Primal-Dual Stability in Local Optimality

Matúš Benko, [1] ***R. Tyrrell Rockafellar*** [2]

Dedicated to Boris Mordukhovich for his 75th birthday

Abstract

Much is known about when a locally optimal solution depends in a single-valued Lipschitz continuous way on the problem's parameters, including tilt perturbations. Much less is known, however, about when that solution and a uniquely determined multiplier vector associated with it exhibit that dependence as a primal-dual pair. In classical nonlinear programming, such advantageous behavior is tied to the combination of the standard strong second-order sufficient condition (SSOC) for local optimality and the linear independent gradient condition (LIGC) on the active constraint gradients. But although second-order sufficient conditons have successfully been extended far beyond nonlinear programming, insights into what should replace constraint gradient independence as the extended dual counterpart have been lacking.

The exact answer is provided here for a wide range of optimization problems in finite dimensions. Behind it are advances in how coderivatives and strict graphical derivatives can be deployed. New results about strong metric regularity in solving variational inequalities and generalized equations are obtained from that as well.

Keywords: *second-order variational analysis, local optimality, primal-dual stability, tilt stability, full stability, metric regularity, Kummer's inverse theorem, implicit mapping theorems, graphically Lipschitzian mappings, crypto-continuity, strict graphical derivatives, coderivatives, variational sufficiency.*

Version of 28 December 2023

[1] Johann Radon Institute for Computational and Applied Mathematics, Linz, Austria;
E-mail: *Matus.Benko@oeaw.ac.at*

[2] University of Washington, Department of Mathematics, Box 354350, Seattle, WA 98195-4350;
E-mail: *rtr@uw.edu*, URL: sites.math.washington.edu/∼rtr/mypage.html

1 Introduction

Stability of solutions is a topic of fundamental importance in optimization. One reason is practical and numerical. The results obtained from an algorithm can be influenced by approximations or by inaccuracies in the data input as well as errors in computation that feed into a stopping criterion. In effect, a method might yield a solution to a slightly different problem than the one intended. How serious might that be? Another reason is the role that stability plays in validating even the formulation of a problem. There is a time-tested tradition in applied mathematics of assessing whether a problem is "well posed" or not, and response to possible perturbations in parameters is a big part of that. What good is a mathematical model if its solutions are too delicate to be determined reliably?

In optimization, the issues come sharply into focus already in the very simple setting where the problem is to

$$\text{find a local minimizer } \bar{x} \text{ of } f_0(x) \text{ on } I\!R^n \tag{1.1}$$

for a $\mathcal{C}^2$ function f_0. Necessarily $\nabla f_0(\bar{x}) = 0$, so an algorithm might take solving the equation $\nabla f_0(x) = 0$ as a surrogate for (1.1) — and yet only be able to reach some $\tilde{x}$ such that $\nabla f_0(\tilde{x})$ is "small." That would correspond to having $\nabla f_0^v(\tilde{x}) = 0$ for the tilted function $f_0^v(x) = f_0(x) - v{\cdot}x$ for small v. The question then is how well $\tilde{x}$ might be expected to approximate a local minimizer of f_0 itself. For instance, when might there be "tilt stability" in the following sense at a point $\bar{x}$ where $\nabla f_0(\bar{x}) = 0$? *Whether there exist neighborhoods $\mathcal{X}$ of $\bar{x}$ and $\mathcal{V}$ of 0 such that, for each $v \in \mathcal{V}$ there is a unique $x \in \mathcal{X}$ having $\nabla f_0^v(x) = 0$, with x moreover minimizing f_0^v over $\mathcal{X}$ and depending in a Lipschitz continuous way on v.* As shown in [13], this holds if and only if the Hessian $\nabla^2 f_0(\bar{x})$ is positive-definite. Without that Hessian property, therefore, such tilt stability is missing and $\bar{x}$ falls short of being a "good" solution that can surely be trusted in environment of potential inexactness.

Similar considerations where constraints are present enter as in classical nonlinear programming, where the problem is to

$$\begin{array}{c}\text{find a local minimizer } \bar{x} \text{ of } f_0(x,\bar{p}) \text{ subject to} \\ f_i(x,\bar{p}) \begin{cases} \leq 0 & \text{for } i = 1,\dots,s, \\ = 0 & \text{for } i = s+1,\dots,m, \end{cases}\end{array} \tag{1.2}$$

where each $f_i(x,p)$ is jointly $\mathcal{C}^2$ in $x \in I\!R^n$ and p as a parameter vector in $I\!R^d$. An issue then is how well a solution $\tilde{x}$ to a nearby problem, in which the designated $\bar{p}$ is replaced by a nearby p, might be expected to approximate a desired $\bar{x}$. For instance, might there be "full stability" in the following sense at a point $\bar{x}$ where the first-order Karush-Kuhn-Tucker (KKT) conditions are satisfied? *Whether there exist neighborhoods $\mathcal{X}$ of $\bar{x}$ and $\mathcal{P}$ of $\bar{p}$ such that, for each $p \in \mathcal{P}$ there is a unique $x \in \mathcal{X}$ satisfying the KKT conditions in the modified problem (with p in place of $\bar{p}$), with x moreover giving the minimum relative to $\mathcal{X}$ and depending in a Lipschitz continuous way on p.* Note that tilt stability is covered within this, because tilts could be part of the way that $f_0(x,p)$ is influenced by p.

In nonlinear programming theory, Lagrange multipliers have customarily been added to this by looking not just at $\bar{x}$, but at a KKT pair $(\bar{x},\bar{y})$ and a neighborhood $\mathcal{Y}$ of $\bar{y}$. Stability is seen not only in terms of local optimality of the primal component, but with $p \in \mathcal{P}$ yielding a unique KKT pair $(x,y) \in \mathcal{X} \times Y$ that depends in a Lipschitz continuous way on p. This can be called "primal-dual full stability." For that there is a known criterion from [3], as long as the parameterization is ample from the angle that the $(m+1) \times n$ matrix with the gradients $\nabla_x f_i(\bar{x},\bar{p})$ as its rows has rank $m+1$. It's present under that if and only if the standard strong second-order sufficient condition for local optimality holds for $(\bar{x},\bar{y})$ and, in addition, the gradients $\nabla_x f_i(\bar{x},\bar{p})$ for the active constraints (the ones with $f_i(\bar{x},\bar{p}) = 0$) are linearly independent: SSOC+LIGC.

Our goal in this paper is extending primal-dual full stability in nonlinear programming to a much broader optimization format where the targeted problem is to

$$\bar{\mathcal{P}} \qquad \text{minimize } \varphi(x,0) \text{ with respect to } x$$

for a closed proper function φ on $I\!R^n \times I\!R^m$, but is viewed as embedded in a parameterized family of problems

$$\mathcal{P}(v,u) \qquad \text{minimize } \varphi(x,u) - v{\cdot}x \text{ with respect to } x$$

as variants, so that

$$\bar{\mathcal{P}} = \mathcal{P}(\bar{v},\bar{u}) \text{ for } (\bar{v},\bar{u}) = (0,0).$$

We fix a locally optimal solution $\bar{x}$ to $\bar{\mathcal{P}}$ and investigate it with the help of subgradients of φ in the sense of variational analysis [18, 8B]. We suppose at $\bar{x}$ that the *basic constraint qualification* in terms of horizon subgradients is satisfied:

$$(0,y) \in \partial^{\infty}\varphi(\bar{x},\bar{u}) \implies y = 0. \tag{1.3}$$

This guarantees by [18, 10.11+10.12] that

$$\bar{v} \in \partial_x\varphi(\bar{x},\bar{u}) \implies \exists\, \bar{y} \text{ such that } (\bar{v},\bar{y}) \in \partial\varphi(\bar{x},\bar{u}) \tag{1.4}$$

and more broadly through [18, 10.16] that the closed-valued mapping

$$\begin{aligned} &Y : (x,u,v) \in \operatorname{gph}\partial_x\varphi \mapsto \{\, y \mid (v,y) \in \partial\varphi(x,u)\} \text{ is nonempty-valued and locally} \\ &\text{bounded when } (x,u,v) \text{ and } \varphi(x,u) \text{ are near enough to } (\bar{x},\bar{u},\bar{v}) \text{ and } \varphi(\bar{x},\bar{u}). \end{aligned} \tag{1.5}$$

The elementary necessary condition for local optimality in $\mathcal{P}$ on the left of (1.4) is thereby augmented by multiplier vectors $\bar{y}$ on the right, which can be partnered with $\bar{x}$ for the analysis of primal-dual stability that will be undertaken.

We assume for technical support, including assurance of local closedness of graphs of subgradient mappings where needed, that

$$\text{for every } \bar{y} \in Y(\bar{x},\bar{u},\bar{v}),\ \varphi \text{ is continuously prox-regular at } (\bar{x},\bar{u}) \text{ for } (\bar{v},\bar{y}). \tag{1.6}$$

The "continuous prox-regularity" means two things. First (prox-regularity), there are neighborhoods $\mathcal{X} \times \mathcal{U}$ of $(\bar{x},\bar{u})$ and $\mathcal{V} \times \mathcal{Y}$ of $(\bar{v},\bar{y})$ and $r > 0$ such that

$$\begin{aligned} &\varphi(x',u') \geq \varphi(x,u) + (v,y){\cdot}(x'-x,u'-u) - \tfrac{r}{2}|(x'-x,u'-u)|^2 \\ &\text{for } (x,u),\,(x',u') \in \mathcal{X} \times \mathcal{U} \text{ and } (v,y) \in \partial\varphi(x,u) \cap [\mathcal{V} \times \mathcal{Y}], \end{aligned} \tag{1.7}$$

and second (subdifferential continuity), the function

$$(x,u,v,y) \in \operatorname{gph}\partial\varphi \mapsto \varphi(x,u) \text{ is continuous in } \mathcal{X} \times \mathcal{U} \times \mathcal{V} \times \mathcal{Y}. \tag{1.8}$$

Because the continuous prox-regularity (a shortened term for prox-regularity plus subdifferential continuity) being assumed for every $\bar{y}$ in $Y(\bar{x},\bar{u},\bar{v})$, which is compact, the combination (1.7)+(1.8) actually holds for a set $\mathcal{Y}$ having all of $Y(\bar{x},\bar{u},\bar{v})$ in its interior. Moreover, we then have by way of (1.5) a far stronger "multiplier rule" than (1.4),

$$\begin{aligned} &\text{for } (x,u,v) \text{ in some neighborhood of } (\bar{x},\bar{u},\bar{v}), \\ &v \in \partial_x\varphi(x,u) \iff \exists\, y \text{ such that } (v,y) \in \partial\varphi(x,u). \end{aligned} \tag{1.9}$$

As a prominent example available from [18, 13.32], φ satisfies (1.6) when it is *strongly amenable* at $(\bar{x},0)$ in the sense of having, in a neighborhood of $(\bar{x},0)$, a representation as the composition of a closed proper convex function with a $\mathcal{C}^2$ mapping under a standard constraint qualification [18, 10F]. A case of strong amenability will be the structure of φ adopted in our final Section 4.

With this as the platform, we pair the designated locally optimal solution $\bar{x}$ in $\bar{\mathcal{P}}$ with a particular multiplier vector $\bar{y}$ in (1.4) and aim at understanding the potential "stability" of $(\bar{x},\bar{y})$ with respect to possible shifts of (v,u) away from $(0,0)$ as seen through behavior of the set-valued mappings

$$\begin{aligned} M(v,u) &= \{\, x \mid v \in \partial_x\varphi(x,u)\}, \ \text{ with } \ \bar{x} \in M(\bar{v},\bar{u}), \\ \bar{M}(v,u) &= \{\, (x,y) \mid (v,y) \in \partial\varphi(x,u)\}, \ \text{ with } \ (\bar{x},\bar{y}) \in \bar{M}(\bar{v},\bar{u}). \end{aligned} \tag{1.10}$$

Note that the graphs of these mappings are closed locally around the elements associated with $\bar{\mathcal{P}}$ by virtue of (1.8).[3]

Stability just of $\bar{x}$ without $\bar{y}$ has already been studied in this general setting in [8]. The focus there was on the generally set-valued mapping

$$M_\delta(v,u) = \operatorname*{argmin}_{|x-\bar{x}|\le\delta} \{\, \varphi(x,u) - v{\cdot}x\}, \ \text{ where } \ \bar{x} \in M_\delta(\bar{v},\bar{u}), \tag{1.11}$$

in its relationship to M and the localized optimal value function

$$m_\delta(v,u) = \min_{|x-\bar{x}|\le\delta} \{\, \varphi(x,u) - v{\cdot}x\}. \tag{1.12}$$

The following concept was introduced.

Definition 1.1 (primal full stability [8])**.** *The locally optimal solution $\bar{x}$ to $\bar{\mathcal{P}}$ is fully stable if there is a neighborhood $\mathcal{V}\times\mathcal{U}$ of $(\bar{v},\bar{u}) = (0,0)$ such that, for $\delta > 0$ sufficiently small, the mapping M_δ is single-valued and Lipschitz continuous on $\mathcal{V}\times\mathcal{U}$. Also, m_δ is Lipschitz continuous on $\mathcal{V}\times\mathcal{U}$.*

The demand for local Lipschitz continuity of the function m_δ is in fact superfluous as part of this definition, because it follows from the properties demanded of M_δ under the constraint qualification (1.3), according to [18, 10.14(a)].

