CAREER: Performance Verification of Machine Learning Models Used in Power and Energy System Applications

1. Overview

Machine Learning (ML) is well-positioned to help accelerate the renewable energy transition and manage the increasing complexity of modern power systems. To ensure network safety, ML models used in the electric power grid will require rigorous validation prior to deployment, according to federal reports [1]. However, state-of-the-art ML verification tools are still orders of magnitude behind some of the largest ML models [2] and power grid application needs [3], and existing verification tools are not generally applicable to the actual problems that power engineers need solved [4]. To overcome these hurdles and educate the next generation of ML-aware grid engineers, this CAREER project will design algorithms and computational tools which rigorously verify the performance of ML models built for use in electric power grids. Through (i) a fused modeling paradigm which exploits advances in both ML verification and power system optimization, (ii) creative projections and cuts in the primal and dual verification search spaces, (iii) continuous feedback and collaboration with industry, and (iv) active student education via competition and software development, this project will scale up ML verification state-of-the-art by orders of magnitude while increasing its relevance to actual power industry problems.

1.1 Motivation: Coal, oil and gas met a combined 81% of the global energy demand in 2023 [5]. If global reliance on carbon-based energy is not drastically curtailed, catastrophic consequences could soon emerge [6]. Total decarbonization of the electrical power grid, however, will require an unprecedented deployment of stochastic renewable energy resources [7]. Consequently, the need for fast and secure grid asset coordination and management will present serious algorithmic challenges for the human operators who run the grid. With its ability to make fast, accurate and insightful predictions, Machine Learning (ML) is well-positioned to help overcome these challenges and profoundly disrupt the operation of electric power systems in the coming decades [1, 8, 9]. While this disruption can bring about profoundly positive outcomes for society (e.g., decarbonization), the infusion of ML technologies into the operation of large-scale power networks can also have devastating consequences if the technology is deployed unchecked (akin to the FirstEnergy software bug that drove the 2003 blackout [10], causing 100 deaths).

An emerging solution to this problem is known as **formal verification**. In verification, the properties of an ML model are rigorously tested to ensure the model behaves as its users expect *prior* to deployment [11, 12]. Within the context of power systems, formal verification has focused on numerically proving that an ML model output will always satisfy some minimally acceptable worst-case error bound or constraint violation metric, as compared to the ground-truth physics [4, 13]. There is an emerging desire from regulators (at the Department of Energy (DOE) [1] and the National Institute of Standards and Technology (NIST) [14]) to foster trustworthy and verifiable ML in "safety-critical" contexts. Unfortunately, verification problems can be very hard to formulate (from a modeling perspective) and even harder to solve (from a computational one) [15, 16]. In light of these challenges, this CAREER proposal seeks to answer fundamentally important questions emerging in the verification research community:

- **Question 1:** Can ML models and power grid physics be jointly modeled in a tractable verification framework which exploits advances from both research communities?
- **Question 2:** Has the size of verifiable ML models already saturated? Or can creative and learning-boosted tightening approaches push verification solver scalability to the next orders of magnitude?
- **Question 3:** Can powerful verification tools provide actionable value to the power system industry?

These questions will be investigated, and answered, primarily through methodological innovations which

allow for the solving of large-scale, power industry-relevant verification problems. Accordingly:

This CAREER project will design algorithmic and computational tools which rigorously verify the performance of ML models used in electrical power system operation.

These tools will help the power engineering research and regulatory communities better understand the computational limits of verification, and they will also help exploit key synergies between algorithmic theory and engineering practice (i.e., by aligning what can be done with what needs to be done). Through research advancement and educational outreach, this CAREER project will help unlock new frontiers of research and practice within the growing field of ML model verification for power systems. In synergy with the proposed research, my education plan aims to unite and inspire the next generation of data-driven engineers, who will run tomorrow's power grid, through accessible scientific computing libraries, engaging competitions, and innovative course material.

1.2 Long-Term Career Vision: Industry adoption of ML technologies in safety-critical spaces is still in its infancy, so verification standards and technologies have yet to emerge in full force [17, 18]. Like other emerging fields of research, ML model verification is thus caught in a catch-22, where verification advances depend on the desired applications, but future desired applications also depend on the computational possibilities (for example, if it was *possible* to guarantee that a self-driving vehicle would never hit a pedestrian, then regulators would require all vehicle manufacturers to meet this standard). As depicted in Fig. 1, **my long-term vision is to advance ML model verification state-of-the-art (SoA) by exploiting the synergistic tensions between theory and practice.** In doing so, my research lab at the University of Vermont (UVM) will unlock new frontiers of both fundamental research and practical application, driving actual deployment of verification tools in the power grid and beyond. Advances from my lab will then inform a series of educational and public outreach opportunities, ultimately driving new research ideas and stronger public+policy support for verification technologies. I plan to have my lab emerge as a leader within the field of safety-critical model verification, driving advanced research which pushes US power grids towards a safe and trustworthy embrace of verifiable ML technologies.

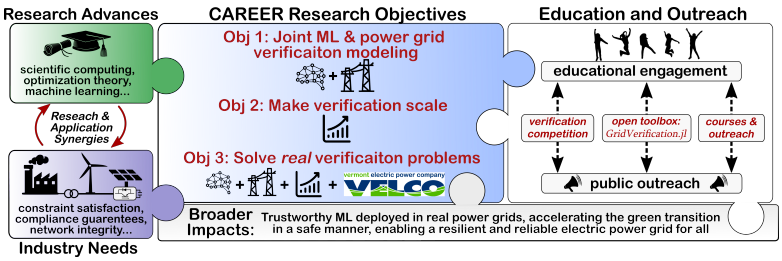


Figure 1: *CAREER research, education, and impact objectives.*

1.3 Why me: ML verification for power systems sits at the nexus of power engineering, mathematical optimization, and ML. Accordingly, this problem cannot be solved by the ML community alone: it will require deep power system modeling intuition and knowledge of actual industry needs. Sitting at the intersection of ML and power system research, I am uniquely well positioned to be successful in this project. As a Marie Curie postdoctoral fellow at the Technical University of Denmark, I pioneered the application of Neural Network (NN) verification to operational power system problems, laying an important foundation of work to build off [4,13,19–21]. Prior to this, I received my PhD from MIT, where my research focused on data-driven modeling and inverse problem theory applied to operational power system problems [22–26]. Finally, my successful participation in the latest ARPA-E Grid Optimization (GO) competition has given me the experience needed to tackle massive-scale network optimization problems [27]. In this competition, my QuasiGrad solver [27] had the 4^{th} highest market surplus function on the largest test case (containing billions of variables), where 8 of the 14 teams could not even find a feasible solution [28]. This experience has inspired me to run my own verification competition ("VerifyGridML") as a key pillar of my educational plan. As an assistant professor at the UVM and a member of the new Center for Energy and Autonomy (CREATE), I have access to world-leading

experts in the fields of power grid engineering, complex systems, and scientific computing. I will also be able to leverage tight regional connections through the Vermont Clean and Resilient Energy Consortium (VCREC) [29], a private-public partnership between UVM and a plethora of local power industry players.

2. Background & Research Gap

Due to the size and complexity of both modern power systems and present-day ML models, emerging verification challenges are inherently *computational* in nature. Accordingly, the verification community has recently pivoted from formal logic and brute-force SAT-solver approaches and instead embraced traditional optimization-based methods (utilizing, e.g., duality, branch-and-cut, semidefinite programming, etc.) [12]. To optimize power system operation and perform network security assessment, grid operators have been using similar optimization tools for decades [30, 31]. The organic connection between the secure and optimal operation of power systems, and the scalable verification approaches emerging from the ML community, has been recently observed [4, 12, 13, 32] but remains under-explored. This CAREER project will exploit the growing connection between these fields and the emerging needs of industry.

2.1 Operational and Computational Power Systems Challenges: Electric power grids are undergoing a rapid transition [33]. Pressing challenges associated with (i) climate change, (ii) affordability, (iii) distributed energy resource (DER) proliferation, and (iv) grid feedback-loop acceleration are driving unprecedented changes in power systems across many timescales [33, 34]. Due to its capacity to learn in complex environments and provide fast predictive solutions, ML (and Artificial Intelligence (AI) in general) is well-posed to help transform power systems in coming decades [35]. Fig. 2 illustrates several *emerging* uses of learning within an ML-aided power system [35–39], including, for example, Optimal Power Flow (OPF) [40, 41] and security assessment [42] prediction. Infusing ML into grid operation has many potential benefits, but it also greatly increases the complexity of the entire system.

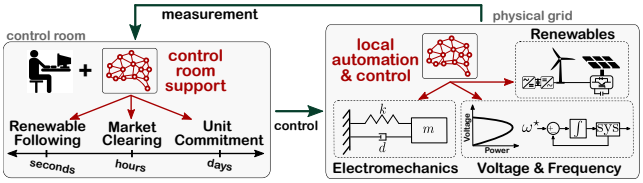


Figure 2: *Emerging uses of ML (red).*

Even without ML integration, computational power system problems can be intractably complex [43]. Power system optimization is rooted in the foundational challenge of reliable and economical grid operation [31], and this field has enjoyed a wonderfully symbiotic relationship with the mathematical programming community. From pushing Mixed Integer solver commercialization (saving billions [44]), to the development of sparse nonlinear network solvers (e.g., KNitro's IPSO tool [45]), power system optimization has a history of pushing SoA forward. Researchers have also developed a suite of fundamental research innovations specifically targeted at globally solving and bounding large-scale power network optimization problems [46], including tight semidefinite programming (SDP) relaxations of OPF [47], Quadratic Convex (QC) relaxations with bound tightening [48, 49], and many others [50–55]. As demonstrated by ARPA-E's GO competitions [56], however, the problem of power system optimization is still largely *unsolved*: reliably finding near-globally optimal and "secure" dispatch solutions on large systems has not yet been achieved (see individual test case results [28] from third GO competition [57]). Given this existing level of computational intractability, verifying the impact of ML-based controllers infused into power system operational loops, as in Fig. 2, may become more challenging still (even if the model's goal was to *decrease* computational burden of system operation).

2.2 Optimization-Based Formal Verification of ML Models: While ML has already enjoyed successful disruption of low-safety-risk industries (e.g., social media), its application in safety-critical systems (e.g., electrical power grids, and other industries where human safety is at stake [58, 59]) has been rightly constrained due to its inherent black-box nature [60]. To overcome these risks, a burgeoning body of research has focused on advancing ML trustworthiness through formal verification and adver-

sarial robustness [11, 12, 15]. Results from such tests present the *strongest* possible claims that can be made about an ML model, definitively answering useful (engineering) questions like, "Does an ML-based controller ever violate known operational limits?", or "Can an ML-based classification (e.g., threat level) change if a situation is perturbed slightly?"

