Journal of Optimization Theory and Applications (2025) 206:11
https://doi.org/10.1007/s10957-025-02683-8

A Regularized Variance-Reduced Modified Extragradient Method for Stochastic Hierarchical Games

Shisheng Cui[1] · Uday V. Shanbhag[2] · Mathias Staudigl[3]

Received: 25 January 2024 / Accepted: 30 March 2025

Abstract

We consider an N-player hierarchical game in which the ith player's objective comprises of an expectation-valued term, parametrized by rival decisions, and a hierarchical term. Such a framework allows for capturing a broad range of stochastic hierarchical optimization problems, Stackelberg equilibrium problems, and leader-follower games. We develop an iteratively regularized and smoothed variance-reduced modified extragradient framework for iteratively approaching hierarchical equilibria in a stochastic setting. We equip our analysis with rate statements, complexity guarantees, and almost-sure convergence results. We then extend these statements to settings where the lower-level problem is solved inexactly and provide the corresponding rate and complexity statements. Our model framework encompasses many game theoretic equilibrium problems studied in the context of power markets. We present a realistic application to the study of virtual power plants, emphasizing the role of hierarchical decision making and regularization. Preliminary numerics suggest that empirical behavior compares well with theoretical guarantees.

Keywords Hierarchical games · Variational inequality problems · Tikhonov regularization · Iterative smoothing · Variance-reduction · Virtual power plants (VPPs)

Mathematics Subject Classification 65K15 · 90C15 · 90C30 · 90C33

Communicated by Nadia Oudjane.

✉ Mathias Staudigl
mathias.staudigl@uni-mannheim.de

Shisheng Cui
css@bit.edu.cn

Uday V. Shanbhag
udaybag@umich.edu

1 School of Automation, Beijing Institute of Technology, Beijing 100081, China

2 University of Michigan, Ann Arbor, MI 48109, USA

3 Department of Mathematics, University of Mannheim, B6, 68159 Mannheim, Germany

Published online: 10 May 2025

1 Introduction

In this paper we consider a class of stochastic hierarchical optimization problems and games, generalizing many learning problems involving sequential optimization. Consider a collection of N-agents, where the ith agent solves the optimization problem parametrized by rival decisions $\mathbf{x}_{-i}$:

$$\min_{\mathbf{x}_i,\mathbf{y}_i} \{\ell_i(\mathbf{x}_i, \mathbf{x}_{-i}, \mathbf{y}_i) \triangleq f_i(\mathbf{x}_i, \mathbf{x}_{-i}) + g_i(\mathbf{x}_i, \mathbf{y}_i)\}, \text{ s.t. } \mathbf{x}_i \in \mathcal{X}_i, \mathbf{y}_i \in \mathrm{SOL}(\phi_i(\mathbf{x}_i, \cdot), \mathcal{Y}_i). \tag{P}$$

We let $i \in \mathcal{I} \triangleq \{1, \ldots, N\}$ represent a set of *leaders*, characterized by two loss functions: (i) $f_i(\mathbf{x}) \triangleq \mathbb{E}_\xi[F_i(\mathbf{x}, \xi)]$, depending on the entire action profile $\mathbf{x} \triangleq (\mathbf{x}_i, \mathbf{x}_{-i}) = (\mathbf{x}_1, \ldots, \mathbf{x}_N) \in \mathcal{X} \triangleq \prod_{i\in\mathcal{I}} \mathcal{X}_i$; (ii) $g_i(\mathbf{x}_i, \mathbf{y}_i)$ is a deterministic function, jointly controlled by leader i's decision variable $\mathbf{x}_i \in \mathcal{X}_i$ and a follower's decision variable $\mathbf{y}_i \in \mathcal{Y}_i$. Each leader's optimization problem exhibits two sets of private constraints, the first given by $\mathbf{x}_i \in \mathcal{X}_i \subseteq \mathbb{R}^{n_i}$, while the second are equilibrium constraints represented by the solution set of a parameterized variational inequality $\mathrm{VI}(\phi_i(\mathbf{x}_i, \cdot), \mathcal{Y}_i)$, which reads as

$$\text{Find } \mathbf{y}_i(\mathbf{x}_i) \in \mathcal{Y}_i \text{ satisfying } \langle \phi_i(\mathbf{x}_i, \mathbf{y}_i(\mathbf{x}_i)), \mathbf{y}_i - \mathbf{y}_i(\mathbf{x}_i)\rangle \geq 0 \quad \forall \mathbf{y}_i \in \mathcal{Y}_i. \tag{1}$$

We denote the set of points $\mathbf{y}_i$ satisfying this condition by $\mathrm{SOL}(\phi_i(\mathbf{x}_i, \cdot), \mathcal{Y}_i)$. This VI is defined in terms of a closed convex set $\mathcal{Y}_i \subseteq \mathbb{R}^{m_i}$ and an expectation-valued mapping $\phi_i(\mathbf{x}, \mathbf{y}_i) \triangleq \mathbb{E}_\xi[\Phi_i(\mathbf{x}, \mathbf{y}_i, \xi)]$. All the problem data are affected by random noise represented by a random variable $\xi : \Omega \to \Xi$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in a measurable space Ξ. Such hierarchical optimization problems traditionally play a key role in operations research and engineering, where they are deeply connected to bilevel programming [19] and mathematical programs under equilibrium constraints (MPEC) [51]. In fact, the canonical MPEC formulation is obtained from (P) when $N = 1$. The multi-agent formulation (P) also relates to leader-follower and Stackelberg games (cf. [3, 17, 45, 59, 64]) which are a traditional model in economics, and also have received increased attention in machine learning recently [5, 9, 25, 50]. Economic equilibria in power markets have been extensively studied using a complementarity framework (cf. [27, 32–34]). More recently, stochastic generalizations have been examined where uncertainty in price and cost functions have been addressed [42, 63]. A recent survey on this topic is [7]. As an immediate application of our algorithmic framework, we present in Sect. 5 a model inspired by Hobbs and Pang [34], but suitably modified to account for uncertainty in prices and costs, multi-period settings with ramping constraints, and the incorporation of virtual power plants (VPPs) (see [22, 53] for a review of VPPs and power markets).

1.1 Our Contributions and Related Work

1.1.1 Hierarchical Optimization, Games and Uncertain Generalizations

To date, hierarchical optimization has been studied under the umbrella of bilevel programming [19, 20] and mathematical programs with equilibrium constraints (MPECs) [51, 57]. Algorithmic schemes for resolving MPECs where the lower-level problem is an optimization problem, or a variational inequality, have largely emphasized either implicit approaches [57] or regularization/penalization-based techniques [36, 49, 51]. Yet, there appears to have been a glaring lacuna in non-asymptotic rate and complexity guarantees for resolving hierarchical optimization and their stochastic and game-theoretic variants. This gap has been partially addressed in the recent papers [15, 16]. Both papers present variance-reduced solution strategies for various versions of hierarchical optimization problems and games, respectively, relying on variance reduction via a sequence of increasing mini-batches. Finite-time and almost sure convergence to solutions is proved under convexity/monotonicity assumptions on the problem data. However, the framework for monotone games in [15] requires exact solutions of lower-level problems, significantly impacting its efficient implementation in large-scale settings. We complement this literature by developing a novel regularized smoothed variance reduction method for the family of hierarchical games (P), building on a disciplined operator splitting approach. Notably, we provide an inexact generalization allowing for random error-afflicted lower-level solutions, addressing a significant shortcoming in [15]. Specifically, in this paper we improve [15] along two important dimensions: (i) First, we allow for inexact resolution of lower-level problems to accommodate large-scale stochastic follower problems; (ii) Second, we provide a novel variance-reduction framework for addressing this problem. Despite the need for inexactness, our statements match the state-of-the-art both in terms of rate and oracle complexity. Algorithmically, these advancements are achieved via a novel stochastic operator splitting approach that combines ideas from iterative smoothing and regularization [72], with modern variance reduction approaches originating in machine learning [31].

1.1.2 Zeroth-Order Optimization, Smoothing, and Regularization

Zeroth-order (gradient-free) optimization is being increasingly embraced for solving machine learning problems where explicit expressions of the gradients are difficult or infeasible to obtain. In hierarchical optimization problems, this is particularly relevant when solutions of the lower level problem are injected into the leader's upper-level problem. In machine learning, this problem is known as approximating the hypergradient. Various techniques for estimating this object have been studied recently, ranging from truncated von Neuman series [29] and fully first-order methods [46]. Instead of computationally expensive first-order (or higher) information about the problem data, we develop an online stochastic approximation approach based solely on function evaluations in order to approximate the directional derivative of the coupling function between the upper and the lower level. Approximating the directional derivative of the thus obtained implicit function has a long history [65] and has been employed for resolving stochastic optimization [47, 69, 70] and variational inequality problems [71, 72].

1.1.3 Variance Reduction

Variance reduction is a commonly employed method exploiting the finite-sum structure of variational problems arising in machine learning and engineering. The classical stochastic variance reduced gradient (SVRG) [39] is embedded within a double loop structure and tailored to the prototypical finite-sum structure in empirical risk minimization. Indeed, in the classical SVRG formulation full gradients are computed "from time-to-time" in the outer loop while cheap variance reduced gradients are used in the frequently activated inner loop subroutine. This construction has been extended to saddle-point problems and stochastic monotone inclusions in [58]. Extensions to monotone *mixed*-variational inequality problems were recently provided in [1, 2, 12]. In contrast, we consider stochastic hierarchical games over general sample spaces, complicated by the presence of nested optimization problems embodied by the interaction between leaders and followers. While the assumptions we make in this work allow us to recast the problem as a mixed variational inequality, several challenges persist. First, subgradients of $g_i(\cdot, \mathbf{y}_i(\cdot))$ are unavailable; Second, enlisting smoothing approaches requires $\mathbf{y}_i(\mathbf{x}_i)$, unavailable in closed form; Third, we are not restricted to finite-sum regimes and allow for general sample spaces by employing increasingly large batch-sizes to approximate the gradient in the outer loop.

2 Preliminaries

In this section, we articulate the standing assumptions employed in this paper and introduce our notation. The decision set of leader i is a subset $\mathcal{X}_i$ in $\mathbb{R}^{n_i}$. We let $\mathcal{X} \triangleq \prod_{i\in\mathcal{I}} \mathcal{X}_i$ represent the set of strategy profiles of the leaders and identify it with a subset of $\mathbb{R}^n$, where $n \triangleq \sum_{i\in\mathcal{I}} n_i$. For any $d \geq 1$, we let $\mathbb{B}_d \triangleq \{\mathbf{x} \in \mathbb{R}^d \mid \|\mathbf{x}\| \leq 1\}$ denote the unit ball in $\mathbb{R}^d$. We start by introducing a basic assumption on the follower's problem. It bears reminding that the VI representation of the follower problem allows for capturing a range of problems, ranging from smooth convex optimization problems to more intricate smooth convex games and equilibrium problems; see [24].[1]

Assumption 1 For $i \in \mathcal{I}$, $\mathcal{Y}_i \subseteq \mathbb{R}^{m_i}$ is closed and convex set, and for all $\mathbf{x}_i \in \mathcal{X}_i$, the mapping $\phi_i(\mathbf{x}_i, \cdot) : \mathcal{Y}_i \to \mathbb{R}^{m_i}$ is strongly monotone and Lipschitz continuous.

By Assumption 1, for any $i \in \mathcal{I}$, the set of solutions to $\mathrm{VI}(\phi_i(\mathbf{x}_i, \cdot), \mathcal{Y}_i)$, denoted by $\mathrm{SOL}(\phi_i(\mathbf{x}_i, \cdot), \mathcal{Y}_i)$, is single-valued with unique element $\mathbf{y}_i(\mathbf{x}_i)$. Moreover, $\phi_i(\mathbf{x}_i, \mathbf{y}_i) = \mathbb{E}_\xi \left[\Phi_i(\mathbf{x}_i, \mathbf{y}_i, \xi)\right]$ for any $(\mathbf{x}_i, \mathbf{y}_i) \in \mathcal{X}_i \times \mathcal{Y}_i$.

Assumption 2 The following assumptions hold for each leader $i \in \mathcal{I}$:

(i) The set $\mathcal{X}_i \subset \mathbb{R}^{n_i}$ is nonempty, compact, and convex. In particular, there exists $C_i > 0$ such that $\sup_{\mathbf{x}_i, \mathbf{x}_i' \in \mathcal{X}_i} \left\|\mathbf{x}_i - \mathbf{x}_i'\right\| \leq C_i$ for all $i \in \mathcal{I}$.

[1] Appendix A.2 explains the terminology related to VIs.

(ii) For some $\delta_0 > 0$, the mapping $\mathbf{x}_i \mapsto g_i(\mathbf{x}_i, \mathbf{y}_i(\mathbf{x}_i))$ is $L_{1,i}$-Lipschitz on $\mathcal{X}_{i,\delta_0} \triangleq \mathcal{X}_i + \delta_0\mathbb{B}_{n_i}$.
(iii) $\mathbf{x}_i \mapsto F_i((\mathbf{x}_i, \mathbf{x}_{-i}), \xi)$ is convex and continuously differentiable over an open set containing $\mathcal{X}_i$, uniformly for all $\mathbf{x}_{-i} \in \mathcal{X}_{-i}$ and almost every $\xi \in \Xi$.
(iv) The mapping $\mathbf{y}_i \mapsto g_i(\mathbf{x}_i, \mathbf{y}_i)$ is $L_{2,i}$-Lipschitz continuous for all $\mathbf{x}_i \in \mathcal{X}_{i,\delta_0}$.
(v) The operator $V : \mathbb{R}^n \to \mathbb{R}^n$, defined by $V(\mathbf{x}) = (V_i(\mathbf{x}))_{i\in\mathcal{I}}$ and $V_i(\mathbf{x}) \triangleq \nabla_{\mathbf{x}_i} f_i(\mathbf{x})$, is L_f-Lipschitz continuous and monotone on $\mathcal{X}_{\delta_0} = \prod_{i\in\mathcal{I}} \mathcal{X}_{i,\delta_0}$.

Assumption 1 is commonly employed in hierarchical optimization problems. Indeed, in the special case where the VI captures the optimality conditions of a parametrized convex optimization problem solved by the follower, then strong monotonicity of $\phi_i(\mathbf{x}_i, \cdot)$, is equivalent to strong convexity of the follower's cost function, an assumption that is known in bilevel optimization literature as the *lower level uniqueness property* [50].

Given Assumption 1 we define the implicit loss function $L_i : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ by

$$L_i(\mathbf{x}_i, \mathbf{x}_{-i}) \triangleq \ell_i(\mathbf{x}_i, \mathbf{x}_{-i}, \mathbf{y}_i(\mathbf{x}_i)) = \mathbb{E}_\xi[F_i((\mathbf{x}_i, \mathbf{x}_{-i}), \xi)] + h_i(\mathbf{x}_i), \tag{2}$$

where $h_i(\mathbf{x}_i) \triangleq g_i(\mathbf{x}_i, \mathbf{y}_i(\mathbf{x}_i))$. In terms of the implicit loss function (2), we convert the hierarchical game (P) into a *stochastic Nash equilibrium problem* in which each player solves the loss minimization problem

$$(\forall i \in \mathcal{I}) : \quad \min_{\mathbf{x}_i \in \mathcal{X}_i} L_i(\mathbf{x}_i, \mathbf{x}_{-i}). \tag{3}$$

We refer to (3) as the *upper-level problem*, and summarize it as the tuple $\mathcal{G}^{\text{upper}} \triangleq \{L_i, \mathcal{X}_i\}_{i\in\mathcal{I}}$. Let $\mathbf{NE}(\mathcal{G}^{\text{upper}})$ denote the set of Nash equilibria of the game $\mathcal{G}^{\text{upper}}$ and $\mathcal{Y} = \prod_{i\in\mathcal{I}} \mathcal{Y}_i$.

Definition 2.1 A $2N$-tuple $(\mathbf{x}^*, \mathbf{y}^*) \in \mathcal{X}\times\mathcal{Y}$ is called a *hierarchical equilibrium* if $\mathbf{x}^* \in \mathbf{NE}(\mathcal{G}^{\text{upper}})$ and, for all $i \in \mathcal{I}$, $\mathbf{y}_i^* = \mathbf{y}_i(\mathbf{x}_i^*)$ the unique solution of $\mathrm{VI}(\phi_i(\mathbf{x}_i^*, \cdot), \mathcal{Y}_i)$.

Typical online learning approaches in game theory employ stochastic approximation (SA) for iteratively approaching a Nash equilibrium of the game $\mathcal{G}^{\text{upper}}$. These iterative methods rely on the availability of a stochastic oracle revealing (noisy) first-order information about the operators involved (i.e. samples of pseudo-gradients of the objective for each individual player). Such direct methods are complicated in hierarchical optimization since the required subgradient is an element of the subdifferential of the sum of two Lipschitz continuous functions. In the current setting, this task is even more complicated since the upper level objective is defined by a function which is available only in an implicit form, as it depends on the solution of the lower level problem $\mathbf{y}_i(\mathbf{x}_i)$, and another function given in terms of an expected value. Thus, even if a sum-rule for a subdifferential applies [13], it would read as $\partial_{\mathbf{x}_i} L(\mathbf{x}_i, \mathbf{x}_{-i}) = \partial_{\mathbf{x}_i} f_i(\mathbf{x}) + \partial_{\mathbf{x}_i}(g_i \circ (\mathrm{Id}, \mathbf{y}_i(\cdot)))(\mathbf{x}_i)$. Hence we would need to invoke a non-smooth chain rule for our chosen version of a subdifferential, in order to evaluate $\partial_{\mathbf{x}_i}(g_i \circ (\mathrm{Id}, \mathbf{y}_i(\cdot)))(\mathbf{x}_i)$. In addition, we would require access to the subdifferential of $\mathbf{y}_i(\bullet)$ at $\mathbf{x}_i$. We circumvent this computationally challenging step by developing a random search procedure based on a finite difference approximation. To develop such scheme, recall that we defined $h_i(\mathbf{x}_i) \triangleq g_i(\mathbf{x}_i, \mathbf{y}_i(\mathbf{x}_i))$ as the loss function coupling the leader and the follower. The following fact can be found in Proposition 1 in [16].

Lemma 2.1 *Let Assumptions 1–2 hold. Then $h_i(\mathbf{x}_i) = g_i(\mathbf{x}_i, \mathbf{y}_i(\mathbf{x}_i))$ is L_{h_i}-Lipschitz continuous on $\mathcal{X}_i$ and directionally differentiable.*

To proceed, we impose a convexity requirement on $h_i(\cdot, \mathbf{y}_i(\cdot))$.

Assumption 3 The implicitly defined function $\mathbf{x}_i \mapsto h_i(\mathbf{x}_i) \triangleq g_i(\mathbf{x}_i, \mathbf{y}_i(\mathbf{x}_i))$ is convex on $\mathcal{X}_i$.

Remark 2.1 Several papers in the literature provided conditions under which the implicit function h_i is indeed convex in hierarchical settings. A structural model framework where convexity provably holds is described in [17, 35].

Remark 2.2 Under Assumptions 2 and 3, an equilibrium exists using classical results. Indeed, since the lower level problem is assumed to have a unique solution $\mathbf{y}_i(\mathbf{x}_i)$ and Lemma 2.1 guarantees that the noncooperative game $\mathcal{G}^{\text{upper}}$ satisfies conditions of [52, Th. I.4.1], so that $\mathbf{NE}(\mathcal{G}^{\text{upper}}) \neq \varnothing$.

Assumptions 2(v), 3 and Lemma 2.1, yield a variational characterization of elements of $\mathbf{NE}(\mathcal{G}^{\text{upper}})$ in terms of an expectation-valued *mixed-variational inequality* (cf. Appendix A.2).

Lemma 2.2 *Let $h : \mathcal{X} \to \mathbb{R}$ be defined by $h(\mathbf{x}) \triangleq \sum_{i\in\mathcal{I}} h_i(\mathbf{x}_i)$. Then, $\mathbf{x}^* \in \mathbf{NE}(\mathcal{G}^{upper})$ if and only if $\mathbf{x}^*$ solves the mixed variational inequality* MVI(V, h)*:*

$$\langle V(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + h(\mathbf{x}) - h(\mathbf{x}^*) \geq 0 \qquad \forall \mathbf{x} \in \mathcal{X}. \tag{4}$$

Solution methods to mixed VIs with expectation-valued operators have been developed recently in cases where the random variable takes values in a finite set, and/or when the VI is derived from a zero-sum game displaying a finite-sum structure [1, 2]. The standard algorithmic approach to iteratively approximate a solution to such structured VIs are extragradient type of methods. A direct application of these methods to the mixed VI (4) is complicated because of the following facts:

(i) L-smoothness of h_i. The assumptions made thus far do not guarantee the differentiability of h_i with a Lipschitz continuous gradient. Hence, a direct application of gradient, or extragradient methods, is a difficult task in our setting. To cope with this technical difficulty, we develop a smoothing approach, yielding a family of approximating models enjoying the typical Lipschitz smoothness requirements.
(ii) Randomness in the operator V: Since the operator V is only available in terms of an expected value, in general, we cannot tractably evaluate it.

Instead we have to use simulation-based methods to obtain random estimators of this mathematical expectation. To keep this simulation task within a feasible computational budget, variance reduction methods are used in iterative methods for generating the input data. However, standard variance reduction techniques rely on smoothness of the data. In our case, non-smoothness is present in terms of the implicit function $h_i(\cdot)$. In principle, one could apply splitting techniques to deal with the non-smooth function via a proximal smoothing. However, this approach requires $h_i(\cdot)$ to be prox-friendly for which we have no a-priori guarantee since it is the value function of the leader, derived from the solution of the follower. With these preparatory remarks in mind, we now explain the design of our algorithmic solution strategy for the hierarchical game problems (P).

3 A Variance Reduced Forward–Backward–Forward Algorithm for Hierarchical Games

In this section, we present our algorithm for computing an equilibrium of the hierarchical game (P). As in the seminal SVRG formulation, our method runs in two loops. Each loop requires as inputs data that are computed in the outer loop. The inputs of the inner and outer loops are constructed as follows. For $\eta > 0$, we define the (Tikhonov) regularized vector field $V^{\eta} : \mathcal{X} \to \mathbb{R}^n$ by

$$V^{\eta}(\mathbf{x}) = (V_i^{\eta}(\mathbf{x}))_{i\in\mathcal{I}}, \text{ where } V_i^{\eta}(\mathbf{x}) \triangleq V_i(\mathbf{x}) + \eta\mathbf{x}_i \quad \forall i \in \mathcal{I}. \tag{5}$$

Tikhonov regularization is a classical tool to obtain stronger convergence results in numerical schemes. It has been examined for deterministic [40] and stochastic equilibrium problems [43, 72].

Our next assumption is concerned with the nature of the stochastic oracle which generated random estimators on the expectation-valued operator V when queried at a given point $\mathbf{x}$.

Assumption 4 The operator V has a stochastic oracle $\hat{V}(\cdot, \xi)$ that is

1. unbiased: $V(\mathbf{x}) = \mathbb{E}_{\xi}[\hat{V}(\mathbf{x}, \xi)]$ for all $\mathbf{x} \in \mathcal{X}$;
2. $\mathcal{L}_f(\xi)$-Lipschitz for almost every $\xi \in \Xi$: $\left\|\hat{V}(\mathbf{x}', \xi) - \hat{V}(\mathbf{x}, \xi)\right\| \leq \mathcal{L}_f(\xi)\left\|\mathbf{x}' - \mathbf{x}\right\|$ for all $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$. The random variable $\mathcal{L}_f(\xi)$ is positive and integrable with $\mathbb{E}_{\xi}[\mathcal{L}(\xi)] = L_f$.

To retrieve in-play information about the value of the implicit loss function $h_i(\cdot)$, we employ a smoothing-based approach, which necessitates defining another sampling mechanism. We follow the gradient sampling strategy of [26], though alternative random estimation strategies are certainly possible (see, e.g. [8, 23, 44]). Specifically, given $\delta > 0$, we denote the finite difference approximation of the directional derivative of h_i in direction $\mathbf{w}_i \in \mathbb{R}^{n_i}$ as

$$\nabla_{(\mathbf{w}_i,\delta)} h_i(\mathbf{x}_i) \triangleq \frac{h_i(\mathbf{x}_i + \delta\mathbf{w}_i) - h_i(\mathbf{x})}{\delta}.$$

Let $\mathbf{W}_i$ be a random vector uniformly distributed on the unit sphere $\mathbb{S}_i \triangleq \{\mathbf{x}_i \in \mathbb{R}^{n_i} : \|\mathbf{x}_i\| = 1\}$.[2] We then define the random vector $H^{\delta}_{i,\mathbf{x}_i}(\mathbf{W}_i)$ as a randomized and suitably

[2] See Appendix A.5 for the explicit construction of such an oracle.

rescaled version of the finite difference approximator reading as

$$H^{\delta}_{i,\mathbf{x}_i}(\mathbf{W}_i) \triangleq n_i \mathbf{W}_i \nabla_{(\mathbf{W}_i,\delta)} h_i(\mathbf{x}_i) \in \mathbb{R}^{n_i}. \tag{6}$$

From Eq. (40) in Appendix A.4, we know that $H^{\delta_i}_{i,\mathbf{x}_i}(\cdot)$ is an unbiased estimator of the gradient of the smoothed function

$$h^{\delta}_i(\mathbf{x}_i) \triangleq \frac{1}{\mathbf{Vol}_n(\delta\mathbb{B}_{n_i})} \int_{\delta\mathbb{B}_{n_i}} h_i(\mathbf{x}_i + \mathbf{u})\mathrm{d}\mathbf{u},$$

where $\mathbb{B}_d \triangleq \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| \leq 1\}$ for any dimension $d \geq 1$. Furthermore, we discuss in Appendix A.4 that the function h^{δ}_i is continuously differentiable with gradient

$$\nabla h^{\delta}_i(\mathbf{x}_i) = \frac{n_i}{\delta}\mathbb{E}_{\mathbf{W}_i\sim\mathsf{U}(\mathbb{S}_{n_i})}[\mathbf{W}_i\,(h_i(\mathbf{x}_i + \delta\mathbf{W}_i) - h_i(\mathbf{x}_i))] = \mathbb{E}_{\mathbf{W}_i\sim\mathsf{U}(\mathbb{S}_{n_i})}[H^{\delta}_{i,\mathbf{x}_i}(\mathbf{W}_i)],$$

and

$$\left\|\nabla h^{\delta}_i(\mathbf{x}_i) - \nabla h^{\delta}_i(\mathbf{y}_i)\right\| \leq \frac{L_{h_i} n_i}{\delta}\|\mathbf{x}_i - \mathbf{y}_i\|, \qquad \forall \mathbf{x}_i, \mathbf{y}_i \in \mathbb{R}^{n_i}, \delta > 0,$$

where $\mathbf{W}_i \sim \mathsf{U}(\mathbb{S}_{n_i})$ indicates that $\mathbf{W}_i$ is uniformly distributed on the surface of a unit sphere $\mathbb{S}_{n_i}$.

3.1 Iterative Regularization Methods

Lemma 2.2 shows that the equilibria of our hierarchical game are entirely captured by the solution set of problem $\mathrm{MVI}(V, h)$. This is a rich class of variational problems for which the number of contributions is so numerous that we just point the reader to [24]. Deducing convergence results on the *last iterate* for standard algorithmic schemes is an important requirement for game-theoretic learning algorithms, but typically is a rare commodity: Despite some special classes of games [4, 30], first-order methods give only guarantees on a suitably constructed ergodic average. To obtain last iterate convergence results, we develop an iterative Tikhonov regularization approach. This leads us to consider the regularized problem $\mathrm{MVI}(V^{\eta}, h)$, which requires to find $\mathbf{s}(\eta) \in \mathcal{X}$ satisfying

$$\langle V^{\eta}(\mathbf{s}(\eta)), \mathbf{x} - \mathbf{s}(\eta)\rangle + h(\mathbf{x}) - h(\mathbf{s}(\eta)) \geq 0 \qquad \forall \mathbf{x} \in \mathcal{X}. \tag{7}$$

Naturally, we would like to understand the nature of the accumulation points of the sequence $\{\mathbf{s}_t\}_{t\in\mathbb{N}}$, where $\mathbf{s}_t \equiv \mathbf{s}(\eta_t)$ and $\eta_t \downarrow 0$. This sequence can be studied in quite some detail, and we summarize some well-known facts in Proposition 3.1 below. As those results are rather scattered in the literature, we provide a self-contained proof in Appendix A.3.

Proposition 3.1 *Let Assumptions 1–3 hold true. Consider the problem* $\mathrm{MVI}(V, h)$ *with nonempty solution set* $\mathrm{SOL}(V, h)$. *Then the following apply:*

(a) *For all* $\eta > 0$*, the set* $\mathrm{SOL}(V^{\eta}, h)$ *is a singleton with unique element denoted by* $\mathbf{s}(\eta)$*;*
(b) $(\forall \eta > 0) : \|\mathbf{s}(\eta)\| \leq \inf\{\|\mathbf{x}\| : \mathbf{x} \in \mathrm{SOL}(V, h)\}$*;*
(c) *Let* $\{\eta_t\}_{t\in\mathbb{N}}$ *be a positive sequence with* $\eta_t \downarrow 0$*. Then, the sequence* $\{\mathbf{s}(\eta_t)\}_{t\in\mathbb{N}}$ *converges to the least norm solution* $\arg\min\{\|\mathbf{x}\| : \mathbf{x} \in \mathrm{SOL}(V, h)\}$*;*
(d) *For any positive sequence* $\{\eta_t\}_{t\in\mathbb{N}}$ *satisfying* $\eta_t \downarrow 0$*, we have*

$$\left(\frac{\eta_t - \eta_{t-1}}{\eta_t}\right) \min_{\mathbf{x}\in\mathrm{SOL}(V,h)} \|\mathbf{x}\| \geq \|\mathbf{s}(\eta_t) - \mathbf{s}(\eta_{t-1})\| . \tag{8}$$

3.2 The Algorithm

Our algorithm for the hierarchical game setting consists of a double-loop structure: The outer loop allows the N players to make multiple independent queries of the stochastic oracle $\hat{V}_1(\cdot, \xi), \ldots, \hat{V}_N(\cdot, \xi)$, and draw multiple independent samples from the surface of a unit sphere, allowing for the simulation of the random estimator (6). However, since the multiple calls are a negative entry on the oracle complexity of the method, we impose some control on the number of mini-batches to be constructed by the agents. Within the inner-loop subroutine, the agents only receive single samples from their stochastic oracles, and employ this new information in an extragradient-type algorithm. We give a precise construction in the following paragraphs.

