Efficient Compressive Phase Retrieval with Constrained Sensing Vectors

Authored by:

Sohail Bahmani
Justin Romberg

Abstract

We propose a robust and efficient approach to the problem of compressive phase retrieval in which the goal is to reconstruct a sparse vector from the magnitude of a number of its linear measurements. The proposed framework relies on constrained sensing vectors and a two-stage reconstruction method that consists of two standard convex programs that are solved sequentially.In recent years, various methods are proposed for compressive phase retrieval, but they have suboptimal sample complexity or lack robustness guarantees. The main obstacle has been that there is no straightforward convex relaxations for the type of structure in the target. Given a set of underdetermined measurements, there is a standard framework for recovering a sparse matrix, and a standard framework for recovering a low-rank matrix. However, a general, efficient method for recovering a jointly sparse and low-rank matrix has remained elusive.Deviating from the models with generic measurements, in this paper we show that if the sensing vectors are chosen at random from an incoherent subspace, then the low-rank and sparse structures of the target signal can be effectively decoupled. We show that a recovery algorithm that consists of a low-rank recovery stage followed by a sparse recovery stage will produce an accurate estimate of the target when the number of measurements is \$mathsf{O}(k,logfrac{d}{k})\$, where \$k\$ and \$d\$ denote the sparsity level and the dimension of the input signal. We also evaluate the algorithm through numerical simulation.

1 Paper Body

Problem setting

The problem of Compressive Phase Retrieval (CPR) is generally stated as the problem of estimating a k-sparse vector x? ? Rd from noisy measurements of the form 2

yi = —hai , x? i— + zi

(1)

for i = 1, 2, . . . , n, where ai is the sensing vector and zi denotes the additive noise. In this paper, we study the CPR problem with specific sensing vectors ai of the form ai = ? T wi ,

(2)

where ? ? Rm?d and wi ? Rm are known. In words, the measurement vectors live in a fixed low-dimensional subspace (i.e, the row space of ?). These types of measurements can be applied in imaging systems that have control over how the scene is illuminated; examples include systems that use structured illumination with a spatial light modulator or a scattering medium [1, 2]. 1

By a standard lifting of the signal x? to X ? = x? x?T , the quadratic measurements (1) can be expressed as E D

? ? + zi . (3) yi = ai aT + zi = ? T wi wT i ,X i ?, X With the linear operator W and A defined as

n W :B 7? wi wT i , B i=1

and

A : X 7? W ? X? T ,

we can write the measurements compactly as y = A (X ?) + z. Our goal is to estimate the sparse, rank-one, and positive semidefinite matrix X ? from the measurements (3), which also solves the CPR problem and provides an estimate for the sparse signal x? up to the inevitable global phase ambiguity. Assumptions We make the following assumptions throughout the paper. A1. The vectors wi are independent and have the standard Gaussian distribution on Rm : wi ? N (0, I) . A2. The matrix ? is a restricted isometry matrix for 2k-sparse vectors and for a constant ?2k ? [0, 1]. Namely, it obeys 2

(1 ? ?2k) kxk2 ? k? xk22 ? (1 + ?2k) kxk22 ,

(4)

for all 2k-sparse vectors x ? Rd . A3. The noise vector z is bounded as kzk2 ? ?. As will be seen in Theorem 1 and its proof below, the Gaussian distribution imposed by the assumption A1 will be used merely to guarantee successful estimation of a rank-one matrix through trace norm minimization. However, other distributions (e.g., uniform distribution on the unit sphere) can also be used to obtain similar guarantees. Furthermore, the restricted isometry condition imposed by the assumption A2 is not critical and can be replaced by weaker assumptions. However, the guarantees obtained under these weaker assumptions usually require more intricate derivations, provide weaker noise robustness, and often do not hold uniformly for all potential target signals. Therefore, to keep the exposition simple and straightforward we assume (4) which is known to hold (with high probability) for various ensembles of random matrices (e.g., Gaussian, Rademacher, partial Fourier, etc). Because in many scenarios we have the flexibility of selecting ? , the assumption (4) is realistic as well. Notation Let us first set the notation used throughout the paper. Matrices and vectors are denoted by bold capital and small letters, respectively. The set of positive integers less than or equal to n is denoted by [n]. The notation f = O (g) is used when f = cg for some absolute constant c ¿ 0. For any matrix M , the Frobenius norm, the nuclear norm, the entrywise ‘1 -norm, and the largest entrywise absolute value of the entries are denoted by kM kF , kM k? , kM k1

