2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Input-to-State Stability of Newton Methods for Generalized Equations in Nonlinear Optimization*

Torbjørn Cunis[1,2] and Ilya Kolmanovsky[2]

***Abstract*— We show that Newton methods for generalized equations are input-to-state stable with respect to perturbations such as due to inexact computations. We then use this result to obtain convergence and robustness of a multistep Newton-type method for multivariate generalized equations. We demonstrate the usefulness of the results with other applications to nonlinear optimization. In particular, we provide a new proof for (robust) local convergence of the augmented Lagrangian method.**

I. Introduction

Generalized equations are set-valued inequalities

$$f(z, v) + F(z) \ni 0 \tag{1}$$

where f is a function, F is a set-valued mapping, and v is a perturbation term representing, e.g., inexactness of the solution or incomplete information. In nonlinear optimization, generalized equations appear frequently as the Karush–Kuhn–Tucker (KKT) system of necessary conditions [1] with f being the gradient of the Lagrangian, F the normal cone to the constraint set, and z aggregating primal and dual decision variables. Optimization algorithms can often be interpreted as solving (1) for its root $\bar{z}$, the optimal solution. A common technique is Newton's method for generalized equations, which yields the iteration

$$f(z_k, v) + \nabla f(z_k, v)(z_{k+1} - z_k) + F(z_{k+1}) \ni 0 \tag{2}$$

and which, when applied to the KKT system, is better known as sequential quadratic programming [2]. If the gradient of f in (2) does not exist or is unknown, Newton's method can be extended to the broader class of quasi-Newton and Josephy–Newton methods which include projected gradient descent and sequential convex (linear) programming. Robust local convergence properties of Newton methods to constant perturbations have been studied under metric regularity assumptions [3–6].

In recent years, properties of optimization algorithms have been studied when interconnected with dynamic systems and used to generate control actions [7–11]. A common framework here is the one of input-to-state stability (ISS) which combines concepts of robust stability and asymptotic gains with dissipation theory [12]. Previously, ISS was proven for classical iterative methods for linear equations [13, 14], gradient descent [15], and proximal gradient descent [16]. In addition, the result in [5] on stability of the sequence generated by (2) can be considered as an ISS-like result. On the other hand, previous works on perturbed Newton methods for generalized equations such as [4, 5] treated the input as a static deviation of the limit point. Establishing ISS of Newton methods for generalized equations enables the treatment of dynamic and time-varying perturbations, which are common in, e.g., the analysis of interconnected optimization algorithms and optimization-based feedback.

In this paper, we investigate input-to-state stability of the classical Josephy–Newton method and a new multistep Newton-type method subject to nonconstant perturbations. To that extent, we build upon inverse function theorems for set-valued mappings from variational analysis [17] under strong regularity and strong subregularity. Since our results depend on properties of the underlying generalized equations, in particular, the KKT conditions or approximations thereof, they are applicable to a large class of iterative optimization algorithms including sequential programming and augmented Lagrangian methods.

The contributions of our paper are threefold: Firstly, we formally prove local ISS of Newton methods for generalized equations in the presence of generic perturbations including due to inexact computations or erroneous gradients. Secondly, we propose a multistep Newton-type method for multivariate generalized equations, which allows for lower-dimensional partial updates, and prove its robust local convergence using the ISS property. We then demonstrate that the result of [5] follows immediately from ISS. Thirdly, we apply our framework and ISS results to approximate sequential programming and the augmented Lagrangian method.

II. Preliminaries

If not noted otherwise, all spaces are considered either finite-dimensional or complete (Banach) vector spaces with norm $\|\cdot\|$. A set-valued mapping F between vector spaces X and Y, denoted $F : X \rightrightarrows Y$, takes values $F(x) \subset Y$ for any $x \in X$. We define the domain and graph of F as $\operatorname{dom} F = \{x \in X \,|\, F(x) \neq \varnothing\}$ and $\operatorname{graph} F = \{(x, y) \,|\, y \in F(x)\}$, respectively. For a closed convex set $C \subset X$, the normal cone mapping is $N_C(\bar{x}) = \{y \in X^* \,|\, \forall x \in C, \langle y, x - \bar{x}\rangle \leq 0\}$ if $\bar{x} \in C$ and $N_C(\bar{x}) = \varnothing$ else, where X^* is the dual space of X. The gradient of a function $f : X \to Y$ at $\bar{x} \in X$, if existing, is a linear operator $\nabla f(\bar{x}) : X \to Y$; and we will assume that any gradient, if existing, is Lipschitz continuous around $\bar{x}$.

*This research is supported in part by AFOSR grant number FA9550-20-1-0385.

[1]Institute of Flight Mechanics and Control, University of Stuttgart, 70569 Stuttgart, Germany. tcunis@ifr.uni-stuttgart.de

[2]Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48104, USA. tcunis@umich.edu, ilya@umich.edu

TABLE I

JOSEPHY–NEWTON METHODS TO SOLVE (5).

Algorithm	Choice of $H(z_k)$	Interpretation
Sequential Quadratic Programming	$\begin{pmatrix} \nabla^2(h(x_k) + \langle g(x_k), y_k\rangle) & \nabla g(x_k)^* \\ \nabla g(x_k) & 0 \end{pmatrix}$	$(x, y)_{k+1}$ is the primal-dual solution to a quadratic program with linear constraints.
Sequential Linear Programming	$\begin{pmatrix} 0 & \nabla g(x_k)^* \\ \nabla g(x_k) & 0 \end{pmatrix}$	$(x, y)_{k+1}$ is the primal-dual solution to a linear program.
Projected Gradient Descent	$\alpha^{-1}\mathbb{I}$ with $\alpha > 0$	x_{k+1} is the projection of $\big(x_k - \alpha\nabla h(x_k) - \alpha g(x_k)^* y_k\big)$ onto C; and $y_{k+1} = y_k - \alpha g(x_k)$.

