Exploiting sparsity for CMB data analysis

Jason McEwen

http://www.jasonmcewen.org/

University College London (UCL)

London Cosmology Discussion Meeting (LCDM) :: April 2013

Outline

1. Sparsity: what is it all about?

2. Wavelets on the sphere for CMB data analysis
   - Motivation
   - Continuous wavelets
   - Scale-discretised wavelets

3. Cosmological applications
   - Exploiting sparsity
   - CMB inpainting
   - Cosmic strings

Outline

1. Sparsity: what is it all about?

2. Wavelets on the sphere for CMB data analysis
   - Motivation
   - Continuous wavelets
   - Scale-discretised wavelets

3. Cosmological applications
   - Exploiting sparsity
   - CMB inpainting
   - Cosmic strings

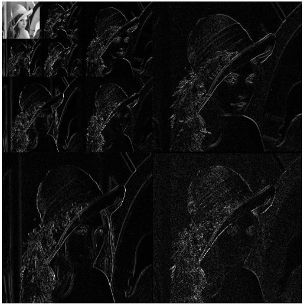
What is sparsity?

What is sparsity?

— representation of data in such a way that many data points are zero.

What is sparsity?

What is sparsity?



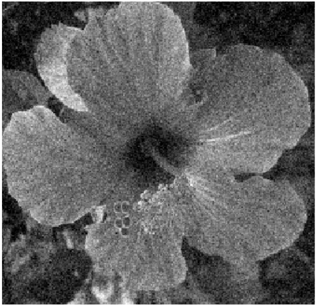
Why is sparsity useful?

Why is sparsity useful?

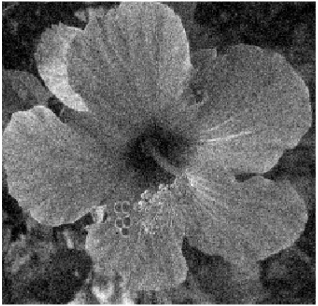
— efficient characterisation of structure.

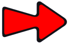
Why is sparsity useful?

Add noise

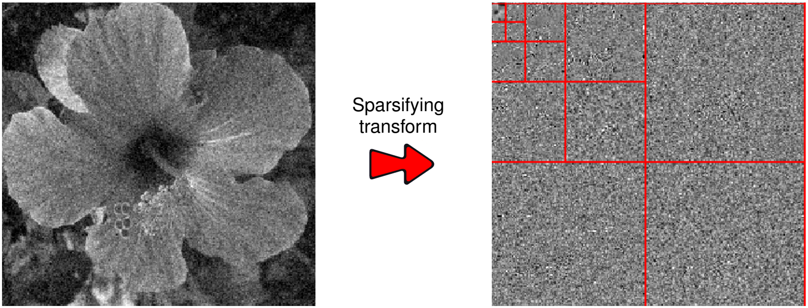


Why is sparsity useful?

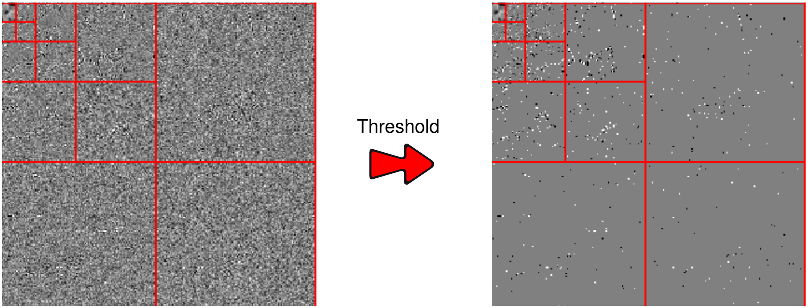




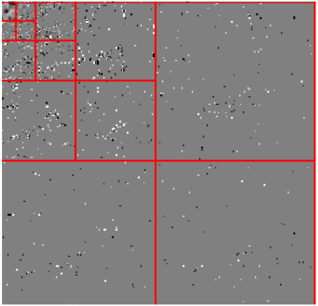
Why is sparsity useful?



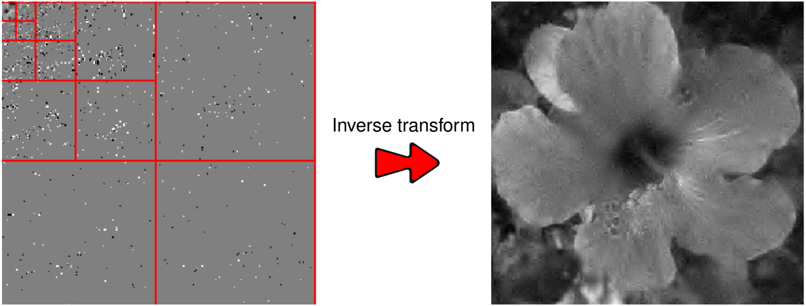
Why is sparsity useful?



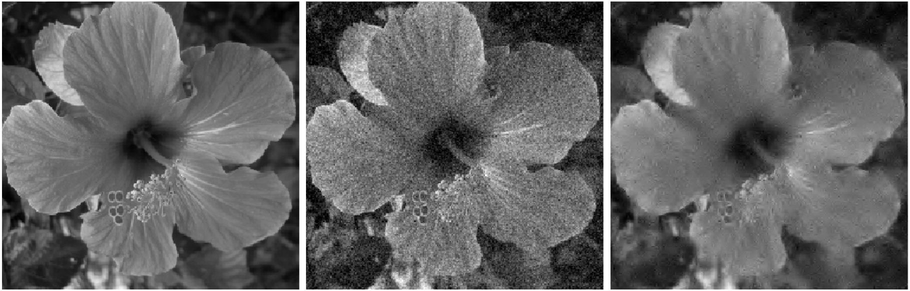
Why is sparsity useful?



Why is sparsity useful?



Why is sparsity useful?



(a) Original (b) Noisy (c) Denoised

[Credit: http://www.ceremade.dauphine.fr/~peyre/numerical-tour/tours/denoisingwav_2_wavelet_2d/]

How can we construct sparsifying transforms?

How can we construct sparsifying transforms?

— many signals in nature have spatially localised, scale-dependent features.

How can we construct sparsifying transforms?

Fourier (1807)

Haar (1909)

Morlet and Grossman (1981)

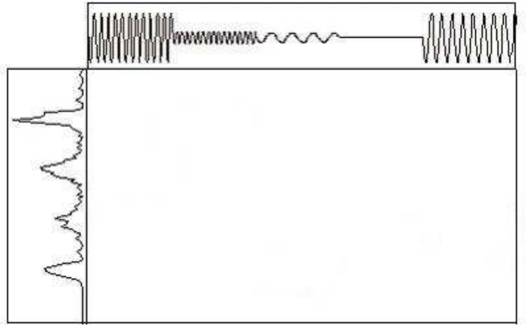


Figure: Fourier vs wavelet transform [Credit: `http://www.wavelet.org/tutorial/`]

How can we construct sparsifying transforms?

Fourier (1807)

Haar (1909)

Morlet and Grossman (1981)

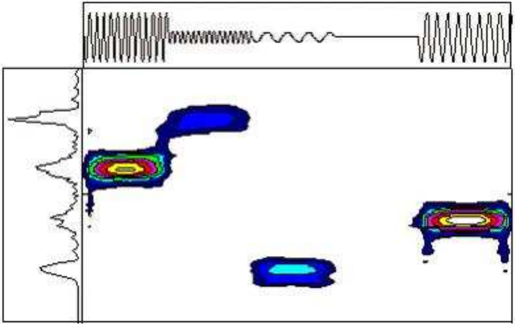


Figure: Fourier vs wavelet transform [Credit: `http://www.wavelet.org/tutorial/`]

How can we construct sparsifying transforms?

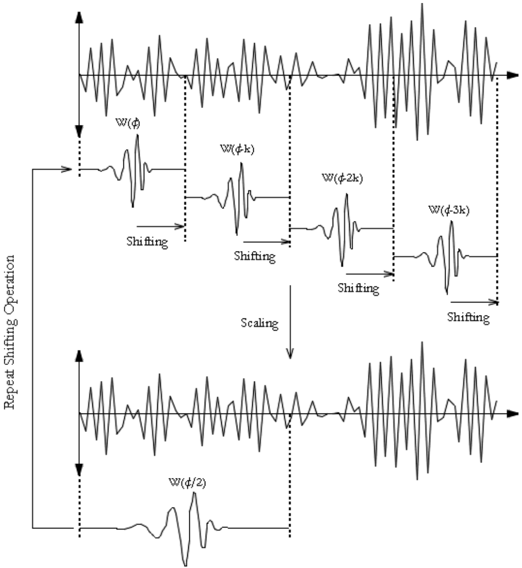


Figure: Wavelet scaling and shifting [Credit: `http://www.wavelet.org/tutorial/`]

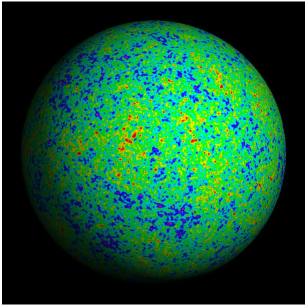
Outline

1. Sparsity: what is it all about?

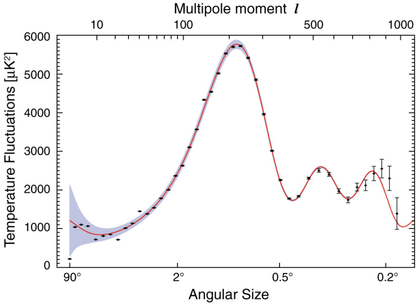
2. Wavelets on the sphere for CMB data analysis
   - Motivation
   - Continuous wavelets
   - Scale-discretised wavelets

3. Cosmological applications
   - Exploiting sparsity
   - CMB inpainting
   - Cosmic strings

CMB in real and harmonic space



(a) Temperature anisotropies



(b) Power spectrum

Figure: CMB observations [Credit: WMAP Science Team]

Spherical harmonic transform

- Spherical harmonics are the eigenfunctions of the Laplacian on the sphere: $\Delta_{\mathbb{S}^2} Y_{\ell m} = -\ell(\ell+1) Y_{\ell m}$.
- Spherical harmonics have global support over the entire sphere.

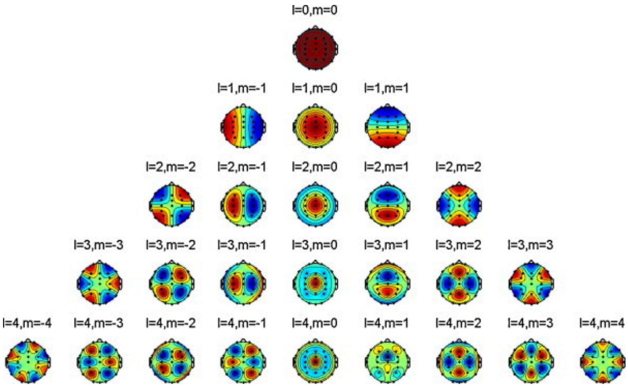


Figure: Spherical harmonic functions.

Spherical harmonic transform

- A function (*i.e.* data) on the sphere $f \in \mathrm{L}^2(\mathbb{S}^2)$ may be represented by its spherical harmonic expansion:

$$f(\theta, \varphi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{\ell m} \, Y_{\ell m}(\theta, \varphi) \; .$$

- The spherical harmonic coefficients are given by the projection onto the basis functions

$$f_{\ell m} = \langle f, \, Y_{\ell m} \rangle = \int_{\mathbb{S}^2} \mathrm{d}\Omega(\theta, \varphi) \, f(\theta, \varphi) \, Y_{\ell m}^*(\theta, \varphi) \; .$$

- In harmonic space we lose all spatial localisation since the spherical harmonics have global support.

- $\Rightarrow$ Wavelets: simultaneous scale and spatial localisation.

