Almost-Zero Duality Gaps in Model-Free Resource Allocation for Wireless Systems

Dionysios S. Kalogerias, Mark Eisen, George J. Pappas, and Alejandro Ribeiro

***Abstract*—We investigate optimal resource management in wireless systems, directly in the model-free setting. Starting with a generic resource allocation task formulated as a variational program with nonconvex stochastic constraints, we leverage classical results on Gaussian smoothing to formulate a finite dimensional, smoothed problem surrogate, effectively solvable in a model-free fashion, without the need of a baseline system model. Further assuming a near-universal policy parameterization, we present explicit upper and lower bounds on the gap between the optimal value of the original variational problem, and the dual optimal value of the smoothed surrogate. In fact, we show that this duality gap depends linearly on smoothing and near-universality parameters, and therefore, it can be made arbitrarily small at will. Our results effectively quantify the effects of both policy parameterization and smoothing on approximating both the value and optimal solution of the original variational program via surrogate dualization, and provide explicit near-optimality guarantees in the model-free regime. We also provide empirical illustration via indicative numerical simulations.**

***Index Terms*—Wireless Systems, Resource Allocation, Zeroth-order Learning, Reinforcement Learning, Lagrangian Duality.**

I. Introduction

This paper considers *ergodic networking and communications resource allocation problems* of the generic form [1]

$$\infty > \mathscr{P}^* \triangleq \underset{\boldsymbol{x},\boldsymbol{p}(\cdot)}{\text{maximize}} \quad g^o(\boldsymbol{x}) \quad \text{subject to} \quad \boldsymbol{x} \le \mathbb{E}\{\boldsymbol{f}(\boldsymbol{p}(\boldsymbol{H}),\boldsymbol{H})\}\,(\in \mathbb{R}), \quad \boldsymbol{g}(\boldsymbol{x}) \ge \boldsymbol{0},\ (\boldsymbol{x},\boldsymbol{p}) \in \mathcal{X}\times\mathcal{P} \tag{1}$$

where $\boldsymbol{H} \in \mathcal{H} \subseteq \mathbb{R}_+^{N_{\boldsymbol{H}}}$ denotes a collection of $N_{\boldsymbol{H}}$ random wireless fading channels, $g^o : \mathbb{R}^{N_S} \to \mathbb{R}$ and $\boldsymbol{g} : \mathbb{R}^{N_S} \to \mathbb{R}^{N_{\boldsymbol{g}}}$ are fixed utility functions, $\boldsymbol{p} : \mathcal{H} \to \mathbb{R}^{N_R}$ is the resource allocation policy, and $\boldsymbol{f} : \mathbb{R}^{N_R} \times \mathcal{H} \to \mathbb{R}^{N_S}$ is an instantaneous performance or service level metric, measuring the efficiency of policy $\boldsymbol{p}$ at each realization of $\boldsymbol{H}$. Also, the set of admissible policies $\mathcal{P}$ enforces pointwise constraints on every Borel measurable candidate policy from $\mathcal{H}$ to $\mathbb{R}^{N_R}$, and $\emptyset \neq \mathcal{X} \subseteq \mathbb{R}^{N_S}$.

Although problem (1) is modular and matches many scenarios in wireless communications and networking [1]–[7], it is rather challenging to tackle. First, the policy $\boldsymbol{p}$ is a function, rendering (1) a *variational (infinite dimensional) stochastic program*. Second, although the utilities g^o and $\boldsymbol{g}$ may be typically chosen as concave functions and $\mathcal{X}$ as a convex set, in most existing and practically relevant wireless system models the performance metric $\boldsymbol{f}$ (describing, e.g., information rates, capacities, or signal-to-noise ratios) is not only naturally nonconcave [1], but also partially or even completely unknown.

Lack of convexity is an inherent challenge and it is accepted that we settle for locally optimal solutions, heuristics, or relaxations. To some extent, the same counts for dimensionality and model availability. However, the recent advent of machine learning for wireless communications [6]–[10] has dawned the realization that both these challenges can be ameliorated with the incorporation of learning parameterizations [6], [7]. To see this, introduce a parameterization $\boldsymbol{\phi} : \mathcal{H} \times \mathbb{R}^{N_{\boldsymbol{\phi}}} \to \mathbb{R}^{N_R}$, and restrict resource allocations as $\boldsymbol{p}(\cdot) \equiv \boldsymbol{\phi}(\cdot,\boldsymbol{\theta}), \boldsymbol{\theta} \in \mathbb{R}^{N_{\boldsymbol{\phi}}}$. Then, the base problem (1) may be relaxed as

$$\infty > \mathscr{P}^*_{\boldsymbol{\phi}} \triangleq \underset{\boldsymbol{x},\boldsymbol{\theta}}{\text{maximize}} \quad g^o(\boldsymbol{x}) \quad \text{subject to} \quad \boldsymbol{x} \le \mathbb{E}\{\boldsymbol{f}(\boldsymbol{\phi}(\boldsymbol{H},\boldsymbol{\theta}),\boldsymbol{H})\}, \quad \boldsymbol{g}(\boldsymbol{x}) \ge \boldsymbol{0},\ (\boldsymbol{x},\boldsymbol{\theta}) \in \mathcal{X}\times\Theta \tag{2}$$

where $\Theta \subseteq \{\boldsymbol{\theta} \in \mathbb{R}^{N_{\boldsymbol{\phi}}} | \boldsymbol{\phi}(\cdot,\boldsymbol{\theta}) \in \mathcal{P}\}$ is nonempty and closed. Problem (2) resembles a *multi-objective learning problem scalarized by* g^o, where each entry of $\boldsymbol{x}$ is associated with an expected reward, with the difference of the two formulating a stochastic constraint. Each expected reward has the form of the objective of a single-stage reinforcement learning problem [11], where $\boldsymbol{H}$ and $\boldsymbol{\phi}(\boldsymbol{H},\boldsymbol{\theta})$ correspond to the state and the control, respectively.