The main result of [8] characterized primal full stability in terms of *coderivatives* of the partial subgradient mapping $\partial_x\varphi$. Such coderivatives, and the strict graphical derivatives that are about to appear as well, will be explained in Section 2; for additional background, see [18].

Theorem 1.2 (criterion for primal full stability [8, Theorem 2.3])**.** *For the full stability of $\bar{x}$ as defined in terms of M_δ, the following combination of conditions is both necessary and sufficient:*

(a) $(v',y') \in D^*[\partial_x\varphi](\bar{x},\bar{u}\,|\,\bar{v})(x'),\ x' \neq 0 \implies x'{\cdot}v' > 0$,

(b) $(0,y') \in D^*[\partial_x\varphi](\bar{x},\bar{u}\,|\,\bar{v})(0) \implies y' = 0$.

Moreover then, for (v,u) in a sufficiently small neighborhood of $(\bar{v},\bar{u}) = (0,0)$,

$$M_\delta(v,u) = M(v,u) \cap \{\, x \mid |x-\bar{x}| < \delta\}. \tag{1.13}$$

The assumptions made in [8] in proving Theorem 1.2 were slightly weaker than our (1.6), having in place of that the continuous prox-regularity of the functions $\varphi(\cdot,u)$ with local uniformity (same r) with respect to u in a neighborhood of $\bar{u}$. That's implied here by (1.7) with $u' = u$.

[3] The subgradient definitions [18, 8B] generally require attention being paid not only to limits of sequences of vectors, but also the limits of function values associated with them. But (1.8) makes the latter be automatic for φ, ensuring closedness of $\operatorname{gph}\partial\varphi$, $\operatorname{gph} Y$ and $\operatorname{gph}\bar{M}$, and then $\operatorname{gph}\partial_x\varphi$ and $\operatorname{gph} M$ inherit closeness as projections inherit closeness as projections under the local boundedness in (1.5).

The final assertion of Theorem 1.2 is significant because, in principle, even if $M_\delta(v,u)$ consists of a single $x \in M(v,u)$, there might conceivably be other points in $M(v,u)$ that don't report the minimum value $m_\delta(v,u)$ over the δ-ball around $\bar{x}$, yet signal a more local minimum at a higher level. This leads to an observation that before now hasn't explicitly been recorded (where it should be kept in mind that the property of m_δ in Definition 1.1 is automatic from that of M_δ).

Corollary 1.3 (alternative portrayal of primal full stability)**.** *Full stability of $\bar{x}$ as a local minimizer in $\bar{\mathcal{P}}$ is equivalent to the existence of neighborhoods $\mathcal{X}$ of $\bar{x}$ and $\mathcal{V}\times\mathcal{U}$ of $(\bar{v},\bar{u}) = (0,0)$ such that the mapping*

$$(v,u) \in \mathcal{V}\times\mathcal{U} \;\mapsto\; x \in \mathcal{X} \cap M(v,u) \tag{1.14}$$

is single-valued and Lipschitz continuous, with x always a local minimizer in $\mathcal{P}(v,u)$ relative to $\mathcal{X}$.

Work on primal full stability has continued in [12], [9], and elsewhere, but here we are aiming instead at a counterpart to Theorem 1.2 that is instead primal-dual. Parallel to M_δ we define

$$\bar{M}_\delta(v,u) = \{\,(x,y)\,|\, x \in M_\delta(v,u),\; y \in Y(x,u,v)),\; |y-\bar{y}| \le \delta\}, \quad \delta > 0 \tag{1.15}$$

for the mapping Y in (1.5) and inquire about its single-valuedness and relationship to $\bar{M}(v,u)$.

Definition 1.4 (primal-dual full stability)**.** *The primal-dual pair $(\bar{x},\bar{y})$ is fully stable in problem $\bar{\mathcal{P}}$ if there is a neighborhood $\mathcal{V}\times\mathcal{U}$ of $(\bar{v},\bar{u}) = (0,0)$ such that, for $\delta > 0$ sufficiently small, the mapping $\bar{M}_\delta$ is single-valued and Lipschitz continuous on $\mathcal{V}\times\mathcal{U}$, and the function m_δ is likewise Lipschitz continuous on $\mathcal{V}\times\mathcal{U}$.*

Again, as noted after Definition 1.1, the property of m_δ in Definition 1.4 follows from the ones demanded of $\bar{M}_\delta$ and thus doesn't really need to be included as a seemingly separate demand.

A key new result, to be confirmed in Section 3,[4] imitates Theorem 1.2 at this different level. Rather than coderivatives of $\partial_x\varphi$, it uses the strict graphical derivatives of $\partial\varphi$ itself.

Theorem 1.5 (criterion for primal-dual full stability)**.** *For the full stability $(\bar{x},\bar{y})$ in Definition 1.4, the combination of (a) and (b) of Theorem 1.1 with following condition is both necessary and sufficient:*

(c) $(0,y') \in D_*[\partial\varphi](\bar{x},\bar{u}\,|\,\bar{v})(0,0) \quad\Longrightarrow\quad y' = 0.$

Moreover then, for (v,u) in a sufficiently small neighborhood of $(\bar{v},\bar{u}) = (0,0)$,

$$\bar{M}_\delta(v,u) = \bar{M}(v,u) \cap \{\,(x,y)\,|\,|x-\bar{x}| < \delta,\; |y-\bar{y}| < \delta\}. \tag{1.16}$$

The result reveals at the end once more that the stability property in question is equivalent to a property that, on the surface, appears to be stronger.

Corollary 1.6 (alternative portrayal of primal-dual full stability)**.** *The full stability of the primal-dual pair $(\bar{x},\bar{y})$ in Definition 1.4 is equivalent to the existence of neighborhoods $\mathcal{X}\times\mathcal{Y}$ of $(\bar{x},\bar{y})$ and $\mathcal{V}\times\mathcal{U}$ of $(\bar{v},\bar{u}) = (0,0)$ such that the mapping*

$$(v,u) \in \mathcal{V}\times\mathcal{U} \;\mapsto\; (x,y) \in \bar{M}(v,u) \cap [\mathcal{X}\times\mathcal{Y}] \tag{1.17}$$

is single-valued and Lipschitz continuous, with x always a local minimizer in $\mathcal{P}(v,u)$ relative to $\mathcal{X}$.

Theorems 1.2 and 1.5 offer a definitive answer in theory to key questions about solution stability, but the partial coderivatives in conditions (a) and (b) can be hard to connect with the specifics of a problem's structure, despite advances in second-order calculus such as in [11]. Often only an inclusion

[4] An immediate consequence of Theorem 3.1, as indicated after its proof.

rule can be invoked, and that means in practice that the necessary and sufficient pair (a)+(b) is replaced by something that's only sufficient.

The strict graphical derivatives in (c) of Theorem 1.5 mark a turn in this subject, for which new support will be provided in Section 2. They will enable us to develop a different approach to stability which can bypass (a)+(b) and take advantage of recent ideas involving variational convexity [15] in second-order conditions for local optimality. Our main result, Theorem 3.5, reveals that the primal-dual full stability in Theorem 1.5 follows from combining (c) with the strong variational sufficient condition of [16], [17], in place of (a)+(b).

The key attraction to strict graphical derivatives is their essential role in the inverse mapping theorem of Kummer [7]; see also [18, 9.54]. They are employed there in a necessary and sufficient condition for the generally set-valued inverse of mapping from a space $\mathbb{R}^N$ into itself to have a single-valued localization that is Lipschitz continuous. The usefulness of that result has been hampered, however, by two things. Strict graphical derivatives can be difficult to determine, and more crucially, the mapping in Kummer's theorem is asked to have a kind of inner continuity property that itself may be difficult to verify.

We get around the latter difficulty through a new observation, namely that the inner continuity assumption can be dropped if the graph of the mapping is locally a "continuous manifold," which is true in particular for graphically Lipschitzian mappings derived from subgradient mappings of prox-regular functions or from local maximal monotonicity. We ameliorate the former difficulty over knowing strict graphical derivatives by reducing the need for that to relatively elementary cases where they are more accessible.

Strict graphical derivatives and graphically Lipschitzian mappings have also been featured, although differently, in recent work of Gfrerer and Outrata [5] and Hang and Sarabi [6].

Before addressing the stability issues in optimization, we lay out, in Section 2, a fresh line of results in implicit mapping theory that relies on our new observation about Kummer's inverse function theorem. Applications to optimization theory in general are taken up after that in Section 3 and specialized to nonlinear programming and its extensions in Section 4.

2 Lipschitzian localizations from strict graphical derivatives

For the time being, we put optimization aside and deal with deeper issues of when a generally set-valued mapping $S : \mathbb{R}^N \rightrightarrows \mathbb{R}^M$ with closed graph

$$\operatorname{gph} S = \{ (x,v) \,|\, v \in S(x) \} \tag{2.1}$$

might have a single-valued localization that is Lipschitz continuous. A single-valued localization of S at $\bar{x}$ relative to an element $\bar{v} \in S(\bar{x})$ is a single-valued mapping $s : \mathcal{X} \to \mathcal{V}$ with $s(\bar{x}) = \bar{v}$ that has $\operatorname{gph} s = [\mathcal{X} \times \mathcal{V}] \cap \operatorname{gph} S$ for a neighborhood $\mathcal{X} \times \mathcal{V}$ of $(\bar{x}, \bar{v})$. The existence of a single-valued localization that is Lipschitz continuous corresponds to the inverse mapping S^{-1} being *strongly metrically regular* at $\bar{v}$ for $\bar{x}$; see [4], [18].

The graphical derivative of S at $\bar{x}$ for $\bar{v}$ is the set-valued mapping $DS(\bar{x} \,|\, \bar{v}) : \mathbb{R}^N \rightrightarrows \mathbb{R}^M$ having as its graph the tangent cone to $\operatorname{gph} S$ at $(\bar{x}, \bar{v})$. In terms of difference quotient mappings

$$\Delta_t S(x \,|\, v)(x') = \frac{1}{t}[S(x + tx') - v] \;\text{ for }\; v \in S(x),\; t > 0, \tag{2.2}$$

this means in set limits that

$$\begin{aligned} DS(\bar{x}\,|\,\bar{v})(\bar{x}') &= \limsup_{t\searrow 0,\, x'\to\bar{x}'} \Delta_t S(\bar{x}\,|\,\bar{v})(x'), \quad\text{or} \\ \operatorname{gph} DS(\bar{x}\,|\,\bar{v}) &= \limsup_{t\searrow 0} \operatorname{gph} \Delta_t S(\bar{x}\,|\,\bar{v}). \end{aligned} \tag{2.3}$$

In contrast, the strict graphical derivative $D_*S(\bar{x}\,|\,\bar{v}) : I\!R^N \rightrightarrows I\!R^M$, having as its graph the paratingent cone to $\operatorname{gph} S$ at $(\bar{x}, \bar{v})$, is defined by

$$\begin{aligned} D_*S(\bar{x}\,|\,\bar{v})(\bar{x}') &= \limsup_{\substack{t\searrow 0,\, x'\to\bar{x}' \\ (x,v)\to(\bar{x},\bar{v}) \text{ in } \operatorname{gph} S}} \Delta_t S(x\,|\,v)(x'), \quad\text{or} \\ \operatorname{gph} D_*S(\bar{x}\,|\,\bar{v}) &= \limsup_{\substack{t\searrow 0 \\ (x,v)\to(\bar{x},\bar{v}) \text{ in } \operatorname{gph} S}} \operatorname{gph} \Delta_t S(x\,|\,v). \end{aligned} \tag{2.4}$$

Both derivative mappings are positively homogeneous; their graphs are closed cones in $I\!R^N \times I\!R^M$. The strict graphical derivative is symmetric in having $-v' \in D_*S(\bar{x}\,|\,\bar{v})(-x')$ when $v' \in D_*S(\bar{x}\,|\,\bar{v})(x')$. The mappings $D[S^{-1}](\bar{v}\,|\,\bar{x})$ and $D_*[S^{-1}](\bar{v}\,|\,\bar{x})$ are the inverses of the mappings $DS(\bar{x}\,|\,\bar{v})$ and $D_*S(\bar{x}\,|\,\bar{v})$.

Differentiability of S at $\bar{x}$ is by definition the case where $S(\bar{x})$ is just $\{\bar{v}\}$ and the "limsup" in (2.3) is a "lim" with $DS(\bar{x}\,|\,\bar{v})$ being a linear mapping. The corresponding version of this in (2.4) is *strict* differentiability. A function (single-valued) is strictly differentiable on an open set if and only if it is $\mathcal{C}^1$ there. Here's a basic rule which takes advantage of that.

