LA-UR- 10-07890

Approved for public release; distribution is unlimited.

Title: Randomized Discrepancy Bounded Local Search for Transmission Expansion Planning

Author(s): Russell Bent
William Daniel

Intended for: Submission to the Power Engineering Society Conference

Los Alamos NATIONAL LABORATORY EST. 1943

Randomized Discrepancy Bounded Local Search for Transmission Expansion Planning

Russell Bent and W. Brent Daniel

Abstract—**In recent years the transmission network expansion planning problem (TNEP) has become increasingly complex. As the TNEP is a non-linear and non-convex optimization problem, researchers have traditionally focused on approximate models of power flows to solve the TNEP. Existing approaches are often tightly coupled to the approximation choice. Until recently these approximations have produced results that are straight-forward to adapt to the more complex (real) problem. However, the power grid is evolving towards a state where the adaptations are no longer easy (e.g. large amounts of limited control, renewable generation) and necessitates new approaches. Recent work on deterministic Discrepancy Bounded Local Search (DBLS) has shown it to be quite effective in addressing this question. DBLS encapsulates the complexity of power flow modeling in a black box that may be queried for information about the quality of proposed expansions. In this paper, we propose a randomization strategy that builds on DBLS and dramatically increases the computational efficiency of the algorithm.**

Index Terms—**Transmission Expansion Planning, TNEP, Local Search.**

I. Introduction

RECENT years have seen an increase in awareness that one of the major challenges of the 21st century is the problem of how to provide clean, sustainable, and cheap energy to the world's rising population [1], [2]. To address this challenge, the United States Department of Energy released a report in 2008 that stated the goal of having 20% of the U.S.'s energy come from wind by 2030 [3]. An important aspect of this report is the question of how to best upgrade and expand the electric power transmission grid to incorporate sustainable, renewable energy sources that are located in transmission deficient areas. This optimization problem has been well-studied as Transmission Network Expansion Planning (TNEP) [4], [5], [6], [7]; however, the requirements for the future grid raise a number of challenges, including:

1) The inability of expansion plans based on linearized DC models of power flows to guarantee an easy modification to account for nonlinearities and AC power flows under conditions imposed by uncontrollable generation [8].
2) With a few exceptions (for example, [9]), expansion algorithms are designed for specific models of power flow.

In previous work [10], we presented a novel approach to address these challenges. This approach, Discrepancy-Bounded Local Search (DBLS), embedded ideas from simulation optimization [11], [12], [13], [14], [15] in a local search procedure that generalizes constructive heuristics [16], [17], constraint-based local search [18], [19], and is related to global search techniques such as limited discrepancy search [20], [21], [22]. The key idea of the approach is the encapsulation of the power model within a simulation black box. The DBLS is allowed to query the black box for power flow information about proposed expansion plans. Unlike traditional simulation optimization that typically uses the "black box" only for evaluation (objective function) or feasibility checking, our approach uses information (i.e. flows) from the simulation to help influence the choices of the DBLS algorithm. While powerful, the approach can require large amounts of computation to solve large scale problems with complex power flow models. This paper considers this computational challenge and demonstrations how the introduction of randomization can significantly boost the computational performance of DBLS.

The authors are with the Decision Applications Division of Los Alamos National Laboratory.

In short, the key contributions of this paper include:

- An approach for randomizing DBLS to improve computational performance.
- A TNEP approach that is decoupled from the details of how power flows are modeled.
- A TNEP approach that handles non-linear models of power flow.
- An algorithm that generalizes existing TNEP heuristics.
- An algorithm that scales to large scale realistic problems.
- A framework for supporting multi-objective expansion planning.
- A coupling of simulation and optimization that allows the simulation results to influence the optimization choices.

Literature Review The literature on TNEP is extensive and references [4], [5], [6], [7] provide excellent surveys of the field. In general, existing approaches have focused on modeling power flows with transportation models [23] or the linearized DC model in order to reduce computational overhead. Until recently it has been been easy to adapt plans derived from these models to more realistic conditions (see [24], [25]). The approaches tend to fall into three categories, complete search based upon mixed-integer program (MIP) formulations [24], [26], locally optimal search such as constructive heuristics [16], [17], and meta-heuristics [27], [28], [29], [30], [31], [32].

One of the most relevant papers related to the work presented here is that of [33] which presents an expansion planning scenario where generation is fixed (also studied in [17]). In these papers, generation is fixed in order to model the challenges of economic dispatching, whereas the primary motivation for fixing generation in this paper is renewable energy sources. This is a pessimistic view of how power systems operate, but is useful in terms of understanding how worst case dispatching impacts expansion planning.

Reference [9] also shares a number of interesting similarities with this paper. It presents a tree-based local search procedure which contains a truncation criteria not unlike the discrepancy parameter of DBLS. Their approach utilizes combinations of the transportation model and the linearized DC model for modeling power flows. The paper does state that the approach is generalizable to more complex models of power flow, but this was not tested. It is important to note that their search procedure is primarily guided by cost, whereas DBLS is guided by both feasibility and cost.

Also of interest is the work of [34], [35] which is the basis for many of the results contained in [3]. These papers provide the fundamental motivations for the work of this paper. They studied how to best integrate large amounts of wind power into power grids spread over large geographic areas using transportation models of power transmission. We seek to address the question of how to account for non linear models of power flow into such planning scenarios as considerable effort is required to adapt models derived from transportation models in this context [8].

The remainder of this paper is organized as follows. Section II formally defines the TNEP. Section III describes the algorithm used to generate expansion plans and the heuristics used to guide the algorithm to reduce physical violations and cost. Section IV describes how randomization is incorporated into the DBLS. Section V discusses the experimental results and Section VI concludes this paper.

II. Problem Definition

Buses The TNEP problem is described in terms of a set of buses, $\mathcal{B}$, that represent geographically located nodes in a power network e.g. generators, loads, and substations. Each bus, i, is defined by parameters g_i, l_i, ι_i^-, ι_i^+, which represent its generation, load (demand for power), minimum voltage (per unit) and maximum voltage (per unit). $P(g_i)$ and $Q(g_i)$ are used to denote the real and reactive components of generation. Similarly, $P(l_i)$ and $Q(l_i)$ are used to denote real and reactive components of load. For simplicity, $P_i = P(g_i) - P(l_i)$ and $Q_i = Q(g_i) - Q(l_i)$ are used to denote the real and reactive power injected at bus i. The decision variable c_i is used to define the number of control components at i (in this paper, shunt capacitors for regulating reactive power). c_i has discrete domain $\{c_i^-, c_i^- + 1, \ldots, c_i^+ - 1, c_i^+\}$. c_i^- is defined as the number of control elements i starts with, ensuring that existing controls are included.