Exploiting the representation power of *(near-)universal* policy parameterizations, such as Deep Neural Networks (DNNs), in conjunction with (2), a *model-free* primal-dual algorithm was recently proposed in [6]. This method approximates the gradients of g^o and $\boldsymbol{g}$ via stochastic finite differences, and advocates a policy gradient approach for estimating the composite gradient of $\boldsymbol{f}(\boldsymbol{\phi}(\boldsymbol{H},\cdot),\boldsymbol{H})$, based only on probing $\boldsymbol{f}$; no model of $\boldsymbol{f}$ or its gradient is required. However, although this approach has been shown to work well in some examples [6], [7], issues associated with model-free operation are not addressed. As is the case with policy gradient, the algorithm of [6] requires use of randomized policies. Not only these are inefficient as compared with optimal deterministic policies, but, additionally, we lack understanding of the loss of optimality associated with specific randomization choices, with respect to both the parameterized problem (2), and the initial, infinite dimensional problem (1).

Rigorously treating these fundamental issues is the subject of this paper. In a nutshell, our main contribution is to put forth a *principled approach* for solving the PFA formulation (2) via

Work supported by the NSF under grant CPS 1837253, and also by the Intel Science and Technology Center for Wireless Autonomous Systems. Kalogerias, Pappas, and Ribeiro are with the Department of Electrical & Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA (e-mails: {dionysis, pappasg, aribeiro}@seas.upenn.edu). Eisen is with Intel Corporation, Hillsboro, OR 97124, USA (e-mail: mark.eisen@intel.com).

model-free training. We do so by completely avoiding the use of randomized policies, and instead relying on appropriately constructed, smoothed surrogates to (2), which enable exact zeroth-order gradient representations [12]. This approach not only yields an efficient *and* technically grounded model-free training algorithm (see [13]), but also enables detailed analysis, which effectively quantifies the relation of *both* problems (1) *and* (2) to the smoothed surrogate and, in fact, provides explicit near-optimality guarantees in the model-free regime.

More specifically, assuming an ϵ-universal policy parameterization and under mild regularity conditions, we take [6] strictly one step further by completely characterizing the duality gap between the optimal value of the variational problem (1) and the dual value of the proposed smoothed surrogate. Specifically, we show that the aforementioned duality gap is, in absolute value, of the order of $\mathcal{O}(\mu_S\sqrt{N_S}+\mu_R\sqrt{N_\phi}+\epsilon)$ (Theorem 8), where $\mu_S \geq 0$ and $\mu_R \geq 0$ are *user-prescribed* smoothing parameters associated with the decision variables $\boldsymbol{x}$ and $\boldsymbol{\theta}$ of (2), respectively, and which control the quality of the approximation. If $\mu_S \equiv \mu_R \equiv 0$ (no smoothing), our duality result recovers exactly previous results developed in [6], whereas, for $\mu_S > 0$ and $\mu_R > 0$, it explicitly quantifies the combined effects of *both* policy parameterization *and* smoothing on approximating the optimal value of the original problem (1) via surrogate dualization. Our results are also confirmed via numerical simulations on a standard sumrate maximization problem for a multiuser networking scenario.

Note: For proofs, as well as detailed versions of several of the results presented in the paper, the reader is referred to [13].

II. Gaussian Smoothing & Its Properties

This section introduces Gaussian smoothing and its properties, and follows the corresponding treatment in [12]. Let $f:\mathbb{R}^N \rightarrow \mathbb{R}$ be Borel. Also, for any $\mathbb{R}^N$-valued random element $\boldsymbol{U} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_N)$, and for $\mu \geq 0$, consider another Borel function $f_\mu : \mathbb{R}^N \rightarrow \mathbb{R}$, defined, for $\boldsymbol{x} \in \mathbb{R}^N$, as $f_\mu(\boldsymbol{x}) \triangleq \mathbb{E}\{f(\boldsymbol{x}+\mu\boldsymbol{U})\}$, provided that it is well-defined.

In many cases, f_μ turns out to be everywhere differentiable on $\mathbb{R}^N$, *even if* f *is not*, whereas the gradient of f_μ admits a *zeroth-order representation*, as the next result suggests.

Lemma 1. (Properties of f_μ [12]) *Let $\boldsymbol{U} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_N)$, and consider any globally L-Lipschitz function $f:\mathbb{R}^N \rightarrow \mathbb{R}$. Then, for any $\mathcal{F} \subseteq \mathbb{R}^N$, the following statements are true:*

S1 *For every $\mu \geq 0$, f_μ is well-defined and finite on $\mathcal{F}$, and*

$$\sup_{\boldsymbol{x}\in\mathcal{F}} |f_\mu(\boldsymbol{x}) - f(\boldsymbol{x})| \leq \mu L\sqrt{N}. \tag{3}$$

S2 *If f is convex on $\mathbb{R}^N$, so is f_μ, and $f_\mu \geq f$ on $\mathcal{F}$.*

S3 *For every $\mu > 0$, f_μ is differentiable on $\mathcal{F}$, and its gradient $\nabla f_\mu : \mathbb{R}^N \rightarrow \mathbb{R}^N$ admits the zeroth-order representation*

$$\nabla f_\mu(\boldsymbol{x}) \equiv \mathbb{E}\left\{\frac{f(\boldsymbol{x}+\mu\boldsymbol{U}) - f(\boldsymbol{x})}{\mu}\boldsymbol{U}\right\}, \tag{4}$$

for all $\boldsymbol{x} \in \mathcal{F}$. Further, it is true that

$$\sup_{\boldsymbol{x}\in\mathcal{F}} \mathbb{E}\left\{\left\|\frac{f(\boldsymbol{x}+\mu\boldsymbol{U}) - f(\boldsymbol{x})}{\mu}\boldsymbol{U}\right\|_2^2\right\} \leq L^2(N+4)^2. \tag{5}$$

