k-bit Hamming Compressed Sensing

Tianyi Zhou
Centre for Quantum Computation & Intelligent Systems
FEIT, University of Technology Sydney, Australia
Email: tianyi.david.zhou@gmail.com

Dacheng Tao
Centre for Quantum Computation & Intelligent Systems
FEIT, University of Technology Sydney, Australia
Email: dacheng.tao@uts.edu.au

***Abstract*—We consider recovering d-level quantization of a signal from k-level quantization of linear measurements. This problem has great potential in practical systems, but has not been fully addressed in compressed sensing (CS). We tackle it by proposing k-bit *Hamming compressed sensing* (HCS). It reduces the decoding to a series of hypothesis tests of the bin where the signal lies in. Each test equals to an independent nearest neighbor search for a histogram estimated from quantized measurements. This method is based on that the distribution of the ratio between two random projections is defined by their intersection angle. Compared to CS and 1-bit CS, k-bit HCS leads to lower cost in both hardware and computation. It admits a trade-off between recovery/measurement resolution and measurement amount and thus is more flexible than 1-bit HCS. A rigorous analysis shows its error bound. Extensive empirical study further justifies its appealing accuracy, robustness and efficiency.**

I. Introduction

Recently, prosperous works in compressed sensing (CS) [1][2] show that an accurate recovery can be achieved by sampling signal at a rate proportional to its underlying "information content" rather than bandwidth. The key improvement of CS is that the sampling rate can be significantly reduced below Nyquist rate by replacing the uniform sampling with linear measurement, when the signals are sparse or compressible on certain dictionary. In particular, CS dedicates to rebuild a sparse signal $x \in \mathbb{R}^n$ from its linear measurements by solving an underdetermined system.

$$\min_{x\in\mathbb{R}^n} \|x\|_p \quad s.t. \quad y = \Phi x, \quad (0 \leq p < 2)$$

where $\Phi \in \mathbb{R}^{m\times n}$ is the sensing matrix allowing $m \ll n$ and fulfilling the restricted isometry property (RIP) or incoherence condition, and ℓ_p norm encourages sparsity. It has been proved that a number of random matrix ensembles satisfy RIP well.

However, in practical digital systems the measurements have to be discretized to a finite number of bits by quantization, which merely gives an interval each measurement lies in. If we use the centroid of the given interval to approximate the linear measurement in traditional CS methods, the distortion caused by quantization can be ignored only when the quantization is of high-resolution. But this requires expensive ADCs. Thus several recent studies develop CS methods treating each measurement as an uncertain value distributed within the given interval. They observe by this mean CS can be successfully done even with very coarse quantized measurements.

$$\min_{x\in\mathbb{R}^n} \|x\|_p \quad s.t. \quad u \leq \Phi x \leq v, \quad (0 \leq p < 2)$$

An extreme case is $y = \mathrm{sign}(\Phi x)$. 1-bit CS [3][4][5] ensures consistent reconstructions of signals on the unit ℓ_2 sphere [6]. The 1-bit measurements lead to a low-cost and robust hardware implementation.

Nevertheless, the signal rebuilt by 1-bit CS and other quantized compressed sensing methods [7][8] is still continuous, so it has to be discretized when stored or transmitted in practical systems. Although additional ADCs can be employed for this task, they require extra cost. Moreover, time consuming iterative optimization in CS and 1-bit CS limits their efficiency. Furthermore, system could be more flexible in accuracy-efficiency trade-off if we have the freedom to adjust the number of bits for the measurement and recovery. Thus recent 1-bit HCS [9] considers to directly recover d-level quantization of signal rather than signal itself. In 1-bit HCS, the 1-bit measurements generate i.i.d. samples of a random variable, whose distribution's nearest neighbor among certain reference distributions indicates the quantized signal. Since quantization is irreversible with information loss, 1-bit HCS weakly relies on the sparse assumption.