A. Continuity & Regularity

A set-valued mapping $F : X \rightrightarrows Y$ is said to be *Lipschitz continuous* on $D \subset X$ with constant $\kappa \geq 0$ if D is nonempty, $F(x)$ is closed, and

$$F(x') \subset \{y' \in Y \mid \exists y \in F(x), \|y - y'\| \leq \kappa \|x - x'\|\} \quad (3)$$

for all $x, x' \in D$. The condition (3) reduces to the classical Lipschitz continuity of functions if F is single-valued on D. Let $U \subset X$ and $V \subset Y$ be neighbourhoods of $(\bar{x}, \bar{y}) \in \operatorname{graph} F$; the mapping F has the *isolated calmness* property at $\bar{x}$ for $\bar{y}$ with constant κ if $F(\bar{x}) \cap V = \{\bar{y}\}$ and $x \mapsto F(x) \cap V$ satisfies (3) for $x = \bar{x}$ and all $x' \in U$. Moreover, a function $f : X \to Y$ which is Lipschitz continuous in a neighbourhood of $\bar{x}$ with constant κ is a *(Lipschitz) continuous single-valued localization* of F at $\bar{x}$ with constant κ for $\bar{y}$ if $F(x) \cap V = \{f(x)\}$ for all $x \in U$.

We now define notions of regularity by continuity of the inverse $F^{-1} : y \mapsto \{x \in X \mid y \in F(x)\}$.

Definition 1: Take $(\bar{x}, \bar{y}) \in \operatorname{graph} F$; the mapping F is *strongly regular* at $\bar{x}$ for $\bar{y}$ with constant κ if and only if F^{-1} has a Lipschitz continuous single-valued localization at $\bar{y}$ for $\bar{x}$ with constant κ.

Definition 2: Take $(\bar{x}, \bar{y}) \in \operatorname{graph} F$; the mapping F is *strongly subregular* at $\bar{x}$ for $\bar{y}$ with constant κ if and only if F^{-1} has the isolated calmness property at $\bar{y}$ for $\bar{x}$ with constant κ.

We say that F is strongly regular (subregular) with constant $\kappa \geq 0$ if κ is the constant of the Lipschitz continuous localization (isolated calmness property) of F^{-1}.

Proposition 1: Let $F : X \rightrightarrows Y$ be strongly regular (subregular) at $\bar{x}$ for $\bar{y}$ with constant $\kappa \geq 0$ and $(\bar{x}, \bar{y}) \in \operatorname{graph} F$; if $g : X \to Y$ is Lipschitz continuous with constant $\mu \in [0, \kappa^{-1})$, then $(g + F)$ is strongly regular (subregular) at $\bar{x}$ for $g(\bar{x}) + \bar{y}$ with constant $\kappa/(1 - \kappa\mu)$.

Proof: See [17, Theorems 8.6 and 12.2]. ■

We next give an interpretation of strong regularity in the context of nonlinear optimization.

B. Nonlinear Optimization

Consider a nonlinear program,

$$\min_{x \in C} h(x) \quad \text{subject to } g(x) = 0 \quad (4)$$

with primal variable $x \in X$, cost function $h : X \to \mathbb{R}$, constraint $g : X \to Y$, and closed convex set $C \subset X$. If $\bar{x}$ is an optimal solution of (4), $\nabla h(\bar{x})$ and $\nabla g(\bar{x})$ exist, and a suitable constraint qualification such as the linear independence constraint qualification (LICQ) holds, then the KKT system

$$\begin{pmatrix} \nabla h(\bar{x}) + \nabla g(\bar{x})^* \bar{y} \\ g(\bar{x}) \end{pmatrix} + N_{C \times X^*}\big((\bar{x}, \bar{y})\big) \ni 0 \quad (5)$$

is satisfied for some dual variable $\bar{y} \in X^*$ [see, e.g., 18, Theorem 5.7]. Eq. (5) is a generalized equation in the form of (1) with $z = (x, y) \in Z = X \times Y$, where the left-hand side is a set-valued maping due to the normal cone mapping. In finite dimensions, this mapping is strongly regular at $(\bar{x}, \bar{y})$ for 0 if and only if LICQ holds (a fortiori, $\bar{y}$ is unique) and $\bar{x}$ is a strongly-stable stationary solution[1] of (4). Optimization algorithms which rely on solving (5), such as Newton-type methods, typical require strong regularity to guarantee that the result is in fact a locally optimal solution of (4).

In Section IV, we will consider a perturbed version of (4) where $h(\cdot, v)$ and $g(\cdot, v)$ are Lipschitz continuous functions of $v \in V$. If the left-hand side of the *parametrized* KKT system (5) is strongly regular at $\bar{v} \in V$, then its solution mapping $S : V \rightrightarrows X \times Y$ has a Lipschitz continuous single-valued localization; this is the result of Robinson's implicit function theorem [17, Theorem 8.5], which we extend to the case of multivariate mappings in the appendix.

C. Newton Methods

To solve the generalized equation (1) with $f : Z \to Z'$ and $F : Z \rightrightarrows Z'$, we define the iteration

$$z_{k+1} \in z_k - \big(H(z_k) + F\big)^{-1} f(z_k) \quad (6)$$

where, broadly speaking, $H(z_k) : Z \to Z'$ helps to approximate f around z_k. Table I reviews some common choices of $H(z_k)$ for the KKT system (5) with $z = (x, y)$ and the resulting optimization algorithms. Eq. (6) can be interpreted as a generalized equation parametrized in the previous solution z_k. In general, the sequences generated by (2) or (6) are not unique, nor is a solution guaranteed to exist. Under regularity assumptions, however, a sequence exists and converges to a root of (1).

Theorem 1 ([17, Theorem 15.2]): Let $\bar{z}$ be a solution to (1) such that $\nabla f(\bar{z})$ exists. If $f + F$ is strongly subregular at $\bar{z}$ for 0, then for any z_0 sufficiently close to $\bar{z}$ there exists a sequence $\{z_k\}_{k=0}^{\infty}$ generated by (2) which converges

[1] See, e.g., [19, Definition 2.7] for a definition of a strongly-stable stationary solution.

quadratically to $\bar{z}$. Moreover, if $f + F$ is strongly regular, then this sequence is unique. $\lhd$

In fact, any sequence that stays sufficiently close to $\bar{z}$ converges quadratically. The following result provides sufficient conditions for convergence if $H(z_k)$ is not a derivative of f; we define $f_H(z, \zeta) = f(\zeta) + H(\zeta)(z - \zeta)$ and assume that $f_H(\cdot, \zeta)$ is Lipschitz continuous uniformly[2] in ζ at $(\bar{z}, \bar{z})$.

