Longitudinal Spacekime Analytics:
Time Complexity & Inferential Uncertainty

Ivo D. Dinov

Statistics Online Computational Resource
Health Behavior & Biological Sciences
Computational Medicine & Bioinformatics
Neuroscience Graduate Program
Michigan Institute for Data Science
University of Michigan
http://SOCR.umich.edu

Joint work with Milen V. Velev (BTU)

Based on an upcoming book “Data Science: Time Complexity and Inferential Uncertainty”

Outline

- Big Biomedical/Health Data Analytic Challenges
- Complex-Time \((kime)\)
- Spacekime Mathematical Foundation
- Statistical Implications of Spacekime Analytics
  - Inferential Uncertainty
  - Bayesian Inference Representation
- Applications – Longitudinal Spacekime Data Analytics
  - Neuroimaging (UKBB, fMRI)
  - Air quality (UCI ML Air Quality Dataset)
Big Biomedical & Health
Data Analytic Challenges

Data Analytics = Information Compression

- From 23 ... to ... $2^{23}$ (10M) \( \left( \frac{23}{2^{13}} \Rightarrow \frac{2^{23}}{8^{13}} \right) \)
- Data Science: 1798 vs. 2019
- In the 18th century, Henry Cavendish used just 23 observations to answer a fundamental question – “What is the Mass of the Earth?” He estimated very accurately the mean density of the Earth/H$_2$O (5.483±0.1904 g/cm$^3$)
- In the 21st century to achieve the same scientific impact, matching the reliability and the precision of the Cavendish’s 18th century prediction, requires a monumental community effort using massive and complex information perhaps on the order of 10M ($2^{23}$) bytes

Dinov (2016) JSMI
Characteristics of Big Biomed Data

IBM Big Data 4V’s: Volume, Variety, Velocity & Veracity

<table>
<thead>
<tr>
<th>Big Bio Data Dimensions</th>
<th>Tools</th>
<th>Example: analyzing observational data of 1,000’s Parkinson’s disease patients based on 10,000’s signature biomarkers derived from multi-source imaging, genetics, clinical, physiologic, phenomics and demographic data elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>Harvesting and management of vast amounts of data</td>
<td>Software developments, student training, service platforms and methodological advances associated with the Big Data Discovery Science all present existing opportunities for learners, educators, researchers, practitioners and policy makers</td>
</tr>
<tr>
<td>Complexity</td>
<td>Wranglers for dealing with heterogeneous data</td>
<td></td>
</tr>
<tr>
<td>Incongruency</td>
<td>Tools for data harmonization and aggregation</td>
<td></td>
</tr>
<tr>
<td>Multi-source</td>
<td>Transfer and joint modeling of disparate elements</td>
<td></td>
</tr>
<tr>
<td>Multi-scale</td>
<td>Macro → meso → micro → nano scale observations</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Techniques accounting for longitudinal effects</td>
<td></td>
</tr>
<tr>
<td>Incomplete</td>
<td>Reliable management of missing data</td>
<td></td>
</tr>
</tbody>
</table>

Dinov, GigaScience (2016) PMID:26918190
Data Science & Predictive Analytics

- **Data Science**: an emerging extremely transdisciplinary field - bridging between the theoretical, computational, experimental, and applied areas. Deals with enormous amounts of complex, incongruent and dynamic data from multiple sources. Aims to develop algorithms, methods, tools, and services capable of ingesting such datasets and supplying semi-automated decision support systems.

- **Predictive Analytics**: process utilizing advanced mathematical formulations, powerful statistical computing algorithms, efficient software tools, and distributed web-services to represent, interrogate, and interpret complex data. Aims to forecast trends, cluster patterns in the data, or prognosticate the process behavior either within the range or outside the range of the observed data (e.g., in the future, or at locations where data may not be available).

http://DSPA.predictive.space

Longitudinal Data Analytics

- **Neuroimaging**:
  - **4D fMRI**: time-series, represents measurements of hydrogen atom densities over a 3D lattice of spatial locations (1 ≤ x, y, z ≤ 64 pixels), about 3 × 3 millimeters² resolution. Data is recorded longitudinally over time (1 ≤ t ≤ 180) in increments of about 3 seconds & post-processing.
  - **State-of-the-art Approaches**: Time-series modeling or Network analysis.
  - **Spacekime Analytics**: 5D fMRI kime-series, representing the of hydrogen atom densities over the same 3D lattice of spatial locations, longitudinally over the 2D kime space, \( \kappa = re^{ip} \in \mathbb{C} \).

- **Differences**: Spacekime analytics estimate and utilize the kime-phases.

