▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

# A Proximal Bundle Variant with Optimal Iteration-Complexity for a Large Range of Prox Stepsizes

OP21 – MS56 Recent Developments in First-Order Methods for Composite Optimization

# Jiaming Liang

School of Industrial and Systems Engineering Georgia Institute of Technology

Joint work with Renato Monteiro

July 21, 2021

- - Assumptions
  - Literature review

#### The RPB method 2

- Review of the composite subgradient method
- Bundle method
- RPB
- CS as an instance of RPB

# 3 Main results

- Upper complexity bounds
- Complexity bounds for another proximal bundle variant

# 4 Optimal complexity

- Lower complexity bounds
- Optimal complexity bounds

| Introduction | The RPB method | Main results | Optimal complexity | Conclusion |
|--------------|----------------|--------------|--------------------|------------|
|              |                |              |                    |            |
|              |                |              |                    |            |

# Introduction

Main problem:

$$\phi_* := \min \left\{ \phi(x) := f(x) + h(x) : x \in \mathbb{R}^n \right\}$$
(1)

▲□▶ ▲□▶ ▲□▶ ▲□▶ ▲□ ● ● ●

# Main goal:

To show the iteration-complexity of the relaxed proximal bundle (RPB) method is optimal.

# Main techniques:

- Inexact proximal point framework
- Bundle method

| Introduction   | The RPB method   | Main results            | Optimal complexity | Conclusio |
|--|--|-------------------------|--------------------|-----------|
| Assumptions  | oduction<br>Assumptions<br>iterature review  |                         |                    |           |
| <ul> <li>2 The</li> <li>• F</li> <li>• F</li> <li>• C</li> </ul> | e RPB method<br>Review of the compos<br>Bundle method<br>RPB<br>CS as an instance of F | ite subgradient<br>RPB  | method             |           |
| (3) Mai<br>● L<br>● C  | in results<br>Ipper complexity bou<br>Complexity bounds for                            | nds<br>r another proxir | nal bundle variant |           |
| ④ Opt<br>● L<br>● C  | i <mark>mal complexity</mark><br>ower complexity bou<br>Optimal complexity bo          | nds<br>ounds            |                    |           |
| 5 Cor  | clusion  |                         |                    |           |

◆□▶ ◆□▶ ◆三▶ ◆三▶ ◆□▶

| Introduction | The RPB method | Main results | Optimal complexity | Conclusion |
|--------------|----------------|--------------|--------------------|------------|
| Assumptions  |                |              |                    |            |
|              |                |              |                    |            |

## Convex nonsmooth problem

Consider (1), where

- (A1) functions  $f, h \in \overline{\text{Conv}}(\mathbb{R}^n)$  are such that  $\operatorname{dom} h \subset \operatorname{dom} f$  and function  $f' : \operatorname{dom} h \to \mathbb{R}^n$  is such that  $f'(x) \in \partial f(x)$  for all  $x \in \operatorname{dom} h$ ;
- (A2) the set of optimal solutions  $X^*$  of problem (1) is nonempty;
- (A3) *h* is  $\mu$ -convex and  $||f'(x)|| \le M_f$  for all  $x \in \text{dom } h$ ;
- (A4) h is  $M_h$ -Lipschitz continuous on dom h, i.e.,

$$|h(u) - h(v)| \le M_h ||u - v|| \quad \forall u, v \in \operatorname{dom} h.$$

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ● ●

| Introduction   | The RPB method   | Main results            | Optimal complexity | Conclusio |
|--|--|-------------------------|--------------------|-----------|
| Literature reviev                                      | v  |                         |                    |           |
| 1 In<br>•<br>•   | troduction<br>Assumptions<br>Literature review   |                         |                    |           |
| (2) TH<br>•<br>•<br>•                                  | ne RPB method<br>Review of the compose<br>Bundle method<br>RPB<br>CS as an instance of 1 | site subgradient<br>RPB | : method           |           |
| 3 M<br>•   | ain results<br>Upper complexity bou<br>Complexity bounds fo                              | inds<br>r another proxi | mal bundle variant |           |
| <ul> <li>4</li> <li>0</li> <li>•</li> <li>•</li> </ul> | ptimal complexity<br>Lower complexity bou<br>Optimal complexity b                        | nds<br>ounds            |                    |           |
| <b>5</b> Co  | onclusion  |                         |                    |           |

