THE UNIVERSITY OF CHICAGO

ESSAYS ON IMPLIED SPHERICAL SPACE FORMS IN STATISTICS AND ECONOMETRICS: FROM APPLICATIONS OF ALGEBRAIC TOPOLOGY TO INTERACTIVE GRAPHICAL TOOLS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE UNIVERSITY OF CHICAGO
BOOTH SCHOOL OF BUSINESS
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
XIAOLONG WU

CHICAGO, ILLINOIS
AUGUST 2018

To CBH

Are the Renaissance painters heroes and gods,

or are they manual workers?

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

The author is in debt to his mentor and the dissertation committee chair, Ruey S. Tsay, for introducing him to the study of time series and providing him with invaluable support throughout the project. He also gratefully acknowledges the guidance from other dissertation committee members, Christian B. Hansen, Robert E. McCulloch, and Drew D. Creal. The author is thankful to the discussions with Hedibert Freitas Lopes, Benson Farb, Amie Wilkinson, Danny Calegari, Matthew Emerton, Richard J. Hill, Jeffrey Harvey, and Carlos E.M. Wagner; the comments from fellow students and seminar participants in the University of Chicago community; the friendship from Victor O. Lima, Francesco Trebbi, Che-Lin Su †, Amanda J. Sharkey, Malaina Brown, and David B. Small; as well as the financial support from the University of Chicago Booth School of Business. He would like to give special thanks to Alexander K. Zentefis, William Fisher, Lucas Hiller, and Joseph R. King, who painstakingly proofread drafts of the manuscript and offered constant encouragement throughout the project.

ABSTRACT

Rao (1945) pioneered the use of differential geometry in statistics by treating the Fisher information as a Riemannian metric. Its theoretical decedents, information geometers, Amari (1968), Costa et al. (2015), and Chen (2016), through explicit calculations, demonstrated the induced statistical manifolds have hyperbolic space forms. This manuscript opens an alternative line of geometric research and calls attention to the ubiquitous phenomena of implied spherical space forms in statistics, which act as benchmark manifolds for positive curvature bounds. By developing novel algebraico-topological techniques suitable to handle these manifolds intrinsically, it introduces significant improvements to two prominent, seemingly unrelated, statistical work: (a) For Andrews and Mikusheva (2016a,b), it provides a radically simplified proof and a succinct new theorem on the distance bound between a random vector and a manifold with a curvature bound; (b) For Huang et al. (1996); Kizhner et al. (2005), it offers rigorous definitions of related concepts and a new fundamental theory that allows smooth transforms between the Hilbert-Huang transform and the Fourier transform of a time series. Through deep engagement with concrete statistical applications, it offers timely remarks about the methodological limitations of information geometry. In lieu of a grand synthesis unattainable at present, as empirical

patches to the methodological wounds inflicted, it offers a suite of new interactive graphical tools based on the foregoing geometric discussions, in particular, the link between the spherical geometry and correlations, for the exploration of dynamic harmonic structures in high-dimensional time series.

KEYWORDS: Fisher information, information geometry, differential geometry, data visualization, financial econometrics, multivariate time series

SUBJECT CATEGORY: Economics (0501), Statistics (0463)

PART I

THEORETICAL LIMITATIONS

«La mathématique est l'art de donner le même nom à des choses différentes.[1]*»*

— *Henri Poincaré, Science et Méthode (1908)*

[1] "Mathematics is the art of giving the same name to different things." G.B. Halsted's translation.

CHAPTER 1

MOTIVATION

The Fisher information was a Riemannian metric before it became the celebrated matrix prescribing the lower bound on the variance of any unbiased estimator: At 25, still a student to Fisher who had already published the *Mathematician Foundations of Theoretical Statistics* 23 years earlier, Rao (1945, 1997 reprint) was immediately canonized; its pioneering use of the differential geometric method, however, is largely ignored. Though differential geometry has inevitably prevailed in other branches of pure mathematics and theoretical physics that need to deal with nonlinear objects, more than seven decades after Rao (1945), it is still not a *lingua franca* in statistics—for good reasons. Information geometers blame it on the inherent difficulty of the mathematics (Amari, 2016), echoing Hilbert's declaration that "physics is becoming too difficult for the physicists."[1] This is rather presumptuous, tinged with condescension: unlike what it has contributed for physics, differential geometry (and by that token modern mathematics in general) hasn't proven its value in statistics. Indeed, the Fisher information, as a Riemannian metric, brings with it the whole Riemannian geometric model which consists of a dazzling array of interactive objects, manifolds, spaces,

[1] The quote itself is in dispute but another documented quote from Hilbert shares the same sentiment: "Every kind of science, if it has only reached a certain degree of maturity, automatically becomes a part of mathematics. (Ewald, 2007)"

bundles, connections, forms, and more (Section 4). Their interactions, though meticulously developed and elegantly stated, are nevertheless not a statistical theory *per se*: "A theory has only the alternative of being right or wrong. A model has a third possibility: it may be right, but irrelevant (Manfred, 1973)."

Any interdisciplinary researcher makes this risky proposal vis-à-vis relevance: he must brazenly burden his colleagues with outlandish constructions, with nothing but an empty promise that all undue discomfort and efforts would pay off; this in turn builds up the expectation that is hard to fulfill because of the very nascent nature of the work. The self-defeating proposition dooms many a researcher in search of a grand unified theory, as evidenced by the deep and ever growing schism between statistics and differential geometry. An outcast of his own volition, he, in spite of the most genuine intentions, inevitably succumbs to the desire for relevance and regresses to a stale state of appeasement: as a rarefied geometer in residence, slowly drifting away to an ever more distant orbit; or worse still, as a mangler who spins mathematical concepts for theoretical expediency. "The doer alone learneth:" it is high time to take out these intricate mathematical artifacts, still mint, from their sophisticated contextual frames, and use these power tools strictly for utilitarian purposes. "*Immer mit den einfachsten Beispielen anfangen*!"[2] Following Hilbert's advice, I will present two modest but cogent examples where differential geometry (and its algebraico-topological extensions) elucidates statistical and econometric practices, as a prolegomenon to a more unified mathematical-statistical theory imagined by Hilbert, information geometers, and mathematical statisticians (Čencov, 1978).

[2] Another Hilbert quote, popularized in German by Artin's *Algebra*, which translates to: "Always start with the simplest examples."

Since differential geometry is the study of manifolds, the simplest examples are space forms, which are manifolds with constant sectional curvatures K (Figure 1.1). Information geometers traditionally focus on the hyperbolic spaces ($\mathbf{H}^n$, $K < 0$). Through explicit calculations, Amari (1968), Costa et al. (2015), and Chen (2016) have exhibited that the manifolds implied by the Fisher information—these are exactly the manifolds Rao imagined, now rebranded as the statistical manifolds—of common distributions (normal, Cauchy, and t-distributions) all have constant negative sectional curvatures. Since analytical statistical practices tacitly take place in the Euclidean spaces ($\mathbf{R}^n$) with vanishing curvatures ($K = 0$), we here provide examples of the remaining third, the spherical space forms ($\mathbf{S}^n$) which have constant positive curvatures ($K > 0$).

This is not to reinvent wheels: indeed, as a most elementary geometric form, spheres are commonplace in statistics, especially in empirical disciplines like geophysics statistics.[3] Our focus is methodological; the thematic focus of spherical spaces, albeit deliberate, is purely instrumental. In analytical statistics, a geometric object, such as a sphere, plays these two, not mutually exclusive, roles—(a) geometric: it can be a real geometric representation of the physical phenomenon (*e.g.*, the Earth in geodetic statistics); or (b) statistical: it can be used to place parametric restrictions on variables to facilitate the estimation process

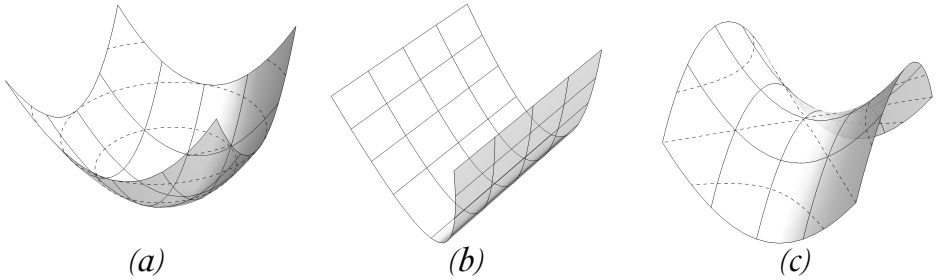
[3] Familiarity doesn't imply understanding. Indeed, the Cartan–Hadamard theorem shows that the universal covering space of a connected complete Riemannian manifold of non-positive sectional curvature is diffeomorphic to $\mathbf{R}^n$. But little can be said about manifolds of non-negative sectional curvature. Even the 3-dimensional case has puzzled mathematicians for nearly a century. Poincaré conjecture, which states "every simply connected, closed 3-manifold is homeomorphic to the 3-sphere," widely known since the start of differential geometry at the beginning of the 20th century, was not proven until Perelman's singular efforts in 2002–2003. He does so by first providing an astonishingly concise proof of a known theorem that reduces the study of a complete manifold of non-negative sectional curvature to that of the normal bundle of a compact manifold, so called the "soul" of the manifold.

(*e.g.*, a vector with a fixed norm in all dimensions). In our examples, the spherical space forms are a bridge object between geometry and statistics: they act as neither the static background spaces on which parameters and observations are grounded (as in the first example about the curvature bound) nor useful functional forms through which extra degrees of freedom can be trimmed (as in the second example about Hilbert-Huang Transform). Instead, they are random geometric objects accompanying each set of observations, encoding and decomposing their nonlinearities: for this reason, we here call these parastatistical geometric objects, or for short, "co-objects" for random variables. This crucial distinction, though seemingly superficial and even cryptic at this point, will be made clear through the concrete examples: it is precisely their hybrid geometric-statistical nature that resists full analytic description and affords deep connections to algebraic topology, whose *raison d'être* after all is to keep track of geometric invariance, independent of coordinate computations. I now refrain from commenting further on the methodological significance of uncovering and analyzing parastatistical spherical spaces, or cospheres, through algebraic topology until explicitly demonstrating its power in these two concrete statistical examples.

This paper is organized as follows. The two examples, the curvature cosphere theorem and the IMF decomposition theorem, will be presented sequentially (Sections 2 and 3) as mostly self-contained case studies, with an eye to showcase the prevalence of parastatistical spherical space forms in statistics, often in disguised forms, and how modern algebraico-topological methods fare comparing to the existing analytical method. Each example starts with (1) a brief introduction to propel the discussion; continues with (2) necessary mathematical groundwork before formally presenting its main findings in the form of new theorems; and ends with (3) a short recapitulation highlighting how these new results contribute to the community research efforts. The HHT discussion, since it's about

an empirical algorithm, includes an additional section offering further comments on the empirical implications of the new findings (Section 3.4). We end the paper with general remarks about the contextual efforts reintroducing geometry to statistics, in particular the paper's relation to the ongoing project of information geometry (Section 4). Section 4.1 provides a brief literature review about landmark papers in the area. Section 4.2 contains definitions of relevant key concepts in differential geometry.

Figure 1.1 Space forms: manifolds with constant sectional curvatures.



(a)	*(b)*	*(c)*
$K = 1$, $\mathbf{S}^n$	$K = 0$, $\mathbf{R}^n$	$K = -1$, $\mathbf{H}^n$
Spherical geometry	Euclidean geometry	Hyperbolic geometry
(This paper)	(Classical statistics)	(Information geometry)

Note. Riemannian manifolds with constant sectional curvature are called space forms. All other connected complete constant curvature manifolds are quotients of those up to isometry. All marked lines are geodesics, *i.e.*, straight lines with respect to their respective geometry. Local charts shown.

CHAPTER 2

CURVATURE COSPHERE THEOREM

In the first example, Andrews and Mikusheva (2016a,b) derived an upper bound on the distance of a random variable to a known manifold. Spherical spaces naturally emerge due to the curvature bound placed on the manifold. The authors take an analytic approach and attempt to fully describe all objects in the coordinates of the ambient space. The laborious process, its manifested rigor notwithstanding, relies on a daunting display of ad-hoc set constructions and optimization problems, in addition to a structural assumption presupposing certain properties of the coordinate system, thus acknowledging the capability limit of their coordinated approach. This is precisely the premise of differential geometry: that a global coordinate system is not always tractable and structures and techniques must be developed to handle charts that only exist locally. This paper takes this modern approach. Instead of pinning down all points in the ambient vector space, geometric objects are constructed hierarchically in a coordinate-free (the so-called intrinsic) manner. In lieu of an uncountable set of auxiliary spherical spaces in the original paper, a single cosphere with desired properties, freed from the ensnarement of the global coordinates, is exposed for each observation. The novel change of perspective, along with two geometric lemmas formally establishing the symmetries of the cospheres and the coordinates enveloping them,

yields an incisive proof without relying on any assumption about coordinates and an elegantly stated new theorem about the distance bound (Theorem 4).

2.1. BACKGROUND

Often credited as the first modern geometer-statistician, Mahalanobis (1936) defines the measure of distance of an observation $x = (x_1, x_2, \ldots, x_N)'$ of a multivariate normal distribution with mean μ and covariance matrix Σ,

$$\rho_M^2 = (x-\mu)'\Sigma^{-1}(x-\mu),$$

now known as the Mahalanobis distance. In econometric, albeit rarely taken in strictly geometric terms, it's often used, to test hypotheses with nuisance parameters for example. Specifically, let $\widehat{\theta}$ be a p-dimensional reduce form estimator with a known (or estimable) covariance matrix Σ and an unknown mean θ_0. We are interested in testing if the model is correctly specified via the restriction function $g(\theta) = 0$ (or equivalently, via the link function $\theta(\beta)$, where β is a p-dimensional unknown parameter). The asymptotic hypothesis test of the restriction is often based on the minimum Mahalanobis distance from the true parameter θ_0:

$$\min_{g(\theta)=0} (\widehat{\theta}-\theta_0)'\Sigma^{-1}(\widehat{\theta}-\theta_0).$$

Since the nonlinear constraint $g(\theta) : \mathbf{R}^k \to \mathbf{R}^{k-p}$ is defined in the full parameter space and has a p-dimensional kernel containing the true value θ_0, Andrews and Mikusheva (2016a,b, hereafter, AM) treats the kernel of the map as a p-dimensional manifold

$$S = \{x : x = \Sigma^{-1/2}(\theta - \theta_0),\ g(\theta) = 0\}$$

and converts the functional restriction to its geometric root, *i.e.*, the minimum distance between a normalized random vector $\xi = \Sigma^{-1/2}(\widehat{\theta} - \theta_0) \sim N(0, I_k)$ and the manifold S:

$$\rho^2(\xi, S) \coloneqq \min_{x \in S} (\xi - x)'(\xi - x).$$

Hypothesizing a curvature bound on the manifold, AM utilizes differential-geometric methods and offers a remarkable theorem relating the geometric property of a manifold to the limiting distribution of an econometric statistic. The theorem reduces the bound on the minimum distance from

$$\rho^2(\xi, S) = \min_{x \in S} (\xi - x)'(\xi - x) \leq (\xi - 0)'(\xi - 0) \sim \chi^2_k \tag{1}$$

prescribed by the predominant "projection method," *e.g.*, Dufour and Jasiak (2001) and Dufour and Taamouti (2005), to a tighter χ^2_p distribution, thereby presenting a new test that is "always more powerful than those based on the projection method."

2.1.1 AM's Theorem

Since various novel constructions are defined in AM's original theorem, for convenience, I am keeping their notations and quoting their theorem in full, with minor adjustments for consistency.

THEOREM 1 (Andrews and Mikusheva). *Let S be a regular p-dimensional manifold in $\mathbf{R}^k$ passing through zero. Assume that the tangent space $T_0(S)$ is spanned by the first p basis vectors. Assume that for some constant $C > 0$, we have $\kappa_q(S) \le \frac{1}{C}$ for all points $q \in S_C$.*[1]

Then:

a. Manifold S_c lies insides the set $\mathscr{M} \cup D_C$, where

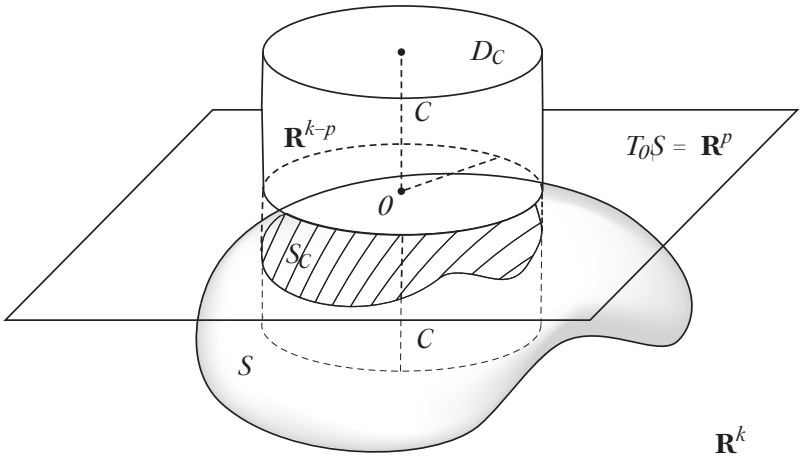
$$\mathscr{M} := \{\|x^{(1)}\|^2 + (C - \|x^{(2)}\|)^2 \ge C^2\}.$$

[1] Before stating the theorem, the authors defined earlier that each vector $x \in \mathbf{R}^k$ has the coordinates $(x^{(1)}, x^{(2)})$ where $x^{(1)} = (x_1, x_2, \ldots, x_p) \in \mathbf{R}^p$ contains the first p coordinates and $x^{(2)} = (x_{p+1}, x_{p+2} \ldots x_k) \in \mathbf{R}^{k-p}$ contains the last $k-p$. They also restrict their attention to "points on the manifold [S] that lie inside a (large) finite cylinder"

$$D_c = \{x = (x^{(1)}, x^{(2)}) : \ \|x^{(1)}\| \le C, \ \|x^{(2)}\| \le C, \ x^{(1)} \in \mathbf{R}^p, \ x^{(2)} \in \mathbf{R}^{k-p}\} \subset \mathbf{R}^k.$$

and define S_C to be "the intersection $S \cap D_C$ if it is connected or the connected parts of $S \cap D_C$ that passes through zero (*i.e.*, the part s of $S \cap D_C$ which can be reached by continuous paths lying in $S \cap D_C$ which pass through zero) otherwise."

Figure 2.1 Basic setup of AM's theorem.



Note. For illustrative purpose, S here is a closed 2-manifold with everywhere positive sectional curvature; T_0S is a tangent space at point "0", a 2-dimensional vector space; D_C is a cylinder of radius C. S_c is the shaded area on the manifold.

b. If [for any $y^{(1)}$ *in* $\mathbf{R}^p$ *with* $\|y^{(1)}\| \le C$*, there exists a point* $x \in S_C$ *such that* $x^{(1)} = y^{(1)}$*], then, for any point* $\xi \in \mathbf{R}^k$*, we have almost surely [for brevity, denoted hereafter by "*a.s.*"].*

$$\rho(\xi, S) \le \max_{u \in \mathbf{R}^{p-k},\ \|u\|=1} \rho(\xi, N_u), \qquad \textit{where}$$

$$N_u := \left\{ \begin{array}{c} x \in \mathbf{R}^k,\ x^{(1)} \in \mathbf{R}^p,\ z \in \mathbf{R}: \\ x = (x^{(1)}, zu),\ \|x^{(1)}\|^2 + \|C - z\|^2 = C^2 \end{array} \right\}.$$

c.

$$\max_{u \in \mathbf{R}^{p-k},\ \|u\|=1} \rho(\xi, N_u) \overset{a.s.}{=\!=} \rho(\xi, N_{\widetilde{u}}), \quad \textit{where } \widetilde{u} := \xi^{(2)}/\|\xi^{(2)}\|.$$

d. If $\xi \sim N(0, I_k)$*, we have for all x, y:*

$$\Pr\{\max_{u \in \mathbf{R}^{p-k},\ \|u\|=1} \rho(\xi, N_u) \le x,\ \|\xi\| \le y\} = \Pr\{\rho_2^2(\eta, N_2^C) \le x,\ \|\eta\| \le y\},$$

where the coordinates of the two-dimensional random vector

$$\eta := (\sqrt{\chi_p^2}, \sqrt{\chi_{k-p}^2}) \in \mathbf{R}^2$$

are independently distributed,

$$N_2^C := \left\{(z_1, z_2) \in \mathbf{R}^2 : z_1^2 + (C + z_2)^2 = C^2\right\}$$

[] and ρ_2 *is [the] Euclidean distance.*

2.1.2 Strategy of AM's Proof

AM's theorem relies on a series of inequalities to refine the existing bound. It accomplishes this by a few ad hoc geometric constructions: D_C is a cylinder with radius C; S_C is its connected intersection with the manifold; $\mathscr{M}$ is the region outside of a p-sphere of radius C, with respect to the ambient space. N_u is a collection of p-spheres, indexed but a unit position vector $u \in \mathbf{R}^{p-k}$ on the p-sphere. $N_{\widetilde{u}}$ is the p-sphere that the random vector ξ passes through. N_2^C is a 1-sphere with radius 1 centered at $(0, -C)$.

2.1 Background

Even though the theorem is geometric, its proof is not strictly so. Proof for (a) relies on its Lemma 1, which "projects" S_C to its tangent space and defines

$$M_v = \{x : \langle x, v\rangle^2 + (C - \|x - \langle x, v\rangle v\|)^2 \geq C^2\}$$

to construct $\mathscr{M}$ as the union of all such sets:

$$\mathscr{M} = \cup_{v \in T_0 S,\ \|v\|=1} M_v.$$

Proof for (b) considers the $k-p$-dimensional linear space $R_r = \{x \in \mathbf{R}^k : x^{(1)} = \tau^{(1)}\}$ for point $\tau \in N_{\widetilde{u}}$; reduces

$$\begin{aligned} R_r \cap \mathscr{M} \cap D_c &= \{x = (\tau^1,\ x^{(2)}) \in D_C :\ \|\tau^{(1)}\|^2 + (C - \|x^{(2)}\|)^2 \geq C^2\} \quad \text{to} \\ &= \{x = (\tau^1,\ x^{(2)}) :\ x^{(2)}\| \leq C - \sqrt{C^2 - \|\tau^{(1)}\|^2}\}; \end{aligned}$$

and then solves the maximization problem "$\rho^2(\xi, x) = \|\xi^{(1)} - \tau^{(1)}\|^2 + \|\xi^{(2)} - x^{(2)}\|^2 \to \max$ [sic.] s.t. $x \in R_r \cap \mathscr{M} \cap D_c$" to obtain $x = \tau$. Statements 1(c) and (d) are straightforward to prove, once (a) and (b) are established. For (c), the proof defines

$$f(u) = \min_{\substack{x^{(1)} \in \mathbf{R}^p,\ z \in \mathbf{R}_+, \\ \|x^{(1)}\|^2 + (c-z)^2 = C^2}} \|\xi^{(1)} - x^{(1)}\|^2 + \|\xi^{(2)} - zu\|^2,$$

and differentiation with respect to u together with the "envelope theorem" yields $\xi^{(2)} = zu$. Statement (d) gives a 2-dimensional reduction of (c) through the definition of η.

We can therefore summarize their strategy to establish the bound on $\rho^2(\xi, S)$ through the following diagram (Figure 2.2). Our job now is to establish a direct link from $\rho^2(\xi, S)$ to $\rho_2^2(\eta, N_c^C)$ (indicated by the wavey arrow).

2.1.3 Clarifications about Curvatures

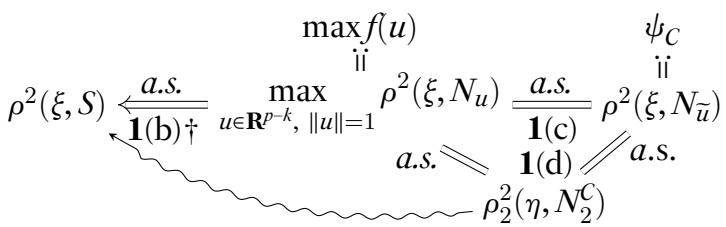
Let's fix the terminology about curvature. There are many different and useful ways to describe the curvature a Riemannian manifold. We use R to denote the CURVATURE TENSOR, which in essence measures the noncommutativity of the covariant derivative,

$$R(u, v)w = \nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{[u,v]} w,$$

where ∇_u is the Levi-Civita connection (covariant differentiation along u) and $[\cdot, \cdot]$ is the Lie bracket of vector fields

$$[X, Y](f) = X(Y(f)) - Y(X(f)) \quad \text{for all } f \in C^\infty(M).$$

Figure 2.2 Diagram of AM's strategy of prove.



Note. The first step (marked by †) further relies on this lemma (their Lemma 1): "Assume $g(0) = 0$. For some $C > 0$, assume that $\frac{\partial}{\partial x'} g(x)$ is full rank for all $x \in S_C$. If the maximal curvature over S_C is not larger than $1/C$, then the projection of S_C on the tangent space $T_0(S_C)$ covers the ball of radius C centered at zero."

2.1 Background

Given two linearly independent tangent vectors at the same point, u and v, the SECTIONAL CURVATURE curvature K is defined to be:

$$K(u,v) = \frac{\langle R(u,v)v, u\rangle}{\langle u,u\rangle\langle v,v\rangle - \langle u,v\rangle^2.}$$

where $\langle \cdot, \cdot \rangle$ is the inner product on the tangent space induced by the metric tensor g introduced earlier. In particular, if u and v are orthonormal vectors, we have the simplification

$$K(u,v) = \langle R(u,v)v, u\rangle.$$

A Riemannian manifold is a space form if its sectional curvature is equal to a constant K. The Riemann tensor of a space form is given by

$$R_{abcd} = K(g_{ac}g_{db} - g_{ad}g_{cb}),$$

The GAUSSIAN CURVATURE, K, is the sectional curvature of a surface, used in the context of the Theorema Egregium which establishes that the measure of the curvature of 2-surface is intrinsic—the Gaussian curvature of a surface does not change if one bends the surface without stretching it. It is the product of PRINCIPAL CURVATURES, often denoted by κ or k, which are curvatures in two orthogonal directions, known as principal directions, given by the second fundamental form.

AM proposes the following definition of the (maximal) curvature at a point q on the manifold:

$$\kappa_q(S) = \sup_{\substack{X \in T_q(S), \\ \dot{\gamma}(0)=X}} \kappa_q(\gamma, S) = \sup_{\substack{X \in T_q(S), \\ \dot{\gamma}(0)=X}} \frac{\|(\ddot{\gamma}(0))^{\perp}\|}{\|\dot{\gamma}(0)\|}, \tag{2}$$

where $\gamma : (\epsilon, \epsilon) \to S$ is a curve on S passing through q at $\gamma(0)$ and $(W)^{\perp}$ is the projection of W onto the space orthogonal to $T_q(S)$ and in adddition suggesting the following scheme to calculate curvature in practice:

$$\kappa_q(S) = \sup_{\substack{u=(u_1,\ldots,u_p)\in\mathbf{R}^p, \\ \|\sum_{i=1}^p u_i u_j\|=1}} \left\| \sum_{i,j=1}^{p} u_i u_j V_{ij}^{\perp} \right\| = \sup_{w=(w_1,\ldots,w_p)\in\mathbf{R}^p} \frac{\left\|\sum_{i=1,j}^{p} w_i w_j V_{ij}^{\perp}\right\|}{\left\|\sum_{i=1}^{p} w_i v_i\right\|}. \tag{3}$$

where V_{ij} is the second derivative $\partial^2_{y_i y_j}\mathbf{x}(y^*)$—this is the supremum of sectional curvatures in all directions u_i, u_j, near a point u. Since the definition of curvature is unsigned, we can assume the sectional curvature is bounded from both sides: $0 < |K| < \frac{1}{C^2}$.

2.2. MAIN RESULTS

As demonstrated in the previous section, AM's four-pronged theorem relies on many ad-hoc constructions that do not always have easy geometric interpretations. Their original proof, though rigorous, is rather cumbersome: it relies on even more intermediary constructions of sets and side optimization problems. I now present a drastically simplified proof and a succinctly stated theorem introducing the novel idea of cospheres, with the helping hand from algebraic topology. Section 2.2.1 develops two theoretical results to deal with high dimensional objects, thereby warranting the seemingly audacious step in the main proof that reduces all irrelevant dimensions. Section 2.2.2 presents the new proof and an im-

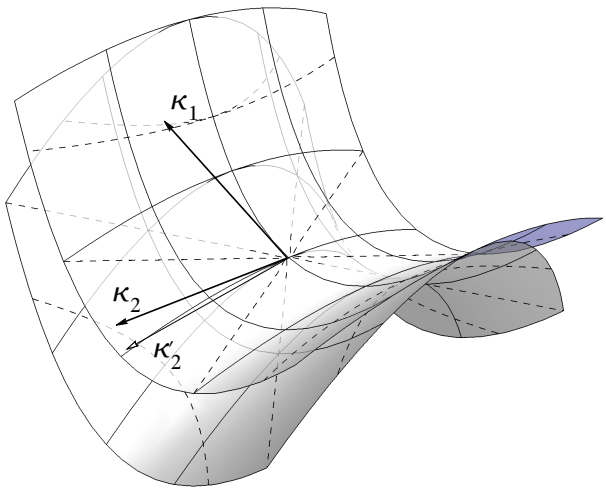
proved theorem. Readers can start directly with the proof and refer to Section 2.2.1 for further explanations on the geometry of hypersphere and factor bundles.

2.2.1 Two Geometric Lemmas

Lemma 2 (Spherical shell). *Let S be a regular p-manifold ($p \geq 2$) embedded in the ambient vector space $\mathbf{R}^k$, where $k \geq p+1$. Assume S has the sectional curvature $|K| \leq \frac{1}{C^2}$, $C > 0$ in each section at each point $s \in S$. Then there exists a p-sphere, such that:*

a. $\mathbf{S}^p$ is centered at $c \in \mathbf{R}^k$, with a radius $r = C$;

Figure 2.3 Principal curvature comparison.



Note. Gauss curvature K of a surface at a point is the product of the principal curvatures, κ_1 and κ_2. It is an intrinsic measure of curvature, not depending on the geodesics chosen. We can make curvature comparisons of different surfaces at a point by aligning them together so that one principal curvature is identified $\kappa_1 = \kappa_1'$.

b. $\mathbf{S}^p$ *intersects with* S *at the point* $0 \in S^p$*; and*

c. For any point on $s \in S$*, we have* $|s - c| \geq r$*.*

Proof. Let's first show a n-sphere $\mathbf{S}^n$ has constant sectional curvature $K = \frac{1}{r^2}$ at each point: one can deduce this from the typological fact that $\mathbf{S}^n = \times^n \mathbf{S}$. Consider $\mathbf{S}^n$ in a $\mathbf{R}^{n+1}$ Euclidean space, so that the spherical center coincides with 0. For each point $p \in \mathbf{S}^n$, represented by the position vector $v(x_1, \ldots, x_{n+1})$ normal to the tangent space, we can form an orthonormal basis $\{e_i\}_{i=1}^j$ at the tangent space $T_p\mathbf{S}^n$ such that $e_i = \lambda_i d_{e_i} v$, where λ_i is a scaling factor. From the definition of sectional curvature earlier or using the more specialized tangential curvature equation, see *e.g.*, Petersen (2006), we then have $K_{\mathbf{R}^{n+1}}(e_i, e_j) = K_{\mathbf{S}^n}(e_i, e_j) - (\lambda_i \lambda_j)^{-1}$, with respect to their respective metrics $g_{\mathbf{R}^{n+1}}$ and $g_{\mathbf{S}^n}$. Using spherical coordinates on the n-sphere, we have

$$g_{\mathbf{S}^n} = dr^2 + r^2 \sum_{i=1}^{n} d\phi_i^2 \quad \text{and} \quad g_{\mathbf{R}^{n+1}} = dr^2 + g_{\mathbf{S}^n}.$$

This therefore shows $\lambda_i = \lambda_j = r$ and $K_{\mathbf{S}^n}(e_i, e_j) = \frac{1}{r^2}$, since the Euclidean space is flat and has 0 sectional curvature.

We now use the method to prove the main claim of the lemma. For each point $q \in S$ and any section given by the orthonormal vectors in its tangent space, $u, v \in T_q S$, there exists a circle $\mathbf{S}^1$ in the same plane $\mathbf{E}$ as u, intersecting S at point q. Under the curvature hypothesis of the lemma, $K(u, v) = \langle R(u, v)v, u \rangle \leq \frac{1}{C^2}$. Elementary geometry on the plane $\mathbf{E}$ shows: for any $\mathbf{S}^1$ with a radius $r = C$, we have for each point s on the intersection of S with $\mathbf{E}$ $s \in S|_{\mathbf{E}}$, we have $|s - c| \geq r = C$. The center c of $\mathbf{S}^1$ is said to be on the curvature center side of S, if $K \geq 0$.

Since it's true for each section and since the manifold is finite-dimensional, induction shows one can at least find an open neighborhood of a point q on the S, such that the p-sphere touches the manifold at q with its center on the curvature center side of S, i.e,

$$\text{for all } s \in U \subset S,\ |s - c| \geq r,$$

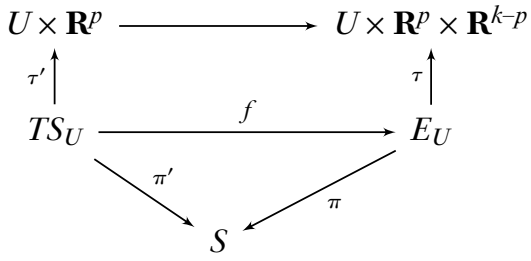
without loss of generality, call this point 0. Suppose the opposite: there exists a point s' on the manifold S such that $|s' - c| < r$. Let γ be a geodesic connecting 0 to s'. Since the manifold is regular, γ is compact (both in S and the ambient vector field), there is a finite set of open covers for γ: each with the desired property. In particular, we have such an open cover for s, U_s. Using the manifold hypothesis, there are local charts (ψ, V_0) and (ϕ, V_s) at points 0 and s respective. Therefore, for all points $s \in V_s \cap U_s$, we have $|s - c| \geq r$. But $s' \in V_s \cap U_s$: a contradiction.

Remark. This is a purely geometric fact. Since the curvature of an n-sphere is reciprocal to its radius squared, the lemma formally establishes the intuition that a mildly deformed spherical object can still contain a smaller sphere with the same dimension, see Figure **??**. One can readily generalize the result, but let's keep our focus on the econometric problem at hand.

> LEMMA 3 (Split of vectors bundles). *Let S be a p-manifold embedded in a $k \geq p+1$-dimensional Euclidean space. The fiber-wise quotient vector spaces $\mathbf{R}^k / T_x S$ form a factor bundle over S.*

Proof. The ambient vector field gives a trivial vector bundle over S, E, that is fiber-wise $\mathbf{R}^k$. Let $\pi' : TS \to S$ be the tangent bundle of S and Let $\pi : E \to S$ be the vector bundle

over S. Let $f : TS \to E$ be the fiber morphism. Locally we have the trivialization for the tangent bundle $\tau' : TS_U \to U \times \mathbf{R}^p$ and the corresponding trivialization for the vector bundle $\tau : E_U \to U \times \mathbf{R}^k$. Since we have the split, viewed as vector field decomposition, $\mathbf{R}^k = \mathbf{R}^p \times \mathbf{R}^{k-p}$, along with the following commutative diagram:



We have the local exact sequences:

$$0 \to \pi' \xrightarrow{f} \pi \quad \text{and} \quad 0 \to TB_U \xrightarrow{f} E_U$$

We can take the disjoint union of all factor fibers E_x/T_xS to form the factor bundle.

Remark. The proof is constructive and technical but the idea is elementary. Each point on a manifold comes with a chart $\psi : U_i \to R^n$ that reduces it to an Euclidean space. Higher constructions, like bundles, inherit the chart, through similar trivializations. The Lemma does not depend on the dimensions and one can replace the tangent bundle of S with any other vector bundle. Lang (2001) provides more details about vector bundles and metric bundles. Exact sequence method is ubiquitous in modern mathematics. Here are two illustrative examples to apply the method to statistics.

EXAMPLE 1. Let $\theta \in \Theta$ be a k-dimensional variable in the full parameter space Θ; $\beta \in \mathrm{B}$ be a p-dimensional explanatory variables; $\theta : U \subset \mathrm{B} \to \Theta$ be a link

function; and $g : \theta \to \mathrm{A}$ be a restriction function. A parameter restriction is valid (in the sense that model is corrected specified and could be identified) if and only if the short sequence

$$0 \longrightarrow \mathrm{B} \xrightarrow{\theta} \boldsymbol{\theta} \xrightarrow{g} \mathrm{A} \longrightarrow 0 \qquad (\mathbf{4})$$

is exact. The following statements are equivalent: θ/B is irrelevant; the structural form $g(\theta) = 0$ and the reduce form $\theta = \theta(\beta)$ are just identified; $\operatorname{Im}\theta := \{\theta : \theta = \theta(\beta)\} == \{\theta : g(\theta) = 0\} =: \operatorname{Ker} g$.

2.2.2 Main Proof and Cosphere Theorem

Proof (Proof to Theorem 1). (a) $S_C \subset M$ by Lemma 2. $S_C \subset D_C$ by construction.

(b) and (c) If $\xi \in S$, (b) and (c) already true by Lemma 2 and if $\xi \in T_0S$, (c) in not defined since $\|\xi^{(2)}\| = 0$. Let's now assume $\xi \notin S$ and $\xi \notin T_0S$ and let p be the projection of ξ on T_0S. Let c be the center of the p-sphere described in Lemma 2 touching S at a fixed point, "0". The three distinct points ξ, p, and 0 determines a unique plane $\mathbf{E} \simeq \mathbf{R}^2$. Since $\mathbf{E}$ is by definition perpendicular to T_0S, it contains a curve γ from S connecting 0 and q,[2] the projection of ξ on γ, a circle centered at c on the curvature center side of γ touching it at 0, and another circle centered at c' also with radius r and touching γ at 0, such that c,

[2] This is true because $\mathbf{E} = (e, e^{\perp})$, by virtue of being perpendicular to T_0S, can be made to have its first coordinate coinciding with the first coordinate of $g(x)$, after a change of coordinates, if necessary. Specifically, let $\phi : \mathbf{R}^k \to \mathbf{R}^k$ be such a change of coordinates, which makes the desired change $(x_1, \ldots, x_k) \mapsto (e, x_2', \ldots, x_k')$, seen as a $SO(p)$ rotation of the tangent space around 0. The curve $\gamma \in \mathbf{E} \cap S$ connecting 0 to q, not necessarily a geodesic, is the composite map $g\phi^{-1}i_1$, where $i_1 : [0, a] \to \mathbf{R}^k$ is the injection into the first coordinate, while holding all other coordinates at 0.

