IEEE CONTROL SYSTEMS LETTERS, VOL. 8, 2024

On the Set of Possible Minimizers of a Sum of Convex Functions

Moslem Zamani, François Glineur, and Julien M. Hendrickx

***Abstract*—Consider a sum of convex functions, where the only information known about each individual summand is the location of a minimizer. In this letter, we give an exact characterization of the set of possible minimizers of the sum. Our results cover several types of assumptions on the summands, such as smoothness or strong convexity. Our main tool is the use of necessary and sufficient conditions for interpolating the considered function classes, which leads to shorter and more direct proofs in comparison with previous work. We also address the setting where each summand minimizer is assumed to lie in a unit ball, and prove a tight bound on the norm of any minimizer of the sum.**

***Index Terms*—Optimization, minimization, convex functions, mathematical programming.**

I. Introduction

Consider the following optimization problem

$$\min_{x\in\mathbb{R}^n} f(x) := \sum_{i\in[m]} f_i(x), \tag{1}$$

with convex functions $f_i : \mathbb{R}^n \to (-\infty, \infty]$, $i \in [m]$. Problem (1) appears naturally in a variety of settings, including distributed optimization; see [1], [2] and references therein. Following [3], [4], [5], we consider the problem of determining the set of possible minimizers of (1) based only on the location of each summand's minimizer $x_i^\star$, that is knowing

$$x_i^\star \in \operatorname*{argmin}_{x\in\mathbb{R}^n} f_i(x), \qquad i \in [m],$$

and some general properties of the functions, such as their smoothness or strong convexity.

Next to its theoretical interest, this natural question is also of practical relevance. As pointed in [5], suppose a model is trained separately on two datasets that should not be shared, which corresponds to minimizing functions f_1 and f_2. Then the knowledge of the set of possible minimizers of the sum $f_1 + f_2$ allows inferring the robustness of the individually trained models, and assessing the relevance of re-training the model based on the combined data. More generally, in the settings of federated learning or decentralized optimization, each node i owns a function f_i from Problem (1) but cannot share it with other nodes, e.g., due to practical or privacy reasons [6], [7]. Determining a priori the set of possible minimizers of Problem (1) is then helpful to compute this minimizer with a distributed optimization algorithm, since it can provide (a) a reasonable starting iterate (using, e.g., some well-centered point in the set), and (b) an upper bound on the distance between this first iterate and the minimizer, which in turn can be used to estimate how many of steps of the algorithm are required, or to tune algorithm parameters. Finally, in some works on resilient distributed optimization (e.g., [8]), algorithms only allow convergence to certain regions such as the convex hull of the minimizers $x_i^\star$. Comparing this zone to the set of possible minimizers is then also of direct interest.

In a recent line of work, Kuwaranancharoen and Sundaram [3], [4], [5] study this problem for $m = 2$ differentiable and strongly convex functions, with the additional knowledge of an upper bound on the norm of the gradient of each summand at the minimizer $x^\star$ (see Proposition (3) for our result in a similar setting). Using a geometric approach, they provide a near-exact characterization of the set of possible minimizers, using *ad hoc* coordinates (more precisely, [5, Th. 6.2] characterizes the boundary of the region of possible minimizers, which coincides with its interior). Assuming that the summands are smooth (i.e., that their gradient is Lipschitz) is also mentioned in [5] as a possibility for further work.

Hendrickx and Rabbat consider a sum of smooth strongly convex functions in the context of open multi-agent systems [9]. They assume all summand minimizers lie in a ball centered at the origin, and derive an upper bond for the norm of a possible minimizer of the sum, expressed in terms of the condition number of the summands [9, Th. 1].

Despite the fundamental nature of the problem, we are not aware of other works characterizing the set of possible minimizers of a sum of convex functions.

In this letter, we show how the use of the interpolation conditions [10], [11] described in Section II, provides a more direct and algebraic approach to tackle this problem. Section III gives an exact characterization of the set of possible minimizers of a sum of two smooth strongly convex functions, and provides an alternative, simpler characterization of set of possible minimizers of a sum of two strongly convex functions

Manuscript received 8 March 2024; revised 10 May 2024; accepted 28 May 2024. Date of publication 13 June 2024; date of current version 10 July 2024. This work was supported in part by the Fonds de la Recherche Scientifique - FNRS and the Fonds Wetenschappelijk Onderzoek - Vlaanderen under EOS Project under Grant 30468160, and in part by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme under Grant 788368. Recommended by Senior Editor S. Dey. *(Corresponding author: Julien M. Hendrickx.)*

Moslem Zamani and Julien M. Hendrickx are with the ICTEAM/INMA, UCLouvain, 1348 Louvain-la-Neuve, Belgium (e-mail: julien.hendrickx@uclouvain.be).

François Glineur is with the ICTEAM/INMA, and the CORE, UCLouvain, 1348 Louvain-la-Neuve, Belgium.

Digital Object Identifier 10.1109/LCSYS.2024.3414378

in the setting of [5]. Section IV extends our results to sums involving arbitrary numbers of functions, and to a variation of the problem where all but one function is smooth. Finally, Section V provides a tight bound for the setting where each summand minimizer lies in a ball.

II. Preliminaries and Interpolation Conditions

Throughout this letter, $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$ denote the Euclidean norm and the dot product, respectively. We also use iff as an abbreviation for if and only if. The set of the first m natural numbers is denoted by $[m]$.

Let $f:\mathbb{R}^n\to(-\infty,\infty]$ be an extended real valued convex function. Function f is proper if its effective domain, defined as $\operatorname{dom} f=\{x: f(x)<\infty\}$, is non-empty. It is closed if its epigraph $\{(x,r): f(x)\le r\}$ is a closed subset of $\mathbb{R}^{n+1}$. The set of subgradients at $x\in\operatorname{dom} f$ (subdifferential) is

$$\partial f(x)=\big\{g: f(y)\ge f(x)+\langle g, y-x\rangle, \forall y\in\mathbb{R}^n\big\}.$$

Let $L\in(0,\infty]$ and $\mu\in[0,\infty)$. A convex function $f:\mathbb{R}^n\to(-\infty,\infty]$ is said to be L-smooth if for any $x_1,x_2\in\mathbb{R}^n$,

$$\|g_1-g_2\|\le L\|x_1-x_2\|\quad \forall g_1\in\partial f(x_1),\ g_2\in\partial f(x_2).$$

If $L<\infty$, then f must be differentiable on $\mathbb{R}^n$. In addition, any convex function is ∞-smooth. Furthermore, a function $f:\mathbb{R}^n\to(-\infty,\infty]$ is said to be μ-strongly convex if the function $x\mapsto f(x)-\frac{\mu}{2}\|x\|^2$ is convex. Note that any convex function is 0-strongly convex. We denote the set of closed proper convex functions that are L-smooth and μ-strongly convex by $\mathcal{F}_{\mu,L}(\mathbb{R}^n)$. In addition, $\mathcal{M}_{\mu,L}(x^\star)$ stands for the set of functions in $\mathcal{F}_{\mu,L}(\mathbb{R}^n)$ for which $x^\star$ is a minimizer. We make the convention $\frac{a}{\infty}=0$ for $a\in\mathbb{R}$.

