International Workshop on Numerical Methods and Scientific Computing

6 - 9 May 2026

The Hong Kong Polytechnic University

Blockwise Direct Search Methods

Haitian Li, Zaikun Zhang

Derivative-free optimization (DFO): what and when?

What is DFO?

Solve an optimization problem

$$\min_{x \in \mathbb{R}^n} f(x)$$

using function values but not derivatives (classical or generalized).

When do we use DFO?

- Derivatives are not available even though f may be smooth.
- not available: the evaluation is impossible or too expensive.

A Gap Between Performance and Simplicity in DFO Methods

	Model-based	Direct-search
Performance	good	less satisfactory
Implementation	complicated	relatively simple

- Goal: simple direct search with satisfactory performance.

Blockwise Direct Search Framework

Algorithm: Blockwise Direct Search (BDS)

Input: $x_0 \in \mathbb{R}^n$, $0 < \theta < 1 \le \gamma$, $\alpha_0^1, \dots, \alpha_0^m \in (0, \infty)$, a forcing function ρ, and a blockwise searching sets $D^1, \dots, D^m$.

for $k = 0, 1, \dots$ **do**

Select one block, whose index is denoted by r.

if *there exists* $d \in D^r$ *such that* $f(x_k + \alpha_k^r d) \le f(x_k) - \rho(\alpha_k^r)$ **then**

Set $x_{k+1} = x_k + \alpha_k^r d$ and $\alpha_{k+1}^r = \gamma \alpha_k^r$.

else

Set $x_{k+1} = x_k$ and $\alpha_{k+1}^r = \theta \alpha_k^r$.

Set $\alpha_{k+1}^i = \alpha_k^i$ for all unvisited blocks i.

Choosing Directions and Blocks for Convergence

- For convergence, it is sufficient to work with arbitrary basis
$$B = [b_1, \dots, b_n].$$
- Let $I_1, \dots, I_m$ be a partition of $\{1, \dots, n\}$, and define
$$D^r = \{\pm b_i : i \in I_r\}, \qquad r = 1, \dots, m.$$

A General Condition for Choosing Blocks

- BDS allows several ways to choose the next block.
- For convergence, we only need that each block is selected infinitely times.
- This includes cyclic, random, and many other rules.

Assumptions for Convergence

- f is continuously differentiable and bounded below.
- ∇f is L-Lipschitz continuous.
- Strongly convex case: f is μ-strongly convex.
- Convex case: f is convex, and the initial sublevel set $\mathcal{L}_0$ is bounded.

Theorem (Global Convergence for Convex Problems)

Suppose the assumptions above hold and each block is selected infinitely times. Let $f^\star = \inf_x f(x)$. *For convex* f, *BDS satisfies*

$$f(x_k) \to f^\star, \qquad \nabla f(x_k) \to 0.$$

Why Convergence Holds (Contradiction Argument)

- Let $\bar{x}$ be a cluster point with $\nabla f(\bar{x}) \ne 0$, set $\ell_{\bar{x}}(x) = \nabla f(\bar{x})^\top (x - \bar{x})$.
- Prove $\ell_{\bar{x}}(x_k) \to 0$.
- Strongly convex:
$$\frac{\mu}{2}\|x - \bar{x}\|^2 \le \big(f(x) - f(\bar{x})\big) - \ell_{\bar{x}}(x),$$
so $x_k \to \bar{x}$.
- Convex: Bregman distance gives $\nabla f(x_k) \to \nabla f(\bar{x})$.
- Both contradict $\alpha_k^r \to 0$ for some block r, so $\nabla f(\bar{x}) = 0$.

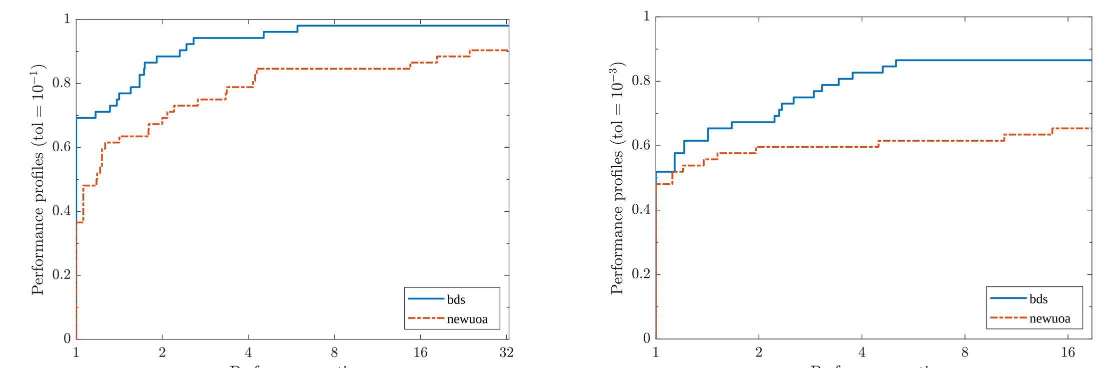
Nonconvex Case: Step-Size Stability

- In nonconvex analysis, convergence requires a balance between step-size expansion γ, contraction θ, and block visits.
- This balance is required in addition to infinite block visits.

Simple implementation

- BDS can be implemented in one A4 page.
- The code has fewer than 40 lines.

Satisfactory performance



(a) $\tau = 10^{-1}$ (b) $\tau = 10^{-3}$

Unconstrained problems, $6 \le n \le 50$

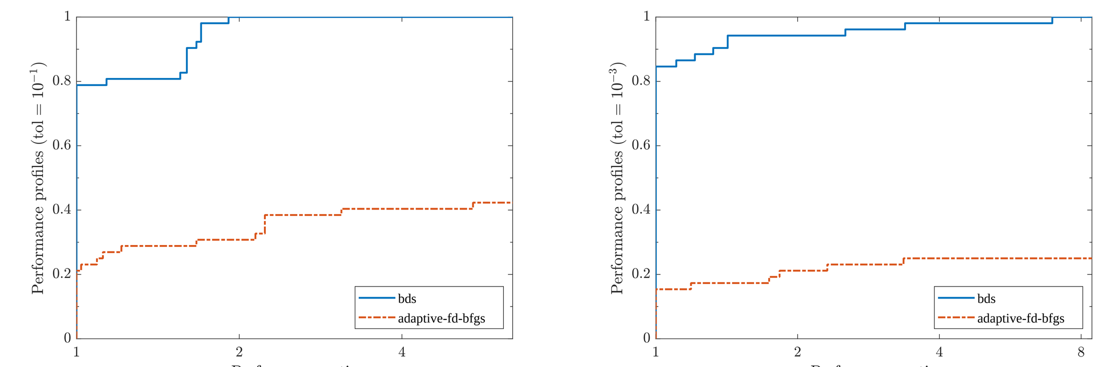
- BDS shows satisfactory performance.

Noise robustness

Observed function value:

$$\tilde{f}(x) = f + \max(|f|, 1)\sigma r(x),$$

where $r(x) \sim \mathcal{N}(0, 1)$ and $\sigma = 10^{-3}$.



(a) $\tau = 10^{-1}$ (b) $\tau = 10^{-3}$

Unconstrained problems, $6 \le n \le 50$

Adaptive stepsize for FD-BFGS: $h = \sqrt{\sigma \max(|f|, 1)}$

The Full BDS Solver

BDS on GitHub

- BDS solver scales the initial step-size using x_0.
- BDS solver handles invalid function values robustly.
- BDS solver avoids wasted evaluations.

haitian-li@connect.polyu.hk; zhangzaikun@mail.sysu.edu.cn