III. Smoothed Constrained Surrogates

Let $\mu_S \geq 0$ and $\mu_R \geq 0$, and consider random elements $\boldsymbol{U}_S \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_{N_S})$ and $\boldsymbol{U}_R \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_{N_\phi})$, the latter taken independent of $\boldsymbol{H}$. Driven by the results of Section II, we define smoothed versions of g^o, $\boldsymbol{g}$ and $\mathbb{E}\{\boldsymbol{f}(\boldsymbol{\phi}(\boldsymbol{H},\cdot),\boldsymbol{H})\} \triangleq \overline{\boldsymbol{f}}^{\phi}$ as

$$g^o_{\mu_S}(\boldsymbol{x}) \triangleq \mathbb{E}\{g^o(\boldsymbol{x}+\mu_S\boldsymbol{U}_S)\}, \quad \boldsymbol{x} \in \mathcal{X}, \tag{6}$$

$$\boldsymbol{g}_{\mu_S}(\boldsymbol{x}) \triangleq \mathbb{E}\{\boldsymbol{g}(\boldsymbol{x}+\mu_S\boldsymbol{U}_S)\}, \quad \boldsymbol{x} \in \mathcal{X} \quad \text{and} \tag{7}$$

$$\overline{\boldsymbol{f}}^{\phi}_{\mu_R}(\boldsymbol{\theta}) \triangleq \mathbb{E}\{\boldsymbol{f}(\boldsymbol{\phi}(\boldsymbol{H},\boldsymbol{\theta}+\mu_R\boldsymbol{U}_R),\boldsymbol{H})\}, \quad \boldsymbol{\theta} \in \Theta, \tag{8}$$

where, temporarily, we arbitrarily assume that all expectations are well-defined and finite on $\mathcal{X}$ and Θ. Then, we may formulate a (hopefully) smoothed version of problem (2) as

$$\begin{array}{ll} \underset{\boldsymbol{x},\boldsymbol{\theta}}{\text{maximize}} & g^o_{\mu_S}(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} + \mathtt{S}(\mu_R) \leq \overline{\boldsymbol{f}}^{\phi}_{\mu_R}(\boldsymbol{\theta}) \\ & \boldsymbol{g}_{\mu_S}(\boldsymbol{x}) \geq \boldsymbol{0},\ (\boldsymbol{x},\boldsymbol{\theta}) \in \mathcal{X}\times\Theta \end{array}, \tag{9}$$

where $\mathtt{S}:\mathbb{R}_+ \rightarrow \mathbb{R}^{N_S}_+$ is a nonnegative *feasibility slack*, with properties to be determined. Formulation of the *smoothed surrogate* (9) is well-motivated. This is because (9) constitutes a natural *zeroth-order proxy* for dealing with (2) in the model-free setting, i.e., when the functions g^o, $\boldsymbol{g}$ and $\boldsymbol{f}$ are unknown, and may be only observed through *limited probing*. Indeed, whenever the objective g^o and all entries of the vector functions $\boldsymbol{g}$ and $\boldsymbol{f}(\boldsymbol{\phi}(\boldsymbol{H},\cdot),\boldsymbol{H})$ are sufficiently well-behaved, Lemma 1 would imply that the smoothed functions $g^o_{\mu_S}$, $\boldsymbol{g}_{\mu_S}$ and $\overline{\boldsymbol{f}}^{\phi}_{\mu_R}$ are differentiable, and that the respective gradients may be represented as averages of suitably defined finite differences. To ensure that this is actually the case, we impose minimal appropriate structure on g^o, $\boldsymbol{g}$ and $\boldsymbol{f}(\boldsymbol{\phi}(\boldsymbol{H},\cdot),\boldsymbol{H})$. Hereafter, the i-th entries of $\boldsymbol{g}$ ($\boldsymbol{g}_{\mu_S}$) and $\boldsymbol{f}$ ($\overline{\boldsymbol{f}}^{\phi}_{\mu_R}$) are denoted as g^i ($g^i_{\mu_S}$), $i\in\mathbb{N}^+_{N_{\boldsymbol{g}}}$ and f^i ($\overline{f}^{\phi,i}_{\mu_R}$), $i\in\mathbb{N}^+_{N_S}$, respectively.

Assumption 1. The following conditions are satisfied:

C1 For every $i\in\{o,\mathbb{N}^+_{N_{\boldsymbol{g}}}\}$, g^i is L^i_g-Lipschitz on $\mathbb{R}^{N_S}$.

C2 For every $i\in\mathbb{N}^+_{N_S}$, there is $L^i_f < \infty$, such that

$$\begin{aligned} &\|f^i(\boldsymbol{\phi}(\boldsymbol{H},\boldsymbol{\theta}_1),\boldsymbol{H}) - f^i(\boldsymbol{\phi}(\boldsymbol{H},\boldsymbol{\theta}_2),\boldsymbol{H})\|_{\mathcal{L}_2} \\ &\quad \leq L^i_f\|\boldsymbol{\theta}_1-\boldsymbol{\theta}_2\|_2,\ \forall(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2) \in \mathbb{R}^{N_\phi}\times\mathbb{R}^{N_\phi}. \end{aligned} \tag{10}$$

Next, for $\boldsymbol{x} \in \mathcal{X}, \mu_S > 0$ and for $i \in \{o, \mathbb{N}^+_{N_{\boldsymbol{g}}}\}$, let us define

$$\Delta^i_g(\boldsymbol{x},\mu_S,\boldsymbol{U}_S) \triangleq \mu_S^{-1}(g^i(\boldsymbol{x}+\mu_S\boldsymbol{U}_S) - g^i(\boldsymbol{x})). \tag{11}$$

Similarly, for $\boldsymbol{\theta} \in \Theta, \mu_R > 0$ and for $i\in\mathbb{N}^+_{N_S}$, define

$$\begin{aligned} &\Delta^i_f(\boldsymbol{\theta},\mu_R,\boldsymbol{U}_R,\boldsymbol{H}) \\ &\triangleq \mu_R^{-1}(f^i(\boldsymbol{\phi}(\boldsymbol{H},\boldsymbol{\theta}+\mu_R\boldsymbol{U}_R),\boldsymbol{H}) - f^i(\boldsymbol{\phi}(\boldsymbol{H},\boldsymbol{\theta}),\boldsymbol{H})). \end{aligned} \tag{12}$$

Assumption 1 may be exploited to establish well-definiteness and basic properties of $g^o_{\mu_S}$, $\boldsymbol{g}_{\mu_S}$ and $\overline{\boldsymbol{f}}^{\phi}_{\mu_R}$, and their gradients.