Interpolation constraints are necessary and sufficient conditions for a finite set of vectors and values to be consistent with an actual function in a given function class. For example, we will heavily rely on the following interpolation result for the class of smooth strongly convex functions.

Theorem 1 (See [10, Th. 4]): Consider the class of functions $\mathcal{F}_{\mu,L}(\mathbb{R}^n)$ with $0\le\mu<L\le\infty$. Given a set of triplets $\{(x_i;g_i;\theta_i)\}_{i\in[m]}$ with $x_i,g_i\in\mathbb{R}^n$ and $\theta_i\in\mathbb{R}$, there exists a function $f\in\mathcal{F}_{\mu,L}(\mathbb{R}^n)$ satisfying

$$f(x_i)=\theta_i,\quad g_i\in\partial f(x_i),\ i\in[m],$$

i.e., a function that interpolates the function and (sub)gradient values as prescribed in $(x_i;g_i;\theta_i)$, if and only if the following set of interpolating conditions is satisfied for $i,j\in[m]$:

$$\begin{aligned}\theta_i-\theta_j-\langle g_j,x_i-x_j\rangle\ge\Big(1-\frac{\mu}{L}\Big)^{-1}\Big(&\frac{\mu}{2}\|x_i-x_j\|^2\\ &+\frac{1}{2L}\|g_i-g_j\|^2-\frac{\mu}{L}\langle g_i-g_j,x_i-x_j\rangle\Big).\end{aligned}\tag{2}$$

Interpolation conditions provide an exact characterization of the triplets $\{(x_i;g_i;\theta_i)\}_{i\in[m]}$ that are consistent with actual functions, and as such, are instrumental in the automated tight analysis of black-box optimization algorithm [10].

Our analysis will rely on the following central Lemma, which provides an interpolation condition for a minimizer $x^\star$ and an arbitrary point x, only involving those two points and their gradients, but not their function values.

Lemma 1: Consider the function class $\mathcal{F}_{\mu,L}(\mathbb{R}^n)$ with $0\le\mu<L\le\infty$. Given two points $x,x^\star\in\mathbb{R}^n$ and a (sub)gradient $g\in\mathbb{R}^n$, there exists a function $f\in\mathcal{M}_{\mu,L}(x^\star)$ with $g\in\partial f(x)$ if and only if

$$\langle g,x-x^\star\rangle\ge\Big(1+\frac{\mu}{L}\Big)^{-1}\Big(\frac{1}{L}\|g\|^2+\mu\|x-x^\star\|^2\Big).\tag{3}$$

Proof: The existence of a function f satisfying the assumptions in the Lemma is equivalent to asking that the two triplets $\{(x;g;f_x),(x^\star;0;f^\star)\}$ are interpolable by a function in $\mathcal{F}_{\mu,L}(\mathbb{R}^n)$ for some values of f_x and $f^\star$. Indeed, we have $f_x=f(x)$, $f^\star=f(x^\star)$ and the optimality condition for a minimizer of a convex function gives $g^\star=0\in\partial f(x^\star)$. Theorem 1 states that it is equivalent to the following pair of conditions (2), first for $(x_i,x_j)=(x,x^\star)$ then for $(x^\star,x)$:

$$\begin{aligned}f_x-f^\star&\ge\Big(1-\frac{\mu}{L}\Big)^{-1}E\big(x,g,x^\star\big)\\ f^\star-f_x-\langle g,x^\star-x\rangle&\ge\Big(1-\frac{\mu}{L}\Big)^{-1}E\big(x,g,x^\star\big)\end{aligned}$$

where $E(x,g,x^\star)=\frac{1}{2L}\|g\|^2+\frac{\mu}{2}\|x-x^\star\|^2-\frac{\mu}{L}\langle g,x-x^\star\rangle$. Hence we need to show that these two conditions holds for some f_x and $f^\star$ if and only if inequality (3) is satisfied.

We first establish the only if part. Summing the above two conditions leads to $-\langle g,x^\star-x\rangle\ge2(1-\frac{\mu}{L})^{-1}E(x,g,x^\star)$ which is equivalent to

$$\Big(1-\frac{\mu}{L}\Big)\langle g,x-x^\star\rangle\ge\frac{1}{L}\|g\|^2+\mu\|x-x^\star\|^2-\frac{2\mu}{L}\langle g,x-x^\star\rangle$$

and can be rearranged as inequality (3) in the Lemma.

To prove the if part, we assume inequality (3) holds and show that the above two conditions are satisfied with the following

$$f_x=\Big(1-\frac{\mu}{L}\Big)^{-1}E\big(x,g,x^\star\big),\quad f^\star=0$$

It is easy to check that this choice of f_x and $f^\star$ leads to an equality in the first condition. Furthermore, as the sum of the two inequalities was shown above to be equivalent to condition (3), we can conclude that the second inequality is satisfied. This concludes the proof of the equivalence. ■

Note that condition (3) can be equivalently rewritten as

$$\Big\|g-\frac{L+\mu}{2}\big(x^\star-x\big)\Big\|\le\frac{L-\mu}{2}\big\|x^\star-x\big\|\tag{4}$$

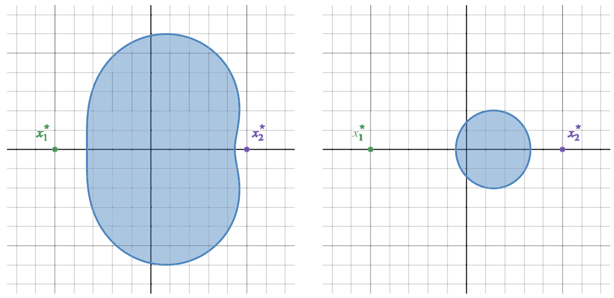
when $L<\infty$, which is easier to interpret geometrically: it states that g must belong to a ball centered at $\frac{L+\mu}{2}(x^\star-x)$ with radius equal to $\frac{L-\mu}{2}\|x^\star-x\|$.

