

Computing an Entire Solution Path of a Nonconvexly Regularized Convex Sparse Model



Yi Zhang, Isao Yamada (email: {yizhang,isao}@sp.ce.titech.ac.jp)

Department of Information & Communications Engineering, Tokyo Institute of Technology

Tokyo Tech

Introduction

1. Sparse least-squares problems:

$$\underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize}} \quad J(\mathbf{x}) := \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \Psi(\mathbf{x}). \quad (1)$$

$\mathbf{y} \in \mathbb{R}^m$: measurement $\mathbf{A} \in \mathbb{R}^{m \times n}$: sensing matrix Ψ : sparse regularizer

2. Conventional convex/nonconvex sparse regularizers:

- (1) ℓ_1 -norm (LASSO [1]): **convex** 😊 **biased** 😞
- (2) the minimax concave (MC [2]) penalty: **less biased** 😊 **nonconvex** 😞

3. The generalized minimax concave (GMC [3]) penalty:

$$\Psi(\mathbf{x}) = \|\mathbf{x}\|_1 - \min_{\mathbf{z} \in \mathbb{R}^n} \left(\|\mathbf{z}\|_1 + \frac{1}{2} \|\mathbf{B}(\mathbf{x} - \mathbf{z})\|_2^2 \right)$$

The diagram illustrates the GMC penalty as the difference between the ℓ_1 -norm (represented by a diamond-shaped contour plot) and a debiasing function (represented by a smooth, elliptical contour plot). The resulting GMC penalty contour plot shows the intersection of these two functions.

- (1) Overall convexity condition: if $\mathbf{B} \in \mathbb{R}^{p \times n}$ satisfies $\mathbf{A}^\top \mathbf{A} \geq \lambda \mathbf{B}^\top \mathbf{B}$, then $J(\cdot)$ in (1) is **convex** despite **nonconvexity** of the GMC penalty.
- (2) We study the scaled GMC (sGMC) model: $\mathbf{B} := \sqrt{\frac{\rho}{\lambda}} \mathbf{A}$ with $\rho \in [0, 1)$.

4. Key contributions:

we prove that the minimum ℓ_2 -norm solution path of the sGMC model is **piecewise linear** with respect to λ , and can be computed by an iterative algorithm **within finite steps**.

Theoretical Results

1. Optimality condition of the sGMC model:

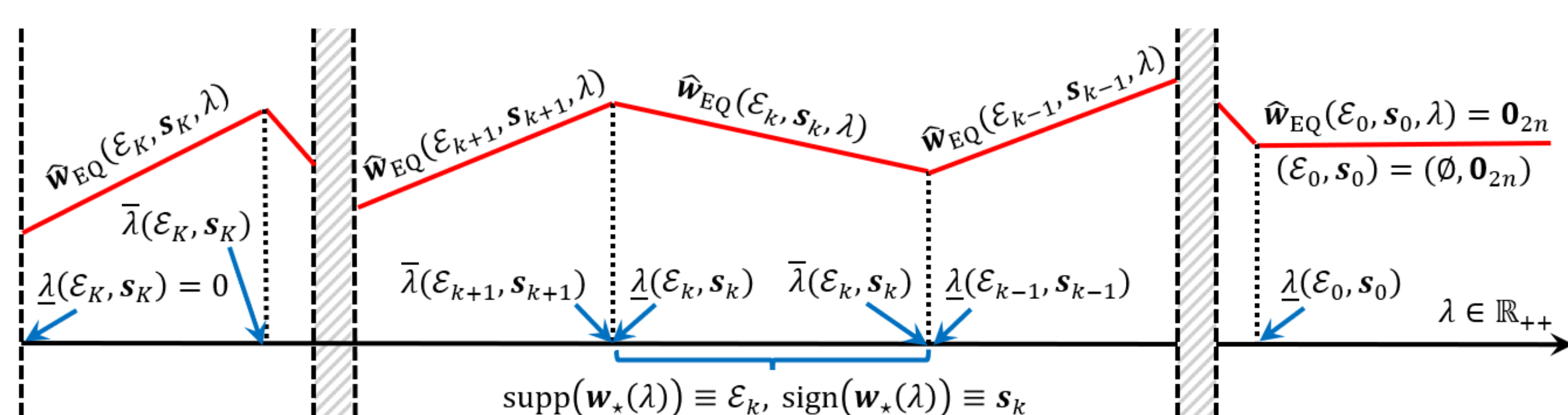
$\mathbf{x} \in \mathbb{R}^n$ is a solution of the sGMC model if and only if there exists $\mathbf{z} \in \mathbb{R}^n$, such that $\mathbf{w} := [\mathbf{x}^\top \quad \mathbf{z}^\top]^\top$ satisfies

