Filter-Based Zeroth-Order Methods for Voltage Control in Distribution Grids: A Practical Approach

Subir Majumder, *Member, IEEE,* Xin Chen, *Member, IEEE,* Le Xie, *Fellow, IEEE*

***Abstract*—Voltage regulation in power distribution systems with renewable distributed energy resources (DERs) faces significant scalability challenges. Model-free zeroth-order (ZO) optimization methods offer a promising solution by generating control signals through direct interaction with the system. However, practical deployment is hindered by noisy gradient estimates, which stem from both the necessary controller-injected disturbances and the impracticality of sampling random variables from standard distributions used in ZO literature. To address these issues, we propose a novel ZO projected gradient method that integrates high-pass and low-pass filters. Specifically, a residual feedback mechanism emulates high-pass filtering, while the 'momentum' term, usually used to accelerate the convergence of a gradient-based method, has been utilized to mimic the low-pass filter. Additionally, as an alternative to traditional random sampling, we introduce a sinusoidal-based sampling technique inspired by extremum-seeking control theory and demonstrate that under this new setting, existing zeroth order estimators still provide gradient estimates of the controller objective. Simulation results on a large-scale Austin Grid test feeder, using realistic datasets, demonstrate that the proposed controller not only integrates seamlessly with existing slow-acting voltage regulation devices but also exhibits robustness against cyber-attacks. These findings suggest that the controller can be reliably deployed in modern power systems, improving scalability and resilience.**

***Index Terms*—Model-free control, projected gradient descent, voltage control, zeroth-order methods.**

I. Introduction

With the increasing integration of distributed renewable energy resources (DERs) into distribution systems, the need for scalable algorithms to manage voltage control has become crucial. The future power distribution grid would often include diverse devices, vendors, technologies, and ownership structures and require solutions that support plug-and-play operations while minimizing communication needs [1]. Local control solutions may offer these advantages, but vanilla local control approaches may not always guarantee optimality due to their primarily droop-based nature. However, there are other classes of local control approaches that estimate gradients based on measurements from the power systems. These algorithms do not require coordination with other controllers, do not require model information, but guarantee optimality [2]–[4].

The widespread deployment of smart meters and information and communication technologies (ICTs) paves the way for determining control actions based on broadcasted signals. One alternative to using data-driven/model-free methods for estimating gradients will be using Reinforcement learning (RL) to generate the control signal. While RL methods have gained significant attention in recent years (see [5] for a thorough literature review), these algorithms face several fundamental limitations, including safety requirements, scalability issues, training efficiency, and limited theoretical guarantees. On the other hand, zeroth-order (ZO) methods (also known as extremum-seeking control) are data-driven local control methods that learn the gradient of an objective function only based on function evaluations or system measurements [6]. Therefore, these algorithms are free from the limitations of RL-based control and have gained considerable attention for solving optimal model-free control problems [7]. Moreover, ZO control methods can self-adapt to dynamic environments, such as fluctuating renewable generation and changing power network topology, and operate in a plug-and-play fashion, with theoretical performance guarantees on stability and robustness [8]. ZO control has been applied in various domains, such as distribution network voltage control [2], [4], [9]–[11], maximum power point tracking in solar systems [12], building HVAC system control [11], continuous bandit games [13] and hardware-in-the-loop systems [14].

Other approaches comparable to local control approaches with advantages such as plug-and-play capability, privacy preservation, etc., are distributed approaches [15]. Unlike zeroth-order approaches that estimate the gradient using only power system measurements, distributed methods are model-based. In these approaches, each controller partaking in distributed control possesses information about a specific part of the power network and coordinates with other controllers to generate the final control signal. However, the accuracy of power network models is often limited. Recent feedback-based algorithms [16]–[19] can help mitigate this limitation by combining data-driven feedback with model-based control. The three approaches are compared in the following table:

TABLE I: Comparing Zeroth-Order Methods, Reinforcement Learning, and Distributed Control.

Feature/Method	Zeroth-Order Methods	Reinforcement Learning	Distributed Control
Model-free operation	✓	✓	
Scalability	✓		✓
Theoretical guarantee	✓		✓
Plug-and-play	✓		✓

Based on the number of function evaluations made during each iteration, ZO methods can be classified into two categories: single-point and multi-point [20]. To compare these two methods, consider a typical minimization problem [21]: $\min_{x\in\mathbb{R}^n} h(x)$, where the function $h(\cdot)$ is differentiable, and typical way of solving this unconstrained minimization

S. Majumder, X. Chen, and L. Xie are with the Department of Electrical Engineering and Computer Sciences, Texas A&M University, USA. (Corresponding author e-mail: subir.majumder@tamu.edu).

This work is supported in part by the Texas A&M Engineering Smart Grid Center and Texas A&M Energy Institute.

problem is using gradient descent (GD) approach, defined as, $x_{k+1} = x_k - \eta_k \nabla h(x_k)$, where, η_k is the step size at iteration k. If the gradient is unknown and only function evaluations are available, then the gradient could be estimated using single-point and two-point estimators — given as $G_h^{(1)}(x;\delta,u) = \frac{n}{\delta} h(x+\delta u)u$ and $G_h^{(2)}(x;\delta,u) = \frac{n}{\delta}\left(h(x+\delta u) - h(x)\right)u$, respectively. Here, u is a random variable drawn from the unit sphere or normal distribution, and δ is the smoothing radius. While second and multi-point methods are better because the variances in ZO estimators are small with small δ, it is difficult to obtain function evaluation at two or more points in an online optimization and dynamic control setting. Nevertheless, algorithms close to two-point zeroth-order methods have been applied in [11], [22].

First-order methods also carry a resemblance with extremum seeking (ES) control, where, based on averaging theory, the controller estimates true gradient. ES control often utilizes filters to obtain the gradient with higher accuracy, and in a recent work [21] borrows the idea of high-pass and low-pass filters and develops a novel SZO method called HLF-SZO (High/Low-pass Filter SZO). Note that our controller still operates in discrete time-space. The HLF-SZO method achieves the fastest convergence in the category of SZO methods to date, with both theoretical guarantees and empirical demonstration for unconstrained cases. However, their performances are unknown in constrained optimization problems, an essential requirement to apply the HLF-SZO method to the power distribution system voltage control. For convergence analysis, the authors in [21] make a few assumptions in the problem settings. These theoretical assumptions are rarely present in their entirety in the real world, and therefore, this research bridges the gap between theoretical work and practical applications for the proposed HLF-SZO-based voltage control considering a realistic distribution network.

Other Relevant Literature. In the algorithmic development side, there have been recent advances in gray-box algorithm [23], which utilizes zeroth-order methods in conjunction with model-based algorithms. There have also been advances in mixing distributed approaches utilizing gradients estimated through zeroth-order dynamics [24]. There have been developments in ensuring the safety of extremum-seeking methods utilizing the practical positivity of an unknown barrier function [25]. To the best of our knowledge, the convergence guarantee of constrained HLF-SZO methods in a practical setting has largely been absent.

Contributions. Accurate real-time power distribution network models, including switch status, transformer tap information, and locations of generating resources, are not always available because of Critical Energy/Electric Infrastructure Information (CEII) limitations. In this paper, we adapt the HLF-SZO method developed in [21] for solving the optimal voltage control (OVC) problem in realistic distribution systems in a model-free manner. The main contributions of this paper are summarized as follows:

(1) From an algorithmic development perspective, we present two main contributions. First, in ZO algorithms, random numbers are typically selected either from a Gaussian distribution or from points on the unit sphere, both of which pose implementation challenges. To address this, we draw inspiration from ES control and use sinusoidal functions as random number generators. However, we find that the resulting joint distribution is primarily concentrated on the surface of a unit hypercube. When using random numbers drawn from this surface, the expected value and variance of the ZO estimator become difficult to characterize. Second, the HLF-SZO method introduced in [21] addresses only simple unconstrained optimization problems. Extending it to *constrained* settings is non-trivial due to the lack of convergence guarantees for projected gradient descent under a ZO framework. This paper establishes the convergence properties for both the SZO-PGD algorithm and SZO-PGD variants incorporating high-pass filters.

(2) While we have shown that both the SZO-PGD algorithm and its variants converge, comparing their performances was challenging without making several assumptions about the properties of the objective function. To address this issue, we conducted extensive numerical experiments to evaluate the performance of the algorithms under different conditions. To further validate controller performance in a realistic setting, we developed a high-fidelity testbed simulator using OpenDSS [26] and implemented a realistic three-phase unbalanced 127-node distribution feeder model based on data from Austin, Texas, with realistic load and generation profiles. The code and data will be made available as open-source[1]. Additionally, we conducted extensive tests to demonstrate the robustness of our algorithm against various scenarios, including time-varying system dynamics, noisy measurements and cyber-attacks such as Denial of Service (DoS) attacks. We also show that the proposed algorithm can self-adapt to changing system conditions and operate in a plug-and-play manner.

The remainder of this paper is structured as follows. Section II formulates the voltage control problem and develops the SZO-based voltage control algorithms. Section **??** describes the testbed and the configuration of the Austin distribution feeder. Section III-A presents the numerical experiment results, and Section IV concludes this work.

Notation. Unbolded lower-case letters are used for scalars, and bolded lower-case letters are for vectors. $||(\cdot)||$ denotes the L_2 norm of a vector.

II. Problem Formulation and Algorithms

In this section, we introduce the OVC problem formulation and present the zeroth-order control algorithms with high-pass and low-pass filters.

A. Optimal Voltage Control in Power Distribution System

Consider a three-phase unbalanced power distribution network with the set of monitored bus $\mathcal{M}$ and the set of controllable bus $\mathcal{C}$. Each bus $j \in \mathcal{M}$ has real-time voltage measurement $v_{j,\phi}^{meas}$ at each available phase $\phi \in \{A, B, C\}$. The reactive power injection q_i^ϕ of phase ϕ at each controllable bus $i \in \mathcal{C}$ is the decision variable and can be controlled separately across phases for voltage regulation. Let $\boldsymbol{q}_i := (q_i^\phi)_{\forall \phi}$ be

[1]Not included to preserve anonymity per the journal's policy.

the vector that collects the per-phase reactive power injections at bus i. The OVC problem is formulated as model (1):

$$\min_{\boldsymbol{q}} h(\boldsymbol{q}) := c \sum_{\forall j \in \mathcal{M}} \sum_{\forall \phi} (v_{j,\phi}(\boldsymbol{q}) - v_j^{sp})^2 + \sum_{\forall i \in \mathcal{C}} \sum_{\forall \phi} d_i (q_i^{\phi})^2, \tag{1a}$$

$$\text{s.t. } \boldsymbol{q}_i \in \mathcal{Q}_i, \quad i \in \mathcal{C}. \tag{1b}$$

Here, c and d_i (typically ≥ 0) are the cost parameters associated with voltage deviation and for providing reactive power support. v_j^{sp} is the specified desired voltage level of bus j (this is often set close to 1 *pu*, but real-world scenarios might involve varying setpoints based on operational needs, loads, or renewable generation variability), around which the voltage should be maintained. The objective function (1a) encompasses two components: the deviation of the observed voltage magnitudes at the monitored buses $j \in \mathcal{M}$, and the opportunity cost associated with participating in voltage regulation (cost for not providing active power) [1] at the controllable buses $i \in \mathcal{C}$. The decision variable $\boldsymbol{q}_i$ at each bus $i \in \mathcal{C}$ in (1b) is constrained by the feasibility set $\mathcal{Q}_i$ (could be coupled across phases), which represents the power capacity of individual devices and is defined by (2)-(3):

$$\boldsymbol{q} := (\boldsymbol{q}_i)_{\forall i \in \mathcal{C}}, \quad \boldsymbol{\mathcal{Q}} := \prod_{\forall i \in \mathcal{C}} \mathcal{Q}_i.$$

There are two types of controllable devices whose reactive power injections can be adjusted for voltage control.