III. Minimizers of a Sum of Two Convex Functions

We first study the case of a sum of two smooth (strongly) convex functions. More precisely, given two points $x_1^\star,x_2^\star$ along with parameters μ_1,μ_2,L_1,L_2, we want to characterize the set of possible minimizers of a sum f_1 and f_2 where $f_1\in\mathcal{M}_{\mu_1,L_1}(x_1^\star)$ and $f_2\in\mathcal{M}_{\mu_2,L_2}(x_2^\star)$. Note that we only have information about the class of functions and their minimizers, but not about the functions themselves.

Proposition 1: Let $0\le\mu_1<L_1<\infty$ and $0\le\mu_2<L_2<\infty$. Given two points $x_1^\star,x_2^\star\in\mathbb{R}^n$, a point $x^\star\in\mathbb{R}^n$ is a minimizer of the sum $f_1(x)+f_2(x)$ for some functions $f_1\in\mathcal{M}_{\mu_1,L_1}(x_1^\star)$ and $f_2\in\mathcal{M}_{\mu_2,L_2}(x_2^\star)$ if and only if

$$\begin{aligned}&\big\|(L_1+\mu_1)\big(x^\star-x_1^\star\big)+(L_2+\mu_2)\big(x^\star-x_2^\star\big)\big\|\\ &\le(L_1-\mu_1)\big\|x^\star-x_1^\star\big\|+(L_2-\mu_2)\big\|x^\star-x_2^\star\big\|.\end{aligned}\tag{5}$$



(a) $\mu_1, \mu_2 = 1, L_1 = 5, L_2 = 15$ (b) $\mu_1 = 1, \mu_2 = 4, L_1, L_2 = 5$

Fig. 1. Sets of possible minimizers $x^\star$ from Proposition 1 with $x_1^\star = (-1, 0)$ and $x_2^\star = (1, 0)$.

Proof: Optimality conditions imply that $x^\star$ is a minimizer of the sum $f_1(x)+f_2(x)$ if and only if there exists subgradients of f_1 and f_2 at $x^\star$ that sum to zero, i.e., if and only if there exists g_1 and g_2 such that

$$g_1 \in \partial f_1\left(x^\star\right), g_2 \in \partial f_2\left(x^\star\right), g_1 + g_2 = 0.$$

Lemma 1 with the geometrical inequality (4) applied to each function f_1 and f_2 implies that $x^\star$ is a possible minimizer iff the following system has a solution,

$$\begin{aligned}&\left\|g_1 - \tfrac{L_1+\mu_1}{2}\left(x^\star - x_1^\star\right)\right\| \leq \tfrac{L_1-\mu_1}{2}\left\|x^\star - x_1^\star\right\|,\\&\left\|g_2 - \tfrac{L_2+\mu_2}{2}\left(x^\star - x_2^\star\right)\right\| \leq \tfrac{L_2-\mu_2}{2}\left\|x^\star - x_2^\star\right\|,\\&g_1 + g_2 = 0.\end{aligned}$$

By replacing g_2 by $-g_1$, this system is equivalent to

$$\begin{aligned}&\left\|g_1 - \tfrac{L_1+\mu_1}{2}\left(x^\star - x_1^\star\right)\right\| \leq \tfrac{L_1-\mu_1}{2}\left\|x^\star - x_1^\star\right\|,\\&\left\|g_1 + \tfrac{L_2+\mu_2}{2}\left(x^\star - x_2^\star\right)\right\| \leq \tfrac{L_2-\mu_2}{2}\left\|x^\star - x_2^\star\right\|.\end{aligned}$$

Geometrically, this system is asking for g_1 to belong to the intersection of two balls. This is possible if and only if the distance between their centers is less than or equal to the sum of their radii, which gives the desired inequality. ∎

One can verify that this set is always bounded provided $\mu_1 > 0$ or $\mu_2 > 0$. Indeed, as the norm of $x^\star$ tends to infinity, $x_1^\star$ and $x_2^\star$ become negligible in (5), leading in the limit to condition $(L_1+\mu_1+L_2+\mu_2)\|x^\star\| \leq (L_1-\mu_1+L_2-\mu_2)\|x^\star\|$ which cannot be satisfied unless $\mu_1+\mu_2 = 0$. On the other hand, when $\mu_1 = \mu_2 = 0$, condition (5) becomes

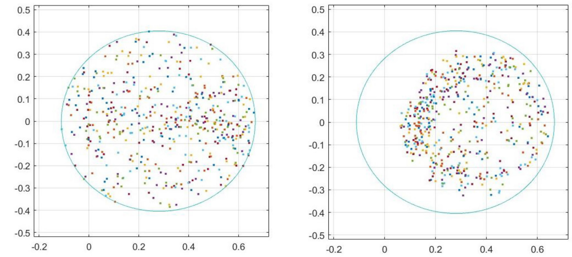
$$\left\|L_1\left(x^\star - x_1^\star\right) + L_2\left(x^\star - x_2^\star\right)\right\| \leq L_1\left\|x^\star - x_1^\star\right\| + L_2\left\|x^\star - x_2^\star\right\|$$

which is the triangle inequality. This implies that any point $x^\star \in \mathbb{R}^n$ can be a minimizer (which can also be derived from the choice $f_1(x) = f_2(x) = 0$). Observe that the following point

$$x^\star = \tfrac{L_1+\mu_1}{L_1+\mu_1+L_2+\mu_2}x_1^\star + \tfrac{L_2+\mu_2}{L_1+\mu_1+L_2+\mu_2}x_2^\star$$

satisfies inequality (5) strictly when $x_1^\star \neq x_2^\star$, from which one can infer that the set of possible minimizers has a nonempty interior. Figure 1 shows the set of possible minimizer $x^\star$ from Proposition 1 for two different sets of parameters in $\mathbb{R}^2$.