◆□ ▶ ◆□ ▶ ◆ 臣 ▶ ◆ 臣 ▶ ○ 臣 ○ のへぐ

| Introduction      | The RPB method | Main results | Optimal complexity | Conclusion |
|-------------------|----------------|--------------|--------------------|------------|
| Literature review |                |              |                    |            |
|                   |                |              |                    |            |

# Lower complexity bound results

• Convex, unconstrained,  $\min f(x)$ 

$$\Omega\left(\frac{M_f^2 d_0^2}{\bar{\varepsilon}^2}\right)$$

where  $d_0 := \inf\{\|x_0 - x_*\| : x_* \in X_*\}$  and  $\bar{\varepsilon}$  is the tolerance.

• Strongly convex, unconstrained,  $\min f(x)$ 

$$\Omega\left(\frac{M_f^2}{\mu\bar{\varepsilon}}\right)$$

where  $\mu$  is the strong convexity of f.

**Drawback**: bounds are inconsistent when  $\mu \rightarrow 0$ 

# Upper bound complexity results

- Subgradient, Mirror descent and Bundle-level method are optimal.
- Bundle method
  - convex, Kiwiel 2000

$$\mathcal{O}_1\left(rac{ ilde{M}^2 ilde{D}^4}{\lambdaar{arepsilon}^3}
ight)$$

where 
$$\tilde{D} = \tilde{D}[\tilde{f}] := \sup\{d(x_j, X^*) : j \ge 0\},\$$
  
 $\tilde{M} = \tilde{M}[\tilde{f}] := \sup\{\|\tilde{f}'(x_j)\| : j \ge 0\}.$ 

• *µ*-strongly convex, Du and Ruszczyński 2017

$$\tilde{\mathcal{O}}_1\left(\frac{\tilde{M}^2\lambda}{\alpha^2\bar{\varepsilon}}\right)$$

where  $\alpha := \min\{\lambda \mu, 1\}.$ 

**Drawback**: bounds are not optimal in general (i.e., for a large range of prox stepsizes  $\lambda$ ) 

| Introduction       | The RPB method   | Main results         | Optimal complexity | Conclusion |
|--------------------|--|----------------------|--------------------|------------|
| Review of the c    | omposite subgradient method  |                      |                    |            |
| 1 In<br>•<br>•     | troduction<br>Assumptions<br>Literature review   |                      |                    |            |
| 2 T<br>•<br>•<br>• | he RPB method<br>Review of the composi<br>Bundle method<br>RPB<br>CS as an instance of F | te subgradient<br>PB | method             |            |
| 3 M                | ain results  |                      |                    |            |

- Upper complexity bounds
- Complexity bounds for another proximal bundle variant

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ 三 のへぐ

# Optimal complexity

- Lower complexity bounds
- Optimal complexity bounds
- 5 Conclusion

| Introduction          | The RPB method        | Main results | Optimal complexity | Conclusion |
|-----------------------|-----------------------|--------------|--------------------|------------|
| Review of the composi | te subgradient method |              |                    |            |
|                       |                       |              |                    |            |

### Composite subgradient (CS) method

$$x_{j} = \operatorname*{argmin}_{u \in \mathbb{R}^{n}} \left\{ f(x_{j-1}) + \langle f'(x_{j-1}), u - x_{j-1} \rangle + h(u) + \frac{1}{2\lambda} \|u - x_{j-1}\|^{2} \right\}$$

#### Theorem

For any given universal constant C > 1, CS with any stepsize  $\lambda$  such that  $\bar{\varepsilon}/(CM_f^2) \leq 4\lambda \leq \bar{\varepsilon}/M_f^2$  has  $\bar{\varepsilon}$ -iteration complexity bound given by

$$\mathcal{O}_1\left(\min\left\{\frac{M_f^2 d_0^2}{\bar{\varepsilon}^2}, \left(\frac{M_f^2}{\mu \bar{\varepsilon}} + 1\right)\log\left(\frac{\mu d_0^2}{\bar{\varepsilon}} + 1\right)\right\}\right)$$
(2)  
where  $d_0 = \inf\{\|x_0 - x^*\| : x^* \in X^*\} = \|x_0 - x_0^*\|.$ 