Proposition 2: Let $\bar{z}$ be a solution to (1) such that $f_H(x, \cdot)$ is Lipschitz continuous with constant γ uniformly in x at $(\bar{z}, \bar{z})$; if $f_H(\cdot, \bar{z}) + F$ is strongly subregular at $\bar{z}$ for 0 with constant κ and $\kappa\gamma < 1$, then for any z_0 sufficiently close to $\bar{z}$ there exists a sequence generated by (6) which converges linearly to $\bar{z}$. Moreover, if $f_H(\cdot, \bar{z}) + F$ is strongly regular, the sequence is unique.

Proof: This is a consequence of [17, Theorems 12.4 and 8.5] with $h = f_H(\cdot, \bar{z})$. ■

If f is continuously differentiable at $\bar{z}$ and $H(z) = \nabla f(z)$, then $f_H(\cdot, \bar{z})$ is the linearization of f around $\bar{z}$ and $f_H(x, \cdot)$ is Lipschitz continuous uniformly in x with constant 0.

D. Input-to-state Stability

We now consider the robustness of the sequences generated by (2) or (6) under perturbations. To that extent, we consider a perturbed dynamic system

$$z_{k+1} = \Phi(z_k, v_k) \tag{7}$$

for all $k \in \mathbb{N}$, where $\mathbf{v} = (v_0, v_1, \dots) \subset V$ is a sequence of perturbations with $\|\mathbf{v}\|_\infty := \sup_{k \in \mathbb{N}} \|v_k\| < \infty$. The following definition makes use of the comparison function classes $\mathcal{KL}$ and $\mathcal{K}_\infty$ of monotonic functions; see [20] for details.

Definition 3: The system (7) is *locally input-to-state stable* around $\bar{z}$ if and only if there exist $\epsilon, \delta > 0$ and functions $\beta \in \mathcal{KL}$ and $\gamma \in \mathcal{K}_\infty$ such that any sequence $\{z_k\}_{k=0}^\infty$ generated under perturbation $\|\mathbf{v}\|_\infty < \delta$ satisfies

$$\|z_k - \bar{z}\| \le \beta(\|z_0 - \bar{z}\|, k) + \gamma(\|\mathbf{v}\|_\infty)$$

for all $k \in \mathbb{N}$, provided that $\|z_0 - \bar{z}\| < \epsilon$.

The definition implies that the solution of (7) converges to a ball $\mathcal{B}_{\gamma,\mathbf{v}}(\bar{z})$ around $\bar{z}$ with radius given by the gain $\gamma(\|\mathbf{v}\|_\infty)$. The system (7) is locally input-to-state stable around $\bar{z}$ if (and only if) there exists a continuous, positive definite function $V : Z \to \mathbb{R}_{\ge 0}$, constants $\epsilon > 0$ and $\delta > 0$, and functions $\alpha, \gamma \in \mathcal{K}_\infty$ such that $\alpha < \mathrm{id}$ and [21]

$$V(\Phi(z, v)) \le \alpha V(z) + \gamma \|v\| \tag{8}$$

for all $(z, v) \in Z \times V$ with $\|z - \bar{z}\| < \epsilon$ and $\|v\| < \delta$.

III. Main Results

We prove local input-to-state stability of perturbed Newton methods in the form of (6), also referred to as Josephy–Newton method, and propose a new multistep Newton-type method for multivariate generalized equations. A discussion of related results concludes this section.

[2]We say that a function $f(x, y)$ is Lipschitz continuous with respect to x *uniformly* in y at $(\bar{x}, \bar{y})$ if $f(\cdot, y)$ is Lipschitz continuous at $\bar{x}$ with some constant $\kappa < \infty$ for all y in a neighbourhood of $\bar{y}$.

A. Josephy–Newton Method

Our first result concerns the Josephy–Newton method under time-varying perturbation,

$$f(z_k, v_k) + H(z_k, v_k)(z_{k+1} - z_k) + F(z_{k+1}) \ni 0 \tag{9}$$

with $f : Z \times V \to Z'$ and $H(z, v) : Z \to Z'$. As before, define $f_H(z, \zeta, v) = f(\zeta, v) + H(\zeta, v)(z - \zeta)$. Here, the perturbation $v_k \in V$ might model, e.g., the inexact evaluation of the gradient $\nabla f(z)$ or a nonzero remainder in solving (1). We make the following assumptions.

Assumption 1: Let $\bar{z} \in Z$ and $\kappa, \gamma_z, \gamma_v > 0$ satisfy:

(a) $\bar{z}$ is a solution of $f(\cdot, 0) + F \ni 0$;
(b) $f_H(\cdot, \zeta, v)$ is Lipschitz continuous uniformly in (ζ, v) at $(\bar{z}, \bar{z}, 0)$;
(c) $f_H(z, \cdot, v)$ is Lipschitz continuous with constant γ_z uniformly in (z, v) at $(\bar{z}, \bar{z}, 0)$;
(d) $f_H(z, \zeta, \cdot)$ is Lipschitz continuous with constant γ_v uniformly in (z, ζ) at $(\bar{z}, \bar{z}, 0)$;
(e) $f_H(\cdot, \bar{z}, 0) + F$ is strongly regular with constant κ at $\bar{z}$ for 0;

and $\kappa\gamma_z < 1$.

To satisfy these assumptions, H must be chosen such that f_H provides a 'sufficiently good' approximation of f around $\bar{z}$, measured by Lipschitz continuity, strong regularity, and the condition that $\kappa\gamma_z < 1$. Our result is based on an extension of Robinson's implicit function theorem for generalized equations with multiple parameters, given in the appendix.

Theorem 2: Under Assumption 1, the iteration in (9) is unique for (z_k, v_k) sufficiently close to $(\bar{z}, 0)$ and satisfies

$$\|z_{k+1} - \bar{z}\| \le \kappa\gamma_z \|z_k - \bar{z}\| + \gamma \|v_k\|$$

that is, (9) is locally input-to-state stable around $\bar{z}$.

