $S_t^x$  is the post-decision state variable, which for our problem is given by

$$S_t^x = (R_{tB}^x, E_t, D_t, P_t^n)$$

where

$$R_{tB}^x = R_{tB} + \rho x_{t,GB} + \rho x_{t,WB} - x_{t,BD}.$$

Before discussing the specific basis functions, we describe the two methods for computing the regression vector  $\theta$  in more detail.

1) Approximate Policy Iteration – With this classical algorithmic strategy, we fix  $\theta = \theta^{n-1}$  which is the parameter vector determined at iteration  $n-1$ , and then use this policy to generate a series of observations  $\hat{v}_t^n$  for different post-decision states  $S_t^{x,n}$ . We then use recursive least squares to estimate a new regression vector  $\theta^n$ . This strategy is well known in the ADP community (see Bertsekas and Tsitsiklis [35], Bertsekas [36], Powell [37]).

2) Direct Policy Search – Let the policy be given by

$$X^\pi(S_t | \theta) = \arg \max_x \left( C(S_t, x) + \sum_{f \in \mathcal{F}} \theta_f \phi_f(S_t^x) \right).$$

Now write the value of a policy using

$$F^\pi(\theta, W) = \sum_{t=0}^T C(S_t, X^\pi(S_t | \theta)),$$

where the state variable evolves according to  $S_{t+1} = S^M(S_t, X^\pi(S_t), W_{t+1}(\omega))$ . We can now pose the problem of finding  $\theta$  in terms of classical stochastic search, given by

$$\max_{\theta} \mathbb{E}F(\theta, W)$$

There are a number of algorithmic strategies that we can use to find  $\theta$  which recognize that we cannot compute the expectation, and have to use sample realizations of  $F^\pi(\theta, W)$  (see [31]). We did not assume that we could compute gradients  $\nabla_{\theta} F(\theta, W)$ .

To solve the policy search problem, we used a relatively new algorithm based on the concept of the knowledge

gradient, which chooses a  $\theta$  to test the value that gives us the highest improvement in expectation from a single measurement. Fig. 8 is a graph of the knowledge gradient surface after four measurements for a problem with two policy parameters. It is beyond the scope of this paper to describe the knowledge gradient in greater depth, but it has been proven to

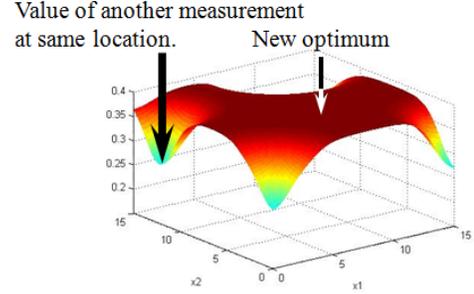


Fig. 8. Knowledge gradient surface for a two-dimensional parameter vector. Each dip corresponds to a previous measurement. The knowledge gradient policy requires finding the maximum of the surface (hill climbing).

asymptotically converge to the best possible policy given the basis functions (developed for discrete alternatives in [32,33]). Since  $\theta$  is continuous, we used a recent adaptation of the knowledge gradient given in [34] for continuous parameters.

In both strategies, we are using the same policy structure. For this reason, this is a pure test of the ability of each algorithm to find the best regression vector. There is a convergence theory surrounding approximate policy iteration, but this requires that the basis functions span the true value function. In practice, there is no guarantee of this, and if we have chosen our basis functions poorly, then estimating the regression coefficients based on sample observations becomes dependent on how we handle issues such as exploration.

### C. Experimental work

We created a battery of test problems based on five attributes: whether demand is deterministic or stochastic, the capacity of the battery, the round trip efficiency, the level of wind, and the price of electricity from the grid. The set of test problems are shown in Table I. For our choice of basis functions, we tested both a rich set of functions with linear, quadratic and cross terms, and a second, simpler set of functions to test the robustness of the two algorithms in the presence of a poor set of basis functions.