Transmission Corridors The TNEP is also described by a set of edges, $\mathcal{E}$, called transmission corridors, connecting pairs of nodes. A transmission corridor i,j between buses i and j has a decision variable $c_{i,j}$ that defines the number of circuits (power lines) in the corridor. The variable has discrete domain $\{c_{i,j}^-, c_{i,j}^- + 1, \ldots, c_{i,j}^+ - 1, c_{i,j}^+\}$ where $c_{i,j}^-$ is defined as the number of circuits the corridor starts with. $c_{i,j}^+ = c_{i,j}^-$ when no circuits may be added to a corridor. A circuit is also defined by parameter $\psi_{i,j}$ which denotes the capacity of a single circuit in the corridor. Similarly, $r_{i,j}$, $x_{i,j}$, $b_{i,j}$ denote the resistance, reactance, and line charging of a single circuit in the corridor.

TNEP Solution A transmission network solution, σ, is defined as a set of variable assignments $\bigcup_{i\in\mathcal{B}} c_i \leftarrow d_i \cup \bigcup_{i,j\in\mathcal{E}} c_{i,j} \leftarrow d_{i,j}$, where d_i is drawn from the domain of c_i and $d_{i,j}$ is drawn from the domain of $c_{i,j}$[1]. By convention, unassigned variables are assumed to be c_i^- and $c_{i,j}^-$, respectively. $\sigma(c_i)$ and $\sigma(c_{i,j})$ are used to denote the variable assignments for σ.

Simulation TNEP algorithms have at their disposal a simulator $\mathcal{S}$ for determining the behavior of power for σ. $\mathcal{S}(\sigma)$ returns true when it is able to compute the behaviors. This is important as some implementations of $\mathcal{S}$ use convergence approaches (e.g. Newton's method) that do not have guarantees on whether or not they are able to obtain a solution. $\mathcal{S}_{f_{i,j}}(\sigma)$ denotes the flow in corridor i,j and $\mathcal{S}_{v_i}(\sigma)$ the voltage at bus i. For simplicity, this notation is shortened to $f_{i,j}$ and v_i when $\mathcal{S}(\sigma)$ is understood from context. The following sets of equations provide an example of $\mathcal{S}$ for the linear DC model where $f_{i,j} = -f_{j,i}$.

$$\begin{array}{llll} \forall_{i\in\mathcal{B}} & P_i & = & \sum_{j\in\mathcal{B}} f_{i,j} \quad (1) \\ \forall_{i,j\in\mathcal{E}} & f_{i,j} & = & \lambda_{i,j} c_{i,j}(\theta_i - \theta_j) \quad (2) \end{array}$$

In this model, $\lambda_{i,j} = \frac{-x_{i,j}}{r_{i,j}^2 + x_{i,j}^2}$ is the susceptance of a circuit in corridor i,j and θ_i is the phase angle at bus i. The first constraint ensures conservation of flow (Kirchoff's current law) and constraint 2 expresses the relationship between phase angle and DC power (Ohm's law). This model does not use control components and does not calculate voltages (assumed to be 1). Implementing $\mathcal{S}$ as this set of equations allows the incorporatation of the traditional TNEP power flow model.

A more advanced implementation of $\mathcal{S}$ uses the following decoupled model of power flow.

$$\forall_{i\in\mathcal{B}} P_i = \sum_{j\in\mathcal{B}} v_i v_j (\gamma_{i,j} c_{i,j} \cos(\theta_i - \theta_j) + \lambda_{i,j} c_{i,j} \sin(\theta_i - \theta_j)) \quad (1)$$
$$\forall_{i\in\mathcal{B}} Q_i = \sum_{j\in\mathcal{B}} v_i v_j (\gamma_{i,j} c_{i,j} \sin(\theta_i - \theta_j) + \lambda_{i,j} c_{i,j} \cos(\theta_i - \theta_j)) \quad (2)$$

where $\gamma_{i,j} = \frac{r_{i,j}}{r_{i,j}^2 + x_{i,j}^2}$ is the conductance of a circuit in i,j.

A TNEP solution σ is feasible when the following constraints are satisfied, i.e.

$$\begin{cases} c_{i,j}^- \le c_{i,j} \le c_{i,j}^+ & (i,j \in \mathcal{E}) \quad (1) \\ c_i^- \le c_i \le c_i^+ & (i \in \mathcal{B}) \quad (2) \\ \mathcal{S}(\sigma) = \text{true} & \quad (3) \end{cases}$$

Physical constraints are relaxed and incorporated into the objective function in order to keep the search space connected (similar to Lagrangian Relaxation). The overload of σ is calculated as the sum of flow that exceeds the capacity of the circuits, i.e. $\eta(\sigma) = \sum_{i,j\in\mathcal{E}} \max(0, f_{i,j} - \psi_{i,j} c_{i,j})$. The voltage violation of σ is calculated as the sum of voltages that fall below ι_i^- or above ι_i^+, i.e. $\nu(\sigma) = \sum_{i\in\mathcal{B}} \max(0, \iota_i^- - v_i, v_i - \iota_i^+)$. Finally, the cost of σ is defined by $\kappa(\sigma) = \sum_{i,j\in\mathcal{E}} c_{i,j}\kappa_{i,j} + \sum_{i\in\mathcal{B}} c_i \kappa_i$, where κ_i is the cost of putting a control at bus i and $\kappa_{i,j}$ is the cost of putting a circuit in corridor i,j. The objective function, $f(\sigma)$, is then a lexicographic multi-objective function of the form $\min f(\sigma) = \langle \eta(\sigma), \nu(\sigma), \kappa(\sigma) \rangle$

III. DBLS Algorithm

In reference [10], a deterministic branch and bound algorithm is presented for the TNEP. This algorithm builds on

[1]This formulation can be generalized for multiple types of control components and circuits.

simulation optimization ideas by encapsulating the behavior of the network into a "black box" that may be queried by the algorithm for information about how a TNEP solution behaves (i.e. $\mathcal{S}(\sigma)$) and embedding it in a discrepancy bounded local search (DBLS) that limits the full exploration of the branch and bound search tree. The intuition behind DBLS is to generalize heuristics that make good decisions on how to build solutions, but make a few bad decisions from time-to-time. DBLS embeds the heuristic in a search tree as the branching heuristic and explores those solutions that are within δ violations (discrepancies) of the heuristic, where δ is a user-specified parameter. DBLS provides a natural way to incorporate constructive heuristics from the TNEP literature, e.g. [16], [17], into a more general framework and is related to the approach of [9]. The formal model of DBLS for TNEP is presented in Figure 1.