$$\mathbf{0} \in \mathbf{C}^\top (\mathbf{D}\mathbf{C}\mathbf{w} - \mathbf{b}) + \lambda \partial(\|\cdot\|_1)(\mathbf{w}), \quad (2)$$

where $\mathbf{C} := \text{blkdiag}(\mathbf{A}, \mathbf{A})$, $\mathbf{b} := [\mathbf{y}^\top \quad \mathbf{0}^\top]^\top$, $\mathbf{D} := \begin{bmatrix} (1-\rho)\mathbf{I} & \rho\mathbf{I} \\ -\rho\mathbf{I} & \rho\mathbf{I} \end{bmatrix}$.

We call $\mathbf{w} \in \mathbb{R}^{2n}$ an *extended solution* if it satisfies (2). In this paper, we focus on the minimum ℓ_2 -norm extended solution $\mathbf{w}_*(\lambda)$.

2. Properties of the minimum ℓ_2 -norm extended solution:



- (1) $\mathbf{w}_*(\lambda)$ is **piecewise linear** in λ with **finite** ($\leq 3^{2n}$) linear pieces.
- (2) Within every linear piece of $\mathbf{w}_*(\lambda)$, there exist uniquely a combination of $\mathcal{E} \subset \{1, 2, \dots, 2n\}$ and $\mathbf{s} \in \{-1, 0, 1\}^{2n}$ such that:
 - $\text{supp}(\mathbf{w}_*(\lambda)) \equiv \mathcal{E}$, $\text{sign}(\mathbf{w}_*(\lambda)) \equiv \mathbf{s}$ are **constant**.
 - $\mathbf{w}_*(\lambda) \equiv \hat{\mathbf{w}}_{\text{EQ}}(\mathcal{E}, \mathbf{s}, \lambda)$, the latter is the least-squares solution of

$$\begin{cases} (\forall i \in \mathcal{E}) & \mathbf{c}_i^\top (\mathbf{b} - \mathbf{D}\mathbf{C}\mathbf{w}) = \lambda s_i, \\ (\forall j \in -\mathcal{E}) & w_j = 0, \end{cases} \quad (\text{EQ-a})$$

$$(\forall j \in -\mathcal{E}) \quad w_j = 0, \quad (\text{EQ-b})$$

- (3) The duration of λ for a linear piece is the set of $\lambda > 0$ satisfying

$$\begin{cases} (\forall i \in \mathcal{E}) & s_i w_i \geq 0. \\ (\forall j \in -\mathcal{E}) & |\mathbf{c}_j^\top (\mathbf{b} - \mathbf{D}\mathbf{C}\mathbf{w})| \leq \lambda, \end{cases} \quad (\text{NQ-a})$$

$$(\forall j \in -\mathcal{E}) \quad |\mathbf{c}_j^\top (\mathbf{b} - \mathbf{D}\mathbf{C}\mathbf{w})| \leq \lambda, \quad (\text{NQ-b})$$

Algorithmic Results

1. Least Angle Regression for the sGMC model:

The minimum ℓ_2 -norm sGMC solution path $\mathbf{w}_*(\lambda)$ can be computed by an extension of the well-known least angle regression (LARS [4]) algorithm:

- (1) In **each iteration**, the proposed LARS-sGMC algorithm computes $(\mathcal{E}, \mathbf{s})$ corresponding to **one linear piece** of $\mathbf{w}_*(\lambda)$.
- (2) Let $(\mathcal{E}_k, \mathbf{s}_k)$ corresponds to the k th linear piece, then:
 - The LARS-sGMC iteration computes the lower breakpoint $\underline{\lambda}(\mathcal{E}_k, \mathbf{s}_k)$ of the current linear piece by solving (NQ).
 - We conduct certain deletion and insertion operations on indices in $\mathcal{E}_k$ to obtain $\mathcal{E}_{k+1}$, and change the components of $\mathbf{s}_k$ in response to the change of $\mathcal{E}_k$ to yield $\mathbf{s}_{k+1}$.

