Rui LIU, Wei-chu SUN, Tao HOU, Chun-hong HU, Lin-bo QIAO, 2019. Block coordinate descent with time perturbation for nonconvex nonsmooth problems in real-world studies. *Frontiers of Information Technology & Electronic Engineering*, 20(10):1390-1403. https://doi.org/10.1631/FITEE.1900341

Block coordinate descent with time perturbation for nonconvex nonsmooth problems in real-world studies

Key words: Convergence analysis; Asynchronous block coordinate descent method; Time perturbation; Nonconvex nonsmooth optimization; Real-world study

Corresponding author: Chun-hong HU
E-mail: huchunhong@csu.edu.cn
ORCID: http://orcid.org/0000-0003-3857-4598

Motivation

- The era of big data in healthcare is here, and this era will significantly improve medicine and especially oncology.
- Traditional machine learning algorithms need to be promoted to solve such large-scale real-world problems due to a large amount of data that needs to be analyzed and the difficulty in solving problems with nonconvex nonlinear settings.

Main idea

- We propose a new asynchronous block coordinate descent (BCD) with path-following and time perturbation method, ATP for short, for solving the problem.
- We present analyses for the theoretical convergence and iteration complexity of the proposed ATP method.
- Experimental results demonstrate the convergence behavior of the proposed algorithm for the randomized block variable selection rule with time perturbation.

Method

1. Randomized block variable selection rule
2. Line search criterion
3. Stopping criterion
4. Time perturbation

Major results

Experiments conducted on the real-world applications of constrained folded concave penalized linear regression show that ATP can escape from saddle points and local optima. Besides, ATP shows strong scalability in handling large-scale machine learning problems in a distributed setting.

Major results (Con’d)

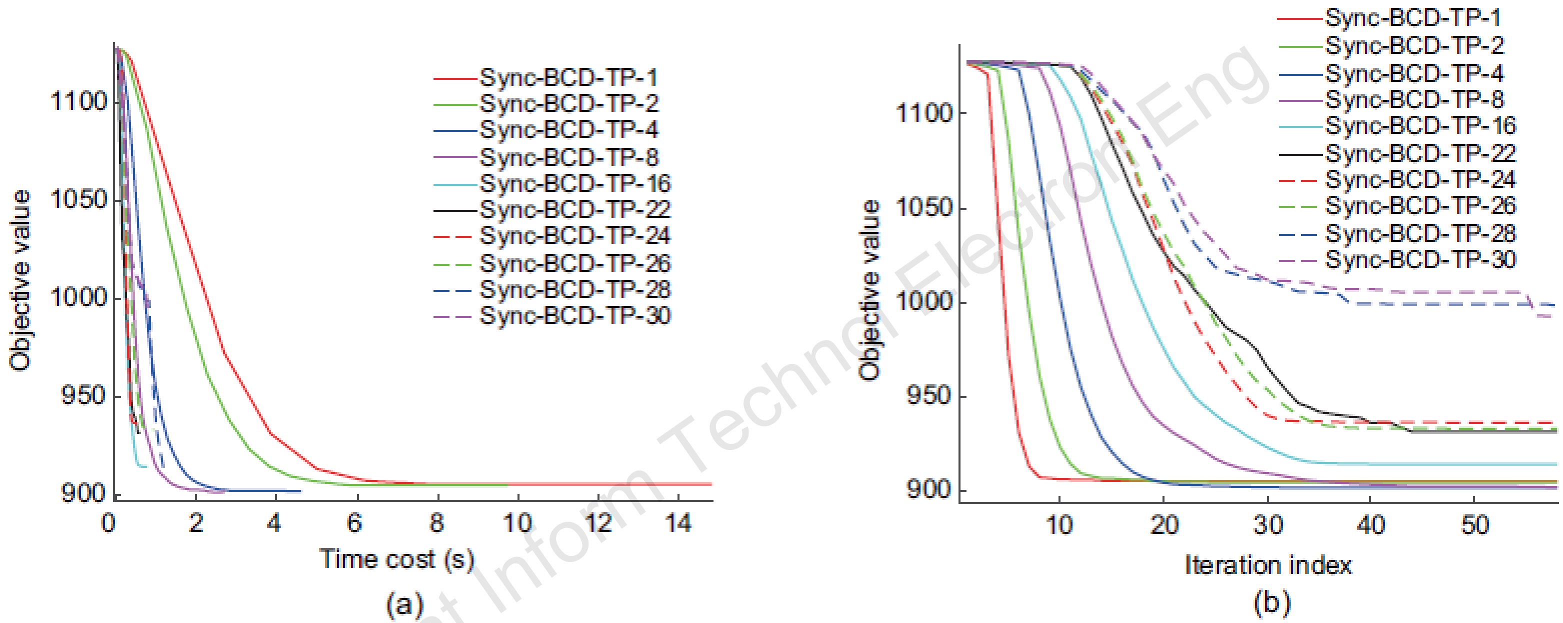

Fig. 2 Convergence behavior of synchronous block coordinate descent with time perturbation (Sync-BCD-TP) with the randomized selection rule: (a) objective value as a function of time cost with different numbers of threads; (b) objective value as a function of the iteration index with different numbers of threads. References to color refer to the online version of this figure