(i) Devices that only inject reactive power: This type includes devices such as Static Var Compensators (SVCs) and Distribution Static Synchronous Compensators (D-STATCOMs), and their capabilities are bounded by the device per phase power rating $[\underline{q}_i^{\phi}, \bar{q}_i^{\phi}]$. The feasibility set for these devices is defined as:

$$\mathcal{Q}_i := \{\boldsymbol{q}_i | \, \underline{q}_i^{\phi} \leq q_i^{\phi} \leq \bar{q}_i^{\phi}\}. \tag{2}$$

(ii) Devices that inject both real and reactive power: Here, devices are capable of injecting both real and reactive power into the grid. Their reactive power injection is further constrained by the per phase apparent power rating $\bar{s}_i^{\phi}$, and other algorithms control active power injection $p_i^{o,\phi}$. Storages also fall under this category. While both active and reactive power could contribute to voltage control [1], [3], [19], this work is limited to reactive power only. The associated feasibility set is given by:

$$\mathcal{Q}_i := \{\boldsymbol{q}_i | \, \underline{q}_i^{\phi} \leq q_i^{\phi} \leq \bar{q}_i^{\phi}, (p_i^{o,\phi})^2 + (q_i^{\phi})^2 \leq (\bar{s}_i^{\phi})^2\}. \tag{3}$$

Next, we present key definitions required for the formal analysis of the proposed algorithms designed to solve the OVC problem..

Definition 1. *(Convexity). A continuously differentiable function $h : C \to \mathbb{R}$ defined on a convex set C is called **convex** if, for any $x, y \in C$ and $\theta \in [0, 1]$, the following inequality holds:*

$h(\theta x + (1-\theta)y) \leq \theta h(x) + (1-\theta)h(y)$.

Definition 2. *(Lipschitz Continuity). A function $h : \mathbb{R}^n \to \mathbb{R}$ is said to be **Lipschitz continuous** on a set $C \subseteq \mathbb{R}^n$ if there exists a constant $L \geq 0$ such that, for all$x, y \in C$:*

$|h(x) - h(y)| \leq L\|x - y\|$,

where $\|x - y\|$ is a norm (typically the Euclidean norm) that measures the distance between x and y.

Definition 3. *(Proper). A function $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is called **proper** if it satisfies the following conditions:*

1) *$h(x) < +\infty$ for at least one $x \in \mathbb{R}^n$.*
2) *$h(x) > -\infty$ for all $x \in \mathbb{R}^n$.*

Definition 4. *(Closed). A function $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is said to be **closed** if its **epigraph** is a closed set. The **epigraph** of h is defined as:*

$epi(h) = \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid h(x) \leq \alpha\}$.

Assumption 1. *[27, Theorem 8.7]*

1) *The function $h : \mathbb{R}^n \to (-\infty, \infty]$ is convex, closed, and proper.*
2) *The set $C \subseteq \mathbb{R}^n$ is convex, closed, and nonempty.*
3) *The set C is contained within the interior of the domain of h, i.e., $C \subseteq \text{int}(\text{dom}(h))$.*
4) *The optimal solution set of $\min\{h(x) : x \in C\}$ is nonempty and denoted by X^*, with the optimal value represented as h_{opt}.*

Now, we refer back to the OVC problem in (1). Note that the functional form $\boldsymbol{v}(\boldsymbol{q}) := (v_{j,\phi}(\boldsymbol{q}))_{\forall j \in \mathcal{M}, \forall \phi}$ captures how reactive power injections influence system-wide voltages without implicitly solving the non-linear and non-convex power flow equations. These voltages can be directly *monitored* from the power distribution network as *feedback*. The incorporation of this functional form also helps us capture the impacts of all uncontrollable loads and renewable generations [28], including discrete voltage control devices, such as on-load tap changing (OLTCs) devices, or static voltage regulators (SVRs).

To solve the OVC problem (1), we first introduce the *Projected Gradient Descent (PGD)* algorithm [28], which under Assumption 1 converges to the global optima [29]:

$$\boldsymbol{q}_i(t+1) = \mathcal{P}_{\mathcal{Q}_i}\Big[\boldsymbol{q}_i(t) - k\frac{\partial h(\boldsymbol{q})}{\partial \boldsymbol{q}_i}\Big], \quad i \in \mathcal{C}, \tag{4}$$

where k is the stepsize and each component in $\frac{\partial h(\boldsymbol{q})}{\partial \boldsymbol{q}_i}$ is

$$\frac{\partial h(\boldsymbol{q})}{\partial q_i^{\phi}} := 2c \sum_{\forall j \in \mathcal{M}} \sum_{\forall \phi'} (v_{j,\phi'}(\boldsymbol{q}) - v_j^{sp}) \frac{\partial v_{j,\phi'}(\boldsymbol{q})}{\partial q_i^{\phi}} + 2d_i q_i^{\phi}. \tag{5}$$

PGD iteratively updates $\boldsymbol{q}_i(t+1)$ while ensuring feasibility as in (1b) via the projection $\mathcal{P}_{\mathcal{Q}_i}(\cdot)$ operator, converging to an optimal solution. If we refer to the first component of OVC problem in (1), the necessary gradient $\frac{\partial h(\boldsymbol{q})}{\partial q_i^{\phi}}$ at each discrete time step t in (4) is not available. For the second component, as shown in (5), the gradient could be calculated based on local information. We utilize ZO methods for estimating the gradient as a part of *model-free controller implementation*.

Remark 1. Although the model in (1) uses a quadratic function to regulate the voltages at specific buses in the power distribution system, it is also possible to use other convex functions for this purpose. Similarly, constraints can extend beyond voltage to regulate parameters like power flow through network branches. □

In (4), the *projection operator* is defined as a mapping $\mathcal{P}_C : \mathbb{R}^n \to C$ that assigns each point $x \in \mathbb{R}^n$ to its nearest point in a given nonempty, closed, and convex set $C \subseteq \mathbb{R}^n$. Specifically, the projection of x onto C, denoted as $\mathcal{P}_C(x)$, is given by:

$$\mathcal{P}_C(x) := \arg\min_{y \in C} \|y - x\|^2. \tag{6}$$

Next, we define the following property of the projection operators, which will be useful for later analysis.

Theorem 1. *[27, Theorem 6.42] Let $C \subseteq \mathbb{R}^n$ be a nonempty, closed, and convex set. For any $x, y \in \mathbb{R}^n$, the projection operator $\mathcal{P}_C$ satisfies:*

1) ***(Firm Nonexpansivity)*** $\langle x - y, \mathcal{P}_C(x) - \mathcal{P}_C(y)\rangle \geq \|\mathcal{P}_C(x) - \mathcal{P}_C(y)\|^2$.
2) ***(Nonexpansivity)*** $\|\mathcal{P}_C(x) - \mathcal{P}_C(y)\| \leq \|x - y\|$.

We also formally state the following assumption as a part of the problem statement:

Assumption 2. *Underlying input-output map $v_{j,\phi}(\boldsymbol{q})$ and its derivatives of any order are unavailable, and only the values of $v_{j,\phi}(\boldsymbol{q})$, $j \in \mathcal{M}$, can be accessed by observing the system outputs. Additionally, both objective function $h(x)$ and the feasible set $\mathcal{Q}$ are known.*

B. Utilizing Zeroth Order methods for OVC

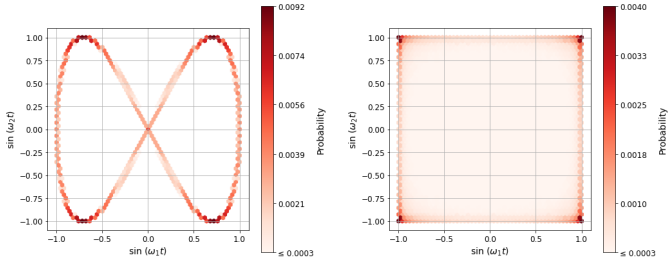
Classical single-point zeroth order (ZO) gradient estimator $g(x_k)$ has the form:

$$g(x_k) := \frac{n}{\delta} h(x_k + \delta u_k) u_k. \tag{7}$$

Here, δ represents the smoothing radius. Typically, u_k is an independent and identically distributed (i.i.d.) random variable at k^{th} iteration, which can be sampled either from a unit sphere [30] or from a standard normal distribution [31] in $\mathbb{R}^n$. However, when sampled from the unit sphere, the components of u_k are correlated, making it challenging to implement a controller that operates without requiring communication among components. Conversely, when sampled from a normal distribution, the components of u_k are independent. However, this may lead to unbounded exploration, which becomes problematic if the function is defined only on a compact domain.

An alternative approach to generating independent random variables is to use sinusoidal signals, where signals of different frequencies are employed to construct the components of u_k. Injection of perturbation signal in the form of a sinusoid [32] is widely used in extremum-seeking (ES) control methods, which can be viewed as the continuous-time counterpart of discrete-time ZO optimization algorithms. The relationship between ES methods and ZO algorithms will be further elaborated in the next subsection. For ES control involving multi-variable (multi-input-multi-output) systems, it is crucial to assign a distinct frequency parameter, denoted as $\boldsymbol{\omega} := (\omega_i)_{i \in \mathcal{C}}$, to each control input. These frequency parameters must satisfy the following conditions for any distinct indices $i, j, k \in \mathcal{C}$: $\omega_i \neq \omega_j$ and $\omega_i + \omega_j \neq \omega_k$ [6]. This ensures that the probing signals do not interfere with each other.

When sinusoidal signals of similar magnitudes and phases are used, a potential issue arises if any two signals are integer multiples of each other, as they form Lissajous figures. Lissajous figures tend to exhibit non-ergodic behavior, repeatedly tracing the same patterns, and thus do not resemble random variables. In contrast, if the frequencies are irrational multiples of each other, the resulting signals may seem to cover a unit hypercube over a large number of iterations.



(a) Frequencies are integer multiple. (b) Frequencies are not integer multiple.

Fig. 1: Example characteristics of random numbers generated through sinusoidal functions of different frequencies.

First, we introduce the following two definitions.

Definition 5. *(Hypercube and Surface of a Hypercube). Define the hypercube centered at the origin with side length 2δ:*

$$\mathbb{H}_\delta = \left\{ v \in \mathbb{R}^n \;\middle|\; |v_i| \leq \delta \text{ for } i = 1, 2, \ldots, n \right\}.$$

Also, $\mathbb{S}_\delta$ denote the surface of $\mathbb{H}_\delta$, consisting of $2n$ number of $(n-1)$-dimensional faces, defined as:

$$\mathbb{S}_\delta = \left\{ v \in \mathbb{R}^n \;\middle|\; \exists i \in \{1, 2, \cdots, n\} \text{ such that } |v_i| = \delta, \text{ and } |v_j| \leq \delta \;\; \forall\, j \neq i \right\}.$$

We have the following two lemmas:

Lemma 1. *As $n \to \infty$, the probability that the random vector $u = [u_1, u_2, \cdots, u_j, \cdots, u_n]$, where each coordinate u_j is independently sampled from the arcsine distribution, lies on the $\mathbb{S}_\delta$ approaches 1.*

Proof. Please see Appendix A for details. □

Lemma 2. *As $n \to \infty$, the probability that the random vector $u = [u_1, u_2, \cdots, u_j, \cdots, u_n]$, where each coordinates u_j is independently sampled from the arcsine distribution, lie near all $2d$ facets (surfaces) of $\mathbb{S}_\delta$ with equal probability.*

Proof. Please see Appendix B for details. □

Arcsine distribution in the previous two lemmas originate from the use od sinusoidal perturbation signal. For the practical distributed implementation of ZO-control, our first objective will be *to investigate whether the gradient estimator in* (7) *can accurately estimate the gradient when the random variables are constrained to the surface of a hypercube*. Additionally, we aim to analyze the variance of the resulting gradient estimate under the said condition.

