

# “Compressive Sampling” 学习笔记

黄冬勃

2015 年 3 月 10 日

2015 年 3 月 10 日

## 目录

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Distributed source coding . . . . .	2
1.2	Compressive sensing (CS) . . . . .	2
1.3	Distributed compressive sensing (DCS) . . . . .	2
<b>2</b>	<b>Compressive Sensing Background</b>	<b>3</b>
2.1	Transform coding . . . . .	3
2.2	Incoherent projections . . . . .	3
2.3	Signal recovery via $\ell_0$ -norm minimization . . . . .	3
2.4	Signal recovery via $\ell_1$ - norm minimization . . . . .	4
2.5	Signal recovery via greedy pursuit . . . . .	4
2.6	Properties of random measurements . . . . .	4
<b>3</b>	<b>Joint Sparsity Signal Models</b>	<b>5</b>
3.1	Notation . . . . .	5
3.2	General framework for joint sparsity . . . . .	5
3.3	Example joint sparsity models . . . . .	6
3.3.1	JSM-1: Sparse common component + innovations](共性部分和个性部分是可稀疏的)	7
3.3.2	JSM-2: Common sparse supports . . . . .	7
3.3.3	JSM-3: Nonsparse common component + sparse innovations . . . . .	7
<b>4</b>	<b>Theoretical Bounds on Measurement Rates</b>	<b>8</b>
4.1	Modeling dependencies using bipartite graphs . . . . .	8
4.2	Quantifying redundancies . . . . .	9
	<b>参考文献</b>	<b>9</b>

## 摘要

压缩感知是一个信号采集构架，它的出现基于一种启示，即一个稀疏信号的线性投影集包含了足够多可以使该信号稳定重构的信息。本文我们介绍了一个新的分布式压缩感知（DCS）理论，该理论使总信号集的

新的分布式编码算法可以有效利用信号内以及信号间的相关性结构。DCS 理论基于一种新的概念，我们称它为一个信号集的“联合稀疏”。我们的理论贡献在于，在无噪声干扰的测量设定下，特征化了联合稀疏信号集 DSC 恢复的基本性能限制；我们的研究结果联合了单信号，联合信号和分布式（多个编码端）信号的压缩感知。为了体现该构架的功效以及说明其可以处理一些额外的挑战，例如计算的简易性，我们在三个举例模型中详细研究了联合稀疏信号的性能。在这些模型中，我们开发了从非相干投影中联合恢复多个信号的实践算法。在其中的两个模型中，试验结果表明，我们的实践算法的性能与上限和下限都匹配，即是渐进最佳的。此外，通过仿真表明，仅通过中等数量的信号就可以使渐进性生效。DCS 可以直接在一系列的场景中进行应用，例如多传感器阵列和网络。

## 1 Introduction

A core tenet of signal processing and information theory is that signals, images, and other data often contain some type of structure that enables intelligent representation and processing. Exploiting signal *correlations* for the purpose of *compression* are focused on in this paper.

现有的压缩算法使用一种去相关性变换的方法，例如使用精确的或者近似的 **Karhunen-Loève transform (KLT)** 变换，对相关信号的能量压缩至一些必要的参数。这种 *transform coders* 利用很多信号都存在一种 *sparse*（稀疏）表示，该表示根据一些基本原则，即以  $K$  数值大小的自适应选择变换系数进行采样传输，而不是远大于  $K$  的  $N$  信号采样。例如，连续信号在傅里叶基中是稀疏的，而在小波变换基中，分段连续信号是稀疏的<sup>[1]</sup>。

### 1.1 Distributed source coding

在 *Slepian – Wolf* 无损分布式编码构架中<sup>[2–5]</sup>，在译码端相关边信息的可用，只要每个传感器节点的条件熵加上独立上大于联合熵时，就可以使其以它的条件熵率而不是其独立熵率进行通信。SW 编码明显的优势是，传感器在对其各自测量编码时不需要相互协作，这样就节省了宝贵的通信数据负载。但不幸运的是，大多数现有的算法 [4, 5] 仅利用了信号间的相关性而没有利用信号内的相关性。至今为止对于分布式编码的研究的进展仅限于所谓的“sources with memory”。直接使用记忆源的话则需要大量的查表工作 [2]。此外，一些方法结合了数据的预处理和后处理泪去除信号内相关性，再结合 SW 编码来利用信号间相关性的做法，应用范围有限，是因为如此处理数据，可能会对某个数据产生其他节点并不可知的改变。最后，尽管最近有研究 [6–8] 提供了有记忆的空间相关源的压缩方法，其解决方案也仅适用于无损分布式压缩而不能轻易拓展应用到有损压缩中。我们得出，在很多潜在的应用中，建设性地，同时利用信号内与信号间相关性的分布式编码的设计很具有挑战性。