Theorem 2.1 (graphical derivative calculus)**.** *Let $S(x) = F(x) + S_0(G(x))$ for $\mathcal{C}^1$ functions $F : I\!R^N \to I\!R^M$, $G : I\!R^N \to I\!R^{N_0}$, and a closed-graph mapping $S_0 : I\!R^{N_0} \rightrightarrows I\!R^M$. Let $\bar{v} \in S(\bar{x})$, so that $\bar{v} - F(\bar{x}) \in S_0(G(\bar{x}))$. Then*

$$\begin{aligned} DS(\bar{x}\,|\,\bar{v})(x') &\subset \nabla F(\bar{x})x' + DS_0(G(\bar{x})\,|\,\bar{v} - F(\bar{x}))(\nabla G(\bar{x})x'), \\ D_*S(\bar{x}\,|\,\bar{v})(x') &\subset \nabla F(\bar{x})x' + D_*S_0(G(\bar{x})\,|\,\bar{v} - F(\bar{x}))(\nabla G(\bar{x})x'), \end{aligned} \tag{2.5}$$

with the inclusions becoming equations when the Jacobian $\nabla G(\bar{x})$ has full rank N_0.

Proof. This rule is an easy consequence of the definitions, but it doesn't seem to have been written up in such generality, although the case when G is the identity mapping is well known [18, 10.43]. We therefore sketch the proof, focusing on the strict graphical derivative as furnishing the general pattern. The difference quotients $\Delta_t F(x)(x') = t^{-1}[F(x+tx') - F(x)]$ and $\Delta_t G(x)(x') = t^{-1}[G(x+tx') - G(x)]$ converge to $\nabla F(\bar{x})\bar{x}'$ and $\nabla G(\bar{x})\bar{x}'$ as $t\searrow 0$, $x \to \bar{x}$ and $x' \to \bar{x}'$. Since $G(x + tx')$ can be written as $G(x) + t\Delta_t G(x)(x')$, and since having $v \in S(x)$ corresponds to having $v - F(x) \in S_0(G(x))$, we get

$$\frac{1}{t}\Big[S_0(G(x + tx')) - (v - F(x))\Big] = \Delta_t S_0\Big(G(x)\,|\,v - F(x)\Big)(\Delta_t G(x)(x')).$$

Therefore

$$\begin{aligned} D_*S(\bar{x}\,|\,\bar{v})(\bar{x}') - \nabla F(\bar{x})\bar{x}' &= \limsup_{\substack{x'\to\bar{x}',\, t\searrow 0 \\ (x,v)\to(\bar{x},\bar{v}) \text{ in } \operatorname{gph} G}} \Delta_t S_0(G(x)\,|\,v - F(x))(\Delta_t G(x)(x')) \\ &\subset \limsup_{\substack{t\searrow 0,\, u'\to\nabla G(\bar{x})\bar{x}' \\ (u,w)\to(G(\bar{x}),\bar{v}-F(\bar{x})) \text{ in } \operatorname{gph} S_0}} \Delta_t S_0(u\,|\,w)(u'). \end{aligned}$$

The rank condition on $\nabla G(\bar{x})$ ensures that everything reachable in the second limsup can be reached in the first limsup. □

Our attention now will be directed at the strict graphical derivative condition

$$D_*S(\bar{x}\,|\,\bar{v})(0) = \{0\}, \text{ or equivalently, } 0 \in D_*[S^{-1}](\bar{v}\,|\,\bar{x})(v') \implies v' = 0. \tag{2.6}$$

In Kummer's inverse function theorem there is a mapping T, which here is S^{-1}, so that $S = T^{-1}$, and $I\!R^M$ is $I\!R^N$. The condition in (2.6) is shown to suffice for T^{-1} to have a single-valued Lipschitz continuous localization at $\bar{v}$ for $\bar{x}$ under a further assumption on T (which Kummer takes already to be single-valued and Lipschitz continuous). That assumption is the inner semicontinuity of T^{-1} at $\bar{x}$ for $\bar{v}$, according to which there exists for every neighborhood $\mathcal{V}$ of $\bar{v}$ a neighborhood $\mathcal{X}$ of $\bar{x}$ such that $T^{-1}(x) \cap \mathcal{V} \neq \emptyset$ for all $x \in \mathcal{X}$. Unfortunately, there is little help from variational analysis for confirming such inner semicontinity for mappings T derived from other mappings in complicated ways, as is typical of the applications where Kummer's theorem might be invoked. So, beyond pure theory, it hasn't found much employment.

In fact, the condition in (2.6) all by itself already provides very important information that has largely been overlooked:

> (2.6) is necessary and sufficient for $(\bar{x}, \bar{v})$ to have a neighborhood $\mathcal{X} \times \mathcal{V}$ such that the mapping s given by $\operatorname{gph} s = [\mathcal{X} \times \mathcal{V}] \cap \operatorname{gph} S$ is single-valued and Lipschitz continuous with respect to $\operatorname{dom} s = \{\, x \in \mathcal{X} \,|\, S(x) \cap \mathcal{V} \neq \emptyset \}$. (2.7)

This was noted in the first part of the proof of [18, Theorem 9.54]. Inner semicontinuity of S at $\bar{x}$ for $\bar{v}$ has the effect of eliminating the possibility of $\operatorname{dom} s$ failing to be a neighborhood of $\bar{x}$ and thus, in combination with (2.6), guarantees through (2.7) that, for $\mathcal{X}$ and $\mathcal{V}$ small enough, s is a single-valued Lipschitz continuous mapping from $\mathcal{X}$ into $\mathcal{V}$. However, inner semicontinuity turns out not to be the only criterion for eliminating empty-valuedness locally.

Definition 2.2 (crypto-continuity of a mapping)**.** *The mapping $S : I\!R^N \rightrightarrows I\!R^M$ will be called crypto-continuous at $\bar{x}$ for $\bar{v} \in S(\bar{x})$ if there is a neighborhood $\mathcal{X} \times \mathcal{V}$ of $(\bar{x}, \bar{v})$ such that the set $C = [\mathcal{X} \times \mathcal{V}] \cap \operatorname{gph} S$ is an N-dimensional continuum in the sense of having a continous parameterization $C = \{\, (x, v) = (f(w), g(w)) \,|\, w \in W \}$ with respect to w in an open subset W of $I\!R^N$ such that the inverse from $(x, v) \in C$ to $w \in W$ is likewise single-valued and continuous.*[5]

In particular S is crypto-continuous at $\bar{x}$ for $\bar{v}$ if it is graphically Lipschitzian there as defined in [18, 9.66]. Maximal monotone mappings, which include the subgradient mappings for closed proper convex functions of that dimension, have that property. But so too, in a local manner, do many subgradient mappings for nonconvex functions, as seen through prox-regularity [14, Theorem 4.7] or variational convexity [15]. Furthermore, crypto-continuity carries forward under many constructions in which there is simply a continuous change or extension of coordinates to transform one graph into another.

Theorem 2.3 (single-valued Lipschitz continuous localizations)**.** *Suppose S is crypto-continuous at $\bar{x}$ for $\bar{v}$. Then (2.6) furnishes a criterion both necessary and sufficient for S to have a single-valued localization at $\bar{x}$ for $\bar{v}$ that is Lipschitz continuous.*

Proof. If there is such a localization, $\operatorname{gph} S$ is in a direct sense graphically Lipschitzian around $(\bar{x}, \bar{v})$, hence crypto-continuous there. Thus, quite apart from (2.6), crypto-continuity is surely always necessary. On the other hand, under (2.6) we have the single-valuedness described in (2.7). Then, with respect to the parameterization in the definition of crypto-continuity, the domain of the mapping

[5]The reason for demanding N-dimensionality is that it's the natural characteristic when S happens to be single-valued, and that's what we are wishing for in a localization.

s in (2.7) is homeomorphic to the open set W in $I\!R^N$. But in that case, by Brouwer's theorem on the invariance of domains (cf. [20]), dom s must be an open set and therefore a neighborhood of $\bar{x}$. □

Corollary 2.4 (new version of Kummer's inverse function theorem)**.** *Let $T : I\!R^N \rightrightarrows I\!R^N$ be crypto-continuous at $\bar{v}$ for $\bar{x}$. Then, in order for T^{-1} to have a single-valued localization at $\bar{x}$ for $\bar{v}$, it is both necessary and sufficient that*

$$0 \in D_*T(\bar{v}\,|\,\bar{x})(v') \quad\Longrightarrow\quad v' = 0. \tag{2.8}$$

Proof. Apply Theorem 2.3 to $S = T^{-1}$, using the fact that crypto-continuity of T at $\bar{v}$ for $\bar{x}$ is equivalent to crypto-continuity of S at $\bar{x}$ for $\bar{v}$. □

Corollary 2.5 (corresponding new implicit function theorem)**.** *Let $R : I\!R^d \times I\!R^n \rightrightarrows I\!R^m$ be a closed-graph mapping, and let*

$$S(p, v) = \{\, x \,|\, R(p, x) \ni v\}. \tag{2.9}$$

Let $\bar{x} \in S(\bar{p}, \bar{v})$ and suppose that R is crypto-continuous at $(\bar{p}, \bar{x})$ for $\bar{v}$. Then for S to have a single-valued localization at $(\bar{p}, \bar{v})$ for $\bar{x}$ it is both necessary and sufficient that

$$0 \in D_*R(\bar{p}, \bar{x}\,|\,\bar{v})(0, x') \quad\Longrightarrow\quad x' = 0. \tag{2.10}$$

Proof. Because $(p, x, v) \in \operatorname{gph} S$ corresponds to $(p, v, x) \in \operatorname{gph} R$, we have

$$(p', x', v') \in \operatorname{gph} D_*R(\bar{p}, \bar{x}\,|\,\bar{v}) \quad\Longleftrightarrow\quad (p', v', x') \in \operatorname{gph} D_*S(\bar{p}, \bar{v}\,|\,\bar{x}).$$

Therefore, the condition in (2.10) translates to $D_*S(\bar{p}, \bar{v}\,|\,\bar{x})(0, 0) = \{0\}$. That corresponds in applying Theorem 2.3 to S to the existence of a single-valued Lipschitz continuous localization, because the crypto-continuity of R at $(\bar{p}, \bar{x})$ for $\bar{v}$ is equivalent to the crypto-continuity of S at $(\bar{p}, \bar{v})$ for $\bar{x}$. □

Corollary 2.6 (application to generalized equations)**.** *For a closed-graph mapping $H : I\!R^n \rightrightarrows I\!R^n$ and a $\mathcal{C}^1$ function $h : I\!R^d \times I\!R^n \to I\!R^n$, let*

$$S(p, v) = \{\, x \,|\, h(p, x) + H(x) \ni v\}. \tag{2.11}$$

Let $\bar{x} \in S(\bar{p}, \bar{v})$, so that $\bar{v} - h(\bar{p}, \bar{x}) \in H(\bar{x})$, and suppose that that H is crypto-continuous at $\bar{x}$ for $\bar{v} - h(\bar{p}, \bar{x})$. Then S has a single-valued Lipschitz continuous localization at $(\bar{p}, \bar{v})$ for $\bar{x}$ if and only if

$$\nabla_x h(\bar{p}, \bar{x})x' + D_*H(\bar{x}\,|\,\bar{v} - h(\bar{p}, \bar{x}))(x') \ni 0 \quad\Longrightarrow\quad x' = 0. \tag{2.12}$$

Proof. This takes $R(p, x) = h(p, x) + H(x)$ in Corollary 2.4. We can view this as

$$R(p, x) = h(p, x) + \bar{H}(p, x) \;\text{ for }\; \bar{H}(p, x) = H(x) \tag{2.13}$$

to recognize that the crypto-continuity of H at $\bar{x}$ for $\bar{v} - h(\bar{p}, \bar{x})$ is inherited by $\bar{H}$ at $(\bar{p}, \bar{x})$, with the dimensionality n in that property becoming $d + n$. The assumptions of Corollary 2.5 are thus satisfied. All that's left is tying the resulting condition in (2.10) to the given structure of R. That can be accomplished by applying to (2.13) the calculus rule in Theorem 2.1 with the mapping G there being the identity. We get

$$D_*R(\bar{p}, \bar{x}\,|\,\bar{v})(p', x') = \nabla_p h(\bar{p}, \bar{x})p' + \nabla_x h(\bar{p}, \bar{x})x' + D_*\bar{H}(\bar{p}, \bar{x}\,|\,\bar{v} - h(\bar{p}, \bar{x}))(p', x')$$

where $D_*\bar{H}(\bar{p}, \bar{x}\,|\,\bar{v} - h(\bar{p}, \bar{x}))(p', x')$ is just $D_*H(\bar{x}\,|\,\bar{v} - h(\bar{p}, \bar{x}))(x')$. That translates the criterion in (2.10) into the one in (2.12). □

Strict graphical derivatives are in the spotlight here because of their enhanced importance in getting single-valued Lipschitz continuous localizations through Theorem 2.3 and its corollaries. However, there is a set-valued kind of Lipschitz-like localization called the Aubin property, which can be explored in comparison, and it is characterized by coderivatives instead of strict graphical derivatives. That property is in principle weaker in lacking a general guarantee of single-valuedness, but in fact single-valuedness is automatic from it in some important circumstances where local maximal monotonicity of a mapping is present. Then coderivatives in effect are already enough to produce a single-valued Lipschitz continuous localization. This is what we take up next.