Definition 6. *(Gaussian Smoothed function over the surface of $\mathbb{H}_\delta$). Let us define the smoothed function $h_\delta(x)$ as the average of h over $\mathbb{H}_\delta$:*
$h_\delta(x) = \frac{1}{(2\delta)^n} \int_{\mathbb{H}_\delta} h(x+v)\, dv = \mathbb{E}_{v\in\mathbb{H}_\delta}[h(x+v)]$.

Next, we propose the following lemma:

Lemma 3. *Let $h : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function, and let $\delta > 0$. The expected value of $h(x+\delta u)\, u$, where u is sampled equiprobably from all facets of $\mathbb{S}_\delta$, is related to the gradient of h_δ:*

$$\mathbb{E}_{u\in\mathbb{S}_\delta}[h(x+\delta u)\, u] = \frac{\delta}{n}\nabla h_\delta(x). \tag{8}$$

Proof. The proof is inspired by [30, Lemma 2.1]. Please see Appendix C for details. □

Lemma 4. *If a vector u is equiprobably sampled from all facets of $\mathbb{S}_\delta$ and within a facet, all other coordinates are independently sampled from the arcsine distribution, then the expected value of the squared norm of u is given by:*

$$\mathbb{E}_{u\in\mathbb{S}_\delta}\left[\|u\|^2\right] = \frac{n+1}{2}\delta^2. \tag{9}$$

Proof. Please see Appendix D for details. □

To analyze the convergence of the PGD under $g(x_k)$ in (7), we first present the following Lemma for the bound of the second moment of the gradient estimator.

Lemma 5. *We have $\mathbb{E}_{\mathbb{S}_1}\left[\|g(x_k)\|^2\right] \le \frac{n^2(n+1)}{1\delta^2}\left(h_C^{\max}\right)^2$ under the PGD update rule; for all $x_k \in \mathbb{R}^n$.*

Proof. Suppose $h_C^{\max}$ is the maximum value of the function $h(x)$ within region C. Note that,

$$\|g(x_k)\|^2 = \|\frac{n}{\delta}h(x_k+\delta u_k)u_k\|^2 \overset{x_k+\delta u_k\in C}{\le} \frac{n^2}{\delta^2}\left(h_C^{\max}\right)^2\|u_k\|^2.$$

Since $\|g(x_k)\|^2 \ge 0$, and $\|g(x_k)\|^2 \le \frac{n^2}{\delta^2}\left(h_C^{\max}\right)^2\|u_k\|^2$ almost surely, we can take expectation on both sides,

$$\mathbb{E}_{\mathbb{S}_1}\left[\|g(x_k)\|^2\right] \le \frac{n^2}{\delta^2}\left(h_C^{\max}\right)^2\mathbb{E}_{\mathbb{S}_1}\left[\|u_k\|^2\right].$$

Since u_k is uniformly sampled from the surface of a unit hypercube, we get the desired bound on the expected value utilizing Lemma 4. □

Theorem 2. *[27, Theorem 8.35] Suppose that Assumption 1 holds. Also, let, $g(x_k)$ be an unbiased estimate of $\nabla h_\delta(x)$, and $\mathbb{E}_{\mathbb{S}_1}\left[\|g(x_k)\|^2\right]$ is bounded. Let $\{x_k\}_{k\ge0}$ be the sequence generated by the stochastic projected subgradient method with positive stepsizes $\{\eta_k\}_{k\ge0}$, and let $\{h_k^{best}\}_{k\ge0}$ be the sequence of best-achieved values. Then,*

1) *If $\frac{\sum_{n=0}^k t_n^2}{\sum_{n=0}^k t_n} \to 0$ as $k \to \infty$, then $\mathbb{E}\left[h_k^{best}\right] \to h_{opt}$ as $k \to \infty$.*
2) *Assume that C is compact. Let B be the positive constant representing the second moment of the unbiased gradient, and let Θ be an upper bound on the half-squared diameter of C in (6). If $t_k = \frac{\sqrt{2\Theta}}{\sqrt{B}\sqrt{k+1}}$, then for all $k \ge 2$, $\mathbb{E}\left[f_k^{best}\right] - f_{opt} \le \mu\sqrt{B}\frac{\sqrt{2\Theta}}{\sqrt{k+2}}$, where $\mu = 2(1+\log(3))$.*

Remark 2. All the assumptions for Theorem 2 are satisfied, where the gradient is estimated using a zeroth-order (ZO) approach on the surface of an n-dimensional hypercube. As a result, a stochastic projected gradient descent algorithm will generally converge. However, the convergence rate is limited to an order of $n^{\frac{3}{2}}$. As discussed in Section III-C, this implies that larger systems tend to exhibit slower convergence rates. The variance of the gradient estimate is bounded in inverse proportion to the perturbation signal. Also, The variance estimate is bounded in squared proportion to the maximum function value within the feasible region. □

The SZO-PGD algorithm is formally provided below, which presents the implementation of a model-free ZO algorithm for the OVC problem. Note that the *distribution network behaves like a multi-port system, which allows for straightforward implementation of zeroth-order (ZO) control. This also means that the controlling phase of each node of the system requires a unique probing signal dictated by the frequency parameter, ω_i^ϕ.*

Algorithm 1 SZO-PGD-Based Model-Free Optimal Voltage Control (OVC) Algorithm.

At each time step t, do the following steps:

- **Each monitored node** $j \in \mathcal{M}$ measures local voltage magnitude $(v_{j,\phi}^{meas}(t))_{\forall\phi}$ and sends it to all controllable nodes.
- **Each controllable node** $i \in \mathcal{C}$ calculates:

$$\tilde{h}_t^i = c\sum_{j\in\mathcal{M}}\sum_{\phi}(v_{j,\phi}^{meas}(t) - v_j^{sp})^2, \tag{10a}$$

$$\boldsymbol{q}_i(t+1) = \mathcal{P}_{\mathcal{Q}_i}\Big[\boldsymbol{q}_i(t) - k\big(\frac{n}{\delta}\tilde{h}_t^i\sin(\boldsymbol{\omega}_i t) + 2d_i\boldsymbol{q}_i(t)\big)\Big], \tag{10b}$$

then deploy the reactive power injection $\hat{\boldsymbol{q}}_i(t+1)$:

$$\hat{\boldsymbol{q}}_i(t+1) = \boldsymbol{q}_i(t+1) + \delta\sin(\boldsymbol{\omega}_i(t+1)), \tag{11}$$

where $\boldsymbol{\omega}_i := (\omega_i^\phi)_{\forall\phi}$ and $\sin(\boldsymbol{\omega}_i t) := (\sin(\omega_i^\phi t))_{\forall\phi}$.

Recall that the OVC objective function in (1) consists of two components, where, as highlighted in (5), gradient estimates are required only for the first component. Therefore, in (10a), we isolate the first component using the captured measurements, while the term $\frac{2}{\delta}h_t^i\sin(\boldsymbol{\omega}_i t)$ represents its gradient estimate. For the second component, the variable $d_i\boldsymbol{q}_i(t)$, which is locally known, is used to directly compute the corresponding gradient. The local knowledge of the feasibility set $\mathcal{Q}_i$ further allows for the reactive power injection updates as part of the PGD algorithm, as shown in (10b). Additionally, we assume voltage measurements are noise-free, although this may not hold in practical applications. The impact of noisy measurements on algorithm performance is analyzed in Section III-C.

C. Zeroth-Order Control with High- and Low-Pass Filters

Extremum-seeking (ES) control can be viewed as the continuous-time equivalent of solving optimization problems using the gradient descent method. The similarity between ES control and ZO methods is illustrated in Fig. 2. In both approaches, the goal is to solve an optimization problem of the form $\min_x h(x)$, where the objective function $h(x)$ is unknown, and only its output, $y = h(x)$, can be observed.

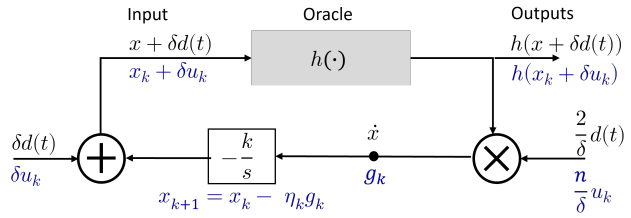


Fig. 2: Relationship between zeroth order optimization and classic extremum seeking control algorithms. Signals corresponding to ZO methods are marked in blue.

As shown in the block diagram, a small dither or probing signal $\delta d(t)$ is added to the state x ($\delta > 0$ is the amplitude of the probing signal), and another signal $\frac{2}{\delta}d(t)$ is multiplied to the function evaluation. $y = h(x)$ represents the plant, while $\frac{1}{s}$ and $-k$ are the integrator and gain, respectively. Considering sinusoidal probing signals $d(t) := \sin(\omega t)$ with the frequency parameter ω, the dynamics of the closed-loop system in Fig. 2 is formulated as:

$$\dot{x} = -k \cdot \frac{2}{\delta} h\left(x + \delta d(t)\right) d(t). \tag{12}$$

Using the Taylor expansion and the averaging theory, the average system dynamics is derived as:

$$\begin{aligned}\dot{x} &= -\frac{k}{T}\int_0^T \frac{2}{\delta} h\left(x + \delta \sin(\omega t)\right)\sin(\omega t) dt \\ &= -\frac{k}{T}\int_0^T 2\sin^2(\omega t) h'(x) + \mathcal{O}(\delta) dt = -k h'(x) + \mathcal{O}(\delta).\end{aligned} \tag{13}$$

Comparing (13) and Fig. 2 we see that the average dynamics of $\dot{x}$ is indeed gradient estimator $g(x_k)$, and, in the steady state, x reaches the optimal solution x^*. See [4] for a detailed analysis of the extremum-seeking system.

Recall that in the SZO-PGD method, the use of a single-point gradient estimate can lead to large oscillations in the independent variables, resulting in slower convergence. To address this, we take inspiration from the high-pass and low-pass filters used in the ES control method, as discussed previously, to enhance the convergence behavior of SZO methods. In ES control, the plant output $h(x + \delta\sin(\omega t))$ can be approximated using a Taylor series expansion as $h(x) + h'(x)\sin(\omega t)$. By applying a high-pass filter to the plant output, we can effectively eliminate the DC component, isolating the oscillatory term $h'(x)\sin(\omega t)$. When this filtered output is then multiplied by the probing signal $\sin(\omega t)$, the resulting expression is $h'(x)\sin^2(\omega t)$, which simplifies to $h'(x) - h'(x)\cos(2\omega t)$. Over multiple iterations, the oscillatory component $h'(x)\cos(2\omega t)$ averages out, leaving only the gradient information $h'(x)$. To further suppress these oscillations, we pass the signal $\dot{x}$ through a low-pass filter, which helps reduce the high-frequency components, allowing for a smoother convergence.