0, and c' are on the same line: this is precisely the basic case for $k = 2$ and $p = 1$, with the exception that the projection of ξ on S, where the minimal distance between ξ and the manifold is obtained, might be somewhere else not on $\mathbf{E}$. But this is not a concern. Possibly after switching the names of c and c', we can assume, without loss of generality, $|\xi - c| \geq |\xi - c'|$ and $|\xi - c| - r \geq |\xi - q'|$, where q' is the intersection of $\xi - c$ on γ.[3] The direction of $\widetilde{u}$ is uniquely given by the direction of $-(\xi - c)$ and we have

[3] Though not strictly necessary, Lemma 3 allows us to further explicitly factor out all irrelevant $(k-2)$-dimensions on each fiber, leaving the plane containing the geometric objects of interest (Figure 2.5). I removed more explicit algebraic constructions here to avoid causing any possible confusion, as the ambient space $\mathbf{R}^k$ needs to be taken as the manifold (instead of the manifold S) for the lemma and each fiber is identified by ξ.

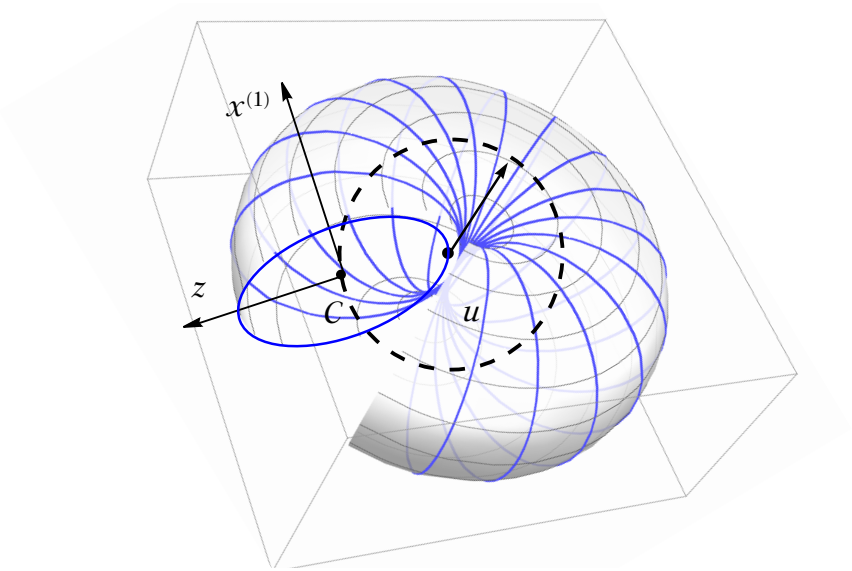


Figure 2.4 N_u as a toru
$\mathbf{S}^p \times \mathbf{S}^{k-p}$.

Note. The collection of p-sphere defined in AM's theorem, N_u, can be seen as a torus $\mathbf{S}^p \times \mathbf{S}^{k-p}$. Each vector u on the $(k-p)$-sphere, $u \in \mathbf{S}^{k-p}$, pins down a section containing a $\mathbf{S}^p$ sphere. The new proof presented exploits the fact it's a product space with the factor $\mathbf{S}^p$ being the relevant one for the distance comparison.

$$\rho(\xi, S) \leq |\xi - q'| \leq |\xi - c| - r = \rho(\xi, N_{\tilde{u}}).$$

The "almost surely" part is redundant.

(d) This is only dimension counting with the degree of freedom of the χ^2 distribution and geometry of η is true by construction. Decompose $\mathbf{R}^k = \mathbf{R}^p \times \mathbf{R}^{k-p}$. Since $\xi \in N(0, I_k)$, each component is therefore an independent variable from standard normal. It's clear its norms in each component vector space, $|\eta_1|^2 = \sum_{i=1}^{p} \xi_i^2$ and $|\eta_2|^2 = \sum_{i=p+1}^{k} \xi_i^2$, follow χ_p^2 and χ_{k-p}^2 distributions respectively. By (c), the fixed center of N_2^C gives the desired geometry. In fact, η can be explicitly written as the $\eta = (\xi - c)$.

Stripping away all auxiliary constructions in the original theorem yields the following statement:

THEOREM 4 (Cosphere). *Let M be a p-manifold embedded ($p \geq 2$) in the ambient vector space $\mathbf{R}^k$, where $k \geq p+1$. Assume S has the sectional curvature $|K| \leq \frac{1}{C^2}, C \geq 0$ in each section, at each point $s \in S$.*

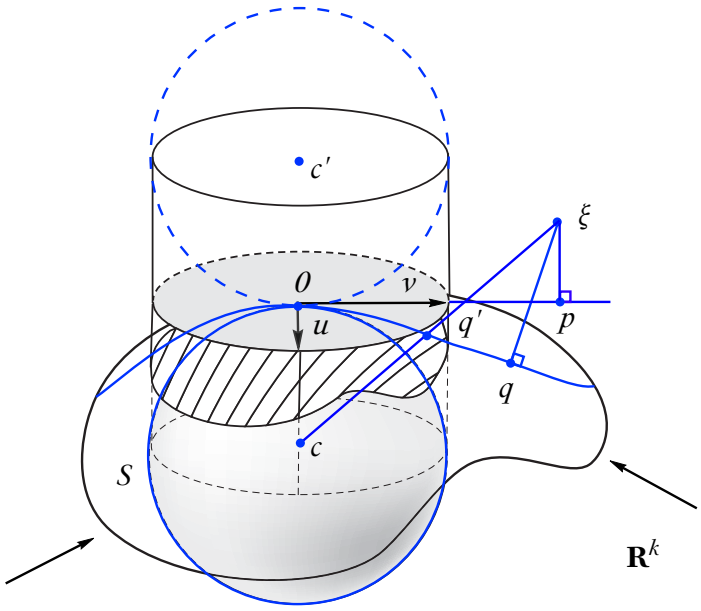
Fix any point m on the manifold as the point of econometric interest. Let ξ be a standard normal random vector: $\xi \sim N(m, I_k)$. There exists a p-sphere, $\mathbf{S}_{\xi}^{p}$, with radius $r = C$ touching M at q, centered at c, such that

$$\rho^2(\xi, M) \leq \rho^2(\xi, \mathbf{S}_{\xi}^{p}). \tag{5}$$

Any such p-sphere $\mathbf{S}_{\xi}^{p}$ defined above is called a COSPHERE of M for ξ.

Proof. Almost identical to our new proof for Theorem 1(b) and (c). Suffice to show such cospheres exist for any ξ, not necessarily uniquely. If $\xi \in M$ or $\xi \in T_mM$, the statement is trivial, any sphere in Lemma 2 is a cosphere. Let's now assume $\xi \notin M$ and $\xi \notin T_mM$. Let p be its projection on T_mM. Let c be the center of the p-sphere described in Lemma 2 touching M at a fixed point, m. The three distinct points ξ, p, and 0 determines a unique plane (Figure 2.5). The plane contains a curve γ from S connecting m and q, the projection of ξ on γ, a circle centered at c on the curvature center side of γ touching it at m, and another circle centered at c' also with radius r and touching γ at m, such that c, m, and c' are on the same line: this is precisely the basic case for $k = 2$ and $p = 1$. Possibly after

Figure 2.5 Sectional reduction of the curvature bound theorem.



Note. Since the curvature bound $K(u,v) = \langle R(u,v)v, u\rangle \leq \frac{1}{C^2}$ holds for any orthonormal section and by the spherical symmetry of $\mathbf{S}^p$, we can reduce the higher-dimensional problem to the most basic 1-dimensional scenario along v (relevant curves shown in blue). Lemma 2 fails near the bottom of the manifold because of areas with large curvatures (indicated by arrows).

switching the names of c and c', we can assume, without loss of generality, $|\xi-c| \geq |\xi-c'|$ and $|\xi-c|-r \geq |\xi-q'|$, where q' is the intersection of $\xi-c$ on γ. By construction, $\rho(\xi, \mathbf{S}^p_\xi) = |\xi-c|-r$. This therefore proves the claim, since $\rho(\xi, M)$, possibly obtained somewhere else on M, is assumed to be smaller than $|\xi - q'|$ and

$$\rho(\xi, M) \leq |\xi - q'| \leq |\xi - c| - r = \rho(\xi, \mathbf{S}^p_\xi).$$

The geometry of η follows by constructing the coordinate system dynamically for each ξ. One can always choose orthonormal basis of $\mathbf{R}^k$ for each ξ, $\sum_{i=1}^k e^i_\xi$, such that the cospheres are always centered at

$$c = -r\sum_{i=1}^{p} e^i_\xi + 0 \sum_{i=p+1}^{k} e^i_\xi .$$

This gives the center of the cosphere coordinates $c = -r\mathbf{1}_p \oplus \mathbf{0}_{k-p} = (-r, \ldots, -r, 0, \ldots, 0)$, where we use the short hand $\mathbf{1}_p$ and $\mathbf{0}_{k-p}$ to keep track of the dimensions where the center of cosphere is shifted by $-r$ and unchanged with respect to m. Note this coordinate system centers the distribution of ξ at 0_k, $m = \mathbf{0}_k$. Since the tangent plane separates c and ξ, ξ always has positive coordinates dynamically in the $-r\mathbf{1}_p$ dimensions. $\rho(\xi, \mathbf{S}^p_\xi)|_{\mathbf{1}_p} = |\xi-c|-r = \xi + r - r = |\xi|$ and $\rho(\xi, \mathbf{S}^p_\xi)|_{\mathbf{0}_{k-p}} = |\xi - c| - r = |\xi| - r$. Therefore $\eta_1 = \rho^2(\xi, \mathbf{S}^p_\xi)|_{\mathbf{1}_p} \sim \chi^2_p$ and $\eta_2 = \rho^2(\xi, \mathbf{S}^p_\xi)|_{\mathbf{0}_{k-p}} \sim \chi^2_{k-p}$ if the $\mathbf{0}_{k-p}$ dimensions are shifted by $-r$: this precisely places η on a circle with the center at $(0, -r)$ and η^2 has the χ^2_k distribution, predicted by the "projection method."

2.3. CONTRIBUTIONS

This paper contributes to the ongoing efforts reintroducing geometry to statistics by bringing significant improvements to an established result. Building on the remarkable work of Andrews and Mikusheva (2016a,b) which relate a purely geometric concept, sectional curvature, to the limiting distribution of an estimator, this paper drastically simplifies the process by freeing the geometric objects from the background coordinate system and encoding the curvature information on the counterfactual object, cospheres. The idea of a co-object is worth noting: it is a geometric object accompanying each random observation and acts as the intermediary between statistics and geometry. The new proof along with the accompanying cosphere theorem contributes to the existing literature in several significant ways.

First, it yields a more direct proof. Our approach, focusing on the geometric intrinsics, provides a new and drastically simplified proof. Since the manifold in question is the kernel of some restriction map on the parameter space $g(\theta)$ and $\theta = (\theta_1, \theta_2, \ldots, \theta_k)$ has the natural coordinates and the curvature is obtained in these coordinates via Equations **2** and **3**, AM describes all geometric constructions with this global coordinate system. This classical analytic approach integrates well with set theory (by describing set intersections with the coordinates) and optimization theory (through differentiation with respect to these coordinates). However, since geometric objects are obtained through coordinate-based calculations, they tend to be stripped of geometric intuitions and as a result become rather cumbersome to manipulate.

Second, the resulting new theorem is more elegant and does not rely on additional assumptions about the coordinate system: one cosphere represents all the curvature in-

formation needed for the distance bound. In AM's original theorem, the validity of the theorem rests on the extra assumption about the power of the background coordinate system: "for any $y^{(1)}$ in $\mathbf{R}^p$ with $\|y^{(1)}\| \leq C$, there exists a point $x \in S_C$ such that $x^{(1)} = y^{(1)}$," their Assumption 1. It further comments: "Lemma 1 shows that Assumption 1 holds quite generally for implicitly defined manifolds." However, this goes against the fundamental assumption of a manifold: a chart is only supposed to be given locally. There is a well-known exponentiation map descending from the tangent space from the manifold $\exp : T_pM \to M$ that yields local isomorphism between vectors in the tangent space and points on the manifold, a map they are in substance utilizing in proving their Lemma 1 (and we will use in the next section), but it is only given locally (unless we are assuming in additional the manifold is a Lie group) from $U \times \mathbf{R}^p \to U$, where U is an open neighborhood on the manifold containing p. The extra metric bound on the open neighborhood is not warranted.

Third, it fixes minor technical errors in AM's proof. For example, in AM's Theorem 1(c) the vector identifying the desired maximizing sphere from their collection of p-dimensional spheres N_u, $\widetilde{u}$, is defined to be $\widetilde{u} = -\frac{1}{\|\xi^{(2)}\|}\xi^{(2)}$. This is invalid if ξ lies in $\mathbf{R}^p$ subspace, *i.e.*, ξ falls in the tangent space T_0S with $\xi^{(2)} = \mathbf{0}_{k-p}$.

Most importantly, I hope to call attention to the existence of implied spherical space forms in statistics and demonstrate the superiority of the differential geometric methods in manipulating these objects. I will make explicit these larger points about dealing with nonlinearity from the intrinsic geometric perspective and the deep connection between statistics and Riemannian metrics in the concluding remarks, after the presentation of the next more powerful example. There, algebraic topology lends support to an empirically effective, yet theoretical groundless statistical practice but at the same time prescribes a stern limit of what it could achieve.

CHAPTER 3

HILBERT-HUANG TRANSFORM

In the second example, our goal is to lend theory to a theory-less practice. Statisticians at the National Aeronautics and Space Administration (NASA), who work with predominantly geophysical data, through practice found a data transform algorithm (officially termed the Hilbert-Huang Transform by NASA, or HHT, since Hilbert Transform is often used in the subsequent analyses) decomposing a time series into several simpler summable parts (Huang et al., 1996; Kizhner et al., 2005). Each part, called an intrinsic mode function (IMF), has more regular geometric shapes and is claimed to be easier for scientific interpretations. The algorithm consists of several intuitive geometric steps but these mathematically nonstandard steps lack formal definitions and resist clear-cut characterizations. This paper answers the call for theoretical clarity initiated by its proponents more than two decades ago. It establishes a direct link between the algorithm to the Fourier transform (Theorem 7), by demonstrating how each IMF can be smoothly transformed into a Fourier basis (Lemma 6). As the proof demonstrates, the Fourier transform can be seen as a canonical curvature decomposition, from an arbitrary square-integrable function into a series of constant functions on the circle, each identified by its wrapping number. Just like the first example, accompanying each time series, there is an implied spherical space

(1-dimensional), a cocircle, where an observation on the time series is compared to a point on the cocircle indexed by its multiplicity, or the wrapping number of the path between the point to a base point. Since NASA researchers in practice work with time series of superimposed cyclic functions, the algorithm therefore provides a heuristic way to identify relevant wrapping numbers.

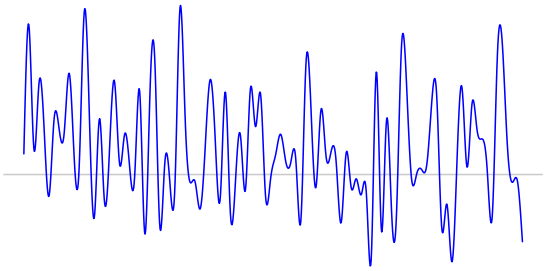
3.1. BACKGROUND

Dissatisfied with the Fourier Transform and the computational companion, the Fast Fourier Transform (FFT), due to their strong assumptions about the source data,[1] researchers at NASA Goddard Space Flight Center (GSFC), developed and commercialized a novel data transform algorithm termed officially, the Hilbert-Huang Transform (HHT). All HHT's empirical successes only make it more urgent to investigate its mathematical limits. Huang is forthcoming about the method's shortcoming: by his own admission, we have no theory to support the adaptive data analysis methodology (Huang and Pan, 2006). Since the Hilbert spectral analysis step of the HHT is well understood mathematically, we focus our attention on the sifting process, which is an empirical procedure to extract temporal features, represented from the IMFs, from the graph of a time series. This is not a straightforward task: since the HHT makes no refutable claims, there is no claim to prove nor test to run. It's tempting to dismiss the HHT as statistical tea-leaf reading but it doesn't offer help to the task at hand—to elucidate and guide the common practice from a theoretical point of view.

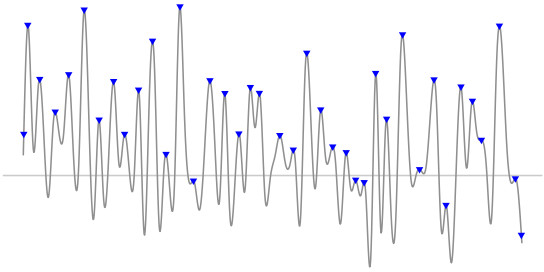
[1] "[S]uch as linearity, of being stationary, and of satisfying the Dirichlet conditions (Kizhner et al., 2005)."

I present a novel application of an algebraico-topological idea, homotopy, to the setting and an elegant theory to reduce the HHT to the Fourier Transform by homotopically transforming each IMF to a basis in the Hilbert space. This not only lends a solid theo-

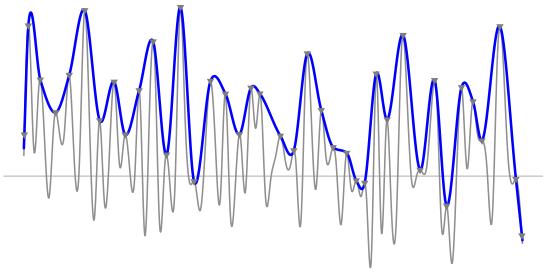
Figure 3.1 An example of the sifting process.



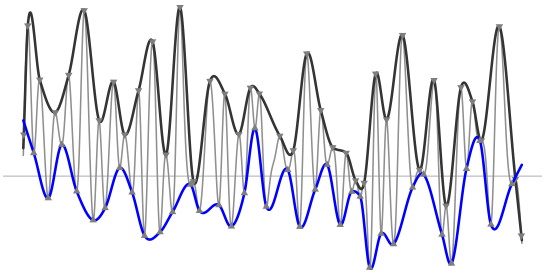
Step 1. Start from an IMF candidate function h.



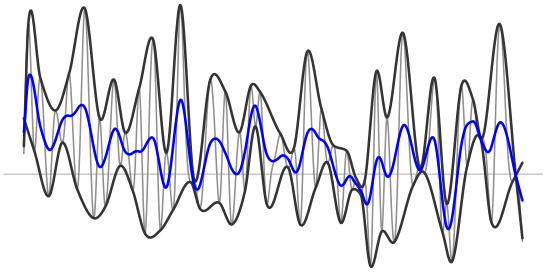
Step 2. Identify the set of local maxima $\{h\}_+$. Indicated by triangles.



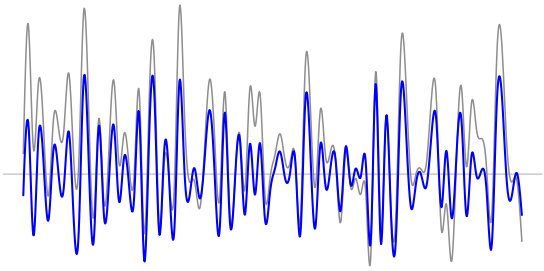
Step 3. Generate a cubic spline $\mathscr{S}\{h\}_+$ that sequentially connects the points of $\{h\}_+$.



Step 4. Repeat Steps 2 and 3 for local minima $\{h\}_-$.



Step 5. Obtain the mean curve by simple averaging: $m(t) = \frac{1}{2}(\mathscr{S}\{h\}_+ + \mathscr{S}\{h\}_-)$. The candidate function h is declared an IMF if $m(t)$ is zero.



Step 6. Obtain an update for the IMF candidate function h' by subtracting the mean: $h'(t) = h(t) - m(t)$. Go to Step 1 or stop if some chosen stoppage criterion is met.

retic support to the serendipitous procedure but also provides practical guide as to where the process might be most useful. It might not seem immediately obvious how the HHT is at all related to the foregoing econometric discussions of the curvature bound. This is precisely the point of the paper—that the phenomenon of curvature bounds, in spherical space forms, arises naturally in many problems, often in disguised forms, due to the fundamental nature of the geometric restrictions placed on the parameter space, intentionally or inadvertently. Same geometric constructions will reappear in this section and I will make the larger point in the concluding remarks.

The HHT is the two-step process introduced by Huang et al. (1996) and subsequently popularized by the author and his collaboration for applying the Hilbert transform on the intrinsic mode functions (IMF), which are obtained recursively through the sifting process (also known as the empirical mode decomposition method) from a time series of observations, usually 1-dimensional, up to some stoppage convention. The HHT method trumps traditional Fourier and wavelet transforms in analyzing spatial-frequency data of mostly nonlinear and non-stationary[2] data where the Fourier transform, notwithstanding the full backing of mathematical rigor, does not yield satisfactory empirical results, according to its proponents. Since its introduction, the HHT has been applied to time series in diverse disciplines from empirical signal-processing fields like imaging processing (Hariharan et al., 2006) and speech recognition (Huang and Pan, 2006) to more theoretically driven fields like financial time-series modeling (Huang et al., 2003; Li and Huang, 2014). Various modifications and extensions, building on the sifting process, have been proposed: the HHT

[2] Stationarity, the common terminology used in the HHT literature, means the amplitudes of the empirical modes implied change over time in this context, in contrast to the basis of the Fourier transform.

method has continuously generated active research interests (Chen and Feng, 2003; Parey and Pachori, 2012; Guang et al., 2014).

The sifting procedure is intuitive and easy to illustrate. Figure 3.3 outlines the algorithm and Figure 3.1 illustrates the process. Readers can refer to Huang et al. (1996) for more detailed description of each step.

Figure 3.2 Comparison between Fourier, wavelet, and HHT analyses.

	Fourier	Wavelet	HHT
Basis	fixed, infinite		adaptive, finite
Theoretical base	mathematical		empirical
Linearity	yes		no
Stationarity	yes	no	
Presentation	frequency	time-frequency	

Note. Adopted from comparison table from Huang and Wu (2008).

Figure 3.3 The sifting algorithm.

$$\boxed{h_0 \equiv x} \rightarrow \boxed{h_1} \rightarrow \boxed{h_2} \rightarrow \ldots \rightarrow \boxed{h_{k-1}} \rightarrow \boxed{h_k \equiv c}$$

a. Start from an IMF candidate function h, *e.g.*, a time series.
b. Identify the set of local maxima $\{h\}_+$.
c. Generate a cubic spline $\mathscr{S}\{h\}_+(t)$ that sequentially connects the points of $\{h\}_+$.
d. Repeat (2) and (3) for local minima $\{h\}_-$.
e. Obtain the mean curve by simple averaging: $m(t) = \frac{1}{2}(\mathscr{S}\{h\}_+(t) + \mathscr{S}\{h\}_-(t))$.
f. Declare the candidate function $h(t)$ an IMF c, if $m(t)$ is zero.
g. Obtain an update for the IMF candidate function h' by subtracting the mean: $h'(t) = h(t) - m(t)$.

3.2. MAIN RESULTS

I now present homotopic theory to smoothly retract the HHT to the Fourier transform, which is the only available convergence theory on the functional space. Section 3.2.1 reframes the setting in mathematically and derives a simple identification criterion for the IMF, crucial for subsequent proofs. Section 3.2.2 introduces the idea of homotopy, commonly known in the algebraic topology community and uses it proves a technical lemma that shows each IMF can be smooth transformed to a unique Fourier basis, identified by an integer. I do so geometrically and in a self-contained exposition. Section 3.2.3 at last puts the elements in a deceptively simple theorem that shows the deep connection between the HHT and the Fourier Transform.

One regularization assumption. For convenience, we can assume $f(0) = f(T) = 0$ and $\#\{h\}_+ = \#\{h\}_-$. This is not a strong assumption for the time series under investigation. Since we are interested in obtaining the *intrinsic* empirical modes, whose amplitudes and frequencies are driven by the physics of the underlying process, we may trim the end points (by discarding a few observations) without affecting the sifting process. The empirical example in Section 3.4 further justifies the assumption: in practice, the time series interested is typically long and the low-frequency (comparing to the observations) IMFs generally do not have any scientific significance.

3.2.1 IMF Identification Criterion

Huang et al. (1996) defines a smooth function to be an intrinsic mode function (IMF) if it satisfies the following two conditions:

a. In the whole data set, the number of extrema and the number of zero-crossings must either be equal or differ at most by one; and

b. At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

This is the standard definition used by the HHT literature but not rigorous enough for further discussions. Given any smooth function $h(t)$, denote the set of its local maxima by $\{h\}_+$; local minima by $\{h\}_-$, and zero *crossings* by $\{h\}_0$. Let $\{h\}_\pm$, the set of local extrema, be the union of $\{h\}_+$ and $\{h\}_-$: $\{h\}_\pm = \{h\}_+ \cup \{h\}_-$. Given an ordered set M, a spline $\mathscr{S}M$ of order-n is a smooth function, defined piecewise as polynomials of order n. Given any smooth function $h(t)$, we may form splines $\mathscr{S}\{h\}_+$ and $\mathscr{S}\{h\}_-$ from the sets of function $h(t)$'s local maxima $\{h\}_+$ and its local minima, respectively. Call these splines, the upper envelope spline and the lower envelope spline, respectively. There is no canonical way to form splines of a given degree of polynomials. A cubic (*i.e.*, a order-3) spline is commonly used via the cubic B-spline and the cubic Bézier spline method. Indeed, it remains an open question of the HHT research to determine the best among these spline methods (Huang and Shen, 2005). For demonstrating purpose, we use the monotone cubic Hermite spline according to the method of Fritsch and Carlson, implemented as `splinefun` in R (Fritsch and Carlson, 1980).

We can simplify the two conditions of the IMF with the notations presented, which shows the first condition, though useful as an easier necessary condition to check, is redundant in the definition.

3.2 Main Results

THEOREM 5 (IMF Identification Criterion). *A function $h(t)$is an IMF if and only if it has symmetric envelope splines:*

$$|\mathscr{S}\{h\}_+| = |\mathscr{S}\{h\}_-|.$$

Proof. This is the second and the only other condition of the definition of an IMF. Given a smooth function h, the first condition of an IMF can be written as:

$$|\#\{h\}_+ + \#\{h\}_- - \#\{h\}_0| \leq 1.$$

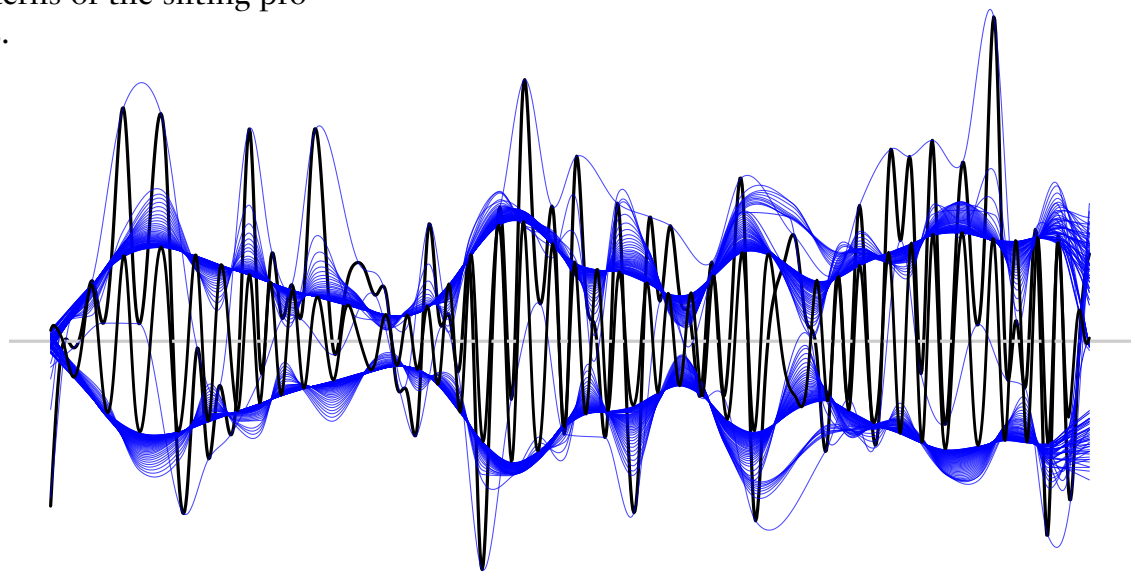
Suffice and easy to show by induction, this inequality holds when $\mathscr{S}\{h\}_+ = \mathscr{S}\{h\}_-$.

The extrema counting condition is a necessary condition for the second condition. As a weaker criterion, it characterizes the basic shape of the envelope splines. Declare functions violating the condition functions with degenerate envelope splines. Given a smooth function h, the second condition in our notation requires $\mathscr{S}\{h\}_+ = -\mathscr{S}\{h\}_-$. This is a very strong condition: given an upper envelope spline, the lower envelope spline is unique defined. Call a function conforming to the condition a function with symmetric envelope splines. One may deduce the non-degenerate envelope condition from the symmetric envelope condition easily for example through induction. Yang and Yang (2009) makes a similar point.

3.2.2 Homotopy

The condition specified in Theorem 5 is rather stringent. Since the function under consideration is assumed to be arbitrary, there is no a priori theoretical assurance of the needed symmetry which requires its upper envelope spline to match its lower envelope spline. In order to obtain an IMF for a function, a heuristic stoppage condition must be employed to declare the envelope splines are symmetric "enough"; the hypothesized condition of the definition is met; and an IMF is produced. In each failed iteration where no IMF is produced, an asymmetric part of the function—that is the mean of the upper and the lower envelopes—is subtracted from the function and moved to its residue. Since all IMFs have

Figure 3.4 Convergence patterns of the sifting process.



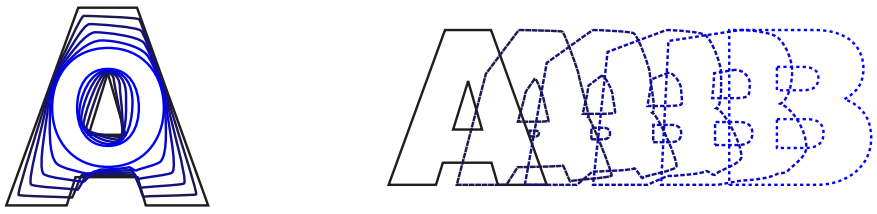
Note. Blue lines converging inwards are enveloping splines generated in the sifting process (see Theorem 5). The black line with greater amplitudes is the original time series. The IMF identified, by virtue of having symmetric envelops, is the inner most black line.

symmetric upper and lower envelopes, the sifting process can be though as a method to "symmetrify" a function, by discarding asymmetric parts iteration by iteration until the resulting envelopes are symmetric enough according to the stoppage condition. Figure 3.4 illustrates a typical run of the sifting process. Blue curves are the envelope splines wrapping the function. The sifting process warps the function, until it is sufficiently symmetric, shown as the blue curve contained in the clusters of symmetric envelope splines. In geometry, the continuous warping process is called a homotopy.

DEFINITION 1. A HOMOTOPY between two continuous functions f and g from a topological space X to a topological space Y is defined to be a continuous function $H : X \times [0,1] \to Y$ from the product of the space X with the unit interval $[0,1]$ to Y such that, if $x \in X$ then $H(x,0) = f(x)$ and $H(x,1) = g(x)$.

Continuous functions f and g are said to be HOMOTOPIC if and only if there is a homotopy H taking f to g.

Figure 3.5 The letter forms A and O are homotopic, but not with B.



Note. Contracting the outer edge and expanding the inner edge gives a homotopy from A to O. But there is no homotopy from A to B.

Figure 3.5 further illustrates the idea. It's clear that homotopy is a equivalence class. Let $f, g : [0,1] \to X$ be two paths in X. We can compose the paths by letting transverses first f and then g. Consider in particular a function f starting and ending at the same point x: its path is a loop and x is the basepoint. The set of all homotopy $[f]$ of loops at the basepoint x is call the FUNDAMENTAL GROUP of X at , denoted by $\pi_1(X,x)$. Let h be a path from x to x'. For each loop $f \in \pi_1(X,x)$, the conjugacy by h, hfh^{-1}, gives a loop in $\pi_1(X,x')$. Since the conjugacy is isomorphic for simply-connected spaces, we shall drop the basepoint from the notation.

LEMMA 6 (IMF Homotopy). *Every IMF is homotopic to a constant function wrapped on a circle n times in either the clockwise or the counterclockwise direction.*

Proof. The proof breaks into two parts: (**1**) I first show any loop on $\mathbf{S}^1$ is homotopic to the composite of single loops—this step is technical but standard; and (**2**) I then show any IMF is homotopic to a loop on a circle.

(**1**) Let p be the canonical descending map that sends a point on the real line to a point on the circle:

$$p : \mathbf{R} \to \mathbf{S}^1, \quad s \mapsto e^{i2\pi s}.$$

Define $\omega_n(s)$ to be the n-times self-winding map:

$$\omega_n(s) : \mathbf{S}^1 \to \mathbf{S}^1, \quad s \mapsto (\cos 2\pi ns, \sin 2\pi ns), \tag{1}$$

which gives a loop at a basepoint say $(1,0)$ on the complex plane. Observe $[\omega_1]^n = [\omega_n]$, since the homotopic equivalence stated earlier. Let $\widetilde{\omega}_n : I \to \mathbf{R}$ be the map $s \mapsto ns$, such that $\omega_n = p\widetilde{\omega}_n$. The map $\widetilde{\omega}_n$ is said to be the lift of ω_n and $\mathbf{R}$ the covering space of $\mathbf{S}^1$.

For each neighborhood U_s of a point on the circle, $s \in \mathbf{S}^1$, p^{-1} lifts U_s to a disjoint union of n open sets in the covering space. If we fix both x and its lift $\widetilde{x} \in p^{-1}(x)$, we can *uniquely* lift each loop f starting at x in the base space to a path $\widetilde{f}$ starting at that specified lifted point $\widetilde{x}$ in the covering space.[3]

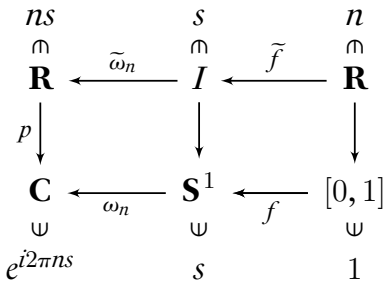
Let $f : [0,1] \to \mathbf{S}^1$ be a loop starting at the basepoint s and let $\widetilde{f}$ be the unique lift starting at 1. By definition, $f \in \pi_1(\mathbf{S}^1, x)$ and by construction, $p\widetilde{f}(1) = f(1) = s$, so s is listed at some integer n. But $\widetilde{\omega}_n$ is the loop constructed above in $\mathbf{R}$ from 0 to n. Therefore $(1-t)\widetilde{f} + t\widetilde{\omega}_n$ gives a homotopy from $\widetilde{f}$ to $\widetilde{\omega}_n$ and composing it with p gives a homotopy in the base space from f to ω_n. This therefore shows $[f] = [\omega_n]$.

To show $[f]$ uniquely determines n. Suppose the contrary: f is homotopic to both ω_n and ω_m, for some $m \neq n$. But homotopy is transitive, so lifting the homotopy yields $\widetilde{\omega}_n \simeq \widetilde{f} \simeq \widetilde{\omega}_n$ in the covering space. Yet $\widetilde{f}(1)$ is uniquely lifted at n, so $m = n$, a contradiction.

This therefore shows every loop in $\mathbf{S}^1$ at the same basepoint is homotopic to ω_n for a unique $n \in \mathbf{Z}$ and $\widetilde{f}$ be a lift at 0.

---

[3] The proof follows the introduction of the isomorphism between the first fundamental group of a circle to the infinite cyclic group generated by the homotopy class of the single loop on the circle (*i.e.*, $\pi_1(\mathbf{S}^1) \simeq \mathbf{Z}$) in Hatcher (2001). Readers can refer to the famed exposition for more relevant technical details.



(**2**) Let $h(x)$ be an IMF. By Theorem 5, it has symmetric envelope splines $g(x) := \mathscr{S}\{h\}_+ = -\mathscr{S}\{h\}_-$. The map $F_t = (1-t)h(x) + t(g^{-1}h)(x)$ gives a homotopy from $h(x)$ to a loop wrapped around a circle of radius 1. Changing of parameter $t = x/T$ gives the map $h(Tt)$ hypothesized in (**1**). This therefore proves the claim.

The proof is technical but the intuition is clear (Figure 3.6). The sifting process iteratively sifts for components with symmetric splines—these are precisely functions that are homotopic to loops on a circle. The significance of these loops is established in the following theorem.

3.2.3 IMF Decomposition

Theorem 7 (IMF Decomposition). *Let $f(t)$ be any square-integrable function. There exists a countable set of IMFs $\{\omega_i\}_{i\in\mathbf{Z}}$ such that*

$$f(t) = \sum_{i=-\infty}^{\infty} c_i \omega_i(t),$$

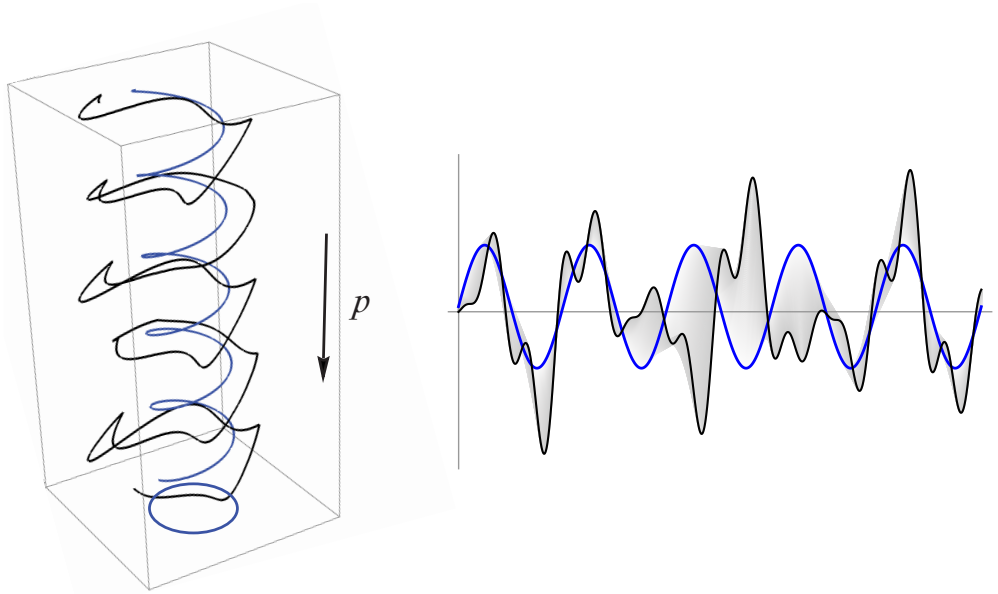
where $c_i(f, \omega_i)$ is a constant over t, for each $i \in \mathbf{Z}$.