Dinov & Velev (2020)
Preview: Get to UKBB Big Data Analytics Study, but first we need some background …

9,914 UKBB participants; 7,614 clinical measurements, phenotypic features, and derived neuroimaging biomarkers. Supervised Decision Tree (binary) Classification
The Fourier Transform

By separability, the classical **spacetime Fourier transform** is just four Fourier transforms, one for each of the four spacetime dimensions, \((x, t) = (x, y, z, t)\). The FT is a function of the **angular frequency** \(\omega\) that propagates in the wave number direction \(k\) **(space frequency)**. Symbolically, the forward and inverse Fourier transforms of a 4D \((n = 4)\) spacetime function \(f\), are defined by:

\[
FT(f) = \hat{f}(k, \omega) = \frac{1}{(2\pi)^2} \int f(x, t) e^{i(\omega t - kx)} dt dx,
\]

\[
IFT(\hat{f}) = \hat{f}(x, t) = \frac{1}{(2\pi)^2} \int \hat{f}(k, \omega) e^{-i(\omega t - kx)} d\omega dk.
\]

\[
[f(x, t) = IFT(\hat{f}) = IFT(FT(f)) = f(x, t)]
\]

---

**1D Fourier Transform Example**

SOCR 1D Fourier / Wavelet signal decomposition into magnitudes and phases (Java applet)

Top-panel: original signal (image), white-color curve drawn manually by the user and the reconstructed synthesized (IFT) signal, red-color curve, computed using the user modified magnitudes and phases

Bottom-panels: the Fourier analyzed signal (FT) with its magnitudes and phases

[http://www.socr.ucla.edu/htmls/game/Fourier_Game.html](http://www.socr.ucla.edu/htmls/game/Fourier_Game.html) (Java Applet)
2D Fourier Transform – The Importance of Magnitudes & Phases

Fourier Analysis
(real part of the Forward Fourier Transform)

<table>
<thead>
<tr>
<th>Square Image Shape</th>
<th>Disk Image Shape</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="square" alt="2D image 1 (square)" /></td>
<td><img src="disc" alt="2D image 2 (disk)" /></td>
</tr>
<tr>
<td>Re(FT(square)) Magnitude</td>
<td>Re(FT(disc)) Magnitude</td>
</tr>
<tr>
<td>FT(square) Phase</td>
<td>FT(disc) Phase</td>
</tr>
</tbody>
</table>

Fourier Synthesis
(real part of the Inverse Fourier Transform)

<table>
<thead>
<tr>
<th>Square Image Shape</th>
<th>Disk Image Shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>IFT(FT(square)) ≡ square</td>
<td>IFT(FT(disc)) ≡ disc</td>
</tr>
<tr>
<td>IFT using square-magnitude &amp; disc-phase</td>
<td>IFT using disc-magnitude &amp; square-phase</td>
</tr>
<tr>
<td>IFT using square-magnitude &amp; nil-phase</td>
<td>IFT using disc-magnitude &amp; nil-phase</td>
</tr>
</tbody>
</table>

Kaluza-Klein Theory

- Theodor Kaluza developed (1921) an extension of the classical general relativity theory to 5D. This included the metric, the field equations, the equations of motion, the stress-energy tensor, and the cylinder condition. Oskar Klein (1926) interpreted Kaluza’s 3D+2D theory in quantum mechanical space and proposed that the fifth dimension was curled up and microscopic.

- The topology of the 5D Kaluza-Klein spacetime is $K_2 \cong M^4 \times S^1$, where $M^4$ is a 4D Minkowski spacetime and $S^1$ is a circle (non-traversable).

$$K_2 \cong M^4 \times S^1$$
Complex-Time (Kime)

- At a given spatial location, \( x \), complex time (kime) is defined by \( k = re^{\imath \phi} \in \mathbb{C} \), where:
  - the magnitude represents the longitudinal events order \( (r > 0) \) and characterizes the longitudinal displacement in time, and
  - event phase \( (\pi \leq \phi < \pi) \) is an angular displacement, or event direction
- There are multiple alternative parametrizations of kime in the complex plane
- Space-kime universe \( (\mathbb{R}^3 \times \mathbb{C}) \):
  - \((x, k1)\) and \((x, k4)\) have the same spacetime representation, but different space-kime coordinates,
  - \((x, k1)\) and \((y, k1)\) share the same kime, but represent different spatial locations,
  - \((x, k2)\) and \((x, k3)\) have the same spatial locations and kime-directions, but appear sequentially in order

The Spacekime Manifold

- Spacekime: \((x, k) = (x^1, x^2, x^3, ck_1 = x^4, ck_2 = x^5) \in \mathcal{X} \), \( c \sim 3 \times 10^8 \text{ m/s} \)
- Kevent (complex events): points (or states) in the spacekime manifold \( X \). Each kevent is defined by where \((x = (x, y, z))\) it occurs in space, what is its causal longitudinal order \((r = \sqrt{(x^4)^2 + (x^5)^2})\), and in what kime-direction \((\phi = \text{atan2}(x^5, x^4))\) it takes place.
- Spacekime interval \((ds)\) is defined using the general Minkowsk 5 × 5 metric tensor \((\lambda_{ij})_{i,j=1,5}\), which characterizes the geometry of the curved spacekime manifold:
  \[
  ds^2 = \sum_{i=1}^{5} \sum_{j=1}^{5} \lambda_{ij} dx^i dx^j
  \]
  Euclidean (flat) spacekime metric corresponds to the tensor: \((\lambda_{ij}) = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & -1
\end{bmatrix}\)
- Spacelike intervals correspond to \(ds^2 > 0\), where an inertial frame can be found such that two kevents \(a, b \in \mathcal{X}\) are simultaneous. An object can’t be present at two kevents which are separated by a spacelike interval.
- Lightlike intervals correspond to \(ds^2 = 0\). If two kevents are on the line of a photon, then they are separated by a lightlike interval and a ray of light could travel between the two kevents.
- Kimelike intervals correspond to \(ds^2 < 0\). An object can be present at two different kevents, which are separated by a kimelike interval.
**Kime Parameterizations**