3.2.1 The Outer Loop

Let $t = 0, \ldots, T-1$ be the iteration counter for the outer loop. We denote by $b_t \in \mathbb{N}$ the pre-defined sample rate defining the number of random variables each player is allowed to generate in round t. Specifically, each player generates an iid sample $\xi_{i,t}^{1:b_t} \triangleq \{\xi_{i,t}^{(s)}; 1 \leq s \leq b_t\}$ and constructs the mini-batch estimator

$$\bar{V}_i^t \triangleq \frac{1}{b_t}\sum_{s=1}^{b_t} \hat{V}_i(\mathbf{x}^t, \xi_{i,t}^{(s)}), \tag{9}$$

Let $\bar{V}^t = (\bar{V}_1^t, \ldots, \bar{V}_N^t)$. A similar assumption is made to obtain point-estimators of the gradient of the implicit function $h_i(\cdot)$. Hence, $\mathbf{W}_{i,t}^{1:b_t} \triangleq \{\mathbf{W}_{i,t}^{(s)}; 1 \leq s \leq b_t\}$ denotes an i.i.d. sample of b_t random vectors drawn uniformly at random from $\mathbb{S}_i$ and define the mini-batch estimator $H_{i,\mathbf{x}_i}^{\delta_t,b_t}$ as

$$H_{i,\mathbf{x}_i}^{\delta_t,b_t} \triangleq \frac{1}{b_t}\sum_{s=1}^{b_t} H_{i,\mathbf{x}_i}^{\delta_t}(\mathbf{W}_{i,t}^{(s)}), \quad H_{\mathbf{x}^t}^{\delta_t,b_t} \triangleq \left(H_{1,\mathbf{x}_1^t}^{\delta_t,b_t}, \ldots, H_{N,\mathbf{x}_N^t}^{\delta_t,b_t}\right), \tag{10}$$

where δ_t denotes a positive smoothing parameter. Equipped with these estimators, each player enters the procedure $\texttt{SFBF}(\mathbf{x}^t, \bar{V}^t, H_{\mathbf{x}^t}^{\delta_t,b_t}, \gamma_t, \eta_t, \delta_t, K)$, that relies on steplength γ_t and regularization parameter η_t, whose role is explained in the description of the inner loop.

3.2.2 The Inner Loop

Given the inputs $(\mathbf{x}^t, \bar{V}^t, H^{\delta_t,b_t}_{\mathbf{x}_t}, \gamma_t, \eta_t)$ prepared in the outer loop, the inner loop of our method is based on a stochastic version of Tseng's modified extragradient method [66], using one-shot estimators of the relevant data. To be precise, given the current iterate $\mathbf{x}^t$, each player i produces a trajectory $\{\mathbf{z}^{(t)}_{i,k}\}_{k\in\{0,1/2,1,\dots,K\}}$. These interim strategy profiles are updated recursively by the procedure $\texttt{SFBF}(\mathbf{x}^t, \bar{V}^t, H^{\delta_t,b_t}_{\mathbf{x}^t}, \gamma_t, \eta_t, \delta_t, K)$ described in Algorithm 1. Starting with the strategy profile $\mathbf{x}^t$, we choose the initial conditions $\mathbf{z}^{(t)}_{i,0} = \mathbf{x}^t_i$ for all $i \in \mathcal{I}$. Then, for each $k \in \{0, 1, \dots, K-1\}$ each player queries the stochastic oracle to obtain the feedback signal

$$\hat{V}^{\eta_t}_{i,t,k+1/2}(\mathbf{z}^{(t)}_{k+1/2}) \triangleq \hat{V}_i(\mathbf{z}^{(t)}_{k+1/2}, \xi_{i,t,k+1/2}) + \eta_t \mathbf{z}^{(t)}_{i,k+1/2}. \tag{11}$$

Similarly, each player obtains the random information $H^{\delta_t}_{\mathbf{z}^{(t)}_{i,k+1/2}}(\mathbf{W}_{i,t,k+1/2})$ and $H^{\delta_t}_{\mathbf{x}^t_i}(\mathbf{W}_{i,t,k+1/2})$, as defined in (6). These random variables are used to generate the parallel updates

$$\begin{aligned}
\mathbf{z}^{(t)}_{i,k+1/2} &= \Pi_{\mathcal{X}_i}[\mathbf{z}^{(t)}_{i,k} - \gamma_t(\bar{V}^t_i + \eta_t \mathbf{x}^t_i + H^{\delta_t,b_t}_{\mathbf{x}^t_i})], \text{ and} \\
\mathbf{z}^{(t)}_{i,k+1} &= \mathbf{z}^{(t)}_{i,k+1/2} - \gamma_t \Big(\hat{V}^{\eta_t}_{i,t,k+1/2}(\mathbf{z}^{(t)}_{k+1/2}) \\
&\quad + H^{\delta_t}_{\mathbf{z}^{(t)}_{i,k+1/2}}(\mathbf{W}_{i,t,k+1/2}) - \hat{V}^{\eta_t}_{i,t,k+1/2}(\mathbf{x}^t) - H^{\delta_t}_{\mathbf{x}^t_i}(\mathbf{W}_{i,t,k+1/2})\Big)
\end{aligned}$$

for all $i \in \mathcal{I}$. These iterations correspond to a stochastic approximation variant of Tseng's forward-backward-forward method [10] for solving the time-varying stochastic variational inequality

$$0 \in V(\bar{\mathbf{x}}) + \nabla h^{\delta_t}(\bar{\mathbf{x}}) + \eta_t \bar{\mathbf{x}} + \mathsf{NC}_{\mathcal{X}}(\bar{\mathbf{x}}),$$

with

$$\mathsf{NC}_{\mathcal{X}}(\mathbf{x}) = \begin{cases} \varnothing & \text{if } \mathbf{x} \notin \mathcal{X}, \\ \{p \mid \sup_{\mathbf{z}\in\mathcal{X}} \langle p, \mathbf{z} - \mathbf{x}\rangle \leq 0\} & \text{if } \mathbf{x} \in \mathcal{X}, \end{cases}$$

the normal cone of $\mathcal{X}$ at $\mathbf{x}$.

Discussion The key innovation of the scheme VRHGS lies in the combination of smoothing (to allow for hierarchy), regularization (to contend with ill-posedness), and variance-reduction (to mitigate bias) within a stochastic forward-backward-forward framework. Our double-loop solution strategy mimics the computational architecture of SVRG, which takes a full gradient sample of the finite sum problem "once in a while", while performing frequent single-sample updates in between. Our method, adapted to general probability spaces, proceeds similarly: The "shadow sequence" $\mathbf{z}^{(t)}_{i,k+1/2}$ uses costly mini-batch estimators computed in the outer loop; these are maintained in memory while executing the inner loop (i.e. only "once in a while" updated). The additional forward steps to obtain the iterates $\mathbf{z}^{(t)}_{i,k+1}$ make use of fresh one-shot

Algorithm 1: $\mathtt{SFBF}(\bar{\mathbf{x}}, \bar{\mathbf{v}}, \bar{H}, \gamma, \eta, \delta, K)$

Result: Iterate $\mathbf{z}_K$
Set $\mathbf{z}_0 = \bar{\mathbf{x}}$;
for $k = 0, 1, \ldots, K-1$ **do**
Update $\mathbf{z}_{k+1/2} = \Pi_{\mathcal{X}}[\mathbf{z}_k - \gamma(\bar{\mathbf{v}} + \eta\bar{\mathbf{x}} + \bar{H})]$ with $\Pi_{\mathcal{X}}$ the orthogonal projector onto $\mathcal{X}$ (cf. Appendix A.1);
Obtain $\hat{V}^{\eta}_{k+1/2}(\mathbf{z}_{k+1/2})$ and $\hat{V}^{\eta}_{k+1/2}(\bar{\mathbf{x}})$ as defined in eq. (11);
Draw iid direction vectors $\mathbf{W}_{k+1/2} = \{\mathbf{W}_{i,k+1/2}\}_{i\in\mathcal{I}}$, with each $\mathbf{W}_{i,k+1/2} \sim \mathsf{U}(\mathbb{S}_i)$. ;
Obtain $H^{\delta}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2})$ and $H^{\delta}_{\bar{\mathbf{x}}}(\mathbf{W}_{k+1/2})$;
Update

$$\mathbf{z}_{k+1} = \mathbf{z}_{k+1/2} - \gamma\left(\hat{V}^{\eta}_{k+1/2}(\mathbf{z}_{k+1/2}) + H^{\delta}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - \hat{V}^{\eta}_{k+1/2}(\bar{\mathbf{x}}) - H^{\delta}_{\bar{\mathbf{x}}}(\mathbf{W}_{k+1/2})\right).$$

end

Algorithm 2: Variance Reduced Hierarchical Game Solver ($\mathtt{VRHGS}$)

Data: $\mathbf{x}$, T, $\{\gamma_t\}_{t=0}^{T}$, $\{b_t\}_{t=0}^{T}$, $\{\eta_t\}_{t=0}^{T}$, $\{\delta_t\}_{t=0}^{T}$, K
Set $\mathbf{x}^0 = \mathbf{x}$.
for $t = 0, 1, \ldots, T-1$ **do**
For each $i \in \mathcal{I}$ receive the oracle feedback $\bar{V}^t$ defined by $\bar{V}^t_i \triangleq \frac{1}{b_t}\sum_{s=1}^{b_t} \hat{V}_i(\mathbf{x}^t, \xi^{(s)}_{i,t})$. ;
For each $i \in \mathcal{I}$ construct the estimator $H^{\delta_t, b_t}_{\mathbf{x}^t}$ defined by $H^{\delta_t, b_t}_{i,\mathbf{x}_i} \triangleq \frac{1}{b_t}\sum_{s=1}^{b_t} H^{\delta_t}_{i,\mathbf{x}_i}(\mathbf{W}^{(s)}_{i,t})$. ;
Update $\mathbf{x}^{t+1} = \mathtt{SFBF}(\mathbf{x}^t, \bar{V}^t, H^{\delta_t, b_t}_{\mathbf{x}^t}, \gamma_t, \eta_t, \delta_t, K)$.
end

estimators of the payoff gradient and the finite difference estimator. All steps are overlaid by a Tikhonov regularization, while smoothing facilitates accommodation with hierarchical objectives. From a computational perspective, our scheme performs a single projection onto the leaders' feasible set $\mathcal{X}_i$. This can save considerably on computational time in cases where the projection operator is costly to evaluate, and constitutes a major difference compared to viable alternative algorithmic schemes like the extragradient or optimistic mirror descent. Hence, our method reduces the sample-complexity of recent mini-batch variance reduction techniques for stochastic VIs [10], while concomitantly reducing the computational bottlenecks of double-call algorithms [37, 41] by lifting one projection step. Finally, similar rate and complexity statements emerge when allowing for inexact generalizations that allow for ϵ-approximate solutions of lower-level problem.

4 Main Results

In this section we state the main results on the asymptotic convergence of scheme $\mathtt{VRHGS}$. All technical and lengthy proofs are collected in Sect. 7.

In the inner and outer loops of $\mathtt{VRHGS}$, we have two sources of randomness at each iteration: (i) the sequence of mini-batches $\xi^{1:b_t}_t \triangleq \{\xi^{1:b_t}_{i,t}\}_{i\in\mathcal{I}}$ and $\mathbf{W}^{1:b_t}_t \triangleq \{\mathbf{W}^{1:b_t}_{i,t}\}_{i\in\mathcal{I}}$, which are used to perform the opening forward-backward step in Algorithm 1;

(ii) the sequences $\xi_{t,k+1/2} = \{\xi_{i,t,k+1/2}\}_{i\in\mathcal{I}}$ and $\mathbf{W}_{t,k+1/2} = \{\mathbf{W}_{i,t,k+1/2}\}_{i\in\mathcal{I}}$ for $k \in \{0, 1, \ldots, K-1\}$, which are employed in constructing the iterate $\mathbf{z}_{k+1}^{(t)}$ in Algorithm 1 during the outer epoch t. To keep track of the information structure of the outer and inner loops, we introduce the filtrations $\mathcal{F}_t \triangleq \sigma(\mathbf{x}^0, \ldots, \mathbf{x}^t)$ for $0 \le t \le T$, and $\mathcal{A}_{t,0} \triangleq \sigma(\mathbf{x}^t, \xi_t^{1:b_t}, \mathbf{W}_t^{1:b_t})$ as well as $\mathcal{A}_{t,k} \triangleq \sigma(\mathcal{A}_{t,0} \cup \sigma(\xi_{t,1/2}, \mathbf{W}_{t,1/2}, \ldots, \xi_{t,k-1/2}, \mathbf{W}_{t,k-1/2}))$ for $k \in \{1, 2, \ldots, K-1\}$. By construction, the iterates $\mathbf{z}_k^{(t)}$ and $\mathbf{z}_{k+1/2}^{(t)}$ are both $\mathcal{A}_{t,k}$-measurable.

4.1 Error Structure of the Estimators

We impose a uniform variance bound on the random vector field $\hat{V}$ over the set $\mathcal{X}$. Compactness of $\mathcal{X}$ implies that such an assumption comes without loss of generality, and the proof of the variance bound in Lemma 4.1 is simple to obtain and thus omitted; See [10].

Lemma 4.1 *There exists $M_V > 0$ such that $\mathbb{E}_\xi[\left\|\hat{V}(\mathbf{x}, \xi) - V(\mathbf{x})\right\|^2] \le M_V^2$ for all $\mathbf{x} \in \mathcal{X}$. Additionally, let $b_t \ge 1$ and $\xi_t^{1:b_t} = \{\{\xi_{i,t}^{(1)}\}_{i\in\mathcal{I}}, \ldots, \{\xi_{i,t}^{(b_t)}\}_{i\in\mathcal{I}}\}$ denote an i.i.d sample of the random variable ξ. Then, for $\varepsilon_V(\xi_t^{1:b_t}) \triangleq \frac{1}{b_t}\sum_{s=1}^{b_t} \hat{V}(\mathbf{x}, \xi_t^{(s)}) - V(\mathbf{x})$, we have $\sqrt{\mathbb{E}\left[\left\|\varepsilon_V(\xi_t^{1:b_t})\right\|^2\right]} \le \frac{M_V}{\sqrt{b_t}}$.*

Concerning the estimator of the gradient of the smoothed lower level function $h^\delta(\cdot)$, we can report the following bounds, which are derived in Appendix A.5 and proved in Lemma A.2.

Lemma 4.2 *Let Assumptions 1 and 2 hold. Define $\mathbb{e}_{\mathbf{x}_i}(\mathbf{W}_i^{1:b}) \triangleq H_{i,\mathbf{x}_i}^{\delta,b} - \nabla h_i^\delta(\mathbf{x}_i)$, where $\{\mathbf{W}_i^{(s)}\}_{s=1}^{b_t}$ is an i.i.d sample drawn uniformly from the unit sphere $\mathbb{S}_i$, i.e. $\mathbf{W}_i^{1:b} \sim \mathsf{U}(\mathbb{S}_i)^{\otimes b}$. Then for all $i \in \mathcal{I}$,*

(a) $\mathbb{E}_{\mathbf{W}_i^{1:b}\sim\mathsf{U}(\mathbb{S}_i)^{\otimes b}}[\mathbb{e}_{\mathbf{x}_i}(\mathbf{W}_i^{1:b})] = 0$;

(b) $\left\|H_{i,\mathbf{x}_i}^{\delta}(\mathbf{w}_i)\right\|^2 \le L_{h_i}^2 n_i^2$ *for all* $\mathbf{w}_i \in \mathbb{S}_i$;

(c) $\mathbb{E}_{\mathbf{W}_i^{1:b}\sim\mathsf{U}(\mathbb{S}_i)^{\otimes b}}[\left\|\mathbb{e}_{\mathbf{x}_i}(\mathbf{W}_i^{1:b})\right\|^2] \le \frac{n_i^2 L_{h_i}^2}{b_i}$.

4.2 Almost Sure Convergence of the Last Iterate

Our analysis of `VRHGS` relies on the following energy inequality, proved in Sect. 7.1.1.

Lemma 4.3 *Let Assumptions 1–4 hold true. Let $\{\mathbf{x}^t\}_{t=0}^{T-1}$ be generated by* `VRHGS` *and denote by $\{\mathbf{z}_k^{(t)}\}_{k\in\{0,1/2,\ldots,K\}}$ the sequence obtained by executing* `SFBF`$(\mathbf{x}^t, \bar{V}^t, H_{\mathbf{x}^t}^{\delta_t,b_t}, \gamma_t, \eta_t, \delta_t, K)$. *Set $L_h \triangleq \sum_{i\in\mathcal{I}} L_{h_i}$. Then for all $\mathbf{x} \in \mathcal{X}$ and $k \in \{0, 1, \ldots, K-1\}$, we have*

$$\begin{aligned}
\left\|\mathbf{z}_{k+1}^{(t)}-\mathbf{x}\right\|^2 \leq (1-\gamma_t\eta_t)\left\|\mathbf{z}_k^{(t)}-\mathbf{x}\right\|^2 - (1-2\gamma_t\eta_t)\left\|\mathbf{z}_{k+1/2}^{(t)}-\mathbf{z}_k^{(t)}\right\|^2 \\
+ 8\gamma_t^2\sum_{i\in\mathfrak{I}} L_{h_i}^2 n_i^2 + 4\gamma_t^2(\mathcal{L}_f(\xi_{t,k+1/2})^2+\eta_0^2)\left\|\mathbf{z}_{k+1/2}^{(t)}-\mathbf{x}^t\right\|^2 \\
- 2\gamma_t\langle \hat{V}_{t,k+1/2}(\mathbf{z}_{k+1/2}^{(t)}) + H^{\delta_t}_{\mathbf{z}_{k+1/2}^{(t)}}(\mathbf{W}_{t,k+1/2}) \\
- V(\mathbf{z}_{k+1/2}^{(t)}) - \nabla h^{\delta_t}(\mathbf{z}_{k+1/2}^{(t)}), \mathbf{z}_{k+1/2}^{(t)}-\mathbf{x}\rangle \\
- 2\gamma_t\langle \big(V(\mathbf{x}^t)+\nabla h^{\delta_t}(\mathbf{x}^t)\big) \\
- \Big(\hat{V}_{t,k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{t,k+1/2})\Big), \mathbf{z}_{k+1/2}^{(t)}-\mathbf{x}\rangle \\
- 2\gamma_t\langle \varepsilon_V^t(\xi_t^{1:b_t}) + \varepsilon_h^t(\mathbf{W}_t^{1:b_t}), \mathbf{z}_{k+1/2}^{(t)}-\mathbf{x}\rangle \\
- 2\gamma_t\Big(\langle V^{\eta_t}(\mathbf{x}), \mathbf{z}_{k+1/2}^{(t)}-\mathbf{x}\rangle + h(\mathbf{z}_{k+1/2}^{(t)}) - h(\mathbf{x})\Big) + 2\gamma_t\delta_t L_h.
\end{aligned}$$

We next prove a.s. convergence of $\{\mathbf{x}^t\}_{t=0}^T$ to the least-norm solution of $\mathrm{MVI}(V,h)$ as $T\to\infty$. The proof rests on a fine comparison between the algorithmic sequence $\{\mathbf{x}^t\}$ and the sequence of solutions of the regularized problems $\mathrm{MVI}(V^{\eta_t},h)$, denoted as $\{\mathbf{s}_t\}_{t=0}^T$.

Theorem 4.1 *Let Assumptions 1–4 hold. Suppose we are given sequences $\{\gamma_t\}_{t\in\mathbb{N}}$, $\{\delta_t\}_{t\in\mathbb{N}}$ and $\{\eta_t\}_{t\in\mathbb{N}}$, satisfying the following conditions:*

(a) $\lim_{t\to\infty}\frac{\gamma_t}{\eta_t} = \lim_{t\to\infty}\frac{\delta_t}{\eta_t} = 0$, *and* $\sum_{t=0}^\infty \gamma_t^2 < \infty$, $\sum_{t=0}^\infty \gamma_t\eta_t = \infty$;
(b) $\gamma_t\eta_t \in (0,1/2)$ *and* $\lim_{t\to\infty}\eta_t = 0$;
(c) $\sum_{t=0}^\infty\left(\frac{\eta_t-\eta_{t-1}}{\eta_t}\right)^2(1+\frac{1}{\gamma_t\eta_t}) < \infty$ *and* $\lim_{t\to\infty}\left(\frac{\eta_t-\eta_{t-1}}{\eta_t}\right)^2\left(\frac{1+\frac{1}{\gamma_t\eta_t}}{\gamma_t\eta_t}\right) = 0.$

Then $\mathbb{P}(\lim_{t\to\infty}\left\|\mathbf{x}^t-\mathbf{x}^*\right\|=0)=1$, where $\mathbf{x}^*$ denotes the unique least norm solution of $\mathrm{MVI}(V,h)$; i.e. $\{(\mathbf{x}^t,\mathbf{y}(\mathbf{x}^t))\}_{t\in\mathbb{N}}$ converges almost surely to a hierarchical equilibrium of the game (P).

The proof of this Theorem can be found in Sect. 7.1.2.

Remark 4.1 The following sequences satisfy the conditions of Theorem 4.1: Let $A\in(0,1/2)$ and $p,q\in(0,1)$ so that $p<q, 0<p+q<1$ and $q>1/2$. Let $\gamma_t = \frac{A}{(t+1)^q} = \delta_t$ and $\eta_t = \frac{1}{(t+1)^p}$. With this choice, it is clear that $\frac{\gamma_t}{\eta_t} = \frac{\delta_t}{\eta_t} = A(t+1)^{p-q}\to 0$ as $t\to\infty$, and $\sum_t\gamma_t^2<\infty$. Additionally, $\gamma_t\eta_t = A(t+1)^{-(p+q)}$ is not summable as $p+q<1$. Finally, define the continuous paths $\eta:\mathbb{R}_{\geq 0}\to\mathbb{R}_{\geq 0}$ and $\gamma:\mathbb{R}_{\geq 0}\to\mathbb{R}_{\geq 0}$ by

$$\eta(t) := \frac{1}{(t+1)^p}, \text{ and } \gamma(t) := \frac{A}{(t+1)^q},$$

so that $\eta(\lfloor t\rfloor) = \eta_{\lfloor t\rfloor}$ and $\gamma(\lfloor t\rfloor) = \gamma_{\lfloor t\rfloor}$. Using these continuous paths, we see that for t sufficiently large

$$\left(\frac{\eta_t - \eta_{t-1}}{\eta_t}\right)^2 (1 + \frac{1}{\gamma_t \eta_t}) \approx \left(\frac{\eta'(t)}{\eta(t)}\right)^2 (1 + \frac{1}{\gamma_t \eta_t}) = \frac{p^2}{(t+1)^2}(1 + A(t+1)^{p+q})$$

Hence, also this term is summable, as $p + q < 1$. The last requirement can be verified in the same way.

4.3 Finite-Time Complexity

The convergence measure usually employed for $\mathrm{MVI}(V, h)$ is the gap function

$$\Gamma(\mathbf{x}) \triangleq \sup_{\mathbf{z} \in \mathcal{X}} (\langle V(\mathbf{z}), \mathbf{x} - \mathbf{z} \rangle + h(\mathbf{x}) - h(\mathbf{z})), \tag{12}$$

Since we work in probabilistic setting, naturally our convergence measure will be based on $\mathbb{E}[\Gamma(\mathbf{x})]$. Our main finite-time iteration complexity result in terms of this performance measure is the next Theorem, whose proof is detailed in Sect. 7.1.3.

Theorem 4.2 *Let Assumptions 1–4 hold and fix $T \in \mathbb{N}$. Consider Algorithm* `VRHGS` *with the inputs $\gamma_t = \eta_t = \delta_t = 1/T$, as well as $K = T$ and batch size $b_t \geq T^2$. Then, $\mathbb{E}[\Gamma(\bar{\mathbf{z}}^T)] = \mathcal{O}\left(\frac{C\sigma}{T}\right)$, where $\sigma \triangleq \sqrt{2M_V^2 + 2\sum_{i \in \mathcal{I}} L_{h_i}^2 n_i^2}$, $C = \max_{i \in \mathcal{I}} C_i$ (cf. Assumption (2.i)), and $\bar{\mathbf{z}}^T \triangleq \frac{\sum_{t=0}^{T-1} \gamma_t \bar{\mathbf{z}}^t}{\sum_{t=0}^{T-1} \gamma_t}$ for $\bar{\mathbf{z}}^t \triangleq \frac{1}{K} \sum_{k=0}^{K-1} \mathbf{z}_{k+1/2}^{(t)}$.*

We next evaluate the oracle complexity of `VRHGS`. To be precise, let $\mathsf{OC}(T, K, \{b_t\}_{t=0}^{T-1})$ the number of random variables method `VRHGS` generates in the inner and outer loop until we achieve a solution that pushes the expected gap below a target value ε.

Remark 4.2 We point out that this measure of oracle complexity ignores the computational effort arising from solving the lower level problem attached with player $i \in \mathcal{I}$. This is consistent as we assume that the solution map is provided to us in terms of an oracle. A full-fledged complexity analysis can be done, and will appear in a future publication.

Proposition 4.1 *Let $\varepsilon > 0$ be given, and set $T = \lceil 1/\varepsilon \rceil$. If we choose the same sequences as in Theorem 4.2, we have $\mathsf{OC}(T, K, \{b_t\}_{t=0}^{T-1}) = \mathcal{O}(2N/\varepsilon^3)$.*

Proof The number of random variables generated in each inner loop iteration is $2K \times N$. In each round of the outer loop we sample $2b_t \times N$ random variables. Hence, the total oracle complexity is $\mathsf{OC}(T, K, \{b_t\}_{t=0}^{T}) = 2KTN + 2N\sum_{t=0}^{T-1} b_t$. For the specific values of $T, K, \gamma_t, \delta_t, \eta_t$ defined in Theorem 4.2 and $b_t = T^2$, it clearly follows $\mathsf{OC}(T, K, \{b_t\}_{t=0}^{T-1}) = \mathcal{O}(2N/\varepsilon^3)$. □

4.4 Inexact Generalization

A key shortcoming in the implementation of `VRHGS` is the need for exact solutions of the lower-level problem. Naturally, when the solution map $\mathbf{y}_i(\cdot)$ corresponds to the solution of a large-scale stochastic optimization/VI problem, this claim is hard to justify. In this section, we allow for an inexact solution $\mathbf{y}_i^\varepsilon(\mathbf{x})$ associated with an error level ε, defined as

$$\mathbb{E}[\left\|\mathbf{y}^\varepsilon(\mathbf{x}) - \mathbf{y}(\mathbf{x})\right\| |\mathbf{x}] \leq \varepsilon \quad \text{a.s.} \tag{13}$$

Under the inexact lower level solution $\mathbf{y}_i^\varepsilon$, we let $h_i^\varepsilon(\mathbf{x}_i) \triangleq g_i(\mathbf{x}_i, \mathbf{y}_i^\varepsilon(\mathbf{x}_i))$.

Remark 4.3 We can obtain the inexact solution $\mathbf{y}_i^\varepsilon(\mathbf{x}_i)$ with rather efficient numerical methods. First, we can parallelize the computation since the problems $\mathrm{VI}(\phi_i(\mathbf{x}_i, \cdot), \mathcal{Y}_i)$ are uncoupled. Second, the mapping $\phi_i(\mathbf{x}_i, \cdot)$ is assumed to be strongly monotone. Hence, we can solve the VI to ε-accuracy with exponential rate using for instance the method in [14].

As in the exact regime, we assume that player i has access to an oracle with which she can construct a spherical approximation of the gradient of the implicit function h_i^ε. Hence, for given $\delta > 0$, we let $h_i^{\varepsilon,\delta}(\mathbf{x}_i) \triangleq \int_{\mathbb{B}_{n_i}} h_i^\varepsilon(\mathbf{x}_i + \delta\mathbf{w}) \frac{\mathbf{dw}}{\mathbf{Vol}_n(\mathbb{B}_{n_i})}$. We denote the resulting estimators by $H_{i,\mathbf{x}_i}^{\delta,\varepsilon}(\mathbf{W}_i) = n_i \mathbf{W}_i \nabla_{(\mathbf{W}_i,\delta)} h_i^\varepsilon(i, \mathbf{x}_i)$, and the mini-batch versions $H_{i,\mathbf{x}_i}^{\delta,\varepsilon,b} \triangleq \frac{1}{b}\sum_{s=1}^b H_{i,\mathbf{x}_i}^{\delta,\varepsilon}(\mathbf{W}_i^{(s)})$. With these concepts in hand, we can adapt `VRHGS` to run exactly the same way as described in Algorithms 1 and 2, replacing the appearance of quantities involving h_i with its inexact version h_i^ε; see Sect. 7.2 for a precise formulation of the method.

