Formulations and Valid Inequalities for Optimal Black Start Allocation in Power Systems

Georgios Patsakis
University of California Berkeley, gpatsakis@berkeley.edu

Ignacio Aravena
Lawrence Livermore National Laboroatory, aravenasolis1@llnl.gov

Deepak Rajan
University of California Berkeley, rdeepak@berkeley.edu

Shmuel Oren
University of California Berkeley, shmuel@berkeley.edu

The restoration of a power system after an extended blackout starts around units with enhanced technical capabilities, referred to as black start units (BSUs). We examine the planning problem of optimally allocating these units on the grid subject to a budget constraint. We present a mixed integer programming model based on current literature in power systems. Binary variables are associated with the allocation of BSUs and with the energization state of buses, branches, and generators of the power system over a time horizon. We extract a substructure of the feasible region which imposes the requirement that each island that appears during the restoration process must contain at least one operational generator. We discuss reformulations for this requirement and introduce a family of exponentially many, polynomially separable, valid inequalities to strengthen the formulation. Under simplifying assumptions, we show that some of the constraints we introduced are facet defining. We perform experiments to examine the difference in strength between formulations and the computational times to solve the problem to near optimality for multiple synthetic power system instances with up to 2000 buses. We illustrate a use case of the model. We conclude by suggesting extensions of the current work for future research.

Subject classifications: Energy, Integer Programming Applications

Area of review: Environment, Energy, and Sustainability

1. Introduction

Electricity supply is such a ubiquitous commodity that the extent of the consequences from a sustained electricity outage are often not fully comprehensible. These vary from the interruption of house and public lighting, refrigeration, and the internet, to the shutdown of banking and ATM systems, water supply, raw sewage dumping, cell phone communications, and farming operations.

All of the above occurred during the Northeast blackout of August 2003 (see Minkel 2008), which affected tens of millions of people for weeks. Such concerns are amplified due to a rapid transformation in the power grid: The generation mixture shifts away from the easily controllable, traditional paradigm because of the introduction of renewables, distributed generation, electric vehicles, microgrids, and storage. The power system infrastructure, built to a large extent about half a century ago, is aging. The climate is changing and extreme events (wildfires, hurricanes) are becoming more frequent. Cyber-security threats are becoming an important concern for a power system that relies increasingly more on the information infrastructure and automation. All of these factors have increased the probability of large scale outages and led to a widespread interest to enhance power system resilience.

Resilience refers to "the ability of the system to withstand and reduce the magnitude or duration of disruptive events" (National Infrastructure Advisory Council 2009, Phillips et al. 2016, Wang et al. 2015), such as natural disasters. A typical example is the ability of the system to restart after an extended blackout (black start). Most of the system generators are not able to start without connecting to an energized grid, since they are unable to setup and support the system frequency by themselves or they require external power to restart their auxiliary equipment. As a result, the system operator relies on a few units with the ability to start autonomously, called black start units (BSUs), to restore the power system. Allocating and maintaining these units is costly and the resulting allocation can severely impact the restoration security and time.

The black start allocation (BSA) problem is a planning problem solved in advance to allocate BSUs and prepare for a possible outage. In most systems, BSUs are sparsely located across the grid, and the allocation relies mainly on the start-up characteristics of the units - for example, hydro power is almost always a preferred choice. As a result, it is often the case that the BSUs are very few and not strategically placed on the grid, and restoration plans can take days or weeks to bring the system back to normal operation, while important grid operations are compromised. As resilience considerations are becoming increasingly important, it is expected that the budget for

commissioning black start resources will increase. The California Independent System Operator (CAISO), for example, has already identified a need for procuring additional black start resources (see California ISO 2017). In this new era, planning tools that connect the budget availability with restoration goals and suggest grid enhancements are becoming increasingly attractive and the scientific interest for these models in the power systems literature is high.

The BSA problem is only the first step in a multi-step planning process for Power System Restoration (PSR). The final outcome of this process is a complete restoration plan, i.e. a succession of switching and energization actions for branches, generators, and load pickup to restore the system to normal operation after an outage (often a complete blackout). The plan has to be tested for feasibility through ac power flow studies, dynamic stability simulations, island synchronization feasibility tests, cold load pickup simulations, protection relay simulations, as well as field tests for the black start units - each of these studies employs different types of software and equipment and needs specialized expertise to perform. Even after constructing a complete plan, it is unlikely that this is going to be faithfully reproduced in practice – an actual blackout involves many unplanned events, including complete failure of system components, loss of power in control centers, inability for remote control of branches or generators (which will require dispatch of personnel for manual operation), as well as unobservable state of a large part of the grid. However, since actual blackouts are rare events, we can only rely on these planning and simulation tools to prepare for them and to detect and fix potential vulnerabilities of the power system.

With this motivation, our goal in this paper is to present an optimization model for BSA based on a typical level of simplification employed for this step of restoration planning and discuss computational techniques that enable the solution of the problem for systems of realistic size. The model accommodates for major restoration concerns (see PJM 2019, Section 10) such as: selecting (additional) BSUs with suitable start-up characteristics, well positioned on the underlying graph of the power system to support the parallel restoration from many starting points, and close to load centers that could ensure the technical minima of generators are satisfied, while respecting a predefined allocation budget.

In order to improve the computational performance of the model, we discuss reformulations of a subset of the problem constraints and valid inequalities to further strengthen the formulation. The reformulations we suggest are relatively easy to implement and yield significant computational benefits in terms of speed and accuracy of the optimization model. We also include a theoretical study to motivate the use of our reformulations (validity of constraints, formulation strength, efficient constraint separation, polyhedral analysis). Additionally, we draw a connection of a substructure of the BSA problem, as well as other problems that employ grid reconfiguration in power systems, with the maximum weight connected subgraph problem. We perform comparative computations for instances of the IEEE-39, IEEE-118, Illinois-200, WECC-225, IEEE-300, South Carolina-500, and Texas-2000 synthetic power systems. We describe a use case for our model for implementing power system resilience upgrades subject to a budget.

This paper also aspires to serve as a reference to introduce a problem from power systems to the Operations Research community and initiate further research on the topic. Our model can be extended in multiple ways to accommodate for more general assumptions and there are many computational challenges to be tackled. For this purpose, we devote a significant part of the paper discussing the relevant literature in power systems as well as providing suggestions for future research, including possible modifications to our formulation at the end of this paper.

2. Literature Review

2.1. Black Start Allocation and Power System Restoration

There is a rich literature that attempts to model different aspects of the PSR problem. As mentioned in the introduction, creating restoration plans is a multi-step process with many considerations (see Jiang et al. 2017, Yutian et al. 2016, and Liu et al. 2012). These steps are not always distinct but they often include, according to a literature review by Yutian et al. 2016: selection of the BSUs (BSA problem), sectionalization (dividing the power system into sections corresponding to each BSU for parallel system restoration), unit start-up sequence optimization (determining the generator start-up order for each section), restoration path optimization (discovering the energization sequence

of all system components - buses, branches, generators), and synchronization (connecting islands of the system that are restored in parallel). Further considerations include load pickup (cold load simulations), damaged component repairs, field testing, and protection and dynamics for black starts. In this review, we mainly cover publications relevant to the BSA problem. We also refer to publications on unit start-up and restoration path optimization that formulate an optimization problem over candidate power system restoration sequences. Specifically, these publications use binary variables to model the energization state (energized or de-energized) of different system components (generators, branches, buses) over a time horizon and impose power systems constraints to model feasible transitions of the power system from a blackout (where all components are de-energized) to normal operation (where the grid is energized) over time.

For the BSA problem, a number of heuristic guidelines is commonly used to select suitable black start units in the literature (see Saraf et al. 2009, Sun et al. 2011a, Kafka 2008), as well as in restoration manuals (e.g., PJM 2017). These heuristics mainly select units with favorable start-up profiles (small start-up times and start-up power requirements). A smaller body of research attempts to allocate these units through an optimization problem. The motivation behind this research is that an optimal black start allocation needs to take into consideration the performance of the resulting restoration. However, even by including a moderate amount of detail, the optimization models become very large even for small systems and their discontinuous nature makes them often intractable. As a result, major simplifications are usually employed in this phase and the models are refined in subsequent planning steps to yield feasible restoration sequences.

Specifically, in Qiu et al. (2016) and Qiu and Li (2017) an optimal BSA problem is formulated by considering simplified active power considerations of the restoration sequence. In Jiang et al. (2017) an aggregate reactive power constraint is also utilized, but the grid flows are not. In Patsakis et al. (2018), we formulate a black start allocation problem with further considerations for thermal line limits and reactive power compensation. Finally, in Patsakis et al. (2019b), a reformulation of part of the problem constraints is employed for a simplified version of the black start allocation problem

with aggregate constraints. In all of the cases, the resulting restoration sequences generated do not necessarily correspond to feasible restoration plans. These are obtained in subsequent steps of the process. In case this proves impossible or if the resulting restoration plan is not satisfactory, the initial optimization problem can be modified accordingly, for instance by removing the previously found allocation from the feasible set via integer cuts, and resolved to obtain new allocations.

We now turn to literature on the restoration problem after a total blackout, where the BSUs are known and models that optimize restoration sequences over the time horizon are used. In Sun et al. (2011b) an optimization problem that includes the generator active power capabilities is considered, without active and reactive flows. A different modeling approach that includes reactive power considerations is adopted in Castillo (2013), aiming to motivate the use of microgrids for the BSA procedure, which is applied to a 6-bus system. A mixed-integer nonlinear program is formulated in Chou et al. (2013) for which the authors find feasible solutions using Ant Colony Optimization. Aravena et al. (2019) use a mixed integer programming approximation of the ac power flows, together with a decomposition scheme to obtain ac feasible restoration plans for utility-scale systems.

The use of BSUs is mainly relevant for large scale outages - smaller outages can be addressed using the surviving energized system. A relevant line of work is that of Van Hentenryck and Coffrin (2015) and Coffrin and Van Hentenryck (2015), where the authors consider models for outages of smaller scale but extend them to accommodate for component repairs and employ a decomposition scheme to tackle the complexity of the problem. An affiliated model appears in Arab et al. (2016), where Benders decomposition is utilized to enable scalability. There is also ample research on prevention of a complete blackout using integer programming models for network planning (e.g., Bienstock and Mattia 2007) or islanded operation (e.g. by utilizing microgrids, as in Gholami and Sun (2018)). Along the lines of enhancing resiliency, Byeon et al. (2020) propose a mixed integer programming model for employing distributed generation and remotely controlled switches in distribution systems. A model that employs generator and line switching to address resiliency concerns due to geomagnetic disturbances appears in Lu et al. (2019).

2.2. Island Energization Constraints

One requirement imposed during the restoration process is that each electrical island (an energized connected component of the power grid) that appears during the restoration needs to contain at least one operational generator. We refer to this requirement as the Island Energization (IE) constraint and properly define it in section 4. We have already studied reformulations of this constraint in Patsakis et al. (2019b), but in this paper we further expand on these results (i) by employing a new family of valid inequalities, (ii) by including a polyhedral analysis of the feasible region in a simplified case, (iii) by performing bound computations to compare the strength of the different formulations and (iv) by experimenting with test systems of a larger scale. In the appendix of this paper, we also describe how the feasible region defined by the IE constraints can be transformed into the feasible region of the Rooted Maximum Weight Connected Subgraph Problem (RMWCS), with weights on edges and nodes. This problem appears in Lee and Dooly (1998), where a decomposition algorithm is devised to solve it. The statement of the problem is as follows: consider an undirected graph with node set $\hat{N}$, edge set $\hat{E}$, and a given node $r \in \hat{N}$. Weights are associated with nodes and edges of the graph. We seek to find a maximum weight connected subgraph that contains r. This problem is a modification of the maximum weight connected subgraph problem.

Integer programming formulations for the RMWCS have been studied in Dilkina and Gomes (2010) for the directed graph version of the problem. In this version, each node in the subgraph must be reachable from the root via a directed path. The authors present a single-commodity and a multi-commodity flow formulation, similar to the ones in our work. They also present a cut-set formulation, but with the further restriction that the subgraph forms an arborescence: for every node there exists a unique directed path from the root node.

For the case of only node weights, Dittrich et al. (2008) transform the undirected, unrooted, connected subgraph problem to the Steiner tree problem, using the fact that the subgraph can, without loss of generality, be chosen to be a tree. Using this transformation, they employ algorithms for the Steiner tree problem to solve instances in protein–protein interaction networks. In Biha

et al. (2015) the feasible region for the unrooted, undirected version of the problem is formulated in the space of only edges. A convex hull description is given for cycles and trees.