To further illustrate this characterization, we perform a numerical experiment where we draw random functions f_1 and f_2 from two specific classes corresponding to the parameters of Fig. 1(b), and compare the location of the minimizers of f_1+f_2 to the set predicted by Proposition 1. We first generate 500 pairs of quadratic functions



(a) Quadratic problems (b) Logistic regression problems

Fig. 2. Illustration of Proposition 1: sets of possible minimizers $x^\star$ (delimited by a cyan curve) along with location of 500 minimizers (colored dots) corresponding to $x_1^\star = (-1, 0)$, $x_2^\star = (1, 0)$, $\mu_1 = 1$, $\mu_2 = 4$, $L_1 = L_2 = 5$ for 500 random quadratic functions (left) and 500 random logistic+quadratic functions (right).

$$q_i(x) = \tfrac{1}{2}\left(x - x_i^\star\right)^T Q_i\left(x - x_i^\star\right) \quad i = 1, 2,$$

(using random symmetric matrices with $I \preceq Q_1 \preceq 5I$ and $4I \preceq Q_2 \preceq 5I$), then we consider logistic + quadratic terms

$$L_j(x) = \sum_{i=1}^{2} \log\left(1 + e^{\langle w_j^i, x\rangle}\right) + \tfrac{\alpha_j}{2}\|x\|^2 + \langle c_j, x\rangle, \quad j = 1, 2,$$

where we generate 500 random instances of parameters $w_j^1, w_j^2, c_j, \alpha_j$ $(j = 1, 2)$ for which $L_1 \in \mathcal{M}_{1,5}(\begin{pmatrix}-1 & 0\end{pmatrix})$ and $L_2 \in \mathcal{M}_{4,5}(\begin{pmatrix}1 & 0\end{pmatrix})$. Figure 2 displays the 500 minimizers of $q_1(x) + q_2(x)$ (left) and $L_1(x) + L_2(x)$ (right), along with the predicted set of possible minimizers.

We observe that all minimizers belong to the predicted set, with quadratic minimizers densely filling the whole set, while logistic minimizers only occupy a smaller subset, probably due to some additional properties that differentiate them from general smooth strongly convex functions.

We now consider a situation where one of the functions is not smooth, assuming without loss of generality $L_2 = \infty$.

Proposition 2: Let $0 \leq \mu_1 < L_1 < \infty$ and $0 \leq \mu_2$. Given two points $x_1^\star, x_2^\star \in \mathbb{R}^n$, a point $x^\star \in \mathbb{R}^n$ is a minimizer of the sum $f_1(x)+f_2(x)$ for some functions $f_1 \in \mathcal{M}_{\mu_1,L_1}(x_1^\star)$ and $f_2 \in \mathcal{M}_{\mu_2,\infty}(x_2^\star)$ if and only if the following holds

$$\begin{aligned}\mu_2\left\|x^\star - x_2^\star\right\|^2 + \tfrac{L_1+\mu_1}{2}\langle x^\star - x_1^\star, x^\star - x_2^\star\rangle \leq&\\ \tfrac{L_1-\mu_1}{2}\left\|x^\star - x_1^\star\right\|\left\|x^\star - x_2^\star\right\|.&\end{aligned} \tag{6}$$

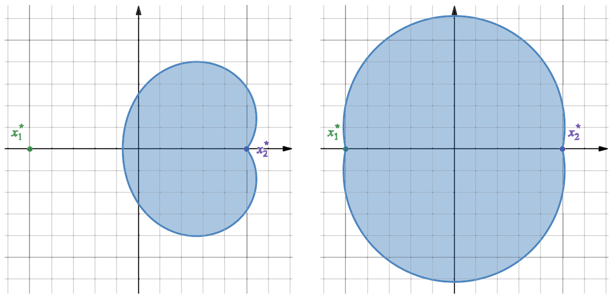
Proof: Due to Lemma 1 and the optimality conditions for the sum $f_1(x)+f_2(x)$, $x^\star$ is a possible minimizer iff the following system has a solution (using both (3) and (4))

$$\begin{aligned}&\left\|g_1 - \tfrac{L_1+\mu_1}{2}\left(x^\star - x_1^\star\right)\right\| \leq \tfrac{L_1-\mu_1}{2}\left\|x^\star - x_1^\star\right\|,\\&\langle g_2, x^\star - x_2^\star\rangle \geq \mu_2\left\|x^\star - x_2^\star\right\|^2,\\&g_1 + g_2 = 0.\end{aligned}$$

Replacing g_2 by $-g_1$, the above system becomes

$$\begin{aligned}&\left\|g_1 - \tfrac{L_1+\mu_1}{2}\left(x^\star - x_1^\star\right)\right\| \leq \tfrac{L_1-\mu_1}{2}\left\|x^\star - x_1^\star\right\|,\\&\langle -g_1, x^\star - x_2^\star\rangle \geq \mu_2\left\|x^\star - x_2^\star\right\|^2.\end{aligned}$$

The desired inequality follows from the necessary and sufficient condition for the intersection between a ball and a half-space to be nonempty. ∎



(a) $\mu_1 = 1, \mu_2 = 3, L_1 = 4$. (b) $\mu_1 = 1, \mu_2 = 0, L_1 = 1.5$.

Fig. 3. Sets of possible minimizers $x^\star$ from Proposition 2 with $x_1^\star = (-1, 0)$ and $x_2^\star = (1, 0)$.

Proposition 2 also corresponds to Proposition 1 in the limit when L_2 tends to infinity; see the discussion before Corollary 1. Figure 3 shows the set of possible minimizer $x^\star$ from Proposition 2 for two different sets of parameters in $\mathbb{R}^2$. As before, this set is bounded when $\mu_1 + \mu_2 > 0$.

When none of the functions is smooth, i.e., when $L_1 = L_2 = \infty$, the situation becomes qualitatively different. Indeed, as shown in [5], the set of possible minimizers becomes unbounded and its closure is equal to the whole $\mathbb{R}^n$ when $n \geq 2$. Hence, similarly to [5], we consider an additional condition on the minimizer: we assume that the subgradients of f_1 and f_2 at the minimizer $x^\star$ have their norm bounded by a constant B (note that since those subgradients sum to zero, they have the same norm).

Since the case $\mu_1 = \mu_2 = 0$ is uninteresting (all points can be minimizers in that case), we assume from here that $\mu_1 + \mu_2 > 0$, i.e., that function f is strongly convex. Using condition (3) from Lemma 1, we have that the subgradient g_1 at a possible minimizer $x^\star$ must satisfy $\|g_1\| \geq \mu_1 \|x_1^\star - x^\star\|$, hence the upper bound B on its norm must satisfy $B \geq \mu_1 \|x_1^\star - x^\star\|$. Similarly, we must have $B \geq \mu_2 \|x_2^\star - x^\star\|$ (also showing that the set of possible minimizers is bounded). Combining with $\|x_1^\star - x^\star\| + \|x_2^\star - x^\star\| \geq \|x_1^\star - x_2^\star\|$ (triangle inequality), these two inequalities imply that

$$B \geq \tfrac{\mu_1 \mu_2}{\mu_1 + \mu_2} \left\|x_1^\star - x_2^\star\right\|. \tag{7}$$

Define the focal point $x_f = \frac{\mu_1}{\mu_1+\mu_2} x_1^\star + \frac{\mu_1}{\mu_1+\mu_2} x_2^\star$, corresponding to the minimizer of the sum of $f_1(x) = \frac{\mu_1}{2}\|x - x_1^\star\|^2 \in \mathcal{M}_{\mu_1,\infty}(x_1^\star)$ and $f_2(x) = \frac{\mu_2}{2}\|x - x_2^\star\|^2 \in \mathcal{M}_{\mu_2,\infty}(x_2^\star)$. Since $\|\nabla f_1(x_f)\| = \frac{\mu_1\mu_2}{\mu_1+\mu_2}\|x_1^\star - x_2^\star\|$, bound (7) is necessary and sufficient for the set of minimizers to be nonempty.