, and kM k? , respectively. To indicate that a matrix M is positive semidefinite we write M ¡ 0. 1.2

Contributions

The main challenge in the CPR problem in its general formulation is to design an accurate estimator that has optimal sample complexity and computationally tractable. In this paper we address this challenge in the special setting where the sensing vectors can be factored as (2). Namely, we propose an algorithm that

? provably produces an accurate estimate of the lifted target X ? from only n = O k log kd measurements, and ? can be computed in polynomial time through efficient convex optimization methods. 2

1.3

Related work

Several papers including [3, 4, 5, 6, 7] have already studied the application of convex programming for (non-sparse) phase retrieval (PR) in various settings and have established estimation accuracy through different mathematical techniques. These phase retrieval methods attain nearly optimal sample complexities that scales with the dimension of the target signal up to a constant factor [4, 5, 6] or at most a logarithmic factor [3]. However, to the best of our knowledge, the exiting methods for CPR either lack accuracy and robustness guarantees or have suboptimal sample complexities. The problem of recovering a sparse signal from the magnitude of its subsampled Fourier transforms is cast in [8] as an ‘1 -minimization with non-convex constraints. While [8] shows that a sufficient number of measurements would grow quadratically in k (i.e., the sparsity of the signal), the numerical simulations suggest that the non-convex method successfully estimates the sparse signal with only about k log kd measurements. Another non-convex approach to CPR is considered in [9] which poses the problem as finding a k-sparse vector that minimizes the residual error that takes a quartic form. A local search algorithm called GESPAR [10] is then applied to (approximate) the solution to the formulated sparsity-constrained optimization. This approach is shown to be effective through simulations, but it also lacks global convergence or statistical accuracy guarantees. An alternating minimization method for both PR and CPR is studied in [11]. This method is appealing in large scale problems because of computationally inexpensive iterations. More importantly, [11] proposes a specific initialization using which the alternating minimization method is shown to converge linearly in noise-free PR and CPR. However, the number of measurements required to establish this convergence is effectively quadratic in k. In [12] and [13] the ‘1 -regularized form of the trace minimization argmin X¡0

trace (X) + ? kXk1 (5)

subject to A (X) = y is proposed for the CPR problem. The guarantees of [13] are based on the restricted isometry propn erty of the sensing operator X 7? [hai a?i , Xi]i=1 for sparse matrices. In [12], however, the analysis is based on construction of a dual certificate through an adaptation of the golfing scheme [14]. Assuming standard Gaussian sensing vectors ai and with appropriate choice of the regularization parameter ?, it is shown in [12] that (5) solves the CPR when n = O k 2 log d . Furthermore, this method fails to recover the target sparse and rank-one matrix if n is dominated by k 2 . Estimation of simultaneously structured matrices through convex relaxations similar to (5) is also studied in [15] where it is shown that these methods do not attain optimal sample complexity. More recently, assuming that the sparse target has a Bernoulli-Gaussian distribution, a generalized approximate message passing framework is proposed in [16] to solve the CPR problem. Performance of this method is evaluated through numerical simulations for standard Gaussian sensing matrices which show the empirical phase transition for successful estimation occurs at n = O k log kd and also the algorithms can have a significantly lower runtime compared to some of the competing algorithms including GESPAR [10] and CPRL [13]. The PhaseCode algorithm is proposed in [17] to solve the CPR problem with sensing vectors designed using sparse graphs and techniques adapted from coding theory. Although PhaseCode is shown to achieve the optimal sample complexity, it lacks robustness guarantees. While preparing the final version of the current paper, we became aware of [18] which has independently proposed an approach similar to ours to address the CPR problem.