Álvarez-Miranda et al. (2013b) formulate the RMWCS for the directed and undirected case in the space of only nodes and present a polyhedral analysis as well as a comparison of the strength of the formulation with the edge-based formulation. The formulation in the space of nodes can be strengthened by lifting the node separator inequalities and the in-degree inequalities (see Álvarez-Miranda et al. 2013a, Wang et al. 2017). In Carvajal et al. (2013) the model appears in a forestry planning application.

In most of the aforementioned literature, the problem is formulated either exclusively in the space of edges or in the space of nodes to reduce the number of variables. In our application, however, we need both nodes and edges, since both types of variables appear in constraints other than the ones regarding connectivity. For example, the node energization variable is necessary to also indicate if a bus can support load or not. Edges appear in the switching models of the power flow equations, to regulate when branches are open/closed. Edges also appear in the reactive power equations, since they are associated with the reactive power generation from long transmission lines. Finally, generator variables appear in the generator start-up curves. Also, the RMWCS formulation differs from the formulations that impose the requirement for a spanning tree of the subgraph to ensure connectivity, such as in the vast literature on the prize collecting Steiner tree, in that the edge variables are allowed to form cycles in our case. Of course, the feasible region projected in the space of node variables is going to be the same since connectivity of the subgraph is equivalent to the existence of a spanning tree.

3. Notation

Let (N, E) be the undirected graph derived from the physical graph of the power system, where buses correspond to nodes (set N) and branches to undirected edges (set E). Let A be a directed edge set corresponding to E, with the direction of each edge defined by arbitrarily picking a "From Bus" and "To Bus" for every system branch, as is common in the power systems literature. Note

that A and E have the same cardinality, i.e. $|A| = |E|$. Let G be the set of all generators, $G(i)$ the set of generators that are connected to bus $i \in N$, and $G(S)$ the set of generators connected to some bus in S, where $S \subseteq N$. For the case where a single generator is connected to i, we denote this generator by $g(i)$, i.e. $G(i) = \{g(i)\}$. If S is a subset of the nodes $S \subseteq N$, the undirected cut-set $\delta(S)$ is defined as the set that contains all the edges in E with one node in S and one node in $N \setminus S$. We denote $d = |N| + |E| + |G|$.

Let $t \in T$ denote an integer time instance (T being a set of consecutive integer time instances starting at 0), g denote a generator in set G, i a bus (node) in set N, and (ij) a branch (edge) in set A (or in set E by a slight abuse of notation, since there is a bijective correspondence between E and A). We use the binary variable u_{BS_g} to indicate, when it is set to 1, the allocation of generator $g \in G$ as a black start unit. The binary variables u_g^t, u_i^t, u_{ij}^t denote the energization status of generator g/bus i/branch (ij) at time t (1 indicates energized, 0 de-energized, for each element). Generator g is also associated with the binary cranking variable $u_{CR_g}^t$, which is 1 while the generator is cranking, and with auxiliary variables w_g^t corresponding to $u_g^t(1 - u_{BS_g})$. We use auxiliary variables $\boldsymbol{f}$, corresponding to energization network flows (not power flows), in the definition of the single- and multi-commodity flow formulations. The variable p_g^t is the active generation of generator g, p_{ij}^t is the active power flow on branch (ij), and $p_{SH_i}^t$ is the load shed at bus i. Finally, the parameters of the problem are the cost C_{BS_g} of allocating generator g to be a BSU, the total allocation budget B, the number of allowable branch energizations per unit time K_{crew}, the generator active power maximum $P_g^{\max}$, active power minimum $P_g^{\min}$ (we assume $P_g^{\max} \geq P_g^{\min} \geq 0$), cranking time T_{CR_g} (assumed to be a positive integer), cranking power P_{CR_g}, minimum reactive power capability $\underline{Q}_g$, the bus load $P_{D_i} \geq 0$, angle ϕ_{D_i}, shunt reactance Q_{SH_i}, the branch shunt susceptance $B_{SH_{ij}}$, and branch power limit $\bar{S}_{ij}$.

4. Black Start Allocation Model

In this section, we present a version of the BSA problem that retains the basic structure of the problem, while permitting good computational performance for large-scale systems. The model

is largely based on previous work by Patsakis et al. (2018), Aravena et al. (2019), Jiang et al. (2017), Sun et al. (2011b), Qiu and Li (2017), Qiu et al. (2016). Some deviations from the previous literature are subsequently discussed. The full model is presented below.

$$\text{maximize} \quad \sum_{(ij)\in A, t\in T} u_{ij}^t + \sum_{i\in N, t\in T} u_i^t + \lambda_G \sum_{t\in T} h(\boldsymbol{u}^t)$$

subject to

$$\sum_{g\in G} C_{BS_g} u_{BS_g} \leq B \tag{1a}$$

$$u_i^0 = 0,\ i\in N,\ u_g^t = 0,\ g\in G, t\in\{-T_{CR_g},\dots,-1\} \tag{1b}$$

$$u_g^t \geq u_g^{t-1},\ g\in G, t\in T\setminus\{0\} \tag{1c}$$

$$u_{CR_g}^t = u_g^t - u_g^{t-T_{CR_g}},\ g\in G, t\in T \tag{1d}$$

$$u_{ij}^t \leq u_i^{t-1} + u_j^{t-1},\ (ij)\in A,\ t\in T\setminus\{0\} \tag{1e}$$

$$\sum_{(ij)\in A} (u_{ij}^t - u_{ij}^{t-1})^+ \leq K_{\text{crew}},\ t\in T\setminus\{0\} \tag{1f}$$

$$u_{CR_g}^t \leq u_i^t + u_{BS_g},\ g\in G(i), i\in N, t\in T \tag{1g}$$

$$u_g^t - u_{CR_g}^t \leq u_i^t,\ i\in N,\ g\in G(i),\ t\in T \tag{1h}$$

$$u_{ij}^t \leq u_i^t,\ u_{ij}^t \leq u_j^t,\ (ij)\in A,\ t\in T \tag{1i}$$

$$0 \leq f_g^t \leq u_g^t - u_{CR_g}^t,\ g\in G,\ t\in T, \tag{1j}$$

$$-u_{ij}^t \leq f_{ij}^t \leq u_{ij}^t,\ (ij)\in A,\ t\in T, \tag{1k}$$

$$\sum_{j:(ji)\in A} f_{ji}^t - \sum_{j:(ij)\in A} f_{ij}^t + \sum_{g\in G(i)} f_g^t = \frac{1}{|N|} u_i^t,\ i\in N,\ t\in T \tag{1l}$$

$$w_g^t + u_{BS_g} \geq u_g^t,\ g\in G, t\in T \tag{1m}$$

$$w_g^t \leq u_g^t,\ g\in G, t\in T \tag{1n}$$

$$w_g^t + u_{BS_g} \leq 1,\ g\in G, t\in T \tag{1o}$$

$$p_g^t \geq -P_{CR_g} w_g^t + (P_{CR_g} + P_g^{\min}) w_g^{t-T_{CR_g}},\ g\in G, t\in T \tag{1p}$$

$$p_g^t \leq P_g^{\max}(u_g^t - u_{CR_g}^t) - P_{CR_g}(w_g^t - w_g^{t-T_{CR_g}}),\ g \in G, t \in T \tag{1q}$$

$$-u_{ij}^t \bar{S}_{ij} \leq p_{ij}^t \leq \bar{S}_{ij} u_{ij}^t,\ (ij) \in A,\ t \in T \tag{1r}$$

$$\sum_{j:(ji)\in A} p_{ji}^t - \sum_{j:(ij)\in A} p_{ij}^t + \sum_{g\in G(i)} p_g^t = P_{D_i} - p_{SH_i}^t,\ i \in N,\ t \in T \tag{1s}$$

$$(1-u_i^t)P_{D_i} \leq p_{SH_i}^t \leq P_{D_i},\ i \in N,\ t \in T \tag{1t}$$

$$\sum_{i\in N}\sum_{g\in G(i)} \underline{Q}_g u_g^{\max\{0,t-T_{CR_g}-1\}} + \sum_{(ij)\in A} B_{SH_{ij}} u_{ij}^t +$$
$$\sum_{i\in N} Q_{SHi} u_i^t - \sum_{i\in N} P_{D_i}\tan(\phi_{D_i}) u_i^t \leq 0,\ t \in T \tag{1u}$$

The objective is a measure of the energization state of the system at different time instances. The first two objective terms encourage energization of grid branches and buses. The third term encourages the energization of generators. We define $h(\boldsymbol{u}^t) = \min(\alpha_L \sum_{i\in N} P_{D_i}, \sum_{g\in G} u_g^t P_g^{\max})$. This definition is motivated by the fact that we prioritize the energization of larger generating units to support the power system (larger units usually have larger inertia constants, making the system more stable), but since real power systems operate with large excess generation capacity, we do not want to encourage unit energization after the total capacity of energized units exceeds a multiple $\alpha_L (\geq 1)$ of the total load. The parameter λ_G is a trade-off coefficient. Variations of this objective have been used in Patsakis et al. (2019b), Aravena et al. (2019), and Sun et al. (2011b), but for a real-world application the objective chosen should additionally reflect any specific restoration priorities of that system.

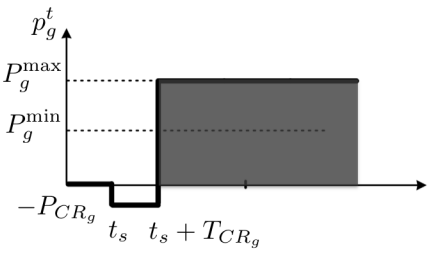
Constraint (1a) imposes the black start allocation budget on the selection of black start units. The restoration of the system starts around these units. One simplified way to visualize the underlying optimization is as follows: We have a given budget to pick a few generators on the system (corresponding to particular nodes). We can then expand the system energization around these nodes in the power system graph over time. The objective metric corresponds roughly to how much of the system we were able to energize, measured at given time instances after the start of the process. Of course, the energization process needs to obey certain rules, such as operator

switching limits, start-up curves for the generators, power constraints for every island. We include these constraints in the optimization problem.

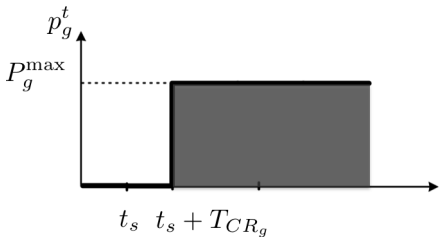
Constraint (1b) initializes the node and generator variables to zero (total blackout). The model could be generalized in a straightforward way to accommodate for partial blackouts as well (see Patsakis et al. 2019a). Constraint (1c) stipulates that a generator will only be energized and not shut down. This simplifying assumption is common in the literature (e.g., Jiang et al. 2017, Sun et al. 2011b) - after the restoration has concluded costly units that were only energized for restoration purposes will be turned off again based on optimal dispatch models. (1d) defines the cranking variable (a generator is cranking if and only if it is currently energized but was not energized T_{CR_g} time units in the past).

(1e) and (1f) impose a limit on the rate of the energization. (1e) imposes that we can only energize a branch, if at least one of its nodes was energized at the previous time step (see Jiang et al. 2017). (1f) is a physical constraint that comes from the fact that the operator has finite resources (people) to restore the system. We assume that each branch energization takes some time to happen (time that has to do with ensuring the feasibility of the energization as well as performing the necessary operations to implement the switching). K_{crew} indicates the number of branch energizations per time step. De-energizations of branches are usually faster, thus we do not limit the number of de-energizations per time step. Note that $(x)^+ = \max\{0, x\}$.

Constraint (1g) stipulates that nodes corresponding to generators that are cranking must be energized (except if the generator is a black start). Constraints (1h) -(1l) impose what we refer to as the IE requirement. In particular, (1h) imposes that a generator that has finished cranking energizes the node it connects to. (1i) imposes that an energized branch must have both its endpoints energized. Constraints (1j)-(1l) impose the following requirement: Any energized bus must be connected to an energized generator that has completed cranking through a path of energized edges. This constraint comes from the fact that to support the voltage and frequency of every electric island we need at least one operational generator (often a black start unit, but not necessarily) on



(a) Start-up curve for a non black-start unit ($u_{BS_g} = 0$) in the simplified model. After the unit starts at time t_s, it needs to crank by absorbing active power P_{CR_g} from the system for time T_{CR_g}. Following that, it can move at any point between the minimum and maximum generation, $P_g^{\min}$ and $P_g^{\max}$ respectively.



(b) Start-up curve for a black-start unit ($u_{BS_g} = 1$) in the simplified model. After the unit starts at time t_s, it needs to crank for time T_{CR_g}. However, the cranking power does not need to be provided by the system. Following that, generation can move at any non negative point below the maximum generation $P_g^{\max}$.