For a closed-graph mapping $S : I\!\!R^N \rightrightarrows I\!\!R^M$, the coderivative mapping $D^*S(\bar{x}\,|\,\bar{v}) : I\!\!R^M \rightrightarrows I\!\!R^N$ at $\bar{x}$ for $\bar{v}$ is obtained from the normal cone $N_{\mathrm{gph}\,S}(\bar{x},\bar{v})$ in the sense of variational analysis by

$$v' \in D^*S(\bar{x}\,|\,\bar{v})(x') \iff (v',-x') \in N_{\mathrm{gph}\,S}(\bar{x},\bar{v}). \tag{2.14}$$

The passage from $(v',-x')$ to (x',v') may be puzzling, but it has a key motivation. This way, if S happens to be single-valued and $\mathcal{C}^1$, so that graphical derivative $DS(\bar{x}\,|\,\bar{v})$ is the linear mapping $x' \to \nabla S(\bar{x})x'$, the coderivative $D^*S(\bar{x},\bar{v})$ is the adjoint linear mapping $v' \mapsto \nabla S(\bar{x})^*v'$.

The Aubin property of S at $\bar{x}$ for $\bar{v}$ is the existence of neighborhoods $\mathcal{X}$ of $\bar{x}$ and $\mathcal{V}$ of $\bar{v}$ such that

$$\mathrm{dist}(S(x),v) \text{ is a Lipschitz continuous function of } x \in \mathcal{X} \text{ for all } v \in \mathcal{V}. \tag{2.15}$$

It obviously makes a single-valued localization, if available, be Lipschitz continuous, but in general doesn't even imply the existence of a continuous selection $s(x) \in S(x)$ around $\bar{x}$. It's equivalent to the *metric regularity* of S^{-1} at $\bar{v}$ for $\bar{x}$, according to which there are neighborhoods $\mathcal{X}$ of $\bar{x}$ and $\mathcal{V}$ of $\bar{v}$ along with a constant $\kappa > 0$ such that

$$\mathrm{dist}(S^{-1}(v),x) \leq \kappa\,\mathrm{dist}(S(x),v) \text{ when } x \in \mathcal{X},\ v \in \mathcal{V} \tag{2.16}$$

[18, 9.43], this being an estimate of significance in numerical analysis. Moreover it's especially useful because of the Mordukhovich criterion for it in [18, 9.40], namely

$$S \text{ has the Aubin property at } \bar{x} \text{ for } \bar{v} \iff D^*S(\bar{x}\,|\,\bar{v})(0) = \{0\}. \tag{2.17}$$

Recall now that, in the case of $S : I\!\!R^N \to I\!\!R^M$ with $M = N$, S is called *monotone locally at $\bar{x}$ for $\bar{v} \in S(\bar{x})$* if there is a neighborhood $\mathcal{X} \times \mathcal{V}$ such that

$$(x_1 - x_0)\cdot(v_1 - v_0) \geq 0 \text{ when } (x_i,v_i) \in [\mathcal{X} \times \mathcal{V}] \cap \mathrm{gph}\,S, \tag{2.18}$$

and this monotonicity is *maximal* if there is no mapping S' with the same local property such that $[\mathcal{X} \times \mathcal{V}] \cap \mathrm{gph}\,S' \supset [\mathcal{X} \times \mathcal{V}] \cap \mathrm{gph}\,S$ and $[\mathcal{X} \times \mathcal{V}] \cap \mathrm{gph}\,S' \neq [\mathcal{X} \times \mathcal{V}] \cap \mathrm{gph}\,S$. When (2.18) is strengthened to

$$(x_1 - x_0)\cdot(v_1 - v_0) \geq \sigma|x_1 - x_0|^2 \text{ when } (x_i,v_i) \in [\mathcal{X} \times \mathcal{V}] \cap \mathrm{gph}\,S, \text{ with } \sigma > 0, \tag{2.19}$$

the maximal mononicity is *strong* with modulus σ. The subgradient mappings associated with closed proper convex functions are maximal monotone globally, but the local version of the property prevails often for nonconvex functions in connection with them being *variationally convex* locally [15]; strong monotonicity corresponds in this to strong convexity.

The convexity aspects of monotonicity will prominently enter the stability analysis of local optimality in the next section. For now, it's most important that

$$\text{local maximal monotocity of } S \text{ at } \bar{x} \text{ for } \bar{v} \text{ implies crypto-continuity there,} \tag{2.20}$$

since local maximal monotonicity makes the S be graphically Lipschitzian at $\bar{x}$ for $\bar{v}$ through Minty parameterization [18, 12.15]. Such mappings thus furnish prime territory for directly applying Theorem 2.3. It turns out, though, that the strict graphical derivatives in that result aren't fully essential then.

Theorem 2.7 (coderivatives replacing strict graphical derivatives). *For a mapping $S : I\!\!R^N \to I\!\!R^N$ that is maximal monotone locally at $\bar{x}$ for $\bar{v}$, the Aubin property automatically reduces to providing a single-valued Lipschitz continuous localization. In that setting, the Mordukhovich criterion (2.17) is thus equivalent to the strict graphical derivative condition (2.6).*

Proof. The local max monotonicity of S at $\bar{x}$ for $\bar{v}$ passes over to the local max monotonicity of S^{-1} at $\bar{v}$ for $\bar{x}$, and in that transformation the Aubin property of S turns into metric regularity while the single-valued version of it turns into strong metric regularity. It's known from [4, 3G.5] that metric regularity and strong metric regularity are equivalent in the presense of local max monotonicity. Therefore, the Aubin property of S and its single-valued version are equivalent in these circumstances, and the criteria for them come out to mean the same thing. □

The equivalence between the two conditions in Theorem 2.7 is somewhat mysterious, since one refers to a "primal" mapping and the other to a "dual" mapping. It's not easy to see how the graph of one relates in general to the graph of the other; the equivalence certainly doesn't say that they are the same. This is a topic for further research, which might build on relationships betwen coderivatives and ordinary (not strict) graphical derivatives that have been uncovered in [19] and [18, 9.62].

Local maximal monotonicity can, of course, be important in other ways than Theorem 2.7 in connection with getting a single-valued Lipschitz continuous localization, because the crypto-continuity it furnishes can be preserved in mapping constructions that fail to preserve monotonicity itself.

Proposition 2.8 (monotonicity implications for graphical derivatives and coderivatives). *If S is maximal monotone locally at $\bar{x}$ for $\bar{v}$, then*

$$\begin{array}{lcl} v' \in D_*S(\bar{x}\,|\,\bar{v})(x') & \Longrightarrow & v'{\cdot}x' \geq 0, \\ v' \in D^*S(\bar{x}\,|\,\bar{v})(x') & \Longrightarrow & v'{\cdot}x' \geq 0. \end{array} \tag{2.20}$$

If the monotonicity is strong with modulus $\sigma > 0$, then

$$\begin{array}{lcl} v' \in D_*S(\bar{x}\,|\,\bar{v})(x') & \Longrightarrow & v'{\cdot}x' \geq \sigma|x'|^2, \\ v' \in D^*S(\bar{x}\,|\,\bar{v})(x') & \Longrightarrow & v'{\cdot}x' \geq \sigma|x'|^2. \end{array} \tag{2.21}$$

Proof. For D_* these properties are evident from the definitions: having $v' \in \Delta_t S(x\,|\,v)(x')$ means having $v + tv' \in S(x + tx')$ together with $v \in S(x)$, where $t > 0$, and then locally by monotonicity

$$0 \leq ([x + tx'] - x){\cdot}([v + tv'] - v) = t^2(x'{\cdot}v'), \;\text{ hence }\; x'{\cdot}v' \geq 0,$$

and so forth. For D^* the properties aren't so immediate but were established in [13, Theorem 2.1]. □

3 General consequences for stability in optimization

The facts laid out in Section 2 have major implications for the answering the stability questions in Section 1. We return now to that framework of a parameterized family of problems $\mathcal{P}(v, u)$ with their local minimizers x and associated multiplier vectors y, the focus being on problem $\bar{\mathcal{P}} = \mathcal{P}(\bar{v}, \bar{u}) = \mathcal{P}(0, 0)$. We re-examine the primal and primal-dual solution mappings M and $\bar{M}$ in (1.10) and the multiplier mapping Y in (1.5) in the light of strict graphical derivatives of the mappings $\partial_x\varphi$ and $\partial\varphi$ on which they are based. Our assumption (1.6), as spelled out in (1.7) and (1.8), ensures that

$$\begin{array}{l} \text{gph}\, M \text{ is closed relative to a neighborhood of } (\bar{v}, \bar{u}, \bar{x}), \text{ while} \\ \text{gph}\, \bar{M} \text{ is closed relative to a neighborhood of } \{(\bar{v}, \bar{u}, \bar{x})\} \times Y(\bar{x}, \bar{u}, \bar{v}), \end{array} \tag{3.1}$$

because such local closedness holds for the graphs of $\partial_x\varphi$ and $\partial\varphi$. According to (1.9) we have

$$\begin{array}{c}\text{for } (v,u,x) \text{ in some neighborhood of } (\bar v,\bar u,\bar v),\\ x\in M(v,u) \iff \exists\, y \text{ such that } (x,y)\in \bar M(v,u).\end{array} \tag{3.2}$$

Another important fact coming from assumption (1.6) and put at our disposal by [14, Theorem 4.7], is that

$$\begin{array}{l}\text{for all } \bar y\in Y(\bar x,\bar u,\bar v), \text{ the mapping } \partial\varphi \text{ is } n+m \text{ dimensionally graphically}\\ \text{graphically Lipschitzian at } (\bar x,\bar u) \text{ for } (\bar v,\bar y), \text{ hence crypto-continuous there.}\end{array} \tag{3.3}$$

This applies also then to the mapping $\bar M$ at $(\bar v,\bar u)$ for $(\bar x,\bar y)$.

Theorem 3.1 (strict derivative criterion for Lipschitzian localization)**.** *The primal-dual mapping $\bar M$ has a single-valued Lipschitz continuous localization at $(\bar v,\bar u)$ for $(\bar x,\bar y)$ if and only if*

$$(0,y')\in D_*[\partial\varphi](\bar x,\bar u\,|\,\bar v,\bar y)(x',0) \implies (x',y')=(0,0). \tag{3.4}$$

If the primal mapping M is already known to have a single-valued Lipschitz continuous localization at $(\bar v,\bar u)$ for $\bar x$, this criterion simplifies to

$$(0,y')\in D_*[\partial\varphi](\bar x,\bar u\,|\,\bar v,\bar y)(0,0) \implies y'=0, \tag{3.5}$$

which is a condition both necessary and sufficient for the mapping Y to have a localization at $(\bar x,\bar u,\bar v)$ for $\bar y$ that is single-valued and Lipschitz continuous relative to $\operatorname{dom} Y$*.*

Proof. The crypto-continuity of $\bar M$ allows us to apply Theorem 2.3 to get the condition

$$(x',y')\in D_*\bar M(\bar v,\bar u\,|\,\bar x,\bar y)(0,0) \implies (x',y')=(0,0) \tag{3.6}$$

as necessary and sufficient for the localization in question. But this condition is equivalent to (3.4) because the permutation of elements that relates the graph of M to that of $\partial\varphi$ affects the graphical derivatives in the same way; details are laid out in Proposition 3.2 below.

The graph of $D_*\bar M(\bar v,\bar u\,|\,\bar x,\bar y)$ is, by the definition recalled in (2.4), the outer limit of the graphs of the difference quotient mappings $\Delta_t\bar M(v,u\,|\,x,y)$ as $t\searrow 0$ and $(v,u,x,y)\to(\bar v,\bar u,\bar x,\bar y)$, and likewise in parallel for $D_*M(\bar v,\bar u\,|\,\bar x)$ and $\Delta_t M(v,u\,|\,x)$. On the other hand, we know from (3.2) that, locally,

$$(v',u',x')\in \operatorname{gph}\Delta_t M(v,u\,|\,x) \iff \exists\, y',y \text{ such that } (v',u',x',y')\in\operatorname{gph}\Delta_t\bar M(v,u\,|\,x,y).$$

In the limits, therefore,

$$(v',u',x',y')\in\operatorname{gph} D\bar M(\bar v,\bar u\,|\,\bar x,\bar y) \implies (v',u',x')\in\operatorname{gph}\Delta_t M(\bar v,\bar u\,|\,\bar x),$$

so that

$$(x',y')\in D_*\bar M(\bar v,\bar u\,|\,\bar x,\bar y)(0,0) \implies x'\in D_*M(\bar v,\bar u\,|\,\bar x)(0,0). \tag{3.7}$$

Although we can't apply Theorem 2.3 directly to M, out of a lack of assurance about the crypto-continuity of M, we do have from (2.7) that the condition

$$x'\in D_*M(\bar v,\bar u\,|\,\bar x)(0,0) \implies x'=0 \tag{3.8}$$

is necessary and sufficient for M to have a truncation that is single-valued and Lipschitz continuous *with respect to its domain.* Thus, if M is already known to have a single-valued Lipschitz continuous localization, (3.8) would hold. In that case, having $(x',y')\in D_*\bar{M}(\bar{v},\bar{u}\,|\,\bar{x},\bar{y})(0,0)$ would entail $x'=0$ by (3.7), and the criterion in (3.4) would reduce to the one in (3.5), as claimed. The description of (3.5) as a condition on Y is based on the fact recalled in (2.7). Strict graphical derivatives of Y correspond to those of $\partial\varphi$ through the obvious permutations of arguments, so (2.6) for Y comes out as (3.5). □

Proof of Theorem 1.5. This is the case of the simplification in Theorem 3.1 being based on the result in Theorem 1.2. □

In Section 4, we'll be able to bring the strict graphical derivative criteria in Theorems 1.5 and 3.1 down to specifics in terms of a given structure of φ. For the criteria (a)+(b) in Theorem 1.2 that enter Theorem 1.5, some calculus is available, for instance in [11], but it typically only furnishes inclusions to serve as estimates for the coderivatives of $\partial_x\varphi$ rather than exact expressions. However, it will be shown below that coderivatives of $\partial\varphi$ itself can, in part, be used instead. Exact formulas for those are easier to obtain, in the experience of [11], and that will be confirmed in Section 4 as well.