Theorem 4.3 *Let Assumptions 2 hold and fix $T \in \mathbb{N}$. Consider Algorithm I-*`VRHGS`*, defined in Sect. 7.2, with the sequence $\gamma_t = \eta_t = \delta_t = 1/T$, as well as $K = T$, the batch size $b_t \geq T^2$ and inexactness regime $\varepsilon_t = 1/T^2$. Then $\mathbb{E}[\Gamma(\bar{\mathbf{z}}^T)] = \mathcal{O}\left(\frac{C\sigma}{T}\right)$, where $\sigma \triangleq \sqrt{2M_V^2 + 2\sum_{i\in\mathcal{I}} L_{h_i}^2 n_i^2}$, $C = \max_{i\in\mathcal{I}} C_i$ (cf. Assumption (2.i)).*

Notably, tractable resolution of the proposed stochastic hierarchical game is possible in inexact regimes and such practically motivated schemes are not adversely affected in terms of either the rate or complexity guarantees.

5 Hierarchical Games in Power Markets

In this section we present a model inspired by Hobbs and Pang [34], but suitably modified to account for uncertainty in prices and costs, multi-period settings with ramping constraints, and the incorporation of virtual power plants (VPPs) (see [22, 53] for a review of VPPs and power markets). The model we present below is at this stage an academic example that demonstrates the modelling power of our hierarchical games approach. In future studies we aim for numerical implementations of this model.

Consider a set of nodes $\mathcal{N}$ of a network and a set of time periods $\mathcal{T} \triangleq \{1, 2, \ldots, T\}$. A generation firm is indexed by f, where f belongs to the finite set $\mathcal{F}$ and each firm is assumed to have an associated VPP. At a node i in the network, a firm f may generate $g_{f,i,t}$ units via conventional generation in period t and sell $s_{f,i,t}$ units during the same period. In addition, at time period t, firm f may generate $P_{f,t}^{\mathrm{pv,S}} + P_{f,t}^{\mathrm{pv,L}}$ units of power via PV capacity, of which $P_{f,t}^{\mathrm{pv,S}}$ is sold and $P_{f,t}^{\mathrm{pv,L}}$ is employed for meeting load. The total amount of power sold at node i during period t by all generating firms is represented by $S_{i,t}$, i.e. $S_{i,t} = \sum_{f \in \mathcal{F}} s_{f,i,t}$. If the nodal power price at the ith node during period t is a random function given by $p_{i,t}(\bullet, \xi)$, where $p_{i,t}(\bullet, \xi)$ is a decreasing function of aggregate nodal sales $S_{i,t}$ for any $\xi \in \Xi$. It follows that firm f's revenue from non-PV power sales at node i during period t under realization ξ is $p_{i,t}(S_{i,t}, \xi)s_{f,i}$. Sales of PV output by firm f at time t is priced using a function $p^{\mathrm{R}}_{f,t}$, earning a revenue given by $p^{\mathrm{R}}_{f,t}\left(P^{\mathrm{pv,S}}_{f,t} + P^{\mathrm{pv,S,V}}_{f,t}\right) P^{\mathrm{pv,S}}$, where $P^{\mathrm{pv,S,V}}_{f,t}$ denotes the sales of firm f's associated VPP (whose problem is described later in this section). We observe that renewable power is priced using this price function, distinct from conventional sources, and is designed to provide incentives for renewable expansion [62]. The costs incurred by firm f at node i during period t are given by the sum of the cost of generating $g_{f,i,t}$ and the cost of transmitting the excess $(s_{f,i,t} - g_{f,i,t})$. Let the random cost function of generation associated with firm f at node i be given by $c_{fi}(\bullet, \zeta)$ while the cost of transmitting power from an arbitrary node (referred to as the hub) to node i is given by w_i. The constraint set incorporates a balance between aggregate sales, aggregate generation, and power injection into the VPP at all nodes for every time period t. In addition, we impose nonnegativity bound on sales and generation at any time period t, enforce a capacity limit on generation levels, and introduce ramping constraints on the change in generation levels. The resulting problem faced by generating firm f, denoted by ($\mathbf{Firm}_f$), requires minimizing generation cost less revenue from conventional and PV sales by optimizing sales $s_{f,i,t}$ and generation $g_{f,i,t}$ at every node i and every time period t as well as load-directed PV output $P^{\mathrm{pv,L}}_{f,t}$ and PV sales $P^{\mathrm{pv,S}}_{f,t}$ at time t. If $P^{\mathrm{pv,S,V,}\epsilon}_{f,t}(\cdot)$ denotes a component of the single-valued solution map of the ϵ-regularized problem of the VPP associated with firm f, denoted by ($\mathbf{VPP}_f(P^{\mathrm{pv,S}}_f)$), then firm f's problem is defined as follows, where $\widehat{\mathcal{T}} = \{1, \ldots, T-1\}$.

$$
\begin{aligned}
&\underset{s_{f,i,t},\, g_{f,i,t},\, P^{\mathrm{pv,S}}_{f,t}\, P^{\mathrm{pv,L}}_{f,t}}{\textbf{maximize}} && \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\left(p_{i,t}(S_{i,t},\xi)s_{f,i,t} - c_{fi}(g_{fi,t},\zeta) - (s_{fi,t} - g_{fi,t})w_{i,t}\right)\right] \\
& && +\sum_{t\in\mathcal{T}}\left(p^{\mathrm{R}}_{f,t}\left(P^{\mathrm{pv,S}}_{f,t} + P^{\mathrm{pv,S,V,}\epsilon}_{f,t}(P^{\mathrm{pv,S}}_{f,t})\right)P^{\mathrm{pv,S}}\right) \\
&\textbf{subject to} && \left\{\begin{array}{l} 0 \le g_{f,i,t} \le \mathrm{cap}_{fi} \\ 0 \le s_{f,i,t} \end{array}\right\} \qquad \forall t \in \mathcal{T}, \quad \forall i \in \mathcal{N} \\
& && -\mathrm{RR}^{\mathrm{down}}_{f,i} \le g_{f,i,t} - g_{f,i,t-1} \le \mathrm{RR}^{\mathrm{up}}_{fi} \quad \forall t \in \widehat{\mathcal{T}}, \forall i \in \mathcal{N} \\
& && P^{\mathrm{pv,L}}_{f,t} + P^{\mathrm{pv,S}}_{f,t} \le \mathrm{cap}^{\mathrm{pv}}_{f,t} \qquad \forall t \in \mathcal{T} \\
& && 0 \le P^{\mathrm{pv,L}}_{f,t}, P^{\mathrm{pv,S}}_{f,t} \qquad \forall t \in \mathcal{T} \\
&\text{and} && \sum_{i\in\mathcal{N}}(s_{f,i,t} - g_{f,i,t}) - P^{\mathrm{pv,L}}_{f,t} = 0 \qquad \forall t \in \mathcal{T}.
\end{aligned}
\tag{$\mathbf{Firm}_f(s_{-f}, g_{-f})$}
$$

It bears reminding that the last set of constraints specified in $\left(\mathbf{Firm}_f(s_{-f}, g_{-f})\right)$ are parametrized by rival decisions and can be relaxed with Lagrange multiplier

$\lambda_{f,t}$, leading to the following *relaxed* problem $\left(\mathbf{Firm}_f^{\mathrm{rel}}(s_{-f}, g_{-f})\right)$, defined as follows.

$$\begin{aligned}
&\underset{s_{f,i,t},\, g_{f,i,t},\, P_{f,t}^{\mathrm{pv,S}}\, P_{f,t}^{\mathrm{pv,L}}}{\mathbf{maximize}} \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i\in\mathcal{N}} \left(p_{i,t}(S_{i,t},\xi)s_{f,i,t} - c_{f,i}(g_{f,i,t},\zeta) - (s_{f,i,t} - g_{f,i,t})w_{i,t} \right)\right] \\
&\quad + \sum_{t\in\mathcal{T}} \left(p_{f,t}^{\mathrm{R}} \left(P_{f,t}^{\mathrm{pv,S}} + P_{f,t}^{\mathrm{pv,S,V}} \right) P_{f,t}^{\mathrm{pv,S}} \right) - \sum_{t=1}^{T} \lambda_{f,t}^{\top} \left(\sum_{i\in\mathcal{N}} \left((s_{f,i,t} - g_{f,i,t}) - P_{f,t}^{\mathrm{pv,L}} \right) \right) \\
&\mathbf{subject\ to} \left\{ \begin{array}{l} 0 \le g_{fi,t} \le \mathrm{cap}_{fi} \\ 0 \le s_{fi,t} \end{array} \right\} \qquad \forall t \in \mathcal{T}, \quad \forall i \in \mathcal{N} \\
&-\mathrm{RR}_{f,i}^{\mathrm{down}} \le g_{f,i,t} - g_{f,i,t-1} \le \mathrm{RR}_{f,i}^{\mathrm{up}} \quad \forall t \in \widehat{\mathcal{T}}, \quad \forall i \in \mathcal{N} \\
&P_{f,t}^{\mathrm{pv,L}} + P_{f,t}^{\mathrm{pv,S}} \le \mathrm{cap}_{f,t}^{\mathrm{pv}} \qquad \forall t \in \mathcal{T} \\
&0 \le P_{f,t}^{\mathrm{pv,L}}, P_{f,t}^{\mathrm{pv,S}} \qquad \forall t \in \mathcal{T}.
\end{aligned} \qquad (\mathbf{Firm}_f(s_{-f}, g_{-f}))$$

In addition, we introduce a pricing player $\left(\mathbf{Price}\left(s_f, g_f, P_{f,t}^{\mathrm{pv,L}}\right)\right)$ corresponding to the determination of $\lambda_{f,t}$ for $f \in \mathcal{F}$ and $t \in \mathcal{T}$, defined as follows.

$$\underset{\lambda}{\mathbf{minimize}} \sum_{f\in\mathcal{F}}\sum_{t\in\mathcal{T}} \lambda_{f,t}^{\top} \left(\sum_{i\in\mathcal{N}} (s_{f,i,t} - g_{f,i,t}) - P_{f,t}^{\mathrm{pv,L}} \right). \qquad \left(\mathbf{Price}\left(s, g, P^{\mathrm{pv,L}}\right)\right)$$

Note that, the generating firm sees the transmission fee $w_{i,t}$ and the rival firms' sales $s_{-fi,t} \equiv \{s_{hi,t} \;:\; h \neq f\}$ as exogenous parameters to its optimization problem even though they are endogenous to the overall equilibrium model as we will see shortly. The ISO sees the transmission fees $w = (w_{i,t})_{i\in\mathcal{N},t\in\mathcal{T}}$ as exogenous and prescribes flows $y = (y_{i,t})_{i\in\mathcal{N},t\in\mathcal{T}}$ as per a solution of the following linear program

$$\begin{aligned}
&\underset{y}{\mathbf{maximize}} \sum_{i\in\mathcal{N}}\sum_{t\in\mathcal{T}} y_{i,t} w_{i,t} \\
&\mathbf{subject\ to} \sum_{i\in\mathcal{N}} \mathrm{PDF}_{ij} y_{i,t} \le \hat{T}_j \qquad \forall j \in \mathcal{K},\ \forall t \in \mathcal{T},
\end{aligned} \qquad (\mathbf{ISO}(w))$$

where $\mathcal{K}$ is the set of all arcs or links in the network with node set $\mathcal{N}$, $\hat{T}_j$ denotes the transmission capacity of link j, $y_{i,t}$ represents the transfer of power (in MW) by the system operator from a hub node to node node i and PDF_{ij} denotes the power transfer distribution factor, which specifies the MW flow through link j as a consequence of unit MW injection at an arbitrary hub node and a unit withdrawal at node i. Finally, to clear the market, the transmission flows y_i must balance the net sales at each node, as specified next.

$$y_{i,t} = \sum_{f\in\mathcal{F}} \left(s_{f,i,t} - g_{f,i,t} \right) \qquad \forall i \in \mathcal{N}, \quad \forall t \in \mathcal{T}. \tag{14}$$

In fact, this constraint can be recast as a collection of pricing players, denoted by ($\mathbf{Flow}^{\text{price}}(g, s, y)$).

$$\underset{\beta}{\textbf{minimize}} \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}}\beta_{i,t}\left(y_{i,t} - \sum_{f\in\mathcal{F}}\left(s_{f,i,t} - g_{f,i,t}\right)\right). \qquad (\mathbf{Flow}^{\text{price}}(g, s, y))$$

We now extend the scope of the framework of power markets by incorporating virtual power plants. A virtual power plant (VPP) represents a collection of distributed energy resources (DERs) (e.g., batteries, smart thermostats, controllable water heaters, and rooftop solar) that can be coordinated to enhance the reliability and sustainability of the electric grid. To satisfy the short-term goals for clean energy technology (CET) deployment, it has been estimated that U.S. VPP capacity must triple by 2030, leading to potential savings of \$10 billion in annual grid costs [22]. Without loss of generality, we assume that any firm $f \in \mathcal{F}$ has a collection of components, which collectively provide "virtual power" in addition to conventional generation. Before proceeding, we model three components in such a VPP, akin to approaches employed in [11, 28, 38].

(a) *Battery storage* Suppose the storage unit associated with firm f has an associated state of charge (SOC) level at time t by $\mathrm{SOC}_{f,t}$.

$$\mathrm{SOC}_{f,t+1} = \mathrm{SOC}_{f,t} + \frac{\eta_f^{\mathrm{b,ch}}\Delta t}{Q_f^{\mathrm{b}}} P_{f,t}^{\mathrm{b,ch}} - \frac{\Delta t}{\eta_f^{\mathrm{b,ds}} Q_f^{\mathrm{b}}} P_{f,t}^{\mathrm{b,ds}} \; \forall\, t \in \mathcal{T}, \tag{15}$$

where $P_{f,t}^{\mathrm{b,ch}}$ and $P_{f,t}^{\mathrm{b,ds}}$ represent charging and discharging power-levels at time t, $\eta_f^{\mathrm{b,ch}}$ and $\eta_f^{\mathrm{b,ds}}$ represent charging and discharging efficiencies at time t, while Q_f^{b} and Δt denote the battery capacity and time interval, respectively. In addition, $\mathrm{SOC}_{f,t}$ is bounded between a minimum value $\mathrm{SOC}_f^{\min}$ and maximum value $\mathrm{SOC}_f^{\max}$ while at any time t, charging and discharging rates cannot exceed $P_f^{\mathrm{b,ch,mx}}$ and $P_f^{\mathrm{b,ds,mx}}$, respectively, as captured by the following bounds.

$$\mathrm{SOC}_f^{\min} \le \mathrm{SOC}_{f,t} \le \mathrm{SOC}_f^{\max} \qquad \forall t \in \mathcal{T} \tag{16}$$

$$0 \le \frac{P_{f,t}^{\mathrm{b,ch}}}{P_f^{\mathrm{b,ch,mx}}} \le 1 \qquad \forall t \in \mathcal{T} \tag{17}$$

$$0 \le \frac{P_{f,t}^{\mathrm{b,ch}}}{P_f^{\mathrm{b,ds,mx}}} \le 1 \qquad \forall t \in \mathcal{T}. \tag{18}$$

(b) *Intermittent resources* We now model intermittency by considering a photovoltaic (PV) array associated with firm f, where at time t, $P_{f,t}^{\mathrm{pv,L,V}}$ and $P_{f,t}^{\mathrm{pv,S,V}}$ denote the PV output employed for meeting load and for deriving sales revenue, respectively. Further, $P_f^{\mathrm{pv,max}}$ represents maximum PV power at time t. Consequently, PV output is modeled as

$$P^{\mathrm{pv,L,V}}_{f,t} + P^{\mathrm{pv,S,V}}_{f,t} = (1 - U^{\mathrm{pv}}_{f}) P^{\mathrm{pv,max}}_{f,t} E_{f,t} \qquad \forall t \in \mathcal{T} \tag{19}$$

$$0 \le U^{\mathrm{pv}}_{f,t} \le 1 \qquad \forall t \in \mathcal{T} \tag{20}$$

$$0 \le P^{\mathrm{pv,L,V}}_{f,t}, P^{\mathrm{pv,S,V}}_{f,t} \qquad \forall t \in \mathcal{T}, \tag{21}$$

where $U^{\mathrm{pv}}_{f,t}$ denotes the PV curtailment employed by firm f at time t while $P^{\mathrm{pv,max}}_{f,t}$ scales with the solar irradiance at time t as seen by firm f, denoted by $E_{f,t}$. We observe that $P^{\mathrm{pv,S,V}}_{f,t}, P^{\mathrm{pv,L,V}}_{f,t} \ge 0$ for any $f \in \mathcal{F}$ and any $t \in \mathcal{T}$.

(c) *Thermal onsite generation* Often VPPs may incorporate onsite thermal generation that can be employed. For any $f \in \mathcal{F}$, suppose the generation capacity is denoted by $\mathrm{Cap}^{\mathrm{onsite}}_{f}$ while the upward and downward ramping rates are given by $\mathrm{RR}^{\mathrm{up}}_{f}$ and $\mathrm{RR}^{\mathrm{down}}_{f}$, respectively. Consequently, if the generation output at time t is denoted by $P^{\mathrm{onsite}}_{f,t}$, then for any $t \in \mathcal{T}$, we have

$$0 \le P^{\mathrm{onsite}}_{f,t} \le \mathrm{Cap}^{\mathrm{onsite}}_{f}. \qquad \forall t \in \mathcal{T} \tag{22}$$

Furthermore, changes in generation level are bounded by ramping rates, as captured by the following set of two-sided constraints.

$$-\mathrm{RR}^{\mathrm{down}}_{f} \le P^{\mathrm{onsite}}_{f,t+1} - P^{\mathrm{onsite}}_{f,t+1} \le \mathrm{RR}^{\mathrm{up}}_{f}. \quad \forall t \in \mathcal{T} \tag{23}$$

VPPs are characterized by an idiosyncratic load profile that cannot be controlled; specifically, $P^{\mathrm{L,V}}_{f,t}$ denotes the load associated with VPP f at time t. In more comprehensive models, we may incorporate HVAC and water heater components that allow for more fine-grained control of such loads but for purposes of simplicity, we omit such a discussion here. In the current setting, the effective load emerging from managing the VPP associated with firm f and time t is given by the sum of the uncontrollable load and the battery load (charging less discharging level) less the sum of onsite generation and load-directed PV output is required to be nonpositive, as specified next.

$$P^{\mathrm{L,V}}_{f,t} + \left(P^{\mathrm{b,ch}}_{f,t} - P^{\mathrm{b,ds}}_{f,t} \right) - P^{\mathrm{onsite}}_{f,t} - P^{\mathrm{pv,L,V}}_{f,t} \le 0. \qquad \forall t \in \mathcal{T} \tag{24}$$

Note that the satisfaction of this constraint relies on appropriate sizing of the battery capacity Q^{b}_{f} and the onsite generation capacity $\mathrm{Cap}^{\mathrm{onsite}}_{f}$. Suppose the decision vector of firm f's VPP is denoted by y^{vpp}_{f}, defined as

$$y^{\mathrm{vpp}}_{f} = \left(\mathrm{SOC}_{f}; P^{\mathrm{b,ch}}_{f}; P^{\mathrm{b,ds}}_{f}; P^{\mathrm{pv,L,V}}_{f}; P^{\mathrm{pv,S,V}}_{f}; U^{\mathrm{pv}}_{f}; P^{\mathrm{onsite}}_{f} \right).$$

The profit function associated with firm f's VPP is the revenue obtained by sales revenue derived from PV sales less the VPP's operational cost (given by the sum of the costs of onsite generation and the (converted) cost of PV curtailment), defined as

$$\mathrm{r}_f^{\mathrm{vpp}}(y_f^{\mathrm{vpp}}; P_f^{\mathrm{pv,S}}) \triangleq \sum_{t \in \mathcal{T}} \left(\underbrace{p_{f,t}^{\mathrm{R}} \left(P_{f,t}^{\mathrm{pv,S,V}} + P_{f,t}^{\mathrm{pv,S}} \right) P_{f,t}^{\mathrm{pv,S,V}}}_{\text{VPP revenue from PV sales}} - \underbrace{c_f^{\mathrm{onsite}}(P_{f,t}^{\mathrm{onsite}})}_{\text{Cost of onsite gen.}} - \underbrace{\beta U_{f,t}^{\mathrm{pv}} P_{f,t}^{\max}}_{\text{Env. cost of PV curtailment}} \right),$$

where $p_{f,t}^{\mathrm{R}}(\cdot)$ denotes the price function of renewables seen at firm f at time t, while the revenue obtained is given by $p_{f,t}^{\mathrm{R}} \left(P_{f,t}^{\mathrm{pv,S,V}} + P_{f,t}^{\mathrm{pv,S}} \right) P_{f,t}^{\mathrm{pv,S,V}}$. We may then formally define the optimization problem faced by the VPP associated with firm f, where the polyhedral constraints are captured by $\left\{ y_f^{\mathrm{vpp}} \mid A_f y_f^{\mathrm{vpp}} \le d_f \right\}$ where $A_f \in \mathbb{R}^{m \times n}$ and $d_f \in \mathbb{R}^m$.

$$\begin{aligned} &\underset{y_f^{\mathrm{vpp}}}{\textbf{maximize}}\ r_f^{\mathrm{vpp}}(y_f^{\mathrm{vpp}}; P_f^{\mathrm{pv,S}}) \\ &\textbf{subject to } (15) - (24) \equiv \left\{ y_f^{\mathrm{vpp}} \mid A_f y_f^{\mathrm{vpp}} \le d_f \right\}. \end{aligned} \qquad (\mathbf{VPP}_f(P_f^{\mathrm{pv,S}}))$$

We now observe that the resulting equilibrium problem comprises of a collection of firms, each of which has a single follower as captured by a VPP, in addition to the ISO and a set of players that determine prices. To facilitate analysis of the necessary and sufficient equilibrium conditions of this hierarchical game, we approximate $\left(\mathbf{VPP}_f(P_f^{\mathrm{pv,S}})\right)$ by employing a smooth (exact) penalized approximation; this latter formulation is of particular relevance in deriving the concavity of the function $p_{f,t}^{\mathrm{R}} \left(P_{f,t}^{\mathrm{pv,S}} + P_{f,t}^{\mathrm{pv,S,V},\epsilon}(P_{f,t}^{\mathrm{pv,S}}) \right) P_{f,t}^{\mathrm{pv,S}}$ in $P_{f,t}^{\mathrm{pv,S}}$, where $P_{f,t}^{\mathrm{pv,S,V},\epsilon}(P_{f,t}^{\mathrm{pv,S}})$ is a component of the single-valued solution map $y_f^{\mathrm{vpp},\epsilon}(P_{f,t}^{\mathrm{pv,S}})$, a solution of the ϵ-regularized and the ϵ-smoothed (exact) penalized approximation of $\left(\mathbf{VPP}_f(P_f^{\mathrm{pv,S}})\right)$. To this end, we define the exact penalty function φ and its smoothed counterpart φ_ϵ as

$$\begin{aligned} \varphi(A_f y_f^{\mathrm{vpp}} - d_f) &\triangleq \sum_{i=1}^m \max\{a_{fi}^\top y_f^{\mathrm{vpp}} - d_{fi}, 0\}, \quad \varphi_\epsilon(A_f y_f^{\mathrm{vpp}} - d_f) \\ &\triangleq \sum_{i=1}^m \psi_\epsilon(a_{fi}^\top y_f^{\mathrm{vpp}} - d_{fi}) \end{aligned} \tag{25}$$

$$\text{and } \psi_\epsilon(t) \triangleq \begin{cases} 0 & \text{if } t \le 0 \\ \frac{\mu t^2}{2\epsilon} & \text{if } 0 \le t \le \epsilon \\ \mu(t - \frac{\epsilon}{2}) & \text{if } t \ge \epsilon. \end{cases} \tag{26}$$

This penalty function and its smoothed counterpart are employed in formally defining the exact penalty reformulation of (VPP$_f$) and its smoothed counterpart.

$$\left\{\begin{array}{c} \max\limits_{y_f^{\text{vpp}}} r^{\text{vpp}}(y_f^{\text{vpp}}, P_f^{\text{pv,S}}) \\ \text{subject to } Ay_f^{\text{vpp}} \le d_f \end{array}\right\} \equiv \left\{ \max_{y_f^{\text{vpp}}} \left(r^{\text{vpp}}(y_f^{\text{vpp}}, P_f^{\text{pv,S}}) + \mu\varphi(Ay_f^{\text{vpp}} - d_f) \right) \right\}$$

$$\approx \left\{ \max_{y_f^{\text{vpp}}} \left(r^{\text{vpp}}(y_f^{\text{vpp}}, P_f^{\text{pv,S}}) + \mu\varphi_\epsilon(Ay_f^{\text{vpp}} - d_f) + \tfrac{\epsilon}{2}\|y_f^{\text{vpp}}\|^2 \right) \right\}. \tag{$\mathbf{VPP}_f^\epsilon(P_f^{\text{pv,S}})$}$$

We observe that the resulting game can be viewed as a noncooperative hierarchical game, defined by upper-level player problems given by firm players $\left(\mathbf{Firm}_f^{\text{rel}}(s_{-f}, g_{-f}, \lambda_f)\right)_{f \in \mathcal{F}}$, pricing players $\left(\mathbf{Price}\left(s, g, P^{\text{pv,firm}-\text{load}}\right)\right)$, the ISO as denoted by (**ISO**), and the transmission pricing player ($\mathbf{Flow}^{\text{price}}(g, s, y)$). In addition, the set of regularized lower-level VPP problems is given by $\left(\mathbf{VPP}_f^\epsilon(P_f^{\text{pv,S}})\right)_{f\in\mathcal{F}}$. We succinctly represent this noncooperative game as an $N+2$ player game, in which the first N players correspond to firm f's problem for $f \in \mathcal{F}$ while the last two correspond to pricing players.

$$\left\{ \min_{\mathbf{z}^1 \in \mathbf{Z}^1} \left(h_1(\mathbf{z}^1, \mathbf{z}^{-1}) + g_1(\mathbf{z}^1, \mathbf{u}^1(\mathbf{z}^1)) \right) \right\} \quad \cdots \quad \left\{ \min_{\mathbf{z}^N \in \mathbf{Z}^N} \left(h_N(\mathbf{z}^N, \mathbf{z}^{-N}) + g_N(\mathbf{z}^N, \mathbf{u}^N(\mathbf{z}^N)) \right) \right\},$$

$$\left\{ \min_{\mathbf{z}^{N+1} \in \mathbf{Z}^{N+1}} \left(h_{N+1}(\mathbf{z}^{N+1}, \mathbf{z}^{-(N+1)}) \right) \right\} \quad \left\{ \min_{\mathbf{z}^{N+2} \in \mathbf{Z}^{N+2}} \left(h_{N+2}(\mathbf{z}^{N+2}, \mathbf{z}^{-(N+2)}) \right) \right\}. \tag{Hier-game$^{\text{vpp}}$}$$

Note that the first N players' objectives are characterized by hierarchical terms; specifically, the hierarchical terms $g_j(\mathbf{z}^j, \mathbf{u}^j(\mathbf{z}^j))$ for any $j \in \{1, \ldots, N\}$ correspond to the hierarchical terms in firm f's problem given by $\sum_{t\in\mathcal{T}} \left(p_{f,t}^{\text{R}} \left(P_{f,t}^{\text{pv,S}} + P_{f,t}^{\text{pv,S,V}} \right) P_{f,t}^{\text{pv,S}} \right)$ for any $f \in \mathcal{F}$.

While convexity of player problems follows in a straightforward fashion from the definition of firm problems and suitable convexity requirements on the cost functions as well as affineness requirements on the price functions. Single-valuedness of the solution map $\mathbf{u}^j(\bullet)$ follows from the observation that the regularized VPP profit function is strongly concave. Additionally, convexity of $g_j(\bullet, \mathbf{u}^j(\bullet))$ is a consequence of analogous analysis for Stackelberg leadership (cf. [17, 67]). Finally, monotonicity of F can be derived in a fashion similar to that considered in [32]. Existence of an equilibrium can then be derived in a fashion similar to that employed in [17]. A comprehensive analysis of this model is left to future work.

