A Unified Framework for Uniform Signal Recovery in Nonlinear Generative Compressed Sensing

Junren Chen, Jonathan Scarlett, Michael K. Ng, Zhaoqiang Liu

HKU, NUS, HKBU, UESTC

Highlights

- **Nonlinear GCS**: In generative compressed sensing (GCS) we seek to recover a signal $\mathbf{x}$ that lies in a bounded k-input L-Lipschitz generative models $G(\cdot):\mathbb{B}_2^k(r)\to\mathbb{R}^n$. We deal with a nonlinear model $y_i=f_i(\mathbf{a}_i^\top\mathbf{x})$ with possibly discontinuous/unknown f_i, which captures 1-bit/multi-bit (dithered) quantization models and single index model.
- **A Uniform Recovery Framework**: We build a unified framework to establish uniform recovery guarantee for generalized Lasso.
- **Near-Optimal Rate**: Our main theorem shows that typically $O(\frac{k}{\epsilon^2}\log P(L))$ ($P(L)$ is a polynomial on L) measurements suffice for uniform recovery of all $\mathbf{x}\in\text{Range}(G)$ up to ϵ-ℓ_2-error, improving on (Genzel and Stollenwerk, FOCM, 2023) for classical compressed sensing (e.g., with sparse prior).

Problem Setup

- **Nonlinear GCS model: (Assump. 1)** $G:\mathbb{B}_2^k(r)\to\mathbb{R}^n$ is L-Lipschitz continuous, we observe $y_i=f_i(\mathbf{a}^\top\mathbf{x})$, $i=1,...,m$ with $\mathbf{a}\sim\mathcal{N}(0,\mathbf{I}_n)$.
- **Discontinuous f_i: (Assump. 2)** We handle possibly unknwon f_i with countably infinite jump discontinuities that is piece-wisely Lipschitz continuous, including (but far beyond) various quantization models.
- **Generalized Lasso:** We achieve uniform recovery via

$$\hat{\mathbf{x}}=\text{arg min}\,\frac{1}{2}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2,\ \text{s.t. }\mathbf{x}\in T\cdot\mathcal{K}\qquad(1)$$

where $\mathcal{K}=G(\mathbb{B}_2^k(r))$, T is a rescaling factor (to be chosen).

Technical Ingredients Needed For Uniform Recovery

- **Lipschitz Approximation: (Assump. 3)**
  - We handle discontinuous f_i by constructing its Lipschitz approximation $f_{i,\beta}$