Proof. This is immediate from Lemma 6. The n-times self-winding maps $\{\omega_n = e^{i2\pi nx} : n \in \mathbf{Z}\}$, constructed in the lemma are all IMFs with constant envelope splines. Define $\overline{\omega}(t) = e^{-i2\pi nx}$ and the inner product between two functions:

$$\langle f, g \rangle := \int_0^{2\pi} f(x)\overline{g(x)}\, dx.$$

Notice ω_n are precisely the orthonormal basis for the Hilbert space $\mathscr{L}^2([0, 2\pi])$ with the inner product defined above, which admits the Fourier series decomposition:

Figure 3.6 Homotopic reduction of a function to a loop on a circle.



Note. The descending map p is represented as the helix in 3-d space. An arbitrary function of is shown in black and the n-times self-winding map is shown in blue. In this illustration, $n = 5$, if we define clockwise winding as positive. We make the map $p : \mathbf{R} \to \mathbf{S}^1$ clear by first embedding it in $\mathbf{R}^3$ via $s \in \mathbf{S}^1 \mapsto (\cos 2\pi s, \sin 2\pi s, s)$ and then projecting it down to $\mathbf{R}^2$ via $(x, y, z) \mapsto (x, y)$.

$$f = \sum_{n=-\infty}^{\infty} \langle f, \omega_n \rangle \omega_n.$$

with the Fourier series defining to be

$$c_i := \langle f, \omega_n \rangle = \int_0^{2\pi} f(x) e^{-i2\pi nx}\, dx.$$

3.3. CONTRIBUTIONS

This paper contributes to the HHT literature in several significant ways. First, it formalizes a new identification criterion for the IMF that is more precise and conducive for theoretical discussions (Theorem 5). Second, it introduces the idea of homotopy to the nascent theoretical research of the HHT. The idea is instrumental in mathematics to prove various geometric invariances and fundamental for the constructions of other crucial modern mathematical objects like homology and, I believe, will set further theoretic discussions of the HHT on a more rigorous footing. Third, by integrating these elements, it provides an answer to the call for theoretic justification from the HHT literature, *e.g.*, Huang et al. (1996); Huang and Pan (2006); Kizhner et al. (2005), in the form of a new equivalence theorem (Theorem 7, along with Lemma 6) which establishes HHT's deep connection with the Fourier transform.

The Fourier transform can be seen as the sifting process in the limit; the sifting process, a bastardized Fourier transform around the "adaptive" basis. The n-times self-winding maps, $\{\omega_n = e^{i2\pi nx}\}$, give a basis to the infinite-dimensional functional space and allow all square-integrable functions to be decomposed in these basis. In practice, however, most time series HHT researchers interested in come from physical processes and can be represented by

a small set of basis—the task now becomes to identify these basis. Since each of Fourier basis is an IMF and can be represented by a loop starting at the same point around a circle, the symmetric envelope condition in Theorem 5 sifts out relevant loops from the infinite possible set of loops. If we allow the radius of the said circle to expand and contract according to the envelope spline $g(x) := \mathscr{S}\{h\}_+ = -\mathscr{S}\{h\}_-$, any IMF is homotopic to the composite loop $g\omega_n$. Since in Fourier theory, all loop types in both looping directions are needed to guarantee convergence for an arbitrary function, this in particular shows any IMF "convergence" is constructed by allowing the basis to vary with the observations. This in particular shows any general functional decomposition theory on IMFs (obtained through the sifting process or otherwise) is equivalent to the Fourier transform.

We are now in the position to address the unanswered theoretical questions posed by the originators of the sifting process in Kizhner et al. (2005) without relying on further conjectures about the statistical properties of the process:

> "[(a)] Why is the fastest changing component of a composite signal being sifted out first in the EMD sifting process? [(b)] Why does the EMD sifting process seemingly converge and why does it converge rapidly? [(c)] Does an IMF have a distinctive structure? (d) Why are the IMFs near orthogonal?"

a. Lemma 6 shows any IMF is homotopic to a loop on a circle. Fast changing IMFs have higher wrapping numbers. If we instead fix the wrapping number and allow the radius of the circle to expand and contract, a fast changing IMF is wrapped around a smaller circle. Since the sifting process is starting from the outermost envelopes, what have been enveloped are the remaining IMFs looped around smaller circles. This therefore explains why fastest changing components are sifted out first.

b. Convergence is true by construction, especially after the employment of a slack stoppage condition. We observe faster convergence comparing to the Fourier transform because of the particular nature of the time series under consideration. Fourier transform requires basis of all wrapping numbers in order to achieve convergence for an arbitrary function. Yet the time series under consideration are observations from cyclic physical processes. One can easily construct counterexample where the sifting process fails to give rapid convergence, or any meaningful convergence at all.

c. This is the content of Lemma 6: all IMFs can be homotopically transformed into Fourier basis.

d. This is due to the fact the Fourier basis are orthogonal:

$$\langle \omega_m, \omega_n \rangle = 0, \quad \text{for all } m \neq n.$$

3.4. EMPIRICAL EXAMPLE

We are now ready to address the empirical application of the HHT through a well-studied example. I have argued in the previous section that an alternative general convergence theory of functional decomposition through the sifting process is not possible. Since the HHT makes no statistical claim, any refinement of the theory must be grounded on a scientific theory. Indeed, in the hope to verify the usefulness of the HHT, evangelists of the theory have tested it on many time series with distinct features and well-understood data generating process. Among them, the LOD data is a most celebrated example.

Some information about the data source and the scientific context. The Earth Orientation Center of the IERS (International Earth Rotation and Reference Systems Service),

located at the Paris Observatory, provides to the geodesy community the international reference time series for the Earth orientation parameters (EOP), called "IERS C04" (Combined 04). The Universal Time (UT1) parameter contained tracks the Earth's rotation in time. Because the Earth's rotation is influenced by large-mass events, *e.g.*, the sea currents, UT1 is not linear with respect to Coordinated Universal Time (*i.e.*, the atomic time). The excess revolution time, measured in milliseconds, is called length of day (LOD). We obtained the data through Paris Observatory IERS ICRS Center's website and replicated the result of Huang and Pan (2006) in Figure 3.7. As an example of the power of the HHT, Huang and Pan (2006) documents the following data features extracted by the IMFs:

a. IMF1 has a 14-day period and a 19-year modulation, representing the semimonthly tidal cycle and the Metonic cycle;[4]

b. IMF2 is mostly high-frequency weather storms; and

c. IMF3 has a 28-day period, representing the monthly tidal cycle, and smaller amplitudes in El Niño years.[5]

From a statistician's point of view, these features could be readily seen in the original time series and made obvious through standard spectral analyses. Any proponent of an empirical method needs to confront the confirmation bias. Huang and Pan (2006) though documents noticeably sharper changes in amplitudes after early 1980s in for example IMF1

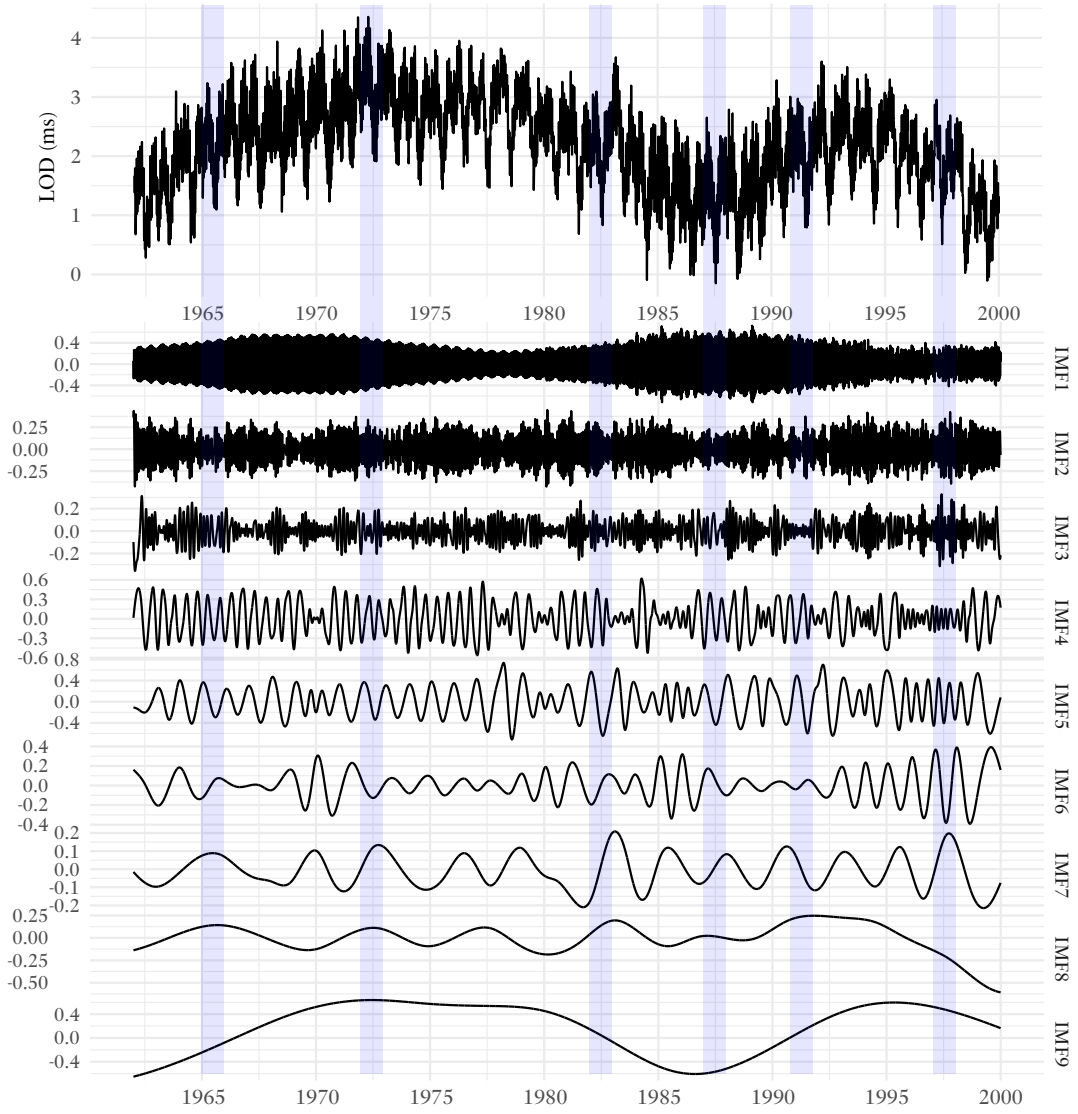
[4] A period of 19 solar years is almost exactly equal to 235 synodic (lunar) months, first noted by the polymath Meton of Athens.

[5] The Oceanic Niño Index (ONI) is one of the primary indices used to monitor the El Niño-Southern Oscillation (ENSO). NOAA (National Oceanic and Atmospheric Administration) uses it to identify El Niño (warm) and La Niña (cool) events in the tropical Pacific, though the agency uses the HHT to construct the index and amends the classifications.

and hypothesizes that this "may be attributed to the change in the density of data," does not however offer explanations to other apparent data features. Since the first two IMFs are dense with diverse fluctuation patterns over time, we can iterate the sifting process multiple times to investigate their empirical mode composition, see Figure 3.8. These iterated IMFs consistently contain Gaussian wave patterns at late 1980s and mid 1990s: these patterns are not immediately apparent in the original time series. However, whether they are important data features or spurious patterns generated by the choice of splines and ending conditions can only be answered by earth scientists.

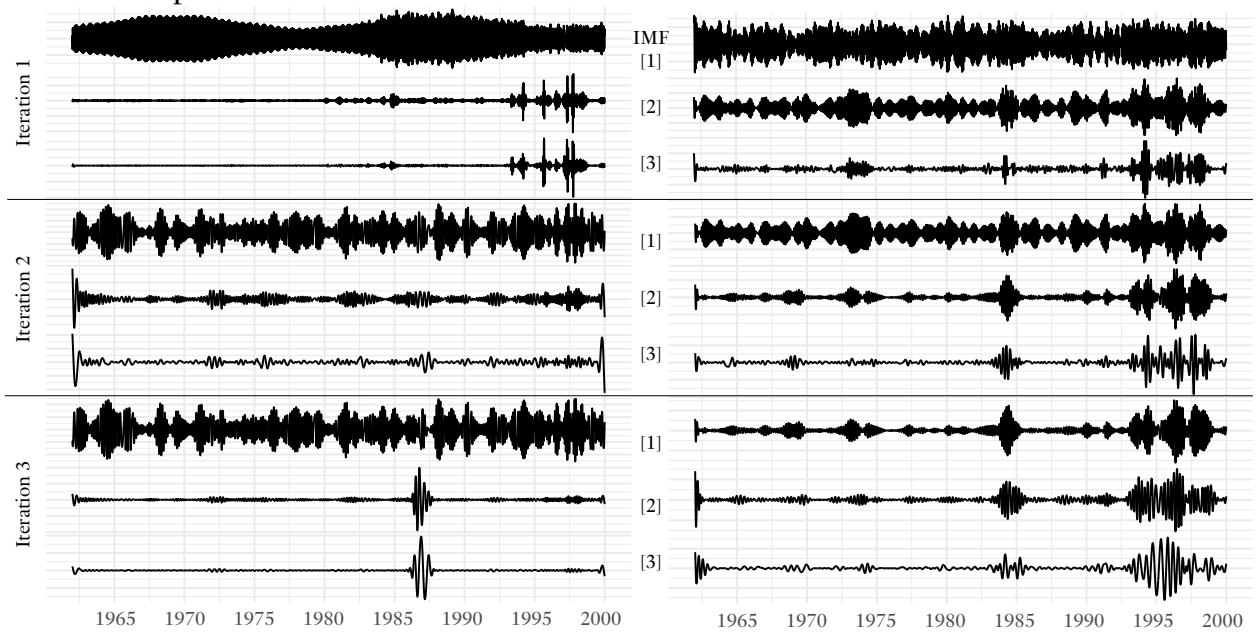
This paper provides a rigorous theory explaining why the transform yields useful results for its proponents but at the same time yields a stern limit of what it can achieve for an arbitrary time series. So when will the HHT be useful? Until it is not. This is not to dismiss the empirical value of the HHT. However, our job as mathematical statisticians ended after characterizing the mathematical nature of the method in the preceding section. A shortcut needs no theory to be useful and no theory can make a shortcut more correct without overparameterization, which only defeats its purpose qua shortcut. Since the HHT is a heuristic process without substantive claims, it is far more fruitful to take the handy hints of its results to theorize the underlying physics rather than elaborating further on the algorithm itself in the hope to make its results more credible. There is no magic algorithm for science. Box's warning about overelaboration applies: "Since all models are wrong the scientist cannot obtain a 'correct' one by excessive elaboration. On the contrary following William of Occam he should seek an economical description of natural phenomena. Just as the ability to devise simple but evocative models is the signature of the great scientist so overelaboration and overparameterization is often the mark of mediocrity. (Box, 1976)"

Figure 3.7 Empirical mode decomposition of the length-of-day data.



Note. The original time series of the IERS LOD observations from 1964 to 1999 is shown in the top panel. The remaining nine time series are IMFs identified through the sifting process. The plots of the IMFs are not to the same scale (actual scales noted on *y*-axes). The parameter, $UT1 - UTC$, is part of the Bulletin B EOP Combined Series C04, under 2014 International Terrestrial Reference System (ITRS2014). El Niño years shaded in blue (ONI>1.5), see Climate Prediction Center (2018).

Figure 3.8 Three more iterations of the empirical mode decomposition.



Note. We iterate the sifting process three times (from top to bottom) on IMF1 (left panel) and IMF2 (right panel), identified in Figure 3.7. For convenience, only the first three IMFs are included for each iteration, since the rest of IMFs do not have significant amplitudes.

CHAPTER 4

DISCUSSIONS

The spherical space form is ubiquitous in statistics, though often in disguised forms. As a manifold with a constant sectional curvature, it serves as an indispensable geometric reference to study nonlinear problems in statistics. In the first example, it emerges naturally as the benchmark space, due to the curvature bound placed on the manifold under investigation. In the second example, it provides the geometric representation of the basis of the Hilbert space as well as their empirical counterparts, the IMFs. Unlike it is in a Euclidean space, a vector in a curved space cannot be moved about freely through the usual arithmetic of coordinates. The innocuous statement has serious consequences. Losing sight of this fundamental restriction leads to a cumbersome proof that structurally depends on an unproven assumption alleviating this restriction in the first example and a theoryless algorithm that essentially attempts to bypass this restriction in the second example. Differnetial geometry is the solution, even though our interest is strictly statistical.

The greatest lesson of differential geometry is the idea of local trivialization. A manifold, with its nonlinear global properties, is always reduced to Euclidean spaces locally with the so-called charts. Even though charts collectively are known as the atlas of the manifold and we are assumed to be able to travel over a manifold freely with the atlas by stitching

up patches of the charts together, the manifold however is not assumed to have global coordinates. In other words, differential geometry provides the needed tools to study a manifold without relying on any particular embedding in a vector space. This idea extends to higher constructions like tangent bundles. A vector bundle, when needed to be expressed in coordinates, is always done so through "trivializations"—no matter how complicated it transforms over the whole manifold, locally it is trivialized as $U \times \mathbf{E}$, where U is an open neighborhood on the manifold and $\mathbf{E}$ is a vector space. Algebraic topology takes this step even further and completely dispenses with charts. The most important properties of

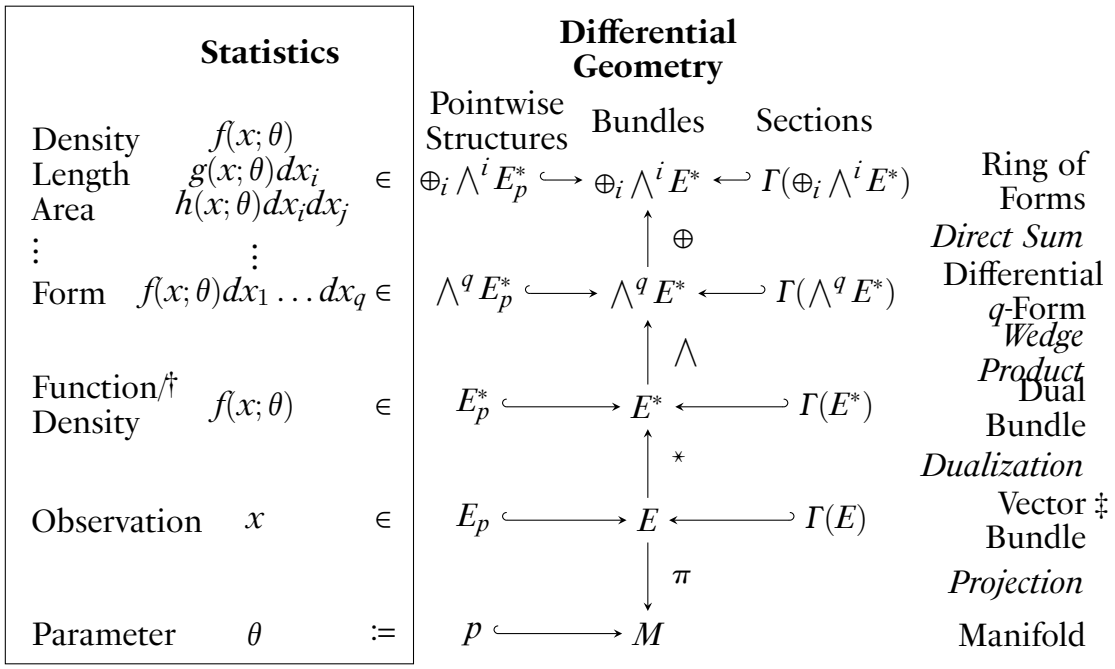


Figure 4.1 Statistics from differential geometric point of view

† 1-dimensional objects shown, higher dimensional objects can be constructed analogously, *e.g.*, with Cartesian product $\times$ and tensor product $\otimes$.

‡ The tangent bundle TM is the most important example of a C^∞-vector bundle.

geometric objects and invariances of smooth maps between them can often be described algebraically (so-called intrinsically) without referencing any coordinate, local or global.

As Rao (1945) demonstrates, statistical inference exploits the nonlinearity of a functional form. Since differential geometry is the mathematical language of nonlinearity, to any statistician with modern geometry background, the lure for an integrated geometric statistical theory is immense. Almost all statistical objects have straightforward counterparts in differential geometry: Parameters under restriction can be thought of as a regular manifold; observations as points in some vector bundle over the manifold; probability densities as differential forms; expectations as integrations over vector fields; parameter estimation as fiber identification (Figure 4.1). If only there were a unified geometric statistical theory, any statistical confusion of a geometric sort, as demonstrated in this paper, could be completely avoided! Indeed, this parallel construction project is the basic research agenda of information geometry. The vision is clear; the impact would be monumental; and even the journey seems noble and idyllic. But in spite of Rao's pioneering work, Hilbert's spiritual guidance, and the dedicated and persistent work of talented researchers, these efforts by all measures have stalled.

In my opinion, the history of Esperanto, the ill-fated universal language, provides a cautionary tale. Zamenhof, Esperanto's creator, shares information geometers' vision: "Were there but an international language, all translations would be made into it alone [...] and all nations would be united in a common brotherhood. (Zamenhof, 1889)" Yet in spite of generations' efforts, no work of cultural significance has been produced, save translated work that the language is designed to avoid. Statistics and differential geometry are in essence two different languages. Even though the latter has more natural expressions for nonlinear phenomena and affords valuable geometric insights through its lexicon, the quixotic at-

tempt of interdisciplinary researchers in information geometry ignores the historical contingencies of scientific research qua human endeavor: Fisher (1922) has set statistics on its own drift; occasional inconveniences cannot turn the tide on the parting disciplines. Hilbert's declaration that any mature science automatically becomes integrated with mathematics might very well be true but maturity is a natural process and it comes with growing pains. This paper remedies these pains rather than offering a prescription for premature integration.

4.1. LITERATURE REVIEW

Though Mahalanobis (1936) first gives the measure of distance of an observation $x = (x_1, x_2, \ldots, x_N)'$ of a multivariate normal distribution with mean μ and covariance matrix Σ,

$$\rho_M^2 = (x-\mu)'\Sigma^{-1}(x-\mu),$$

now canonized as the Mahalanobis distance and Bhattacharyya (1943, 1946) extend the geometric idea of the Mahalanobis distance to a measure of divergence between two population, it is Rao (1945) who first explicitly introduces to statistics the idea of a Riemannian metric (in the form of Fisher information) and the associated geodesic distance (called the Rao distance) on the parameter space (*i.e.*, the Rao space) viewed as a differential manifold. The choice of metric breaks off from the geometry literature and is influenced by Fisher, who presented a full account of the foundations of theoretical statistics in an extraordinary exposition, Fisher (1922), wherein the expected information quantity was used and formally presented shortly after (Fisher, 1925, 1990 reprint). Established as "an unbroken link in the continuing evolution of modern statistics (Pathak, 1997)," Rao essentially ex-

tends the asymptotic concepts of consistency, efficiency, sufficiency, and information loss introduced in Fisher (1922) to finite samples.

Efron (1975) revives his geometric line of investigation by introducing the concept of "statistical curvature," which quantifies how similar an arbitrary one-parameter family of curves are to the exponential family: the quantity is of statistical interest because the MLE for a vector parameter is a sufficient statistic only for multiparameter exponential families (Fisher, 1922, 1934). As its accompanying discussant paper, Reeds (1975), similarly argues, since Rao (1945) establishes the Fisher information as a Riemannian metric, earlier results (Huzurbazar, 1950, 1956; Mitchell, 1962; Holland, 1973) in their collaboration constructing explicit matrix expressions of the transformations of the Fisher information in orthogonal parameters, in fact implicitly seek to construct an affine connection in a moving orthonormal frame—an archetypal differential geometer question—and Efron (1975) is the "logical successor" to present a more cogent differential geometric picture to statistical estimation. Dawid (1975, 1977), citing Hicks's (1965) emphatic caution against explicit coordinate-dependent approach to apply differential geometry to the study of statistics,[1] sketches out Efron's main arguments in coordinate-free languages and gives the definition of the Efron connection, implicitly used in Efron (1975). Applying the idea of "statistical curvature," Efron (1978) and Efron and Hinkley (1978) describe some geometric results relating the observed and the expected parameter spaces of the multivariate exponential family in the econometric context of assessing the accuracy of the MLE $\widehat{\theta}$: to what extent

[1] I couldn't find relevant quotes from the source and can't independently verify the claim but as the Huzurbazar-Mitchell-Holland explicit construction of the Fisher information matrix transformations illustrates, a coordinate-free approach of differential geometry could avoid cumbersome calculus computations caused by the need to keep track of coordinate changes and as a result often gives cleaner proofs to questions of existence and invariance.

can the Fisher information $I(\theta)$ be replaced by $-\partial^2_{\theta^2} \log f(x;\theta)|_{\theta=\widehat{\theta}}$ in the variance bound for $\widehat{\theta}$?

Encouraged by the success of Efron and his collaborators, Madsen's thesis (1978; 1979) discusses their implications on assessing the second order estimation accuracy of the MLE and competing estimators; Kass's thesis (1980) develops similar geometric ideas for the model space by focusing on the Jeffreys priors, which is proportional to $\sqrt{\det I(\theta)}$ and by construction invariant under reparameterization of the parameter vector θ (Jeffrerys, 1998 reprint); and Amari (1982a,b), synthesizing his earlier notes (1968; 1980) and Efron's work on one-parameter family of curves, presents to a larger audience a full multiparametric theory of curved exponential families with new forms of "curvatures." Barndorff-Nielsen et al. (1986) gives a nontechnical account of the role of differential geometry in statistical theory.

4.2. GLOSSARY

Notation and conventions largely follow Lang (2001). A bijection $f : X \to Y$ is called a **diffeomorphism** if both f and f^{-1} are smooth, *i.e.*, all partial derivatives exist. We often say **locally**, meaning there is a open neighborhood with the desired property for each point. A set is called a **manifold** if locally it is diffeomorphic to $\mathbf{R}^m$. We will call it for short a m-manifold. We use **E** for a Euclidean space of an arbitrary dimension. The diffeomorphism $\phi : U_x \to \mathbf{E}$ is called a **chart**, or local coordinates.

You are familiar with tangent vectors. In differential geometry language, a **tangent vector** is an equivalency class. Let v be a vector in E, we identify v with all vectors w such that there exists another compatible chart ψ:

$$(\psi\phi^{-1})'(\phi x)v = w.$$

Observations x are vectors *without* coordinates. All tangent vectors at a point x form a vector space called the **tangent space**, denoted by T_xX. Disjoint union of spaces gives the idea of a **bundle**. The disjoint union of tangent spaces $T_x(X)$ for all points on manifold is called a **tangent bundle**. We can similarly form a **vector bundle** over a manifold with the map $Y\,\pi : E \to Y$, by associating a vector space **E** with each point of Y and disjoining these copies of spaces together. The vector space $\pi^{-1}y$ is often denoted by E_y and it's called a **fiber** at y.

Let $\pi : E \to Y$ be a vector bundle and $f : X \to Y$ a map. Then the vector bundle at Y can be pulled back at X by associating x with the vector space $(f^*E)_x := E_{f(x)}$ and this induces the vector bundle $f^*\pi$ (and the bundle map $\pi^* f$):

$$f^*\pi : \quad f^*E \to X.$$

We call f^*E the **pullback** of E by f. (We can analogously form the **pushforward** bundles and these induced maps are denoted by f_*.) Precomposition with a function provides the intuition of a pullback.

Because we are dealing various types of statistical objects, it is needed to introduce the category language—this allows us to describe maps between these objects without having to construct them formally. A **category** is any collection of objects with associative composition law with an identity element for each object, called **morphisms**, identified by the objects $[\![X, Y]\!]$. Sets, groups, rings, manifolds, metric spaces, vector spaces are all example

of categories (with their corresponding morphisms and additional structures). A function is a morphism between sets: the converse is not true.

$[\![X,X]\!]$ is called an **endomorphism**. $[\![X,Y]\!]$ is called an isomorphism if each morphism has a inverse. An **isomorphic** endomorphism is an **automorphism**. An isomorphism in the category of manifolds is called a **homeomorphism**. Functors is a meta-category: a category of categories. A functor gives a map between categories that is also a morphism between these categories—by definition, it respects their respective identities. Let F be a functor and $p,\ q$ be morphisms: if $F(pg) = F(p)F(q)$, the fuctor is said to be **covariant**; if $F(pg) = F(q)F(p)$, it is called **contravariant**.

Derivatives generally exist without calculus. We often say locally, meaning there is an open neighborhood with the desired property for each point. Let $p : \mathbf{E} \to \mathbf{F}$ be a continuous map. If locally there exists another map $dp : \mathbf{E} \to \mathbf{F}$,

$$p(\theta_0 + y) = p(\theta_0) + dp\, y + \psi(y)$$

for a small y. dp is called a **derivative** and ψ is tangent to 0. Let $p_i : \mathbf{E} \to \mathbf{F}_i$ be continuous maps. Partial derivatives df_i by holding all but the i-th component fixed. Functor morphisms induce **natural** transformations of the objects.

Let $f : X \to Y$ be a map between vector spaces (or groups in general). The **image** of a map is the subspace $f(X) \subset Y$, denoted by $\operatorname{Im} f$ and the **kernel** of it is the subspace $\operatorname{Ker} f \subset X$, such that for each element $x \in X$, we have $f(x) = 0$. The quotient $Y/\operatorname{Im} f$ is the **cokernel** fo

the map f. The sequence

$$X \xrightarrow{f} Y \xrightarrow{g} Z$$

is called **exact** if the image of f is equal to the kernel of the g. If follows these definitions that:

$$\begin{cases} 0 \to X \xrightarrow{f} Y \\ \mathrm{Y} \xrightarrow{g} Z \xrightarrow{0} \end{cases} \text{ is exact, if and only } \begin{cases} f \text{ is injective;} \\ g \text{ is surjective.} \end{cases}$$

The short exact sequence

$$0 \to X \xrightarrow{f} Y \xrightarrow{g} Z \to 0$$

is called **split** if there exists a homomorphism $h : Z \to Y$ such that the composition gh is the identity map on Z. For any map $f : X \to Y$, the following exact sequence connects the kernel with its cokernel:

$$0 \to \operatorname{Ker} f \to X \xrightarrow{f} Y \to \operatorname{Coker} f \to 0.$$

Some terminology to facilitate the discussion of maps between manifolds. Let $f : X \to Y$ be a map between two manifolds and (U_x, ϕ) and $(V_{f(x)}, \psi)$ be charts at x and $f(x)$ correspondingly. Consider the map on a product space $U_1 \times U_2$ (as an open mapping):

$$f_{V,\,U} \coloneqq \psi f \phi^{-1} : U_1 \times U_2 \to V_{f(x)}.$$

f is an **immersion** if and only we can find charts that makes $f'_{V,\,U}$ injective; f is an **submersion** if and only we can find charts that makes $f'_{V,\,U}$ surjective. Both idea are intuitive as

descriptions of isomorphism classes with respect to the larger product space: an immersion gives an isomorphism from U to a subspace in $U_1 \times U_2$; submersion gives an isomorphism from U to the whole product space $U_1 \times U_2$. An injective immersion is called a **embedding**. An important example of an immersion that is not an embedding is a flat surface to a self-intersecting surface.

PART II

EMPIRICAL REMEDIES

„Das Bild ist eine Tatsache.[2]*“*

Abschnitt 2.141,

Tractatus Logico-Philosophicus (1921)

— Ludwig Wittgenstein

[2] “A picture is a fact.” D.F. Pears and B.F. McGuinness’s translation.

CHAPTER 5

GRAPHIC TOOLS FOR HARMONIC ANALYSES OF TIME SERIES

The preceding part draws attention to certain hardwired theoretical limitations of handling spherical space forms in statistics with classical mathematical tools and calls for algebraico-topological extensions of existing theories: in Chapter 2, the theoretical limitation is manifested in the unjustified assumption (their Assumption 1) about the efficacy of the very analytico-geometric approach undertaken and we transcend from the dialectical contradiction by reframing the problematics with coordinate-free differential-geometric languages; in Chapter 3, the limitation is recast as the inability of a *prima facie* empirically useful statistical procedure to make any refutable statements and we put its mathematical core on display under the penetrating lens of modern algebraic topology.

These theoretical discussions are in essence a methodological manifesto—but without offering any concrete statistical procedures, albeit of long-term theoretical interest, they remain abstract and far removed from the technocratic ethos of the statistical practice today. To alleviate this shortcoming and further ground these lofty discussions in the current communal research efforts, I now offer a collection of novel interactive graphic tools for the empirical study of time series, in lieu of a grand synthesis unattainable at the moment. These tools, in spite of their diverse appearances, are in fact the progeny of the foregoing

discussions on the geometric interpretations of harmonic decompositions of time series. Thematically, Chapter 2 lays out the foundation in the form of a generalized curvature comparison theorem to use spherical space forms in statistics, *qua* the model space of a constant curvature. Chapter 3 further demonstrates how these spherical space forms, as basis for the Fourier transform, are used to reconstruct a time series and cautions against intuitionists' attempt to formulate more expedient decompositions. Chapter 2 sets the stage; Chapter 3 points to a void; and this chapter remedies this theoretical void with a suite of new graphic tools to help practitioners explore the dynamic harmonic structures of high-dimensional time series.

5.1. BACKGROUND

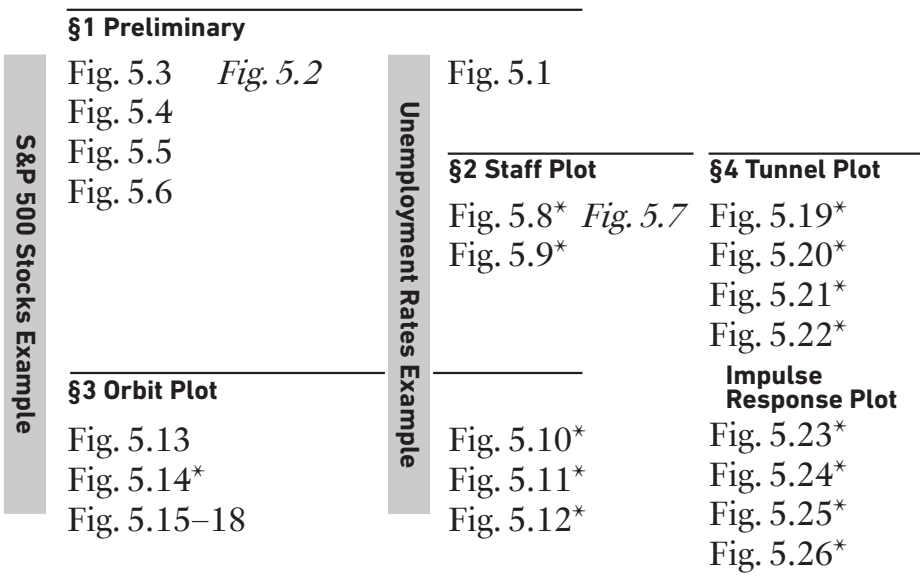
Seeing is believing. Yet unlike performing statistical tests, the simple task of seeing is strictly confined to 2 dimensions, in print and on screen. Even imagination can only extends it into the third. Adding time, it is the complete enumeration of the space-time. But the universe, according to the M-theory, has 11 dimensions—we are all low-dimensional creatures in a high dimensional universe! To visually express high-dimensional data relations—dynamically in the context of time series—with a plot of by comparison negligibly low dimensions is part data science, part art of (mis)direction. Thanks to the ever cheaper computational power, manipulating high-dimensional objects numerically has become a routine task of the trade. Yet without compatible graphical tools to visualize these high-dimensional objects, statistical research is essentially blind.

We provide a suite of four new graphical tools to help researchers explore the dynamic structures of high-dimensional time series: (a) the staff plot; (b) the orbit plot; (c) the tunnel

plot; and (d) the impulse response plot. Unlike traditional plots designed for print, these plots presented are a new breed of screen-first graphical applications with rich interactive features. Utilizing latest programing libraries for high-performance graphic rendering and runtime manipulation of graphic elements, these new tools, though equally at ease on paper as figures in this chapter demonstrate, are designed to be workhorse data exploration utilities. The staff and the orbit plots are animated 3-dimensional models: they are designed to address the shortcomings of scatter plots. The tunnel and the impulse response plots are interactive 2-dimensional plots: they are improved heat maps.

Figure 5.0 outlines the organization of the plots presented in this chapter. Typically, in papers debuting static plots, authors showcase the versatility of a new plot with diverse

Figure 5.0 Interdependency of figures in this chapter.



Note. Key plots presented in this chapter are marked by an asterisk (*). Illustrative figures not directly related to the examples are in italic. Figures in this chapter are optimized for screen or *color printing at a resolution of at least 300dpi.*

examples. We however take more of a case study approach with two involved examples, for two simple reasons. First, we have an array of unconventional tools to introduce. The sheer number of new plots demand our focus. Second, our goal is to exhibit intricate dynamic structures hidden in high-dimensional time series. The suite of tools are meant to be used in tandem. Focusing on the same datasets allows us to tease out these threads from different angles. Since scatter plots and heat maps are standard statistic tools, readers are invited to start from Section 5.2 and refer to Section 5.1.2 for background information on the examples.