Proof: By virtue of Corollary 2 (in the Appendix) with $h = f_H(\cdot, \bar{z}, 0)$ and $\mu = 0$, the Josephy–Newton step (9) has a locally unique solution $s : Z \times V \to Z$ for (z_k, v_k) in the neighbourhood of $(\bar{z}, 0)$ satisfying

$$\|s(z, v) - s(z', v')\| \le \kappa\gamma_z \|z - z'\| + \kappa\gamma_v \|v - v'\|$$

for $(z, v), (z', v')$ around $(\bar{z}, 0)$. Taking $\alpha = \kappa\gamma_z < 1$ and $\gamma = \kappa\gamma_v$ as well as noting that $z_{k+1} = s(z_k, v_k)$ and $\bar{z} = s(\bar{z}, 0)$ leads to the desired result. ■

We can further strengthen this result if the perturbation affects the gradient of f, that is, $f(\cdot, v) = f$ and $H(\zeta, v) = \nabla f(\zeta) + v$, assuming continuous differentiability of f. Note that this implies Lipschitz continuity of f_H with arbitrarily small constants γ_z and γ_v; moreover, $f_H(\cdot, \bar{z}, 0) + F$ is strongly regular if and only if $f + F$ is. In this case, (9) resembles a *quasi-Newton* method. Specialising Theorem 2 to the quasi-Newton method, we obtain quadratic convergence to $\mathcal{B}_{\gamma,\mathbf{v}}(\bar{z})$ where the gain γ vanishes close to $\bar{z}$.

Corollary 1: Under Assumption 1, not only is

$$f(z_k) + (\nabla f(z_k) + v_k)(z_{k+1} - z_k) + F(z_{k+1}) \ni 0 \tag{10}$$

locally input-to-state stable around $\bar{z}$ but the generated sequence $\{z_k\}_{k=0}^{\infty}$ satisfies

$$\|z_{k+1} - \bar{z}\| \le \kappa L \|z_k - \bar{z}\|^2 + \gamma_k \|v_k\| \tag{11}$$

for all $k \in \mathbb{N}$ with $\gamma_k = \kappa \|z_k - \bar{z}\|$.

Proof: Let $\{z_k\}_{k=0}^{\infty}$ be the sequence generated by (10) which, by virtue of Theorem 2, exists, is unique, and remains in the neighbourhood of $\bar{z}$ for z_0 and $\mathbf{v}$ close to $(\bar{z}, 0)$. We now argue with [17, Proof of Theorem 15.2] that

$$\begin{aligned}\|f(z_k) - f(\bar{z}) - (\nabla f(z_k) + v_k)(z_k - \bar{z})\| \\ \le L\|z_k - \bar{z}\|^2 + |\langle v_k, z_k - \bar{z}\rangle|\end{aligned}$$

where $2L$ is the Lipschitz constant of $\nabla f(z)$ around $\bar{z}$, guaranteed to exist by Assumption 1 and $H(z, v) = \nabla f(z) + v$, and hence (11) holds by strong regularity of f_H. ■

B. Multistep Newton-type Method

For our second result, we consider the perturbed multivariate generalized equation

$$f(x, y, v) + F(x, y) \ni 0 \tag{12}$$

with $f : X \times Y \times V \to Z'$ and $F : X \times Y \rightrightarrows Z'$. We propose to solve (12) by the multistep Newton-type method

$$\tilde{f}(x_{k+1}, y_k, v_k) + \tilde{F}(x_{k+1}) \ni 0 \tag{13a}$$

$$f_{Hy}(x_{k+1}, y_{k+1}, y_k, v_k) + F(x_{k+1}, y_{k+1}) \ni 0 \tag{13b}$$

where, in the first step, $\tilde{f} : X \times Y \times V \to Z''$ and $\tilde{F} : X \rightrightarrows Z''$ provide a (possibly lower-order) generalized equation for x parametrized in y; in the second step, $f_{Hy}(\xi, y, \eta, v) = f_H(\xi, \xi, y, \eta, v)$ with

$$f_H : (x, \xi, y, \eta, v) \mapsto f(\xi, \eta, v) + H(\xi, \eta, v)(x - \xi, y - \eta)$$

and operator $H(\xi, \eta, v) : X \times Y \to Z'$ is a perturbed approximation of $f(\xi, y, 0)$ with respect to y around η.

The intuition behind the multistep Newton-type method is that the inclusion (13a) is solved separately, e.g., through a nested sequence of Newton-type iterations, to obtain x_{k+1}; whereas (13b) represents a Josephy–Newton step partially in y. The multistep Newton-type method is a useful framework to study bilevel optimization problems, where (13a) corresponds to the KKT system of a lower-level parametrized optimization problem. If the inclusion (13a) is solved inexactly, e.g., by limiting the number of nested Newton-type steps, the error would be reflected by the perturbation. We will demonstrate the usefulness of the multistep framework on the example of the augmented Lagrangian method which can be interpreted as solving the dual problem of (4), which leads to a bilevel optimization [22].

Assumption 2: Let $\bar{z} = (\bar{x}, \bar{y}) \in X \times Y$ and $\tilde{\kappa}, \kappa, \gamma_y, \gamma_v, \gamma_w > 0$ satisfy:

(a) $(\bar{x}, \bar{y})$ is a solution of (12);
(b) $\tilde{f}(\cdot, \bar{y}, 0) + \tilde{F}$ is strongly subregular with constant $\tilde{\kappa}$ at $\bar{x}$ for 0;
(c) $\tilde{f}(\cdot, y, v)$ is Lipschitz continuous uniformly in (y, v) at $(\bar{x}, \bar{y}, 0)$;
(d) $\tilde{f}(x, \cdot, \cdot)$ is Lipschitz continuous with constant γ_w uniformly in x at $(\bar{x}, \bar{y}, 0)$;
(e) $f_H(\cdot, \xi, \cdot, \eta, v)$ is Lipschitz continuous uniformly in (ξ, η, v) at $(\bar{x}, \bar{x}, \bar{y}, \bar{y}, 0)$;
(f) $f_H(x, \cdot, y, \cdot, v)$ is Lipschitz continuous with constant γ_y uniformly in (x, y, v) at $(\bar{x}, \bar{x}, \bar{y}, \bar{y}, 0)$;
(g) $f_H(x, \xi, y, \eta, \cdot)$ is Lipschitz continuous with constant γ_v uniformly in (x, ξ, y, η) at $(\bar{x}, \bar{x}, \bar{y}, \bar{y}, 0)$;
(h) $f_H(\cdot, \bar{x}, \cdot, \bar{y}, 0) + F$ is strongly regular with constant κ at $(\bar{x}, \bar{y})$ for 0;

and $\kappa\gamma_y < 1$.