This paper expands 1-bit HCS to k-bit measurement case on both algorithm and theory. Different from 1-bit HCS bridging the signal and 1-bit measurements [4] via distribution of sign for the product of two random projections, we investigate the distribution of the ratio between two random projections. It is a Cauchy distribution uniquely parameterized by the signal's one dimension, whilst its histogram can be estimated from k-level quantization of linear measurements. Interestingly, the Bernoulli distribution in 1-bit HCS is a special case of 2-bin histogram in k-bit HCS. In recovery, for each dimension of the signal, k-bit HCS searches the nearest neighbor of the estimated histogram among d predefined reference histograms, which corresponds to the d bins for signal quantization. This can be seen as a hypothesis test theoretically supported by concentration inequality for random functions. Compared to 1-bit HCS, the signal quantization in k-bit HCS is more flexible because the d bins can be chosen independently with the d reference histograms.

The primary contributions of k-bit HCS are 1) its direct recovery of quantized signal from quantized measurements largely saves the hardware cost in practical system; 2) its recovery of each dimension is an independent nearest neighbor search among d histograms, and thus is considerably more efficient than CS and 1-bit CS. Moreover, it is direct to further accelerate it by parallel computing and fast nearest neighbor search; 3) it offers trade-off among measurement number m, measurement resolution k and recovery resolution d, which lead to a more flexible system than 1-bit HCS; 4) it has sample complexity $\mathcal{O}(\log n)$ and weak dependence on sparsity.

II. Quantization in k-bit HCS

There are three types of quantization need to be defined in k-bit HCS, i.e., the k-level quantization to the linear measurement, the b-level quantization for reference histogram, and the d-level quantization to the signal. The goal of k-bit HCS is to recover the last from the first via comparison with the second. Different from 1-bit HCS deriving the signal quantization from the reference distributions, the three types of quantization can be independently selected in k-bit HCS.

In this paper, we simply choose the linear uniform quantization that partitions given range $[a, b]$ into k bins with average widths. Thus a real value in any of the k bins could be represented by a $\log_2 k$-bit binary code. In particular, a k-level scalar quantizer $Q : \mathbb{R} \rightarrow \mathbb{R}$ is formally defined as a composition of two mappings $Q = \beta \circ \alpha$, where $\alpha : \mathbb{R} \rightarrow \{1, 2, \dots, k\}$ is the lossy encoder and $\beta : \{1, 2, \dots, k\} \rightarrow \{(q_i + q_{i+1})/2\}_{i=1}^{k}$ is the midpoint decoder. Thus we have

$$\alpha(x) = i \quad \text{if} \quad x \in [q_i, q_{i+1}) \,. \tag{1}$$

We use different notations to represent the three types of quantization. In k-level quantization to the linear measurements $y = \Phi x$, we use $\{q_i\}_{i=1}^{k+1}$ to represent the associated $k + 1$ boundaries defining the k bins, and α_y to represent the corresponding encoder. In b-level quantization to r_j, the boundaries are $\{s_i\}_{i=1}^{b+1}$ and the encoder is α_r. In d-level quantization to the signal x, $\{p_i\}_{i=1}^{d+1}$ denote the boundaries and α_x denotes the encoder.

III. Distribution of Random Projection Ratio

We are interested in inferring the signal from statistics to the measurements. However, directly estimating distribution of the linear measurements (or their quantization) does not provide any useful information, because the randomness of measurements is caused by the sensing matrix rather than the signal. So we need to build a random variable whose i.i.d. samples can be computed from the measurements, and which obeys a distribution defined by the signal. Similar to 1-bit HCS that leverages the sign distribution of random projection product between $\langle x, \phi\rangle$ and $\langle e_j, \phi\rangle$ (e_j is for reference and ϕ is a random vector), we adopts the latter quantity as reference in k-bit HCS. However, rather than studying the sign of their product as a new random variable, we explore the ratio between the two random projections.

$$r_j = \frac{\langle x, \phi\rangle}{\langle e_j, \phi\rangle} = \frac{\sum_{i/j} \phi_i x_i}{\phi_j} + x_j. \tag{2}$$

Here ϕ is a row in the sensing matrix Φ that is sampled from the standard Gaussian ensemble in this paper. We restrict the normalized signal x on the unit sphere (we will show this restriction can be eliminated in journal version). Based on these settings, we have the distribution of r_j below.