5.1.1 Problems with Existing Plots

The scatter plot, simple in its construction as the direct spatial representation of a 2- to 3-dimensional dataset, is still the definitive tool to visualize complex patterns of bivariate dependency. Its strength is its faithfulness: viewed as a map, a scatter plot gives an isomorphic representation to a low-dimensional dataset; one can theoretically reconstruct the dataset from the plot, given a measurement device of arbitrary precision. We will emphasize this conceptual point throughout the chapter and treat it as the guiding design principle—for each plot, we will discuss at the start whether it represents the data faithfully and; if certain data have been modded out, what are the justifications for the equivalence relation implicitly declared. This might seem at first a pedantic exercise but, as the chi-plot example demonstrates, essential to establish plotting as a scientific discipline.

Faithfulness comes at the cost of clutter, which is the obvious shortcoming of scatter plots. Since all data are transcribed literally as points in a Euclidean space, the scatter plot is a low-dimensional graphic tool and researchers would have to examine scatter plots for all

variable pairs in order to piece together the correlation structures [1] of a high-dimensional dataset. A natural solution is to shrink its size and present many scatter plots in a matrix (Hartigan, 1975; Cleveland and McGill, 1984). R alone has numerous implementations of this basic idea, some with additional complications: `pairs` (R-Core), `lattice:splom` (Sarkar, 2008, Trellis Graphics for R), `DescTools:PlotMatrix` (et al., 2018, Tools for Descriptive Statistics), and `Deducer:ggcorplot` (Fellows, 2012, A Data Analysis GUI for R). However, these plot matrices only confound the cluttering problem and become impossible to read for high-dimensional data (Friendly, 2012).

Plots like the chi-plot[2] (Fisher and Switzer, 2001, based on the chi-squared statistics) and the K-plot (Genest and Boies, 2003, based on Kendall's tau) are essentially transformed scatter plots and, as we argued above, cannot beat the scatter plot in its faithfulness. Take the chi-plot for example. It is said to reveal far richer bivariate dependence relations to formal statistical tests such as Wilbert C. M. Kallenberg and Rafajłowicz (1997); Kallen-

[1] We use the term "correlation" for Pearson's correlation coefficient throughout the chapter. When we say the correlations of a multivariate time series, we broadly include autocorrelations and cross-correlations of all pairs of variables up to a given lag, unless stated otherwise.

[2] Consider a n-sample of 2-dimensional observations $\{(x_i, y_i)\}_{i=1}^n$. For any given pair (x_i, y_i), we can define the following empirical c.d.f.s:

$$\begin{cases} F_i = \frac{1}{n-1}\#\{x_j : \ x_j < x_i\}; \\ G_i = \frac{1}{n-1}\#\{y_j : \ y_j < y_i\}; \\ H_i = \frac{1}{n-1}\#\{(x_i, y_j) : \ x_j < x_i, \ y_j < y_i\}. \end{cases} \quad \text{and} \quad \begin{cases} F_i^\bullet = F_i - 1/2; \\ G_i^\bullet = G_i - 1/2, \end{cases}.$$

The chi-plot is the graph of (χ_i, λ_i), where

$$\chi_i = \frac{H_i - F_i G_i}{\sqrt{F_i(1-F_i)G_i(1-G_i)}} \quad \text{and} \quad \lambda_i = 4\text{sgn}(F_i^\bullet G_i^\bullet)\max(|F_i^\bullet|^2, |G_i^\bullet|^2). \tag{1}$$

berg and Ledwina (1999), suggesting a deeper link between the chi-plot and the copula functions. The "richness of the graphs" is however the consequence of a poor choice of the distance function λ. The distinct lobed-structures in many chi-plots,[3] which are assumed in the literature to suggest mixing of different distributional families, are in fact largely spurious features of the jumping points in the signed distance function λ and simply disappear if we choose a more conventional smooth distance function.[4] Fisher and Switzer (1985, 2001) does concede there is no good reason to choose any particular functional form of the distance function λ and offers several alternative distance functions. But in Fisher and Switzer (1985, 2001) and related literature, the distance function in Equation (**1**) is almost used exclusively.

The heat map takes a different approach. It does away with geometric representations of variables but instead encodes the value of research interest, usually correlations, with colored cells. This new approach, thanks to the invention of high-resolution color display monitors, allows researchers to effectively visualize the correlations of a high dimensional dataset. Yet in practice the heat map is more of an impressionist painting than a precision instrument: it leaves researchers with a hazy impression of the overall correlation

[3] **Figure 5.1***(a)* reproduces the example of Fisher and Switzer (2001). The original dataset (not publicly available) come from Griffin et al. (1999), which analyzes the element compositions of 13,317 individual grains of mantle-derived peridotic garnet.

[4] **Figures 5.1***(b)–(d)* illustrate the point. Panel *(b)* gives the scatter plot of a randomly generated normal sample with a weak positive correlation. To examine how the chi-plot remaps each point, we introduce a color gradient to encode each point's 2-dimensional position. Panel *(c)* illustrates Fisher and Switzer's λ_i (Equation **1**) and panel *(d)* gives the corresponding chi-plot, with the signature arch. Note how originally adjacent points are redistributed according to the idiosyncratic lobed-structures of the chosen distance function. Panel *(d′)* demonstrates these artificial features disappear with a more mild choice of the distance function, *e.g.*, the linear distance function in *(c′)*.

distribution but makes identifying each cell by variables difficult. Clustering can help. But it doesn't solve either the precision or the navigational problems, while creating its own problems in the process: for example, clustering depends on the method of seriation and is computationally expensive (McKenna et al., 2016). Wilkinson and Friendly (2012) provides a historical review of the heat map. In the context of high-dimensional time series, we are often interested in studying how correlations change along lags. Heat maps for different lags like **Figure 5.2** can readily show the overall decaying of correlations as we gradually increase the number of lags. However, they offer little help to answering simple observational questions like these: a) how fast does the correlation of a given pair decay along lags? (b) are there any outliers? or (c) what variables do the cells of a certain color represent?

5.1.2 Persistence of Tidal-Locked Periods

We will use the suite of new tools to visualize the intricate dynamic structures of two well-studied datasets. The first example is the monthly unemployment rates (seasonally adjusted) of 50 U.S. States and the District of Columbia (D.C.) from January 1976 to May 2018 released by the U.S. Bureau of Labor Statistics, as reported by the Federal Reserve Bank of St. Louis. The second example is the 512 constituent stocks of the Standard and Poor's 500 (the S&P) index from February 8, 2013 to February 7, 2018, using the Daily Stock File database from the Center for Research in Security Prices (CRSP) at The University of Chicago Booth School of Business. **Figure 5.3** gives the complete list of these stocks and their Standard Industrial Classification (SIC) numbers: manufacturing (173 stocks, 33.8% of the sample); finance, insurance and real estate (96, 18.8%); services and public administration (74, 14.5%); transportation, communications, electric, gas and san-

itary service (68, 13.3%); wholesale and retail trade (58, 11.3%); mining and construction (31, 6.1%); and the rest do not have an official SIC code assigned (12, 2.3%).

It's not common practice to talk about phases of time series outside of the context of macroeconomics. Before arguing formally in Section 5.2 that phases are simply correlations on a sphere, see Equations (**2**) and (**3**), we first provide some elementary empirical evidences on the harmonic aspects of stock prices. This is of course not to present a new theory about stock returns, which is beyond the scope of this modest exposition on purely descriptive visualization tools. In fact, the very premise of studying stock prices, rather than their returns, is very much against the established conventions in the field of financial statistics, for sound financial and statistical reasons. However, since we don't engage with financial theories here and price levels in general have profiles more of the desired smooth wave-like forms (details see Section 5.3.1), let's treat them naively as anonymized time series stripped of far-reaching financial implications. Curious readers are invited to redo these exercises with stock returns with the open-source tools provided (Section 5.5).

We first rescale all prices from 0 and 1, since we are only interested in the shapes of these line plots. **Figure 5.4***(a)* shows these prices are very weakly correlated but with non-trivial dynamics as evidenced by the composite plot in Panel *(b)*. We see darkened areas in the bottom left corner (2003–2014), the upper right corner (mid 2017–2018), as well as a handful of darkened strands in the middle (*e.g.*, 2015): this suggests even the overall correlations are very weak, there are prolonged periods when a large subset of the time series move in synchrony. In particular, we notice two defined half circle patterns in the bottom left corner.

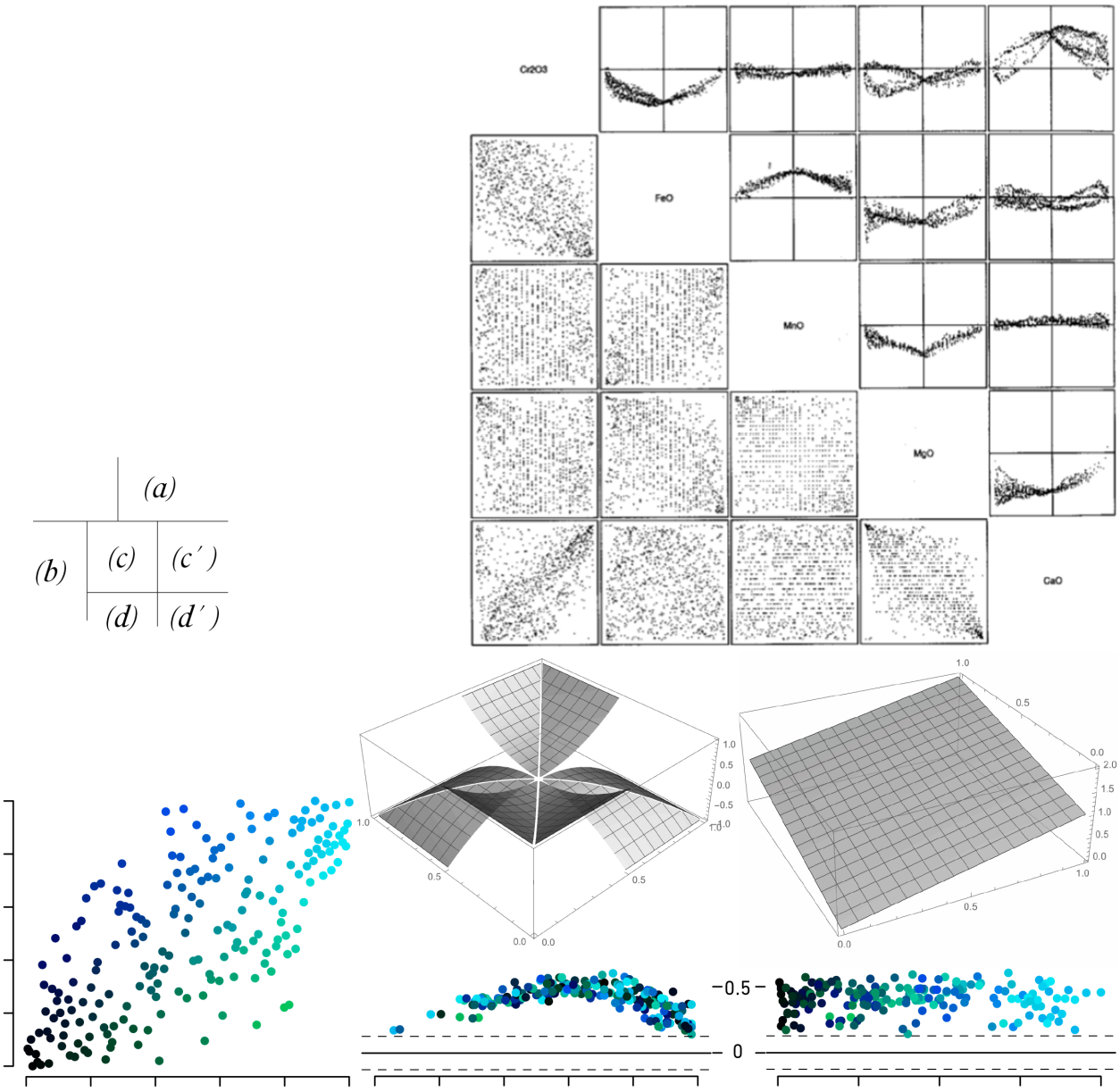
5.1 Background

To further analyze the phasal dynamics, we wrap the time series on a unit circle and track their phases, so that local maxima always give phase-0 and local minima always give phase-π (details see Section 5.3.1). **Figure 5.5***(a)* gives the phasal probability density across time. Panel *(b)* gives the corresponding contours. We observe large elliptical regions in magenta, indicating probability concentrations in some phases over time. Indeed, we see the three phasal density peaks at π at the start of the time series, corresponding to the half circles pattern described. To better visualize these phasal dynamics, we introduce a density-contour hybrid plot in Panel *(c)*, where only high density regions are highlighted. As we expect, there are periods when these time series are out of phase and fluctuate in their own cycles: *e.g.*, from 2017 to 2018, even though we observe dark bands of price levels in **Figure 5.4***(b)*, the phases of these series are evenly distributed; this is reflected in **Figure 5.5***(c)* as flat probability density curves from 2017 to 2018. However, there are also periods when these series move together and certain phases have elevated probability densities in *(c)*: *e.g.*, in late 2013 (around Period 150), we observe the probability density peak gradually shifts from phase-π to phase-0; in early 2014 (around Period 210), the peak moves from $3\pi/2$ to π; and from mid 2015 to early 2016 (Periods 600–800), there are significant and irregular phasal density shifts. What is happening? What components are moving in sync?

Figure 5.6*(b)* displays the composite phasal plot. Comparing to the composite level plot in **Figure 5.4***(b)*, we observe distinct patterns of phasal integration: when phases are out of sync, the composite phasal plot displays a blurry image of curves (*e.g.*, 2017–2018); when they come in sync, however, sharp sinusoid shapes emerge (*e.g.*, late 2013 and late 2015). Often, phases do not overlap but move in sync: we call the period when time series share the same harmonic frequency, their tidal-locked period. The intuition will become clear in Section 5.2 when we introduce the orbit plot. **Figure 5.6***(a)* counts the length

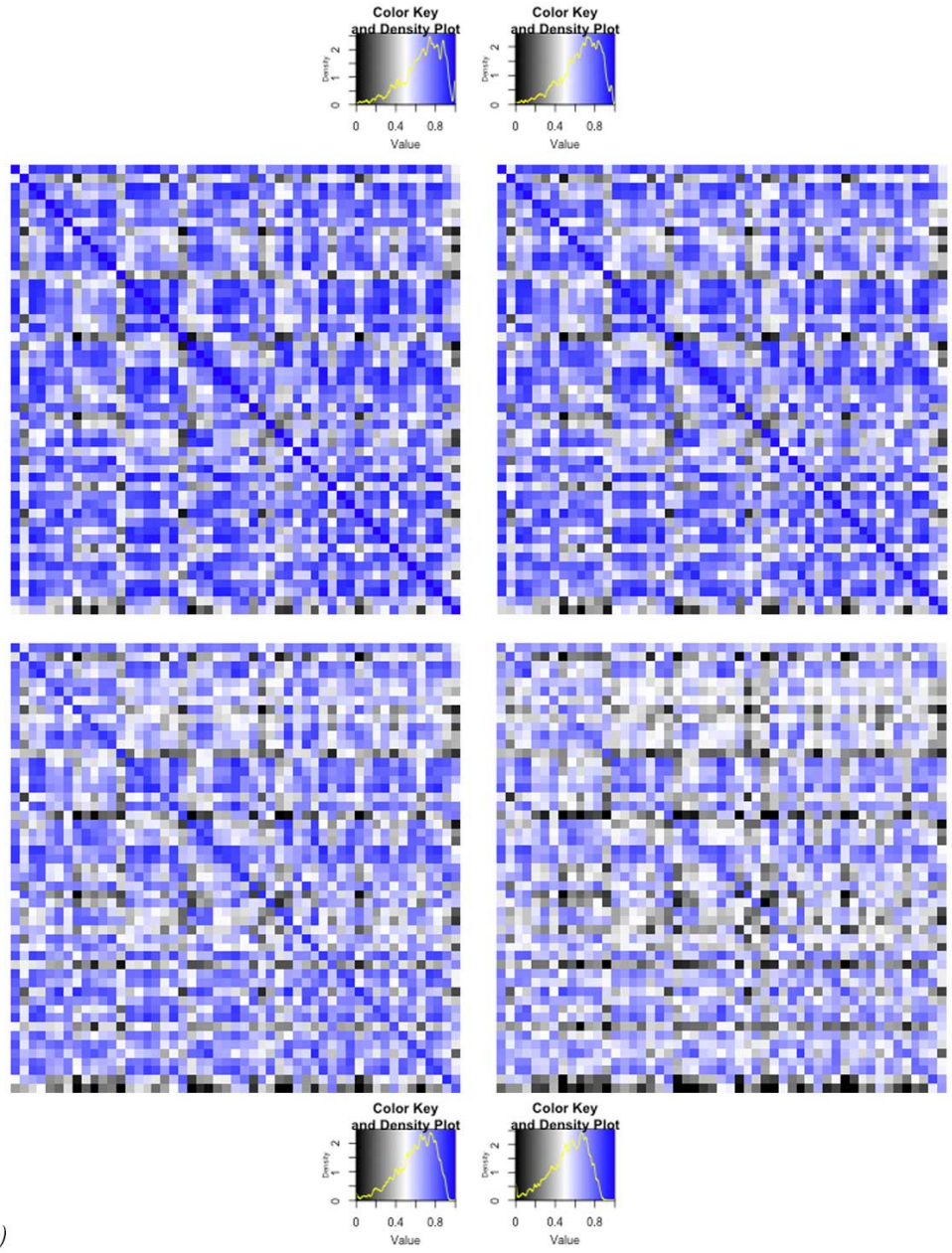
of each tidal-locked period for each pair of the first 40 time series. Since the correlation matrix is symmetric, we stack the entries of its upper triangle row by row, differentiated by background color blocks. We visualize the lengths of tidal-locked periods by colorizing the horizontal stripes with the yellow-red gradient. Note tidally locked periods widely exist for all pairs, though of varying lengths and frequencies.

Figure 5.1 Chi-plots for bivariate dependence screening.



Note. *(a)* Ranked scatter plots and corresponding chi-plots, replicated from Fisher and Switzer (2001); *(b)* sample of randomly generated normal scatter plot; *(c)* Fisher and Switzer's λ_i (Equation **1**) and *(d)* corresponding chi-plot; as well as *(c′)* linear $\lambda_i = H_i$ and *(d′)* corresponding chi-plot.

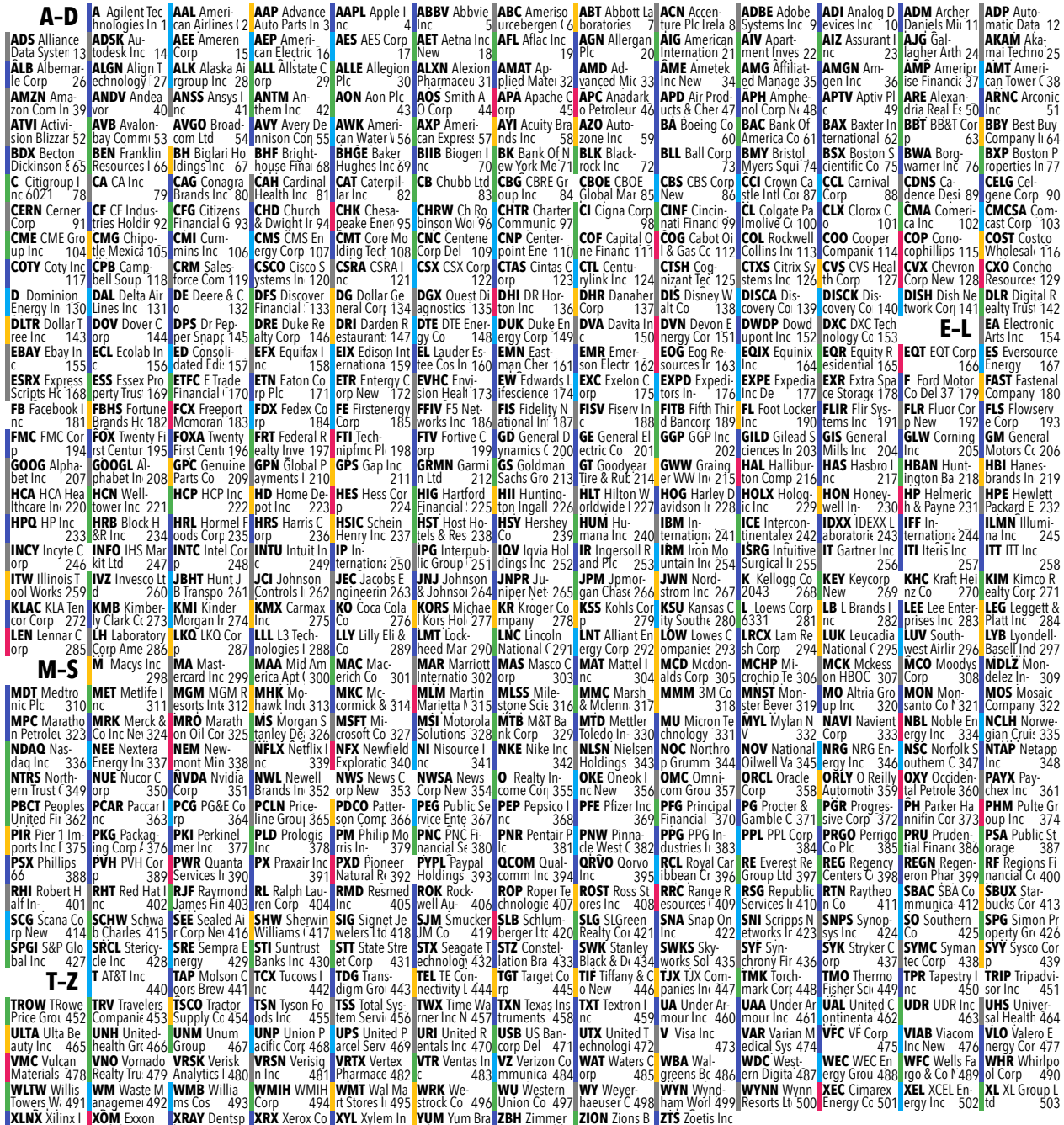
Figure 5.2 Heat map of cross-correlation of different lags.



(a)	*(b)*
(c)	*(d)*

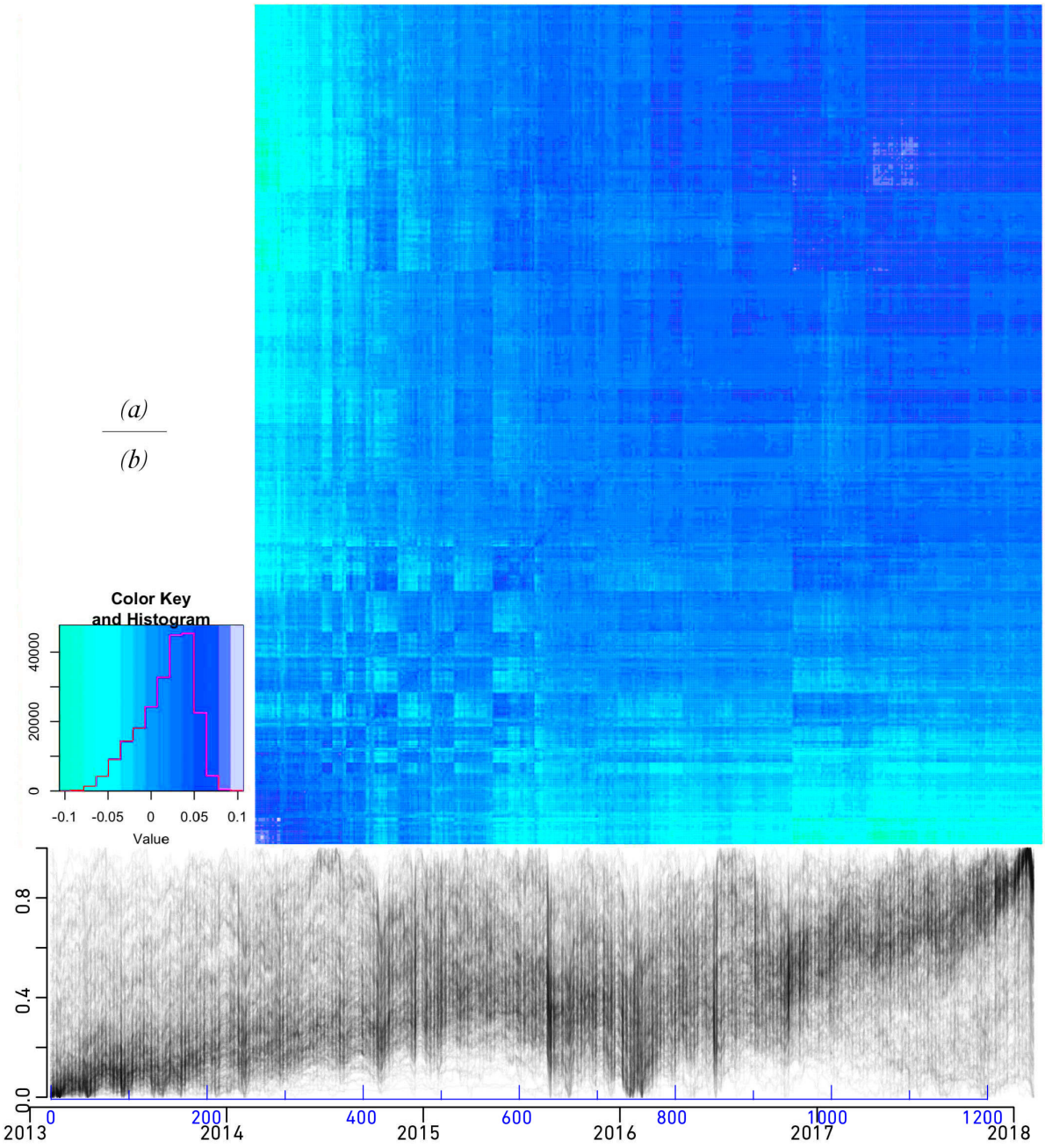
Note. Monthly unemployment rates (seasonally adjusted) of 50 states from January 1976 to May 2018, U.S. Bureau of Labor Statistics, retrieved from Federal Reserve Bank of St. Louis. States sorted by postal codes. *(a)* Lag 0; *(b)* Lag 4; *(c)* Lag 8; and *(d)* Lag 12.

Figure 5.3 List of all stocks tracked.



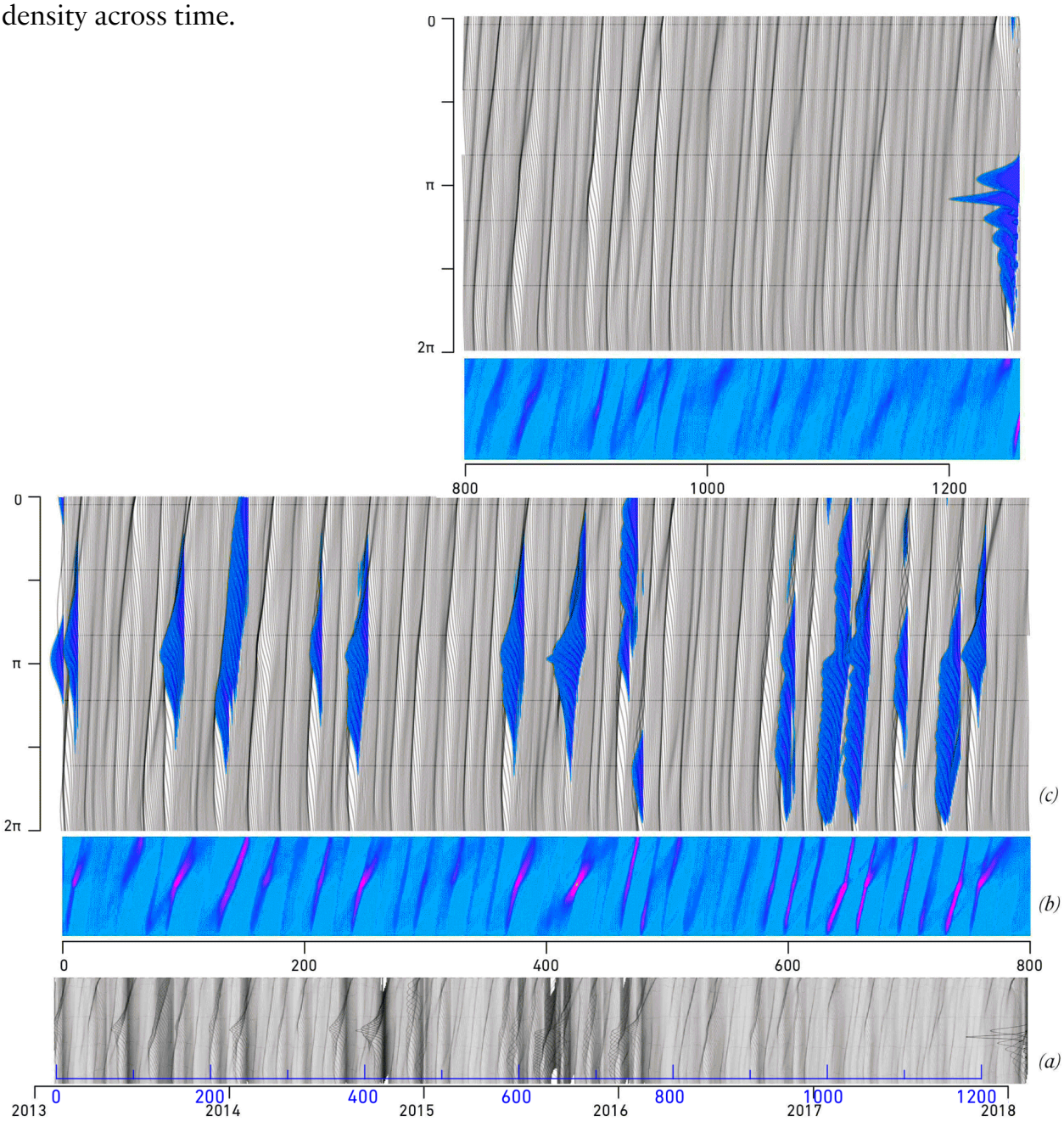
Note. S&P 500 Index stocks from February 8, 2013 to February 7, 2018, CRSP.

Figure 5.4 Cross correlations of S&P 500 stocks.



Note. Daily closing price of S&P 500 Index stocks from February 8, 2013 to February 7, 2018, Daily Stock File, Center for Research in Security Prices (CRSP), The University of Chicago Booth School of Business. *(a)* Heat map of cross correlations; and *(b)* composite plot of closing prices of all stocks tracked, rescale from 0 to 1.

Figure 5.5 Phasal kernel density across time.



Note. *(a)* Empirical phasal kernel density over time (*i.e.*, each vertical curve, an empirical p.d.f. of phases of all stocks tracked); *(b)* corresponding contour plot, lighter color for higher probability density; and *(c)* zoomed hybrid plot, higher density region colored for clarity.

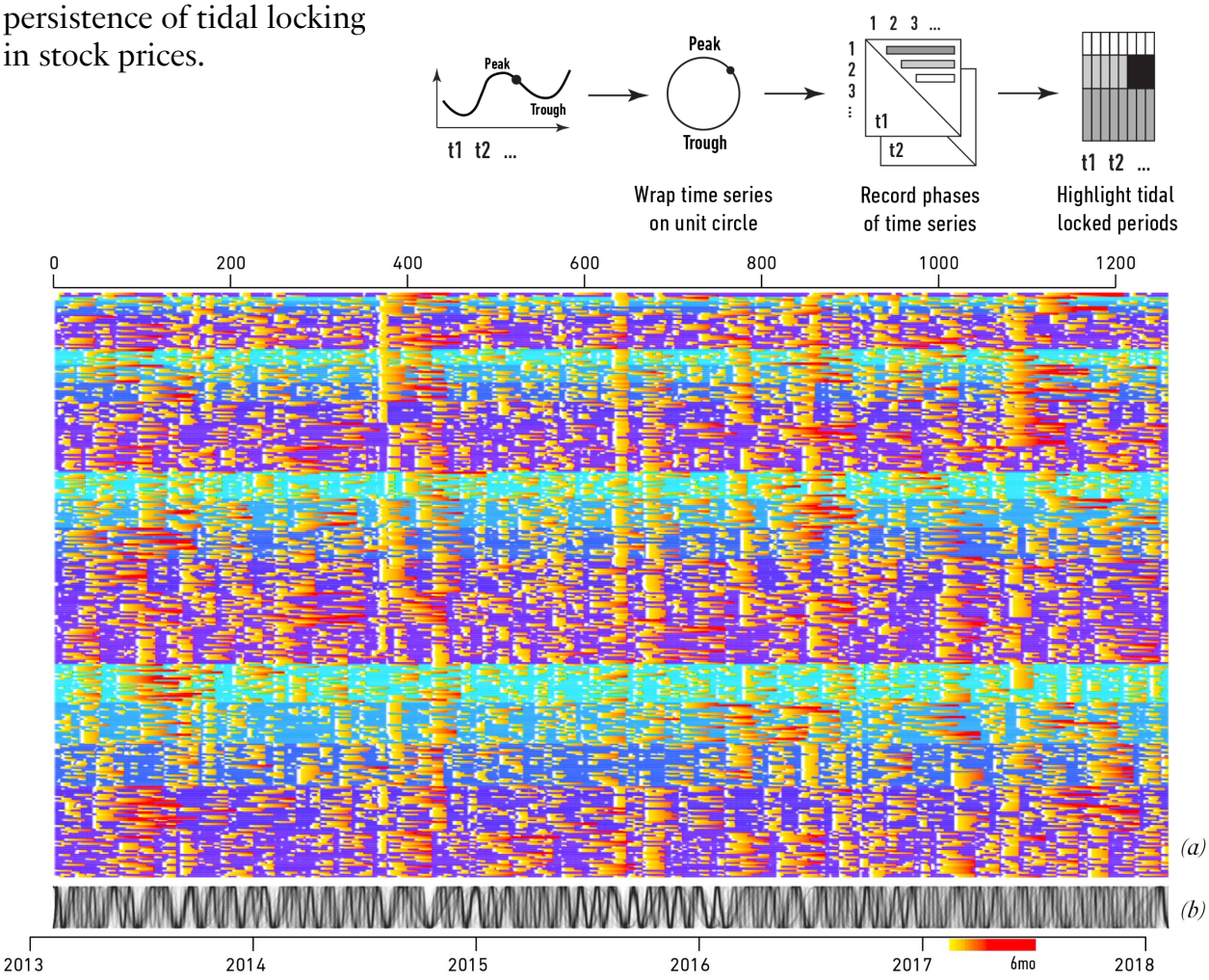


Figure 5.6 Prevalence and persistence of tidal locking in stock prices.

Note. *(a)* Tidally locked periods, all 820 combinatorial pairs of first 40 stocks shown; and *(b)* composite phasal plot of stock prices, cf., composite price level plot in **Figure 5.4***(b)*.

Let $p_t^1,\ p_t^2$ be the phases of two time series. Define their phase difference at t to be $d_t = p_t^1 - p_t^2$. Period t is declared to be a tidal-locked period if $\Delta_t = |d_t - d_{t-1}| < 5°$, where we take into consideration the spherical geometry so that $\Delta_t \in [0, 2\pi)$. Let $I_{t,q} = \{t, t+1, \ldots, t+q-1\}$ be a consecutive segment of tidally locked periods, so that p_t^1 and p_t^2 are not tidally locked at $t-1$ and $t+q$. We call $q = \#I_{t,q}$ the length of the tidal-locked periods. To avoid trivially tidal-locked periods, we are only interested in $I_{t,q}$, that is, periods tidally locked for at least 20 consecutive trading days.

5.2. STAFF PLOT

We first introduce the dynamic staff plot to visualize lagged correlations of high-dimensional time series. The link between the correlation and spherical geometry might not be immediately obvious—the former, after all, is used to measure linear dependence and often fails to detect the presence of nonlinear causal relationships which include, *inter alia*, spherical dependence. Yet there exists a canonical isomorphism between the correlation and the angle of two variables. The staff plot exploits the property. To exhibit this simple isomorphism, consider two univariables of length T, $x = (x_1,\ x_2,\ \ldots,\ x_T)$ and $y = (y_1,\ y_2,\ \ldots,\ y_T)$. Their correlation (more accurately, their lag-0 correlation)

$$r(x,y) = \frac{\sum_{t=1}^{T}(x_t - \overline{x})\sum_{t=1}^{T}(y_t - \overline{y})}{\sqrt{\sum_{t=1}^{T}(x_t - \overline{x})^2}\sqrt{\sum_{t=1}^{T}(y_t - \overline{y})^2}}$$

can be written as

$$r(x,y) = \frac{x^\bullet}{\|x^\bullet\|} \cdot \frac{y^\bullet}{\|y^\bullet\|} = \cos\langle x^\bullet,\ y^\bullet\rangle, \tag{2}$$

where the 0-centered observations $x^\bullet = x - \overline{x}\,1_T = (x_1 - \overline{x},\ x_2 - \overline{x},\ \ldots,\ x_T - \overline{x})$ and $y^\bullet = y - \overline{y}1_T = (y_1 - \overline{y},\ y_2 - \overline{y},\ \ldots,\ y_T - \overline{y})$ are the univariables with their respective means subtracted. Equation (**2**) therefore gives the isomorphism between the correlation of the two univariables $r(x,y)$ and their centered angle $\langle x^\bullet,\ y^\bullet\rangle$, using the principal value of arccosine. If we further define the standardization procedure as

$$x^\circ = x^\bullet/\|x^\bullet\| \quad \text{and} \quad y^\circ = y^\bullet/\|y^\bullet\|,$$

their correlation $r(x,y)$ can be conveniently written as

$$r(x,y) = x^\circ \cdot y^\circ = \cos\langle x^\circ, y^\circ\rangle.$$

Given an n-component time series $x_{1\cdot}$, $x_{2\cdot}$, ..., $x_{n\cdot}$ of equal length, its components after standardization, $x^\circ_{1\cdot}$, $x^\circ_{2\cdot}$, ..., $x^\circ_{n\cdot}$, can be conveniently represented as points on an n-sphere, S^n. Consequently, the pair-wise correlations of these components can be seen under the isomorphism as their angles on the big circles, which for brevity we will call their phasal differences.[5] If we further project all these points onto any one of the big circles formed by connecting two noncollinear points on a hypersphere, x and y, and without loss of generality, move x to the north pole, the correlations of all components with x can be plainly represented by the latitudes of these projected points.[6] This simple geometric fact is the motivation for correlation plots based on spherical projection, *e.g.*, the s-CorrPlots (McKenna et al., 2016). We will briefly restate their construction procedures in the following section but readers can find more detailed explanations in McKenna et al. (2016).