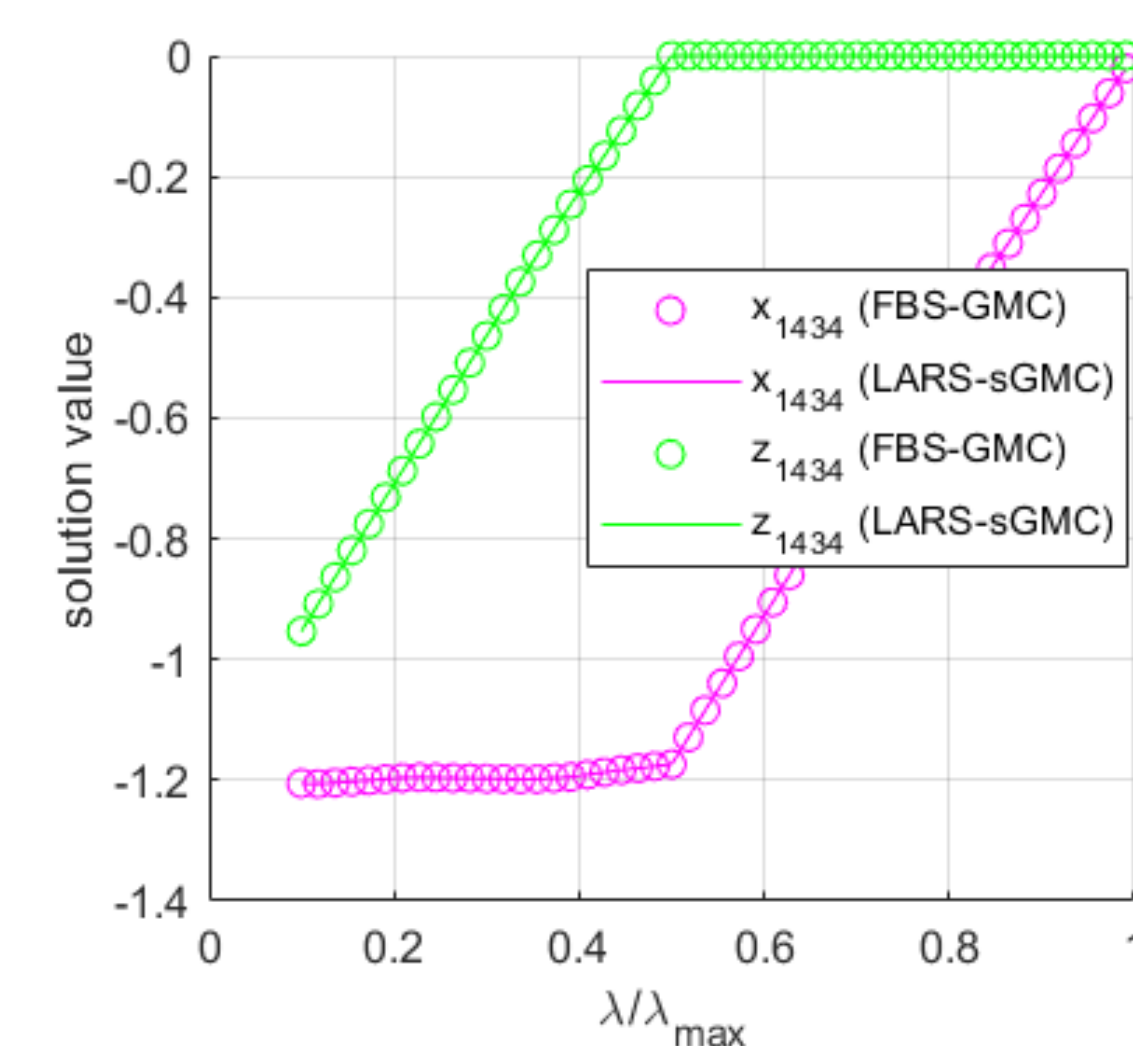
See Algorithm 1 in the **extended version of this paper** [5] for details.