Conjugate Pairs \( \{x, \bar{z} \in \mathbb{C} \} \)

\( x = (x + \bar{x})/2 \)
\( y = -(x - \bar{x})/2 \)
\( z = x + iy \)
\( \bar{z} = x - iy \)

\( d\ell \) is Lorentz transform invariant

\( r = \sqrt{x^2 + y^2} \)
\( \phi = \arccos \left( \frac{z + \bar{z}}{2(x + \bar{x})} \right) \)

**Spacekime Math Generalizations**

- Kime (Wirtinger) derivative & acceleration (second order kime-derivative at \( k = (r, \phi) \)):

\[
\begin{align*}
\frac{\partial^2 f}{\partial r^2} - \frac{2}{r} \sin(2\phi) \frac{\partial^2 f}{\partial r \partial \phi} - \frac{1}{r^2 \cos(2\phi)} \frac{\partial^2 f}{\partial \phi^2} \\
+ i \left( \sin(2\phi) \frac{\partial^2 f}{\partial r^2} + \frac{2}{r} \cos(2\phi) \frac{\partial^2 f}{\partial r \partial \phi} - \frac{1}{r^2 \sin(2\phi)} \frac{\partial^2 f}{\partial \phi^2} \right)
\end{align*}
\]

- Kime generalization of Lorentz transform between two reference frames, \( K' \) & \( K'' \):

\[
\begin{pmatrix}
\zeta & 0 & 0 & -\frac{v_1^2}{v_2} \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
-\frac{1}{v_2} & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
k_1 \\
k_2
\end{pmatrix} =
\begin{pmatrix}
x' \\
y' \\
k_1' \\
k_2'
\end{pmatrix}
\]

- Derived other spacekime concepts: calculus of differentiation & integration, law of addition of velocities, energy-momentum conservation law, stability conditions for particles moving in spacekime, conditions for nonzero rest particle mass, and causal structure of spacekime...

Dinov & Velev (2020)
Spacekime Foliations

Manifold foliation (spacekime slicing) is a covering decomposition into hypersurfaces of lower dimension (e.g., n-1) paired with a smooth scalar field (regular with non-trivial gradient), so that each hypersurface (leaf) is a level surface of the scalar field.

- Space (x) Foliation of Spacekime:
- (Radial, t) Time-Foliation of Spacekime:
- (Angular, φ) Phase-Foliation of Kime:

Spacekime Connection to Data Analytics?
### Mathematical-Physics ⟹ Data Science

<table>
<thead>
<tr>
<th>Mathematical-Physics</th>
<th>Data Science</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A particle</strong> is a small localized object that permits observations and characterization of its physical or chemical properties</td>
<td>An <strong>object</strong> is something that exists by itself, actually or potentially, concretely or abstractly, physically or incorporeal (e.g., person, subject, etc.)</td>
</tr>
<tr>
<td>An <strong>observable</strong> is a dynamic variable about particles that can be measured</td>
<td>A <strong>feature</strong> is a dynamic variable or an attribute about an object that can be measured</td>
</tr>
<tr>
<td>Particle <strong>state</strong> is an observable particle characteristic (e.g., position, momentum)</td>
<td><strong>Datum</strong> is an observed quantitative or qualitative value, an instantiation, of a feature</td>
</tr>
<tr>
<td>Particle <strong>system</strong> is a collection of independent particles and observable characteristics, in a closed system</td>
<td><strong>Problem</strong>, aka Data System, is a collection of independent objects and features, without necessarily being associated with apriori hypotheses</td>
</tr>
<tr>
<td><strong>Wave-function</strong></td>
<td><strong>Inference-function</strong></td>
</tr>
<tr>
<td>Reference-Frame transforms (e.g., Lorentz)</td>
<td>Data transformations (e.g., wrangling, log-transform)</td>
</tr>
<tr>
<td><strong>State of a system</strong> is an observed measurement of all particles – wavefunction</td>
<td><strong>Dataset</strong> (data) is an observed instance of a set of datum elements about the problem system, ( O = {X,Y} )</td>
</tr>
</tbody>
</table>

A particle system is **computable** if (1) the entire system is logical, consistent, complete and (2) the unknown internal states of the system don’t influence the computation (wavefunction, intervals, probabilities, etc.)

**Wave-function**

- **Inference-function** - describing a solution to a specific data analytic system (a problem). For example:
  - A linear (GLM) model represents a solution of a prediction inference problem, \( Y = X\beta \), where the inference function quantifies the effects of all independent features \( X \) on the dependent outcome \( Y \), data: \( O = \{X,Y\} \):
    \[
    \psi(O) = \psi(X,Y) = \hat{\beta} = \hat{\beta}_{\text{OLS}} = (X'X)^{-1}X'Y.
    \]
  - A non-parametric, non-linear, alternative inference is SVM classification. If \( \psi_x \in H \) is the lifting function \( \psi: \mathbb{R}^q \rightarrow \mathbb{R}^d \) (\( \psi: x \in \mathbb{R}^q \rightarrow x = \psi_x \in H \)), where \( q < d \), the kernel \( \psi(x,y) = \langle x,y \rangle \). If \( O \rightarrow H \) is a transformation, the SVM prediction operator is the weighted sum of the kernel functions at \( \psi_{o_i} \), where \( \beta^* \) is a solution to the SVM regularized optimization:
    \[
    \langle \psi_o | \beta^* \rangle_H = \sum_{i=1}^{n} p^*_i (\langle \psi_o | \psi_{o_i} \rangle_H).
    \]

The linear coefficients, \( p^*_i \), are the dual weights that are multiplied by the label corresponding to each training instance, \( y_i \).