6 Preliminary Numerics

In this section, we examine the performance of the proposed schemes via an instance of a two-stage hierarchical game. To this end, consider an N-player hierarchical game where each player has a single follower and the follower problem is parametrized by uncertainty. This represents a two-stage variant of the class of hierarchical games presented in this paper. In particular, the ω-specific lower-level problem corresponding to the follower i, parametrized by leader decision $\mathbf{x}_i$, is

$$\max_{\mathbf{y}_i \geq 0} p(\mathbf{y}_i + \mathbf{x}_i, \omega)\mathbf{y}_i - c_i(\mathbf{y}_i). \tag{Follower$_i(\mathbf{x}_i, \omega)$}$$

Suppose the inverse demand function $p(\cdot, \omega)$ is defined as

$$\mathbb{R} \times \Omega \ni (u, \omega) \mapsto p(u, \omega) = a(\omega) - b(\omega)u.$$

Under this condition, the follower's objective can be shown to be strictly concave in $\mathbf{y}_j$. Consequently, the necessary and sufficient conditions of optimality are given by the following complementarity problem.

$$0 \leq \mathbf{y}_i \perp c_i'(\mathbf{y}_i) - p(\mathbf{x}_i + \mathbf{y}_i, \omega) - p'(\mathbf{x}_i + \mathbf{y}_i, \omega)\mathbf{y}_i \geq 0. \tag{Opt$_{\text{foll}}(\mathbf{x}^i, \omega)$}$$

We observe that the optimality conditions of (Opt$_{\text{foll}}(\mathbf{x}^i, \omega)$) correspond to a strongly monotone variational inequality problem for $\mathbf{x}_i \geq 0$ and for every $\omega \in \Omega$. Consequently, $\mathbf{y}_i : \mathbb{R}_+ \times \Omega \to \mathbb{R}_+$ is a single-valued map and is convex in its first argument for every ω if c_i is quadratic and convex for $j = 1, \ldots, N$ (see [18, Prop. 4.2]). In fact, it can be claimed that $\mathbf{y}_i(\cdot, \omega)$ is a piecewise C^2 and non-increasing function with $\partial_{\mathbf{x}_i} \mathbf{y}^i(\mathbf{x}_i, \omega) \subset (-1, 0]$ for $\mathbf{x}_i \geq 0$. Consider the ith leader's problem, defined as

$$\max_{\mathbf{x}^i \geq 0} \left[\mathbb{E}\left[p(\mathbf{x}^i + X^{-i} + \mathbf{y}^i(\mathbf{x}^i, \omega), \omega)\mathbf{x}^i\right] - C_i(\mathbf{x}_i)\right], \tag{Leader$_i(\mathbf{x}_{-i})$}$$

where $X \triangleq \sum_{i=1}^{N} \mathbf{x}_i$ and $X^{-i} \triangleq \sum_{j \neq i} \mathbf{x}_j$. Therefore, under suitable convexity and smoothness assumptions on C_i, we have that

$$\mathbb{R}_+ \ni \mathbf{x}_i \perp \mathbb{E}\left[-p(\mathbf{x}_i + X^{-i} + \mathbf{y}_i(\mathbf{x}_i, \omega), \omega) + (1 + \partial_{\mathbf{x}_i} \mathbf{y}_i(\mathbf{x}_i, \omega))b(\omega)\mathbf{x}_i\right] + C_i'(\mathbf{x}_i) \in \mathbb{R}_+,$$

where $\partial_{\mathbf{x}_i} \mathbf{y}_i(\mathbf{x}_i, \omega)$ is analyzed in [18, Elec.Comp.]. By concatenating the problems for players $1, \ldots, N$, we obtain the following complementarity problem.

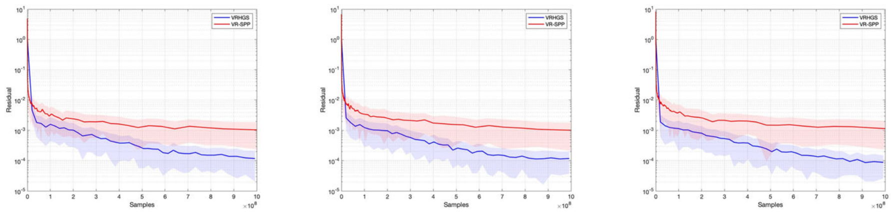

Fig. 1 Comparison of (VRHGS) and (VR-SPP) (left: $N = 13$, middle: $N = 23$, right: $N = 33$)

$$\mathbb{R}_+^N \ni \mathbf{x} \perp \prod_{i=1}^{N} \left\{ \mathbb{E}\left[-p(X + \mathbf{y}^i(\mathbf{x}_i, \omega), \omega) \right] \right\} + \left(C_i'(\mathbf{x}_i) \right)_{i=1}^N$$
$$+ \prod_{i=1}^{N} \{ \mathbb{E}[(1 + \partial_{\mathbf{x}^i} \mathbf{y}^i(\mathbf{x}_i, \omega)) b(\omega) \mathbf{x}_i] \} \in \mathbb{R}_+^N.$$

I. Problem parameters and Algorithm specifications. We consider a setting with N leaders and let C_i be a quadratic function, where $C_i(\mathbf{x}_i) \triangleq \tilde{C}_i \cdot (\mathbf{x}_i)^2$. Further, $\tilde{C}_i$ is generated from the distribution $\mathsf{U}(0, 100)$ for $i = 1, \ldots, N$, where $\mathsf{U}(l, u)$ denotes the uniform distribution on the interval $[l, u]$. In addition, $c_i = 50$, for $i = 1, \ldots, N$, $b = 7$, and $a(\omega) \sim \mathsf{U}(33, 37)$. Next, we define the parameters employed in our implementation of (**VRHGS**) and a variance-reduced inexact proximal-point framework (**VR-SPP**), defined in [15, Section 3.2.3]. Note that the solution quality is compared by estimating the expectation of the gap function evaluated at $\bar{\mathbf{z}}^T$, i.e. as given by $\mathbb{E}\left[G(\bar{\mathbf{z}}^T) \right]$.

(i) (**VRHGS**). At iteration t, we run $T = 1000$ steps and use $b_t = T^2$ samples for $t > 0$. In addition, we assume step-length $\gamma_t = \frac{1}{T}$, regularization parameter $\eta_t = \frac{1}{T}$, and smoothing parameter $\delta_t = \frac{0.1}{T}$ for $t = 0, 1, \ldots, T - 1$.

(ii) (**VR-SPP**). In our implementation of (**VR-SPP**), we employ $N_k = \lfloor 1.1^{k+1} \rfloor$, a proximal parameter $\lambda = 0.1$, and a diminishing step-length $\frac{\alpha_0}{k}$ with $\alpha_0 = 0.1$ to approximate the resolvent via the (SA) scheme. (also presented in [15, Section 3.2.3]).

II. Comparison between (VRHGS) and (VR-SPP). We compare the performance of (**VRHGS**) with (**VR-SPP**) on problems with increasing number of players N. Several aspects are apparent based on these trajectories. First, these preliminary tests suggest the empirical superiority of (**VRHGS**) with respect to (**VR-SPP**) in terms of the residual. Second, the trajectories also appear to align with the $\mathcal{O}(1/T)$ rate guarantees in terms of the expectation of the gap function (Fig. 1).

III. Inexact generalization of (VRHGS). We notice that (**VR-SPP**) requires exact lower-level follower solutions, such information may often be unavailable. However, an inexact counterpart of (**VRHGS**), referred to as (**I-VRHGS**), is suitable for such problems and its performance is now tested. We choose the same parameters as employed in (**VRHGS**) and further assume that the inexactness level $\varepsilon_t = 0.001$ In Fig. 2, we compare the plots for (**I-VRHGS**) across different regularization sequences. While all of the trajectories suggest that the schemes behave well for different choices

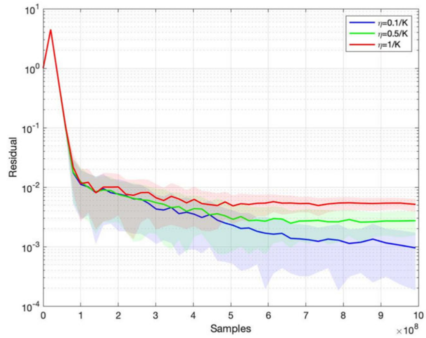

Fig. 2 Trajectories for (**I-VRHGS**) with different regularization parameters

Table 1 Errors and time of (**I-VRHGS**) with various N and K

N	$K = 500$		$K = 1000$		$K = 2000$	
	$\texttt{res}(\mathbf{x}^k)$	Time	$\texttt{res}(\mathbf{x}^k)$	Time	$\texttt{res}(\mathbf{x}^k)$	Time
13	1.5e−3	81	1.1e−3	33	2.3e−3	39
23	1.2e−3	175	9.2e−4	74	2.2e−3	112
33	1.0e−3	342	8.4e−4	119	1.8e−3	161

The errors and time represent the average over 20 runs

of η, we do observe that the trajectories display better empirical performance when $\eta = \frac{0.1}{T}$. In addition, we examine the sensitivity of the schemes to various parameters such as N and K in Table 1. First, we note that as N increases, we do not see a profound deterioration in the empirical behavior (in terms of the residual) in terms of N. Second, recall that K denotes the number of inner steps in the inner loop. It is seen that when $K = 1000$, the scheme produce better results while maintaining the same run-time with the same number of samples.

7 Proof of the Main Theorem

7.1 Analysis of the Exact Scheme

The proof on the finite time-complexity estimate starts by a technical derivation of an energy-type inequality that gives us an upper bound on the change of the energy function $\frac{1}{2}\left\|\mathbf{z}_{k+1}^{(t)} - \mathbf{x}\right\|^2$, computed within an arbitrary inner loop evaluation, and for an arbitrary anchor point $\mathbf{x} \in \mathcal{X}$. Via a sequence of tedious, but otherwise straightforward, manipulations we arrive out our first main result, Lemma 4.3. From there, we proceed as in the standard analysis of stochastic approximation schemes [54], and derive an upper bound on the gap function of the mixed variational inequality.

7.1.1 Proof of Lemma 4.3

To simplify notation we omit the dependence on the outer iteration loop t, and thus simply write $\mathbf{z}_k$ for $\mathbf{z}_k^{(t)}$. The same notational simplification will be used in all variables that are computed within the inner loop executed in the t-round of the outer loop procedure. With the hope that the reader agrees that this reduces notational complexity a bit, we proceed to derive the the postulated energy inequality. To start, we observe that for each $\mathbf{x} \in \mathcal{X}$ we have

$$
\begin{aligned}
\|\mathbf{z}_{k+1} - \mathbf{x}\|^2 &= \left\|\mathbf{z}_{k+1} - \mathbf{z}_{k+1/2} + \mathbf{z}_{k+1/2} - \mathbf{z}_k + \mathbf{z}_k - \mathbf{x}\right\|^2 \\
&= \left\|\mathbf{z}_{k+1} - \mathbf{z}_{k+1/2}\right\|^2 - \left\|\mathbf{z}_{k+1/2} - \mathbf{z}_k\right\|^2 + \|\mathbf{z}_k - \mathbf{x}\|^2 \\
&\quad + 2\langle \mathbf{z}_{k+1} - \mathbf{z}_{k+1/2}, \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&\quad + 2\langle \mathbf{z}_{k+1/2} - \mathbf{z}_k, \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&= \|\mathbf{z}_k - \mathbf{x}\|^2 - \left\|\mathbf{z}_{k+1/2} - \mathbf{z}_k\right\|^2 \\
&\quad + \left\|\gamma_t \left(\hat{V}^{\eta_t}_{k+1/2}(\mathbf{z}_{k+1/2}) - \hat{V}^{\eta_t}_{k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2})\right)\right\|^2 \\
&\quad - 2\gamma_t \langle \hat{V}^{\eta_t}_{k+1/2}(\mathbf{z}_{k+1/2}) - \hat{V}^{\eta_t}_{k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) \\
&\quad - H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2}), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&\quad + 2\langle \mathbf{z}_{k+1/2} - \left(\mathbf{z}_k - \gamma_t(\bar{V}^t + H^{\delta_t, b_t}_{\mathbf{x}^t} + \eta_t \mathbf{x}^t)\right), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&\quad - 2\gamma_t \langle \bar{V}^t + H^{\delta_t, b_t}_{\mathbf{x}^t} + \eta_t \mathbf{x}^t, \mathbf{z}_{k+1/2} - \mathbf{x}\rangle.
\end{aligned}
$$

Lemma A.1(i) gives $2\langle \mathbf{z}_{k+1/2} - \left(\mathbf{z}_k - \gamma_t(\bar{V}^t + H^{\delta_t, b_t}_{\mathbf{x}^t} + \eta_t \mathbf{x}^t\right), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \leq 0$. Additionally, Assumption 4 and Lemma A.2(b) gives

$$
\begin{aligned}
&\left\|\gamma_t \left(\hat{V}^{\eta_t}_{k+1/2}(\mathbf{z}_{k+1/2}) - \hat{V}^{\eta_t}_{k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2})\right)\right\|^2 \\
&\leq 2\gamma_t^2 \left\|\hat{V}^{\eta_t}_{k+1/2}(\mathbf{z}_{k+1/2}) - \hat{V}^{\eta_t}_{k+1/2}(\mathbf{x}^t)\right\|^2 + 2\gamma_t^2 \left\|H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2})\right\|^2 \\
&\leq 4\gamma_t^2 \left\|\hat{V}_{k+1/2}(\mathbf{z}_{k+1/2}) - \hat{V}_{k+1/2}(\mathbf{x}^t)\right\|^2 + 4\eta_t^2\gamma_t^2 \left\|\mathbf{z}_{k+1/2} - \mathbf{x}^t\right\|^2 \\
&\quad + 4\gamma_t^2 \left(\left\|H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2})\right\|^2 + \left\|H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2})\right\|^2\right) \\
&\leq 4\gamma_t^2 \left(\mathcal{L}_f(\xi_{k+1/2})^2 + \eta_t^2\right) \left\|\mathbf{z}_{k+1/2} - \mathbf{x}^t\right\|^2 + 8\gamma_t^2 \sum_{i \in \mathcal{I}} L^2_{h_i} n_i^2.
\end{aligned}
$$

It follows

$$
\begin{aligned}
\|\mathbf{z}_{k+1} - \mathbf{x}\|^2 &\leq \|\mathbf{z}_k - \mathbf{x}\|^2 - \left\|\mathbf{z}_{k+1/2} - \mathbf{z}_k\right\|^2 + 8\gamma_t^2 \sum_{i \in \mathcal{I}} L^2_{h_i} n_i^2 \\
&\quad + 4\gamma_t^2(\mathcal{L}_f(\xi_{k+1/2})^2 + \eta_0^2) \left\|\mathbf{z}_{k+1/2} - \mathbf{x}^t\right\|^2 \\
&\quad - 2\gamma_t \langle \hat{V}^{\eta_t}_{k+1/2}(\mathbf{z}_{k+1/2}) + H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - \hat{V}^{\eta_t}_{k+1/2}(\mathbf{x}^t) \\
&\quad - H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2}), z_{k+1/2} - \mathbf{x}\rangle \\
&\quad - 2\gamma_t \langle \bar{V}^t + \eta_t \mathbf{x}^t + H^{\delta_t, b_t}_{\mathbf{x}^t}, \mathbf{z}_{k+1/2} - \mathbf{x} \\
&= \|\mathbf{z}_k - \mathbf{x}\|^2 - \left\|\mathbf{z}_{k+1/2} - \mathbf{z}_k\right\|^2
\end{aligned}
$$

$$
\begin{aligned}
&+ 8\gamma_t^2 \sum_{i\in\mathfrak{I}} L_{h_i}^2 n_i^2 + 4\gamma_t^2(\mathcal{L}_f(\xi_{k+1/2})^2 + \eta_0^2) \left\|\mathbf{z}_{k+1/2} - \mathbf{x}^t\right\|^2 \\
&- 2\gamma_t \langle \hat{V}_{k+1/2}(\mathbf{z}_{k+1/2}) + H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - V(\mathbf{z}_{k+1/2}) \\
&- \nabla h^{\delta_t}(\mathbf{z}_{k+1/2}), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&+ 2\gamma_t \langle \hat{V}_{k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2}) - \bar{V}^t - H^{\delta_t,b_t}_{\mathbf{x}^t}, \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&- 2\gamma_t \langle V(\mathbf{z}_{k+1/2}) + \nabla h^{\delta_t}(\mathbf{z}_{k+1/2}) + \eta_t \mathbf{z}_{k+1/2}, \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
= &\|\mathbf{z}_k - \mathbf{x}\|^2 - \left\|\mathbf{z}_{k+1/2} - \mathbf{z}_k\right\|^2 + 8\gamma_t^2 \sum_{i\in\mathfrak{I}} L_{h_i}^2 n_i^2 \\
&+ 4\gamma_t^2(\mathcal{L}_f(\xi_{k+1/2})^2 + \eta_0^2) \left\|\mathbf{z}_{k+1/2} - \mathbf{x}^t\right\|^2 \\
&- 2\gamma_t \langle \hat{V}_{k+1/2}(\mathbf{z}_{k+1/2}) + H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - V(\mathbf{z}_{k+1/2}) \\
&- \nabla h^{\delta_t}(\mathbf{z}_{k+1/2}), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&+ 2\gamma_t \langle \hat{V}_{k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2}) - \bar{V}^t - H^{\delta_t,b_t}_{\mathbf{x}^t}, \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&- 2\gamma_t \langle V^{\eta_t}(\mathbf{z}_{k+1/2}) + \nabla h^{\delta_t}(\mathbf{z}_{k+1/2}) - V^{\eta_t}(\mathbf{x}) - \nabla h^{\delta_t}(\mathbf{x}), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&- 2\gamma_t \langle V^{\eta_t}(\mathbf{x}) + \nabla h^{\delta_t}(\mathbf{x}), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle.
\end{aligned}
$$

Since the operator $\mathbf{x} \mapsto V^{\eta_t}(\mathbf{x}) + \nabla h^{\delta_t}(\mathbf{x})$ is η_t-strongly monotone, we can further bound the expression above as

$$
\begin{aligned}
\|\mathbf{z}_{k+1} - \mathbf{x}\|^2 \leq &\|\mathbf{z}_k - \mathbf{x}\|^2 - \left\|\mathbf{z}_{k+1/2} - \mathbf{z}_k\right\|^2 \\
&+ 8\gamma_t^2 \sum_{i\in\mathfrak{I}} L_{h_i}^2 n_i^2 + 4\gamma_t^2(\mathcal{L}_f(\xi_{k+1/2})^2 + \eta_0^2) \left\|\mathbf{z}_{k+1/2} - \mathbf{x}^t\right\|^2 \\
&- 2\gamma_t \langle \hat{V}_{k+1/2}(\mathbf{z}_{k+1/2}) + H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) \\
&- V(\mathbf{z}_{k+1/2}) - \nabla h^{\delta_t}(\mathbf{z}_{k+1/2}), \mathbf{z}_{k+1/2} - \mathbf{x} \\
&+ 2\gamma_t \langle \hat{V}_{k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2}) - \bar{V}^t - H^{\delta_t,b_t}_{\mathbf{x}^t}, \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&- 2\gamma_t \eta_t \left\|\mathbf{z}_{k+1/2} - \mathbf{x}\right\|^2 \\
&- 2\gamma_t \langle V^{\eta_t}(\mathbf{x}) + \nabla h^{\delta_t}(\mathbf{x}), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle.
\end{aligned}
$$

Next, we split the mini-batch estimator $\bar{V}^t$ into its mean component and its error component so that

$$\bar{V}^t = V(\mathbf{x}^t) + \varepsilon_V^t(\xi_t^{1:b_t}). \tag{27}$$

Similarly, we write

$$\varepsilon_h^t(\mathbf{W}_t^{1:b_t}) \triangleq H^{\delta_t,b_t}_{\mathbf{x}^t} - \nabla h^{\delta_t}(\mathbf{x}^t). \tag{28}$$

Using these error terms, we may further bound the right hand side of the penultimate display as

$$
\begin{aligned}
\left\|\mathbf{z}_{k+1}-\mathbf{x}\right\|^2 \leq \|\mathbf{z}_k-\mathbf{x}\|^2 &- \left\|\mathbf{z}_{k+1/2}-\mathbf{z}_k\right\|^2 \\
&+ 8\gamma_t^2\sum_{i\in\mathfrak{I}} L_{h_i}^2 n_i^2 + 4\gamma_t^2(\mathcal{L}_f(\xi_{k+1/2})^2+\eta_0^2)\left\|\mathbf{z}_{k+1/2}-\mathbf{x}^t\right\|^2 \\
&- 2\gamma_t\langle \hat{V}_{k+1/2}(\mathbf{z}_{k+1/2}) + H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - V(\mathbf{z}_{k+1/2}) \\
&- \nabla h^{\delta_t}(\mathbf{z}_{k+1/2}), \mathbf{z}_{k+1/2}-\mathbf{x}\rangle \\
&- 2\gamma_t\langle \left(V(\mathbf{x}^t)+\nabla h^{\delta_t}(\mathbf{x}^t)\right) - \left(\hat{V}_{k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2})\right), \mathbf{z}_{k+1/2}-\mathbf{x}\rangle \\
&- 2\gamma_t\eta_t\left\|\mathbf{z}_{k+1/2}-\mathbf{x}\right\|^2 - 2\gamma_t\langle \varepsilon_V^t(\xi_t^{1:b_t}) + \varepsilon_h^t(\mathbf{W}_t^{1:b_t}), \mathbf{z}_{k+1/2}-\mathbf{x}\rangle \\
&- 2\gamma_t\langle V^{\eta_t}(\mathbf{x}) + \nabla h^{\delta_t}(\mathbf{x}), \mathbf{z}_{k+1/2}-\mathbf{x}\rangle.
\end{aligned}
$$

A simple application of the triangle inequality shows

$$
-2\gamma_t\eta_t\left\|\mathbf{z}_{k+1/2}-\mathbf{x}\right\|^2 \leq 2\gamma_t\eta_t\left\|\mathbf{z}_{k+1/2}-\mathbf{z}_k\right\|^2 - \gamma_t\eta_t\|\mathbf{z}_k-\mathbf{x}\|^2.
$$

Using this bound, we continue with the derivations above to arrive at

$$
\begin{aligned}
\left\|\mathbf{z}_{k+1}-\mathbf{x}\right\|^2 \leq (1-\gamma_t\eta_t)\|\mathbf{z}_k-\mathbf{x}\|^2 &- (1-2\gamma_t\eta_t)\left\|\mathbf{z}_{k+1/2}-\mathbf{z}_k\right\|^2 \\
&+ 8\gamma_t^2\sum_{i\in\mathfrak{I}} L_{h_i}^2 n_i^2 + 4\gamma_t^2(\mathcal{L}_f(\xi_{k+1/2})^2+\eta_0^2)\left\|\mathbf{z}_{k+1/2}-\mathbf{x}^t\right\|^2 \\
&- 2\gamma_t\langle \hat{V}_{k+1/2}(\mathbf{z}_{k+1/2}) + H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - V(\mathbf{z}_{k+1/2}) \\
&- \nabla h^{\delta_t}(\mathbf{z}_{k+1/2}), \mathbf{z}_{k+1/2}-\mathbf{x}\rangle \\
&- 2\gamma_t\langle \left(V(\mathbf{x}^t)+\nabla h^{\delta_t}(\mathbf{x}^t)\right) - \left(\hat{V}_{k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2})\right), \mathbf{z}_{k+1/2}-\mathbf{x}\rangle \\
&- 2\gamma_t\langle \varepsilon_V^t(\xi_t^{1:b_t}) + \varepsilon_h^t(\mathbf{W}_t^{1:b_t}), \mathbf{z}_{k+1/2}-\mathbf{x}\rangle \\
&- 2\gamma_t\langle V^{\eta_t}(\mathbf{x}) + \nabla h^{\delta_t}(\mathbf{x}), \mathbf{z}_{k+1/2}-\mathbf{x}\rangle.
\end{aligned}
$$

By convexity of the application $\mathbf{x}\mapsto h^{\delta_t}(\mathbf{x})$, we have

$$
h^{\delta_t}(\mathbf{z}_{k+1/2}) \geq h^{\delta_t}(\mathbf{x}) + \langle \nabla h^{\delta_t}(\mathbf{x}), \mathbf{z}_{k+1/2}-\mathbf{x}\rangle.
$$

Hence, the penultimate display turns into

$$
\begin{aligned}
\left\|\mathbf{z}_{k+1}-\mathbf{x}\right\|^2 \leq (1-\gamma_t\eta_t)\|\mathbf{z}_k-\mathbf{x}\|^2 &- (1-2\gamma_t\eta_t)\left\|\mathbf{z}_{k+1/2}-\mathbf{z}_k\right\|^2 \\
&+ 8\gamma_t^2\sum_{i\in\mathfrak{I}} L_{h_i}^2 n_i^2 + 4\gamma_t^2(\mathcal{L}_f(\xi_{k+1/2})^2+\eta_0^2)\left\|\mathbf{z}_{k+1/2}-\mathbf{x}^t\right\|^2 \\
&- 2\gamma_t\langle \hat{V}_{k+1/2}(\mathbf{z}_{k+1/2}) + H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - V(\mathbf{z}_{k+1/2}) \\
&- \nabla h^{\delta_t}(\mathbf{z}_{k+1/2}), \mathbf{z}_{k+1/2}-\mathbf{x}\rangle \\
&- 2\gamma_t\langle \left(V(\mathbf{x}^t)+\nabla h^{\delta_t}(\mathbf{x}^t)\right) - \left(\hat{V}_{k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2})\right), \mathbf{z}_{k+1/2}-\mathbf{x}\rangle
\end{aligned}
$$

$$
\begin{aligned}
&- 2\gamma_t \langle \varepsilon_V^t(\xi_t^{1:b_t}) + \varepsilon_h^t(\mathbf{W}_t^{1:b_t}), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&- 2\gamma_t \left(\langle V^{\eta_t}(\mathbf{x}), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle + h^{\delta_t}(\mathbf{z}_{k+1/2}) - h^{\delta_t}(\mathbf{x}) \right).
\end{aligned}
$$

From Lemma 2 in [69], we know for $L_h \triangleq \sum_{i\in\mathfrak{I}} L_{h_i}$ that

$$
h^{\delta_t}(\mathbf{z}_{k+1/2}) - h^{\delta_t}(\mathbf{x}) \geq h(\mathbf{z}_{k+1/2}) - h(\mathbf{x}) - \delta_t L_h.
$$

Therefore,

$$
\begin{aligned}
\left\|\mathbf{z}_{k+1} - \mathbf{x}\right\|^2 \leq\ & (1-\gamma_t\eta_t)\left\|\mathbf{z}_k - \mathbf{x}\right\|^2 - (1-2\gamma_t\eta_t)\left\|\mathbf{z}_{k+1/2} - \mathbf{z}_k\right\|^2 \\
&+ 8\gamma_t^2 \sum_{i\in\mathfrak{I}} L_{h_i}^2 n_i^2 + 4\gamma_t^2(\mathcal{L}_f(\xi_{k+1/2})^2 + \eta_0^2)\left\|\mathbf{z}_{k+1/2} - \mathbf{x}^t\right\|^2 \\
&- 2\gamma_t \langle \hat{V}_{k+1/2}(\mathbf{z}_{k+1/2}) + H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - V(\mathbf{z}_{k+1/2}) \\
&- \nabla h^{\delta_t}(\mathbf{z}_{k+1/2}), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&- 2\gamma_t \langle \left(V(\mathbf{x}^t) + \nabla h^{\delta_t}(\mathbf{x}^t)\right) - \left(\hat{V}_{k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2})\right), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&- 2\gamma_t \langle \varepsilon_V^t(\xi_t^{1:b_t}) + \varepsilon_h^t(\mathbf{W}_t^{1:b_t}), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle \\
&- 2\gamma_t \left(\langle V^{\eta_t}(\mathbf{x}), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle + h(\mathbf{z}_{k+1/2}) - h(\mathbf{x}) \right) + 2\gamma_t\delta_t L_h.
\end{aligned}
$$

which is what has been claimed.