Figure 1 Generator active power curve for a black-start and a non black-start unit over time. The curves are jointly described by equations (1m)-(1q). For both figures, assuming a generator starts at time $t_s \in T$, the binary variable u_g^t (indicating the start-up of a generator) would be 0 for $t < t_s$ and 1 for $t \geq t_s$. The binary variable $u_{CR_g}^t$ (indicating a generator is cranking, i.e. in preparation to become operational) would be 1 for $t_s \leq t < t_s + T_{CR_g}$, and 0 otherwise.

the island. This single-commodity flow formulation is adapted in many power systems publications to impose a connectivity requirement (e.g. Demetriou et al. 2018, Teymouri et al. 2017, Ding et al. 2015, Ding et al. 2017, Golari et al. 2014, Fan et al. 2012, Qiu et al. 2016). We discuss reformulations of the IE requirement in the next section.

Constraints (1m)-(1q) define the generator model, as described in Figure 1. Constraints (1m)-(1o) define the variable w_g^t, which is equal to $u_g^t(1 - u_{BS_g})$, using the McCormick relaxation. This start-up curve is slightly simplified compared to the start-up curve often found in the literature (e.g., Sun et al. 2011b, Patsakis et al. 2018) in that the ramping constraints are omitted. This change comes from the fact that for large-scale systems we are using a coarse time resolution in our optimization problem. As a result, most of the generating units that are considered for the first steps of restoration are capable to ramp up or down quickly enough to render the ramping constraints redundant, which in turn allows us to use a simpler model based on step functions. We directly model the start-up curves for black start and non black-start units using binary variables.

Constraints (1p)-(1t) define a transportation model for the active power flow. This is in fact a major simplification made for the sake of computational tractability, since the ac transmission switching model that is instead imposed by physics in this problem is omitted. As a result, Kirchhoff's voltage law might not be respected by the solution. However, the set of constraints used ensures, at least, that every island that appears during the restoration process satisfies active power balance (i.e. the total island generation equals the island demand). This incorporates two important restoration considerations: first, for every generator that gets cranked, there is sufficient active power generation in the island to support its cranking power; second, every generator that has finished cranking and needs to keep its generation above the minimum generation point $P_g^{\min}$ can find enough load in the island to do that. Finally, (1u) imposes an aggregate reactive power constraint (the reactive power absorption capability of shunt reactances, reactive load, and generators after they crank is enough to overcome the reactive power generation induced by the branch shunt susceptances of large transmission lines).

As a justification to the power flow constraint simplifications, we should note that we do not expect the BSA solution to generate an ac-feasible restoration sequence that captures all the necessary switching operations. Rather, the goal is to identify good locations for black start units in the graph, while picking black start units with suitable start-up characteristics. For that purpose, a simplified model that captures the main problem characteristics can be sufficient for a planning tool, while retaining tractability. Indeed, in most of the literature for BSA the active power flow constraints in the optimization models are initially either completely omitted and aggregate power models are used (e.g., Sun et al. 2011b, Jiang et al. 2017, Patsakis et al. 2019b), or simplified network flow models are used to ensure consistent partitioning of the graph in sections corresponding to black start units (e.g., Qiu et al. 2016). Generating the actual detailed restoration sequences that respect ac power flow constraints is part of a refinement process in subsequent steps (e.g., Qiu and Li 2017).

5. Reformulating the Island Energization Constraints

5.1. Reformulations

This section studies reformulations and valid inequalities for the IE constraints (1h) -(1l), which were discussed in section 4. Note that due to constraint (1d), the difference $u_g^t - u_{CR_g}^t$ equals $u_g^{t-T_{CR_g}}$. Therefore, we can replace the expression $u_g^t - u_{CR_g}^t$ with $u_g^{t-T_{CR_g}}$ in (1h) and (1j) . The analysis of this section studies the set of constraints for a given $t \in T$. The binary variables involved in these constraints are $\{u_i^t\}_{i\in N}$, $\{u_g^{t-T_{CR_g}}\}_{g\in G}$, and $\{u_{ij}^t\}_{(ij)\in E}$. For convenience, we drop all indices referring to time. We refer to the vector of these variables with $\boldsymbol{u}$.

We denote the feasible region of the points satisfying the IE constraint by H. Constraints (1h) and (1i) are included in all the formulations and are rewritten based on the discussion in the previous paragraph as:

$$u_g \leq u_i,\ g \in G(i),\ i \in N \tag{2a}$$

$$u_{ij} \leq u_i, u_{ij} \leq u_j,\ (ij) \in E \tag{2b}$$

The first formulation (F_1) is given by:

$$F_1 = \{\mathbf{u} \in [0,1]^d : \exists \{f_g\}_{g\in G}, \{f_{ij}\}_{(ij)\in A} : \text{(2a), (2b), (4a)-(4c)}\} \tag{3}$$

where the constraints (1j) -(1l) are used, rewritten as:

$$0 \leq f_g \leq u_g,\ g \in G \tag{4a}$$

$$-u_{ij} \leq f_{ij} \leq u_{ij},\ (ij) \in A \tag{4b}$$

$$\sum_{j:(ji)\in A} f_{ji} - \sum_{j:(ij)\in A} f_{ij} + \sum_{g\in G(i)} f_g = \frac{1}{|N|} u_i,\ i \in N \tag{4c}$$

Given the definition of F_1 above, we define the feasible region of the IE problem to be $H = F_1 \cap \mathbb{B}^d$. In the definition of F_1, a set of auxiliary network flow variables f_g, $g \in G$ and f_{ij}, $(ij) \in A$, are employed and projected out. An energized node (i.e. $u_i = 1$) acts as a sink of $\frac{1}{|N|}$ amount of network flow, captured in the right hand side of (4c). Network flow can only be generated from energized

generators, due to (4a). It can only flow through energized branches due to (4b). This ensures that there is a path from any energized node to an energized generator that uses only energized branches (this is a path that the network flow follows to move from the energized generator to the energized node). The normalization constant $\frac{1}{|N|}$ allows one generator to energize all nodes. Adaptations of the single-commodity flow formulation are the ones most commonly used in power systems applications (e.g. Demetriou et al. 2018, Teymouri et al. 2017, Ding et al. 2015, Ding et al. 2017, Golari et al. 2014, Fan et al. 2012, Qiu et al. 2016).

An alternative formulation approach, following the same logic, would be to consider a different type of flow corresponding to the energization of each node. This would lead to the following multi-commodity flow formulation:

$$F_2 = \{\mathbf{u} \in [0,1]^d : \exists \{f_g^k\}_{k\in N, g\in G}, \quad \{f_{ij}^k\}_{k\in N,(ij)\in A} : \text{(2a), (2b), (6a)-(6d)}\} \tag{5}$$

where:

$$0 \leq f_g^k \leq u_g, \ k \in N, g \in G \tag{6a}$$

$$-u_{ij} \leq f_{ij}^k \leq u_{ij}, \ k \in N, (ij) \in A \tag{6b}$$

$$\sum_{j:(ji)\in A} f_{ji}^k - \sum_{j:(ij)\in A} f_{ij}^k + \sum_{g\in G(i)} f_g^k = u_i, \ k \in N, i \in N : i = k \tag{6c}$$

$$\sum_{j:(ji)\in A} f_{ji}^k - \sum_{j:(ij)\in A} f_{ij}^k + \sum_{g\in G(i)} f_g^k = 0, \ k \in N, i \in N : i \neq k \tag{6d}$$

The idea behind this formulation is that each node is treated separately and is associated with its own type of network flow and constraints. If node $k \in N$ is energized, then one or more of the energized generators needs to generate the type k network flow, which has to pass through energized branches. The only sink for that type of network flow is node k, which means that the network flow of type k is preserved at every other node $i \neq k$. In this case, the normalization of $\frac{1}{|N|}$ is not necessary in (6c), since a single energized generator can generate all $|N|$ types of network flows to energize all the nodes. This formulation is rarely used in power systems due to its size.

A third formulation is a cut-set formulation that only employs the binary variables. Specifically,

$$F_3 = \{\mathbf{u} \in [0,1]^d : (2a),\ (2b),\ (8)\} \tag{7}$$

where:

$$\sum_{(ij)\in\delta(S)} u_{ij} + \sum_{g\in G(S)} u_g \geq u_n,\ n \in S, S \subseteq N \tag{8}$$

We refer to the family of inequalities (8) as *Type I* constraints. The idea behind this formulation is that, given any subset S of the nodes, if any node in that subset is energized (i.e. if u_n in the right hand side of (8) is equal to 1 for some $n \in S$), then an energized generator must be providing the energizing flow. Therefore, either one generator within the set S should be energized (i.e. $\sum_{g\in G(S)} u_g \geq 1$ in the left hand side of (8)), so that the energizing flow comes from that generator, or at least one edge in the cut-set should be energized (i.e. $\sum_{(ij)\in\delta(S)} u_{ij} \geq 1$ in the left hand side of (8)), so that the energizing flow comes from a generator outside the set S. Formulation F_3 employs an exponential number of cuts, but can be separated in polynomial time (see, for example, Patsakis et al. (2019b)).

The three formulations restricted to $\mathbb{B}^d$ represent the same region, i.e. $H = F_1 \cap \mathbb{B}^d = F_2 \cap \mathbb{B}^d = F_3 \cap \mathbb{B}^d$ (see Patsakis et al. (2019b)). When comparing two MILP formulations F_i and F_j, we say that F_i is (strictly) stronger than F_j if F_i is a (strict) subset of F_j. A stronger formulation yields better bounds in the execution of branch and bound (B&B). We have the following result for the relative strength of the three formulations (see Patsakis et al. (2019b)):

PROPOSITION 1. *F_2 is strictly stronger than F_1 and F_2 is as strong as F_3, i.e. $F_2 = F_3 \subset F_1$.*

5.2. NP-hardness

One important question after defining the feasible region of the IE constraints H is whether we can efficiently optimize over that region. The following proposition illustrates that this is hard to know. The proof in appendix EC.1 also draws the connection to RMWCS.

PROPOSITION 2. *The following optimization problem ($\boldsymbol{c} \in \mathbb{Q}^{\boldsymbol{d}}$) is NP-hard:*

$$\begin{aligned} &\underset{\boldsymbol{u}}{\text{maximize}} && \boldsymbol{c}^T\boldsymbol{u} \\ &\text{subject to} && \boldsymbol{u} \in H \end{aligned} \tag{9}$$

5.3. A New Family of Valid Inequalities

After examining convex hull formulations generated by optimization software for power systems with a few buses, we devised the set of valid inequalities of this section to strengthen our formulation. All the statements of this section are shown in appendix EC.2. Define the following function of the binary vector $\boldsymbol{u}$ and of the subset $S \subseteq N$:

$$f_S(\boldsymbol{u}) = \sum_{(ij)\in E: i,j\in S} u_{ij} + \sum_{(ij)\in\delta(S)} u_{ij} + \sum_{g\in G(S)} u_g - \sum_{i\in S} u_i \tag{10}$$

Now, we define the inequalities, which we refer to as *Type II* inequalities, as

$$f_S(\boldsymbol{u}) \geq 0,\ S \subseteq N \tag{11}$$

PROPOSITION 3. *Constraints* (11) *are valid for the feasible region* H.

We also show the following remark:

REMARK 1. The continuous relaxation region defined by the family of Type II cuts (11) is neither stronger nor weaker than that of the family of Type I cuts (8).

Constraints (11) are exponentially many, but fortunately we can efficiently separate them exploiting submodularity, as the following results suggest.

PROPOSITION 4. *For a given, possibly fractional,* $\boldsymbol{u} \in [0,1]^d$, $f_S(\boldsymbol{u})$ *is a submodular function of* $S \subseteq N$.

COROLLARY 1. *Constraints* (11) *can be separated in polynomial time.*

5.4. Polyhedral Analysis

For the analysis in this subsection we assume for simplicity:

ASSUMPTION 1. *The underlying (undirected) graph of the power system* (N, E) *is complete and each node has a single generator.*

This assumption is not realistic for actual power systems, but it allows for an easier analysis, with the aim of getting some insights on the strength of the inequalities introduced in the last two sections. We use these insights in the computations of section 7. The proofs of statements of this section are provided in appendix EC.3.

PROPOSITION 5. *Let assumption 1 hold. The convex hull of the feasible region H, conv(H), is a full-dimensional polyhedron.*

PROPOSITION 6. *Let assumption 1 hold. Constraints* (8) *with* $|S|=1$, *i.e.:*

$$\sum_{j\in N\setminus\{n\}} u_{nj} + u_{g(n)} \geq u_n, \quad n\in N \tag{12}$$

are facet defining for conv(H).

PROPOSITION 7. *Let assumption 1 hold. Constraints* (8) *with* $S=N$, *i.e.:*

$$\sum_{i\in N} u_{g(i)} \geq u_n, \quad n\in N \tag{13}$$

are facet defining for conv(H).