Spherical harmonic transform

- A function (*i.e.* data) on the sphere $f \in \mathrm{L}^2(\mathbb{S}^2)$ may be represented by its spherical harmonic expansion:

$$f(\theta, \varphi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{\ell m} \, Y_{\ell m}(\theta, \varphi) \, .$$

- The spherical harmonic coefficients are given by the projection onto the basis functions

$$f_{\ell m} = \langle f, \, Y_{\ell m} \rangle = \int_{\mathbb{S}^2} \mathrm{d}\Omega(\theta, \varphi) \, f(\theta, \varphi) \, Y_{\ell m}^*(\theta, \varphi) \, .$$

- In harmonic space we lose all spatial localisation since the spherical harmonics have global support.

- $\Rightarrow$ Wavelets: simultaneous scale and spatial localisation.

Spherical harmonic transform

- A function (*i.e.* data) on the sphere $f \in \mathrm{L}^2(\mathbb{S}^2)$ may be represented by its spherical harmonic expansion:

$$f(\theta, \varphi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{\ell m} \, Y_{\ell m}(\theta, \varphi) \, .$$

- The spherical harmonic coefficients are given by the projection onto the basis functions

$$f_{\ell m} = \langle f, \, Y_{\ell m} \rangle = \int_{\mathbb{S}^2} \mathrm{d}\Omega(\theta, \varphi) \, f(\theta, \varphi) \, Y_{\ell m}^*(\theta, \varphi) \, .$$

- In harmonic space we lose all spatial localisation since the spherical harmonics have global support.

- $\Rightarrow$ Wavelets: simultaneous scale and spatial localisation.

Wavelet transform in Euclidean space

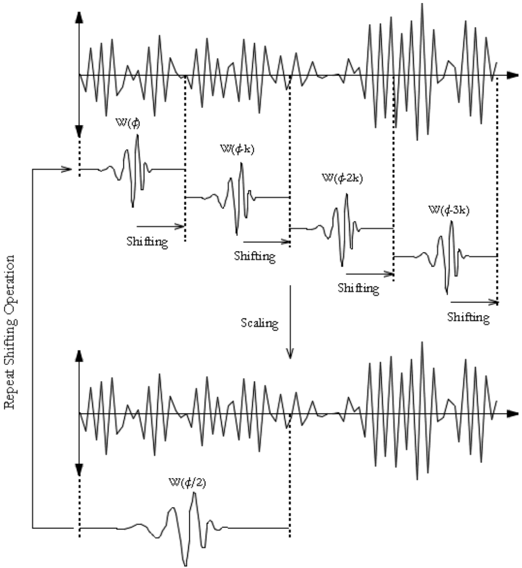


Figure: Wavelet scaling and shifting [Credit: `http://www.wavelet.org/tutorial/`]

Continuous wavelets on the sphere

- First natural wavelet construction on the sphere was derived in the seminal work of Antoine and Vandergheynst (1998) (reintroduced by Wiaux 2005).
- Construct wavelet atoms from affine transformations (dilation, translation) on the sphere of a mother wavelet.
- The natural extension of translations to the sphere are rotations. Rotation of a function f on the sphere is defined by

$$[\mathcal{R}(\rho)f](\omega) = f(\rho^{-1}\omega), \quad \omega = (\theta, \varphi) \in \mathbb{S}^2, \quad \rho = (\alpha, \beta, \gamma) \in \mathrm{SO}(3) \; .$$

- How define dilation on the sphere?
- The spherical dilation operator is defined through the conjugation of the Euclidean dilation and stereographic projection Π:

$$\mathcal{D}(a) \equiv \Pi^{-1} \, d(a) \, \Pi \; .$$

Continuous wavelets on the sphere

- First natural wavelet construction on the sphere was derived in the seminal work of Antoine and Vandergheynst (1998) (reintroduced by Wiaux 2005).
- Construct wavelet atoms from affine transformations (dilation, translation) on the sphere of a mother wavelet.
- The natural extension of translations to the sphere are rotations. Rotation of a function f on the sphere is defined by

$$[\mathcal{R}(\rho)f](\omega) = f(\rho^{-1}\omega), \quad \omega = (\theta, \varphi) \in \mathbb{S}^2, \quad \rho = (\alpha, \beta, \gamma) \in \mathrm{SO}(3) .$$

- How define dilation on the sphere?
- The spherical dilation operator is defined through the conjugation of the Euclidean dilation and stereographic projection Π:

$$\mathcal{D}(a) \equiv \Pi^{-1}\, d(a)\, \Pi .$$

Continuous wavelets on the sphere

- First natural wavelet construction on the sphere was derived in the seminal work of Antoine and Vandergheynst (1998) (reintroduced by Wiaux 2005).

- Construct wavelet atoms from affine transformations (dilation, translation) on the sphere of a mother wavelet.

- The natural extension of translations to the sphere are rotations. Rotation of a function f on the sphere is defined by

$$[\mathcal{R}(\rho)f](\omega) = f(\rho^{-1}\omega), \quad \omega = (\theta, \varphi) \in \mathbb{S}^2, \quad \rho = (\alpha, \beta, \gamma) \in \mathrm{SO}(3) \ .$$

- How define dilation on the sphere?

- The spherical dilation operator is defined through the conjugation of the Euclidean dilation and stereographic projection Π:

$$\mathcal{D}(a) \equiv \Pi^{-1}\, d(a)\, \Pi \ .$$

Continuous wavelets on the sphere

- First natural wavelet construction on the sphere was derived in the seminal work of Antoine and Vandergheynst (1998) (reintroduced by Wiaux 2005).
- Construct wavelet atoms from affine transformations (dilation, translation) on the sphere of a mother wavelet.
- The natural extension of translations to the sphere are rotations. Rotation of a function f on the sphere is defined by

$$[\mathcal{R}(\rho)f](\omega) = f(\rho^{-1}\omega), \quad \omega = (\theta, \varphi) \in \mathbb{S}^2, \quad \rho = (\alpha, \beta, \gamma) \in \mathrm{SO}(3) .$$

- How define dilation on the sphere?
- The spherical dilation operator is defined through the conjugation of the Euclidean dilation and stereographic projection Π:

$$\mathcal{D}(a) \equiv \Pi^{-1}\, d(a)\, \Pi .$$

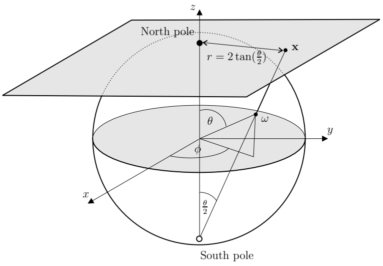


Figure: Stereographic projection.

Continuous wavelet analysis

- Wavelets on the sphere constructed from rotations and dilations of a mother spherical wavelet Ψ:

$$\{\Psi_{a,\rho} \equiv \mathcal{R}(\rho)\mathcal{D}(a)\Psi : \rho \in \mathrm{SO}(3),\, a \in \mathbb{R}_*^+\}.$$

- The forward wavelet transform is given by

$$W_\Psi^f(a,\rho) = \langle f, \Psi_{a,\rho}\rangle = \int_{\mathbb{S}^2} \mathrm{d}\Omega(\omega)\, f(\omega)\, \Psi_{a,\rho}^*(\omega)\,,$$

  where $\mathrm{d}\Omega(\omega) = \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\varphi$ is the usual invariant measure on the sphere.

- Transform general in the sense that all orientations in the rotation group $\mathrm{SO}(3)$ are considered, thus directional structure is naturally incorporated.

- Fast algorithms essential (for a review see Wiaux, McEwen & Vielva 2007)
  - Factoring of rotations: McEwen *et al.* (2007), Wandelt & Gorski (2001)
  - Separation of variables: Wiaux *et al.* (2005)

- FastCSWT code available to download: `http://www.jasonmcewen.org/`

Continuous wavelet analysis

- Wavelets on the sphere constructed from rotations and dilations of a mother spherical wavelet Ψ:

$$\{\Psi_{a,\rho} \equiv \mathcal{R}(\rho)\mathcal{D}(a)\Psi : \rho \in \mathrm{SO}(3),\, a \in \mathbb{R}^+_*\}.$$

- The forward wavelet transform is given by

$$W^f_\Psi(a, \rho) = \langle f, \Psi_{a,\rho}\rangle = \int_{\mathbb{S}^2} \mathrm{d}\Omega(\omega)\, f(\omega)\, \Psi^*_{a,\rho}(\omega)\ ,$$

where $\mathrm{d}\Omega(\omega) = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi$ is the usual invariant measure on the sphere.

- Transform general in the sense that all orientations in the rotation group $\mathrm{SO}(3)$ are considered, thus directional structure is naturally incorporated.

- Fast algorithms essential (for a review see Wiaux, McEwen & Vielva 2007)
  - Factoring of rotations: McEwen *et al.* (2007), Wandelt & Gorski (2001)
  - Separation of variables: Wiaux *et al.* (2005)

- FastCSWT code available to download: http://www.jasonmcewen.org/

Continuous wavelet analysis

- Wavelets on the sphere constructed from rotations and dilations of a mother spherical wavelet Ψ:

$$\{\Psi_{a,\rho} \equiv \mathcal{R}(\rho)\mathcal{D}(a)\Psi \,:\, \rho \in \mathrm{SO}(3),\, a \in \mathbb{R}^+_*\}.$$

- The forward wavelet transform is given by

$$W^f_\Psi(a,\rho) = \langle f, \Psi_{a,\rho} \rangle = \int_{\mathbb{S}^2} \mathrm{d}\Omega(\omega)\, f(\omega)\, \Psi^*_{a,\rho}(\omega) \,,$$

where $\mathrm{d}\Omega(\omega) = \sin\theta \,\mathrm{d}\theta \,\mathrm{d}\varphi$ is the usual invariant measure on the sphere.

- Transform general in the sense that all orientations in the rotation group $\mathrm{SO}(3)$ are considered, thus directional structure is naturally incorporated.

- Fast algorithms essential (for a review see Wiaux, McEwen & Vielva 2007)
  - Factoring of rotations: McEwen *et al.* (2007), Wandelt & Gorski (2001)
  - Separation of variables: Wiaux *et al.* (2005)

- FastCSWT code available to download: `http://www.jasonmcewen.org/`

Continuous wavelet synthesis (reconstruction)

- The inverse wavelet transform given by

$$f(\omega) = \int_0^\infty \frac{\mathrm{d}a}{a^3} \int_{\mathrm{SO}(3)} \mathrm{d}\varrho(\rho) W_\Psi^f(a,\rho) \, [\mathcal{R}(\rho)\widehat{L}_\Psi \Psi_a](\omega) \; ,$$

where $\mathrm{d}\varrho(\rho) = \sin\beta \, \mathrm{d}\alpha \, \mathrm{d}\beta \, \mathrm{d}\gamma$ is the invariant measure on the rotation group $\mathrm{SO}(3)$.

- Perfect reconstruction is ensured provided wavelets satisfy the admissibility property:

$$0 < \widehat{C}_\Psi^\ell \equiv \frac{8\pi^2}{2\ell+1} \sum_{m=-\ell}^{\ell} \int_0^\infty \frac{\mathrm{d}a}{a^3} \mid (\Psi_a)_{\ell m} \mid^2 < \infty, \quad \forall \ell \in \mathbb{N}$$

where $(\Psi_a)_{\ell m}$ are the spherical harmonic coefficients of $\Psi_a(\omega)$.

- Continuous wavelets used effectively in many cosmological studies, for example:
  - Non-Gaussianity (*e.g.* Vielva *et al.* 2004; McEwen *et al.* 2005, 2006, 2008)
  - ISW (*e.g.* Vielva *et al.* 2005, McEwen *et al.* 2007, 2008)

- BUT...

Continuous wavelet synthesis (reconstruction)

- The inverse wavelet transform given by

$$f(\omega) = \int_0^\infty \frac{\mathrm{d}a}{a^3} \int_{\mathrm{SO}(3)} \mathrm{d}\varrho(\rho) W_\Psi^f(a, \rho) \, [\mathcal{R}(\rho) \widehat{L}_\Psi \Psi_a](\omega) \; ,$$

  where $\mathrm{d}\varrho(\rho) = \sin\beta \, \mathrm{d}\alpha \, \mathrm{d}\beta \, \mathrm{d}\gamma$ is the invariant measure on the rotation group $\mathrm{SO}(3)$.