2 2.1

Main Results Algorithm

We propose a two-stage algorithm outlined in Algorithm 1. Each stage of the algorithm is a convex program for which various efficient numerical solvers exists. In the first stage we solve (6) to obtain b which is an estimator of the matrix a low-rank matrix B B? = ? X ?? T. 3

b is used in the second stage of the algorithm as the measurements for a sparse estimation Then B expressed by (7). The constraint of (7) depends on an absolute constant C ¿ 0 that should be sufficiently large. Algorithm 1: input : the measurements y, the operator W, and the matrix ? c output: the estimate X 1 Low-rank estimation stage: b ? argmin B trace (B) B¡0

subject to 2

(6) kW (B) ? yk2 ? ?

Sparse estimation stage: c ? argmin X X

subject to

kXk1

C?

b

??

? X? T ? B n F

(7)

Post-processing. The result of the low-rank estimation stage (6) is generally not rank-one. Simic that is k ? k-sparse (i.e., it has larly, the sparse estimation stage does not necessarily produce a X at most k nonzero rows and columns) and rank-one. In fact, since we have not imposed the posic is not even guaranteed to be tive semidefiniteness constraint (i.e., X ¡ 0) in (7), the estimate X positive semidefinite (PSD). However, we can enforce the rank-one or the sparsity structure in postprocessing steps simply by projecting the produced estimate on the set of rank-one or k ? k-sparse c onto the desired sets at PSD matrices. The simple but important observation is that projecting X most doubles the estimation error. This fact is shown by Lemma 2 in Section 4 in a general setting. Alternatives. There are alternative convex relaxations for the low-rank estimation and the sparse estimation stages of Algorithm (1). For example, (6) can be replaced by its regularized least squares analog b ? argmin 1 kW (B) ? yk2 + ? kBk , B 2 ? 2 B¡0 for an appropriate choice of the regularization parameter ?. Similarly, instead of (7) we can use an ‘1 -regularized least squares. Furthermore, to perform the low-rank estimation and the sparse estimation we can use non-convex greedy type algorithms that typically have lower computational costs. For example, the low-rank estimation stage can be performed via the Wirtinger flow method proposed in [19]. Furthermore, various greedy compressive sensing algorithms such as the Iterative Hard Thresholding [20] and CoSaMP [21] can be used to solve the desired sparse estimation. To guarantee the accuracy of these compressive sensing algorithms, however, we might need to adjust the assumption A2 to have the restricted isometry property for ck-sparse vectors with c being some small positive integer. 2.2

Accuracy guarantees

The following theorem shows that any solution of the proposed algorithm is an accurate estimator of X ? . Theorem 1. Suppose that the assumptions A1, A2, and A3 hold with a sufficiently small constant ?2k . Then, there exist positive absolute constants C1 , C2 , and C3 such that if n ? C1 m, (8) c of the Algorithm 1 obeys then any estimate X

C2 ?

c

X ? X ? ? ? , n F 4

for all rank-one and k ? k-sparse matrices X ? ¡ 0 with probability exceeding 1 ? e?C3 n . The proof of Theorem 1 is straightforward and is provided in Section 4. The main idea is first to show the low-rank estimation stage produces an accurate estimate of B ? . Because this stage can be viewed as a standard phase retrieval through lifting, we can simply use accuracy guarantees that are already established in the literature (e.g., [3, 6, 5]). In particular, we use [5, Theorem 2] which established an error bound that holds uniformly for all valid B ? . Thus we can ensure that X ? is feasible in the sparse estimation stage. Then the accuracy of the sparse estimation stage can also be established by a simple adaptation of the analyses based on the restricted isometry property such as [22]. The dependence of n (i.e., the number of measurements) and k (i.e., the sparsity of the signal) is not explicit in Theorem 1. This dependence is absorbed in m which must be sufficiently large for Assumption A2 to hold. Considering a Gaussian matrix ? , the following corollary gives a concrete example where the dependence of non k through m is exposed. Corollary 1. Suppose that the assumptions of Theorem 1 including (8) hold. Furthermore, suppose 1 that ? is a Gaussian matrix with iid N 0, m entries and d m ? c1 k log , k c for some absolute constant c1 ¿ 0. Then any estimate X produced by Algorithm 1 obeys

C2 ?

c

X ? X ? ? ? , n F

(9)

for all rank-one and k ? k-sparse matrices X ? ¡ 0 with probability exceeding 1 ? 3e?c2 m for some constant c2 ¿ 0.

