Dr. Sidelobes, Or How I Learned To Stop Worrying & Love Simulations

O. Smirnov (Rhodes University & SKA SA)
Radio telescopes, apropos of nothing
Introduction

“A high quality radio map is a lot like a sausage, you might be curious about how it was made, but trust me you really don't want to know.” – Jack Hickish

- Interferometry is hard and unintuitive, but we've carved out some comfort zones
- New observatories and new surveys will be [are] boldly going where no man has made sausage before
- ...as we plunge into the SKA/exascale era
Radio Interferometer...

(In celebration of the passing of an extremely lame but blissfully short-lived internet meme)
Radio Interferometer...
(In celebration of the passing of an extremely lame but blissfully short-lived internet meme)

What lay people think I do
Radio Interferometer...

(In celebration of the passing of an extremely lame but blissfully short-lived internet meme)

What lay people think I do

What funding agencies think I do
Radio Interferometer...

(In celebration of the passing of an extremely lame but blissfully short-lived internet meme)

What lay people think I do

What funding agencies think I do

What cosmologists & astrophysicists think I do
Radio Interferometer...

(In celebration of the passing of an extremely lame but blissfully short-lived internet meme)

What lay people think I do

What funding agencies think I do

What cosmologists & astrophysicists think I do

What my engineers think I do
Radio Interferometer...

(In celebration of the passing of an extremely lame but blissfully short-lived internet meme)

What lay people think I do

What funding agencies think I do

What cosmologists & astrophysicists think I do

What my engineers think I do

What I actually do
How To Make An Interferometer 1

- Start with a normal reflector telescope....
- Then break it up into sections...
How To Make An Interferometer 3

- Replace the optical path with electronics
How To Make An Interferometer 4

- Move the electronics outside the dish
- ...and add cable delays
Why not drop the pieces onto the ground?
How To Make An Interferometer 6

- ...all of them
How To Make An Interferometer 7

- And now replace them with proper radio dishes.
- ...and that's all! (?)
- Well almost, what about the other pixels?
How Does Optical Imaging Do It?

This bit sees the EMF from all directions, added up together.

\[ \iint S(l, m) e^{i(ul+vm)} \, dl \, dm \]

This bit sees the EMF from all parts of the dish surface, added up together.
Fourier Transforms

- An optical imaging system implicitly performs two Fourier transforms:
  - 1. Aperture EMF distribution = FT of the sky
  - 2. Focal plane = FT\(^{-1}\) of the aperture EMF

- A radio interferometer array measures (1)
  - Then we do the second FT in software
  - Hence, “aperture synthesis” imaging
The $uv$-Plane

- In a sense, the two are entirely equivalent

One baseline samples one visibility at a time
Where's The Catch?

- We don't measure the full $uv$-plane, thus we can never recover the image fully (missing information).

- Every visibility measurement is distorted (complex receiver gains, etc.), needs to be calibrated.
Catch 1: Missing Information

- Response to a point source: Point Spread Function (PSF)
- PSF = FT(uv-coverage)
- Observed “dirty image” is convolved with the PSF
- Structure in the PSF = uncertainty in the flux distribution (corresponding to missing data in the uv-plane)

PSF of WSRT. The regular rings are due to the regular spacing of its antennas in the East-West direction.

PSF of MeerKAT
Dr. Sidelobes
The Villain of The Piece

PSF *sidelobes* extend over the entire sky
Deconvolution: from dirty to clean images

- Dirty image dominated by PSF **sidelobes** from the stronger sources
- Deconvolution required to get at the faint stuff underneath.

- A whole continuum of skies fits the dirty image (pick any value for the missing $uv$-components)
- Deconvolution picks one = interpolates the missing info from extra assumptions (e.g.: “sources are point-like”).
Catch 2: Measurement Errors

- Incoming signal is subject to distortions (refraction, delay, amplitude loss)
  - atmospheric and electronic
An Uncalibrated Interferometer

- Complex gain error: signal multiplied by a amplitude and phase delay term
- Delay errors correspond to differences in arrival time, i.e. random shifts of antennas towards and away from the source
- Amplitude errors = different sensitivities
...And Its Optical Equivalent
And The Result...

- One point-like source, but observed with phase errors
- In the $uv$-plane, phase encodes information about location
- **Phase errors** tend to spread the flux around
- **Amplitude errors** distort structure
- And **Dr Sidelobes** ensures that the damage is distributed democratically
How To Deal With This?