・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・

| Introduction   | The RPB method  | Main results          | Optimal complexity | Conclusio |
|--|---|-----------------------|--------------------|-----------|
| Bundle method  |   |                       |                    |           |
| <ul><li>Introd</li><li>Ass</li><li>Lite</li></ul>                                    | duction<br>sumptions<br>erature review  |                       |                    |           |
| <ul> <li>2 The F</li> <li>• Rev</li> <li>• Bu</li> <li>• RP</li> <li>• CS</li> </ul> | <b>RPB method</b><br>view of the composi<br><b>ndle method</b><br>'B<br>as an instance of R | te subgradient<br>RPB | method             |           |
| <ul><li>3 Main</li><li>• Up</li><li>• Contract</li></ul>                             | results<br>per complexity bour<br>mplexity bounds for                                       | nds<br>another proxin | nal bundle variant |           |
| <ul><li>④ Optin</li><li>● Lov</li><li>● Op</li></ul>                                 | nal complexity<br>wer complexity bour<br>timal complexity bc                                | nds<br>ounds          |                    |           |

5 Conclusion

| Introduction  | The RPB method | Main results | Optimal complexity | Conclusion |
|---------------|----------------|--------------|--------------------|------------|
| Bundle method |                |              |                    |            |

## Bundle method

Solving the proximal problem

$$x^{+} \leftarrow \min_{u \in \mathbb{R}^{n}} \left\{ \phi(u) + \frac{1}{2\lambda} \|u - x\|^{2} \right\}$$
(3)

▲□▶▲□▶▲≡▶▲≡▶ ≡ めぬぐ

can be as difficult as solving  $\min\{\phi(u) : u \in \mathbb{R}^n\}$ .

Bundle method approximately solves (3) and recursively builds up a model by using a standard cutting-plane approach.

| Introduction  | The RPB method | Main results | Optimal complexity | Conclusion |
|---------------|----------------|--------------|--------------------|------------|
| Bundle method |                |              |                    |            |

## Bundle method

The bundle method solves a sequence of prox subproblems of the form

$$x_j = \operatorname*{argmin}_{u \in \mathbb{R}^n} \left\{ \Gamma_j^{\lambda}(u) := f_j(u) + h(u) + \frac{1}{2\lambda} \|u - x_{j-1}^c\|^2 \right\}, \qquad (4)$$

where  $x_{j-1}^{c}$  is the **prox-center**,  $f_{j}$  is the **cutting-plane** model defined as

$$f_j(u) = \max\{f(x) + \langle f'(x), u - x \rangle : x \in C_j\} \quad \forall u \in \mathbb{R}^n.$$



▲□▶ ▲□▶ ▲□▶ ▲□▶ ▲□ ● ● ●

| Introduction  | The RPB method | Main results | Optimal complexity | Conclusion |
|---------------|----------------|--------------|--------------------|------------|
| Bundle method |                |              |                    |            |
|               |                |              |                    |            |

# Bundle method

The **bundle method** solves a sequence of prox subproblems of the form

$$x_j = \operatorname*{argmin}_{u \in \mathbb{R}^n} \left\{ \Gamma_j^{\lambda}(u) := f_j(u) + h(u) + \frac{1}{2\lambda} \|u - x_{j-1}^c\|^2 \right\},$$

where  $x_{i-1}^{c}$  is the **prox-center**,  $f_{j}$  is the **cutting-plane** model defined as

$$f_j(u) = \max\{f(x) + \langle f'(x), u - x \rangle : x \in C_j\} \quad \forall u \in \mathbb{R}^n,$$

and decides to perform a **serious** or **null** iteration based on the **descent** condition  $\phi(x_j) \leq (1 - \gamma)\phi(x_{j-1}^c) + \gamma(f_j + h)(x_j)$  for some  $\gamma \in (0, 1)$ .