1 Proof. It is well-known that if ? has iid N 0, m and we have (9) then (4) holds with high probability. For example, using a standard covering argument and a union bound [23] shows that if (9) holds for a sufficiently large constant c1 ¿ 0 then we have (4) for a sufficiently small constant ?2k with probability exceeding 1 ? 2e?cm for some constant c ¿ 0 that depends only on ?2k . Therefore, Theorem 1 yields the desired result which holds with probability exceeding 1 ? 2e?cm ? e?C3 n ? 1 ? 3e?c2 m for some constant c2 ¿ 0 depending only on ?2k .

3

Numerical Experiments

We evaluated the performance of Algorithm 1 through some numerical simulations. The low-rank estimation stage and the sparse estimation stage are implemented using the TFOCS package [24]. We considered the target k-sparse signal x? to be in R256 (i.e., d = 256). The support set of of the target signal is selected uniformly at random and the entry values on this support are drawn independently from N (0, 1). The noise vector z is also Gaussian with independent N 0, 10?4 . The operator W and the matrix ? are drawn from some Gaussian ensembles as described in Corollary ? b kX?X k 1. We measured the relative error kX ? k F of achieved by the compared methods over 100 trials F with sparsity level (i.e., k) varying in the set {2, 4, 6, . . . , 20}. In the first experiment, for each value of k, the pair (m, n) that determines the size W and ? are selected from {(8k, 24k) , (8k, 32k) , (12k, 36k) , (12k, 48k) , (16k, 48k)}. Figure 1 illustrates the 0.9 quantiles of the relative error versus k for the mentioned choices of m. In the second experiment we compared the performance of Algorithm 1 to the convex optimization methods that do not exploit the structure of the sensing vectors. The setup for this experimentis the same as in the first experiment except for the size of W and ? ; we chose m = 2k 1 + log kd and n = 3m, where dre denotes the smallest integer greater than r. Figure 2 illustrates the 0.9 quantiles of the measured relative errors for Algorithm 1, the semidefinite program (5) for ? = 0 and ? = 0.2, and the ‘1 -minimization argmin kXk1 X

subject to A (X) = y, 5

Figure 1: The empirical 0.9 quantile of the relative estimation error vs. sparsity for various choices of m and n with d = 256.

Figure 2: The empirical 0.9 quantile of the relative estimation error vs. sparsity for Algorithm 1

and different trace- and/or ‘1 - minimization methods with d = 256, m = 2k 1 + log kd , and n = 3m.

which are denoted by 2-stage, SDP, SDP+‘1 , and ‘1 , respectively. The SDP-based method did not perform significantly different for other values of ? in our complementary simulations. The relative error for each trial is also overlaid in Figure 2 visualize its empirical distribution. The empirical performance of the algorithms are in agreement with the theoretical results. Namely in a regime where n = O (m) = O k log kd , Algorithm 1 can produce accurate estimates whereas while the other approaches fail in this regime. The SDP and SDP+‘1 show nearly identical performance. The ‘1 -minimization, however, competes with Algorithm 1 for small values of k. This observation can be

explained intuitively by the fact that the ‘1 -minimization succeeds with n = O k 2 measurements

which for small values of k can be sufficiently close to the considered n = 3 2k 1 + log kd measurements. 6

4

Proofs

Proof of Theorem 1. Clearly, B ? = ? X ? ? T is feasible in 6 because of A3. Therefore, we can b of (6) accurately estimates B ? using existing results on nuclear-norm show that any solution B minimization. In particular, we can invoke [5, Theorem 2 and Section 4.3] which guarantees that for some positive absolute constants C1 , C20 , and C3 if (8) holds then