The international Verification of Neural Networks Competition (VNN-Comp)[12, 61] has synergistically inspired a number of breathtakingly successful verification algorithms, e.g., α, β-CROWN [11, 62], Multi-Neuron Guided Branch-and-Bound [63], DeepPoly [64], etc. The winningest methods serve as the bellwether for SoA within the NN verification community. In the literature, verification problems are posed as the minimization of some metric $m(\cdot)$ wrapped around a NN model $\mathrm{NN}(\cdot)$ [11, 16].

$$\gamma = \min_{x \in \mathcal{C}} \; f(x), \; f(x) \triangleq m(\mathrm{NN}(x)). \quad (1)$$

As depicted in Fig. 3, if γ is proved to be everywhere non-negative, the associated problem is **verified** (e.g., the ML is proven to never violate some desirable property). However, if an adversarial counter example is found, then the model fails the verification test. Varying verification methodologies have utilized, e.g., activation function convex relaxations (SDP, Lagrangian, etc.), mixed-integer reformulations, and norm-bounding propagation [15, 65–69]. The winningest algorithms in VNN-COMP, however, have embraced specialized, Graphics Processing Unit (GPU)-accelerated dual-based branch-and-bound formulations [12, 61]. While these methods have scaled to NNs with hundreds of millions of parameters, they are not yet applicable to Large Language Model (LLM) sized-systems [2]. More fundamentally, however, existing verification methodologies are not well suited for solving power system verification problems since physics-based network constraints are not enforceable, network switching cannot be modeled, and existing verifiers have not exploited the many advances made by the power flow optimization community [46].

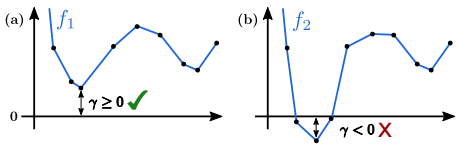


Figure 3: *Verified* $(\mathbf{a})$ *vs falsified* $(\mathbf{b})$.

While emerging verification algorithms have been successful within the domain of ML itself (and have even spun off successful "AI Audit" companies [70]), integration of this work into the verification of realistically sized cyber-physical systems remains nascent [4, 12]. In power systems, NN verification was first considered in [71, 72], where the NNs trained on power system data were reformulated as Mixed Integer Linear Programs (MILPs), via [69], and then compared to the ground truth model, via optimization, to assess constraint violations. These methods were extended to the ACOPF problem in [73], where small verification problems on a 39-bus grid could take up to 5 hours to solve. Recent approaches have thus targeted scalability explicitly. My work has used targeted SDP-based tightening cuts of relaxed NN power flow models [13] and has exploited GPU-based verification routines (i.e., α, β-CROWN [11]) to verify over a massive number of network constraints simultaneously [4]. Others have used GPU-based verifiers to tighten big-M bounds for a CPU-based solver (Gurobi) [74] and exploited scalable optimality verification [32] via projected gradient attack. Despite these efforts, the literature still has not scaled beyond ∼1,000 buses, and to my knowledge, no verification approach has actually been used by the power system industry yet. This project will pioneer scalable grid-aware verification tools which target direct industry engagement and uptake.

2.3 Emerging Policy Directives Related to ML Verification: Governmental organizations and standard setting bodies are beginning to scrutinize the safety of ML and AI-based systems [75]. The Biden Administration recently announced its Executive Order (EO) on Trustworthy AI, which seeks to "establish a plan for global engagement on promoting and developing AI standards", specifically targeting "trustworthiness, verification, and assurance of AI systems" [18]. The EU, through its new AI Act [76], aims to safeguard consumers and boost AI trustworthiness via a wide variety of standards. The DOE, with its recent "AI for Energy Report" [1], is embracing similar priorities, aiming to ensure that ML modes behave "reliably and safely when applied to power grid operations."

Just because governments are waking up to the potential dangers of AI does not mean the existing

technologies are capable of properly safeguarding society. In response to the EO, the National Institute of Standards and Technologies (NIST) has recently kicked off its AI Testing, Evaluation, Validation and Verification (TEVV) program [14], which will help create "guidelines and benchmarks" for AI systems, especially in safety-critical settings where there is potential for "harm" (i.e., to humans or infrastructure) [18, 77]. This has also led to the formation of the Artificial Intelligence Safety Institute Consortium (AISIC) [77], an assembly of industry and research groups want to see AI standards emerge which are both effective and reasonable. The recent EO, for example, indicates that AI may need to meet certification criteria *before* it can be applied in certain safety-critical applications. While this sounds prudent, **formally verifying the performance of ML models embedded within large, multiscale power systems is computationally impossible, given current SoA.** This CAREER project seeks to design tools capable of solving the complex verification problems envisaged by regulators.

2.4 Summarized Research Gap: As demonstrated in the previous three subsections, (i) power system operational and computational challenges are increasing rapidly, and ML infusion will make this complexity increase even more-so, exposing new vulnerabilities; (ii) ML verification methodologies are also maturing rapidly, but they do not yet suit the needs of large-scale power system verification; and (iii) governmental regulators are beginning to strongly push for trustworthy AI standards and regulations, especially within the context of safety-critical applications. Taken together, there clearly exists a need for new, rigorous, use-inspired verification tools which scale to the size of the oncoming challenge.

3. Research Plan

Through this research plan, **I seek to develop a computationally tractable verification framework which will be able to verify if an electrical power system, infused with various levels of ML-based control, can satisfy key security, stability, and operational metrics.** While ML may be used in many facets of emerging power systems, this research plan will focus on steady state operational uses; this includes problems related to state estimation, OPF, voltage stability, congestion management, adversarial threat detection, contingency screening, and so on. The algorithmic innovations advanced in this project, however, will be generally applicable to a much wider class of NN verification problems.

prove: **(a) safety metric**
st: **(b) neural network**
(c) power grid physics

Figure 4: *Verification problem.*

The performance verification problem targeted in this CAREER proposal is a problem of proving some **(a) safety metric**, subject to both **(b) NN mapping** and **(c) power grid physics** constraints, as stated in Fig. 4. As a concrete example, ISO New England, Amazon, and Arizona State researchers have recently developed an ML-based voltage control agent which makes transformer tap and switched shunt control decisions [78]. Before deployment, a pressingly relevant verification problem follows: can the ML agent ever cause a severe network voltage violation? Mathematically, this would be posed by finding, via optimization, the bounded loading condition which yields the worst network **voltage violation**, subject to NN voltage control decisions $G, B = \mathrm{NN}(v, p, q)$ and **AC power flow** constraints.

In the verification literature, the problem posed in Fig. 4 turn into problems of proving *lower performance bounds* [11, 16]. The formulation illustrated in Fig. 3, for example, does not need to be solved to global optimality to be conclusive. It needs to, either, *prove* (e.g., via convex relaxation) that $\gamma \geq 0$, meaning safety is guaranteed, or find a counter example (adversarial input) which disproves the safety metric. To prove lower bounds, this project will exploit optimally tight convex relaxations of (1) via creative projections in the dual variable space, and lifted cuts in the primal space. To apply these advances, I will pioneer a novel power system + NN verification framework, enabling the "sound" (i.e., correct) and "complete" (i.e., conclusive) [16] verification of large-scale, industry relevant problems.

As illustrated in Fig. 5, the research plan itself is divided into three major objectives. **Objective 1** will focus on fusing power system modeling and ML verification into a single coherent framework. **Ob-**

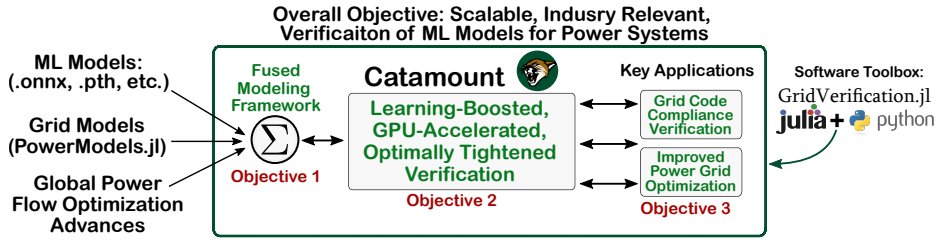


Figure 5: Illustration of key research objectives, all wrapped inside of `GridVerification.jl`.

jective 2 will focus on the algorithmic advances needed to scale power system and ML model verification capabilities to the next orders of magnitude. Finally, **Objective 3** will focus on key applications and use cases, related especially to grid code compliance verification and improved power system optimization. As depicted, these advances will be actively integrated into an accessible software toolbox called `GridVerification.jl` which will be used in my SafeML class to educate students. The verification advances themselves will be rolled into a single, coherent overarching framework called **Catamount**, which will act as the algorithmic heart of the `GridVerification.jl` toolbox. As a competitive researcher and as a curious scholar, extreme-scale verification of power grid models is a challenge which resonates deeply with my passion. I will use this passion to inspire the graduate and undergraduate researchers who I recruit to help drive these advances forward, and in my collaboration with the optimization industry to advance these methods (see letter of collaboration from Hassan Hijazi at Gurobi).

Objective 1: Fusing Power System Optimization and Neural Network Verification.

Modern verification solvers have become increasingly powerful [12], but the generality of the types of problems they are able to solve is highly limited. In other words, ML verification tools are good at solving problems emerging from *within* the ML community, but their direct application to, e.g., the engineering domain remains limited [74, 79, 80]. Thus, modern verifiers cannot solve the full set of problems which power system engineers need solved. More fundamentally, since electric power systems have not been the target application for verification solvers, the verification community has not exploited the large set of innovations pioneered by the power flow research community. This objective, therefore, asks:

Can ML and power grids be jointly modeled in a tractable verification framework?

To answer this question, this objective introduces a verification modeling paradigm which will suit the needs of the power system engineering community (**Obj 1.1**), and it will directly exploit the many global optimization advances made by the power system community in recent years (**Obj 1.2**).

Objective 1.1: Capturing General Power System Models and Constraints in Verification. Many key problems from power systems cannot be directly solved by top-performing [11, 63] verifiers (i.e., ones in VNN-COMP [61]). For example, the following three verification problems cannot be solved:

$$\begin{array}{lll} \textbf{(a) hard constraints:} & \textbf{(b) switching behaviour:} & \textbf{(c) bilevel program:} \\ \min \ f(x) & \min \ f(x,b) & \min \ f(x,y) \\ \text{s.t. } h_1(x) = h_2(x), \text{ etc.} & \text{s.t. } b \in \{0,1\} & \text{s.t. } y = \min \ c(x), \text{ etc.} \end{array} \tag{2}$$

Problem **(a)** enables the verification of NN-based dispatch and control laws within the context of enforced power flow physics; problem **(b)** is crucial for power system models constraining switchable transmission lines or generators [81]; and problem **(c)** enables the computation of optimality gaps associated with NN optimization surrogates [32]. Recent work (via INVPROP [82]) has reverse propagated bounds associated with NN output constraints back to the input. Other work has included AC power flow mappings as an additional set of NN layers [4, 12, 83], thus enabling direct verification. These efforts represent special edge case formulations, however, and do not enable general solutions for (2).

$$\min_{x_0 \in \mathcal{X}} \; f(x_n) \tag{3a}$$
$$\text{s.t.} \; \boldsymbol{h(x_n) = 0} \tag{3b}$$
$$x_n = \text{NN}(x_0). \tag{3c}$$

In this objective, I will design a tractable and flexible framework for solving these problems. In particular, I will build off my approach in [4], which proposed, at its core, "dualizing" various constraints (see problem (**a**)) into a set of generalized NN layers. Consider, for example, the minimization of $f(x)$ subject to an equality constraint $\boldsymbol{h(x_n) = 0}$, followed by a series of sequential NN layer mappings. Without (3b), (3) could be solved using standard verification solver methods [11] due to its sequential structure. To deal with the given constraint, the associated Lagrangian, $\mathcal{L}(x_n, \lambda) = f(x_n) + \lambda^T h(x_n)$, can be treated as an augmented NN output layer, mapping some primal variables through a generalized layer $g(\cdot) = f(\cdot) + \lambda^T h(\cdot)$.