PROPOSITION 8. *Let assumption 1 hold. Constraints* (11) *with* $S=N$, *i.e.:*

$$\sum_{g\in G} u_g + \sum_{(ij)\in E} u_{ij} \geq \sum_{i\in N} u_i \tag{14}$$

are facet defining for conv(H).

6. Reformulating the Active Power Constraints

We next use a reformulation of constraints (1p)-(1t) to project the variables $p_g^t, p_{ij}^t, p_{SH_i}^t$ out of the formulation and hence reduce the number of variables in the optimization. The resulting formulation is equally strong, but in our implementations resulted in a slightly better computational performance for the instances considered. It employs the following set of constraints. For all $t\in T$, $S\subseteq N$:

$$\sum_{i\in S} P_{D_i} u_i^t + \sum_{g\in G(S)} (P_{CR_g} w_g^t - (P_{CR_g}+P_g^{\min}) w_g^{t-T_{CR_g}}) + \sum_{(ij)\in\delta(S)} \bar{S}_{ij} u_{ij}^t \geq 0 \tag{15a}$$

$$\sum_{g\in G(S)} (P_g^{\max}(u_g^t - u_{CR_g}^t) - P_{CR_g}(w_g^t - w_g^{t-T_{CR_g}})) + \sum_{(ij)\in\delta(S)} \bar{S}_{ij} u_{ij}^t \geq 0 \tag{15b}$$

The following proposition ensures the equivalence of the reformulation. The proposition follows from Hoffman's circulation theorem. We present a proof based on the minimum cut - maximum flow theorem in the Appendix instead, in order to also illustrate the process we employ to separate constraints (15) in our experimental results.

PROPOSITION 9. *For a given* $t \in T$*, assume* (1d) *and* (1n) *hold and that the variables* $\{u_g^t\}_{g\in G}, \{u_{CR_g}^t\}_{g\in G}, \{w_g^t\}_{g\in G}, \{w_g^{t-T_{CR_g}}\}_{g\in G}$*,* $\{u_i^t\}_{i\in N}, \{u_{ij}^t\}_{(ij)\in E}$ *are in* $[0,1]$*. Then, the system* (1p)-(1t) *has a solution with respect to* $\{p_g^t\}_{g\in G}$ *,* $\{p_{ij}^t\}_{(ij)\in A}$*,* $\{p_{SH_i}^t\}_{i\in N}$ *if and only if* (15) *hold for all* $S \subseteq N$*.*

7. Experimental Results

All optimization problems in this section were formulated using Python 3 and solved using Gurobi as MILP engine. Each simulation was executed at a single node of the Lawrence Livermore National Laboratory quartz server (2 × Intel Xeon E5-2695, 36 cores, 128 GB RAM), running Linux Red Hat Enterprise 7.8. All the data used in these simulations are available as an online dataset at Patsakis et al. (2020). We experiment with the impact that reformulating the constraints of the problem, as described in the previous two sections, has on multiple instances. We perform two types of simulations. First, we want to evaluate the bound quality due to the difference in the strength of the formulations. Second, we want to compare how the difference in the formulation of the problem impacts the solution times.

7.1. Formulation Strength

As far as the quality of bound is concerned, we compare four possible implementations and solve the continuous relaxation in each case. Since the reformulation of section 6 leads to the same linear relaxation, these experiments compare only the different ways to impose the IE constraint. These implementations are the following:

(i) Formulation F_1.

(ii) Formulation F_3 (Type I cuts), which is equivalent to formulation F_2.

BSA Root	F_1 [%]	F_3 (Type I) [%]	F_1 & Type II [%]	F_3 & Type II [%]	Best Bound
IEEE-39	21.75	15.59	15.35	15.35	2199.15
IEEE-118	6.57	1.89	1.41	1.23	5051.41
Illinois-200	8.45	3.37	2.43	2.43	6480.59
WECC-225	4.13**	1.48	1.47	1.47	16598.51
IEEE-300	16.19	6.47	4.85	4.81	12684.21
South Carolina-500	16.99	4.19	1.07	0.95	13778.61

Table 1 Optimal BSA root node relaxations for synthetic test systems, employing the different formulations. The last column corresponds to the best bound found by the B&B tree in the results of Table 2. Columns 2-5 are percentages above the best upper bound found from B&B (column 6). (** indicates suboptimal termination due to numerical errors in Gurobi.)

(iii) Formulation F_1 with Type II cuts.

(iv) Formulation F_3 (Type I cuts) with Type II cuts.

For the separation of the Type I constraints we use an implementation of the push-relabel algorithm for the maximum flow in the python-igraph package (Csardi et al. 2006). For the separation of the Type II constraints, we implement a cutting plane algorithm for submodular minimization based on minimizing the Lovasz extension (see Atamturk and Narayanan 2019), using linear programming. An accuracy of 0.001 to identify violated constraints and a time limit of 20000s was used - hence, the values presented should be treated as upper bounds of the actual relaxation values.

The results are presented on Table 1. Our first observation is that there is a significant improvement in the quality of the lower bound between using the single-commodity flow formulation F_1 and the formulation with all Type I and II cuts (column F_3 & Type II). A second observation is that the best bound obtained from B&B (last column) in most cases is actually fairly close to the bound obtained by including all Type I and Type II cuts in the continuous relaxation. Finally, the Type II cuts seem to only make a noticeable difference in the strength of the lower bound in

	Gap [%]	UB	LB	Time [s]
IEEE-39				
(A)	Optimal	2221.04	2199.15	2.6
(B)	Optimal	2210.59	2199.15	10
(C)	Optimal	2199.15	2199.15	3.9
(D)	Optimal	2199.15	2199.15	3.2
(E)	Optimal	2206.75	2199.15	1.7
IEEE-118				
(A)	Optimal	5066.87	5018.69	50
(B)	24.9	5101.76	4084.56	2000
(C)	Optimal	5062.64	5018.16	47
(D)	Optimal	5051.41	5019.56	55
(E)	Optimal	5053.76	5018.56	120
Illinois-200				
(A)	Optimal	6480.59	6460.99	147
(B)	–	6517.80	0	2000
(C)	Optimal	6494.89	6460.99	53
(D)	Optimal	6510.60	6446.90	42
(E)	Optimal	6494.54	6460.79	44

	Gap [%]	UB	LB	Time [s]
WECC-225				
(A)	Optimal	16598.51	16574.62	96
(B)*	–	16696.60	0	2000
(C)	Optimal	16604.51	16444.93	58
(D)	Optimal	16610.33	16516.27	71
(E)	Optimal	16610.56	16510.42	53
IEEE-300				
(A)	8.69	12916.52	11883.79	2000
(B)*	–	16989.79	0	2000
(C)	3.98	12807.97	12317.66	2000
(D)	2.02	12773.35	12520.20	2000
(E)	1.18	12684.21	12536.08	2000
South Carolina-500				
(A)	156.47	13987.39	5453.70	2000
(B)*	–	19009.15	0	2000
(C)	8.55	14262.80	13138.36	2000
(D)	4.35	13778.61	13203.37	2000
(E)	5.22	13782.59	13098.70	2000

Table 2 Optimal BSA results for synthetic test systems. Depicting the gap at termination, upper bound (UB), lower bound (LB), and total time.

the larger test systems (IEEE-300 and South Carolina-500), compared to simply using just Type I cuts.

7.2. Solving to Near Optimality

For the second part of simulations, we solve for points over the same feasible region in $\mathbb{Z}^b$, implemented using the following alternatives for imposing the problem constraints:

(A) Formulation F_1 ((1j)-(1l)) imposes the IE constraint and (1p)-(1t) impose the active power constraints.

(B) Formulation F_2 imposes the IE constraint and (1p)-(1t) impose the active power constraints.

(C) Type I constraints (8) are used to impose the IE constraint. Specifically, Type I constraints for $|S| \in \{1, |N|\}$ are added a priori and the remaining constraints are imposed through an integer callback (i.e. whenever an integer point is found by the solver). Since the points examined are integer, separation of Type I cuts can be done simply through a graph traversal (we can find islands that violate the IE constraints in linear time). (1p)-(1t) impose the active power constraints.

(D) Type I constraints (8) and Type II constraints (11) are used to impose the IE requirement. Type I and II constraints for $|S| \in \{1, |N|\}$ are added a priori. We use integer callbacks to lazily add the remaining Type I constraints, only if they are violated at a candidate integer feasible point. For the islands that are identified to violate a constraint, we also add the Type II cuts.

(E) Type I and II constraints are used as in (D) to impose the IE requirement. The system (15) is used to impose the active power constraints. The constraints (15a) and (15b) for $|S| \in \{1, |N|\}$ are added a priori, and the remaining constraints are imposed through an integer callback if violated. Note that for the implementations where Type I and II cuts are used, we always include the constraints for $|S| \in \{1, |N|\}$ a priori in the formulation. This is justified, since (at least for the simplified case considered in the polyhedral analysis of section 5.4), these constraints are facet defining.

A 2000s time limit was set to the solver and a 1% optimality gap termination was considered (i.e. a solution guaranteed to be within 1% of the optimal is denoted as optimal) for all simulations. Presolve was enabled in all simulations. Some instances caused the B&B tree to run out of memory due to the size of the problem. For these instances (indicated with *), we present results using settings that limit the solver's memory use by restricting the number of threads to 2 and using the file system as temporary storage. The results are shown on Table 2.

A first observation is that implementation (B), which uses formulation F_2, is unable to compete with the other implementations due to its size. In fact, for the larger systems (IEEE-300 and South Carolina-500), not even the root node relaxation can be solved within the 2000s time limit. All the other implementations are able to solve the four smaller power systems fairly easily. Additionally,

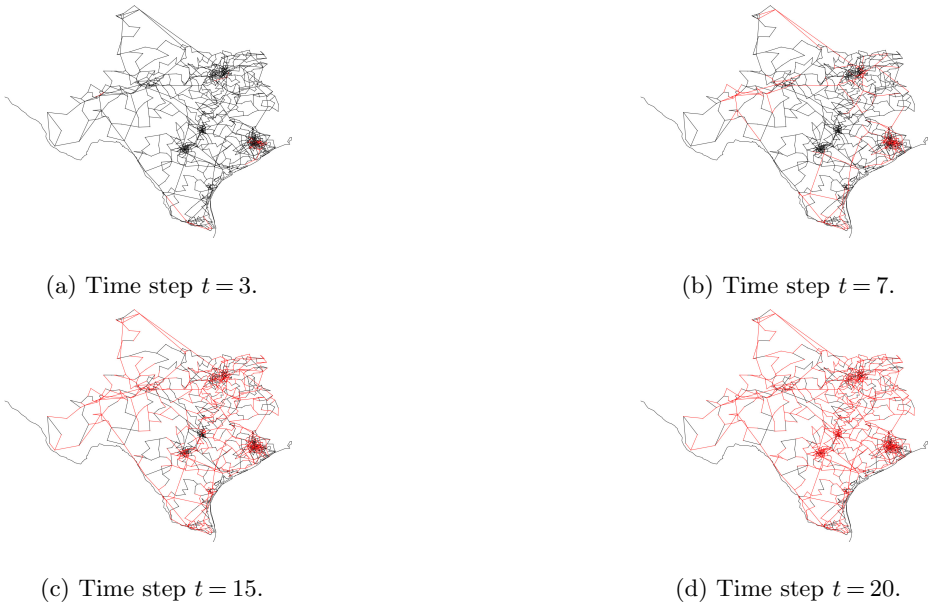
	UB after root	Time after root [s]	UB at time limit	LB at time limit	Gap at time limit
(A)	178217	3519	178217	164112	8.59%
(C)	175326	2429	173725	164112	5.85%
(D)	171850	5667	171479	164112	4.49%
(E)	171805	669	170775	164112	4.06%

Table 3 Simulations in Gurobi for the different implementations for the Texas system BSA. The second column corresponds to the best upper bound found by Gurobi after solving the root node relaxation. Note that presolve is on. The third column corresponds to the time for solving the root node relaxation. The last three columns give the final upper bound, lower bound, and gap at the time limit of 20000s.

we observe that formulation F_1 is outperformed by the cut-based formulations for imposing the IE constraint - it is slightly slower in the Illinois-200 system, and significantly outperformed in the IEEE-300 and South Carolina-500 systems. Type II cuts, employed in (D) and (E), seem to help for the largest two systems in obtaining a better bound within the time limit. Finally, for the same systems, the use of (15) seems to help further close the gap by approximately 1%.

7.3. The Texas System Test Case

The synthetic Texas system consists of 2 000 buses, 3 206 branches, 544 generators. We used a time horizon of 40 time steps, with one time step corresponding to an hour. This instance is the largest one considered in this paper, and in order to solve it we use additional techniques to the ones described in the previous subsections. The system data (buses, branches, generators) and the power flow characteristics were obtained from the data sets of Texas A&M University (Birchfield et al. 2016). The remaining data used are available at Patsakis et al. (2020). The resulting optimization problem (A) has 1 830 671 constraints, 779 178 variables, of which 505 114 are continuous and 274 064 are binary, and a total of 4 977 963 nonzero elements in the constraint matrix.