DBLS takes as arguments a starting solution σ, (often the current state of the network, i.e. $\bigcup_{i \in \mathcal{B}} c_i \leftarrow c_i^- \cup \bigcup_{i,j \in \mathcal{E}} c_{i,j} \leftarrow c_{i,j}^-$), a set of variables, $\mathcal{X}$, drawn from $\bigcup_{i \in \mathcal{B}} c_i \cup \bigcup_{i,j \in \mathcal{E}} c_{i,j}$, a heuristic discrepancy parameter, δ, a worsening discrepancy parameter α, and a divergence discrepancy parameter β. The δ parameter is used to control the number of times the branching heuristic may be violated in the search and is decremented in line 16. As the TNEP has the property that $f(\sigma)$ is non-monotonic, i.e. adding components can make $\eta(\sigma)$ and $\nu(\sigma)$ rise or fall (sometimes referred to as Braess's paradox). The parameter α is used to limit the number of times in a row that $f(\sigma)$ may worsen (lines 8-10). A similar parameter is used in [9]. Finally, as it is possible for $\mathcal{S}(\sigma)$ to fail (diverge) for a given σ, a parameter β is introduced to limit the number of times in a row that $\mathcal{S}(\sigma)$ may fail (lines 11-13). All three parameters are checked for violation in lines 1 and 2.

Line 4 chooses a variable to explore based upon the results provided by $\mathcal{S}$. It is here that the results of $\mathcal{S}$ drive the search and represent the largest departure from traditional simulation optimization. Line 5 provides the heuristic for ordering the domain of x. When $\eta(\sigma) > 0$ or $\nu(\sigma) > 0$ the domain is ordered by component additions, no change ($\sigma(x)$), and component removals, i.e.

$$\langle \sigma(x)+1, \ldots, x^+, \sigma(x), \sigma(x)-1, \ldots, x^- \rangle$$

otherwise it is ordered in reverse, i.e.

$$\langle \sigma(x)-1, \ldots, x^-, \sigma(x), \sigma(x)+1, \ldots, x^+ \rangle$$

Line 5 unassigns the current variable assignment of x and lines 6-16 iterate over the ordered domain of the variable. It is worth noting that line 7 implicitly updates attributes associated with the new σ and is where $\mathcal{S}$ is executed.

One of the challenges of this approach is that the performance of DBLS is highly dependent on the quality of early decisions. It can take a considerable amount of time to revisit those choices due the amount of backtracking that is required, especially on large scale problems. From a scalability perspective in reference [10] it was found to be productive to keep δ small when executing DBLS and iteratively restart DBLS with improving starting solutions. Restarts were also found to be productive in the algorithm presented by [32]. The restart procedure is described in the function OPTIMIZETNEP, where the algorithm is continuously restarted until the solution no longer improves.

OPTIMIZETNEP($\sigma, \mathcal{X}, \delta, \alpha, \beta$)
repeat
    $\sigma^* \leftarrow \sigma$;
    $\sigma \leftarrow$ DBLS($\sigma, \mathcal{X}, \delta, \alpha, \beta$);
  until $f(\sigma) \geq f(\sigma^*)$;
return σ^*;

DBLS($\sigma, \mathcal{X}, \delta, \alpha, \beta$)
1 **if** $\delta \leq 0$ **or** $\alpha \leq 0$ **or** $\beta \leq 0$
2   **then return** σ;
3 $x \leftarrow$ CHOOSEVARIABLE($\mathcal{X}, \sigma$);
4 $\langle d_1, d_2, \ldots, d_k \rangle \leftarrow$ ORDERDOMAIN(x);
5 $\sigma \leftarrow \sigma \setminus [x \leftarrow \sigma(x)]$;
6 **for** $i \leftarrow 1 \ldots k$
7 **do** $\sigma_i \leftarrow \sigma \cup [x \leftarrow d_i]$;
8   **if** $f(\sigma_i) < f(\sigma)$
9   **then** $\alpha_i \leftarrow 0$;
10   **else** $\alpha_i = \alpha - 1$;
11   **if** $\mathcal{S}(\sigma_i)$
12   **then** $\beta_i \leftarrow 0$;
13   **else** $\beta_i = \beta - 1$;
14   **if** $f(\sigma_i) \leq f(\sigma^*)$ **and** $\mathcal{S}(\sigma_i)$
15   **then** $\sigma^* \leftarrow \sigma_i$;
16   DBLS($\sigma_i, \mathcal{X} \setminus x, \delta - i, \alpha_i, \beta_i$);
17 **return** σ^*;

Fig. 1. Discrepancy Bounded Local Search

In this paper two implementations of CHOOSEVARIABLE are used. For ease of presentation, $\mathcal{E}(\mathcal{X})$ is used to denote those corridors that have circuit variables in $\mathcal{X}$, i.e. $\bigcup i,j \in \mathcal{E} \mid c_{i,j} \in \mathcal{X}$. $\mathcal{B}(\mathcal{X})$ is used to denote those buses that have capacitor variables in $\mathcal{X}$, i.e. $\bigcup i \in \mathcal{B} \mid c_i \in \mathcal{X}$.

The first implementation is described in Figure 2 and is based upon the constructive heuristic presented in [16]. It first chooses the variable corresponding to the corridor that is most overloaded (lines 1-3), if one exists. Otherwise the heuristic chooses the corridor within $n = 1$ hops of an overload that decreases an overloaded the most (lines 7-16). It then iteratively tries $n = 2, 3, 4, \ldots$ up to a user specified maximum until it finds a decreasing $f(\sigma)$ circuit addition (lines 6-17). If there are no corridors that satisfy this criteria, the heuristic selects the bus with the lowest voltage for adding shunt compensation (lines 18-19). This heuristic is used when $\eta(\sigma) > 0$ or $\nu(\sigma) > 0$.

The second heuristic is based upon the standard cost reduction stages of constructive heuristics [16], [17] and chooses to explore those variables whose removal of components will decrease the cost the most (lines 1-2 of Figure 3). It is used when $\eta(\sigma) = \nu(\sigma) = 0$.