TABLE I  
CHARACTERISTICS OF TEST PROBLEMS

	Demand	Storage Size	Round Trip Efficiency	Wind Turbines	Grid Costs
1	Det.	500	Average	Low	Average
2	Det.	500	High	Low	Average
3	Det.	50	Average	Low	Average
4	Det.	50	High	Low	Average
5	Rand.	500	Average	Low	Average
6	Rand.	500	High	Low	Average
7	Rand.	50	High	Low	Average
8	Rand.	50	Average	Low	High
9	Rand.	500	Average	Low	High
10	Rand.	500	Average	High	High
11	Rand.	5000	Average	High	High

Including size of battery, conversion losses, level of energy from wind, and electric power prices from the grid.

For these algorithmic comparisons, it is very important to have a benchmark. For this reason, modifications were made to all the problems so that they could be solved optimally using classical value iteration. This streamlined problem assumes that demands are deterministic or follow a zeroth order Markov process, where the sequence  $\hat{D}_t$  is independently and identically distributed (i.i.d.). We made a similar assumption about the real-time price process. This new problem, then, has only a two-dimensional post-decision state vector  $S_t^x = (R_t^x, E_t)$  (the pre-decision state has four dimensions). We then created a new problem where  $x_t$  is discretized, as is the energy process  $\hat{E}_t$ . This model can then be solved using the classical methods of Markov decision processes. We use this problem to test our approximation strategies so that we can quantify the error precisely.

We compared the performance of the policy where the regression parameters  $\theta$  were estimated using approximate policy iteration to those obtained using direct search with the knowledge gradient. Approximate policy iteration is provably convergent under certain assumptions, the most important being that with the right value of  $\theta$ , the basis functions are such that the approximate value function matches the true value function. Since we cannot guarantee that we have chosen a good set of basis functions, it is important to test both algorithms using “good” and “bad” sets of basis functions. All results are evaluated as a fraction of the value of the optimal policy, where we took advantage of our ability to solve the simplified version of the problem. The results are shown in Table II. When using the large set of basis functions, both direct search and approximate policy iteration work extremely well, producing results that are all within five percent of optimal, and seven out of the eleven within two percent. When we use the reduced set of basis functions, approximate policy iteration produces very poor results on three of the eight datasets. For one dataset (#11), approximate policy iteration produced an objective function that was only 35

percent of the optimal policy. By contrast, the worst result produced by direct policy search was 94.5 percent of optimal, and nine out of 11 were within four percent of optimal.

These results are important. They show that policies based on approximations of value functions can produce very high quality solutions. This conclusion is supported by figure 6, when we compare against an optimal, deterministic benchmark, and table II where we use policies based on value function approximations. However, we have also shown that we can get very poor results, even if we use an algorithm such as approximate policy iteration, which is supported by the

TABLE II  
RESULTS OF DIRECT SEARCH AND APPROXIMATE POLICY ITERATION

	Large basis set		Small basis set	
	Direct search	API	Direct search	API
1	99.9	99.2	99.8	99.9
2	99.8	98.7	99.7	99.9
3	99.9	99.9	99.8	99.8
4	99.5	99.7	99.9	99.7
5	98.4	96.1	96.0	95.9
6	98.3	99.7	94.8	94.7
7	96.3	99.0	96.7	96.8
8	95.0	99.3	94.5	94.6
9	97.1	94.7	98.4	73.9
10	95.2	94.9	97.8	85.3
11	100.0	99.7	99.7	34.9

Using a large and small set of basis functions, expressed as a percent of the optimal policy.

strongest convergence proofs in the theoretical literature [36, 38, 39]. The problem is that our algorithms do not satisfy all the assumptions that are required by existing rigorous convergence proofs. By contrast, direct policy search was found to be more robust. However, additional research will be needed to extend this strategy to high-dimensional vectors  $\theta$ .

## VII. EFFICIENT FRONTIER

Only now are Adaptive Stochastic Control systems familiar to other industries beginning to be used by utilities. These controllers form a system of systems that integrates simulation models, machine learning, ADP, statistical diagnostics, capital asset planning, and contingency analysis tools to consider both the *next worst* and the *next most likely* events that might occur to the electric grid now and into the short term future. Exogenous drivers must be matched with cost/benefit analyses so that capital asset (CAPEX) and operations, and maintenance (OPEX) budgets are properly allocated in order to make sure the system is working reliably and economically, as well as efficiently, at all times. The Smart Grid is so new that few quantitative cost/benefit analyses yet exist, but McDonald [45] has made a start.

