A new nonsmooth optimal control framework for wind turbine power systems☆

Hesham Abdelfattah [a,b], Sameh A. Eisa [a],*, Peter Stechlinski [c]

[a] Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati, Cincinnati, 45221, OH, USA
[b] Department of Mathematical Sciences, University of Cincinnati, Cincinnati, 45221, OH, USA
[c] Department of Mathematics and Statistics, University of Maine, Orono, 04469, ME, USA

ARTICLE INFO

ABSTRACT

Optimal control theory extending from the calculus of variations has not been used to study the wind turbine power system (WTPS) control problem, which aims at achieving two targets: (i) maximizing power generation in lower wind speed conditions; and (ii) maintaining the output power at the rated level in high wind speed conditions. A lack of an *optimal control* framework for the WTPS (i.e., no access to actual optimal control trajectories) reduces optimal control design potential and prevents competing control methods of WTPSs to have a reference control solution for comparison. In fact, the WTPS control literature often relies on reduced and linearized models of WTPSs, and avoids the nonsmoothness present in the system during transitions between different conditions of operation. In this paper, we introduce a novel optimal control framework for the WTPS control problem. We use in our formulation a recent accurate, nonlinear differential–algebraic equation (DAE) model of WTPSs, which we then generalize over all wind speed ranges using nonsmooth functions. We also use developments in nonsmooth optimal control theory to take into account nonsmoothness present in the system. We implement this new WTPS optimal control approach to solve the problem numerically, including (i) different wind speed profiles for testing the system response; (ii) real-world wind data; and (iii) a comparison with smoothing and naive approaches. Results show the effectiveness of the proposed approach.

1. Introduction

According to the U.S. Department of Energy, as can be seen in this report [1], wind energy is recognized as a prominent contributor to electric-power capacity growth in the U.S. Consequently, considerable efforts have been dedicated in the past two decades to enhancing the performance of wind turbine power systems (WTPSs) and integrating them with traditional power systems and the power grid. However, given their distinct nature compared to traditional power systems, integrating WTPSs with the existing power grid remains a significant challenge [2] that is being studied by national labs, industry and by academics [3–6]. Challenges in WTPSs include inherent fluctuations, limited predictability, reduced grid stabilization inertia, and heightened sensitivity to uncontrolled factors such as weather conditions. Consequently, research endeavors encompassing control systems, optimization, and power generation/storage have significantly surged over the past two decades [5–19], where major efforts have been made to have a better understanding of the response of the WTPS to sudden changes of uncontrolled parameters, such as the wind speed, power grid loads and the bus frequency or voltage [4–6,9,11,20–22].

1.1. Modeling and control of WTPSs

Traditionally, there are two control inputs/strategies for WTPSs [5,10]: (i) pitch angle control, and (ii) torque/tip-ratio control. Using the first strategy (i.e., pitch control), we can control the power produced from wind by changing the pitch angle [4–6]. In the second strategy, the torque/tip-ratio based control methods enable changing the rotor speed of the WTPS to maximize power production; a common strategy/method for torque control is known as maximum power point tracking (MPPT) where the objective is to maximize power production through changing the speed of the turbine [10,15,23,24]. In this paper, we focus mainly on pitch angle control. Pitch control allows one to change the pitch angle in order to control the aerodynamic loads on the rotor. This affects the power coefficient, known as C_p, of the wind turbine, which in turn changes the amount of power produced from the WTPS [4–6]. Active pitch control methods are used [4–6,10] to control the pitch angle typically in Region 3, where the generator torque is usually kept constant and the pitch angle is varied to keep the generated electric power at the rated level. Pitch control has been traditionally implemented and designed using proportional–integral–derivative (PID) controllers [4–6,10,11]. There are some studies aimed at improving the efficiency of the PID controllers, such as using estimation and optimization techniques [12] and genetic algorithm techniques [25]. Additionally, feedforward, Model predictive control (MPC), sliding mode and other controllers have been also considered (e.g., [14,18,26,27]). It is worth noting that some works in literature have considered both pitch and torque/tip-ratio controls simultaneously such as [28,29].

In WTPSs control literature, many studies use linearized/reduced and/or models in the frequency/s-domain (e.g., [3,4,30–34]) while testing and analyzing newly developed control methods. Recently, significant progress has been made in WTPS modeling [5,6]. In said recent works, WTPSs have been rigorously studied and converted into a full scale nonlinear time-domain differential–algebraic-equation (DAE) model based on physical/industrial literature such as, but not limited to [3,4,11]. This nonlinear time domain WTPS DAE model [5,6] will be used in the control study of this paper (in Section 3).

1.2. Developments in nonsmooth ODE/DAE Optimal Control

Traditional optimal control methods require differentiability of the system's dynamics as well as the objective functional. However, many applications, such as the one this paper is concerned about, i.e., WTPSs, exhibit nonsmoothness in the dynamics and the objective functional due to transitions between different Regions of operation (see Fig. 3 in Section 3); in fact, a detailed study on nonsmoothness appearing in the WTPS parameters can be found in [35]. Therefore, an extension of traditional optimal control methods is desired to overcome this issue (i.e., study nonsmooth problems). Two popular methods in the literature used to resolve the nonsmoothness issue include using slack variables and smoothing approximations [36–40]. Both the slack variables and the smoothing approximation methods have limitations. Using the slack variables method increases the size of the problem which results in a slower convergence and higher complexity especially for large systems, while using a smoothing approximation fundamentally changes the optimization problem with the accuracy of approximation depending on the choice of the smoothing parameters, which adds an element of arbitrariness; in fact, this has been observed directly with relevance to the WTPS [35].

Unlike slack variables or smoothing approaches, frameworks now exist that enable the direct treatment of nonsmoothness. For example, there are methods that extend standard indirect methods (in which calculus of variations is used to furnish a boundary value problem that can be solved to obtain a trajectory that minimizes/maximizes a given objective functional [36,41]) for optimal control to nonsmooth problems, including constrained nonsmooth dynamical systems (see Ch.5 in [42] and the work in [43]). However, these results generally come with restrictive conditions, and so are practically implementable typically only in special cases, such as linear complementarity systems [44]. As for nonsmooth direct optimal control methods, lexicographic directional derivative (LD-derivative) [45] based nonsmooth optimal control methods for ODE systems and DAE systems were introduced recently in [46,47], respectively. The advantage of such methods is that LD-derivatives satisfy sharp calculus rules, which makes them a practical tool for calculating computationally relevant generalized gradient elements of the objective functional. LD-derivatives based nonsmooth optimal control methods will be presented in Section 2 as we will use it in the optimal control study of this paper. Moreover, nonsmooth optimization methods [48] can utilize LD-derivatives based methods. Some studies have applied LD-derivative based methods to solve nonsmooth optimal control problems involving nonsmooth systems with constraints, including nonsmooth DAE systems in [47], mixed nonlinear complementarity systems [49], and DAEs with optimality criteria [50].

1.3. Motivation and contributions

Despite the significant progress made in WTPS control strategies (as discussed in Section 1.1), current research efforts often suffer from three major issues. First, model linearization and reduction that can lead to inaccurate conclusions about the control method or its performance (see relevant discussions on this in [5,16,51]). Second, it is quite common for WTPSs to experience nonsmooth (non-differentiable) profiles in their dynamics, parameters and/or the objective functional (see for example [35]); however, to the authors' best knowledge, no work has been done that directly addresses the nonsmoothness present in the WTPS model for control studies. Overcoming nonsmoothness by smoothing has been shown to be less effective in studying WTPSs as shown in [35]. Third, there is a lack of optimal control methods extending from the calculus of variations in the sense of [36,41,52] (such as direct methods) that find the optimal feasible trajectory for use as a reference control solution/design tool, which can be used by different control methods for realization and/or comparison efforts. We are motivated to provide a new *theory and framework* to address these issues. To overcome the first issue, the recent nonlinear time domain WTPS DAE model introduced in [5,6] (see relevant discussion in Section 1.1) will be used for the following reasons: (i) to the best of the authors' knowledge, it is the only WTPS model that has

been shown to be mathematically well-posed (the existence and uniqueness properties have been proved in [51]); (ii) this model is shown to be amenable to available, non-commercial, stiff numerical solvers of DAE systems [5]; (iii) it has been shown in [53] that reductions based on linear/multiple-scale approximations can be made, however, the separation between fast and slow variables is quite complicated and is not easy to construct, suggesting against reduction/linearization; and (iv) it is accurate when validated against real-world measured data (see [5,6,53]) and outperformed some of the most cited models that existed in the literature as asserted in [5,6,53]. The second and third issues can be addressed using the developments in nonsmooth optimal control theory for dynamical systems with DAE dynamics [47] (as discussed in Section 2). These developments with relevance to the work of this paper will be re-introduced in Section 2. A simpler case study for a reduced version of the mentioned WTPS model [5,6] in a particular wind speed range and smooth setting has been conducted successfully by the authors in [16], which further suggests that the work of this paper is with significant potential and is highly needed.

In this paper, we make the following contributions:

1. We unify and generalize the modeling and control framework for WTPSs which includes/covers all wind speed ranges: low, moderate and high (Regions 1–3, see Fig. 3). This generalized framework is achieved by making use of nonsmooth functions which enable continuous transition between all different operational conditions (i.e., Regions 1–3) of the WTPS. Our generalized framework is based on the recently developed time-domain, nonlinear, validated against real-world data, DAE WTPS models in [5,6]. We did not perform any linearization/reduction or smoothing approximation to have the generalized framework as accurate, and reflective to WTPS dynamics, as possible.

2. We provide a novel optimal control method that is suitable for (and specialized to) this *nonsmooth* WTPS generalized modeling and framework using tools from generalized derivatives theory. We thus provide, for the first time, a mathematically rigorous framework (extended from the calculus of variations) that aims to find the feasible optimal trajectory for the pitch angle controller that achieves the WTPS objectives in all Regions of wind conditions simultaneously (i.e., in Regions 1–3, see Fig. 3): (i) maximizing the power generation (in lower wind speed conditions in Regions 1 and 2); and (ii) keeping the output power at the rated level (in high wind speed conditions in Region 3). For that, we use the established results in [47] to formulate a nonsmooth sequential optimal control (i.e., a direct method) approach for the WTPS. Furthermore, we test our approach and present the optimal control response of WTPSs when exposed to wind variations with different wind profiles (Gaussian and ramp profiles). We also apply our approach to, and present results of, the highly turbulent, real-world wind data that was used in modeling validation in [5,6,53].

3. We demonstrate the importance of this novel approach, which treats the nonsmoothness directly and is based on state-of-the-art accurate models of WTPSs via: (i) comparison with a smoothing approach/approximation (which is shown to introduce error); and (ii) comparison with commercial smooth optimal control solver, namely GPOPS-II [54], where the nonsmoothness is ignored (which is shown to fail sometimes).

Given all the contributions listed above, we believe that the optimal control response our generalized framework produces should provide a benchmark (reference for comparison) to assess the performance of any control strategy applied to WTPSs.

1.4. Organization of the paper

In Section 2, we provide a background and mostly self-contained material on solving nonsmooth optimal control via lexicographic differentiation. What is provided in Section 2 is customized to the class of DAE systems this paper is concerned with. Then, in Section 3, we provide our first contribution (unification and generalization of WTPS DAE models using nonsmooth functions). Following that, in Section 4, we provide our second contribution (novel nonsmooth optimal control approach). In Section 5, we provide our third contribution, as we compare the novel nonsmooth approach derived and tested in Section 4 with a smoothing approximation and a smooth optimal control solver, to demonstrate the effectiveness of our new approach. We conclude the paper in Section 6.

