Split Block-Coordinate Frank-Wolfe Algorithm

Nicholas Harsell (Advised by Dr. Zev Woodstock)

James Madison University: Optimization and Algorithmic Theory Lab

NSF Grant #DMS-2532423

JMU Mathematics and Statistics

Problem Definition

We consider the minimization of a smooth objective function $f(x)$ over a Hilbert space $\mathcal{H}$, where the solution must lie in the intersection of multiple compact, convex, and nonempty constraint sets C_i:

$$\min_{x\in\mathcal{H}} f(x) \quad \text{subject to} \quad x \in \bigcap_{i=1}^{m} C_i. \qquad (*)$$

Finding a projection onto the full intersection is often computationally prohibitive, motivating a decoupled approach. We make use of newer research on **Frank-Wolfe Splitting Algorithms (FWSAs)**.

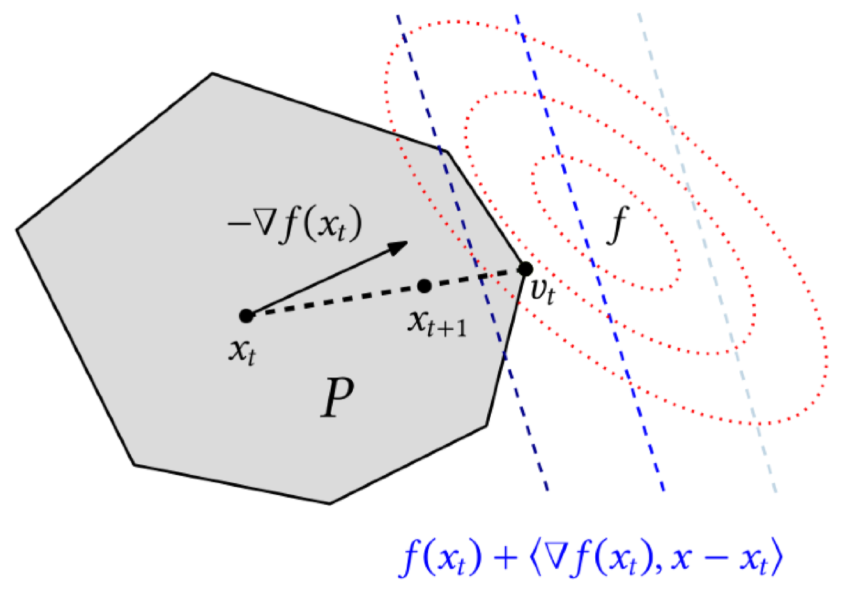

Fig. 1: Standard Frank-Wolfe Visualization [2]

Product Space Reformulation

To decouple the constraints, we lift the problem into a product space:

$$\boldsymbol{x} = (x_1, x_2, \dots, x_m) \in \bigtimes_{i=1}^{m} C_i \subset \mathcal{H}^m.$$

Further define $A : \mathcal{H}^m \to \mathcal{H}$ by

$$A(x_1, x_2, \dots, x_m) = \sum_{i=1}^{m} \omega_i x_i$$

where $\sum_{i=1}^{m} \omega_i = 1$ and $\omega_i \in (0,1]$ for all $1 \le i \le m$. Then minimize a penalized objective over the Cartesian product:

$$\min_{\boldsymbol{x}\bigtimes_{i=1}^{m} C_i} f(A\boldsymbol{x}) + \frac{1}{2}\lambda \operatorname{dist}^2_{\boldsymbol{\mathcal{D}}}(\boldsymbol{x}),$$

where $\boldsymbol{\mathcal{D}}$ is the diagonal subspace and λ is a smoothing parameter.

$\boldsymbol{x} = (x^*, x^*, \dots, x^*) \in \boldsymbol{\mathcal{D}}$ and $A\boldsymbol{x} \in \arg\min f(x)$ gives a point x^* that is a solution to $(*)$.

SCG Algorithm [4]

The Split Conditional Gradient (SCG) Algorithm (Woodstock 2025) Uses this product space reformulation and has been proven to converge for convex and non-convex objectives.

LMO Evaluation Barrier

SCG requires m LMO evaluations per iteration for m constraint sets. We employ a block-coordinate update scheme which has been proven to converge for convex and non-convex objectives [1] to form the **Split Block-Coordinate Frank-Wolfe (SBCFW) Algorithm**, requiring only 1 LMO evaluation per iteration to show convergence.