CHOOSEVARIABLE-FEASIBLE$(\mathcal{X}, \sigma)$

1 $i, j \leftarrow \arg\max_{i,j \in \mathcal{E}(\mathcal{X})} |f_{i,j}| - \psi_{i,j}\sigma(c_{i,j})$;
2 **if** $|f_{i,j}| - \psi_{i,j}\sigma(c_{i,j}) > 0$
3   **then return** $c_{i,j}$;
4 $\hat{\mathcal{E}} \leftarrow \mathcal{E}$;
5 **while** $|\hat{\mathcal{E}}| > 0$
6 **do for** $k \leftarrow 1 \ldots n$
7   **do** $i, j \leftarrow \arg\max_{i,j \in \hat{\mathcal{E}}} |f_{i,j}| - \psi_{i,j}\sigma(c_{i,j})$;
8     $\mathcal{B} \leftarrow$ NEIGHBORS$(i, n) \cup$ NEIGHBORS(j, n);
9     $\hat{\mathcal{E}}_{i,j} \leftarrow (\mathcal{B} \times \mathcal{B}) \cap \mathcal{E}(\mathcal{X})$;
10   **for** $\hat{i,j} \in \hat{\mathcal{E}}_{i,j}$
11   **do** $\hat{\sigma} \leftarrow \sigma \setminus [c_{\hat{i,j}} \leftarrow \sigma(c_{\hat{i,j}})]$;
12       $\hat{\sigma} \leftarrow \hat{\sigma} \cup [c_{\hat{i,j}} \leftarrow \min(c^{+}_{\hat{i,j}}, \sigma(c_{\hat{i,j}}) + 1)]$;
13       $\perp_{\hat{i,j}} \leftarrow \mathcal{S}_{f_{i,j}}(\hat{\sigma})$;
14     $\hat{i,j} \leftarrow \arg\max_{\hat{i,j} \in \hat{\mathcal{E}}_{i,j}} \perp_{\hat{i,j}}$;
15   **if** $\perp_{\hat{i,j}} \leq \mathcal{S}_{f_{i,j}}(\sigma)$
16   **then return** $c_{\hat{i,j}}$;
17   $\hat{\mathcal{E}} \leftarrow \hat{\mathcal{E}} \setminus i, j$;
18 $i \leftarrow \arg\min_{i \in \mathcal{B}(\mathcal{X})} v_i$;
19 **return** c_i;

Fig. 2. Feasibility Branching Heuristic

CHOOSEVARIABLE-COST$(\mathcal{X}, \sigma)$

1 $i, j \leftarrow \arg\max_{i,j \in \mathcal{E}(\mathcal{X}) \ | \ \sigma(c_{i,j}) > c^{-}_{i,j}} \kappa_{i,j}$;
2 $i \leftarrow \arg\max_{i \in \mathcal{B}(\mathcal{X}) \ | \ \sigma(c_i) > c^{-}_{i}} \kappa_i$;
3 **if** $\kappa_{i,j} \geq \kappa_i$
4   **then return** $c_{i,j}$;
5 **return** c_i;

Fig. 3. Cost Reduction Branching Heuristic

IV. Randomized DBLS Algorithm

We now present two approaches for introducing randomness into DBLS that dramatically improves its computationally efficiency. The first approach considers the discrepancy parameter for worsening solutions. While a useful pruning mechanism, it does not use information about the degree of worsening. To address this limitation, we introduce a simulated annealing like acceptance criteria for exploring a worsening solution as seen in Line 10 of Figure 4, where T is the temperature parameter. Like simulated annealing T is cooled by parameter t in line 12.

The second approach adds randomness in how the discrepancy search tree is explored as seen in Figure 5. The search chooses P random paths in the tree (lines 4-8 of OPTIMIZETNEP). The random paths are guided by the branching heuristic using determinism parameters ω and ζ (lines 4 and 8 of DPLS). When ω and ζ are near 1, the choice of variable and variable assignment nearly follow the branching heuristic. The larger ω and ζ get, the more the branching follows a uniform distribution for selecting x (line 5) and d_i (line 9). The parameters control the influence of the branching heuristic. This approach is referred to as Discrepancy Probe Local Search (DPLS) throughout the rest of the paper.

DBLS$(\sigma, \mathcal{X}, \delta, T)$

1 **if** $\delta \leq 0$
2   **then return** σ;
3 $x \leftarrow$ CHOOSEVARIABLE$(\mathcal{X}, \sigma)$;
4 $\langle d_1, d_2, \ldots, d_k \rangle \leftarrow$ ORDERDOMAIN(x);
5 $\sigma \leftarrow \sigma \setminus [x \leftarrow \sigma(x)]$;
6 **for** $i \leftarrow 1 \ldots k$
7 **do** $\sigma_i \leftarrow \sigma \cup [x \leftarrow d_i]$;
8   **if** $f(\sigma_i) \leq f(\sigma^*)$ **and** $\mathcal{S}(\sigma_i)$
9   **then** $\sigma^* \leftarrow \sigma_i$;
10   **if** RANDOM$([0,1]) \leq e^{-\frac{f(\sigma_i) - f(\sigma)}{T}}$
11   **then** DBLS$(\sigma_i, \mathcal{X} \setminus x, \delta - i, T \times t)$;
12 **return** σ^*;

Fig. 4. Randomized Discrepancy Bounded Local Search

OPTIMIZETNEP$(\sigma, \mathcal{X}, T)$

1 **repeat**
2     $\sigma^* \leftarrow \sigma$;
3     $\hat{\sigma} \leftarrow \sigma$;
4   **for** $i \leftarrow 1 \ldots P$
5   **do** $\sigma_i \leftarrow$ DPLS$(\sigma, \mathcal{X}, T)$;
6   **if** $f(\sigma)_i < f(\hat{\sigma})$
7   **then** $\hat{\sigma} \leftarrow \sigma_i$;
8     $\sigma \leftarrow \hat{\sigma}$;
9   **until** $f(\sigma) \geq f(\sigma^*)$;
10 **return** σ^*;

DPLS$(\sigma, \mathcal{X}, T)$

1 **if** $\delta \leq 0$
2   **then return** σ;
3 $\langle x_1, x_2, \ldots, x_k \rangle \leftarrow$ ORDERVARIABLES$(\mathcal{X}, \sigma)$;
4 $j \leftarrow \lfloor$RANDOM$([0,1])^{\omega} \times |\mathcal{X}|\rfloor$;
5 $x \leftarrow x_j$;
6 $\langle d_1, d_2, \ldots, d_k \rangle \leftarrow$ ORDERDOMAIN(x);
7 $\sigma \leftarrow \sigma \setminus [x \leftarrow \sigma(x)]$;
8 $i \leftarrow \lfloor$RANDOM$([0,1])^{\zeta} \times (x^{+} - x^{-})\rfloor$;
9 $\sigma_i \leftarrow \sigma \cup [x \leftarrow d_i]$;
10 **if** $f(\sigma_i) \leq f(\sigma^*)$ **and** $\mathcal{S}(\sigma_i)$
11   **then** $\sigma^* \leftarrow \sigma_i$;
12 **if** RANDOM$([0,1]) \leq e^{-\frac{f(\sigma_i) - f(\sigma)}{T}}$
13   **then** DPLS$(\sigma_i, \mathcal{X} \setminus x, T \times t)$; **return** σ^*;