Lagged correlations, as correlations of shifted variables, inherit the same spherical geometry. The lag-ℓ correlation between x and y,

[5] This geometry fact gives the intuitions behind the unconventional notations: the centered observations $x^\bullet$ and $y^\bullet$ are interior points of a disk of some radius and their normalized counterparts x° and x° lie on the boundary of the unit disk, *i.e.*, on a unit circle, S^1.

[6] Strictly speaking, we should use the geographic terms, such as the poles, the equator, and the latitude circles, for a 3-dimensional sphere, S^2, only. However, the terms are clear when they are used to describe the projection on the xy-circle, since one can extend into the additional (and irrelevant) dimension by choosing an arbitrary third univariable, z, provided it's collinear with neither x nor y.

$$r_\ell(x,y) = \frac{\sum_{t=\ell+1}^{T}(x_t - \overline{x})\sum_{t=1}^{T-\ell}(y_t - \overline{y})}{\sqrt{\sum_{t=\ell+1}^{T}(x_t - \overline{x})^2}\sqrt{\sum_{t=\ell}^{T-\ell}(y_t - \overline{y})^2}}.$$

can be written as

$$r_\ell(x,y) = x^\circ_{+\ell} \cdot y^\circ_{-\ell} = \cos\langle x^\circ_{+\ell}, y^\circ_{-\ell}\rangle, \tag{3}$$

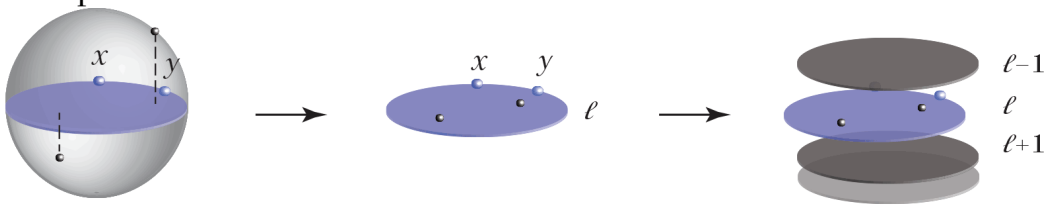
if we use $x_{+\ell}$ to denote the ℓ-truncated x (i.e., a univariable with its first ℓ observations removed) and $y_{-\ell}$ to denote the ℓ-lagged y (i.e., a univariable with its last ℓ observations removed). Both shifted observations $x_{+\ell}$ and $y_{-\ell}$ now have length $T-\ell$.[7]

5.2.1 Construction

The staff plot is straightforward to construct, at least conceptually (**Figure 5.7**)—it is simply the s-CorrPlots (McKenna et al., 2016) of multiple lags stacked together in a 3-dimensional space. The added complication does require a brand new implementation

[7] We have of course overloaded the notations for $\overline{x}$ and $\overline{y}$ for convenience and it's clear from the context that $\overline{x} \equiv \frac{1}{T-\ell}\sum_{t=\ell+1}^{T} = \overline{x}_{+\ell}$ and $\overline{y} \equiv \frac{1}{T-\ell}\sum_{t=\ell+1}^{T} = \overline{y}_{-\ell}$.

Figure 5.7 Construction procedure of staff plots.



and lead to engineering complications with the graphic tools at the moment. However, this is a minor concern here: I will address related computational issues in the technical note at the end. The staff plot can be constructed as follows.

(a) We have just demonstrated in the previous section that standardized components of a multivariate time series (of a given lag ℓ) can be represented as points on a hypersphere. For clarity, we now suppress the subscript for the lag.

(b) For any two noncollinear variables x and y, the Gram–Schmidt procedure gives two orthonormal basis:

$$x^\circ \quad \text{and} \quad y^\circ_\perp = \frac{x^\circ - (x^\circ \cdot y^\circ)y^\circ}{\|x^\circ - (x^\circ \cdot y^\circ)y^\circ\|}.$$

Simple projection of any component z° onto the orthonormal basis x° and $y^\circ_\perp$ gives the coordinates of z° on the unit xy-disk. More specifically, let

$$P_{xy} \equiv [x^\circ,\ y^\circ_\perp]^\mathsf{T}$$

be the $2 \times (T - \ell)$ projection matrix. The desired coordinates of z° (or more strictly, $z^\circ_{-\ell}$, if we wish to express the fixed lag explicitly) on the reference xy-disk are simply $P_{xy}z^\circ$. This therefore projects all points on a hypersphere onto a unit correlation disk.

We name the component x the pivot component: all points produced on the graph here and in the subsequent procedures represent either the lagged autocorrelations of this component or the lagged cross-correlations of this component with the other components. For easy identification, we shall always place the pivot component at the north pole. We call the component y the secondary component.

(c) We draw latitude guidelines at the equator and various significant levels for the correlations with the pivot component according to Equation (**2**): $0 = 0°$; $\pm 0.4 = \pm 53.13°$; $\pm .6 = \pm 66.42°$; $\pm .8 = \pm 78.46°$; and the perfectly correlated points (± 1) are located at the antipodal polar points $(\pm 90°)$.

(d) Repeat the previous steps for all lags of interest and stack these correlation disks indexed by the lag numbers sequentially with lag-0 correlation disk at the top.

(e) Finally we draw a line stringing together points from the same component and for easier identification, color adjacent components differently.

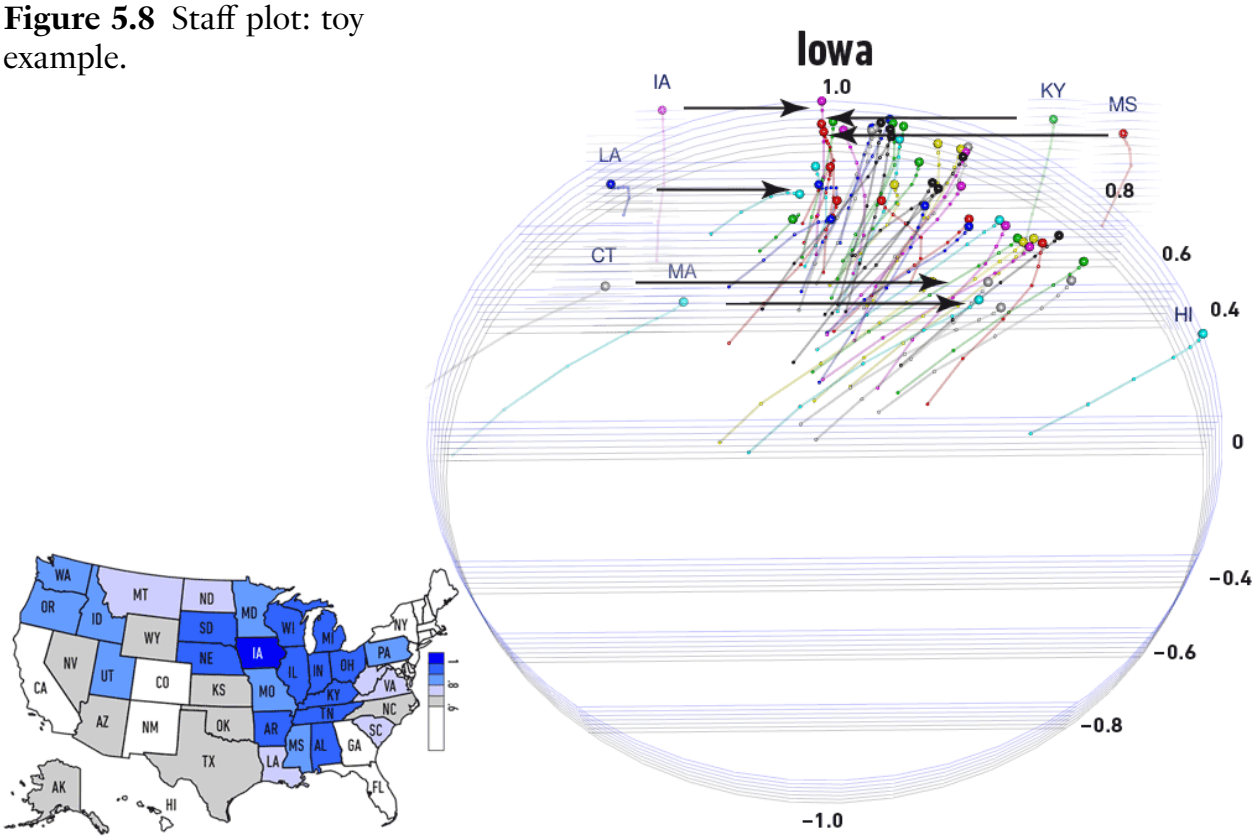
5.2.2 Examples

Figure 5.8 provides a tutorial to the staff plot. Comparing to the heat maps in **Figure 5.2**, the staff plot gives a far more intelligible visualization to lagged correlations of high-dimensional time series. The example plot includes the lagged correlation of Iowa with all 50 states, including itself. To avoid cluttering, only correlations of lags- 0, 1, 3, 6, 12, 18, 24 are displayed. Lagged correlations from the same pairs are connected with line segments, giving the appearance of strands of beads. We have also colored correlation strands differently to distinguish them from one another.

Here is a list of all graphic elements and their statistical interpretations. Readers should bear in mind that the staff plot in our implementation is a fully interactive 3-dimensional model: users have the option to zoom and rotate their viewing perspectives and highlight a specific correlation strand of interest.

(1) *Plot name.* The variable name of the pivot component is prominently displayed at the top of the staff plot. It clearly indicates to users that this staff plot is about the lagged correlations of Iowa. Users can alternatively identify the pivot component by locating the correlation strand with the first node at the north pole, since the lag-0 correlation of any component with itself is by definition 1.

Figure 5.8 Staff plot: toy example.



Note. Monthly unemployment rates (seasonally adjusted) of 50 states from January 1976 to May 2018, U.S. Bureau of Labor Statistics, retrieved from Federal Reserve Bank of St. Louis. For clarity, only correlations of lags 0, 1, 3, 6, 12, 18, 24 with Iowa are plotted. Note the staff plot is a 3-dimensional model: users have option to change their viewing perspectives and highlight a chosen correlation strand. Details see Section 5.2.2.

(2) *Layer.* The staff plot contains several identical layers of guidelines, with each layer consisting of a circle and several horizontal lines. Guideline layers are colored with a subtle blue-to-black color gradient, representing different lags from lag-0 to the last lag. Since correlations for lags 0, 1, 3, 6, 12, 18, 24 are provided, the staff plot displayed contains 7 layers. Each layer is a s-CorrPlot for the correlations of that lag.

(3) *Guideline.* Horizontal guidelines indicate different significant levels for correlations with the pivot component. From the poles to the equator, the staff plot includes guidelines for ±0.8, ±0.6, and ±0.4 correlations, with the northern hemisphere representing positive correlations. These guidelines resemble music staves, the plot's namesake.

(4) *Strand.* Strands of beads of different colors are scattered across the plot. Each strand represents the lagged correlations between a pair of components: at least one of these components is always the pivot component.[8] For comparison, in **Figure 5.19**, it is represent by a column of colored cells. Note again, the strand for self-correlations, *i.e.*, the lagged correlation of the pivot component with itself, always starts at the north pole.

(5) *Bead.* Each bead represents the correlation of a specific lag for a given pair of variables. Readers can tell which variable pair each bead represents from the color and the mouse-over tooltip. Beads on each strand are equally spaced vertically, since they represent different lags and are therefore embedded in different layers, by (2). The first bead of each strand, representing the lag-0 correlation, is larger than the rest.

[8] Mouse hover each strand gives the name of the other component.

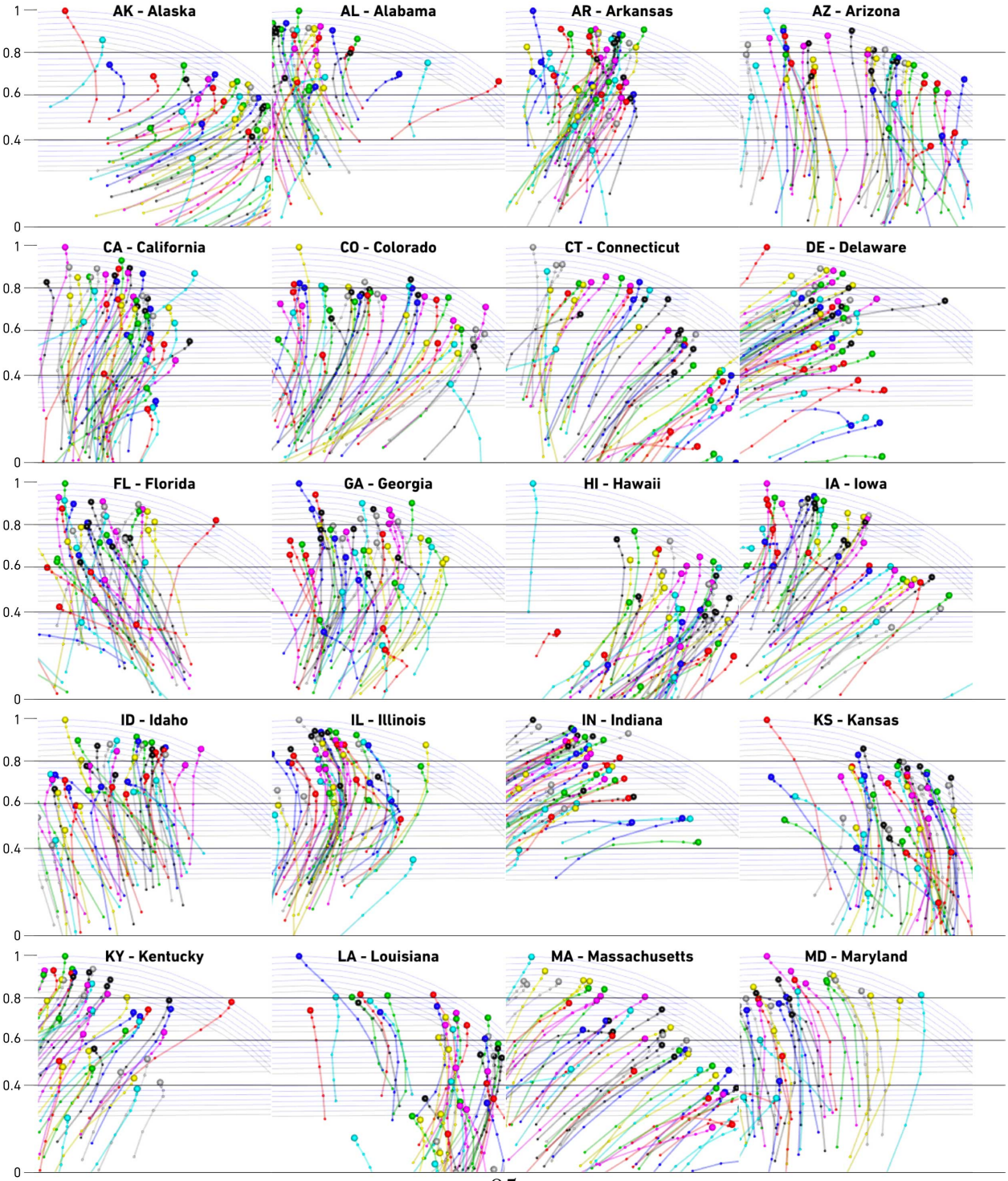
Readers can screen bivariate dependence of a high-dimensional time series by noticing the shape of each strand and the locations of the beads on the strand. **Figure 5.8** in addition contains some examples of possible shapes and their statistical interpretations.

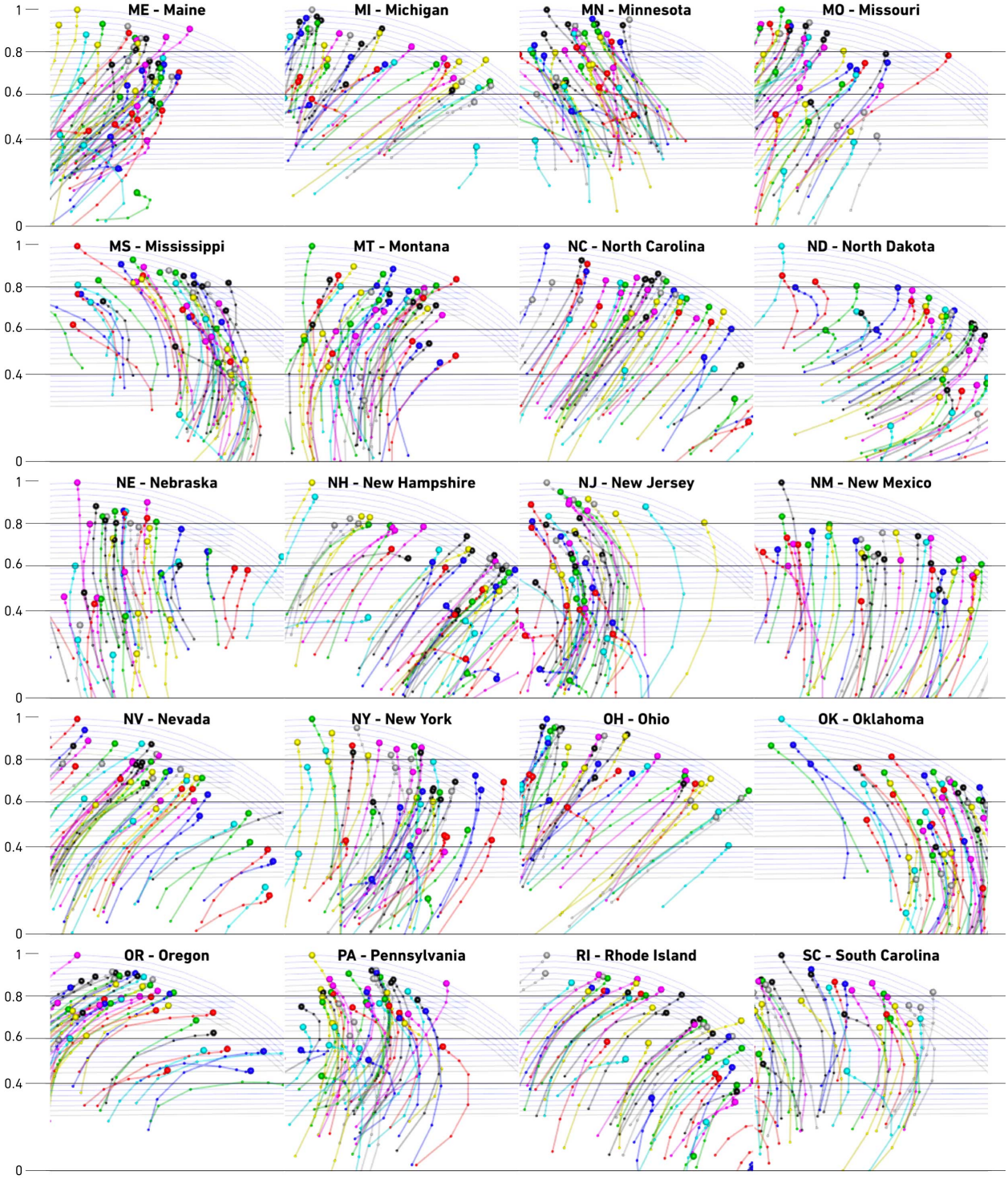
(1) *Strand location.* Strands starting from higher latitudes have higher correlations with the pivot component. The guidelines provide a convenient visual tool to sift out significantly correlated pairs. In the example, we can quickly tell that Kentucky (KY) and Mississippi (MS) have lag-0 correlations with Iowa ($\rho_0 > .9$); Louisiana (LA) has a moderate lag-0 correlation with Iowa ($\rho_0 \approx .7$); and Connecticut (CT), Massachusetts (MA), and Hawaii (HI) all have weak lag-0 correlations with Iowa ($\rho_0 < .4$). Similarly, by interactively exploring the 3-dimensional model, we can read the correlations of later lags off the locations of the smaller beads.

(2) *Strand length.* The length of each strand illustrates how fast the correlations between the pivot component and that variable decay along the lags chosen. We can tell from its long and straight strand that Massachusetts (MA) has not only a relatively weak lag-0 correlation with Iowa but also the fastest correlation decay along the lags: its lag-24 correlation with Iowa is close to 0. The Louisiana (LA) correlations with Iowa, however, have a very different profile; its short and bended strand shows their correlations actually increase along early lags (lags 1, 3, 6) before eventually decaying slowly ($\rho_0 \approx \rho_{24} \approx .7$).

(3) *Outlier.* Outliers quickly emerge from the plot. The strand for Hawaii (HI) correlations with Iowa stands apart from the rest. Not only can we quickly identify the outliers by correlation levels, we can also visually identify variables with irregular correlation decaying patterns by noticing the unusual shapes of the strands. Louisiana (LA) is a such example.

Figure 5.9 Staff plots for correlation screening.

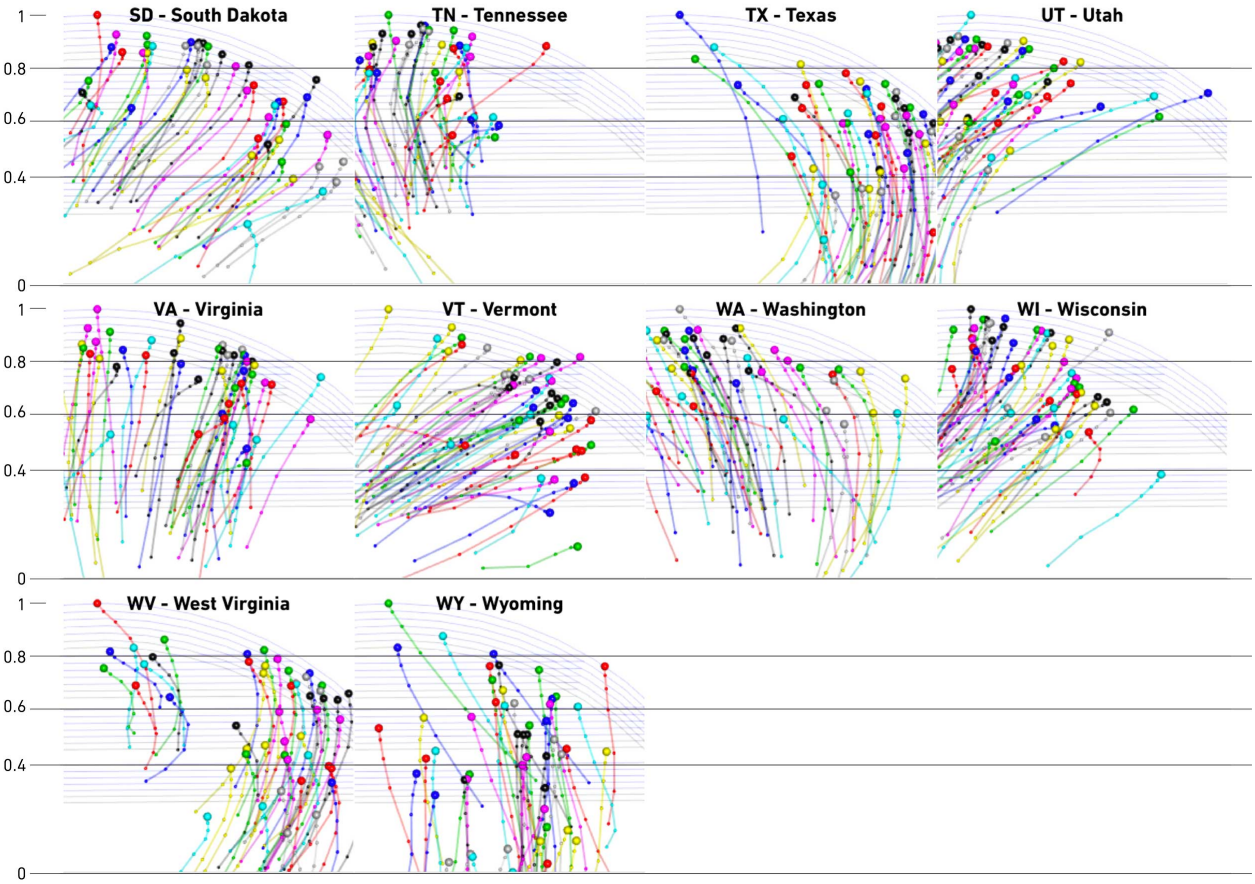




Note. **Figure 5.9** continued.

5.2 Staff Plot

The strength of the staff plot is its compactness. **Figure 5.9** gives the complete guide to the lagged correlation dynamics of the unemployment rates of 50 U.S. states. Readers can follow the tutorial of **Figure 5.8** and glean a wealth of information (about bivariate dependence, correlation decay patterns, as well as clusters and outliers) by simply studying these intuitive plots.

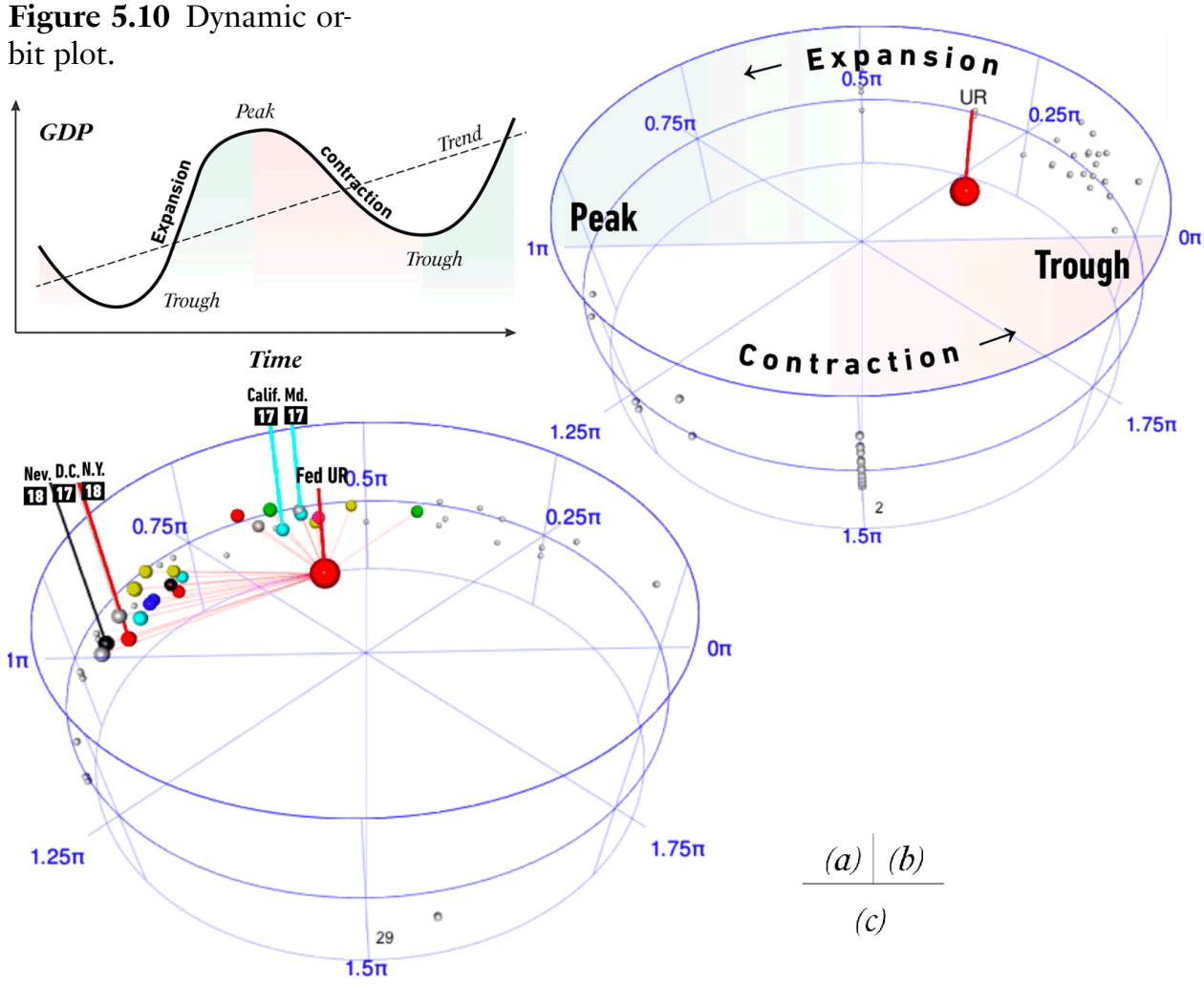


Note. **Figure 5.9** continued.

5.3. ORBIT PLOT

In the previous section on the staff plot, we have demonstrated and made use of the canonical isomorphism between the correlation of a pair of univariables and their phasal differ-

Figure 5.10 Dynamic orbit plot.



Note. *(a)* Business cycles are of pivotal importance in macroeconomics. *(b)* The orbit plot proposed illustrates nuanced dynamic harmonic relations of high-dimensional time series. *(c)* Additional complications are added to facilitate macroeconomic analyses, see Section 5.3.1 for details.

ence. As a consequence, we can further deduce from the fact that lagged correlations can alternatively be faithfully represented by phasal changes of time series. This awareness of changes in phasal differences is of critical importance in macroeconomics, in particular the study of business cycles, as **Figure 5.10***(a)* illustrates, where phasal changes of economic indicators can often lead to impactful changes of economic policies and investment decisions. Yet in spite of their intuitive affinity to harmonic analyses, practitioners nevertheless rely on observing time series line plots, often component by component, to detect peaks and troughs, expansions and contractions, and use these single-dimensional fragments to construct vague mental depictions of these high-dimensional time series, sharpened only by trial and error experiences.

We take the extra step and make explicit the phasal changes of time series in the new orbit plot. Like the companion staff plot just introduced, the orbit plot materializes a simple motivational idea: we want to make apparent the dynamic phasal transitions of a high-dimensional time series. Intuitively, the orbit plot can be viewed as the HHT in the limit, where the time series itself is interpreted as an empirical mode. Rather than introspectively decomposing a time series into dubious intrinsic modes (see theoretical discussions in Chapter 3), we turn outward and depict all such empirical modes, now phenomenologically observed, on a single dynamic plot.

Some words about terminology before we start. Since terms used to describe macroeconomic time series features—such as, expansions and contractions, peaks and troughs,—have become common parlance, we use them without further explanations. Readers can turn to standard textbooks for reference, *e.g.*, Abel et al. (2016) or Mankiw (2015). To avoid confusions, however, we use these macroeconomic descriptors exclusively to characterize

the underlying business cycles. Since indicators such as the unemployment rates are countercyclical, these terms don't describe their time series features.

5.3.1 Construction

The construction of the orbit plot is analogous to the HHT (see Chapter 3). Our goal is to infer the phases of a time series from its levels observed. Since the concept of phases can only be suitably applied to time series with smooth sinusoid-like shapes, before we start the construction procedures, some regularization procedures might be needed to denoise the time series, depending on the scientific content and the measurement method of the time series: in general, time series rapidly transitioning between local extrema on a line plot are candidates for additional regularization procedures. Some simple regularization ideas include: averaging multiple measurements; smooth curve fitting; and changing observation frequencies. Since regularization is not the focus of the paper, we henceforth assume the time series under investigation have been denoised and have well-behaved wave-like forms.

Like the HHT, the construction procedures for the orbit plot are natural but somewhat cumbersome to describe—**Figure 5.10** should give readers a straightforward depiction of the motivational ideas. Take any component $x \equiv x_{i\cdot} \in \mathbf{R}^T = (x_1, x_2, \ldots, x_T)$ of an n-dimensional time series $x_{1\cdot}$, $x_{2\cdot}$, $\ldots, x_{n\cdot}$, each component with length T. Suffice to give the map $h : \mathbf{R}^T \to S^{T-2}$, which in general is not an isomorphism. We start with a naive piecewise linear mapping: this corresponds to the piecewise uniform angular motion in the orbit plot; that is, all components orbit in the same direction (say counterclockwise) with step-function velocities jumping only at phases 0 and π. Using the notations introduced in Section 3.2.1, let $\{x\}_+$ be the set of local maxima of x and $\{x\}_-$, its local minima. Since

we use arccos as the isomorphic map between the two topologies in the staff plot, declare $h : \{x\}_+ \mapsto 0$ and $h : \{x\}_- \mapsto \pi$, that is, points in $\{x\}_+$ are mapped to 0—strictly speaking, $\#\{x\}_+$ copies of 0—under h and $\{x\}_-$, π. We now only need to linearly interpolate the images of x for points other than 0 and π: the procedure is as follows.

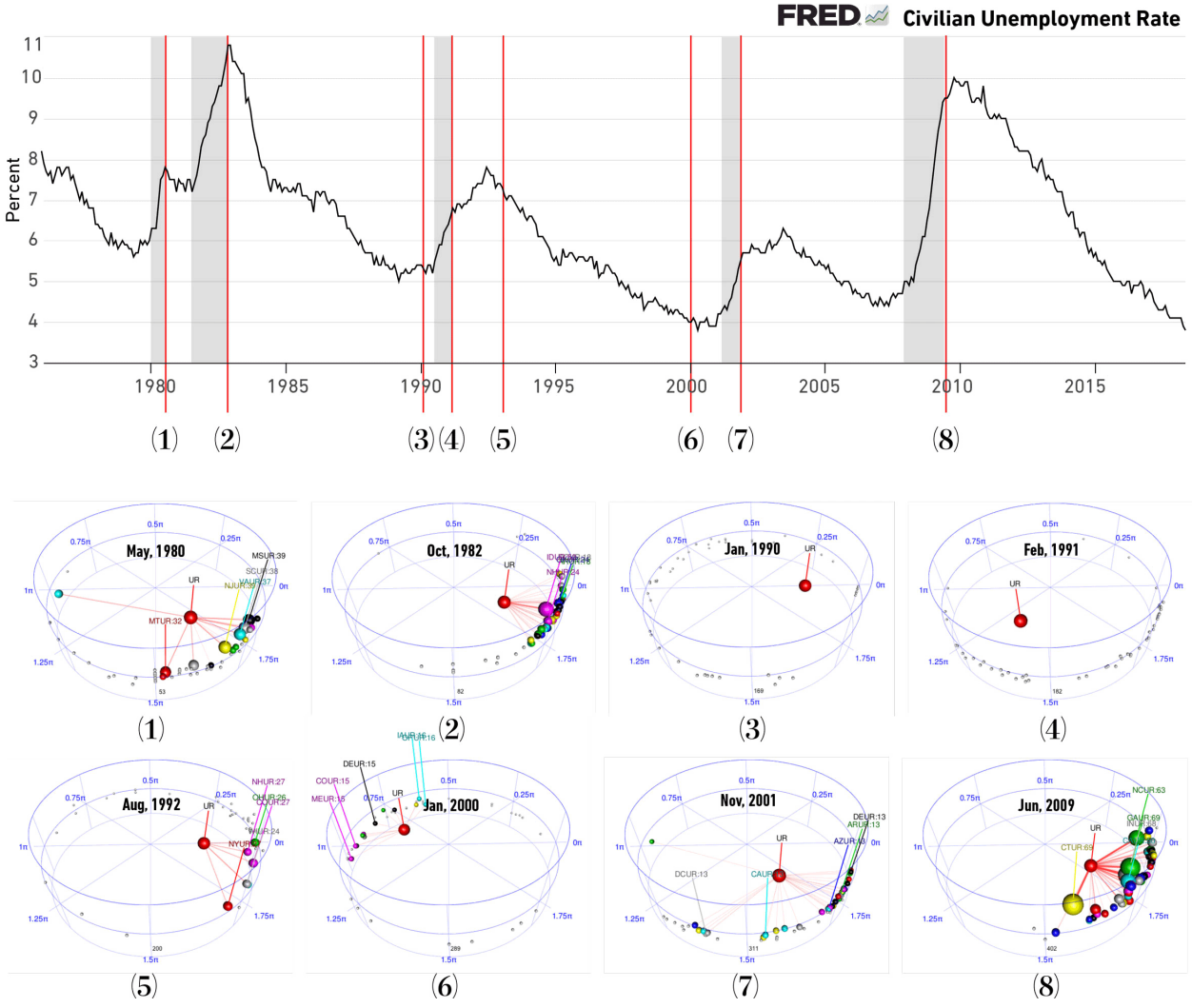
It is clear all other points $\{x_i\}/\{x\}_\pm$ lie either between a local maximum (on its left) and a local minimum (on its right) or *vice versa*. Note the end points of the line plot at 0 and T are by construction extrema: since this is an artificial fact by construction and does not yield meaningful physical interpretation (again like what they are in the HHT), we in practice drop a couple of cycles at the beginning and the end of a time series to effectively give the dynamic orbit plot time to boot up and wind down. Let x_i and x_j, $j > i+1$) be an ordered pair of adjacent local maximum (*i.e.*, $hx_i = 0$) and local minimum ($hx_j = \pi$), we then have $hx_p = \frac{p-i}{j-i}\pi,\ i < p < j$. Similarly, if we instead have $hx_i = \pi$ and $hx_j = 0$, define $hx_p = \pi + \frac{p-i}{j-i}\pi,\ i < p < j$, where we use the spherical topology property $2n\pi = 0$. This completes the map h: readers can verify all points of x have been assigned an image on a circle.

5.3.2 Examples

Figures 5.11 and **5.12** continue with the foregoing example of unemployment rates. The center sphere in red is the federal civilian unemployment rate (UR) and the spheres orbiting it are the 50 state unemployment rates. We draw additional guidelines between the federal sphere and state spheres to help readers visualize how strong the tidal locks between these pairs are—*viz.*, how many consecutive months each pair has been in phase locked positions. Instead of summarizing intricate dynamics with correlations as we do in

the staff plot, the orbit plot provides animated representations of how these time series interact month by month. It provides insights into the following questions that existing graphic tools and the staff plot cannot answer, questions nevertheless of great importance

Figure 5.11 Orbit plots and U.S. business cycles.