### 1.2 Compressive sensing (CS)

### 1.3 Distributed compressive sensing (DCS)

In a typical DCS scenario, a number of sensors measure signals that are each individually sparse in some basis and also correlated from sensor to sensor. Each sensor separately encodes its signal by projecting it onto another, incoherent basis (such as a random one) and then transmits just a few of the resulting coefficients to a single collection point.

## 2 Compressive Sensing Background

### 2.1 Transform coding

Consider a real-valued signal  $x \in \mathbb{R}^N$  indexed as  $x(n)$ ,  $n \in \{1, 2, \dots, N\}$ . Suppose that the basis  $\Psi = [\psi_1, \dots, \psi_N]$  provides a  $K$ -sparse representation of  $x$ ; that is,

$$x = \sum_{n=1}^N \vartheta(n) \psi_n = \sum_{k=1}^K \vartheta(n_k) \psi_{n_k}, \quad (2.1)$$

where  $x$  is a linear combination of  $K$  vectors chosen from  $\Psi$ ,  $\{n_k\}$  are the indices of those vectors, and  $\{\vartheta(n)\}$  are the coefficients; the concept is extendable to tight frames<sup>[1]</sup>.

### 2.2 Incoherent projections

在压缩感知中，并不对  $K$  个重要系数  $\vartheta(n)$  直接测量或编码。而是将  $M < N$  个信号  $y(m) = \langle x, \phi_m^T \rangle$  进行编码投影到第二组函数集  $\phi_m, m = 1, 2, \dots, M$  中，其中  $\phi_m^T$  表示  $\phi_m$  的转置， $\langle \cdot, \cdot \rangle$  表示内积。以矩阵形式，计算  $y = \Phi x$ ，其中  $y$  是一个  $M \times 1$  列向量，测量矩阵  $\Phi$  是  $M \times N$  阶且其每一行都由测量向量  $\phi$  构成。由于  $M < N$ ，一般而言，从  $y$  中恢复  $x$  是非良置性的；然而额外的信号稀疏度的假设，使信号恢复称为可能且可实现。

### 2.3 Signal recovery via $\ell_0$ -norm minimization

通过在  $y$  中的  $M$  观测测量相符的稀疏集系数向量  $\{\hat{\vartheta}(n)\}$  来搜索信号，进行最优化可恢复重要协同系数  $\{\vartheta(n)\}$  的稀疏集。恢复过程依赖于以  $\Psi, \Phi$  作为温和条件的关键观测值。协同系数向量  $\vartheta$  是  $\ell_0$ -norm 最小化的唯一解：

$$\hat{\vartheta} = \arg \min \|\vartheta\|_0 \quad \text{s.t. } y = \Phi \Psi \vartheta \quad (2.2)$$

**定理 2.1.** *Let  $\Psi$  be an orthonormal basis for  $\mathbb{R}^N$ , and let  $1 \leq K < N$ . Then:*

1. *Let  $\Phi$  be an  $M \times N$  measurement matrix with i.i.d. Gaussian entries with  $M \geq 2K$ . Then all signals  $x = \Psi \vartheta$  having expansion coefficients  $\vartheta \in \mathbb{R}^N$  that satisfy  $\|\vartheta\|_0 = K$  can be recovered uniquely from the  $M$ -dimensional measurement vector  $y = \Phi x$  via the  $\ell_0$ -norm minimization Equation (2.2) with probability one over  $\Phi$*
2. *Let  $x = \Psi \vartheta$  such that  $\|\vartheta\|_0 = K$ . Let  $\Phi$  be an  $M \times N$  measurement matrix with i.i.d. Gaussian entries (notably, independent of  $x$ ) with  $M \geq K + 1$ . Then  $x$  can be recovered uniquely from the  $M$ -dimensional measurement vector  $y = \Phi x$  via the  $\ell_0$ -norm minimization with probability one over  $\Phi$ .*
3. *Let  $\Phi$  be an  $M \times N$  measurement matrix, where  $M \leq K$ . Then, aside from pathological cases (specified in the proof), no signal  $x = \Psi \vartheta$  with  $\|\vartheta\|_0 = K$  can be uniquely recovered from the  $M$ -dimensional measurement vector  $y = \Phi x$ .*

*Remark 1.* The second statement of the theorem differs from the first in the following respect: when  $K < M < 2K$ , there will necessarily exist  $K$ -sparse signals  $x$  that cannot be uniquely recovered from the  $M$ -dimensional measurement vector  $y = \Phi x$ . However, these signals form a set of measure zero within

the set of all  $K$ -sparse signals and can safely be avoided with high probability if  $\Phi$  is randomly generated independently of  $x$ .

