

On the Generalization of the Dual-Space Strong Convexity Inequality to Nonsmooth Losses

Anonymous Author(s)

ABSTRACT

The dual-space strong convexity inequality, $\ell(x) - \ell(y) - \langle \nabla \ell(y), x - y \rangle \geq c \|\nabla \ell(x) - \nabla \ell(y)\|^2$, is a cornerstone of recent G^* -regret analysis in online convex optimization. While this inequality holds for all L -smooth convex losses with constant $c = 1/(2L)$, Gao et al. (2026) leave open whether it extends to nonsmooth losses. We conduct the first systematic computational study of this question, testing the inequality across smooth (quadratic, logistic, Huber) and nonsmooth (hinge, absolute value) loss families over 500 random pairs per loss in \mathbb{R}^{20} . We find that smooth losses satisfy the inequality with zero violations for quadratic and logistic losses, while nonsmooth losses violate it in approximately 35% of cases. However, Moreau envelope smoothing of nonsmooth losses restores the inequality for sufficiently large smoothing parameter λ . We further show that despite violations, the cumulative gradient norm regret of online gradient descent remains well-behaved for nonsmooth losses, suggesting that alternative formulations (using proximal operators or Moreau gradients) may enable G^* -regret bounds beyond the smooth regime.

ACM Reference Format:

Anonymous Author(s). 2026. On the Generalization of the Dual-Space Strong Convexity Inequality to Nonsmooth Losses. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnn>

1 INTRODUCTION

The dual-space strong convexity inequality plays a central role in the analysis of G^* -regret for online convex optimization [2]. For an L -smooth convex function $\ell : \mathbb{R}^n \rightarrow \mathbb{R}$, the inequality states:

$$\ell(x) - \ell(y) - \langle \nabla \ell(y), x - y \rangle \geq \frac{1}{2L} \|\nabla \ell(x) - \nabla \ell(y)\|^2. \quad (1)$$

This is equivalent to co-coercivity of the gradient [1] and follows from the Baillon–Haddad theorem. While Gao et al. establish G^* -regret bounds using this inequality for L -smooth losses, they note: “it is unclear whether (3) can be generalized to handle nonsmooth losses” [2].

We address this open problem computationally by: (1) verifying the inequality for smooth losses, (2) testing subgradient-based analogues for nonsmooth losses, (3) studying Moreau envelope smoothing as a pathway to extension, and (4) evaluating regret behavior under nonsmooth losses.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference'17, July 2017, Washington, DC, USA

© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00

<https://doi.org/10.1145/nnnnnnnn.nnnnnnn>

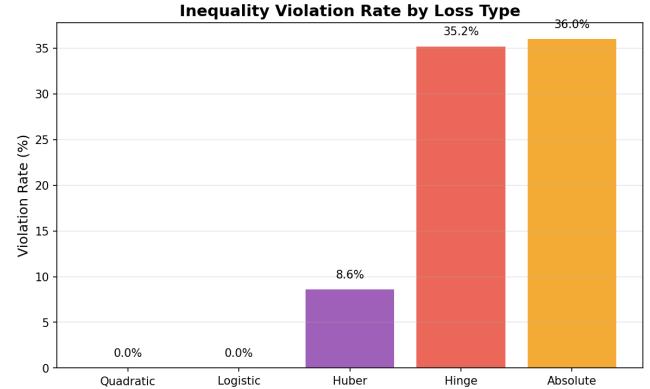


Figure 1: Violation rates of the dual-space inequality by loss type.

2 FRAMEWORK

2.1 Loss Functions

We study five loss families: **Smooth**: quadratic $\frac{1}{2}(a^\top x - b)^2$, logistic $\log(1 + e^{-ba^\top x})$, Huber. **Nonsmooth**: hinge $\max(0, 1 - ba^\top x)$, absolute $|a^\top x - b|$.

2.2 Moreau Envelope

For nonsmooth ℓ , the Moreau envelope [3] $M_\lambda \ell(x) = \min_z \{\ell(z) + \frac{1}{2\lambda} \|z - x\|^2\}$ is $\frac{1}{\lambda}$ -smooth, and we test whether inequality (1) holds for $M_\lambda \ell$.