**Inference always depends on the (input) data; however, it does not have 1-1 and onto bijective correspondence with the data, since the inference function quantifies predictions in a probabilistic sense.**

Let's assume that we have:
(1) Kime extension of Time, and
(2) Parallels between wavefunctions ↔ inference functions

Often, we can’t directly observe (record) data natively in 5D spacekime. Yet, we can measure quite accurately the kime-magnitudes ($r$) as event orders, “times”.

To reconstruct the 2D spatial structure of kime, borrow tricks used by crystallographers to resolve the structure of atomic particles by only observing the magnitudes of the diffraction pattern in k-space. This approach heavily relies on (1) prior information about the kime directional orientation (that may be obtained from using similar datasets and phase-aggregator analytical strategies), or (2) experimental reproducibility by repeated confirmations of the data analytic results using longitudinal datasets.

---

### 2D Image Analysis / Character Recognition

<table>
<thead>
<tr>
<th>Kime-direction (Phase) Synthesis</th>
<th>Correct Phase</th>
<th>Swapped Phase</th>
<th>Nil-Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D Images</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyrillic Alphabet</td>
<td>A Б В Г Д Е Ж З И Й К Л М Н О П Р С Т У Ф Х Ц Ч Ш Ь Ь ЮЯ</td>
<td>A Б В Г Д Е Ж З И Й К Л М Н О П Р С Т У Ф Х Ц Ч Ш Ь Ь ЮЯ</td>
<td>A Б В Г Д Е Ж З И Й К Л М Н О П Р С Т У Ф Х Ц Ч Ш Ь Ь ЮЯ</td>
</tr>
<tr>
<td>Observed Data</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dinov & Velev (2020)
Back to fMRI (4D spacetime data)

3D rendering of 3 time cross-sections of the fMRI series over a 2D spatial domain