2. Solving nonsmooth optimal control problems via LD-Derivatives

As discussed in [5,35], a WTPS can be formulated as a DAE system in the semi-explicit form. Therefore, we turn to the work in [47], which provides the generalized gradient elements of the objective functional for nonsmooth optimal control problems with (possibly nonsmooth) DAE systems embedded. In particular, the DAE optimal control problem of interest is as follows:

$$\min_{\mathbf{u}} \quad \phi(\mathbf{u}) = \phi(t_f, \mathbf{u}(t_f), \mathbf{x}(t_f), \mathbf{y}(t_f)), \tag{1a}$$

$$\text{s.t.} \quad \dot{\mathbf{x}}(t) = \mathbf{h}(t, \mathbf{u}(t), \mathbf{x}(t), \mathbf{y}(t)), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \tag{1b}$$

$$\mathbf{0}_{n_y} = \mathbf{g}(t, \mathbf{x}(t), \mathbf{y}(t)), \tag{1c}$$

with differential state variables $\mathbf{x}(t) \in D_x \subseteq \mathbb{R}^{n_x}$ (with initial condition $\mathbf{x}_0 \in D_x$), algebraic state variables $\mathbf{y}(t) \in D_y \subseteq \mathbb{R}^{n_y}$, control $\mathbf{u}(t) \in D_u \subseteq \mathbb{R}^{n_u}$, and finite time interval $[t_0, t_f] \subset D_t \subseteq \mathbb{R}$, and RHS functions $\mathbf{h} : D_t \times D_u \times D_x \times D_y \to \mathbb{R}^{n_x}$ and $\mathbf{g} : D_t \times D_x \times D_y \to \mathbb{R}^{n_y}$, where D_t, D_u, D_x, D_y are open and connected sets. In this paper we use the sequential method (also known as direct single shooting) to solve the optimal control problem in (1); in the sequential method, the continuous-time problem is converted into an nonlinear programming (NLP) by discretizing the control input to produce a finite set of decision variables — the trajectory of the embedded dynamical system (in this case, the DAE system) is approximated via numerical simulation, and so the decision variables in the NLP

come from the parameterization of the control [36,52]. Namely, the optimal control problem in (1) seeks to minimize the objective functional over some function space, i.e., $\min_{\mathbf{u} \in \mathcal{U}} \phi(\mathbf{u})$ (where, e.g., $\mathcal{U}$ could be the space of Lebesgue integrable functions on some finite time horizon). Instead, we convert this problem into a finite-dimensional problem by parametrically discretizing the control into a finite truncation of basis functions, using the control parameters $\mathbf{p} \in D_p \subseteq \mathbb{R}^{n_p}$, and then we instead solve $\min_{\mathbf{p} \in D_p} \phi(\mathbf{p})$. In this paper, we focus on a piecewise constant control discretization on $n_s \in \mathbb{N}$ subintervals $[\tau_0, \tau_1], \dots, [\tau_{n_s-1}, \tau_{n_s}]$:

$$\mathbf{u}(t) = \mathbf{u}_{(i)} \quad \text{if } t \in (\tau_{i-1}, \tau_i], \tag{2}$$

with pre-fixed times τ_i satisfying $t_0 = \tau_0 < \tau_1 < \dots < \tau_{n_s-1} < \tau_{n_s} = t_f$, with $\tau_i - \tau_{i-1} = (t_f - t_0)/n_s$, and with control parameters $\mathbf{p} = (\mathbf{u}_{(1)}, \dots, \mathbf{u}_{(n_s)}) \in D_p = D_u^{n_s} \subseteq \mathbb{R}^{n_u n_s}$ (i.e., $\mathbf{u}_{(i)} \in D_u$ for each i, and $n_p = n_u n_s$ here). The sequential approach uses the gradient information $\nabla \phi(\mathbf{p})$ to update the control parameters in the spirit of $\mathbf{p}_{k+1} = \mathbf{p}_k - \gamma \nabla \phi(\mathbf{p}_k)$ for some step size γ and repeats the process until a termination criteria is satisfied (so that $\nabla \phi(\mathbf{p}_k) \approx \mathbf{0}$). Moreover, the gradient information is obtained using the sensitivity functions $\frac{\partial \mathbf{x}}{\partial \mathbf{p}}$ and $\frac{\partial \mathbf{y}}{\partial \mathbf{p}}$ associated with the embedded DAE system via

$$\nabla \phi(\mathbf{p}) = \frac{\partial \phi}{\partial \mathbf{p}} + \frac{\partial \phi}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{p}} + \frac{\partial \phi}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \mathbf{p}}. \tag{3}$$

However, motivated by WTPSs, we are interested in the case where the maps $\phi, \mathbf{h}, \mathbf{g}$ are assumed to be nonsmooth and thus not necessarily differentiable. Consequently, the solutions $\mathbf{x}$ and $\mathbf{y}$ of the DAE system may be nonsmooth with respect to $\mathbf{u}$, invalidating (3). To overcome this issue, we turn our attention to generalized derivatives theory and lexicographic differentiation to compute (generalized) derivative information of ϕ for use in a nonsmooth sequential direct method to solve (1).

2.1. Lexicographic directional derivatives

Let $X \subseteq \mathbb{R}^n$ be open and $\mathbf{f} : X \to \mathbb{R}^m$ be locally Lipschitz on X. Clarke's generalized derivative [42] of $\mathbf{f}$ at $\mathbf{x}_0 \in X$ is given by the convex hull of the Bouligand subdifferential of $\mathbf{f}$, i.e.,

$$\partial_C \mathbf{f}(\mathbf{x}_0) := \text{conv } \partial_B \mathbf{f}(\mathbf{x}_0), \tag{4}$$

where the Bouligand subdifferential is defined as:

$$\partial_B \mathbf{f}(\mathbf{x}_0) := \left\{ \mathbf{F} \in \mathbb{R}^{m \times n} : \exists \mathbf{x}_i \to \mathbf{x}_0 \text{ s.t. } \mathbf{x}_i \in X \setminus Z_{\mathbf{f}} \, \forall i \in \mathbb{N} \text{ and } \mathbf{Jf}(\mathbf{x}_i) \to \mathbf{F} \right\},$$

where $Z_{\mathbf{f}}$ is the zero measure subset on which $\mathbf{f}$ is not differentiable and $\mathbf{Jf}(\mathbf{x}_i)$ is the Jacobian matrix of $\mathbf{f}$ at $\mathbf{x}_i$. Note that in case $\mathbf{f}$ is C^1 at $\mathbf{x}_0$, then $\partial_B \mathbf{f}(\mathbf{x}_0) = \partial_C \mathbf{f}(\mathbf{x}_0) = \{\mathbf{Jf}(\mathbf{x}_0)\}$.

Although elements of Clarke's generalized derivative are useful for nonsmooth numerical methods, it does not satisfy sharp calculus rules (e.g. sum rule — see [55]), which makes it difficult to calculate said elements of Clarke's generalized derivative for complicated functions. The lexicographic directional derivative (LD-derivative) [45], which is based on lexicographic differentiation [56], was proposed to overcome this limitation (and other limitations [55]), which is applicable to lexicographically smooth (L-smooth) functions [56]. The class of L-smooth functions are locally Lipschitz and directionally differentiable to arbitrary order, which includes all C^1 functions, PC^1 functions (hence abs-value, min, max), and convex functions (hence any p-norm), as well as the compositions of these functions. Following from the definition of L-smooth functions, Nesterov [56] defined the lexicographic derivative (L-derivative) for L-smooth functions which can be seen as a Jacobian-like object, as discussed in [45,56,57], and is just as useful in nonsmooth numerical methods as Clarke generalized derivative elements.

An LD-derivative, denoted $\mathbf{f}'(\mathbf{x}_0; \mathbf{M})$, can be used to calculate an L-derivative of an L-smooth function $\mathbf{f}$ if the so-called "directions matrix" $\mathbf{M} = [\mathbf{m_1} \quad \mathbf{m_2} \quad \dots \quad \mathbf{m_k}] \in \mathbb{R}^{n \times k}$ has full row rank. Roughly, the columns in $\mathbf{M}$, which span the domain space if $\mathbf{M}$ has full row rank, act as probing directions. The main advantage of the LD-derivative formulation is that it satisfies sharp calculus rules and that there are closed-form expressions available for nonsmooth elemental functions such as max, min, absolute-value, etc. (see [45,55]):

- smooth rule: $\mathbf{f} \in C^1 \Rightarrow \mathbf{f}'(\mathbf{x}_0; \mathbf{M}) = \mathbf{Jf}(\mathbf{x}_0)\mathbf{M}$, e.g. $\sin'(x_0; \mathbf{M}) = \cos(x_0)\mathbf{M}$.
- abs rule: $\text{abs}'(x_0; \mathbf{M}) = \text{fsign}(x_0, \mathbf{M})\mathbf{M}$ where the first-sign function $\text{fsign}(\cdot)$ returns the sign of the first nonzero element, or zero if its input is zero.
- min rule: for $\mathbf{M}_1, \mathbf{M}_2 \in \mathbb{R}^{1 \times k}$,

$$\min{}'(x_0, y_0; (\mathbf{M}_1, \mathbf{M}_2)) = \mathbf{slmin}([x_0 \quad \mathbf{M}_1], [y_0 \quad \mathbf{M}_2])$$

$$= \begin{cases} \mathbf{M}_1 & \text{if } \text{fsign}\left(\begin{bmatrix} x_0 \\ \mathbf{M}_1^\mathsf{T} \end{bmatrix} - \begin{bmatrix} y_0 \\ \mathbf{M}_2^\mathsf{T} \end{bmatrix}\right) \leq 0, \\ \mathbf{M}_2 & \text{if } \text{fsign}\left(\begin{bmatrix} x_0 \\ \mathbf{M}_1^\mathsf{T} \end{bmatrix} - \begin{bmatrix} y_0 \\ \mathbf{M}_2^\mathsf{T} \end{bmatrix}\right) > 0, \end{cases}$$

  i.e., the shifted-lexicographic-minimum $\mathbf{slmin}$ returns the lexicographic minimum of the two vector arguments, left-shifted by one element.
- componentwise rule: $\mathbf{f}'(\mathbf{x}_0; \mathbf{M}) = (f_1'(\mathbf{x}_0; \mathbf{M}), f_2'(\mathbf{x}_0; \mathbf{M}), \dots, f_n'(\mathbf{x}_0; \mathbf{M}))$.
- chain rule: $[\mathbf{g} \circ \mathbf{f}]'(\mathbf{x}_0; \mathbf{M}) = \mathbf{g}'(\mathbf{f}(\mathbf{x}_0); \mathbf{f}'(\mathbf{x}_0; \mathbf{M}))$.
- sum rule: $[\mathbf{f} + \mathbf{g}](\mathbf{x}_0; \mathbf{M}) = \mathbf{f}'(\mathbf{x}_0; \mathbf{M}) + \mathbf{g}'(\mathbf{x}_0; \mathbf{M})$.
- product rule: $[fg](\mathbf{x}_0; \mathbf{M}) = f'(\mathbf{x}_0; \mathbf{M})g(\mathbf{x}_0) + f(\mathbf{x}_0)g'(\mathbf{x}_0; \mathbf{M})$.

Because of these sharp calculus rules, the LD-derivative can be computed accurately, hence, making it a computationally-relevant tool that can be effectively used to furnish an L-derivative from solving a system of linear equations. Namely,

$$\mathbf{f}'(\mathbf{x}_0; \mathbf{M}) = \mathbf{J}_L \mathbf{f}(\mathbf{x}_0; \mathbf{M})\mathbf{M}, \tag{5}$$

where $\mathbf{J}_L \mathbf{f}(\mathbf{x}_0; \mathbf{M})$ is the L-derivative of $\mathbf{f}$ at $\mathbf{x}_0$ in the directions $\mathbf{M}$, a computationally relevant generalized derivative object. For more details on LD-derivative theory and some of its applications, the reader may refer to [45,55].

2.2. Generalized gradients of control problems with nonsmooth DAEs embedded

Returning to the optimal control problem in (1), we recall some results from the literature that enable the use of LD-derivatives in nonsmooth optimal control settings: For nonsmooth optimal control with ODE systems, the framework presented in [46] provides the generalized gradient elements of the objective functional with nonsmoothness present in the objective functional and/or the system dynamics, where such description of generalized gradient elements of the objective functional is based on the sensitivity analysis results in [57] concerning generalized derivatives of ODE solutions with respect to the parameters of the system. The authors in [58] established a similar sensitivity analysis theory for semi-explicit DAEs with "generalized differentiation index one", which can be exploited to furnish generalized gradient elements of the objective functional of an optimal control problem with an embedded DAE [47], in the same spirit as [46]. In particular, a generalized gradient element of the objective function in (1) can be obtained via the following result, which specializes Theorem 3.2 in [47].

Theorem 1. *Suppose that the following conditions hold:*

(i) *Assume that $\phi, \mathbf{h}, \mathbf{g}$ are L-smooth functions.*
(ii) *Assume that $\mathbf{u}$ is parameterized according to (2).*
(iii) *Assume that there exists a solution $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ of (1b)–(1c) on $[t_0, t_f]$ through the initial data $\{(t_0, \mathbf{p}_0, \mathbf{x}_0, \mathbf{y}_0)\}$, for some $\mathbf{p}_0 \in D_p = D_u^{n_s}$, such that*

$$\pi_{\mathbf{y}} \partial \mathbf{g}(t, \tilde{\mathbf{x}}(t), \tilde{\mathbf{y}}(t)) = \{\mathbf{G}_{\mathbf{y}} \in \mathbb{R}^{n_y \times n_y} : \exists [\mathbf{G}_t \ \mathbf{G}_{\mathbf{x}} \ \mathbf{G}_{\mathbf{y}}] \in \partial \mathbf{g}(t, \tilde{\mathbf{x}}(t), \tilde{\mathbf{y}}(t))\} \tag{6}$$

contains no singular matrices for any $t \in [t_0, t_f]$.