- Perfect reconstruction is ensured provided wavelets satisfy the admissibility property:

$$0 < \widehat{C}_\Psi^\ell \equiv \frac{8\pi^2}{2\ell+1} \sum_{m=-\ell}^{\ell} \int_0^\infty \frac{\mathrm{d}a}{a^3} \mid (\Psi_a)_{\ell m} \mid^2 < \infty, \quad \forall \ell \in \mathbb{N}$$

  where $(\Psi_a)_{\ell m}$ are the spherical harmonic coefficients of $\Psi_a(\omega)$.

- Continuous wavelets used effectively in many cosmological studies, for example:
  - Non-Gaussianity (*e.g.* Vielva *et al.* 2004; McEwen *et al.* 2005, 2006, 2008)
  - ISW (*e.g.* Vielva *et al.* 2005, McEwen *et al.* 2007, 2008)

- BUT...

Continuous wavelet synthesis (reconstruction)

- The inverse wavelet transform given by

$$f(\omega) = \int_0^\infty \frac{\mathrm{d}a}{a^3} \int_{\mathrm{SO}(3)} \mathrm{d}\varrho(\rho) W_\Psi^f(a,\rho) \, [\mathcal{R}(\rho)\widehat{L}_\Psi \Psi_a](\omega) \; ,$$

where $\mathrm{d}\varrho(\rho) = \sin\beta \, \mathrm{d}\alpha \, \mathrm{d}\beta \, \mathrm{d}\gamma$ is the invariant measure on the rotation group $\mathrm{SO}(3)$.

- Perfect reconstruction is ensured provided wavelets satisfy the admissibility property:

$$0 < \widehat{C}_\Psi^\ell \equiv \frac{8\pi^2}{2\ell+1} \sum_{m=-\ell}^{\ell} \int_0^\infty \frac{\mathrm{d}a}{a^3} \mid (\Psi_a)_{\ell m} \mid^2 < \infty, \quad \forall \ell \in \mathbb{N}$$

where $(\Psi_a)_{\ell m}$ are the spherical harmonic coefficients of $\Psi_a(\omega)$.

- Continuous wavelets used effectively in many cosmological studies, for example:
  - Non-Gaussianity (*e.g.* Vielva *et al.* 2004; McEwen *et al.* 2005, 2006, 2008)
  - ISW (*e.g.* Vielva *et al.* 2005, McEwen *et al.* 2007, 2008)

- BUT... **exact reconstruction not feasible in practice!**

Scale-discretised wavelets on the sphere

- **Exact reconstruction not feasible in practice with continuous wavelets!**
- Wiaux, McEwen, Vandergheynst, Blanc (2008)
  Exact reconstruction with directional wavelets on the sphere
  S2DW code

- Dilation performed in harmonic space.
  Following McEwen *et al.* (2006), Sanz *et al.* (2006).
- The scale-discretised wavelet $\Psi \in \mathrm{L}^2(\mathrm{S}^2, \mathrm{d}\Omega)$ is defined in harmonic space:

$$\widehat{\Psi}_{\ell m} = \tilde{K}_\Psi(\ell) S^\Psi_{\ell m} \ .$$

- Construct wavelets to satisfy a resolution of the identity for $0 \leq \ell < L$:

$$\tilde{\Phi}^2_\Psi(\alpha^J \ell) + \sum_{j=0}^{J} \tilde{K}^2_\Psi(\alpha^j \ell) = 1 \ .$$

Scale-discretised wavelets on the sphere

- **Exact reconstruction not feasible in practice with continuous wavelets!**
- Wiaux, McEwen, Vandergheynst, Blanc (2008)
  Exact reconstruction with directional wavelets on the sphere
  S2DW code

- Dilation performed in harmonic space.
  Following McEwen *et al.* (2006), Sanz *et al.* (2006).
- The scale-discretised wavelet $\Psi \in \mathrm{L}^2(\mathrm{S}^2, \mathrm{d}\Omega)$ is defined in harmonic space:

$$\widehat{\Psi}_{\ell m} = \tilde{K}_\Psi(\ell) S^\Psi_{\ell m} \, .$$

- Construct wavelets to satisfy a resolution of the identity for $0 \leq \ell < L$:

$$\tilde{\Phi}^2_\Psi(\alpha^J \ell) + \sum_{j=0}^{J} \tilde{K}^2_\Psi(\alpha^j \ell) = 1 \, .$$

Scale-discretised wavelets on the sphere

- **Exact reconstruction not feasible in practice with continuous wavelets!**
- Wiaux, McEwen, Vandergheynst, Blanc (2008)
  Exact reconstruction with directional wavelets on the sphere
  S2DW code

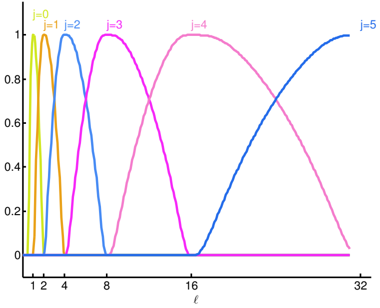


Figure: Harmonic tiling on the sphere.

- Dilation performed in harmonic space.
  Following McEwen *et al.* (2006), Sanz *et al.* (2006).
- The scale-discretised wavelet $\Psi \in \mathrm{L}^2(\mathrm{S}^2, \mathrm{d}\Omega)$ is defined in harmonic space:
  $$\widehat{\Psi}_{\ell m} = \tilde{K}_\Psi(\ell) S^\Psi_{\ell m} \; .$$
- Construct wavelets to satisfy a resolution of the identity for $0 \leq \ell < L$:
  $$\tilde{\Phi}^2_\Psi(\alpha^J \ell) + \sum_{j=0}^{J} \tilde{K}^2_\Psi(\alpha^j \ell) = 1 \, .$$

Scale-discretised wavelets on the sphere

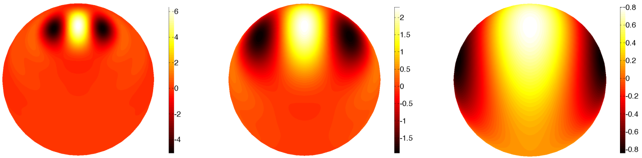


Figure: Spherical scale-discretised wavelets.

- Construct directional and steerable wavelets.
- The scale-discretised wavelet transform is given by the usual projection onto each wavelet:

$$W_{\Psi}^{f}(\rho, \alpha^{j}) = \langle f, \Psi_{\rho,\alpha^{j}} \rangle = \int_{\mathbb{S}^2} \mathrm{d}\Omega(\omega)\, f(\omega)\, \Psi^{*}_{\rho,\alpha^{j}}(\omega) \, .$$

- The original function may be recovered exactly in practice from the wavelet (and scaling) coefficients:

$$f(\omega) = \left[\Phi_{\alpha^{J}} f\right](\omega) + \sum_{j=0}^{J} \int_{\mathrm{SO}(3)} \mathrm{d}\varrho(\rho)\, W_{\Psi}^{f}\left(\rho, \alpha^{j}\right) \left[R(\rho)\, L^{\mathrm{d}} \Psi_{\alpha^{j}}\right](\omega) \, .$$

Scale-discretised wavelets on the sphere

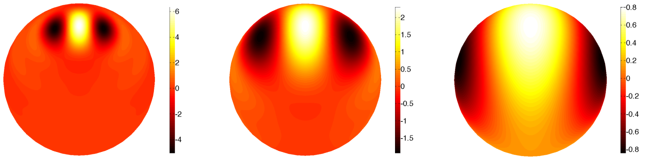


Figure: Spherical scale-discretised wavelets.

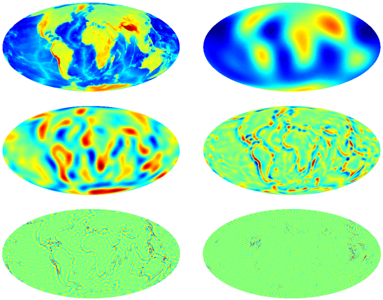
- Construct directional and steerable wavelets.
- The scale-discretised wavelet transform is given by the usual projection onto each wavelet:

$$W^{f}_{\Psi}(\rho, \alpha^{j}) = \langle f, \Psi_{\rho,\alpha^{j}} \rangle = \int_{\mathbb{S}^2} \mathrm{d}\Omega(\omega)\, f(\omega)\, \Psi^{*}_{\rho,\alpha^{j}}(\omega)\ .$$

- The original function may be recovered exactly in practice from the wavelet (and scaling) coefficients:

$$f(\omega) = \left[\Phi_{\alpha^J} f\right](\omega) + \sum_{j=0}^{J} \int_{\mathrm{SO}(3)} \mathrm{d}\varrho(\rho)\, W^{f}_{\Psi}\left(\rho, \alpha^{j}\right) \left[R(\rho)\, L^{\mathrm{d}} \Psi_{\alpha^{j}}\right](\omega)\ .$$

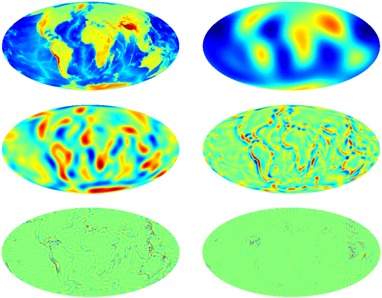
Scale-discretised wavelet transform of the Earth



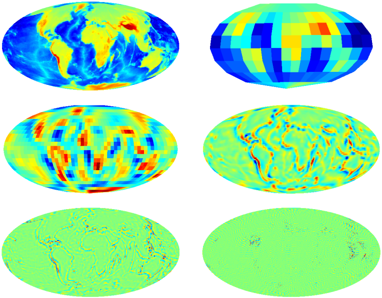
(a) Undecimated

Figure: Scale-discretised wavelet transform of a topography map of the Earth.

Scale-discretised wavelet transform of the Earth



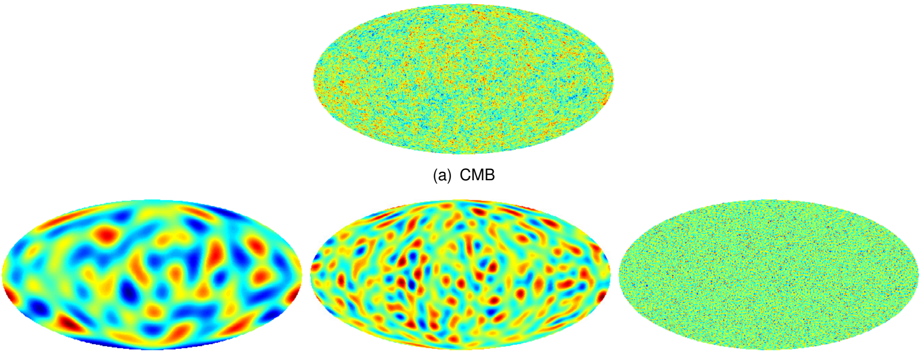
(a) Undecimated



(b) Multi-resolution

Figure: Scale-discretised wavelet transform of a topography map of the Earth.

Scale-discretised wavelet transform of the CMB



(a) CMB

(b) Wavelet coeff. (large scale) (c) Wavelet coeff. (intermediate scale) (d) Wavelet coeff. (fine scale)

Figure: Scale-discretised wavelet transform of a simulated CMB map.