The updated problem in (4) can now be solved using standard bound propagation and dual norm advances [11], which I will exploit. To successfully apply this approach, however, nonconvex constraints (e.g., power flow constraints) will be tightly bounded using optimally tight projections (see next Obj 2.1), and convex and conic constraints will be dealt with via dual cone projections (see next Obj 1.2). Discrete constraints (see problem (**b**)) will be continuously relaxed and successively branched on, and I will explore the application of tightening Gomory cuts. Bilevel optimization constraints (see problem (**c**)) will be dealt with in principally the same manner. The lower level of a bilevel problem can be recast as a set of algebraic constraints using KKT conditions [84], where the associated dual variables become part of the primal variable set. These new KKT *constraints* can be dealt with in the same way that I dealt with $\boldsymbol{h(x_n) = 0}$: introduce new dual variables (μ_*, λ_*), and dualize the KKT constraints.

$$\boldsymbol{\max_{\lambda}} \; \min_{x_0 \in \mathcal{X}} \; f(x_n) + \boldsymbol{\lambda^T h(x_n)} \tag{4a}$$
$$\text{s.t.} \; x_n = \text{NN}(x_0). \tag{4b}$$

The approach used to deal with these three challenging problems (reformulate, dualize, etc.) can be generalized to a very wide class of formulations, allowing me to tackle a diverse variety of power grid verification problems. Most importantly, however, this approach allows me to directly build off highly successful verification methods pioneered by the ML community in recent years [61] (Obj 2.1), *and* the power flow tightening tools that have been developed by power system researchers (Obj 1.2) [46].

Objective 1.2: GPU-Friendly Tightening of Advanced Power Flow Relaxations. The power flow optimization community has seen a number of fantastic advances in recent years, especially within the realm of globally tightening nonconvex OPF-type problems [47–49, 51–55]. While these methods can suffer serious challenges related to primal feasibility [47], they are a perfect match for the verification domain, since *proving lower bounds* is the primary endeavor of verification. Existing methods, however, are typically deigned to run on CPUs (via, e.g., interior point or barrier-based methods [85].)

Accordingly, this objective will leverage the dualized modeling framework from Obj 1.1 to exploit highly targeted global optimization methods on the power flow side. The resulting formulations will be runnable on GPUs, and they will not replace bound propagation-based methods [11, 86] for optimizing over the NN itself (see Obj 2.1). Instead, these methods will help tighten the power flow formulation [13], which will show up as an essential portion of many verification problems (i.e., the **power grid physics** constraint in Fig. 4). Since verification claims require global optimality, I will start by performing convex relaxation of embedded AC power flow equations, followed by a chordal decomposition of the power network [51,52]. I will apply highly selective second-order-cone (SOC), semidefinite programming (SDP), and polynomial cuts [50]; due to the computational expense associated with SDP constraints, efficient "determinant cuts" [54] will be used for further relaxation. In each case, resulting convex constraints "$x \in K$" (where K is a convex cone) will be dualized, such that the associated dual variables become projected into the dual cone K^* [84, 87] and the inner primal incurs a "linear" penalty ($-x^T s$):

$$\textbf{orignal primal:} \; \min_{x} \; f(x) \quad \text{s.t.} \; x \in K \quad \Leftrightarrow \quad \textbf{dualized primal:} \; \max_{s \in K^*} \; \min_{x} \; f(x) - x^T s. \tag{5}$$

Critically, the linear penalty term is fully consistent with backward bound propagation [11], which Objective 2.1 will utilize. I have recently deployed the dual cone projection in (5) to solve large-scale DC-OPF problems with quadratic cost functions [88], and the method exhibits rapid convergence (see Fig. 6). As the dualized verification problem (5) tends towards convergence (i.e., as the gradient-based iterations begin converging towards some unknown upper bound), power flow tightness will be assessed. For non-tight power flow expressions, my previous work on sequentially targeted tightening (STT) [13] will be leveraged to tighten the formulation. In this work, SDP-based RLT cuts were iteratively added to a NN power flow verification problem, sequentially tightening the formulation. After several rounds of tightening, STT could provide a tighter error bound (blue curve in Fig. 7) than Gurobi's nonconvex Mixed Integer Quadratic Program (MIQP) solver (red curve). To further tighten the power flow problem, this objective will explore the use of sequentially lifted variables (selected from Lasserre hierarchies) [89]. This will occur by identifying looseness within in the formulation and dynamically adding new lifted variables (e.g., $z = V_1V_2V_3$), which can yield tightening cuts.

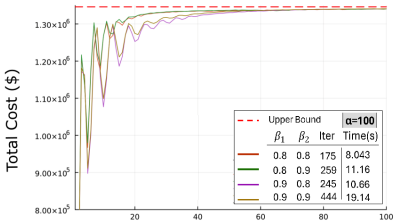


Figure 6: *Gradient convergence of dualized DCOPF. 10k-buses.*

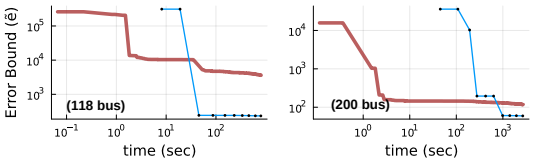


Figure 7: *Performance guarantees by Gurobi (red) vs my STT alg. (blue) [13].*

GPU Implementation: As with other highly successful verification approaches, I will exploit the parallelization capabilities of GPU programming to solve the described formulations. Via dual norm bounding [4, 11, 62, 63], I will transform (5) into a pure dual problem (e.g., $\max_{\tau \in \mathcal{T}} \; d(\tau)$); I will solve this dual with an iterative, gradient-based routine, where parallelized gradient computation occurs on powerful GPUs. At each gradient step, I will enforce dual feasibility $\tau \in \mathcal{T}$ via projection methods (Objective 2.1), thus ensuring a valid lower bound. Ultimately, this framework will allow me to exploit any SDP/polynomial/SOC tightening innovation from the power flow community for the updated purpose of verification.

• **Objective 1: Expected Outcomes and Challenge Mitigation.** This framework should be flexible enough to verify over a broad class of industry-relevant grid verification problems (e.g., verification of NN-based ACOPF [40, 41, 90]), and it will be shared widely with the power research community, via e.g., the Federal Energy Regulatory Commission's Software Workshop. I will also teach this framework in my SafeML class, showing students how to verify network constrained problems. If power flow formulations become overly intractable, I will explore decomposition via Bender's [91] and ADMM [92] approaches. If pure gradient-based routines are converging slowly, I will explore other GPU-amenable approaches, like quasi-Newton (LBFGS) or iterative conjugate gradient [93], which I have used previously [27, 94, 95].

Objective 2: Scalable Verification via Creative Tightening, Learning, and Exploring.

While verification capabilities have scaled rapidly in recent years [61], they are still unable to verify over the largest of NN models [2]. Modern security constrained OPF problems, meanwhile, may contain billions of variables, as with ARPA-E's recent GO competition [3]. Using Obj 1's framework, Obj 2 will focus on the **(b) neural network** constraint of Fig. 4. Here, I will develop fundamental innovations which enable next generation verifiers to scale by orders of magnitude, answering the following question:

Has the size of verifiable ML models saturated? Or can creative and learning-boosted tightening approaches push verification solvers to the next orders of magnitude?

At its core, efficient verification depends on fast methods for proving or disproving lower bounds. To this end, **Objective 2.1** will focus on new verification methodologies which optimally tighten various nonlinear activation function transformations (for bound *proving*), and **Objective 2.2** will design tree-

search methods for identifying adversarial inputs which violate a safety metric (for bound *disproving*). Finally, due to the emerging complexity of the task, **Objective 2.3** will pioneer new, self-supervised learning-based algorithms to accelerate the verification process. All innovations will be fused into a self-consistent algorithmic structure called Catamount; this fusion is depicted in Fig. 8. Catamount will be a general purpose, "sound and complete" [16], GPU-based verifier, and it will Branch & Bound (B&B) over nonlinear activation functions in search of the lower true bound.

Objective 2.1: Optimal Tightening of Non-Convex Activation Functions. Despite verification advances, relatively simple methods (e.g., static linear cuts) are still being used to bound the behavior of nonlinear, non-ReLU activation functions (e.g., sigmoid, tanh) in dual formulations [11, 96]. This objective will employ creative dual variable projections to exploit the *tightest possible* convex relaxations of nonlinear activation functions. Building off of the approach proposed in [11, 86], I will parameterize nonlinear, convexified spaces using some set of tunable parameters τ:

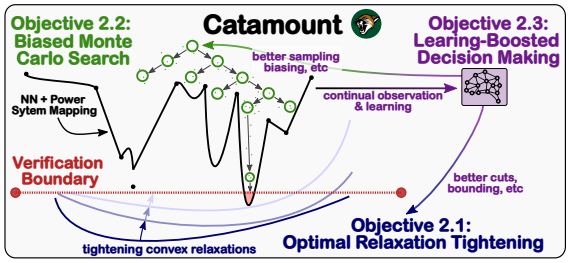


Figure 8: *Objective 2 innovations.*

$$\textbf{Max-Min Verification problem:} \quad \max_{\overline{\tau} \le \tau \le \underline{\tau}} \; \min_{x} \; f(x, \tau), \tag{6}$$

where maximizing over τ gives the *tightest lower bound*. For example, the ubiquitous **sigmoid** activation function $y = \sigma(x) = 1/(1+e^{-x})$ can be upper (or lower) bounded by the linear function $y = \alpha x + \beta$. To ensure that this linear function acts as a true bound, we may analytically solve the pair of equations $\sigma(x) = \alpha x + \beta$ (for interception) and $\sigma'(x) = \alpha$ for slope. In doing so, we get a fascinating, unique relationship between α and β (derivation not shown, nor yet published):

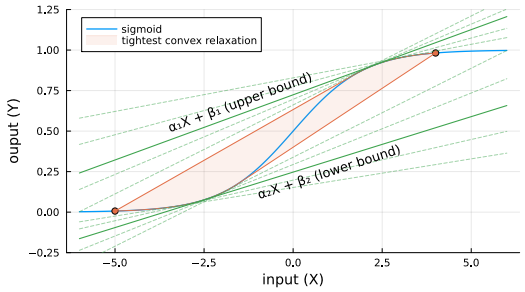


Figure 9: *Tight sigmoid relaxation.*

$$\beta = \frac{1}{1 + e^{-\cosh^{-1}\left(\frac{1}{2\alpha} - 1\right)}} - \alpha \cosh^{-1}\left(\frac{1}{2\alpha} - 1\right). \tag{7}$$

In (7), we have a direct relationship between slope α and intercept β, yielding the green cuts in Fig 9, which smoothly "rotate" around sigmoid's upper epigraph. The dual problem can maximize over α, β, subject to projections of (7), to yield the tightest bound on the sigmoid function in the dual space. Enforcing (7) via projection ensures that $\alpha x + \beta$ acts as a true upper (or when modified, lower) bound on the sigmoid activation function. Interestingly, capturing this tightest convex relaxation of the sigmoid (shaded orange region) using conventional optimization solvers cannot be done without, either, the introduction of integer variables, or approximate surrogate modeling of the sigmoid function. Projected gradient methods, however, are *highly flexible* and can enable optimal tightening of nonconvex spaces. This creative tightening approach can be generalized to many other nonlinear activation function relaxations. For example, the **softmax** operator is a key building block of modern transformers (i.e., attention mechanisms) [97] and can be written as $z = \sigma(x)_i = e^{x_i} / \sum e^{x_j}$. Meanwhile, the **bilinear product**, which is crucially important within power flow modeling, is given as $z = x_i x_j$. Both of these multi-variate functions will be bounded using generalized, tunable hyperplanes given by

$$z_b = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n + \beta, \quad \{\alpha_1, \alpha_2, \ldots, \alpha_n, \beta\} \in \mathcal{F}, \tag{8}$$

where $\mathcal{F}$ represents the feasible set of hyperplane parameters, depending on the nonlinear function in consideration. In some cases, the structure of the set $\mathcal{F}$ will enable explicit projection of α, β into the feasible bounding space (e.g., with the sigmoid (7)). Other cases will require implicit projection.