(a) Time step $t = 3$.

(b) Time step $t = 7$.

(c) Time step $t = 15$.

(d) Time step $t = 20$.

Figure 2 Visualization of four snapshots from the Texas system restoration in the 4.06% optimal solution found. The energized part of the system is indicated with red and the de-energized with black. A system aggregation is used before plotting for better visualization. A total of 9 black starts were assigned.

Using an adaptation of the heuristic strategy described in Patsakis et al. (2018), we were able to obtain a feasible solution with objective 164 113, which was fed as a MIP start in Gurobi. For all simulations the "Method" parameter in Gurobi was set to 3. The "Heuristics" parameter was increased to 0.3 Due to an observed better computational performance, we also set the number of threads to 8 and we introduced constraint (1f) as a lazy constraint, by simultaneously adding a priori the weaker constraint $\sum_{(ij)\in A}(u_{ij}^t - u_{ij}^{t-1}) \leq K_{\text{crew}}$ for each time step t.

Simulation results are shown on Table 3. The implementation that achieves the best optimality guarantee within the time limit is (E). We should note that, in the case of the formulations with exponentially many constraints, these constraints are added through integer callbacks (and not at fractional points), so the linear programming relaxation values provided do not reflect the full strength of the formulation. Finally, a visualization of the solution (for which a 4.06% optimality guarantee was obtained by the best performing (E) implementation) is shown in Figure 2.

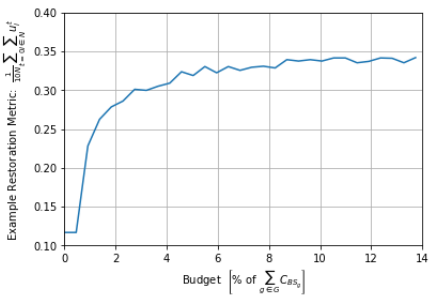


Figure 3 Plot of an example metric as a function of the budget allocated to enhance the black start capability of the simplified WECC system.

7.4. Black Start Capability Upgrade Example

In this section, we provide one potential use of the model we presented for black start planning. We use the WECC-225 system as an example. We assume there is already one black start unit in the system, namely generator 81. The system operator considers potential upgrades to the black start capability of the system. The operator may be interested in different metrics. In our example, we assume one of them is the average of the buses restored over the first 10 time steps, namely $\sum_{t=0}^{9} \sum_{i \in N} u_i^t$. Following that, we can construct a plot of the metric as a function of the allocated budget for black start upgrades - i.e. the optimization model is solved for different values of the parameter B (the unit already allocated is excluded from the budget constraint and set to a black start). The outcome is shown in Figure 3.

Based on that plot, we can see that by investing 2.74% of the total investment cost to convert all units to BSUs ($\sum_{g \in G} C_{BS_g}$), we can get a 157% estimated improvement on the metric of interest. This corresponds to allocating a total of 4 new black start units. Also, based on the plot we notice that, for this particular metric, there seems to be not much additional improvement above investing approximately 4% of $\sum_{g \in G} C_{BS_g}$. We should of course notice that the realization of this improvement needs to be verified with more detailed models and simulations of the restoration process. However, this test case illustrates how the model can serve as the first step of a planning tool to enhance the restoration capability of the power system.

8. Conclusions and Future Research

In this paper we examined the problem of optimal BSA and studied a particular substructure that arises, namely the IE constraints. The allocation of black start generators is the first step in power system restoration planning for large scale outages. It involves evaluating suitable candidates for black starts based on the location and characteristics of the units on the grid, as well as an available budget for the allocation. We presented a model for this problem based on current literature. We then discussed reformulations and valid inequalities for the feasible region of the IE constraints. We provided theoretical results regarding the strength of some of the cuts we introduced, under simplifying assumptions.

We proceeded to perform computations for the optimal BSA using our reformulations on multiple power system instances. We compared the strength of the root node relaxations of the problem under different formulations. We performed computations to solve the systems to near optimality. We observed that the best computational performance comes from a priori including in the formulation selected constraints (which in our simplified polyhedral analysis were shown to be faced defining) and lazily generating the remaining constraints through integer callbacks. Finally, we discussed a use case for the model in planning black start capability upgrades given a budget.

The IE constraints we examined in this paper are implied by a detailed ac model of the power flows (since shunt elements would force node voltages to zero if no generation exists in an electrical island). As a result, they are valid in other power systems applications and can be used to potentially strengthen the formulation and speed up computations, especially if the objective of the problem is similar to the one in our application. Additionally, these constraints are favorable to use in a branch and cut framework - they are easy to separate (linear time for integral points) and have small integral coefficients. For example, we have used these constraints as part of a constraint generation algorithm for the power system restoration problem described in Aravena et al. (2019), employing a detailed approximation of the ac power flows.

There are multiple extensions to our model that would constitute interesting subsequent research directions. One extension would be to upgrade the modeling of the power flow equations to the ac

transmission switching model, or a suitable approximation. In particular, reactive power compensation is one of the most important considerations during the actual restoration process, mainly because of the Ferranti rise effect caused by reactive power injections from large transmission lines in combination with low load conditions. There are two important points when modeling reactive power in restoration problems. First, reactive power is mainly local, since reactive power losses are very high due to large line reactances. Therefore, a lossless reactive power model would not be a suitable choice. On the other hand, real power systems are actually equipped with reactive power compensation devices placed strategically on the grid exactly to alleviate local overvoltages where they are expected to occur. Second, convex relaxations of the power flow equations are often not suitable for modeling reactive power under extreme operational conditions. This is because they lead to fictitious reactive power compensation, as observed in Aravena et al. (2018). Dedicated models that more accurately capture the non-linear nature of reactive power (such as the one discussed in Aravena et al. 2018) are more suitable. Ideas from Aravena et al. (2019) are likely to successfully extend to the framework of this paper. Another model extension would be to relax some of the generator model assumptions. For example, we neglected the generator ramping constraints and assumed generators will not shutdown after their start-up. It is expected that new techniques will be necessary to handle these model refinements while maintaining scalability.

Acknowledgments

Georgios Patsakis is a Fellow of the Onassis Foundation. Ignacio Aravena was supported by the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Professor Shmuel Oren was supported by the ARO grant W911NF-17-1-0555.

References

Álvarez-Miranda E, Ljubić I, Mutzel P (2013a) The maximum weight connected subgraph problem. *Facets of Combinatorial Optimization*, 245–270 (Springer).

Álvarez-Miranda E, Ljubić I, Mutzel P (2013b) The rooted maximum node-weight connected subgraph problem. *International Conference on AI and OR Techniques in Constriant Programming for Combinatorial Optimization Problems*, 300–315 (Springer).

Arab A, Khodaei A, Khator SK, Han Z (2016) Electric power grid restoration considering disaster economics. *IEEE Access* 4:639–649.

Aravena I, Rajan D, Patsakis G (2018) Mixed-integer linear approximations of ac power flow equations for systems under abnormal operating conditions. *2018 IEEE Power & Energy Society General Meeting (PESGM)*, 1–5 (IEEE).

Aravena I, Rajan D, Patsakis G, Oren S (2019) A scalable mixed-integer decomposition approach for optimal power system restoration. Technical report, Optimization Online.

Atamturk A, Narayanan V (2019) Submodular function minimization and polarity. *arXiv preprint arXiv:1912.13238* .

Bienstock D, Mattia S (2007) Using mixed-integer programming to solve power grid blackout problems. *Discrete Optimization* 4(1):115–141.

Biha MD, Kerivin HL, Ng PH (2015) Polyhedral study of the connected subgraph problem. *Discrete Mathematics* 338(1):80–92.

Birchfield AB, Xu T, Gegner KM, Shetye KS, Overbye TJ (2016) Grid structural characteristics as validation criteria for synthetic networks. *IEEE Transactions on power systems* 32(4):3258–3265.

Byeon G, Van Hentenryck P, Bent R, Nagarajan H (2020) Communication-constrained expansion planning for resilient distribution systems. *INFORMS Journal on Computing* .

California ISO (2017) Black start and system restoration Phase 2. `https://www.caiso.com/informed/Pages/StakeholderProcesses/Blackstart_SystemRestorationPhase2.aspx`, [Online; accessed Aug-2017].

Carvajal R, Constantino M, Goycoolea M, Vielma JP, Weintraub A (2013) Imposing connectivity constraints in forest planning models. *Operations Research* 61(4):824–836.

Castillo A (2013) Microgrid provision of blackstart in disaster recovery for power system restoration. *Smart Grid Communications, 2013 IEEE International Conference on*, 534–539 (IEEE).

Chou YT, Liu CW, Wang YJ, Wu CC, Lin CC (2013) Development of a black start decision supporting system for isolated power systems. *IEEE Transactions on Power Systems* 28(3):2202–2210.

Coffrin C, Van Hentenryck P (2015) Transmission system restoration with co-optimization of repairs, load pickups, and generation dispatch. *International Journal of Electrical Power & Energy Systems* 72:144–154.

Csardi G, Nepusz T, et al. (2006) The igraph software package for complex network research. *InterJournal, complex systems* 1695(5):1–9.

Demetriou P, Asprou M, Kyriakides E (2018) A real-time controlled islanding and restoration scheme based on estimated states. *IEEE Transactions on Power Systems* .

Dilkina B, Gomes CP (2010) Solving connected subgraph problems in wildlife conservation. *International Conference on Integration of Artificial Intelligence (AI) and Operations Research (OR) Techniques in Constraint Programming*, 102–116 (Springer).

Ding T, Sun K, Huang C, Bie Z, Li F (2015) Mixed-integer linear programming-based splitting strategies for power system islanding operation considering network connectivity. *IEEE Systems Journal* .

Ding T, Sun K, Yang Q, Khan AW, Bie Z (2017) Mixed integer second order cone relaxation with dynamic simulation for proper power system islanding operations. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 7(2):295–306.

Dittrich MT, Klau GW, Rosenwald A, Dandekar T, Müller T (2008) Identifying functional modules in protein–protein interaction networks: an integrated exact approach. *Bioinformatics* 24(13):i223–i231.

Fan N, Izraelevitz D, Pan F, Pardalos PM, Wang J (2012) A mixed integer programming approach for optimal power grid intentional islanding. *Energy Systems* 3(1):77–93.

Gholami A, Sun XA (2018) Towards resilient operation of multimicrogrids: an misocp-based frequency-constrained approach. *IEEE Transactions on Control of Network Systems* 6(3):925–936.

Golari M, Fan N, Wang J (2014) Two-stage stochastic optimal islanding operations under severe multiple contingencies in power grids. *Electric Power Systems Research* 114:68–77.

Ideker T, Ozier O, Schwikowski B, Siegel AF (2002) Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* 18(suppl_1):S233–S240.

Jiang Y, Chen S, Liu CC, Sun W, Luo X, Liu S, Bhatt N, Uppalapati S, Forcum D (2017) Blackstart capability planning for power system restoration. *International Journal of Electrical Power & Energy Systems* 86:127–137.

Kafka R (2008) Review of PJM restoration practices and NERC restoration standards. *Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE*, 1–5 (IEEE).

Lee HF, Dooly DR (1998) Decomposition algorithms for the maximum-weight connected graph problem. *Naval Research Logistics (NRL)* 45(8):817–837.

Liu S, Podmore R, Hou Y (2012) System restoration navigator: A decision support tool for system restoration. *Power and Energy Society General Meeting, 2012 IEEE*, 1–5 (IEEE).

Lu M, Eksioglu SD, Mason SJ, Bent R, Nagarajan H (2019) Distributionally robust optimization for a resilient transmission grid during geomagnetic disturbances. *arXiv preprint arXiv:1906.04139* .

Minkel J (2008) The 2003 Northeast Blackout–Five Years Later. *Scientific American* 13.

National Infrastructure Advisory Council (2009) *Critical infrastructure resilience: Final report and recommendations* (National Infrastructure Advisory Council (US)).

Patsakis G, Aravena I, Oren S (2020) Black Start Allocation for Power System Restoration Data. URL `http://dx.doi.org/10.5281/zenodo.3842167`.

Patsakis G, Aravena I, Rajan D (2019a) A stochastic program for black start allocation. *Proceedings of the 52nd Hawaii International Conference on System Sciences*.

Patsakis G, Rajan D, Aravena I, Oren S (2019b) Strong mixed-integer formulations for power system islanding and restoration. *IEEE Transactions on Power Systems* 34(6):4880–4888.

