Robust 1-Bit Compressive Sensing: How the Sign of Random Projections Distinguishes Sparse Vectors

Laurent Jacques (ICTEAM/ELEN, UCL)

Joint work with:

Jason Laska (Rice U., TX, USA), Petros Boufounos (MERL, USA) and Richard Baraniuk (Rice U., TX, USA)

> FNRS Contact Group "Wavelets and applications" Dec. 21, 2011



1-bit Sampling?





1-bit Sampling?





1-bit Sampling?



1-bit Sampling?

epl

Why 1-bit? Very Fast Quantizers! $V_{\rm out}$ Theoretical slope = 1/3 b/dB 25 Flash -1 **Actual Slope** Folding 23 = 1/2.3 b/dB Half-Flash 21 × Pipelined X SAR Stated Number of Bits (N) 19 Sigma-Delta × XX Unknown 17 册 15 13 х 11 9 ж 7 P Degradation 5 10 20 30 40 50 60 70 80 90 100 0 $10\log(f_s)$ (dBsps)

[FIG1] Stated number of bits versus sampling rate.

Université catholique de Louvain ер

[From "Analog-to-digital converters" B. Le, T.W. Rondeau, J.H. Reed, and C.W.Bostian, IEEE Sig. Proc. Magazine, Nov 2005]

1

[FIG1] Stated number of bits versus sampling rate.

[From "Analog-to-digital converters" B. Le, T.W. Rondeau, J.H. Reed, and C.W.Bostian, IEEE Sig. Proc. Magazine, Nov 2005]

- Dec. 21, 2011

Few slides to understand Compressed Sensing

Compressed Sensing

Compressed Sensing

Compressed Sensing

Université de louisie epile de louisie et lo

Université (i) ept victeam ELEN

Université (i) Epl 🚽 icteam ELEN

Compressed Sensing Limits

* CS measurements $\in \mathbb{R}^{M}$, what happens if quantized ? $\mathcal{Q}[\Phi x]$

victeam ELEN

* "Ok" for high resolution quantization (\sim noise):

[Candes, Tao, 2004] [Jacques, Hammond, Fadili 2009, 2011] [Laska et al. 2009] [Dai, Pham, Milenkovic, 2009] ...

$$\|\boldsymbol{x} - \boldsymbol{x}^*\| \leq C \|\boldsymbol{y} - \boldsymbol{\Phi} \boldsymbol{x}\|$$

meas. distortion

* But is it still valid for extreme quantization? 1-bit?

1-bit Compressed Sensing

1-bit Compressed Sensing

M-bits! But, which information inside y_s ?

M-bits! But, which information inside y_s ?

Intuitively ... x on S^2 , M vectors $\{\varphi_i : 1 \leq i \leq M\}$

UCL Université activitation

Université 🔞 EPL 🧹 icteam ELEN

Université (i) EDI dicteam ELEN

Université de lovaire de lovaire

Starting point: Hamming/Angle Concentration

* Metrics of interest:

$$d_{H}(\boldsymbol{u}, \boldsymbol{v}) = \frac{1}{M} \sum_{i} (u_{i} \oplus v_{i}) \quad \text{(norm. Hamming)}$$
$$d_{\text{ang}}(\boldsymbol{x}, \boldsymbol{s}) = \frac{1}{\pi} \arccos(\langle \boldsymbol{x}, \boldsymbol{s} \rangle) \quad \text{(norm. angle)}$$

Starting point: Hamming/Angle Concentration

* Metrics of interest:

random plane

-icteam ELEN

UCL Université catholique de Laurain $d_{H}(\boldsymbol{u}, \boldsymbol{v}) = \frac{1}{M} \sum_{i} (u_{i} \oplus v_{i}) \quad \text{(norm. Hamming)}$ $d_{\text{ang}}(\boldsymbol{x}, \boldsymbol{s}) = \frac{1}{\pi} \arccos(\langle \boldsymbol{x}, \boldsymbol{s} \rangle) \quad \text{(norm. angle)}$

* Known fact: if $\Phi \sim \mathcal{N}^{M \times N}(0, 1)$ [*e.g.*, Goemans, Williamson 1995]

Let
$$\boldsymbol{\Phi} \sim \mathcal{N}^{M \times N}(0, 1), A(\cdot) = \operatorname{sign}(\boldsymbol{\Phi} \cdot) \in \{-1, 1\}^M \text{ and } \epsilon > 0.$$

For any $\boldsymbol{x}, \boldsymbol{s} \in S^{N-1}$, we have

$$\mathbb{P}_{\Phi}\left[\left|d_{H}(A(\boldsymbol{x}), A(\boldsymbol{s})) - d_{\mathrm{ang}}(\boldsymbol{x}, \boldsymbol{s})\right| \leq \epsilon\right] \geq 1 - 2e^{-2\epsilon^{2}M}.$$

Thanks to A(.), Hamming distance concentrates around vector angles!