Theorem 1. (**Random Projection Ratio**) *Let $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$ denote the unit sphere in $\mathbb{R}^n$ and $e_j \in \{0, 1\}^n$ which is 1 in entry j and 0 elsewhere, for signal $x \in \mathbb{S}^{n-1}$, and random vector $\phi \in \mathbb{S}^{n-1}$ uniformly distributed on $\mathbb{S}^{n-1}$ (i.e., each element of ϕ is firstly drawn i.i.d. from the standard Gaussian $\mathcal{N}(0, 1)$ and then ϕ is normalized as $\phi/\|\phi\|_2$), the ratio r_j between x's projection on ϕ and e_j's projection on ϕ is a Cauchy random variable whose distribution is parameterized by x_j, such that*

$$r_j = \frac{\langle x, \phi\rangle}{\langle e_j, \phi\rangle} \sim \mathfrak{C}(x_j, \sqrt{1 - x_j^2}). \tag{3}$$

Proof: The numerator $\sum_{i/j} \phi_i x_i$ of the first term in (2) follows a Gaussian distribution $\mathcal{N}(0, 1 - x_j^2)$, and the denominator ϕ_j simply obeys a standard Gaussian $\mathcal{N}(0, 1)$. We normalize the numerator to standard Gaussian and thus r_j can be written as a linear transformation of a ratio between two i.i.d Gaussian such that

$$r_j = \sqrt{1 - x_j^2}\,\frac{\sum_{i/j} \phi_i x_i/\sqrt{1 - x_j^2}}{\phi_j} + x_j. \tag{4}$$

Since the ratio between two independent standard Gaussian random variables follows the standard Cauchy distribution $\mathfrak{C}(0, 1)$, and according to the linear transformation property of Cauchy distribution, r_j obeys a Cauchy distribution of $\mathfrak{C}(x_j, \sqrt{1 - x_j^2})$. ■

Hence the random variable r_j completely meets our former requirements: its distribution is merely determined by x_j, the dimension j of the signal x, and its m i.i.d. samples can be computed from m linear measurements Φx.

Note the available information in k-bit HCS is the k-level quantization of Φx rather than the linear measurements themselves. Hence we can only obtain the number of r_j's samples in arbitrary pre-defined bin. This gives an estimation to the histogram of r_j. In particular, assume bin i is defined by its boundaries as $[s_i, s_{i+1})$, by the CDF of Cauchy distribution, the proportion of r_j's samples in bin i is an estimation to

$$\begin{aligned}&\Pr\left(r_j \in [s_i, s_{i+1}) \middle| x_j\right) = \\ &\frac{1}{\pi}\left[\arctan\left(\frac{s_{i+1} - x_j}{\sqrt{1 - x_j^2}}\right) - \arctan\left(\frac{s_i - x_j}{\sqrt{1 - x_j^2}}\right)\right].\end{aligned} \tag{5}$$

The above equation bridges the histogram of random variable r_j with the signal x_j. We will show the histogram can be estimated from quantization of the linear measurements according to (2). So (5) provides critical clue for quantization recovery by hypothesis tests.

Remark: It is not hard to verify that 1-bit HCS can be seem as a special case of k-bit HCS when 2-bin histogram ($b = 1$) is applied. Another critical observation of (5) is that it results in a 1-D maximum likelihood estimator of x_j. If a prior distribution encouraging sparsity (zero) is applied to x_j, we can further obtain a maximum a posteriori (MAP) estimator. A favorable property of these estimators is that the entries in x can be independently estimated in parallel manner. Since this paper focuses on digital system that prefers directly processing quantization of x rather than x, we will leave these to the journal version.

IV. Nearest Neighbor Search based Recovery

In this section, we introduce the recovery of signal quantization in k-bit HCS. According to the Cauchy distribution of the random projection ratio given in (2), each x_j on the real axis uniquely defines a Cauchy distribution $\mathfrak{C}(x_j, \sqrt{1-x_j^2})$ whose histogram can be estimated from the quantized measurements. In order to determine which bin x_j belongs to, we compare its associated histogram with some reference histograms associated with pre-defined bins. The nearest reference histogram then indicates the bin x_j lies in. In order to see this, we firstly give the following definitions.