Then a generalized gradient of ϕ in (1) can be calculated according to

$$\boldsymbol{\mu} = \phi'(t_f, \mathbf{u}_{(n_s)}, \tilde{\mathbf{x}}(t_f), \tilde{\mathbf{y}}(t_f); \mathbf{0}_{1 \times n_u n_s}, \mathbf{e}_{(n_s)}^{\mathrm{T}} \otimes \mathbf{I}_{n_u}, \tilde{\mathbf{X}}(t_f), \tilde{\mathbf{Y}}(t_f))^{\mathrm{T}} \in \partial \phi(\mathbf{p}_0), \tag{7}$$

where the LD-derivative sensitivities $(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ solve the following for $i = 1, \ldots, n_s$:

$$\dot{\mathbf{X}}(t) = \mathbf{h}'(t, \mathbf{u}_{(i)}, \tilde{\mathbf{x}}(t), \tilde{\mathbf{y}}(t); (\mathbf{0}_{1 \times n_s n_u}, \mathbf{e}_{(i)}^{\mathrm{T}} \otimes \mathbf{I}_{n_u}, \mathbf{X}(t), \mathbf{Y}(t))), \quad \forall t \in (\tau_{i-1}, \tau_i],$$

$$\mathbf{0}_{n_y \times n_s n_u} = \mathbf{g}'(t, \tilde{\mathbf{x}}(t), \tilde{\mathbf{y}}(t); (\mathbf{0}_{1 \times n_s n_u}, \mathbf{X}(t), \mathbf{Y}(t))), \quad t \in [t_0, t_f],$$

$$\mathbf{X}(t_0) = \mathbf{0}_{n_x \times n_s n_u}, \tag{8}$$

$$\mathbf{0}_{n_y \times n_s n_u} = \mathbf{g}'(t_0, \mathbf{x}_0, \mathbf{y}_0; (\mathbf{0}_{1 \times n_s n_u}, \mathbf{0}_{n_x \times n_s n_u}, \mathbf{Y}(t_0))),$$

where $\otimes$ is the Kronecker product and $\mathbf{e}_{(i)} \in \mathbb{R}^{n_s}$ denotes the ith standard basis vector.

Proof. The DAE system in (1b)–(1c), with controls $\mathbf{u}$ parameterized according to (2), is in the form of the embedded DAE system in Eq. (6) in [47], with $\mathbf{f}$ replaced by $\mathbf{h}$, $\mathbf{f}_0(\mathbf{p}) = \mathbf{x}_0$ a constant, the control $\mathbf{v}$ absent here, and extra problem parameters absent here. Hence, the L-smoothness of $\phi, \mathbf{h}, \mathbf{g}$ here guarantees conditions (i)-(ii) and (v) of Assumption 3.1 in [47] hold. Moreover, we can rewrite $\mathbf{u}$ in (2) as

$$\mathbf{u}(t) = \mathbf{u}(t; \mathbf{p}) = \sum_{i=1}^{n_s} \mathbf{u}_{(i)} \mathcal{X}_{(\tau_{i-1}, \tau_i]}(t),$$

i.e., $\Psi_i(t) = \mathcal{X}_{(\tau_{i-1}, \tau_i]}(t)$ is the indicator function, so that condition (iii) of Assumption 3.1 in [47] holds. (Condition (iv) holds vacuously.) The assumptions on the RHS functions $\mathbf{h}, \mathbf{g}$ and the piecewise constant control $\mathbf{u}$ imply that the solution $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ on $[t_0, t_f]$ is unique, and the assumption that (6) contains no singular matrices guarantees that said solution is regular in the terminology of [47]. Hence, Theorem 3.2 in [47] is applicable, and the conclusions in this theorem follow immediately from it, and Example 3.4 in [47], with directions matrix $\mathbf{M} = \mathbf{I}_{n_s n_u}$, since in that case

$$\boldsymbol{\Psi}(t) = [\psi_{(1)}(t) \quad \cdots \quad \psi_{(n_s)}] \otimes \mathbf{I}_{n_u} = \mathbf{e}_{(i)}^{\mathrm{T}} \otimes \mathbf{I}_{n_u} \quad \forall t \in (\tau_{i-1}, \tau_i]. \quad \square$$

Remark 2.1. The assumptions in Theorem 1, including existence and regularity of a solution of (1b)–(1c) on $[t_0, t_f]$, are in line with direct optimal control methods for smooth systems, except of course for the allowance of $\phi, \mathbf{h}, \mathbf{g}$ to be nonsmooth. In fact, the function $\mathbf{h}$ is permitted to be discontinuous with respect to t in the theory in [47], and both $\mathbf{h}$ and $\mathbf{g}$ may depend on extra problem
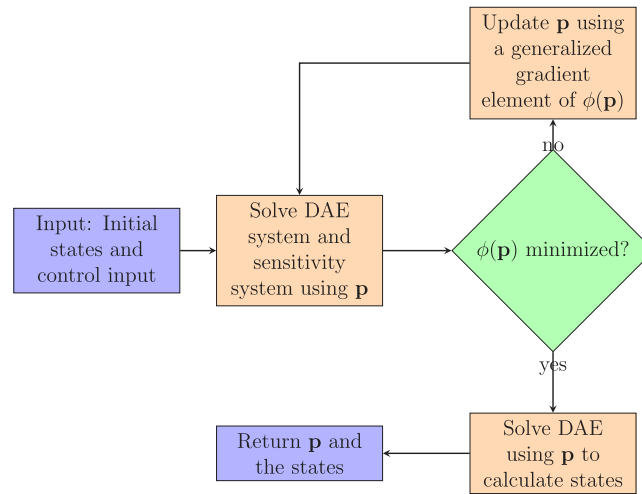
Fig. 1. Simple flow chart clarifying the sequential optimal control process.

parameters, although we omit these generalizations (and use $\mathbf{M} = \mathbf{I}_{n_s n_u}$) to make the exposition clearer. Note also that if $\phi, \mathbf{h}, \mathbf{g}$ are C^1, $\tilde{\mathbf{X}}(t) = \frac{\partial \tilde{\mathbf{x}}}{\partial \mathbf{p}}(t; \mathbf{p}_0)$, $\tilde{\mathbf{Y}}(t) = \frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{p}}(t; \mathbf{p}_0)$, $\boldsymbol{\mu} = \nabla \phi(\mathbf{p}_0)$ in Theorem 1 (hence classical results are recovered). Lastly, we note that the objective functional ϕ in (1a) above is in Mayer form, but Bolza/Lagrange forms of the objective functional are also permitted with the theory in Theorem 1 since Bolza/Lagrange can be transformed into Mayer form (see, e.g., [59]), as we demonstrate in the next example.

For an intuitive illustration of the sequential optimal control process, the reader can refer to the simple flow chart in Fig. 1. We also give an illustration of how the sequential LD-derivatives based optimal control method can be applied through the following example from [36] (in this case, using the ODE theory in [46] for comparative purposes).

Example 2.2. Consider a block that moves between the two points $x_1 = 0$ and $x_1 = 1$, where x_1 is the position of the block, within the time interval $[0, 1]$, with the block's velocity, x_2, being zero at both the initial time, $t_0 = 0$, and the final time, $t_f = 1$. The control input u to the system is the force applied to the left side of the block and no friction is assumed. We consider a nonsmooth objective functional as used in [36] that minimizes the absolute work done. The optimal control problem can be then described as:

$$\min_u \phi(u) = \int_0^1 |u(\tau) x_2(\tau)| d\tau, \tag{9a}$$

$$\text{s.t. } \dot{x}_1 = x_2 =: h_1, \qquad x_1(0) = 0, \qquad x_1(1) = 1, \tag{9b}$$

$$\dot{x}_2 = u =: h_2, \qquad x_2(0) = 0, \qquad x_2(1) = 0, \tag{9c}$$

In using the sequential method, the continuous-time problem is transformed into an NLP by discretizing the control according to Eq. (2) and having the system dynamics in Eqs. (9b)–(9c) converted to a set of constraints applied to the states and control at discrete time points. A numerical ODE solver is used to compute the values of the states at these time points and, in the final step, the NLP solver is provided with an initial guess in the form of reference control and states.

In order to apply the framework presented in [46], we need the objective functional (9a) to be in the Mayer form. This can be done by introducing a new state variable $x_3(t)$, and equating the time derivative of that new variable to the integrand of the objective functional, i.e.

$$\dot{x}_3(t) = |u(t) x_2(t)| =: h_3, \tag{10}$$

with $x_3(0) = 0$. So $n_x = 3$ and $n_u = 1$ and now the objective functional becomes

$$\min_u \phi(u) = \min_u x_3(t_f), \tag{11}$$

which is in Mayer form. Parameterizing u as a piecewise constant according to (2), i.e., $u(t) = u_{(i)}$ on $(\tau_{i-1}, \tau_i]$, $i = 1, \ldots, n_s$, with $n_s = 100$, the LD-derivative sensitivity function $\tilde{\mathbf{X}}(t) = (\tilde{\mathbf{X}}_1(t), \tilde{\mathbf{X}}_2(t), \tilde{\mathbf{X}}_3(t)) \in \mathbb{R}^{n_x \times n_s n_u} = \mathbb{R}^{3 \times 100}$ solves

$$\dot{\mathbf{X}}_1(t) = \mathbf{X}_2(t), \quad \forall t \in [0, 1],$$

$$\dot{\mathbf{X}}_2(t) = \mathbf{e}_{(i)}^{\mathrm{T}}, \quad \forall t \in (\tau_{i-1}, \tau_i],$$

$$\dot{\mathbf{X}}_3(t) = \mathrm{fsign}\big(u_{(i)} \tilde{\mathbf{X}}_2(t), \tilde{x}_2(t) \mathbf{e}_{(i)}^{\mathrm{T}} + u_{(i)} \mathbf{X}_2(t)\big) \times \big(\tilde{x}_2(t) \mathbf{e}_{(i)}^{\mathrm{T}} + u_{(i)} \mathbf{X}_2(t)\big), \quad \forall t \in (\tau_{i-1}, \tau_i],$$

$$\mathbf{X}_1(0) = \mathbf{0}_{1 \times n_s}, \quad \mathbf{X}_2(0) = \mathbf{0}_{1 \times n_s}, \quad \mathbf{X}_3(0) = \mathbf{0}_{1 \times n_s},$$
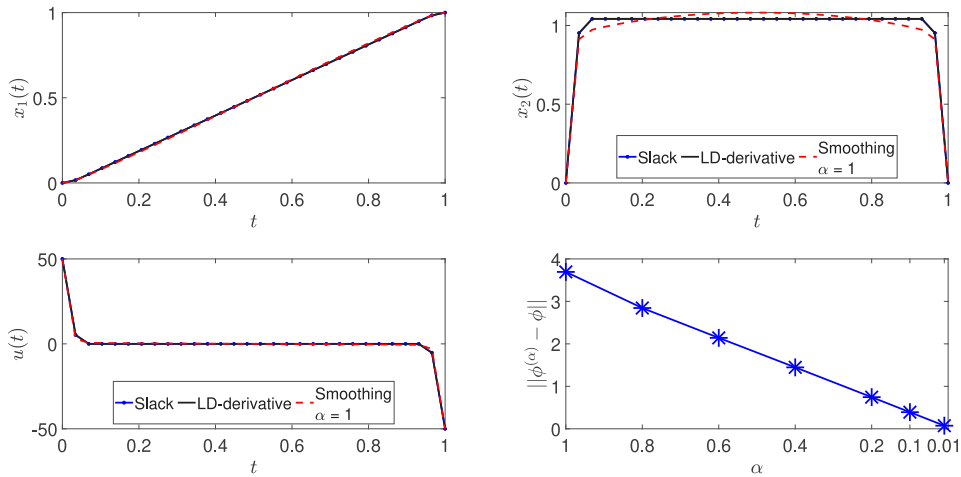
Fig. 2. This figure shows the results of Example 2.2 using the objective functional in (11) (with piecewise constant control on $n_s = 100$ subintervals of $[0, 1]$) via slack variables, smoothing approximation and LD-derivatives based sequential method. In the left panel, we see the optimal trajectory for the position $x_1(t)$ (upper panel) and the velocity $x_2(t)$ (lower panel). In the right panel, we show the control input signal $u(t)$ (upper panel) similar to the states in the left panel. In addition, we compare the 2-norm difference between the smoothed objective functional, $\phi^{(\alpha)}$, evaluated at different values of the smoothing parameter α, and the nonsmooth objective functional ϕ.

where the LD-derivative rule

$$[\text{abs} \circ f]'(z; \mathbf{M}) = \text{abs}'(f(z); f'(z; \mathbf{M})) = \text{fsign}(f(z), f'(z; \mathbf{M})) \times f'(z; \mathbf{M})$$

has been used with $z = (u, x_2)$ and $f(z) = ux_2$ here. Then, according to Theorem 1 (with algebraic equations absent, i.e., using [46, Theorem 6 and Corollary 7]) for some reference control parameters $\mathbf{p}_0 = (u_{(1)}, \ldots, u_{(n_s)})$,

$$\boldsymbol{\mu} = \left(\tilde{\mathbf{X}}_3(t_f) \right)^{\mathrm{T}} \in \partial \phi(\mathbf{p}_0). \tag{12}$$

In Fig. 2, the results of the LD-derivatives based sequential method (above) is compared to/validated against the slack variables approach used in [36] for this same problem, and a smoothing approach using the following approximation of the absolute-value function (as provided in [36]):

$$|x| \approx x \tanh\left(\frac{x}{\alpha}\right), \tag{13}$$

for small $\alpha > 0$.