Note. U.S. Bureau of Labor Statistics, Civilian Unemployment Rate [UNRATE], retrieved from FRED, Federal Reserve Bank of St. Louis; `fred.stlouisfed.org/series/UNRATE`, June 18, 2018. Shaded areas indicate U.S. recessions. Details see **Figure 5.12**.

to macroeconomic policy makers. Since economic recessions are a well-researched subject, we use established facts about recessions as a benchmark to validate the orbit plot: we will demonstrate how key features of recessions are effectively depicted in the plot before moving on to a more stylized example.

(1) *Expansion and contraction.* All spheres in the upper cylinder represent economies in expansion (decreasing unemployment) and the rest in the lower cylinder are economies in contraction. The plot shows how long and in what sequential orders economies transition between the two states. The 2009 recession stands out from the rest in that all economies are clustered in the fourth quadrant in **Figure 5.12**: it suggests that economies have become more integrated since 1980 and the 2009 recession is more widespread than the rest. This is also reflected in the unusually long periods of tidal-locked rotations: we observe at the end of the recession, economies have been tidally locked for 5 and half years, comparing to 3 years in 1980, 2 years in 1982, 2 years in 1992, and 1 year in 2001.

(2) *Procyclical and anticyclical.* The plot gives a vivid recount of how each economy behaves in each business cycle by noticing its relative phase with regard to other states. Take the 2009 recession for example, we see industrial states like Indiana (`IN`) move into the recession before the overall U.S. economy but states with heavy financial sectors like Connecticut (`CT`) drag in recovery, confirming the fact the 2009 recession is the result of the financial crisis of 2007–2008.

(3) *Classification of recessions.* Classification of a general economic recession is a difficult question since each state behaves differently: The Federal Reserve frequently revises its official definitions of recessional periods (shown in gray area in **Figures 5.11**). There

is a natural definition of recession from the orbit plot: if and only if we see economies are strongly tidal locks in the lower quadrants, the overall U.S. economy is in a recession. There is a one-to-one mapping between the occurrences of these patterns and the recessions officially declared. In fact, **Figure 5.12** shows the official declaration of the ending of a recession precisely coincides with that prescribed by the orbit plot: note they do not always coincide with the peaks of federal unemployment rate. The only exception is the 1991 recession: the Fed declares the recession ends in February 1991 but accordingly to the orbit plot, most tidally locked economies do not begin the recovery process until August 1992, a year and a half later. This might be political compromise: the reunification of Germany in October, 1990, the end of Gulf War in February 1991, and events leading up to the official dissolution of the Soviet Union in December 1991, all make it the perfect timing for recovery.

Figures 5.13 to **5.16** provide an in-depth case study on how to use the orbit plot to examine the dynamic relations of a given component of a high-dimensional time series *in situ*, that is, with respect to other components of the time series. Taking on the role of a stock analyst of, say Apple Inc. (AAPL), we want to understand precisely how AAPL fluctuates with respect to other stocks in the market, similar to the exercise we did for the unemployment rates of 50 states. Standard portfolio theory shows stock returns are correlated with the return of the market portfolio. **Figure 5.13***(b)* documents the intricate dynamic relations. We have introduced the phenomenon of tidal locking and exhibited its prevalence in S&P 500 stocks in Section 5.1.2. The plot overlays the tidally locked segments of all 511 stocks with AAPL: there are periods when a large subset of stocks are persistently tidally locked with AAPL (*e.g.*, around Period 800, in April 2016) and there are periods when almost none does (*e.g.*, around Period 300, in March 2014). Panel *(c)* shows this cannot be

simply explained away with market volatility. We plot the market volatility levels in gray so that the gray region is wider when the market volatility is high; we similarly plot the tidal lock strength in red so that the red region is wider when the time series are weakly locked. We would expect to observe weak tidal locks when the market volatility is high and vice versa but this is seldom the case. Our task is then to explore their precise relations with the orbit plot. Since the orbit plot is dynamic, to facilitate further discussions, let's take snapshot plots (see **Figure 5.14**) from four periods indicated in **Figure 5.13**: (1) Period 139: August 26, 2013; (2) Period 285: March 27, 2014; (3) Period 799: April 11, 2016; and (4) Period 1201: November 10, 2017. These particular periods are chosen because at these four representative time slices, the animated orbit plot exhibits four distinct visual patterns, which cannot be explained through fundamental analyses or event studies.

Since the orbit plot is a 3-dimensional model, without accessing the interactive tools needed for exploring 3-dimensional models, these static snapshots are difficult to read in print. To alleviate the problem, **Figure 5.15** lists out the top tidally locked stocks for each of the snapshots, except for Period (2) when all stocks are weakly tidally locked with AAPL. For comparison, **Figure 5.16** lists out the overall top tidally locked stocks. To further provide the context of these plot, **Figure 5.17** provides the top news feed for AAPL from the Bloomberg Terminal on or shortly before these dates and **Figure 5.18** provides the corresponding top news on the overall economy.

(1) *Period 139: August 26, 2013.* After a period of jittery sideway movements, AAPL climbs to phase-0 but most of its top tidally locked stocks are in the second and the third quadrants, counting clockwise. Apple is testing components for iPhone 5S due to release in a month, a small annual refresh to its predecessor. Ballmer is leaving Microsoft in a

year. U.S. economy continues with the slow recovery from the financial crisis 5 years ago.

(2) *Period 285: March 27, 2014.* Apple sideway movement continues. Apple fights with U.S. over encryption and iPhnone sales lose momentum as the market pushes for low-cost smart phones. U.S. economy expands more than expect. AAPL phasally detaches from other stocks.

(3) *Period 799: April 11, 2016.* Microsoft Office for iPad is released and Apple increases battery orders. Obama presides over steady recovery. AAPL rotates in sync with its top tidally locked stocks.

(4) *Period 1201: November 10, 2017.* Apple iPhone X was released a week earlier: market reception is warm. US economy rebounds under Trump. AAPL becomes tidally locked with a wide array of stocks.

Analogous to the previous example on unemployment rates, analysts can use the orbit plot to perform macroeconomic studies on AAPL. The orbit plot, in essence, offers a first graphical tool to allow analysts to study the business cycles of AAPL *vis-à-vis* those of the market. This consists of various aspects of macroeconomic analysis, as we carried out before in the previous example. For example, a typical stock analysis question asks how much of AAPL's gain during a certain period is due to the overall bullish market. Without a scientific tool to study the phenomenon of tidal locking, the answer is a judgment call, resting on an analyst's ability sensing the market sentiments from various fundamental indicators, technical signals, and words on the street. We can analogously propose the natural criterion for market growth—this plays an important role later in the discussion of the beta of a stock—if and only if we observe a cluster of tidally locked stocks in the fourth

quadrant. This is exactly what we observe in Snapshot (3). By highlighting certain stocks for comparison (what practitioners call "comps"), analysts can further customize the orbit plot and use it to answer comparative questions like how AAPL performs with respect to its comps.

This in effect gives a dynamic representation of the beta of AAPL. In finance, the expected stock return ($\mathbf{E}r$) in excess of the risk-free rate of interest r_f, the "risk premium," is postulated to be

$$\mathbf{E}r - r_f = \beta(\mathbf{E}r_m - r_f),$$

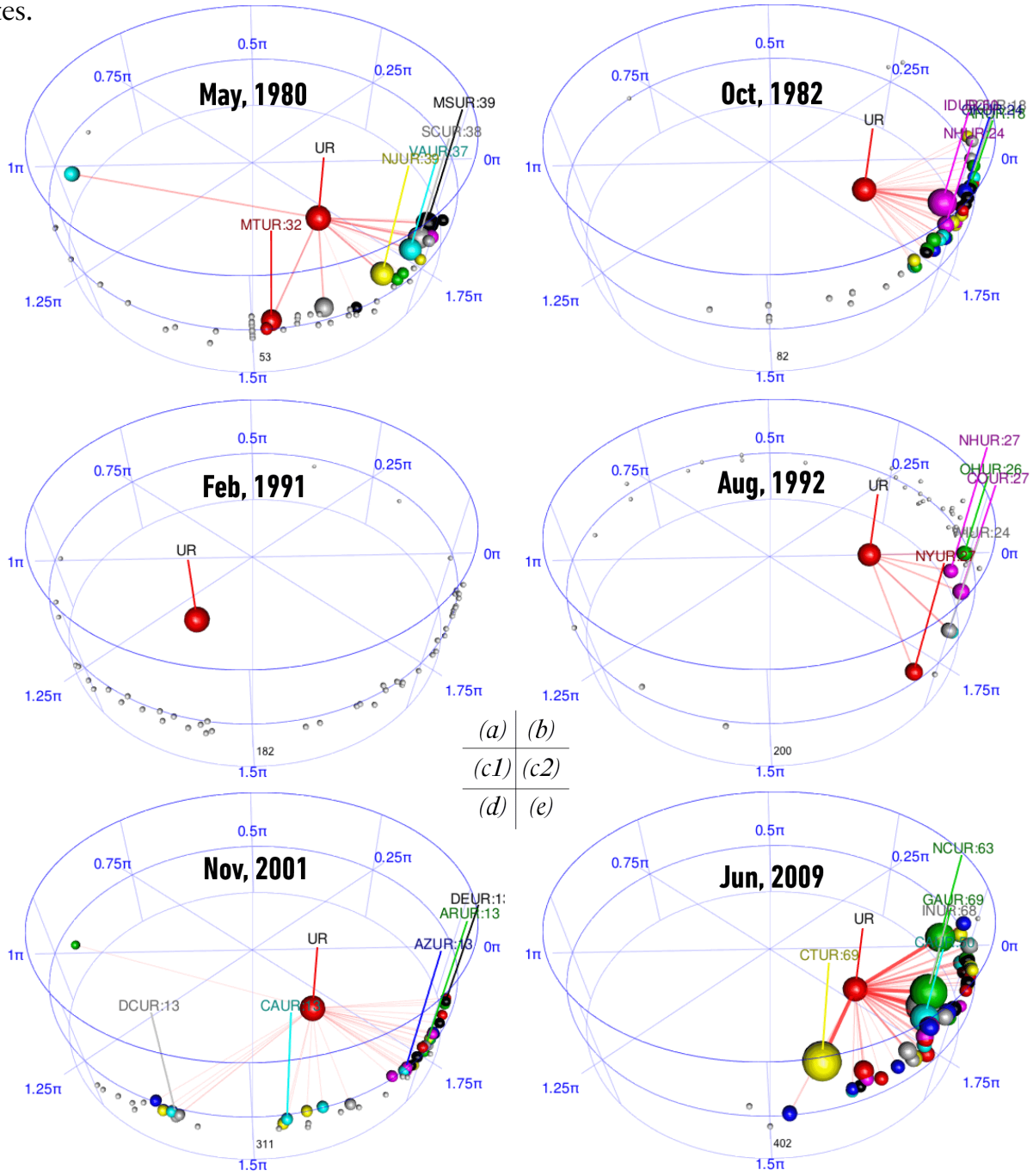
where $\mathbf{E}r_m$ is the expected market return. This is widely celebrated as the capital asset pricing model (CAPM).[9] In the U.S., published betas typically use a stock market index such as the S&P500 as the market portfolio and are reported as objective measurements along with fundamental variables like Open Price, 52-Week Range, Earnings Per Share (EPS), and others.[10] The regression coefficient, β, reflects the bivariate dependence of the stock return and the market return, is often taken as the risk measure and called the reward-to-risk ratio. Comparing to a single number reported, the orbit plot gives animated representations of how β changes overtime. In Snapshot (3), we see AAPL is of the same phase with tightly

[9] You can control for more factors, *e.g.*, market capitalization (SMB, small minus big), book-to-market ratio (HML, high minus low), profitability (RMW, robust minus week), investment (CMA, conservative minus aggressive), and momentum (MOM), in the regression but most reported betas do not. In any case, it's tangent to the point that factor loading are dynamic.

[10] These calculated betas vary widely from venue and venue: For example, the beta for AAPL is reported to be 1.01 (Marketwatch), 1.02 (Nasdaq), 1.097 (Investopedia), 1.14 (Yahoo Finance), 1.15 (Reuters), 1.15 (CNBC), FT (1.1521), and 1.23 (MSN), on July 17, 2018. Since the convention for calculating beta is not clear, we refrain from further discussions on the effects of choosing different frequencies and lengths for the returns as well as the distinction between equity and asset betas.

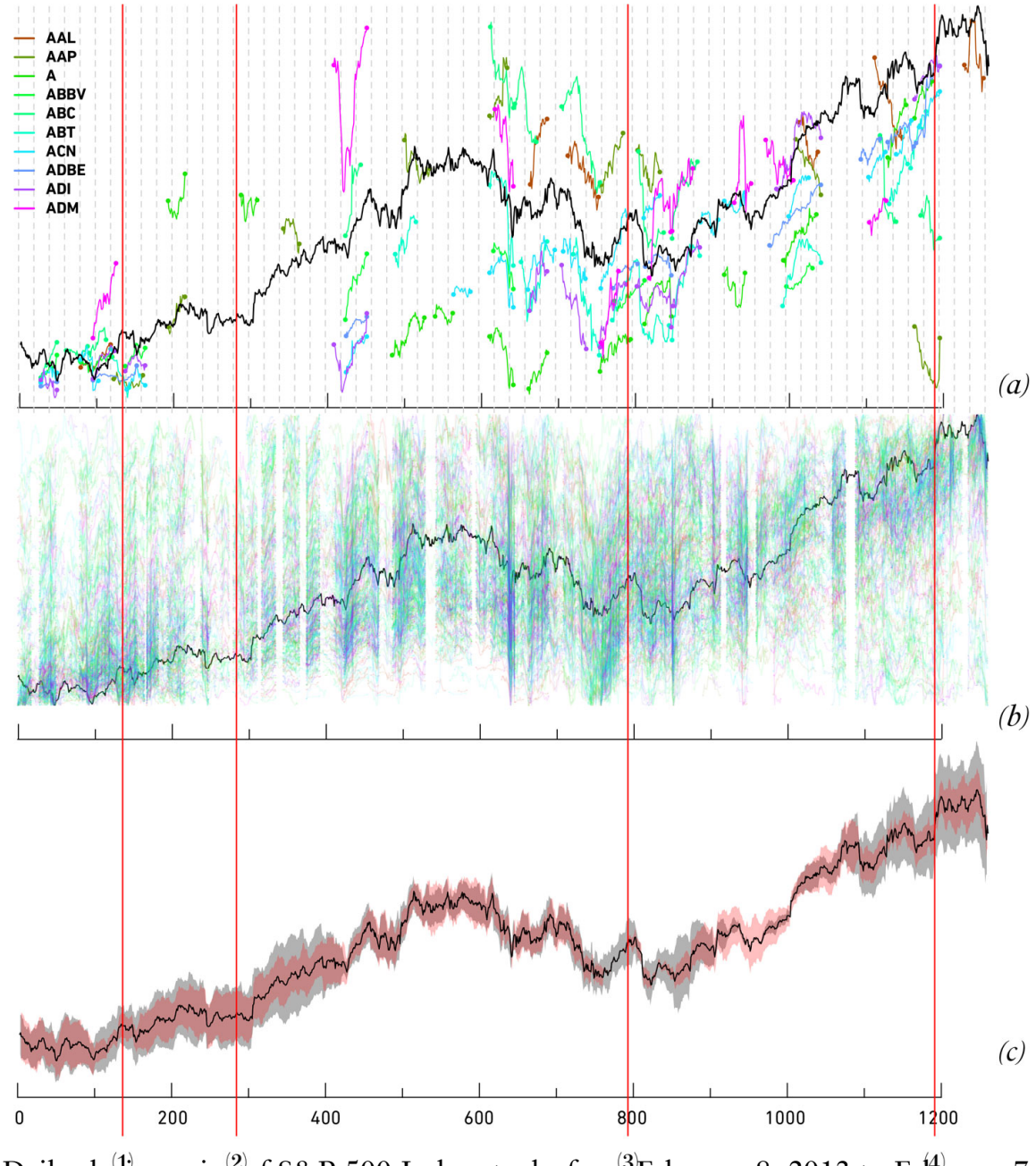
tidally locked stocks: this is evidence for an increased β. Contrary, in Snapshot (1), we see AAPL is out of phase with tidally locked stocks: this is evidence for an decreased β. The orbit plot demonstrates clearly that β of a stock varies widely over time and cautions against a common practice using the beta of a stock as the one-for-all measurement of its relative risk to the market. These already insightful conclusions are drawn from our glancing over four snapshots, observant readers can surely make better use of the dynamic plot on screen, a 1258-frame animation for a 512-dimensional time series, especially after paired with existing stock investment tools.

Figure 5.12 Closeup orbit plots of unemployment rates.



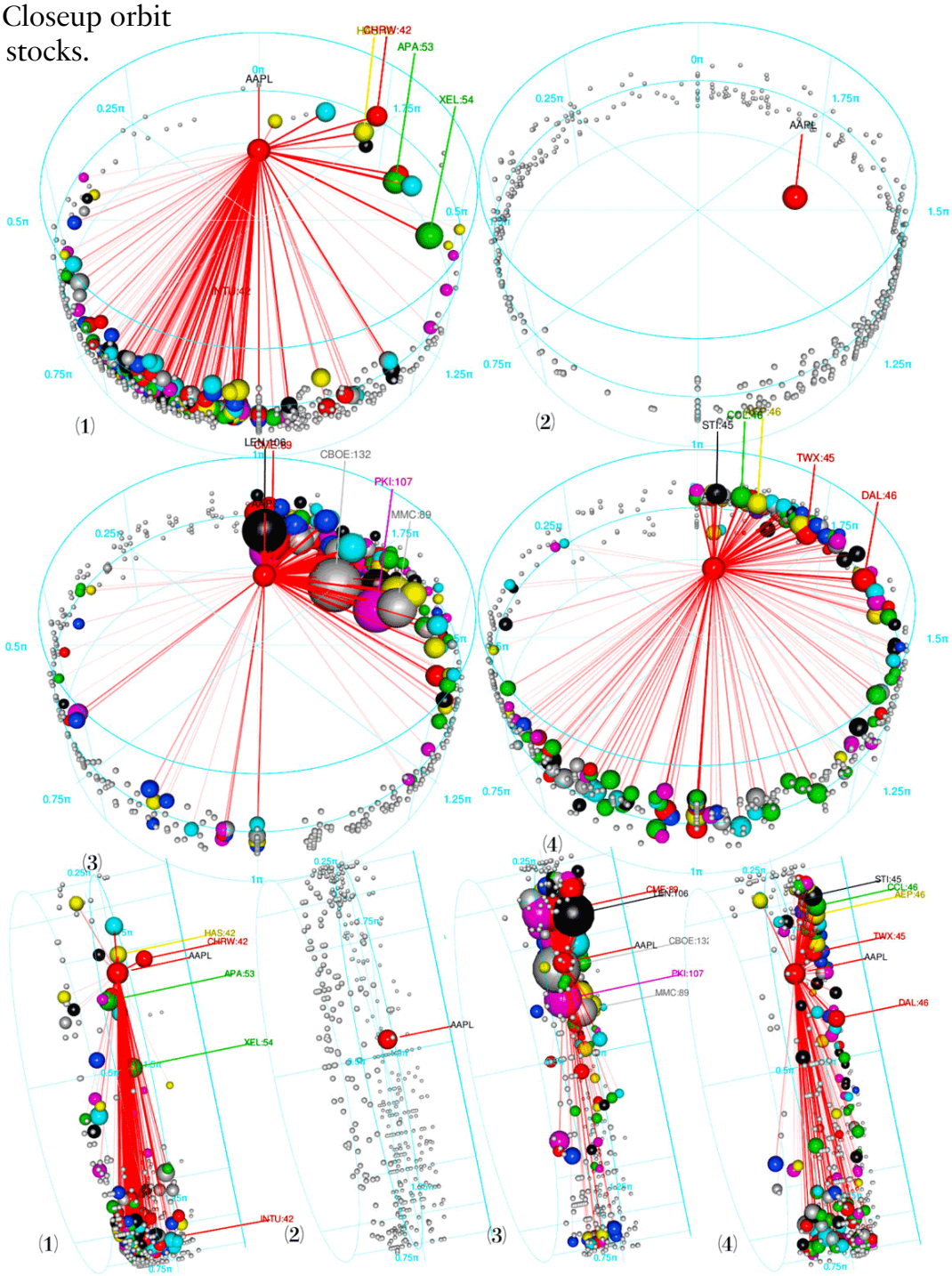
Note. Snapshots of animated orbit plot of monthly unemployment rates of 50 states from January 1976 to May 2018, time stamps indicated in **Figure 5.11**.

Figure 5.13 Dynamic tidal locking of S&P stocks.



Note. Daily closing price of S&P 500 Index stocks from February 8, 2013 to February 7, 2018, CRSP. *(a)* Tidally locked periods of AAPL and first 10 stocks, monthly bins indicated; *(b)* tidally locked periods of AAPL and all other S&P stocks; and *(c)* market volatility and tidal strength (wider gray area for higher volatility, *narrower* red area for stronger tidal influence).

Figure 5.14 Closeup orbit plots of S&P stocks.



Note. Snapshots of animated orbit plot of AAPL and S&P stocks, time stamps indicated in **Figure 5.13**. Side views shown in bottom row. Top 25 tidally locked stocks for each slice profiled in **Figure 5.15**.

Figure 5.15 Lists of top tidally locked stocks for given slices

Slice 139: 8/26/2013

#1 XEL 53
Xcel Energy Inc. is a utility holding company based in Minneapolis, Minnesota, serving more than 3.3 million electric customers and 1.8 million natural gas customers i…

#2 APA 52
Apache Corporation is an American petroleum and natural gas exploration and production company incorporated in Delaware and headquartered in Houston.

#3 CHRW 41
C.H. Robinson is a Fortune 500 provider of multimodal transportation services and third-party logistics. The company offers freight transportation, transport…

#4 HAS 41
Hasbro, Inc. is an American multinational toy and board game company. It is the largest toy maker in the world in terms of stock market value, and third largest with rev…

#5 INTU 41
Intuit Inc. is a business and financial software company that develops and sells financial, accounting, and tax preparation software and related services for small…

#6 LRCX 41
Lam Research Corporation is an American corporation that engages in the design, manufacture, marketing, and service of semiconductor processing equipme…

#7 MHK 41
Mohawk Industries is an American flooring manufacturer based in Calhoun, Georgia, United States. Mohawk produces floor covering products for residential and co…

#8 WMT 41
Walmart Inc. is an American multinational retail corporation that operates a chain of hypermarkets, discount department stores, and grocery stores.

#9 MCHP 32
Microchip Technology is an American manufacturer of microcontroller, memory and analog semiconductors.

#10 ARE 31
Alexandria Real Estate Equities is a major United States real estate investment trust.

#11 MA 31
Mastercard Incorporated is an American multinational financial services corporation headquartered in the Mastercard International Global Headquarter…

#12 ZTS 31
Zoetis, Inc. is the world's largest producer of medicine and vaccinations for pets and livestock.

#13 NUE 30
Nucor Corporation is a producer of steel and related products headquartered in Charlotte, North Carolina. It is the largest steel producer in the United States of Ameri…

#14 UHS 30
Universal Health Services is an American Fortune 500 company based in King of Prussia, Pennsylvania. It is one of the largest hospital management companies i…

#15 EBAY 28
eBay Inc. is a multinational e-commerce corporation based in San Jose, California that facilitates consumer-to-consumer and business-to-consumer sales through…

Slice 799: 4/11/2016

#1 CBOE 132
CBOE Global Markets is an American company that owns the Chicago Board Options Exchange and the stock exchange operator BATS Global Markets

#2 PKI 107
PerkinElmer, Inc., is an American multinational corporation focused in the business areas of human and environmental health.

#3 LEN 106
Lennar Corporation is a home construction and real estate company based in Miami, Florida. In 2017 the company was the largest home construction company in…

#4 CME 89
CME Group Inc. is an American financial market company operating an options and futures exchange. It owns and operates large derivatives and futures ex-…

#5 MMC 89
Marsh & McLennan Companies, Inc. is a global professional services firm, headquartered in New York City with businesses in insurance brokerage, risk manage-…

#6 NFX 87
Newfield Exploration Company is a petroleum, natural gas, and natural gas liquids exploration and production company organized in Delaware and headquartered in…

#7 WBA 71
Walgreens Boots Alliance, Inc. is an American holding company headquartered in Deerfield, Illinois that owns Walgreens, Boots, and a number of pharmaceutic…

#8 LNC 70
Lincoln National Corporation is a Fortune 250 American holding company, which operates multiple insurance and investment management businesses through s…

#9 CSCO 69
Cisco Systems, Inc. is an American multinational technology conglomerate headquartered in San Jose, California, in the center of Silicon Valley, that develops, m…

#10 SNA 69
Snap-on Incorporated is a designer, manufacturer and marketer of high-end tools and equipment for professional use in the transportation industry including the auto…

#11 BRK.B 68
Berkshire Hathaway Inc. is an American multinational conglomerate holding company headquartered in Omaha, Nebraska, United States.

#12 VRSK 67
Verisk Analytics, Inc. is an American data analytics and risk assessment firm based in Jersey City, New Jersey, United States, serving customers worldwide in insur-…

#13 NIKE 59
Nike, Inc. is an American multinational corporation that is engaged in the design, development, manufacturing, and worldwide marketing and sales of footwear, ap…

#14 CTSH 57
Cognizant is a multinational corporation that provides IT services, including digital, technology, consulting, and operations services. It is headquartered in Teaneck, N…

#15 IVZ 57
Invesco Ltd. is an American independent investment management company that is headquartered in Atlanta, Georgia, United States, and has branch offices in 20…

Slice 1201: 11/10/2017

#1 AEP 46
American Electric Power is a major investor-owned electric utility in the United States of America, delivering electricity to more than five million customers in 11…

#2 CCL 46
Carnival Corporation & plc is a United States-based cruise company and the world's largest travel leisure company, with a combined fleet of over 100 vessels across …

#3 DAL 46
Delta Air Lines, Inc., commonly referred to as Delta, is a major American airline, with its headquarters and largest hub at Hartsfield–Jackson Atlanta In-…

#4 STI 45
SunTrust Banks, Inc., is an American bank holding company. The largest subsidiary is SunTrust Bank

#5 TWX 45
Warner Media, LLC, doing business as WarnerMedia, and previously known as Time Warner Inc., is an American multinational mass media and entertainment…

#6 BEN 43
Franklin Resources Inc. is an American holding company that, together with its subsidiaries, is referred to as Franklin Templeton Investments; it is a global inves…

#7 A 32
Agilent Technologies is an American public research, development and manufacturing company established in 1999 as a spin-off from Hewlett-Packard. The resul…

#8 AKAM 32
Akamai Technologies, Inc. is an American content delivery network and cloud service provider headquartered in Cambridge, Massachusetts, in the United…

#9 ALB 32
Albemarle Corporation is a chemical company with corporate headquarters in Charlotte, North Carolina. It is a specialty chemical manufacturing enterprise.

#10 ANTM 32
Anthem, Inc. is an American health insurance company founded in the 1940s, prior to 2014 known as WellPoint, Inc. It is the largest for-profit managed healt…

#11 CAT 32
Caterpillar Inc. is an American Fortune 100 corporation which designs, develops, engineers, manufactures, markets and sells machinery, engines, financial prod…

#12 CEL 32
Celgene Corporation is an American biotechnology company that discovers, develops and commercializes medicines for cancer and inflammatory disorders.

#13 CSCO 32
Cisco Systems, Inc. is an American multinational technology conglomerate headquartered in San Jose, California, in the center of Silicon Valley, that develops, m…

#14 DIS 32
The Walt Disney Company, commonly known as Disney, is an American diversified multinational mass media and entertainment conglomerate, headquartered a…

#15 DOV 32
Dover Corporation is an American conglomerate manufacturer of industrial products. Founded in 1955 in New York City, Dover is now based in Downers Grove, Il…

Standard Industrial Classification (SIC) Code:

- 0100-0999 Agriculture, Forestry and Fishing
- 1000-1499 Mining
- 1500-1799 Construction
- 2000-3999 Manufacturing
- 4000-4999 Transportation, Communications, Electric, Gas and Sanitary service
- 5000-5199 Wholesale Trade
- 5200-5999 Retail Trade
- 6000-6799 Finance, Insurance and Real Estate
- 7000-8999 Services
- 9100-9729 Public Administration
- 1800-1999 not used

Note. Company blurbs retrieved from main entries on Wikipedia on June 19, 2018, with minor editorial changes. Corresponding orbit plots shown in **Figure 5.14**.

Figure 5.16 Lists of top tidally locked stocks.

Note. Company blurbs retrieved from main entries on Wikipedia on June 19, 2018, with minor editorial changes.

Figure 5.17 News feed for AAPL.

Slice 139: 8/26/2013		Slice 285: 3/27/2014		Slice 799: 4/11/2016		Slice 1201: 11/10/2017	
Cheaper iPhone seeks to retain core values	FTI	Apple Fails To Knock Out Ringer-Silencing Pate	BLW	Microsoft unveils Office reboot	FTI	AppleInsider: Review: Fitbit Ionic aims at Apple	BLG
IPhone Trade-In Program May Start Next Monti	BFW	Andy Grove and the line between 'good' and 'ba	FTI	Court Denies Stay to Apple Workers in FLSA Ca	BLW	AppleInsider: This week on AI: Apple AR glasse	BLG
Apple looks to build market share with the cheat	FTI	PC Shipments 60.6m in 1Q, Down 11.5% Y/y,	BFW	Microsoft boss launches Office for iPad	FTI	Fox Business: Why Won't Apple Inc. Talk About	NS1
Apple's iWork for iCloud Service May Be Limite	BFW	The Apple Watch One Year Later: Success or Dud	TST	MORE: Apple to Start Global Sales of New IPhc	BFW	Apple Has Several Big Strengths It Can Leverage	TST
Apple Suppliers Including CRUS, OVTI, AVGO (	BFW	Apple, FBI, Encryption: Does Safety Mean No Se	BN	Microsoft to Offer Office for iPad, Maybe a Bit La	NYT	Apple's Booming Stock Price Could Gain Anothe	TST
Apple Said to Be Readying 'Graphite' IPhone 5	BFW	Apple (AAPL) Stock Climbs as Barron's Sees 40%	TST	IN JAPANESE MEDIA: New Apple IPhone, SoftE	BFW	TheStreet.com: Apple's Booming Stock Price Cc	NS1
Top Tech Analyst: This Week Will Offer an 'Early	PRN	Apple IPhone SE Expands Addressable Market by	BI	Microsoft's Office Apps for IPad Ushers in New	APW	Forbes: Carl Icahn Sold Apple Too Soon & It Cos	FOR
Apple Top Holding Among Hedge Funds, Surp	BFW	IPhone Loses Momentum as Growth Shifts to Low	BI	FTC Approves Final Order in Apple Case Over K	BFW	Forbes: TripAdvisor Brands Hotels With Sexual /	FOR
U.S. TMT Pre-Market: DDD, SSYS New Buy at Ci	BFW	Israel to Levy New Taxes on Google, Facebook in	BN	*FED TRADE CMSN: FTC OKS FINAL ORDER IN (	BFW	Forbes: Apple Loop: Apple Confirms iPhone X F	FOR
Microsoft New CEO Could Tap $77 Billion of Cash	BI	Mississippi Can Resume Google Investigation;	WPT	Microsoft opens the door to a world beyond Win	FTI	MacRumors: Apple's iPhone X vs. Google's Pixe	BLG
Steve Ballmer's Exit May Be Another Canary in Co	BI	[Delayed] Cutting Apple Estimates And Target C	BTG	Microsoft CEO Satya Nadella Unveils Office for IP	BN	AppleInsider: Video: Apple's Clips 2.0 puts you	BLG
Pandora Scraps Mobile Limit on $15 Billion Radic	BI	[Delayed] Smartphone Replacement Cycles Are	BTG	Global Mobile Phones: Analyze Industry Earning	BI	New York Post: Some iPhone X buyers report ar	NYP
Microsoft New CEO May Cede Consumer to Apple	BI	U.S. vs. Apple: New Battles in NY, Massachusett	BLC	BlackBerry, Microsoft-Nokia Key in Handset Earni	BI	CBC: Apple admits some iPhone X models free	CBC
DoCoMo Rises on Speculation Carrier to Release	BN	MAZ Launches Home: Connected TV Platform	PRN	MORE: Microsoft Office for iPad Will Have Free	BFW	Smarter Analyst: 2 Sectors, 2 Top Stocks for 201	BLG
SEC Proposal, CFTC Rules, BB&T, Monte Paschi, A	BN	It's Time the World Learned How to Say H-U-A-\	BBO	Microsoft's Nadella Unveils Office for IPad in Mo	BN	The Top 10 Songs And Albums on the ITunes S	APW
Apple, Google, Bosch, CIT Group: Intellectual Pr	BN	Moscovici Says Google, Facebook, Amazon, Mi	BFW	*MSFT REPORTS DEVICE MANAGEMENT SERVI	BFW	Next Web: Some iPhone X displays plagued by	BLG
Sony Seen Growing Smartphone Profits Amid !	BFW	The Top 10 Songs And Albums on the ITunes S	APW	Microsoft Shows Office Software for Apple's IPa	BFW	Jekyll or Hyde? Does It Even Matter in High-Gr	BFW
Bloomberg Industries Most Read: E-Cigarette Mo	BI	U.S. EQUITY PREVIEW: AAPL, BSX, CLBS, DIS, O	BFW	Apple to Start Global Sales of New IPhone in S	BFW	Rolling Stone: Gift Guide: The Best Smartwatch	NS1
Lenovo Takes Page From Apple in Chasing Sams	BN	Streaming Lifts Music Sales Higher for First Tim	SYH	PREVIEW BLACKBERRY 4Q: Focus on Cash Burı	BFW	MacRumors: MacRumors Giveaway: Win Custo	BLG
Ballmer resignation unlikely to quell unease ove	FTI	U.S. Presses Bid to Force Apple to Unlock iPhon	NYT	*SHARP, JAPAN DISPLAY, LG DISPLAY TO SUPPI	BFW	SlashGear: If the 2018 iPad with Face ID looks t	BLG
Apple Said Tested 64-Bit Chips for iPhone 5S: '	BFW	Op-Ed Contributors: Why Apple's Stand Against	NYT	*APPLE TO START GLOBAL SALES OF NEW IPHO	BFW	IBD: Stocks Down, Apple Still Solid; Will These :	IBD
Silicon Valley's Older Workers Fear Discriminat	BFW	BARRON'S ROUNDUP: Gross Calls for Rate Rise	BFW	BlackBerry Is More Than Just a Handset Vendor, (	BI	San Jose Bus Jrn: Here's a look at 8 Bay Area st	NS1
Global corporate bond issuance at lowest level ir	FTI	Why Apple's Stand Against the F.B.I. Hurts Its O	NYT	Microsoft, Nokia $7.5 Billion Integration Key to M	BI	Senate Plan Better for Business, But Corporate Ta	BI
Rewards await Corporate America if it's canny wi	FTI	Brooklyn Case Takes Front Seat in Apple Encryp	APW	NRG Pursuing Rooftop Solar to Avoid 'White Eler	BN	Apple Clips Selfie Scenes lets users put selfies o	FEX
Cheaper local mobiles beat Apple,Samsung sale	PTI	Justice Dept. Says It Still Wants to Force Apple	WPT	Android Dominates Global Market as IOS Fights f	BI	IBD: OLED Stock Displays Continued Strength, H	IBD
Here's What Steve Ballmer Didn't Get About th	WPT	US Judge Ordered Apple to Help Retrieve Data	APW	Microsoft Buys Nokia, While Apple, Google Go fo	BI	MacRumors: Apple's Extended 2017 Holiday R	BLG
BARRON'S ROUNDUP: Puerto Rico's Debt Woe	BFW	Apple's Fight With U.S. Over Privacy Enters a Nev	BN	New Smartphones, Related Hardware Among Top	BI	AppleInsider: Some iPhone X owners report my	BLG
Paul Krugman: On The Symmetry Between Mic	NYT	Gadgetwise: How to Switch to iPhone From Anc	NYT	Low-End Phones, Wearables on Stage at 2014 M	BI	AppleInsider: Deals: 9.7" iPads for $299, 10.5"	BLG
Mobile Internet Innovation Highlights Macwor	PRN	Apple Resisting a February IPhone Search Order	BN	BlackBerry Leads Connected Cars Google Covets	BN	Forbes: Apple iPhone X: Fix For Cold Weather P	FOR
CEO Ballmer Exits After Failing to Take Microsoft	BN	Apple iPhone SE, iPad Pro and new Watch buyir	FEX	BlackBerry Slump, Microsoft-Nokia Deal Key in Ha	BI	AppleInsider: Apple Watch gets special Veteran	BLG
Ballmer will leave behind an unfinished agenda	FTI	Judge Orders Apple to Assist FBI in Massachus	BFW	59% Sales Drop Puts BlackBerry Lowest of Hands	BI	Gizmodo: iPhone X Doesn't Work Right in the C	BLG
Why Jobs worked as a manager but not in a mov	FTI	Apple, FBI headed for another battle over drug	DPA	Apple EPS Estimates Rise 0.3%, Trail Sony, BlackB	BI	Apple Could Gain Another 12% From Here, The	TST
Nasdaq defends handling of longest outage in it	FTI	US keeps Apple encryption battle alive in drug c	PTI	Handset Vendor Margins in Focus as Average Sell	BI	IBD: Netflix, Apple Supplier Lead 5 Stocks Still B	IBD
Financial, Research Information Not Used In Ap	BLW	US pushes Apple to unlock iPhone used in New '	FTI	Canon to Showcase Print and Scan Solutions at	BUS	Fortune: You're Not Alone: Some Users Report t	FOT
Petrobras Outspends Exxon Researching Next Oi	BN	A real-world solution to the tax repatriation ruck	FTI	Spotify Said to Plan IPO in 3Q: Quartz	BFW	MediaTek Diversifies Sales to IoT Even Amid Weal	BI
At Apple Inc., birth of iPhone 5S, 5C set to heral	FEX	Apple's Fight With U.S. Over Privacy Enters a Nev	BN	Apple Patent Filing May Enable Transparent Te	BFW	AppleInsider: Extreme test shows OLED iPhone	BLG
Microsoft's 9% Enterprise Growth Suggests M&A	BI	FastFT: US fight with Apple over privacy continue	FTI	Height: Senate Judiciary to Resume Patent Refo	HTA	MacRumors: Apple Community Envisions Bette	BLG
Microsoft's Next CEO Needs to Shift Focus as PCs	BI	Apple's Fight With U.S. Over Privacy Enters a Nev	BN	*SHARP TO MAKE LCD PANEL FOR NEW IPHON	BFW	TheStreet.com: Why Apple Could Gain Another	NS1
Ballmer to leave Microsoft within a year	FTI	U.S. Presses Ahead With Appeal in Brooklyn iP	BLW	U.S. PRE-MARKET MOVERS: ADXS BAX C CAMT	BFW	How the 'Warren Buffett of Arabia' Built His Fortu	BN
US proposes shorter e-books injunction for Appl	PTI	Profitable Share Gain Remains HP's PC Segment	BI	*CORRECT:SHARP TO PROVIDE LCD PANEL FOF	BFW	Investopedia: iPhone X Will Help Apple Beat Sa	NS1
Apple Duels With U.S. Over E-Book Price-Fixing F	BN	How Silicon Valley – Not Just Apple – Became	WPT	King and Quercus: what about the windfall	FTI	Fast Company: Could The iPhone X's Most Inan	NS1
Nasdaq halt puts pressure on Greifeld	FTI	IPhone Backdoors Would Pose a Threat, French F	BN	Apple Patent Filing Suggests Sapphire Display	BFW	The Fly: Canalys: iPhone 8 Plus out-ships iPhon	NS1
Microsoft Needs a Tech Visionary to Jump-Start G	BI	Apple iPhone SE, iPad Pro and Watch available i	FEX	BlackBerry Falls; SocGen Cuts, Sees $6/Share S	BFW	[Delayed] Morning Research Summary	OPY
Ballmer Retirement Announcement a Surprise, G	BI	United States Software and Information Techno	ACQ	Device Shipments to Rise 6.9% Worldwide This	BFW	The Fly: Apple acknowledges iPhone X becomir	NS1
DOJ Proposes Shorter E-books Injunction for A	APW	Put Away Your Keyboard: It's Time to Talk to Our	TEL	IN CHINESE MEDIA: 'Negative List' for Brokera	BFW	Apple Finally Fixes Annoying IPhone Autocorrec	TEL
Even Cord Cutters Will Have to Pay the Cable Bill	BN	Comment: The clampdown on tax inversions is o	FTI	Apple Store Workers Seek Stay In FLSA Case Ur	BLW	Will Qualcomm Agree to a Deal With Broadcom	NYT
Apple Nears Golden Cross, Monsanto Forms D	BFW	Tim Bradshaw: Apple iPad Pros New tablets are b	FTI	Apple Buys Hyundai Bonds as Investor Pool Wid	BN	Next Web: Apple's latest acquisition could enat	BLG
Internet launches fightback against state snoope	FTI	FBI Continues to Debate Sharing IPhone Hack	APW	Twitter Preparing New Music Strategy, WSJ Say	BFW	Barron's: Apple Supplier AAC Technologies Jun	NS1
Ballmer's Exit Leaves Microsoft Searching for He	BN	Why BTIG Is Cutting Apple's Earnings Estimates	BLC	Labor Groups Challenge Apple on Chinese Plan	BLW	Fanuc's Rosy 2018 Order Outlook Amid Smartphc	BI
Ballmer's Smartphone Misdial Put Microsoft Behi	BI	Apple Among 10 Cos. That Screen Well for Est.	BFW	Apple to Increase Battery Orders to China's Des	BFW	California Governor Avoids Criticizing U.S. Tech	BLW
Apple Objects to U.S. Revised E-Books Remedy P	BN	Blueshift Research's PayPal Idea Proposal	SFT	*APPLE TO INCREASE BATTERY ORDERS TO CHI	BFW	MacRumors: Apple Launches New 'This Weeker	BLG
Why Steve Jobs worked as a manager but not in	FTI	FBI Bought Tool to Break Into IPhone Used in Ter	BN	Daily Briefing: Apple-Samsung, King Digital, D	BLW	Apple reiterated its commitment to diversity – I	BDR
Mobile Handset Makers: The Eight Most Critical T	BI	FBI's IPhone Hack Doesn't Work on Newer Mo	BFW	Global Mobile Phones: Assess Industry Valuation	BI	Patently Apple: Apple Acquired InVisage with w	BLG
Apple, Samsung Win Appeals Ruling to Keep Fir	BN	FBI Debates Sharing IPhone Hacking Details W	APW	Mobile Handset Industry Valuation Assessment: I	BI	HTC Vive VR Headset Waxes as Smartphones War	BI
*APPLE AND SAMSUNG DON'T HAVE TO REVEA	BFW	Tech Stocks Dominate List of 53 Cos. W/ Tax Ra	BFW	EU, China Deal to Lower Threat of Tariffs on Hu	BFW	Trending: Nutella Fans Flip Over Recipe Tweak;	WPT
Windows Runs on 90% of Notebooks as Ballmer I	BI	Apple (AAPL) Stock Slumps as BTIG Warns of Lor	TST	BlackBerry delays executive's Apple move	FTI	AppleInsider: Apple acknowledges iPhone X be	BLG
High-End, Low-End Unit Gulf a Key Handset Them	BI	Dustin Volz: FBI director Comey says iPhone ha	TWT	Comcast's Power Unveiled Courtesy of Apple R	BBO	Apple and Tesla Make Low-Key Buys; Small Caps	TST
Microsoft Windows 3.6% of Smartphones as Ballı	BI	Some For-the-moment Final Thoughts on Appl	WPT	Microsoft's Mobile Suite May Challenge SAP, VM	BI	Fox Business: Better Buy: Corning Incorporated	NS1
U.S. Files Revised Proposal for Apple E-Books C	BFW	Free Data Streamed for Consumers by T-Mobile at	BI	Russia Govt Switches From Apple to Samsung	BFW	MacRumors: Apple Working on Fix for Bug Cau	BLG
Microsoft's Ballmer Failed to Capitalize on Smart	BI	Oversupply, Weak Pricing Squeeze Memory-Chip	BI	*APPLE LITTLE CHANGED, GIVES UP MOST OF E	BFW	TheStreet.com: Apple and Tesla Make Low-Key I	NS1
Microsoft Total Return Under Ballmer Trails S&I	BFW	Hard to Say When Apple Tax Case Will End: EU	BFW	U.S. Stock Options With Biggest Changes in Imp	BN	TheStreet.com: Apple Has Several Big Strength	NS1
PANDORA STREET WRAP: Analysts Negative on	BFW	Apple Talks Up Services, But It's Still a Device C	BBO	JAMF Software Enables Apple's New Enterpris	MWR	Wells Fargo Clearing Services Adds Baker Hughe	BN
Google Buys Wearable Technology Patents From	BN	White House Won't Support Encryption Bill; FE	WPT	Makor - TECH VIEW AAPL US (545 last) - close s	MKR	San Jose Bus Jrn: Apple buys Newark startup th	NS1
P: Beat Our Forecast On Strong Results, Guidan	ABF	CORRECT: Apple Ests., PT Cut at BTIG on Hands	BFW	Gamco's Ward Says Biggest Problem for Apple Is	BN	SiliconANGLE: Apple has quietly acquired quai	WE1
U.S. TMT Pre-Market: P 3Q View Misses; Jefferi	BFW	Why One Analyst Thinks Shares of Tech Giant Ap	TST	U.S. PRE-MARKET MOVERS: APP BODY CXDC F	BFW	We now know who makes the $14,000 chairs c	BDR