Definition 1. (**Histogram**) *Given the cumulative density function (CDF) $F(\cdot)$ of random variable r, the histogram of r over b bins $[s_i, s_{i+1})_{i=1}^b$ is defined as a vector $h\left(F, \{s_i\}_{i=1}^{b+1}\right) \in \mathbb{R}^b$ such that*

$$h_i\left(F, \{s_i\}_{i=1}^{b+1}\right) = F(s_{i+1}) - F(s_i). \tag{6}$$

Given m i.i.d. samples $\{R^j\}_{j=1}^n$ of r, the resulting estimated histogram $\hat{h}\left(F, \{s_i\}_{i=1}^{b+1}\right) \in \mathbb{R}^b$ such that

$$\hat{h}_i\left(F, \{s_i\}_{i=1}^{b+1}\right) = \frac{1}{m}\sum_{j=1}^{m} I\left(R^j \in [s_i, s_{i+1})\right). \tag{7}$$

In this paper, we only consider the histogram of the Cauchy distribution $\mathfrak{C}(c, \sqrt{1-c^2})$, where c is the unique parameter of the corresponding CDF. In addition, the histogram bins in this paper are fixed as $[s_i, s_{i+1})_{i=1}^b$. Hence we simply write $h\left(F, \{s_i\}_{i=1}^{b+1}\right)$ and $\hat{h}\left(F, \{s_i\}_{i=1}^{b+1}\right)$ as h^c and $\hat{h}^c$, respectively.

Definition 2. (**Reference Histograms**) *For random variable r_j with arbitrary j, its histogram h^{x_j} (see (5)) is defined over b bins with boundaries $\{s_i\}_{i=1}^{b+1}$. Given the boundaries $\{p_i\}_{i=1}^{d+1}$ for d-level quantization of x_j, and the distance measure $\mathcal{D}(\cdot,\cdot)$ between two histograms, we define d reference histograms $\{h^{c_i}\}_{i=1}^d$ corresponding to the d bins such that*

$$\mathcal{D}\left(h^{c_i}, h^{p_{i+1}}\right) = \mathcal{D}\left(h^{c_{i+1}}, h^{p_{i+1}}\right), c_1 = p_1. \tag{8}$$

According to the above definition, the d reference histograms can be computed in sequence from h^{c_1} to h^{c_d}. Specifically, given h^{c_i}, c_{i+1} can be obtained by solving the first equation in (8) numerically or analytically. In this paper, we use numerical results, which is good enough for producing reliable recovery.

In recovery of k-bit HCS, the first step is to estimate the histogram of r_j from k-level quantization of m linear measurements $y = \Phi x$. According to (2), the m i.i.d. samples of r_j, which is denoted by $\{R_j^i\}_{i=1}^m$, can be computed from y as

$$R_j^i = \frac{y_i}{\Phi_{ij}}, \quad \text{for} \quad i = 1, 2, \ldots, m. \tag{9}$$

The k-level quantization of y_i is $\alpha_y(y_i)$, which is the index of the bin that y_i belongs to. Thus from quantization of y we can obtain the range of R_j^i

$$R_j^i \in \left[\frac{q_{\alpha_y(y_i)}}{\Phi_{ij}}, \frac{q_{\alpha_y(y_i)+1}}{\Phi_{ij}}\right), \quad \text{for} \quad i = 1, 2, \ldots, m. \tag{10}$$

Statistics of the above ranges for all $\{R_j^i\}_{i=1}^m$ produce an estimation to the histogram of r_j. In particular, we assume each sample R_j^i is uniformly distributed in its range, whose intersection part with a bin $[s_l, s_{l+1})$ indicates the probability that the sample lies in bin l. For each bin, we sum up such probabilities contributed by all the samples. Then the estimated histogram of r_j is obtained after normalization.

The above procedure is detailed in pseudo code below.