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ● ●

| Introduo |   | The RPB method   | Main results             | Optimal complexity | Conclusio |
|----------|---|--|--------------------------|--------------------|-----------|
| RPB      | Introduc<br>• Assur<br>• Litera                 | ction<br>nptions<br>ature review                               |                          |                    |           |
| 2        | The RP<br>• Revie<br>• Bund<br>• RPB<br>• CS as | B method<br>w of the compos<br>le method<br>s an instance of I | site subgradient<br>RPB  | method             |           |
| 3        | Main re<br>• Uppe<br>• Comp                     | <mark>sults</mark><br>r complexity bou<br>plexity bounds fo    | ınds<br>r another proxin | nal bundle variant |           |
| 4        | Optima<br>• Lowe                                | l complexity<br>r complexity bou                               | inds                     |                    |           |

- Optimal complexity bounds
- **5** Conclusion



0. Let  $x_0 \in \text{dom } h$ ,  $\lambda > 0$  and  $\overline{\varepsilon} > 0$  be given, and set  $x_0^c = x_0$ ,  $C_1 = \{x_0\}$ , and j = 1;

1. Compute

$$x_j = \operatorname*{argmin}_{u \in \mathbb{R}^n} \left\{ \Gamma_j^{\lambda}(u) := f_j(u) + h(u) + \frac{1}{2\lambda} \|u - x_{j-1}^c\|^2 \right\}, \ m_j = \Gamma_j^{\lambda}(x_j).$$

Moreover, consider the function

$$\phi_j^{\lambda} = \phi + \frac{1}{2\lambda} \| \cdot -x_{j-1}^c \|^2,$$
 (5)

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

and let  $\tilde{x}_i$  be such that

$$\tilde{x}_j \in \operatorname{Argmin}\left\{\phi_j^{\lambda}(u) : u \in \{x_j, \tilde{x}_{j-1}\}\right\};$$
 (6)

| Introduction | The RPB method | Main results | Optimal complexity | Conclusion |
|--------------|----------------|--------------|--------------------|------------|
| RPB          |                |              |                    |            |
|              |                |              |                    |            |

2. **If** 

$$t_j = \phi_j^\lambda(\widetilde{x}_j) - m_j \leq rac{\overline{arepsilon}}{2},$$
 (7)

- 2.a) then perform a serious iteration, i.e., set x<sub>j</sub><sup>c</sup> = x<sub>j</sub>, choose an arbitrary finite set C<sub>j+1</sub> such that {x<sub>j</sub>} ⊂ C<sub>j+1</sub>;
- 2.b) else perform a null iteration, i.e., set  $x_j^c = x_{j-1}^c$ , choose  $C_{j+1}$  such that

$$A_j \cup \{x_j\} \subset C_{j+1} \subset C_j \cup \{x_j\}$$
(8)

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

where

$$A_j = \{x \in C_j : f(x) + \langle f'(x), x_j - x \rangle = f_j(x_j)\}$$
(9)

set  $f_{j+1} = \max\{f(x) + \langle f'(x), \cdot - x \rangle : x \in C_{j+1}\};$ 

3. Set  $j \leftarrow j + 1$  and go to step 1.



#### RPB vs. standard bundle method

- introduce an auxiliary iterate  $\tilde{x}_j$ , convergence in  $\{\tilde{x}_j\}$
- null/serious iterate decision making based on t<sub>j</sub>
- motivation for  $\tilde{x}_j$  and  $t_j$ : define  $m_j^* := \min\{\phi_j^{\lambda}(u) : u \in \mathbb{R}^n\}$ , then we have

$$m_j \leq m_j^* \leq \phi_j^\lambda( ilde x_j),$$

and hence

$$\phi_j^\lambda( ilde x_j) - m_j^* \leq t_j \leq rac{arepsilon}{2}$$

where  $t_j = \phi_j^{\lambda}(\tilde{x}_j) - m_j$ .