定理 1 提出了一个强逆测量区域，在一定程度上与文献<sup>[2]</sup>提出的强信道编码逆定理相似。解决  $\ell_0 - norm$  最小化问题是过于复杂的，它需要将  $\binom{N}{K}$  个可能的稀疏子空间进行组合枚举。

## 2.4 Signal recovery via $\ell_1 - norm$ minimization

支持新的压缩感知理论的现实启示是，要恢复关键参数集  $\vartheta(n)$ ，并不需要解决  $\ell_0 - norm$  最小化问题。实际上，一个简化很多的问题可以得出的一个相同解（同样感谢基之间的不相干性）；我们仅需要解决与测量值  $y$  相符合  $\ell_1 - norm$  系数矢量  $\vartheta$ <sup>[9,10]</sup>：

$$\hat{\vartheta} = \arg \min \|\vartheta\|_1 \quad s.t. y = \Phi \Psi \vartheta \quad (2.3)$$

这个最优化问题，也被称为**基本追踪**，可以通过传统线性编程技术，使计算复杂度呈  $N$  阶多项式形式，使其更易实现。

**定理 2.2.**<sup>[11-13]</sup> Set  $K=SN$  with  $0 < S \ll 1$ . Then there exists an overmeasuring factor  $c(S)=O(\log(1/S))$ ,  $c(S)>1$ , such that, for a  $K$ -sparse  $x$  in basis  $\Psi$ , the following statements hold:

1. The probability of recovering  $x$  via  $\ell_1 - norm$  minimization from  $(c(S)+\epsilon)K$  random projections,  $\epsilon > 0$ , converges to one as  $\rightarrow \infty$ .
2. The probability of recovering  $x$  via  $\ell_1 - norm$  minimization from  $(c(S)-\epsilon)K$  random projections,  $\epsilon > 0$ , converges to zero as  $\rightarrow \infty$ .

在一系列启发性的文章中，Donoho 和 Tanner<sup>[12-14]</sup> 精确特征化了过测量系数  $c(S)$ 。在我们的研究工作中，发现过测量系数非常接近于  $\log_2(1 + S^{-1})$ 。我们发现该表达式的一个有用的经验法则，以近似精确地获得过测量率。额外的过测量被证明可以提供抵抗测量噪声以及量化误差的鲁棒性<sup>[9]</sup>。

## 2.5 Signal recovery via greedy pursuit

为了从  $y$  测量值中恢复  $x$ ，迭代贪婪算法也得到了发展。例如正交匹配跟踪 (**Orthogonal Matching Pursuit OMP**) 算法，该算法从矩阵  $\Phi \Psi$  中迭代选取一些矢量，这些矢量包括大多数测量矢量  $y$  的能量。每个迭代选择设定是基于  $\Phi \Psi$  的列与残差的内积；该残差反映了  $y$  的组成是与前一列成正交的。收到 OMP 算法的启发，例如有序正交匹配追踪 (**regularized orthogonal matching pursuit**)<sup>[15]</sup>，**CoSaMP**<sup>[16]</sup>，和子空间追踪 (**Subspace Pursuit**)<sup>[17]</sup> 算法，已被证明可以为基于最优化副本提供类似的保障。本文为了从非相干测量中恢复联合稀疏信号，应用了基本追踪和贪婪算法。

## 2.6 Properties of random measurements

1. if a better sparsity-inducing basis is found for the signals, then the same measurements can be used to recover a more accurate view of the environment.
2. Random coding is also robust: the measurements coming from each sensor have equal priority, unlike Fourier or wavelet coefficients in current coders.
3. random measurements allow a progressively better recovery of the data as more measurements are obtained; one or more measurements can also be lost without corrupting the entire recovery