We only require that **after some $K \ge 1$ iterations, all constraint sets have had their respective LMO evaluated at least once.**

The SBCFW Algorithm

Algorithm 1: Split-Block-Coordinate Frank-Wolfe

Require: $(L_f + \lambda)$-Smooth function f, weights $\{\omega_i\}_{i\in I}$ where $|I| = m$ s.t. $\sum_{i\in I}\omega_i = 1$ and $\omega_i > 0\ \forall i \in I$, initial $\boldsymbol{x}_0 \in \bigtimes_{i\in I} C_i$, access to $(\text{LMO}_i)_{i\in I}$, and $\bigcup_{k=0}^{K-1} I_{t+k} = I$

$x_0 \leftarrow \sum_{i\in I} \omega_i \boldsymbol{x}_0^i$
for $t = 0, 1$ **to** $\dots$ **do**
  Choose a block $I_t \subseteq I$, s.t. $I_t \neq \emptyset$
  Choose penalty parameter $\lambda_t \in \,]0, +\infty[$ (f is $(L_f + \lambda_t)$-smooth)
  $\boldsymbol{g}_t \leftarrow \nabla f(\boldsymbol{x}_t)$
  for $i = 1$ **to** m **do**
    if $i \in I_t$ **then**
      $\boldsymbol{v}_t^i \leftarrow \text{LMO}_i(\boldsymbol{g}_t)$
      $\gamma_t^i = \min\left\{1, \frac{\langle \boldsymbol{g}_t | \boldsymbol{x}_t^i - \boldsymbol{v}_t^i\rangle}{(L_f+\lambda_t)\|\boldsymbol{v}_t^i - \boldsymbol{x}_t^i\|^2}\right\}$
      $\boldsymbol{x}_{t+1}^i \leftarrow \boldsymbol{x}_t^i + \gamma_t^i(\boldsymbol{v}_t^i - \boldsymbol{x}_t^i)$
    else
      $\boldsymbol{x}_{t+1}^i \leftarrow \boldsymbol{x}_t^i$
    end if
  end for
  $\boldsymbol{x}_{t+1} \leftarrow \sum_{i\in I} \omega_i \boldsymbol{x}_{t+1}^i$
end for

Notation and Definitions

Let R be the diameter of $\bigtimes_{i=1}^{m} C_i$. After choosing λ_0, define $\lambda_n := \lambda_0(n+1)^q$ with $q \in (0, \frac{1}{2})$. Then, define $H_0 = F_{\lambda_0}(\boldsymbol{x}_0) - \min_{\boldsymbol{x}\in\bigtimes_{i=1}^{m} C_i} f(A\boldsymbol{x})$, and

$$E_n := 2H_0 + D(\lambda_n - \lambda_0),$$
$$V_n := K(L_f + \lambda_n)R^2,$$
$$\bar{A}_n := \frac{1}{n}\sum_{p=0}^{n-1} A_{pK,\lambda}.$$

Convergence and Rates

Using our defined terms, we find that for sufficiently large n our average Frank-Wolfe gap is bounded as follows:

$$\frac{1}{n}\sum_{p=0}^{n-1} G_{I,\lambda_{pK}}(\boldsymbol{x}_{pK}) \le \begin{cases} \dfrac{E_n}{n} + \dfrac{V_n}{2} - \bar{A}_n & \text{if } n \le \dfrac{2E_n}{V_n}, \\ \sqrt{\dfrac{2E_nV_n}{n}} - \bar{A}_n & \text{otherwise.} \end{cases}$$

Note: All analysis is for the non-convex case, and therefore holds for the convex case as well. Using this block-coordinate update scheme, analysis reveals that the SBCFW algorithm yields a convergence rate of $\mathcal{O}(t^{-1/4})$ on the average smoothed Frank-Wolfe gap in the non-convex case.

Preliminary analysis using [3] suggests that a convergence rate of $\mathcal{O}(t^{-3/8})$ on a subsequence of smoothed Frank-Wolfe gaps can be achieved, and that a schedule for λ_t can be adjusted to allow for a convergence rate of $O(t^{-1/3})$ on the product space objective.

References & Auxiliaries

References

[1] Gábor Braun, Jannis Halbey, Sebastian Pokutta, and Zev Woodstock. Flexible Block-Iterative Analysis For the Frank-Wolfe Algorithm. *Journal on Optimization Theory and Applications (to appear)*, 2026.

[2] Sebastian Pokutta. The Frank-Wolfe Algorithm: A Short Introduction. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 126(1):3–35, Dec 2023.

[3] Antonio Silveti-Falls, Cesare Molinari, and Zev Woodstock. Frank-Wolfe with Moreau Envelope Smoothing for Nonsmooth Nonconvex Problems, 2026.

[4] Zev Woodstock and Sebastian Pokutta. Splitting the Conditional Gradient Algorithm. *SIAM Journal on Optimization*, 35(1):347–368, 2025.