Spacekime Analytics: fMRI Example

- 3D isosurface Reconstruction of (space=2, time=1) fMRI signal

<table>
<thead>
<tr>
<th>4D spacetime: Reconstruction using trivial phase-angle; kime=time=(magnitude, 0)</th>
<th>5D Spacekime: Reconstruction using correct kime=(magnitude, phase)</th>
</tr>
</thead>
<tbody>
<tr>
<td>f = ( \hat{h}(x_1, x_2, \frac{t}{\text{time}}) )</td>
<td>3D pseudo-spacetime reconstruction:</td>
</tr>
</tbody>
</table>

```markdown
\[
f = \hat{h}(x_1, x_2, \frac{t}{\text{time}})
\]```

3D rendering of 3 time cross-sections of the fMRI series over a 2D spatial domain.
**Spacekime Analytics:**

Kime-series = Surfaces (not curves)

In the 5D spacekime manifold, time-series curves extend to kime-series, i.e., surfaces parameterized by kime-magnitude ($t$) and the kime-phase ($\phi$).

Kime-phase aggregating operators that can be used to transform standard time-series curves to spacekime kime-surfaces, which can be modeled, interpreted, and predicted using advanced spacekime analytics.

---

**Spacekime Analytics: fMRI kime-series**

fMRI kime-series at a single spatial voxel location (rainbow color represents fMRI kime intensities)

Kime-Foliation

Specific 1D time-series are projections of kime-series (red & blue curves)

Top view

Side view
Spacekime Analytics: fMRI Example

Reconstruction of the fMRI timeseries at a single spatial voxel location

\[ \text{Cor(Orig, Nil-Phase)} = 0.16 \]
\[ \text{Cor(Orig, Estim-Phase)} = 0.79 \]

Statistical Implications of Spacekime Analytics
Uncertainty

- **Quantum Mechanics**: \[ | \frac{\hbar}{i} \partial_x u | i \partial_x u \rangle = \frac{\hbar}{2} | u \rangle \] is non-commutation of the unbounded operators \( D_x = \frac{\hbar}{i} \partial_x \) and \( x \), (multiplication by \( x \)).

- **Signal processing**: Functions can’t be time-limited and band-limited. Alternatively, a function and its Fourier transform cannot both have bounded domains \( \sigma_t \times \sigma_\omega \geq 1/(4\pi) \), where \( \sigma_t, \sigma_\omega \) are the time and frequency SDs.

- **Entropic uncertainty**: Entropy can be used just like the SD to quantify distribution structure. For instance, for angular, bimodal, or divergent-variance distributions, Entropy may be a better measure of dispersion than SD. For \( F{T}(f)(\omega) = \hat{f}(\omega) \) and \( I\hat{F}(\hat{f}(\omega)) = \hat{f}(x) \), the Shannon information entropies:
  \[ H_x = \int \hat{f}(x) \log \left( \hat{f}(x) \right) dx \] and \( H_\omega = \int \hat{f}(\omega) \log \left( \hat{f}(\omega) \right) d\omega \).

  \[ H_x + H_\omega \leq \log(e/2). \]

- **\( L^2(\mathbb{R}) \)** uncertainty: It is impossible for \( f \in L^2 \) and \( \hat{f} \) to both decrease extremely rapidly. If both have rapidly decreasing tails: \( |f(x)| \leq C(1 + |x|)^n e^{-ax^2} \) and \( |\hat{f}(\omega)| \leq C(1 + |\omega|)^n e^{-b\omega^2} \), for some constant \( C \), polynomial power \( n \), and \( a, b \in \mathbb{R} \), then \( f = 0 \) (when \( ab > 1 \)); \( f(x) = P_n(x) e^{-ax^2} \) and \( \hat{f}(\omega) = \hat{P}_n(\omega) e^{-b\omega^2} \), where \( \deg(P_n) \leq n \) (when \( ab = 1 \)); or (when \( ab < 1 \)).

Heisenberg’s Uncertainty in Spacekime?

- **Heisenberg’s uncertainty is resolved in 5D spacekime**

  - We can derive the classical 4D spacetime Heisenberg uncertainty as a reduction of Einstein-like 5D deterministic dynamics:

  - The math is terse – it involves deriving the equations of motion by maximizing the distance (integral along the geodesic) between two points in 5D spacekime

  - The inner product \( du^\mu dx_\mu = \frac{du^\mu dx_\mu}{\sqrt{-g}} \) since \( \frac{du^\mu}{\sqrt{-g}} = 1 \) near the leaf membrane, \( du^\mu dx_\mu = L = \frac{\hbar}{mc} \).

  - Replacing the change in velocity \( \frac{du^\mu}{\sqrt{-g}} \) by the change in momentum \( dp^\mu \) yields: \( dp^\mu dx_\mu = \hbar \).

  - This relation is similar to the quantum mechanics uncertainty principle in 4D Minkowski spacetime; however, it is obtained from 5D Einstein deterministic dynamics. In other words, in spacetime, Heisenberg’s uncertainty principal manifests simply because of the one degree of freedom (kime-phase), i.e., lack of sufficient information about the second kime dimension.

  - In 5D spacetime, the conventional geodesic motion is perturbed by an extra force \( f^\mu \) that can be split into two parts \( f^\mu = f^\mu_{\parallel} + f^\mu_{\perp} \). The normal component \( f^\mu_{\parallel} \) is similar to other conventional forces and obeys the usual orthogonality condition \( f^\mu_{\parallel} u_\mu = 0 \). However, the parallel component \( f^\mu_{\parallel} \) has no analog in 4D spacetime. In general, it has a non-trivial inner product with the 4-velocity \( u^\mu, f^\mu_{\parallel} u^\mu \neq 0 \).

  - In Minkowski 4D spacetime, the lack of kime-phase data naturally leaves one degree of freedom in the system causing Heisenberg’s uncertainty. However, the latter can be explicated by information knowledge of the fifth component (kime-phase).
Bayesian Inference Representation

- Suppose we have a single spacetime observation \( X = \{x_i\} \sim p(x | y) \) and \( y \sim p(y | \varphi = \text{phase}) \) is a process parameter (or vector) that we are trying to estimate.

- Spacekime analytics aims to make appropriate inference about the process \( X \).