Patsakis G, Rajan D, Aravena I, Rios J, Oren S (2018) Optimal black start allocation for power system restoration. *IEEE Transactions on Power Systems* 33(6):6766–6776.

Phillips J, Finster M, Pillon J, Petit F, Trail J (2016) State energy resilience framework. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States).

PJM (2017) *PJM Manual: System Restoration*. Rev. 24.

PJM (2019) *PJM Manual 14D:Generator Operational Requirements*. Rev. 51.

Qiu F, Li P (2017) An integrated approach for power system restoration planning. *Proceedings of the IEEE* 105(7):1234–1252.

Qiu F, Wang J, Chen C, Tong J (2016) Optimal black start resource allocation. *IEEE Transactions on Power Systems* 31(3):2493–2494.

Saraf N, McIntyre K, Dumas J, Santoso S (2009) The annual black start service selection analysis of ERCOT grid. *IEEE Transactions on Power Systems* 24(4):1867–1874.

Sun W, Liu CC, Liu S (2011a) Black start capability assessment in power system restoration. *Power and Energy Society General Meeting, 2011 IEEE*, 1–7 (IEEE).

Sun W, Liu CC, Zhang L (2011b) Optimal generator start-up strategy for bulk power system restoration. *IEEE Transactions on Power Systems* 26(3):1357–1366.

Teymouri F, Amraee T, Saberi H, Capitanescu F (2017) Towards controlled islanding for enhancing power grid resilience considering frequency stability constraints. *IEEE Transactions on Smart Grid* .

Van Hentenryck P, Coffrin C (2015) Transmission system repair and restoration. *Mathematical Programming* 151(1):347–373.

Wang Y, Buchanan A, Butenko S (2017) On imposing connectivity constraints in integer programs. *Mathematical Programming* 166(1-2):241–271.

Wang Y, Chen C, Wang J, Baldick R (2015) Research on resilience of power systems under natural disasters—a review. *IEEE Transactions on Power Systems* 31(2):1604–1613.

Yutian L, Rui F, Terzija V (2016) Power system restoration: a literature review from 2006 to 2016. *Journal of Modern Power Systems and Clean Energy* 4(3):332–341.

Proofs of Statements

The appendix provides the proofs of propositions, corollaries, and remarks in the paper. To avoid confusion, the statements are repeated for the reader's convenience.

EC.1. Proof of Proposition 2.

PROPOSITION 2. *The following optimization problem ($\boldsymbol{c} \in \mathbb{Q}^{\boldsymbol{d}}$) is NP-hard:*

$$\begin{aligned} &\underset{\boldsymbol{u}}{\text{maximize}} \quad \boldsymbol{c}^T \boldsymbol{u} \\ &\text{subject to} \quad \boldsymbol{u} \in H \end{aligned} \tag{9}$$

Proof of Proposition 2. The Rooted Maximum Weight Connected Subgraph Problem (RMWCS), with weights on edges and nodes, known to be NP-hard (see Ideker et al. 2002), can be reduced to (9). The statement of RMWCS is as follows: consider an undirected graph $\hat{G}$ with node set $\hat{N}$, edge set $\hat{E}$, and a given node $r \in \hat{N}$. Weights are associated with nodes and edges of the graph. We seek to find a maximum weight connected subgraph that contains r.

We partition the set of edges $\hat{E}$ into two sets, the set of neighboring edges to r, $\hat{E}(r)$, and $\hat{E} \setminus \hat{E}(r)$. Now consider a power system with generator set $\hat{E}(r)$, bus set $\hat{N} \setminus \{r\}$, and branch set $\hat{E} \setminus \hat{E}(r)$. The generator objective weights are set equal to the weights of edges $\hat{E}(r)$, the branch weights equal to the weights of the set $\hat{E} \setminus \hat{E}(r)$, and the bus weights equal to the weights of $\hat{N} \setminus \{r\}$ accordingly. We establish a 1-1 mapping between rooted connected subgraphs of $\hat{G}$ and feasible points of (9), with the same objective value. Due to (2a)-(2b), the solution of (9) corresponds to a subgraph of the original graph $\hat{G}$ (if we select an edge, we select both its endpoints). The all-zero point in (9) corresponds to the case where the subgraph of $\hat{G}$ chosen is the singleton $\{r\}$. In any nonzero feasible solution to (9), there exists a path from every energized node to at least one generator in $\hat{E}(r)$, which means that the subgraph induced in the original graph is connected and contains the root. Conversely, every connected subgraph of the original graph corresponds to a subgraph of $(\hat{N} \setminus \{r\}, \hat{E} \setminus \hat{E}(r))$, which consists of connected components each one of which has to be connected to r through one of its neighbors $\hat{E}(r)$ - this means that each induced island has at least one energized generator. This completes the reduction. $\square$

EC.2. Proofs of Statements of Section 5.3.

EC.2.1. Proof of Proposition 3.

PROPOSITION 3. *Constraints* (11) *are valid for the feasible region* H.

Proof of Proposition 3. Let $\boldsymbol{u}$ be an feasible point for H (binary vector). We show that for all $S \subseteq N$, constraints (11) are satisfied at $\boldsymbol{u}$. The proof is by induction on the cardinality of the set S. For $|S| = 1$, the constraints reduce to the Type I constraints (8) with $|S| = 1$, which are valid. Now assume that all constraints for sets with $|S| = k$ are valid. Take an arbitrary set of cardinality $k+1$ and assume it has the form $\hat{S} \cup \{i_0\}$, where $i_0 \notin \hat{S}$ is an arbitrary node in that set. By the induction hypothesis we have for the set $\hat{S}$: $f_{\hat{S}}(\boldsymbol{u}) \geq 0$. It suffices to show that $f_{\hat{S}\cup\{i_0\}}(\boldsymbol{u}) \geq 0$.

We consider two cases. First, assume that $f_{\hat{S}}(\boldsymbol{u}) \geq 1$. Using definition (10), we have:

$$f_{\hat{S}\cup\{i_0\}}(\boldsymbol{u}) - f_{\hat{S}}(\boldsymbol{u}) = \sum_{j \in N\setminus(\hat{S}\cup\{i_0\}):(i_0 j)\in E} u_{i_0 j} + \sum_{g\in G(\{i_0\})} u_g - u_{i_0} \tag{EC.1}$$

Due to the binary nature of the variables in the right hand side of equation EC.1 we have $f_{\hat{S}\cup\{i_0\}}(\boldsymbol{u}) - f_{\hat{S}}(\boldsymbol{u}) \geq -1$. Since $f_{\hat{S}}(\boldsymbol{u}) \geq 1$, we obtain $f_{\hat{S}\cup\{i_0\}}(\boldsymbol{u}) \geq 0$.

Now consider the case that $f_{\hat{S}}(\boldsymbol{u}) = 0$. If $u_{i_0} = 0$, from the right hand side of (EC.1), we obtain $f_{\hat{S}\cup\{i_0\}}(\boldsymbol{u}) - f_{\hat{S}}(\boldsymbol{u}) \geq 0$, which yields $f_{\hat{S}\cup\{i_0\}}(\boldsymbol{u}) \geq 0$. If $u_{i_0} = 1$, define $\sum_{i\in\hat{S}} u_i = \bar{k}$. Consequently, the following equality holds using (10) and $f_{\hat{S}}(\boldsymbol{u}) = 0$:

$$\sum_{(ij)\in E: i,j\in\hat{S}} u_{ij} + \sum_{(ij)\in\delta(\hat{S})} u_{ij} + \sum_{g\in G(\hat{S})} u_g = \bar{k} \tag{EC.2}$$

Denote $\bar{S}$ the subset of $\hat{S}$ corresponding to energized nodes: $\bar{S} = \{i \in \hat{S} : u_i = 1\}$, with $|\bar{S}| = \bar{k}$. Define the auxiliary undirected graph $\bar{G} = (V_{\bar{G}}, E_{\bar{G}})$. The node set is $V_{\bar{G}} = \bar{S} \cup \{i_0\} \cup \{s\}$, where s is a dummy super-node which corresponds to an aggregation of all generators and of all nodes $N \setminus (\hat{S} \cup \{i_0\})$. To get the edge set $E_{\bar{G}}$: We add an edge between $i \in \bar{S} \cup \{i_0\}$ and $j \in \bar{S} \cup \{i_0\}$ if $u_{ij} = 1$. Between $i \in \bar{S} \cup \{i_0\}$ and s we add $\sum_{j\in N\setminus(\hat{S}\cup\{i\}):(ij)\in E} u_{ij} + \sum_{g\in G(\{i\})} u_g$ edges. Now observe that the value of the left hand side of equation (EC.2) is exactly the number of edges of $\bar{G}$ that have at least one endpoint in the set $\bar{S}$ (note that even though $\bar{G}$ contains no edges associated with

generator variables or edge variables connected to nodes in $\hat{S} \setminus \bar{S}$, these variables are all zero due to inequalities (2a) and (2b) because the nodes in $\hat{S} \setminus \bar{S}$ are de-energized).

Since the vector $\boldsymbol{u}$ is feasible for H, from every energized node i (i.e., node with $u_i = 1$) on graph (N, E), there is a path of energized edges in E leading to a node with an energized generator in G. In the context of graph $\bar{G}$, that means there exists a path from every node in $\bar{S} \cup \{i_0\}$ to node s. That implies that graph $\bar{G}$ is connected. The graph has $\bar{k}+2$ nodes, so it must have at least $\bar{k}+1$ edges. Exactly $\bar{k}$ edges have at least one endpoint in $\bar{S}$, due to (EC.2). Therefore, there is at least one edge with no endpoints in $\bar{S}$, i.e. there is an edge between i_0 and the dummy node s. By the construction of the edge set of the graph $\bar{G}$, this implies $\sum_{j \in N \setminus (\hat{S} \cup \{i_0\}):(i_0 j) \in E} u_{i_0 j} + \sum_{g \in G(\{i_0\})} u_g \geq 1$, which together with (EC.1) and $f_{\hat{S}}(\boldsymbol{u}) = 0$ leads to $f_{\hat{S} \cup \{i_0\}}(\boldsymbol{u}) \geq 0$. This concludes the proof by induction. $\square$

EC.2.2. Proof of Remark 1.

REMARK 1. The continuous relaxation region defined by the family of Type II cuts (11) is neither stronger nor weaker than that of the family of Type I cuts (8).

Proof of Remark 1. The family of Type II cuts is neither stronger nor weaker than the family of Type I cuts we introduced before. To see that, first consider a graph with two nodes $N = \{1, 2\}$, one branch $E = \{(1, 2)\}$, and two generators, $g(1)$ connected to node 1 and $g(2)$ connected to node 2. Consider the point $(u_1, u_2, u_{12}, u_{g(1)}, u_{g(2)}) = (1, 1, 1/2, 1/2, 1/2)$. This point satisfies all the Type I constraints. However, it violates the Type II constraint for $S = \{1, 2\}$, i.e. $u_{g(1)} + u_{g(2)} + u_{12} \geq u_1 + u_2$.

Conversely, consider the graph with $N = \{1, 2, 3\}$, edges $E = \{(1, 2), (1, 3), (2, 3)\}$ and one generator $g(1)$ connected to node 1. The point: $(u_1, u_2, u_3, u_{12}, u_{13}, u_{23}, u_{g(1)}) = (1/2, 1/2, 1/2, 1/2, 1/2, 1/2, 0)$ satisfies all constraints of Type II, but violates $u_{g(1)} \geq u_1$, the Type I constraint for $S = N$ for node $1 \in S$. $\square$

EC.2.3. Proof of Proposition 4.

PROPOSITION 4. *For a given, possibly fractional, $\boldsymbol{u} \in [0,1]^d$, $f_S(\boldsymbol{u})$ is a submodular function of $S \subseteq N$.*

Proof of Proposition 4. Consider two sets $A, B \subseteq N$ with $A \subseteq B$ and $i_0 \notin B$. We show that $f_{A\cup\{i_0\}}(\boldsymbol{u}) - f_A(\boldsymbol{u}) \geq f_{B\cup\{i_0\}}(\boldsymbol{u}) - f_B(\boldsymbol{u})$. Using (EC.1), the above inequality reduces to:

$$\sum_{j\in N\setminus(A\cup\{i_0\}):(i_0 j)\in E} u_{i_0 j} + \sum_{g\in G(\{i_0\})} u_g - u_{i_0} \geq \sum_{j\in N\setminus(B\cup\{i_0\}):(i_0 j)\in E} u_{i_0 j} + \sum_{g\in G(\{i_0\})} u_g - u_{i_0} \qquad \text{(EC.3)}$$

which is satisfied, since $A \subseteq B \implies N \setminus (A \cup \{i_0\}) \supseteq N \setminus (B \cup \{i_0\})$ and the variables u_{ij} are non negative. □

EC.2.4. Proof of Corollary 1.