Fig. 5. Discrepancy Probe Local Search

V. Experimental Results

In order to evaluate our approach we considered four expansion planning benchmarks from the TNEP literature [33]. The approach was also tested on an expansion scenario for the electric power grid of the state of New Mexico based upon load and wind generation growth projections found in [3]. The commercial electric power simulation package T2000 [36] and

Bus	G1 $Q(g)$	G2 $Q(g)$	G3 $Q(g)$	G4 $Q(g)$	$Q(g)_{max}$	$Q(g)_{min}$
1	94.43	76.24	94.43	85.25	240.0	-150.0
2	46.8	46.8	46.8	42.32	240.0	-150.0
7	193.5	155.23	193.5	174.58	540.0	0.0
13	758.8	609.43	623.55	684.32	720.0	0.0
14	41.1	41.1	41.1	41.1	200.0	-150.0
15	0.15	0.15	0.08	0.13	330.0	0.0
16	75.66	75.66	45.88	68.17	240.0	-150.0
18	412.2	412.2	207.13	246.63	600.0	-150.0
21	324.6	324.6	257.24	291.32	600.0	-150.0
22	-89.28	-89.28	-89.28	-89.28	288.0	-180.0
23	64.6	195.45	406.08	287.94	930.0	-375.0

TABLE I
AC GENERATION

Bus	$Q(g)$	Bus	$Q(g)$	Bus	$Q(g)$	Bus	$Q(g)$	Bus	$Q(g)$
1	66	5	42	8	105	13	162	18	204
2	60	6	84	9	108	15	192	19	111
3	111	7	75	10	120	16	60	20	78
4	45								

TABLE II
AC LOAD

Bus	Bus	b	Bus	Bus	b	Bus	Bus	b
1	8	0.043	13	14	0.088	19	23	0.122
2	8	0.034	14	23	0.14	16	23	0.179
6	7	0.052						

TABLE III
NEW CORRIDOR LINE CHARGING

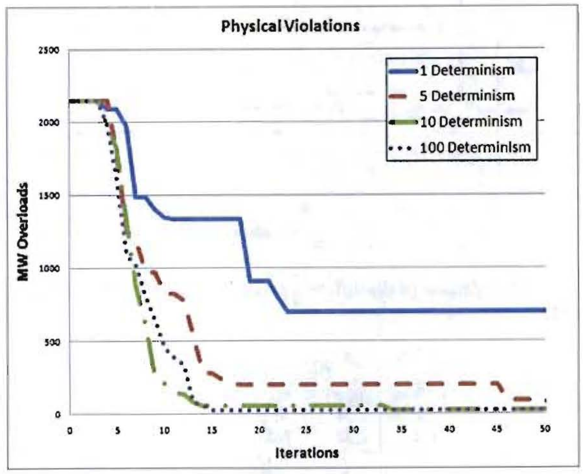

Fig. 6. A comparison of the different parameter settings of DPLS on problem G2 for η.

the linearized DC flow model are used as implementations of $\mathcal{S}$. It is important to note that since T2000 uses convergence methods for solving the power flow equations, there is no guarantee of a unique solution. Thus, it is possible that a stable flows exists for a σ that achieves a better value η than the one returned by $\mathcal{S}$. However, as the approach is not tied to a particular choice of $\mathcal{S}$, a user may supply a simulation model that either returns a unique solution or the best of a set of solutions, if desired.

The four benchmarks proposed in reference [33] are based on the RTS-79 and RTS-96 problems of [37], [38]. [33] grew demand and generation of the RTS by 200-300%. The problems allow up to 3 additional circuits in the 34 existing corridors and up to 3 circuits in each of 7 new corridors (thus, the domain of each circuit variable has size 4). The benchmarks pessimistically assume that generation cannot be dispatched. This provides worst case scenarios, e.g. all generation is wind-based. The approach described here does not depend on this property, as dispatching and/or optimal power flow are definable within $\mathcal{S}$, when appropriate.

The definition of the original RTS problems provide all the parameters for solving AC and DC power flows, however, as [33] used DC power flows, some information was not provided in the new problems, namely growth in AC generation and demand and line charging for circuits in new corridors. To overcome this limitation the AC load and generation were scaled by the same factors as [33]. We also modeled the generators as "voltage" controlled, thereby allowing $\mathcal{S}$ to adjust reactive generation to achieve certain voltage levels. This makes the problems easier, as the intent of the benchmarks is to make generation fixed. However, allowing reactive generation to fluctuate does provide a fairer comparison with results based on DC flows (as the behavior of the AC flows can be improved with flexible AC generation). The AC generation parameters for problems G1, G2, G3, and G4 are in Tables I, II, and III.

DC Power Flow Expansions We first test the approach on the benchmarks of [33] using the linearized DC power flow equations and compare the algorithm described in this paper with existing approaches in the literature.

The first results are described in Figure 6 and consider some of the parameter settings of DPLS on problem G2 for η. It shows the performance of varying $\omega = \zeta = \{1, 5, 10, 100\}$ (determinism) and keeping the other parameters fixed. The figure plots the best η found during the course of the search on average for $P = 100$ (the number of nodes visited in the search tree (iterations)). It is clear that some randomness around the branching heuristic improves the efficiency of the search. However, with too much randomness (like $\omega = \zeta = 100$), the search quality begins to degrade, thereby providing an indication of the value of the branching heuristic as a guide for the search. These results are further confirmed in Figure 7 which plots the best κ found during the search. Once again, a determinism factor of between 5 and 10 brings the most benefit.

In Figure 7 a comparison of DPLS and its randomized counterpart is presented on problem G2. The figure plots κ as a function of the number of nodes visited in the search tree. The figure plots the best performance of DBLS from its set of possible parameter settings (between 1 and 5 for δ, α, and β). The DPLS parameters are set statically to be $\omega = \zeta = 10$, $P = 100$, $T = 1$ and $t = .25$ and the results are an average of multiple runs. In this figure, DBLS initially performs better, however it is very quickly outpaced by DPLS. This provides evidence that while the branching heuristic is a good guide, the deterministic version of the search can spend time in unproductive regions of the search tree. The randomized version is able to more quickly probe other areas

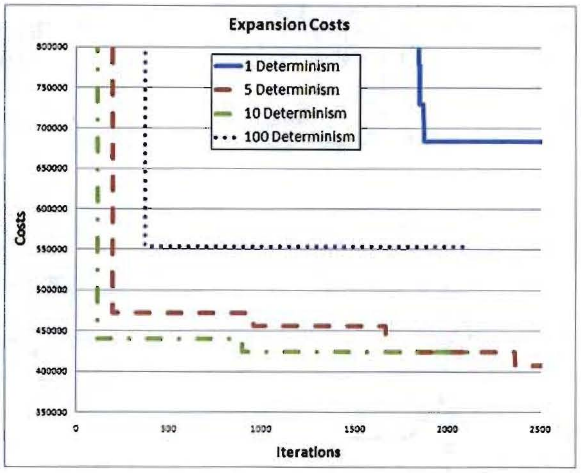

Fig. 7. A comparison of the different parameter settings of DPLS on problem G2 for κ.