| Introduction   | The RPB method  | Main results            | Optimal complexity | Conclusion |
|--|---|-------------------------|--------------------|------------|
| CS as an instance of I   | RPB   |                         |                    |            |
| • Ass<br>• Lite  | sumptions<br>erature review   |                         |                    |            |
| <ul> <li>2 The F</li> <li>Rev</li> <li>Bu</li> <li>RP</li> <li>CS</li> </ul> | RPB method<br>view of the compos<br>ndle method<br>B<br>as an instance of F | ite subgradient<br>RPB  | method             |            |
| <ul><li>Main</li><li>Up</li><li>Con</li></ul>                                | results<br>per complexity bou<br>mplexity bounds for                        | nds<br>r another proxin | nal bundle variant |            |

▲□▶ ▲□▶ ▲ □▶ ▲ □▶ □ のへぐ

# Optimal complexity

- Lower complexity bounds
- Optimal complexity bounds

5 Conclusion

| Introduction             | The RPB method | Main results | Optimal complexity | Conclusion |
|--------------------------|----------------|--------------|--------------------|------------|
| CS as an instance of RPB |                |              |                    |            |
|                          |                |              |                    |            |
|                          |                |              |                    |            |

# Exploration-exploitation trade-off

- RPB can be viewed as an inexact proximal point method that consists of a number of stages (exploration) and each stage aims to solve approximately a proximal subproblem by an iterative procedure (exploitation).
- inner complexity:  $\mathcal{O}_1(\lambda M_f^2/\bar{\varepsilon})$ , outer complexity:  $\mathcal{O}_1(d_0^2/(\lambda \bar{\varepsilon}))$ .
- smaller  $\lambda \implies$  less work done inside stages, and more number of stages.
- CS only conducts exploration but no exploitation.
- If  $\lambda = \bar{\varepsilon}/M_f^2$ , then it can be shown that every iteration index of RPB is a serious one, and RPB reduces to CS.

| Introduct  | ion The RPB method  | Main results                     | Optimal complexity | Conclusior |
|------------|---|----------------------------------|--------------------|------------|
| Upper comp | exity bounds  |                                  |                    |            |
|            | <ul><li>Assumptions</li><li>Literature review</li></ul>                                       |                                  |                    |            |
| 2          | The RPB method<br>• Review of the compo<br>• Bundle method<br>• RPB<br>• CS as an instance of | site subgradient<br>RPB          | method             |            |
| 3          | Main results <ul> <li>Upper complexity boo</li> <li>Complexity bounds for</li> </ul>          | <b>unds</b><br>or another proxir | nal bundle variant |            |
| 4          | Optimal complexity<br>• Lower complexity bou  | unds                             |                    |            |

- Optimal complexity bounds
- **5** Conclusion

| Introduction            | The RPB method | Main results | Optimal complexity | Conclusion |
|-------------------------|----------------|--------------|--------------------|------------|
| Upper complexity bounds |                |              |                    |            |
|                         |                |              |                    |            |

# A general complexity bound

#### Theorem

The total number of iterations performed until RPB obtains a  $\bar{\varepsilon}\mbox{-solution}$  is bounded by

$$\mathcal{O}\left(\left[\frac{M_f\min\{\lambda M,\lambda_{\mu}M_f+d_0\}}{\bar{\varepsilon}}+1\right]\left[\min\left\{\frac{d_0^2}{\lambda\bar{\varepsilon}},\frac{1}{\lambda_{\mu}\mu}\log\left(\frac{\mu d_0^2}{\bar{\varepsilon}}+1\right)\right\}+1\right]\right)$$

where

$$M=M_f+M_h, \quad \lambda_\mu=rac{\lambda}{1+\lambda\mu}.$$

▲□▶▲□▶▲□▶▲□▶ □ のQの

| Introduction            | The RPB method | Main results | Optimal complexity | Conclusion |
|-------------------------|----------------|--------------|--------------------|------------|
| Upper complexity bounds |                |              |                    |            |
|                         |                |              |                    |            |