Algorithm 1 Histogram Estimator from Scaled Quantization

Input: Ranges $\left\{R_j^i \in [a_i, b_i)\right\}_{i=1}^m$ for m samples of r_j, bins $\left\{[s_i, s_{i+1})\right\}_{i=1}^b$ for histogram of r_j.
Output: Estimated histogram $\hat{h}^{x_j}$ for r_j.
Initialize: $\hat{h}^{x_j} := 0$ for all i.
for $i = 1$ to m **do**
  $u := \alpha_r(a_i);\ v := \alpha_r(b_i);\ c = b_i - a_i;$
  if $u == v$ **then**
    $\hat{h}_u^{x_j} := h_u^{x_j} + 1$; break;
  else
    $\hat{h}_u^{x_j} := h_u^{x_j} + (s_{u+1} - a_i)/c$; $\hat{h}_v^{x_j} := h_v^{x_j} + (b_i - s_v)/c$;
    if $v - u > 1$ **then**
      for $l = u + 1$ to $v - 1$ **do**
        $\hat{h}_l^{x_j} := h_l^{x_j} + (s_{l+1} - s_l)/c$;
      end for
    end if
  end if
end for
$\hat{h}^{x_j} := \hat{h}^{x_j}/m$.

Now we can estimate the bin for x_j from the d bins by comparing the distances between the estimated histogram $\hat{h}^{x_j}$ of r_j and the d reference histograms $\{h^{c_i}\}_{i=1}^d$. The index of the reference histogram with the smallest distance to $\hat{h}^{x_j}$ indicates the bin that x_j belongs to, a.k.a.

$$\hat{\alpha}_x(x_j) = \arg\min_i \mathcal{D}\left(h^{c_i}, \hat{h}^{x_j}\right). \tag{11}$$

In this paper, we set $\mathcal{D}(\cdot,\cdot)$ as ℓ_2 distance. However, many other distance metrics have also proved effectiveness in our study.

We summarize the recovery algorithm of k-bit HCS in Algorithm 2. Since the recovery for each dimension is independent, the procedure can be computed in parallel, and each thread merely includes computations of distances.

V. Recovery Error Bound

We investigate the recovery error bound of k-bit HCS in this section. Since former recovery in k-bit HCS is essentially n times of hypothesis tests, each of which aims at selecting the correct bin that x_j belongs to out of the d bins, we measure the recovery error by the probability that k-bit HCS fails to

Algorithm 2 k-bit HCS decoder

Input: Quantization $\{\alpha_y(y_i)\}_{i=1}^m$ to measurements $y = \Phi x$, sensing matrix Φ, reference histograms $\{h^{c_i}\}_{i=1}^d$, boundaries $\{q_i\}_{i=1}^{k+1}$ for signal quantization.
Output: Estimated d-level quantization $\{\hat{\alpha}_x(x_j)\}_{j=1}^n$.
for $j = 1$ to n **do**
  Compute ranges $\left\{R_j^i \in [a_i, b_i)\right\}_{i=1}^m$ according to (10);
  Estimate histogram of r_j as $\hat{h}^{x_j}$ by Algorithm 1;
  Estimate quantization of x_j as $\hat{\alpha}_x(x_j)$ by (11);
end for

choose the correct bin for x_j. Hence we need to study the upper bound for the following probability:

$$\begin{aligned}&\Pr\left(\hat{\alpha}_x(x_j) \neq \alpha_x(x_j)\right) = \\ &\Pr\left(\mathcal{D}\left(h^{c_{i:i\neq l}}, \hat{h}^{x_j}\right) \leq \mathcal{D}\left(h^{c_l}, \hat{h}^{x_j}\right)\middle| x_j \in [q_l, q_{l+1})\right).\end{aligned} \tag{12}$$

However, it is impractical to analyze the failure probability (12) for all cases $i \in \{i : i \neq l\}$, especially in general d is not fixed to a specific number. So we need to find an equivalent form of (12) that is analyzable.

Recall Theorem 1 k-bit HCS is based on the bijection between x_j and the Cauchy distribution of r_j. Thus a small distance between two points in the signal space leads to big similarity between their associated Cauchy distributions. This isometry makes the exact recovery from the Cauchy distribution possible. In k-bit HCS, the quantization of measurements produce an estimated histogram for the Cauchy distribution. Thus we can reduce the comparison with $d-1$ reference histograms in the right side of (12) to the comparison with the two nearest ones $h^{c_{l\pm1}}$. This leads to the following Lemma.