Meanwhile, we record for clear reference some connections that come up in this context which, especially for coderivatives, can sometimes get confusing because of switches in signs.

Proposition 3.2 (derivative and coderivative relations). *Strict graphical derivatives and coderivatives of $\partial\varphi$ are related to those of the primal-dual mapping $\bar{M}$ and its inverse $\bar{M}^{-1}$ by*

$$\begin{aligned}(v',y')\in D_*[\partial\varphi](\bar{x},\bar{u}\,|\,\bar{v},\bar{y})(x',u') &\iff (x',y')\in D_*\bar{M}(\bar{v},\bar{u}\,|\,\bar{x},\bar{y})(v',u')\\ &\iff (v',u')\in D_*\bar{M}^{-1}(\bar{x},\bar{y}\,|\,\bar{v},\bar{u})(x',y'),\\ (v',y')\in D^*[\partial\varphi](\bar{x},\bar{u}\,|\,\bar{v},\bar{y})(x',u') &\iff (-x',y')\in D^*\bar{M}(\bar{v},\bar{u}\,|\,\bar{x},\bar{y})(-v',u')\\ &\iff (v',-u')\in D^*\bar{M}^{-1}(\bar{x},\bar{y}\,|\,\bar{v},\bar{u})(x',-y').\end{aligned}\tag{3.9}$$

Proof. The elements (x,u,v,y) of $\operatorname{gph}\partial\varphi$ permute to the elements (v,u,x,y) of $\operatorname{gph}\bar{M}$, and in the same way the elements (x',u',v',y') of the graph of $D_*[\partial\varphi](\bar{x},\bar{u}\,|\,\bar{v},\bar{y})$ permute to the elements (v',u',x',y') of the graph of $D_*\bar{M}(\bar{v},\bar{u}\,|\,\bar{x},\bar{y})$. That explains the first equivalence in (3.9), from which the second is obvious. The explanation of the coderivative relations is similar but must cope with how signs enter in the definition of coderivatives, seen in (2.14). The elements (x',u',v',y') in the graph of $D^*[\partial\varphi](\bar{x},\bar{u}\,|\,\bar{v},\bar{y})$ are the elements $(v',y',-x',-u')$ in the normal cone to to the graph of $\partial\varphi$ at $(\bar{x},\bar{u},\bar{v},\bar{y})$. Those are the elements $(-x',y',v',-u')$ in the normal cone to the graph of $\bar{M}$ at $(\bar{v},\bar{u},\bar{x},\bar{y})$ and indicate that $(-v',u',-x',y')$ belongs to the graph of $D^*\bar{M}(\bar{v},\bar{u}\,|\,\bar{x},\bar{y})$. In passing to inverses all signs in coderivatives get reversed, so this completes the derivation of (3.9). □

Theorem 1.5, as a product of Theorem 3.1, is definitive in its way, as an answer to what to add to Theorem 1.2 to also handle multiplier vectors. But Theorem 3.1 suggests that in the primal-dual setting there might be a shortcut to full stability that doesn't have to pass through the hard-won coderivative conditions in Theorem 1.2. All that's needed on top of a single-valued Lipschitz continuous localization of the mapping $\bar{M}$ is something to ensure that the x components stay locally optimal.

A route towards that is opened by the recently developed concept of *variational convexity* [15], according to which subgradients and associated function values behave locally in a manner indistinguishable from those coming from a convex function. Under the simplifying assumption of subdifferential continuity in (1.8) that we're operating under, φ is variationally convex at $(\bar{x},\bar{u})$ for $(\bar{v},\bar{y})$ when

$$\begin{aligned}&\exists\text{ neighborhoods } \mathcal{X}\times\mathcal{U} \text{ of } (\bar{x},\bar{u}) \text{ and } \mathcal{V}\times\mathcal{Y} \text{ of } (\bar{v},\bar{y}) \text{ along with}\\ &\text{a closed proper convex function } \psi\le\varphi \text{ on } \mathcal{X}\times\mathcal{U} \text{ such that,}\\ &\text{within those neighborhoods, } \operatorname{gph}\partial\varphi \text{ coincides with } \operatorname{gph}\partial\psi \text{ and,}\\ &\text{for elements } (x,u;v,y) \text{ in the common graph, } \varphi(x,u)=\psi(x,u).\end{aligned}\tag{3.10}$$

More about this can be seen in [15], where *variational strong convexity* is also considered as the case where ψ is not just convex but strongly convex. Strong convexity of ψ on a convex neighborhood $\mathcal{X}\times\mathcal{U}$ has several equivalent descriptions, one of them being that the function

$$\psi(x,u)-\frac{s}{2}|(x,u)-(\bar{x},\bar{u})|^2 \text{ is convex on } \mathcal{X}\times\mathcal{U} \text{ for some } s>0. \tag{3.11}$$

Definition 3.3 (variational sufficiency [16, 17])**.** *The variational sufficient condition for $\bar{x}$ to be a local minimizer in problem $\bar{\mathcal{P}}$ with multiplier $\bar{y}$ satisfying the first-order condition*

$$(\bar{v},\bar{y})\in\partial\varphi(\bar{x},\bar{u}), \text{ where } (\bar{v},\bar{u})=(0,0),$$

is the existence of $r>0$ such that the function

$$\varphi_r(x,u)=\varphi(x,u)+\frac{r}{2}|u|^2, \text{ having } \partial\varphi_r(\bar{x},\bar{u})=\partial\varphi(\bar{x},\bar{u}), \tag{3.12}$$

is variationally convex at $(\bar{x},\bar{u})$ for $(\bar{v},\bar{y})$. The strong variational sufficient condition is the same as this, but asks for variational strong convexity of φ_r.

That the condition in question guarantees local optimality can directly be appreciated from the property that defines of variational convexity in (3.10) in reducing the case of a convex function. But the impact of variational sufficiency on our topic is much bigger than just that.

Theorem 3.4 (parametric local optimality from variational sufficiency)**.** *Under the variational sufficient condition for the local optimality of $\bar{x}$ in $\bar{\mathcal{P}}$ with multiplier $\bar{y}$, there exist neighborhoods $\mathcal{X}\times\mathcal{U}$ of $(\bar{x},\bar{u})$ and $\mathcal{V}\times\mathcal{Y}$ of $(\bar{v},\bar{y})$ such that*

$$\text{for } (v,u)\in\mathcal{V}\times\mathcal{U} \text{ and } (x,y)\in\bar{M}(v,u)\cap[\mathcal{X}\times\mathcal{Y}], \; x \text{ minimizes over } \mathcal{X} \text{ in } \mathcal{P}(v,u), \tag{3.13}$$

and there are no local minimizers over $\mathcal{X}$ in $\mathcal{P}(v,u)$ other than such x. Under the strong variational sufficient condition, x is furthermore the unique minimizer over $\mathcal{X}$ in $\mathcal{P}(u,v)$.

Proof. The variationally convex function φ_r provided by the assumed condition, as in (3.12) with $r>0$, has $\partial\varphi_r(x,u)=\partial\varphi(x,u)+(0,ru)$. Like φ, it can be viewed as furnishing a parameterized family of problems

$$\mathcal{P}_r(v,u) \qquad \text{minimize } \varphi_r(x,u)-v{\cdot}x \text{ with respect to } x,$$

where $\mathcal{P}_r(\bar{v},\bar{u})=\mathcal{P}(\bar{v},\bar{u})$. The associated primal-dual mapping is

$$\bar{M}_r(v,u)=\{\,(x,y_r)\,|\,(v,y_r)\in\partial\varphi_r(x,u)\}, \text{ with } (\bar{x},\bar{y})\in\bar{M}_r(\bar{v},\bar{u}), \tag{3.14}$$

where

$$(x,y)\in\bar{M}_r(v,u) \iff (x,y-ru)\in\bar{M}(v,u). \tag{3.15}$$

Because $\varphi_r(x,u)$, as a function of x, is the same as $\varphi(x,u)$ except for a constant term, local minimizers in $\mathcal{P}_r(v,u)$ are the same as in $\mathcal{P}(v,u)$. Therefore, by confirming the version of (3.13) for $\bar{M}_r(v,u)$ and $\mathcal{P}_r(v,u)$, we can can confirm (3.13) itself; only an adjustment in neighborhoods based on (3.15) separates the two.

The variational convexity of φ_r provides a convex function $\psi\le\varphi_r$ locally such that, for some neighborhoods $\mathcal{X}\times\mathcal{U}$ of $(\bar{x},\bar{u})$ and $\mathcal{V}\times\mathcal{Y}$ of $(\bar{v},\bar{y})$, that may as well be convex, we have

$$\begin{aligned}&\bar{M}_r(v,u)=\{\,(x,y)\,|\,(v,y)\in\partial\psi(x,u)\} \text{ for } (v,u)\in\mathcal{V}\times\mathcal{U},\; (x,y)\in\bar{M}_r(v,u)\cap[\mathcal{X}\times\mathcal{Y}],\\ &\text{and there furthermore } \varphi_r(x,u)=\psi(x,u), \text{ while } \varphi_r\ge\psi \text{ elsewhere on } \mathcal{X}\times\mathcal{U}.\end{aligned} \tag{3.16}$$

The convexity of ψ implies from $(v,y)\in\partial\psi(x,u)$ that

$$\psi(x',u')\geq\psi(x,u)+v\cdot(x'-x)+y\cdot(u'-u)\ \text{ for }\ (x',u')\in\mathcal{X}\times\mathcal{U},$$

and in particular that

$$\psi(x',u)-v\cdot x'\geq\psi(x,u)-v\cdot x\ \text{ for }\ x\in\mathcal{X}. \tag{3.17}$$

Through the relationships in (3.16) we then have the same inequality for φ_r in place of ψ, so that x minimizes over $\mathcal{X}$ in $\mathcal{P}_r(v,u)$. Because ψ is convex (and $\mathcal{X}$ can be taken to be open), there can't be local minimizers within $\mathcal{X}$ other than these.

When ψ is strongly convex, as called for by the strong version of variational sufficiency, there exists $s>0$ such that (3.17) holds with the term $\frac{s}{2}|x'-x|^2$ added on the right. Then there can't be more than one minimizer. □

The strong variational sufficient condition for local optimality has attracted the most interest until now. Its interpretation for special structures of φ has been investigated at length in [17]. For the choice of φ that corresponds to classical nonlinear programming with its canonical perturbations, for example, it corresponds exactly to the standard strong second-order sufficient condition (SSOC). We'll engage in such specialization shortly in Section 4.

Our main result that appeals to variational sufficiency doesn't need the strong version, however. Just the ordinary version provides enough assistance and furthermore facilitates bringing back coderivatives, but this time of $\partial\varphi$ itself instead of those of $\partial_x\varphi$ in Theorems 1.2 and 3.1.

Theorem 3.5 (primal-dual stability under variational sufficiency)**.** *Under the variational sufficient condition for local optimality of $\bar{x}$ in $\bar{\mathcal{P}}$ with multiplier $\bar{y}$, primal-dual full stability holds if and only if the strict graphical derivative criterion in (3.4) is fulfilled. Moreover, that criterion can be replaced by the parallel coderivative criterion*

$$(0,y')\in D^*[\partial\varphi](\bar{x},\bar{u}\,|\,\bar{v},\bar{y})(x',0)\quad\Longrightarrow\quad(x',y')=(0,0). \tag{3.18}$$

Under the strong variational sufficient condition for local optimality, (3.4) can be simplified to (3.5), which can in turn be replaced by

$$(0,y')\in D^*[\partial\varphi](\bar{x},\bar{u}\,|\,\bar{v},\bar{y})(0,0)\quad\Longrightarrow\quad y'=0. \tag{3.19}$$

Proof. With Theorem 1.5 having been derived from Theorem 3.1, we also have its Corollary 1.6 describing primal-dual full stability. Looking at that from the angle of Theorem 3.4, we see the stability can be guaranteed by combining variational sufficiency with the Lipschitz localization criterion (3.4) in Theorem 3.1.