We believe the following characterization is easier to state and slightly more complete than the one proposed in [5].

Proposition 3: Let $\mu_1, \mu_2 \geq 0$ such that $\mu_1 + \mu_2 > 0$. Given two points $x_1^\star \neq x_2^\star \in \mathbb{R}^n$ and a bound B satisfying (7), a point $x^\star \in \mathbb{R}^n$ is a minimizer of the sum $f_1(x) + f_2(x)$ for some functions $f_1 \in \mathcal{M}_{\mu_1,\infty}(x_1^\star)$, $f_2 \in \mathcal{M}_{\mu_2,\infty}(x_2^\star)$ with $\|g\| \leq B$ for some $g \in \partial f_1(x^\star)$ if and only if it satisfies

$$\mu_1 \left\|x^\star - x_1^\star\right\| \leq B \text{ and } \mu_2 \left\|x^\star - x_2^\star\right\| \leq B$$

and at least an additional inequality among the following:

$$(i) \quad \mu_1 \langle x_1^\star - x^\star, x^\star - x_2^\star\rangle \geq \mu_2 \left\|x^\star - x_2^\star\right\|^2$$

$$(ii) \quad \mu_2 \langle x_2^\star - x^\star, x^\star - x_1^\star\rangle \geq \mu_1 \left\|x^\star - x_1^\star\right\|^2$$

$$(iii) \quad \det\Big(M\big(x^\star, x_1^\star, x_2^\star, B^2\big)\Big) \geq 0.$$

where matrix $M(x^\star, x_1^\star, x_2^\star, \alpha)$ be given as follows

$$\begin{pmatrix} 1 & \frac{\langle x^\star - x_1^\star, x^\star - x_2^\star\rangle}{\|x^\star - x_1^\star\|\|x^\star - x_2^\star\|} & \mu_1 \left\|x^\star - x_1^\star\right\| \\ \frac{\langle x^\star - x_1^\star, x^\star - x_2^\star\rangle}{\|x^\star - x_1^\star\|\|x^\star - x_2^\star\|} & 1 & -\mu_2 \left\|x^\star - x_2^\star\right\| \\ \mu_1 \left\|x^\star - x_1^\star\right\| & -\mu_2 \left\|x^\star - x_2^\star\right\| & \alpha \end{pmatrix}.$$

Proof: Using Lemma 1, we see that $x^\star$ is a possible minimizer if and only if the optimal value of the following convex problem is less than or equal to B^2,

$$\begin{aligned} \min \ \|g\|^2 \text{ s.t. } & \langle g, x_1^\star - x^\star\rangle \leq -\mu_1 \left\|x^\star - x_1^\star\right\|^2 \\ & \langle g, x^\star - x_2^\star\rangle \leq -\mu_2 \left\|x^\star - x_2^\star\right\|^2 \end{aligned} \tag{8}$$

It is seen that the problem has a unique solution. Due to the optimality condition, the feasible point g is the optimal solution of problem (8) iff there exist $\lambda_1, \lambda_2 \geq 0$ with

$$\begin{aligned} g &= \lambda_1\big(x^\star - x_1^\star\big) + \lambda_2\big(x_2^\star - x^\star\big), \\ &\lambda_1\Big(\langle g, x_1^\star - x^\star\rangle + \mu_1 \left\|x^\star - x_1^\star\right\|^2\Big) = 0 \\ &\lambda_2\Big(\langle g, x^\star - x_2^\star\rangle + \mu_2 \left\|x^\star - x_2^\star\right\|^2\Big) = 0. \end{aligned}$$

As $x_1^\star \neq x_2^\star$, the optimal value cannot be zero, and at least one constraint is active at the optimal solution. If the first constraint is active, the optimal value is less that or equal to B^2 iff

$$\mu_1 \left\|x^\star - x_1^\star\right\| \leq B, \mu_1 \langle x_1^\star - x^\star, x^\star - x_2^\star\rangle \geq \mu_2 \left\|x^\star - x_2^\star\right\|^2.$$

The case that the second is active is shown analogously. Now, we consider the case that both constraints are active. Without loss of generality, we assume that $x^\star - x_1^\star$ and $x^\star - x_2^\star$ are not parallel. The case that these vectors are parallel is covered by the former cases. Consider the Gram matrix composed of $x^\star - x_1^\star$, $x^\star - x_2^\star$ and g. In this case, problem (8) and the following semi-definite program share the same optimal value,

$$\min \ \alpha \text{ s.t. } M\big(x^\star, x_1^\star, x_2^\star, \alpha\big) \succeq 0. \tag{9}$$

Note that we factor $\|x^\star - x_1^\star\|^2$ and $\|x^\star - x_2^\star\|^2$ in problem (9) as $x^\star \neq x_1^\star, x_2^\star$. The optimal value of semi-definite program (9) is less than or equal to B^2 iff $\det(M(x^\star, x_1^\star, x_2^\star, B^2)) \geq 0$. This follows from the fact that a symmetric matrix is positive definite iff its leading principal minors are positive. To complete the proof, we need to show that if $x^\star - x_1^\star$ and $x^\star - x_2^\star$ are parallel and $\det(M(x^\star, x_1^\star, x_2^\star, B^2) \geq 0$, then $x^\star$ is a possible minimizer. In this case, we have

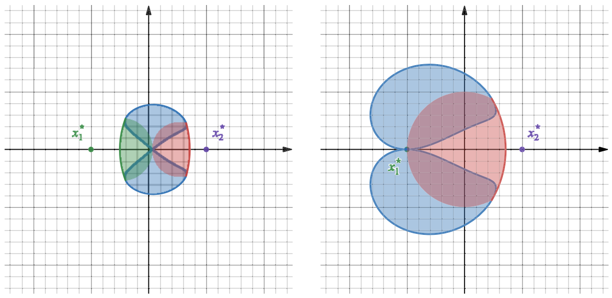
$$\det\big(M\big(x^\star, x_1^\star, x_2^\star, \alpha\big)\big) = -\left\|\mu_1\big(x^\star - x_1^\star\big) + \mu_2\big(x^\star - x_2^\star\big)\right\|^2,$$

which implies $x^\star = x_f$ and the proof is complete. ■

Closedness of the set of possible minimizers under the assumptions of Proposition 3 follows from closedness of the intersection of the feasible region of (8) and $\|g\| \leq B$.