- The sampling distribution, \( p(x | y) \), is the distribution of the observed data \( X \) conditional on the parameter \( y \) and the prior distribution, \( p(y | \varphi) \), of the parameter \( y \) before the data \( X \) is observed, \( \varphi = \text{phase aggregator} \).

- Assume that the hyperparameter (vector) \( \varphi \), which represents the kime-phase estimates for the process, can be estimated by \( \varphi = \varphi' \).

- Such estimates may be obtained from an oracle, approximated using similar datasets, acquired as phases from samples of analogous processes, or derived via some phase-aggregation strategy.

- Let the posterior distribution of the parameter \( y \) given the observed data \( X = \{x_i\} \) be \( p(y | X, \varphi') \) and the process parameter distribution of the kime-phase hyperparameter vector \( \varphi \) be \( y \sim p(y | \varphi) \).

Bayesian Inference Representation

- We can formulate spacekime inference as a Bayesian parameter estimation problem:

\[
\frac{p(y | X, \varphi)}{p(X, \varphi)} = \frac{p(y | X, \varphi')}{p(X, \varphi')} = \frac{p(X | y, \varphi') \times p(y, \varphi')}{p(X | \varphi') \times p(\varphi')} = \frac{p(X | y, \varphi') \times p(y, \varphi')}{p(X | \varphi') \times p(\varphi')} = \frac{p(X | y, \varphi') \times p(y, \varphi')}{p(X | \varphi') \times p(\varphi')}. 
\]

- In Bayesian terms, the posterior probability distribution of the unknown parameter \( y \) is proportional to the product of the likelihood and the prior.

- In probability terms, the posterior = likelihood times prior, divided by the observed evidence, in this case, a single spacetime data point, \( x_{i0} \).
Bayesian Inference Representation

- Spacekime analytics based on a single spacetime observation $x_{i_0}$ can be thought of as a type of Bayesian prior-predictive or posterior-predictive distribution estimation problem.

- Prior predictive distribution of a new data point $x_{i_0}$, marginalized over the prior - i.e., the sampling distribution $p(x_{i_0}|\gamma)$, weight-averaged by the pure prior distribution:

$$p(x_{i_0}|\varphi) = \int p(x_{i_0}|\gamma) \times p(\gamma|\varphi) \, d\gamma.$$ 

- Posterior predictive distribution of a new data point $x_{i_0}$, marginalized over the posterior; i.e., the sampling distribution $p(x_{i_0}|\gamma)$, weight-averaged by the posterior distribution:

$$p(x_{i_0}|x_{i_0}, \varphi) = \int p(x_{i_0}|\gamma) \times p(\gamma|x_{i_0}, \varphi) \, d\gamma.$$ 

- The difference between these two predictive distributions is that
  - the posterior predictive distribution is updated by the observation $X = \{x_{i_0}\}$ and the hyperparameter, $\varphi$ (phase aggregator),
  - whereas the prior predictive distribution only relies on the values of the hyperparameters that appear in the prior distribution.

Bayesian Inference Representation

- The posterior predictive distribution may be used to sample or forecast the distribution of a prospective, yet unobserved, data point $x_{i_0}$.

- The posterior predictive distribution spans the entire parameter state-space (Domain($\gamma$)), just like the wavefunction represents the distribution of particle positions over the complete particle state-space.

- Using maximum likelihood or maximum a posteriori estimation, we can also estimate an individual parameter point-estimate, $\gamma_0$. In this frequentist approach, the point estimate may be plugged into the formula for the distribution of a data point, $p(x|\gamma_0)$, which enables drawing IID samples or individual outcome values.
Bayesian Inference Simulation

- Simulation example using 2 random samples drawn from mixture distributions each of $n_A = n_B = 10K$ observations:
  - $\{X_{A,i}\}_{i=1}^{n_A}$, where $X_{A,i} = 0.3U_i + 0.7V_i, U_i \sim N(0,1)$ and $V_i \sim N(5,3)$, and
  - $\{X_{B,i}\}_{i=1}^{n_B}$, where $X_{B,i} = 0.4P_i + 0.6Q_i, P_i \sim N(20,20)$ and $Q_i \sim N(100,30)$.

- The intensities of cohorts $A$ and $B$ are independent and follow different mixture distributions. We’ll split the first cohort ($A$) into training ($C$) and testing ($D$) subgroups, and then:
  - Transform all four cohorts into Fourier k-space,
  - Iteratively randomly sample single observations from cohort $C$,
  - Reconstruct the data into spacetime using a single kime-magnitude value and alternative kime-phase estimates derived from cohorts $B$, $C$, and $D$, and
  - Compute the classical spacetime-derived population characteristics of cohort $A$ and compare them to their spacekime counterparts obtained using a single $C$ kime-magnitude paired with $B$, $C$, or $D$ kime-phases.

Bayesian Inference Simulation

Summary statistics for the original process (cohort $A$) and the corresponding values of their counterparts computed using the spacekime reconstructed signals based on kime-phases of cohorts $B$, $C$, and $D$. The estimates for the latter three cohorts correspond to reconstructions using a single spacetime observation (i.e., single kime-magnitude) and alternative kime-phases (in this case, kime-phases derived from cohorts $B$, $C$, and $D$).

<table>
<thead>
<tr>
<th>Summaries</th>
<th>Spacetime (A)</th>
<th>Spacekime Reconstructions (single kime-magnitude)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Original</td>
<td>Phase-True</td>