COROLLARY 1. *Constraints* (11) *can be separated in polynomial time.*

Proof of Corollary 1. Given a fractional point $\boldsymbol{u}$, minimizing $f_S(\boldsymbol{u})$ over $S \subseteq N$ can be achieved in polynomial time due to submodularity (see Atamturk and Narayanan 2019). If the optimal objective is negative, the minimizer S^* yields a violated constraint from (11). If the minimizer is non-negative, (11) is satisfied for all $S \subseteq N$. □

EC.3. Proof of Propositions of Section 5.4.

For the results of this section, we assume that the graph (N, E) is complete and there is one generator on every bus (assumption 1). For notational convenience, we assume an arbitrary ordering of the nodes $\{1, \ldots, |N|\}$ (where the corresponding generators are $\{g(1), \ldots, g(|N|)\}$).

EC.3.1. Proof of Proposition 5.

PROPOSITION 5. *Let assumption 1 hold. The convex hull of the feasible region H, conv(H), is a full-dimensional polyhedron.*

Proof of Proposition 5. The convex hull is a polyhedron since the set of feasible points is finite. We show that its dimension is $d = |N| + |E| + |G| = 2|N| + |E|$ by identifying $d+1$ affinely independent points. Since the point $\boldsymbol{u} = 0$ is feasible, it suffices to find d linearly independent feasible points. Consider the following points:

1. $\boldsymbol{u}^{a_i}, i \in N$: Only node i and the corresponding generator $g(i)$ are energized. ($|N|$ points)

2. $\boldsymbol{u}^{b_{ij}}, (ij) \in E$: Only two nodes i, j, the corresponding generators, and the edge connecting them are energized. ($|E|$ points)

3. $\boldsymbol{u}^{c_{1j}}, j \in N \setminus \{1\}$: Only node 1 with generator $g(1)$, node j, and edge $(1j) \in E$ are energized. ($|N|-1$ points)

4. $\boldsymbol{u}^{d}$: Only node 1, node 2 with its generator, and edge (12) are energized. (1 point)

We show that the standard basis for $\mathbb{R}^d$ can be generated using linear combinations of the points above. Hence, the points span $\mathbb{R}^d$, and since there are d of them, they are linearly independent. To generate a vector $\boldsymbol{u}$ with only nonzero coordinate $u_g = 1$, for $g \in G(i), i \in N$: For $i = 1$, we simply need to subtract $\boldsymbol{u}^{d}$ from $\boldsymbol{u}^{b_{12}}$. For $i \neq 1$, we need to subtract $\boldsymbol{u}^{c_{1i}}$ from $\boldsymbol{u}^{b_{1i}}$. To generate a vector $\boldsymbol{u}$ with only nonzero coordinate $u_i = 1$ for $i \in N$, we can simply subtract the standard basis vector with only $u_g = 1, g \in G(i)$, generated previously, from $\boldsymbol{u}^{a_i}$. Finally, to generate a vector with only nonzero coordinate $u_{ij} = 1$, $(ij) \in E$, we form $\boldsymbol{u}^{b_{ij}} - \boldsymbol{u}^{a_i} - \boldsymbol{u}^{a_j}$. This completes the proof. $\square$

EC.3.2. Proof of Proposition 6.

PROPOSITION 6. *Let assumption 1 hold. Constraints* (8) *with* $|S| = 1$*, i.e.:*

$$\sum_{j \in N \setminus \{n\}} u_{nj} + u_{g(n)} \geq u_n, \quad n \in N \tag{12}$$

are facet defining for conv(H)*.*

Proof of Proposition 6. Since the convex hull of H is full-dimensional, it suffices to find d affinely independent points in H that satisfy (12) with equality. One of them is the zero vector, so it suffices to find $d-1$ linearly independent points in H that satisfy (12) with equality. These are:

1. We first consider points that have $u_n = 0$, $u_{g(n)} = 0$, and $u_{nj} = 0, j \in N$. Since these are the only variables in (12), the constraint is satisfied with equality. The variables corresponding to nodes $N \setminus \{n\}$, their generators, and the edges with both their endpoints between them (we refer to these variables collectively by $\boldsymbol{u}_R$) can be chosen freely in a way that satisfies the IE constraints of a graph with $(|N|-1)$ nodes and generators and $(|E|-|N|+1)$ edges. Using the argument in the

proof of proposition 5, we can find $2(|N|-1)+(|E|-|N|+1)=|N|+|E|-1$ feasible linearly independent points in this reduced space, which correspond to linearly independent points in the full space (where we have set the remaining variables equal to zero).

For the remaining points, the variables $\boldsymbol{u}_R$ are set equal to 1. We then choose:

2. For $j \in N \setminus \{n\}$, set $u_n = 1$, $u_{g(n)} = 0$, $u_{nj} = 1$, $u_{nk} = 0, k \in N \setminus \{n, j\}$. This yields $|N|-1$ points that are linearly independent from each other and from all previous points since each one has a nonzero value at a position all previous ones did not have (i.e., u_{nj}).

3. Set $u_n = 1$, $u_{g(n)} = 1$, $u_{nj} = 0, j \in N \setminus \{n\}$. This yields 1 point linearly independent from the previous ones, since it is the only one that has $u_{g(n)} = 1$.

Therefore, we generated $(|N|+|E|-1)+(|N|-1)+1 = 2|N|+|E|-1 = d-1$ linearly independent points in H, which concludes the proof. $\square$

EC.3.3. Proof of Proposition 7.

PROPOSITION 7. *Let assumption 1 hold. Constraints* (8) *with* $S = N$*, i.e.:*

$$\sum_{i \in N} u_{g(i)} \geq u_n, \quad n \in N \tag{13}$$

are facet defining for conv(H)*.*

Proof of Proposition 7. We prove the statement with $n = 1$, for notational convenience. The zero vector satisfies (13) with equality, so we only need to find $d-1$ linearly independent points in H that satisfy (13) with equality. We generate d points (that are linearly dependent) and we show they span a $d-1$ dimensional subspace - hence there exists a subset of $d-1$ points that are linearly independent. These points are:

1. $\boldsymbol{u}^{e_i}, i \in N$: Generator $g(1)$, nodes $1, 2, \ldots, i$, and edges $(1,2), (2,3), \ldots, (i-1, i)$ are energized. The rest of the system is de-energized. Note that $\boldsymbol{u}^{e_1}$ refers to the case where only $\boldsymbol{u}^{e_1}_{g(1)} = 1, \boldsymbol{u}^{e_1}_1 = 1$, and the rest of the network is de-energized. ($|N|$ points)

2. $\boldsymbol{u}^{f_i}, i \in N$: All nodes and edges are energized. Only the generator $g(i)$ (of node i) is energized and the others are not. ($|N|$ points)

3. $\boldsymbol{u}^{h_{ij}}, (ij) \in E$: All nodes are energized. All edges are energized, except edge (ij). Only generator $g(1)$ is energized and the rest are not. ($|E|$ points)

We show the points above span a $d-1$ dimensional space by generating the columns of the following $d \times (d-1)$ full rank matrix:

$$\begin{array}{c} \\ \boldsymbol{u}_G \\ u_1 \\ \boldsymbol{u}_{N\setminus\{1\}} \\ \boldsymbol{u}_E \end{array} \begin{array}{c} \begin{array}{ccc} |G| & |N|-1 & |E| \end{array} \\ \begin{pmatrix} \boldsymbol{I}_{|G|} & \boldsymbol{0} & \boldsymbol{0} \\ \mathbb{1}_{1\times|G|} & \boldsymbol{0} & \boldsymbol{0} \\ 0 & \boldsymbol{I}_{|N|-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_{|E|} \end{pmatrix} \end{array},$$

where $\boldsymbol{I}$ denotes the identity matrix and $\mathbb{1}$ the matrix of all ones. More specifically: to generate a vector with only nonzero coordinate $u_{ij} = 1$, for $(ij) \in E$, we form $\boldsymbol{u}^{f_1} - \boldsymbol{u}^{h_{ij}}$. To generate a vector with only nonzero coordinate $u_i = 1$, $i \in N \setminus \{1\}$, we form the vector $\boldsymbol{u}^{e_i} - \boldsymbol{u}^{e_{i-1}} + \boldsymbol{u}^{h_{i-1,i}} - \boldsymbol{u}^{f_1}$. Finally, we can generate a vector with only nonzero entries $u_{g(i)} = 1, u_1 = 1$, for some $i \in N$, by forming: $\boldsymbol{u}^{e_1} + \boldsymbol{u}^{f_i} - \boldsymbol{u}^{f_1}$. This completes the proof. $\square$

EC.3.4. Proof of Proposition 8.

PROPOSITION 8. *Let assumption 1 hold. Constraints* (11) *with* $S = N$, *i.e.:*

$$\sum_{g\in G} u_g + \sum_{(ij)\in E} u_{ij} \geq \sum_{i\in N} u_i \tag{14}$$

are facet defining for conv(H).

Proof of Proposition 8. We assume without loss of generality that when we refer to an edge $(ij) \in E$, we have $i < j$. The zero vector satisfies (14) with equality, so we only need to find $d-1$ linearly independent points in H that satisfy (14) with equality. These are:

1. $\boldsymbol{u}^{a_i}, i \in N$: Only node i and the corresponding generator $g(i)$ are energized. ($|N|$ points)
2. $\boldsymbol{u}^{k_{ij}}, (ij) \in E$: Only nodes i, j, edge (ij) and generator $g(i)$ are energized. ($|E|$ points)
3. $\boldsymbol{u}^{l_i}, i \in N \setminus \{1\}$: Only nodes $1, i$, edge $(1i)$ and generator $g(i)$ are energized. ($|N|-1$ points)

We show these points span a $d-1$ dimensional subspace by showing we can use them to generate all columns of the full rank $d \times (d-1)$ matrix:

$$\begin{array}{c} \\ \boldsymbol{u}_G \\ u_1 \\ \boldsymbol{u}_{N\setminus\{1\}} \\ \boldsymbol{u}_E \end{array} \begin{array}{c} \begin{array}{ccc} |G| & |N|-1 & |E| \end{array} \\ \begin{pmatrix} \boldsymbol{I}_{|G|} & \boldsymbol{0} & \boldsymbol{0} \\ \mathbb{1}_{1\times|G|} & -\mathbb{1}_{1\times|N|-1} & \mathbb{1}_{1\times|E|} \\ 0 & \boldsymbol{I}_{|N|-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_{|E|} \end{pmatrix} \end{array}.$$

To generate a column with $u_{g(i)} = 1, u_1 = 1$, for some $i \in N$, we form: $\boldsymbol{u}^{l_i} - \boldsymbol{u}^{k_{1i}} + \boldsymbol{u}^{a_1}$ for $i \neq 1$ and $\boldsymbol{u}^{a_1}$ for $i = 1$. To generate a column with $u_1 = -1, u_i = 1$, for $i \in N \setminus \{1\}$, we form $\boldsymbol{u}^{a_i} - \boldsymbol{u}^{a_1} - \boldsymbol{u}^{l_i} + \boldsymbol{u}^{k_{1i}}$. Finally, to generate a column with $u_1 = 1, u_{ij} = 1$, for some $(ij) \in E$, we form $\boldsymbol{u}^{k_{ij}} - \boldsymbol{u}^{a_i} - \boldsymbol{u}^{a_j} + \boldsymbol{u}^{a_1} + \boldsymbol{u}^{l_j} - \boldsymbol{u}^{k_{1j}}$. This concludes the proof. □

EC.4. Proof of Proposition 9.

PROPOSITION 9. *For a given $t \in T$, assume* (1d) *and* (1n) *hold and that the variables* $\{u_g^t\}_{g\in G}, \{u_{CR_g}^t\}_{g\in G}, \{w_g^t\}_{g\in G}, \{w_g^{t-T_{CR_g}}\}_{g\in G}$, $\{u_i^t\}_{i\in N}, \{u_{ij}^t\}_{(ij)\in E}$ *are in* $[0,1]$. *Then, the system* (1p)-(1t) *has a solution with respect to* $\{p_g^t\}_{g\in G}$, $\{p_{ij}^t\}_{(ij)\in A}$, $\{p_{SH_i}^t\}_{i\in N}$ *if and only if* (15) *hold for all* $S \subseteq N$.

Proof of Proposition 9. We will take take an indirect approach in showing the result in order to illustrate the separation algorithm we used in our implementation, which is based on minimum cut. The theorem can alternatively be derived from Hoffman's circulation theorem.