Lemma 1. (**Equivalence**) *The probability of estimating $\alpha_x(x_j)$ as any wrong bin in $\{i : i \neq l\}$ is upper bounded by the probability of mistakenly estimating $\alpha_x(x_j)$ as the left and right neighbors $l \pm 1$ of the correct bin l, such that*

$$\begin{aligned}&\Pr\left(\mathcal{D}\left(h^{c_{i:i\neq l}}, \hat{h}^{x_j}\right) < \mathcal{D}\left(h^{c_l}, \hat{h}^{x_j}\right)\middle| x_j \in [q_l, q_{l+1})\right) \leq \\ &\max_{\pm}\left\{\Pr\left(\mathcal{D}\left(h^{c_{l\pm1}}, \hat{h}^{x_j}\right) < \mathcal{D}\left(h^{c_l}, \hat{h}^{x_j}\right)\middle| x_j \in [q_l, q_{l+1})\right)\right\}\end{aligned}$$

By Lemma 1, the d reference histograms involved in failure probability (12) can be reduced to three, i.e., $\{h^{c_{l\pm1}}, h^{c_l}\}$. However, their distances to the unique random vector $\hat{h}^{x_j}$ in (12) are unclear. So we need to firstly study the concentration inequality of $\hat{h}^{x_j}$. Based on Boucherons's inequality and Talagrand's inequality [10], which are concentration inequalities for random functions, we have the following concentration inequality for $\hat{h}^{x_j}$.

Theorem 2. (**Histogram estimation**) *Let r_j be a random variable that lies in bin i with probability $h_i^{x_j}$ for $i = 1, \ldots, b$. If $\hat{h}^{x_j}$ is the estimated histogram of r_j from m i.i.d. samples, then for $t > 0$*

$$\Pr\left(\left\|\hat{h}^{x_j} - h^{x_j}\right\|_2 \geq \frac{1}{\sqrt{m}} + \frac{t}{\sqrt{m}}\right) \leq e^{\frac{-t^2}{4}}. \tag{13}$$

$$\Pr\left(\left\|\hat{h}^{x_j} - h^{x_j}\right\|_2 \geq \frac{1}{\sqrt{m}} + t\sqrt{\frac{1}{m} + \frac{2\sqrt{2}}{m^{2/3}} + \frac{2t^2}{3m}}\right) \leq e^{\frac{-t^2}{2}} \tag{14}$$

By using Lemma 1 and Theorem 2, we can now obtain the following conclusion of the failure probability in (12).

Theorem 3. (**Quantized recovery bound**) *The failure probability (12) of k-bit HCS is upper bounded as*

$$\Pr\left(\hat{\alpha}_x(x_j) \neq \alpha_x(x_j)\right) \leq \max_{\pm}\left\{e^{-(\sqrt{m}t_{\pm}(x_j)-1)^2}\right\}, \tag{15}$$

$$t_{\pm}(x_j) = \mathcal{D}\left(h^{c_{l\pm1}}, h^{x_j}\right) - \mathcal{D}\left(h^{c_l}, h^{x_j}\right). \tag{16}$$

The minimum amount of k-bit measurements that ensures the successful quantized recovery in k-bit HCS is then directly obtained from Theorem 3.

Corollary 1. (**Amount of measurements**) *k-bit HCS successfully reconstructs the n-dimensional signal x with probability exceeding $1-\eta$ ($0 \leq \eta \leq 1$) if the number of measurements*

$$m \geq \frac{1}{\min_{j,\pm} t_{\pm}^2(x_j)}\left(\sqrt{\log\frac{n}{\eta}} + 1\right)^2. \tag{17}$$

Remark: A trade-off in k-bit HCS is between the measurement amount m and the level d (or bits) for recovery. According to the definition of $t_{\pm}(x_j)$ in (16), both the upper bound for failure probability in (15) and the least amount of measurements in (17) will become smaller if the distances between x_j and its two boundaries q_l and q_{l+1} increase. Therefore, when we increase the level d (resolution), x_j will be closer to its boundaries, which leads to more measurements for achieving a successful recovery.