It can be seen from the results in the upper right panel of Fig. 2 that the LD-derivatives based method produces an accurate solution when compared with the solutions from the slack variables and the smoothing approximation method (with a proper choice of the smoothing parameter α), while avoiding the limitations of both of these methods (computational complexity of the slack variables and arbitrariness of smoothing approximation). We also used different values of the smoothing parameter α to show the dependency of this method's accuracy on the smoothing parameter, e.g., choosing $\alpha = 1$ provides a much more accurate solution than choosing $\alpha = 5$. In the bottom of the right panel of Fig. 2, we see the 2-norm of difference between the objective functional, $\phi^{(\alpha)}$, evaluated at different values of the smoothing parameter α, and the nonsmooth objective functional ϕ. As can be seen, a decrease in the value of α results in a better accuracy for the smoothing approximation. However, it remains unclear how to choose α to achieve a desired error, and it remains unclear in general whether such a smoothing approximation converges to the "true" solution. This, in part, motivates us to treat the nonsmooth problem directly (i.e., an LD-derivatives based approach).

3. Unifying and generalizing WTPS modeling and control in Regions 1–3 using nonsmooth functions

Nonsmoothness in the WTPS arises in three ways: (i) nonsmoothness in the parameters can be present due to extreme variations or situations, which can be modeled as nonsmooth profiles [35]; (ii) nonsmoothness in the control input; and (iii) nonsmoothness from unifying and generalizing the WTPS modeling framework. In this part, we provide a unified and generalized WTPS control problem and modeling framework that is operable for all Regions 1–3 (as illustrated in Fig. 3), by introducing nonsmooth RHS functions and objective functional to the model. Afterward, in Section 4, we treat the introduced nonsmoothness directly using the LD-derivatives based sequential optimal control method outlined in Section 2.

3.1. WTPS DAE framework

First we provide the WTPS DAE model we use in this study (from [5,6]) for the reasons (i)-(iv) we stated in Section 1.3. The reader is directed to the book chapter [6] for very detailed steps on the model derivations based on accurate physical/industrial considerations. Additionally, in said book chapter [6], it is mentioned how different control conditions and design considerations can amend the model for the user's need.

For example, one can take different settings for the reactive power control design, among many other options. However, with the generalization and unification in mind, and the steps provided in this section, one should be able to build on the results here to construct other particular frameworks customized to their problem of interest. For our control problem, we consider the control input $\mathbf{u}$ to be the pitch angle θ, i.e. $u = \theta$. Additionally, we consider the two mass model representation of electric and mechanical torques, (14a)–(14c), as this is known, physically, to provide more accurate representation [5,6] of the conversion of mechanical torque (due to turbine blades rotation via wind) into electrical torque (provided to the generator). Moreover, we consider the power control (14d)–(14e) to operate without optional/extra [5,6] inertia controls, frequency-bus control, or power storage. Also, the reactive power control is represented by the power factor mode (14f)–(14g) [5,6]. Furthermore, the electric variables representing the active and reactive voltage/current delivered to the grid are represented by (14h)–(14j) [5,6]. Lastly, the terminal voltage (at the connection between the wind turbine and the grid) is represented by the infinite-bus model [5,6], commonly used in literature for general testing of power sources interaction with the grid; this can be written algebraically based on Kirchhoff's laws as in (14k). Hence, the WTPS DAE system is given as follows:

$$\dot{w}_g = \frac{1}{2H_g}\left[-\frac{P_{elec}}{w_g + w_0} - D_{tg}(w_g - w_t) - K_{tg}\Delta\theta_m\right] =: h_1, \tag{14a}$$

$$\dot{w}_t = \frac{1}{2H}\left[\frac{P_{mech}}{w_t + w_0} + D_{tg}(w_g - w_t) + K_{tg}\Delta\theta_m\right] =: h_2, \tag{14b}$$

$$\Delta\dot{\theta}_m = w_{base}(w_g - w_t) =: h_3, \tag{14c}$$

$$\dot{f}_1 = w_g + w_0 - w_{ref} =: h_4, \tag{14d}$$

$$\dot{P}_{inp} = \frac{1}{T_{pc}}\left[(w_g + w_0)(K_{ptrq}(w_g + w_0 - w_{ref}) + K_{itrq}f_1) - P_{inp}\right] =: h_5, \tag{14e}$$

$$\dot{P}_{1elec} = [P_{elec} - P_{1elec}]/T_{pwr} =: h_6, \tag{14f}$$

$$\dot{V}_{ref} = K_{Qi}[Q_{cmd} - Q_{gen}] =: h_7, \tag{14g}$$

$$\dot{E}_{qcmd} = K_{vi}[V_{ref} - V] =: h_8, \tag{14h}$$

$$\dot{E}_q = [E_{qcmd} - E_q]/0.02 =: h_9, \tag{14i}$$

$$\dot{I}_{plv} = [P_{inp}/V - I_{plv}]/0.02 =: h_{10}, \tag{14j}$$

$$0 = \left(V^2\right)^2 - \left[2(P_{elec}R + Q_{gen}X) + E^2\right]V^2 + \left(R^2 + X^2\right)\left(P_{elec}^2 + Q_{gen}^2\right) =: g, \tag{14k}$$

where $Q_{gen} = V(E_q - V)/X_{eq}$, $Q_{cmd} = \tan(PFE)P_{1elec}$, PFE is small angle, w_{ref} is given by (18), $P_{elec} = I_{plv}V$,

$$P_{mech} = \frac{1}{2}C_p(\lambda, \theta)\rho A_r v_{wind}^3 = \frac{1}{2}\left(\sum_{i=0}^{4}\sum_{j=0}^{4}\alpha_{i,j}\theta^i\lambda^j\right)\rho A_r v_{wind}^3, \tag{15}$$

where v_{wind} is the wind speed and $\lambda = \frac{(w_t + w_0)}{v_{wind}}$ is the tip ratio. The system in (14) has differential state variables $\mathbf{x} = (w_g, w_t, \Delta\theta_m, f_1, P_{inp}, P_{1elec}, V_{ref}, E_{qcmd}, E_q, I_{plv})$ (i.e., $n_x = 10$) and algebraic state variable $y = V$ (i.e., $n_y = 1$) representing the terminal voltage, where $w_g, w_t, \Delta\theta_m, f_1, P_{inp}, P_{1elec}, V_{ref}, E_{qcmd}, E_q,$ and I_{plv} are the generator speed, the dynamic turbine speed, the integral of difference between w_t and w_g, the integral of differences of speeds, the power order, the filtered electrical power, the reference voltage, the reactive voltage command, and the generator reactive variable, respectively. Denotation of all of the state variables and the parameters, as well as the values of the parameters are discussed and taken from [5,6], see Appendix.

3.2. Unifying regions: Nonsmooth wind profiles and shaft speed

We now introduce the wind profiles, representing wind speed variations, that we use to test/analyze the optimal control response in studying WTPSs. For that, we first consider a nonsmooth ramp wind profile which represents changes in the wind speed level (low slope for slow transitions and high slope, near discontinuity, for extreme and fast changes). This nonsmooth generalized ramp wind speed profile was also introduced in [35] as:

$$v_{wind,1}(t) = v_0 + m(\max(0, t - t_{on}) - \max(0, t - t_{off})), \tag{16}$$

where v_0, m, t_{on}, t_{off} are the initial wind speed before the start of the ramp, the slope of the ramp, the time at which the ramp starts, and the time at which the ramp ends, respectively. For wind gusts or disturbances, Gaussian wind speed profiles can be used (with
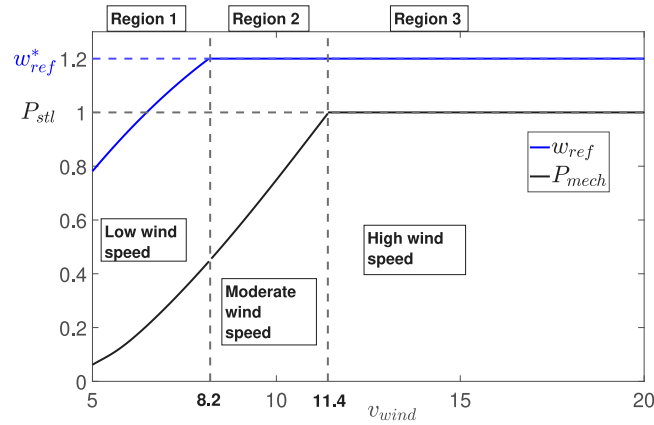
Fig. 3. How w_{ref} and P_{mech} change as v_{wind} changes. There are three Regions of operation corresponding to different wind speed ranges. w_{ref} transitions through nonsmoothness between Regions 1 and 2 at $v_{wind} = 8.2$ m/s [5,6] and P_{mech} transitions through nonsmoothness between Regions 2 and 3 at about $v_{wind} = 11.4$ m/s [5,6], respectively.

large variance for calm gusts and very small variance, such that the wind profile will be close to a delta function, for extreme gusts), and can be expressed as:

$$v_{wind,2}(t) = v_0 + \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}(\frac{t-\mu}{\sigma})^2}, \qquad (17)$$

where v_0 is the initial wind speed, μ and σ are the mean value and the standard deviation of the normal distribution, respectively.

Next, we note that having the system (1b)–(1c) applicable to all Regions of operations (with Regions 1–2 representing low and moderate wind speeds that are below the rated wind speed, and Region 3 representing high wind speeds above the rated wind speed [60], as illustrated in Fig. 3) will require some unification and generalization of conditions relevant to Regions 1–3. This unification and generalization necessarily introduces nonsmoothness in the dynamics and the objective functional of the control problem. We start with the reference shaft speed, w_{ref}, which controls the rotor. It can be modeled as a quadratic polynomial function of the generated electric power, $w_{ref} = -0.75P_{elec}^2 + 1.59P_{elec} + 0.63$, in Region 1 of operations (low wind speeds) due to physical considerations [5,6]. However, once the wind speeds are high enough (Regions 2–3), w_{ref} should be kept at its rated value (constant) $w_{ref} = w_{ref}^*$, e.g. $w_{ref}^* = 1.2$ for $v_{wind} \geq 8.2$ m/s [5,6]. Thus, unifying w_{ref} in Regions 1–3 can be done by introducing the nonsmooth min function:

$$w_{ref} = \min(-0.75P_{elec}^2 + 1.59P_{elec} + 0.63, w_{ref}^*). \qquad (18)$$

We can see from Eq. (18) that for lower wind speeds ($v_{wind} < 8.2$ m/s), the quadratic expression in the left argument of the min function is less than w_{ref}^*, i.e., $-0.75P_{elec}^2 + 1.59P_{elec} + 0.63 < w_{ref}^*$, so the output of the min function will be $w_{ref} = -0.75P_{elec}^2 + 1.59P_{elec} + 0.63$. At $v_{wind} = 8.2$ m/s, we have the quadratic expression exactly equal to w_{ref}^*, i.e., $-0.75P_{elec}^2 + 1.59P_{elec} + 0.63 = w_{ref}^*$, hence, we have $w_{ref} = -0.75P_{elec}^2 + 1.59P_{elec} + 0.63 = w_{ref}^*$. Finally, at $v_{wind} > 8.2$ m/s, the quadratic expression gives a value higher than w_{ref}^*, hence the output of the min function becomes $w_{ref} = w_{ref}^*$. It is common to take $w_{ref}^* = 1.2$ [5,6].