We equivalently transform the network to one with non-negative flows. For that, we define for a given $t \in T$:

$$\hat{P}_g^{t,\max} = P_g^{\max}(u_g^t - u_{CR_g}^t) - P_g^{\min} w_g^{t-T_{CR_g}},\ g \in G \tag{EC.4a}$$

$$\hat{p}_g^t = p_g^t + P_{CR_g} w_g^t - (P_{CR_g} + P_g^{\min}) w_g^{t-T_{CR_g}},\ g \in G \tag{EC.4b}$$

$$\hat{P}_{SH_i}^{t,\max} = P_{D_i} u_i^t,\ i \in N \tag{EC.4c}$$

$$\hat{p}^t_{SH_i} = p^t_{SH_i} - (1-u^t_i)P_{D_i},\ i \in N \tag{EC.4d}$$

$$\hat{P}^t_{D_i} = P_{D_i}u^t_i + \sum_{g\in G(i)} (P_{CR_g}w^t_g - (P_{CR_g}+P^{\min}_g)w^{t-T_{CR_g}}_g),\ i \in N \tag{EC.4e}$$

Since (1d) and (1n) hold, and $0 \le P^{\min}_g \le P^{\max}_g$, it directly follows that $\hat{P}^{t,\max}_g \ge 0$. Now, recall that the edge set A is a set of directed edges that correspond to the branches of the power system by arbitrarily picking a "from" and "to" direction for each branch. We also define a (directed) edge set $\tilde{A}$ based on the edge set A which contains both the arcs (ji) and (ij) for every arc $(ij) \in A$. That way, every flow p^t_{ij} satisfying (1r) in A can be decomposed into two non negative flows in $\tilde{A}$ with $\hat{p}^t_{ij} = \max(0, p^t_{ij})$ and $\hat{p}^t_{ji} = \max(0, -p^t_{ij})$ (and conversely we can obtain a flow for A by setting $p^t_{ij} = \hat{p}^t_{ij} - \hat{p}^t_{ji}$). Using the above definitions, equations (1p)-(1t) are equivalently expressed:

$$0 \le \hat{p}^t_g \le \hat{P}^{t,\max}_g,\ g \in G \tag{EC.5a}$$

$$0 \le \hat{p}^t_{ij} \le \bar{S}_{ij}u^t_{ij},\ 0 \le \hat{p}^t_{ji} \le \bar{S}_{ij}u^t_{ij},\ (ij) \in \tilde{A} \tag{EC.5b}$$

$$0 \le \hat{p}^t_{SH_i} \le \hat{P}^{t,\max}_{SH_i},\ i \in N \tag{EC.5c}$$

$$\sum_{j:(ji)\in\tilde{A}} \hat{p}^t_{ji} - \sum_{j:(ij)\in\tilde{A}} \hat{p}^t_{ij} + \sum_{g\in G(i)} \hat{p}^t_g + \hat{p}^t_{SH_i} = \hat{P}^t_{D_i},\ i \in N \tag{EC.5d}$$

Note the variables $p^t_g, p^t_{ij}, p^t_{SH_i}$ are uniquely defined by the variables $\hat{p}^t_g, \hat{p}^t_{ij}, \hat{p}^t_{SH_i}$ and vice versa via the transformations given above. If we apply the same definitions (EC.4) to (15), we equivalently obtain the following system of equations: For all $S \subseteq N$,

$$\sum_{i\in S} \hat{P}^t_{D_i} + \sum_{(ij)\in\delta(S)} \bar{S}_{ij}u^t_{ij} \ge 0 \tag{EC.6a}$$

$$\sum_{i\in S} \hat{P}^{t,\max}_{SH_i} + \sum_{g\in G(S)} \hat{P}^{t,\max}_g + \sum_{(ij)\in\delta(S)} \bar{S}_{ij}u^t_{ij} \ge \sum_{i\in S} \hat{P}^t_{D_i} \tag{EC.6b}$$

All that remains to show now is that the system (EC.5) is feasible for $\hat{p}^t_g, \hat{p}^t_{ij}, \hat{p}^t_{SH_i}$ if and only if (EC.6) holds for all $S \subseteq N$.

PROPOSITION EC.1. *Under the assumptions of proposition 9, system* (EC.5) *has a solution with respect to* $\{\hat{p}^t_g\}_{g\in G}$, $\{\hat{p}^t_{ij}\}_{(ij)\in A}$, $\{\hat{p}^t_{SH_i}\}_{i\in N}$ *if and only if system* (EC.6) *is satisfied.*

Proof of Proposition EC.1. Assume (EC.5) has a solution with respect to $\{\hat{p}^t_g\}_{g\in G}$, $\{\hat{p}^t_{ij}\}_{(ij)\in A}$, $\{\hat{p}_{SH_i}\}_{i\in N}$. Given a set $S\subseteq N$, by summing equations (EC.5d) for $i\in S$, we obtain:

$$\sum_{(ji)\in\tilde{A}:j\in N\setminus S,i\in S}\hat{p}^t_{ji}+\sum_{g\in G(S)}\hat{p}^t_g+\sum_{i\in S}\hat{p}^t_{SH_i}=\sum_{i\in S}\hat{P}^t_{D_i}+\sum_{(ij)\in\tilde{A}:i\in S,j\in N\setminus S}\hat{p}^t_{ij}$$

Due to (EC.5a)-(EC.5c), the left hand side of the above equation is non negative and the right hand side is less than or equal to $\sum_{i\in S}\hat{P}^t_{D_i}+\sum_{(ij)\in\delta(S)}\bar{S}_{ij}u^t_{ij}$. This yields equation (EC.6a). Furthermore, again due to (EC.5a)-(EC.5c), the right hand side of the above equation is greater than or equal to $\sum_{i\in S}\hat{P}^t_{D_i}$ and the left hand side is less than or equal to $\sum_{i\in S}\hat{P}^{t,\max}_{SH_i}+\sum_{g\in G(S)}\hat{P}^{t,\max}_g+\sum_{(ij)\in\delta(S)}\bar{S}_{ij}u^t_{ij}$. This yields equation (EC.6b).

In order to show the reverse implication, we construct the graph of Figure EC.1. The node set consists of the power system nodes N, a source node n_s, a sink node n_t, and a node corresponding to generators n_G. We define $N_{\text{pos}}=\{i\in N:\hat{P}^t_{D_i}>0\}$ and $N_{\text{neg}}=\{i\in N:\hat{P}^t_{D_i}<0\}$. Note that due to (EC.6a) for $S=N$, we have that $\sum_{i\in N}\hat{P}^t_{D_i}\geq 0$.

We add the following directed capacitated edges:

1. From n_s to n_G, with capacity $\sum_{i\in N}\hat{P}^t_{D_i}$.
2. From n_s to $i\in N_{\text{neg}}$, with capacity $-\hat{P}^t_{D_i}$.
3. From $j\in N_{\text{pos}}$ to n_t, with capacity $\hat{P}^t_{D_j}$.
4. From $i\in N$ to $j\in N$ if $(ij)\in\tilde{A}$, with capacity $\bar{S}_{ij}u^t_{ij}$
5. From n_G to $i\in N$ with capacity $P^{t,\max}_{SH_i}+\sum_{g\in G(i)}\hat{P}^{t,\max}_g$.

We now show that, if (EC.6) holds, the minimum n_s-n_t cut in this graph is exactly equal to $\sum_{i\in N_{\text{pos}}}\hat{P}^t_{D_i}$. In order to show that, we show that every cut is greater than or equal to $\sum_{i\in N_{\text{pos}}}\hat{P}^t_{D_i}$. Since the cut $S=\{n_s,n_G,N\}$, $T=\{n_t\}$ has exactly that capacity, it is a minimum cut.

Indeed, we have two cases: If node n_G belongs to T, the capacity of the cut is:

$$\sum_{i\in N}\hat{P}^t_{D_i}+\sum_{(ij)\in\delta(S)}\bar{S}_{ij}u^t_{ij}+\sum_{i\in N_{\text{neg}}\cap T}(-\hat{P}^t_{D_i})+\sum_{i\in N_{\text{pos}}\cap S}\hat{P}^t_{D_i},$$

which is greater than or equal to $\sum_{i\in N_{\text{pos}}}\hat{P}^t_{D_i}$ due to (EC.6a) applied to set S. If node n_G belongs to S, the capacity of the cut is:

$$\sum_{i\in T}P^{t,\max}_{SH_i}+\sum_{g\in G(T)}\hat{P}^{t,\max}_g+\sum_{(ij)\in\delta(T)}\bar{S}_{ij}u^t_{ij}+\sum_{i\in N_{\text{neg}}\cap T}(-\hat{P}^t_{D_i})+\sum_{i\in N_{\text{pos}}\cap S}\hat{P}^t_{D_i},$$

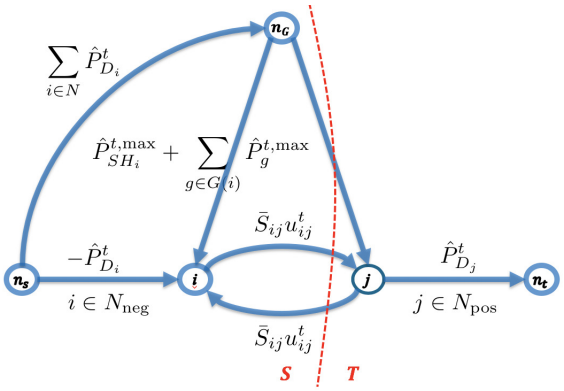


Figure EC.1 Graph used to separate constraints (EC.6).

which is greater than or equal to $\sum_{i\in N_{\text{pos}}} \hat{P}^t_{D_i}$ due to (EC.6b) applied to set T.

Since the minimum cut exactly equals $\sum_{i\in N_{\text{pos}}} \hat{P}^t_{D_i}$, there exists a corresponding maximum flow of the same value on the graph. Since the flows from $i \in N_{\text{pos}}$ to n_t are non negative and upper bounded by $\hat{P}^t_{D_i}$, and the total flow into n_t has to equal $\sum_{i\in N_{\text{pos}}} \hat{P}^t_{D_i}$, we have that the flow from any given $i \in N_{\text{pos}}$ to n_t exactly equals $\hat{P}^t_{D_i}$. Using a similar argument, since $\sum_{i\in N_{\text{pos}}} \hat{P}^t_{D_i} = \sum_{i\in N} \hat{P}^t_{D_i} + \sum_{i\in N_{\text{neg}}} (-\hat{P}^t_{D_i})$, the flow from n_s to $i \in N_{\text{neg}}$ is equal to $(-\hat{P}^t_{D_i})$ and the flow from n_s to n_G is equal to $\sum_{i\in N} \hat{P}^t_{D_i}$. The remaining flows on the graph yield feasible values for $\hat{p}^t_g$, $\hat{p}^t_{ij}$, and $\hat{p}^t_{SH_i}$ in the system (EC.5). Specifically, the flow on $(ij) \in \tilde{A}$ corresponds to $\hat{p}^t_{ij}$ and the flow on the edge from n_G to i corresponds to $\sum_{g\in G(i)} \hat{p}^t_g + \hat{p}^t_{SH_i}$. The flow balance equation for node $i \in N$ yields (EC.5d). □

The constraint set (EC.6) can be separated efficiently using minimum cut on the graph in Figure EC.1 (the details of its construction are provided in the proof of proposition EC.1) . This separation can be used in a Benders-like scheme and is faster in general than solving the linear program (EC.5). Given a point $\{\hat{P}^t_{D_i}\}_{i\in N}$, $\{\hat{P}^{t,\max}_{SH_i}\}_{i\in N}$, $\{u^t_{ij}\}_{(ij)\in E}$, $\{\hat{P}^{t,\max}_g\}_{g\in G}$:

1. If $\sum_{i\in N} \hat{P}^t_{D_i} < 0$, add cut (EC.6a) for $S = N$.
2. If $\sum_{i\in N} \hat{P}^t_{D_i} \geq 0$, create the graph of Figure EC.1:

- If the minimum $n_s - n_t$ cut is less than $\sum_{i\in N_{\text{pos}}} \hat{P}^t_{D_i}$ and n_G belongs to the sink set T, add cut (EC.6a) for the set T.

- If the minimum $n_s - n_t$ cut is less than $\sum_{i \in N_{\text{pos}}} \hat{P}^t_{D_i}$ and n_G belongs to the source set S, add cut (EC.6b) for the source set S.
- If the minimum $n_s - n_t$ cut is equal to $\sum_{i \in N_{\text{pos}}} \hat{P}^t_{D_i}$, all the constraints (EC.6) are satisfied.

The other advantage of having an explicit form for the reformulation (as opposed to only a cut generation scheme) is that we can add some of the constraints a priori, and observe if this improves the quality of bounds. Finally, note that the variables $\hat{P}^{t,\max}_g$, $\hat{P}^{t,\max}_{SH_i}$, $\hat{P}^t_{D_i}$ are only present as parameters in the auxiliary separation problem (EC.6) - the constraints added to the BSA optimization problem are (15), i.e. in the space of the original variables. $\square$