2. Properties of LARS-sGMC:

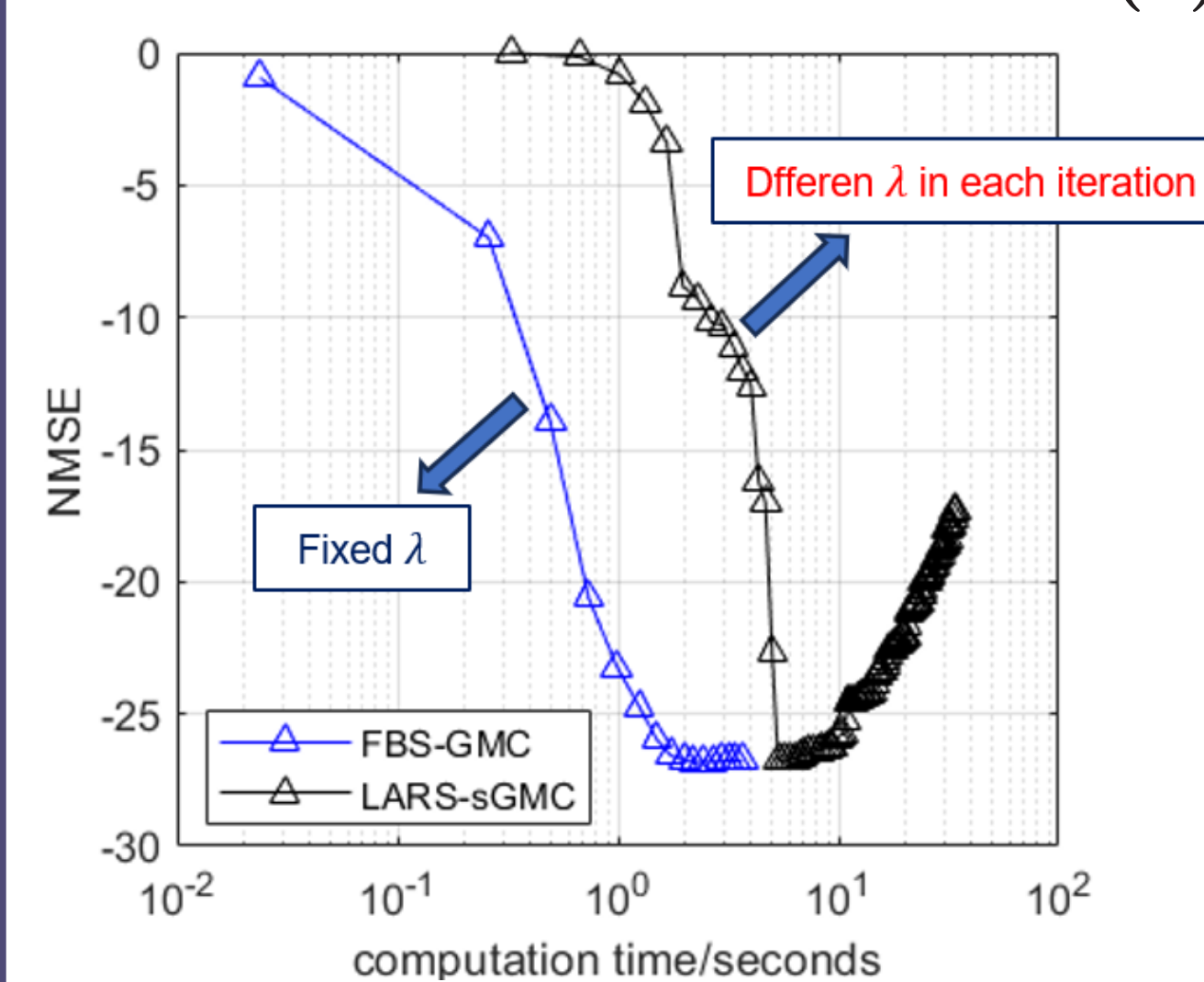
- (1) We prove the correctness and finite termination of LARS-sGMC under a mild assumption.
- (2) The complexity of the k th iteration is $\mathcal{O}(m|\mathcal{E}_k|^2 + |\mathcal{E}_k|^3)$.
- (3) If we set $\rho = 0$ in the sGMC model, then LARS-sGMC reduces to the conventional LARS [4] algorithm for LASSO.