</tr>
<tr>
<td>Min</td>
<td>-2.38798</td>
<td>-2.98116</td>
</tr>
<tr>
<td>1st Quartile</td>
<td>-0.89359</td>
<td>-0.76765</td>
</tr>
<tr>
<td>Median</td>
<td>0.03111</td>
<td>-0.05982</td>
</tr>
<tr>
<td>Mean</td>
<td>0.00000</td>
<td>0.00000</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>0.75772</td>
<td>0.72795</td>
</tr>
<tr>
<td>Max</td>
<td>3.61346</td>
<td>3.64800</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.348269</td>
<td>0.2372526</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.68176</td>
<td>-0.4452207</td>
</tr>
</tbody>
</table>
Bayesian Inference Simulation

The correlation between the original data \(A\) and its reconstruction using a single kime magnitude and the correct kime-phases \(C\) is \(\rho(A, C) = 0.89\).

This strong correlation suggests that a substantial part of the \(A\) process energy can be recovered using only a single observation. In this case, to reconstruct the signal back into spacetime and compute the corresponding correlation, we used a single kime-magnitude (sample-size=1) and process \(C\) kime-phases.

Let's demonstrate the Bayesian inference corresponding to this spacekime data analytic problem using a simulated bimodal experiment:

\[ X_A = 0.3U + 0.7V, \text{ where } U \sim N(0,1) \text{ and } V \sim N(5,3) \]

Specifically, we will illustrate the Bayesian inference using repeated single spacetime observations from cohort \(A\), \(X = \{x_i\}\), and varying kime-phase priors \((\theta = \text{phase aggregator})\) obtained from cohorts \(B\), \(C\), or \(D\), using different posterior predictive distributions.

Relations between the empirical data distribution (dark blue) and samples from the posterior predictive distribution, representing Bayesian simulated spacekime reconstructions (light-blue). The derived Bayesian estimates do not perfectly match the empirical distribution of the simulated data, yet there is clearly information encoding that is captured by the spacekime data reconstructions.

This signal compression can be exploited by subsequent model-based or model-free data analytic strategies for retrospective prediction, prospective forecasting, ML classification, derived clustering, and other spacekime inference methods.
Bayesian Inference Simulation

Applications – Longitudinal Spacekime Data Analytics
Exogenous Feature Time-series Analysis

ARIMAX modeling of UCI ML Air Quality Dataset (9,358 hourly-averaged CO responses from an array of sensors). Demonstrate the effect of kime-direction on the analysis of the longitudinal data.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Nil</th>
<th>Average</th>
<th>True=original</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Estimate</td>
<td>ARIMA(2,0,1)</td>
<td>ARIMA(2,0,3)</td>
<td>ARIMA(1,1,4)</td>
</tr>
<tr>
<td>AIC</td>
<td>13179</td>
<td>14183</td>
<td>10581</td>
</tr>
<tr>
<td>ar1</td>
<td>1.11406562</td>
<td>0.329482302</td>
<td>0.275312</td>
</tr>
<tr>
<td>ar2</td>
<td>-0.14565048</td>
<td>0.238363531</td>
<td>0.238363531</td>
</tr>
<tr>
<td>ma1</td>
<td>-0.78919188</td>
<td>0.267295185</td>
<td>-0.88913497</td>
</tr>
<tr>
<td>ma2</td>
<td>-0.006079386</td>
<td>0.12679494</td>
<td>0.12679494</td>
</tr>
<tr>
<td>ma3</td>
<td>0.13726556</td>
<td>0.03043726</td>
<td>0.03043726</td>
</tr>
<tr>
<td>ma4</td>
<td>-0.17655728</td>
<td>-0.017655728</td>
<td>-0.017655728</td>
</tr>
<tr>
<td>intercept</td>
<td>742.800113</td>
<td>-0.08035744</td>
<td>-0.08035744</td>
</tr>
<tr>
<td>xreg1</td>
<td>-0.40283891</td>
<td>0.58379483</td>
<td>0.58379483</td>
</tr>
<tr>
<td>xreg2</td>
<td>0.13656613</td>
<td>0.28093931</td>
<td>6.14947902</td>
</tr>
<tr>
<td>xreg3</td>
<td>-0.51457636</td>
<td>-0.64972755</td>
<td>0.09859223</td>
</tr>
<tr>
<td>xreg4</td>
<td>1.09611981</td>
<td>1.239910298</td>
<td>0.01634736</td>
</tr>
<tr>
<td>xreg5</td>
<td>1.21946209</td>
<td>-0.026110332</td>
<td>-0.04816591</td>
</tr>
<tr>
<td>xreg6</td>
<td>1.30628469</td>
<td>1.08177756</td>
<td>-0.01104142</td>
</tr>
<tr>
<td>xreg7</td>
<td>1.2086397</td>
<td>0.254018471</td>
<td>0.1832854</td>
</tr>
<tr>
<td>xreg8</td>
<td>1.14905809</td>
<td>0.306524131</td>
<td>0.17648482</td>
</tr>
<tr>
<td>xreg9</td>
<td>0.48235756</td>
<td>-0.40524908</td>
<td>6.53739782</td>
</tr>
<tr>
<td>xreg10</td>
<td>0.03145281</td>
<td>0.351063312</td>
<td>1.79388326</td>
</tr>
<tr>
<td>xreg11</td>
<td>-0.46395772</td>
<td>-0.457689796</td>
<td>-12.06965578</td>
</tr>
</tbody>
</table>

ARIMAX (p, d, q)

- p = order (# of time lags of the AR part)
- d = differencing (# of past values subtractions)
- q = order of MA part
Exogenous Feature Timeseries Forecasting

CO-ARIMAX models derived on 3 different signal reconstructions based on alternative kine-direction estimates

Exogenous Feature Time-series Analysis

<table>
<thead>
<tr>
<th>Synthesis Approach</th>
<th>Nil-Phase</th>
<th>Correct (True) Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Nonzero (Active) LASSO Coefficients</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
<tr>
<td>LASSO Mean Square Error CV Error of Model Coefficients</td>
<td><img src="image3.png" alt="Graph" /></td>
<td><img src="image4.png" alt="Graph" /></td>
</tr>
<tr>
<td>LASSO Regression Model Coefficients</td>
<td><img src="image5.png" alt="Graph" /></td>
<td><img src="image6.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