Remark 1: The mapping $f_H(\cdot, \bar{x}, \cdot, \bar{y}, 0) + F$ is strongly regular if $f(\cdot, \cdot, 0) + F$ is strongly regular with constant $\tilde{\kappa}$ and $f_H(\cdot, \xi, \cdot, \eta, v) - f(\cdot, \cdot, 0)$ is Lipschitz continuous with constant smaller than $\tilde{\kappa}^{-1}$ uniformly in (ξ, η, v) [17, Theorem 8.6].

Remark 2: An immediate consequence of Assumption 2 is that the solution mapping $S : Y \times V \rightrightarrows X$ of (13a) has the isolated calmness property with constant $\tilde{\kappa}\gamma_w$ at $(\bar{y}, 0)$ for $\bar{x}$ by virtue of [17, Theorem 12.4].

Remark 3: If F is a piecewise polyhedral mapping, then strong subregularity is equivalent to $\bar{x}$ being an isolated point in $S(\bar{y}, 0)$, a consequence of outer Lipschitz continuity of piecewise polyhedral mappings [17, Theorem 12.5], which is again equivalent to a unique solution of (12).

For the following result, we impose the norm on $X \times Y$ as $\|(x, y)\| := \|x\| + \|y\|$; recall that $\mathbf{v} = (v_0, v_1, \dots) \subset V$.

Theorem 3: Under Assumption 2, there exists a sequence $\{(x_k, y_k)\}_{k=0}^{\infty}$ generated by (13) for (x_0, y_0) and $\mathbf{v} \subset V$ sufficiently close to $(\bar{x}, \bar{y}, 0)$ such that $\{y_k\}$ is unique and

$$\|(x_k, y_k) - (\bar{x}, \bar{y})\| \le \alpha_k \|(x_0, y_0)\| + \kappa\gamma_v \|\mathbf{v}\|_\infty \tag{14}$$

for all $k \in \mathbb{N}$ with $\alpha_k \to 0$ as $k \to \infty$, that is, (13) is locally input-to-state stable with gain $\kappa\gamma_v$.

Proof: Take $(y_k, v_k) \in Y \times V$ close to $(\bar{y}, 0)$; by strong subregularity (Assumption 2-b), there exists a solution x_{k+1} of (13a) close to $\bar{x}$. Let y_{k+1} solve (13b) and observe that

$$\begin{aligned}0 \in f(x_{k+1}, y_k, v_k) + F(x_{k+1}, y_{k+1}) \\ + H(x_{k+1}, y_k, v_k)(x_{k+1} - x_{k+1}, y_{k+1} - y_k)\end{aligned}$$

in other words, (x_{k+1}, y_{k+1}) is a Josephy–Newton step in the sense of (9) for (12) with $z_k = (x_{k+1}, y_k)$. By virtue of Theorem 2, the point y_{k+1} is unique, satisfies

$$\|(x_{k+1}, y_{k+1}) - \bar{z}\| \le \kappa\gamma_y \|(x_{k+1}, y_k) - \bar{z}\| + \kappa\gamma_v \|v_k\|$$

and hence,

$$\|y_{k+1} - \bar{y}\| \le \kappa\gamma_y \|y_k - \bar{y}\| + \kappa\gamma_v \|v_k\| \tag{15a}$$

by choice of $\|\cdot\|$ on $X \times Y$ and $\kappa\gamma_y < 1$ by Assumption 2. Moreover, the solution map S of (13a) has the isolated calmness property (Remark 2) and thus,

$$\|x_{k+1} - \bar{x}\| \le \tilde{\kappa}\gamma_w \|y_k - \bar{y}\| + \tilde{\kappa}\gamma_w \|v_k\|. \tag{15b}$$

Combining (15a) and (15b), we obtain (14) with $\alpha_k = (\kappa\gamma_y)^{k-1}(\kappa\gamma_y + \tilde{\kappa}\gamma_w)$, the desired result. ■

Remark 4: It should be noted that is never evaluated for a nonzero argument $x - \xi$ but in the theoretical analysis and hence can freely be chosen to satisfy the strong regularity condition in Assumption 2. In particular, a possible choice for H is

$$H(\xi, \eta, v) : (d_x, d_y) \mapsto f(\xi + d_x, \eta, v) + H_y(\xi + d_x, \eta, v)d_y - f(\xi, \eta, v)$$

with $H_y(\xi, \eta, v) : Y \to Z'$, that is,

$$f_H(x, \xi, y, \eta, v) \equiv f(x, \eta, v) + H_y(x, \eta, v)(y - \eta)$$

and regularity and continuity of f_H depend on f_{Hy} only.

We present applications of these results in nonlinear optimization in the next section.

C. Related Results

Previous works studied a Newton-type iteration of the form of (9) with $H(z, p) = \nabla_x f(z, p)$ to solve parametrized generalized equations, assuming Fréchet differentiability of f with respect to z and continuity of f and $\nabla_x f$. Under strong regularity assumptions similar to Assumption 1, the authors of [4] concluded that the sequence $\{z_k\}_{k=0}^{\infty}$ is locally unique and convergent to a solution $z(p)$ for any constant p sufficiently close to 0, and $\|z(p) - z(0)\| \leq \mu\|p\|$ for some constant $\mu > 0$. Furthermore, in [5], it was proven that the sequence satisfies

$$\sup_{k\in\mathbb{N}_+} \|z_k - \bar{z}\| \leq \alpha\|z_0 - \bar{z}\| + \gamma\|p\|$$

for some $\alpha < 1$ and $\gamma > 0$; this result is both necessary and sufficient for local input-to-state stability in the sense of (8).

Another classical topic in the study of Newton-type methods is the convergence of the iteration (2) or (6) if the right-hand side is a nonzero remainder, viz.

$$f(z_k) + H(z_k)(z_{k+1} - z_k) + F(z_k) \ni e_k$$

typically corresponding to solving inexactly the underlying linear equations (see, e.g., [23]). Using local input-to-state stability properties, we can immediately retrieve the desired convergence of $\{z_k\}_{k=0}^{\infty}$ to $\bar{z}$ if $\|e_k\| \to 0$.