Again, recall the two-point implementations of zeroth-order (ZO) methods, where one such implementation relies on an oracle that provides both $h(x + \delta u)$ and $h(x)$. However, this approach is not practical in many real-world scenarios because of inherent measurement noises. One possible alternative is to utize past measurements from the system $\{\cdots, h_i(\boldsymbol{q}(k) + \delta\sin(\boldsymbol{\omega}(k))), h_i(\boldsymbol{q}(k+1) + \delta\sin(\boldsymbol{\omega}(k+1))), \cdots\}$, which may already be available to a controlling node. Using this sequence, we define $\gamma_t^i = h_t^i - h_{t-1}^i$ for gradient estimation. From the theory of discrete time control, it is well-known that first-order differencing acts as a high-pass filter that removes the DC component present in the plant output. The use of first-order differencing within ZO methods has been previously introduced in the literature and is referred to as 'residual feedback' [31]. Zhang et al. [31] have also demonstrated that when random perturbations are sampled from a unit sphere, the expected value of the single-point residual feedback scheme provides a gradient estimate of the Gaussian-smoothed version of the true objective function, while ensuring that the second moment remains bounded.

Now, we define the gradient estimation as follows:

$$\tilde{g}(x_k) := \frac{n}{\delta}\left(h(x_k + du_k) - h(x_{k-1} + du_{k-1})\right) u_k. \tag{14}$$

Lemma 6. *Let $h : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function, and let $\delta > 0$. The expected value of $(h(x_k + \delta u_k) - h(x_{k-1} + \delta u_{k-1}))\, u_k$, where both u_k and u_{k-1} are independently sampled uniformly from $\mathbb{S}_1$ at k^{th} step, is related to the gradient of h_δ:*

$$\mathbb{E}_{u\in\mathbb{S}_1}[(h(x_k + \delta u_k) - h(x_{k-1} + \delta u_{k-1}))\, u_k] = \frac{\delta}{n}\nabla h_\delta(x_k).$$

Proof. This proof is relatively straightforward because u_{k-1} and u_k are independently sampled, and its expected value has a zero mean, implying $\mathbb{E}_{u\in\mathbb{S}_1}[h(x_{k-1} + \delta u_{k-1})\, u_k] = 0$. Utilizing Lemma 3 then completes the proof. □

We need the following lemma before we discuss the variance of the expected gradient:

Lemma 7. *Let $u \in \mathbb{R}^n$ be a random vector uniformly sampled from $\mathbb{S}_\delta$. Then, the expectation of the fourth power of the norm of u, $\mathbb{E}[\|u\|^4]$, is given by:*

$$\mathbb{E}_{\mathbb{S}_\delta}[\|u\|^4] = \frac{2n^2 + 5n + 1}{8}\delta^4.$$

Proof. Please see Appendix E for details. □

Lemma 8. *Under the assumption of L-Lipschutz continuity of h and the PGD update rule with stepsize η_k; for all $x_k \in \mathbb{R}^n$, we have:*

$$\begin{aligned}\mathbb{E}_{\mathbb{S}_1}\left[\|\tilde{g}(x_k)\|^2\right] \le \frac{(n+1)L_0^2}{\delta^2}\eta_{k-1}^2\mathbb{E}[\|\tilde{g}(x_{k-1})\|^2] \\ + \frac{L^2}{2}\left(4n^2 + 9n + 3\right).\end{aligned}$$

Proof. Please see Appendix F for details. □

Next, in regards to filtering, recursive difference equation $x(t+1) = x(t) + \alpha\left(x(t) - x(t-1)\right)$ effectively behaves as a low pass filter, where for small α values, the filter has a lower cutoff frequency, implying only very low-frequency signals pass through, and with large α values, the filter has a higher cutoff frequency. Note that in an optimization setting, this recursive term is similar to "momentum" terms [33] especially

heavy-ball method, which can accelerate the convergence of an optimization algorithm [21].

Remark 3. While Lemma 8 shows that the variance of the gradient estimate $\tilde{g}$ is bounded, which is an essential condition for the convergence of the PGD algorithm (see Theorem 2), without additional restrictions on h, it is difficult to analytically compare the convergence rate with the incorporation of residual feedback term. Analyzing convergence gets further complex with the incorporation of momentum terms in the PGD technique. Here, we will solely rely on numerical simulation to compare algorithm performance, and the theoretical convergence guarantee of ZO PGD with momentum term will be a part of future work. □

Both of these differencing and momentum terms as high- and low-pass filters are suitably incorporated in the SZO-PGD-based model, and the overall algorithm is presented in Algorithm 2.

Algorithm 2 HLF-SZO-PGD-Based Model-Free Optimal Voltage Control (OVC) Algorithm.

At each time step t, do the following steps:

- **Each monitored node** $j \in \mathcal{M}$ measures local voltage magnitude $(v_{j,\phi}^{meas}(t))_{\forall \phi}$ and sends it to all controllable nodes.
- **Each controllable node** $i \in \mathcal{C}$ calculates:

$$\text{Equation(10a)},\ \gamma_t^i = \tilde{h}_t^i - \tilde{h}_{t-1}^i, \tag{15a}$$

$$\boldsymbol{q}_i(t+1) = \mathcal{P}_{\mathcal{Q}_i}\Big[\boldsymbol{q}_i(t) - k\big(\frac{n}{\delta}\gamma_t^i \sin(\boldsymbol{\omega}_i t) + 2d_i\boldsymbol{q}_i(t)\big) + \alpha\big(\boldsymbol{q}_i(t) - \boldsymbol{q}_i(t-1)\big)\Big], \tag{15b}$$

then deploy the reactive power injection $\hat{\boldsymbol{q}}_i(t+1)$ (11).

Compared with Algorithm 1, Algorithm 2 replaces the reactive power update rule (10b) with (15). Equation (15a) results from the integration of a high-pass filter, where γ_t^i is the residual difference between the function evaluation h_t^i at time t and the function evaluation h_{t-1}^i one time step earlier. In (15b), an additional momentum term $\alpha\big(\boldsymbol{q}_i(t) - \boldsymbol{q}_i(t-1)\big)$ is added, where $\alpha \in [0, 1]$ is an tunable parameter.

Remark 4. In contrast to the HLF-SZO method in [21], our proposed HLF-SZO-PGD method (Algorithm 2) significantly extends its capabilities, making it suitable for constrained optimization problems, and easy to deploy in realistic environments by utilizing sinusoids as random number generators. Moreover, compared to [3], this paper offers three key advantages: 1) We utilize both high-pass and low-pass filters for improved performance, while [3] uses only a low-pass filter; 2) Our method is iterative, making it more practical than the continuous-time dynamics used in [3]; and 3) We extend the voltage control application to more realistic three-phase unbalanced distribution systems with comprehensive testing. □

D. Practical Implementation

The implementation of both the SZO-PGD and HLF-SZO-PGD algorithms, as outlined above, is depicted in Fig. 3. The controllable device comprises an energy source, power electronic switches, measurement units for determining control references, and a pulse width modulation (PWM) generator to drive the converters. These controllable devices are linked to the grid via the point of common coupling (PCC). The inverter control receives both active and reactive power setpoints, denoted as Pset and Qset, respectively. In our implementation, Q_{set} is regulated by the ZO controller. Since Algorithm 2 operates in discrete time steps, Q_{set}, represented as $\hat{\boldsymbol{q}}_i(t)$, remains constant between consecutive time intervals. *This behavior is reflected in the staircase function depicted in the diagram. The assumption is that the system's operating conditions do not change within each time step.* As a result, the probing signal used for Qset takes the form of a discrete-time sinusoidal signal, where t represents the integer time-step. As shown in both the diagram and Algorithm 2, the controller relies solely on the broadcasted unidirectional voltage magnitudes across the system, significantly reducing communication overhead.

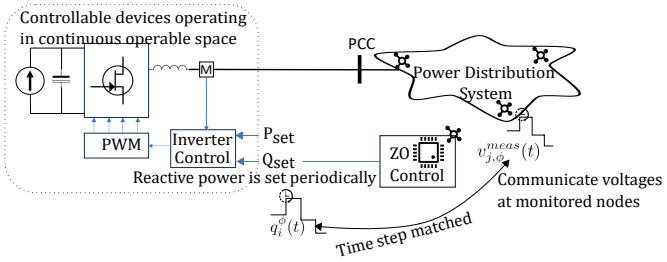


Fig. 3: Discrete time control action.

Furthermore, at the time of integration, the controllers require various input parameters, including system model parameters such as c and d, specified voltages v_j^{sp}, and algorithm-specific parameters like δ, k, and α. Additionally, the controllers must negotiate values for ω_i^{ϕ} during the connection process and update the value of n with the addition or removal of controllers.

In both algorithms, the complexity per controller node exhibits similar asymptotic characteristics. Each controller node receives voltage measurements from all monitored nodes ($|\mathcal{M}|$) across multiple phases ($|\phi|$), resulting in comparable time, memory, and computational requirements. For the *SZO-PGD* algorithm, each controller node computes a cost function by aggregating voltage deviations across all monitored nodes and updates the reactive power using zeroth-order gradient descent. This process results in a time and computational complexity of $O(|\mathcal{M}| \times |\phi|)$ per controller, driven mainly by the cost function evaluation. The memory complexity is also $O(|\mathcal{M}| \times |\phi|)$, as each controller stores the voltage measurements and corresponding reactive power values. The *HLF-SZO-PGD* algorithm adds the computation of the residual difference between consecutive cost values and incorporates a momentum term. While these additions introduce a few extra arithmetic operations, the overall time and computational complexities remain similar. Likewise, the memory complexity is unchanged since the controller now stores the cost function values from the previous iteration in addition to the existing data.

Based on the scope and control implementation, a taxonomy of the above SZO control algorithms is presented in Fig. 4, which summarizes the key features of these algorithms [1]. In particular, the proposed HLF-SZO-PGD algorithm is

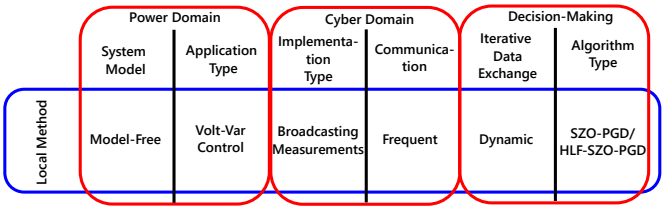


Fig. 4: Taxonomy of the **SZO-PGD/HLF-SZO-PGD** algorithm.

fully model-free. The algorithm also enables plug-and-play operation and does not require a control center to coordinate a large number of distributed energy devices. This taxonomy helps us characterize the suitability of deployment of this algorithm in the real world.

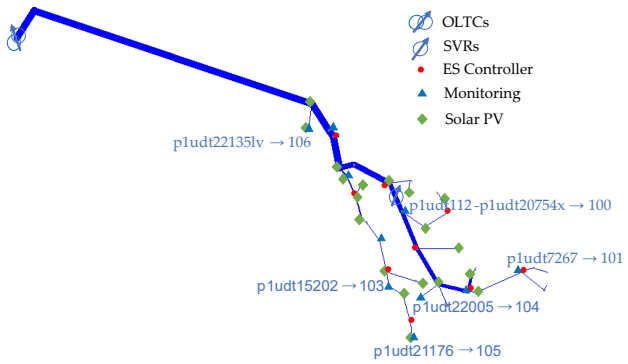


Fig. 5: The test feeder from the synthetic Austin Grid.