Another trade-off is between measurement amount m and level k (or bits) for measurement quantization. In particular, since $t_{\pm}(x_j)$ becomes lager with the increasing of k, the required measurement amount will reduce in this case according to (17). Comparing to 1-bit HCS, k-bit HCS can freely choose to either increase the number or improve the resolution of measurements for achieving the same accuracy.

VI. Empirical Study

This section gives an empirical study of k-bit HCS via comparison with "1-bit CS+quantizer" and 1-bit HCS, and error trade-off between k and d. We use average quantized recovery error $\sum_{j=1}^n |\hat{\alpha}_x(x_j) - \alpha_x(x_j)|/nd$ to measure the recovery accuracy and CPU seconds to measure the efficiency. We simply set $b = 50$. In each trial, we draw a normalized Gaussian random matrix $\Phi \in \mathbb{R}^{m\times n}$ as sensing matrix and a signal of length n and cardinality K, whose K nonzero entries are drawn uniformly at random on the unit ℓ_2 sphere.

A. Phase diagram for error and speed

We first plot the phase diagrams of 1-bit CS, 1-bit HCS and k-bit HCS for quantized recovery error and time in noisy case, based on 10^5 trials of "BIHT[4](1-bit CS)+quantizer", 1-bit HCS and k-bit HCS algorithms. In particular, given fixed n, d (and k for k-bit HCS), we uniformly choose 50 K/n

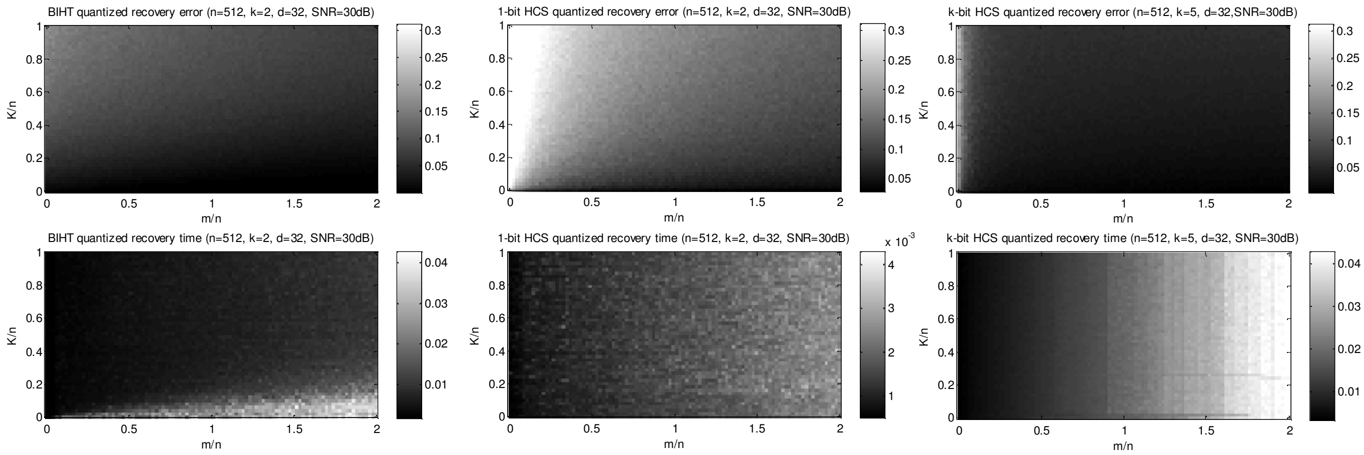

Fig. 1. Phase diagrams of quantized recover error and time costs of "1-bit CS+quantizer", 1-bit HCS and k-bit HCS in noisy case.

values between 0 and 1, and 100 m/n values between 0 and 2. For each $\{K/n, m/n\}$ pair, we conduct 20 trials of the three methods. The average quantized recovery errors and average CPU seconds of the three methods on overall 5000 $\{K/n, m/n\}$ pairs are shown in Figure 1.