Figure 4 shows the set of possible minimizer $x^\star$ from Proposition 3 for two different sets of parameters in $\mathbb{R}^2$ (colors green and red correspond to the first two inequalities, color blue corresponds to the third determinant condition). Note that a possible minimizer may satisfy more than one system among the three systems given in Proposition 3.

Our last theorem bounds the distance between any minimizer in Proposition 3 and the focal point x_f.



(a) $\mu_1 = 1.75, \mu_2 = 2, B = 3$. (b) $\mu_1 = 1.75, \mu_2 = 0, B = 3$.

Fig. 4. Sets of possible minimizers $x^\star$ from Proposition 3 with $x_1^\star = (-1, 0)$ and $x_2^\star = (1, 0)$.

Theorem 2: Any minimizer $x^\star$ verifying the assumptions of Proposition 3 must satisfy

$$\left\|x^\star - x_f\right\|^2 \le \min\Big(\tfrac{B}{\mu_1+\mu_2}\left\|x_1^\star - x_2^\star\right\| - \tfrac{\mu_1\mu_2}{(\mu_1+\mu_2)^2}\left\|x_1^\star - x_2^\star\right\|^2, \tfrac{B^2}{(\mu_1+\mu_2)^2}\Big).$$

Proof: We may assume w.l.o.g. that $x_1^\star = \frac{\mu_2\|x_1^\star - x_2^\star\|}{\mu_1+\mu_2}e_1$ and $x_2^\star = -\frac{\mu_1\|x_1^\star - x_2^\star\|}{\mu_1+\mu_2}e_1$, which implies that $x_f = 0$ and

$$\mu_1\left\|x_1^\star\right\|^2 + \mu_2\left\|x_2^\star\right\|^2 = \tfrac{\mu_1\mu_2}{\mu_1+\mu_2}\left\|x_1^\star - x_2^\star\right\|^2. \tag{10}$$

First, we establish

$$\left\|x^\star - x_f\right\|^2 \le \tfrac{B}{\mu_1+\mu_2}\left\|x_1^\star - x_2^\star\right\| - \tfrac{\mu_1\mu_2}{(\mu_1+\mu_2)^2}\left\|x_1^\star - x_2^\star\right\|^2.$$

By (3), we have

$$\langle g, x^\star - x_1^\star\rangle \ge \mu_1\left\|x^\star - x_1^\star\right\|^2,$$
$$\langle -g, x^\star - x_2^\star\rangle \ge \mu_2\left\|x^\star - x_2^\star\right\|^2.$$

By summing these inequalities, we get

$$\langle g, x_2^\star - x_1^\star\rangle \ge \mu_1\left\|x_1^\star\right\|^2 + \mu_2\left\|x_2^\star\right\|^2 + (\mu_1+\mu_2)\left\|x^\star\right\|^2.$$

and we conclude with Cauchy–Schwarz inequality and (10)

$$(\mu_1+\mu_2)\left\|x^\star\right\|^2 \le B\left\|x_1^\star - x_2^\star\right\| - \tfrac{\mu_1\mu_2}{\mu_1+\mu_2}\left\|x_1^\star - x_2^\star\right\|^2.$$

Now, we prove that $\|x^\star - x_f\|^2 \le \frac{B^2}{(\mu_1+\mu_2)^2}$. Suppose that $\gamma = \frac{1}{(\mu_1+\mu_2)^2}$. Consider the following inequalities,

$$\gamma\Big(B^2 - \|g\|^2\Big) \ge 0,$$
$$2\mu_1\gamma\Big(\langle g, x^\star - x_1^\star\rangle - \mu_1\left\|x^\star - x_1^\star\right\|^2\Big) \ge 0,$$
$$2\mu_2\gamma\Big(\langle -g, x^\star - x_2^\star\rangle - \mu_2\left\|x^\star - x_2^\star\right\|^2\Big) \ge 0.$$

By summing these inequalities, we get

$$\gamma B^2 - \left\|x^\star\right\|^2 - \gamma\left\|\mu_1 x_1^\star - \mu_2 x_2^\star + (\mu_2 - \mu_1)x^\star + g\right\|^2 \ge 0,$$

which implies the desired inequality. ■

IV. Minimizers of a Sum of m Convex Functions

We now extend the results of Section III to sums of more than two functions, and start with the smooth case.

Theorem 3: Let $0 \le \mu_i < L_i < \infty$ for all $i \in [m]$. Given m points $x_i^\star \in \mathbb{R}^n$, $i \in [m]$, a point $x^\star \in \mathbb{R}^n$ is a minimizer of the sum $\sum_{i=1}^m f_i(x)$ for some functions $f_i \in \mathcal{M}_{\mu_i,L_i}(x_i^\star)$, $i \in [m]$ if and only if

$$\left\|\sum_{i\in[m]}(L_i+\mu_i)\big(x^\star - x_i^\star\big)\right\| \le \sum_{i\in[m]}(L_i-\mu_i)\left\|x^\star - x_i^\star\right\|. \tag{11}$$

Proof: The proof is analogous to that of Proposition 1. Due to the optimality conditions, $x^\star$ is a minimizer of $\sum_{i\in[m]} f_i(x)$ iff there exist $g_i \in \partial f_i(x^\star)$, $i \in [m]$ such that $\sum_{i\in[m]} g_i = 0$. Hence, by Lemma 1, $x^\star \in \operatorname{argmin}\sum_{i\in[m]} f_i$ with $f_i \in \mathcal{M}_{\mu_i,L_i}(x_i^\star)$ for $i \in [m]$ iff there exists $g_i, i \in [m]$ satisfying $\sum_{i\in[m]} g_i = 0$ and

$$\left\|g_i - \tfrac{L_i+\mu_i}{2}\big(x^\star - x_i^\star\big)\right\| \le \tfrac{L_i-\mu_i}{2}\|x^\star - x_i^\star\|,\ i \in [m]$$