For example, an explicit solution for the rotatable hyperplane which bounds the softmax function in Fig. 10 is not yet generally known [98]. In these cases, I will use a regularized Newton solve, followed by clipping-based feasibility restoration [99], to approximate an explicit projection. Fig. 10 illustrates my rotatable (green) hyperplanes applied to bilinear and softmax functions, which I am proposing for the first time. Crucially, as long hyperplane *angles* stays within some precomputed bounds, the hyperplane will be a valid over/under-approximator, meaning no part of the nonconvex verification space is lost.

Branch and Bound (B&B): As famously shown by Salman et al., [100], verification relaxation faces a "convex relaxation barrier", where relaxed solutions are typically not strong enough to prove a desired lower bound. To overcome this, I will employ spatial B&B with dual bounding of convex subdomains. This search will solve nonlinear convex Lagrange dual problems via specialized gradient ascent, run in parallel on GPUs. Similar in spirit to recent CROWN work which B&Bs over nonlinear activations with static cuts [96], my work will be the first to solve truly nonlinear, optimally tightened sub-problems in the B&B search tree, massively propelling verification scalability.

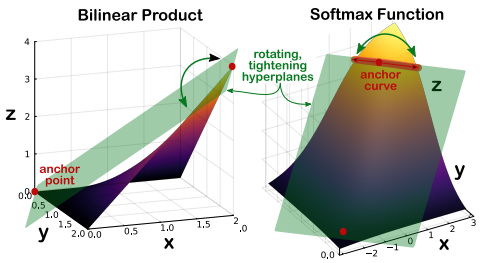


Figure 10: *Bilinear/softmax bounds.*

Objective 2.2: Stochastically-Biased Tree Search for Adversarial Threat Detection. Broadly speaking, Obj 2.1 focuses on raising the lower bound, from below, on verifiable performance, iteratively proving that a relaxed system will never violate a metric. Inspired by recent adversarial attack methods [32, 101], this objective will approach the problem from the opposite end: by finding successively deeper adversarial inputs which systematically approach the true lower bound from above. To execute this attack, a stochastically-biased Monte Carlo Tree Search (MCTS) [102] algorithm will be designed.

On its own, MCTS has enabled some of the most widely notarized ML success stories, from drug discovery to chess/Go victories [103]. At its core, MCTS exploits stochastic sampling of a tree-structured search space, and it takes actions based on the most likely outcome after a series of "playouts". This thrust will fuse the stochastic MCTS method with gradient+bound propagation-based search routines (i.e., search routines which use local information, derived from Obj 2.1, to guide the "optimal" selection of adversarial inputs). Gradient information will bias the MCTS sampling routines toward regions of the search tree which show greatest promise. Enabling this fusion of (a) randomized search with (b) gradient based guiding will help Catamount find dangerous threats, or build confidence in their nonexistence.

Objective 2.3: Learning-Boosted Verification via Self-Supervised Decision Making. Gurobi provides the fastest optimizers in the world, but their documentation admits that successful optimization is "more of an art than a science" [104]; many advances are due to better heuristics, learned via trial and error. Recent work has thus used ML to select better cutting planes for MILP optimization [105], and fused traditional cutting planes (RLT, Gomory, Clique cuts, etc.) with GPU-based verification [74, 106].

Effective verification of ML-infused power systems will require much more than just linear cutting plane selection – it will require a symphony of choices related to variable lifting, targeted bound tightening, determinant cut selection, and branching decisions. Each time a verification problem is solved, or even iterated on, something can be learned. This objective will employ *self-supervised learning agents* to (i) observe mathematical programming decisions (in the way that Gurobi human engineers do), (ii) learn from the successes and failures of different approaches, and then (iii) help drive the decision making process. This approach mimics how a human might build better intuition from on-the-job training, rather than studying from a pre-collected supervised dataset. The potential benefits of this approach are bolstered by inherent risk neutrality: while an agent may make *good* decisions or *bad* decisions, the automated selection process will be engineered such that there are no **wrong** decisions (e.g., only give the learning agent a set of valid cuts to choose from). This objective, therefore, will "fight fire

with fire" by infusing ML deep within the verification problem itself. The agents trained to solve this problem will jointly benefit from the latest in power system-based self-supervised learning [107, 108], and they will use optimization heuristics (e.g., "smart branching", "strong branching" [109], etc.) as a prior foundation, biasing decisions toward known rules until self-supervision suggests otherwise.

• **Objective 2: Expected Outcomes and Challenge Mitigation.** Catamount should be able to verify over NNs with 10s of billions of parameters (i.e., an order of magnitude+ larger than the largest VNN-COMP test cases) and 10s of billions of power grid variables (on par with the largest GO test cases [3]), assuming ∼100GB of GPU memory. If this scalability cannot be achieved, I will exploit observations from my previous work, which showed that small changes in the verification problem can lead to much faster run times [13]. Thus, when a hard problem cannot be solved, I will use dual-based regularization to help accelerate convergence, homotopically "searching" for easier problem versions.

Objective 3: Grid Code Compliance Verification and Improved Optimization Routines.

This third objective will apply Catamount's advanced bound-proving capabilities in a plethora of useful contexts, including for the explicit purpose of dis/proving "grid code" compliance of ML models (**Obj 3.1**), and for the more general purpose of enhancing conventional power system optimization problems via rapid, scalable bound tightening (**Obj 3.2**). This objective will target the question:

Can powerful verification tools provide actionable value to the power systems industry?

Objective 3.1: Grid Code Compliance Verification. Domestic power grids are governed by a plethora of mandated grid codes and standards, developed, by the North American Electric Reliability Corporation (NERC), the IEEE, and other regulators. For example,

- NERC's FAC-011 [110] establishes the need for voltage stability System Operating Limits;
- PRC-006 [111, 112] ensures Under Frequency Load Shedding, to prevent frequency collapse;
- IEEE 1547-2018 [113, 114]) standardizes DER+grid interaction via, e.g, droop voltage control.

These grid codes keep the grid safe, but they also represent some of the *barriers* which keep ML out of the control room. Verifying that an ML-based model does not somehow violate these codes is a nontrivial task, both from a computational perspective, and from a modeling one (e.g., how does one "prove" compliance?). This objective will engage with industry to identify and solve pressing ML-related grid code compliance challenges, thus *ensuring* that ML models adhere to relevant grid codes.

Power Industry Engagement: I will actively seek feedback from utility and system operator partners (see letters of collaboration from Frankie Zhang of ISO New England, Dan Kopin of the Vermont transmission operator (VELCO), and Cyril Brunner of the Vermont Electric Co-Op (VEC)). Each semester, I will invite representatives from these groups to a meeting, where I will demonstrate the verification progress that my lab has made. I will then ask them to answer key questions, like "What are your pressing concerns regarding ML model integrity?" and, "What sort of verification technologies would be most useful for you?" Answers to these questions will help drive applications for `GridVerification.jl`, but they will also help drive its *core algorithmic advances, too*. For example: "ensure that an ML-based Decision Support System [115] suggestion will never violate a transmission line limit constraint."

To pose and solve grid code compliance problems, my approach will leverage Catamount's enhanced proclivity for *bound proving*. Minimum viable targets (i.e., constraint margins) will be mathematically codified and added to a "library" of compliance targets, allowing other users of the tool to test similar compliance problems. Once mathematically codified (via, e.g., conditional value at risk, or standard violation functions), `GridVerification.jl` will use its parallelized verification capabilities to prove that relevant compliance metrics can be mathematically satisfied. Applications related to critical contingency selection (i.e., finding worst-case operating conditions), threat detection (i.e., determining system weaknesses), and cyberattacks (i.e., finding potential min-max attack vectors) will also be explored.

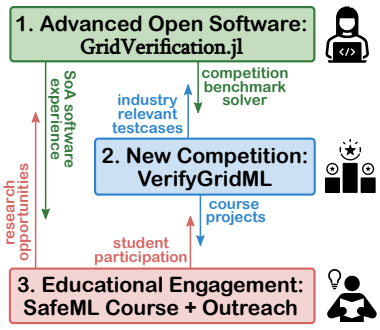
Objective 3.2: Improved Grid Optimization via Parallelized Bound Tightening. Pre-tightening variable bounds has a number of key uses in power system optimization [116]. This includes McCormick envelope relaxation tightening for globally optimal ACOPF [117] and state estimation [118]; or big-M coefficient tightening for e.g., transmission switching [119] or network reduction [120, 121]. Pre-tightening, however, is computationally expensive and can require massive parallelization for concurrent tightening; such parallelization is generally not available on CPUs. Recent work has thus explored using verification solvers for rapidly approximating variable bounds in hard optimization problems [74, 122].

This objective will focus on leveraging `GridVerification.jl`'s inherent parallelization capabilities, along with its flexible modeling architecture, to enable rapid bound tightening of large-scale power-system optimization problems (verification related, or not). Functionally, I will treat a power grid model like a NN mapping, where currents and powers flow through a power network, based on voltage differentials, akin to how information flows through a NN based on activation function status. I will then define a generalized optimization modeling framework which can, in parallel, solve for upper and lower variable bounds associated with pre-selected tightening variables v_i via, e.g., $\max\ v_i$, subject to $h(x) = 0$ and $g(x) \leq 0$. Given problem size, GPU memory availability, and time budget (e.g., 5 minutes), this bound tightening regime will dynamically manage the amount of time spent on each problem, branching and bounding on variables and function spaces as time permits. Bound tightening solutions need not be optimal, only valid. Accordingly, I will explore the fundamental trade-offs between "loosely tightening" all variables, versus "tightly tightening" a select few key variables. I will also test `GridVerification.jl`'s ability to bound sub-problems within a B&B search tree (e.g., each subproblem of DC Unit Commitment is just a DCOPF problem). My work (see Fig. 6, [88]) has successfully explored bounding large-scale DCOPF objectives, solved on GPUs, using explicit verification approaches.

Gurobi Collaboration: Gurobi is designing solutions for embedding ML models inside of otherwise conventional optimization problems (see "Gurobi Machine Learning" [123]). However, these problems can be very hard to solve, requiring hours to optimize over NNs with only 3,000 nonlinear activation functions [124]. I will work with Gurobi engineers to explore importing verification problems into the Gurobi ML package, and I will collaboratively test how their CPU-based solve routines, mixed with my parallelized GPU-based formulations, can accelerate the solve times of their hardest NN-embedded optimization problems (see letter of collaboration from Hassan Hijazi, who is a senior R&D scientist at Gurobi and a world-leading expert in security constrained OPF [48,50,125,126] and verification [12,122]).