For more insights, fix $r>0$ such that φ_r in (3.12) is variationally convex, as in Definition 3.3, and observe that since $\partial\varphi(x,u)=\partial\varphi(x,u)+(0,ru)$, the derivatives of $\partial\varphi_r$ are related to those of $\partial\varphi$ by

$$D_*[\partial\varphi_r](x,u\,|\,v,y)(x',u')=D_*[\partial\varphi](x,u\,|\,v,y-ru)(x',u')+(0,ru') \tag{3.20}$$

as a case of Theorem 2.1 in which the G mapping is the identity. In particular then, from having $(\bar{v},\bar{u})=(0,0)$, we have

$$\begin{aligned}(0,y')\in D_*[\partial\varphi](\bar{x},\bar{u}\,|\,\bar{v},\bar{y})(x',0)\quad&\Longleftrightarrow\quad(0,y')\in D_*[\partial\varphi_r](\bar{x},\bar{u}\,|\,\bar{v},\bar{y})(x',0)\\&\Longleftrightarrow\quad(x',y')\in D_*\bar{M}_r(\bar{v},\bar{u}\,|\,\bar{x},\bar{y})(0,0)\end{aligned} \tag{3.21}$$

for the mapping $\bar{M}_r$ in (3.14), as an echo of the relations in Proposition 3.2(a). The condition in (3.4) can therefore be expressed equivalently by

$$(x', y') \in D_*\bar{M}_r(\bar{v}, \bar{u} \,|\, \bar{x}, \bar{y})(0, 0) \quad\Longrightarrow\quad (x', y') = (0, 0). \tag{3.22}$$

Note next that the variational convexity of φ_r makes $\partial\varphi_r$ be graphically Lipschitzian around $(\bar{x}, \bar{u}; \bar{v}, \bar{y})$, inasmuch as the graph coincides locally in Definition 3.3 with that of $\partial\psi$ for a convex function ψ. Not just that, $\partial\varphi_r$ is also then maximal monotone locally at $(\bar{x}, \bar{y})$ for $(\bar{v}, \bar{y})$. But that carries over to $\bar{M}_r$ being maximal monotone locally at $(\bar{v}, \bar{u})$ for $(\bar{x}, \bar{y})$. Theorem 2.7 tells us that, with such monotonicity at hand, the strict graphical derivative in the criterion in (3.22) can be replaced by the coderivative. Maneuvering that back to a statement in terms of the mapping $\partial\varphi_r$, we get the criterion in the form

$$(0, y') \in D^*[\partial\varphi_r](\bar{x}, \bar{u} \,|\, \bar{v}, \bar{y})(x', 0) \quad\Longrightarrow\quad (x', y') = (0, 0).$$

That's identical to (3.18) because coderivatives obey the same rule as in (3.20), see [18, 10.43].

In the case of strong variational sufficiency, we have φ_r variationally strongly convex, and that implies the strong monotonicity of a localization of $\partial\varphi_r$ around the point

$$(\bar{x}, \bar{u}; \bar{v}, \bar{y} + r\bar{u}) = (\bar{x}, 0; 0, \bar{y}) = (\bar{x}, \bar{u}; \bar{v}, \bar{y})$$

in its graph. By Proposition 2.8, we have $\sigma > 0$ such that, in general,

$$\begin{aligned} (v', y') \in D_*[\partial\varphi_r](\bar{x}, \bar{u} \,|\, \bar{v}, \bar{y})(x', u') &\quad\Longrightarrow\quad (v', y')\cdot(x', u') \geq \sigma|(x', u')|^2, \\ (v', y') \in D^*[\partial\varphi_r](\bar{x}, \bar{u} \,|\, \bar{v}, \bar{y})(x', u') &\quad\Longrightarrow\quad (v', y')\cdot(x', u') \geq \sigma|(x', u')|^2, \end{aligned}$$

but more particularly when invoking (3.4) or (3.18) get

$$\begin{aligned} (0, y') \in D_*[\partial\varphi_r](\bar{x}, \bar{u} \,|\, \bar{v}, \bar{y})(x', 0) &\quad\Longrightarrow\quad (0, y')\cdot(x', 0) \geq \sigma|(x', 0)|^2 \quad\Longrightarrow\quad x' = 0, \\ (0, y') \in D^*[\partial\varphi_r](\bar{x}, \bar{u} \,|\, \bar{v}, \bar{y})(x', 0) &\quad\Longrightarrow\quad (0, y')\cdot(x', 0) \geq \sigma|(x', 0)|^2 \quad\Longrightarrow\quad x' = 0, \end{aligned}$$

where furthermore

$$\begin{aligned} D_*[\partial\varphi_r](\bar{x}, \bar{u} \,|\, \bar{v}, \bar{y})(x', 0) &= D_*[\partial\varphi](\bar{x}, \bar{u} \,|\, \bar{v}, \bar{y})(x', 0), \\ D^*[\partial\varphi_r](\bar{x}, \bar{u} \,|\, \bar{v}, \bar{y})(x', 0) &= D^*[\partial\varphi](\bar{x}, \bar{u} \,|\, \bar{v}, \bar{y})(x', 0). \end{aligned}$$

Then (3.4) and (3.18) reduce to (3.5) and (3.19), as claimed. □

Although the coderivative versions in Theorem 3.5 are equivalent to the strict graphical derivative versions of the conditions in question, as seen through Theorem 2.7, the calculus that may be applicable to them can be quite different. Which one is more advantageous will likely therefore depend on particular circumstances.

4 Specializing the optimization structure

To get a better feeling for how the general stability results in Section 3 relate to particular situations already familiar from the history of nonlinear programming and its variants, we now specialize to having

$$\begin{aligned} &\varphi(x, u) = f_0(x) + g(F(x) + u) \quad \text{for} \quad F(x) = (f_1(x), \ldots, f_m(x)), \\ &\text{where } g \text{ is closed proper convex on } I\!R^m \text{ and each } f_i \text{ is } \mathcal{C}^2 \text{ on } I\!R^n. \end{aligned} \tag{4.1}$$

The problem at hand is then to

$\bar{\mathcal{P}}_*$ $\qquad$ minimize $f_0(x) + g(F(x))$ with respect to x, where the set of feasible solutions is $\{ x \,|\, F(x) \in \operatorname{dom} g \}$. We are viewing it as embedded in the family of problems

$$\mathcal{P}_*(v,u) \qquad \text{minimize } f_0(x) - v{\cdot}x + g(F(x)+u) \text{ with respect to } x.$$

This long-standing model in composite optimization, which aims at a simple yet amply versatile balance between the smoothness in the functions f_i and "controlled" nonsmoothness induced from g, was dubbed *generalized nonlinear programming* (GNLP) in its role as a platform in [17] for investigating variational sufficient conditions for local optimality.

"Conic programming" is the case where g is the indicator of a closed convex cone K, while classical nonlinear programming is the special case of that cone that makes the g term represent constraints $f_i(x)+u_i \le 0$ or $f_i(x)+u_i = 0$, namely where K is the product of intervals $(-\infty, 0]$ and $[0,0]$. But many "nonconic" formats fit into it as well, and also variants in which the canonical u parameterization can be replaced by a parameterization $f_i(x,p)$ extending the one in (1.2). It's tempting to add a further term in x, such as an indicator δ_X, but we hold off from that because it could get in the way of comparisons that need to be made. It can ultimately be handled anyway by augmenting g and F.

From the structure in (4.1), φ enjoys properties that were foundational in [17]. It is everywhere strongly amenable with

$$\begin{array}{lcl} (v,y) \in \partial\varphi(x,u) & \Longleftrightarrow & y \in \partial g(F(x)+u) \text{ and } \nabla f_0(x) + \nabla F(x)^* y = v, \\ (v,y) \in \partial^\infty\varphi(x,u) & \Longleftrightarrow & y \in \partial g(F(x)+u) \text{ and } \nabla F(x)^* y = v, \end{array} \tag{4.2}$$

where

$$\nabla F(x)^* y = \sum_{i=1}^m y_i \nabla f_i(x) \text{ for } (y_1,\dots,y_m) = y. \tag{4.3}$$

Therefore, in focusing on $(\bar v, \bar y) \in \partial\varphi(\bar x, \bar u)$ at which the basic constraint qualification in (1.3) is satisfied, we are supposing that

$$(y_1,\dots,y_m) \in \partial g(F(\bar x)), \quad \textstyle\sum_{i=1}^m y_i \nabla f_i(x) = 0 \quad \Longrightarrow \quad y_i = 0, \forall i. \tag{4.3}$$

Since strong amenability of φ provides the continuous prox-regularity assumed in (1.6) by [18, 13.32], we are solidly in position to apply the results in Section 3 and see what they reveal.

The relation $y \in \partial(F(x)+u)$ can just as well be written as $F(x)+u \in \partial g^*(y)$ for the closed proper convex function g^* conjugate to g. This provides a big boost because, in terms of the Lagrangian function

$$L(x,y) = f_0(x) + y{\cdot}F(x) = f_0(x) + y_1 f_1(x) + \cdots + y_m f_m(x), \tag{4.4}$$

we have

$$(v,y) \in \partial\varphi(x,u) \quad \Longleftrightarrow \quad \nabla_x L(x,y) = v \text{ with } \nabla_y L(x,y) + u \in \partial g^*(y), \tag{4.5}$$

where

$$\nabla_x L(x,y) = \nabla f_0(x) + \textstyle\sum_{i=1}^m y_i \nabla f_i(x), \qquad \nabla_y L(x,y) = F(x). \tag{4.6}$$

First-order optimality in $\bar{\mathcal{P}}_*$ has the equivalent expressions

$$\begin{array}{lcl} (\bar v, \bar y) \in \partial\varphi(\bar x, \bar u) & \Longleftrightarrow & \nabla_x L(\bar x, \bar y) = 0, \;\; \nabla_y L(\bar x, \bar y) \in \partial g^*(\bar y) \\ & \Longleftrightarrow & \nabla f_0(\bar x) + \sum_{i=1}^m y_i \nabla f_i(\bar x) = 0, \;\; F(\bar x) \in \partial g^*(\bar y) \end{array} \tag{4.7}$$

describing the primal-dual pairs $(\bar x, \bar y)$ in $\bar{\mathcal{P}}^*$. The primal-dual mapping $\bar M$ in (1.10) at the center of our stability study has that set of such pairs as $\bar M(\bar v, \bar u)$, and more generally

$$(x,y) \in \bar M(v,u) \quad \Longleftrightarrow \quad (v,u) \in (\nabla_x L(x,y), -\nabla_y L(x,y)) + (0, \partial g^*(y)). \tag{4.8}$$

The valuable insight is that the inverse mapping $\bar{M}^{-1}$ on the right of (4.8) has highly favorable structure as the sum of a convex-type subdifferential mapping and a $\mathcal{C}^1$ mapping. This leads to the following calculations.

Theorem 4.1 (formulas for derivatives and coderivatives in GNLP)**.** *Strict graphical derivatives of φ are given in the GNLP setting by*

$$\begin{array}{l} (v',y') \in D_*[\partial\varphi](\bar{x},\bar{u}\,|\,\bar{v},\bar{y})(x',u') \\ \iff \left\{ \begin{array}{l} v' = \nabla^2_{xx}L(\bar{x},\bar{y})x' + \nabla F(\bar{x})^* y' \ \textit{ with} \\ y' \in D_*[\partial g](F(\bar{x})\,|\,\bar{y})(\nabla F(\bar{x})x' + u'), \end{array}\right. \end{array} \tag{4.9}$$

whereas the corresponding coderivatives of φ are given by

$$\begin{array}{l} (v',y') \in D^*[\partial\varphi](\bar{x},\bar{u}\,|\,\bar{v},\bar{y})(x',u') \\ \iff \left\{ \begin{array}{l} v' = \nabla^2_{xx}L(\bar{x},\bar{y})x' + \nabla F(\bar{x})^* y' \ \textit{ with} \\ y' \in D^*[\partial g](F(\bar{x})\,|\,\bar{y})(\nabla F(\bar{x})x' + u'), \end{array}\right. \end{array} \tag{4.10}$$

In both cases the conditions entail having

$$x'\cdot v' - x'\cdot\nabla^2_{xx}L(\bar{x},\bar{y})x' = y'\cdot\nabla F(\bar{x})x' \geq -y'\cdot u'. \tag{4.11}$$