Major results (Con'd)

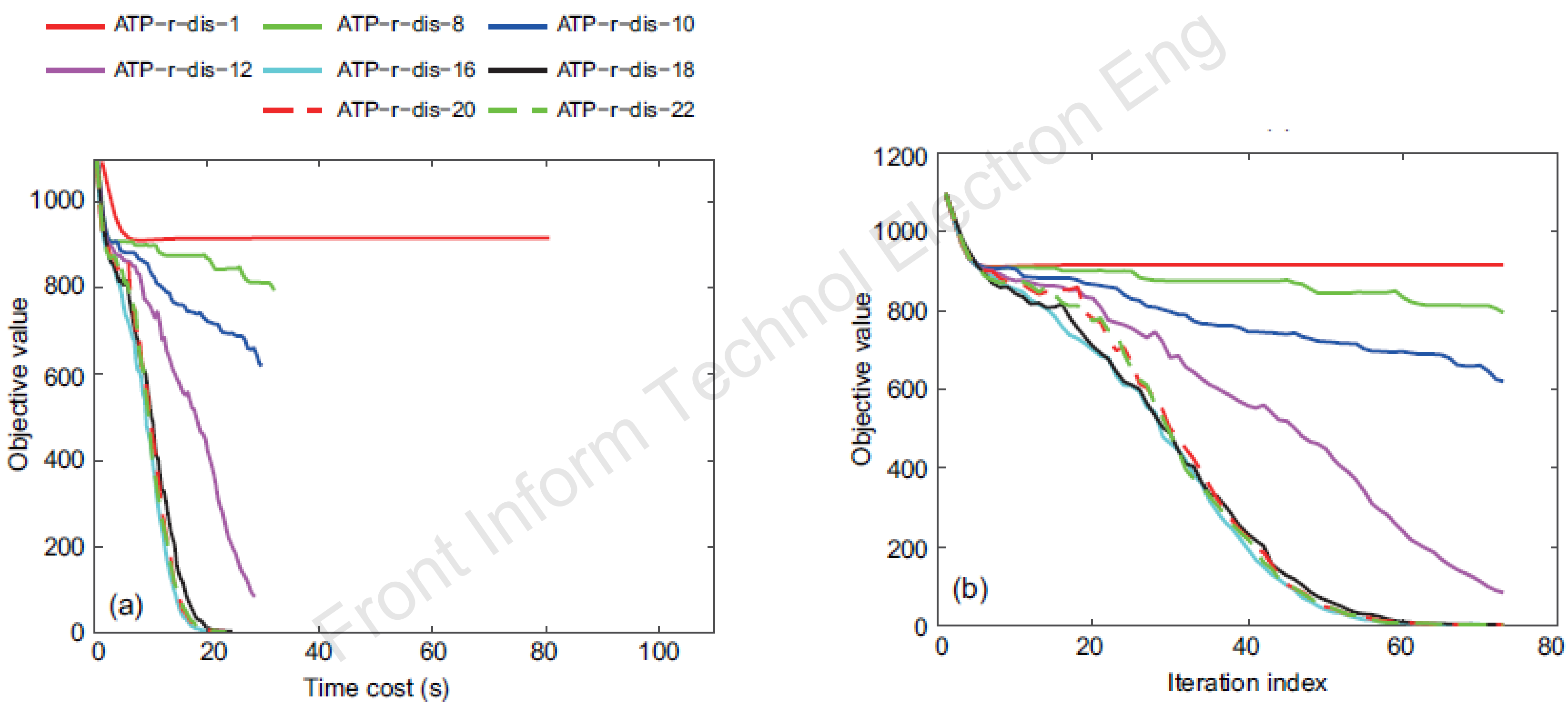

Fig. 3 Convergence behavior of ATP escaping from local optima and saddle points: (a) objective value as a function of time cost; (b) objective value as a function of the iteration index. References to color refer to the online version of this figure

Conclusions

- Experimental results demonstrate that time perturbation enables ATP to escape from saddle points and sub-optimal points, providing a promising way to handle nonconvex optimization problems with inequality constraints employing asynchronous block coordinate descent.
- The asynchronous parallel implementation on shared memory multi-core platforms indicates that the proposed algorithm, ATP, has strong scalability.