C0 ?

b

B ? B ? ? ?2 , n F holds for all valid B ? , thereby for all valid X ? , with probability exceeding 1 ? e?C3 n . Therefore, with C = C20 , the target matrix X ? would be feasible in (7). Now, it suffices to show that the sparse estimation stage can produce an accurate estimate of X ? . Recall that by A2, the matrix ? is restricted isometry for 2k-sparse vectors. Let X be a matrix that is 2k ? 2k-sparse, i.e., a matrix whose entries except for some 2k ? 2k submatrix are all zeros. Applying (4) to the columns of X and adding the inequalities yield 2

(1 ? ?2k) kXkF ? k? Xk2F ? (1 + ?2k) kXk2F . T

(10)

T

Because the columns of X ? are also 2k-sparse we can repeat the same argument and obtain

2

2

2

(1 ? ?2k) X T ? T ? (11)

? X T ? T ? (1 + ?2k) X T ? T . F F

F

Using the facts that X T ? T = k? XkF and ? X T ? T = ? X? T , the inequalities (10) F F F and (11) imply that

2 2 2

2 2 (1 ? ?2k) kXkF ? (12)

? X? T ? (1 + ?2k) kXkF . F

The proof proceeds with an adaptation of the arguments used to prove accuracy of ‘1 -minimization c ?X ? . in compressive sensing based on the restricted isometry property (see, e.g., [22]). Let E = X ? Furthermore, let S0 ? [d] ? [d] denote the support set of the k ? k-sparse target X . Define E 0 to be a d ?

d matrix that is identical to E over the index set S0 and zero elsewhere. By optimality of c and feasibility of X ? in (7) we have X

c kX ? k1 ? X

= kX ? + E ? E 0 + E 0 k1 ? kX ? + E ? E 0 k1 ? kE 0 k1 1

= kX ? k1 + kE ? E 0 k1 ? kE 0 k1 , where the last line follows from the fact that X ? and E ? E 0 have disjoint supports. Thus, we have kE ? E 0 k1 ? kE 0 k1 ? k kE 0 kF . (13) Now consider a decomposition of E ? E 0 as the sum J X E ? E0 = Ej , (14) j=1

such that for j ? 0 the d ? d matrices E j have disjoint support sets of size k ? k except perhaps for the last few matrices that might have smaller supports. More importantly, the partitioning matrices E j are chosen to have a decreasing Frobenius norm (i.e., kE j kF ? kE j+1 kF) for j ? 1. We have

X

J J X

J

1 1X

? kE j?1 k1 ? kE ? E 0 k1 ? kE 0 kF ? kE 0 + E 1 kF , (15) E kE k ? j j F

k j=2 k

j=2 j=2 F

where the chain of inequalities follow from the triangle inequality, the fact that kE j k? ? 1 k2 kE j?1 k1 by construction, the fact that the matrices E j have disjoint support and satisfy (14), the bound (13), and the fact that E 0 and E 1 are orthogonal. Furthermore, we have ? ? * + J

2 X

T T T Ej ? ?

? (E 0 + E 1) ? = ? (E 0 + E 1) ? , ? ?E ? F

j=2 1 X J D

E X

? ? (E 0 + E 1) ? T

? E? T + ? Ei? T, ? Ej ? T , F

F

i=0 j=2

(16) 7

where the first term is obtained by the Cauchy-Schwarz inequality and the summation is obtained by c ? X ? by definition, the triangle inequality and the fact that the triangle inequality. Because E = X

c T b

? b c are feasible in (7) imply that ? B + ? X ? ? T ? B X and X

?

? E? T ? ? X? F

F

F

2C? ? . n

Furthermore, Lemma 1 below which is adapted from [22, Lemma 2.1] guarantees that for D E i ? {0, 1} and j ? 2 we have ? E i ? T , ? E j ? T ? 2?2k kE i kF kE j kF . Therefore, we obtain

2

2 2 (1 ? ?2k) kE 0 + E 1 kF ? ? (E 0 + E 1) ? T

F

1 X J

X 2C?