	DBLS		DPLS	
Problem	$\eta(\sigma)$	$\kappa(\sigma)$	$\eta(\sigma)$	$\kappa(\sigma)$
G1	52	431	8	8
G2	32	7881	11	1028
G3	7	1813	5	65
G4	8	35	7	70

TABLE IV
NODE COUNTS FOR ACHIEVING THE BEST QUALITY SOLUTION FOR DIFFERENT PORTIONS OF THE OBJECTIVE FUNCTION.

of the search tree, biased by the guidance of the branching heuristic. This observation is reinforced by the results in Table V. This table shows the best performance of DPLS and DBLS in terms of number of search tree nodes explored to find the best values for η and κ. Note that on problems G3 and G4, the best result of DBLS is worse than the best result for DPLS for κ.

Finally, Table V provides the best solutions that we aware of in the literature for the benchmarks and the best solution found by the approaches presented in this paper. In the table

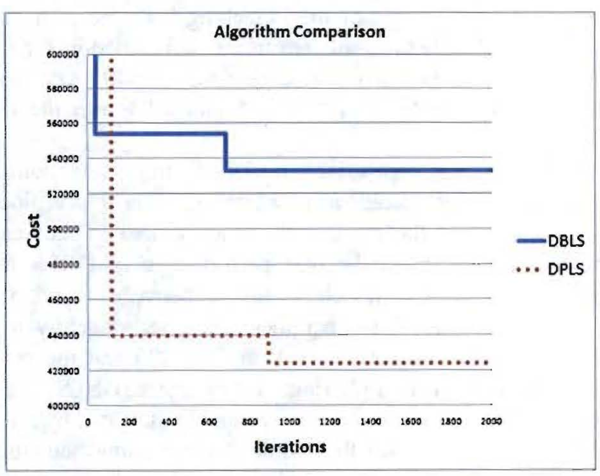

Fig. 8. A comparison between the deterministic and randomized discrepancy bounded local search.

Problem	Best Known	Ref	Best Found
G1	438K	RRMS	390K
G2	451K	FH	392K
G3	218K	RRMS	272K
G4	376K	FH	341K

TABLE V
BEST SOLUTIONS TO BENCHMARKS OF [33]

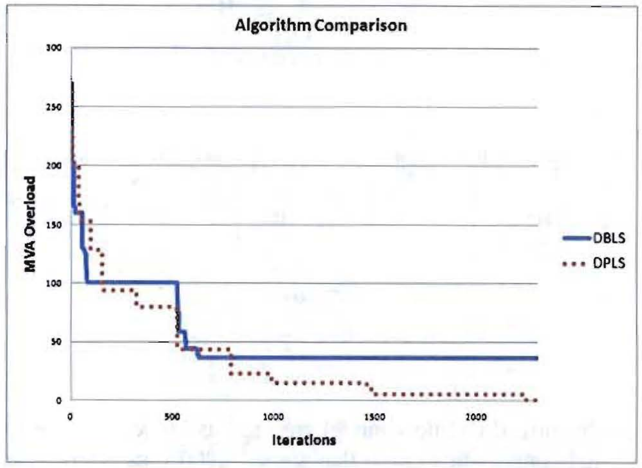

Fig. 9. A comparison of η on DBLS and DPLS when using nonlinear AC for $\mathcal{S}$.

RRMS refers to [17] and FH refers to [33]. In three cases, improved solutions are discovered. The solution to G1 adds the following circuits: [7,8](2), [16,17](2), [16,19], [6,10], [17,18](2), [14,16], [1,5], [15,24], [3,24]. The solution to G2 adds the following circuits: [10,11], [3,24], [14,16], [7,8], [16,17](2), [6,10], [1,5], [15,24], [17,18](2). The solution to G4 adds the following circuits: [15,24], [14,16], [7,8](2), [6,10], [16,17], [3,9], [3,24], [10,12].

AC Power Flow Expansions The second test of the DPLS algorithm uses nonlinear AC for $\mathcal{S}$ on the benchmarks of [33]. $\mathcal{S}$ is implemented using T2000 [36]. The behavior of DBLS and DPLS for nonlinear AC is illustrated in Figure 9. The figure plots the best case performance of DBLS for $\alpha = \beta = \delta = \{1,2,3,4,5\}$ and the average case performance of DPLS for $\omega = \zeta = 10, P = 100, T = 1$ and $t = .1$. For the most part, the algorithms behave roughly the same early in the search procedure. However, as the search proceeds, DBLS begins to explore poorer regions of the search space whereas DPLS is able to more quickly sample solutions from other areas of the search space.

The solutions obtained under nonlinear AC power flow models are interesting to compare with the solutions obtained under linearized DC power flow models. For example, the solution to G1 require 1316K in expansion costs. The DC solution required 390K in expansion costs. In other words, the AC solution is more than 3 times as expensive. This provides evidence that in planning scenarios where generation is not dispatchable, plans obtained using DC power flow models do not necessarily approximate the expansions required for AC based expansion very well. For reference, the solution to G1 is [1,5](2), [1,2], [1,3], [1,8], [2,4],

$[2,6]$, $[2,8]$, $[3,24]$, $[5,10](2)$, $[6,10]$, $[7,8]$, $[6,7](3)$, $[8,10]$, $[8,9]$, $[9,12](2)$, $[9,11]$, $[12,13]$, $[14,16](2)$, $[15,21]$, $[15,16]$, $[15,24](2)$, $[16,17]$, $[16,19]$, $[17,18]$, $[21,22]$.

New Mexico Expansions. While the algorithm is very effective in solving the benchmark problems, it is also important test its effectiveness on real networks. To perform this test, we took the transmission system for the state of New Mexico and modified the peak power demands according to the 2020 projections of [3]. We also added the generation that is scheduled to be built by 2020, which includes wind generation in the eastern part of the state. Under this planning scenario, if the grid is not upgraded, there is roughly 1700 MVA of overloads (spread over 31 transmission corridors) in the system as highlighted in Figure 10 (a). In order to resolve the physical violations, DPLS finds a solution within 100 search tree nodes that eliminates all physical violations at a cost of about $300,000,000, using the transmission expansion cost estimates of [39]. This solution adds 30 circuits in 28 existing corridors.