IV. Applications

We apply the results of Theorems 2 and 3 to derive new robust convergence properties for nonlinear optimization algorithms. We consider a perturbed variant of (4), viz.

$$\min_{x\in C} h(x, v) \quad \text{subject to } g(x, v) = 0 \tag{16}$$

where we assume that h and g are twice continuously differentiable in x at $(\bar{x}, 0)$ uniformly in v; and the derivatives are Lipschitz continuous in v uniformly in x.

A. Approximate Sequential Quadratic Programming

A classical approach to sequential quadratic programming is the approximation of the Hessian of the (perturbed) Lagrangian $L(x, y, v) := h(x, v) + \langle g(x, v), y\rangle$ for (16), which appears in the upper-left block of the gradient when computing the (exact) Newton step for the perturbed variant of (5). Approximating the Hessian by a positive definite matrix B_{k+1} at step $k \in \mathbb{N}$, the Newton step then becomes equivalent to solving the convex program [17, Theorem 11.1]

$$\min_{x\in C} \frac{1}{2}\langle B_{k+1}(x - x_k), x - x_k\rangle + \nabla h(x_k, v_k)(x - x_k) \tag{17a}$$

$$\text{subject to } g(x_k, v_k) + \nabla g(x_k, v_k)(x - x_k) = 0 \tag{17b}$$

and taking $z_{k+1} = (x, y)_{k+1}$ as (unique) primal-dual solution of (17). Popular algorithms to compute the approximation B_{k+1} along the solution $\{(x, y)_k\}_{k\in\mathbb{N}}$ include the BFGS and DFP methods (named, respectively, for its discoverers), which belong to the larger Broyden class of Hessian update formulas and often provide superlinear convergence of the quasi-Newton iteration [24].

Assumption 3: Eq. (4) has an optimal solution $(\bar{x}, \bar{y}) \in X \times Y$ such that (5) is strongly regular at $(\bar{x}, \bar{y})$ for 0; the update $B_{k+1} = \Psi(B_k, z_{k+1}, v_k)$ is locally input-to-state stable around $\nabla^2 L(\bar{x}, \bar{y}, 0)$ with inputs (z_{k+1}, v_k).

Hessian approximations such as BFGS and DFP often require additional conditions to ensure that $B_k \to \nabla^2 L(\bar{x}, \bar{y}, 0)$. Here, however, we neglect the intricacies of the approximation and instead focus on the interplay between quasi-Newton step and Hessian update.

Proposition 3: Under Assumption 3, the quasi-Newton step of (17) with Hessian update $B_{k+1} = \Psi(B_k, z_{k+1}, v_k)$ is locally input-to-state stable.

Proof: Note that the KKT system of (17) can be written in the form of (9) with $z_k = (x_k, y_k)$,

$$H_k = \begin{pmatrix} \nabla^2 L(x_k, y_k, v_k) + w_k + e_k & \nabla g(x_k, v_k)^* \\ \nabla g(x_k, v_k) & 0 \end{pmatrix}$$

with $w_k = B_{k+1} - \nabla^2 L(x_k, y_k, 0)$ and $e_k = \nabla^2 L(x_k, y_k, 0) - \nabla^2 L(x_k, y_k, v_k)$. By virtue of Corollary 1, Assumption 3, and Lipschitz continuity of $\nabla^2 L$, we have that

$$\|z_{k+1} - \bar{z}\| \leq \alpha_1\|z_k - \bar{z}\| + \gamma_w\|w_k\| + \gamma_v\|v_k\|$$

and

$$\|w_{k+1}\| \leq \alpha_2\|w_k\| + \gamma_{Bz}\|z_{k+1} - \bar{z}\| + \gamma_{Bv}\|v_k\|$$

with $\alpha_1, \alpha_2 \in [0, 1)$, $\gamma_w, \gamma_v, \gamma_{Bz}, \gamma_{Bv} > 0$, and $\alpha_1, \gamma_w \to 0$ as $z_k \to \bar{z}$. Combining these results, we obtain

$$\|z_{k+1} - \bar{z}\| + \|w_{k+1}\| \leq \bar{\alpha}(\|z_k - \bar{z}\| + \|w_k\|) + \bar{\gamma}\|v_k\| \tag{18}$$

with $\bar{\alpha} = \max\{\alpha_1(1 + \gamma_B), \alpha_2 + \gamma(1 + \gamma_B)\}$; assuming that z_k is sufficiently close to $\bar{z}$ such that $\bar{\alpha} < 1$ gives the desired result. ■

B. Augmented Lagrangian Method

The augmented Lagrangian method solves the perturbed nonlinear program (16) by iterating over

$$x_{k+1} \in \arg\min_{x \in C} \left\{ h(x, v_k) + \langle y_k, g(x, v_k)\rangle + \frac{\varrho}{2}\|g(x, v_k)\|^2 \right\} \tag{19a}$$

$$y_{k+1} = y_k + \varrho g(x_{k+1}, v_k) \tag{19b}$$

for some penalty $\varrho > 0$ and perturbation $v_k \in V$. The cost function in (19a) is the titular *augmented Lagrangian*, parametrized in the dual variable y_k, and the necessary conditions can be written as a parametrized generalized equation

$$\nabla h(x, v_k) + \nabla g(x, v_k)^* y_k + \varrho \nabla g(x, v_k)^* g(x, v_k) + N_C(x) \ni 0 \tag{20}$$

provided that h and g are continuously differentiable. A classical result [22] states that, under mild assumptions and for sufficiently large (but finite) value of ϱ, the function minimized in (19a) with $v_k = 0$ becomes locally strictly convex and (19b) can be interpreted as gradient ascent for the dual problem. This also corresponds to strong regularity of (20) for all y_k around $\bar{y}$.

Assumption 4: Eq. (4) with $v = 0$ has an optimal solution $(\bar{x}, \bar{y}) \in X \times Y$ and (5) is strongly regular at $(\bar{x}, \bar{y})$ for 0.

An immediate consequence is strong subregularity of (20) for sufficiently large penalties; to that extent, we introduce

$$f_\varrho(x, y, y_k, v_k) = \begin{pmatrix} \nabla h(x, v_k) + \nabla g(x, v_k)^* y \\ g(x, v_k) + \varrho^{-1}(y_k - y) \end{pmatrix}$$

and study the augmented KKT system as follows.