? ? ? (E 0 + E 1) ? T + 2?2k kE i kF kE j kF n F i=0 j=2 1

J

XX 2C? ? ? (1 + ?2k) kE 0 + E 1 kF + 2?2k kE i kF kE j kF n i=0 j=2 2C? ? ? (1 + ?2k) kE 0 + E 1 kF + 2?2k (kE 0 kF + kE 1 kF) kE 0 + E 1 kF n

? 2C? ? kE 0 + E 1 kF ? (1 + ?2k) + 2 2?2k kE 0 + E 1 kF n where the chain of inequalities follow from the lower bound in (12),?the bound (16), the upper bound in (12), the bound (15), and the fact that kE 0 kF + kE 1 kF ? 2 kE 0 + E 1 kF . If ?2k ¡ p ? ? ? 2 1 + 2 1 ? 1 + 2 ? 0.216, then we have ? := (1 ? ?2k) ? 2 2?2k ¿ 0 and thus kE 0 + E 1 kF ?

2C (1 + ?2k) ? ? . ? n

Adding the above inequality to (13) and applying the triangle then yields the desired result. Lemma 1. Let ? be a matrix obeying (4). Then for any pair of k ? k-sparse matrices X and X 0 with disjoint supports we have D E

? X? T , ? X 0 ? T ? 2?2k kXkF X 0 F . Proof. Suppose that X and X 0 have unit Frobenius norm. Using the identity

2

2 D E

? X? T , ? X 0 ? T = 41 ? X + X 0 ? T ? ? X ? X 0 ? T and the fact that F

F

X and X 0 have disjoint supports, it follows from (12) that 2 2 D E (1 + ?)2 ? (1 ? ?)2 (1 ? ?2k) ? (1 + ?2k) 2k 2k ? ? X? T , ? X 0 ? T ? = 2?2k . ?2?2k = 2 2 The general result follows immediately as the desired inequality is homogeneous in the Frobenius norms of X and X 0 . Lemma 2 (Projected estimator). Let S be a closed nonempty subset of a normed vector space b ? V, not necessarily in S, that obeys (V, k?k). Suppose that for v ? ? S we have an estimator v e denotes a projection of v b onto S, then we have ke kb v ? v ? k ? . If v v ? v ? k ? 2. e ? argminv?S kv ? v bk . Therefore, because v ? ? S we have Proof. By definition v bk ? 2 kb ke v ? v ? k ? kb v ? v ? k + ke v?v v ? v ? k ? 2.

Acknowledgements This work was supported by ONR grant N00014-11-1-0459, and NSF grants CCF-1415498 and CCF-1422540. 8

2 References

[1] Jacopo Bertolotti, Elbert G. van Putten, Christian Blum, Ad Lagendijk, Willem L. Vos, and Allard P. Mosk. Non-invasive imaging through opaque