Note. Retrieved from Bloomberg Terminal for AAPL top news on or shortly before stated dates, with minor editorial changes. Bloomberg news sources in bold.

Figure 5.18 News feed for U.S. Economy.

Slice 139: 8/26/2013		Slice 285: 3/27/2014		Slice 799: 4/11/2016		Slice 1201: 11/10/2017	
China Profits Jump as Yi Sees Limited Effect From	BN	Asia Rates/Credit Week Ahead: RBA, RBI Rate D	BFW	U.S. President Barack Obama is happy with th	SWM	Pacific Nations Get Framework to Salvage Trade I	BN
Asian Stocks Fall as Kerry Says U.S. to Hold Syria	BN	Fed's Evans Says He Would Wait Until Early 20	BFW	Is the U.S. Strong Enough for Two Rate Hikes?	BLC	APEC Ministers Vow to Fight Protectionism as Tru	BN
FX DAYBOOK EUROPE: German Ifo Survey, U.S	BFW	Don't raise rates and keep inflation party going	SYH	Missouri Removing Box on Convictions From J	APW	Top Predictions for U.S. Home Prices In 2018	TST
Bitcoin Meeting, OTC Derivatives, Trump Univers	BN	*FED'S EVANS SAYS U.S. 1Q GROWTH PROBABI	BFW	What a Stronger Yen Means for Investors	BLC	Trump Tours Asia; Brexit Debate; Art Sale: Week	BN
Insider Trading; Galleon's Collapse: Top Busines	BN	Evans Sees Fed Raising Rates in Second Half of 2	BN	Econs - US-The forgotten Unemployment Cohor	MIF	[Delayed] Large Cap Banks: Fed Weekly: C&I A	JPM
N.J. Jail Is Home for Husband as Lifetime of Alin	BN	Global markets paused amidst better than expe	KCF	US economic data docket is very light today (1	UOB	US/EU/UK economy: Subsidence	EIU
Oil Gains With Yen on Syria as Asian Stocks, Treas	BN	*EVANS: TIMING OF 1ST RATE INCREASE NOT A	BFW	ECB Counts 500-Euro Cost Even as Death of Cash	BN	Predictions for 2018 U.S. Home Prices	TST
NORDIC DAYBOOK: Fed's Williams, ECB's Coeu	BFW	Evans Sees Fed Rates Near Zero 'Well Into' Next '	BN	IN FOCUS: BOJ's Harada, India Inflation; Stock	BFW	Bloomberg Markets: NYU's Lustbader on Crispr a	BN
U.S. Natural Gas Declines After Gaining Third Tim	BN	*EVANS SAYS HE EXPECTS 1.25% FED FUNDS F	BFW	Fed's Global Focus Keeps U.S. 10-Year Yields Nea	BN	Bloomberg Markets: Dave Wilson's Stock of the I	BN
China a `Cloud' Over Emerging Markets, Condor	BN	*FED'S EVANS SAYS HE'D WAIT UNTIL EARLY 20	BFW	SCOTIABANK ECONOMICS: Closing Points (Apri	SCO	President Trump on China Trade Deficit; Intervi	APW
*CHINA CAN REACH 7.5% GROWTH TARGET, K	BFW	*EVANS SAYS FED WILL DO WHAT IT TAKES FOR	BFW	Missouri Removing Box on Convictions From J	APW	Bloomberg Markets: Cordaro on Market Valuatic	BN
U.S. Must Consider Spillover Effect of QE Exit, (	BFW	Evans Sees Fed Raising Interest Rates in Second	BN	DOE Coal Prices by Region for the Week Ended A	BN	Nobel Laureate Phelps on Inflation and Fed Poli	BLC
*CHINA'S MONETARY POLICY TO REMAIN PRU	BFW	*FED'S EVANS SAYS INCREASING QE PROBABL	BFW	U.S. Gasoline Prices for the Week of April 11: Su	BN	Business Insider: The Fed could be tightening r	BLG
*CHINA TO MAINTAIN PROACTIVE FISCAL POLI	BFW	*FED WOULD HAVE MADE INTEREST RATES NE	BFW	Obama, Fed Chair Yellen Discuss Outlook for E	APW	U.S. Treasury Statement and Cash Balance for Nc	BN
130827 Technology Sector: Slightly improved s	OIR	For the US data docket today (28 March), the U	UOB	Former Yellen Adviser Unveils Plan for Fed Refo	BLC	Bloomberg Markets: Bond Report, Eco Brief for M	BN
*US 'MUST' CONSIDER SPILLOVER EFFECTS OF	BFW	Citigroup 'Stress Test' Said to Send Corbat Scra	BFW	Missouri Removing Box on Convictions From J	APW	Forbes: Simple, Bilateral Policy In A Complicate	FOR
*BROOKFIELD AUSTRALIA CEO POWELL SPEAK	BFW	Fed Retreat From Mortgages Nears Tipping Poi	BLW	Fed's Kaplan on U.S. Economy, Policy Outlook (A	BN	Dollar Drifts Lower in Week Focused on Tax Chan	BN
Loans Deflecting Bond Rout Lure Western Asset:	BN	Pending Sales of Existing Homes In U.S. Declir	BLW	The Fed's Balancing Act With Inflation	BLC	Federal Retirement Plan Rose to $511.7 Billion i	BN
*BRICS COUNTRIES WILL DISCUSS FX RESERVE	BFW	Jobless Claims in U.S. Unexpectedly Decreasec	BLW	U.S. Foreign Exchange Rates for the Week Endec	BN	Mnuchin: 'Minor Differences' Between House	BFW
G-20 to Discuss Possible Impact From QE Exit,	BFW	Economy in U.S. Expands More Than Previousl	BLW	Fed's Kaplan Says Weak Data Show There's No N	BN	Central Bank Watch: Countries, Rates, Changes (	BN
America Most Resilient Five Years From Worst GI	BN	China Faces `Mini' Debt Crisis, Rabobank's Every	BN	Which Party Will Benefit From the Weak Obam	WPT	Fed's Bullard Sceptical Low Unemployment Wi	BFW
Dealbook: Five Years After TARP, Misgivings on I	NYT	FED'S EVANS SPEAKS ON U.S. ECONOMIC POLIC	BN	Obama Is `Pleased' With Yellen Amid Signs of SI	BN	U.S. INDUSTRIAL AGENDA: GE Investor Day, Go	BFW
Won Best in Asia as Traders Flee Rupee-to-Real R	BN	*FED'S EVANS SEES ZERO INTEREST RATE `WEL	BFW	Stock Rally Stalls as Earnings Season Kicks Off	BLC	Consumer Sentiment Slips in U.S., Yet Tax-Cut Ho	BI
Falling aircraft demand hits US durable goods or	FTI	Fed's Evans Sees Interest Rates Near Zero 'Well	BFW	White House Readout of Obama Meeting with	BFW	Gallup Poll: Trump Approval Rating 37%; Disapp	BN
Mortgage Plunge on Fed Taper Limiting Economi	BN	Closing Bell: TSX, Wall Street fall amid mixed i	NPW	*OBAMA, YELLEN DISCUSSED NEAR, LONG-TEF	BFW	Fed's Bullard: U.S. Growth Outlook Brighter, In	BFW
Pimco's Hodge Says Demand for Bonds to Recov	BN	Fed's Evans Set to Speak on U.S. Economy, Poli	BFW	Lew: I Hope We've Stopped Inversions Pipelin	BFW	Saudi Shakeup Drives Oil Gains, Tax Reform Hope	BI
ASIA RATES/CREDIT DAYBOOK: China Profits, P	BFW	Fed's stress tests set bar high for Europe	FTI	U.S. Raw Steel Production for the Week Ending A	BN	BI Economics: Other Regional Week Ahead Sumr	BI
Commodities Daybook: Corn, Soy Surge Most in	BN	Fed under fire on 'opaque' stress tests	FTI	*LEW: WE'LL HAVE EXCITING ANNOUNCEMEN	BFW	Fed's Bullard Says Gov. Powell Is Effective Cons	BFW
Will Obama Make the Fed Even Worse With Su	BBO	Initial Jobless Claims Fell Unexpectedly In Late	BLW	USDA Crop Progress by State for the Week of Apr	BN	*24 COUNTERPARTIES TAKE $39.7B AT FED'S F	BFW
Hungary Set to Slow Rate-Cut Pace After Fed Sig	BN	Fed's Evans Sees Rates Rising in 2H 2015, Wo	BFW	*LEW: PEOPLE SHOULDN'T BE SURPRISED WE	BFW	*FED'S BULLARD SAYS FED POLICY RATE ABOU	BFW
The Strong Case for Optimism About U.S. Grow	BBO	*EVANS SEES U.S. GROWTH RUN RATE ABOUT	BFW	USDA Crop Conditions by State for the Week of A	BN	Canada Rig Count Rose by 11 to 203 Week Endi	BN
All About the Velocity of Corporate Income Grov	SBY	*FED'S EVANS SEES RATES RISING IN 2H 2015	BFW	U.S. Crop Progress for April 10: Statistical Summ	BN	U.S. Rig Count Rose by 9 to 907 Week Ending N	BN
Closing Bell: TSX little changed, durable good	NPW	Paul Krugman: America's Taxation Tradition	NYT	Which Party Will Benefit From the Weak Obam	WPT	Senate Plan Better for Business, But Corporate Ta	BI
Lew Tells Congress Treasury to Hit Debt Limit in M	BN	U.S. Economy Grew 2.6 Percent in Fourth Quar	WPT	U.S. Treasury Statement and Cash Balance for Ap	BN	*NY FED BEGINS DAILY OVERNIGHT REVERSE R	BFW
Mexican Peso Plunges Most Among Major Curre	BN	*CORRECT:TARULLO COMPLACENCY OVER RISI	BFW	Fed's Global Focus Keeps U.S. 10-Year Yields Nea	BN	Economics: Consumer Sentiment Slips, Tax-Cut	BBF
U of Chicago's Kashyap Discusses Key Jackson H	BN	*TARULLO: MORE WORK NEEDED TO ADDRESS	BFW	*LEW: NO ONE ON HORIZON TO TAKE OVER U.	BFW	U.S. Repo Close: Old 3-Year Note at Lowest Rate,	BN
SCOTIABANK ECONOMICS: CLOSING POINTS, A	SCO	Tarullo Defends Fed's Move to Supervise Foreign	BN	What Corporate Credit Says About Health of Cor	BLC	Bloomberg Intelligence FICC Weekly Strategy Bri	BI
Breakeven Inflation Rate: Five-to-Ten-Year Forwa	BN	*TARULLO DEFENDS FED'S MOVE TO SUPERVIS	BFW	Bloomberg Economic Evaluation of States (Table	BN	St. Louis Fed's GDP Model Sees Q4 U.S. GDP at 2	BN
N.J. Jail Is Home for Husband as Lifetime of Alin	BN	Fed of 1970s Shows Capacity Clues May Mislead	BN	Central Bank Watch: Countries, Rates, Changes (	BN	St. Louis Fed Real GDP Nowcast Model Sees U.	BFW
The Insiders: Bad News Is Good News for the O	WPT	Best Start Since '09 Defies Forecast of Annual Lo	BN	White House Briefing: Obama-Yellen Meeting O	BN	Global Inflation Watch: World Inflation at 3.7% (	BN
U.S. Gasoline Prices for the Week of Aug. 26: Su	BN	Clarification: Federal Reserve stress tests	FTI	Top Forecasters of the U.S. Economy Q1 2016: R	BN	Global Inflation Watch: Economies Sorted by Inf	BN
Durable-Goods Drop Imperils Outlook for U.S. Pi	BN	INSIDE AUSTRALIA: AUD Strong Before U.S. GD	BFW	Kaplan Says Weak First Quarter Means Now Nc	BFW	Citi Economic Surprise Comparison by Region	BN
TSX little changed, durable goods data raises q	CNP	USA economy: Quick View - GM CEO to testify a:	EIU	*KAPLAN SAYS HE'S OPEN-MINDED ABOUT PO	BFW	U.S. Oct. ISM Regional Purchasers Index Compar	BN
New York City Water Reservoirs Above Normal Ca	BN	ASIA RATES/CREDIT DAYBOOK: Japan Inflation	BFW	*KAPLAN: DATA DOESN'T SUPPORT A FED RATE	BFW	New York Fed's GDP Model Sees 4Q U.S. GDP at	BN
Lipsky, Taylor Weigh in on Jackson Hole Del (Au	BN	Economists: Texas Economy Strong, Getting St	APW	Obama Is 'Pleased' With Yellen, White House Sa	BN	U.S. REACT: Consumer Sentiment Slips, Yet Tax-Cu	BI
BofA's Harris Says Fed Won't Start Taper in Septe	BN	Banks Lending Like It's 2007 Belied by Deposits	BN	U.S. Diesel Prices for the Week of April 11: Sumr	BN	Houston Purchasing Index Rises for Second Con	BN
Fed Asks Judge to Leave Swipe Fee Rules Alone I	BN	U.S. Reports Modestly Better Economic Growth	NYT	Bloomberg Advantage: Todd on Housing, Healt	BBR	U.S. Consumer Sentiment Unexpectedly Falls Fro	BN
Lew Tells Congress Treasury to Hit Debt Limit in M	BN	US economy shows signs of strength	SYH	Kaplan Says Fed Should Be 'Cautious, Patient'	BFW	U.S. October Wages Rose 3.4% Y/y: Atlanta Fed (	BN
U.S. Foreign Exchange Rates for the Week Endec	BN	Yellen Might Help Asia Kick the Easy-Money Ha	BBO	Obama `Cares Deeply' About Preserving Fed In	BFW	FastFT: US consumer sentiment gauge cools in N	FTI
Ford's Fusion Output Boost Tests $2,300 Premiu	BN	Wal-Mart Sues Visa Claiming Card Transaction Fe	BN	Premature to Rule Out 'Helicopter Drops,' Bern	BFW	U-MICH ECONOMIST CURTIN ON NOV. PRELIM S	BN
U.S. Crop Progress and Conditions for Aug. 25: S	BN	Top 300 Billionaires' Wealth Falls to $3.581 Trilli	BN	Obama to Meet With Fed Chair Yellen to Discu	APW	U.S. Consumer Sentiment Unexpectedly Drops A	BN
Lew Tells Congress Treasury Will Hit Debt Limit	BFW	Euro Drops to 3-Week Low Against Pound on Ou	BN	Is the U.S. Economy About to Go Bankrupt? Here	TST	Nov. Preliminary Univ. of Michigan Sentiment Re	BN
U.S. Treasury Statement and Cash Balance for Au	BN	Dick Bove Blasts Citi's 'Horrendous' Error in Mex	BN	*EARNEST SAYS DOESN'T EXPECT OBAMA TO U	BFW	Preliminary Nov. Michigan Sentiment Fell to 9	BFW
U.S. Diesel Prices for the Week of Aug. 26: Sumr	BN	Action Economics's Englund Says Fed Focus on R	BN	Obama 'Pleased' With Way Yellen Has Done Jo	BFW	*MICHIGAN PRELIM. NOV. CONSUMER SENTIM	BFW
Yellen Has 45% Chance to Replace Bernanke, S	BFW	Contrarian Corner's Eyes Results of 7-Year Note S	BN	Fed's Kaplan Says Sub-Zero Rates an Option, P	BFW	U.S. Nov. Prel. Michigan Consumer Sentiment (T	BN
Fed Seeks Federal Circuit Writ To Block Deposit	BLW	COULD THE BOC SERIOUSLY LAG THE FED	SCO	WHAT'S PRICED IN: No Changes For BOE, BOC	BFW	U.S. ECO PREVIEW: Univ. Mich Consumer Sent	BFW
*U.S. TREASURY TO REACH DEBT LIMIT IN MID-	BFW	St. Louis Federal Reserve Money Multiplier (Tabl	BN	*FED'S KAPLAN: NEGATIVE RATES HAVE NUMB	BFW	GSC COMITÉ DIARIO DE ESTRATEGIA 10 NOVIEM	GST
U.S. June Federal Reserve Finance Companies R	BN	Fewer Firings a Sign U.S. to Regain Growth Mom	BN	*KAPLAN: GREATER DEBT IN ADVANCES ECON	BFW	Powell Says Plan to Replace Libor Should Work	BFW
U.S. June Homebuyer Affordability Index: Sumn	BN	Maryland Business Activity for March (Table)	BN	*EARNEST SAYS OBAMA `PLEASED' BY WAY YEI	BFW	Donald Trump lauds PM Narendra Modi's econc	FEX
Fed Asks Judge to Leave Swipe Fee Rules in Plac	BN	Breakeven Inflation Rate: Five-to-Ten-Year Forwa	BN	What's Behind the Disconnect Between Stocks,	BLC	Quartz: What's left to explain Janet Yellen's disi	NS1
U.S. Raw Steel Production for the Week Ending A	BN	Carolinas Business Activity Decreased in March (	BN	Bloomberg Advantage: Penner on Positives of I	BBR	GSC INFORME DIARIO DE ESTRATEGIA 10 NOVII	GST
U.S. Poultry Condemned by Inspectors for July (T	BN	Federal Reserve Balance Sheet: Snapshot (Table)	BN	The Bloomberg Advantage: Cleveland on new I	BBR	Fed May Have No Choice But to Accelerate Rate	BBO
U.S. Chilled and Frozen Ready-To-Cook Poultry fo	BN	Fed Balance-Sheet Assets Rise $4.9b to $4.227	BFW	Fed's Fixed-Rate Reverse Repo Facility Draws $	BFW	Quarles in Charge? No, But He'll Dilute Dodd-Fra	BI
U.S. Poultry Slaughter and Live Weight for July (	BN	U.S. Money Supply Components for Week Endin	BN	Fed's Kaplan to Speak at Community Forum in	BFW	UniCredit Global Economic Forecasts as of Nov. 1	BN
U.S. Poultry Slaughter for July: Summary (Table)	BN	U.S. Mortgage-Backed Securities Purchase Progr	BN	Gallup Economic Confidence Tracking Poll Rises	BN	How will US companies use additional profits?	IXS

Note. Retrieved from Bloomberg Terminal for U.S. Economy top news on or shortly before stated dates, with minor editorial changes. Bloomberg news sources in bold.

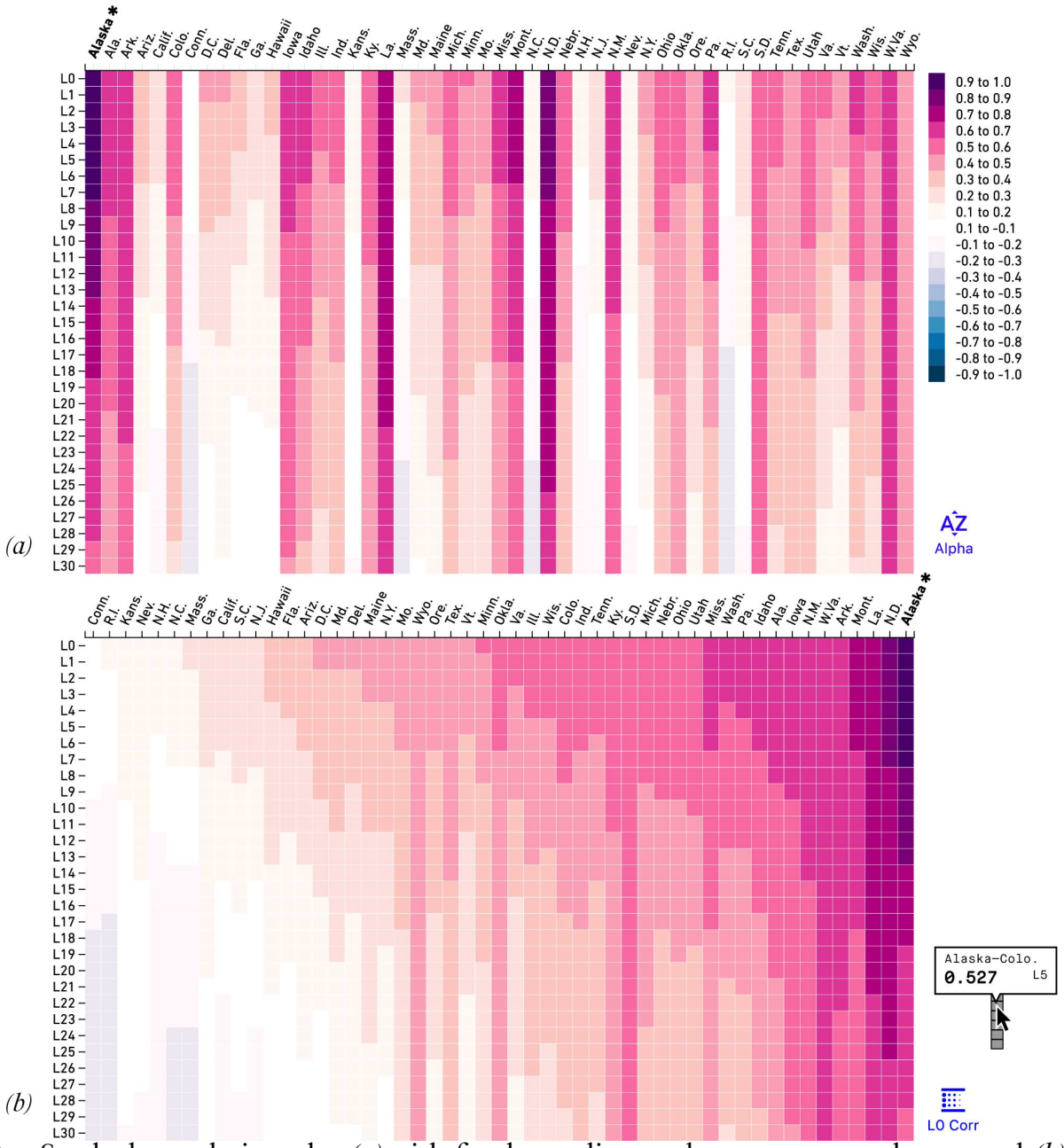
5.4. INTERACTIVE PLOTS

We now present the last set of interactive plots to help researchers exploring dynamic phasal relations of high-dimensional time series. These plots take on the familiar appearance of heat maps—that is, they all use color gradients to encode level values—and therefore do not need elaborations on their design. What's new here are their interactive features and the novel use of programming libraries not specifically developed for statistical research. Since these libraries are bleeding-edge graphic toolkits for the Internet, through the example of these elementary plots, we hope to motivate future research and development of web-based interactive graphic tools to explore the dynamic structures of high-dimensional time series. Since our focus here is to get familiar with novel plot features, to prevent us from being distracted by the expository need to set up other examples of high-dimensional time series which by their high-dimensionality need more space to contextualize and motivate, let's continue with the example of state unemployment rates introduced in Section 5.3.

Figure 5.19 presents a plot for lagged cross-correlations in the layout of a heat map, where we plot out the lagged correlations of `Alaska` with all states (including itself and `D.C.`) As before, the pivot component is marked by an asterisk (*). To better differentiate color shades, we utilize the color schemes from ColorBrewer (the so-named 9-class `RdPu` scale for positive correlations and `PuBu` for negative correlations), designed by Cynthia A. Brewer at the GeoVISTA Center at Pennsylvania State University.

Since a major drawback of heat maps is that they do not scale well for panels of large dimensions (see Section 5.1), in our example, rather than laying out all $51^2 \times 60 = 156,060$ colored cells, we only display a small block of the cells (2% of the total number of cells) at a time representing all autocorrelations and cross-correlations related to the pivot compo-

Figure 5.19 Two modes of interactive heat maps for lagged correlations.



Note. Stacked correlation plot *(a)* with fixed coordinates, by state names shown; and *(b)* with dynamic coordinates, by L0 correlation shown. Both plots truncated at Lag 30. Users can toggle between the two modes by clicking the blue icons displayed. Mouse over each cell gives correlation details.

nent (`Alaska`) and instead allow users to interact with the plot and explore other blocks of cells of their choosing. Users can switch between the pivot component by clicking on an arbitrary colored cell on the plot: on mouse release, the plot will display a subtle transition animation originated from the click point to remind users of the click event before presenting a new plot for the pivot component chosen.

Another problem of heat maps is that because of the grid layout, ironically it is often difficult to tell what coordinates each cell has, especially for those lie towards the center of a heat map, far away from both axes. This navigational problem is exacerbated for high-dimensional datasets. Heat maps like **Figures 5.2** and **5.4** give intuitive presentations of the overall levels of cross-correlations in the dataset but are impossible to navigate due to the high dimensionality. Our new plot offers an easy solution. Since each cell is rendered as a distinct node on the Document Object Model (DOM) tree for the plot (see Section 5.5), with simple JavaScripts to manipulate the DOM structure, we can provide additional interactive features to help users navigate the plot.

A tooltip automatically appears above the mouse cursor if the user's mouse hovers over a cell for a specified duration of time: the tooltip element dynamically displays detailed information about the cell, including the correlation pair (*e.g.*, "`Alaska-Colo.`", the former is the pivot component), the lag level (*e.g.*, "`L5`"), and the correlation value calculated (*e.g.*, "`.527`"). In addition, to easily sift out components with significant correlations, users have the option to sort the cells by correlations. To toggle between the two sort modes, users can click either the button labeled `Alpha` (by alphabetical order of a predefined array of variable names, in the current example, the two-letter postal abbreviations of all U.S. states and D.C.) or the one labeled `L0 Corr` (by Lag-0 Correlations). Both interactive features are implemented for all applicable plots presented in this section.

5.4.1 Tunnel Plot

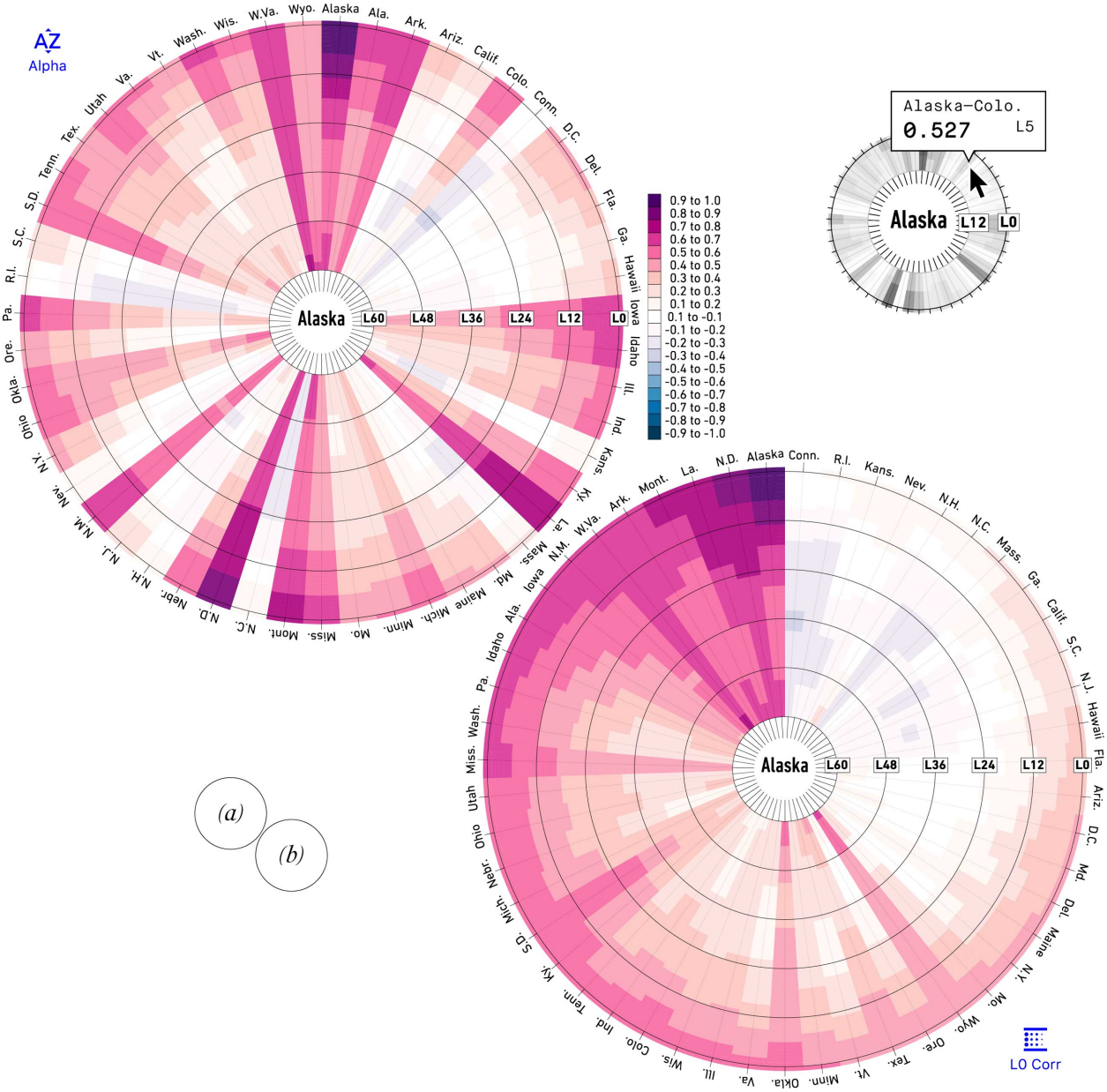
Building upon the classical heat map design, we now introduce an improved lagged correlation plot, the tunnel plot (see **Figure 5.20**). Its atypical circular look is the natural solution to this classical infographic design problem: we want a long axis to arrange the large collection of variables but we also desire a large plotting canvas to fit in all the colored cells for different lags. Since the circle is the plane curve enclosing the maximum area for a given arc length, it is natural to arrange the collection of variables around a circle. This also gives a clear indication that all variables are *a priori* equally important, as there is no preferred radial direction on a circle. And since correlations typically decay along the lags, the first a few lags are usually more informative: we arrange them on the larger rings towards the circular boundary. The end result resembles the view looking along the interior of a tunnel of correlations from the lag-0 correlations into the correlations of later lags.

Besides the general interactive features described above (mouse-over tooltips, mouse-click refresh of the pivot component, and sort-mode toggles), the tunnel plot has two more slider controls. Users can increase or decrease the view depth of the tunnel plot with the maximum lag slider control (labeled `Max Lag`: **Figure 5.21** illustrates the behavior. By decreasing the maximum lag and focusing on the earlier lags, users can perform correlation screenings with smaller tunnel plots, similar to the procedure introduced with the staff plots (**Figure 5.9**). We have also included the minimum correlation slider control (labeled `Min Corr`, see **Figure 5.22** for its behavior) that allows users to place a minimum correlation threshold on the plot. This removes unwanted cells representing insignificant correlations from the plot and declutters the plot for faster correlation screenings. Again, since correlations typically decay along the lags, later lags usually become unnecessary for

higher minimum correlation thresholds: the tunnel plot automatically decreases the view depth as users increase the minimum correlation threshold sliders, though users can override the automatic adjustment by manually sliding the maximum lag control again.