### 3 Joint Sparsity Signal Models

#### 3.1 Notation

- $\Lambda := 1, 2, \dots, J$  denote the set of indices for the  $J$  signals in the ensemble.  $\Lambda$  表示全体信号中的  $J$  个信号的索引集。
- Denote the *signals* in the ensemble by  $x_j$ ,  $j \in \Lambda$  and assume that each signal  $x_j \in \mathbb{R}^N$ .
- $x_j(n)$  denotes sample  $n$  in signal  $j$ , and assume for the sake of illustration – but without loss of generality – that these signals are sparse in the *canonical basis*, i.e.,  $\Psi = \mathbf{I}$ .
- $\Phi_j$  is denoted the measurement matrix for signal  $j$ ;  $\Phi_j$  is  $M_j \times N$
- in general, the entries of  $\Phi_j$  are different for each  $j$
- Thus,  $y_j = \Phi_j x_j$  consists of  $M_j < N$  random measurements of  $x_j$ .
- Random i.i.d. (*independent identically distributed, IID, 独立同分布*) Gaussian matrices  $\Phi_j$  will be emphasized in the following, but other schemes are possible, including  $\pm 1$  Bernoulli/Rademacher matrices, and so on.
- $\bar{M} = \sum_{j \in \Lambda} M_j$  and define  $X \in \mathbb{R}^{JN}$ ,  $Y \in \mathbb{R}^{\bar{M}}$ , and  $\Phi \in \mathbb{R}^{\bar{M} \times JN}$  as

$$X = \begin{bmatrix} cx_1 \\ x_2 \\ \vdots \\ x_J \end{bmatrix}, Y = \begin{bmatrix} cy_1 \\ y_2 \\ \vdots \\ y_J \end{bmatrix}, \Phi = \begin{bmatrix} \Phi_1 & 0 & \cdots & 0 \\ 0 & \Phi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Phi_J \end{bmatrix} \quad (3.1)$$

以上公式表明分离测量矩阵，当稀疏矢量的输入通过信号进行分组时，具有特征块对角结构的性质。

#### 3.2 General framework for joint sparsity

该框架基于信号集的因素化表达 ([因式分解表达?](#))，去除了其位置与数值信息的耦合。一开始，本文通过检查稀疏信号结构来促使该因素化表达，where  $x \in \mathbb{R}_N$  with  $K \ll N$  nonzero entries(非零元素)。可以通过  $x = P\theta$  来去除出  $x$  中位置和数值信息的耦合。其中  $\theta \in \mathbb{R}_N$  仅包含  $x$  中的非零元素,  $P$  是一个单位子矩阵, 换言之,  $P$  包含  $M \times N$  单位矩阵  $\mathbf{I}$  中的  $K$  列。任何  $K$ -稀疏的信号都可以以相似的形式来表达。为了模型化所有可能的稀疏信号集，这里令  $\mathcal{P}$  表示所有可能的、大小为  $N \times K'$ ,  $1 \leq K' \leq N$ , 的单位子矩阵集。这里参考  $\mathcal{P}$  作为稀疏模型。一个信号是否充分稀疏，可以通过这个模型背景下来判断：给定一个信号  $x$ ，可以把其所有可能的分解因子看作  $x = P\theta, P \in \mathcal{P}$ 。在这些分解因子中， $\theta$  最小维度的唯一表达等于信号  $x$  在模型  $\mathcal{P}$  之中的**稀疏等级**。

在信号全集的情况下，我们将该形式的分解因子看作  $X = P\Theta$ ，其中  $X \in \mathbb{R}^{JN}$ ,  $P \in \mathbb{R}^{JN}$ ，对于变化整数  $\delta$ ， $\Theta \in \mathbb{R}^\delta$ 。这里  $P$  和  $\Theta$  分别指的是**位置矩阵**和**数值向量**。根据可采纳的变化列数量的位置矩阵  $P$  的一个  $\mathcal{P}$  集，定义了一个联合稀疏模型 (*joint sparsity model (JSM)*)；这里指定在额外的条件下，位置矩阵  $P$  必须达到每个模型的要求。对于一个给定的  $X$  全集，这里使  $\mathcal{P}_F(X) \subseteq \mathcal{P}$  意指可用位置矩阵的集合  $P \in \mathcal{P}$ ，存在一个因子分解  $X = P\Theta$ 。这里定义信号全集的**联合稀疏等级**为：

**定义 3.1.** The joint sparsity level  $D$  of the signal ensemble  $X$  is the number of columns of the smallest matrix  $P \in \mathcal{P}_F(X)$

相对于单信号情况，对于多个矩阵  $P$  应该是联合稀疏模型  $\mathcal{P}$  的成员，有多个自然选择。在后续研究中注意力将集中在称为 *common/innovation component JSMs* 中。在这些模型中每个信号  $x_j$  由两个成分联合构成：

- a common component  $z_C$ , which is present in all signals, and
- an innovation component  $z_j$ , which is unique to each signal.