VI. Conclusion

As discussed, the electric power system is currently undergoing a revolutionary transformation that requires new approaches for solving the TNEP. The increased desire and need to incorporate sustainable power generation (wind and solar) that is less controllable has created a situation where nonlinear flows must be accounted for when evaluating expansion plans. Prior work has shown that DBLS is a powerful approach for solving problems with non-linear representations. This paper has shown that randomization strategies for the DBLS vastly improves its computational efficiency and thereby scaling the approach to larger problem instances. Furthermore, the approach relies on encapsulating portions of the problem's model as a black box similar to simulation optimization. The strength of this approach is that it uses the black box for more than just an evaluation criteria, but to direct the search procedure itself. A core contribution of the algorithms is a general search procedure that is *decouples* the model used for flows from the search and achieves solutions to the TNEP using non-linear flow equations.

Given the success of the approach described in this paper, it will be interesting to explore how to generalize this approach to more types of expansion options such as generation expansion, voltage upgrades, and other types of control components. The randomization strategies should help improve computational efficiencies when the number of variables is increased to this extent. Second, it will also be important to account for uncertainty in the planning process as described in [40], [41], in particular as it relates to the intermittent output of renewable energy. Once again, this increases the scale of the problems and DPLS should help make this problem tractable. Finally, it will be important to the study the effects on expansion solution quality when dispatchable generation or load management is included in S. This will allow and understanding of when the DC power flow model is a good approximation of the more complex power flows in expansion planning. The approach suggested in [42] for comparing DC models with AC models for different applications may be a good starting point for this sort of analysis.

Acknowledgments

This work was supported by the Los Alamos National Laboratory LDRD project Optimization and Control Theory for Smart Grids. We also acknowledge Eli Ben-Naim's insightful comments on incorporating randomness into the deterministic version of this algorithm.

References

[1] J. Lawhorn, D. Osborn, J. Caspary, B. Nickel, D. Larson, W. Lasher, and M. E. Rahman, "The view from the top," *IEEE Power and Energy Magazine*, vol. november/december, pp. 76–88, 2009.

[2] J. Achenbach, "The 21st century grid: Can we fix the infrastructure that powers our lives?" *National Geographic Magazine*, pp. 118–140, July 2010.

[3] United States Department of Energy, "20% wind energy by 2030: Increasing wind energy's contribution to U.S. electricity supply," United States Department of Energy, Tech. Rep., 2008.

[4] B. Hobbs, "Optimization methods for electric utility resource planning," *European Journal of Operational Research*, vol. 83, pp. 1–20, 1995.

[5] J. McCalley, R. Kumar, O. Volij, V. Ajjarapu, H. Liu, L. Jin, and W. Zhang, "Models for transmission expansion planning based on reconfigurable capacitor switching," in *Electric Power Networks, Efficiency, and Security*. John Wiley and Sons, 2006.

[6] K. Nara, "State of the arts of the modern heuristics application to power systems," *IEEE Power Engineering Society Winter Meeting, 2000*, vol. 2, pp. 1279–1283, 2000.

[7] G. Latorre, R. D. Cruz, J. M. Areiza, and A. Villegas, "Classification of publications and models on transmission expansion planning," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 938–946, 2003.

[8] L. Toole, M. Fair, A. Berscheid, and R. Bent, "Electric power transmission network design for wind generation in the western united states: Algorithms, methodology, and analysis," in *2010 IEEE Power Engineering Society Transmission and Distribution Conference & Exposition*, New Orleans, Louisiana, 2010, pp. 1–8.

[9] G. Latorre and I. J. Perez-Arriaga, "Chopin, a heuristic model for long term transmission expansion planning," *IEEE Transactions on Power Systems*, vol. 9, no. 4, pp. 1886–1894, 1994.

[10] R. Bent, A. Berscheid, and G. L. Toole, "Transmission network expansion planning with simulation optimization," in *Proceedings of Twenty-Fourth AAAI Conference on Artificial Intelligence*, Atlanta, Georgia, 2010.

[11] M. Fu, "Optimization for simulation: Theory vs. practice," *INFORMS Journal on Computing*, vol. 14, no. 3, pp. 192–215, 2002.

[12] J. April, F. Glover, J. Kelly, and M. Laguna, "Practical introduction to simulation optimization," in *Proceedings of the 2003 Winter Simulation Conference*, New Orleans, LA, 2003, pp. 71–78.

[13] M. Fu and C.-H. Chen, "Some topics for simulation optimization," in *Proceedings of the 2008 Winter Simulation Conference*, Miami, FL, 2008, pp. 27–38.

[14] Y. Guo, W. Liao, X. Cheng, and L. Liu, "Simopt: A new simulation optimization system based virtual simulation for manufacturing system," *Simulation Modelling Practice and Theory*, vol. 14, pp. 577–585, 2006.

[15] Q. Zeng and Z. Yang, "Integrating simulation and optimization to schedule loading operations in container terminals," *Computers and Operations Research*, vol. 36, pp. 1935–1944, 2009.

[16] E. Bustamante-Cedeno and S. Arora, "Multi-step simultaneous changes constructive heuristic algorithm for transmission network expansion planning," *Electric Power Systems Research*, vol. 79, pp. 586–594, 2009.

[17] R. Romero, C. Rocha, J. Mantovani, and I. Sanchez, "Constructive heuristic algorithm for the dc model in network transmission expansion planning," in *IEEE Proceedings of Generation, Transmission and Distribution*, vol. 152, 2005, pp. 277–282.

[18] P. Shaw, "Using constraint programming and local search methods to solve vehicle routing problems," in *Proceedings of the Fourth International Conference on the Principles and Practice of Constraint Programming (CP)*, Pisa, Italy, 1998, pp. 417–431.

[19] L. Michel and P. Van Hentenryck, "A constraint-based architecture for local search," in *OOPLSA'02*, Seattle, WA, 2002.