• **Objective 3: Expected Outcomes and Challenge Mitigation.** I expect for users outside the verification community to adopt and use the variable bounding tool. Students in my SafeML class will use these bounding tools to quickly tighten convexified model formulations, and to test for grid code compliance. I also expect for `GridVerification.jl` to be used to solve actual industry-inspired problems within 1 year after its release. If industry groups are not able to articulate their grid code compliance needs, I will work with National Lab collaborators at Los Alamos National Lab (LANL) and Pacific Northwest National Lab (PNNL) to create a "verification manifesto" to share with industry.

4. Educational and Outreach Plan



Safety verification of ML models in power systems is a fairly new research area. Therefore, my education and outreach plan will **expose** a new generation of young learners and researchers to this material through 1) open-software software, 2) a grid verification competition, and 3) student-focused course and outreach material (see Fig on right).

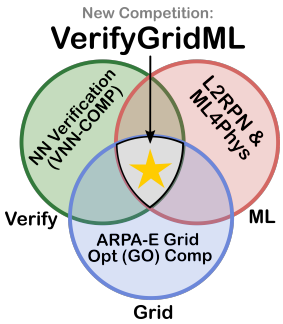
★ **Outreach Pillar 1: Open Source Software.** A primary enabler of power system research has been the proliferation of open source power system software toolboxes (e.g., the DOE-funded PowerMod-

els.jl family [127–129]) and test case libraries [130, 131]. The fusion of advanced modeling libraries with realistically-sized grid models has enabled groups from outside of power, like DeepMind, to impact the field [41], and they have allowed curious students to have sophisticated modeling tools at their fingertips.

Given the popularity and usefulness of emerging generic NN verification libraries (e.g., α, β-CROWN [132], Auto LiRPA [133]), this pillar will focus on fusing my research advances into an accessible, open source, Julia-based [134] "package" for verifying NN models used in power systems (extendable to Python via JuliaPy [135]). The resulting package, called `GridVerification.jl`, will exploit powerful GPU APIs (e.g., CUDA.jl, Metal.jl [136]), and it will fuse seamlessly with, and will be built around, the PowerModels.jl packages. This will allow `GridVerification.jl` to directly exploit the many advanced parsing, modeling, and optimization tools already built into the mature ecosystem of PowerModels.jl. Dr. Carleton Coffrin, chief architect of PowerModels.jl, will support me in my effort to integrate `GridVerification.jl` (see letter of collaboration from Carleton). `GridVerification.jl` will be perpetually advertised on email list-serves (e.g., PowerGlobe) and social media venues to attract users and developers from highly diverse backgrounds.

`GridVerification.jl` be continuously developed as a package which actively reflects the SoA within my research group, but it will be built with *accessibility* in mind (existing verification toolboxes are highly specialized, requiring advanced knowledge of tool functionality and specialized operating systems, which can be overly cumbersome for power engineers). `GridVerification.jl` will be built with a number of challenging, industry-relevant test case examples, allowing users to practice with the software package. Tutorials will show beginners how the tool can be used, and I plan to build several "games" into these tutorials, which challenge users to swiftly and accurately verify network models or find adversarial inputs. This package will be integrated into my SafeML class (Pillar 3) for verifying over advanced ML models, and it will also enable students to have direct research involvement (e.g., by helping build and maintain the package) and SoA software experience. Specifically, students will learn skills like GPU programming and optimized pre-allocation of memory.

• **Pillar 1 Evaluation:** Success will be achieved if the package can (a) steadily reflect SoA within my research group, (b) continuously involve 2-3 undergraduates in its development and testing, (c) and successfully integrate into my SafeML class. **Impact:** This toolbox will accessibly open up advanced verification methodologies to all power system researchers, and it will help the power community take one step closer to actual deployment of data-driven models in the safety-critical realm of power systems. Students in my class at UVM will learn how to verify over extremely large-scale ML models, and they will get to see how physics-based models can be fused with data-driven ones for the purpose of verification.



★ **Outreach Pillar 2: Power System Model Verification Competition.** Across the power system, verification, and ML research communities, competitions have been a successful driver of research innovation. Engaging students, researchers, and industry, these competitions include (1) the ARPA-E GO competition [56, 137], for grid optimization; VNN-Comp [12, 61], for NN verification; and "Learning 2 Run a Power Network" (L2RPN) [138–140]/"ML4Physics" [141], for ML control of power systems. The missing link is a competition which focuses on ML verification for power grids.

Leveraging my experience with these existing competitions, this pillar will design a new competition, called the Verification of Grid ML (**VerifyGridML**) competition, which fills this gap. Designed as a tool to advance academic research, this competition will also excite the growing group of students and scholars who want to work on safe-ML topics for power grid applications. Furthermore, it will focus on solving industry relevant ML verification tasks which explicitly incorporate grid physics. Competition test cases will be formulated with feedback from industry, and the competition framework will serve as a useful tool for designing educational course projects (next pillar).

Finally, `GridVerification.jl` will serve as the *benchmark solver* for the competition. To design the competition, I will solicit help and formal feedback from Dr. Steve Elbert, who led PNNL's effort to design and run the GO competitions [57] (see letter of collaboration from Steve).

The VerifyGridML competition itself with have two tiers: an educational tier, aimed at undergrad and graduate students, and a more advanced research tier, aimed at researchers in the field. The educational tier, which will target 10-15 small teams of students, will be advertised to STEM undergraduates at UVM via the PI's SafeML course, via UVM's local IEEE undergraduate club, and via the local Green Mountain Section chapter of the IEEE Power and Energy Society (I am the secretary). To reach a more inclusive set of students from underrepresented backgrounds, I will also advertise, and seek to actively recruit, through UVM's Mosaic Center for Students of Color, the Prism Center, and the Women & Gender Equity Center. In all advertising, I will draw strong connections between ML verification and the increased deployment of renewable energy (which UVM students are highly passionate about). Winners of the competition will have the chance to present their solution methodologies in front of industry via VCREC [29] locally, and at an invited session of the PES General Meeting. Research shows that competitions can be hugely beneficial for student learning [142]; I will try to tailor the educational tier of the competition towards maximally capturing these benefits (i.e., teamwork, problem solving, and bridging the theory/practice gap [142]). The more competitive research tier will be advertised through PowerGlobe and my connections with organizers of GO, VNN, and L2RPN. This tier will focus on realistically sized, industry motivated challenges *beyond* SoA. For both competitions, a set of rules will be published, along with sample test cases, and teams will submit their verification software solutions prior to a deadline, mimicking the ARPA-E GO competition structure [57].

• **Pillar 2 Evaluation:** The competition will be run annually, and it will be tethered to a domestic summertime conference (most likely, the PES General Meeting). At the associated conference session, results from the competition will be announced, and participating groups will have opportunities to share their solutions. **Impact:** This competition will excite students and actively spur competitive research advances (see, e.g., GO competition innovations [27, 92, 92, 125]). It will also help stimulate a coherent verification research community, and it will show industry our progress towards ML model verification.

⋆ **Outreach Pillar 3: SafeML Course and Rural Outreach.** While there is much excitement about ML at universities across the US, there are few (if any) ML courses which focus on the intersection of engineering application and ML verification. This third pillar will focus on developing a UVM course which will sit at the intersection of ML, power/energy systems, and verification. It will exploit both the `GridVerification.jl` toolbox (for software and algorithm learning) and the VerifyGridML competition framework (for student engagement and project opportunities). Entitled "Safe ML for Engineering" (SafeML), this course will be experiential and project-based, where students have ample opportunity to apply key algorithmic tools on various engineering applications. I will carefully curate all course material (lecture notes, etc.) for productive dispersion across the power teaching communities (similar to courses by Hao Zhu [143], Baosen Zhang [144]). For help with course design, I will work with Dr. Kieran Killeen, Associate Dean for the College of Education at UVM. Kieran will guide me on curriculum development, equitable student engagement, and course evaluation metrics (see letter of collaboration from Kieran). To facilitate interaction with industry, I will recruit speakers from local engineering companies who use ML, so students can hear their challenges and embrace use-inspired projects.

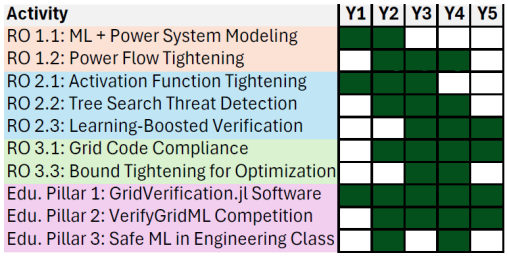
To have impact beyond the classroom, I also plan to interact with UVM Extension's 4-H program [145], which engages young (K-12) and rural audiences from across the state of Vermont. Working with Sarah Kleinman, who is the local 4-H director (see letter of collaboration), I will participate in 4-H's Discover Engineering workshops, where I can showcase the importance of safe ML for, e.g., renewable energy deployment. I will also interact with 4-H's farm safety, robotics, and automation program to showcase the general principles behind ML verification in a fun and exciting way.

• **Pillar 3 Evaluation:** I will use frequent formative assessments (short, in-class quizzes and coding assignments) to track student learning and engagement in the SafeML course, and I will use UVM's new survey-based "ticket home" program to get frequent feedback on course pace, pedagogy and inclusion. **Impact:** The SafeML course will force students to think about verified model performance and trustworthiness, not just as an afterthought, but as an integrated part of the design and testing processes. This course, and associated materials that are shared with 4-H, will engage a diverse array of early learners, and it will expose them to entirely new and exciting ML-related challenges.

5. Research and Education Plan Synergies

In this project, research and education are inextricably linked .The SafeML class will train up a cohort of students to participate in verification research and actively contribute to `GridVerification.jl`. To spur active student engagement with both industry needs and technical algorithmic advances, I will design capstone course projects based on grid code compliance verification. Specifically, I will solicit general needs from local power system industry groups (e.g., ISONE), and I will spin these needs into related capstone course projects. Results from these projects will be built into a special "dev" (i.e., development) corner of the `GridVerification.jl` software package, allowing project results to be shared widely with potential users and developers. Exceptional projects will be encouraged for submission to a relevant conference (e.g., the Conference on Computer Aided Verification). Core algorithmic advances from my lab will be fused into Catamount. The Catamount, which is UVM's mascot, is a sleek and powerful cat, and I plan to use its name and representation to excite and recruit Vermont students into this work. Through this CAREER project, I plan to position Catamount to be competitive with verification solvers and approaches from top-level companies (e.g., DeepMind, LatticeFlow) and universities (e.g., MIT, ETH). Showing students that advanced, ML-safety technology can emerge from the most rural state in the country [146], led by myself, who was born-and-raised in Vermont, will excite and motivate students from diverse and underserved backgrounds to push the boundaries of science and technology.

The VerifyGridML competition will both push SoA and engage a key, emerging group of researchers and students who sit at the intersection of power and ML. I will intentionally market the competition toward students who want to have real-world engineering impact, but who also want to be involved with the push towards ML. The Gantt chart shows the research plan and educational+outreach timeline.