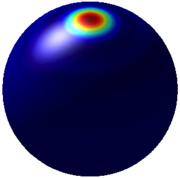
Codes to compute scale-discretised wavelets on the sphere

S2DW code

Exact reconstruction with directional wavelets on the sphere

Wiaux, McEwen, Vandergheynst, Blanc (2008)

- Fortran
- Parallelised
- Supports directional, steerable wavelets



S2LET code

S2LET: A code to perform fast wavelet analysis on the sphere

Leistedt, McEwen, Vandergheynst, Wiaux (2012)

- C, Matlab, IDL, Java
- Support only axisymmetric wavelets at present
- Future extensions:
  - Directional, steerable wavelets
  - Faster algorithms to perform wavelet transforms
  - Spin wavelets

All codes available from: `http://www.jasonmcewen.org/`

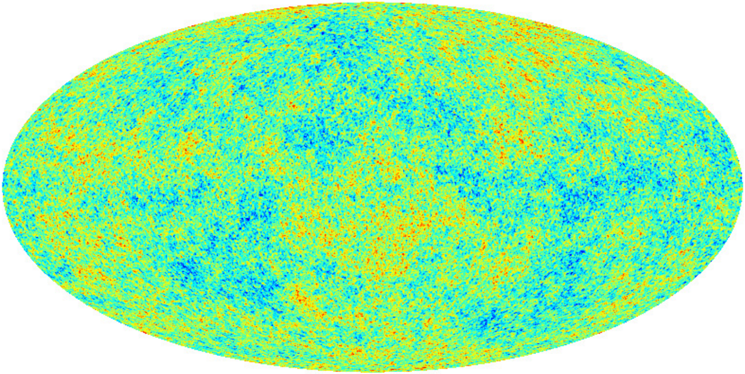
Outline

1. Sparsity: what is it all about?

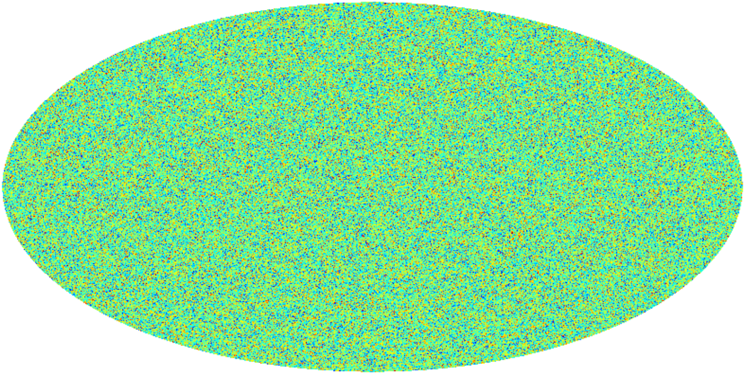
2. Wavelets on the sphere for CMB data analysis
   - Motivation
   - Continuous wavelets
   - Scale-discretised wavelets

3. Cosmological applications
   - Exploiting sparsity
   - CMB inpainting
   - Cosmic strings

Exploiting sparsity for CMB data analysis
CMB

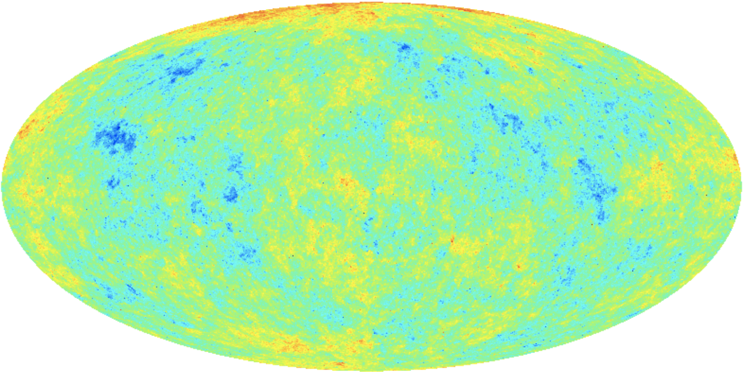


Exploiting sparsity for CMB data analysis
Wavelet coefficients of CMB



CMB is *not* sparse!

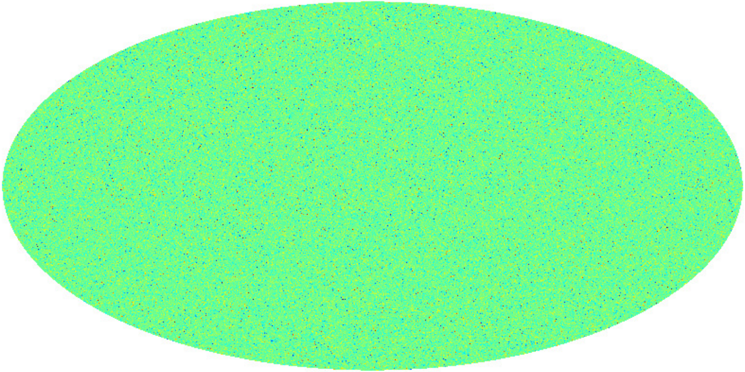
Exploiting sparsity for CMB data analysis
CMB contribution due to cosmic strings



[Credit: Ringeval *et al.* (2012)]

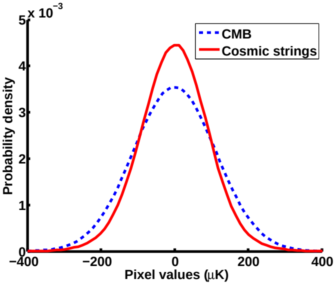
Exploiting sparsity for CMB data analysis

Wavelet coefficients of CMB contribution due to cosmic strings



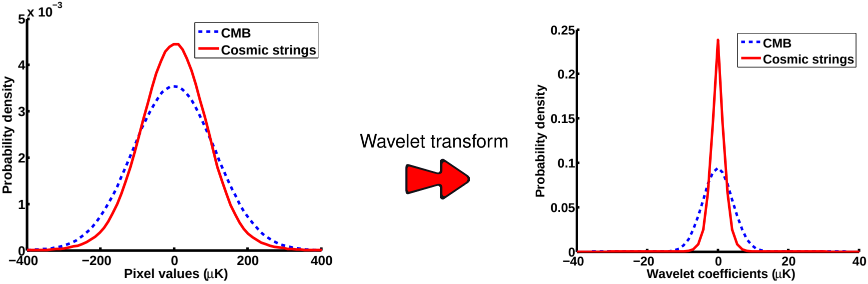
Other cosmological signals *are* sparse!

Exploiting sparsity for CMB data analysis
Correct approach



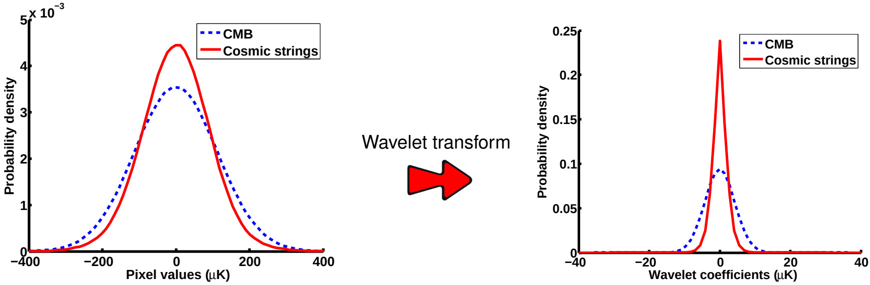
- While the CMB is not sparse, it may contain sparse contributions.
- Correct way to exploit sparsity is to treat, say, the CMB as (non-sparse) noise, and exploit sparsity of other cosmological or astrophysical signals.
- Not always the approach taken in the literature.

Exploiting sparsity for CMB data analysis
Correct approach



- While the CMB is not sparse, it may contain sparse contributions.
- Correct way to exploit sparsity is to treat, say, the CMB as (non-sparse) noise, and exploit sparsity of other cosmological or astrophysical signals.
- Not always the approach taken in the literature.

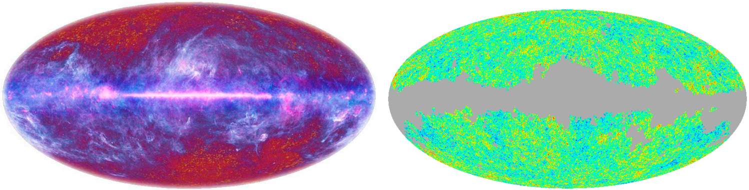
Exploiting sparsity for CMB data analysis
Correct approach



- While the CMB is not sparse, it may contain sparse contributions.
- Correct way to exploit sparsity is to treat, say, the CMB as (non-sparse) noise, and exploit sparsity of other cosmological or astrophysical signals.
- Not always the approach taken in the literature.

CMB inpainting

- Incomplete observations of the CMB on the full-sky due to Galactic contamination.



(a) Galactic contamination (b) Excise galaxy

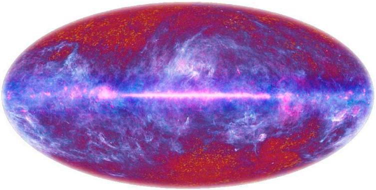
- Model observations by $y = \Phi x = \Phi \Lambda \hat{x}$ where Λ represents the inverse spherical harmonic transform and $\hat{x}$ harmonic coefficients.
- Inpainting problem solved in harmonic space (Starck *et al.* 2012):

$$\hat{x}^\star = \arg\min_{\hat{x}} \|\hat{x}\|_1 \text{ such that } y = \Phi \Lambda \hat{x} .$$

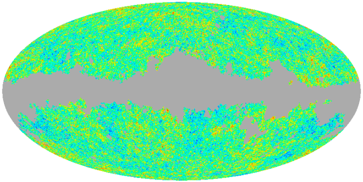
- Imposes sparsity of the spherical harmonic coefficients of the CMB!

CMB inpainting

- Incomplete observations of the CMB on the full-sky due to Galactic contamination.



(a) Galactic contamination



(b) Excise galaxy

- Model observations by $\boldsymbol{y} = \Phi \boldsymbol{x} = \Phi \Lambda \hat{\boldsymbol{x}}$ where Λ represents the inverse spherical harmonic transform and $\hat{\boldsymbol{x}}$ harmonic coefficients.

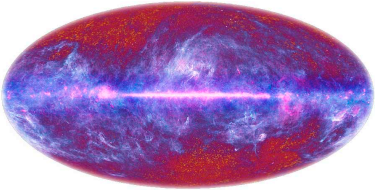
- Inpainting problem solved in harmonic space (Starck *et al.* 2012):

$$\hat{\boldsymbol{x}}^\star = \arg\min_{\hat{\boldsymbol{x}}} \|\hat{\boldsymbol{x}}\|_1 \text{ such that } \boldsymbol{y} = \Phi \Lambda \hat{\boldsymbol{x}} .$$

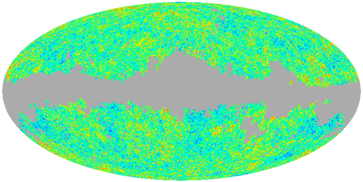
- Imposes sparsity of the spherical harmonic coefficients of the CMB!

CMB inpainting

- Incomplete observations of the CMB on the full-sky due to Galactic contamination.



(a) Galactic contamination



(b) Excise galaxy

- Model observations by $\boldsymbol{y} = \Phi \boldsymbol{x} = \Phi \Lambda \hat{\boldsymbol{x}}$ where Λ represents the inverse spherical harmonic transform and $\hat{\boldsymbol{x}}$ harmonic coefficients.
- Inpainting problem solved in harmonic space (Starck *et al.* 2012):

$$\hat{\boldsymbol{x}}^\star = \underset{\hat{\boldsymbol{x}}}{\arg\min} \, \|\hat{\boldsymbol{x}}\|_1 \text{ such that } \boldsymbol{y} = \Phi \Lambda \hat{\boldsymbol{x}} \, .$$

- Imposes sparsity of the spherical harmonic coefficients of the CMB!

CMB inpainting

- **BUT** we have a very strong physical prior. . . the CMB is very close to Gaussian!
- Solving the CMB inpainting problem in this manner is equivalent to assuming harmonic coefficients are independent and Laplacian $\rightarrow$ not a good prior.
- Furthermore, for an isotropic random field, the harmonic coefficients are independent if and only if they are Gaussian distributed.
- We can see this intuitively since a rotation in harmonic space may be written

$$(\mathcal{R}(\alpha, \beta, \gamma) a)_{\ell m} = \sum_n D^{\ell}_{mn}(\alpha, \beta, \gamma)\, a_{\ell n} \, .$$

- Sparse CMB inpainting breaks statistical isotropy!