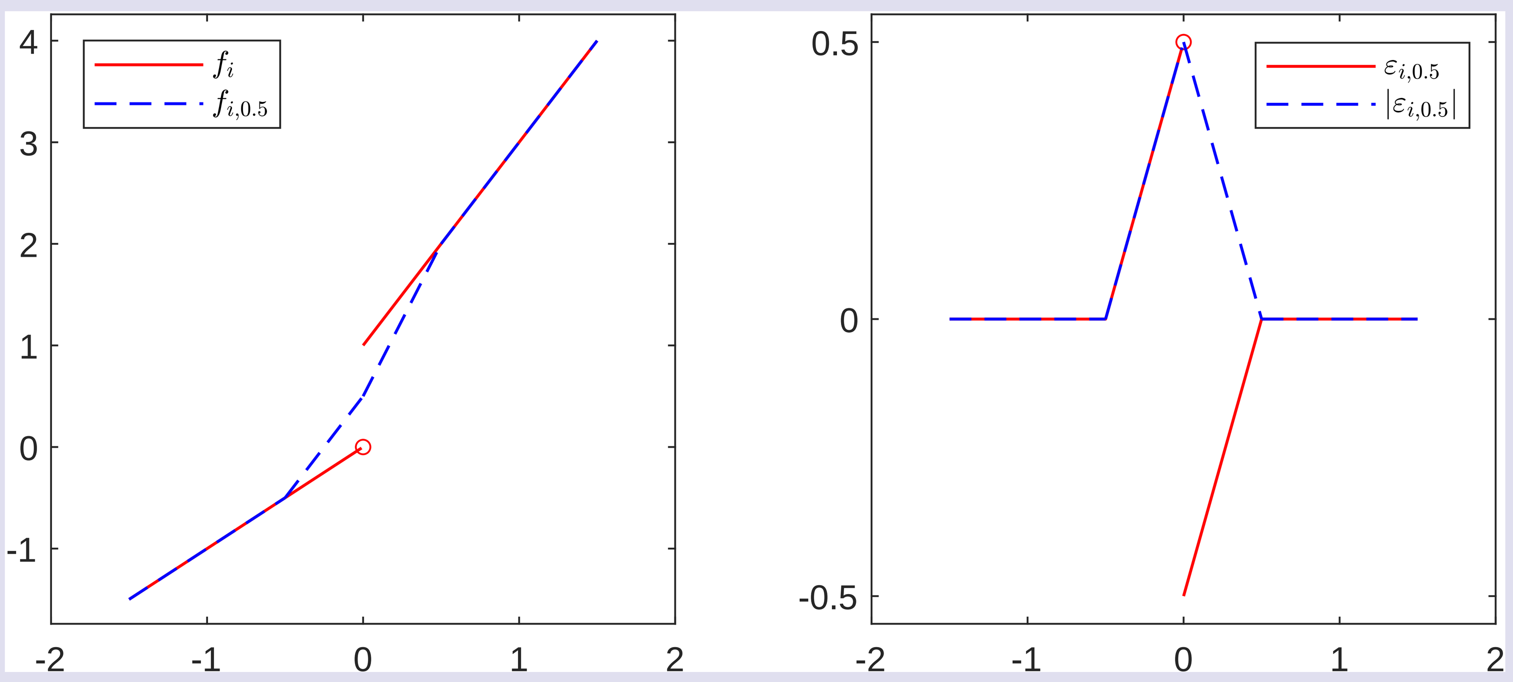

Figure 1:(Left): f_i and its approximation $f_{i,0.5}$; (Right): approximation error $\varepsilon_{i,0.5}$, $|\varepsilon_{i,0.5}|$.

  - We define $\xi_{i,\beta}(a)=f_{i,\beta}(a)-Ta$, $\varepsilon_{i,\beta}(a)=f_{i,\beta}(a)-f_i(a)$ and require the bounds $\sup_{\mathbf{x}\in\mathcal{K}}\|\xi_{i,\beta}(\mathbf{a}^\top\mathbf{x})\|_{\psi_2}\le A_g^{(1)}$ and $\sup_{\mathbf{x}\in\mathcal{K}}\|\varepsilon_{i,\beta}(\mathbf{a}^\top\mathbf{x})\|_{\psi_2}\le A_g^{(2)}$.
  - We also require bounds on $\sup_{\mathbf{x}\in\mathcal{K}}|\xi_{i,\beta}(\mathbf{a}^\top\mathbf{x})|$ and $\sup_{\mathbf{x}\in\mathcal{K}}|\varepsilon_{i,\beta}(\mathbf{a}^\top\mathbf{x})|$, but they can be crude ones and totally unproblematic.
- **Small Mismatch: (Assump. 4)**
  - The mismatch associated with the nonlinearity f_i, defined as $\rho(\mathbf{x})=\|E[f_i(\mathbf{a}_i^\top\mathbf{x})]-T\mathbf{x}\|_2$
  - The mismatch induced by the Lipschitz approximation $f_{i,\beta}$, defined as $\mu_\beta(\mathbf{x})=P(\mathbf{a}^\top\mathbf{x}\in\mathscr{D}_{f_i}+[-\frac{\beta}{2},\frac{\beta}{2}])$, where $\mathscr{D}_{f_i}$ is the set of discontinuities of f_i.
  - We require $\sup_{\mathbf{x}\in\mathcal{K}}\rho(\mathbf{x})$ and $\sup_{\mathbf{x}\in\mathcal{K}}\sqrt{\mu_\beta(\mathbf{x})}$ to be $O((A_g^{(1)}+A_g^{(2)})\sqrt{\frac{k}{m}})$