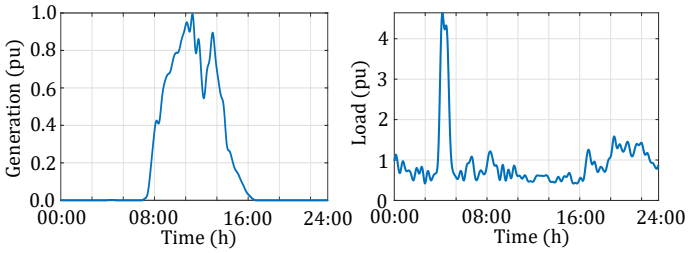


Fig. 6: The solar and load profiles in a typical day.

III. Case Study

To understand the scalability of our algorithms, we evaluated and compared the performance of the proposed algorithms using two different systems: (1) the IEEE 13-bus system and (2) a realistic 3-phase unbalanced test feeder comprising of 127 nodes modeled after an actual distribution network located in Austin, Texas. This test feeder is a part of the large T&D simulation model developed in [34] utilizing only publicly available data with parameters tuned based on historical dataset. Note that the load profiles used to tune the network model may be generic, which limits the accuracy of the synthetic feeder with the actual power grid. The feeder's topology is shown in Fig. 5, where we have highlighted specific nodes used in the controller performance analysis. To enhance the readability of the figures, we have renamed these nodes.

There were a few challenges when using the synthetic test feeder to evaluate our algorithm's accuracy. First, the distribution feeder did not specify which nodes had solar PV installations. To overcome this, we randomly selected a few nodes as potential locations for solar PVs. The nodes interfacing with the solar PV installations and their respective controllers are illustrated in Fig. 5. The PV generation profiles used for testing were derived from real-world data provided by Pecan Street [35], representative of the Austin, Texas region. Second, the initial load profiles were generic and lacked specificity. To address this, we incorporated realistic load profiles sampled from Pecan Street's dataset [35], ensuring the loading conditions reflected real-world scenarios. In particular, we selected profiles that demonstrate the controller's performance under stressed operating conditions. Please note that, as shown in Fig. 6, the inclusion of sudden spikes in load and generation profiles in time-varying scenarios aids in evaluating the controller's robustness. Finally, we enhanced the synthetic feeder by adding a substation with a 12.47 kV/7.20 kV, 7 MVA transformer equipped with on-load tap changers (OLTCs) to evaluate our algorithm's performance in the presence of other devices not directly controlled by it. We also integrated a static voltage regulator (SVR), as shown in Fig. 5, in this regard. We assumed that the pre-calculated setpoints for OLTCs and SVRs were already in place.

Other than these, as shown in Fig. 5, we have included a few nodes containing energy storages, where active power setpoints are predefined. The set of monitored nodes where the voltages are to be maintained around set points is also highlighted. We employed *OpenDSS* [26] as the system simulator, accurately capturing the full nonlinear, non-convex power flow equations. We assumed that the pre-calculated setpoints for OLTCs and SVRs were already in place. Here, we assume that our power system simulation process is quasi-static, where we run power flow simulations considering datasets 2 seconds apart. We also assume that within this time interval, the voltages are measured at the monitoring nodes, broadcasted, and received by the controller nodes. Controller complexity was provided earlier, which demonstrates that low-cost computing devices could be utilized in field deployment. Given the reliance on broadcasting from monitoring nodes, challenges such as data packet corruption, missing measurements, noisy voltage data, and potential cyber-attacks are significant concerns. In particular, we focus on denial-of-service (DoS) attacks as a key cyber threat.

A. Implementation and Results

For simulation purposes, we have strategically selected probing signal frequencies based on prime numbers, and the exact calculation methodology could be easily accessed from the shared repository[2]. We define this set of frequencies as ω (with units rad/s). In the experiments to analyze the convergence of the algorithms, we considered a fixed PV injection of 0 *pu* and a system-wide load of 3 *pu* to understand the convergence characteristics of the controllers. The simulation results are shown in Fig. 7. Notably, both controllers ensured that the steady-state voltage profile remained within acceptable bounds (see Figs. 7(a.2-3) and 7(b.2-3)).

[2]Not shared yet to preserve anonymity.

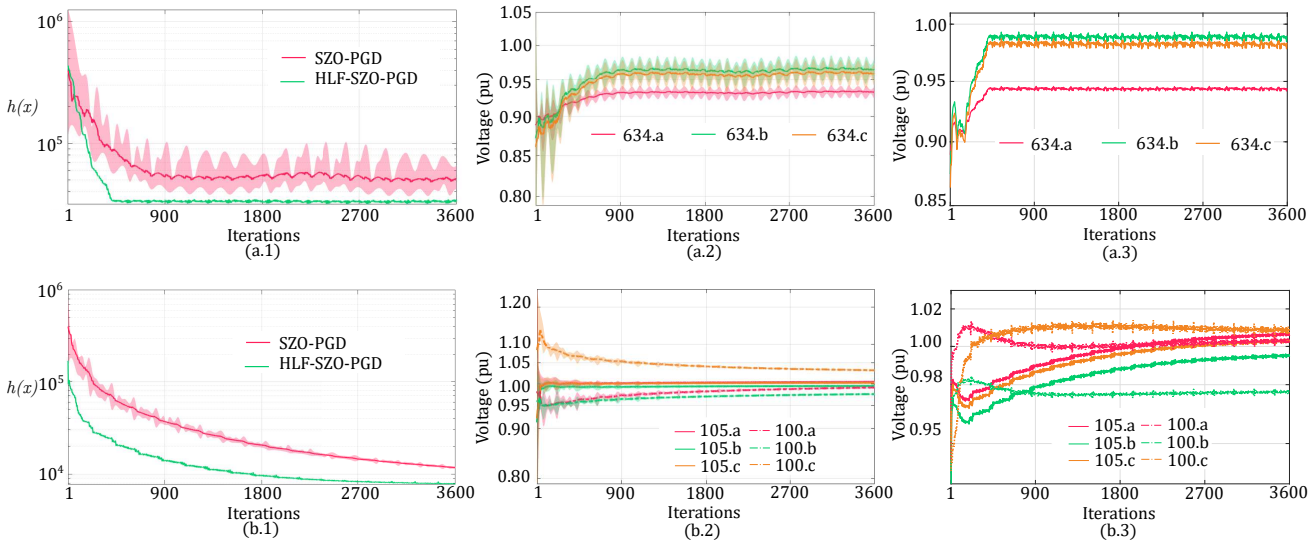


Fig. 7: Algorithm performance under constant load for both IEEE 13-node (a.1-a.3) and synthetic Austin grid (b.1-b.3).

We make two observations here. First, with both SZO-PGD and HLF-SZO-PGD, as the problem size grew, the convergence time grew from 900 iterations in IEEE 13-bus system to 1800 iterations for 127 node austin grid test system (it is not the system size that increases the convergence time, rather it is number of controllers n in (7)). Second, given the voltage measurements with algorithm *SZO-PGD* shown in Figs. 7(a.2) and 7(b.2), it can be seen that the introduction of filters into the *SZO-HLF-PGD* algorithm notably attenuated high-frequency perturbations in voltage measurements. Additionally, due to low variance in the gradient estimate in the filter-based technique, one could choose a higher step size, which can lead to rapid convergence of the algorithm, as demonstrated in Figs. 7(a.1) and 7(b.1). For example, for the IEEE 13-node case, with filters, $\delta = 0.1$ and ω, one could increase step-size from 0.42×10^{-6} to 0.90×10^{-6} with $\alpha = 0.9$ to reduce the convergence time from 900 iterations to $\sim$ 450 iterations. For the Austin grid case, with filters, $\delta = 0.03$ and ω, the step-size could be increased from 0.3×10^{-7} to 10^{-7} with $\alpha = 0.6$ to reduce the convergence time to 1800 iterations (for voltages). While the number of iterations reported here is attributed to the cold start operation, in practical deployment with the ZO controller continuously operating, the algorithm would converge much faster. Note that in all cases the objective function asymptotically converges to the true optima.

Given the iterative nature of the algorithms, there can be multiple of these hyperparameters (δ, k, α, and ω) that dictate the convergence. In the next subsection, we will analyze the same while limiting our focus to the Austin grid case. Note that these hyperparameters could be made adaptive to improve the convergence speed.

B. *Parametric Analysis of* SZO-HLF-PGD *algorithm*

1) Impact of Probing signal Amplitude (δ) and step size (k) on Convergence: Examining Lemma 6 and 8, we ob-

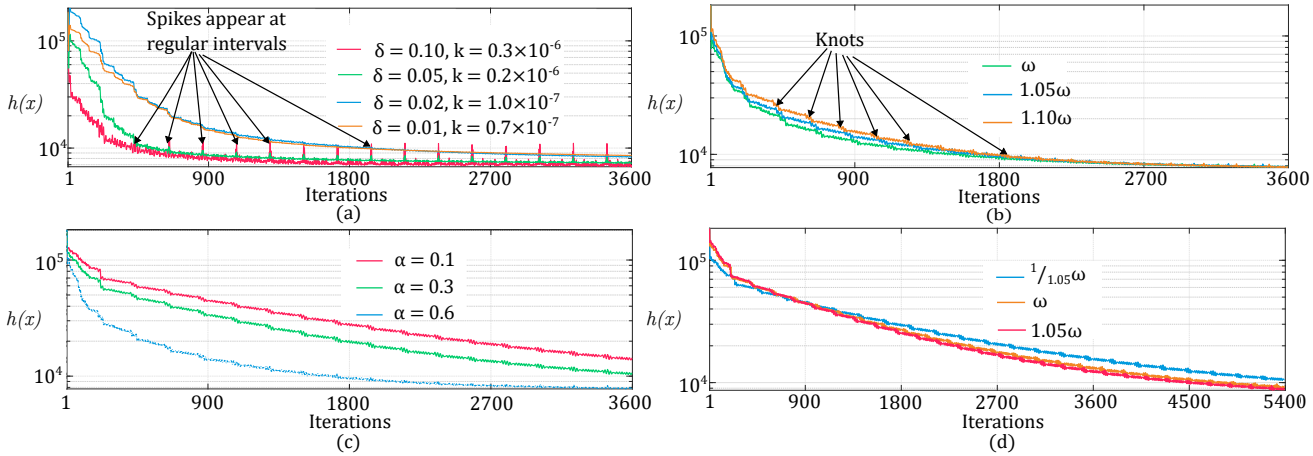


Fig. 8: Parametric analysis of *SZO-HLF-PGD* algorithm considering synthetic Austin grid with constant load demand.

serve that a higher value of δ generally results in a more accurate gradient estimate with lower variance. Consequently, as shown in Fig. 8(a), a higher δ enables the use of a larger k, leading to faster convergence without causing system instability. However, it is important to note that $x_k + \delta u_k$ is injected into the system, which implies that increasing δ will inherently produce a larger perturbation in the function evaluation. Despite this, Fig. 7(b.2) reveals the emergence of spikes or knots in the objective function after a certain number of iterations, indicating that additional analysis is required to understand their origin. Thus, δ and k cannot be set too high for practical implementation. All experiments were conducted with $\alpha = 0.6$ and ω.