- Radio interferometers can be calibrated
- This is done by modelling the instrument and fitting this to the observed visibilities
  - Even possible to fit for the instrument and the sky together (in alternating steps) = selfcal
- Need a parameterized model, i.e. an equation describing the measurement
Radio Interferometer Measurement Equation (RIME)

\[
R_j R_k^* = G_{Rj} G_{Rk}^* \left[ E_{Rj} E_{Rk}^* e^{-i(\phi_j - \phi_k)} + D_{Rk}^* E_{Rj} E_{Lk}^* e^{-i(\phi_j + \phi_k)} + \\
D_{Rj} E_{Lj} E_{Rk}^* e^{i(\phi_j + \phi_k)} + D_{Rj} D_{Rk}^* E_{Lj} E_{Lk}^* e^{i(\phi_j - \phi_k)} \right]
\]

\[
R_j L_k^* = G_{Rj} G_{Lk}^* \left[ E_{Rj} E_{Lk}^* e^{-i(\phi_j + \phi_k)} + D_{Lk}^* E_{Rj} E_{Rk}^* e^{-i(\phi_j - \phi_k)} + \\
D_{Rj} E_{Lj} E_{Lk}^* e^{i(\phi_j - \phi_k)} + D_{Rj} D_{Lk}^* E_{Lj} E_{Rk}^* e^{i(\phi_j + \phi_k)} \right].
\]
Radio Interferometer Measurement Equation (RIME)

\[ R_j R_k^* = G_R_j G_R_k^* \left[ E_R_j E_R_k^* e^{-i(\phi_j - \phi_k)} + D_{R_k}^* E_R_j E_{L_k}^* e^{-i(\phi_j + \phi_k)} + D_{R_k} E_R_j E_{L_k}^* e^{i(\phi_j - \phi_k)} \right] \]

\[ D_{R_j} E_{L_j} E_{R_k}^* e^{i(\phi_j + \phi_k)} + D_{R_j} D_{R_k} E_{L_j} E_{L_k}^* e^{i(\phi_j - \phi_k)} \]

\[ R_j L_k^* = G_R_j G_L_{L_k} \left[ E_R_j E_{L_k}^* e^{-i(\phi_j + \phi_k)} + D_{L_k}^* E_R_j E_{R_k}^* e^{-i(\phi_j - \phi_k)} + D_{L_k} E_R_j E_{R_k}^* e^{i(\phi_j + \phi_k)} \right] \]

\[ V_{pq} = G_p E_p X E_q^H G_q^H \]
Radio Interferometer Measurement Equation (RIME)

\[
R_{jk} R_{kj}^* = G_{Rj} G_{Rk}^* \left[ E_{Rj} E_{Rk}^* - i(\phi_j - \phi_k) \right]
\]

\[
D_{Rj} E_{Lj} E_{Rk}^* \left[ E_{Rj} E_{Lk}^* - i(\phi_j + \phi_k) \right] + D_{Lk} E_{Rk} E_{Lj}^* \left[ E_{Rj} E_{Lk}^* - i(\phi_j - \phi_k) \right]
\]

\[
V_{\rho q} = G_{\rho p} E_{\rho p} X E_{q}^H G_{q}^H
\]

Jones matrices
(these describe the instrument)

...more Jones matrices

Coherency matrices
(these describe the sky)
The Devil Is In the Details!

“The Devil In E-Jones”

A lot of instrumental subtlety

Also need more sophisticated sky models

\[ R_j R^*_k = G_{Rj} G^*_{Rk} \left[ E_{Rj} E^*_{Rk} \right] \]

\[ D_{Rj} E_{Lj} E^*_{Rk} e^{i(\phi_j - \phi_k)} \]

\[ R_j L^*_k = G_{Rj} G^*_{Lk} \left[ E_{Rj} E^*_{Lk} \right] \]

\[ D^*_{Lk} E_{Lk} e^{i(\phi_j - \phi_k)} \]

\[ V_{pq} = G_p E_p X E^H_q G^H_q \]
The Past: Paranoid Overengineering ("Instrumental stability at any cost!")
The Present: Four Sticks In The Ground ("We'll just fix the rest in software")
The Future: Telescopes Made Of Cheap Junk
The Future: Telescopes Made Of Cheap Junk

The lonesome death of instrumental stability
Why MeqTrees

- Future instruments need more flexible and subtle instrumental models
- MeqTrees is a software system for implementing and running Measurement Equations (hence “Meq-”)
- Interferometry simulations
  - Science skies + instrumental model $\rightarrow$ simulated visibilities
- Calibration = simulator plugged into a solver
  - Fit parameterized models to observed visibilities
Trees = Expression Trees

- Any mathematical expression can be represented by a tree:

\[ f = \alpha \sin(b \times x + c \times y + 1) \]
US Navy-Approved

First use: Hopper 1948

“Nobody believed that I had a running compiler and nobody would touch it. They told me computers could only do arithmetic.”
Grace Hopper
MeqTrees Architecture

TDL modules
(libraries)
(Python)

Config
file(s)

Definition of computational structure

Control commands

Results

Meqserver
(compute engine)
(mostly C++)

data
(MS)

TDL script
(Python)

meqbrowser (Python GUI) or Python session or script

α

sin

b

x

c

y

1

x

* 

+ 

* 

α

* 

sin
A Library Of RIME Components

- TDL scripts essentially specify the structure of RIMEs on the meqserver side
- Modules provide many “pre-cooked” components
Instrumental subtleties can be implemented very rapidly (10s of lines of Python vs. 100s of lines of C++)

Without necessarily sacrificing performance

For extra flexibility, particularly tricky nodes can be prototyped in Python

(and historically, none of these “prototypes” have ever needed to be rewritten)
"But What About The Overhead?"