Lemma 2. (Existence & Properties of $g^o_{\mu_S}$ and $\boldsymbol{g}_{\mu_S}$) *Suppose that Assumption 1 is in effect. Then, for every $i \in \{o,\mathbb{N}^+_{N_{\boldsymbol{g}}}\}$*

and for every $\mu_S > 0$*, each* $g^i_{\mu_S}$ *is well-defined, finite, concave, differentiable everywhere on* $\mathcal{X}$*, and*

$$\sup_{\boldsymbol{x}\in\mathcal{X}} |g^i_{\mu_S}(\boldsymbol{x}) - g^i(\boldsymbol{x})| \le \mu_S L^i_g \sqrt{N_S}, \tag{13}$$

$$\sup_{\boldsymbol{x}\in\mathcal{X}} \mathbb{E}\{\|\Delta^i_g(\boldsymbol{x},\mu_S,\boldsymbol{U}_S)\boldsymbol{U}_S\|^2_2\} \le (L^i_g)^2(N_S+4)^2 \tag{14}$$

$$\text{and} \quad \mathbb{E}\{\Delta^i_g(\boldsymbol{x},\mu_S,\boldsymbol{U}_S)\boldsymbol{U}_S\} \equiv \nabla g^i_{\mu_S}(\boldsymbol{x}),\ \forall \boldsymbol{x}\in\mathcal{X}. \tag{15}$$

Lemma 3. (Existence & Properties of $\overline{\boldsymbol{f}}^{\boldsymbol{\phi}}_{\mu_R}$) *Suppose that Assumption 1 is in effect. Then, for every* $i \in \mathbb{N}^+_{N_S}$ *and for every* $\mu_R > 0$*, each* $f^i_{\mu_R}$ *is well-defined, finite, differentiable everywhere on* Θ*, and*

$$\sup_{\boldsymbol{\theta}\in\Theta} |\overline{f}^{\boldsymbol{\phi},i}_{\mu_R}(\boldsymbol{\theta}) - \overline{f}^{\boldsymbol{\phi},i}(\boldsymbol{\theta})| \le \mu_R L^i_f \sqrt{N_{\boldsymbol{\phi}}}, \tag{16}$$

$$\sup_{\boldsymbol{\theta}\in\Theta} \mathbb{E}\{\|\Delta^i_f(\boldsymbol{\theta},\mu_R,\boldsymbol{U}_R,\boldsymbol{H})\boldsymbol{U}_R\|^2_2\} \le (L^i_f)^2(N_{\boldsymbol{\phi}}+4)^2 \tag{17}$$

$$\text{and} \quad \mathbb{E}\{\Delta^i_f(\boldsymbol{\theta},\mu_R,\boldsymbol{U}_R,\boldsymbol{H})\boldsymbol{U}_R\} \equiv \nabla \overline{f}^{\boldsymbol{\phi},i}_{\mu_R}(\boldsymbol{\theta}), \forall \boldsymbol{\theta}\in\Theta. \tag{18}$$

We see that, under Assumption 1, the objective and constraints of (2) are all approximated *uniformly* by those of the smoothed surrogate (9). Further, the gradients of all functions involved in (9) exist and admit *exact zeroth-order* representations as *stable* integrated finite differences.

We now investigate conditions ensuring feasibility of (9), *but which are on the original parameterized problem* (2). This is useful from a practical perspective, as the exact form of (9) is, in principle, unknown. The relevant result follows after we define vectors $\boldsymbol{c}_S \triangleq [L^1_g \ \dots \ L^{N_{\boldsymbol{g}}}_g]^{\boldsymbol{T}}$ and $\boldsymbol{c}_R \triangleq [L^1_f \ \dots \ L^{N_S}_f]^{\boldsymbol{T}}$, and under the following assumption.

Assumption 2. The feasibility slack $\mathtt{S}_{\boldsymbol{f}}$ is increasing around the origin, and $\lim_{\mu_R \downarrow 0} \mathtt{S}_{\boldsymbol{f}}(\mu_R) \equiv \mathtt{S}_{\boldsymbol{f}}(0) \equiv \boldsymbol{0}$.

Theorem 4. (Surrogate Strict Feasibility) *Let Assumptions 1 and 2 be in effect, and suppose that* $(\boldsymbol{x}^\dagger, \boldsymbol{\theta}^\dagger) \in \mathbb{R}^{N_S} \times \mathbb{R}^{N_{\boldsymbol{\phi}}}$ *is a strictly feasible point of problem* (2)*. Then, there exist* $\mu^\dagger_S > 0$*,* $\mu^\dagger_R > 0$*, such that, for every* $0 \le \mu_S \le \mu^\dagger_S$ *and* $0 \le \mu_R \le \mu^\dagger_R$*, the same point* $(\boldsymbol{x}^\dagger, \boldsymbol{\theta}^\dagger)$ *is strictly feasible for the smoothed surrogate* (9)*.*

Theorem 4 confirms strict feasibility for problem (9), *uniformly in* μ_S *and* μ_R. A byproduct of Theorem 4 is that (9) is a feasible problem. Another question is how much the constraints of (2) are violated at every feasible point of (9).

Theorem 5. (PFA Constraint Violation) *Let Assumption 1 be in effect. Then, for every* $\mu_R \ge 0$ *such that* $\mathtt{S}(\mu_R) - \mu_R \boldsymbol{c}_R \sqrt{N_{\boldsymbol{\phi}}} \ge \boldsymbol{0}$*, and for every* $\mu_S \ge 0$*, every feasible point of* (9) *is also feasible for* (2)*. Otherwise, if the stated condition fails to hold, then the negative values of its left-hand-side correspond to the maximal constraint violation for* (2)*.*

As an example, one can set $\mathtt{S}(\mu_R) \equiv \mu_R \boldsymbol{c}_R \sqrt{N_{\boldsymbol{\phi}}}$, readily satisfying Assumption 2. However, this might not be feasible in practice, since $\boldsymbol{c}_R$ will probably be unknown. Still, Theorem 5 provides a principle for choosing $\mathtt{S}$. For instance, the choice $\mathtt{S}(\mu_R) \equiv \boldsymbol{C}\mu_R\sqrt{N_{\boldsymbol{\phi}}}$ would work fine, for an appropriate constant vector $\boldsymbol{C} > \boldsymbol{0}$, which may chosen experimentally.