2) Impact of Filters: High-pass filters are embedded into the algorithm through (15a) and, as a result, cannot be directly controlled. Impacts of low-pass filters are illustrated in Fig. 8(c), where the sensitivity analysis indicates that the algorithm converges faster with increasing values of α (where $\alpha \in [0, 1]$). Although the exact relationship between α and the convergence rate of the algorithm is not yet established (and will be addressed in future work), it is anticipated that excessively high values of α could cause the control algorithm to diverge. This behavior may arise because momentum terms tend to accumulate over iterations. The experiments presented were conducted using $\delta = 0.03$, $k = 1.0 \times 10^{-7}$, and ω.

3) Impact of Probing Signal Frequency (ω) on Convergence: Here, the probing signals are a part of random number generation; therefore, their impacts are largely absent in the gradient and its variance estimation. However, we can rely on ES control literature to understand the impacts of ω on model convergence. First, intuitively speaking, large ω covers $\mathbb{S}_\delta$ faster, and therefore, probing sinewave will mimic random numbers at a quicker rate. Second, drawing inspiration from ES theory, typically, larger ω would lead to better time-scale separation property [36], leading to faster convergence. However, its impact could be limited by the choice of low-pass filters. In this regard, we have tested the algorithm against $\alpha = 0.6$ with three sets of probing signals, ω, 1.05ω, and 1.10ω and the results are provided in Fig. 8(b). We have also performed experiments with $\frac{1}{1.05}\omega$, ω, and 1.05ω with $\alpha = 0.1$. In both cases, we have chosen $\delta = 0.03$, and $k = 1 \times 10^{-7}$ (see Fig. 8(d)).

With the lower value of α, the convergence of the algorithm gets enhanced with increasing probing signal frequency; however, with higher α, convergence deteriorates with increasing probing signal frequency. Selected probing signal frequencies do not exert a discernible influence on the optimal solutions and could be a future research direction.

C. Algorithmic Robustness

1) Considering Cyber-Anomalies: Here, we systematically assess the performance of the *SZO-HLF-PGD* algorithm under the distinct test conditions within the synthetic Austin grid while considering consistent probing signals and hyper-parameter settings ($\delta = 0.03$, ω, $\alpha = 0.6$, $k = 1\times10^{-7}$). Here, we examined the algorithm convergence considering (a) noisy measurements, (b) delayed measurements, and (c) unavailable data. In scenarios where data are unavailable, the controller is assumed to rely on a default measurement value of 1 *pu*. The simulation results are illustrated in Fig. 9. We have visualized only the phase-a voltages for brevity.

For the noisy measurement test case, we introduce measurement noise levels of 1%, 5%, and 7% across all available measurements. Trajectories of objective functions (see Fig. 9(a.1)) and the voltages in the presence of varying measurement noise (see Fig. 9(a.2)) reveal that increasing measurement noise levels appear to exert no discernible impact on convergence. This robustness could be attributed to the filters incorporated as part of the *SZO-HLF-PGD* algorithm. Zhang et al. have shown in [31, Lemma 12] that in the Zero-Order (ZO) optimization with residual feedback, where random numbers are sampled from the unit sphere and the function estimation noise is bounded, the squared moment of the ZO estimator is upper-bounded as a function of the noise. A similar property could be applicable to our problem and warrants further investigation. It could also be expected that the controller may diverge with large measurement noises.

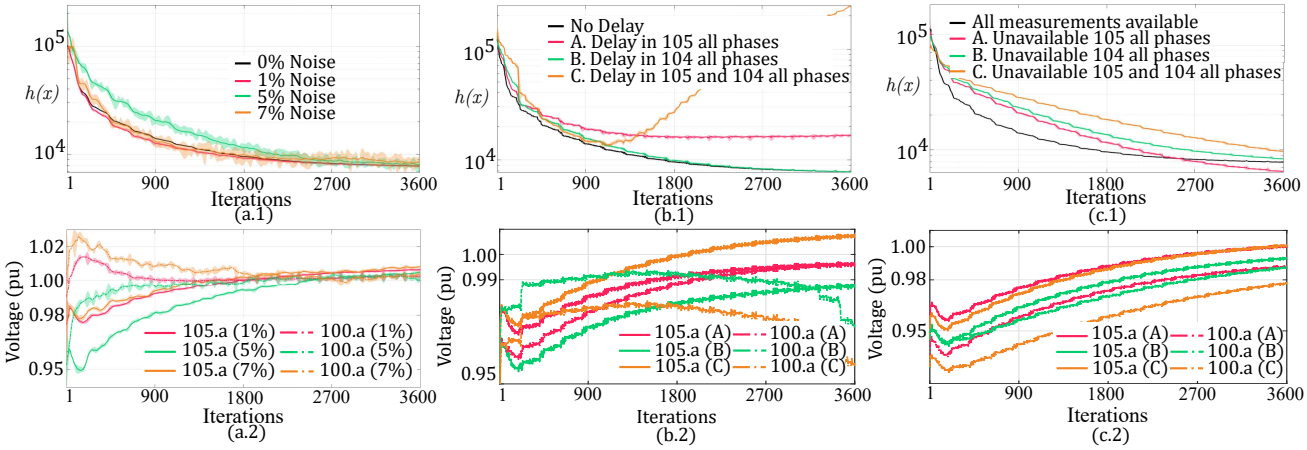


Fig. 9: Robustness of *HLF-SZO-PGD* algorithm with constant load under different scenarios: (a) noise, (b) delay in measurements, (c) missing measurements. The voltages shown in this figure correspond to Phase A only.

In the delayed measurements scenario, we introduce a fixed 3-sample delay in voltage measurements across all phases from a designated node. We examine three distinct cases: (i) measurement delay from node '105,' (ii) measurement delay from node '104,' and (iii) simultaneous measurement delay from both nodes '105' and '104.' As shown in Fig. 9(b.1), delays in different nodes manifest varying effects on convergence. All the phase voltages can be equally affected but not shown for brevity.

Concerning the replacement of unavailable measurements with pseudo-measurements of 1 *pu*, we conduct three experiments: (i) unavailability of measurements from node '105,' (ii) unavailability of measurements from node '104,' and (iii) unavailability of measurements from both nodes '105' and '104.' Comparative analysis between Figs. 9(b.1) and 9(c.1) suggest that the absence of measurements from multiple nodes appears to have a comparatively lesser impact. This could be because the unavailability of a measurement from a node implies the absence of a corresponding control objective in (1). This is also the reason the controllers reach a new operating point with a better control objective than that of with all measurements available, as shown in Fig. 9(c.1). Notably, in contemporary IP-based communication systems, measurements may include timestamps – which could be utilized before deploying the delayed signals with significant improvement in a performance gain. While severe data loss can be detrimental to the overall system performance, it does not significantly impair the controller's performance, especially when the controller observes severe data loss or longer delays.

2) Time-Varying Case for Synthetic Austin Grid: Here, we evaluate the performance of the *SZO-HLF-PGD* algorithm under time-varying load and generation conditions, as depicted in Fig. 6. The evaluation is conducted in the presence of traditional voltage control devices, such as OLTCs and SVRs, which operate using predefined setpoints. The impacts of these traditional controllers are shown in Fig. 10. The OLTCs and SVRs adjust their setpoints several times throughout the day—specifically around 08:00, 10:00, 12:00, 15:00, and 17:00 hours—marked by vertical dotted lines. Meanwhile, the ZO controllers operate independently using local measurements, with parameters set to $\delta = 0.03$, ω, $\alpha = 0.6$, $k = 1 \times 10^{-7}$. Notably, despite having no communication among the controllers, the ZO controllers are able to respond to exogenous sharp voltage change induced by the traditional controllers while limiting the voltages within reliability limits. Although detailed results are omitted for brevity, our analysis indicates that the introduction of additional control actions significantly reduces reactive power demand from the ZO controllers. This suggests that optimizing the setpoints of traditional controllers is crucial for minimizing the reactive power output of the ZO controllers, thereby enhancing overall voltage regulation. Instead of simultaneous control of multiple tap positions, tap operations could be coordinated considering the presence of ZO voltage controller dynamics for an even better voltage profile. Developing such a hybrid control methodology represents a promising direction for future research.

IV. Conclusion

This paper overcomes practical challenges in deploying zeroth-order (ZO) methods for real-world voltage control by conducting extensive tests using the 3-phase unbalanced IEEE 13-node and synthetic Austin grid feeders, along with theoretical guarantees. In the context of ZO methods, for the first time, we have employed sinusoidal waveforms to create pseudo-random numbers, showing that existing zeroth-order estimators can accurately approximate the system gradient needed for generating control signals. Additionally, we have utilized methods from classical optimization literature with filtering properties to improve controller performance. For comparison, we adopted a projected gradient descent method and developed two algorithms, *SZO-PGD* and *HLF-SZO-PGD*. Our results demonstrate that the HLF-SZO-PGD algorithm, which integrates filtration, enhances gradient tracking without compromising the end result. The mathematical guarantees provided confirm the controller's effectiveness in terms of convergence, robustness, and resilience, not only in simulations but also under realistic operational conditions. We see that our approach works seamlessly alongside traditional voltage regulators and dynamically adjusts to fluctuating grid conditions, even without detailed system models. The insights from this study lay a robust groundwork for future enhancements of ZO methods. Future research will aim to reduce reliance on system perturbations, incorporate advanced predictive analytics via machine learning for proactive grid management, and further streamline the algorithms to lower computational requirements.

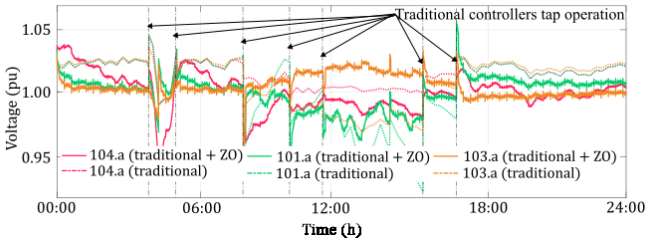


Fig. 10: Voltage profile with time-varying load profile in synthetic Austin grid feeder with OLTCs, SVRs, with diverse PVs, numerous controllers, and monitoring nodes to demonstrate *HLF-SZO-PGD* controller algorithm scalability. The voltages shown in this figure correspond to Phase A only.