scattering layers. Nature, 491(7423):232?234, Nov. 2012. [2] Antoine Liutkus, David Martina, S?bastien Popoff, Gilles Chardon, Ori Katz, Geoffroy Lerosey, Sylvain Gigan, Laurent Daudet, and Igor Carron. Imaging with nature: Compressive imaging using a multiply scattering medium. Scientific Reports, volume 4, article no. 5552, Jul. 2014. [3] Emmanuel J. Cand?s, Thomas Strohmer, and Vladislav Voroninski. PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming. Communications on Pure and Applied Mathematics, 66(8):1241?1274, 2013. [4] Emmanuel J. Cand?s and Xiaodong Li. Solving quadratic equations via PhaseLift when there are about as many equations as unknowns. Foundations of Computational Mathematics, 14(5):1017?1026, 2014. [5] R. Kueng, H. Rauhut, and U. Terstiege. Low rank matrix recovery from rank one measurements. Applied and Computational Harmonic Analysis, 2015. In press. Preprint arXiv:1410.6913 [cs.IT]. [6] Joel A. Tropp. Convex recovery of a structured signal from independent random linear measurements. Preprint arXiv:1405.1102 [cs.IT], 2014. [7] Ir?ne Waldspurger, Alexandre d?Aspremont, and St?phane Mallat. Phase recovery, MaxCut and complex semidefinite programming. Mathematical Programming, 149(1-2):47?81, 2015. [8] Matthew L. Moravec, Justin K. Romberg, and Richard G. Baraniuk. Compressive phase retrieval. In Proceedings of SPIE Wavelets XII, volume 6701, pages 670120 1?11, 2007. [9] Yoav Shechtman, Yonina C. Eldar, Alexander Szameit, and Mordechai Segev. Sparsity based subwavelength imaging with partially incoherent light via quadratic compressed sensing. Optics Express, 19(16):14807?14822, Aug. 2011. [10] Yoav Shechtman, Amir Beck, and Yonina C. Eldar. GESPAR: Efficient phase retrieval of sparse signals. Signal Processing, IEEE Transactions on, 62(4):928?938, Feb. 2014. [11] Praneeth Netrapalli, Prateek Jain, and Sujay Sanghavi. Phase retrieval using alternating minimization. In Advances in Neural Information Processing Systems 26 (NIPS 2013), pages 2796?2804, 2013. [12] Xiaodong Li and Vladislav Voroninski. Sparse signal recovery from quadratic measurements via convex programming. SIAM Journal on Mathematical Analysis, 45(5):3019?3033, 2013. [13] Henrik Ohlsson, Allen Yang, Roy Dong, and Shankar Sastry. CPRL?an extension of compressive sensing to the phase retrieval problem. In Advances in Neural Information Processing Systems 25 (NIPS 2012), pages 1367?1375, 2012. [14] David Gross. Recovering low-rank matrices from few coefficients in any basis. Information Theory, IEEE Transactions on, 57(3):1548?1566, Mar. 2011. [15] Samet Oymak, Amin Jalali, Maryam Fazel, Yonina Eldar, and Babak Hassibi. Simultaneously structured models with application to sparse and low-rank matrices. Information Theory, IEEE Transactions on, 61(5):2886?2908, 2015. [16] P. Schniter and S. Rangan. Compressive phase retrieval via generalized approximate message passing. Signal Processing, IEEE Transactions on, 63(4):1043?1055, February 2015. [17] Ramtin Pedarsani, Kangwook Lee, and Kannan Ramchandran. Phasecode: Fast and efficient compressive phase retrieval based on sparse-graph codes. In Communication, Control, and Computing (Allerton), 52nd Annual Allerton Conference on, pages 842?849, Sep. 2014. Extended preprint arXiv:1408.0034 [cs.IT]. [18] Mark Iwen, Aditya Viswanathan, and Yang Wang. Robust sparse phase retrieval made easy. Applied and Computational Harmonic Analysis, 2015. In press. Preprint arXiv:1410.5295 [math.NA]. [19] Emmanuel J. Cand?s, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval via Wirtinger flow: Theory and algorithms. Information Theory, IEEE Transactions on, 61(4):1985?2007, Apr. 2015. [20] Thomas Blumensath and Mike E. Davies. Iterative hard thresholding for compressed sensing. Applied and Computational Harmonic Analysis, 27(3):265?274, 2009. [21] Deanna Needell and Joel A. Tropp. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. Applied and Computational Harmonic Analysis, 26(3):301?321, 2009. [22] Emmanuel J. Cand?s. The restricted isometry property and its implications for compressed sensing. Comptes Rendus Mathematique, 346(9-10):589?592, 2008. [23] Richard Baraniuk, Mark Davenport, Ronald DeVore, and Michael Wakin. A simple proof of the restricted isometry property for random matrices. Constructive Approximation, 28(3):253?263, 2008. [24] Stephen R. Becker, Emmanuel J. Cand?s, and Michael C. Grant. Templates for convex cone problems with applications to sparse signal recovery. Mathematical Programming Computation, 3(3):165?218, 2011.

9