IV. Lagrangian Duality

A promising approach for dealing with the explicit constraints of either problems (2) or (9) is by exploiting *Lagrangian Duality* [14], [15]. However, since both problems (2) and (9) are nonconvex, standard results in convex Lagrangian Duality do not apply automatically.

Our treatment is based on results reported in [6], which rely on earlier results in [1]. In fact, our purpose will be to explicitly link the smoothed surrogate (9) to the parameterized problem (2), and ultimately to the policy search problem (1), *in the dual domain*, effectively characterizing the resulting *end-to-end* duality gap. Our results essentially provide a technically grounded path to dealing with the constrained problem (1) in the model-free setting, through the zeroth-order proxy (9).

To this end, and for every $\mu_S \ge 0$ and $\mu_R \ge 0$, consider the *Lagrangian function* $\mathscr{L}_{\boldsymbol{\phi},\mu} : \mathcal{X} \times \Theta \times \mathbb{R}^{N_{\boldsymbol{g}}} \times \mathbb{R}^{N_S} \to \mathbb{R}$ associated with the smoothed surrogate (9) as

$$\mathscr{L}_{\boldsymbol{\phi},\mu}(\boldsymbol{x},\boldsymbol{\theta},\boldsymbol{\lambda}) \triangleq g^o_{\mu_S}(\boldsymbol{x}) + \langle \boldsymbol{\lambda}_S, \boldsymbol{g}_{\mu_S}(\boldsymbol{x}) \rangle + \langle \boldsymbol{\lambda}_R, \overline{\boldsymbol{f}}^{\boldsymbol{\phi}}_{\mu_R}(\boldsymbol{\theta}) - \boldsymbol{x} - \mathtt{S}(\mu_R) \rangle, \tag{19}$$

where $\boldsymbol{\lambda} \equiv (\boldsymbol{\lambda}_S, \boldsymbol{\lambda}_R) \in \mathbb{R}^{N_{\boldsymbol{g}}} \times \mathbb{R}^{N_S}$ are multipliers tied to the constraints of the *primal problems* (2) and (9). Then the *dual function* $\mathscr{D}_{\boldsymbol{\phi},\mu} : \mathbb{R}^{N_{\boldsymbol{g}}} \times \mathbb{R}^{N_S} \to (-\infty, \infty]$ is defined as

$$\mathscr{D}_{\boldsymbol{\phi},\mu}(\boldsymbol{\lambda}) \triangleq \sup_{(\boldsymbol{x},\boldsymbol{\theta})\in\mathcal{X}\times\Theta} \mathscr{L}_{\boldsymbol{\phi},\mu}(\boldsymbol{x},\boldsymbol{\theta},\boldsymbol{\lambda}). \tag{20}$$

Since $\mathscr{P}^*_{\boldsymbol{\phi},\mu} \le \mathscr{D}_{\boldsymbol{\phi},\mu}$ *on* $\mathbb{R}^{N_{\boldsymbol{g}}}_+ \times \mathbb{R}^{N_S}_+$, it is most reasonable to consider the *dual problem* $\inf_{\boldsymbol{\lambda}\ge\boldsymbol{0}} \mathscr{D}_{\boldsymbol{\phi},\mu}(\boldsymbol{\lambda})$, whose value

$$\mathscr{D}^*_{\boldsymbol{\phi},\mu} \triangleq \inf_{\boldsymbol{\lambda}\ge\boldsymbol{0}} \mathscr{D}_{\boldsymbol{\phi},\mu}(\boldsymbol{\lambda}) \in (-\infty, \infty] \tag{21}$$

serves as the tightest over-estimate of the optimal value of (9), $\mathscr{P}^*_{\boldsymbol{\phi},\mu}$, when knowing only $\mathscr{D}_{\boldsymbol{\phi},\mu}$. Note that $\mathscr{D}_{\boldsymbol{\phi},\mu}$ is convex on $\mathbb{R}^{N_{\boldsymbol{g}}} \times \mathbb{R}^{N_S}$, as a pointwise supremum of affine functions.

Hereafter, we assume that both $\mathscr{P}^*$ and $\mathscr{P}^*_{\boldsymbol{\phi}}$ are attained. We also exploit another standard assumption, as follows.

Assumption 3. Problem (2) is strictly feasible.

Under Assumption 3, it is true that problem (1) is strictly feasible as well; its feasible set contains that of (2).

A. The Dual Smoothed Surrogate

First, we provide fundamental upper and lower bounds on the on the gap of dual optimal values of (2) and (9), as follows.

Theorem 6. (Dual Value Approximation) *Let Assumptions 1, 2 and 3 be in effect, choose* $\mathtt{S}(\mu_R) \equiv \boldsymbol{C}\mu_R\sqrt{N_{\boldsymbol{\phi}}}, \boldsymbol{C} \ge \boldsymbol{0}$*, and suppose that* $\mathscr{D}^*_{\boldsymbol{\phi}} < \infty$*. Then there exist* $\mu^\dagger_S > 0$ *and* $\mu^\dagger_R > 0$*, such that, for every* $0 \le \mu_S \le \mu^\dagger_S$ *and* $0 \le \mu_R \le \mu^\dagger_R$*,*

$$\boxed{-(\mu_S \Sigma^l_S + \mu_R \Sigma^l_R) \le \mathscr{D}^*_{\boldsymbol{\phi},\mu} - \mathscr{D}^*_{\boldsymbol{\phi}} \le \mu_R \Sigma^r_R,} \tag{22}$$

where $0 \le \Sigma^l_S \equiv \mathcal{O}(\sqrt{N_S})$*,* $0 \le \Sigma^l_R \equiv \mathcal{O}(\sqrt{N_{\boldsymbol{\phi}}})$ *and* $\mathbb{R} \ni \Sigma^r_R \equiv \mathcal{O}(\sqrt{N_{\boldsymbol{\phi}}})$ *are finite constants, problem dependent but independent of* μ_S *and* μ_R*. Additionally, whenever* $\boldsymbol{C} \ge \boldsymbol{c}_R$*, then the right-hand-side of* (22) *is nonpositive, and may be replaced by zero.*

B. Approximate Strong Duality

By exploiting the conclusions of by Theorem 6, it is possible to relate $\mathscr{D}^*_{\phi,\mu}$ to the optimal value of the base problem (1). In particular, we would like to characterize the *end-to-end duality gap between the primal problem* (1) *and the dual to* (9).