CMB inpainting

- **BUT** we have a very strong physical prior. . . the CMB is very close to Gaussian!
- Solving the CMB inpainting problem in this manner is equivalent to assuming harmonic coefficients are independent and Laplacian $\rightarrow$ not a good prior.
- Furthermore, for an isotropic random field, the harmonic coefficients are independent if and only if they are Gaussian distributed.
- We can see this intuitively since a rotation in harmonic space may be written

$$(\mathcal{R}(\alpha, \beta, \gamma)a)_{\ell m} = \sum_n D^{\ell}_{mn}(\alpha, \beta, \gamma)\, a_{\ell n}\ .$$

- Sparse CMB inpainting breaks statistical isotropy!

Cosmic strings

- Symmetry breaking phase transitions in the early Universe → topological defects.
- Cosmic strings well-motivated phenomenon that arise when axial or cylindrical symmetry is broken → line-like discontinuities in the fabric of the Universe.
- Although we have not yet observed cosmic strings, we have observed string-like topological defects in other media, e.g. ice and liquid crystal.
- Cosmic strings are distinct to the fundamental superstrings of string theory.
- However, recent developments in string theory suggest the existence of macroscopic superstrings that could play a similar role to cosmic strings.
- The detection of cosmic strings would open a new window into the physics of the Universe!

Figure: Optical microscope photograph of a thin film of freely suspended nematic liquid crystal after a temperature quench. [Credit: Chuang *et al.* (1991).]

Cosmic strings

- Symmetry breaking phase transitions in the early Universe $\rightarrow$ topological defects.
- Cosmic strings well-motivated phenomenon that arise when axial or cylindrical symmetry is broken $\rightarrow$ line-like discontinuities in the fabric of the Universe.
- Although we have not yet observed cosmic strings, we have observed string-like topological defects in other media, e.g. ice and liquid crystal.
- Cosmic strings are distinct to the fundamental superstrings of string theory.
- However, recent developments in string theory suggest the existence of macroscopic superstrings that could play a similar role to cosmic strings.
- The detection of cosmic strings would open a new window into the physics of the Universe!

Figure: Optical microscope photograph of a thin film of freely suspended nematic liquid crystal after a temperature quench. [Credit: Chuang *et al.* (1991).]

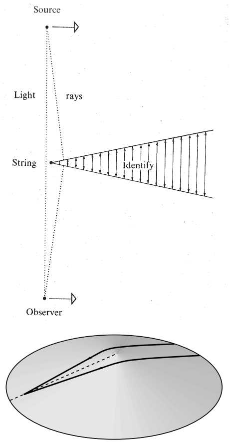
Cosmic strings

- Symmetry breaking phase transitions in the early Universe → topological defects.
- Cosmic strings well-motivated phenomenon that arise when axial or cylindrical symmetry is broken → line-like discontinuities in the fabric of the Universe.
- Although we have not yet observed cosmic strings, we have observed string-like topological defects in other media, e.g. ice and liquid crystal.
- Cosmic strings are distinct to the fundamental superstrings of string theory.
- However, recent developments in string theory suggest the existence of macroscopic superstrings that could play a similar role to cosmic strings.
- The detection of cosmic strings would open a new window into the physics of the Universe!

Figure: Optical microscope photograph of a thin film of freely suspended nematic liquid crystal after a temperature quench. [Credit: Chuang *et al.* (1991).]

Observational signatures of cosmic strings

- Spacetime about a cosmic string is canonical, with a three-dimensional wedge removed (Vilenkin 1981).
- Strings moving transverse to the line of sight induce line-like discontinuities in the CMB (Kaiser & Stebbins 1984).
- The amplitude of the induced contribution scales with $G\mu$, the string tension.

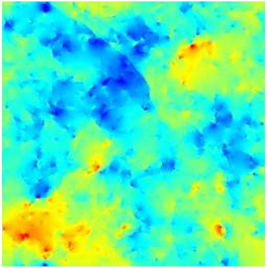


Spacetime around a cosmic string. [Credit: Kaiser & Stebbins 1984, DAMTP.]

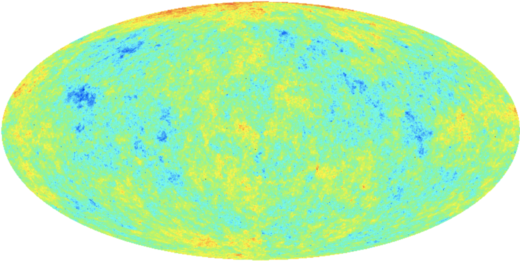
Observational signatures of cosmic strings

- Make contact between theory and data using high-resolution simulations.
- Amplitude of the signal is given by the string tension $G\mu$.
- Search for a weak string signal s embedded in the CMB c, with observations d given by

$$d = c + s \, .$$



(a) Flat patch (Fraisse *et al.* 2008)



(b) Full-sky (Ringeval *et al.* 2012)

Figure: Cosmic string simulations.

Using wavelets to detect cosmic strings

- Ongoing work of McEwen, **Feeney**, Peiris, Wiaux, Ringeval & Bouchet.
- Adopt the scale-discretised wavelet transform on the sphere (Wiaux, McEwen *et al.* 2008), where we denote the wavelet coefficients of the data d by $W^d_{j\rho} = \langle d,\ \Psi_{j\rho} \rangle$ for scale $j \in \mathbb{Z}^+$ and position $\rho \in \mathrm{SO}(3)$.

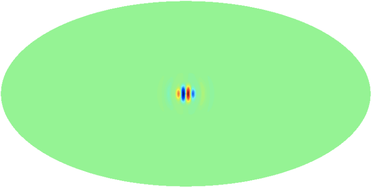


Figure: Example wavelet.

- Wavelet transform yields a sparse representation of the string signal $\rightarrow$ hope to effectively separate the CMB and string signal in wavelet space.

Using wavelets to detect cosmic strings

- Ongoing work of McEwen, **Feeney**, Peiris, Wiaux, Ringeval & Bouchet.
- Adopt the scale-discretised wavelet transform on the sphere (Wiaux, McEwen *et al.* 2008), where we denote the wavelet coefficients of the data d by $\boxed{W^d_{j\rho} = \langle d,\ \Psi_{j\rho} \rangle}$ for scale $j \in \mathbb{Z}^+$ and position $\rho \in \mathrm{SO}(3)$.

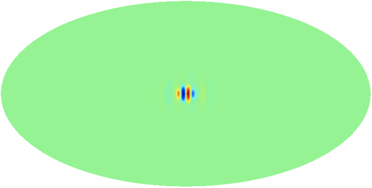


Figure: Example wavelet.

- Wavelet transform yields a sparse representation of the string signal $\rightarrow$ hope to effectively separate the CMB and string signal in wavelet space.

Using wavelets to detect cosmic strings

- Ongoing work of McEwen, **Feeney**, Peiris, Wiaux, Ringeval & Bouchet.
- Adopt the scale-discretised wavelet transform on the sphere (Wiaux, McEwen *et al.* 2008), where we denote the wavelet coefficients of the data d by $W^d_{j\rho} = \langle d, \; \Psi_{j\rho} \rangle$ for scale $j \in \mathbb{Z}^+$ and position $\rho \in \mathrm{SO}(3)$.

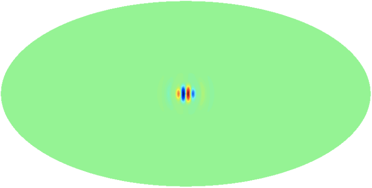


Figure: Example wavelet.

- Wavelet transform yields a sparse representation of the string signal → hope to effectively separate the CMB and string signal in wavelet space.

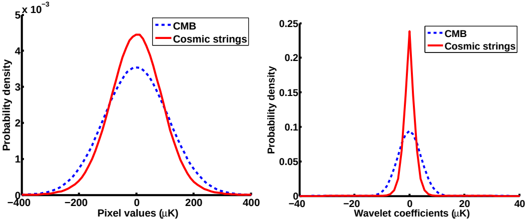


Figure: Distribution of CMB and string signal in pixel (left) and wavelet space (right).

Learning the statistics of the CMB and string signals in wavelet space

- Need to determine statistical description of the CMB and string signals in wavelet space.
- Calculate analytically the probability distribution of the CMB in wavelet space:

$$\mathrm{P}_j^c(W_{j\rho}^c) = \frac{1}{\sqrt{2\pi(\sigma_j^c)^2}} \mathrm{e}^{\left(-\frac{1}{2}\left(\frac{W_{j\rho}^c}{\sigma_j^c}\right)^2\right)} , \quad \text{where} \quad (\sigma_j^c)^2 = \langle W_{j\rho}^c \, W_{j\rho}^{c\ *} \rangle = \sum_{\ell m} C_\ell \, |(\Psi_j)_{\ell m}|^2 .$$

- Fit a generalised Gaussian distribution (GGD) for the wavelet coefficients of a string training map (*cf.* Wiaux *et al.* 2009):

$$\mathrm{P}_j^s(W_{j\rho}^s \mid G\mu) = \frac{\upsilon_j}{2G\mu\nu_j\Gamma(\upsilon_j^{-1})} \mathrm{e}^{\left(-\left|\frac{W_{j\rho}^s}{G\mu\nu_j}\right|^{\upsilon_j}\right)} ,$$

with scale parameter ν_j and shape parameter υ_j.

Learning the statistics of the CMB and string signals in wavelet space

- Need to determine statistical description of the CMB and string signals in wavelet space.
- Calculate analytically the probability distribution of the CMB in wavelet space:

$$\mathrm{P}_j^c(W_{j\rho}^c) = \frac{1}{\sqrt{2\pi(\sigma_j^c)^2}}\,\mathrm{e}^{\left(-\frac{1}{2}\left(\frac{W_{j\rho}^c}{\sigma_j^c}\right)^2\right)}, \quad \text{where} \quad (\sigma_j^c)^2 = \langle W_{j\rho}^c\, W_{j\rho}^{c\,*}\rangle = \sum_{\ell m} C_\ell \,|(\Psi_j)_{\ell m}|^2 \,.$$

- Fit a generalised Gaussian distribution (GGD) for the wavelet coefficients of a string training map (*cf.* Wiaux *et al.* 2009):

$$\mathrm{P}_j^s(W_{j\rho}^s \mid G\mu) = \frac{\upsilon_j}{2G\mu\nu_j\Gamma(\upsilon_j{}^{-1})}\,\mathrm{e}^{\left(-\left|\frac{W_{j\rho}^s}{G\mu\nu_j}\right|^{\upsilon_j}\right)},$$

with scale parameter ν_j and shape parameter υ_j.

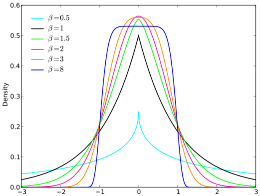
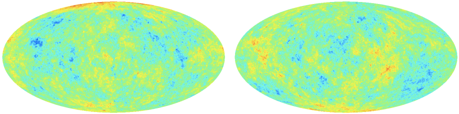


Figure: Generalised Gaussian distribution (GGD).

Learning the statistics of the CMB and string signals in wavelet space

- Require two simulated string maps: one for training; one for testing.



(a) String1 (b) String2

Figure: Cosmic string simulations.

- Compare distribution learnt from the training simulation (string2) with the distribution of the testing simulation (string1).
- Distributions in close agreement.