Master Theorem: Uniform Recovery with Sharp Rate

- **Theorem 1 (Main Thm.):** Under Assump. 1-4, given any $\epsilon\in(0,1)$, if $m\gtrsim(A_g^{(1)}+A_g^{(2)})\frac{k}{\epsilon^2}P(L)$, then w.h.p. on a single draw of $(\mathbf{a}_i,f_i)_{i=1}^m$, we have $\|\hat{\mathbf{x}}-T\mathbf{x}\|_2\le\epsilon$ for all $\mathbf{x}\in\mathcal{K}$, where $\hat{\mathbf{x}}$ is as per (1).
- **Implications:** We check Assump. 1-4 for specific models to get the uniform sharp ℓ_2 error rate $\tilde{O}(\sqrt{\frac{k\log P(L)}{m}})$:
  - **1-bit GCS:** $f_i(\cdot)=\text{sign}(\cdot)$, recovering result from (Liu and Scarlett, NeurIPS, 2020) without using local embedding property
  - **1-bit Dithered GCS:** $f_i(\cdot)=\text{sign}(\cdot+\tau_i)$ with uniform dither τ_i, yielding more general results with guarantee comparable to (Qiu et al., ICML, 2020)
  - **Lipschitz-continuous SIM:** $f_i(\cdot)$ is possibly unknown, random, and Lipschitz continuous, improving result from (Liu and Scarlett, NeurIPS, 2020) without using local embedding property
  - **Multi-bit Dithered GCS:** $f_i(\cdot)=\mathcal{Q}_\delta(\cdot+\tau)$ with uniform dither τ_i, yielding new result not available in the literature.

Prove Sharp Rate by Tighter Concentration Inequality

- **Technical Challenges**:
  - Compared to non-uniform guarantee, proving a uniform guarantee is much more challenging. In particular, we need to bound the product process taking the form

$$\sup_{\mathbf{x}\in\mathcal{X}}\sup_{\boldsymbol{v}\in\mathcal{V}}\left[h(\mathbf{a}_i^\top\mathbf{x})\mathbf{a}_i^\top\boldsymbol{v}-\mathbb{E}\big(h(\mathbf{a}_i^\top\mathbf{x})\mathbf{a}_i^\top\boldsymbol{v}\big)\right]\qquad(2)$$

  - By Lipschitz approximation, we manage to render $h(\cdot)$ Lipschitz continuous.
- **The key to get sharp rate:**
  - It's natural to use the concentration inequality due to (Mendelson, 2016) to bound (2), but this in general does not yield a sharp rate but a rate of $m^{-1/4}$ instead, as per (Genzel and Stollenwerk, FOCM, 2023)
  - Observe that in the setting of GCS, $\mathcal{X}$ and $\mathcal{V}$ in (2) both possess low metric entropy. By covering argument, we develop a concentration inequality for product process that yields essentially tighter bound in such setting.
- **Theorem 2: (Tighter Bound on (2), informal and simplified)** Let $\mathscr{H}(\mathcal{X},r)=\log\mathscr{N}(\mathcal{X},r)$ be the metric entropy. Suppose that $\mathscr{H}(\mathcal{X},r)$ and $\mathscr{H}(\mathcal{V},r)$ only logarithmically depend on r, then if $\|h(\mathbf{a}_i^\top\mathbf{x})\|_{\psi_2}\le A_1$, $\|\mathbf{a}_i^\top\|_{\psi_2}\le A_2$, then w.h.p. we can bound (2) as $A_1A_2\sqrt{\frac{\mathscr{H}(\mathcal{X},r_1)+\mathscr{H}(\mathcal{V},r_2)}{m}}$. (We omit r_1,r_2 since they have logarithmic impact on the bound)

Numerical Results: Recovering Multiple Signals With One Design

- Reconstructed images and quantitative results of the MNIST dataset for uniformly quantized CS with dithering measurements.