7.1.2 Proof of Theorem 4.1

Let $\{\eta_t\}_{t\in\mathbb{N}}$ be a positive sequence with $\eta_t \downarrow 0$. Let $\mathbf{s}_t \equiv \mathbf{s}(\eta_t)$ denote the corresponding sequence of solutions to $\mathrm{MVI}(V^{\eta_t}, h)$. Set $q_t \equiv 1 - \gamma_t\eta_t \in (0, 1/2)$. Then, iterating the energy inequality established in Lemma 4.3, we have for $\mathbf{x} = \mathbf{s}_t$:

$$
\begin{aligned}
&\left\|\mathbf{z}_K^{(t)} - \mathbf{s}_t\right\|^2 \leq q_t^K \left\|\mathbf{z}_0^{(t)} - \mathbf{s}_t\right\|^2 \\
&+ \sum_{k=0}^{K-1} q_t^{K-k+1}\gamma_t^2 \left(8\left(\sum_{i\in\mathfrak{I}} L_{h_i}^2 n_i^2\right) + 4(\mathcal{L}_f(\xi_{t,k+1/2})^2 + \eta_0^2)\left\|\mathbf{z}_{k+1/2}^{(t)} - \mathbf{x}^t\right\|^2 \right) \\
&- 2\gamma_t \sum_{k=0}^{K-1} q_t^{K-k+1} \langle \hat{V}_{t,k+1/2}(\mathbf{z}_{k+1/2}^{(t)}) + H^{\delta_t}_{\mathbf{z}_{k+1/2}^{(t)}}(\mathbf{W}_{t,k+1/2}) - V(\mathbf{z}_{k+1/2}^{(t)}) \\
&- \nabla h^{\delta_t}(\mathbf{z}_{k+1/2}^{(t)}), \mathbf{z}_{k+1/2}^{(t)} - \mathbf{s}_t\rangle \\
&- 2\gamma_t \sum_{k=0}^{K-1} q_t^{K-k+1} \langle \left(V(\mathbf{x}^t) + \nabla h^{\delta_t}(\mathbf{x}^t)\right) - \left(\hat{V}_{t,k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{t,k+1/2})\right), \mathbf{z}_{k+1/2}^{(t)} - \mathbf{s}_t\rangle \\
&- 2\gamma_t \sum_{k=0}^{K-1} q_t^{K-k+1} \langle \varepsilon_V^t(\xi_t^{1:b_t}) + \varepsilon_h^t(\mathbf{W}_t^{1:b_t}), \mathbf{z}_{k+1/2}^{(t)} - \mathbf{s}_t\rangle \\
&- 2\gamma_t \sum_{k=0}^{K-1} q_t^{K-k+1} \left(\langle V^{\eta_t}(\mathbf{s}_t), \mathbf{z}_{k+1/2}^{(t)} - \mathbf{s}_t\rangle + h(\mathbf{z}_{k+1/2}^{(t)}) - h(\mathbf{s}_t) - 2\delta_t L_h \right).
\end{aligned}
$$

By definition of the point $\mathbf{s}_t$, we have $\langle V^{\eta_t}(\mathbf{s}_t), \mathbf{z}^{(t)}_{k+1/2} - \mathbf{s}_t\rangle + h(\mathbf{z}^{(t)}_{k+1/2}) - h(\mathbf{s}_t) \geq 0$. Furthermore, the estimators involved are unbiased, which means

$$\begin{aligned}
&\mathbb{E}[\langle \hat{V}_{t,k+1/2}(\mathbf{z}^{(t)}_{k+1/2}) \\
&\quad + H^{\delta_t}_{\mathbf{z}^{(t)}_{k+1/2}}(\mathbf{W}_{t,k+1/2}) - V(\mathbf{z}^{(t)}_{k+1/2}) - \nabla h^{\delta_t}(\mathbf{z}^{(t)}_{k+1/2}), \mathbf{z}^{(t)}_{k+1/2} - \mathbf{s}_t\rangle|\mathcal{F}_t] = 0,\\
&\mathbb{E}[\big(\langle V(\mathbf{x}^t) + \nabla h^{\delta_t}(\mathbf{x}^t)\big) - \Big(\hat{V}_{t,k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{t,k+1/2})\Big),\\
&\quad \mathbf{z}^{(t)}_{k+1/2} - \mathbf{s}_t\rangle|\mathcal{F}_t] = 0,\\
&\mathbb{E}[\langle \varepsilon^t_V(\xi^{1:b_t}_t) + \varepsilon^t_h(\mathbf{W}^{1:b_t}_t), \mathbf{z}^{(t)}_{k+1/2} - \mathbf{s}_t\rangle|\mathcal{F}_t] = 0.
\end{aligned}$$

To wit, let us focus on the first line of the above display. Using the law of iterated expectations, we have

$$\begin{aligned}
&\mathbb{E}[\langle \hat{V}_{t,k+1/2}(\mathbf{z}^{(t)}_{k+1/2}) + H^{\delta_t}_{\mathbf{z}^{(t)}_{k+1/2}}(\mathbf{W}_{t,k+1/2}) - V(\mathbf{z}^{(t)}_{k+1/2}) - \nabla h^{\delta_t}(\mathbf{z}^{(t)}_{k+1/2}), \mathbf{z}^{(t)}_{k+1/2} - \mathbf{s}_t\rangle|\mathcal{F}_t]\\
&= \mathbb{E}\left[\mathbb{E}\left(\langle \hat{V}_{t,k+1/2}(\mathbf{z}^{(t)}_{k+1/2}) + H^{\delta_t}_{\mathbf{z}^{(t)}_{k+1/2}}(\mathbf{W}_{t,k+1/2}) - V(\mathbf{z}^{(t)}_{k+1/2}) - \nabla h^{\delta_t}(\mathbf{z}^{(t)}_{k+1/2}), \mathbf{z}^{(t)}_{k+1/2} - \mathbf{s}_t\rangle|\mathcal{A}_{t,k}\right)|\mathcal{F}_t\right]\\
&= 0
\end{aligned}$$

This being true because $\mathbf{z}_{k+1/2}$ is $\mathcal{A}_{t,k}$-measurable. The remaining two equalities can be demonstrated in the same way. Since $\mathbf{z}_K = \mathbf{x}^{t+1}$ and $\mathbf{z}_0 = \mathbf{x}^t$, and using the results above, we are left with the estimate

$$\begin{aligned}
\mathbb{E}[\left\|\mathbf{x}^{t+1} - \mathbf{s}_t\right\|^2|\mathcal{F}_t] \leq q^K_t \left\|\mathbf{x}^t - \mathbf{s}_t\right\|^2 + K\gamma^2_t(L^2_f + \eta^2_0)C^2 + 8\gamma^2_t\left(\sum_{i\in\mathcal{I}} L^2_{h_i}n^2_i\right)\\
+2K\gamma_t\delta_t L_h.
\end{aligned}$$

From Proposition 3.1(d), we obtain the estimate

$$\begin{aligned}
\left\|\mathbf{x}^t - \mathbf{s}_t\right\|^2 &\leq (1+\gamma_t\eta_t)\left\|\mathbf{x}^t - \mathbf{s}_{t-1}\right\|^2 + (1+\frac{1}{\gamma_t\eta_t})\left\|\mathbf{s}_t - \mathbf{s}_{t-1}\right\|^2\\
&\leq (1+\gamma_t\eta_t)\left\|\mathbf{x}^t - \mathbf{s}_{t-1}\right\|^2 + (1+\gamma_t\eta_t)\left(\frac{\eta_t - \eta_{t-1}}{\eta_t}\right)^2\mathbf{A}^2_x,
\end{aligned}$$

where $\mathbf{A}_x$ is a constant upper bound of $\inf_{\mathbf{x}\in\mathrm{SOL}(V,h)}\|\mathbf{x}\|$ (cf. Proposition A.4). Moreover,

$$(1+\gamma_t\eta_t)q^K_t = (1+\gamma_t\eta_t)(1-\gamma_t\eta_t)q^{K-1}_t = (1-\gamma^2_t\eta^2_t)q^{K-1}_t < q_t.$$

This allows us to conclude

$$\mathbb{E}[\left\|\mathbf{x}^{t+1} - \mathbf{s}_t\right\|^2|\mathcal{F}_t] \leq q_t\left\|\mathbf{x}^t - \mathbf{s}_{t-1}\right\|^2 + q^K_t(1+\frac{1}{\gamma_t\eta_t})\mathbf{A}^2_x\left(\frac{\eta_t - \eta_{t-1}}{\eta_t}\right)^2$$

$$+\ K\gamma_t^2(L_f^2+\eta_0^2)C^2+8\sum_{i\in\mathcal{J}}L_{h_i}^2n_i^2+2KL_h\gamma_t\delta_t.$$

Define

$$\begin{aligned}\mathtt{a}_t\ &\triangleq q_t^K\left(1+\frac{1}{\gamma_t\eta_t}\right)\mathbf{A}_x^2\left(\frac{\eta_t-\eta_{t-1}}{\eta_t}\right)^2\\&+K\gamma_t^2(L_f^2+\eta_0^2)C^2+8\gamma_t^2\sum_{i\in\mathcal{J}}L_{h_i}^2n_i^2+2KL_h\gamma_t\delta_t,\end{aligned}$$

and $\psi_t\triangleq\mathbb{E}[\left\|\mathbf{x}^t-\mathbf{s}_{t-1}\right\|^2]$, so that we obtain the recursion

$$\psi_{t+1}\leq q_t\psi_t+\mathtt{a}_t.$$

Under the assumptions stated in Theorem 4.1, we have $\sum_{t=0}^{\infty}\gamma_t\eta_t=\infty$ and $\lim_{t\to\infty}\frac{\mathtt{a}_t}{\gamma_t\eta_t}=0$. Using Lemma 3 in [60], it follows $\lim_{t\to\infty}\psi_t=0$, and therefore $\lim_{t\to\infty}\left\|\mathbf{x}^t-\mathbf{s}_t\right\|=0$ almost surely. Now, let $\mathbf{w}\triangleq\inf_{\mathbf{x}\in\mathrm{SOL}(V,h)}\|\mathbf{x}\|$. Then, using the triangle inequality we conclude

$$\left\|\mathbf{x}^{t+1}-\mathbf{w}\right\|\leq\left\|\mathbf{x}^{t+1}-\mathbf{s}_t\right\|+\|\mathbf{s}_t-\mathbf{w}\|\to 0\quad\text{as } t\to\infty,\ \text{a.s.}$$

Since $\mathbf{y}(\cdot)$ is Lipschitz continuous (cf. Fact A.2, Appendix A.2), the claim follows.

7.1.3 Proof of Theorem 4.2

We use the energy estimate formulated in Lemma 4.3 to deduce a bound on the gap function (12) relative to a suitably constructed ergodic average.

Lemma 7.1 *For any* $t\in\{0,1,\ldots,T-1\}$, *define* $\bar{\mathbf{z}}^t\triangleq\frac{1}{K}\sum_{k=0}^{K-1}\mathbf{z}_{k+1/2}^{(t)}$. *Let* $\{\gamma_t\}_t$, $\{\eta_t\}_t$ *be positive sequences satisfying* $0<\gamma_t\eta_t<1/2$. *For* $k\in\{0,1,\ldots,K-1\}$ *define*

$$\mathbf{Y}_{t,k}^1\triangleq\hat{V}_{t,k+1/2}(\mathbf{z}_{k+1/2}^{(t)})+H_{\mathbf{z}_{k+1/2}^{(t)}}^{\delta_t}(\mathbf{W}_{t,k+1/2})-V(\mathbf{z}_{k+1/2}^{(t)})-\nabla h^{\delta_t}(\mathbf{z}_{k+1/2}^{(t)}),\ \textit{and}$$

$$\mathbf{Y}_{t,k}^2\triangleq V(\mathbf{x}^t)+\nabla h^{\delta_t}(\mathbf{x}^t)-\hat{V}_{t,k+1/2}(\mathbf{x}^t)-H_{\mathbf{x}^t}^{\delta_t}(\mathbf{W}_{t,k+1/2}).$$

Under the same Assumptions as in Lemma 4.3, we have for all $\mathbf{x}\in\mathcal{X}$:

$$\begin{aligned}\gamma_t\left(\langle V(\mathbf{x}),\bar{\mathbf{z}}^t-\mathbf{x}\rangle+h(\bar{\mathbf{z}}^t)-h(\mathbf{x})\right)&\leq\frac{1}{2K}\left(\left\|\mathbf{x}^t-\mathbf{x}\right\|^2-\left\|\mathbf{x}^{t+1}-\mathbf{x}\right\|^2\right)\\&+2\frac{\gamma_t^2}{K}(\mathcal{L}_f(\xi_{t,k+1/2})^2+\eta_0^2)\sum_{k=0}^{K-1}\left\|\mathbf{z}_{k+1/2}^{(t)}-\mathbf{x}\right\|^2+4\gamma_t^2\sum_{i\in\mathcal{J}}L_{h_i}^2n_i^2\\&-\gamma_t\langle\varepsilon_V^t(\xi_t^{1:b_t})+\varepsilon_h^t(\mathbf{W}_t^{1:b_t}),\bar{\mathbf{z}}^t-\mathbf{x}\rangle+\gamma_t\delta_tL_h-\gamma_t\eta_t\langle\mathbf{x},\bar{\mathbf{z}}^t-\mathbf{x}\rangle\\&-\frac{\gamma_t}{K}\sum_{k=0}^{K-1}\langle\mathbf{Y}_{t,k}^1,\mathbf{z}_{k+1/2}^{(t)}-\mathbf{x}\rangle-\frac{\gamma_t}{K}\sum_{k=0}^{K-1}\langle\mathbf{Y}_{t,k}^2,\mathbf{z}_{k+1/2}^{(t)}-\mathbf{x}\rangle.\end{aligned}\tag{29}$$

Proof We depart from the energy bound in Lemma 4.3. Rearranging this inequality and using $\gamma_t\eta_t \in (0, 1/2)$, it follows

$$
\begin{aligned}
&2\gamma_t \left(\langle V(\mathbf{x}), \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x}\rangle + h(\mathbf{z}^{(t)}_{k+1/2}) - h(\mathbf{x})\right) \leq \left\|\mathbf{z}^{(t)}_{k} - \mathbf{x}\right\|^2 - \left\|\mathbf{z}^{(t)}_{k+1} - \mathbf{x}\right\|^2 \\
&\quad + 8\gamma_t^2 \sum_{i\in\mathfrak{I}} L^2_{h_i} n_i^2 + 4\gamma_t^2(\mathcal{L}_f(\xi_{t,k+1/2})^2 + \eta_0^2) \left\|\mathbf{z}^{(t)}_{k+1/2} - \mathbf{x}^t\right\|^2 \\
&\quad - 2\gamma_t \langle \hat{V}_{t,k+1/2}(\mathbf{z}^{(t)}_{k+1/2}) + H^{\delta_t}_{\mathbf{z}^{(t)}_{k+1/2}}(\mathbf{W}_{k+1/2}) - V(\mathbf{z}^{(t)}_{k+1/2}) - \nabla h^{\delta_t}(\mathbf{z}^{(t)}_{k+1/2}), \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x}\rangle \\
&\quad - 2\gamma_t \langle \left(V(\mathbf{x}^t) + \nabla h^{\delta_t}(\mathbf{x}^t)\right) - \left(\hat{V}_{t,k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{t,k+1/2})\right), \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x}\rangle \\
&\quad - 2\gamma_t \langle \varepsilon^t_V(\xi^{1:b_t}_t) + \varepsilon^t_h(\mathbf{W}^{1:b_t}_t), \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x}\rangle + 2\gamma_t\delta_t L_h - 2\gamma_t\eta_t\langle \mathbf{x}, \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x}\rangle.
\end{aligned}
$$

Summing from $k = 0, \ldots, K-1$ and calling $\bar{\mathbf{z}}^t \triangleq \frac{1}{K}\sum_{k=0}^{K-1} \mathbf{z}^{(t)}_{k+1/2}$, we get first from Jensen's inequality

$$
\begin{aligned}
&2\frac{\gamma_t}{K}\sum_{k=0}^{K-1} \left(\langle V(\mathbf{x}), \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x}\rangle + h(\mathbf{z}^{(t)}_{k+1/2}) - h(\mathbf{x})\right) \\
&\quad \geq 2\gamma_t \left(\langle V(\mathbf{x}), \bar{\mathbf{z}}^t - \mathbf{x}\rangle + h(\bar{\mathbf{z}}^t) - h(\mathbf{x})\right).
\end{aligned}
$$

Second, telescoping the expression in the penultimate display and using the definitions of the process $\{\mathbf{Y}^{\nu}_{t,k}\}_{k=0}^{K-1}$, we deduce the bound

$$
\begin{aligned}
&\gamma_t \left(\langle V(\mathbf{x}), \bar{\mathbf{z}}^t - \mathbf{x}\rangle + h(\bar{\mathbf{z}}^t) - h(\mathbf{x})\right) \\
&\quad \leq \frac{1}{2K}\left(\left\|\mathbf{x}^t - \mathbf{x}\right\|^2 - \left\|\mathbf{x}^{t+1} - \mathbf{x}\right\|^2\right) \\
&\qquad + 2\frac{\gamma_t^2}{K}\sum_{k=0}^{K-1}(\mathcal{L}_f(\xi_{t,k+1/2})^2 + \eta_0^2)\left\|\mathbf{z}_{k+1/2} - \mathbf{x}\right\|^2 + 4\gamma_t^2\sum_{i\in\mathfrak{I}} L^2_{h_i} n_i^2 \\
&\qquad - \gamma_t\langle \varepsilon^t_V(\xi^{1:b_t}_t) + \varepsilon^t_h(\mathbf{W}^{1:b_t}_t), \bar{\mathbf{z}}^t - \mathbf{x}\rangle + \gamma_t\delta_t L_h - \gamma_t\eta_t\langle \mathbf{x}, \bar{\mathbf{z}}^t - \mathbf{x}\rangle \\
&\qquad - \frac{\gamma_t}{K}\sum_{k=0}^{K-1}\langle \mathbf{Y}^1_{t,k}, \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x}\rangle - \frac{\gamma_t}{K}\sum_{k=0}^{K-1}\langle \mathbf{Y}^2_{t,k}, \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x}\rangle.
\end{aligned}
$$

□

We can now give the proof of Theorem 4.2. Let us introduce the auxiliary processes $\{\mathbf{u}^{\nu}_{t,k}\}_{k=0}^{K-1}$ for $\nu = 1, 2$ defined recursively as

$$
\mathbf{u}^{\nu}_{t,k+1} = \Pi_{\mathcal{X}}(\mathbf{u}^{\nu}_{t,k} - \gamma_t\mathbf{Y}^{\nu}_{t,k}), \quad \mathbf{u}^{\nu}_{t,0} = \mathbf{x}^t. \tag{30}
$$

The definition of the auxiliary sequence gives for $\nu = 1, 2$ (see e.g. [54])

$$
\left\|\mathbf{u}^{\nu}_{t,k+1} - \mathbf{x}\right\|^2 \leq \left\|\mathbf{u}^{\nu}_{t,k} - \mathbf{x}\right\|^2 - 2\gamma_t\langle \mathbf{Y}^{\nu}_{t,k}, \mathbf{u}^{\nu}_{t,k} - \mathbf{x}\rangle + \gamma_t^2\left\|\mathbf{Y}^{\nu}_{t,k}\right\|^2
$$

$$= \left\|\mathbf{u}^{\nu}_{t,k} - \mathbf{x}\right\|^2 + 2\gamma_t \langle \mathbf{Y}^{\nu}_{t,k}, \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x}\rangle$$
$$- 2\gamma_t \langle \mathbf{Y}^{\nu}_{t,k}, \mathbf{u}^{\nu}_{t,k} - \mathbf{z}^{(t)}_{k+1/2}\rangle + \gamma_t^2 \left\|\mathbf{Y}^{\nu}_{t,k}\right\|^2.$$

Rearranging and telescoping shows

$$-2\gamma_t \sum_{k=0}^{K-1} \langle \mathbf{Y}^{\nu}_{t,k}, \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x}\rangle \leq \left\|\mathbf{u}^{\nu}_{t,0} - \mathbf{x}\right\|^2 - \left\|\mathbf{u}^{\nu}_{t,K} - \mathbf{x}\right\|^2 + \gamma_t^2 \sum_{k=0}^{K-1} \left\|\mathbf{Y}^{\nu}_{t,k}\right\|^2$$
$$- 2\gamma_t \sum_{k=0}^{K-1} \langle \mathbf{Y}^{\nu}_{t,k}, \mathbf{z}^{(t)}_{k+1/2} - \mathbf{u}^{\nu}_{t,k}\rangle.$$

Plugging this into eq. (29), we get

$$\gamma_t \left(\langle V(\mathbf{x}), \bar{\mathbf{z}}^t - \mathbf{x}\rangle + h(\bar{\mathbf{z}}^t) - h(\mathbf{x})\right) \leq \frac{1}{2K}\left(\left\|\mathbf{x}^t - \mathbf{x}\right\|^2 - \left\|\mathbf{x}^{t+1} - \mathbf{x}\right\|^2\right)$$
$$+ 2\frac{\gamma_t^2}{K} \sum_{k=0}^{K-1} (\mathcal{L}_f(\xi_{t,k+1/2})^2 + \eta_0^2) \left\|\mathbf{z}^{(t)}_{k+1/2} - \mathbf{x}\right\|^2 + 4\gamma_t^2 \sum_{i\in\mathcal{I}} L_{h_i}^2 n_i^2$$
$$- \gamma_t \langle \varepsilon^t_V(\xi_t^{1:b_t}) + \varepsilon^t_h(\mathbf{W}_t^{1:b_t}), \bar{\mathbf{z}}^t - \mathbf{x}\rangle + \gamma_t \delta_t L_h - \gamma_t \eta_t \langle \mathbf{x}, \bar{\mathbf{z}}^t - \mathbf{x}\rangle$$
$$+ \frac{\gamma_t}{K}\left(\left\|\mathbf{u}^1_{t,0} - \mathbf{x}\right\|^2 - \left\|\mathbf{u}^1_{t,K} - \mathbf{x}\right\|^2 + \gamma_t^2 \sum_{k=0}^{K-1}\left\|\mathbf{Y}^1_{t,k}\right\|^2 - 2\gamma_t \sum_{k=0}^{K-1} \langle \mathbf{Y}^1_{t,k}, \mathbf{z}^{(t)}_{k+1/2} - \mathbf{u}^1_{t,k}\rangle\right)$$
$$+ \frac{\gamma_t}{K}\left(\left\|\mathbf{u}^2_{t,0} - \mathbf{x}\right\|^2 - \left\|\mathbf{u}^2_{t,K} - \mathbf{x}\right\|^2 + \gamma_t^2 \sum_{k=0}^{K-1}\left\|\mathbf{Y}^2_{t,k}\right\|^2 - 2\gamma_t \sum_{k=0}^{K-1} \langle \mathbf{Y}^2_{t,k}, \mathbf{z}^{(t)}_{k+1/2} - \mathbf{u}^2_{t,k}\rangle\right).$$

Summing this expression over the outer-iteration loop and introduce the averaged iterate

$$\bar{\mathbf{z}}^T \triangleq \frac{\sum_{t=0}^{T-1} \gamma_t \bar{\mathbf{z}}^t}{\sum_{t=0}^{T-1} \gamma_t}.$$

Jensen's inequality readily implies

$$\sum_{t=0}^{T-1} \left(\langle V(\mathbf{x}), \bar{\mathbf{z}}^t - \mathbf{x}\rangle + h(\bar{\mathbf{z}}^t) - h(\mathbf{x})\right) \geq \left(\sum_{t=0}^{T-1} \gamma_t\right)\left(\langle V(\mathbf{x}), \bar{\mathbf{z}}^T - \mathbf{x}\rangle + h(\bar{\mathbf{z}}^T) - h(\mathbf{x})\right).$$

Recall that $C = \max_{i\in\mathcal{I}} C_i$ is the upper bound on the diameter of the set $\mathcal{X}_i$ (cf. Assumption 2.(i)). This assumed compactness of the set $\mathcal{X}$, we derive the a-priori bounds

$$\left\|\mathbf{x}^t - \mathbf{x}\right\| \leq C \qquad \forall t = 0, 1, \ldots, T-1,$$
$$\left\|\mathbf{z}^{(t)}_{k+1/2} - \mathbf{x}\right\| \leq C \qquad \forall k = 0, 1, \ldots, K-1, \text{ and } \left\|\bar{\mathbf{z}}^t - \mathbf{x}\right\| \leq C \qquad \forall t = 0, 1, \ldots, T-1.$$

Using these bounds, we conclude

$$
\begin{aligned}
&\langle V(\mathbf{x}), \bar{\mathbf{z}}^T - \mathbf{x}\rangle + h(\bar{\mathbf{z}}^T) - h(\mathbf{x}) \\
&\quad \le \frac{1}{2K\sum_{t=0}^{T-1}\gamma_t}\left(\left\|\mathbf{x}^0 - \mathbf{x}\right\|^2 - \left\|\mathbf{x}^T - \mathbf{x}\right\|^2\right) \\
&\qquad + \frac{2\sum_{t=0}^{T-1}\gamma_t^2\sum_{k=0}^{K-1}(\mathcal{L}_f(\xi_{t,k+1/2})^2 + \eta_0^2)\left\|\mathbf{z}_{k+1/2}^{(t)} - \mathbf{x}\right\|^2}{K\sum_{t=0}^{T-1}\gamma_t} \\
&\qquad + \frac{\sum_{t=0}^{T-1}4\gamma_t^2\sum_{i\in\mathcal{I}}L_{h_i}^2 n_i^2 + L_h\sum_{t=0}^{T-1}\delta_t\gamma_t}{\sum_{t=0}^{T-1}\gamma_t} \\
&\qquad + \frac{1}{K\sum_{s=0}^{T-1}\gamma_s}\sum_{t=0}^{T-1}\gamma_t\left(\left\|\mathbf{u}_{t,0}^1 - \mathbf{x}\right\|^2 + \left\|\mathbf{u}_{t,0}^2 - \mathbf{x}\right\|^2\right) \\
&\qquad + \frac{\sum_{t=0}^{T-1}\gamma_t^3}{K\sum_{s=0}^{T-1}\gamma_s}\sum_{k=0}^{K-1}\left(\left\|\mathbf{Y}_{t,k}^1\right\|^2 + \left\|\mathbf{Y}_{t,k}^2\right\|^2\right) \\
&\qquad + \frac{2\sum_{t=0}^{T-1}\gamma_t^2}{K\sum_{s=0}^{T-1}\gamma_s}\sum_{k=0}^{K-1}\Big(\langle\mathbf{Y}_{t,k}^1, \mathbf{u}_{t,k}^1 - \mathbf{z}_{k+1/2}^{(t)}\rangle \\
&\qquad + \langle\mathbf{Y}_{t,k}^2, \mathbf{u}_{t,k}^2 - \mathbf{z}_{k+1/2}^{(t)}\rangle\Big) \\
&\qquad + \frac{\sum_{t=0}^{T-1}\gamma_t\langle\varepsilon_V^t(\xi_t^{1:b_t}) + \varepsilon_h^t(\mathbf{W}_t^{1:b_t}), \mathbf{x} - \bar{\mathbf{z}}^t\rangle}{\sum_{t=0}^{T-1}\gamma_t} \\
&\qquad + \frac{\sum_{t=0}^{T}\gamma_t\eta_t\langle\mathbf{x}, \mathbf{x} - \bar{\mathbf{z}}^t\rangle}{\sum_{t=0}^{T-1}\gamma_t}.
\end{aligned}
$$

We bound each of the terms above individually as follows:

1. $\frac{1}{2K\sum_{t=0}^{T-1}\gamma_t}\left(\left\|\mathbf{x}^0 - \mathbf{x}\right\|^2 - \left\|\mathbf{x}^T - \mathbf{x}\right\|^2\right) \le \frac{C^2}{2K\sum_{t=0}^{T-1}\gamma_t}$,
2. $\frac{2\sum_{t=0}^{T-1}\gamma_t^2\sum_{k=0}^{K-1}(\mathcal{L}_f(\xi_{t,k+1/2})^2+\eta_0^2)\left\|\mathbf{z}_{k+1/2}^{(t)}-\mathbf{x}\right\|^2}{K\sum_{t=0}^{T-1}\gamma_t} \le 2C^2\frac{\sum_{t=0}^{T-1}\gamma_t^2\sum_{k=0}^{K-1}(\mathcal{L}_f(\xi_{t,k+1/2})^2+\eta_0^2)}{K\sum_{s=0}^{T-1}\gamma_s}$,
3. $\frac{1}{K}\left(\left\|\mathbf{u}_{t,0}^1 - \mathbf{x}\right\|^2 + \left\|\mathbf{u}_{t,0}^2 - \mathbf{x}\right\|^2\right) = \frac{2}{K}\left\|\mathbf{x}^t - \mathbf{x}\right\|^2 \le \frac{2C^2}{K}$,
4. $\frac{\sum_{t=0}^{T}\gamma_t\eta_t\langle\mathbf{x},\mathbf{x}-\bar{\mathbf{z}}^t\rangle}{\sum_{t=0}^{T-1}\gamma_t} \le \frac{3C^2\sum_{t=0}^{T-1}\gamma_t\eta_t}{2\sum_{t=0}^{T-1}\gamma_t}$,
5. $\frac{\sum_{t=0}^{T-1}\gamma_t\langle\varepsilon_V^t(\xi_t^{1:b_t})+\varepsilon_h^t(\mathbf{W}_t^{1:b_t}),\mathbf{x}-\bar{\mathbf{z}}^t\rangle}{\sum_{t=0}^{T-1}\gamma_t} \le \frac{\sum_{t=0}^{T-1}\gamma_t(C\left\|\varepsilon_V^t(\xi_t^{1:b_t})\right\|+C\left\|\varepsilon_h^t(\mathbf{W}_t^{1:b_t})\right\|)}{\sum_{t=0}^{T-1}\gamma_t}$

Plugging all these bounds into the penultimate display gives

$$
\begin{aligned}
&\langle V(\mathbf{x}), \bar{\mathbf{z}}^T - \mathbf{x}\rangle + h(\bar{\mathbf{z}}^T) - h(\mathbf{x}) \\
&\quad \le \frac{C^2}{2K\sum_{t=0}^{T-1}\gamma_t} + 2C^2\frac{\sum_{t=0}^{T-1}\gamma_t^2\sum_{k=0}^{K-1}(\mathcal{L}_f(\xi_{t,k+1/2})^2 + \eta_0^2)}{K\sum_{s=0}^{T-1}\gamma_s}
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{\sum_{t=0}^{T-1} 4\gamma_t^2 \sum_{i\in\mathcal{I}} L_{h_i}^2 n_i^2 + L_h \sum_{t=0}^{T-1}\delta_t\gamma_t}{\sum_{t=0}^{T-1}\gamma_t}\\
&+\frac{2C^2}{K}+\frac{\sum_{t=0}^{T-1}\gamma_t(C\left\|\varepsilon_V^t(\xi_t^{1:b_t})\right\|+C\left\|\varepsilon_h^t(\mathbf{W}_t^{1:b_t})\right\|)}{\sum_{t=0}^{T-1}\gamma_t}\\
&+\frac{\sum_{t=0}^{T-1}\gamma_t^3}{K\sum_{t=0}^{T-1}\gamma_t}\sum_{k=0}^{K-1}\left(\left\|\mathbf{Y}_{t,k}^1\right\|^2+\left\|\mathbf{Y}_{t,k}^2\right\|^2\right)\\
&+\frac{2\sum_{t=0}^{T-1}\gamma_t^2}{K\sum_{t=0}^{T-1}\gamma_t}\sum_{k=0}^{K-1}\left(\langle\mathbf{Y}_{t,k}^1,\mathbf{u}_{t,k}^1-\mathbf{z}_{k+1/2}^{(t)}\rangle+\langle\mathbf{Y}_{t,k}^2,\mathbf{u}_{t,k}^2-\mathbf{z}_{k+1/2}\rangle\right)\\
&+\frac{3C^2\sum_{t=0}^{T-1}\gamma_t\eta_t}{2\sum_{t=0}^{T-1}\gamma_t}.
\end{aligned}
$$

Let $L_f^2 \triangleq \mathbb{E}_\xi[\mathcal{L}_f(\xi)^2]$, and define

$$
\begin{aligned}
\mathcal{D}_{K,T} \triangleq \frac{C^2}{2K\sum_{t=0}^{T-1}\gamma_t}+2C^2(L_f^2+\eta_0^2)\frac{\sum_{t=0}^{T-1}\gamma_t^2}{\sum_{t=0}^{T-1}\gamma_t}+\frac{3C^2\sum_{t=0}^{T-1}\gamma_t\eta_t}{2\sum_{t=0}^{T-1}\gamma_t}\\
+\frac{\sum_{t=0}^{T-1} 4\gamma_t^2 \sum_{i\in\mathcal{I}} L_{h_i}^2 n_i^2 + L_h \sum_{t=0}^{T-1}\delta_t\gamma_t}{\sum_{t=0}^{T-1}\gamma_t}+\frac{2C^2}{K}.
\end{aligned}
\tag{31}
$$