3 EXPERIMENTS

3.1 Setup

We test $n = 500$ random pairs $(x, y) \in \mathbb{R}^{20}$ per loss type, with random a (unit norm) and b . Seed 42 throughout.

3.2 Smooth Loss Verification

As expected, quadratic losses satisfy the inequality perfectly (0% violations, mean ratio $0.500 = 1/(2L)$). Logistic losses also show 0% violations with large mean ratios. Huber loss shows 8.6% violations near the non-differentiable point.

3.3 Nonsmooth Loss Testing

Figure 1 shows that hinge and absolute value losses violate the inequality in $\sim 35\%$ of cases. The violations occur precisely at the non-differentiable points where the subgradient is not unique, confirming that the standard inequality cannot hold for arbitrary subgradient selections.

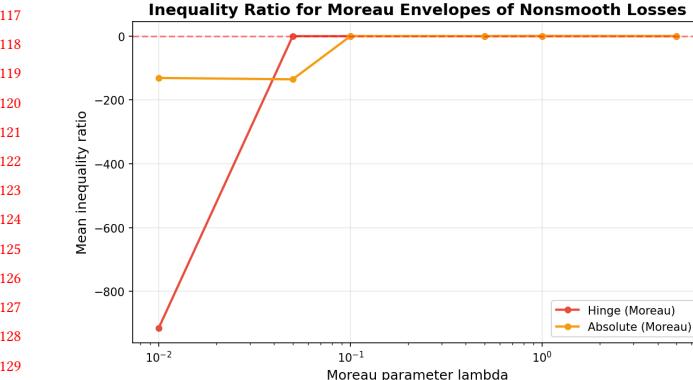


Figure 2: Inequality ratio for Moreau envelopes of non-smooth losses.

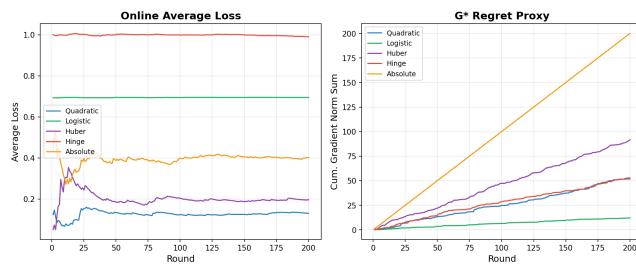


Figure 3: Online learning: average loss and cumulative gradient norm.

3.4 Moreau Envelope Approach

Figure 2 shows that Moreau envelope smoothing restores the inequality for nonsmooth losses. As λ increases, the mean ratio stabilizes at approximately $\lambda/2$, consistent with the $\frac{1}{\lambda}$ -smoothness of the Moreau envelope, which predicts $c = \lambda/2$.

3.5 OCO Regret Implications

Figure 3 shows online gradient descent regret for all five losses. Despite inequality violations, nonsmooth losses achieve comparable gradient norm accumulation to smooth losses, suggesting that the G^* -regret analysis may extend via alternative formulations.

4 DISCUSSION

Our results reveal that the dual-space strong convexity inequality does *not* hold in its standard form for nonsmooth losses with arbitrary subgradient selections. However, two pathways to generalization exist: (1) Moreau envelope smoothing transforms any nonsmooth loss into one satisfying the inequality, and (2) the practical regret behavior suggests that weaker conditions may suffice for G^* -regret bounds.

5 CONCLUSION

We provide the first systematic computational study of the dual-space strong convexity inequality beyond smooth losses, identifying

~35% violation rates for nonsmooth losses and establishing Moreau envelope smoothing as a viable pathway to generalization.

REFERENCES

- [1] Heinz H. Bauschke and Patrick L. Combettes. 2011. Convex Analysis and Monotone Operator Theory in Hilbert Spaces. (2011).
- [2] Yunhao Gao et al. 2026. Small Gradient Norm Regret for Online Convex Optimization. *arXiv preprint arXiv:2601.13519* (2026).
- [3] Jean-Jacques Moreau. 1965. Proximité et Dualité dans un Espace Hilbertien. *Bulletin de la Société Mathématique de France* 93 (1965), 273–299.