3.3. Unifying the control objectives over all regions: A nonsmooth objective functional

As for generalizing an expression for P_{mech} for all wind speeds (Regions 1–3 in Fig. 3), it is not as straightforward as what we did with w_{ref}; P_{mech} is the objective itself that the WTPS should be designed to achieve [5,6]: (i) maximization of P_{mech} when physically it is not possible to produce the rated level, P_{stl}, which is usually normalized to be 1 Per Unit [5,6], due to lower wind speeds (Regions 1–2); and (ii) maintaining P_{mech} at the rated level, i.e., $P_{mech} = P_{stl}$ when physically P_{mech} can exceed P_{stl} (Region 3). Hence, unification of Regions 1–3 can be achieved by generalizing the control objective itself, i.e., the objective functional. In fact, the authors in [61] introduced an objective function that achieves the above-mentioned control objective. Within the context of an optimal control problem, their expression of the objective functional can be expressed as:

$$\max_u \int_{t_0}^{t_f} \min(P_{mech}(\mathbf{x}(t), u(t)), P_{stl})dt. \qquad (19)$$

In (19) whenever P_{mech} is less than the rated power, P_{stl}, the min function will return the P_{mech} argument which leads to maximization of power production; if P_{mech} is more than the rated power P_{stl}, then the min function will return the constant P_{stl} argument. One

major problem with (19) is that it does not have an isolated maximum when the min function returns the constant P_{stl} argument (one side of the function is constant). To resolve this issue, we introduce a well-defined new objective functional which achieves the same objective of (19) but with an isolated maximum. This generalized objective functional that will unify the control objective over Regions 1–3 is expressed as:

$$\max_u \int_{t_0}^{t_f} \omega(\mathbf{x}(t), u(t)) dt = \phi(u), \tag{20}$$

where

$$\omega(\mathbf{x}(t), u(t)) = \min(P_{stl}, P_{mech}(\mathbf{x}(t), u(t))) - \min(0, P_{stl} - P_{mech}(\mathbf{x}(t), u(t))) \times (P_{stl} - P_{mech}(\mathbf{x}(t), u(t))) =: h_{11}. \tag{21}$$

In relevance to our results in the next subsection where we introduce our optimal control formulation, one needs to have the objective functional in the Mayer form. To convert (20) from Lagrange to Mayer form, we introduce an auxiliary variable x_{aux} as:

$$\dot{x}_{aux}(t) = h_{11}, \qquad x_{aux}(t_0) = 0. \tag{22}$$

With this definition, we get $\phi(u) = \int_{t_0}^{t_f} \omega(\mathbf{x}(t), u(t)) dt = x_{aux}(t_f)$, and hence (20) is replaced by

$$\max_u \phi(u) = \max_u x_{aux}(t_f). \tag{23}$$

This objective functional, again, is generalized over Regions 1–3 and unifies both the control objectives (maximizing power production at lower wind speeds and maintaining the power production at the rated level for higher wind speeds). To elaborate further, observe that in (21) whenever P_{mech} is less than the rated power (which happens at low wind speeds, where we need to maximize power production), then what we get from Eq. (21) is $P_{mech} - 0$, which leads to the maximization (minimization with negative sign) of power production. However, if P_{mech} is greater than the rated power P_{stl}, then Eq. (21) returns $P_{stl} - (P_{stl} - P_{mech})^2$, where the function is not constant anymore (hence, we resolve the issue of (20)) and then the difference $(P_{stl} - P_{mech})^2$ is minimized, keeping P_{mech} at the rated level P_{stl}. This generalization of the objective functional over all wind speeds is achieved thanks to the presence of nonsmooth functions in the objective functional.

4. LD-derivatives based method: A novel optimal control approach for WTPSs

Note that the ODE associated with h_2 contains P_{mech}, and thus the piecewise control $u = \theta$ (i.e., $n_u = 1$). We also note that for w_{ref} and the objective functional, we use the nonsmooth expressions in (18) and (23), respectively, that were introduced in the previous section. Recall from (22) that to have the objective functional in Mayer form we need to introduce an extra differential state variable, x_{aux}, to the system of Eqs. in (14). Hence, we get a total of eleven differential state variables and one algebraic state variable (i.e., $n_x = 11$ and $n_y = 1$) and the time derivative of this extra state variable will appear coupled with the system (14) as given in Eq. (22). With the control parameters $\mathbf{p} = (\mathbf{u}_{(1)}, \ldots, \mathbf{u}_{(n_s)})$ in the discretization

$$\mathbf{u}(t) = \mathbf{u}_{(i)} \quad \text{if } t \in (\tau_{i-1}, \tau_i], \tag{24}$$

as the decision variables, we used the MATLAB® built-in NLP solver, fmincon, to solve the NLP problem arising from this problem. (We note that τ_i can also be taken as decision variables for more accuracy but in our case we only consider controlling $u_{(i)}$.) Because participating functions are nonsmooth, conventional methods (such as fmincon) may fail. To overcome this, the LD-derivative sensitivity functions as well as the generalized gradient elements of the objective functional obtained via Theorem 1 are supplied to the solver fmincon (which does not require such inputs in the smooth setting, as this method can generate these objects itself given a smooth ODE/DAE control problem).

Assuming that (14) admits a regular solution on $[t_0, t_f]$, i.e. (6) is satisfied along the solution, it must hold that $\frac{\partial g}{\partial V} \neq 0$ along the solution since in this case g is smooth and we have that $\pi_y \partial g = \{\frac{\partial g}{\partial V}\}$ (see Eq. (8) in [35]). In this case, Theorem 1 yields a generalized gradient of (7) by building the auxiliary system whose solutions are LD-derivative sensitivity functions. This can be accomplished by applying LD-derivative rules, as outlined in Section 2.2. The sensitivity equations associated with g and h_j, $j \notin \{4, 5, 11\}$, are the classical sensitivity equations:

$$\dot{\mathbf{X}}_{x_j}(t) = \frac{\partial h_j}{\partial u} \mathbf{e}_{(i)}^{\mathrm{T}} + \frac{\partial h_j}{\partial \mathbf{x}} \mathbf{X}(t) + \frac{\partial h_j}{\partial V} \mathbf{V}(t), \qquad t \in (\tau_{i-1}, \tau_i], \tag{25}$$

with partial derivatives evaluated at $(t, u_{(i)}, \tilde{\mathbf{x}}(t), \tilde{V}(t))$, e.g.,

$$\dot{\mathbf{X}}_{E_q}(t) = \frac{1}{0.02} \left[\mathbf{X}_{E_{qcmd}}(t) - \mathbf{X}_{E_q}(t) \right].$$

Similarly, the algebraic equation Eq. (14k) gives the following sensitivity equation:

$$\mathbf{0}_{1 \times n_s} = \frac{\partial g}{\partial E_q} \mathbf{X}_{E_q}(t) + \frac{\partial g}{\partial I_{plv}} \mathbf{X}_{I_{plv}}(t) + \frac{\partial g}{\partial V} \mathbf{Y}(t). \tag{26}$$

For the nonsmooth functions h_4, h_5, and $h_{11} := \dot{x}_{aux}$, we apply the theory of LD-derivatives to get the corresponding sensitivity equations, i.e., for h_4 we get:

$$\begin{aligned}
\dot{\mathbf{X}}_{f_1}(t) &= h_4'(t, u(t), \tilde{\mathbf{x}}(t), \tilde{V}(t); (\mathbf{0}_{1 \times n_s}, \mathbf{e}_{(i)}^{\mathrm{T}}, \mathbf{X}(t), \mathbf{Y}(t))) \\
&= \mathbf{X}_{w_g}(t) - [w_{ref}]' \\
&= \mathbf{X}_{w_g}(t) - \left[\min(-0.75 P_{elec}^2 + 1.59 P_{elec} + 0.63, w_{ref}^*) \right]' \\
&= \mathbf{X}_{w_g}(t) - \mathbf{slmin}([q(P_{elec}) \quad (-1.5 P_{elec} + 1.59) P_{elec}'], [w_{ref}^* \quad \mathbf{0}_{1 \times n_s}]), \quad \forall t \in (\tau_{i-1}, \tau_i],
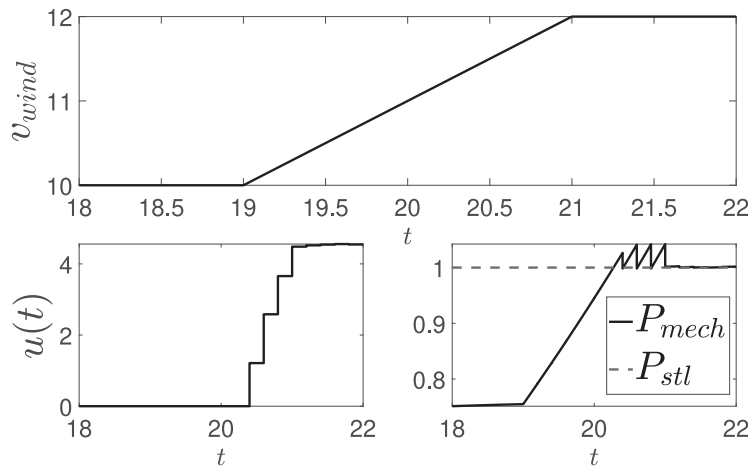\end{aligned} \tag{27}$$

Fig. 4. The upper panel shows the ramp profile of v_{wind} with an initial wind speed $v_0 = 10$ m/s, and a final wind speed $v_f = 12$ m/s. The lower left panel shows the optimal control response, $u(t)$, to the change in v_{wind}. The lower right panel shows the effect of the control response on P_{mech}.

where $q(P_{elec}) = -0.75 P_{elec}^2 + 1.59 P_{elec} + 0.63$ and $P'_{elec} = \frac{\partial P_{elec}}{\partial u} \mathbf{e}_{(i)}^{\mathrm{T}} + \frac{\partial P_{elec}}{\partial \mathbf{x}} \mathbf{X}(t) + \frac{\partial P_{elec}}{\partial V} \mathbf{Y}(t)$. Similarly, for h_5 we have:

$$\dot{\mathbf{X}}_{P_{inp}}(t) = \frac{1}{T_{pc}} \left[K_{itrq}(w_g + w_0) \right] \mathbf{X}_{f_1}(t) - \frac{\mathbf{X}_{P_{inp}}(t)}{T_{pc}} + \left[K_{ptrq}(w_g + w_0 - w_{ref}) + K_{itrq} f_1 + K_{ptrq}(w_g + w_0) \right] \frac{\mathbf{X}_{w_g}(t)}{T_{pc}}$$
$$- \frac{K_{ptrq}(w_g + w_0)}{T_{pc}} \left[w_{ref} \right]'. \tag{28}$$

For h_{11} we get:

$$\dot{\mathbf{X}}_{x_{aux}}(t) = \left[\min(P_{stl}, P_{mech}) - \min(0, P_{stl} - P_{mech})(P_{stl} - P_{mech}) \right]'$$
$$= \mathbf{slmin} \left(\begin{bmatrix} P_{stl} & \mathbf{0}_{1 \times n_s} \end{bmatrix}, \begin{bmatrix} P_{mech} & \frac{\partial P_{mech}}{\partial u} \mathbf{e}_{(i)}^{\mathrm{T}} + \frac{\partial P_{mech}}{\partial \mathbf{x}} \mathbf{X}(t) \end{bmatrix} \right)$$
$$- \mathbf{slmin} \left(\begin{bmatrix} 0 & \mathbf{0}_{1 \times n_s} \end{bmatrix}, \begin{bmatrix} P_{stl} - P_{mech} & -\frac{\partial P_{mech}}{\partial u} \mathbf{e}_{(i)}^{\mathrm{T}} - \frac{\partial P_{mech}}{\partial \mathbf{x}} \mathbf{X}(t) \end{bmatrix} \right) \times (P_{stl} - P_{mech})$$
$$+ \min(0, P_{stl} - P_{mech}) \left(\frac{\partial P_{mech}}{\partial u} \mathbf{e}_{(i)}^{\mathrm{T}} + \frac{\partial P_{mech}}{\partial \mathbf{x}} \mathbf{X}(t) \right). \tag{29}$$

Recall that the initial conditions for $\mathbf{X}(t_0)$ are zero, and it follows from Eq. that $\mathbf{Y}(t_0)$ must satisfy

$$\mathbf{0}_{1 \times n_s n_u} = g'(t_0, \mathbf{x}_0, \mathbf{y}_0; (\mathbf{0}_{1 \times n_s n_u}, \mathbf{0}_{n_x \times n_s n_u}, \mathbf{Y}(t_0))) = \frac{\partial g}{\partial V} \mathbf{Y}(t_0)$$

with $\frac{\partial g}{\partial V} \neq 0$ by regularity, so that $\mathbf{Y}(t_0) = \mathbf{0}_{1 \times n_s n_u}$ must hold. Finally, the generalized gradient $\boldsymbol{\mu}$ of the objective functional in (23) can be calculated as: $\boldsymbol{\mu} = (\mathbf{X}_{x_{aux}}(t_f))^{\mathrm{T}} \in \partial \phi(\mathbf{p}_0)$.

4.1. Simulation results of LD-derivatives based method

In order to test the optimal control formulation presented in the previous subsections, we expose the WTPS to variations in the wind speed, v_{wind}. We consider three simulation cases for v_{wind}: the ramp profile in Eq. (16), the Gaussian profile in Eq. (17), and a wind profile produced from actual wind speed data (real-world measurements) that was used to validate the DAE model (14) as illustrated in [5,6,53]. We note that we have carefully chosen our profiles such that it guarantees the WTPS will transition between Regions 1–3 in Fig. 3. Hence, we test directly the effectiveness of our approach when the system passes through unavoidable nonsmoothness. We note that the reader can refer to this GitHub repository [62] for the MATLAB files used to produce the results discussed in this subsection.

In the upper panel of Fig. 4, we show the first case of a ramp wind profile with an initial wind speed $v_0 = 10$ m/s and a final wind speed $v_f = 12$ m/s. We run the simulation from $t_0 = 18$ s to $t_f = 22$ s, with the number of time steps $n_s = 20$, and with the start and end times of the ramp at $t_{on} = 19$ s and $t_{off} = 21$ s, respectively. The corresponding optimal control response, $u(t)$, to the change in v_{wind} is depicted in the lower left panel, and the effect of the control response on P_{mech} is depicted in the lower right panel. We notice that there is no change in $u(t)$ until about $v_{wind} = 11.4$ m/s, which is the critical value of v_{wind} at which there is nonsmoothness in the objective functional (23) (i.e. $P_{mech} = P_{stl}$). After reaching that point, $u(t)$ increases to keep P_{mech} at the rated power $P_{stl} = 1$.