The Adaptive Stochastic Control framework required for a successful Smart Grid must demonstrably provide ways of treating uncertainty from both operational and financial standpoints, simultaneously. Hopefully, optimal, efficient and

safe operations will result far into the future. Similar systems engineering methodologies have been using these techniques for many years in other industries [46], [47].

An important issue that will arise in the Smart Grid is the handling of different objective functions. For example, we may have to balance decisions which increase the load that stresses portions of the grid (commonly termed a “load pocket”) against recommendations to curtail loads from customer buildings. Alternatively, we may have to balance the environmental cost of using a backup diesel generator against the financial cost of moving a mobile battery into position to be used by a building. These issues arise whenever we use optimization to solve a complex problem.

We anticipate using the classical strategy of introducing a utility function that is a weighted sum of different objectives. Of course, this means that we will need to tune these weights to strike the right balance for an operator, manager, or policy maker. This can be done by simulating different weights, reporting the value of the objectives and then letting an “Efficient Frontier” choose the weights that best reflect the goals of the organization. Such Pareto surfaces can be visualized to structurally understand the cost/benefit gains one gets by playing one objective against another. As part of a portfolio being managed, the Adaptive Stochastic Controller can be configured to compute and output the set of actions that most optimally follow such a Pareto surface (Fig. 9).

Optimal engineering design seeks regions that exhibit robust tradeoffs where the objectives work well with each other for a range of values that satisfy all the objectives. This is akin to the notion of robust policies.

A related issue arises when introducing the issue of risk into the Efficient Frontier. We can consider risk as one of the objectives in a multi-objective formulation (cost, benefit, risk). A common strategy (e.g. in Markowitz portfolio theory) to reduce the problem back to two dimensions (cost and benefit) is to include in the utility a cost for risk measures such as volatility. We can also include a penalty for specific outcomes, such as exceeding environmental regulations in the use of distributed generation.

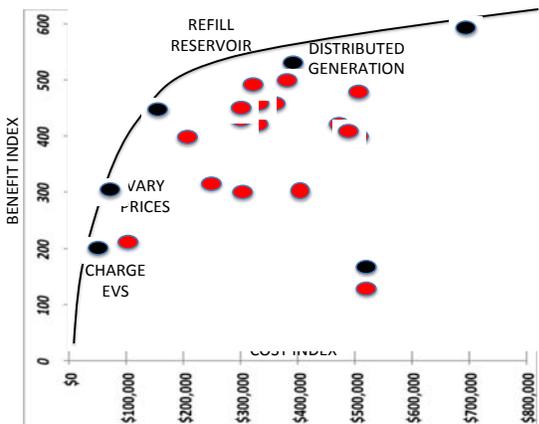


Fig. 9. An example of the Efficient Frontier (black line) for evaluating optimal Load and Source Control actions.

## VIII. CHALLENGES

Challenges to the future success of the Smart Grid come from many fronts, such as the need for more consumer buy-in and cost reduction. Consumers have to see real savings and efficiency improvements. In addition, governmental regulations must stay up to date technologically while at the same time staying in touch with these consumer requirements. Smart Grid components must be individually, as well as systemically, cost effective. Utilities, service companies and universities must produce a new generation of systems engineer savvy in computer sciences as well as the traditional electrical engineering to staff the Smart Grid. In the future, unimaginable products must be easily adopted and adapted into the Smart Grid, since it will evolve over the next 20 to 30 years.

A primary objective of the Smart Grid is to improve our capacity to use more, but cheaper, electricity to power the improvements in the standard-of-living of all people on Earth. The transition must be cost effective, or we will never get there from here. The tracking of key performance metrics that continuously and automatically score improvements generated by the Smart Grid will be required if the effort is to be sustainable.