List of Acronyms

DER	Distributed Energy Resources
OLTC	On-load Tap-Changer
SVR	Static Voltage Regulator
ES	Extremum seeking methods
HLF	High/Low-pass Filter
OVC	Optimal Voltage Control
PGD	Projected Gradient Descent
RL	Reinforcement Learning
(S)ZO	(Single-point) Zeroth-Order
ICT	Information and Communication Technologies
DoS	Denial of Service

Appendix A
Proof of Lemma 1

First, recognize that the arcsine distribution as given in (A.1) describes the probability distribution of the values taken by

the sine function, particularly highlighting how these values are concentrated near the endpoints.

$$f(x) = \frac{1}{\pi\delta\sqrt{1-\left(\frac{x}{\delta}\right)^2}}, \quad \text{for } x \in [-\delta, \delta]. \tag{A.1}$$

In our case, a random vector $u \in \mathbb{H}_\delta$ where each coordinate u_j is independently sampled from the arcsine distribution. Considering interval $u_j \in [-\delta, -\delta-\epsilon] \cup [\delta-\epsilon, \delta]$, substituting $u_j = \delta \sin\theta$, and with symmetry, we get:

$$\begin{aligned} P\left(|u_j| > |\delta-\epsilon| \middle| \forall k \neq j\right) &= \frac{2}{\pi}\int_{\arcsin\left(\frac{\delta-\epsilon}{\delta}\right)}^{\frac{\pi}{2}} d\theta \\ &= \frac{2}{\pi}\left(\frac{\pi}{2} - \arcsin\left(1-\frac{\epsilon}{\delta}\right)\right). \end{aligned}$$

For small ϵ, using the approximation $\arcsin\left(1-\frac{\epsilon}{\delta}\right) \approx \frac{\pi}{2} - \sqrt{2\epsilon/\delta}$, we obtain:

$$P\left(|u_j| \geq |\delta-\epsilon| \middle| \forall k \neq j\right) \approx \frac{2\sqrt{2\epsilon/\delta}}{\pi}.$$

Since the coordinates are independent:

$$P\left(\text{all } |u_j| < |\delta-\epsilon|\right) = \left[P\left(|u_j| < |\delta-\epsilon|\right)\right]^n.$$

For large n, using the approximation $\left(1-\frac{c}{\sqrt{n}}\right)^n \approx e^{-c\sqrt{n}}$, where $c = \frac{2\sqrt{2\epsilon/\delta}}{\pi}$, we have:

$$P\left(\text{all } |u_j| < |\delta-\epsilon|\right) \approx e^{-c\sqrt{n}} \to 0 \quad \text{as } n \to \infty. \tag{A.2}$$

Then, the probability that at least one coordinate satisfies $|u_j| \geq |\delta-\epsilon|$ is:

$$\begin{aligned} P\left(\max_j |u_j| \geq |\delta-\epsilon|\right) &= 1 - P\left(\text{all } |u_j| < \delta-\epsilon\right) \\ &\approx 1 - e^{-c\sqrt{n}} \to 1 \quad \text{as } n \to \infty. \end{aligned}$$

□

Appendix B
Proof of Lemma 2

The facet $F_j^\pm$ of $\mathbb{S}_\delta$ corresponding to fixing the j^{th} coordinate at δ or $-\delta$:

$$F_j^+ = \left\{ v \in H_\delta \;\middle|\; v_j = \delta \right\}, \quad F_j^- = \left\{ v \in H_\delta \;\middle|\; v_j = -\delta \right\}, \quad \text{for } j = 1, 2, \ldots, d. \tag{B.1}$$

Consider the probability that u is near the facet F_j^+, i.e., $u_j \geq \delta - \epsilon$:

$$P(F_j^+(\epsilon)\big|\forall k \neq j) = P(\delta - \epsilon \leq u_j \leq \delta). \tag{16}$$

We took conditional probability, which summed to zero Second, since all coordinates u_j are identically and independently distributed. Note that because of symmetry,

$$P(F_j^+(\epsilon)\big|\forall k \neq j) = P(F_j^-(\epsilon)\big|\forall k \neq j) \approx \frac{\sqrt{2\epsilon/\delta}}{\pi}. \tag{17}$$

The approximation is due to small ϵ. See Appendix A for detailed calculation. Note that the probability is independent of surface j. □

Appendix C
Proof of Lemma 3

The expected value of $h(x+\delta u)\, u$ is given by:

$$\mathbb{E}_{u\in\mathbb{S}_\delta}[h(x+\delta u)\, u] = \frac{1}{\text{SurfaceArea}(\mathbb{S}_\delta)} \int_{\mathbb{S}_\delta} h(x+\delta u)\, u \, d\mathbb{S}(u). \tag{A.1}$$

Here, $\text{SurfaceArea}(\mathbb{S}_\delta)$ is the surface area of the hypercube. Since the surface of $\mathbb{S}_\delta$ consists of $2n$ faces, we can write the integral as a sum over each face:

$$\begin{aligned} \int_{\mathbb{S}_\delta} h(x+\delta u)\, u \, d\mathbb{S}(u) = \sum_{i=1}^n \Bigg(&\int_{F_i^+} h(x+\delta u)\, u_i \, d\mathbb{S}(u) \\ &+ \int_{F_i^-} h(x+\delta u)\, u_i \, d\mathbb{S}(u) \Bigg) \mathbf{e}_i, \end{aligned}$$

where F_i^+ and F_i^- represent the faces where $u_i = 1$ and $u_i = -1$, respectively, and $\mathbf{e}_i$ is the unit vector in the i^{th} coordinate direction. Now let $u_{-i} \in [-1,1]^{n-1}$ denote the components of u excluding u_i. Define:

$$\begin{aligned} \phi_i^\pm(x) &= \int_{u_{-i}\in[-1,1]^{n-1}} h\left(x + \delta\left(u_{-i}, \pm 1\right)\right) d\mathbb{S}(u_{-i}) \\ &\overset{v_{-i}=\delta u_{-i}}{=} \int_{v_{-i}\in[-\delta,\delta]^{n-1}} h\left(x + \left(v_{-i}, \pm 1\delta\right)\right) dv_{-i}. \end{aligned}$$

Then, the integral over each face is given by:

$$I_i = \phi_i^+(x) - \phi_i^-(x). \tag{A.2}$$

Next, consider the partial derivative of the smoothed function $h_\delta(x)$:

$$h_\delta(x) = \frac{1}{(2\delta)^n}\int_{\mathbb{H}_\delta} h(x+v)\, dv.$$

Differentiating $h_\delta(x)$ with respect to x_i:

$$\frac{\partial h_\delta(x)}{\partial x_i} = \frac{1}{(2\delta)^n}\int_{\mathbb{H}_\delta} \frac{\partial h(x+v)}{\partial x_i}\, dv = \frac{1}{(2\delta)^n}\int_{\mathbb{H}_\delta} \frac{\partial h(x+v)}{\partial v_i}\, dv.$$

Applying the fundamental theorem of calculus:

$$\begin{aligned} \int_{\mathbb{H}_\delta} \frac{\partial h(x+v)}{\partial v_i}\, dv. &= \int_{v_{-i}\in[-\delta,\delta]^{n-1}} \int_{v_i=-\delta}^{\delta} \frac{\partial h(x+v)}{\partial v_i}\, dv_i \\ &= \int_{[-\delta,\delta]^{n-1}} \left[h\left(x+v_{-i}+\delta e_i\right) - h\left(x+v_{-i}-\delta e_i\right)\right] dv_{-i}.. \end{aligned} \tag{A.3}$$

From (A.2) and (A.3),

$$I_i = (2\delta)^n \frac{\partial h_\delta(x)}{\partial x_i}.$$

The expected value over the surface from (A.1) will be:

$$\mathbb{E}_{u\in\mathbb{S}_\delta}[h(x+\delta u)\, u_i] = \frac{I_i}{\text{SurfaceArea}(\mathbb{S}_\delta)}.$$

Given, the surface area of the hypercube is $2n(2\delta)^{n-1}$, and due to Lemma 2 each faces are equiprobable:

$$\mathbb{E}_{u\in\mathbb{S}_\delta}[h(x+\delta u)\,u_i] = \frac{(2\delta)^n \frac{\partial h_\delta(x)}{\partial x_i}}{2n(2\delta)^{n-1}} = \frac{\delta}{n}\frac{\partial h_\delta(x)}{\partial x_i}.$$

Finally, in vector form:

$$\mathbb{E}_{u\in\mathbb{S}_\delta}[h(x+\delta u)\,u] = \frac{\delta}{n}\nabla h_\delta(x).$$

□

Appendix D
Proof of Lemma 4

Choose one of the $2n$ facets of $\mathbb{S}_\delta$. The remaining $n-1$ coordinates $\{u_j\}_{j\neq i}$ are independently sampled from the arcsine distribution. Then, the squared norm of u is given by:

$$\|u\|^2 = u_i^2 + \sum_{j\neq i} u_j^2.$$

Since u_i is fixed to either δ or $-\delta$, $\mathbb{E}[u_i^2] = \delta^2$. For the remaining $j \neq i$, u_j is sampled from the arcsine distribution bounded by $[\delta, \delta]$, the expectation of u_j^2 considering independence:

$$\mathbb{E}[u_j^2] = \frac{\delta^2}{\pi}\int_{-1}^{1}\frac{y^2}{\sqrt{1-y^2}}\,dy. \tag{B.1}$$

The above integral was contained through the change of variables. The integral $\int_{-1}^{1}\frac{y^2}{\sqrt{1-y^2}}\,dy$ can be computed by noting that it is equal to $\frac{\pi}{2}$. Thus:

$$\mathbb{E}[u_j^2] = \frac{\delta^2}{\pi}\cdot\frac{\pi}{2} = \frac{\delta^2}{2}.$$

Therefore, the expected value of the squared norm $\|u\|^2$ is

$$\mathbb{E}[\|u\|^2] = \mathbb{E}[u_i^2] + \sum_{j\neq i}\mathbb{E}[u_j^2] = \delta^2 + (n-1)\cdot\frac{\delta^2}{2} = \delta^2\cdot\frac{n+1}{2}. \tag{B.2}$$

Since this calculation holds for any of the $2n$ facets, and each coordinate contributes equally due to Lemma 2, the overall expected value of the squared norm can be subsequently obtained. □

Appendix E
Proof of Lemma 7

Assume without loss of generality that $u_1 = \delta$ (due to symmetry, the result is the same for $u_1 = -\delta$), and the remaining coordinates $u_2, u_3, \ldots, u_n$ are independently selected. The fourth power of the norm is:

$$\|u\|^4 = \left(\sum_{i=1}^n u_i^2\right)^2 = \delta^4 + 2\delta^2\sum_{i=2}^n u_i^2 + \left(\sum_{i=2}^n u_i^2\right)^2.$$

Taking the conditional expectation on the facet:

$$\mathbb{E}[\|u\|^4] = \mathbb{E}[\delta^4] + 2\delta^2\mathbb{E}\left[\sum_{i=2}^n u_i^2\right] + \mathbb{E}\left[\left(\sum_{i=2}^n u_i^2\right)^2\right]. \tag{C.1}$$

Each u_i for $i = 2, 3, \ldots, n$ sampled from the arcsine distribution bounded by $[\delta, \delta]$, and are independent. We calculate the expectation of each term separately:

(i) $\mathbb{E}\left[\sum_{i=2}^n u_i^2\right]$: We have: $\mathbb{E}[u_i^2] = \frac{\delta^2}{2}$ (see, Appendix D). Thus:

$$\mathbb{E}\left[\sum_{i=2}^n u_i^2\right] = (n-1)\cdot\frac{\delta^2}{2} = \frac{(n-1)\delta^2}{2}. \tag{C.2}$$

(ii) $\mathbb{E}\left[\left(\sum_{i=2}^n u_i^2\right)^2\right]$: Expanding the square:

$$\left(\sum_{i=2}^n u_i^2\right)^2 = \sum_{i=2}^n u_i^4 + 2\sum_{2\leq i<j\leq n} u_i^2 u_j^2.$$

Note that: $\mathbb{E}[u_i^4] = \frac{3}{8}\delta^4$,
and, $\mathbb{E}[u_i^2 u_j^2] = \left(\mathbb{E}[u_i^2]\right)^2 = \frac{\delta^2}{2}$. Then, summing over all terms:

$$\mathbb{E}\left[\left(\sum_{i=2}^n u_i^2\right)^2\right] = (n-1)\cdot\frac{3}{8}\delta^4 + (n-1)(n-2)\frac{\delta^4}{4}. \tag{C.3}$$