6. Broader Impacts

Presently, the potential for society to benefit from, or be harmed by, ML technology is equally unbounded. This CAREER project seeks to safeguard safety-critical engineering systems from the unchecked risks posed by ML (i.e., *avoid the harms*), and remove the barriers associated with ML deployment in highly beneficial contexts (i.e., *embrace the benefits*). Accordingly, this work *directly* aligns to the White House's EO on Trustworthy AI [18], which seeks to boost the "verification and assurance" of AI systems and foster safe deployment. Enticing students and young researchers into the exciting field of power system ML verification will help train a generation of people who think critically about the dangers of ML, not just in a hypothetical or "sci-fi" sense, but in a literal, "here's why the model might cause damage" sense. My SafeML class will give engineering students the analytical and software tools they need to have impact in the world of engineering + ML application. My advances will help push verification technology out of academic research confines and into the power industry control rooms. This will contribute to a more reliable, green, and equitable electrical power grid for everyone.

7. Results from Prior NSF Support

I have not received prior support from the NSF.

References

[1] K. J. Benes, J. E. Porterfield, and C. Yang, "Ai for energy: Opportunities for a modern grid and clean energy economy," 2024.

[2] L. Sun, Y. Huang, H. Wang, S. Wu, Q. Zhang, C. Gao, Y. Huang, W. Lyu, Y. Zhang, X. Li, Z. Liu, Y. Liu, Y. Wang, Z. Zhang, B. Kailkhura, C. Xiong, C. Xiao, C. Li, E. Xing, F. Huang, H. Liu, H. Ji, H. Wang, H. Zhang, H. Yao, M. Kellis, M. Zitnik, M. Jiang, M. Bansal, J. Zou, J. Pei, J. Liu, J. Gao, J. Han, J. Zhao, J. Tang, J. Wang, J. Mitchell, K. Shu, K. Xu, K.-W. Chang, L. He, L. Huang, M. Backes, N. Z. Gong, P. S. Yu, P.-Y. Chen, Q. Gu, R. Xu, R. Ying, S. Ji, S. Jana, T. Chen, T. Liu, T. Zhou, W. Wang, X. Li, X. Zhang, X. Wang, X. Xie, X. Chen, X. Wang, Y. Liu, Y. Ye, Y. Cao, Y. Chen, and Y. Zhao, "Trustllm: Trustworthiness in large language models," 2024.

[3] "Grid optimziation competition 3, datasets," https://gocompetition.energy.gov/challenges/600650/datasets, accessed: 2024-07-01.

[4] S. Chevalier, I. Murzakhanov, and S. Chatzivasileiadis, "Gpu-accelerated verification of machine learning models for power systems," *Hawaii International Conference on System Sciences*, 2024.

[5] "2024 statistical review of world energy," https://www.energyinst.org/statistical-review, accessed: 2024-07-09.

[6] N. Wunderling, M. Willeit, J. F. Donges, and R. Winkelmann, "Global warming due to loss of large ice masses and arctic summer sea ice," *Nature Communications*, vol. 11, no. 1, p. 5177, 2020.

[7] M. R. Almassalkhi and S. Kundu, "Intelligent electrification as an enabler of clean energy and decarbonization," *Current Sustainable/Renewable Energy Reports*, vol. 10, no. 4, pp. 183–196, 2023.

[8] Y. Chen, X. Fan, R. Huang, Q. Huang, A. Li, and K. P. Guddanti, "Artificial intelligence/machine learning technology in power system applications," Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), Tech. Rep., 2024.

[9] D. Slate, A. Parisot, L. Min, P. Panciatici, and P. Van Hentenryck, "Adoption of artificial intelligence by electric utilities," 2024.

[10] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, N. Martins, J. Paserba, P. Pourbeik, J. Sanchez-Gasca, R. Schulz, A. Stankovic, C. Taylor, and V. Vittal, "Causes of the 2003 major grid blackouts in north america and europe, and recommended means to improve system dynamic performance," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 1922–1928, 2005.

[11] S. Wang, H. Zhang, K. Xu, X. Lin, S. Jana, C.-J. Hsieh, and J. Z. Kolter, "Beta-crown: Efficient bound propagation with per-neuron split constraints for neural network robustness verification," *Advances in Neural Information Processing Systems*, vol. 34, pp. 29 909–29 921, 2021.

[12] C. Brix, S. Bak, C. Liu, and T. T. Johnson, "The fourth international verification of neural networks competition (vnn-comp 2023): Summary and results," *arXiv preprint arXiv:2312.16760*, 2023.

[13] S. Chevalier and S. Chatzivasileiadis, "Global performance guarantees for neural network models of ac power flow," 2024.

[14] "Ai test, evaluation, validation and verification (tevv)," https://www.nist.gov/ai-test-evaluation-validation-and-verification-tevv, accessed: 2024-06-19.

[15] S. Dathathri, K. Dvijotham, A. Kurakin, A. Raghunathan, J. Uesato, R. R. Bunel, S. Shankar, J. Steinhardt, I. Goodfellow, P. S. Liang, and P. Kohli, "Enabling certification of verification-agnostic networks via memory-efficient semidefinite programming," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., vol. 33. Curran Associates, Inc., 2020, pp. 5318–5331. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2020/file/397d6b4c83c91021fe928a8c4220386b-Paper.pdf

[16] S. Wang, *Efficient Neural Network Verification Using Branch and Bound*. Columbia University, 2022.

[17] "Latticeflow ai joins the u.s. ai safety institute consortium," 2024. [Online]. Available: https://latticeflow.ai/news/latticeflow-ai-joins-us-ai-safety-institute-consortium/

[18] "Executive order on the safe, secure, and trustworthy development and use of artificial intelligence," https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/, accessed: 2024-06-18.

[19] J. Stiasny, S. Chevalier, R. Nellikkath, B. Sævarsson, and S. Chatzivasileiadis, "Closing the loop: A framework for trustworthy machine learning in power systems," 2022. [Online]. Available: https://arxiv.org/abs/2203.07505

[20] J. Stiasny, S. Chevalier, and S. Chatzivasileiadis, "Learning without data: Physics-informed neural networks for fast time-domain simulation," in *2021 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. IEEE, 2021, pp. 438–443.

[21] S. Chevalier, J. Stiasny, and S. Chatzivasileiadis, "Accelerating dynamical system simulations with contracting and physics-projected neural-newton solvers," in *Proceedings of The 4th Annual Learning for Dynamics and Control Conference*, ser. Proceedings of Machine Learning Research, R. Firoozi, N. Mehr, E. Yel, R. Antonova, J. Bohg, M. Schwager, and M. Kochenderfer, Eds., vol. 168. PMLR, 23–24 Jun 2022, pp. 803–816. [Online]. Available: https://proceedings.mlr.press/v168/chevalier22a.html

[22] S. Chevalier, P. Vorobev, and K. Turitsyn, "A bayesian approach to forced oscillation source location given uncertain generator parameters," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1641–1649, 2019.

[23] S. Chevalier, L. Schenato, and L. Daniel, "Accelerated probabilistic state estimation in distribution grids via model order reduction," in *2021 IEEE Power & Energy Society General Meeting (PESGM)*, 2021, pp. 1–5.

[24] S. Chevalier, F. M. Ibanez, K. Cavanagh, K. Turitsyn, L. Daniel, and P. Vorobev, "Network topology invariant stability certificates for dc microgrids with arbitrary load dynamics," *IEEE Transactions on Power Systems*, vol. 37, no. 3, pp. 1782–1797, 2022.

[25] S. Chevalier, L. Schenato, and L. Daniel, "Accelerated probabilistic power flow in electrical distribution networks via model order reduction and neumann series expansion," *IEEE Transactions on Power Systems*, vol. 37, no. 3, pp. 2151–2163, 2022.

[26] S. C. Chevalier, P. Vorobev, and K. Turitsyn, "Using effective generator impedance for forced oscillation source location," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6264–6277, 2018.

[27] S. Chevalier, "A parallelized, adam-based solver for reserve and security constrained ac unit commitment," *arXiv preprint arXiv:2310.06650*, 2023.

[28] "Grid optimziation competition 3, event 4 leaderboard," https://gocompetition.energy.gov/challenges/challenge-3/Leaderboards/Event-4, accessed: 2024-06-19.

[29] "Uvm formally announces consortium promoting clean, renewable energy," https://www.uvm.edu/news/story/uvm-formally-announces-consortium-promoting-clean-renewable-energy, accessed: 2024-07-09.

[30] H. W. Dommel and W. F. Tinney, "Optimal power flow solutions," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-87, no. 10, pp. 1866–1876, 1968.

[31] F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R. E. Bohn, *Spot pricing of electricity*. Springer Science & Business Media, 1988.

[32] W. Chen, H. Zhao, M. Tanneau, and P. V. Hentenryck, "Compact optimality verification for optimization proxies," 2024.

[33] N. Hatziargyriou, J. Milanovic, C. Rahmann, V. Ajjarapu, C. Canizares, I. Erlich, D. Hill, I. Hiskens, I. Kamwa, B. Pal, P. Pourbeik, J. Sanchez-Gasca, A. Stankovic, T. Van Cutsem, V. Vittal, and C. Vournas, "Definition and classification of power system stability – revisited & extended," *IEEE Transactions on Power Systems*, vol. 36, no. 4, pp. 3271–3281, 2021.

[34] F. Milano, F. Dörfler, G. Hug, D. J. Hill, and G. Verbič, "Foundations and challenges of low-inertia systems (invited paper)," in *2018 Power Systems Computation Conference (PSCC)*, 2018, pp. 1–25.

[35] J. L. Cremer, A. Kelly, R. J. Bessa, M. Subasic, P. N. Papadopoulos, S. Young, A. Sagar, and A. Marot, "A pioneering roadmap for ml-driven algorithmic advancements in electrical networks," 2024.

[36] B. Donon, R. Clément, B. Donnot, A. Marot, I. Guyon, and M. Schoenauer, "Neural networks for power flow: Graph neural solver," *Electric Power Systems Research*, vol. 189, p. 106547, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779620303515

[37] W. Chen, M. Tanneau, and P. Van Hentenryck, "End-to-end feasible optimization proxies for large-scale economic dispatch," *IEEE Transactions on Power Systems*, vol. 39, no. 2, pp. 4723–4734, 2024.

[38] J. Garland, K. Baker, and B. Livneh, "Weather-induced power outage prediction: A comparison of machine learning models," in *2023 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2023, pp. 1–6.

[39] R. Harris and D. K. Molzahn, "Detecting and mitigating data integrity attacks on distributed algorithms for optimal power flow using machine learning," in *57th Hawaii International Conference on System Sciences, HICSS 2024, Hilton Hawaiian Village Waikiki Beach Resort, Hawaii, USA, January 3-6, 2024*, T. X. Bui, Ed. ScholarSpace, 2024, pp. 3170–3181. [Online]. Available: https://hdl.handle.net/10125/106765

[40] M. Mohammadian, K. Baker, M. H. Dinh, and F. Fioretto, "Learning solutions for intertemporal power systems optimization with recurrent neural networks," in *2022 17th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, 2022, pp. 1–6.

[41] L. Piloto, S. Liguori, S. Madjiheurem, M. Zgubic, S. Lovett, H. Tomlinson, S. Elster, C. Apps, and S. Witherspoon, "Canos: A fast and scalable neural ac-opf solver robust to n-1 perturbations," 2024.

[42] J.-M. H. Arteaga, F. Hancharou, F. Thams, and S. Chatzivasileiadis, "Deep learning for power system security assessment," in *2019 IEEE Milan PowerTech*, 2019, pp. 1–6.