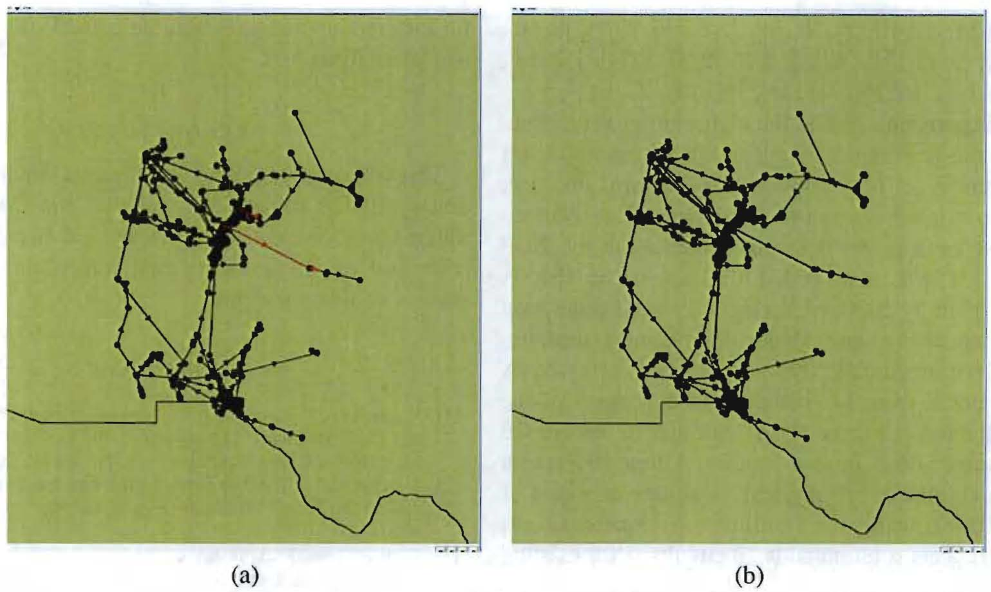

Fig. 10. Grid planning scenario for the state of New Mexico. Picture (a) shows the state of the grid with no expansion. Red indicates physical violations. Picture (b) shows the expansions to alleviate the physical violations (colored blue)

[20] W. Harvey and M. Ginsberg, "Limited discrepancy search," in *Proceedings of the Forteenth International Joint Conference on Artifical Intelligence (IJCAI)*, Montreal, Canada, 1995, pp. 607–615.

[21] R. Korf, "Improved limited discrepancy search," in *Thirteenth National Conference on Artificial Intelligence*, Portland, Oregon, 1996, pp. 286–291.

[22] T. Walsh, "Depth-bounded discrepancy search," in *Fifteenth International Joint Conference on Artifical Intelligence (IJCAI-97)*, Nagoya, Japan, 1997, pp. 1388–1393.

[23] S. Haffner, A. Monticelli, A. Garcia, J. Mantovani, and R. Romero, "Branch and bound algorithm for transmission system expansion planning using a transportation model," in *IEEE Proceedings - Generation, Transmission and Distribution*, 2000, pp. 149–156.

[24] N. Alguacil, A. Motto, and A. Conejo, "Transmission expansion planning: A mixed-integer lp approach," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1070–1077, 2003.

[25] I. Sanchez, R. Romero, J. Mantovani, and M. Rider, "Transmission-expansion planning using the dc model and nonlinear-programming technique," in *IEE Proceedings of Generation, Transmission and Distribution*, 2005, pp. 763–769.

[26] S. Binato, M. V. F. Pereira, and S. Granville, "A new benders decomposition approach to solve power transmission network design problems," *IEEE Transactions on Power Systems*, vol. 16, no. 2, pp. 235–240, 2001.

[27] E. da Silva, H. Gil, and J. Areiza, "Transmission network expansion planning under an improved genetic algorithm," in *Proceedings of the 21st 1999 IEEE International Conference Power Industry Computer Applications (PICA 99)*, 1999, pp. 315–321.

[28] H. Faria, S. Binato, M. Resende, and D. Falcao, "Power transmission network design by greedy randomized adaptive path relinking," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 43–49, 2005.

[29] R. Gallego, A. Alves, A. Monticelli, and R. Romero, "Parallel simulated annealing applied to long term transmission network expansion planning," *IEEE Transactions on Power Systems*, vol. 12, no. 1, pp. 181 – 188, 1997.

[30] R. Gallego, A. Monticelli, and R. Romero, "Comparative studies on non-convex optimization methods for transmission network expansion planning," *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 822–828, 1998.

[31] Mori, Hiroyuki, and Yoshinori, "Transmission network expansion planning with a hybrid meta-heuristic method of parallel tabu search and ordinal optimization," in *Proceedings of the 14th International Conference on Intelligent Systems Applications to Power Systems*, Kaohsiung, Taiwan, 2007, pp. 1–6.

[32] S. Binato, G. Couto de Oliveira, and J. Lizardo de Arajo, "A greedy randomized adaptive search procedure for transmission expansion planning," *IEEE Transactions on Power Systems*, vol. 16, no. 2, pp. 247–253, 2001.

[33] R. Feng and D. Hill, "A new strategy for transmission expansion in competitive electricity markets," *IEEE Transactions on Power Systems*, vol. 18, no. 1, pp. 374–380, 2003.

[34] P. Denholm and W. Short, "Documentation of winds base case data," National Renewable Energy Lab, Tech. Rep., 2006.

[35] NREL Energy Analysis Office, "Winds detailed model description, complete report," Tech. Rep., 2009.

[36] Commonwealth Associates Inc., "Transmission 2000 series power flow," 2005.

[37] Reliability Test System Task Force of the Application of Probability Methods Subcommittee, "IEEE reliability test system," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, no. 6, pp. 2047–2054, 1979.

[38] Reliability Test System Task Force, "The IEEE reliability test system - 1996," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 1010–1020, 1996.

[39] G. L. Toole, M. Ewers, R. Bent, M. Salazar, M. Hinrichs, K. Verma, and R. Ulibarri, "New mexico renewable development study: 5, 10, and 20-year transmission collection," Los Alamos National Laboratory, Tech. Rep. LA-UR 10-6319, 2010.

[40] E. Bustamante-Cedeno and S. Arora, "Stochastic and minimum regret formulations for transmission network expansion planning under uncertainties," *Journal of the Operational Research Society*, vol. 59, pp. 1547–1556, 2008.

[41] V. Miranda and L. M. Proenca, "Probabilistic choice vs. risk analysis - conflicts and synthesis in power system planning," *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 1038–1043, 1998.

[42] T. Overbye, X. Cheng, and Y. Sun, "A comparison of the AC and DC power flow models for LMP calculations," in *Proceedings of the 37th Hawaii International Conference on System Sciences*, Waikoloa Village, Hawaii, 2004, pp. 1–9.

Russell Bent Russell Bent received his PhD in Computer Science from Brown University in 2005 and is currently a staff scientist and a research team leader at Los Alamos National Laboratory in the energy and infrastructure analysis group. Russell is responsible for the development of LogiSims, a project for resource management, planning, and distribution for disasters. He is also co-principal investigator for a project on optimization and control theory for Smart Grid technologies. Russells publications include deterministic optimization, optimization under uncertainty, infrastructure modeling and simulation, constraint programming, algorithms, and simulation. In the last five years, Russell has published 1 book and over 25 articles in peer reviewed journals and conferences in artificial intelligence and operations research. A full list of his publications can be found at http://public.lanl.gov/rbent/.