Hence, using the definition of the gap function (12), we see

$$
\begin{aligned}
\mathbb{E}[\Gamma(\bar{\mathbf{z}}^T)] \leq \mathcal{D}_{K,T} &+ \mathbb{E}\left[\frac{\sum_{t=0}^{T-1}\gamma_t^3}{K\sum_{t=0}^{T-1}\gamma_t}\sum_{k=0}^{K-1}\left(\left\|\mathbf{Y}_{t,k}^1\right\|^2+\left\|\mathbf{Y}_{t,k}^2\right\|^2\right)\right]\\
&+\mathbb{E}\left[\frac{2\sum_{t=0}^{T-1}\gamma_t^2}{K\sum_{t=0}^{T-1}\gamma_t}\sum_{k=0}^{K-1}\left(\langle\mathbf{Y}_{t,k}^1,\mathbf{u}_{t,k}^1-\mathbf{z}_{k+1/2}^{(t)}\rangle+\langle\mathbf{Y}_{t,k}^2,\mathbf{u}_{t,k}^2-\mathbf{z}_{k+1/2}^{(t)}\rangle\right)\right]\\
&+\mathbb{E}\left[\frac{\sum_{t=0}^{T-1}\gamma_t(C\left\|\varepsilon_V^t(\xi_t^{1:b_t})\right\|+C\left\|\varepsilon_h^t(\mathbf{W}_t^{1:b_t})\right\|)}{\sum_{t=0}^{T-1}\gamma_t}\right].
\end{aligned}
$$

Next, observe that

$$
\begin{aligned}
&\left\|\mathbf{Y}_{t,k}^1\right\| \leq \left\|\hat{V}_{t,k+1/2}(\mathbf{z}_{k+1/2}^{(t)})-V(\mathbf{z}_{k+1/2})\right\|+\left\|H_{\mathbf{z}_{k+1/2}^{(t)}}^{\delta_t}(\mathbf{W}_{t,k+1/2})-\nabla h^{\delta_t}(\mathbf{z}_{k+1/2}^{(t)})\right\|, \text{ and}\\
&\left\|\mathbf{Y}_{t,k}^2\right\| \leq \left\|\hat{V}_{t,k+1/2}(\mathbf{x}^t)-V(\mathbf{x}^t)\right\|+\left\|H_{\mathbf{x}^t}^{\delta_t}(\mathbf{W}_{t,k+1/2})-\nabla h^{\delta_t}(\mathbf{x}^t)\right\|.
\end{aligned}
$$

Moreover, using compactness of $\mathcal{X}$,

$$
\langle\mathbf{Y}_{t,k}^\nu,\mathbf{u}_{t,k}^\nu-\mathbf{z}_{k+1/2}^{(t)}\rangle \leq C\left\|\mathbf{Y}_{t,k}^\nu\right\| \quad \forall\nu=1,2. \tag{32}
$$

Lemma A.2(c) (for $b = 1$) gives

$$\mathbb{E}_{\mathbf{W}\sim \mathsf{U}(\mathbb{S}_n)}\left[\left\| H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{t,k+1/2}) - \nabla h^{\delta_t}(\mathbf{x}^t)\right\|^2 |\mathcal{A}_{t,k}\right] \leq \sum_{i\in\mathcal{I}} n_i^2 L_{h_i}^2.$$

Lemma 4.1 in turn implies

$$\mathbb{E}_{\xi}\left[\left\| \hat{V}_{t,k+1/2}(\mathbf{z}^{(t)}_{k+1/2}) - V(\mathbf{z}^{(t)}_{k+1/2})\right\|^2 |\mathcal{A}_{t,k}\right] \leq M_V^2.$$

By Jensen's inequality in tandem with Lemmas 4.1 and 4.2, we conclude from (27) and (28)

$$\mathbb{E}\left[\left\|\varepsilon^t_V(\xi_t^{1:b_t})\right\| |\mathcal{F}_t\right] \leq \sqrt{\mathbb{E}\left[\left\|\varepsilon^t_V(\xi_t^{1:b_t})\right\|^2 |\mathcal{F}_t\right]} \leq \frac{M_V}{\sqrt{b_t}}, \text{ and}$$

$$\mathbb{E}\left[\left\|\varepsilon^t_h(\mathbf{W}_t^{1:b_t})\right\| |\mathcal{F}_t\right] \leq \sqrt{\mathbb{E}\left[\left\|\varepsilon^t_h(\mathbf{W}_t^{1:b_t})\right\|^2 |\mathcal{F}_t\right]} \leq \frac{\left(\sum_{i\in\mathcal{I}} L_{h_i}^2 n_i^2\right)^{1/2}}{\sqrt{b_t}}.$$

This implies $\mathbb{E}[\left\|\mathbf{Y}^{\nu}_{t,k}\right\|^2 |\mathcal{A}_{t,k}] \leq 2\,M_V^2 + 2\sum_{i\in\mathcal{I}} L_{h_i}^2 n_i^2 \equiv \sigma^2$ for all $\nu = 1, 2$, as well as

$$\mathbb{E}[\left\|\mathbf{Y}^{\nu}_{t,k}\right\| |\mathcal{A}_{t,k}] \leq \sqrt{\mathbb{E}[\left\|\mathbf{Y}^{\nu}_{t,k}\right\|^2 |\mathcal{A}_{t,k}]} \leq \sigma.$$

We conclude, via a repeated application of the law of iterated expectations, that

$$\begin{aligned}\mathbb{E}[\varGamma(\bar{\mathbf{z}}^T)] \leq \mathcal{D}_{K,T} &+ \frac{2\sigma^2\sum_{t=0}^{T-1}\gamma_t^3}{\sum_{t=0}^{T-1}\gamma_t} + \frac{4C\sigma\sum_{t=0}^{T-1}\gamma_t^2}{\sum_{t=0}^{T-1}\gamma_t}\\ &+ \frac{\sum_{t=0}^{T-1}\frac{\gamma_t}{\sqrt{b_t}}\left(CM_V + C(\sum_{i\in\mathcal{I}} L_{h_i}^2 n_i^2)^{1/2}\right)}{\sum_{t=0}^{T-1}\gamma_t}.\end{aligned}$$

Using the specification $\gamma_t = 1/T = \eta_t = \delta_t$, as well as $K = T$ and $b_t \geq T^2$ gives

$$\mathcal{D}_{K,T} \leq \frac{5C^2}{T} + \frac{2C^2(L_f^2 + 1/T^2)}{T} + \frac{3C^2}{2T} + \frac{4(\sum_{i\in\mathcal{I}} L_{h_i}^2 n_i^2) + L_h}{T} \equiv \mathtt{c}_T.$$

and consequently,

$$\mathbb{E}[\varGamma(\bar{\mathbf{z}}^T)] \leq \mathtt{c}_T + \frac{2\sigma^2}{T^2} + \frac{C\sigma}{T} + \frac{C(M_V + (\sum_{i\in\mathcal{I}} L_{h_i}^2 n_i^2)^{1/2})}{T} = \mathcal{O}(C\sigma/T)$$

Algorithm 3: $\texttt{ISFBF}(\bar{\mathbf{x}}, \bar{\mathbf{v}}, \bar{H}, \gamma, \eta, \delta, \varepsilon, K)$

Result: Iterate $\mathbf{z}_K$
Set $\mathbf{z}_0 = \bar{\mathbf{x}}$;
for $k = 0, 1, \ldots, K-1$ **do**
  Update $\mathbf{z}_{k+1/2} = \Pi_{\mathcal{X}}[\mathbf{z}_k - \gamma(\bar{\mathbf{v}} + \eta\bar{\mathbf{x}} + \bar{H})]$,;
  Obtain $\hat{V}^{\eta}_{k+1/2}(\mathbf{z}_{k+1/2})$ and $\hat{V}^{\eta}_{k+1/2}(\bar{\mathbf{x}})$ as defined in eq. (11);
  Draw iid direction vectors $\mathbf{W}_{k+1/2} = \{\mathbf{W}_{i,k+1/2}\}_{i\in\mathcal{I}}$, with each $\mathbf{W}_{i,k+1/2} \sim \mathsf{U}(\mathbb{S}_i)$. ;
  Obtain $H^{\delta,\varepsilon}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2})$ and $H^{\delta,\varepsilon}_{\bar{\mathbf{x}}}(\mathbf{W}_{k+1/2})$;
  Update
  $\mathbf{z}_{k+1} = \mathbf{z}_{k+1/2} - \gamma\left(\hat{V}^{\eta}_{k+1/2}(\mathbf{z}_{k+1/2}) + H^{\delta,\varepsilon}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - \hat{V}^{\eta}_{k+1/2}(\bar{\mathbf{x}}) - H^{\delta,\varepsilon}_{\bar{\mathbf{x}}}(\mathbf{W}_{k+1/2})\right)$.
end ;

Algorithm 4: Inexact Variance Reduced Hierarchical Game Solver (I-$\texttt{VRHGS}$)

Data: $\mathbf{x}, T, \{\gamma_t\}_{t=0}^T, \{b_t\}_{t=0}^T, \{\eta_t\}_{t=0}^T, \{\varepsilon_t\}_{t=0}^T, \{\delta_t\}_{t=0}^T, K$
Set $\mathbf{x}^0 = \mathbf{x}$.
for $t = 0, 1, \ldots, T-1$ **do**
  For each $i \in \mathcal{I}$ receive the oracle feedback $\bar{V}^t$ defined by $\bar{V}^t_i \triangleq \frac{1}{b_t}\sum_{s=1}^{b_t} \hat{V}_i(\mathbf{x}^t, \xi^{(s)}_{i,t})$. ;
  For each $i \in \mathcal{I}$ construct the estimator $H^{\delta_t,b_t}_{\mathbf{x}^t}$ defined by $H^{\delta_t,\varepsilon_t,b_t}_{i,\mathbf{x}_i} \triangleq \frac{1}{b_t}\sum_{s=1}^{b_t} H^{\delta_t,\varepsilon_t}_{i,\mathbf{x}_i}(\mathbf{W}^{(s)}_{i,t})$. ;
  Update $\mathbf{x}^{t+1} = \texttt{ISFBF}(\mathbf{x}^t, \bar{V}^t, H^{\delta_t,b_t,\varepsilon_t}_{\mathbf{x}^t}, \gamma_t, \eta_t, \delta_t, \varepsilon_t, b_t, K)$
end

7.2 Analysis of the Inexact Scheme

The inexact version of our method $\texttt{VRHGS}$ is obtained by replacing the estimates for the implicit function using the inexact solution map $\mathbf{y}^{\varepsilon}_i$. The precise implementation is summarized in Algorithms 3 and 4.

The proof of Theorem 4.3 is analogous to the one of Theorem 4.2, with the simple modification due to inexact feedback from the follower's problem. We state the main changes here, leaving the straightforward derivations to the reader. We begin with the modified energy inequality, similar to Lemma 4.3. Here, we also follow the same notational simplification by suppressing the outer iteration counter t from the variables.

Lemma 7.2 *Let Assumptions 1–4 hold true. Then, for all $t \in \{0, 1, \ldots, T-1\}$ and all anchor points $\mathbf{x} \in \mathcal{X}$, we have*

$$
\begin{aligned}
\left\|\mathbf{z}_{k+1} - \mathbf{x}\right\|^2 \leq{}& (1-\gamma_t\eta_t)\left\|\mathbf{z}_k - \mathbf{x}\right\|^2 - (1-2\gamma_t\eta_t)\left\|\mathbf{z}_{k+1/2} - \mathbf{z}_k\right\|^2 \\
&+ 8\gamma_t^2\left(\left\|H^{\delta_t,\varepsilon_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2})\right\|^2 + \left\|H^{\delta_t,\varepsilon_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2})\right\|^2\right) \\
&+ 4\gamma_t^2(\mathcal{L}_f(\xi_{k+1/2})^2 + \eta_0^2)\left\|\mathbf{z}_{k+1/2} - \mathbf{x}^t\right\|^2 \\
&- 2\gamma_t\langle \hat{V}_{k+1/2}(\mathbf{z}_{k+1/2}) + H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - V(\mathbf{z}_{k+1/2}) \\
&- \nabla h^{\delta_t}(\mathbf{z}_{k+1/2}), \mathbf{z}_{k+1/2} - \mathbf{x}\rangle
\end{aligned}
$$

$$
\begin{aligned}
&- 2\gamma_t \langle \left(V(\mathbf{x}^t) + \nabla h^{\delta_t}(\mathbf{x}^t) \right) - \left(\hat{V}_{k+1/2}(\mathbf{x}^t) + H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2}) \right), \mathbf{z}_{k+1/2} - \mathbf{x} \rangle \\
&- 2\gamma_t \langle \varepsilon^t_V(\xi^{1:b_t}) + \varepsilon^t_h(\mathbf{W}^{1:b_t}), \mathbf{z}_{k+1/2} - \mathbf{x} \rangle \\
&- 2\gamma_t \langle H^{\delta_t, b_t}_{\mathbf{x}^t} - H^{\delta_t, b_t, \varepsilon_t}_{\mathbf{x}^t} + H^{\delta_t, \varepsilon_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2}) - H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{k+1/2}), \mathbf{z}_{k+1/2} - \mathbf{x} \rangle \\
&- 2\gamma_t \langle H^{\delta_t, \varepsilon_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}) - H^{\delta_t}_{\mathbf{z}_{k+1/2}}(\mathbf{W}_{k+1/2}), \mathbf{z}_{k+1/2} - \mathbf{x} \rangle \\
&- 2\gamma_t \left(\langle V^{\eta_t}(\mathbf{x}), \mathbf{z}_{k+1/2} - \mathbf{x} \rangle + h(\mathbf{z}_{k+1/2}) - h(\mathbf{x}) \right) + 2\gamma_t \delta_t L_h,
\end{aligned}
$$

where $L_h \triangleq \sum_{i \in \mathcal{I}} L_{h_i}$.

The inexact version of Lemma 7.1 reads then as follows.

Lemma 7.3 *For any* $t \in \{0, 1, \ldots, T-1\}$*, define* $\bar{\mathbf{z}}^t \triangleq \frac{1}{K} \sum_{k=0}^{K-1} \mathbf{z}^{(t)}_{k+1/2}$*. Let* $\{\gamma_t\}_t$, $\{\eta_t\}_t$ *be positive sequences satisfying* $0 < \gamma_t \eta_t < 1/2$*. For* $k \in \{0, 1, \ldots, K-1\}$ *define the process* $\{\mathbf{Y}^{\nu}_{t,k}\}_{k=0}^{K-1}$, $\nu \in \{1, 2\}$ *as in Lemma 7.1. Then, we have for all* $\mathbf{x} \in \mathcal{X}$*:*

$$
\begin{aligned}
&\gamma_t \left(\langle V(\mathbf{x}), \bar{\mathbf{z}}^t - \mathbf{x} \rangle + h(\bar{\mathbf{z}}^t) - h(\mathbf{x}) \right) \le \frac{1}{2K} \left(\left\| \mathbf{x}^t - \mathbf{x} \right\|^2 - \left\| \mathbf{x}^{t+1} - \mathbf{x} \right\|^2 \right) \\
&\quad + \frac{2\gamma_t^2}{K} \sum_{k=0}^{K-1} (\mathcal{L}_f(\xi_{t,k+1/2})^2 + \eta_0^2) \left\| \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x} \right\|^2 - \gamma_t \eta_t \langle \mathbf{x}, \bar{\mathbf{z}}^t - \mathbf{x} \rangle \\
&\quad - \gamma_t \langle \varepsilon^t_V(\xi^{1:b_t}_t) + \varepsilon^t_h(\mathbf{W}^{1:b_t}_t), \bar{\mathbf{z}}^t - \mathbf{x} \rangle + \gamma_t \delta_t L_h \\
&\quad - \frac{\gamma_t}{K} \sum_{k=0}^{K-1} \langle \mathbf{Y}^1_{t,k}, \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x} \rangle - \frac{\gamma_t}{K} \sum_{k=0}^{K-1} \langle \mathbf{Y}^2_{t,k}, \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x} \rangle \\
&\quad + \frac{2\gamma_t^2}{K} \sum_{k=0}^{K-1} \left(\left\| H^{\delta_t, \varepsilon_t}_{\mathbf{z}^{(t)}_{k+1/2}}(\mathbf{W}_{t,k+1/2}) \right\|^2 + \left\| H^{\delta_t, \varepsilon_t}_{\mathbf{x}^t}(\mathbf{W}_{t,k+1/2}) \right\|^2 \right) \\
&\quad - \frac{\gamma_t}{K} \sum_{k=0}^{K-1} \langle H^{\delta_t, b_t}_{\mathbf{x}^t} - H^{\delta_t, b_t, \varepsilon_t}_{\mathbf{x}^t} + H^{\delta_t, \varepsilon_t}_{\mathbf{x}^t}(\mathbf{W}_{t,k+1/2}) - H^{\delta_t}_{\mathbf{x}^t}(\mathbf{W}_{t,k+1/2}), \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x} \rangle \\
&\quad - \frac{\gamma_t}{K} \sum_{k=0}^{K-1} \langle H^{\delta_t, \varepsilon_t}_{\mathbf{z}^{(t)}_{k+1/2}}(\mathbf{W}_{t,k+1/2}) - H^{\delta_t}_{\mathbf{z}^{(t)}_{k+1/2}}(\mathbf{W}_{t,k+1/2}), \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x} \rangle.
\end{aligned}
$$

Using this bound, we conclude in the same way as in the analysis of the exact scheme that

$$
\begin{aligned}
&\langle V(\mathbf{x}), \bar{\mathbf{z}}^T - \mathbf{x} \rangle + h(\bar{\mathbf{z}}^T) - h(\mathbf{x}) \le \frac{1}{2K \sum_{t=0}^{T-1} \gamma_t} \left\| \mathbf{x}^0 - \mathbf{x} \right\|^2 \\
&\quad + \frac{\sum_{t=0}^{T} \gamma_t \eta_t \langle \mathbf{x}, \mathbf{x} - \bar{\mathbf{z}}^t \rangle}{\sum_{t=0}^{T-1} \gamma_t} + \frac{L_h \sum_{t=0}^{T-1} \delta_t \gamma_t}{\sum_{t=0}^{T-1} \gamma_t} \\
&\quad + \frac{2 \sum_{t=0}^{T-1} \gamma_t^2 \sum_{k=0}^{K-1} (\mathcal{L}_f(\xi_{t,k+1/2})^2 + \eta_0^2) \left\| \mathbf{z}^{(t)}_{k+1/2} - \mathbf{x} \right\|^2}{K \sum_{s=0}^{T-1} \gamma_s}
\end{aligned}
$$

$$
\begin{aligned}
&+ \frac{1}{K\sum_{s=0}^{T-1}\gamma_s}\sum_{t=0}^{T-1}\gamma_t\left(\left\|\mathbf{u}_{t,0}^1-\mathbf{x}\right\|^2+\left\|\mathbf{u}_{t,0}^2-\mathbf{x}\right\|^2\right)\\
&+ \frac{\sum_{t=0}^{T-1}\gamma_t^3}{K\sum_{t=0}^{T-1}\gamma_t}\sum_{k=0}^{K-1}\left(\left\|\mathbf{Y}_{t,k}^1\right\|^2+\left\|\mathbf{Y}_{t,k}^2\right\|^2\right)\\
&+ \frac{2\sum_{t=0}^{T-1}\gamma_t^2}{K\sum_{t=0}^{T-1}\gamma_t}\sum_{k=0}^{K-1}\left(\langle \mathbf{Y}_{t,k}^1,\mathbf{u}_{t,k}^1-\mathbf{z}_{k+1/2}^{(t)}\rangle+\langle \mathbf{Y}_{t,k}^2,\mathbf{u}_{t,k}^2-\mathbf{z}_{k+1/2}^{(t)}\rangle\right)\\
&+ \frac{\sum_{t=0}^{T-1}\gamma_t\langle \varepsilon_V^t(\xi_t^{1:b_t})+\varepsilon_h^t(\mathbf{W}_t^{1:b_t}),\mathbf{x}-\bar{\mathbf{z}}^t\rangle}{\sum_{t=0}^{T-1}\gamma_t}\\
&+ \sum_{t=0}^{T-1}\frac{2\gamma_t^2}{K\sum_{s=0}^{T-1}\gamma_s}\sum_{k=0}^{K-1}\left(\left\|H_{\mathbf{z}_{k+1/2}^{(t)}}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})\right\|^2+\left\|H_{\mathbf{x}^t}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})\right\|^2\right)\\
&- \sum_{t=0}^{T-1}\frac{\gamma_t}{K\sum_{s=0}^{T-1}\gamma_s}\sum_{k=0}^{K-1}\langle H_{\mathbf{x}^t}^{\delta_t,b_t}-H_{\mathbf{x}^t}^{\delta_t,b_t,\varepsilon_t}\\
&+ H_{\mathbf{x}^t}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})-H_{\mathbf{x}^t}^{\delta_t}(\mathbf{W}_{t,k+1/2}),\mathbf{z}_{k+1/2}^{(t)}-\mathbf{x}\rangle\\
&- \sum_{t=0}^{T-1}\frac{\gamma_t}{K\sum_{s=0}^{T-1}\gamma_s}\sum_{k=0}^{K-1}\langle H_{\mathbf{z}_{k+1/2}^{(t)}}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})-H_{\mathbf{z}_{k+1/2}^{(t)}}^{\delta_t}(\mathbf{W}_{t,k+1/2}),\mathbf{z}_{k+1/2}^{(t)}-\mathbf{x}\rangle.
\end{aligned}
$$

Performing the same bounding steps as done in the exact case, we readily arrive at the expression

$$
\begin{aligned}
&\langle V(\mathbf{x}),\bar{\mathbf{z}}^T-\mathbf{x}\rangle+h(\bar{\mathbf{z}}^T)-h(\mathbf{x})\le\frac{C^2}{2K\sum_{t=0}^{T-1}\gamma_t}\\
&\quad+2C^2\frac{\sum_{t=0}^{T-1}\gamma_t^2\sum_{k=0}^{K-1}(\mathcal{L}_f(\xi_{t,k+1/2})^2+\eta_0^2)}{K\sum_{t=0}^{T-1}\gamma_t}\\
&\quad+\frac{3C^2\sum_{t=0}^{T-1}\gamma_t\eta_t}{2\sum_{t=0}^{T-1}\gamma_t}+\frac{2C^2}{K}+\frac{L_h\sum_{t=0}^{T-1}\delta_t\gamma_t}{\sum_{t=0}^{T-1}\gamma_t}\\
&\quad+\frac{\sum_{t=0}^{T-1}\gamma_t\left(C\left\|\varepsilon_V^t(\xi_t^{1:b_t})\right\|+C\left\|\varepsilon_h^t(\mathbf{W}_t^{1:b_t})\right\|\right)}{\sum_{t=0}^{T-1}\gamma_t}\\
&\quad+\frac{\sum_{t=0}^{T-1}\gamma_t^3}{K\sum_{t=0}^{T-1}\gamma_t}\sum_{k=0}^{K-1}\left(\left\|\mathbf{Y}_{t,k}^1\right\|^2+\left\|\mathbf{Y}_{t,k}^2\right\|^2\right)\\
&\quad+\frac{2\sum_{t=0}^{T-1}\gamma_t^2}{K\sum_{t=0}^{T-1}\gamma_t}\sum_{k=0}^{K-1}\left(\langle \mathbf{Y}_{t,k}^1,\mathbf{u}_{t,k}^1-\mathbf{z}_{k+1/2}^{(t)}\rangle+\langle \mathbf{Y}_{t,k}^2,\mathbf{u}_{t,k}^2-\mathbf{z}_{k+1/2}^{(t)}\rangle\right)\\
&\quad+\sum_{t=0}^{T-1}\frac{2\gamma_t^2}{K\sum_{s=0}^{T-1}\gamma_s}\sum_{k=0}^{K-1}\left(\left\|H_{\mathbf{z}_{k+1/2}^{(t)}}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})\right\|^2+\left\|H_{\mathbf{x}^t}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})\right\|^2\right)
\end{aligned}
$$

$$+\sum_{t=0}^{T-1}\frac{C\gamma_t}{K\sum_{s=0}^{T-1}\gamma_s}\sum_{k=0}^{K-1}\left\|H_{\mathbf{x}^t}^{\delta_t,b_t}-H_{\mathbf{x}^t}^{\delta_t,b_t,\varepsilon_t}+H_{\mathbf{x}^t}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})-H_{\mathbf{x}^t}^{\delta_t}(\mathbf{W}_{t,k+1/2})\right\|$$
$$+\sum_{t=0}^{T-1}\frac{C\gamma_t}{K\sum_{s=0}^{T-1}\gamma_s}\sum_{k=0}^{K-1}\left\|H_{\mathbf{z}_{k+1/2}}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})-H_{\mathbf{z}_{k+1/2}^{(t)}}^{\delta_t}(\mathbf{W}_{t,k+1/2})\right\|.$$

We next estimate the error terms appearing because of the inexact feedback map in the coupling function. Lemma A.3 yields

$$\mathbb{E}\left[\left\|H_{\mathbf{z}_{k+1/2}}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})\right\|^2+\left\|H_{\mathbf{x}^t}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})\right\|^2|\mathcal{A}_{t,k}\right]$$
$$\leq 6\sum_{i\in\mathcal{I}}\left(\frac{n_i}{\delta_t}\right)^2(2L_{2,i}^2\varepsilon_t^2+L_{1,i}^2\delta_t^2)\equiv\alpha_t^{(1)}.$$

Furthermore, the triangle inequality, Jensen's inequality, and (43) gives

$$\begin{aligned}
&\mathbb{E}\left[\left\|H_{\mathbf{x}^t}^{\delta_t,b_t}-H_{\mathbf{x}^t}^{\delta_t,b_t,\varepsilon_t}+H_{\mathbf{x}^t}^{\delta_t,\varepsilon_t}(\mathbf{W}_{k+1/2})-H_{\mathbf{x}^t}^{\delta_t}(\mathbf{W}_{k+1/2})\right\||\mathcal{A}_{t,k}\right]\\
&\leq\mathbb{E}\left[\left\|H_{\mathbf{x}^t}^{\delta_t,b_t}-H_{\mathbf{x}^t}^{\delta_t,b_t,\varepsilon_t}\right\||\mathcal{A}_{t,k}\right]\\
&\quad+\mathbb{E}\left[\left\|H_{\mathbf{x}^t}^{\delta_t,\varepsilon_t}(\mathbf{W}_{k+1/2})-H_{\mathbf{x}^t}^{\delta_t}(\mathbf{W}_{k+1/2})\right\||\mathcal{A}_{t,k}\right]\\
&\leq\sqrt{\mathbb{E}\left[\left\|H_{\mathbf{x}^t}^{\delta_t,b_t}-H_{\mathbf{x}^t}^{\delta_t,b_t,\varepsilon_t}\right\|^2|\mathcal{A}_{t,k}\right]}+\sqrt{\mathbb{E}\left[\left\|H_{\mathbf{x}^t}^{\delta_t,\varepsilon_t}(\mathbf{W}_{k+1/2})-H_{\mathbf{x}^t}^{\delta_t}(\mathbf{W}_{k+1/2})\right\|^2|\mathcal{A}_{t,k}\right]}\\
&\leq\frac{2\varepsilon_t}{\delta_t}\left(\frac{1}{\sqrt{b_t}}+1\right)\sqrt{\sum_{i\in\mathcal{I}}L_{2,i}^2n_i^2}\equiv\alpha_t^{(2)}.
\end{aligned}$$

Lastly, we bound

$$\mathbb{E}\left[\left\|H_{\mathbf{z}_{k+1/2}}^{\delta_t,\varepsilon_t}(\mathbf{W}_{k+1/2})-H_{\mathbf{z}_{k+1/2}}^{\delta_t}(\mathbf{W}_{k+1/2})\right\||\mathcal{A}_{t,k}\right]\leq\frac{2\varepsilon_t}{\delta_t}\sqrt{\sum_{i\in\mathcal{I}}L_{2,i}^2n_i^2}\equiv\alpha_t^{(3)}.$$

We now set

$$\begin{aligned}
\mathcal{D}_{K,T}\triangleq&\frac{C^2}{2K\sum_{t=0}^{T-1}\gamma_t}+2C^2(L_f^2+\eta_0^2)\frac{\sum_{t=0}^{T-1}\gamma_t^2}{\sum_{t=0}^{T-1}\gamma_t}\\
&+\frac{3C^2\sum_{t=0}^{T-1}\gamma_t\eta_t}{2\sum_{t=0}^{T-1}\gamma_t}+\frac{2C^2}{K}+\frac{L_h\sum_{t=0}^{T-1}\delta_t\gamma_t}{\sum_{t=0}^{T-1}\gamma_t}.
\end{aligned}$$

Using the definition of the gap function (12), we deduce

$$\mathbb{E}[\Gamma(\bar{\mathbf{z}}^T)]\leq\mathcal{D}_{K,T}+\mathbb{E}\left[\frac{\sum_{t=0}^{T-1}\gamma_t^3}{K\sum_{t=0}^{T-1}\gamma_t}\sum_{k=0}^{K-1}\left(\left\|\mathbf{Y}_{t,k}^1\right\|^2+\left\|\mathbf{Y}_{t,k}^2\right\|^2\right)\right]$$

$$
\begin{aligned}
&+\mathbb{E}\left[\frac{2\sum_{t=0}^{T-1}\gamma_t^2}{K\sum_{t=0}^{T-1}\gamma_t}\sum_{k=0}^{K-1}\left(\langle \mathbf{Y}_{t,k}^1, \mathbf{u}_{t,k}^1-\mathbf{z}_{k+1/2}^{(t)}\rangle+\langle \mathbf{Y}_{t,k}^2, \mathbf{u}_{t,k}^2-\mathbf{z}_{k+1/2}^{(t)}\rangle\right)\right]\\
&+\mathbb{E}\left[\frac{\sum_{t=0}^{T-1}\gamma_t(C\left\|\varepsilon_V^t(\xi_t^{1:b_t})\right\|+C\left\|\varepsilon_h^t(\mathbf{W}_t^{1:b_t})\right\|)}{\sum_{t=0}^{T-1}\gamma_t}\right]\\
&+\mathbb{E}\left[\sum_{t=0}^{T-1}\frac{2\gamma_t^2}{K\sum_{s=0}^{T-1}\gamma_s}\sum_{k=0}^{K-1}\left(\left\|H_{\mathbf{z}_{k+1/2}}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})\right\|^2+\left\|H_{\mathbf{x}^t}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})\right\|^2\right)\right]\\
&+\mathbb{E}\left[\sum_{t=0}^{T-1}\frac{C\gamma_t}{K\sum_{s=0}^{T-1}\gamma_s}\sum_{k=0}^{K-1}\left\|H_{\mathbf{x}^t}^{\delta_t,b_t}-H_{\mathbf{x}^t}^{\delta_t,b_t,\varepsilon_t}+H_{\mathbf{x}^t}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})-H_{\mathbf{x}^t}^{\delta_t}(\mathbf{W}_{t,k+1/2})\right\|\right]\\
&+\mathbb{E}\left[\sum_{t=0}^{T-1}\frac{C\gamma_t}{K\sum_{s=0}^{T-1}\gamma_s}\sum_{k=0}^{K-1}\left\|H_{\mathbf{z}_{k+1/2}^{(t)}}^{\delta_t,\varepsilon_t}(\mathbf{W}_{t,k+1/2})-H_{\mathbf{z}_{k+1/2}^{(t)}}^{\delta_t}(\mathbf{W}_{t,k+1/2})\right\|\right]\\
&\leq \mathcal{D}_{K,T}+\frac{2\sigma^2\sum_{t=0}^{T-1}\gamma_t^3}{\sum_{t=0}^{T-1}\gamma_t}+\frac{4C\sigma\sum_{t=0}^{T-1}\gamma_t^2}{\sum_{t=0}^{T-1}\gamma_t}\\
&+\frac{\sum_{t=0}^{T-1}\frac{\gamma_t}{\sqrt{b_t}}C(M_V+(\sum_{i\in\mathcal{I}}L_{h_i}^2n_i^2)^{1/2}}{\sum_{t=0}^{T-1}\gamma_t}\\
&+\sum_{t=0}^{T-1}\frac{2\gamma_t^2\alpha_t^{(1)}}{\sum_{s=0}^{T-1}\gamma_s}+\sum_{t=0}^{T-1}\frac{C\alpha_t^{(2)}\gamma_t}{\sum_{s=0}^{T-1}\gamma_s}+\sum_{t=0}^{T-1}\frac{C\gamma_t\alpha_t^{(3)}}{\sum_{s=0}^{T-1}\gamma_s}.
\end{aligned}
$$

Making the choice $\gamma_t=\eta_t=\delta_t=1/T$ as well as $K=T, b_t\geq T^2$ and $\varepsilon_t=1/T^2$, we see that $\alpha_t^{(1)}=\mathcal{O}(1/T^2), \alpha_t^{(2)}=\mathcal{O}(1/T)$ and $\alpha_t^{(3)}=\mathcal{O}(1/T)$. It follows $\mathbb{E}[\Gamma(\bar{\mathbf{z}}^T)]=\mathcal{O}(C\sigma/T)$, which completes the proof of Theorem 4.3.