Substituting (C.2)-(C.3) to (C.1)

$$\mathbb{E}[\|u\|^4] = \delta^4\left(1 + 2\cdot\frac{n-1}{2} + \frac{3(n-1)}{8} + \frac{(n-1)(n-2)}{4}\right) = \frac{2n^2+5n+1}{8}\delta^4. \tag{18}$$

Since this calculation holds for all the coordinates, and each coordinate contributes equally due to the symmetry of the hypercube, the overall expected value can be subsequently obtained. □

Appendix F
Proof of Lemma 8

Based on [31, Lemma 6] and given that h is a L-Lipschutz continuous, we have

$$\begin{aligned}&\mathbb{E}[\|\tilde{g}(x_k)\|^2]\\ &= \mathbb{E}\left[\frac{1}{\delta^2}(h(x_k+\delta u_k) - h(x_{k-1}+\delta u_{k-1}))^2\|u_k\|^2\right]\\ &\leq \frac{2L_0^2}{\delta^2}\mathbb{E}[\|x_k - x_{k-1}\|^2\|u_k\|^2] + 2L^2\mathbb{E}[\|u_k - u_{k-1}\|^2\|u_k\|^2].\end{aligned} \tag{D.1}$$

Since u_k is independently sampled from $x_k - x_{k-1}$, we have $\mathbb{E}[\|x_k - x_{k-1}\|^2\|u_k\|^2] = \mathbb{E}[\|x_k - x_{k-1}\|^2]\mathbb{E}[\|u_k\|^2]$. For the second part, we get that $\mathbb{E}[\|u_k - u_{k-1}\|^2\|u_k\|^2] \leq 2\mathbb{E}[(\|u_k\|^2 + \|u_{k-1}\|^2)\|u_k\|^2] = 2\mathbb{E}[\|u_k\|^4] + 2\mathbb{E}[\|u_{k-1}\|^2\|u_k\|^2] = 2\left(\mathbb{E}[\|u_k\|^4] + \mathbb{E}[\|u_k\|^2]^2\right) = \frac{1}{4}\left(4n^2+9n+3\right)$.
Plugging these bounds into inequality (A.1), we have that

$$\mathbb{E}[\|\tilde{g}(x_k)\|^2] \leq \frac{2(n+2)L_0^2}{3\delta^2}\mathbb{E}[\|x_k - x_{k-1}\|^2] + \frac{L^2}{2}\left(4n^2+9n+3\right).$$

Now, using the PGD update rule and the non-expansiveness property:

$$
\begin{aligned}
&\mathbb{E}[\|x_k - x_{k-1}\|^2 | x_{k-1}] \\
&= \mathbb{E}[\|P_C(x_{k-1} - \eta_{k-1}\tilde{g}(x_{k-1})) - x_{k-1}\|^2 | x_{k-1}] \\
&\overset{\text{non-expansiveness}}{\leq} \|(x_{k-1} - \eta_{k-1}\tilde{g}(x_{k-1})) - x_{k-1})\|^2 \\
&= \eta_{k-1}^2 \mathbb{E}[\|g(x_{k-1})\|^2]
\end{aligned}
$$

Now from (D.1),

$$
\begin{aligned}
&\mathbb{E}[\|\tilde{g}(x_k)\|^2] \\
&\leq \frac{(n+1)L_0^2}{\delta^2}\eta_{k-1}^2 \mathbb{E}[\|\tilde{g}(x_{k-1})\|^2] + \frac{L^2}{2}\left(4n^2 + 9n + 3\right).
\end{aligned} \tag{D.2}
$$

□

REFERENCES

[1] S. Majumder, A. Vosughi, H. M. Mustafa, T. E. Warner, and A. K. Srivastava, "On the Cyber-Physical Needs of DER-Based Voltage Control/Optimization Algorithms in Active Distribution Network," *IEEE Access*, vol. 11, pp. 64 397–64 429, 2023.

[2] X. Chen, J. I. Poveda, and N. Li, "Model-free feedback constrained optimization via projected primal-dual zeroth-order dynamics," *arXiv preprint arXiv:2206.11123*, 2022.

[3] ——, "Model-Free Optimal Voltage Control via Continuous-Time Zeroth-Order Methods," *arXiv preprint arXiv:2103.14703*, 2021.

[4] ——, "Continuous-time zeroth-order dynamics with projection maps: Model-free feedback optimization with safety guarantees," *arXiv preprint arXiv:2303.06858*, 2023.

[5] X. Chen, G. Qu, Y. Tang, S. Low, and N. Li, "Reinforcement Learning for Decision-Making and Control in Power Systems," in *Women in Power: Research and Development Advances in Electric Power Systems*, ser. Women in Engineering and Science. Cham: Springer International Publishing, 2023, pp. 265–285.

[6] K. B. Ariyur and M. Krstic, *Real-Time Optimization by Extremum-Seeking Control*. John Wiley & Sons, 2003.

[7] A. Hauswirth, Z. He, S. Bolognani, G. Hug, and F. Dörfler, "Optimization algorithms as robust feedback controllers," *Annual Reviews in Control*, vol. 57, p. 100941, 2024.

[8] J. I. Poveda and N. Li, "Robust hybrid zero-order optimization algorithms with acceleration via averaging in time," *Automatica*, vol. 123, p. 109361, 2021.

[9] D. B. Arnold, M. Negrete-Pincetic, M. D. Sankur, D. M. Auslander, and D. S. Callaway, "Model-Free Optimal Control of VAR Resources in Distribution Systems: An Extremum Seeking Approach," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3583–3593, Sep. 2016.

[10] H. Nazaripouya, H. R. Pota, C.-C. Chu, and R. Gadh, "Real-Time Model-Free Coordination of Active and Reactive Powers of Distributed Energy Resources to Improve Voltage Regulation in Distribution Systems," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 3, pp. 1483–1494, Jul. 2020.

[11] C. Hu, X. Zhang, and Q. Wu, "Gradient-free accelerated event-triggered scheme for constrained network optimization in smart grids," *IEEE Transactions on Smart Grid*, 2023.

[12] X. Li, Y. Li, J. E. Seem, and P. Lei, "Maximum power point tracking for photovoltaic systems using adaptive extremum seeking control," in *IEEE Conference on Decision and Control and European Control Conference*. Orlando, FL, USA: IEEE, Dec. 2011, pp. 1503–1508.

[13] Y. Huang and J. Hu, "Zeroth-order learning in continuous games via residual pseudogradient estimates," *arXiv preprint arXiv:2301.02279*, 2023.

[14] J. Johnson, A. Summers, R. Darbali-Zamora, J. Hernandez-Alvidrez, J. Quiroz, D. Arnold, and J. Anandan, "Distribution Voltage Regulation Using Extremum Seeking Control With Power Hardware-in-the-Loop," *IEEE Journal of Photovoltaics*, vol. 8, no. 6, pp. 1824–1832, Nov. 2018.

[15] D. K. Molzahn, F. Dorfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, "A Survey of Distributed Optimization and Control Algorithms for Electric Power Systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2941–2962, Nov. 2017.

[16] H. J. Liu, W. Shi, and H. Zhu, "Hybrid Voltage Control in Distribution Networks Under Limited Communication Rates," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2416–2427, May 2019.

[17] N. Patari, A. K. Srivastava, and N. Li, "Distributed Optimal Voltage Control Considering Latency and Asynchronous Communication for Three Phase Unbalanced Distribution Systems," *IEEE Transactions on Power Systems*, vol. 38, no. 2, pp. 1033–1043, Mar. 2023.

[18] G. Qu and N. Li, "Optimal Distributed Feedback Voltage Control Under Limited Reactive Power," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 315–331, Jan. 2020.

[19] P. S. Sarker, S. Majumder, M. F. Rafy, and A. K. Srivastava, "Impact Analysis of Cyber-Events on Distributed Voltage Control with Active Power Curtailment," in *2022 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*. Jaipur, India: IEEE, Dec. 2022, pp. 1–6.

[20] S. Liu, P.-Y. Chen, B. Kailkhura, G. Zhang, A. O. Hero III, and P. K. Varshney, "A primer on zeroth-order optimization in signal processing and machine learning: Principals, recent advances, and applications," *IEEE Signal Processing Magazine*, vol. 37, no. 5, pp. 43–54, 2020.

[21] X. Chen, Y. Tang, and N. Li, "Improve single-point zeroth-order optimization using high-pass and low-pass filters," in *International Conference on Machine Learning*. PMLR, 2022, pp. 3603–3620.

[22] Z. He, S. Bolognani, J. He, F. Dörfler, and X. Guan, "Model-free nonlinear feedback optimization," *IEEE Transactions on Automatic Control*, 2023.

[23] Z. He, S. Bolognani, M. Muehlebach, and F. Dörfler, "Gray-box nonlinear feedback optimization," *arXiv preprint arXiv:2404.04355*, 2024.

[24] W. Wang, Z. He, G. Belgioioso, S. Bolognani, and F. Dörfler, "Online Feedback Optimization over Networks: A Distributed Model-free Approach," *arXiv preprint arXiv:2403.19834*, 2024.

[25] A. Williams, M. Krstic, and A. Scheinker, "Semi-global practical extremum seeking with practical safety," in *2023 62nd IEEE Conference on Decision and Control (CDC)*. IEEE, 2023, pp. 6774–6779.

[26] R. C. Dugan and T. E. McDermott, "An open source platform for collaborating on smart grid research," in *2011 IEEE Power and Energy Society General Meeting*, Jul. 2011, pp. 1–7.

[27] A. Beck, *First-order methods in optimization*. SIAM, 2017.

[28] V. Kekatos, L. Zhang, G. B. Giannakis, and R. Baldick, "Voltage Regulation Algorithms for Multiphase Power Distribution Grids," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3913–3923, Sep. 2016.

[29] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, Mass: Athena scientific, 1999.

[30] A. D. Flaxman, A. T. Kalai, and H. B. McMahan, "Online convex optimization in the bandit setting: gradient descent without a gradient," *arXiv preprint cs/0408007*, 2004.

[31] Y. Zhang, Y. Zhou, K. Ji, and M. M. Zavlanos, "A new one-point residual-feedback oracle for black-box learning and control," *Automatica*, vol. 136, p. 110006, Feb. 2022.

[32] Y. Tan, D. Nešić, and I. Mareels, "On the choice of dither in extremum seeking systems: A case study," *Automatica*, vol. 44, no. 5, pp. 1446–1450, May 2008.

[33] W. Tao, S. Long, G. Wu, and Q. Tao, "The role of momentum parameters in the optimal convergence of adaptive polyak's heavy-ball methods," *arXiv preprint arXiv:2102.07314*, 2021.

[34] C. Mateo, F. Postigo, T. Elgindy, A. B. Birchfield, P. Dueñas, B. Palmintier, N. Panossian, T. Gómez, F. de Cuadra, T. J. Overbye *et al.*, "Building and validating a Large-Scale combined transmission & distribution synthetic electricity system of Texas," *International Journal of Electrical Power & Energy Systems*, vol. 159, p. 110037, 2024.

[35] "Pecan Street – Pecan Street Inc." [Online]. Available: https://www.pecanstreet.org/about/

[36] Y. Tan, W. H. Moase, C. Manzie, D. Nesic, and I. M. Y. Mareels, "Extremum seeking from 1922 to 2010," in *Proceedings of the 29th Chinese Control Conference*. Beijing, China: IEEE, 2010, pp. 14–26.