These combine additively, giving

$$x_j = z_C + z_j, j \in \Lambda$$

$$z_C = P_C \theta_C, z_j = P_j \theta_j, j \in \Lambda$$

where  $\theta \in \mathbb{R}^{K_C}$  and each  $\theta_j \in \mathbb{R}^{K_j}$  have nonzero entries. Each matrix  $P \in \mathcal{P}$  that can express such signals  $x_j$  has the form

$$P = \begin{bmatrix} P_C & P_1 & 0 & \cdots & 0 \\ P_C & 0 & P_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_C & 0 & 0 & \cdots & P_J \end{bmatrix} \quad (3.2)$$

where  $P_C, P_{j \in \Lambda}$  are identity submatrices.  $\Theta = [\theta_C^T \theta_1^T \theta_2^T \cdots \theta_J^T]^T, \theta_C \in \mathbb{R}^{K_C}$  and each  $\theta_j \in \mathbb{R}^{K_j}$ , to obtain  $X = P\Theta$ . 尽管  $K_C$  和  $K_j$  的值相关于矩阵  $P$ ，在后续的研究中，为了简洁明，除非有需要澄清，将忽略该相关性。

如果一个信号全集  $X = P\Theta$ ,  $\Theta \in \mathbb{R}^\delta$  由  $P_C$  和  $\{P_j\}_{j \in \Lambda}$  产生，其中全部  $J+1$  个单位子矩阵共用一个列向，则  $P$  不会是满秩的。在其他情况，我们可能观测到一个向量  $\Theta$  含有零值输入 (**zero-valued entries**)；换言之，对于一些  $1 \leq k \leq K_j$  以及  $j \in \Lambda$  可能存在  $\theta_j(k) = 0$ ，或者对于一些  $1 \leq k \leq K_C$ ,  $\theta_C(k) = 0$ 。在这两种情况中，通过从任意单位子矩阵中移除一列的实例中，可以得到一个有更少列的  $Q$  矩阵，存在  $\Theta' \in \mathbb{R}^{\delta-1}$  给出  $X = Q\Theta'$ 。如果  $Q \in \mathcal{P}$ ，则称这种现象叫做**稀疏减少**。当稀疏减少出现时，将会减少信号全集的有效联合稀疏度。作为一个稀疏减少的例子，这里看作  $J=2$  各信号的长度为  $N=2$ 。考虑公用部分  $z_C$  的系数  $z_C(1) \neq 0$  和相应的独特部分系数  $z_1(1), z_2(1) \neq 0$ 。假设其他所有系数为零。位置矩阵  $P$  产生：

$$P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

### 3.3 Example joint sparsity models

由于不同的应用场景导致了稀疏全信号之间不同相关形式，我们对联合稀疏模型  $\text{JSM}\mathcal{P}$  考虑几种可行的设计。在我们的三种联合稀疏模型之间，区别主要在于考虑了对共性部分和个性部分的不同稀疏度假设。



### 3.3.1 JSM-1: Sparse common component + innovations](共性部分和个性部分是可稀疏的)

在该模型中，假设每个信号都包含共性部分  $z_C$  和个性部分  $z_j$  且两者都是稀疏的。因此，这联合稀疏模型 ( $JSM-1$ ) $\mathcal{P}$  被表示为公式 eq. (3.2) 所示，其中  $K_C, K_j$  均小于  $N$ 。假设“稀疏减少”不存在，那么联合稀疏度为  $D = K_C + \sum_{j \in \Lambda} K_j$ 。

通过该框架进行的一个练好建模的实践例子是，对室外地点布置一组传感器节点进行全天温度测量。温度采集值  $x_j$  同时具有时间（信号内）和空间（信号间）相关性。全局因素，例如光照和风，这两者对于全部传感器都有影响，进而对  $z_C$  有影响，使其结构足够允许稀疏表示。更多的局部因素，如阴影，水或者动物，贡献了对个性部分  $z_j$  的影响， $z_j$  也是结构性的（因此是稀疏的）。一个相似的情况可以拓展到，利用传感网络记录光照强度，空气气压或其它现象。所有这些场景对应的物理过程的测量特性，在时间和空间上是平滑改变的，因此也是高度相关的。<sup>[18, 19]</sup>