8 Conclusion

In this work, we proposed a new solution approach to solve a fairly large class of stochastic hierarchical games. Using a combination of smoothing, zeroth-order gradient approximation, and iterative regularization, we develop a novel variance reduction method for stochastic VIs affected by general stochastic noise. We demonstrate consistency of the method by proving that solution trajectory converges almost surely to a particular equilibrium of the game and derive a $\mathcal{O}(1/T)$ convergence rate in terms of the expected gap function, using a suitably defined averaged trajectory. This rate result is robust to inexact solutions of the lower level problem of the follower and aligns with state-of-the-art variance reduction methods tailored to finite-sum problems. Furthermore, our approach is based on Tseng's splitting technique, which shares the same number of function calls as the popular extragradient method, but saves on one projection step. This implies that our scheme reduces the oracle complexity relative to vanilla mini-batch approaches and at the same time reduces the computational bottlenecks in every single iteration. This leaves open the door for many future investigations, involving bias and non-convexities, that we leave for future research.

A Auxiliary Facts

A.1 Generalities

Given a closed convex set $\mathcal{X} \subset \mathbb{R}^n$, we denote by $\Pi_{\mathcal{X}} : \mathbb{R}^n \to \mathcal{X}$ the orthogonal projector defined as

$$\Pi_{\mathcal{X}}(\mathbf{w}) := \operatorname*{argmin}_{\mathbf{x}\in\mathcal{X}} \frac{1}{2} \|\mathbf{x} - \mathbf{w}\|^2 .$$

This is the solution map of a strongly convex optimization problem with the following well-known properties.

Lemma A.1 *Let $\mathcal{X} \subset \mathbb{R}^n$ be a nonempty closed convex set. Then:*

(i) *$\Pi_{\mathcal{X}}(\mathbf{w})$ is the unique point satisfying $\langle \mathbf{w} - \Pi_{\mathcal{X}}(\mathbf{w}), \mathbf{x} - \Pi_{\mathcal{X}}(\mathbf{w})\rangle \leq 0$ for all $\mathbf{x} \in \mathcal{X}$;*
(ii) *For all $\mathbf{w} \in \mathbb{R}^n$ and $\mathbf{x} \in \mathcal{X}$, we have $\|\Pi_{\mathcal{X}}(\mathbf{w}) - \mathbf{x}\|^2 + \|\Pi_{\mathcal{X}}(\mathbf{w}) - \mathbf{w}\|^2 \leq \|\mathbf{w} - \mathbf{x}\|^2$;*
(iii) *For all $\mathbf{w}, \mathbf{v} \in \mathbb{R}^n$, $\|\Pi_{\mathcal{X}}(\mathbf{w}) - \Pi_{\mathcal{X}}(\mathbf{v})\|^2 \leq \|\mathbf{w} - \mathbf{v}\|^2$;*

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given probability space carrying a filtration $\mathbb{F} = \{\mathcal{F}_k\}_{k\geq 0}$. We call the tuple $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ a discrete stochastic basis. Given a vector space $\mathcal{K} \subseteq \mathbb{R}^n$ with Borel σ-algebra $\mathcal{B}(\mathcal{K})$, a $\mathcal{K}$-valued random variable is a $(\mathcal{F}, \mathcal{B}(\mathcal{K}))$-measurable map $f : \Omega \to \mathcal{K}$; we write $f \in L^0(\Omega, \mathcal{F}, \mathbb{P}; \mathcal{K})$. For every $p \in [1, \infty]$, define the equivalence class of random variables $f \in L^0(\Omega, \mathcal{F}, \mathbb{P}; \mathcal{K})$ with $\mathbb{E}[\|f\|^p]^{1/p} < \infty$ as $f \in L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathcal{K})$. For $f_1, \ldots, f_k \in L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathcal{K})$, we denote the sigma-algebra generated by these random variables by $\sigma(f_1, \ldots, f_k)$. We denote by $\ell^0_+(\mathbb{F})$ the set of non-negative random variables $\{\xi_k\}_{k\geq 0}$ such that for each $k \geq 0$, we have $\xi_k \in L^0(\Omega, \mathcal{F}_k, \mathbb{P}; \mathbb{R}_+)$. For $p \geq 1$, we set

$$\ell^p_+(\mathbb{F}) = \{\{\xi_k\}_{k\geq 0} \in \ell^0_+(\mathbb{F}) | \sum_{k\geq 0} |\xi_i|^p < \infty \quad \mathbb{P} - \text{a.s.}\}.$$

A.2 Variational Inequalities

In this appendix we summarize the essential parts from the theory of finite-dimensional variational inequalities we use in the paper. A complete treatment can be found in [24].

The data of a variational inequality problem consist of mappings $\phi : \mathbb{R}^n \to \mathbb{R}^n$ and $r : \mathbb{R}^n \to (-\infty, +\infty]$ a proper, convex and lower semi-continuous function. Denote by $\operatorname{dom}(r) = \{\mathbf{x} \in \mathbb{R}^n | r(\mathbf{x}) < \infty\}$. The mixed variational inequality problems associated with (ϕ, r) is

$$\text{find } \mathbf{x} \in \mathbb{R}^n \text{ such that } \langle \phi(\mathbf{x}), \mathbf{y} - \mathbf{x}\rangle + r(\mathbf{y}) - r(\mathbf{x}) \geq 0 \qquad \forall \mathbf{y} \in \mathbb{R}^n. \quad (\text{MVI}(\phi, r))$$

When $r = \delta_{\mathcal{K}}$ for a closed convex set $\mathcal{K} \subset \mathbb{R}^n$, the problem MVI$(\phi, r)$ reduces to the classical variational inequality VI$(\phi, \mathcal{K})$:

$$\text{find } \mathbf{x} \in \mathcal{K} \text{ such that } \langle \phi(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq 0 \qquad \forall \mathbf{y} \in \mathcal{K}. \tag{VI(ϕ, $\mathcal{K}$)}$$

We note in passing that if r is proper, convex and lower semi-continuous, then problem MVI(ϕ, r) is equivalent to the generalized equation

$$0 \in \phi(\mathbf{x}) + \partial r(\mathbf{x}), \tag{33}$$

where $\partial r(\mathbf{x}) \triangleq \{p \in \mathbb{R}^n | r(\mathbf{x}') \geq r(\mathbf{x}) + \langle p, \mathbf{x}' - \mathbf{x} \rangle \quad \forall \mathbf{x}' \in \mathbb{R}^n\}$ is the subgradient of r at $\mathbf{x}$.

In order to measure the distance of a candidate point to the solution set we introduce as a merit function for MVI(ϕ, r) the gap function

$$\Gamma_{\mathcal{K}}(\mathbf{x}) \triangleq \sup_{\mathbf{z} \in \mathcal{K}} \left(\langle \phi(\mathbf{z}), \mathbf{x} - \mathbf{z} \rangle + r(\mathbf{x}) - r(\mathbf{z}) \right),$$

where $\mathcal{K} \subset \mathbb{R}^n$ is a compact subset to handle the possibility of unboundedness of dom(r). As proven in [56], this restricted version of the gap function is a valid measure as long as $\mathcal{K}$ contains any solution of MVI(ϕ, r).

For existence and uniqueness questions of variational problems, we usually rely on monotonicity and continuity properties of the map ϕ.

Definition A.1 A mapping $\phi : \mathbb{R}^n \to \mathbb{R}^n$ is said to be μ-monotone if there exists $\mu \geq 0$ such that

$$\langle \phi(\mathbf{x}) - \phi(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \geq \mu \|\mathbf{x} - \mathbf{y}\|^2 \qquad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

A 0-monotone mapping is called monotone.

Fact A.1 (Solution Convexity of Monotone VIs) *Consider the problem* VI$(\phi, \mathcal{K})$, *where* $\phi : \mathbb{R}^n \to \mathbb{R}^n$ *is monotone on* $\mathcal{K} \subset$ dom(ϕ), *and* $\mathcal{K}$ *is a closed convex set. Then, the solution set*

$$\mathcal{S} = \{\mathbf{x} \in \mathcal{K} | \langle \phi(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq 0 \qquad \forall \mathbf{y} \in \mathcal{K}\}$$

is closed and convex. If ϕ *is* μ*-monotone with* $\mu > 0$ *on* $\mathcal{K}$*, then* $\mathcal{S}$ *is a singleton.*

See [21, Theorem 2F.1, 2F.6] for a proof of this Fact. We now extend the scope of variational inequalities and introduce parameters into the problem data. This is effectively the lower level problem solved by the followers in our hierarchical game model. Let $\mathcal{X} \subset \mathbb{R}^n$ be a nonempty compact convex set, and $\mathcal{Y} \subset \mathbb{R}^m$ a closed convex set. The object of study is the parameterized generalized equation

$$0 \in \phi(\mathbf{x}, \mathbf{y}) + \mathsf{NC}_{\mathcal{Y}}(\mathbf{y}) \tag{34}$$

for a given function $\phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ and the normal cone

$$\mathsf{NC}_{\mathcal{Y}}(\mathbf{y}) = \begin{cases} \varnothing & \text{if } \mathbf{y} \notin \mathcal{Y}, \\ \{\xi \in \mathbb{R}^m \,|\, \sup_{\mathbf{z}\in\mathcal{Y}} \langle \xi, \mathbf{z} - \mathbf{y}\rangle \leq 0\} & \text{if } \mathbf{y} \in \mathcal{Y}. \end{cases} \tag{35}$$

Specifically, we are interested in understanding the properties of the solution mapping

$$\mathcal{S}(\mathbf{x}) \triangleq \{\mathbf{y} \in \mathbb{R}^m | 0 \in \phi(\mathbf{x}, \mathbf{y}) + \mathsf{NC}_{\mathcal{Y}}(\mathbf{y})\}. \tag{36}$$

This is a subclass of classical problems, thoroughly summarized in [21], and dating back to the landmark paper [61]. The interested reader can find proofs of the facts stated below, as well as much more information on this topic in these references. We point out that many of the strong assumption made below can be relaxed, at the price of more complicated verification steps. Our aim is to present a simple and not entirely unrealistic set of verifiable conditions under which our model assumptions provably hold; A more general result can be found in [68, Lemma2.2].

Fact A.2 *Consider problem* (34) *with the following assumptions on the problem data:*

- $\mathcal{X} \subset \mathbb{R}^n$ *is compact and convex,*
- $\mathcal{Y} \subset \mathbb{R}^m$ *is closed convex,*
- $\phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ *is strictly differentiable on* $\mathcal{K} \times \mathcal{Y}$*, where* $\mathcal{K}$ *is an open set containing* $\mathcal{X}$*;*
- $\phi(x, \cdot)$ *is strongly monotone for every* $\mathbf{x} \in \mathcal{X}$*.*

Then $\mathcal{S}(\mathbf{x}) = \{\mathbf{y}(\mathbf{x})\}$*, and* $\mathbf{y}(\cdot)$ *is Lipschitz continuous on* $\mathcal{X}$*.*

A.3 Tikhonov Regularization

Tikhonov regularization is a classical method in numerical analysis aiming for introducing additional stability into a computational scheme. Given problem $\text{MVI}(\phi, r)$ and $\eta > 0$, we define the Tikhonov regularized mixed variational inequality problem as $\text{MVI}(\phi^\eta, r)$, in which the operator is defined as $\phi^\eta(\mathbf{x}) \triangleq \phi(\mathbf{x}) + \eta\mathbf{x}$. It is easy to see that if ϕ is 0-monotone, then ϕ^η is η-monotone (i.e. strongly monotone). Hence, for every $\eta > 0$, problem $\text{MVI}(\phi^\eta, r)$ has a unique solution $\mathbf{x}^*(\eta)$. The first result we are going to demonstrate is that the net $\{\mathbf{x}^*(\eta)\}_{\eta\geq 0}$ is bounded.

Proposition A.1 *Consider problem* $\text{MVI}(\phi, r)$ *admitting a nonempty solution set* $\text{SOL}(\phi, r)$*. Then, for all* $\eta > 0$*, we have*

$$\left\|\mathbf{x}^*(\eta)\right\| \leq \inf_{\mathbf{x}\in\text{SOL}(\phi,r)} \|\mathbf{x}\|.$$

Proof Using the characterization of a point $\mathbf{x}^* \in \text{SOL}(\phi, r)$ as a solution of a monotone inclusion, we have

$$-\phi^\eta(\mathbf{x}^*(\eta)) \in \partial r(\mathbf{x}^*(\eta)) \text{ and } -\phi(\mathbf{x}^*) \in \partial r(\mathbf{x}^*).$$

Since ∂r is maximally monotone, it follows

$$\begin{aligned}&\langle \phi^{\eta}(\mathbf{x}^*(\eta)) - \phi(\mathbf{x}^*), \mathbf{x}^* - \mathbf{x}^*(\eta)\rangle \geq 0\\ &\quad \Leftrightarrow \langle \phi(\mathbf{x}^*) - \phi(\mathbf{x}^*(\eta)), \mathbf{x}^* - \mathbf{x}^*(\eta)\rangle \leq \eta\langle \mathbf{x}^*(\eta), \mathbf{x}^* - \mathbf{x}^*(\eta)\rangle.\end{aligned}$$

Since $\phi(\cdot)$ is monotone, it follows $\langle \phi(\mathbf{x}^*) - \phi(\mathbf{x}^*(\eta)), \mathbf{x}^* - \mathbf{x}^*(\eta)\rangle \geq 0$, so that

$$0 \leq \langle \mathbf{x}^*(\eta), \mathbf{x}^* - \mathbf{x}^*(\eta)\rangle = \langle \mathbf{x}^*(\eta), \mathbf{x}^*\rangle - \left\|\mathbf{x}^*(\eta)\right\|^2.$$

The Cauchy–Schwarz inequality implies $\|\mathbf{x}^*(\eta)\| \leq \|\mathbf{x}^*\|$. Since $\mathbf{x}^*$ has been chosen arbitrarily, the claim follows. □

We next study the asymptotic regime in which $\eta \to 0^+$. Since the net $\{\mathbf{x}^*(\eta)\}_{\eta>0}$ is bounded, the Bolzano–Weierstrass theorem guarantees the existence of a converging subsequence $\eta_t \to 0$ such that $\mathbf{x}^*(t) \equiv \mathbf{x}^*(\eta_t) \to \hat{\mathbf{x}}$. Since ∂r is maximally monotone, the set $\text{graph}(\partial r)$ is closed in the product topology [6]. Hence, $\hat{\mathbf{x}} \in \text{dom}(\partial r)$. Moreover, for all t

$$(\mathbf{x}^*(t), -\phi(\mathbf{x}^*(t)) - \eta_t \mathbf{x}^*(t)) \in \text{graph}(\partial r) \qquad \forall t > 0.$$

Continuity and Proposition A.1, together with the just mentioned closed graph property, yields for $t \to \infty$,

$$(\hat{\mathbf{x}}, -\phi(\hat{\mathbf{x}})) \in \text{graph}(\partial r) \qquad \forall t > 0.$$

The next claim follows.

Proposition A.2 *Every accumulation point of the Tikhonov sequence* $\{\mathbf{x}^*(\eta)\}_{\eta>0}$ *defines a solution of the problem* $\text{MVI}(\phi, r)$.

We next deduce a non-asymptotic estimate of the Tikhonov sequence. Let $\{\eta_t\}_{t\in\mathbb{N}}$ be a positive sequence of regularization parameters satisfying $\eta_t \downarrow 0$. Exploiting again the variational characterization of the unique solutions $\mathbf{x}_t^* \equiv \mathbf{x}^*(\eta_t)$, we have

$$-\phi^{\eta_{t-1}}(\mathbf{x}_{t-1}^*) \in \partial r(\mathbf{x}_{t-1}^*) \text{ and } -\phi^{\eta_t}(\mathbf{x}_t^*) \in \partial r(\mathbf{x}_t^*).$$

By monotonicity of ∂r, we obtain

$$\langle \phi^{\eta_{t-1}}(\mathbf{x}_{t-1}^*) - \phi^{\eta_t}(\mathbf{x}_t^*), \mathbf{x}_t^* - \mathbf{x}_{t-1}^*\rangle \geq 0.$$

Hence, by monotonicity of ϕ, it follows

$$\begin{aligned}0 &\geq \langle \phi(\mathbf{x}_{t-1}^*) - \phi(\mathbf{x}_t^*), \mathbf{x}_t^* - \mathbf{x}_{t-1}^*\rangle \geq \langle \eta_t \mathbf{x}_t^* - \eta_{t-1}\mathbf{x}_{t-1}^*, \mathbf{x}_t^* - \mathbf{x}_{t-1}^*\rangle\\ &= \eta_t\langle \mathbf{x}_t^* - \mathbf{x}_{t-1}^*, \mathbf{x}_t^* - \mathbf{x}_{t-1}^*\rangle + (\eta_t - \eta_{t-1})\langle \mathbf{x}_{t-1}^*, \mathbf{x}_t^* - \mathbf{x}_{t-1}^*\rangle.\end{aligned}$$

Whence,

$$\eta_t \left\|\mathbf{x}_t^* - \mathbf{x}_{t-1}^*\right\|^2 \leq (\eta_t - \eta_{t-1})\langle \mathbf{x}_{t-1}^*, \mathbf{x}_{t-1}^* - \mathbf{x}_t^*\rangle \leq (\eta_t - \eta_{t-1}) \left\|\mathbf{x}_{t-1}^*\right\| \cdot \left\|\mathbf{x}_t^* - \mathbf{x}_{t-1}^*\right\|.$$

The next claim follows:

Proposition A.3 *For any monotonically decreasing sequence $\{\eta_t\}_{t\in\mathbb{N}} \subset (0, \infty)$ satisfying $\eta_t \downarrow 0$ we have*

$$\left(\frac{\eta_t - \eta_{t-1}}{\eta_t}\right) \inf_{\mathbf{x}\in \mathrm{SOL}(\phi,r)} \|\mathbf{x}\| \geq \left\|\mathbf{x}^*(\eta_t) - \mathbf{x}^*(\eta_{t-1})\right\|.$$

Lastly, we provide an exact localization result on the Tikhonov sequence.

Proposition A.4 *Let $\eta_t \downarrow 0$ and $\mathbf{x}_t^* \equiv \mathbf{x}^*(\eta_t)$ the corresponding sequence of solutions to the regularized problem* $\mathrm{MVI}(\phi^{\eta_t}, r)$. *Then,* $\inf_{\mathbf{x}\in \mathrm{SOL}(\phi,r)} \|\mathbf{x}\|$ *exists and is uniquely attained and* $\mathbf{x}_t^* \to \arg\min_{\mathbf{x}\in \mathrm{SOL}(\phi,r)} \|\mathbf{x}\|$.

Proof The set $\mathrm{SOL}(\phi, r)$ agrees with the zeros of the monotone inclusion problem (33). Since ∂r is maximally monotone, and ϕ is continuous and monotone, it follows from Corollary 24.4 in [6] that the set of zeros is closed and convex. Hence, the problem $\inf_{\mathbf{x}\in \mathrm{SOL}(\phi,r)} \|\mathbf{x}\|$ admits a unique solution, proving the first part of the Proposition. For the second part, let $\eta_t \downarrow 0$ and $\mathbf{x}_t^* \equiv \mathbf{x}^*(\eta_t)$ the corresponding sequence of unique solution of $\mathrm{MVI}(\phi^{\eta_t}, r)$. Since $\{\mathbf{x}_t^*\}_t$ is bounded (Proposition A.1), we can pass to a converging subsequence. By an abuse of notation, omitting the relabeling, let us take the full sequence to be converging with limit point $\hat{\mathbf{x}}$. By Proposition A.2, we know $\hat{\mathbf{x}} \in \mathrm{SOL}(\phi, r)$. In particular, $\left\|\hat{\mathbf{x}}\right\| \geq \inf_{\mathbf{x}\in \mathrm{SOL}(\phi,r)} \|\mathbf{x}\|$. But then, in view of Proposition A.1, it follows $\left\|\hat{\mathbf{x}}\right\| = \inf_{\mathbf{x}\in \mathrm{SOL}(\phi,r)} \|\mathbf{x}\|$. Since the accumulation point $\hat{\mathbf{x}}$ is arbitrary, the entire sequence $\{\mathbf{x}_t^*\}_t$ converges with limit $\arg\min_{\mathbf{x}\in \mathrm{SOL}(\phi,r)} \|\mathbf{x}\|$. □

A.4 Smoothing

We let $\mathbb{B}_n \triangleq \{\mathbf{x} \in \mathbb{R}^n | \, \|\mathbf{x}\| \leq 1\}$ denote the unit ball in $\mathbb{R}^n$. The unit sphere is denoted by $\mathbb{S}_n = \{\mathbf{x} \in \mathbb{R}^n | \, \|\mathbf{x}\| = 1\}$. The volume of the unit ball with radius δ with respect to n-dimensional Lebesgue measure is $\mathbf{Vol}_n(\delta\mathbb{B}_n) = \delta^n \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}$, where $\Gamma(\cdot)$ is the Gamma function. Therefore, the measure $\mathrm{d}\mu_n(\mathbf{u}) \triangleq \mathbf{1}_{\{\mathbf{u}\in\mathbb{B}_n\}} \frac{\mathrm{d}\mathbf{u}}{\mathbf{Vol}_n(\mathbb{B}_n)}$ defines a uniform distribution on the unit ball in $\mathbb{R}^n$. Recall that $\mathbf{Vol}_{n-1}(\delta\mathbb{S}_n) = \frac{n}{\delta}\mathbf{Vol}_n(\delta\mathbb{B}_n)$ for all $\delta > 0$. Given $\delta > 0$, and $\mathcal{X}$ be a closed convex set in $\mathbb{R}^n$. We define the set $\mathcal{X}_\delta \triangleq \mathcal{X} + \delta\mathbb{B}_n$.

Definition A.2 Let $h : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. The spherical smoothing of h is defined by

$$h^\delta(\mathbf{x}) \triangleq \frac{1}{\mathbf{Vol}_n(\delta\mathbb{B}_n)} \int_{\delta\mathbb{B}_n} h(\mathbf{x} + \mathbf{u})\mathrm{d}\mathbf{u} = \int_{\mathbb{B}_n} h(\mathbf{x} + \delta\mathbf{w}) \frac{\mathrm{d}\mathbf{w}}{\mathbf{Vol}_n(\mathbb{B}_n)}. \tag{37}$$

The following properties of the spherical smoothing can be deduced from [55, Section 9.3.2]; see also [16, Lemma 1].

Fact A.3 *Let $h : \mathbb{R}^n \to \mathbb{R}$ be a continuous function that is L_h-Lipschitz continuous on $\mathcal{X}_\delta$. Then, for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}$, we have*

$$\left|h^\delta(\mathbf{x}) - h^\delta(\mathbf{y})\right| \leq L_h \left\|\mathbf{x} - \mathbf{y}\right\|, \tag{38}$$

$$\left|h^\delta(\mathbf{x}) - h(\mathbf{x})\right| \leq L_h \delta. \tag{39}$$

Using Stoke's Theorem, one can easily show that

$$\begin{aligned} \nabla h^\delta(\mathbf{x}) &= \frac{n}{\delta} \int_{\mathbb{S}_n} h(\mathbf{x} + \delta \mathbf{v}) \mathbf{v} \frac{\mathrm{d}\mathbf{v}}{\mathbf{Vol}_{n-1}(\mathbb{S}_n)} \\ &= \frac{n}{\delta} \mathbb{E}_{\mathbf{W} \sim \mathsf{U}(\mathbb{S}_n)}[\mathbf{W} h(\mathbf{x} + \delta \mathbf{W})] \\ &= \frac{n}{\delta} \mathbb{E}_{\mathbf{W} \sim \mathsf{U}(\mathbb{S}_n)}[\mathbf{W}\left(h(\mathbf{x} + \delta \mathbf{W}) - h(\mathbf{x})\right)], \end{aligned} \tag{40}$$

where $\mathbf{W} \sim \mathsf{U}(\mathbb{S}_n)$ means that $\mathbf{W}$ is uniformly distributed on $\mathbb{S}_n$. A simple application of Jensen's inequality shows then that the spherical smoothing admits a Lipschitz continuous gradient, whose modulus depends on the smoothing parameter δ.

Fact A.4 *Let $h : \mathbb{R}^n \to \mathbb{R}$ be a continuous function that is L_h-Lipschitz continuous on $\mathcal{X}_\delta$. Then, for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}$, we have*

$$\left\|\nabla h^\delta(\mathbf{x}) - \nabla h^\delta(\mathbf{y})\right\| \leq \frac{L_h n}{\delta} \left\|\mathbf{x} - \mathbf{y}\right\|. \tag{41}$$

The smoothed function and its gradient is used to construct a variance reduced gradient estimator for the implicit cost function of player i in our hierarchical game problem.