Following [6], we exploit the notion of a ϵ-universal policy parameterization, whose definition is presented below.

Definition 7. (ϵ-Universality) Fix $\epsilon \geq 0$, choose a parameterization $\boldsymbol{\phi} : \mathcal{H} \times \mathbb{R}^{N_{\boldsymbol{\phi}}} \to \mathbb{R}^{N_R}$, and let $\Theta \subseteq \mathbb{R}^{N_{\boldsymbol{\phi}}}$ be any parameter subspace. A class of admissible policies

$$\mathcal{P}^{\boldsymbol{\phi}}_{\Theta} \triangleq \{\boldsymbol{p} \in \mathcal{P} \,|\, \boldsymbol{p}(\cdot) \equiv \boldsymbol{\phi}(\cdot, \boldsymbol{\theta}),\ \boldsymbol{\theta} \in \Theta\} \subseteq \mathcal{P} \tag{23}$$

is called ϵ*-universal in* $\mathcal{P}$ if and only if, for every $\boldsymbol{p} \in \mathcal{P}$, there exists $\boldsymbol{\phi}(\cdot, \boldsymbol{\theta} \equiv \boldsymbol{\theta}(\epsilon, \boldsymbol{p})) \in \mathcal{P}^{\boldsymbol{\phi}}_{\Theta}$, such that

$$\mathbb{E}\{\|\boldsymbol{p}(\boldsymbol{H}) - \boldsymbol{\phi}(\boldsymbol{H}, \boldsymbol{\theta})\|_\infty\} < \epsilon. \tag{24}$$

Additionally, also as in [6], we will impose the following additional structural assumptions.

Assumption 4. The Borel measure of $\boldsymbol{H}$, $\mathcal{M}_{\boldsymbol{H}} \in [0,1]$, is nonatomic: For every Borel set $\mathcal{E}$ such that $\mathcal{M}_{\boldsymbol{H}}(\mathcal{E}) > 0$, there is another Borel set $\mathcal{E}^o$ satisfying $\mathcal{M}_{\boldsymbol{H}}(\mathcal{E}) > \mathcal{M}_{\boldsymbol{H}}(\mathcal{E}^o) > 0$.

Assumption 5. For every pair $(\boldsymbol{x}, \boldsymbol{x}') \in \mathcal{X} \times \mathcal{X}$ such that $\boldsymbol{x} \leq \boldsymbol{x}'$, it is true that $g^o(\boldsymbol{x}) \leq g^o(\boldsymbol{x}')$ and $\boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{g}(\boldsymbol{x}')$.

Assumption 6. There exists a number $L^{\boldsymbol{f}}_{\mathcal{P}} < \infty$, such that, for every pair $(\boldsymbol{p}, \boldsymbol{p}') \in \mathcal{P} \times \mathcal{P}$, it is true that

$$\begin{aligned}&\|\mathbb{E}\{\boldsymbol{f}(\boldsymbol{p}(\boldsymbol{H}), \boldsymbol{H})\} - \mathbb{E}\{\boldsymbol{f}(\boldsymbol{p}'(\boldsymbol{H}), \boldsymbol{H})\}\|_\infty \\ &\quad \leq L^{\boldsymbol{f}}_{\mathcal{P}} \mathbb{E}\{\|\boldsymbol{p}(\boldsymbol{H}) - \boldsymbol{p}'(\boldsymbol{H})\|_\infty\}.\end{aligned} \tag{25}$$

Assumptions 4, 5 and 6 are reasonable and are fulfilled by most practically significant wireless resource allocation problems, as clearly explained in ([6], Section III.A).

We are now in position to present our main result. To this end, let $\mathscr{D} : \mathbb{R}^{N_g} \times \mathbb{R}^{N_S} \to (-\infty, \infty]$ be the dual function of the infinite dimensional problem (1).

Theorem 8. (Almost-Zero Duality Gaps) *Let Assumptions 1 – 6 be in effect, choose* $\mathtt{S}(\mu_R) \equiv \boldsymbol{C}\mu_R\sqrt{N_{\boldsymbol{\phi}}}, \boldsymbol{C} \geq \boldsymbol{0}$, *and suppose that, for some* $\epsilon \geq 0$, $\boldsymbol{\phi}$ *is* ϵ*-universal in* $\mathcal{P}$. *Then there exist* $\mu^{\dagger}_S > 0$ *and* $\mu^{\dagger}_R > 0$, *such that, for every* $0 \leq \mu_S \leq \mu^{\dagger}_S$ *and* $0 \leq \mu_R \leq \mu^{\dagger}_R$,

$$\boxed{-(\mu_S\Sigma^l_S + \mu_R\Sigma^l_R + \|\boldsymbol{\lambda}^*\| L^{\boldsymbol{f}}_{\mathcal{P}}\epsilon) \leq \mathscr{D}^*_{\boldsymbol{\phi},\mu} - \mathscr{P}^* \leq \mu_R\Sigma^r_R,} \tag{26}$$

where $\boldsymbol{\lambda}^* \in \arg\min_{\boldsymbol{\lambda} \geq \boldsymbol{0}} \mathscr{D}(\boldsymbol{\lambda}) \neq \emptyset$, *and* $0 \leq \Sigma^l_S \equiv \mathcal{O}(\sqrt{N_S})$, $0 \leq \Sigma^l_R \equiv \mathcal{O}(\sqrt{N_{\boldsymbol{\phi}}})$ *and* $\mathbb{R} \ni \Sigma^r_R \equiv \mathcal{O}(\sqrt{N_{\boldsymbol{\phi}}})$ *are as in Theorem 6. If, additionally,* $\boldsymbol{C} \geq \boldsymbol{c}_R$, *then the right-hand-side* (26) *is nonpositive, and may be replaced by zero.*

Theorem 8 explicitly quantifies the gap between dual optimal value of the smoothed surrogate (9) and the (primal) optimal value of the constrained variational problem (1). What is more, the gap can be made *arbitrarily small at will*, and scales *linearly* relative to the near-universality precision ϵ, *and* the smoothing parameters μ_S and μ_R.