Binary ϵ Stable Embedding (Bese)

A mapping $A : \mathbb{R}^N \to \mathcal{B}^M$ is a **binary** ϵ -stable embedding (B ϵ SE) of order K for sparse vectors if

$$d_{\text{ang}}(\boldsymbol{x}, \boldsymbol{s}) - \epsilon \leqslant d_H(A(\boldsymbol{x}), A(\boldsymbol{s})) \leqslant d_{\text{ang}}(\boldsymbol{x}, \boldsymbol{s}) + \epsilon$$

for all $\boldsymbol{x}, \boldsymbol{s} \in S^{N-1}$ with $\boldsymbol{x} \pm \boldsymbol{s}$ K-sparse.

kind of "binary restricted (quasi) isometry"

Binary ϵ Stable Embedding (B ϵ SE)

A mapping $A : \mathbb{R}^N \to \mathcal{B}^M$ is a **binary** ϵ -stable embedding (B ϵ SE) of order K for sparse vectors if

$$d_{\text{ang}}(\boldsymbol{x}, \boldsymbol{s}) - \epsilon \leqslant d_H(A(\boldsymbol{x}), A(\boldsymbol{s})) \leqslant d_{\text{ang}}(\boldsymbol{x}, \boldsymbol{s}) + \epsilon$$

for all $\boldsymbol{x}, \boldsymbol{s} \in S^{N-1}$ with $\boldsymbol{x} \pm \boldsymbol{s}$ K-sparse.

-icteam ELEN

kind of "binary restricted (quasi) isometry"

- * Corollary: for any algorithm with output \boldsymbol{x}^* jointly K-sparse and consistent (i.e., $A(\boldsymbol{x}^*) = A(\boldsymbol{x})$), $d_{\mathrm{ang}}(\boldsymbol{x}, \boldsymbol{x}^*) \leq 2\epsilon!$
- * If limited binary noise, d_{ang} still bounded
- * If not exactly sparse signals (but almost), d_{ang} still bounded

$B\epsilon SE$ existence? Yes!

Let $\mathbf{\Phi} \sim \mathcal{N}^{M \times N}(0, 1)$, fix $0 \leq \eta \leq 1$ and $\epsilon > 0$. If

$$M \geq \frac{4}{\epsilon^2} \left(K \ln(N) + 2K \ln(\frac{50}{\epsilon}) + \ln(\frac{2}{\eta}) \right),$$

then Φ is a B ϵ SE with Pr > 1 - η .

 $M = O(\epsilon^{-2} K \ln N)$

$B\epsilon SE$ existence? Yes!

Let $\mathbf{\Phi} \sim \mathcal{N}^{M \times N}(0, 1)$, fix $0 \leq \eta \leq 1$ and $\epsilon > 0$. If

$$M \geq \frac{4}{\epsilon^2} \left(K \ln(N) + 2K \ln(\frac{50}{\epsilon}) + \ln(\frac{2}{\eta}) \right),$$

then $\mathbf{\Phi}$ is a B ϵ SE with Pr > 1 - η .

Proof sketch:

1) Generalize

$$\mathbb{P}_{\Phi}\left[\left|d_{H}\left(A(\boldsymbol{x}), A(\boldsymbol{s})\right) - d_{\mathrm{ang}}(\boldsymbol{x}, \boldsymbol{s})\right| \leq \epsilon\right] \geq 1 - 2e^{-2\epsilon^{2}M}.$$

to

$$\mathbb{P}_{\Phi}\left[\left|d_{H}\left(A(\boldsymbol{u}), A(\boldsymbol{v})\right) - d_{\mathrm{ang}}(\boldsymbol{x}, \boldsymbol{s})\right| \leq \epsilon + \left(\frac{\pi}{2}D\right)^{1/2}\delta\right] \geq 1 - 2e^{-2\epsilon^{2}M}.$$

for $\boldsymbol{u}, \boldsymbol{v}$ in a *D*-dimensional neighborhood of width δ around \boldsymbol{x} and \boldsymbol{s} resp. 2) Covers the space of "*K*-sparse signal pairs" in \mathbb{R}^N by

 $O(\binom{N}{K}\delta^{-2K}) = O((\frac{eN}{K\delta^2})^K)$ neighborhoods.