Experiments

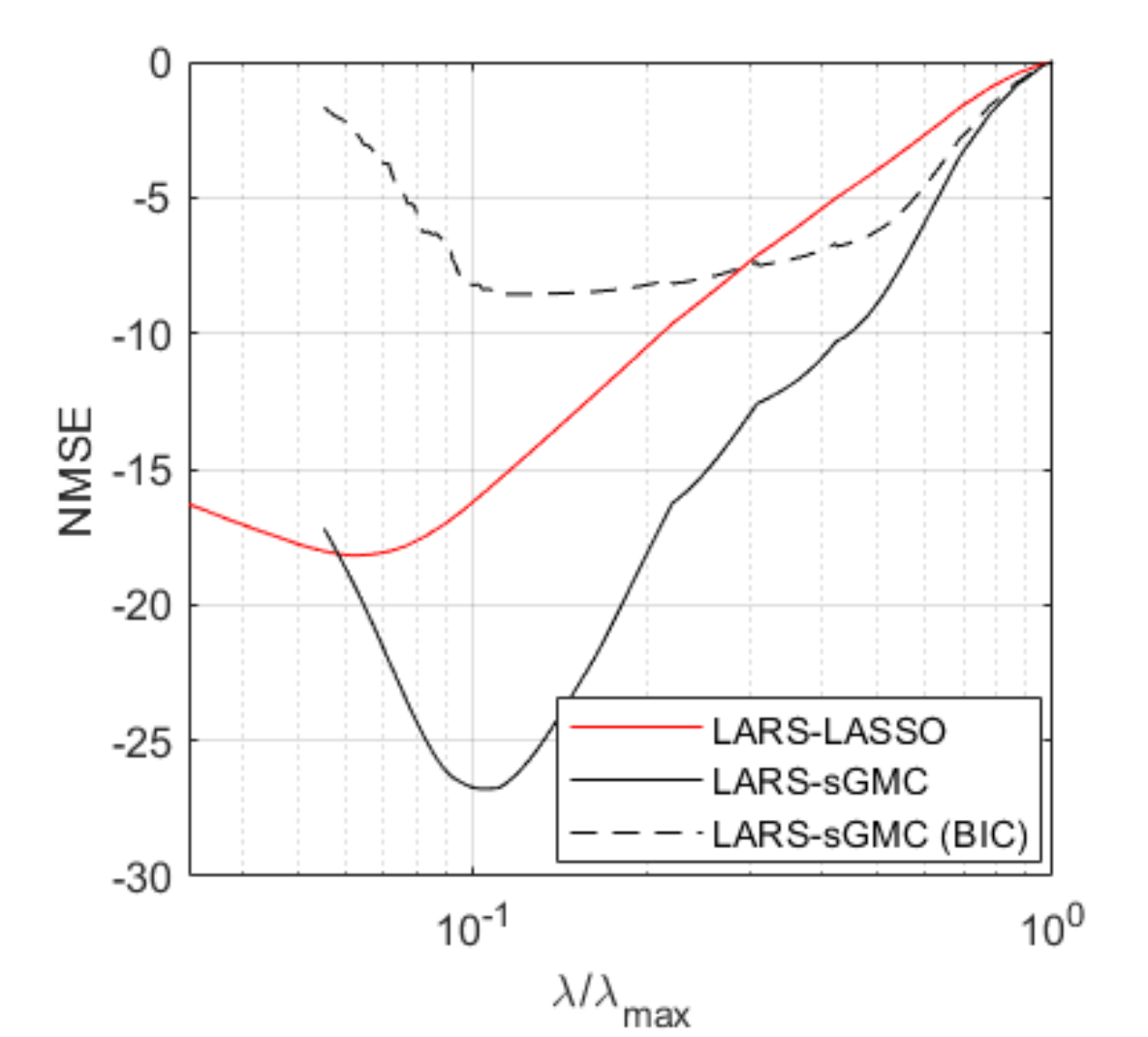
We demonstrate the **correctness**, **efficiency** and **practical utility for regularization parameter tuning** of the LARS-sGMC algorithm.



(a) Correctness.



(b) Efficiency.



(c) Practical utility.

Related Publications

- [1] R. Tibshirani, "Regression Shrinkage and Selection Via the Lasso," *J. R. Stat. Soc., B: Stat. Methodol.*, vol. 58, no. 1, pp. 267–288, 1996.
- [2] C. Zhang, "Nearly unbiased variable selection under minimax concave penalty," *Ann. Stat.*, vol. 38, no. 2, pp. 894–942, 2010.
- [3] I. Selesnick, "Sparse Regularization via Convex Analysis," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4481–4494, 2017.
- [4] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Stat.*, vol. 32, no. 2, pp. 407–499, 2004.
- [5] Y. Zhang and I. Yamada, "Solution-Set Geometry and Regularization Path of a Nonconvexly Regularized Convex Sparse Model," Mar 2024, arXiv: 2311.18438 [math].