The importance of Theorem 8 is *twofold*. *First*, together with Theorems 4 and 5, Theorem 8 provides solid technical evidence justifying the dual of (9) as a proxy for obtaining the optimal value of (1), *and* a corresponding optimal solution. This very useful, since the dual problem embeds the constraints of (9) in its objective, via the Lagrangian formulation. This can be made technically concrete via the following result.

Theorem 9. (Near–Ideal Solutions) *Let Assumptions 1 – 6 be in effect, choose* $\mathtt{S}(\mu_R) \equiv \boldsymbol{C}\mu_R\sqrt{N_{\boldsymbol{\phi}}}, \boldsymbol{C} \geq \boldsymbol{0}$, *and suppose that, for some* $\epsilon \geq 0$, $\boldsymbol{\phi}$ *is* ϵ*-universal in* $\mathcal{P}$. *Then, for any point* $(\boldsymbol{x}^\star, \boldsymbol{\theta}^\star, \boldsymbol{\lambda}^\star) \in \mathcal{X} \times \Theta \times \mathbb{R}^{N_g}_+ \times \mathbb{R}^{N_S}_+$ *such that*

1) $\mathscr{L}_{\boldsymbol{\phi},\mu}(\boldsymbol{x}^\star, \boldsymbol{\theta}^\star, \boldsymbol{\lambda}^\star) \equiv \max_{(\boldsymbol{x},\boldsymbol{\theta}) \in \mathcal{X} \times \Theta} \mathscr{L}_{\boldsymbol{\phi},\mu}(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{\lambda}^\star)$,
2) $(\boldsymbol{x}^\star, \boldsymbol{\theta}^\star)$ *is (primally) feasible for* (9), *and*
3) $\langle \boldsymbol{\lambda}^\star_S, \boldsymbol{g}_{\mu_S}(\boldsymbol{x}^\star) \rangle \equiv \langle \boldsymbol{\lambda}^\star_R, \overline{\boldsymbol{f}}^{\boldsymbol{\phi}}_{\mu_R}(\boldsymbol{\theta}^\star) - \boldsymbol{x}^\star - \mathtt{S}(\mu_R) \rangle \equiv 0$, *i.e., complementary slackness holds for* (9),

the pair $((\boldsymbol{x}^\star, \boldsymbol{\theta}^\star), \boldsymbol{\lambda}^\star)$ *is primal-dual optimal for* (9), *and* (9) *has zero duality gap. If, additionally,* $\boldsymbol{C} \geq \boldsymbol{c}_R$, *then* $(\boldsymbol{x}^\star, \boldsymbol{\theta}^\star)$ *is feasible and almost optimal for* (1), *in the sense that*

$$\boxed{|g^o(\boldsymbol{x}^\star) - \mathscr{P}^*| \leq \mathcal{C}(\mu_S\sqrt{N_S} + \mu_R\sqrt{N_{\boldsymbol{\phi}}} + \epsilon),} \tag{27}$$

where $\mathcal{C} > 0$ *is finite and problem dependent, but independent of the rest of the right-hand-side of* (27). *In other words, the untouchable* ($\mathscr{P}^*$) *becomes almost attainable (by* $(\boldsymbol{x}^\star, \boldsymbol{\theta}^\star)$*).*

Second, solving for (9) in the dual domain can be performed in a *gradient-free (model-free) fashion*, using *only* evaluations of the functions present in *both* the objective *and* constraints of (9). This makes optimal wireless resource allocation in the model-free setting possible, within a non-heuristic and predictable framework. For more details on this, see [13].

V. Simulations & Discussion

To empirically illustrate our results, we consider a simple multiuser networking scenario where each user is given a dedicated channel to communicate, with no channel interference. We wish to allocate power between users in order to maximize the weighted network sumrate, given a total expected power budget p_{max}. In such a scenario, optimal power allocation may be achieved by solving the stochastic program

$$\begin{aligned}\underset{x^i, \boldsymbol{\theta}^i, i \in \mathbb{N}^+_{N_S}}{\text{maximize}} \quad & \sum\nolimits_{i \in \mathbb{N}^+_{N_S}} w^i x^i \\ \text{subject to} \quad & x^i \leq \mathbb{E}\left\{\log\left(1 + \frac{H^i \phi^i(H^i, \boldsymbol{\theta}^i)}{\nu^i}\right)\right\}, \\ & \mathbb{E}\left\{\sum\nolimits_{i \in \mathbb{N}^+_{N_S}} \phi^i(H^i, \boldsymbol{\theta}^i)\right\} \leq p_{max} \\ & (x^i, \boldsymbol{\theta}^i) \in \mathbb{R}_+ \times \mathbb{R}^{N_{\phi^i}},\ \forall i \in \mathbb{N}^+_{N_S}\end{aligned} \tag{28}$$

where $H^i \geq 0$ and $\nu^i > 0$ are the fading power and noise variance experienced by the i-th user, and each parameterization $\phi^i : \mathbb{R}_+ \times \mathbb{R}^{N_{\phi^i}} \to [0, p_{max}]$ is chosen as a DNN with single input, two ReLU hidden layers with eight and four neurons, respectively, and a single sigmoid output, for all $i \in \mathbb{N}^+_{N_S}$. The reason for choosing N_S uncoupled DNNs, one for each user, comes from the structure of the globally optimal solution to