Documenting these improvements requires the establishment of an initial baseline for all major performance components of the existing grid, and then the continuous measurement of the impact of new technologies against that baseline. We predict that a benefit from this “brutally empirical” measurement of performance will be the validation of Adaptive Stochastic Control as an optimal methodology for redirection of load around congestion, management of peak demand, weather vagaries, equipment problems, and other grid uncertainties in ways that will eliminate the need for the purchase of expensive new capital assets like additional power plants, substations, and transformers. That is why we envisage that “Approximate Dynamic Programming driven Adaptive Stochastic Control for the Smart Grid holds the promise of providing the autonomous intelligence required to elevate the electric grid to efficiency and self-healing capabilities more comparable to the Internet” as we stated in our Abstract. Such “Computer-Aided Lean Management”[1], operating at every level of the new Smart Grid, could eventually save the need to build terra-watts of new generation capacity worldwide. This alone would result in a major drop in the generation of greenhouse gases driving global climate concerns.

## ACKNOWLEDGMENT

The demonstration of the Adaptive Stochastic Controller for the Smart Grid of New York City is a key component of the American Reconstruction and Relief Act (ARRA) award won by prime contractor Consolidated Edison of New York, Inc., and we as subawardee’s, the Center for Computational Learning Systems of Columbia University and the CASTLE Laboratory of Princeton University. We thank Con Edison for their material and human support of the work described herein. We also thank the associate editor and referee for

comments that improved the presentation.