### 3.3.2 JSM-2: Common sparse supports

该模型中，共性部分  $z_C$  为 0，每个个性部分  $z_j$  是稀疏的，个性部分  $z_j$  共享同一个稀疏支持（稀疏基？sparse support）但有不同的非零系数。为了在联合稀疏模型 ( $JSM-2$ ) 中形式化描述这种设置，使  $\mathcal{P}$  表示 eq. (3.2) 形式的所有矩阵集，对所有的  $j \in \Lambda$ ， $P_C = \emptyset$  以及  $P_j = \bar{P}$ 。 $\bar{P}$  表示一个大小为  $N \times K$  的任意单位子矩阵，其中  $K \ll N$ 。对一个给定的  $X = P\Theta$ ，这里可能再一次划分值向量  $\Theta = [\theta_1^T \theta_2^T \cdots \theta_J^T]$ ，每个  $\theta \in \mathbb{R}^K$ 。简单可看出  $JSM-2$  中的矩阵集  $\mathcal{P}$  是满秩的。因此，当不可能稀疏减少时，联合稀疏  $D = JK$ 。

可以立即将  $JSM-2$  模型应用在声学 and 射频传感器阵列。在此，每一个传感器都获取（采样）相同的傅立叶稀疏信号副本。但是具有相位移动，以及由于信号传播带来的幅度衰减。在这种情况下，恢复每一个感知信息是非常重要的。对于这个框架，另一个有用的应用场景是  $MIMO$  通信<sup>[20]</sup>。

在同步稀疏逼近 (simultaneous sparse approximation) 领域，考虑了类似的信号模型<sup>[20–22]</sup>。在这种情况下，收集的稀疏信号共享冗余字典中相同的扩展向量。稀疏逼近能通过贪婪算法被恢复，例如，同步正交匹配追踪算法 (Simultaneous Orthogonal Matching Pursuit SOMP)<sup>[20, 21]</sup>。或者 MMV 阶递归匹配追踪算法 (MMV Order Recursive Matching Pursuit(M-ORMP))<sup>[22]</sup>。我们使用 SOMP 算法，（在我们设立情况下，5.2 节），从非相关测量恢复一个共享公用稀疏结构的全信号。

### 3.3.3 JSM-3: Nonsparse common component + sparse innovations

在这个模型中，我们假定每一个信号都包含一个任意的公共部分  $z_C$ ，和一个稀疏个性部分  $z_j$ ；这个模型通过放松假设，即假设公用部分  $z_C$  存在一个稀疏表示，来对  $JSM-1$  模型进行扩展。为了形式化  $JSM-3$  模型中的这种设置，这里令  $\mathcal{P}$  代表 eq. (3.2) 中矩阵的全体，其中  $P_C = I$ ，是  $N \times N$  的单位矩阵。这意味着每一个  $K_j$  小于  $N$  的而  $K_C = N$ ；因此，我们得到  $\theta_C \in \mathbb{R}^N$  和  $\theta_j \in \mathbb{R}^{K_j}$ 。假定“稀疏减少”现象不会发生，联合稀疏度为  $D = N + \sum_{j \in \Lambda} K_j$ 。这里也考虑到了一个特殊情况，所有信号共享个性部分的支持，它扩展了  $JSM-2$ ；在这种情况下，对所有  $j \in \Lambda$ ，这里有  $P_j = \bar{P}$ ， $\bar{P}$  是大小为  $N \times K$  的单位子矩阵。容易看出这种情况下，稀疏减少是可能的，因此，联合稀疏度可以下降到  $D = N + (J-1)K$ 。注意，在  $JSM-3$  情况下，因为公共部分是不可稀疏的，所以当每个传感器的测量值小于  $N$  时，恢复分离 CS 是不可能的。然后，这里将演示联合 CS 恢复可以利用公共结构做到。

实际情况下，这种框架可以被很好的应用的是，几个数据源被不同的传感器一起记录，且背景信号在任何基下都不稀疏。例如，在一个零件生产工厂中，一个验证系统，在那里，为了检查制造缺陷，相