## Reduction to the bound of CS in the strongly convex case

#### Theorem

Let universal constants C, C' > 0 be given and consider an instance  $(x_0, (f, f'; h))$  of (1) which satisfies (A1)-(A4) with parameter triple  $(M_f, M_h, \mu)$  such that

$$\frac{CM_f d_0}{\bar{\varepsilon}} \ge 1, \qquad M_h \in [0, +\infty], \qquad 0 \le \mu \le \frac{C'M_f}{d_0}. \tag{10}$$

Then, RPB with any  $\lambda$  lying in the (nonempty) interval

$$\frac{d_0}{M_f} \le \lambda \le \frac{C d_0^2}{\bar{\varepsilon}} \tag{11}$$

has  $\bar{\varepsilon}$ -iteration complexity bound given by (17).

# Reduction to the bound of CS in the convex case

#### Theorem

Let universal constants C, C' > 0 be given and consider an instance  $(x_0, (f, f'; h))$  of (1) which satisfies (A1)-(A4) with parameter triple  $(\mu, M_f, M_h)$  such that

$$\frac{CM_f d_0}{\bar{\varepsilon}} \ge 1, \qquad M_h \le C' M_f, \qquad \mu = 0.$$
(12)

Then, RPB with any  $\lambda$  lying in the (nonempty) interval

$$\frac{\bar{\varepsilon}}{CM_{f}^{2}} \leq \lambda \leq \frac{Cd_{0}^{2}}{\bar{\varepsilon}}$$
(13)

has  $\bar{\varepsilon}$ -iteration complexity bound  $\mathcal{O}_1(M_f^2 d_0^2/\bar{\varepsilon}^2)$ , and hence agrees with (17).

| Introduction  | The RPB method                                       | Main results    | Optimal complexity | Conclusion |
|---|--|-----------------|--------------------|------------|
| Complexity bounds for   | r another proximal bundle variant                    | :               |                    |            |
| <ol> <li>Introd</li> <li>Ass</li> <li>Lite</li> </ol>           | luction<br>sumptions<br>erature review               |                 |                    |            |
| <ul> <li>The F</li> <li>Rev</li> <li>But</li> <li>RP</li> </ul> | RPB method<br>view of the compos<br>ndle method<br>B | ite subgradient | method             |            |

• CS as an instance of RPB

# 3 Main results

- Upper complexity bounds
- Complexity bounds for another proximal bundle variant

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

# Optimal complexity

- Lower complexity bounds
- Optimal complexity bounds

5 Conclusion

(Kiwiel 2000) and (Du and Ruszczyński 2017) study a proximal bundle variant (PBV) for solving the set constrained problem

$$\min\{\tilde{f}(x): x \in X\}$$
(14)

where X is a nonempty closed convex set and  $\tilde{f}$  is a  $\mu$ -convex ( $\mu \ge 0$ ) finite everywhere function.

• convex, Kiwiel 2000

$$\mathcal{O}_1\left(rac{ ilde{M}^2 ilde{D}^4}{\lambdaar{arepsilon}^3}
ight)$$

where

$$ilde{D}= ilde{D}[ ilde{f}]:=\sup\{d(x_j,X^*):j\geq 0\},\quad ilde{M}= ilde{M}[ ilde{f}]:=\sup\{\| ilde{f}'(x_j)\|:j\geq 0\}.$$

• µ-strongly convex, Du and Ruszczyński 2017

$$\tilde{\mathcal{O}}_1\left(\frac{\tilde{M}^2\lambda}{\alpha^2\bar{\varepsilon}}\right)$$

where  $\alpha := \min\{\lambda \mu, 1\}$ .