- TDL script (Python)
- Config file(s)

Model definition

- Meqserver (compute engine) (mostly C++)

Python MCMC script

- $\alpha$ (compute engine)
- $\sin$
- $b x^c y^n$

- x1,000,000 times
- Parameter updates
- Predicted visibilities
Wither MeqTrees

- There are simulation capabilities in other packages (CASA, MIRIAD, etc.)
- MeqTrees provides unmatched flexibility in terms of specifying instrumental effects
- When the going gets weird, the weird install† MeqTrees

(†and the truly weird install it on a Mac)
Example Applications

- High-DR calibration & imaging (DD calibration, etc.)
- Evaluation of beam-related effects
  - Prime focus vs offset Gregorian performance
  - Element gain drifts in PAFs
  - Fundamental sensitivity limits due to beam instability
- Ionosphere and EoR sims (see talk by H. Shukla)
- Weak lensing simulations (incl. SKA1)
- MC & Bayesian sampling
- Generating training data for ML
Good News First: Radio Interferometry Is Possible...

3C147 @21cm
12h WSRT synthesis
160 MHz bandwidth

22 Jy peak (3C147)
13.5 µJy noise
1,600,000:1 DR
thermal noise-limited

Regular calibration does not reach the noise, leaves off-axis artefacts due to direction-dependent effects (left inset)

Addressed via differential gains (right inset)
Luxury Problem?

- High-DR imaging has been called many names
  - “Marketing gimmick”
  - “Macho imaging”
  - “Bragging rights”
  - “Luxury problem”
    - (...well we only came up with that one to annoy the NRAO folks)
Not Always a Luxury Problem
(Courtesy of Ian Heywood)

EVLA 8 GHz: Looking for sub-mm galaxies and QSOs in the William Herschel Deep Field.

Dominant effect: bright calibrator source rotating through first sidelobe of the primary beam.

(This also has a horrible PSF, being an equatorial field.)

This is your phase calibrator

This is your science (good luck!)

Brightness scale 0~50μJy
Keep Your Friends Close, and your calibrators as far away as you can...

An approximation of the primary beam response, overlaid on top of the image.

As the sky rotates, the sidelobes of the PB sweep over the source, thus making it effectively *time-variable*.

(Brightness scale 0~50μJy)
Deconvolution Doesn't Help...

Residual image, after deconvolution.

The contaminating source cannot be deconvolved away properly, due to its **instrumental** time-variability.

...5 years ago this would observation would probably be a complete write-off.

(Brightness scale 0~50μJy)
Differential Gains To The Rescue

Residual image after applying differential gain solutions to the contaminating source

Brightness scale 0~50μJy
Multi-Band Image

Multi-band residual image: noise-limited, no trace of contaminating source.

Brightness scale 0~50μJy

Phase calibrator used to be here
3C147+JVLA: New DR Record

- Collaboration with Rick Perley (NRAO)
- 3C147
- JVLA-C @1.4 GHz
- ~192 MHz bandwidth used
- Best image after regular selfcal
3C147+JVLA: New DR Record

- Best image after regular selfcal
- DD calibration on a few sources
- DR ~ 3.2 million
3C147+JVLA: New DR Record

- Best image after regular selfcal
- DD calibration on a few sources
- DR ~ 3.2 million
High DR In The SKA Era

- Today's extreme-DR is tomorrow's routine data reduction
- All current showcase images produced in “hero mode”
  - See e.g. talk by Pandey
- This processing model does not scale!
Exascale = Exapostdocs?
Obvious ML Application

Stage 1 calibration: secondary sources generate artefacts

Stage 2: shallow source finding and sky model update
Obvious ML Application

Stage 3: designate troublesome sources, apply direction-dependent solutions

Stage 4: Deeper source finder, update sky model, rinse & repeat...
Moving Forward

- Current implementation is thread-level parallelized
- Relatively efficient on a single machine
  - Can also provide GPU implementations of “critical” nodes
- Primitive MPI support
  - Trees can span clusters
  - Problem is, how to do this efficiently?

\[
\alpha \sin b x + c y + 1
\]
Self-optimizing Computation?

- TDL scripts define the high-level ("human friendly") structure of the computation
- The tree (graph) is a low-level machine-friendly abstraction
  - High parallelism in the tree itself (besides the obvious data parallelism)
- Cost of computation and data transport in any given tree is knowable
- The Holy Grail: automatically map a tree to any given cluster architecture in a reasonably efficient way?
I have no answers, only questions
But they're good questions!

“May you live in interesting times.”
– wrongly attributed to ancient Chinese

“You may not be interested in the polarization,
but the polarization is interested in you.”
– wrongly attributed to Leon Trotsky
Thank You & Goodbye