It shows that 1) In all methods, error increases when augmenting m or reducing K, and it shows phase transition, but 1-bit HCS and k-bit HCS has weaker dependence on K than 1-bit CS; 2) 1-bit HCS and k-bit HCS is significantly efficient due to their hypothesis test based recovery, while k-bit HCS's speed weakly depends on sparsity; 3) k-bit HCS achieves much better accuracy than 1-bit HCS even with very coarse ($k = 5$) quantized measurements (slightly finer than 1-bit measurements of $k = 2$); 4) comparing to 1-bit CS, k-bit HCS takes slightly finer but much fewer measurements to reach the same accuracy with competitive speed. Considering the merits, this trade-off is quite advantageous.

B. Trade-off between k and d

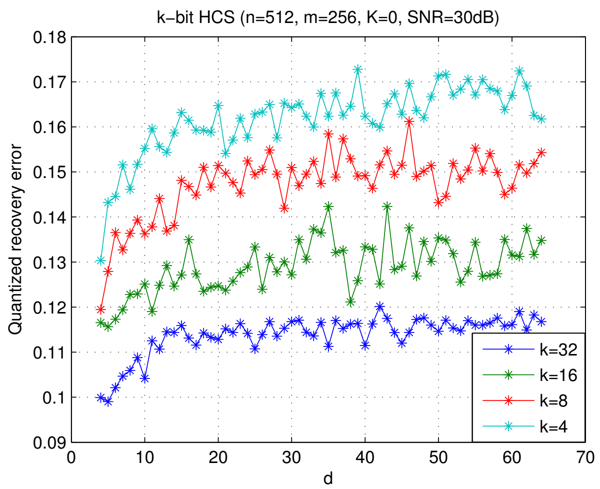

Fig. 2. Measurement resolution k vs. recovery resolution d trade-off.

It is a new and favorable property of k-bit HCS that the resolutions of both measurements and recovery can be freely adjusted according to the demands for computational costs and system accuracy. Figure 2 shows the effect of trade-off between k and d on 2440 trials for recovering noisy signals of fixed n, K and k and signal-to-noise ratio (SNR). For each $\{k, d\}$ pair, the quantized recovery errors of 10 i.i.d. trails are averaged and shown in Figure 2, where each curve corresponds to a fixed k. On all of the four curves, the error firstly increases and then becomes stable along with increasing d. This verifies the robustness and the theory in Section V that the error bound will become worse if d increases. In addition, the curve with larger k is always beneath the curve with smaller k. This fact is also consistent with our analysis that large k can result in small failure probability.

ACKNOWLEDGEMENTS

We would like to thank all the anonymous reviewers on improving this paper. This work is supported by Australian Research Council Discovery Project ARC DP-120103730.

REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
[2] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.
[3] P. T. Boufounos and R. G. Baraniuk, "One-bit compressive sensing," in *CISS*, 2008.
[4] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *IEEE Transactions on Information Theory*, 2012.
[5] M. Yan, Y. Yang, and S. Osher, "Robust 1-bit compressive sensing using adaptive outlier pursuit," *IEEE Transactions on Signal Processing*, 2012.
[6] J. N. Laska, Z. Wen, W. Yin, and R. G. Baraniuk, "Trust, but verify: Fast and accurate signal recovery from 1-bit compressive measurements," *IEEE Transactions on Signal Processing*, 2010.
[7] A. Zymnis, S. Boyd, and E. Candès, "Compressed sensing with quantized measurements," *IEEE Signal Processing Letters*, vol. 17, no. 2, pp. 149–152, 2010.
[8] U. Kamilov, V. Goyal, and S. Rangan, "Message-passing de-quantization with applications to compressed sensing," *IEEE Transactions on Signal Processing*, 2012.
[9] T. Zhou and D. Tao, "1-bit hamming compressed sensing," in *ISIT*, 2012.
[10] V. I. Norkin and R. J.-B. Wets, "Law of small numbers as concentration inequalities for sums of independent random setsand random set valued mappings," in *STOPROG*, 2012.