#### REFERENCES

- [1] Anderson, R., Boulanger, A., Johnson, J., and Kressner, A., Computer-Aided Lean Management in the Energy Industry, *PennWell Press*, 2008.
- [2] Mahmood, A., Aamir, M., and Anis, M., Design and Implementation of AMR Smart Grid System, *IEEE Electrical Power & Energy Conference*, 2008.
- [3] Tsoukalas, L., and Gao, R., From Smart Grids to an Energy Internet Assumptions, Architectures and Requirements, *IEEE DRPT Conference*, 2008.
- [4] Katz, J., Educating the Smart Grid, *IEEE Energy 2030*, 2008.
- [5] Divan, D. and Johal, S., Distributed FACTS—A New Concept for Realizing Grid Power Flow Control, *Power Electronics*, IEEE, 2007.
- [6] Divan, D., (2008) Smart Distributed Control of Power Systems, *Conversion and Delivery of Electrical Energy in the 21st Century*, IEEE.
- [7] Schnurr, N., Weber, T., Wellssow, W., and Wess, T., (2000) Load-Flow Control with FACTS Devices in Competitive Markets, *Electric Utility Deregulation and Restructuring and Power Technologies*, IEEE.
- [8] Steinberger, J., Van Niel, J., Bourg, D., Profiting from Negawatts: Reducing absolute consumption and emissions through a performance-based energy economy, in *Elsevier, Energy Policy*, 2009.
- [9] Boulanger, A., Chu, A., Maxx, S., and Waltz, D., Vehicle Electrification: Status and Issues, *IEEE Proceedings, Special Issue on the Smart Grid*, 2011.
- [10] Dicorato, M., Forte, G., and Trovato, M., A procedure for evaluating technical and economic feasibility issues of MicroGrids, *IEEE Bucharest Power Tech Conference*, 2009.
- [11] Pipattanasomporn A., and Rahman, A., Multi-Agent Systems in a Distributed Smart Grid: Design and Implementation, *Proc. IEEE PES 2009 Power Systems Conference and Exposition*, 2009.
- [12] Liu, X., and Su, B., Microgrids - An Integration of Renewable Energy Technologies, in Protection, Control, Communication and Automation of Distribution Networks, S3-25, CT 1800, CICED, 2008.
- [13] Jiang, Z., Power Management of Hybrid Photovoltaic - Fuel Cell Power Systems, *IEEE paper 1-4244-0493-2*, 2006.
- [14] Chowdhury, A., and Koval, D., Impact of PV Power Sources on a Power System's Capacity reliability Levels, *IEEE I&CPS-05-4*, 2005.
- [15] Swider, D., Compressed Air Energy Storage in an Electricity System, with Significant Wind Power Generation, *IEEE Trans of Energy Conversion*, v. 22, no. 1, 2007, 95-102.
- [16] Anderson, R., Texas Wind Energy Plan, Report to the Texas Energy Planning Council, *Railroad Commission of Texas*, 2004.
- [17] Lerch, E., Storage of Fluctuating Wind Storage: Case for Compressed Air Energy Storage in Germany, *IEEE*, 2008.
- [18] Smalley, R., Our Energy Challenge, at <http://video.google.com/videoplay?docid=-4626573768558163231#>
- [19] Yakobson, A., and Smalley, R., Fullerene Nanotubes: C1,000,000 and Beyond, *American Scientist*, 85-4, 1997, 324-337.
- [20] Anantram, M., and Govindan, T., Transmission through carbon nanotubes with olyhedral caps. *M. P. Phys. Rev. B*, 61(7), 2000, 5020.
- [21] Garrity, T., Innovation and Trends for Future Electric Power Systems, *IEEE Power and Energy*, 2008.
- [22] Momoh, J., Optimal Methods for Power System Operation and Management, *PSCE*, 2006, 179-186.
- [23] Werbos, P., Putting More Brain-Like Intelligence into the Electric Power Grid: What We Need and How to Do It, Proceedings of the 2009 international joint conference on Neural Networks, *IEEE Computational Intelligence*, 2009.
- [24] Chuang, J., and McGranaghan, M., Functions of a Local Controller to Coordinate Distributed Resources in a Smart Grid Angela Chuang, *IEEE*, 2008.
- [25] Hannah, L., Blei, D. and Powell, W., "Dirichlet Process Mixtures of Generalized Linear Models," Working paper, Department of Operations Research and Financial Engineering, *Princeton University*, 2010a.
- [26] Hannah, L., D. Blei, W. B. Powell, "Dirichlet Process Mixtures of Generalized Linear Models," *AISTATS*, 2010b.
- [27] Godfrey, G. and Powell, W., An Adaptive, Distribution-Free Algorithm for the Newsvendor Problem with Censored Demands, with Applications to Inventory and Distribution. *Management Science*, 47(8), 2001, 1101-1112.
- [28] Topaloglu, H., & Powell, W. Dynamic Programming Approximations for Stochastic, Time-Staged Integer Multicommodity Flow Problems. *Inform Journal on Computing*, 18, 2006, 31-42.
- [29] Powell, W., George, A., Lamont, A., & Stewart, J., SMART: A Stochastic Multiscale Model for the Analysis of Energy Resources, Technology and Policy. Working paper, Department of Operations Research and Financial Engineering, *Princeton University*, 2010.
- [30] Powell, W., Ruszczyński, A., & Topaloglu, H., Learning algorithms for separable approximations of discrete stochastic optimization problems. *Math. Oper. Res.*, 29(4), 2004, 814-836.
- [31] Spall, J., Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control. *John Wiley & Sons*, 2003.
- [32] Frazier, P., Powell, W. B., & Dayanik, S., A knowledge gradient policy for sequential information collection. *SIAM Journal on Control and Optimization*, 47(5), 2008, 2410-2439.
- [33] Frazier, P., Powell, W., & Dayanik, S., The Knowledge-Gradient Policy for Correlated Normal Beliefs. *INFORMS Journal on Computing*, 21(4), 2009, 599-613.
- [34] Scott, W., Frazier, P., & Powell, W. B., The Correlated Knowledge Gradient for Maximizing Expensive, Continuous Functions with Noisy Observations using Gaussian Process Regression. Department of Operations