In the upper panel of Fig. 5 we show the second case with a Gaussian wind profile. We run the simulation from $t_0 = 18$ s to $t_f = 22$ s, with the number of time steps $n_s = 20$, $v_0 = 10$ m/s, mean $\mu = 20$, and standard deviation $\sigma = 0.5$. The corresponding
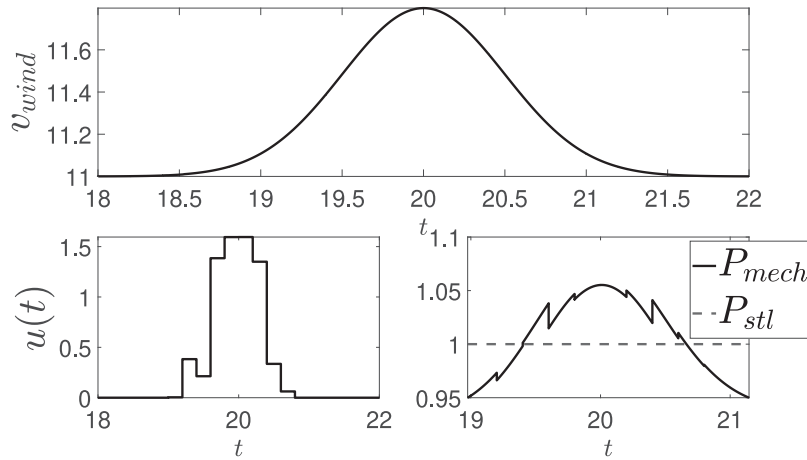
Fig. 5. The upper panel shows the Gaussian profile of v_{wind} with an initial wind speed $v_0 = 10$ m/s, and standard deviation $\sigma = 0.5$. The lower left panel shows the optimal control response, $u(t)$, to the change in v_{wind}. The lower right panel shows the effect of the control response on P_{mech}.
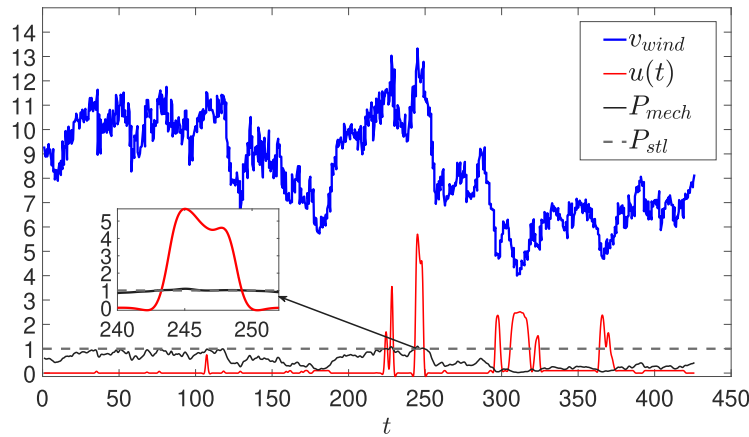

Fig. 6. Our approach handling a real-world, high-turbulent wind data. The wind speed profile is shown as well as the corresponding control response and the resulting P_{mech}. The zoomed-in portion illustrates more clearly how the optimal control solution and the resulting P_{mech}.

optimal control to the change in v_{wind} is depicted in the lower left panel, and the effect of the control response on P_{mech} is depicted in the lower right panel. This change of $u(t)$ keeps the value of P_{mech} close to the value of the rated power $P_{stl} = 1$.

Finally, Fig. 6 shows actual real-world wind speed data used for data validation in [5,6,53]. Here, we take the number of time steps to be $n_s = 425$. In Fig. 6, the wind speed profile, v_{wind}, is shown as well as the corresponding control response and the resulting P_{mech}. We provide a close-up look of one Region of interest with very abrupt changes in wind speed. It is clear that the mechanical power is maximized for low v_{wind} when P_{mech} is less than the rated value P_{stl}. However, for higher wind speeds, the power is maintained at the rated value $P_{stl} = 1$. This shows that the solution obtained using the sequential LD-derivative optimal control method maximizes the nonsmooth objective functional in a highly-turbulent wind profile where the system passes through points of nonsmoothness transitioning between Regions 1–3. Moreover, it can be seen that in this scenario the optimal control implementation was on a large time horizon of varying wind speeds, which demonstrates the effectiveness of the proposed framework and its general applicability to WTPSs.

Having verified the LD-derivatives based sequential nonsmooth optimal control method for the WTPS control problem via simulations, we provide the following observations, including the application potential of our optimal control method:

1. It is noticeable that in general the optimal control trajectory seems to be close/matching in its shape to the wind profile which is useful information in control design.
2. The current framework can adopt more control inputs (e.g., torque/tip-ratio based control inputs).
3. Bounds and limiters as desired by the user, or similar to what is implemented in industry (e.g., [6]), can be easily added as extra constraints to this nonsmooth optimal control formulation.
4. The current work can be extended to involve reactive-power-control using the Q Droop function (see [5]) as a control input in fault cases.

5. For practical purposes, the proposed framework is useful in the design and analysis phase of control designs of WTPSs as it provides the optimal control trajectory (without linearization of the system or approximations of the nonsmooth functions present in the system and/or the objective functional). This can be then used to judge the effectiveness of other control methods (including real-time control methods). That is, the optimal control solution can serve as a reference of performance assessment to developing/proposed control designs.

6. The provided method can be implemented in near-real-time setting with the aid of high processing power and/or parallel processing handling the NLP.

We also provide the following limitations:

1. We are not able to use second-order derivatives information within the optimal control problem due to the current absence of second order LD-derivatives theory in the literature.

2. As is the case with other popular optimal control methods in the same vein, a suitable choice of initial conditions for the states and controls remains a challenge in our formulation.

In the next section, we make the case clearer for why the proposed nonsmooth LD-derivatives based sequential optimal control approach is needed as opposed to other alternatives.

5. Comparison between our proposed LD-derivatives based method and alternative approaches for the WTPS control problem

In this section, we consider alternative approaches to solving the WTPS optimal control problem. In particular, instead of treating the nonsmoothness directly, like the proposed LD-derivative based method above, we consider methods that avoid the nonsmoothness altogether. We broadly group possible alternatives into two approaches: (i) smoothing approaches (where nonsmooth functions are approximated by smooth functions and standard methods are then employed) and (ii) "naive" approaches (where standard methods are immediately employed and any theoretical/computational issues stemming from the nonsmoothness are simply ignored).

Before proceeding, we note that in our comparative results we consider the nonsmoothness present in the objective functional by limiting the wind speed range over Regions 2–3. We also interpolated the piecewise constant control solution for improved visualizations in all figures in this section. We note that for all simulations in this section, we have used the fmincon package in MATLAB®, where we specify the following options:

- We use the sequential quadratic programming (SQP) algorithm to solve the NLP problems in all of the approaches presented below.
- For the LD-derivatives based solutions, we provide LD-derivative information as the gradient information to be used in its corresponding NLP problem.
- We use the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm in all of the approaches presented below, which provides an approximation to the Hessian matrix in the NLP problems. In the case of the LD-derivatives based approach, the BFGS uses the first-order LD-derivative information.

5.1. Comparison with a smoothing approach

We approximate the nonsmooth objective functional in (23) (by replacing the nonsmooth expression ω in (21)) using a known logarithmic approximation for the min function [63]. Thus, the smoothed version is:

$$\omega \approx \widehat{\omega} = \frac{\ln(e^{-NP_{stl}} + e^{-NP_{mech}})}{-N} - \frac{\ln(e^0 + e^{-N(P_{stl} - P_{mech})})}{-N}(P_{stl} - P_{mech}), \tag{30}$$

where N is the smoothing parameter, with $\widehat{\omega} \to \omega$ as $N \to \infty$. Now, we solve the WTPS optimal control problem in a similar manner to the previous section (our LD-derivatives based approach) but with a smooth optimal control solver after using the smoothed objective functional above. Two results are presented in Figs. 7–8. Fig. 7 shows the applied ramp wind profile and how different choices of N affect the accuracy of the solution. The upper panel shows the profile of v_{wind} with an initial wind speed $v_0 = 10$ m/s, a final wind speed $v_f = 12$ m/s, start time and end time of the ramp at $t = 19$ s, $t = 21$ s, respectively. The lower left panel shows the optimal control response, $u(t)$, to the change in v_{wind} using the introduced LD-derivatives based sequential nonsmooth optimal control approach with Eq. (23) as the objective functional, compared with the smoothed objective functional in Eq. (30) with different values of the smoothing parameters N. The lower right panel shows the effect of each of the control response trajectories on P_{mech}. The trajectory of P_{mech} corresponding to the (optimal) trajectory of u obtained using the LD-derivatives based approach effectively drives the dynamics as we expect, based on the nonsmooth objective functional in Eq. (23) (i.e. maximizing P_{mech} when $P_{mech} < P_{stl}$, and maintaining P_{mech} at the rated level, i.e., $P_{mech} = P_{stl}$, when it is physically possible to produce $P_{mech} > P_{stl}$). As for the smoothing approximation approach, it seems that the smoothed trajectories converge to the optimal trajectory as the value of N gets larger. Similarly, in Fig. 8, we show the applied Gaussian wind profile with $v_0 = 11$ m/s, mean $\mu = 20$, and standard deviation $\sigma = 0.5$. Similar to the ramp wind profile case, the optimal trajectory given by the LD-derivatives based approach is effective in steering the system's dynamics as desired, according to objective functional in Eq. (23) and the criteria described above in the ramp wind profile case. In this case also, the effectiveness of the smoothed approach depends on how large N is. We emphasize that, here
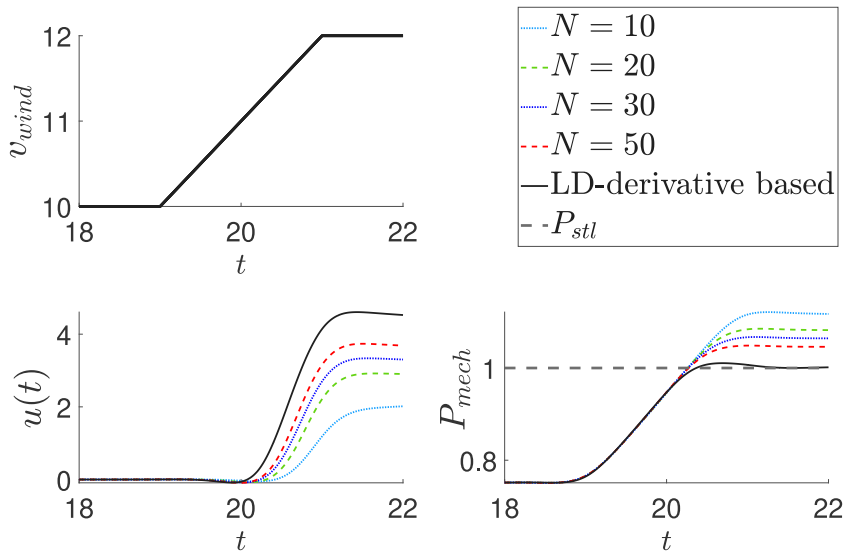
Fig. 7. Smoothing approximation with different values of the smoothing parameters N vs. LD-derivatives based sequential approach with a ramp wind profile.
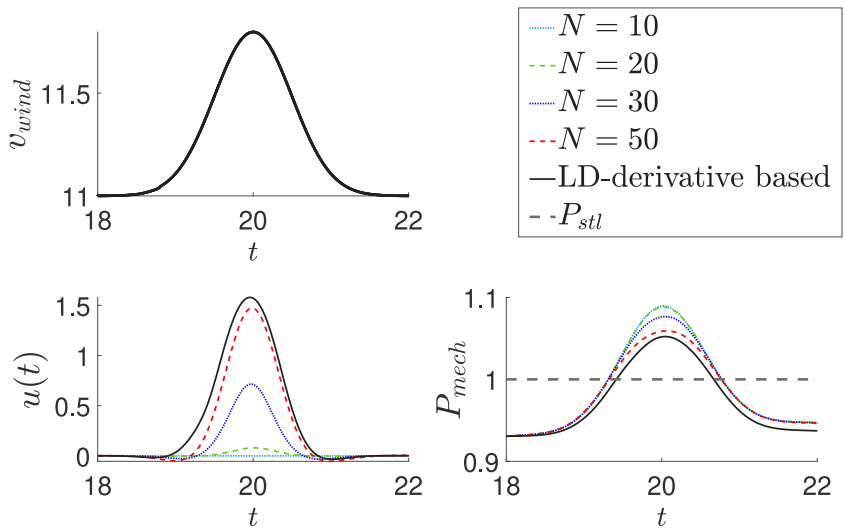

Fig. 8. Smoothing approximation with different values of the smoothing parameters N vs. LD-derivatives based sequential approach with a Gaussian wind profile.

and in general, there are typically no theoretical guarantees that a smoothing approximation approach solution will converge to the true solution as the approximation converges to the nonsmooth function. The LD-derivative based solution, on the other hand, directly treats the nonsmoothness without approximations, and hence only faces other unavoidable numerical errors (e.g., rounding error), which is a clear advantage (see Fig. 9).