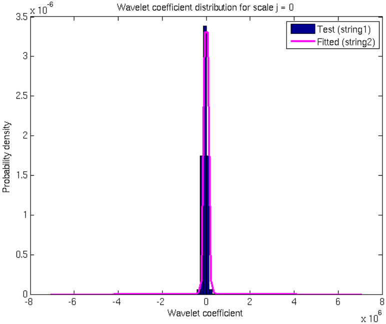
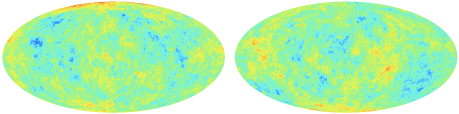


Figure: Distributions for wavelet scale $j = 0$.

Learning the statistics of the CMB and string signals in wavelet space

- Require two simulated string maps: one for training; one for testing.



(a) String1 (b) String2

Figure: Cosmic string simulations.

- Compare distribution learnt from the training simulation (string2) with the distribution of the testing simulation (string1).
- Distributions in close agreement.

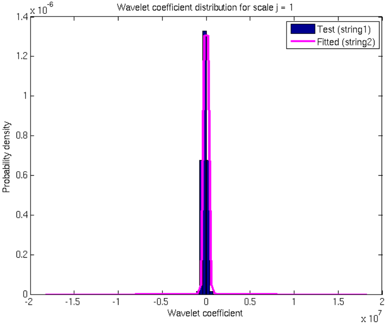
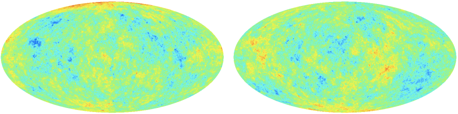


Figure: Distributions for wavelet scale $j = 1$.

Learning the statistics of the CMB and string signals in wavelet space

- Require two simulated string maps: one for training; one for testing.



(a) String1 (b) String2

Figure: Cosmic string simulations.

- Compare distribution learnt from the training simulation (string2) with the distribution of the testing simulation (string1).
- Distributions in close agreement.

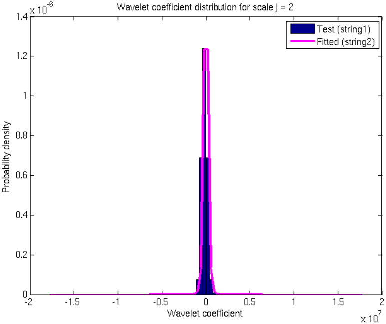
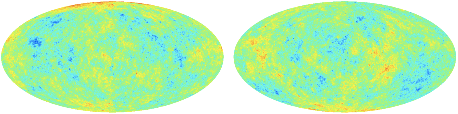


Figure: Distributions for wavelet scale $j = 2$.

Learning the statistics of the CMB and string signals in wavelet space

- Require two simulated string maps: one for training; one for testing.



(a) String1 (b) String2

Figure: Cosmic string simulations.

- Compare distribution learnt from the training simulation (string2) with the distribution of the testing simulation (string1).
- Distributions in close agreement.

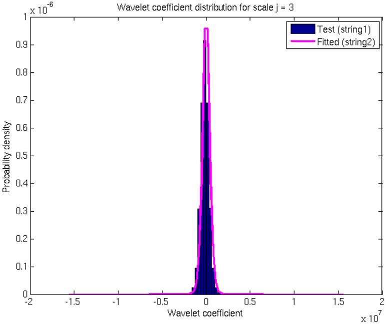
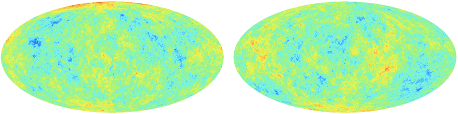


Figure: Distributions for wavelet scale $j = 3$.

Learning the statistics of the CMB and string signals in wavelet space

- Require two simulated string maps: one for training; one for testing.



(a) String1 (b) String2

Figure: Cosmic string simulations.

- Compare distribution learnt from the training simulation (string2) with the distribution of the testing simulation (string1).
- Distributions in close agreement.

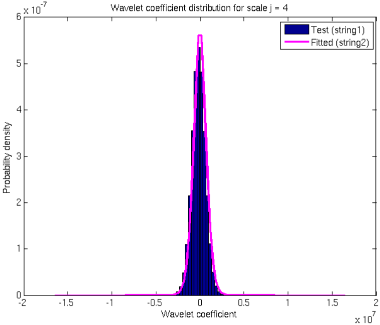
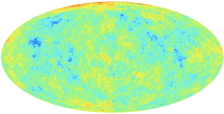


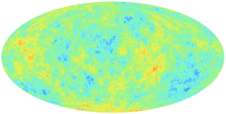
Figure: Distributions for wavelet scale $j = 4$.

Learning the statistics of the CMB and string signals in wavelet space

- Require two simulated string maps: one for training; one for testing.



(a) String1



(b) String2

Figure: Cosmic string simulations.

- Compare distribution learnt from the training simulation (string2) with the distribution of the testing simulation (string1).
- Distributions in close agreement.
- We have accurately characterised the statistics of string signals in wavelet space.

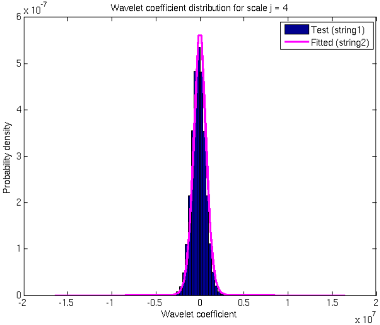


Figure: Distributions for wavelet scale $j = 4$.

Spherical wavelet-Bayesian string tension estimation

- Perform Bayesian string tension estimation in wavelet space, where the CMB and string distributions are very different.
- For each wavelet coefficient the likelihood is given by

$$\mathrm{P}(W^d_{j\rho} \mid G\mu) = \mathrm{P}(W^s_{j\rho} + W^c_{j\rho} \mid G\mu) = \int_{\mathbb{R}} \mathrm{d}W^s_{j\rho}\, \mathrm{P}^c_j(W^d_{j\rho} - W^s_{j\rho})\, \mathrm{P}^s_j(W^s_{j\rho} \mid G\mu) \ .$$

- The overall likelihood of the data is given by

$$\mathrm{P}(W^d \mid G\mu) = \prod_{j,\rho} \mathrm{P}(W^d_{j\rho} \mid G\mu) \ ,$$

where we have assumed independence.

Spherical wavelet-Bayesian string tension estimation

- Perform Bayesian string tension estimation in wavelet space, where the CMB and string distributions are very different.
- For each wavelet coefficient the likelihood is given by

$$\mathrm{P}(W^d_{j\rho} \mid G\mu) = \mathrm{P}(W^s_{j\rho} + W^c_{j\rho} \mid G\mu) = \int_{\mathbb{R}} \mathrm{d}W^s_{j\rho} \, \mathrm{P}^c_j(W^d_{j\rho} - W^s_{j\rho}) \, \mathrm{P}^s_j(W^s_{j\rho} \mid G\mu) \, .$$

- The overall likelihood of the data is given by

$$\mathrm{P}(W^d \mid G\mu) = \prod_{j,\rho} \mathrm{P}(W^d_{j\rho} \mid G\mu) \, ,$$

where we have assumed independence.

Spherical wavelet-Bayesian string tension estimation

- Compute the string tension posterior $\mathrm{P}(G\mu \mid W^d)$ by Bayes theorem:

$$\mathrm{P}(G\mu \mid W^d) = \frac{\mathrm{P}(W^d \mid G\mu)\, \mathrm{P}(G\mu)}{\mathrm{P}(W^d)} \propto \mathrm{P}(W^d \mid G\mu)\, \mathrm{P}(G\mu) \,.$$

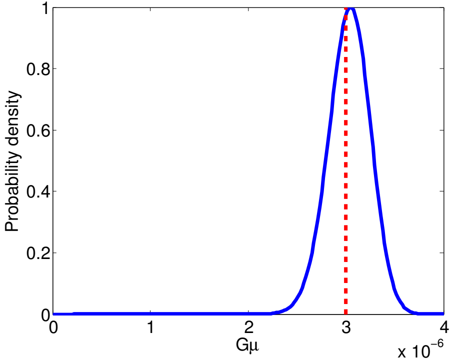


Figure: Posterior distribution of the string tension (true $G\mu = 3 \times 10^{-6}$).

Spherical wavelet-Bayesian string tension estimation

- Compute the string tension posterior $\mathrm{P}(G\mu \mid W^d)$ by Bayes theorem:

$$\mathrm{P}(G\mu \mid W^d) = \frac{\mathrm{P}(W^d \mid G\mu)\,\mathrm{P}(G\mu)}{\mathrm{P}(W^d)} \propto \mathrm{P}(W^d \mid G\mu)\,\mathrm{P}(G\mu)\ .$$

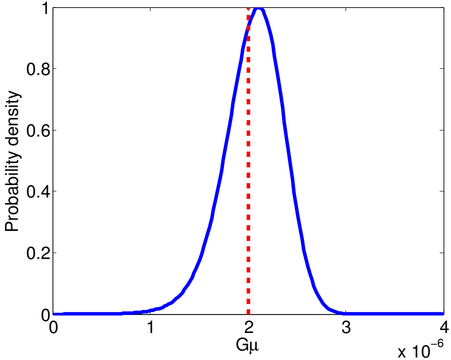


Figure: Posterior distribution of the string tension (true $G\mu = 2 \times 10^{-6}$).

Spherical wavelet-Bayesian string tension estimation

- Compute the string tension posterior $\mathrm{P}(G\mu \mid W^d)$ by Bayes theorem:

$$\mathrm{P}(G\mu \mid W^d) = \frac{\mathrm{P}(W^d \mid G\mu)\, \mathrm{P}(G\mu)}{\mathrm{P}(W^d)} \propto \mathrm{P}(W^d \mid G\mu)\, \mathrm{P}(G\mu) \, .$$

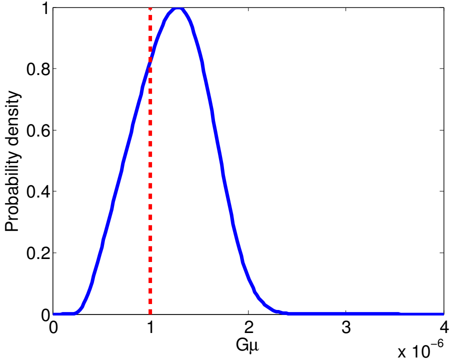


Figure: Posterior distribution of the string tension (true $G\mu = 1 \times 10^{-6}$).

Bayesian evidence for strings

- Compute Bayesian evidences to compare the string model M^s to the alternative model M^c that the observed data is comprised of just a CMB contribution.
- The Bayesian evidence of the string model is given by

$$E^s = \mathrm{P}(W^d \mid \mathrm{M}^s) = \int_{\mathbb{R}} \mathrm{d}(G\mu)\, \mathrm{P}(W^d \mid G\mu)\, \mathrm{P}(G\mu)\ .$$

- The Bayesian evidence of the CMB model is given by

$$E^c = \mathrm{P}(W^d \mid \mathrm{M}^c) = \prod_{j,\rho} \mathrm{P}_j^c(W_{j\rho}^d)\ .$$

- Compute the Bayes factor to determine the preferred model:

$$\Delta \ln E = \ln(E^s / E^c)\ .$$

Bayesian evidence for strings

- Compute Bayesian evidences to compare the string model M^s to the alternative model M^c that the observed data is comprised of just a CMB contribution.

- The Bayesian evidence of the string model is given by

$$E^s = \mathrm{P}(W^d \mid \mathrm{M}^s) = \int_{\mathbb{R}} \mathrm{d}(G\mu)\, \mathrm{P}(W^d \mid G\mu)\, \mathrm{P}(G\mu) \ .$$

- The Bayesian evidence of the CMB model is given by

$$E^c = \mathrm{P}(W^d \mid \mathrm{M}^c) = \prod_{j,\rho} \mathrm{P}_j^c(W_{j\rho}^d) \ .$$

- Compute the Bayes factor to determine the preferred model:

$$\Delta \ln E = \ln(E^s / E^c) \ .$$

Table: Tension estimates and log-evidence differences for simulations.