Below we use $L_i^+ = \frac{L_i+\mu_i}{2}$ and $L_i^- = \frac{L_i-\mu_i}{2}$ for convenience. Replacing g_m with $-\sum_{i\in[m-1]} g_i$, we obtain

$$\left\|g_i - L_i^+\big(x^\star - x_i^\star\big)\right\| \le L_i^-\left\|x^\star - x_1^\star\right\|,\ i \in [m-1]$$
$$\left\|\sum_{i\in[m-1]} g_i + L_m^+\big(x^\star - x_m^\star\big)\right\| \le L_m^-\left\|x^\star - x_m^\star\right\|.$$

We check the consistency of this system by an elimination method. We first investigate the solvability with respect to g_{m-1} while $g_1, \ldots, g_{m-2}$ are fixed. As g_{m-1} only appears in the following two inequalities, defining two balls

$$\left\|g_{m-1} - L_{m-1}^+\big(x^\star - x_{m-1}^\star\big)\right\| \le L_{m-1}^-\left\|x^\star - x_{m-1}^\star\right\|$$
$$\left\|g_{m-1} + \sum_{i\in[m-2]} g_i + L_m^+\big(x^\star - x_m^\star\big)\right\| \le L_m^-\left\|x^\star - x_m^\star\right\|,$$

the system has a solution in terms of g_{m-1} iff distance between the two ball centers is not greater than the sum of the two radii, leading to the inequality

$$\left\|\sum_{i\in[m-2]} g_i + L_m^+\big(x^\star - x_m^\star\big) + L_{m-1}^+\big(x^\star - x_{m-1}^\star\big)\right\|$$
$$\le\ L_m^-\left\|x^\star - x_m^\star\right\| + L_{m-1}^-\left\|x^\star - x_{m-1}^\star\right\|.$$

Consider now at g_{m-2}, which must again lie in the intersection of two balls: applying the same principle leads to an inequality of the same form as above, with one more term with L_{m-2}^+ inside of the left-hand side norm, and one more norm term with L_{m-2}^- on the right-hand side. A simple induction argument then completes the proof. ■

The set of possible minimizers in Theorem 3 is closed and bounded set if $\sum_{i\in[m]}\mu_i > 0$. If $x_i^\star$'s are distinct, inequality (11) is satisfied strictly for the focal point $x_f = \frac{1}{\sum_{i\in[m]} L_i+\mu_i}\sum_{i\in[m]}(L_i+\mu_i)x_i^\star$, and the interior of the minimizer set is non-empty.

We now consider the case where (exactly) one summand is not smooth, using a limiting argument. Rewriting (11) as

$$\frac{1}{2}\sum_{i\in[m]}\mu_i L_i\left\|x^\star - x_i^\star\right\|^2 + \sum_{i<j} L_i^+ L_j^+\langle x^\star - x_i^\star, x^\star - x_j^\star\rangle$$

$$\leq \sum_{i<j,i,j\in[m]} L_i^- L_j^- \left\|x^\star - x_i^\star\right\| \left\|x^\star - x_j^\star\right\| \tag{12}$$

dividing both sides by L_m and taking the limit as $L_m \to \infty$, we obtain the inequality of the following Corollary.

Corollary 1: Let $0 \leq \mu_i < L_i < \infty$ for all $i \in [m-1]$ and $0 \leq \mu_m < L_m = \infty$. Given m points $x_i^\star \in \mathbb{R}^n$, a point $x^\star \in \mathbb{R}^n$ is a minimizer of the sum $\sum_{i=1}^m f_i(x)$ for some functions $f_i \in \mathcal{M}_{\mu_i,L_i}(x_i^\star)$, $i \in [m]$ if and only if

$$\sum_{i\in[m-1]} \tfrac{L_i+\mu_i}{2}\left\langle x^\star - x_i^\star, x^\star - x_m^\star\right\rangle + \mu_m\left\|x^\star - x_m^\star\right\|^2$$
$$\leq \sum_{i\in[m-1]} \tfrac{L_i-\mu_i}{2}\left\|x^\star - x_i^\star\right\|\left\|x^\star - x_m^\star\right\|. \tag{13}$$

Observe that if more than one summand is not smooth, the set of possible minimizers may become very large set, as its closure may be whole space [5].

V. Bounds When Summand Minimizers Lie in a Ball

Hendrickx and Rabbat [9] consider a sum of smooth strongly convex functions f_i that all belong to the same class $\mathcal{F}_{\mu,L}(\mathbb{R}^n)$, and all have minimizers $x_i^\star$ liying in the unit ball centered at the origin, i.e., $\|x_i^\star\| \leq 1$ for all $i \in [m]$. They prove that any minimizer $x^\star$ satisfies $\|x^\star\| \leq 1 + \sqrt{\kappa}$ where $\kappa = \frac{L}{\mu}$, which we now improve below.

Theorem 4: Let $f_i \in \mathcal{M}_{\mu,L}(x_i^\star)$ with $0 < \mu < L < \infty$, $i \in [m]$, and $\|x_i^\star\| \leq 1$. If $x^\star \in \operatorname{argmin} \sum_{i=1}^m f_i(x)$ then

$$\left\|x^\star\right\| \leq \tfrac{1}{2}\left(\sqrt{\kappa} + \tfrac{1}{\sqrt{\kappa}}\right), \quad \text{where } \kappa = \tfrac{L}{\mu}.$$

Proof: Consider inequality (12). Since $L_i = L$ and $\mu_i = \mu$ for $i \in [m]$, $L_i^+ = \frac{L+\mu}{2}$ and $L_i^- = \frac{L-\mu}{2}$. By dividing both sides of the inequality by $\frac{\mu^2}{4}$, we get

$$2\kappa \sum_{i\in[m]} \left\|x^\star - x_i^\star\right\|^2 + (\kappa+1)^2 \sum_{i<j}\left\langle x^\star - x_i^\star, x^\star - x_j^\star\right\rangle$$
$$\leq (\kappa-1)^2 \sum_{i<j,i,j\in[m]} \left\|x^\star - x_i^\star\right\|\left\|x^\star - x_j^\star\right\|.$$