### Compare RPB with PBV

Consider  $\tilde{f} = f + h$  with f satisfying (A1)-(A3),  $h \equiv \mu \| \cdot -x_0 \|^2/2$  and  $X = \mathbb{R}^n$ . When  $\mu = 0$ , PBV has the iteration complexity (Kiwiel)

$$\mathcal{O}_1\left(\frac{M_f^2(d_0+\lambda M_f)^4}{\lambda\bar{\varepsilon}^3}\right),\tag{15}$$

when  $\mu > 0$ , PBV has the iteration complexity (Du and Ruszczyński)

$$\tilde{\mathcal{O}}_1\left(\frac{M_f^2}{\lambda\mu^2\bar{\varepsilon}} + \frac{d_0^2}{\lambda\bar{\varepsilon}}\right).$$
(16)

In general, the above bounds are worse than that of PRB

$$\mathcal{O}_1\left(\min\left\{\frac{M_f^2 d_0^2}{\bar{\varepsilon}^2}, \left(\frac{M_f^2}{\mu\bar{\varepsilon}} + 1\right)\log\left(\frac{\mu d_0^2}{\bar{\varepsilon}} + 1\right)\right\}\right).$$
(17)

#### Proof for the convex case

Note that the arithmetic-geometric mean inequality implies that

$$d_0 + \lambda M_f = rac{d_0}{3} + rac{d_0}{3} + rac{d_0}{3} + \lambda M_f \ge 4 \left(rac{1}{27} d_0^3 \lambda M_f\right)^{1/4},$$

and hence that

$$\mathcal{O}_1\left(rac{M_f^2(d_0+\lambda M_f)^4}{\lambdaar{arepsilon}^3}
ight)$$

is minorized by  $\mathcal{O}_1(M_f^3 d_0^3/\bar{\varepsilon}^3)$ , which in turn is minorized by  $\mathcal{O}_1(M_f^2 d_0^2/\bar{\varepsilon}^2)$  in view of the assumption that  $CM_f d_0/\bar{\varepsilon} \geq 1$ .

・ロト・西・・田・・田・・日・

| Introduct  | ion The RPB method  | Main results             | Optimal complexity | Conclusio |
|------------|---|--------------------------|--------------------|-----------|
| Lower comp | lexity bounds   |                          |                    |           |
| 1          | Introduction <ul> <li>Assumptions</li> <li>Literature review</li> </ul>                         |                          |                    |           |
| 2          | The RPB method<br>• Review of the compose<br>• Bundle method<br>• RPB<br>• CS as an instance of | site subgradient<br>RPB  | method             |           |
| 3          | Main results <ul> <li>Upper complexity bot</li> <li>Complexity bounds for</li> </ul>            | unds<br>or another proxi | mal bundle variant |           |
| 4          | Optimal complexity <ul> <li>Lower complexity bot</li> <li>Optimal complexity b</li> </ul>       | unds<br>bounds           |                    |           |
| 5          | Conclusion  |                          |                    |           |

◆□▶ ◆□▶ ◆三▶ ◆三▶ ◆□▶

| Introduction            | The RPB method | Main results | Optimal complexity | Conclusion |
|-------------------------|----------------|--------------|--------------------|------------|
| Lower complexity bounds |                |              |                    |            |
|                         |                |              |                    |            |

# Define a proper class of instances

#### Definition

Given  $(M_f, \mu, R_0) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_{++}$ , let  $\mathcal{I}_{\mu}(M_f, R_0)$  denote the class consisting of all instances  $(x_0, (f, f'; h))$  satisfying conditions (A1)-(A3) and the condition that  $d_0 \leq R_0$ . Moreover, let  $\mathcal{I}_{\mu}^u(M_f, R_0)$  denote the unconstrained class consisting of all instances  $(x_0, (f, f'; h)) \in \mathcal{I}_{\mu}(M_f, R_0)$  such that  $h \equiv \mu \| \cdot \|^2/2$ .

| Introduction            | The RPB method | Main results | Optimal complexity | Conclusion |
|-------------------------|----------------|--------------|--------------------|------------|
| Lower complexity bounds | s              |              |                    |            |
|                         |                |              |                    |            |
|                         |                |              |                    |            |

# $\bar{\varepsilon}\text{-lower}$ complexity bound

#### Theorem

For any given quadruple  $(M_f, \mu, R_0, \overline{\varepsilon}) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_{++} \times \mathbb{R}_{++}$ , there exists an instance  $(x_0, (f, f'; h))$  such that:

a) 
$$(x_0, (f, f'; h)) \in \mathcal{I}^u_\mu(M_f, R_0);$$

b) it has lower complexity bound with respect to  $\mathcal{A}(\mathcal{I}^u_\mu(M_f, R_0), \bar{\varepsilon})$  given by

$$\left\lfloor \min\left\{\frac{M_f^2 R_0^2}{128\bar{\varepsilon}^2}, \frac{M_f^2}{8\mu\bar{\varepsilon}}\right\} \right\rfloor + 1.$$
(18)