机给每个元件照相。每个照片都是极其复杂，并且是不可稀疏的。由于每一个相机用很小的变化观察相同的元件，所以这些图片之间应该是高度相关的。

JSM-3 也应用于非分布式场合，例如，视频压缩，在这种场景下，即便单个帧不是非常稀疏的，视频帧的个性部分或是差异可能是稀疏的。在这种情况下，JSM-3 建议，我们用 CS 分别编码每个视频帧，然后，联合译码所有的视频帧。这样做的优点是，为视频译码降低大量的计算复杂度。PRISM 系统提出了类似的结果，基于 Wyner-Ziv 分布式编码<sup>[23]</sup>。

存在许多可能的联合稀疏模型超出了以上介绍的模型，也超出了公共部分+个性部分信号模型。进一步的工作将产生新的 JSM 模型，以满足其它应用场景。一个应用例子，包含多个相机，这些相机从各种角度拍同场景的照片<sup>[24]</sup>。扩展和讨论在第 6 节。

## 4 Theoretical Bounds on Measurement Rates

本节提供了一个可视化模型。基于该模型，衍生出了每个传感器测量的数量，以及公式化的合并恢复过程。因此，这里将 theorem 2.1 一般化到分布式情况中，以获得使能恢复全体稀疏信号的测量值数量界线。

基于第三章，恢复  $X$  需要决定向量  $\Theta$  和位置矩阵  $P$ ，以得到  $X = P\Theta$ 。由其自身提出两个挑战：

1. 给定的一个测量仅与一些  $\Theta$  的组成部分有关，而且这些测量值的预算，应根据传感器之间在可搜集到的  $\Theta$  组分信息来进行调整适应。例如，如果一个组分  $\Theta(d)$  不影响任何传感器  $j$  的信号系数  $x_j(\cdot)$ ，相应的测量值  $y_j$  为  $\Theta(d)$  不提供信息。
2. 解码器必须从集  $\mathcal{P}$  中鉴别出位置矩阵  $P \in \mathcal{P}_F(X)$  和测量值  $Y$ 。

### 4.1 Modeling dependencies using bipartite graphs

这里介绍了一个图像化表示方式，体现了测量值  $Y$  和数值向量  $\Theta$  (分别由  $\Phi$  和  $P$  表示) 之间的依赖性。考虑将  $X$  进行可行的分解，称为满秩矩阵  $P \in \mathcal{P}_F(X)$  以及相应的  $\Theta$ ；矩阵  $P$  定义了共性部分  $K_C$  和个性部分  $K_{=j}$  的稀疏度， $1 \leq j \leq J$ ，以及联合稀疏度  $D = K_C + \sum_{j=1}^J K_j$ 。定义了以下顶点集：

- (i) 数值顶点  $V_V$  的元素索引  $d \in \{1, \dots, D\}$  表示数值矩阵  $\Theta(d)$  的输入，
- (ii) 测量集  $y_j(m)$ ， $j \in \Lambda$ ， $m \in \{1, \dots, M_j\}$ 。

这些集的基分别是  $|V_V| = D$ ， $|V_M| = \overline{M}$ 。

接下来引入一个双向图  $G = (V_V, V_M, E)$  来表示数值向量的输入和测量值之间的关系。边集  $E$  定义为：

1. 对每一个  $d \in \{1, 2, \dots, K_C\} \subseteq V_V$  和  $j \in \Lambda$ ，这样  $P_C$  的列  $e$  也不表现为  $P_j$  的列，我们有一个边连接  $d$  到每一个顶点  $(j, m) \in V_M$  for  $1 \leq m \leq M_j$ 。
2. 对每一个  $d \in \{K_C + 1, K_C + 2, \dots, D\} \subseteq V_V$ ，这里认为传感器  $j$  与  $P$  的列  $d$  相联系，可以得到边连接  $d$  到每一个顶点  $(j, m) \in V_M$  for  $1 \leq m \leq M_j$ 。

总结而言，可以说传感器  $j$  的第  $m^{th}$  个测量  $y_j(m)$ ，在图  $G$  中，如果顶点  $d \in V_V$  连接到顶点  $(j, m) \in V_M$ ，测量了  $\Theta(d)$ 。



## 4.2 Quantifying redundancies

为了得到所需测量值数量的精确边界，本文对测量过程的分析必须阐明 (must account for) 共性部分和个性部分中非零系数位置之间的冗余。最后，我们考虑了每个信号中共性和个性部分的叠加。当一个信号  $j$  以及一些索引  $1 \leq n \leq N$ ，有  $z_c(n) \neq 0$  和  $z_j \neq 0$ ，仅使用该信号是不能恢复信号测量中的两个系数，而需要使用叠加不同的其他信号的测量值来对  $z_c(n)$  恢复。这样，如 section 3.2 中描述的，在由  $P$  和  $\Theta$  给出的一个可行表达下，需要量化信号  $\Gamma \subset \Lambda$  的所有子集。