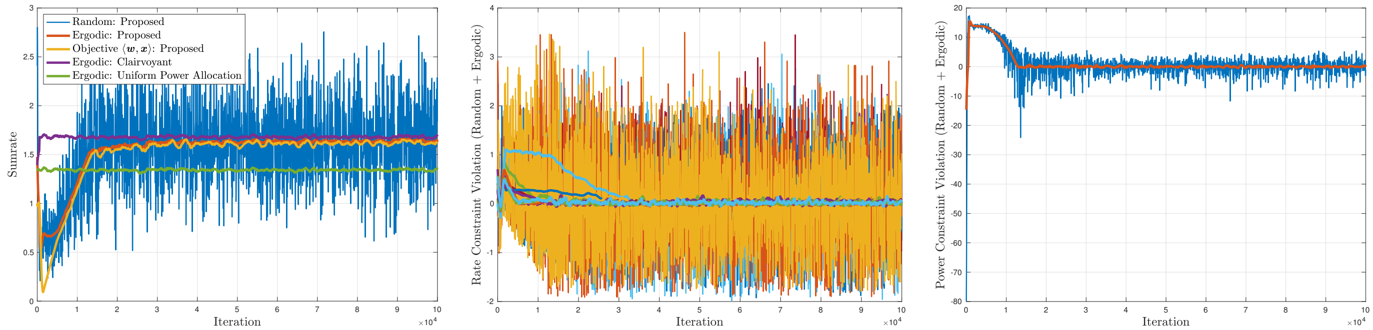

Figure 1: System performance (left), rate constraint violation (center) and power constraint violation (right), both random and ergodic, in the case of the multiuser networking scenario considered.

the most general, functional version of problem (28) (mapping to (1)), which, for this simple stylized networking setting, may be efficiently determined exactly [2].

In our example, we assume $N_S \equiv 10$ users, $p_{max} \equiv 20$, and a randomly generated weight vector $\boldsymbol{w}$. We also assume that $\nu^i \equiv 1$, and that H^i is exponentially distributed with parameter $1/2$, modeling the square of a unit variance Rayleigh channel state, for all $i \in \mathbb{N}^+_{N_S}$. Then, problem (28) is solved in a model-free fashion using the primal-dual algorithm proposed in [13] (i.e., assuming the capacity function in (28) is unknown), with null feasibility slack $\mathtt{S} \equiv \mathbf{0}$, and with a smoothing parameter $\mu_R \equiv 10^{-9}$. Since the objective of (28) is usually known to the designer, we may set $\mu_S \equiv 0$ in this case.

Fig. 1 (left) shows the evolution of objective values, achieved instantaneous sumrates, as well as an approximation of the corresponding ergodic sumrate. The latter quantity is also compared with the (approximate) ergodic performance of the unparameterized, globally optimal power policy solving (1) [2], as well as that of a deterministic uniform power policy across users. All ergodic estimates were computed via simple moving average smoothing of the respective realizations. Fig. 1 (center, right) shows similar type estimates (instantaneous and ergodic) concerning violation of the rate and power constraints of problem (28) (positive values indicate constraint violation).

Fig. 1 (left) demonstrates that the values of the objective of (28) match the values of the estimated ergodic sumrate. Additionally, from Fig. 1 (center, right) we observe that all constraints are active on average, confirming that the model-free method converges to feasible power allocation policies, as desired. At the same time, the ergodic sumrate achieved converges remarkably close to that corresponding to the optimal (model-based) policy (a *benchmark*). Thus, for our example, Theorem 9 may be verified as a *posterior solution certificate* for (28), rigorously implying a near-optimal solution for (28), *as well as* for its infinite dimensional counterpart (cf. (1)).

VI. Conclusions

In this paper, we studied the problem of optimal model-free resource allocation in wireless systems. Starting with a generic but in most cases intractable variational formulation, and by invoking classical results on Gaussian smoothing, we first constructed a finite dimensional, smoothed surrogate to the original variational problem, enabling model-free policy training. Then, assuming near-universal policy parameterizations, we quantified the duality gap between the original problem and the surrogate dual, establishing linear dependence of this gap on smoothing and universality parameters, while providing explicit near-optimality guarantees in the model-free regime.

References

[1] A. Ribeiro, "Optimal Resource Allocation in Wireless Communication and Networking," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, pp. 272, 2012.

[2] X. Wang and N. Gao, "Stochastic Resource Allocation Over Fading Multiple Access and Broadcast Channels," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2382–2391, may 2010.

[3] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An Iteratively Weighted MMSE Approach to Distributed Sum-Utility Maximization for a MIMO Interfering Broadcast Channel," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4331–4340, sep 2011.

[4] K. Gatsis, M. Pajic, A. Ribeiro, and G. J. Pappas, "Opportunistic Control Over Shared Wireless Channels," *IEEE Transactions on Automatic Control*, vol. 60, no. 12, pp. 3140–3155, dec 2015.

[5] J. Rubio, O. Muñoz-Medina, A. Pascual-Iserte, and J. Vidal, "Stochastic Resource Allocation with a Backhaul Constraint for the Uplink," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*. sep 2016, pp. 1–6, IEEE.

[6] M. Eisen, C. Zhang, L. F. Chamon, D. D. Lee, and A. Ribeiro, "Learning Optimal Resource Allocations in Wireless Systems," *IEEE Transactions on Signal Processing*, vol. 67, no. 10, pp. 2775–2790, may 2019.

[7] M. Eisen and A. Ribeiro, "Optimal Wireless Resource Allocation with Random Edge Graph Neural Networks," *arXiv preprint, arXiv:1909.01865*, sep 2019.

[8] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to Optimize: Training Deep Neural Networks for Interference Management," *IEEE Transactions on Signal Processing*, vol. 66, no. 20, pp. 5438–5453, oct 2018.

[9] T. Van Chien, E. Björnson, and E. G. Larsson, "Sum Spectral Efficiency Maximization in Massive MIMO Systems: Benefits from Deep Learning," *arXiv preprint, arXiv:1903.08163*, mar 2019.

[10] W. Cui, K. Shen, and W. Yu, "Spatial Deep Learning for Wireless Scheduling," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 6, pp. 1248–1261, dec 2019.

[11] W. B. Powell, "A Unified Framework for Stochastic Optimization," jun 2019.

[12] Y. Nesterov and V. Spokoiny, "Random Gradient-Free Minimization of Convex Functions," *Foundations of Computational Mathematics*, vol. 17, no. 2, pp. 527–566, apr 2017.

[13] D. S. Kalogerias, M. Eisen, G. J. Pappas, and A. Ribeiro, "Model-Free Optimal Stochastic Resource Allocation in Wireless Communications," *IEEE Trans. on Signal Processing, under review*, 2019.

[14] S. P. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[15] A. P. Ruszczyński, *Nonlinear Optimization*, Princeton University Press, 2006.