Lemma 1: Under Assumption 4, there exist constants $\varrho_0 > 0$ and $k_{\varrho_0} > 0$ such that, for all $\varrho \geq \varrho_0$,

(a) $f_\varrho(\cdot, \cdot, \bar{y}, 0) + N_{C \times X^*}$ is strongly subregular at $(\bar{x}, \bar{y})$ for 0 with constant $k_\varrho \in (0, k_{\varrho_0}]$;

(b) Eq. (20) is strongly subregular at $\bar{x}$ for 0 with constant $k_\varrho \in (0, k_{\varrho_0}]$ if $y_k = \bar{y}$ and $v_k = 0$.

Proof: We observe that since (5) is strongly subregular at $(\bar{x}, \bar{y})$ for 0 with some constant $\kappa > 0$, the set-valued mapping

$$F_\varrho(x, y) := f_\varrho(x, y, \bar{y}, 0) + N_{C \times X^*}((x, y)) \tag{21}$$

is strongly subregular at $(\bar{x}, \bar{y})$ for 0 with constant $k_\varrho = \varrho\kappa/(\varrho - \kappa)$ for all $\varrho > \kappa$ [17, Theorem 12.2]. Note that k_ϱ is strictly decreasing as $\varrho \to \infty$. Substituting $y_x = \bar{y} + \varrho g(x, 0)$, we have that $F_\varrho(x, y_x) \ni (\delta, 0)$ if and only if

$$\nabla h(x, 0) + \nabla g(x, 0)^* \bar{y} + \varrho \nabla g(x, 0)^* g(x, 0) + N_C(x) \ni \delta$$

for all δ around 0. Hence, (20) is strongly subregular at $\bar{x}$ for 0 with constant $k_\varrho \leq k_{\varrho_0}$ if $y_k = \bar{y}$ and $\varrho \geq \varrho_0 > \kappa$. ■

We show that the augmented Lagrangian method is an instance of the multistep Newton-type method (13) for the generalized equation (5) in (x, y), hence proving local input-to-state stability. Note that our approach does not require f to be twice differentiable.

Proposition 4: Under Assumption 4, the iteration (19) is locally input-to-state stable around $(\bar{x}, \bar{y})$ for all $\varrho \geq \bar{\varrho} > 0$.

Proof: Let (x_{k+1}, y_{k+1}) be the result of (19) for a given $(y_k, v_k) \in Y \times V$; then

$$f_\varrho(x_{k+1}, y_{k+1}, y_k, v_k) + N_{C \times X^*}((x_{k+1}, y_{k+1})) \ni 0 \tag{22}$$

for any $\varrho > 0$. Eq. (22) corresponds to a partial Newton step for (5) in the sense of (13b) and Remark 4, where

$$f_{Hy}(\xi, y, \eta, v) = \begin{pmatrix} \nabla h(\xi, v) + \nabla g(\xi, v)^* \eta \\ g(\xi, v) \end{pmatrix} + \begin{bmatrix} \nabla g(\xi, v)^* \\ -\varrho^{-1} \end{bmatrix} (y - \eta)$$

and $f_{Hy}(\xi, y, \cdot, v)$ is Lipschitz continuous with constant ϱ^{-1} uniformly in (ξ, y, v). Moreover, $f_{Hy}(\cdot, \cdot, \bar{y}, 0) + N_{C \times X^*}$ is strongly subregular with constant $k_\varrho \leq k_{\varrho_0}$ at $(\bar{x}, \bar{y})$ for 0 for all $\varrho \geq \varrho_0$ by virtue of Lemma 1. Pick $\bar{\varrho} \geq \varrho_0$ such that $\bar{\varrho}^{-1} k_{\varrho_0} < 1$; the desired result follows from Theorem 3 for any $\varrho \geq \bar{\varrho}$. ■

V. Conclusions

Newton methods for generalized equations play a major role in nonlinear optimization. Our local input-to-state stability result for the perturbed Josephy–Newton method enables the study of optimization algorithms interconnected with dynamic systems, such as in optimization-based control, under perturbations or uncertain conditions. In addition, our locally input-to-state stable multistep Newton-type method allows for advanced optimization techniques as demonstrated on the augmented Lagrangian method. Further work will focus on relaxations of strong regularity and Lipschitz continuity conditions within the general ISS framework.

Appendix

We provide implicit function theorems for generalized equations with multiple parameters, extending [17, Theorems 8.5 and 12.4]. To that extent, define

$$\widehat{\mathrm{lip}}_x(f; (\bar{x}, \bar{p})) = \limsup_{\substack{x_1, x_2 \to \bar{x}, x_1 \neq x_2 \\ p \to \bar{p}}} \frac{\|f(x_1, p) - f(x_2, p)\|}{\|x_1 - x_2\|}$$

for $f : X \times P \to Y$, and note that $f(\cdot, p)$ is Lipschitz continuous with constant γ uniformly in p around $(\bar{x}, \bar{p})$ if and only if $\widehat{\mathrm{lip}}_x(f; (\bar{x}, \bar{p})) \leq \gamma$. We consider the parametrized generalized equation

$$f(x, p_1, p_2) + F(x) \ni 0 \tag{23}$$

with solution map

$$S : p = (p_1, p_2) \mapsto \{x \in X \mid (x, p) \text{ solves } (23)\}$$

and $P = P_1 \times P_2$.

Theorem 4: Let $(\bar{x}, \bar{p}) \in \operatorname{graph} S$ and suppose that $h : X \to Y$ satisfies

(a) $f(\bar{x}, \bar{p}) = h(\bar{x})$;

(b) $h + F$ is strongly subregular with constant κ at $\bar{x}$ for 0;

(c) $f(\cdot, p) - h$ is Lipschitz continuous with constant μ uniformly in p at $(\bar{x}, \bar{p})$;

(d) $f(x, \cdot, p_2)$ and $f(x, p_1, \cdot)$ are Lipschitz continuous uniformly in (x, p) at $(\bar{x}, \bar{p})$;

and $\kappa\mu < 1$; then the solution $S(\cdot)$ of (23) has the isolated calmness property at $\bar{p}$ for $\bar{x}$ satisfying

$$\|x - \bar{x}\| \leq \omega\, \widehat{\mathrm{lip}}_{p_1}(f; (\bar{x}, \bar{p}))\|p_1 - \bar{p}_1\| + \omega\, \widehat{\mathrm{lip}}_{p_2}(f; (\bar{x}, \bar{p}))\|p_2 - \bar{p}_2\|$$

with $\omega = (1 - \kappa\mu)^{-1}\kappa$ for all $(p_1, p_2, x) \in \operatorname{graph} S$ in a neighbourhood of $(\bar{x}, \bar{p})$.