- Research and Financial Engineering, *Princeton*, 2010, [http://www.castlelab.princeton.edu/Papers/ScottPowell-akg\\_2010\\_05\\_11.pdf](http://www.castlelab.princeton.edu/Papers/ScottPowell-akg_2010_05_11.pdf).
- [35] Bertsekas, D., & Tsitsiklis, J., *Neuro-Dynamic Programming*, *Athena Scientific*, 2006.
- [36] Bertsekas, D., *Dynamic Programming and Optimal Control*, Vol. II. *Athena Scientific*, 2007.
- [37] Powell, W. B., *Approximate Dynamic Programming: Solving the curses of dimensionality*. *New York: John Wiley and Sons*, 2007.
- [38] Bertsekas, D., *Approximate Policy Iteration : A Survey and Some New Methods*. *Journal of Control Theory and Applications*, 2010.
- [39] Ma, J., & Powell, W. B., *Convergence Analysis of On-Policy LSPI for Multi-Dimensional Continuous State and Action Space MDPs and Extension with Orthogonal Polynomial Approximation*, Working paper, Department of Operations Research and Financial Engineering, *Princeton University*, 2010.
- [40] Rudin, C., Waltz, D., Anderson, R, Boulanger, A., Salleb-Aouissi, A., Chow, M, Dutta, H., Gross, P., Huang, B., Jerome, S., Isaac, D., Kressner, A., Passonneau, R., Radeva, A., Wu, L., *Machine Learning for the New York City Power1 Grid*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011.
- [41] Anderson, R., Boulanger, A., Waltz, D., Long, P., Arias, M., Gross, P., Becker, H., Kressner, A., Mastrocinque, M., Koenig, M., Johnson, J., *System And Method For Grading Electricity Distribution Network Feeders Susceptible To Impending Failure*, *United States Letters Patent*, <http://www.freepatentsonline.com/y2009/0157573.html> , 2009.
- [42] Ormoneit, D., & Sen, S. (2002). Kernel-based reinforcement learning. *Machine Learning*, 49, 2002, 161-178.
- [43] Nascimento, J. and Powell, W. *An Optimal Approximate Dynamic Programming Algorithm for the Energy Dispatch Problem with Grid- Level Storage*,” working paper, Department of Operations Research and Financial Engineering, *Princeton University*, 2010.
- [44] Powell, W., and Nascimento, J., *An Optimal Approximate Dynamic Programming Algorithm for the Lagged Asset Acquisition Problem*. *Mathematics of Operations Research*, 34, 2009, 210-237.
- [45] McDonald, J., *Leader or Follower: Developing the Smart Grid Business Case*, *IEEE Power &Energy*, 2008, 18-24.
- [46] Schulz, A., *Agile engineering versus Agile Systems Engineering*, *Systems Engineering*, V. 3, Issue 4, 1999, 180-211.
- [47] Lemoine, D., *Valuing Plug-In Hybrid Electric Vehicles' Battery Capacity Using a Real Options Framework*, *Working paper 09-022 of the United States Association for Energy Economics*, 2009.



**Roger N. Anderson, M'09.** Roger has been at Columbia University for 35 years, where he is Senior Scholar at the Center for Computational Learning Systems in the Fu School of Engineering and Applied Sciences (SEAS). Roger is Principal Investigator of a team of 15 scientists and graduate students in Computer Sciences at Columbia who are jointly developing with Con Edison, Boeing and others the York City. Previously at the Lamont-Doherty Earth Observatory of Columbia, Roger founded the Borehole Research, Global Basins, 4D Seismic, Reservoir Simulation, Portfolio Management, and Energy Research Groups. Roger also teaches Planet Earth, a science requirement course in the core curriculum at Columbia College from his position in the Department of Earth and Environmental Sciences. He co-founded the Alternative Energy program at the School of International and Public Affairs at Columbia, and is a director of the Urban Utility Center at the Polytechnic Institute of New York University.