[43] T. Overbye, X. Cheng, and Y. Sun, "A comparison of the ac and dc power flow models for lmp calculations," in *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the*, 2004, pp. 9 pp.–.

[44] M. Cain, R. O'Neill, and A. Castillo, "Optimal power flow papers: Paper 1. history of optimal power flow and formulations," *Federal Energy Regulatory Commission, Tech. Rep*, 2013.

[45] "Optimizing sustainable electric power systems," https://www.artelys.com/solvers/knitro/optimizing-sustainable-electric-power-systems/, accessed: 2024-06-19.

[46] D. K. Molzahn and I. A. Hiskens, "A survey of relaxations and approximations of the power flow equations," *Foundations and Trends® in Electric Energy Systems*, vol. 4, no. 1-2, pp. 1–221, 2019.

[47] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 92–107, 2012.

[48] C. Coffrin, H. L. Hijazi, and P. Van Hentenryck, "The qc relaxation: A theoretical and computational study on optimal power flow," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 3008–3018, 2016.

[49] K. Sundar, H. Nagarajan, S. Misra, M. Lu, C. Coffrin, and R. Bent, "Optimization-based bound tightening using a strengthened qc-relaxation of the optimal power flow problem," in *2023 62nd IEEE Conference on Decision and Control (CDC)*. IEEE, 2023, pp. 4598–4605.

[50] H. Hijazi, C. Coffrin, and P. V. Hentenryck, "Polynomial sdp cuts for optimal power flow," 2015. [Online]. Available: https://arxiv.org/abs/1510.08107

[51] R. A. Jabr, "Exploiting sparsity in sdp relaxations of the opf problem," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 1138–1139, 2012.

[52] A. Eltved, J. Dahl, and M. S. Andersen, "On the robustness and scalability of semidefinite relaxation for optimal power flow problems," *Optimization and Engineering*, vol. 21, no. 2, pp. 375–392, 2020.

[53] D. K. Molzahn and I. A. Hiskens, "Mixed sdp/socp moment relaxations of the optimal power flow problem," in *2015 IEEE Eindhoven PowerTech*, 2015, pp. 1–6.

[54] S. Gopinath, H. Hijazi, T. Weisser, H. Nagarajan, M. Yetkin, K. Sundar, and R. Bent, "Proving global optimality of acopf solutions," *Electric Power Systems Research*, vol. 189, p. 106688, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779620304910

[55] S. Fattahi, M. Ashraphijuo, J. Lavaei, and A. Atamtürk, "Conic relaxations of the unit commitment problem," *Energy*, vol. 134, pp. 1079–1095, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360544217310666

[56] I. Aravena, D. K. Molzahn, S. Zhang, C. G. Petra, F. E. Curtis, S. Tu, A. Wächter, E. Wei, E. Wong, A. Gholami, K. Sun, X. A. Sun, S. T. Elbert, J. T. Holzer, and A. Veeramany, "Recent developments in security-constrained ac optimal power flow: Overview of challenge 1 in the arpa-e grid optimization competition," 2022.

[57] J. T. Holzer, C. J. Coffrin, C. DeMarco, R. Duthu, S. T. Elbert, B. C. Eldridge, T. Elgindy, M. Garcia, S. L. Greene, N. Guo, E. Hale, B. Lesieutre, T. Mak, C. McMillan, H. Mittelmann, H. Oh, R. O'Neill, T. Overbye, B. Palmintier, F. Safdarian, A. Tbaileh, P. Van Hentenryck, A. Veeramany, and J. Wert, "Grid optimization competition challenge 3 problem formulation," Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), Tech. Rep., 2024.

[58] R. kamal Kaur, B. Pandey, and L. K. Singh, "Dependability analysis of safety critical systems: Issues and challenges," *Annals of Nuclear Energy*, vol. 120, pp. 127–154, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S030645491730213X

[59] T. Popik and R. Humphreys, "The 2021 texas blackouts: causes, consequences, and cures," *Journal of Critical Infrastructure Policy*, vol. 2, no. 1, pp. 47–73, 2021.

[60] F. Tambon, G. Laberge, L. An, A. Nikanjam, P. S. N. Mindom, Y. Pequignot, F. Khomh, G. Antoniol, E. Merlo, and F. Laviolette, "How to certify machine learning based safety-critical systems? a systematic literature review," *Automated Software Engineering*, vol. 29, no. 2, Apr. 2022. [Online]. Available: http://dx.doi.org/10.1007/s10515-022-00337-x

[61] C. Brix, M. N. Müller, S. Bak, T. T. Johnson, and C. Liu, "First three years of the international verification of neural networks competition (vnn-comp)," 2023.

[62] Z. Lyu, C.-Y. Ko, Z. Kong, N. Wong, D. Lin, and L. Daniel, "Fastened crown: Tightened neural network robustness certificates," 2019.

[63] C. Ferrari, M. N. Muller, N. Jovanovic, and M. Vechev, "Complete verification via multi-neuron relaxation guided branch-and-bound," *arXiv preprint arXiv:2205.00263*, 2022.

[64] G. Singh, T. Gehr, M. Püschel, and M. Vechev, "An abstract domain for certifying neural networks," *Proceedings of the ACM on Programming Languages*, vol. 3, no. POPL, pp. 1–30, 2019.

[65] E. Wong and Z. Kolter, "Provable defenses against adversarial examples via the convex outer adversarial polytope," in *Proceedings of the 35th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, J. Dy and A. Krause, Eds., vol. 80. PMLR, 10–15 Jul 2018, pp. 5286–5295. [Online]. Available: https://proceedings.mlr.press/v80/wong18a.html

[66] Krishnamurthy, Dvijotham, R. Stanforth, S. Gowal, T. Mann, and P. Kohli, "A Dual Approach to Scalable Verification of Deep Networks," *arXiv e-prints*, p. arXiv:1803.06567, Mar. 2018.

[67] M. Fazlyab, M. Morari, and G. J. Pappas, "Safety verification and robustness analysis of neural networks via quadratic constraints and semidefinite programming," *IEEE Transactions on Automatic Control*, vol. 67, no. 1, pp. 1–15, 2022.

[68] K. D. Dvijotham, R. Stanforth, S. Gowal, C. Qin, S. De, and P. Kohli, "Efficient neural network verification with exactness characterization," in *Proceedings of The 35th Uncertainty in Artificial Intelligence Conference*, ser. Proceedings of Machine Learning Research, R. P. Adams and V. Gogate, Eds., vol. 115. PMLR, 22–25 Jul 2020, pp. 497–507. [Online]. Available: https://proceedings.mlr.press/v115/dvijotham20a.html

[69] V. Tjeng, K. Xiao, and R. Tedrake, "Evaluating Robustness of Neural Networks with Mixed Integer Programming," *arXiv e-prints*, p. arXiv:1711.07356, Nov. 2017.

[70] "Latticeflow," https://latticeflow.ai/, accessed: 2024-01-28.

[71] A. Venzke and S. Chatzivasileiadis, "Verification of neural network behaviour: Formal guarantees for power system applications," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 383–397, 2020.

[72] A. Venzke, G. Qu, S. Low, and S. Chatzivasileiadis, "Learning optimal power flow: Worst-case guarantees for neural networks," in *2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2020, pp. 1–7.

[73] R. Nellikkath and S. Chatzivasileiadis, "Physics-informed neural networks for ac optimal power flow," *Electric Power Systems Research*, vol. 212, p. 108412, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779622005636

[74] R. Nellikkath, M. Tanneau, P. V. Hentenryck, and S. Chatzivasileiadis, "Scalable exact verification of optimization proxies for large-scale optimal power flow," 2024.

[75] A. Azhar, "Ai companies will be required to report safety tests to u.s. government," https://www.enterpriseai.news/2024/02/08/ai-companies-will-be-required-to-report-safety-tests-to-u-s-government/, accessed: 2024-06-19.

[76] "Eu ai act: first regulation on artificial intelligence," https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence, accessed: 2024-06-19.

[77] "Artificial intelligence safety institute consortium (aisic)," https://www.nist.gov/aisi/artificial-intelligence-safety-institute-consortium-aisic, accessed: 2024-06-19.

[78] Y. Wang, V. Vittal, X. Luo, S. Maslennikov, Q. Zhang, M. Hong, and S. Zhang, "Reinforcement learning based voltage control using multiple control devices," in *2023 IEEE Power & Energy Society General Meeting (PESGM)*, 2023, pp. 1–5.

[79] N. Rober, S. M. Katz, C. Sidrane, E. Yel, M. Everett, M. J. Kochenderfer, and J. P. How, "Backward reachability analysis of neural feedback loops: Techniques for linear and nonlinear systems," 2022. [Online]. Available: https://arxiv.org/abs/2209.14076

[80] M. Everett, G. Habibi, C. Sun, and J. P. How, "Reachability analysis of neural feedback loops," *IEEE Access*, vol. 9, pp. 163 938–163 953, 2021.

[81] T. Han, Y. Song, and D. J. Hill, "Ensuring network connectedness in optimal transmission switching problems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 68, no. 7, pp. 2603–2607, 2021.

[82] S. Kotha, C. Brix, Z. Kolter, K. Dvijotham, and H. Zhang, "Provably bounding neural network preimages," 2024.

[83] "ml4acopf benchmark," https://github.com/AI4OPT/ml4acopf_benchmark, 2023.

[84] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[85] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, pp. 25–57, 2006.

[86] K. Xu, H. Zhang, S. Wang, Y. Wang, S. Jana, X. Lin, and C.-J. Hsieh, "Fast and complete: Enabling complete neural network verification with rapid and massively parallel incomplete verifiers," *arXiv preprint arXiv:2011.13824*, 2020.

[87] M. ApS, "Mosek modeling cookbook," 2020.

[88] S. S. Rafiei and S. Chevalier, "Gpu-accelerated dcopf using gradient-based optimization," 2024. [Online]. Available: https://arxiv.org/abs/2406.13191

[89] C. Josz and D. K. Molzahn, "Lasserre hierarchy for large scale polynomial optimization in real and complex variables," *SIAM Journal on Optimization*, vol. 28, no. 2, pp. 1017–1048, 2018. [Online]. Available: https://doi.org/10.1137/15M1034386

[90] M. H. Dinh, F. Fioretto, M. Mohammadian, and K. Baker, "An analysis of the reliability of ac optimal power flow deep learning proxies," in *2023 IEEE PES Innovative Smart Grid Technologies Latin America (ISGT-LA)*, 2023, pp. 170–174.

[91] A. J. Conejo, E. Castillo, R. Minguez, and R. Garcia-Bertrand, *Decomposition techniques in mathematical programming: engineering and science applications*. Springer Science & Business Media, 2006.

[92] A. Gholami, K. Sun, S. Zhang, and X. A. Sun, "An alternating direction method of multipliers-based distributed optimization method for solving security-constrained alternating current optimal power flow," *OPERATIONS RESEARCH*, vol. 71, no. 6, 2023.

[93] J. Nocedal and S. J. Wright, *Numerical optimization*. Springer, 1999.

[94] S. Chevalier, L. Schenato, and L. Daniel, "Accelerated probabilistic power flow in electrical distribution networks via model order reduction and neumann series expansion," *IEEE Transactions on Power Systems*, vol. 37, no. 3, pp. 2151–2163, 2022.