$G\mu/10^{-6}$	0.7	0.8	0.9	1.0	2.0	3.0
$\widehat{G\mu}/10^{-6}$	1.1	1.2	1.2	1.3	2.1	3.1
$\Delta \ln E$	-1.3	-1.1	-0.9	-0.7	5.5	29

Recovering string maps

- Our best inference of the underlying string map is encoded in the posterior probability distribution $\mathrm{P}(W^s_{j\rho} \mid W^d)$.

- Estimate the wavelet coefficients of the string map from the mean of the posterior distribution:

$$\overline{W}^s_{j\rho} = \int_{\mathbb{R}} \mathrm{d}W^s_{j\rho}\, W^s_{j\rho}\, \mathrm{P}(W^s_{j\rho} \mid W^d) = \int_{\mathbb{R}} \mathrm{d}(G\mu)\, \mathrm{P}(G\mu \mid d)\, \overline{W}^s_{j\rho}(G\mu)\ ,$$

where

$$\overline{W}^s_{j\rho}(G\mu) = \int_{\mathbb{R}} \mathrm{d}W^s_{j\rho}\, W^s_{j\rho}\, \mathrm{P}(W^s_{j\rho} \mid W^d_{j\rho}, G\mu) = \frac{1}{\mathrm{P}(W^d_{j\rho} \mid G\mu)} \int_{\mathbb{R}} \mathrm{d}W^s_{j\rho}\, W^s_{j\rho}\, \mathrm{P}^c_j(W^d_{j\rho} - W^s_{j\rho})\, \mathrm{P}^s_j(W^s_{j\rho} \mid G\mu)\ .$$

- Recover the string map from its wavelets (possible since the scale-discretised wavelet transform on the sphere supports exact reconstruction).

- Work in progress. . .

Recovering string maps

- Our best inference of the underlying string map is encoded in the posterior probability distribution $\mathrm{P}(W^s_{j\rho} \mid W^d)$.

- Estimate the wavelet coefficients of the string map from the mean of the posterior distribution:

$$\overline{W}^s_{j\rho} = \int_{\mathbb{R}} \mathrm{d}W^s_{j\rho} \, W^s_{j\rho} \, \mathrm{P}(W^s_{j\rho} \mid W^d)$$
$$= \int_{\mathbb{R}} \mathrm{d}(G\mu) \, \mathrm{P}(G\mu \mid d) \, \overline{W}^s_{j\rho}(G\mu) \, ,$$

  where

$$\overline{W}^s_{j\rho}(G\mu) = \int_{\mathbb{R}} \mathrm{d}W^s_{j\rho} \, W^s_{j\rho} \, \mathrm{P}(W^s_{j\rho} \mid W^d_{j\rho}, G\mu)$$
$$= \frac{1}{\mathrm{P}(W^d_{j\rho} \mid G\mu)} \int_{\mathbb{R}} \mathrm{d}W^s_{j\rho} \, W^s_{j\rho} \, \mathrm{P}^c_j(W^d_{j\rho} - W^s_{j\rho}) \, \mathrm{P}^s_j(W^s_{j\rho} \mid G\mu) \, .$$

- Recover the string map from its wavelets (possible since the scale-discretised wavelet transform on the sphere supports exact reconstruction).

- Work in progress. . .

Conclusions

Sparsity is a powerful concept that can provide new insight and is complementary to a Bayesian approach.

But, as all techniques, sparsity must be exploited in the correct manner.

Great potential for cosmology, leading to the emerging field of **CosmoInformatics**.

Compressive sensing

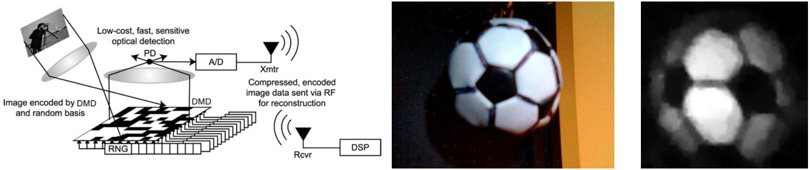
- Next evolution of wavelet analysis $\rightarrow$ wavelets are a key ingredient.
- The mystery of JPEG compression (discrete cosine transform; wavelet transform).
- Move compression to the acquisition stage $\rightarrow$ compressive sensing.
- Acquisition versus imaging.

Compressive sensing

- Next evolution of wavelet analysis $\rightarrow$ wavelets are a key ingredient.
- The mystery of JPEG compression (discrete cosine transform; wavelet transform).
- Move compression to the acquisition stage $\rightarrow$ compressive sensing.
- Acquisition versus imaging.

Compressive sensing

- Next evolution of wavelet analysis $\rightarrow$ wavelets are a key ingredient.
- The mystery of JPEG compression (discrete cosine transform; wavelet transform).
- Move compression to the acquisition stage $\rightarrow$ compressive sensing.
- Acquisition versus imaging.

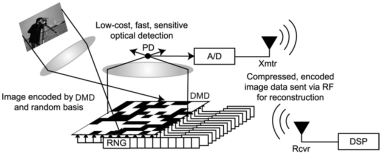


(a) Architecture (b) Scene (c) Recov. (20% meas.)

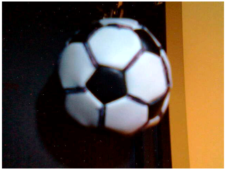
Figure: Single pixel camera

Compressive sensing

- Next evolution of wavelet analysis $\rightarrow$ wavelets are a key ingredient.
- The mystery of JPEG compression (discrete cosine transform; wavelet transform).
- Move compression to the acquisition stage $\rightarrow$ compressive sensing.
- Acquisition versus imaging.



(a) Architecture



(b) Scene

(c) Recov. (20% meas.)

Figure: Single pixel camera

An introduction to compressive sensing

- Linear operator (algebra) representation of signal decomposition (into *atoms* of a *dictionary*):

$$x(t) = \sum_i \alpha_i \Psi_i(t) \quad \rightarrow \quad \boldsymbol{x} = \sum_i \Psi_i \alpha_i = \begin{pmatrix} | \\ \Psi_0 \\ | \end{pmatrix} \alpha_0 + \begin{pmatrix} | \\ \Psi_1 \\ | \end{pmatrix} \alpha_1 + \cdots \quad \rightarrow \quad \boxed{\boldsymbol{x} = \Psi \boldsymbol{\alpha}}$$

- Linear operator (algebra) representation of measurement:

$$y_i = \langle x, \Phi_j \rangle \quad \rightarrow \quad \boldsymbol{y} = \begin{pmatrix} - \ \Phi_0 \ - \\ - \ \Phi_1 \ - \\ \vdots \end{pmatrix} \boldsymbol{x} \quad \rightarrow \quad \boxed{\boldsymbol{y} = \Phi \boldsymbol{x}}$$

- Putting it together: $\boxed{\boldsymbol{y} = \Phi \boldsymbol{x} = \Phi \Psi \boldsymbol{\alpha}}$

An introduction to compressive sensing

- Linear operator (algebra) representation of signal decomposition (into *atoms* of a *dictionary*):

$$x(t) = \sum_i \alpha_i \Psi_i(t) \quad \rightarrow \quad \boldsymbol{x} = \sum_i \Psi_i \alpha_i = \begin{pmatrix} | \\ \Psi_0 \\ | \end{pmatrix} \alpha_0 + \begin{pmatrix} | \\ \Psi_1 \\ | \end{pmatrix} \alpha_1 + \cdots \quad \rightarrow \quad \boxed{\boldsymbol{x} = \Psi \boldsymbol{\alpha}}$$

- Linear operator (algebra) representation of measurement:

$$y_i = \langle x, \Phi_j \rangle \quad \rightarrow \quad \boldsymbol{y} = \begin{pmatrix} - & \Phi_0 & - \\ - & \Phi_1 & - \\ & \vdots & \end{pmatrix} \boldsymbol{x} \quad \rightarrow \quad \boxed{\boldsymbol{y} = \Phi \boldsymbol{x}}$$

- Putting it together: $\boxed{\boldsymbol{y} = \Phi \boldsymbol{x} = \Phi \Psi \boldsymbol{\alpha}}$

An introduction to compressive sensing

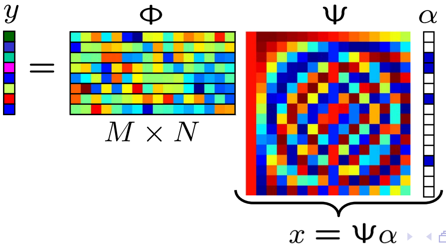
- Linear operator (algebra) representation of signal decomposition (into *atoms* of a *dictionary*):

$$x(t) = \sum_i \alpha_i \Psi_i(t) \quad \rightarrow \quad \boldsymbol{x} = \sum_i \Psi_i \alpha_i = \begin{pmatrix} | \\ \Psi_0 \\ | \end{pmatrix} \alpha_0 + \begin{pmatrix} | \\ \Psi_1 \\ | \end{pmatrix} \alpha_1 + \cdots \quad \rightarrow \quad \boxed{\boldsymbol{x} = \Psi \boldsymbol{\alpha}}$$

- Linear operator (algebra) representation of measurement:

$$y_i = \langle x, \Phi_j \rangle \quad \rightarrow \quad \boldsymbol{y} = \begin{pmatrix} - & \Phi_0 & - \\ - & \Phi_1 & - \\ & \vdots & \end{pmatrix} \boldsymbol{x} \quad \rightarrow \quad \boxed{\boldsymbol{y} = \Phi \boldsymbol{x}}$$

- Putting it together: $\boxed{\boldsymbol{y} = \Phi \boldsymbol{x} = \Phi \Psi \boldsymbol{\alpha}}$



An introduction to compressive sensing

- Ill-posed inverse problem:

$$\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{n} = \Phi \Psi \boldsymbol{\alpha} + \boldsymbol{n} .$$

- Solve by imposing a regularising prior that the signal to be recovered is sparse in Ψ, *i.e.* solve the following ℓ_0 optimisation problem:

$$\boldsymbol{\alpha}^\star = \underset{\boldsymbol{\alpha}}{\arg\min} \|\alpha\|_0 \text{ such that } \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon ,$$

where the signal is synthesising by $\boldsymbol{x}^\star = \Psi\boldsymbol{\alpha}^\star$.

- Recall norms given by:

$$\|\alpha\|_0 = \text{no. non-zero elements} \qquad \|\alpha\|_1 = \sum_i |\alpha_i| \qquad \|\alpha\|_2 = \Big(\sum_i |\alpha_i|^2\Big)^{1/2}$$

- Solving this problem is difficult (combinatorial).

- Instead, solve the ℓ_1 optimisation problem (convex):

$$\boldsymbol{\alpha}^\star = \underset{\boldsymbol{\alpha}}{\arg\min} \|\alpha\|_1 \text{ such that } \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon .$$

An introduction to compressive sensing

- Ill-posed inverse problem:

$$\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{n} = \Phi \Psi \boldsymbol{\alpha} + \boldsymbol{n} .$$

- Solve by imposing a regularising prior that the signal to be recovered is sparse in Ψ, *i.e.* solve the following ℓ_0 optimisation problem:

$$\boldsymbol{\alpha}^\star = \arg\min_{\boldsymbol{\alpha}} \|\alpha\|_0 \text{ such that } \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon ,$$

  where the signal is synthesising by $\boldsymbol{x}^\star = \Psi \boldsymbol{\alpha}^\star$.

- Recall norms given by:

$$\|\alpha\|_0 = \text{no. non-zero elements} \qquad \|\alpha\|_1 = \sum_i |\alpha_i| \qquad \|\alpha\|_2 = \Big(\sum_i |\alpha_i|^2\Big)^{1/2}$$

- Solving this problem is difficult (combinatorial).