As a consequence, (18) is also a  $\bar{\varepsilon}$ -lower complexity bound for any instance class  $\mathcal{I} \supset \mathcal{I}^u_\mu(M_f, R_0)$ .

| Introduc   | tion TI  | ne RPB method                                     | Main results            | Optimal complexity | Conclusion |
|------------|--|---|-------------------------|--------------------|------------|
| Optimal co | mplexity bounds  |   |                         |                    |            |
| 1          | Introductio <ul> <li>Assumpt</li> <li>Literature</li> </ul>  | n<br>ions<br>e review                             |                         |                    |            |
| 2          | The RPB n<br>• Review o<br>• Bundle n<br>• RPB<br>• CS as an | nethod<br>f the compos<br>nethod<br>instance of F | ite subgradient<br>RPB  | method             |            |
| 3          | Main result<br>• Upper co<br>• Complex                       | s<br>mplexity bou<br>ity bounds foi               | nds<br>r another proxir | nal bundle variant |            |
| 4          | Optimal co<br>• Lower co<br>• Optimal                        | mplexity<br>mplexity boun<br>complexity bo        | nds<br>ounds            |                    |            |
| 5          | Conclusion   |   |                         |                    |            |

◆□▶ ◆□▶ ◆三▶ ◆三▶ ◆□▶



#### Optimality in the strongly convex case

Using Theorem 3, we can show that RPB with  $\lambda$  satisfying

$$\frac{R_0}{M_f} \le \lambda \le \frac{CR_0^2}{\bar{\varepsilon}}$$

has a complexity bound

$$\mathcal{O}_1\left(\min\left\{\frac{M_f^2R_0^2}{\bar{\varepsilon}^2},\frac{M_f^2}{\mu\bar{\varepsilon}}\log\left(\frac{\mu R_0^2}{\bar{\varepsilon}}+1\right)\right\}\right),$$

and RPB is optimal for any instance class  ${\cal I}$  and scalar  $\mu \in [0,\, C'M_f/R_0]$  such that

$$\mathcal{I}_{\mu}^{u}(M_{f},R_{0})\subseteq\mathcal{I}\subseteq\mathcal{I}_{\mu}(M_{f},R_{0}). \tag{19}$$

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

## Optimality in the convex case

Using Theorem 4, we can show that RPB with  $\lambda$  satisfying

$$\frac{\bar{\varepsilon}}{CM_f^2} \le \lambda \le \frac{CR_0^2}{\bar{\varepsilon}}$$

has a complexity bound

$$\mathcal{O}_1\left(\frac{M_f^2 R_0^2}{\bar{\varepsilon}^2}\right),$$

and RPB is optimal for any instance class  ${\mathcal I}$  such that

$$\mathcal{I}_0^u(M_f, R_0) \subseteq \mathcal{I} \subseteq \mathcal{I}_0(M_f, R_0; C)$$
(20)

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

where

$$\mathcal{I}_0(M_f, R_0; C) := \{ (x_0, (f, f'; h)) \in \mathcal{I}_0(M_f, R_0) : \exists M_h \le CM_f \text{ such that } h \text{ satisfies (A4)} \}$$

| Introduction | The RPB method | Main results | Optimal complexity | Conclusion |
|--------------|----------------|--------------|--------------------|------------|
|              |                |              |                    |            |
|              |                |              |                    |            |

# **Concluding remarks**

- Iteration-complexity bound for RPB to find a  $\bar{\varepsilon}$ -solution.
- Optimal complexity bounds in both convex and strongly convex settings.
- RPB can be interpreted as an inexact proximal point method.

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

• RPB with sufficiently small constant prox stepsize becomes the composite subgradient method.

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ 三三 - のへぐ

# THE END Thanks!