Using $2\|x^\star - x_i^\star\|\|x^\star - x_j^\star\| \leq \|x^\star - x_i^\star\|^2 + \|x^\star - x_j^\star\|^2$ and further algebraic manipulations, we rewrite this inequality as

$$\kappa\left(\left(\sqrt{\kappa} + \tfrac{1}{\sqrt{\kappa}}\right)^2 - m\left(\sqrt{\kappa} - \tfrac{1}{\sqrt{\kappa}}\right)^2\right)\sum_{i\in[m]}\left\|x^\star - x_i^\star\right\|^2$$
$$+2\kappa\left(\sqrt{\kappa} + \tfrac{1}{\sqrt{\kappa}}\right)^2 \sum_{i<j,i,j\in[m]}\left\langle x^\star - x_i^\star, x^\star - x_j^\star\right\rangle \leq 0. \tag{14}$$

Squaring inequality $\|(x_i^\star - x^\star) + x^\star\| = \|x_i^\star\| \leq 1$, we find

$$\left\|x^\star - x_i^\star\right\|^2 + \left\|x^\star\right\|^2 + 2\left\langle x^\star - x_i^\star, x^\star\right\rangle \leq 1 \text{ for } i \in [m]$$

We now sum these m inequalities multiplied by $\frac{1}{4m}(\sqrt{\kappa} + \frac{1}{\sqrt{\kappa}})^2$, and add inequality (14) multiplied by $\alpha = (\frac{1}{2m\sqrt{\kappa}(\sqrt{\kappa}-\frac{1}{\sqrt{\kappa}})}(\sqrt{\kappa} + \frac{1}{\sqrt{\kappa}}))^2$, and simplifying the resulting expression leads to inequality $\|x^\star\|^2 + \|\frac{1}{2}(\sqrt{\kappa} - \frac{1}{\sqrt{\kappa}})x^\star + \sqrt{\alpha}(\kappa+1) + \sum_{i\in[m]}(x^\star - x_i^\star)\|^2 \leq \frac{1}{4}(\sqrt{\kappa} + \frac{1}{\sqrt{\kappa}})^2$, which concludes the proof. ■

We now address the setting where one of the functions (w.l.o.g. f_m) is an arbitrary closed proper convex function.

Theorem 5: Given $0 < \mu < L < \infty$, let $f_i \in \mathcal{M}_{\mu,L}(x_i^\star)$, $i \in [m-1]$, and let $f_m \in \mathcal{M}_{0,\infty}(x_m^\star)$. Assume that $\|x_i^\star\| \leq 1$ for $i \in [m]$. If $x^\star \in \operatorname{argmin} \sum_{i=1}^m f_i(x)$ then

$$\left\|x^\star\right\| \leq \sqrt{\kappa+1}, \quad \text{where } \kappa = \tfrac{L}{\mu}.$$

Proof: The proof is analogous to that of Theorem 4. By dividing both sides of inequality (13) by $\frac{\mu}{2}$, we get

$$(\kappa+1)\sum_{i\in[m-1]}\left\langle x^\star - x_i^\star, x^\star - x_m^\star\right\rangle - \tfrac{1}{2}(\kappa-1)$$
$$\sum_{i\in[m-1]}\left(\left\|x^\star - x_i^\star\right\|^2 + \left\|x^\star - x_m^\star\right\|^2\right) \leq 0. \tag{15}$$

By the assumptions, we have for $i \in [m-1]$

$$\left\|x_i^\star\right\|^2 = \left\|x^\star - x_i^\star\right\|^2 + \left\|x^\star\right\|^2 + 2\left\langle x^\star - x_i^\star, x^\star\right\rangle \leq 1, \tag{16}$$

and similarly

$$\left\|x^\star - x_m^\star\right\|^2 + \left\|x^\star\right\|^2 + 2\left\langle x^\star - x_m^\star, x^\star\right\rangle \leq 1. \tag{17}$$

Multiplying inequalities (15), (16) and (17) by $\frac{1}{2m-2}(1+\frac{1}{\kappa})$, $\frac{1}{2m-2}(\kappa+1)$ and $\frac{1}{2}(\kappa+1)$, respectively and summing them leads to $\|x^\star\|^2 + \frac{1}{4\kappa(m-1)}\sum_{i\in[m-1]}\|2\kappa x^\star + (\kappa+1)(x^\star - x_m^\star) + (\kappa+1)(x^\star - x_i^\star)\|^2 \leq \kappa+1$. which implies the desired bound. ■

VI. Conclusion

This letter has demonstrated how the use of interpolation constraints simplifies the analysis of questions about convex functions and their minimizers. This framework can in principle allow for even more accurate descriptions of those sets of possible minimizers if additional information on the summands is provided, such as the bounds or exact values of the summands or their gradient at some remarkable points.

References

[1] A. Nedić and J. Liu, "Distributed optimization for control," *Annu. Rev. Control, Robot., Auton. Syst.*, vol. 1, pp. 77–103, May 2018.

[2] T. Yang et al., "A survey of distributed optimization," *Annu. Rev. Control*, vol. 47, pp. 278–305, Jun. 2019.

[3] K. Kuwaranancharoen and S. Sundaram, "On the location of the minimizer of the sum of two strongly convex functions," in *Proc. IEEE Conf. Decision Control (CDC)*, 2018, pp. 1769–1774.

[4] K. Kuwaranancharoen and S. Sundaram, "On the set of possible minimizers of a sum of known and unknown functions," in *Proc. Am. Control Conf. (ACC)*, 2020, pp. 106–111.

[5] K. Kuwaranancharoen and S. Sundaram, "The Minimizer of the sum of two strongly convex functions," 2023, *arXiv:2305.13134*.

[6] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," 2016, *arXiv:1610.05492*.

[7] M. Pirani, A. Mitra, and S. Sundaram, "Graph-theoretic approaches for analyzing the resilience of distributed control systems: A tutorial and survey," *Automatica*, vol. 157, Nov. 2023, Art. no. 111264.

[8] S. Sundaram and B. Gharesifard, "Distributed optimization under adversarial nodes," *IEEE Trans. Autom. Control*, vol. 64, no. 3, pp. 1063–1076, Mar. 2019.

[9] J. M. Hendrickx and M. G. Rabbat, "Stability of decentralized gradient descent in open multi-agent systems," in *Proc. 59th IEEE Conf. Decision Control (CDC)*, 2020, pp. 4885–4890.

[10] A. B. Taylor, J. M. Hendrickx, and F. Glineur, "Smooth strongly convex interpolation and exact worst-case performance of first-order methods," *Math. Program.*, vol. 161, pp. 307–345, Jan. 2017.

[11] A. Rubbens, N. Bousselmi, S. Colla, and J. M. Hendrickx, "Interpolation constraints for computing worst-case bounds in performance estimation problems," in *Proc. 62nd IEEE Conf. Decision Control (CDC)*, 2023, pp. 3015–3022.