A smoothing approach can cause numerical errors that are not traceable. In addition, approximations based on smoothing usually require user-defined parameters that may alter the original problem in an arbitrary manner; in fact, this point was demonstrated in a detailed study by the authors in [35] with relevance to WTPSs.
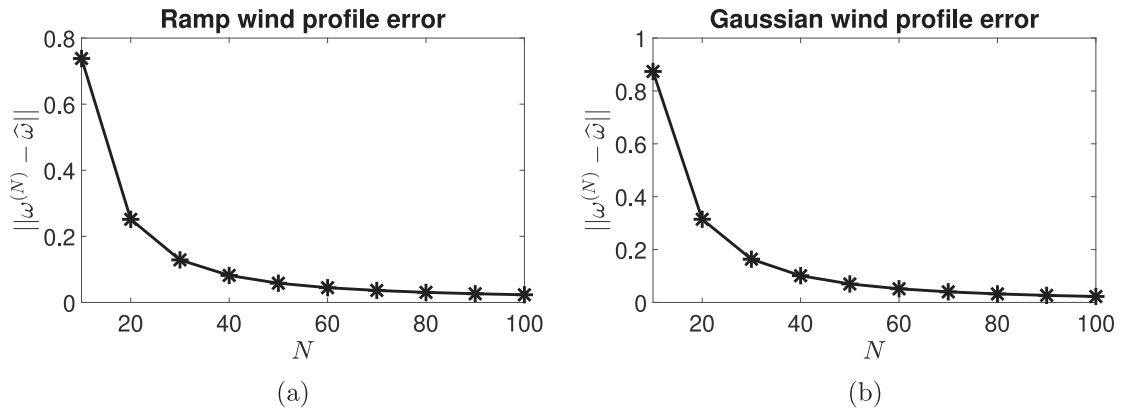
Fig. 9. The 2-norm of error between the nonsmooth function, ω, in (21), and the smoothed function $\widehat{\omega}^{(N)}$ in (30), evaluated at different values of the smoothing parameters N for a ramp wind profile (left panel) and a Gaussian wind profile (right panel).

5.2. Comparison with a naive approach

The second approach is the direct application of an off-the-shelf smooth optimal control solver to the system exhibiting the nonsmoothness without any modifications (i.e., the nonsmoothness is ignored). In particular, we compare our approach vs. solving the nonsmooth WTPS optimal control problem using a popular smooth optimal control solver, GPOPS-II [64], without providing any (generalized) derivative information. That is, we ignore the presence of nonsmoothness completely (hence, "naive approach") and pass the problem to a solver without a smoothing approximation or any other alterations. GPOPS-II [64] is a powerful optimal control solver that is compatible with MATLAB®. In GPOPS-II, hp-adaptive Legendre–Gauss–Radau quadrature orthogonal collocation method is used so that the optimal control problem is transcribed to a large sparse NLP. GPOPS-II uses the simultaneous collocation method, which is another direct optimal control method that also solves a trajectory optimization problem by transforming it into an NLP problem. Both the simultaneous collocation method and the sequential method used in this paper are considered as direct methods as opposed to indirect methods which typically yield an auxiliary set of optimality conditions that can be difficult to solve, especially in a nonsmooth setting. The key difference between the simultaneous collocation method and the sequential method used in this paper is the discretized parameters. In the sequential method, we only discretize the controls, but in the simultaneous collocation method, we discretize both the controls and the states. The reader can refer to [36,52,64,65] for more details. In order to achieve a specified accuracy, GPOPS-II implements an adaptive mesh refinement method for determining the number of mesh intervals and the degree of the approximating polynomial within each mesh interval. Sparse finite-differencing of the optimal control problem functions is used to furnish all derivatives required by the NLP solver. In this subsection, we compare the results of our proposed LD-derivatives based sequential optimal control solver with the results of GPOPS-II. We note that the reader can refer to this GitHub repository [62] for the GPOPS-II implementation used to produce the following results.

Moving forward to the comparison of results, we use the nonsmooth objective functional (23) with a Gaussian wind profile in two different cases: (i) wind speed range that is strictly in Region 2 (medium wind speeds) to avoid any nonsmoothness; and (ii) wind speed range that causes transitions through the nonsmoothness in the objective functional (Regions 2–3). In the first case, when there is no nonsmoothness present, the solutions from our approach and the GPOPS-II are nearly identical; this is expected, but it highlights that our approach does not need a priori knowledge of nonsmoothness and recovers smooth solutions automatically. We show the results of the second case in Fig. 10, where the upper panel shows the Gaussian wind profile with mean $\mu = 20$, standard deviation $\sigma = 0.5$, with different initial wind speeds. The lower left panel shows the optimal control response, $u(t)$, and the lower right panel shows the effect of the optimal control response on P_{mech}. Unlike the first case, we can see from the results of the second case (when nonsmoothness is present) that the solution from GPOPS-II software (which is designed for smooth dynamics/objective functional) gives inaccurate results (breaks) as can be seen in the left panels (with red curves), while the solution from the nonsmooth sequential LD-derivatives based method provides the expected trajectory as seen in the right panels (with black curves). This shows the robustness and generalized nature of the proposed method in all Regions of operation of WTPSs.

The "naive" approach ignores the nonsmoothness completely and hence comes with no guarantees whatsoever of convergence, accuracy, etc. Unsurprisingly then, a method designed for smooth functions, where conventional numerical differentiation tools (such as finite differencing) are used to provide the derivative information required by the solver, may fail. This is what is observed above in Fig. 10, as the solver breaks in some cases. On the other hand, the main advantage of our proposed method is that it directly treats nonsmooth functions in the system's dynamics and/or the objective functional by providing derivative information via LD-derivatives. And hence, while this approach can face other typical numerical issues (e.g., rounding error), we do not face this same potential for failure as the naive approach. Lastly, we note that a performance comparison in terms of computational efficiency between our LD-derivatives based solver and the smooth optimal control solver, GPOPS-II, is not appropriate at this
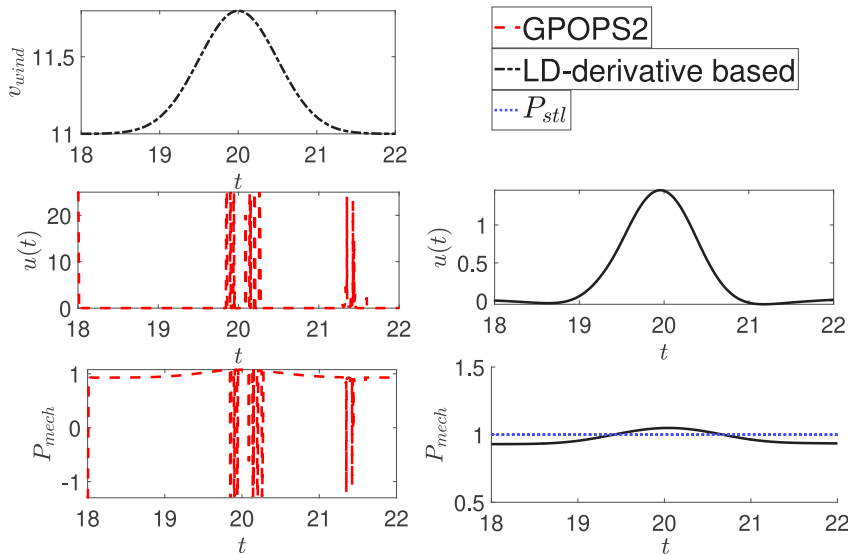
Fig. 10. The collocation optimal control solution using GPOPS-II vs. the LD-derivatives based sequential approach. The upper panel shows a Gaussian wind profile with an initial wind speed $v_0 = 11.0$ m/s, and a standard deviation $\sigma = 0.5$. The left panels (red curves) show the optimal control response $u(t)$ and the effect of the control response on P_{mech} for the smooth solver GPOPS-II. The right panels (in black curves) show the optimal control response and the resulting P_{mech} for the LD-derivatives based sequential nonsmooth optimal control method.

stage; the numerical implementation based on our LD-derivatives based approach is a prototype, and its implementation was not focused on optimizing computational efficiencies, while GPOPS-II is a commercial powerful smooth optimal control solver that uses state-of-the-art techniques for solving a wide range of smooth optimal control problems.

6. Conclusions

In this paper, we provided a framework for a nonsmooth sequential optimal control approach to solve the WTPS control problem. Before summarizing, we note that the MATLAB files used to produce the simulations in Sections 4.1 and 5.2 are uploaded to GitHub as an open-source code [62]. We note that the GPOPS-II software is a commercial software, so, although our implementation is open, the user will require GPOPS-II prior to running the m-file that uses GPOPS-II. The key advantages of the proposed approach are summarized as follows:

- The proposed approach incorporates recent significant developments in WTPS modeling and computationally relevant nonsmooth optimal control theory for DAEs.
- The proposed approach is applicable to WTPSs in all wind speed ranges (Regions 1–3) and achieves two targets simultaneously: (i) maximization of power during lower wind speeds; and (ii) maintaining the power at the rated level during higher wind speeds.
- We demonstrated the effectiveness of the proposed approach via simulations, including real-world data, and a comparison with alternative approaches, such as smoothing (which is lacking) and a naive approach (which fails).
- Having access to the feasible optimal control solution of the WTPS problem extends many possibilities in control developments and, more importantly, provides researchers/industry with a reference solution for comparisons between competing control methods; after all, all control methods should try to realize and be compared with the *optimal control*.
- The current work can inspire similar approaches in other power systems in the form of (1b)–(1c), especially renewable energy power systems which can experience nonsmooth conditions and need substantial control developments.

CRediT authorship contribution statement

Hesham Abdelfattah: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Sameh A. Eisa:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Conceptualization. **Peter Stechlinski:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

Denotations of state and parameters in the WTPS model (14):

ρ, A_r, v_{wind} - air density, rotor area (m^2), wind speed (m/s).

C_p, w_{ref}: aerodynamic power coefficient, reference speed.

$\alpha_{i,j}$, P_{mech}: empirical constants, mechanical power.

w_0, w_{base}: initialized speed constant, base angular frequency.

P_{elec}, V: electrical (active) power, terminal voltage.

R, X, E: infinite bus resistance, reactance, voltage.

Q_{gen}, I_{plv}: reactive power, active current.

λ, θ: tip ratio, pitch angle in degrees.

w_t, w_g: dynamic turbine and generator speeds.

$\Delta\theta_m$: integral of difference between w_t and w_g.

f_1: integral of differences of speeds.

P_{stl}: rated power.

P_{inp}, P_{1elec}: power order, filtered electrical power.

T_{pc}, K_{ptrq}: time constant, torque control proportional.

V_{ref}, E_{qcmd}: reference voltage, reactive voltage command.

H, H_g: turbine inertia constant, generator inertia constant.

D_{tg}, K_{tg} - shaft damping constant, shaft stiffness constant.

K_{itrq}: torque control constant.

K_{Qi}: reference voltage gain.

T_{pwr}, K_{vi}, E_q: filtered electric power time constant, reactive voltage command time constant, generator reactive variable.

The values of the parameters involved, and the C_p coefficients are given in Tables A.1, and A.2, respectively [5,6,51].

Table A.1
Representative values for parameters in the model.

Parameter	Value
w_0	1 (any choice bigger than 0)
X_{eq}, D_{tg}	0.8, 1.5
K_{tg}, w_{base}	1.11, 125.66 respectively
$\frac{1}{2}\rho A_r, K_b$	0.00159 and 56.6 respectively
H	4.94
H_g, K_{itrq}	0.62, 0.6 respectively
T_{pc}, K_{ptrq}	0.05, 3 respectively
T_{pwr}, K_{Qi}	0.05, 0.1 respectively
P_{stl}, w_{ref}	1, 1.2 respectively
K_{vi}, R, E	40, 0.02, 1.0164 respectively
$X = X_l + X_{tr}$	$X_l = 0.0243, X_{tr} = 0.00557$

Table A.2
c_p coefficients $\alpha_{i,j}$.

i	j	$\alpha_{i,j}$	i	j	$\alpha_{i,j}$
4	4	4.9686e−10	4	3	−7.1535e−8
4	2	1.6167e−6	4	1	−9.4839e−6
4	0	1.4787e−5	3	4	−8.9194e−8
3	3	5.9924e−6	3	2	−1.0479e−4
3	1	5.7051e−4	3	0	−8.6018e−4
2	4	2.7937e−6	2	3	−1.4855e−4
2	2	2.1495e−3	2	1	−1.0996e−2
2	0	1.5727e−2	–	–	–
1	4	−2.3895e−5	1	3	1.0683e−3
1	2	−1.3934e−2	1	1	6.0405e−2
1	0	−6.7606e−2	0	4	1.1524e−5
0	3	−1.3365e−4	0	2	−1.2406e−2
0	1	2.1808e−1	0	0	−4.1909e−1

References

[1] R. Wiser, M. Bolinger, B. Hoen, D. Millstein, J. Rand, G. Barbose, N. Darghouth, W. Gorman, S. Jeong, B. Paulos, Land-Based Wind Market Report: 2022 Edition, Tech. Rep., Lawrence Berkeley National Lab (LBNL), 2022.