- Instead, solve the ℓ_1 optimisation problem (convex):

$$\boldsymbol{\alpha}^\star = \arg\min_{\boldsymbol{\alpha}} \|\alpha\|_1 \text{ such that } \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon .$$

An introduction to compressive sensing

- Ill-posed inverse problem:

$$\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{n} = \Phi \Psi \boldsymbol{\alpha} + \boldsymbol{n} .$$

- Solve by imposing a regularising prior that the signal to be recovered is sparse in Ψ, *i.e.* solve the following ℓ_0 optimisation problem:

$$\boldsymbol{\alpha}^\star = \arg\min_{\boldsymbol{\alpha}} \|\alpha\|_0 \text{ such that } \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon ,$$

  where the signal is synthesising by $\boldsymbol{x}^\star = \Psi \boldsymbol{\alpha}^\star$.

- Recall norms given by:

$$\|\alpha\|_0 = \text{no. non-zero elements} \qquad \|\alpha\|_1 = \sum_i |\alpha_i| \qquad \|\alpha\|_2 = \Big(\sum_i |\alpha_i|^2\Big)^{1/2}$$

- Solving this problem is difficult (combinatorial).

- Instead, solve the ℓ_1 optimisation problem (convex):

$$\boldsymbol{\alpha}^\star = \arg\min_{\boldsymbol{\alpha}} \|\alpha\|_1 \text{ such that } \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon .$$

An introduction to compressive sensing

- Ill-posed inverse problem:

$$\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{n} = \Phi \Psi \boldsymbol{\alpha} + \boldsymbol{n} .$$

- Solve by imposing a regularising prior that the signal to be recovered is sparse in Ψ, *i.e.* solve the following ℓ_0 optimisation problem:

$$\boldsymbol{\alpha}^\star = \arg\min_{\boldsymbol{\alpha}} \|\alpha\|_0 \text{ such that } \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon ,$$

  where the signal is synthesising by $\boldsymbol{x}^\star = \Psi \boldsymbol{\alpha}^\star$.

- Recall norms given by:

$$\|\alpha\|_0 = \text{no. non-zero elements} \qquad \|\alpha\|_1 = \sum_i |\alpha_i| \qquad \|\alpha\|_2 = \Big(\sum_i |\alpha_i|^2\Big)^{1/2}$$

- Solving this problem is difficult (combinatorial).

- Instead, solve the ℓ_1 optimisation problem (convex):

$$\boldsymbol{\alpha}^\star = \arg\min_{\boldsymbol{\alpha}} \|\alpha\|_1 \text{ such that } \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon .$$

An introduction to compressive sensing

- The solutions of the ℓ_0 and ℓ_1 problems are often the same.
- Space of sparse vectors given by the union of subspaces aligned with the coordinate axes.

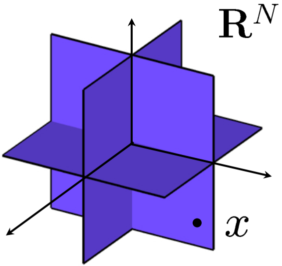


Figure: Space of the sparse vectors [Credit: Baraniuk]

An introduction to compressive sensing

- The solutions of the ℓ_0 and ℓ_1 problems are often the same.
- Restricted isometry property (RIP):

$$(1-\delta_{2K})\|\boldsymbol{x}_1-\boldsymbol{x}_2\|_2^2 \leq \|\Phi\boldsymbol{x}_1-\Phi\boldsymbol{x}_2\|_2^2 \leq (1+\delta_{2K})\|\boldsymbol{x}_! - \boldsymbol{x}_2\|_2^2 \,,$$

  for K-sparse $\boldsymbol{x}$.
- Measurement must preserve geometry of sets of sparse vectors.

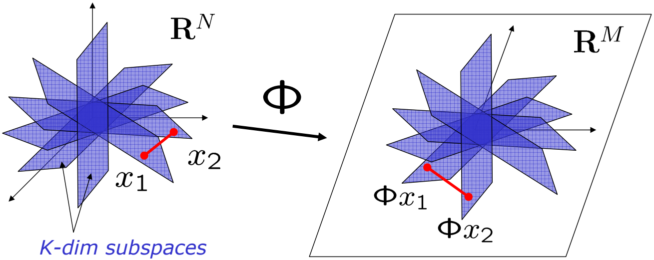


Figure: Measurement must preserve geometry of sets of sparse vectors. [Credit: Baraniuk]

An introduction to compressive sensing

- The solutions of the ℓ_0 and ℓ_1 problems are often the same.
- Geometry of ℓ_2 and ℓ_1 problems.

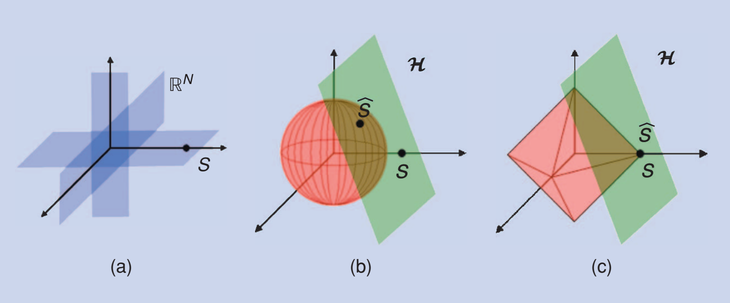


Figure: Geometry of (a) ℓ_0 (b) ℓ_2 and (c) ℓ_1 problems. [Credit: Baraniuk (2007)]

An introduction to compressive sensing

- In the absence of noise, compressed sensing is **exact!**
- Number of measurements required to achieve exact reconstruction is given by

$$M \geq c\mu^2 K \log N ,$$

where K is the sparsity and N the dimensionality.

- The coherence between the measurement and sparsity basis is given by

$$\mu = \sqrt{N} \max_{i,j} |\langle \Psi_i, \Phi_j \rangle| .$$

- Robust to noise.

An introduction to compressive sensing

- In the absence of noise, compressed sensing is **exact!**
- Number of measurements required to achieve exact reconstruction is given by

$$M \geq c\mu^2 K \log N ,$$

where K is the sparsity and N the dimensionality.

- The coherence between the measurement and sparsity basis is given by

$$\mu = \sqrt{N} \max_{i,j} |\langle \Psi_i, \Phi_j \rangle| .$$

- Robust to noise.

An introduction to compressive sensing

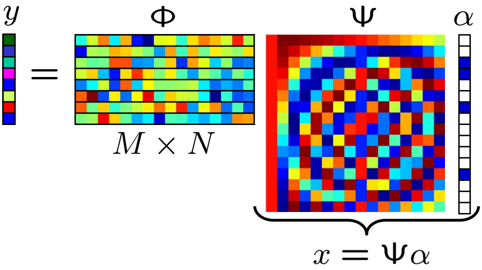
- In the absence of noise, compressed sensing is **exact!**
- Number of measurements required to achieve exact reconstruction is given by

$$M \geq c\mu^2 K \log N \ ,$$

where K is the sparsity and N the dimensionality.

- The coherence between the measurement and sparsity basis is given by

$$\mu = \sqrt{N} \max_{i,j} |\langle \Psi_i, \Phi_j \rangle| \ .$$



- Robust to noise.

An introduction to compressive sensing

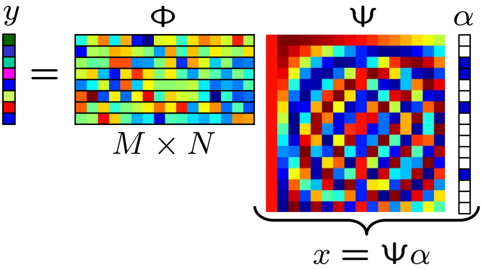
- In the absence of noise, compressed sensing is **exact!**
- Number of measurements required to achieve exact reconstruction is given by

$$M \geq c\mu^2 K \log N \ ,$$

where K is the sparsity and N the dimensionality.

- The coherence between the measurement and sparsity basis is given by

$$\mu = \sqrt{N} \max_{i,j} |\langle \Psi_i, \Phi_j \rangle| \ .$$



- Robust to noise.

A Bayesian perspective

- Consider the inverse problem:

$$y = \Phi \Psi \alpha + n \ .$$

- Assume Gaussian noise, yielding the likelihood:

$$\mathrm{P}(y \mid \alpha) \propto \exp\Big(\|y - \Phi \Psi \alpha\|_2^2 / (2\sigma^2) \Big) \ .$$

- Consider the Laplacian prior:

$$\mathrm{P}(\alpha) \propto \exp\Big(-\beta \|\alpha\|_1 \Big) \ .$$

- The maximum *a-posteriori* (MAP) estimate is then

$$x^\star_{\text{MAP-S}} = \Psi \cdot \arg\max_{\alpha} \mathrm{P}(\alpha \mid y) = \Psi \cdot \arg\min_{\alpha} \|y - \Phi \Psi \alpha\|_2^2 + \lambda \|\alpha\|_1 \ ,$$

  with $\lambda = 2\beta\sigma^2$.

- One possible Bayesian interpretation.

- Recall also that the signal may not be distributed according to the prior but rather ℓ_0-sparse, in which case solving the ℓ_1 problem finds the correct ℓ_0-sparse solution.

A Bayesian perspective

- Consider the inverse problem:
$$\boldsymbol{y} = \Phi \Psi \boldsymbol{\alpha} + \boldsymbol{n} \; .$$

- Assume Gaussian noise, yielding the likelihood:
$$\mathrm{P}(\boldsymbol{y} \mid \boldsymbol{\alpha}) \propto \exp\Bigl(\| \boldsymbol{y} - \Phi \Psi \boldsymbol{\alpha} \|_2^2 / (2\sigma^2) \Bigr) \; .$$

- Consider the Laplacian prior:
$$\mathrm{P}(\boldsymbol{\alpha}) \propto \exp\Bigl(-\beta \| \boldsymbol{\alpha} \|_1 \Bigr) \; .$$

- The maximum *a-posteriori* (MAP) estimate is then
$$\boxed{\boldsymbol{x}^\star_{\text{MAP-S}} = \Psi \cdot \underset{\boldsymbol{\alpha}}{\arg\max}\, \mathrm{P}(\boldsymbol{\alpha} \mid \boldsymbol{y}) = \Psi \cdot \underset{\boldsymbol{\alpha}}{\arg\min}\, \| \boldsymbol{y} - \Phi \Psi \boldsymbol{\alpha} \|_2^2 + \lambda \| \boldsymbol{\alpha} \|_1 \; ,}$$
with $\lambda = 2\beta\sigma^2$.

- One possible Bayesian interpretation.

- Recall also that the signal may not be distributed according to the prior but rather ℓ_0-sparse, in which case solving the ℓ_1 problem finds the correct ℓ_0-sparse solution.

A Bayesian perspective

- Consider the inverse problem:

$$\boldsymbol{y} = \Phi\Psi\boldsymbol{\alpha} + \boldsymbol{n} \ .$$

- Assume Gaussian noise, yielding the likelihood:

$$\mathrm{P}(\boldsymbol{y} \mid \boldsymbol{\alpha}) \propto \exp\left(\|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2^2 / (2\sigma^2) \right) \ .$$

- Consider the Laplacian prior:

$$\mathrm{P}(\boldsymbol{\alpha}) \propto \exp\left(-\beta \|\boldsymbol{\alpha}\|_1 \right) \ .$$

- The maximum *a-posteriori* (MAP) estimate is then

$$\boxed{\boldsymbol{x}^\star_{\text{MAP-S}} = \Psi \cdot \arg\max_{\boldsymbol{\alpha}} \mathrm{P}(\boldsymbol{\alpha} \mid \boldsymbol{y}) = \Psi \cdot \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_1 \ ,}$$

  with $\lambda = 2\beta\sigma^2$.

- One possible Bayesian interpretation.

- Recall also that the signal may not be distributed according to the prior but rather ℓ_0-sparse, in which case solving the ℓ_1 problem finds the correct ℓ_0-sparse solution.

Other Bayesian interpretations

- Other Bayesian interpretations of the synthesis-based approach are also possible (Gribonval 2011).
- Minimum mean square error (MMSE) estimators
  - $\subset$ synthesis-based estimators with appropriate penalty function, *i.e.* penalised least-squares (LS)
  - $\subset$ MAP estimators