Results of regularized linear modeling of CO-concentration using LASSO penalty
Big Data Analytics Study – UKBB

- 9,914 UKBB participants; 7,614 features:
  Features: clinical+phenotypic variables (5K) and derived neuroimaging biomarkers (2.5K)
- Supervised Decision Tree (binary Dx) Classification – Correct Kime-Phase Estimates

### Raw Decision Tree
- Prediction: 0 1
- 0: 362 60
- 1: 79 399
- Accuracy: 0.8456
- 95% CI: (0.82, 0.87)
- No Information Rate: 0.51
- P-Value [Acc > NIR]: <2e-16
- Kappa: 0.6907
- McNemar’s Test P-Value: 0.1268
- Sensitivity: 0.8209
- Specificity: 0.8693
- Detection Rate: 0.6022
- Detection Prevalence: 0.4689
- Balanced Accuracy: 0.8451

### Pruned Decision Tree
- Prediction: 0 1
- 0: 388 127
- 1: 53 332
- Accuracy: 0.8000
- 95% CI: (0.77, 0.83)
- No Information Rate: 0.51
- P-Value [Acc > NIR]: <2e-16
- Kappa: 0.6012
- McNemar’s Test P-Value: 5.295e-08
- Sensitivity: 0.8798
- Specificity: 0.7233
- Detection Prevalence: 0.5722
- Balanced Accuracy: 0.8016


Big DataAnalytics Study – UKBB

- 9,914 UKBB participants (11 epochs of 900 cases); 7,614 clinical measurements, phenotypic features, and derived neuroimaging biomarkers
- Supervised Decision Tree (binary) Classification – Epoch-average Kime-Phases

### Raw Decision Tree
- Prediction: 0 1
- 0: 354 85
- 1: 87 374
- Accuracy: 0.8089
- 95% CI: (0.78, 0.83)
- No Information Rate: 0.51
- P-Value [Acc > NIR]: <2e-16
- Kappa: 0.6176
- McNemar’s Test P-Value: 0.9392
- Sensitivity: 0.8027
- Specificity: 0.8148
- Detection Rate: 0.3933
- Detection Prevalence: 0.4878
- Balanced Accuracy: 0.8088

### Pruned Decision Tree
- Prediction: 0 1
- 0: 190 130
- 1: 251 329
- Accuracy: 0.5767
- 95% CI: (0.54, 0.61)
- No Information Rate: 0.51
- P-Value [Acc > NIR]: 3.501e-05
- Kappa: 0.1484
- McNemar’s Test P-Value: 7.857e-10
- Sensitivity: 0.4308
- Specificity: 0.7168
- Detection Rate: 0.2111
- Detection Prevalence: 0.3556
- Balanced Accuracy: 0.5738
Big Data Analytics Study – UKBB

- 9,914 UKBB participants; 7,614 clinical measurements, phenotypic features, and derived neuroimaging biomarkers
- Supervised Decision Tree (binary) Classification – **Nil-average Kime-Phases**

### Raw Decision Tree

<table>
<thead>
<tr>
<th>Reference</th>
<th>Prediction</th>
<th>0</th>
<th>1</th>
<th>100</th>
<th>373</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>341</td>
<td>86</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>100</td>
<td>373</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Accuracy</strong>:</td>
<td></td>
<td>0.7933</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>95% CI</strong>:</td>
<td></td>
<td>(0.77, 0.82)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No Information Rate</strong>:</td>
<td></td>
<td>0.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>P-Value [Acc &gt; NIR]</strong>:</td>
<td></td>
<td>&lt;2e-16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Kappa</strong>:</td>
<td></td>
<td>0.5862</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mcnemar’s Test P-Value</strong>:</td>
<td></td>
<td>0.3405</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sensitivity</strong>:</td>
<td></td>
<td>0.7732</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Specificity</strong>:</td>
<td></td>
<td>0.8126</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Detection Rate</strong>:</td>
<td></td>
<td>0.4744</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Balanced Accuracy</strong>:</td>
<td></td>
<td>0.7929</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pruned Decision Tree (not shown) was degenerate.

---

Big Data Analytics Study – UKBB

- 9,914 UKBB participants; 7,614 clinical measurements, phenotypic features, and derived neuroimaging biomarkers. Supervised Decision Tree (binary) Classification

Overall feature averages across cases for the 3 complementary kime-reconstruction analytic strategies.
Summary

- Need new methods to tackle substantial Big Biomed/Health Data Challenges
- *Spacekime* representation makes a difference in predictive analytics

- Math models useful for representation & analysis of complex-temporal data
- *Spacekime transform* enables small sample inference

- Optimal *kime-phase aggregators*?
- Spacekime analytics representation has lots of "Open problems" (math, stats, DS)

Interested in Spacekime Analytics?

- Check [www.SpaceKime.org](http://www.SpaceKime.org)
- Contact me
- We have lots of “Open Problems”
Acknowledgments

Funding
NIH: P20 NR015331, U54 EB020406, P50 NS091856, P30 DK089503, UL1TR002240, R01CA233487
NSF: 1916425, 1734853, 1636840, 1416953, 0716055, 1023115

Collaborators
- **SOCR**: Milen Velev, Yonghai Qiu, Zhe Yin, Yufei Yang, Aleksandar Kalinin, Selvam Palaniappan, Syed Husain, Matt Leventhal, Adriana Khan, Rami Eikast, Abhishek Chowdhury, Patrick Tan, Pratyush Pati, Brian Zhang, Juana Sanchez, Dennis Peairs, Kyle Siegrist, Rob Gould, Nicolas Christou, Hanbo Sun, Tui Wang, Yi Wang, Lu Wei, Lu Wang, Simeone Marino
- **LONI**: Arthur Toga, Roger Woods, Jack Van Horn, Zhuowen Tu, Yanggang Shi, David Shattuck, Elizabeth Sowell, Katherine Harr, Anand Joshi, Shartarni Joshi, Paul Thompson, Luminta Vese, Stan Osher, Stefano Soatto, Seok Moon, Junming Li, Young Sung, Cari Kesselman, Fabio Macciardi, Federica Tom

http://SOCR.umich.edu

Slides Online: “SOCR News”