[2] X. Liang, Emerging power quality challenges due to integration of renewable energy sources, IEEE Trans. Ind. Appl. 53 (2) (2016) 855–866.

[3] P. Pourbeik, Specification of the Second Generation Generic Models for Wind Turbine Generators, Report, Electric Power Research Institute, 2014.

[4] K. Clark, N.W. Miller, J.J. Sanchez-Gasca, Modeling of GE Wind Turbine-Generators for Grid Studies, Report, General Electric International, Inc., 2010.

[5] S.A. Eisa, Modeling dynamics and control of type-3 DFIG wind turbines: Stability, Q Droop function, control limits and extreme scenarios simulation, Electr. Power Syst. Res. 166 (2019) 29–42.

[6] S.A. Eisa, Nonlinear modeling, analysis and simulation of wind turbine control system with and without pitch control as in industry, Adv. Control Optim. Paradigms Wind Energy Syst. (2019) 1–40.

[7] A. Kusiak, H. Zheng, Z. Song, Power optimization of wind turbines with data mining and evolutionary computation, Renew. Energy 35 (3) (2010) 695–702.

[8] M.-S. Lu, C.-L. Chang, W.-J. Lee, L. Wang, Combining the wind power generation system with energy storage equipment, IEEE Trans. Ind. Appl. 45 (6) (2009) 2109–2115.

[9] S.A. Eisa, Investigating the problem of oscillatory orbits and attractors in wind turbines system under control limits imposed by industry, Electr. Power Syst. Res. 180 (2020) 106098.

[10] E.J.N. Menezes, A.M. Araújo, N.S.B. Da Silva, A review on wind turbine control and its associated methods, J. Clean. Prod. 174 (2018) 945–953.

[11] G. Tsourakis, B. Nomikos, C. Vournas, Effect of wind parks with doubly fed asynchronous generators on small-signal stability, Electr. Power Syst. Res. 79 (1) (2009) 190–200.

[12] R. Gao, Z. Gao, Pitch control for wind turbine systems using optimization, estimation and compensation, Renew. Energy 91 (2016) 501–515.

[13] J. Lan, R.J. Patton, X. Zhu, Fault-tolerant wind turbine pitch control using adaptive sliding mode estimation, Renew. Energy 116 (2018) 219–231.

[14] H. Zhao, Q. Wu, J. Wang, Z. Liu, M. Shahidehpour, Y. Xue, Combined active and reactive power control of wind farms based on model predictive control, IEEE Trans. Energy Convers. 32 (3) (2017) 1177–1187.

[15] D. Kumar, K. Chatterjee, A review of conventional and advanced MPPT algorithms for wind energy systems, Renew. Sustain. Energy Rev. 55 (2016) 957–970.

[16] H. Abdelfattah, S. Eisa, Novel nonlinear time domain optimal control formulation of wind turbine power system with GPOPS2 software simulations, in: 2023 IEEE Green Technologies Conference (GreenTech), 2023, pp. 1–5.

[17] H.H. Hansen, N. MacDougall, C.D. Jensen, M. Kulahci, B.F. Nielsen, Condition monitoring of wind turbine faults: Modelling and savings, Appl. Math. Model. (2024) 160–174.

[18] F. Golnary, H. Moradi, Dynamic modelling and design of various robust sliding mode controls for the wind turbine with estimation of wind speed, Appl. Math. Model. 65 (2019) 566–585.

[19] M. Kear, B. Evans, R. Ellis, S. Rolland, Computational aerodynamic optimisation of vertical axis wind turbine blades, Appl. Math. Model. 40 (2) (2016) 1038–1051.

[20] M. Rahimi, Dynamic performance assessment of DFIG-based wind turbines: A review, Renew. Sustain. Energy Rev. 37 (2014) 852–866.

[21] S.A. Eisa, K. Wedeward, W. Stone, Sensitivity analysis of a type-3 DFAG wind turbine's dynamics with pitch control, in: 2016 IEEE Green Energy and Systems Conference, 2016, pp. 1–6.

[22] S. Ghosh, S. Kamalasadan, N. Senroy, J. Enslin, Doubly fed induction generator (DFIG)-based wind farm control framework for primary frequency and inertial response application, IEEE Trans. Power Syst. 31 (3) (2015) 1861–1871.

[23] Y. Zou, M. Elbuluk, Y. Sozer, Stability analysis of maximum power point tracking (MPPT) method in wind power systems, in: 2011 IEEE Industry Applications Society Annual Meeting, 2011, pp. 1–8.

[24] M.A. Abdullah, A. Yatim, C.W. Tan, R. Saidur, A review of maximum power point tracking algorithms for wind energy systems, Renew. Sustain. Energy Rev. 16 (5) (2012) 3220–3227.

[25] Z. Civelek, M. Lüy, E. Çam, N. Barışçı, Control of pitch angle of wind turbine by fuzzy PID controller, Intell. Autom. Soft Comput. 22 (3) (2016) 463–471.

[26] A. Koerber, R. King, Combined feedback–feedforward control of wind turbines using state-constrained model predictive control, IEEE Trans. Control Syst. Technol. 21 (4) (2013) 1117–1128.

[27] C.-T. Chang, Fuzzy score technique for the optimal location of wind turbines installations, Appl. Math. Model. 44 (2017) 576–587.

[28] M. Soliman, O. Malik, D. Westwick, Multiple model multiple-input multiple-output predictive control for variable speed variable pitch wind energy conversion systems, IET Renew. Power Gener. 5 (2) (2011) 124–136.

[29] H. Jafarnejadsani, J. Pieper, Gain-scheduled ℓ_1-optimal control of variable-speed-variable-pitch wind turbines, IEEE Trans. Control Syst. Technol. 23 (2015) 372–379.

[30] J. Slootweg, S. De Haan, H. Polinder, W. Kling, General model for representing variable speed wind turbines in power system dynamics simulations, IEEE Trans. Power Syst. 18 (1) (2003) 144–151.

[31] I.A. Hiskens, Dynamics of type-3 wind turbine generator models, IEEE Trans. Power Syst. 27 (1) (2011) 465–474.

[32] A. Lasheen, A.L. Elshafei, Wind-turbine collective-pitch control via a fuzzy predictive algorithm, Renew. Energy 87 (2016) 298–306.

[33] Y. Hu, T.K. Chau, X. Zhang, H.H.-C. Iu, T. Fernando, D. Fan, A novel adaptive model predictive control strategy for DFIG wind turbine with parameter variations in complex power systems, IEEE Trans. Power Syst. (2022) 4582–4592.

[34] Z. Li, B. Wen, Z. Peng, X. Dong, Y. Qu, Dynamic modeling and analysis of wind turbine drivetrain considering the effects of non-torque loads, Appl. Math. Model. 83 (2020) 146–168.

[35] S.A. Eisa, P. Stechlinski, Sensitivity analysis of nonsmooth power control systems with an example of wind turbines, Commun. Nonlinear Sci. Numer. Simul. 95 (2021) 105633.

[36] M. Kelly, An introduction to trajectory optimization: How to do your own direct collocation, SIAM Rev. 59 (2017) 849–904.

[37] M.N. Skandari, A. Kamyad, S. Effati, Smoothing approach for a class of nonsmooth optimal control problems, Appl. Math. Model. 40 (2) (2016) 886–903.

[38] A. Caspari, L. Lüken, P. Schäfer, Y. Vaupel, A. Mhamdi, L.T. Biegler, A. Mitsos, Dynamic optimization with complementarity constraints: Smoothing for direct shooting, Comput. Chem. Eng. 139 (2020) 106891.

[39] A.M. Bagirov, A. Al Nuaimat, N. Sultanova, Hyperbolic smoothing function method for minimax problems, Optimization 62 (6) (2013) 759–782.

[40] E. Amhraoui, T. Masrour, Smoothing approximations for piecewise smooth functions: A probabilistic approach, Numer. Algebra Control Optim. 12 (4) (2022) 745–762.

[41] L.S. Pontryagin, Mathematical Theory of Optimal Processes, CRC Press, 1987.

[42] F.H. Clarke, Optimization and Nonsmooth Analysis, SIAM, 1990.

[43] F. Clarke, M.d.R. De Pinho, Optimal control problems with mixed constraints, SIAM J. Control Optim. 48 (7) (2010) 4500–4524.

[44] L. Guo, J.J. Ye, Necessary optimality conditions for optimal control problems with equilibrium constraints, SIAM J. Control Optim. 54 (5) (2016) 2710–2733.

[45] K.A. Khan, P.I. Barton, A vector forward mode of automatic differentiation for generalized derivative evaluation, Optim. Methods Softw. 30 (6) (2015) 1185–1212.

[46] K.A. Khan, P.I. Barton, Generalized gradient elements for nonsmooth optimal control problems, in: 53rd IEEE Conference on Decision and Control, 2014, pp. 1887–1892.

[47] P. Stechlinski, P.I. Barton, Generalized derivatives of optimal control problems with nonsmooth differential-algebraic equations embedded, in: 55th IEEE Conference on Decision and Control, 2016, pp. 592–597.

[48] A.M. Bagirov, M. Gaudioso, N. Karmitsa, M.M. Mäkelä, S. Taheri, Numerical Nonsmooth Optimization, Springer, 2020.

[49] P. Stechlinski, Dynamic optimization of complementarity systems, IEEE Trans. Autom. Control 68 (2) (2022) 1122–1129.

[50] P. Stechlinski, Optimization-constrained differential equations with active set changes, J. Optim. Theory Appl. 187 (1) (2020) 266–293.

[51] S.A. Eisa, W. Stone, K. Wedeward, Mathematical analysis of wind turbines dynamics under control limits: boundedness, existence, uniqueness, and multi time scale simulations, Int. J. Dyn. Control 6 (3) (2018) 929–949.

[52] A. Rao, A survey of numerical methods for optimal control, Adv. Astronaut. Sci. 135 (01) (2009) 497–528.

[53] S.A. Eisa, K. Wedeward, W. Stone, Wind turbines control system: nonlinear modeling, simulation, two and three time scale approximations, and data validation, Int. J. Dyn. Control (2018) 1–23.

[54] M.A. Patterson, A.V. Rao, GPOPS-II: A MATLAB software for solving multiple-phase optimal control problems using hp-adaptive Gaussian quadrature collocation methods and sparse nonlinear programming, 2014, 41, (1).

[55] P.I. Barton, K.A. Khan, P. Stechlinski, H.A.J. Watson, Computationally relevant generalized derivatives: theory, evaluation and applications, Optim. Methods Softw. 33 (2018) 1030–1072.

[56] Y. Nesterov, Lexicographic differentiation of nonsmooth functions, Math. Program. 104 (2005) 669–700.

[57] K.A. Khan, P.I. Barton, Generalized derivatives for solutions of parametric ordinary differential equations with non-differentiable right-hand sides, J. Optim. Theory Appl. 163 (2) (2014) 355–386.

[58] P. Stechlinski, P.I. Barton, Generalized derivatives of differential–algebraic equations, J. Optim. Theory Appl. 171 (1) (2016) 1–26.

[59] D.E. Kirk, Optimal Control Theory: an Introduction, Courier Corporation, 2004.

[60] J.G. Njiri, D. Söffker, State-of-the-art in wind turbine control: Trends and challenges, Renew. Sustain. Energy Rev. 60 (2016) 377–393.

[61] K.E. Johnson, G. Fritsch, Assessment of extremum seeking control for wind farm energy production, Wind Eng. 36 (6) (2012) 701–715.

[62] H. Abdelfattah, S. Eisa, P. Stechlinski, LD-derivatives based optimal control WTPS, 2024, GitHub,.

[63] J.D. Cook, Basic properties of the soft maximum, 2011, https://www.johndcook.com/soft_maximum.pdf.

[64] M.A. Patterson, A.V. Rao, GPOPS-II: A MATLAB software for solving multiple-phase optimal control problems using hp-adaptive Gaussian quadrature collocation methods and sparse nonlinear programming, ACM Trans. Math. Softw. 41 (1) (2014) 1–37.

[65] L.T. Biegler, An overview of simultaneous strategies for dynamic optimization, Chem. Eng. Process.: Process Intensif. 46 (11) (2007) 1043–1053.