**定义 4.1.** The overlap size for the set of signals  $\Gamma \subset \Lambda$ , denoted  $K_C(\Gamma, P)$ , is the number of indices in which there is overlap between the common and the innovation component supports at all signals  $j \notin \Gamma$ :

$$K_C(\Gamma, P) = |\{n \in \{1, \dots, N\} : z_c(N) \neq 0 \text{ and } \forall j \notin \Gamma, z_j(n) \neq 0\}|. \quad (4.1)$$

We also define  $K_C(\Lambda, P) = K_C(P)$  and  $K_C(\emptyset, P) = 0$ .

## 参考文献

- [1] S. Mallat, "A wavelet tour of signal processing," *Academic Press*, 1999.
- [2] T. M. Cover and J. A. Thomas, "Elements of information theory," *New York: Wiley*, 1991.
- [3] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, pp. 471-480, July 1973.
- [4] S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (discus): Design and construction," *IEEE Trans. Inform. Theory*, vol. 49, pp. 626-643, Mar. 2003.
- [5] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Proc. Mag.* vol. 21, pp. 80-94, Sept. 2004.
- [6] T. Uyematsu, "Universal coding for correlated sources with memory," *Canadian Workshop Inform. Theory. Vancouver, Canada*, June 2001.
- [7] J. Garcia-Frias and W. Zhong, "Ldpc codes for compression of multi-terminal sources with hidden markov correlation," *IEEE Communications Letters*, vol. 7, pp. 115-117, Mar. 2003.
- [8] C. Chen, X. Ji, Q. Dai, and X. Liu, "Slepian-wolf coding of binary finite memory sources using burrows wheeler transform," *Data Compression Conference*, Mar 2009.
- [9] E. J. Candès, "Compressive sampling," *International Congress of Mathematicians*, vol. 3, pp. 1433-1452, 2006.
- [10] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289-1306, Apr 2006.
- [11] E. Candès, M. Rudelson, T. Tao, and R. Vershynin, "Error correction via linear programming," *Annual IEEE Symposium on Foundations of Computer Science, (Pittsburgh, PA)*, pp. 295-308, Oct. 2005.

- [12] D. Donoho and J. Tanner, "Neighborliness of randomly projected simplices in high dimensions," *Proc. National Academy of Sciences*, vol. 102, no. 27, pp. 9452–9457, 2005.
- [13] D. Donoho, "High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension," *Discrete and Computational Geometry*, vol. 35, pp. 617–652, Mar 2006.
- [14] D. L. Donoho and J. Tanner, "Counting faces of randomly projected polytopes when the projection radically lowers dimension," *Journal of the American Mathematical Society*, vol. 22, pp. 1–53, Jan. 2009.
- [15] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Preprint*, Dec. 2007.
- [16] D. Needell and J. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, June 2008.
- [17] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing: Closing the gap between performance and complexity," Mar. 2008.
- [18] D. Estrin, K. P. D. Culler, , and G. Sukhatme, "Connecting the physical world with pervasive networks," *IEEE Pervasive Computing*, vol. 1, no. 1, pp. 59–69, 2002.
- [19] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Comm. ACM*, vol. 43, no. 5, pp. 51–58, 2000.
- [20] J. Tropp, A. C. Gilbert, and M. J. Strauss, "Simultaneous sparse approximation via greedy pursuit," in *IEEE Int. Conf. Acoustics, Speech, Signal Proc. (ICASSP)*, Philadelphia, Mar 2005.
- [21] V. N. Temlyakov, "A remark on simultaneous sparse approximation," *East J. Approx.*, vol. 100, pp. 17–25, 2004.
- [22] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Proc.*, vol. 51, pp. 2477–2488, July 2005.
- [23] R. Puri and K. Ramchandran, "Prism: A new robust video coding architecture based on distributed compression principles," in *Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Mar 2002.
- [24] R. Wagner, R. G. Baraniuk, and R. D. Nowak, "Distributed image compression for sensor networks using correspondence analysis and super-resolution," in *Data Comp. Conf. (DCC)*, Mar 2000.