3) Apply Point 1 with union bound, and "stir until the proof thickens"

 $M = O(\epsilon^{-2} K \ln N)$

random plane

Hope for better bounds? ... Limited

 $B\epsilon SE$ consistency "width":

$$\epsilon = O((K/M)^{(1-\alpha)/2} (\ln N)^{1/2}), \text{ for any } \alpha > 0.$$

Hope for better bounds? ... Limited

 $B\epsilon SE$ consistency "width":

$$\epsilon = O((K/M)^{(1-\alpha)/2} (\ln N)^{1/2}), \text{ for any } \alpha > 0.$$

Forgetting stability, we can prove: for two unit K-sparse signals

$$A(\boldsymbol{x}) = A(\boldsymbol{s}) \quad \Rightarrow \quad \|\boldsymbol{x} - \boldsymbol{s}\| \leq O((K/M)^{1-\alpha} \ln N)$$

 $(\mathit{i.e.}, \ d_{H}=0)$

Hope for better bounds? ... Limited

 $B\epsilon SE$ consistency "width":

$$\epsilon = O((K/M)^{(1-\alpha)/2} (\ln N)^{1/2}), \text{ for any } \alpha > 0.$$

Forgetting stability, we can prove: for two unit K-sparse signals

$$A(\boldsymbol{x}) = A(\boldsymbol{s}) \quad \Rightarrow \quad \|\boldsymbol{x} - \boldsymbol{s}\| \leq O((K/M)^{1-\alpha} \ln N)$$

 $(\mathit{i.e.},\ d_{H}\!=0)$

Lower bound: Worst case distance btw two unit K-sparse vectors:

$$\epsilon_{\text{worst}} = \Omega(K/M)$$
 [P. Boufounos]

1-bit CS Reconstructions ?

Numerical Reconstructions:

* [Boufounos, Baraniuk 2008]

$$\boldsymbol{x}^* = \operatorname*{arg\,min}_{\boldsymbol{u}} \| \boldsymbol{u} \|_1$$
 s.t. $\operatorname{diag}(A(\boldsymbol{x})) \Phi \boldsymbol{u} > 0$ and $\| \boldsymbol{u} \|_2 = 1$

+ other iterative methods: Matching Sign pursuit (MSP), Restricted-Step Shrinkage (RSS)

Numerical Reconstructions:

* [Boufounos, Baraniuk 2008]

$$\boldsymbol{x}^* = \operatorname*{arg\,min}_{\boldsymbol{u}} \|\boldsymbol{u}\|_1$$
 s.t. $\operatorname{diag}(A(\boldsymbol{x})) \Phi \boldsymbol{u} > 0$ and $\|\boldsymbol{u}\|_2 = 1$

+ other iterative methods: Matching Sign pursuit (MSP), Restricted-Step Shrinkage (RSS)

* Binary Iterative Hard Thresholding (BIHT):

Given $\boldsymbol{y}_s = A(\boldsymbol{x})$ and K, set l = 0, $\boldsymbol{x}^0 = 0$: ($\tau > 0$ controls gradient descent)

$$\begin{array}{l} & \boldsymbol{x}^{l+1} = \boldsymbol{x}^l + \frac{\tau}{2} \Phi^T \left(\boldsymbol{y}_s - A(\boldsymbol{x}^l) \right), \\ & \boldsymbol{x}^{l+1} = \eta_K(\boldsymbol{a}^{l+1}), \quad l \leftarrow l+1 \end{array} \quad (\text{irgradient" towards consistency}) \\ \end{array}$$

Stop when $d_H(\boldsymbol{y}_s, A(\boldsymbol{x}^{l+1})) = 0$ or $l = \max$. iter.

with $\eta_K(\boldsymbol{u}) = \text{best } K\text{-term approximation of } \boldsymbol{u}$

····→ minimizes $\mathcal{J}(\boldsymbol{x}) = \|[\operatorname{diag}(\boldsymbol{y}_s)(\Phi \boldsymbol{x})]_-\|_1$, with $(\lambda)_- = (\lambda - |\lambda|)/2$ (connections with ML hinge loss, 1-bit classification)

Simulations

N = 1000, K = 10 Bernoulli-Gaussian model normalized signals 1000 trials

Matching Sign pursuit (MSP) Restricted-Step Shrinkage (RSS) Binary Iterative Hard Thresholding (BIHT)

(tested also for signals sparse in DCT, wavelets, ...)