A.5 Random Sampling

In this section we explain how to construct a random oracle to sample the gradient of the smoothed implicit function h_i^δ. Let $b \in \mathbb{N}$ denote the batch size. In each round of the algorithm, agent i enters the outer loop procedure, which asks this agent to construct a $n_i \times b$ matrix $\mathbf{W}_i^{1:b} = [\mathbf{W}_i^{(1)}; \ldots; \mathbf{W}_i^{(b)}]$ satisfying $\left\|\mathbf{W}_i^{(s)}\right\| = 1$ for all $1 \leq s \leq b$. To construct the uniformly distributed unit vector $\mathbf{W}_i^{(s)}$, we generate n_i random numbers $w_i^s(k) \sim \mathsf{N}(0, 1)$, and then compute

$$\mathbf{W}_i^{(s)} = \frac{1}{\sqrt{\sum_{k=1}^{n_i} w_i^s(k)^2}} [w_i^{(s)}(1); \ldots; w_i^{(s)}(n_i)]^\top \quad s = 1, \ldots, b.$$

The outcome of this procedure is a $n_i \times b$ random matrix $\mathbf{W}_i^{1:b} = [\mathbf{W}_i^{(1)}; \ldots; \mathbf{W}_i^{(b)}] \in \mathbb{R}^{n_i \times b}$ with $\left\|\mathbf{W}_i^{(s)}\right\| = 1$ for all $1 \leq s \leq b$. In fact, since the Gaussian is spherical, the columns of this matrix will be iid uniformly distributed on $\mathbb{S}_{n_i}$. Having constructed these random vectors, each agent constructs a gradient estimator involving the finite-difference approximation of the directional derivative

$$H^{\delta}_{\mathbf{x}_i}(\mathbf{W}_i) \triangleq n_i \mathbf{W}_i \nabla_{(\mathbf{W}_i,\delta)} h_i(\mathbf{x}_i), , \quad \nabla_{(\mathbf{w},\delta)} h_i(\mathbf{x}_i) \triangleq \frac{h_i(\mathbf{x}_i + \delta\mathbf{w}) - h_i(\mathbf{x}_i)}{\delta},$$

as well as its Monte–Carlo variant (with some abuse of notation)

$$H^{\delta,b}_{\mathbf{x}_i} \triangleq \frac{1}{b}\sum_{s=1}^{b} H^{\delta}_{\mathbf{x}_i}(\mathbf{W}_i^{(s)}).$$

Note that, since $h_i(\cdot)$ is $L_{1,i}$-Lipschitz and directionally differentiable on the convex compact set $\mathcal{X}$, we have

$$\lim_{\delta\to 0^+} \nabla_{(\mathbf{w},\delta)} h_i(\mathbf{x}) = h_i^{\circ}(\mathbf{x}_i, \mathbf{w})$$

as well $\left|h_i^{\circ}(\mathbf{x}, \mathbf{w})\right| \leq L_{1,i}$ for all $\mathbf{w} \in \mathbb{S}_{n_i}$. To understand the statistical properties of this estimator, we need the next Lemma. To simplify the notation, we omit the index of player i.

Lemma A.2 *Suppose h is L_h-Lipschitz continuous on $\mathcal{X}_\delta \triangleq \mathcal{X} + \delta\mathbb{B}$. Define $\mathbb{e}_{\mathbf{x}}(\mathbf{W}^{1:b}) \triangleq H^{\delta,b}_{\mathbf{x}} - \nabla h^{\delta}(\mathbf{x})$, where $\mathbf{W}^{(i)}$ is an i.i.d sample drawn uniformly from the unit sphere $\mathbb{S}_n$, i.e. $\mathbf{W}^{1:b} \sim \mathsf{U}(\mathbb{S}_n)^{\otimes b}$. Then*

(a) $\mathbb{E}_{\mathbf{W}^{1:b}\sim\mathsf{U}(\mathbb{S}_n)^{\otimes b}}[\mathbb{e}_{\mathbf{x}}(\mathbf{W}^{1:b})] = 0$*;*

(b) $\left\| H^{\delta}_{\mathbf{x}}(\mathbf{w})\right\|^2 \leq L_h^2 n^2$ *for all* $\mathbf{w} \in \mathbb{S}_n$*;*

(c) $\mathbb{E}_{\mathbf{W}^{1:b}\sim\mathsf{U}(\mathbb{S}_n)^{\otimes b}}[\left\| \mathbb{e}_{\mathbf{x}}(\mathbf{W}^{1:b})\right\|^2] \leq \frac{n^2 L_h^2}{b}$.

Proof By linearity of the expectation operator and independence, we see

$$\mathbb{E}_{\mathbf{W}^{1:b}\sim\mathsf{U}(\mathbb{S}_n)^{\otimes b}}[H^{\delta}_{\mathbf{x}}(\mathbf{W}^{(1)}, \ldots, \mathbf{W}^{(b)})] = \frac{1}{b}\sum_{s=1}^{b} \mathbb{E}_{\mathbf{W}\sim\mathsf{U}(\mathbb{S}_n)}[H^{\delta}_{\mathbf{x}}(\mathbf{W})] = \mathbb{E}_{\mathbf{W}\sim\mathsf{U}(\mathbb{S}_n)}[H^{\delta}_{\mathbf{x}}(\mathbf{W})] \overset{(40)}{=} \nabla h^{\delta}(\mathbf{x}).$$

This proves part (a). Part (b) is a simple consequence of the following Lipschitz argument:

$$\left\| H^{\delta}_{\mathbf{x}}(\mathbf{w})\right\|^2 = \left(\frac{n}{\delta}\right)^2 \|\mathbf{w}\|^2 \left|h(\mathbf{x} + \delta\mathbf{w}) - h(\mathbf{x})\right|^2 \leq L_h^2 n^2,$$

using $\|\mathbf{w}\| = 1$. For part (c), observe that for the random i.i.d. sample $\mathbf{W}^{1:b} = \{\mathbf{W}^{(1)}, \ldots, \mathbf{W}^{(b)}\}$ taking values in $\mathbb{S}_n$, we have

$$\left\|\mathbb{e}_{\mathbf{x}}(\mathbf{W}^{1:b})\right\|^2 = \frac{1}{b^2}\left\|\sum_{i=1}^{b}\left(\frac{n}{\delta} H^{\delta}_{\mathbf{x}}(\mathbf{W}^{(i)}) - \nabla h^{\delta}(\mathbf{x})\right)\right\|^2 = \frac{1}{b^2}\left\|\sum_{i=1}^{b} X_i\right\|^2,$$

where $X_i \triangleq \frac{n}{\delta} H^\delta_{\mathbf{x}}(\mathbf{W}^{(i)}) - \nabla h^\delta(\mathbf{x})$ are i.i.d zero-mean random variables, almost surely bounded in squared norm. By independence, we have $\mathbb{E}[\langle X_i, X_j \rangle] = 0$ for $i \neq j$, so that

$$\mathbb{E}_{\mathbf{W}^{1:b} \sim \mathsf{U}(\mathbb{S}_n)^{\otimes b}} \left[\left\| \mathrm{e}_{\mathbf{x}}(\mathbf{W}^{1:b}) \right\|^2 \right] = \frac{1}{b} \left(\mathbb{E}_{\mathbf{W} \sim \mathsf{U}(\mathbb{S}_n)} \left[\left\| H^\delta_{\mathbf{x}}(\mathbf{W}) \right\|^2 \right] - \left\| \nabla h^\delta(\mathbf{x}) \right\|^2 \right)$$
$$\leq \frac{1}{b} \mathbb{E}_{\mathbf{W} \sim \mathsf{U}(\mathbb{S}_n)} \left[\left\| H^\delta_{\mathbf{x}}(\mathbf{W}) \right\|^2 \right] \leq \frac{n^2 L_h^2}{b}.$$

□

A.6 Inexact Implementation

In the main text we have assumed that the solution of the lower level problem is available exactly. In practice, this is difficult to guarantee, particularly when the lower-level problem is large and possible stochastic. Motivated by this concern, we outline a modification of our hierarchical game solver in this section, reliant on access to an ε-inexact solution of the lower-level problem.

Definition A.3 Let $\delta_0 > 0$ be given and set $\mathcal{X}_{i,\delta_0} = \mathcal{X}_i + \delta_0 \mathbb{B}_{n_i}$ for all $i \in \mathcal{I}$. Given $\varepsilon > 0$ and $i \in \mathcal{I}$, we call $\mathbf{y}^\varepsilon_i : \mathcal{X}_{i,\delta} \to \mathcal{Y}_i$ an ε-solution of the lower level problem $\mathrm{VI}(\phi_i(\mathbf{x}_i, \cdot), \mathcal{Y}_i)$ if

$$\mathbb{E}[\left\| \mathbf{y}^\varepsilon_i(\mathbf{x}_i) - \mathbf{y}_i(\mathbf{x}_i) \right\| |\mathbf{x}_i] \leq \varepsilon \qquad \text{a.s.}$$

We note that such a solution is immediately available by employing $\mathcal{O}(1/\varepsilon^2)$ steps of a single-sample stochastic approximation scheme for resolving $\mathrm{VI}(\phi_i(\mathbf{x}_i, \cdot), \mathcal{Y}_i)$. Similarly, if projection onto $\mathcal{Y}_i$ is a computationally costly operation, then a geometrically increasing mini-batch scheme provides a similar oracle complexity but requires only $\mathcal{O}(\ln(1/\varepsilon^2))$ steps. (cf. [48]).

Under the inexact lower level solution $\mathbf{y}^\varepsilon_i$, we let $h^\varepsilon_i(\mathbf{x}_i) = g_i(\mathbf{x}_i, \mathbf{y}^\varepsilon_i(\mathbf{x}_i))$ denote the resulting implicit function coupling leader i and the associated follower. As in the exact regime, we assume that player i has access to an oracle with which she can construct a spherical approximation of the gradient of the implicit function h^ε_i. Hence, for given $\delta > 0$, we let

$$h^{\varepsilon,\delta}_i(\mathbf{x}_i) \triangleq \int_{\mathbb{B}_{n_i}} h^\varepsilon_i(\mathbf{x}_i + \delta \mathbf{w}) \frac{\mathrm{d}\mathbf{w}}{\mathbf{Vol}_n(\mathbb{B}_{n_i})}.$$

Using the notation for Sect. A.5, we denote the resulting estimators by $H^{\delta,\varepsilon}_{i,\mathbf{x}_i}(\mathbf{W}_i) \triangleq n_i \mathbf{W}_i \nabla_{(\mathbf{W}_i,\delta)} h^\varepsilon_i(\mathbf{x}_i)$, while the corresponding mini-batch counterpart as

$$H^{\delta,\varepsilon,b}_{i,\mathbf{x}_i} \triangleq \frac{1}{b} \sum_{s=1}^{b} H^{\delta,\varepsilon}_{i,\mathbf{x}_i}(\mathbf{W}^{(s)}_i),$$

Next, we derive some bounds of the thus constructed estimator. To reduce notational clutter, we omit the label of player i in the next Lemma.

Lemma A.3 *Let Assumption 2 hold true. Then, for all $\delta, \varepsilon > 0$ and $\mathbf{x} \in \mathcal{X}$, we have*

$$\mathbb{E}\left[\left\| H_{\mathbf{x}}^{\delta,\varepsilon}(\mathbf{W})\right\|^2 \middle| \mathbf{x}\right] \leq 3\left(\frac{n}{\delta}\right)^2 (2L_2^2\varepsilon^2 + L_1^2\delta^2) \text{ and} \tag{42}$$

$$\mathbb{E}\left[\left\| H_{\mathbf{x}}^{\delta,\varepsilon}(\mathbf{W}) - H_{\mathbf{x}}^{\delta}(\mathbf{W})\right\|^2 |\mathbf{x}\right] \leq \left(\frac{2L_{2,h}n\varepsilon}{\delta}\right)^2 \tag{43}$$

almost surely.

Proof We have

$$\begin{aligned}\left\| H_{\mathbf{x}}^{\delta,\varepsilon}(\mathbf{W})\right\|^2 &= \left(\frac{n}{\delta}\right)^2 \left\|\mathbf{W}\left[h^{\varepsilon}(\mathbf{x}+\delta\mathbf{W}) - h(\mathbf{x}+\delta\mathbf{W}) + h(\mathbf{x}+\delta\mathbf{W}) + h(\mathbf{x}) - h(\mathbf{x}) - h^{\varepsilon}(\mathbf{x})\right]\right\|^2 \\ &\leq 3\left(\frac{n}{\delta}\right)^2\left[\left\|\mathbf{W}(h^{\varepsilon}(\mathbf{x}+\delta\mathbf{W}) - h(\mathbf{x}+\delta\mathbf{W})\right\|^2 + \left\|\mathbf{W}(h^{\varepsilon}(\mathbf{x}) - h(\mathbf{x}))\right\|^2 + \|\mathbf{W}(h(\mathbf{x}+\delta\mathbf{W}) - h(\mathbf{x}))\|^2\right].\end{aligned}$$

We bound each of the three terms separately. First, by Assumption 2.(iv), we note

$$\begin{aligned}\left\|\mathbf{W}(h^{\varepsilon}(\mathbf{x}+\delta\mathbf{W}) - h(\mathbf{x}+\delta\mathbf{W})\right\|^2 &\leq \left\|g(\mathbf{x}+\delta\mathbf{W}, \mathbf{y}^{\varepsilon}(\mathbf{x}+\delta\mathbf{W})) - g(\mathbf{x}+\delta\mathbf{W}, \mathbf{y}(\mathbf{x}+\delta\mathbf{W}))\right\|^2 \\ &\leq L_2^2\left\|\mathbf{y}^{\varepsilon}(\mathbf{x}+\delta\mathbf{W}) - \mathbf{y}(\mathbf{x}+\delta\mathbf{W})\right\|^2.\end{aligned}$$

Consequently, using Definition A.3, we obtain

$$\mathbb{E}\left[\left\|\mathbf{W}(h^{\varepsilon}(\mathbf{x}+\delta\mathbf{W}) - h(\mathbf{x}+\delta\mathbf{W})\right\|^2 |\mathbf{x}\right] \leq L_2^2\varepsilon^2.$$

Second,

$$\left\|\mathbf{W}(h^{\varepsilon}(\mathbf{x}) - h(\mathbf{x}))\right\|^2 \leq \left\|g(\mathbf{x}, \mathbf{y}^{\varepsilon}(\mathbf{x})) - g(\mathbf{x}, \mathbf{y}(\mathbf{x})\right\|^2 \leq L_2^2\left\|\mathbf{y}(\mathbf{x}) - \mathbf{y}^{\varepsilon}(\mathbf{x})\right\|^2.$$

Invoking again Definition A.3, it follows

$$\mathbb{E}\left[\left\|\mathbf{W}(h^{\varepsilon}(\mathbf{x}) - h(\mathbf{x}))\right\|^2 |\mathbf{x}\right] \leq L_2^2\varepsilon^2.$$

Third, by Assumption 2.(ii) and since $\mathbf{W} \in \mathbb{S}$, we conclude

$$\|\mathbf{W}(h(\mathbf{x}+\delta\mathbf{W}) - h(\mathbf{x}))\|^2 \leq L_1^2\|\delta\mathbf{W}\|^2 = L_1^2\delta^2.$$

Summarizing all these bounds, we obtain (42). To show (43), we first note

$$\begin{aligned}\left\| H_{\mathbf{x}}^{\delta,\varepsilon}(\mathbf{W}) - H_{\mathbf{x}}^{\delta}(\mathbf{W})\right\| \leq n&\left\|\mathbf{W}\frac{g(\mathbf{x}+\delta\mathbf{W}, \mathbf{y}^{\varepsilon}(\mathbf{x}+\delta\mathbf{W})) - g(\mathbf{x}+\delta\mathbf{W}, \mathbf{y}(\mathbf{x}+\delta\mathbf{W}))}{\delta}\right\| \\ &+ n\left\|\mathbf{W}\frac{g(\mathbf{x}, \mathbf{y}^{\varepsilon}(\mathbf{x})) - g(\mathbf{x}, \mathbf{y}(\mathbf{x}))}{\delta}\right\|\end{aligned}$$

$$\leq \frac{L_{2,h}n}{\delta}\left\|\mathbf{y}^{\varepsilon}(\mathbf{x}+\delta\mathbf{W})-\mathbf{y}(\mathbf{x}+\delta\mathbf{W})\right\|+\frac{nL_{2,h}}{\delta}\left\|\mathbf{y}^{\varepsilon}(\mathbf{x})-\mathbf{y}(\mathbf{x})\right\|.$$

Taking expectations on both sides, it follows

$$\mathbb{E}\left[\left\|H_{\mathbf{x}}^{\delta,\varepsilon}(\mathbf{W})-H_{\mathbf{x}}^{\delta}(\mathbf{W})\right\| \,|\mathbf{x}\right] \leq \frac{2L_{2,h}n\varepsilon}{\delta}.$$

□

Acknowledgements Shisheng Cui acknowledges supported in part by NSFC under Grant 62373050. Uday V. Shanbhag are partially supported by ONR Grant N00014-22-1-2589, DOE Grant DE-SC0023303, and AFOSR Grant FA9550-24-1-0259. The research of MST benefited from the support of the FMJH Program Gaspard Monge for optimization and operations research and their interactions with data science.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

1. Alacaoglu, A., Malitsky, Y.: Stochastic variance reduction for variational inequality methods. In: Conference on Learning Theory, pp. 778–816 (2022)
2. Alacaoglu, A., Malitsky, Y., Cevher, V.: Forward-reflected-backward method with variance reduction. Comput. Optim. Appl. **80**(2), 321–346 (2021)
3. Aussel, D., Egea, C., Schmidt, M.: A tutorial on solving single-leader-multi-follower problems using sos1 reformulations. Int. Trans. Oper. Res. **32**, 1227 (2025)
4. Azizian, W., Iutzeler, F., Malick, J., Mertikopoulos, P.: The last-iterate convergence rate of optimistic mirror descent in stochastic variational inequalities. In: Conference on Learning Theory, pp. 326–358 (2021)
5. Bai, Y., Jin, C., Wang, H., Xiong, C.: Sample-efficient learning of stackelberg equilibria in general-sum games. In: Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P.S., Wortman Vaughan, J. (eds.), Advances in Neural Information Processing Systems, vol. 34, pp. 25799–25811. Curran Associates, Inc. (2021)
6. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, Berlin (2016)
7. Beck, Y., Ljubić, I., Schmidt, M.: A survey on bilevel optimization under uncertainty. Eur. J. Oper. Res. **311**(2), 401–426 (2023)
8. Berahas, A.S., Cao, L., Choromanski, K., Scheinberg, K.: A theoretical and empirical comparison of gradient approximations in derivative-free optimization. Found. Comput. Math. **22**(2), 507–560 (2022)
9. Blau, T., Bonilla, E.V., Chades, I., Dezfouli, A.: Optimizing sequential experimental design with deep reinforcement learning. In: International Conference on Machine Learning, pp. 2107–2128 (2022)
10. Boţ, R.I., Mertikopoulos, P., Staudigl, M., Vuong, P.T.: Minibatch forward-backward-forward methods for solving stochastic variational inequalities. Stoch. Syst. **11**, 112 (2021)
11. Castillo, A., Flicker, J., Hansen, C.W., Watson, J.-P., Johnson, J.: Stochastic optimisation with risk aversion for virtual power plant operations: a rolling horizon control. IET Gener. Transm. Distrib. **13**(11), 2063–2076 (2019)

12. Chavdarova, T., Gidel, G., Fleuret, F., Lacoste-Julien, S.: Reducing noise in gan training with variance reduced extragradient. In: Wallach, H., Larochelle, H., Beygelzimer, A., d' Alché-Buc, F., Fox, E., Garnett, R. (eds.) Advances in Neural Information Processing Systems, vo. 32. Curran Associates, Inc. (2019)
13. Clarke, F.H., Ledyayev, Y.S., Stern, R.J., Wolenski, P.R.: Nonsmooth Analysis and Control Theory. Springer, Berlin (1998)
14. Cui, S., Shanbhag, U., Staudigl, M., Vuong, P.: Stochastic relaxed inertial forward-backward-forward splitting for monotone inclusions in Hilbert spaces. Comput. Optim. Appl. **83**(2), 465–524 (2022)
15. Cui, S., Shanbhag, U.V.: On the computation of equilibria in monotone and potential stochastic hierarchical games. Math. Program. **198**(2), 1227–1285 (2023)
16. Cui, S., Shanbhag, U.V., Yousefian, F.: Complexity guarantees for an implicit smoothing-enabled method for stochastic mpecs. Math. Program. **198**, 1153 (2022)
17. DeMiguel, V., Xu, H.: A stochastic multiple-leader Stackelberg model: analysis, computation, and application. Oper. Res. **57**(5), 1220–1235 (2009)
18. DeMiguel, V., Huifu, X.: A stochastic multiple-leader Stackelberg model: analysis, computation, and application. Oper. Res. **57**(5), 1220–1235 (2009)
19. Dempe, S.: Foundations of Bilevel Programming. In: Nonconvex Optimization and Its Applications, vol. 61. Kluwer, Dordrecht (2002)
20. Dempe, S., Zemkoho, A. (eds.): Bilevel Optimization–Advances and Next Challenges. In: Springer Optimization and Its Applications, vol. 161. Springer, Cham (2020)
21. Dontchev, A.L., Rockafellar, R.T.: Implicit Functions and Solution Mappings: A View from Variational Analysis, vol. 11. Springer, New York (2009)
22. Downing, J., Johnson, N., McNicholas, M., Nemtzow, D., Oueid, R., Paladino, J.: Pathways to Commercial Liftoff and Elizabeth Bellis Wolfe. Virtual Power Plants. Technical Report (2023)
23. Duvocelle, B., Mertikopoulos, P., Staudigl, M., Vermeulen, D.: Multiagent online learning in time-varying games. In: Mathematics of Operations Research (2022)
24. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems , vol. I, II. Springer Series in Operations Research (2003)
25. Fiez, T., Chasnov, B., Ratliff, L.: Implicit learning dynamics in Stackelberg games: equilibria characterization, convergence analysis, and empirical study. In: International Conference on Machine Learning, pp. 3133–3144 (2020)
26. Flaxman, A.D., Kalai, A.T., McMahan, H.B.: Online convex optimization in the bandit setting: gradient descent without a gradient. In: Society for Industrial and Applied Mathematics (2005)
27. Gabriel, S.A., Conejo, A.J., Fuller, J.D., Hobbs, B.F., Ruiz, C., Gabriel, S.A., Conejo, A.J., Fuller, J.D., Hobbs, B.F., Ruiz, C.: Optimality and complementarity. In: Complementarity Modeling in Energy Markets, pp. 31–69 (2013)
28. Garifi, K., Baker, K., Touri, B., Christensen, D.: Stochastic model predictive control for demand response in a home energy management system. In: 2018 IEEE Power and Energy Society General Meeting (PESGM), pp. 1–5 (2018)
29. Ghadimi, S., Wang, M.: Approximation methods for bilevel programming. arXiv preprint arXiv:1802.02246 (2018)
30. Giannou, A., Vlatakis-Gkaragkounis, E.V., Mertikopoulos, P.: Survival of the strictest: stable and unstable equilibria under regularized learning with partial information. In: Conference on Learning Theory pp. 2147–2148 (2021)
31. Gower, R.M., Schmidt, M., Bach, F., Richtárik, P.: Variance-reduced methods for machine learning. Proc. IEEE **108**(11), 1968–1983 (2020)
32. Hobbs, B.: Linear complementarity models of Nash–Cournot competition in bilateral and POOLCO power markets. IEEE Trans. Power Syst. **16**(2), 194–202 (2001)
33. Hobbs, B., Metzler, C.B.: Nash–Cournot equilibria in power markets on a linearized DC network with arbitrage: formulations and properties. Netw. Spat. Econ. **3**, 123–150 (2003)
34. Hobbs, B.F., Pang, J.S.: Nash–Cournot equilibria in electric power markets with piecewise linear demand functions and joint constraints. Oper. Res. **55**(1), 113–127 (2007)
35. Hu, M., Fukushima, M.: Existence, uniqueness, and computation of robust Nash equilibria in a class of multi-leader-follower games. SIAM J. Optim. **23**(2), 894–916 (2013)
36. Hu, X.M., Ralph, D.: Convergence of a penalty method for mathematical programming with complementarity constraints. J. Optim. Theory Appl. **123**(2), 365–398 (2004)

37. Iusem, A.N., Jofré, A., Oliveira, R.I., Thompson, P.: Extragradient method with variance reduction for stochastic variational inequalities. SIAM J. Optim. **27**(2), 686–724 (2017)
38. Jin, X., Baker, K., Christensen, D., Isley, S.: Foresee: a user-centric home energy management system for energy efficiency and demand response. Appl. Energy **205**, 1583–1595 (2017)
39. Johnson, R., Zhang, T.: Accelerating stochastic gradient descent using predictive variance reduction. In: Burges, C.J.C., Bottou, L., Welling, M., Ghahramani, Z., Weinberger, K.Q. (eds.), Advances in Neural Information Processing Systems, vol. 26, pp. 315–323. Curran Associates, Inc. (2013)
40. Kannan, A., Shanbhag, U.V.: Distributed computation of equilibria in monotone Nash games via iterative regularization techniques. SIAM J. Optim. **22**(4), 1177–1205 (2012)
41. Kannan, A., Shanbhag, U.V.: Optimal stochastic extragradient schemes for pseudomonotone stochastic variational inequality problems and their variants. Comput. Optim. Appl. **74**(3), 779–820 (2019)
42. Kannan, A., Shanbhag, U.V., Kim, H.M.: Strategic behavior in power markets under uncertainty. Energy Syst. **2**(2), 115–141 (2011)
43. Koshal, J., Nedić, A., Shanbhag, U.V.: Regularized iterative stochastic approximation methods for stochastic variational inequality problems. IEEE Trans. Automat. Control **58**(3), 594–609 (2013)
44. Kozak, D., Molinari, C., Rosasco, L., Tenorio, L., Villa, S.: Zeroth-order optimization with orthogonal random directions. Math. Program. **199**, 1179 (2022)
45. Kulkarni, A.A., Shanbhag, U.V.: An existence result for hierarchical Stackelberg v/s Stackelberg games. IEEE Trans. Autom. Control **60**(12), 3379–3384 (2015)
46. Kwon, J., Kwon, D., Wright, S., Nowak, R.D.: A fully first-order method for stochastic bilevel optimization. In: PMLR (2023)
47. Lakshmanan, H., de Farias, D.P.: Decentralized resource allocation in dynamic networks of agents. SIAM J. Optim. **19**(2), 911–940 (2008)
48. Lei, J., Shanbhag, U.V.: Distributed variable sample-size gradient-response and best-response schemes for stochastic Nash equilibrium problems. SIAM J. Optim. **32**(2), 573–603 (2022)
49. Leyffer, S., López-Calva, G., Nocedal, J.: Interior methods for mathematical programs with complementarity constraints. SIAM J. Optim. **17**(1), 52–77 (2006)
50. Liu, R., Gao, J., Zhang, J., Meng, D., Lin, Z.: Investigating bi-level optimization for learning and vision from a unified perspective: a survey and beyond. IEEE Trans. Pattern Anal. Mach. Intell. **44**(12), 10045–10067 (2022)
51. Luo, Z.-Q., Pang, J.-S., Ralph, D.: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge (1996)
52. Mertens, J.-F., Sorin, S., Zamir, S.: Repeated Games. Econometric Society Monographs, Cambridge University Press, Cambridge (2015)
53. Naval, N., Yusta, J.M.: Virtual power plant models and electricity markets: a review. Renew. Sustain. Energy Rev. **149**, 111393 (2021)
54. Nemirovski, A.S., Juditsky, A., Lan, G., Shapiro, A.: Robust stochastic approximation approach to stochastic programming. SIAM J. Optim. **19**(4), 1574–1609 (2009)
55. Nemirovski, A.S., Yudin, D.B.: Problem Complexity and Method Efficiency in Optimization. Wiley, New York (1983)
56. Nesterov, Y.: Dual extrapolation and its applications to solving variational inequalities and related problems. Math. Program. **109**(2), 319–344 (2007)
57. Outrata, J., Kočvara, M., Zowe, J.: Nonsmooth Approach to Optimization Problems with Equilibrium Constraints, Nonconvex Optimization and its Applications, Theory, Applications and Numerical Results, vol. 18. Kluwer, Dordrecht (1998)
58. Palaniappan, B., Bach, F.: Stochastic variance reduction methods for saddle-point problems. In: Advances in Neural Information Processing Systems, pp. 1416–1424 (2016)
59. Pang, J.-S., Fukushima, M.: Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games. CMS **2**(1), 21–56 (2005)
60. Polyak, B.T.: Introduction to Optimization. Optimization Software (1987)
61. Robinson, S.M.: Strongly regular generalized equations. Math. Oper. Res. **5**(1), 43–62 (1980)
62. Rui, Z., Peng, W., Qin, X., Wang, J.: Assessing carbon cap-and-trade policies on hybrid renewable energy investments: implications for pricing and capacity decisions. Resour. Policy **86**, 104113 (2023)
63. Shanbhag, U.V., Infanger, G., Glynn, P.W.: A complementarity framework for forward contracting under uncertainty. Oper. Res. **59**(4), 810–834 (2011)
64. Sherali, H.D.: A multiple leader Stackelberg model and analysis. Oper. Res. **32**(2), 390–404 (1984)

65. Steklov, V.A.: Sur les expressions asymptotiques decertaines fonctions dfinies par les quations diffrentielles du second ordre et leers applications au problme du dvelopement d'une fonction arbitraire en sries procdant suivant les diverses fonctions. Commun. Charkov Math. Soc **2**(10), 97–199 (1907)
66. Tseng, P.: A modified forward–backward splitting method for maximal monotone mappings. SIAM J. Control Optim. **38**(2), 431–446 (2000)
67. Xu, H.: An MPCC approach for stochastic Stackelberg–Nash–Cournot equilibrium. Optimization **54**(1), 27–57 (2005)
68. Xu, H.: An implicit programming approach for a class of stochastic mathematical programs with complementarity constraints. SIAM J. Optim. **16**(3), 670–696 (2006)
69. Yousefian, F., Nedić, A., Shanbhag, U.V.: Convex nondifferentiable stochastic optimization: a local randomized smoothing technique. In: Proceedings of the 2010 American Control Conference, pp. 4875–4880 (2010)
70. Yousefian, F., Nedić, A., Shanbhag, U.V.: On stochastic gradient and subgradient methods with adaptive steplength sequences. Automatica **48**(1), 56–67 (2012)
71. Yousefian, F., Nedic, A., Shanbhag, U.V.: A regularized smoothing stochastic approximation (rssa) algorithm for stochastic variational inequality problems. In: 2013 Winter Simulation Conference (WSC), pp. 933–944. IEEE (2013)
72. Yousefian, F., Nedić, A., Shanbhag, U.V.: On smoothing, regularization, and averaging in stochastic approximation methods for stochastic variational inequality problems. Math. Program. **165**(1), 391–431 (2017)