Roger received his Ph.D. from the Scripps Institution of Oceanography, University of California at San Diego. He is the inventor of 16 Patents, and has written 3 books, & more than 200 peer-reviewed scientific papers. In addition to his desk at the Manhattan Electric Control Center of Con Edison for the last 7 years, he has had technical, business, computational, and working collaborations with many other companies, including Baker Hughes, Boeing, BBN, BP, Chevron, IBM Research, KBR, Lockheed Martin, Pennzoil, Schlumberger, Siemens, Shell, United Technologies, and Western GECO. Roger's specialties include the Smart Grid, Optimization of Control Center Operations of Energy Companies, Real Options and Portfolio Management, 4D Reservoir Management, and Alternative Energy Research. His new book on the subject, *Computer-Aided Lean Management*, from PennWell Press, is available on Amazon.com. He has written scientific and opinion pieces for magazines such as CIO Insight, Discover, Economist, EnergyBiz, Forbes, National Geographic, Nature, New York Times, Oil and Gas Journal, Scientific American, Wall Street Journal, and Wired. Roger assisted in the design of the Wiess Energy Hall at the Houston Museum of Natural History, was technical consultant for the NBC News/Discovery Channel documentary "Anatomy of a Blackout," and has been a frequent contributor to business radio and TV.



founding member

**Albert Boulanger.** Albert received a B.S. in physics and the University of Florida, Gainesville, Florida USA in 1979 and a M.S. in computer science at the University of Illinois, Urbana-Champaign, Illinois USA in 1984.

He is a co-founder of CALM Energy, Inc. and a member of the board at the not-for-profit environmental and social organization World Team Now and of World-Team Building, LLC. He is a

Senior Staff Associate at Columbia University's Center for Computational Learning Systems, and before that, at the Lamont-Doherty Earth Observatory. For the past 12 years at Columbia, Albert has been involved in far reaching energy research and development – in oil and gas and electricity He is currently a member of a team of 15 scientists and graduate students in Computer Sciences at Columbia who are jointly developing with Con Edison, Boeing and others the next generation Smart Grid for intelligent control of the electric grid of New York City. He held the CTO position of vPatch Technologies, Inc., a startup company commercializing a computational approach to efficient production of oil from reservoirs based on time-lapse 4D seismic technologies. Prior to coming to Lamont, Albert spent twelve years doing contract R&D at Bolt, Beranek, and Newman (now Raytheon BBN Technologies). His specialties are complex systems integration and intelligent computational reasoning that interacts with humans within large scale systems.



**Warren B. Powell (M'10).** Warren has been a faculty member at Princeton University since 1981. Warren holds a Ph.D. and M.S. in Civil Engineering from MIT and graduated Summa Cum Laude with a B.S.E. from Princeton. He is the founder and director of CASTLE Laboratory, which was created in 1990 to reflect an expanding research program into dynamic resource management. He has

been funded by the Air Force Office of Scientific Research, the National Science Foundation, the Department of Homeland Security, Lawrence Livermore National Laboratory and numerous industrial companies in freight transportation and logistics, including United Parcel Service, Schneider National and Norfolk Southern Railroad. He pioneered the first interactive optimization model for network design in freight transportation, and he developed the first real-time optimization model for the truckload industry using approximate dynamic programming. His research focuses on stochastic optimization problems arising in energy, transportation, health and finance. He pioneered a new class of approximate dynamic programming algorithms for solving very high-dimensional stochastic dynamic programs. He coined the term "three curses of dimensionality," and introduced the concept of the post-decision state variable to eliminate the imbedded expectation. He is also working in the area of optimal learning for the efficient collection of information.

Warren has founded Transport Dynamics, Inc. and the Princeton Transportation Consulting Group. He is the author of Approximate Dynamic Programming: Solving the curses of dimensionality, and co-editor of Learning and Approximate Dynamic Programming: Scaling up to the Real World. The author/coauthor of over 160 publications, he is a recipient of the Informs Fellows Award, has twice been a finalist in the prestigious Franz Edelman Award, in 2009 directed the team that won the Daniel H. Wagner prize. He has served as President of the Transportation Science Section of Informs, in addition to numerous other leadership positions within Informs.



**Warren Scott.** Warren is a Ph.D. student at the Castle Laboratory of the Department of Operations Research and Financial Engineering at Princeton University. His research is in the area of optimal learning, and he has adapted the knowledge gradient to applications with continuous, multidimensional design parameters. He is continuing this research within the contextual domain of energy systems analysis, with a special focus

on optimal control of storage systems and approximate dynamic programming for load and source control.