Proof. Through Proposition 3.2, our duties are reduced to demonstrating that the expressions at the ends of (4.9) and (4.10) describe the derivatives and coderivatives of $\bar{M}^{-1}$ indicated in (3.9). In (4.8) we have $\bar{M}^{-1}(x,y) = q(x,y) + Q(x,y)$ for the $\mathcal{C}^1$ mapping $q(x,y) = (\nabla_x L(x,y), -\nabla_y L(x,y))$ and a closed-graph mapping $Q(x,y) = (0, \partial g^*(y))$. For that sum we have the elementary rules

$$\begin{array}{l} D_*[q+Q](\bar{x},\bar{y}\,|\,\bar{v},\bar{u}) = \nabla q(\bar{x},\bar{y}) + D_*Q(\bar{x},\bar{y}\,|\,(\bar{v},\bar{u}) - q(\bar{x},\bar{y})), \\ D^*[q+Q](\bar{x},\bar{y}\,|\,\bar{v},\bar{u}) = \nabla q(\bar{x},\bar{y})^* + D^*Q(\bar{x},\bar{y}\,|\,(\bar{v},\bar{u}) - q(\bar{x},\bar{y})), \end{array} \tag{4.12}$$

with the Jacobian matrix being

$$\nabla q(x,y) = \begin{bmatrix} \nabla^2_{xx}L(\bar{x},\bar{y}) & \nabla^2_{xy}L(\bar{x},\bar{y}) \\ -\nabla^2_{yx}L(\bar{x},\bar{y}) & -\nabla^2_{yy}L(\bar{x},\bar{y}) \end{bmatrix}$$

so that

$$\nabla q(x,y) = \begin{bmatrix} \nabla^2_{xx}L(\bar{x},\bar{y}) & \nabla F(\bar{x})^* \\ -\nabla F(\bar{x}) & 0 \end{bmatrix}, \qquad \nabla q(x,y)^* = \begin{bmatrix} \nabla^2_{xx}L(\bar{x},\bar{y}) & -\nabla F(\bar{x})^* \\ \nabla F(\bar{x}) & 0 \end{bmatrix}.$$

Because $(\bar{v},\bar{u}) - q(\bar{x},\bar{y}) = (-\nabla_x L(\bar{x},\bar{y}), F(\bar{x}))$, we have

$$\begin{array}{l} D_*Q(\bar{x},\bar{y}\,|\,(\bar{v},\bar{u}) - q(\bar{x},\bar{y}))(x',y') = D_*[\partial g^*](\bar{y}\,|\,F(\bar{x}))(y'), \\ D^*Q(\bar{x},\bar{y}\,|\,\bar{v},\bar{w}) - q(\bar{x},\bar{y})(x',y') = D^*[\partial g^*](\bar{y}\,|\,F(\bar{x}))(y'). \end{array}$$

Therefore, in proceeding from (4.12), where $q + Q = \bar{M}^{-1}$,

$$\begin{array}{l} (v',u') \in D_*\bar{M}^{-1}(\bar{x},\bar{y}\,|\,\bar{v},\bar{u})(x',y') \\ \qquad \iff \left\{ \begin{array}{l} v' = \nabla^2_{xx}L(\bar{x},\bar{y})x' + \nabla F(\bar{x})^* y', \\ u' \in -\nabla F(\bar{x})x' + D_*[\partial g^*](\bar{y}\,|\,F(\bar{x}))(y'), \end{array}\right. \\ (v',-u') \in D^*\bar{M}^{-1}(\bar{x},\bar{y}\,|\,\bar{v},\bar{y})(x',-y') \\ \qquad \iff \left\{ \begin{array}{l} v' = \nabla^2_{xx}L(\bar{x},\bar{y})x' - \nabla F(\bar{x})^* y', \\ -u' \in \nabla F(\bar{x})x' + D^*[\partial g^*](\bar{y}\,|\,F(\bar{x}))(-y'). \end{array}\right. \end{array} \tag{4.13}$$

We have, on the other hand, since $\partial g^* = (\partial g)^{-1}$,

$$\begin{aligned}
&\nabla F(\bar{x})x' + u' \in D_*[\partial g^*](\bar{y} \,|\, F(\bar{x}))(y') \\
&\qquad \Longleftrightarrow \quad y' \in D_*[\partial g](F(\bar{x}) \,|\, \bar{y})(\nabla F(\bar{x})x' + u'), \\
&\nabla F(\bar{x})x' + u' \in -D^*[\partial g^*](\bar{y} \,|\, F(\bar{x}))(-y') \\
&\qquad \Longleftrightarrow \quad y' \in D^*[\partial g](F(\bar{x}) \,|\, \bar{y})(\nabla F(\bar{x})x' + u').
\end{aligned}$$

In using these relations to transform the conditions on the right in (4.13), we arrive at the expressions on the right in (4.9) and (4.10), as was our goal.

In (4.11), the inequality on the right arises from Proposition 2.8 and the maximal monotonicity of ∂g as the subgradient mapping for a closed proper convex function [18, 12.17]. The equation at the left comes from taking the inner product of the first condition in (4.9) and (4.10) with x'. □

Theorem 4.2 (primal-dual full stability in GNLP)**.** *For φ as in (4.1), the primal-dual mapping $\bar{M}$ in (4.8) has a single-valued Lipschitz continous localization at $(\bar{v}, \bar{u})$ for $(\bar{x}, \bar{y})$ if and only if*

$$\left.\begin{array}{c} \nabla^2_{xx} L(\bar{x}, \bar{y})x' + \nabla F(\bar{x})^* y' = 0 \ \textit{ with} \\ y' \in D_*[\partial g](F(\bar{x}) \,|\, \bar{y}))(\nabla F(\bar{x})x') \end{array}\right\} \implies (x', y') = (0, 0). \tag{4.14}$$

When the variational sufficient condition for local optimality holds at $\bar{x}$ in problem $(\bar{P}_*)$, this strict graphical derivative criterion is both necessary and sufficient for primal-dual full stability in $\bar{\mathcal{P}}_*$. Moreover, it can be replaced by the coderivative criterion

$$\left.\begin{array}{c} \nabla^2_{xx} L(\bar{x}, \bar{y})x' + \nabla F(\bar{x})^* y' = 0 \ \textit{ with} \\ y' \in D^*[\partial g](F(\bar{x}) \,|\, \bar{y})(\nabla F(\bar{x})x') \end{array}\right\} \implies (x', y') = (0, 0). \tag{4.15}$$

With strong variational sufficiency, (4.14) can be simplified to

$$\nabla F(\bar{x})^* y' = 0 \ \textit{ with } \ y' \in D_*[\partial g](F(\bar{x}) \,|\, \bar{y})(0) \implies y' = 0, \tag{4.16}$$

while (4.15) can be simplified to

$$\nabla F(\bar{x})^* y' = 0 \ \textit{ with } \ y' \in D^*[\partial g](F(\bar{x}) \,|\, \bar{y})(0) \implies y' = 0. \tag{4.17}$$

Even without variational sufficiency, (4.16) is the GNLP form for condition (c) of Theorem 1.5. In both (4.14) and (4.15), the assumptions on x' and y' entail having

$$-x' \cdot \nabla^2_{xx} L(\bar{x}, \bar{y})x' = \sum\nolimits_{i=1}^m y_i' [\nabla f_i(\bar{x}) \cdot x'] \geq 0. \tag{4.18}$$

Proof. This specializes Theorem 3.5 through Theorem 4.1, while (4.18) specializes (4.11). □

The coderivative condition in (4.17) has some history behind it. In [11, Theorem 3.3] it was employed as a "second-order qualification" needed to get a chain rule for the coderivatives of a composite function that here corresponds to $g \circ F$. That chain rule was of inclusion type, but a follow-up in [11, Theorem 4.3] with the outer function piecewise linear-quadratic was of equation type under the same "qualification." In Theorem 4.5 the sets of vectors y' in (4.16) and (4.17) will be seen to be identical when g is piecewise linear-quadratic.

How "realistic" is the variational sufficiency assumption in this theorem? Sufficient conditions for local optimality have traditionally been tied to second derivatives of some kind, but this is different. That's the topic in [17], where in fact equivalent conditions of derivative type are developed for various important cases, but perhaps more interestingly, variational sufficiency is characterized by an augmented Lagrangian saddle point property. The efforts in [17] on obtaining second-derivative-type characterizations for variational sufficiency in conic programming center on exploiting that property. Here's the result for nonlinear programming itself.

Example 4.3 (strong variational sufficiency in classical NLP [17]). *In the case where g captures inequality and equation constraints on the functions f_i for $i = 1, \dots, m$, namely*

$$g = \delta_K \text{ for } K = (-\infty, 0]^s \times [0, 0]^{m-s}, \tag{4.19}$$

the strong variational sufficient condition combines the first-order condition (4.7) on $(\bar{x}, \bar{y})$ with the standard strong second-order condition SSOC, namely

$$\begin{array}{l} \nabla^2_{xx} L(\bar{x}, \bar{y}) \text{ is positive-definite relative to the linear subspace} \\ \{ x' \mid \nabla f_i(\bar{x}) \cdot x' = 0 \text{ for } i \in [1, s] \text{ with } \bar{y}_i > 0, \text{ and } i \in [s+1, m] \}. \end{array} \tag{4.20}$$

An analogous condition for strong variational sufficiency is verified in [17] for "second-order cone" programming, and there are examples beyond conic programming there as well. Nonlinear semidefinite programming has been addressed successfully by Wang, Ding, Zhang and Zhao in [21].

Of course, Theorem 4.2 still leaves us with having to figure out the strict graphical derivatives or coderivatives of g, but that's a convex function and quite possibly of a simple kind. Block-separability can help to simplify matters further.

Example 4.4 (derivative and coderivative calculus utilizing separability). *Suppose $u = (u_1, \dots, u_q)$ with $u_i \in I\!R^{m_i}$, and*

$$g(u) = g_1(u_1) + \cdots + g_q(u_q) \text{ on } I\!R^m = I\!R^{m_1} \times \cdots \times I\!R^{m_q}. \tag{4.21}$$

Then

$$\begin{array}{lcl} y' \in D_*[\partial g](u \mid y)(u') & \iff & y'_i \in D_*[\partial g_i](u_i \mid y_i)(u'_i), \ \forall i, \\ y' \in D^*[\partial g](u \mid y)(u') & \iff & y'_i \in D^*[\partial g_i](u_i \mid y_i)(u'_i), \ \forall i. \end{array} \tag{4.22}$$

When the block-separability is full separability, the components u_i in (4.21) being one-dimensional, this rule reduces the challenge of applying the criteria in Theorem 4.2 to that of determining derivative and coderivatives of convex functions of a single real variable, which in many cases is easy.

In overview, the calculus of strict graphical derivatives has received less attention than that of coderivatives and remains underdeveloped as well as more difficult. Coderivatives benefit from coming from normal cones, about which lots is known. Normal cones can be approached by dualizing tangent cones and taking limits, and the tangent cone to $\operatorname{gph} \partial\varphi$ at $(x, u; v, y)$ is the graph of the graphical derivative $D[\partial\varphi](x, u \mid v, y)$. Often, as in the case of fully amenable functions [18, 13.50], nice formulas for such derivatives are available and could be put to use.

Luckily there's a major situation where strict graphical derivatives can be discerned without difficulty. That's when the convex function g is *piecewise linear-quadratic* in the sense that $\operatorname{dom} g$ is the union of finitely many polyhedral convex sets on each of which g can be expressed by a polynomial function of degree at most two. For such g on $I\!R^m$, the subgradient mapping ∂g is piecewise polyhedral in being the union of finitely many polyhedral convex sets of dimension m [18, 10E].

Theorem 4.5 (special features of piecewise linear-quadratic functions). *For a convex piecewise linear-quadratic function g on $I\!R^m$, the strict graphical derivatives of ∂g can be obtained from the ordinary graphical derivatives of ∂g by*

$$\operatorname{gph} D_*[\partial g](u \mid y) = \operatorname{gph} D[\partial g](u \mid y) - \operatorname{gph} D[\partial g](u \mid y). \tag{4.23}$$

Furthermore, at the origin the strict graphical derivatives agree with the coderivatives of ∂g:

$$D_*[\partial g](u \mid y)(0) = D^*[\partial g](u \mid y)(0) = \bigcup \{ S \mid S \in \mathcal{S}(u) \}, \tag{4.24}$$

where $\mathcal{S}(u)$ is a finite collection of subspaces. Specifically, these are the subspaces that serve, in every small enough neighborhood $\mathcal{U}$ of u, as the affine hulls of the convex sets $\partial g(u') - \partial g(u')$ (when nonempty) for $u' \in \mathcal{U}$.

Proof. The graph of $D_*[\partial g](u \,|\, y)$ only depends on local aspects of the graph of ∂g around (u, y). There's no loss of generality in taking $(u, y) = (0, 0)$, in which case, due to $\operatorname{gph} \partial g$ being a union of polyhedral convex sets,

$$\operatorname{gph} \partial g = \operatorname{gph} D[\partial g](0 \,|\, 0) \text{ in a neighborhood of } (0, 0). \tag{4.25}$$

Then (4.23) emerges trivially from the definition of strict graphical derivatives, cf. (2.4).