Proof: The proof is analogous to [17, Proof of Theorem 12.4] using that

$$\|f(x, p_1, p_2) - f(x, \bar{p}_1, \bar{p}_2)\| \leq \gamma_1\|p_1 - \bar{p}_1\| + \gamma_2\|p_2 - \bar{p}_2\|$$

for all (x, p) around $(\bar{x}, \bar{p})$ with some constants $\gamma_1, \gamma_2 \geq 0$ by uniform Lipschitz continuity[3] of f. ■

Corollary 2: If the assumptions of Theorem 4 hold with $h + F$ being strongly regular with constant κ at $\bar{x}$ for 0, then the solution $S(\cdot)$ of (23) has a single-valued localization $s : P_1 \times P_2 \to X$ at $\bar{p}$ for $\bar{x}$ satisfying

$$\|s(p_1, p_2) - \bar{x}\| \leq \omega\, \widehat{\mathrm{lip}}_{p_1}(f; (\bar{x}, \bar{p}))\|p_1 - \bar{p}_1\| + \omega\, \widehat{\mathrm{lip}}_{p_2}(f; (\bar{x}, \bar{p}))\|p_2 - \bar{p}_2\|$$

with $\omega = (1 - \kappa\mu)^{-1}\kappa$ for all (p_1, p_2) in a neighbourhood of $\bar{p}$. ◁

[3] In fact, uniform calmness would suffice.

References

[1] G. Giorgi, A. Guerraggio, and J. Thierfelder, *Mathematics of Optimization: Smooth and Nonsmooth Case*. Amsterdam: Elsevier, 2004.

[2] A. F. Izmailov and M. V. Solodov, *Newton-Type Methods for Optimization and Variational Problems*. Cham: Springer International Publishing, 2014.

[3] T. F. Coleman and A. R. Conn, "On the Local Convergence of a Quasi-Newton Method for the Nonlinear Programming Problem," *SIAM Journal on Numerical Analysis*, vol. 21, no. 4, pp. 755–769, 1984.

[4] S. P. Dokov and A. L. Dontchev, "Robinson's Strong Regularity Implies Robust Local Convergence of Newton's Method," in *Optimal Control: Theory, Algorithms, and Applications*, W. H. Hager and P. M. Pardalos, Eds. New York, NY: Springer, 1998, pp. 116–129.

[5] A. L. Dontchev and R. T. Rockafellar, "Newton's method for generalized equations: A sequential implicit function theorem," *Mathematical Programming*, vol. 123, no. 1, pp. 139–159, 2010.

[6] ——, "Convergence of inexact newton methods for generalized equations," *Mathematical Programming*, vol. 139, no. 1-2, pp. 115–137, 2013.

[7] D. Liao-McPherson, M. M. Nicotra, and I. Kolmanovsky, "Time-distributed optimization for real-time model predictive control: Stability, robustness, and constraint satisfaction," *Automatica*, vol. 117, no. October, p. 108973, 7 2020.

[8] J. Leung, D. Liao-McPherson, and I. V. Kolmanovsky, "A Computable Plant-Optimizer Region of Attraction Estimate for Time-distributed Linear Model Predictive Control," in *Proceedings of the American Control Conference*, 2021, pp. 3384–3391.

[9] G. Belgioioso, D. Liao-Mcpherson, M. H. De Badyn, S. Bolognani, J. Lygeros, and F. Dorfler, "Sampled-Data Online Feedback Equilibrium Seeking: Stability and Tracking," in *Proceedings of the IEEE Conference on Decision and Control*, Austin, TX, 2021, pp. 2702–2708.

[10] A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, "Optimization Algorithms as Robust Feedback Controllers," *Annual Reviews in Control*, 2021.

[11] T. Skibik and M. M. Nicotra, "Analysis of Time-Distributed Model Predictive Control When Using a Regularized Primal-Dual Gradient Optimizer," *IEEE Control Systems Letters*, vol. 7, pp. 235–240, 2023.

[12] E. D. Sontag, "Input to State Stability: Basic Concepts and Results," in *Nonlinear and Optimal Control Theory*, P. Nistri and G. Stefani, Eds. Berlin, Heidelberg: Springer, 2004, pp. 163–220.

[13] A. Hasan, E. C. Kerrigan, and G. A. Constantinides, "Control-theoretic forward error analysis of iterative numerical algorithms," *IEEE Transactions on Automatic Control*, vol. 58, no. 6, pp. 1524–1529, 2013.

[14] G. G. Colabufo, P. M. Dower, and I. Shames, "Newton's method: Sufficient conditions for practical and input-to-state stability," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 6334–6339, 2020.

[15] E. D. Sontag, "Remarks on input to state stability of perturbed gradient flows, motivated by model-free feedback control learning," *Systems and Control Letters*, vol. 161, p. 105138, 2022.

[16] T. Cunis and I. Kolmanovsky, "Input-to-State Stability of a Bilevel Proximal Gradient Descent Algorithm*," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 7474–7479, 2023.

[17] A. L. Dontchev, *Lectures on Variational Analysis*, ser. Applied Mathematical Sciences. Cham: Springer, 2021, no. 205.

[18] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation*, ser. Grundlehren der mathematischen Wissenschaften. Berlin, Heidelberg: Springer, 2006, vol. 2, no. 331.

[19] A. Grothey, "Decomposition Methods for Nonlinear Nonconvex Optimization Problems," Ph.D. dissertation, The University of Edinburgh, Edinburgh, 2001.

[20] C. M. Kellett, "A compendium of comparison function results," *Mathematics of Control, Signals, and Systems*, vol. 26, no. 3, pp. 339–374, 2014.

[21] Z.-P. Jiang and Y. Wang, "Input-to-state stability for discrete-time nonlinear systems," *Automatica*, vol. 37, no. 6, pp. 857–869, 2001.

[22] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*. New York: Academic Press, 1982.

[23] R. S. Dembo, S. C. Eisenstat, and T. Steihaug, "Inexact Newton Methods," *SIAM Journal on Numerical Analysis*, vol. 19, no. 2, pp. 400–408, 1982.

[24] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York, NY: Springer, 2006.