Université (a) CPU dicteam ELEN

Simulations: Testing $B\epsilon SE$ $d_{ang}(\boldsymbol{x}, \boldsymbol{x}^*) - \epsilon(M) \leq d_H(A(\boldsymbol{x}), A(\boldsymbol{x}^*)) \leq d_{ang}(\boldsymbol{x}, \boldsymbol{x}^*) + \epsilon(M)$

M/N = 0.7

M/N = 1.5

Universite de louvain ELEN

icteam ELEN

icteam ELEN

Final Comparison: CS vs bits/meas.

icteam ELEN

- * Keeping sign of random measurements distinguishes sparse vectors $(B\epsilon SE)$
- * Algorithms exist to reconstruct good signal estimate (up to amplitude)

- * Keeping sign of random measurements distinguishes sparse vectors $(B\epsilon SE)$
- * Algorithms exist to reconstruct good signal estimate (up to amplitude)
 - * Stability and Convergence Guarantees? Recent success in [Y. Plan, R. Vershynin, 2011A]
- * Generalization (Gaussian mean width) [Y. Plan, R. Vershynin, 2011B]

- * Keeping sign of random measurements distinguishes sparse vectors $(B\epsilon SE)$
- * Algorithms exist to reconstruct good signal estimate (up to amplitude)
 - * Stability and Convergence Guarantees? Recent success in [Y. Plan, R. Vershynin, 2011A]
- * Generalization (Gaussian mean width) [Y. Plan, R. Vershynin, 2011B]
- * What is the link btw Sensing Matrix and Quantization?
 - Ex: a Bernoulli (+/- 1) matrix don't work in 1-bit CS!!!
 [Y. Plan, R. Vershynin, 2011]
 - * Quantization is a sampling of meas. distribution...
- * Linking $B\epsilon SE$ to $\Sigma\Delta$ quantization (1 bit) ?

- * Keeping sign of random measurements distinguishes sparse vectors $(B\epsilon SE)$
- * Algorithms exist to reconstruct good signal estimate (up to amplitude)
 - * Stability and Convergence Guarantees? Recent success in [Y. Plan, R. Vershynin, 2011A]
- * Generalization (Gaussian mean width) [Y. Plan, R. Vershynin, 2011B]
- * What is the link btw Sensing Matrix and Quantization?
 - Ex: a Bernoulli (+/- 1) matrix don't work in 1-bit CS!!!
 [Y. Plan, R. Vershynin, 2011]
 - * Quantization is a sampling of meas. distribution...
- * Linking $B\epsilon SE$ to $\Sigma\Delta$ quantization (1 bit) ?
- * Short BIHT matlab code available: <u>http://perso.uclouvain.be/laurent.jacques/index.php/Main/BIHTDemo</u>
- * 1-Bit CS Resource Page: <u>http://dsp.rice.edu/1bitCS</u>

Subset of References (& Thank you for your attention)

- * Chen, S. S., Donoho, D. L., & Saunders, M. A. "Atomic Decomposition by Basis Pursuit". SIAM Journal on Scientific Computing, 20(1), 33-61, 1998.
- * Candès, E. J., Romberg, J., & Tao, T, "Stable signal recovery from incomplete and inaccurate measurements". Comm. Pure Appl. Math, 59(8), 1207-1223, 2006.
- * Candès, E. J., "The restricted isometry property and its implications for compressed sensing". Compte Rendus de l'Academie des Sciences, Paris, Serie I, 346, 589-592, 2008.
- * Goemans, M. & Williamson, D., "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," Journ. ACM, vol. 42, no. 6, pp. 1145, 1995.
- * Laska, J., Boufounos, P., Davenport, M., Baraniuk, R., "Democracy in action: Quantization, saturation, and compressive sensing," App. Comp. and Harm. Anal. (ACHA), 2011.
- * Boufounos, P., & Baraniuk, R., "1-Bit Compressive Sensing", 42nd annual Conference on Information Sciences and Systems (CISS) (pp. 19-21). Princeton, NJ., 2008.
- * Plan, Y., & Vershynin, R., "One-bit compressed sensing by linear programming", preprint, 2011, <u>http://arxiv.org/abs/1109.4299</u>
- * Plan, Y., & Vershynin, R., "Dimension Reduction by Random Hyperplane Tessellations", Submitted. November 2011.
- * 1-Bit CS Resource Page: <u>http://dsp.rice.edu/1bitCS/</u>

icteam

ELEN