[95] S. Chevalier and R. Parker, "Towards perturbation-induced static pivoting on gpu-based linear solvers," 2024. [Online]. Available: https://arxiv.org/abs/2311.11833

[96] Z. Shi, Q. Jin, J. Z. Kolter, S. Jana, C.-J. Hsieh, and H. Zhang, "Formal verification for neural networks with general nonlinearities via branch-and-bound," 2023.

[97] M. E. Ildiz, Y. Huang, Y. Li, A. S. Rawat, and S. Oymak, "From self-attention to markov models: Unveiling the dynamics of generative transformers," 2024. [Online]. Available: https://arxiv.org/abs/2402.13512

[98] D. Wei, H. Wu, M. Wu, P.-Y. Chen, C. Barrett, and E. Farchi, "Convex bounds on the softmax function with applications to robustness verification," 2023. [Online]. Available: https://arxiv.org/abs/2303.01713

[99] M. Tanneau and P. V. Hentenryck, "Dual lagrangian learning for conic optimization," 2024. [Online]. Available: https://arxiv.org/abs/2402.03086

[100] H. Salman, G. Yang, H. Zhang, C.-J. Hsieh, and P. Zhang, "A Convex Relaxation Barrier to Tight Robustness Verification of Neural Networks," *arXiv e-prints*, p. arXiv:1902.08722, Feb. 2019.

[101] H. Zhang, S. Wang, K. Xu, Y. Wang, S. Jana, C.-J. Hsieh, and Z. Kolter, "A branch and bound framework for stronger adversarial attacks of relu networks," in *International Conference on Machine Learning*. PMLR, 2022, pp. 26 591–26 604.

[102] M. Świechowski, K. Godlewski, B. Sawicki, and J. Mańdziuk, "Monte carlo tree search: A review of recent modifications and applications," *Artificial Intelligence Review*, vol. 56, no. 3, pp. 2497–2562, 2023.

[103] P. Liu, J. Zhou, and J. Lv, "Exploring the first-move balance point of go-moku based on reinforcement learning and monte carlo tree search," *Knowledge-Based Systems*, vol. 261, p. 110207, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095070512201303X

[104] K. Siefen, "Your journey to success with mathematical optimization," https://www.gurobi.com/resources/your-journey-to-success-with-mathematical-optimization-part-i/, accessed: 2024-07-02.

[105] A. Deza and E. B. Khalil, "Machine learning for cutting planes in integer programming: A survey," in *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, ser. IJCAI-2023. International Joint Conferences on Artificial Intelligence Organization, Aug. 2023. [Online]. Available: http://dx.doi.org/10.24963/IJCAI.2023/739

[106] H. Zhang, S. Wang, K. Xu, L. Li, B. Li, S. Jana, C.-J. Hsieh, and J. Z. Kolter, "General cutting planes for bound-propagation-based neural network verification," 2022. [Online]. Available: https://arxiv.org/abs/2208.05740

[107] S. Park and P. Van Hentenryck, "Self-supervised primal-dual learning for constrained optimization," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 4, pp. 4052–4060, Jun. 2023. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/25520

[108] S. Park and P. V. Hentenryck, "Self-supervised learning for large-scale preventive security constrained dc optimal power flow," 2024. [Online]. Available: https://arxiv.org/abs/2311.18072

[109] A. D. Palma, R. Bunel, A. Desmaison, K. Dvijotham, P. Kohli, P. H. S. Torr, and M. P. Kumar, "Improved branch and bound for neural network verification via lagrangian decomposition," 2021. [Online]. Available: https://arxiv.org/abs/2104.06718

[110] N. S. FAC, "Fac-011-4-system operating limits methodology for the operations horizon," https://www.nerc.com/pa/Stand/Reliability%20Standards/FAC-011-4.pdf, North American Electric Reliability Corporation, Tech. Rep.

[111] N. S. PRC, "Nerc standard prc-006-2-automatic underfrequency load shedding," North American Electric Reliability Corporation, Tech. Rep.

[112] X. Xu, R. Yousefian, M. Elkhatib, B. Choi, L. Huang, Y. Mao, and A. Berner, "Automatic underfrequency load shedding study of the pjm system," in *2019 IEEE Power & Energy Society General Meeting (PESGM)*, 2019, pp. 1–5.

[113] "Ieee standard for interconnection and interoperability of distributed energy resources with associated electric power systems interfaces," *IEEE Std 1547-2018 (Revision of IEEE Std 1547-2003)*, pp. 1–138, 2018.

[114] S. Gupta, V. Kekatos, and S. Chatzivasileiadis, "Optimal design of volt/var control rules of inverters using deep learning," *IEEE Transactions on Smart Grid*, pp. 1–1, 2024.

[115] A. M. Prostejovsky, C. Brosinsky, K. Heussen, D. Westermann, J. Kreusel, and M. Marinelli, "The future role of human operators in highly automated electric power systems," *Electric Power Systems Research*, vol. 175, p. 105883, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779619302020

[116] H. Nagarajan, M. Lu, E. Yamangil, and R. Bent, "Tightening mccormick relaxations for nonlinear programs via dynamic multivariate partitioning," 2016. [Online]. Available: https://arxiv.org/abs/1606.05806

[117] M. Bynum, A. Castillo, J.-P. Watson, and C. D. Laird, "Tightening mccormick relaxations toward global solution of the acopf problem," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 814–817, 2019.

[118] P. Sang and A. Pandey, "Circuit-theoretic joint parameter-state estimation – balancing optimality and ac feasibility," 2024. [Online]. Available: https://arxiv.org/abs/2404.10676

[119] S. Pineda, J. M. Morales, Álvaro Porras, and C. Domínguez, "Tight big-ms for optimal transmission switching," 2024. [Online]. Available: https://arxiv.org/abs/2306.02784

[120] S. Chevalier and M. R. Almassalkhi, "Towards optimal kron-based reduction of networks (opti-kron) for the electric power grid," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, 2022, pp. 5713–5718.

[121] O. Mokhtari, S. Chevalier, and M. Almassalkhi, "Enhancing Scalability of Optimal Kron-based Reduction of Networks (Opti-KRON) via Decomposition with Community Detection," *arXiv e-prints*, p. arXiv:2407.02679, Jul. 2024.

[122] H. Zhao, H. Hijazi, H. Jones, J. Moore, M. Tanneau, and P. Van Hentenryck, "Bound tightening using rolling-horizon decomposition for neural network verification," in *Integration of Constraint Programming, Artificial Intelligence, and Operations Research*, B. Dilkina, Ed. Cham: Springer Nature Switzerland, 2024, pp. 289–303.

[123] "Gurobi machine learning," 2024. [Online]. Available: https://github.com/Gurobi/gurobi-machinelearning

[124] "Webinar: Using trained machine learning predictors in gurobi," 2023. [Online]. Available: https://www.youtube.com/watch?v=jaux5Oo4qHU

[125] H. Hijazi, G. Wang, and C. Coffrin, "Gravity: A mathematical modeling language for optimization and machine learning," 2018.

[126] S. Gopinath and H. L. Hijazi, "Benchmarking large-scale acopf solutions and optimality bounds," in *2022 IEEE Power & Energy Society General Meeting (PESGM)*, 2022, pp. 1–5.

[127] C. Coffrin, R. Bent, K. Sundar, Y. Ng, and M. Lubin, "Powermodels.jl: An open-source framework for exploring power flow formulations," 2018.

[128] D. M. Fobes, S. Claeys, F. Geth, and C. Coffrin, "Powermodelsdistribution.jl: An open-source framework for exploring distribution power flow formulations," *Electric Power Systems Research*, vol. 189, p. 106664, Dec. 2020. [Online]. Available: http://dx.doi.org/10.1016/j.epsr.2020.106664

[129] M. Alkhraijah, R. Harris, C. Coffrin, and D. K. Molzahn, "Powermodelsada: A framework for solving optimal power flow using distributed algorithms," 2023.

[130] S. Babaeinejadsarookolaee, A. Birchfield, R. D. Christie, C. Coffrin, C. DeMarco, R. Diao, M. Ferris, S. Fliscounakis, S. Greene, R. Huang, C. Josz, R. Korab, B. Lesieutre, J. Maeght, T. W. K. Mak, D. K. Molzahn, T. J. Overbye, P. Panciatici, B. Park, J. Snodgrass, A. Tbaileh, P. V. Hentenryck, and R. Zimmerman, "The power grid library for benchmarking ac optimal power flow algorithms," 2021.

[131] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Grid structural characteristics as validation criteria for synthetic networks," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3258–3265, 2017.

[132] "alpha-beta-crown," https://github.com/Verified-Intelligence/alpha-beta-CROWN, 2024.

[133] "auto-lirpa," https://github.com/Verified-Intelligence/auto_LiRPA, 2024.

[134] J. Bezanson, S. Karpinski, V. B. Shah, and A. Edelman, "Julia: A fast dynamic language for technical computing," 2012.

[135] C. Rowley, "Pythoncall.jl: Python and julia in harmony," 2022. [Online]. Available: https://github.com/JuliaPy/PythonCall.jl

[136] T. Besard, P. Verstraete, and B. D. Sutter, "High-level gpu programming in julia," 2016.

[137] F. Safdarian, J. Snodgrass, J. H. Yeo, A. Birchfield, C. Coffrin, C. Demarco, S. Elbert, B. Eldridge, T. Elgindy, S. L. Greene, N. Guo, J. Holzer, B. Lesieutre, H. Mittelmann, R. P. O'Neill, T. J. Overbye, B. Palmintier, P. Van Hentenryck, A. Veeramany, T. W. Mak, and J. Wert, "Grid optimization competition on synthetic and industrial power systems," in *2022 North American Power Symposium (NAPS)*, 2022, pp. 1–6.

[138] A. Marot, B. Donnot, C. Romero, L. Veyrin-Forrer, M. Lerousseau, B. Donon, and I. Guyon, "Learning to run a power network challenge for training topology controllers," 2019.

[139] A. Marot, B. Donnot, G. Dulac-Arnold, A. Kelly, A. O'Sullivan, J. Viebahn, M. Awad, I. Guyon, P. Panciatici, and C. Romero, "Learning to run a power network challenge: a retrospective analysis," 2021.

[140] A. Marot, D. Rousseau, and Z. Xu, "Ai competitions and benchmarks: towards impactful challenges with post-challenge papers, benchmarks and other dissemination actions," 2023.

[141] "Machine learning for physical simulation challenge (ml4physics)," https://ml-for-physical-simulation-challenge.irt-systemx.fr/powergrid-challenge/, accessed: 2024-06-19.

[142] M. Zuhrie, I. Buditjahjanto, L. Nurlaela, and I. Basuki, "Do educational robotics competitions impact students' learning?" in *Journal of Physics: Conference Series*, vol. 1810, no. 1. IOP Publishing, 2021, p. 012045.

[143] "Hao zhu: Ee379k/394v: Data analytics in power systems," https://sites.utexas.edu/haozhu/teaching/, accessed: 2024-06-21.

[144] "Baosen zhang: Ee 555 a pmp data science for power systems, spring 2023," https://zhangbaosen.github.io/teaching/EE555, accessed: 2024-06-21.

[145] "Uvm extension: 4-h and youth," https://www.uvm.edu/extension/youth, accessed: 2024-06-21.

[146] K. Barrett, "Nation's urban and rural populations shift following 2020 census," https://www.census.gov/newsroom/press-releases/2022/urban-rural-populations.html, accessed: 2024-07-08.