Moving to the confirmation of (4.24), we first examine the cone there on the left. On the basis of (4.23), we have $(0, y') \in \operatorname{gph} D_*[\partial g](0 \,|\, 0)$ if and only if $(0, y') = (u', y'_+) - (u', y'_-)$ for some u' having both y'_+ and y'_- in $D[\partial g](0 \,|\, 0)(u')$. In other words,

$$D_*[\partial g](0 \,|\, 0)(0) = \bigcup_{u'} \Big[D[\partial g](0 \,|\, 0)(u') - D[\partial g](0 \,|\, 0)(u') \Big].$$

Because we're looking at a cone, we know it includes for each u' the cone generated by the difference set on the right. That cone is the subspace parallel to the affine hull of $D[\partial g](0 \,|\, 0)(u')$. Appealing to (4.25), we see that, locally, the subspace in question is the same as the subspace parallel to the affine hull of $\partial g(u')$. Thus, $D_*[\partial g](0 \,|\, 0)(0)$ is the union of such subspaces for u' in a small enough neighborhood of 0. That's also the description of $D^*[\partial g](0 \,|\, 0)(0)$ furnished by [11, Theorem 4.1] (in taking the mapping h there to be the identity). The equality in (4.24) is therefore correct. ☐

Corollary 4.6 (equivalence of the simplified criteria)**.** *When g is piecewise linear-quadratic, the strict graphical derivative and coderivative simplifications in (4.16) and (4.17) of criteria (4.14) and (4.15) come out to be the identical.*

Example 4.4 and Theorem 4.5 can be combined as follows.

Example 4.7 (full separability with piecewise linear-quadratic functions)**.** *In the fully separable case of Example 4.4, with $m_i = 1$ and $q = m$, suppose each g_i is a piecewise linear-quadratic convex function on $I\!R$. Then*

$$\begin{array}{l} \operatorname{gph} D_*[\partial g](u \,|\, y) = G_1 \times \cdots G_m, \quad \textit{where} \\ G_i = \operatorname{gph} D_*[\partial g_i](u_i \,|\, y_i) = \operatorname{gph} D[\partial g_i](u_i \,|\, y_i) - \operatorname{gph} D[\partial g_i](u_i \,|\, y_i). \end{array} \tag{4.26}$$

The graphs G_i can be understood more explictly through the fact that ∂g_i is a piecewise linear maximal monotone mapping from $I\!R$ to $I\!R$, so $D[\partial g_i](u_i \,|\, y_i)$ has the simple form that its graph is the union of two rays R_i^+ and R_i^- in $I\!R \times I\!R$ with slopes γ_i^+ and γ_i^- in $[0, \infty]$, the first in the northeast quadrant and the second in the southwest quadrant. Then

(a) *if $\gamma_i^- = \gamma_i^+$, G_i is the line with that common slope (vertical if the slope is infinite),*

(b) *if $\gamma_i^- \neq \gamma_i^+$, G_i is the union of the two lines with those different slopes and the wedge $R_i^+ - R_i^-$ and its reflection $R_i^- - R_i^+$, which fill in between those lines towards northeast and southwest.*

The graphs of the coderivative mappings $D^*[\partial g_i](u_i \,|\, y_i)$ in Example 4.7 come out almost the same as the graphs G_i of the mappings $\operatorname{gph} D_*[\partial g_i](u_i \,|\, y_i)$ as described in (a) and (b), except that in (b) only one of the two wedges appears, depending on which of the slopes γ_i^+ or γ_i^- is bigger. But anyway $D^*[\partial g_i](u_i \,|\, y_i)(0) = D_*[\partial g_i](u_i \,|\, y_i)(0)$ here, and this illustrates (4.24).

For the coderivative mapping $D^*[\partial g](u \,|\, y)$ without separability, we aren't aware of a formula having yet been devised in the piecewise linear-quadratic case to compare with the one for the strict graphical derivative mapping in Example 4.4, although some similar close geometric relationship between the graphs is likely. Perhaps a formula could be built on results about normal cones to sets that are unions of finitely many polyhedral convex sets as reviewed and extended by Adam, Červinka and Pištěk [1], since the graph of $D^*[\partial g](u\,|\,y)$ emerges from the normal cone to $\mathrm{gph}\,\partial g$, which is just such a union. Formulas are already available for some subclasses of piecewise linear-quadratic g or g^*, such as the indicator of a polyhedral convex C in [2], the sum of that and a quadratic function in [9], or piecewise linear (polyhedral) function in [10].

Theorem 4.8 (piecewise linear-quadratic criteria for primal-dual stability)**.** *The strict graphical derivative criterion in (4.14) in the case of g being piecewise linear-quadratic can be recast as:*

$$\left.\begin{array}{c} \nabla^2_{xx}L(\bar{x},\bar{y})x' + \nabla F(\bar{x})^*[y'_+ - y'_-] = 0 \\ \text{and } \nabla F(\bar{x})x' = u'_+ - u'_- \text{ with} \\ y'_+ \in D[\partial g](F(\bar{x})\,|\,\bar{y}))(u'_+) \\ y'_- \in D[\partial g](F(\bar{x})\,|\,\bar{y}))(u'_-) \end{array}\right\} \implies x' = 0,\ y'_+ = y'_-,\ u'_+ = u'_-. \tag{4.27}$$

For separable g this specializes in terms of $y' = (y'_1, \ldots, y'_m)$ and the graphs G_i in (4.26) to

$$\left.\begin{array}{l} \nabla^2_{xx}L(\bar{x},\bar{y})x' + \sum_{i=1}^{m} y'_i \nabla f_i(\bar{x}) = 0 \\ \text{with } (\nabla f_i(\bar{x}) \cdot x', y'_i) \in G_i \text{ for all } i \end{array}\right\} \implies x' = 0,\ y'_i = 0, \tag{4.28}$$

with the assumptions on x' and y'_i entailing

$$-x' \cdot \nabla^2_{xx}L(\bar{x},\bar{y})x' = \sum_{i=1}^{m} y'_i[\nabla f_i(\bar{x}) \cdot x'] \ \text{ with } \ y'_i[\nabla f_i(\bar{x}) \cdot x'] \geq 0,\ \forall i. \tag{4.29}$$

The simpler criterion (4.16) specializes to

$$F(\bar{x})^* y' = 0 \ \text{ with } \ y' \in \bigcup \{\, S \,|\, S \in \mathcal{S}(u) \} \implies y' = 0, \tag{4.30}$$

where $\mathcal{S}(u)$ is the finite collection of subspaces described at the end of Theorem 4.5. In the separable case this reduces to saying that

$$\begin{array}{l} \text{the gradients } \nabla f_i(\bar{x}) \text{ are linearly independent} \\ \text{for the set of indices } i \text{ with } f_i(\bar{x}) \notin \mathrm{int\,dom}\, g_i. \end{array} \tag{4.31}$$

Proof. The specialization in (4.27) is directly based on elaborating (4.23) as

$$y' \in D_*[\partial g](u\,|\,y)(u') \iff \left\{ \begin{array}{l} y' = y'_+ - y'_- \text{ and } u' = u'_+ - u'_- \text{ for some} \\ y'_+ \in D[\partial g](u\,|\,y)(u'_+),\ y'_- \in D[\partial g](u\,|\,y)(u'_-). \end{array} \right. \tag{4.34}$$

Likewise for (4.28). In specializing in (4.30) to (4.14), based on (3.5), we take $x' = 0$ in (4.27), which makes $u'_+ = u'_-$. The reductions in (4.30) and (4.31) come from the version of (4.14) in (4.16) and its further simplification in Theorem 4.5. □

Example 4.9 (application to classical NLP)**.** *The case of problem $(\bar{P}_*)$ in which the g term represents constraints $f_i(x) \leq 0$ for $i \in [1,s]$ and $= 0$ for $i \in [s+1,m]$ fits the separability in Theorem 4.8 with*

$$g_i = \delta_{(-\infty,0]} \ \text{ for } i \in [1,s], \qquad g_i = \delta_{[0,0]} \ \text{ for } i \in [s+1,m].$$

The sets $G_i \subset I\!R^2$ in Theorem 4.8 that give the rule in (4.28) specializing the criterion in (4.14) for $\bar{M}$ to have a single-valued Lipschitz continuous localization then have the following description according to whether a constraint is active and the status of its Lagrange multiplier:

(a) $G_i = u_i'$-*axis if* $i \in I_-(\bar{x},\bar{y})$, *meaning* $i \in [1,s]$, $f_i(\bar{x}) < 0$, $\bar{y}_i = 0$.

(b) $G_i = y_i'$-*axis if* $i \in I_+(\bar{x},\bar{y})$, *meaning* $i \in [s+1,m]$ *or* $i \in [1,s]$, $f_i(\bar{x}) = 0$, $\bar{y}_i > 0$.

(c) $G_i = I\!R^2_+ \cup I\!R^2_-$ *if* $i \in I_0(\bar{x},\bar{y})$, *meaning* $i \in [1,s]$, $f_i(\bar{x}) = 0$, $\bar{y}_i = 0$.

Under the strong variational sufficient condition for local optimality in its equivalence to the standard strong second optimality condition SSOC in Example 4.3, Theorem 4.8 allows (4.28) to be replaced by (4.31), which is the linear independence of gradients of the active constraints, LICG.

This recaptures the known result from [3] for classical NLP that primal-dual full stability corresponds to SSOC+LIGC.

Acknowledgement: The work of Matúš Benko was supported by the Austrian Science Fund (FWF) under grant P32832-N as well as by the infrastructure of the Institute of Computational Mathematics, Johannes Kepler University, Linz, Austria.

Data availability: No data was generated or analyzed.

References

[1] Adam, L., Červinka, M., and Pištěk, M., "Normally admissible stratifications and calculation of normal cones to a finite union of polyhedral sets," *Set-valued and Variational Analysis* **24** (2016), 207–229.

[2] Dontchev, A.D., Rockafellar, R.T., "Characterizations of strong regularity for variational inequalities over polyhedral convex sets," *SIAM J. Optimization* **6** (1996), 1087–1105.

[3] Dontchev, A.D., Rockafellar, R.T., "Parametrically robust optimality in nonlinear programming,' *Appl. Comput. Math.* **5** (2006), 32–38.

[4] Dontchev, A.D., Rockafellar, R.T., *Implicit Functions and Solution Mappings,* Springer Verlag, second edition: 2014.

[5] Gfrerer, H., and Outrata, J., "On (local) analysis of multifunctions via subspaces contained in graphs of generalized derivatives," *J. Math. Analysis and Applications* **508** (2022) 125895.

[6] Hang, N.T.V., and Sarabi, E., "Smoothness of subgradient mappings and its applications in parametric optimization" (2023).

[7] Kummer, B., "Lipschitzian inverses, directional derivatives, and applications in $\mathcal{C}^{1,1}$ optimization," *J. Optimization Theory and Applications* **158** (1991), 35–46.

[8] Levy, A.B., Poliquin, R.A., and Rockafellar, R. T., "Stability of locally optimal solutions," *SIAM J. Optimization* **10** (2000), 580–604.

[9] Mordukhovich, B.S., Nghia, T.T.A., and Rockafellar, R.T., "Full stability in finite-dimensional optimization," *Mathematics of Operations Research* **40** (2015), 226–252.

[10] Mordukhovich, B.S., and Sarabi, E., "Variational analysis and full stability of optimal solutions to constrained and minimax problems," *Nonlinear Analysis* **121** (2015), 36-53.

[11] Mordukhovich, B.S., and Rockafellar, R.T., "Second-order subdifferential calculus with applications to tilt stability in optimization," *SIAM J. Optimization* **22** (2012), 953–986.

[12] Mordukhovich, B.S., Rockafellar, R.T., and Sarabi, M.E., "Characterizations of full stability in constrained optimization," *SIAM J. Optimization* **23** (2013), 1810–1849.

[13] Poliquin, R.A., and Rockafellar, R.T., "Tilt stability of a local minimum," *SIAM J. Optimization* **8** (1998), 287–299.

[14] Poliquin, R.A., and Rockafellar, R.T., "Prox-regular functions in variational analysis," *Transactions Amer. Math. Soc.* **348** (1996), 1805–1838.

[15] Rockafellar, R.T., "Variational convexity and the local monotonicity of subgradient mappings," *Vietnam Journal of Mathematics* **47** (2019), 547–561.

[16] Rockafellar, R.T., "Progressive decoupling of linkages in optimization and variational inequalities with elicitable convexity or monotonicity," *Set-Valued and Variational Analysis* **27** (2019), 863–893.

[17] Rockafellar, R.T., "Augmented Lagrangians and hidden convexity in sufficient conditions for local optimality," *Math. Programming* **198** (2023), 159–194.

[18] Rockafellar, R. T., and Wets, R.J-B, *Variational Analysis*, No. 317 in the series *Grundlehren der Mathematischen Wissenschaften*, Springer-Verlag, 1997.

[19] Rockafellar, R.T., and Zagrodny, D., "A derivative-coderivative inclusion in second-order nonsmooth analysis," *Set-Valued Analysis* **5** (1997), 1–17.

[20] Spanier, E.H., *Algebraic Topology*, McGraw Hill, New York, 1966.

[21] Wang, S.W., Ding, C., Zhang, Y.J., and Zhao, X.Y., "Strong variational sufficiency for nonlinear semidefinite programming and its implications," *SIAM Journal on Optimization***33** (2023), 2988–3011.