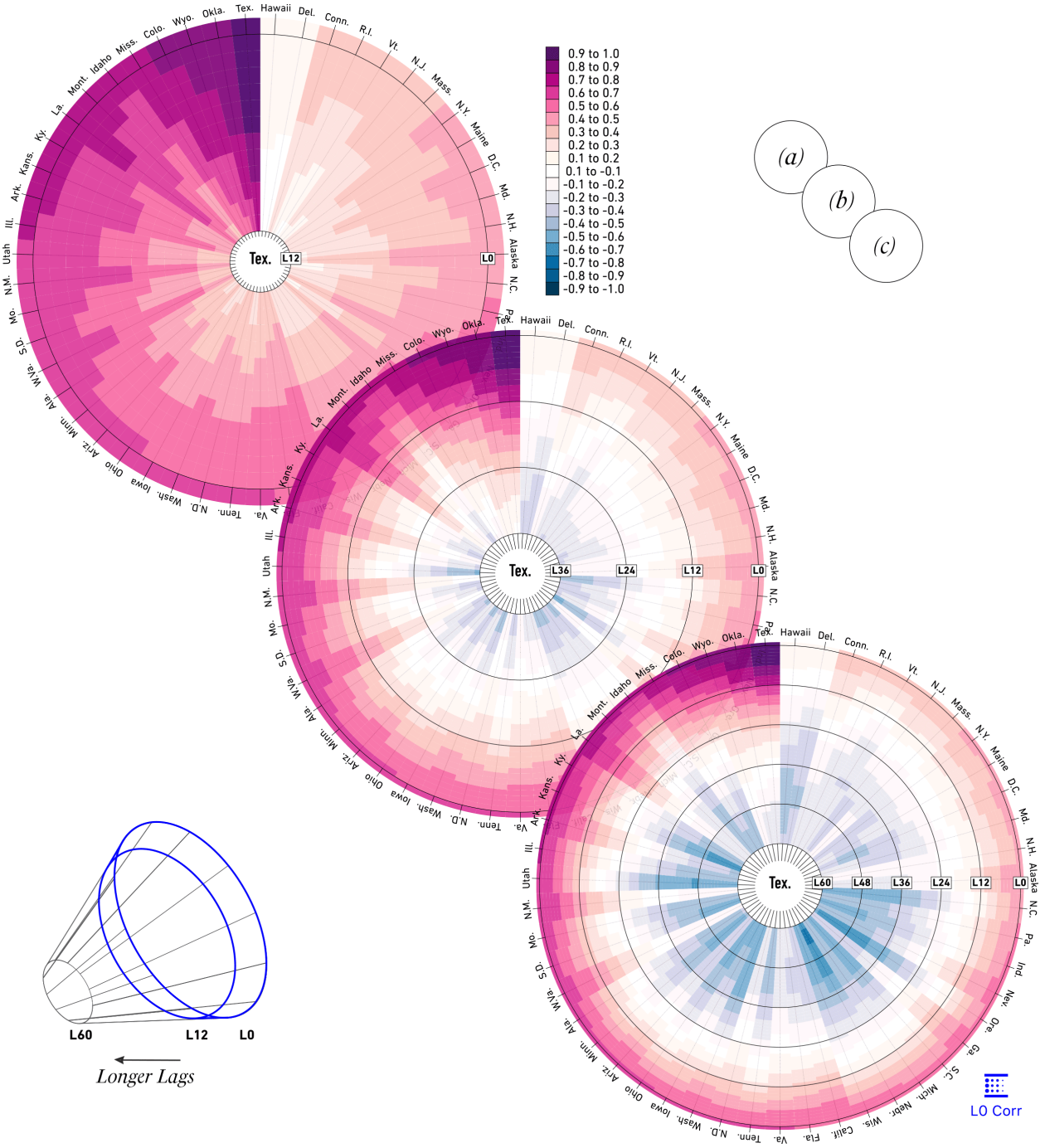
The tunnel plot and the staff plot presented in Section 5.2 are two strikingly different graphic tools for the same purpose. Both are designed to visualize lagged cross-correlations of high-dimensional time series: the latter encodes the data triad (secondary variable–lag–correlation) purely geometrically with 3-dimensional positions, while the former, a 2-dimensional plot, sheds the third dimension altogether thanks to the color coding of correlation levels. We intentionally keep the same unemployment rate example, so that users can compare and pick the tool of their liking. Since we have analyzed the example earlier in our primer to the staff plot, the more exotic plot, further elaboration is not needed. We will present a purely technical comparison of these tools in the concluding section.

Figure 5.20 Two modes of interactive tunnel plots.



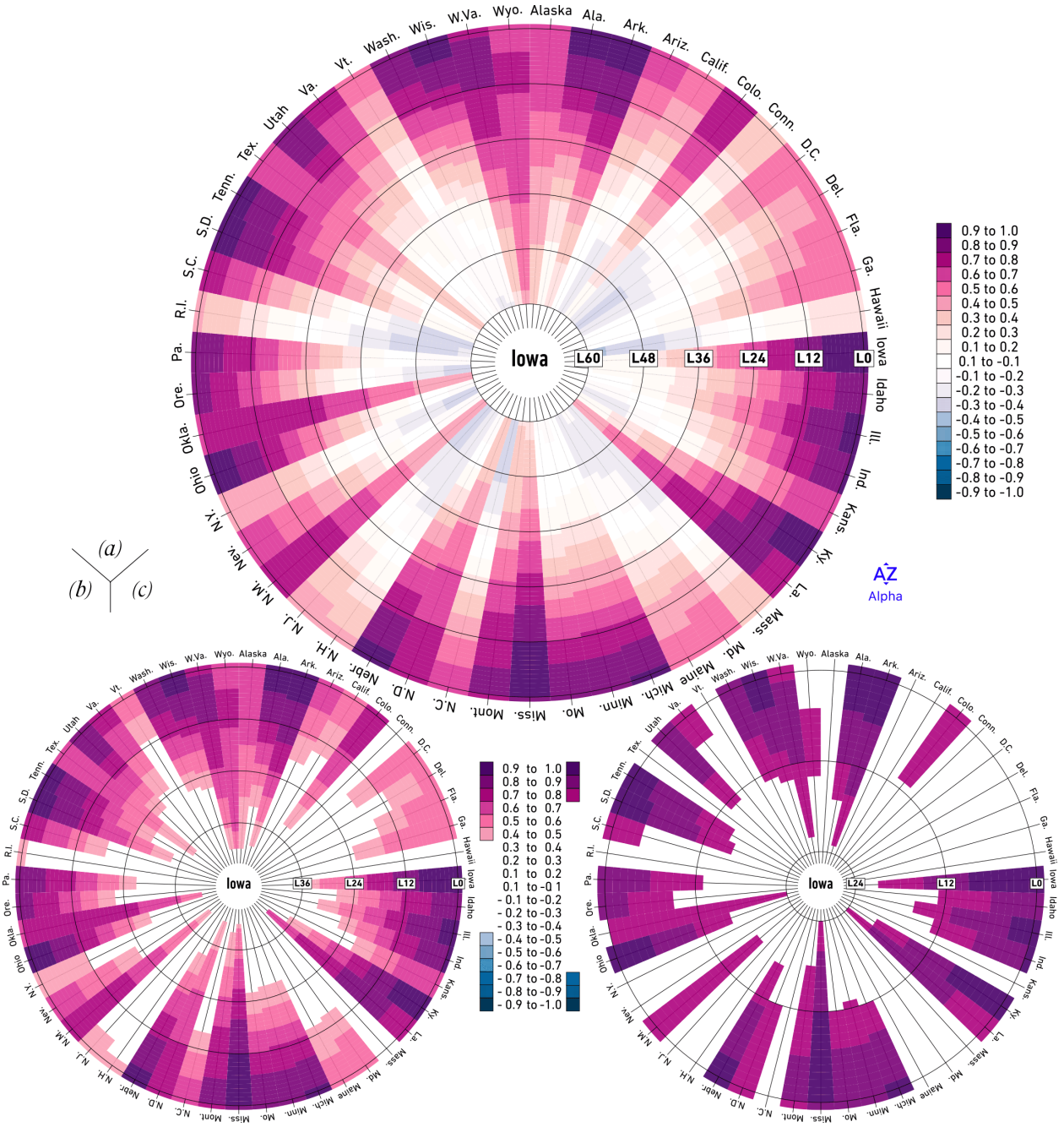
Note. *(a)* Tunnel plot with fixed coordinates, by state names shown; and *(b)* tunnel plot with dynamic coordinates, by L0 correlation shown. Users can toggle between the two modes by clicking the blue icons displayed. Mouse over each cell gives correlation details.

Figure 5.21 Adjusting tunnel plot view depth.



Note. Maximum lag: *(a)* 12; *(b)* 36; and *(c)* 60. Users can adjust the view depth by dragging the slider control labeled `Max Lag`.

Figure 5.22 Adjusting tunnel plot correlation threshold.



Note. Minimum correlation (in absolute value): *(a)* 0; *(a)* 0.5; and *(a)* 0.8. Users can adjust the correlation threshold by dragging the slider control labeled `Min Corr`.

5.4.2 Impulse Response Plot

Having offered a wide array of descriptive graphic tools for high-dimensional time series, we now present a useful plot for model checking, namely the interactive impulse response plot, see **Figure 5.23***(a)* and **5.24***(a)*. With the code base developed for previous plots, the impulse response plot comes for free, since presentation-wise it is simply a stacked correlation plot transposed, cf., **Figure 5.19**, even though the color gradient now encodes different information. To avoid possible confusions, we now adopt a different color scale: this is also suited because unlike correlations, impulse response functions are not bounded between -1 and 1.

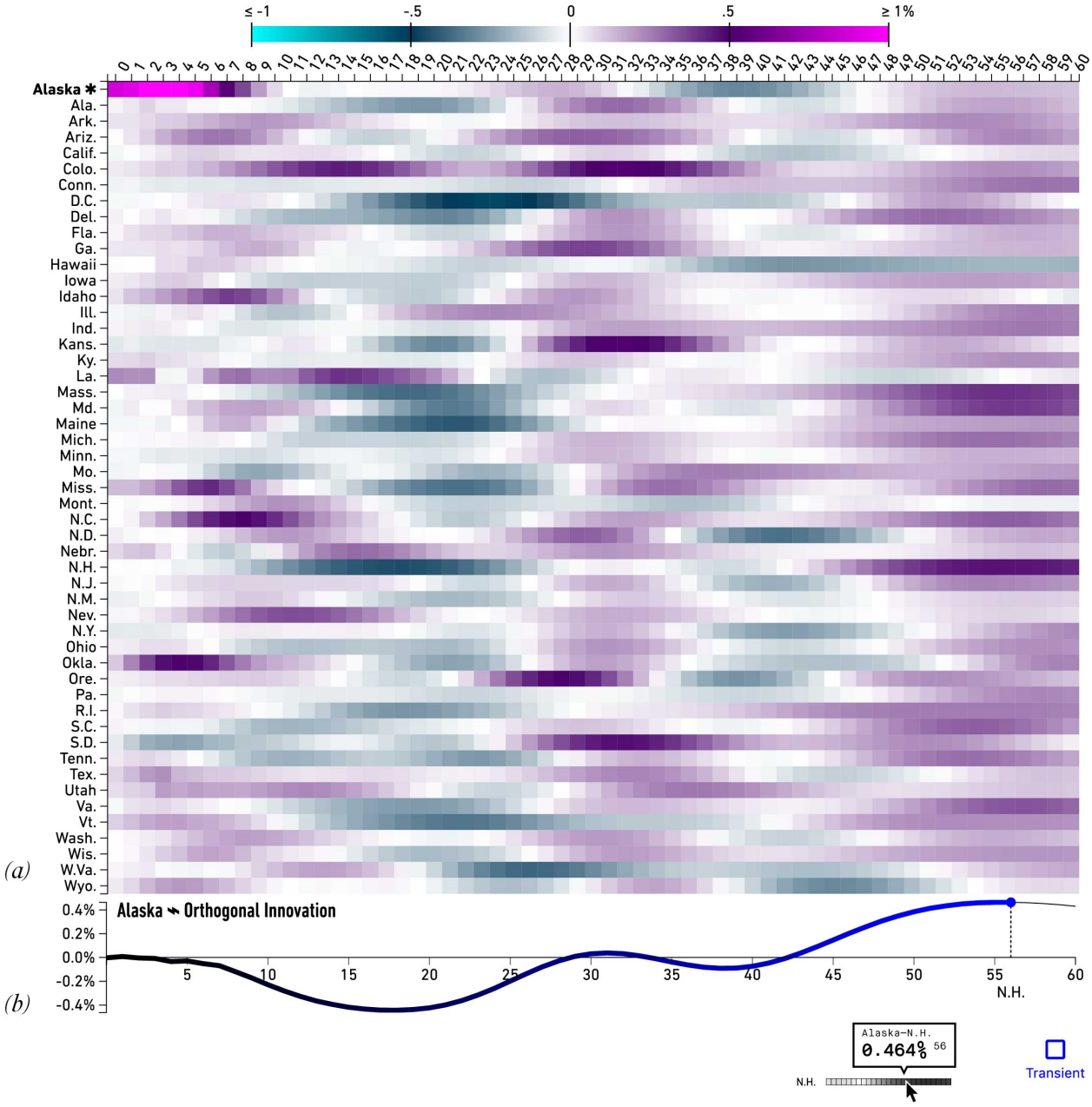
Since our task is to visualize impulse response functions, we assume readers have already built a plausible multivariate time series model; performed basic model checking and refinement; and obtained the impulse response functions as data frames to some given innovations. To continue our example of the U.S. unemployment rates, we first build a VAR(4) model with the `VAR` function in the `MTS` package for R. We perform simple model refinement and simplification with its `refVAR` function and model checking with the `MTSdiag` function. Subsequently, we obtain the impulse response functions with the `VARirf` function with orthogonal unit innovations. Readers can refer to standard textbooks on time series modeling, *e.g.*, Tsay (2013) and Tsay (2014), for details.

As before, users can obtain detailed information for each cell by triggering mouse-over event. However, since the color scale is no longer canonical and changes from case to case, users cannot be reasonably expected to read the new gradient with proficiency. To help them processing the diverse shapes in impulse responses, we also automatically present a small line plot at the bottom panel, triggered by the same mouse-over event, see **Figure**

5.23*(b)* and **5.24***(b)*, in addition to the tooltips displayed above the mouse cursor. Similarly, users can change the innovation by clicking any colored cell on the plot: this changes the innovation to a unit shock in the variable of that row. We also add subtle transition animations as visual feedbacks to different types of mouse events, in order to improve user experience and direct users' attention to the elements updated. In addition to the standard interactive features of the suite, users can also toggle between `Transient` and `Accumulated` responses, provided they have obtained these data frames separately from a third-party tool, for example, the `VARirf` function in the `MTS` package for R.

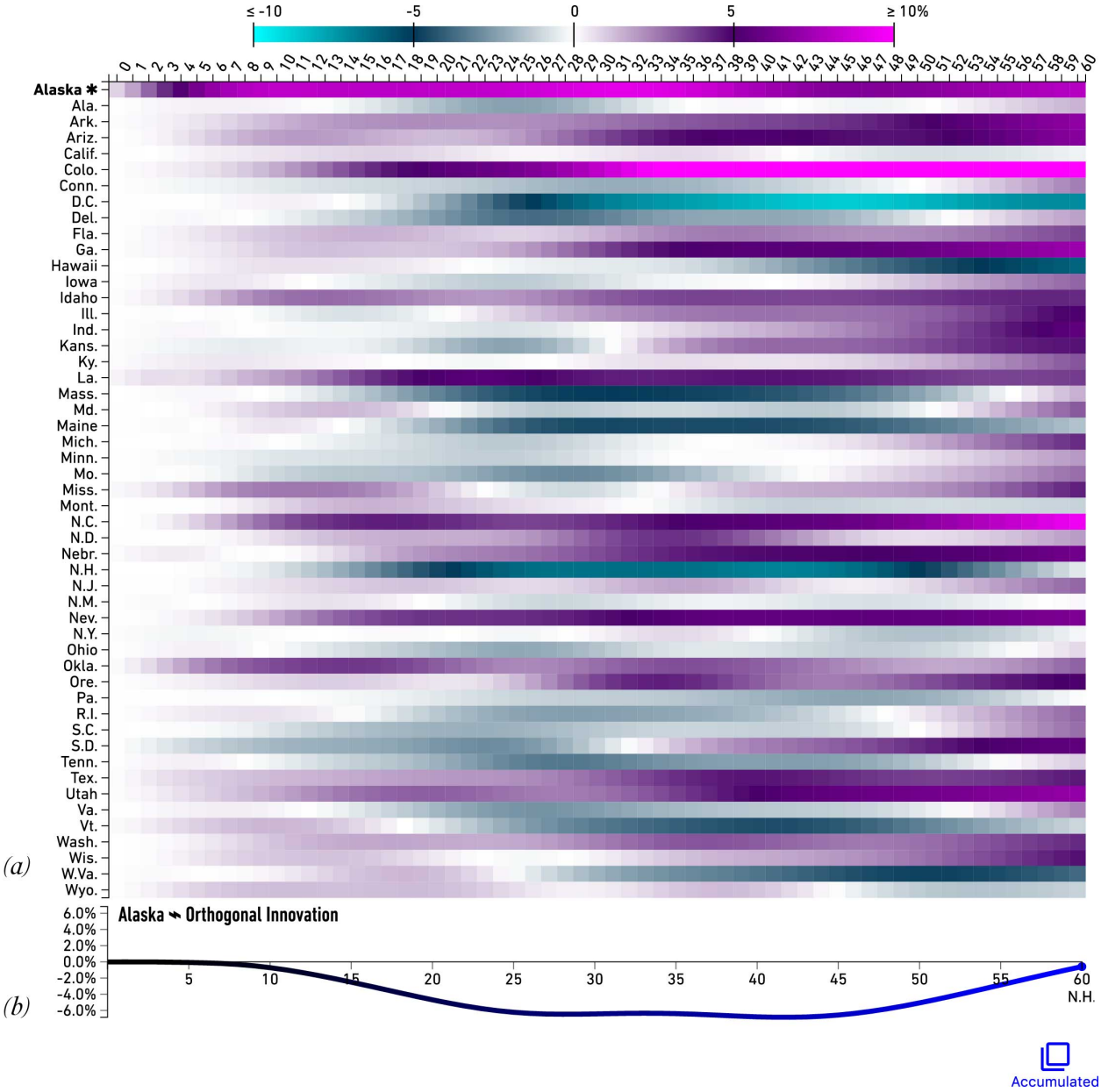
Now that we have presented a suite of interactive graphic tools for both data exploration and model checking, we could juxtapose them and compare how the impulse response functions predicted by a time series model fare against the correlation structures observed in the data. In fact, since all plots are rendered as individual DOM nodes in the same client browser using the identical JavaScript libraries, code savvy researchers can interact with multiple plots at runtime and develop more sophisticated graphic tools by chaining these basic plots. **Figures 5.25** and **5.26** provide a simple example.

Figure 5.23 Interactive impulse response plot: transient mode.



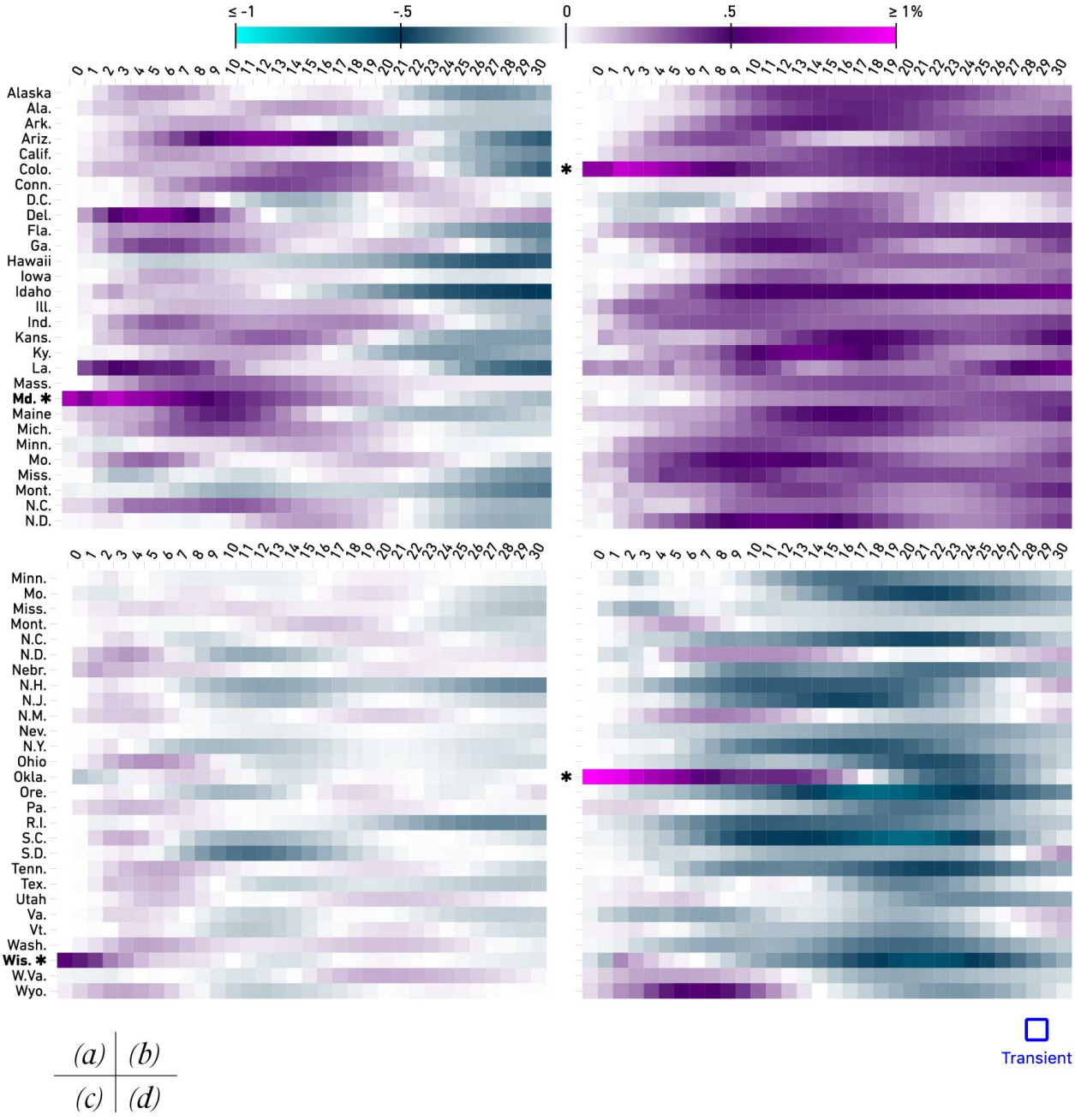
Note. *(a)* Heat map of impulse responses to orthogonal innovation in Alaska, displaying after left mouse click on `Alaska` row; and *(b)* line plot of New Hampshire to innovation in Alaska, displaying on mouse over `N.H.` row. Users can switch on the transient response mode by clicking the blue icon displayed.

Figure 5.24 Interactive impulse response plot: accumulative mode.



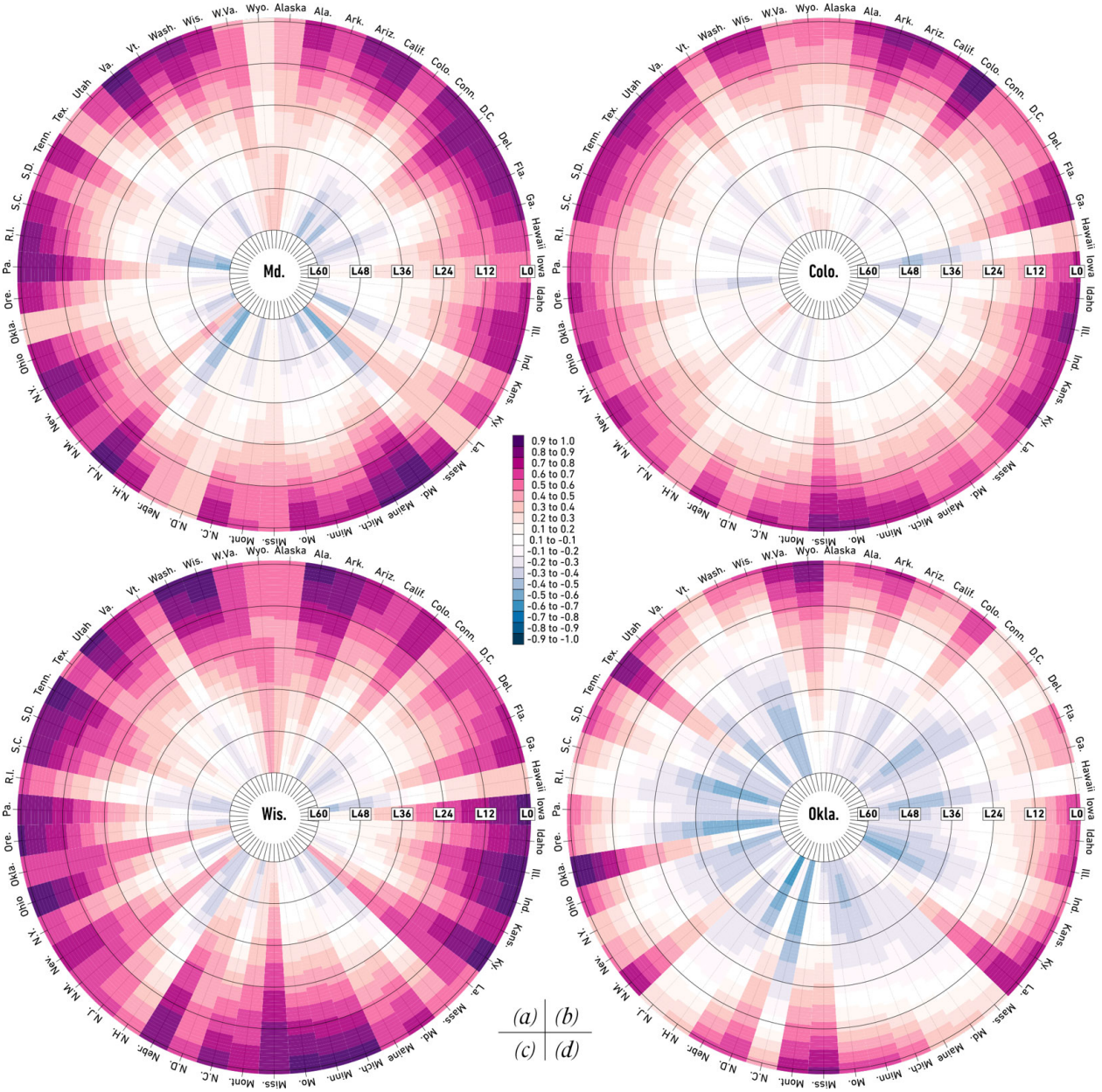
Note. *(a)* Heat map of impulse responses to orthogonal innovation in Alaska, displaying after left mouse click on `Alaska` row; and *(b)* line plot of New Hampshire to innovation in Alaska, displaying on mouse over `N.H.` row. Users can switch on the transient response mode by clicking the blue icon displayed.

Figure 5.25 Examples of different impulse response plots.



Note. Transient responses to orthogonal innovations in *(a)* Maryland; *(b)* Colorado; *(c)* Wisconsin; and *(d)* Oklahoma. The impulse response plot can be used in conjunction with the tunnel plot to compare modeled responses to correlations exhibited by the data, cf. **Figure 5.26**.

Figure 5.26 Examples of different tunnel plots for comparison.



Note. Lagged correlations of *(a)* Maryland; *(b)* Colorado; *(c)* Wisconsin; and *(d)* Oklahoma. The tunnel plot can be used in conjunction with impulse response plot to compare correlations exhibited by the data to modeled responses, cf. **Figure 5.25**.

5.5. TECHNICAL NOTE

The 3-dimensional dynamic staff and orbit plots are implemented in R using the `rgl` package (Adler et al., 2018, 3D Visualization Using OpenGL). To improve readability of the source code, we further use the `findPeaks` and `findValleys` functions in the `quantmod` package (Ryan and Ulrich, 2018, Quantitative Financial Modeling Framework) to obtain local extrema. Scenes of the plots are then exported to the HTML and JavaScript format using WebGL (Web Graphics Library), a JavaScript API for rendering interactive 2- and 3-dimensional graphics within compatible web browsers without the need for additional plug-ins, for post processing. Since the Open Graphics Library (OpenGL) was designed more than a quarter century ago to interact with a graphics processing unit (GPU) for hardware-accelerated rendering and R is not suited for high-performance GPU computing, the rendering process implemented is very slow (20min per frame) on a consumer-grade computer. (In June 2018, Apple deprecated OpenGL APIs on all of their platforms in favor of Metal 2, its own low-level APIs for near-direct access to the GPU.) To circumvent the performance problem, we export all rendered frames as PNGs (Portable Network Graphics) and manipulate these processed images with `jQuery`, a standard JavaScript library designed to simplify the client-side scripting. Performance-conscious researchers can optimize these plots for specific applications and, without relying on proprietary libraries, implement these 3-dimensional plots with basic JavaScript functions or `three.js`, a JavaScript library for creating and displaying animated 3-dimensional computer graphics in a web browser with WebGL. This solves the performance problem. However, since all graphic objects are hiding behind a rasterized canvas DOM using WebGL, limited interactive features can be convincing implemented.

5.5 Technical Note

In order bypass current technological limitations on developing general-purpose 3-dimensional statistical graphic tools, the tunnel plot and the impulse response plot take a completely different approach. To avoid the performance bottleneck of R, we process all statistical computing in R first, for example through the `MTS` package (Tsay, 2015, All-Purpose Toolkit for Analyzing Multivariate Time Series and Estimating Multivariate Volatility Models), and export the correlations and impulse response functions as data frames. This can be easily done in R with the `as.data.frame` and the `write.csv` functions. We render the plots procedurally in the web browser as SVGs (Scalable Vector Graphics), a XML-based plain text format for 2-dimensional vector images, with `D3.js` (Data-Driven Documents), a JavaScript library for producing dynamic and interactive visualizations in web browsers. As before, we use `jQuery` to code additional interactive functionalities.

Supplementary materials to the manuscript, including all source code and live demonstrations of the plots introduced, are available through the online archive of the dissertation project at `home.uchicago.edu/dwood/thesis`. We have packaged the source code for each plot as a standalone JavaScript library. Readers can also find further documentation on how to generate these plots from their own data sets by downloading the source code or through the web interface provided. All source code are provided "as is" and licensed under the MIT License.

REFERENCE

Abel, A. B., B. Bernanke, and D. Croushore (2016). *Macroeconomics* (9th Edition ed.). Pearson.

Adler, D., D. Murdoch, and et al. (2018). *rgl: 3D Visualization Using OpenGL*. R package version 0.99.16.

Amari, S.-I. (1968). Theory of information space: a geometrical foundation of the analysis of communication systems. RAAG Memoirs No. 4, Research Association of Applied Geometry.

Amari, S.-I. (1980). Theory of information space: A differential geometrical foundation of statistics. Post RAAG Report No. 106, Research Association of Applied Geometry.

Amari, S.-I. (1982a). Differential geometry of curved exponential families—curvatures and information loss. *The Annals of Statistic 10*(2), 357–385.

Amari, S.-I. (1982b). Geometrical theory of asymptotic ancillarity and conditional inference. *Biometrika 69*(1), 1–17.

Amari, S.-I. (1985). *Differential-Geometrical Methods in Statistics (Lecture Notes in Statistics)*. Springer-Verlag New York.

Amari, S.-I. (2016). *Information Geometry and Its Applications (Applied Mathematical Sciences)*. Springer Japan.

Andrews, D. W. K. and X. Cheng (2012). Estimation and inference with weak, semi-strong, and strong identification. *Econometrica 80*(5), 2153–2211.

Andrews, I. and A. Mikusheva (2016a). A geometric approach to nonlinear econometric models. *Econometrica 84*(3), 1249–126.

Andrews, I. and A. Mikusheva (2016b). Supplement to "a geometric approach to nonlinear econometric models". *Econometrica Supplemental Material*. Available at `dx.doi.org/10.3982/ECTA12030`.

Arwini, K. A. and C. T. J. Dodson (2008). *Information Geometry: Near Randomness and Near Independence (Lecture Notes in Mathematics)*. Springer-Verlag Berlin Heidelberg.

Atkinson, C. and A. F. S. Mitchell (1981). Rao's distance measure. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002) 43*(3), 345–365.

Barndorff-Nielsen, O., D. Cox, and N. Reid (1986). The role of differential geometry in statistical theory. *International Statistical Review / Revue Internationale de Statistique 54*(1), 83–96.

Barnhart, B. L. (2011). *The Hilbert-Huang Transform: theory, applications, development*. Ph. D. thesis, University of Iowa.

Basu, D. (1955). On statistics independent of a complete sufficient statistic. *Sankhyā: The Indian Journal of Statistics (1933-1960) 15*(4), 377–380.

Bhattacharyya, A. (1943). On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society 35*, 99–109.

Bhattacharyya, A. (1946). On a measure of divergence between two multinomial populations. *Sankhyā: The Indian Journal of Statistics (1933-1960) 7*(4), 401–406.

Box, G. E. (1976). Science and statistics. *Journal of the American Statistical Association 71*, 791–799.

Burbea, J. and J. Oller (1989). On Rao distance asymptotic distribution. Mathematical Preprint Series 67, Universitat de Barcelona.

Casella, G. and R. L. Berger (2001). *Statistical Inference (Duxbury Advanced Series)*. Duxbury Press;.

Čencov, M. N. (1978). Algebraic foundation of mathematical statistics. *Statistics: A Journal of Theoretical and Applied Statistics 9*(2), 267–276.

Chen, W. W. (2016). On the comparison of statistical curvature with gaussian curvature. *International Journal of Mathematics and Statistics Studies 4*(6), 13–23.

Chen, X. and YanqinFan (2006). Estimation of copula-based semiparametric time series models. *Journal of Econometrics 130*(2), 307–335.

Chen, Y. and M. Q. Feng (2003). A technique to improve the empirical mode decomposition in the hilbert-huang transform. *Earthquake Engineering and Engineering Vibration. 2*(1), 75–85.

Cleveland, W. S. and R. McGill (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association 79*(387), 531–554.

Climate Prediction Center (2018). Cold and warm episodes by season (oceanic niño index, ver. 5). Technical report, National Oceanic and Atmospheric Administration National Weather Service.

Costa, S. I., S. A.Santos, and J. E.Strapassonb (2015). Fisher information distance: A geometrical reading. *Discrete Applied Mathematics 197*, 59–69.

Dätig, M. and T. Schlurmann (2004). Performance and limitations of the hilbert–huang transformation (hht) with an application to irregular water waves. *Ocean Engineering 31*, 1783–1834.

Dawid, A. P. (1975). Comments on the paper by efron (1975). *The Annals of Statistics 3*(6), 1231–1234.

Dawid, A. P. (1977). Further comments on some comments on a paper by bradley efron. *The Annals of Statistics 5*(6), 1249.

Doob, J. (1990). *Stochastic Processes* (Revised ed.). Wiley-Interscience;.

Dufour, J.-M. and J. Jasiak (2001). Finite sample limited information inference methods for structural equations and models with generated regressor. *International Economic Review 42*(3), 815–844.

Dufour, J.-M. and M. Taamouti (2005). Projection-based statistical inference in linear structural models with possibly weak instruments. *Econometrica 73*(4), 1351–1365.

Dynken, E. (1951). Neobhodimye i dostatochnye statistiki dlja semejstva raspredelenij verojatnostej (necessary and sufficient statistics for the family of probability distributions). *Uspehi matematicheskih nauk 6*(1), 68–90.

Efron, B. (1975). Defining the curvature of a statistical problem (with applications to second order efficiency). *The Annals of Statistic 3*(6), 1189–1242.

Efron, B. (1978). The geometry of exponential families. *The Annals of Statistics 6*(2), 362–376.

Efron, B. (2016). Curvature and inference for maximum likelihood estimates. *The Annals of Statistic* (Preprint). Available at `statweb.stanford.edu/~ckirby/brad/papers/2016CurvatureInferenceMLEs.pdf`.

Efron, B. and D. Hinkley (1978). Assessing the accuracy of the maximum likelihood estimator: Observed versus expected Fisher information. *Biometrika 65*, 457–487.

Eguchi, S. (1985). A differential geometric approach to statistical inference on the basis of contrast functionals. *Hiroshima Mathematical Journal 15*(2), 341–391.

Elliott, G., U. K. Müller, and M. W. Watson (2015). Nearly optimal tests when a nuisance parameter is present under the null hypothesis. *Econometrica 83*(2), 771–811.

et al., A. S. (2018). *DescTools: Tools for Descriptive Statistics*. R package version 0.99.24.

Ewald, W. B. (2007). *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, Volume 2. Oxford University Press.

Federer, H. (1996). *Geometric Measure Theory (Classics in Mathematics)*. Springer-Verlag Berlin Heidelberg.

Fellows, I. (2012). Deducer: A data analysis gui for R. *Journal of Statistical Software 49*(8), 1–15.

Fisher, N. I. and P. Switzer (1985). Chi-plots for assessing dependence. *Biometrika 72*(2), 253–265.

Fisher, N. I. and P. Switzer (2001). Graphical assessment of dependence: is a picture worth 100 tests? *The American Statistician 55*.

Fisher, R. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical Engineering Sciences 222*(602), 594–604.

Fisher, R. (1934). Two new properties of mathematical likelihood. *Philosophical Transactions of the Royal Society A: Mathematical, Physical Engineering Sciences 144*(852), 285–307.

Fisher, R. A. (1990). *Statistical Methods, Experimental Design, and Scientific Inference: A Reissue of Statistical Methods for Research Workers (1925), The Design of Experiments (1935), and Statistical Methods and Scientific Inference (1956)*. Oxford University Press.

Fourier, J. J. (1822). *Théorie Analytique de la Chaleur*. Paris: Chez Firmin Didot, père et fils.

Fraser, D. (1962). On sufficiency and the exponential family. Technical report, Stanford University.

Fraser, D. (1963). On sufficiency and the exponential family. *Journal of the Royal Statistical Society. Series B (Methodological) 25*(1), 115–123.

Friendly, M. (2012). Corrgrams: Exploratory displays for correlation matrices. *The American Statistician 56*(4), 316–324.

Fritsch, F. N. and R. E. Carlson (1980). Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis 17*(2), 238–246.

Genest, C. and J.-C. Boies (2003). Detecting dependence with kendall plots. *The American Statistician 57*(4), 275–284.

Genest, C. and L.-P. Rivest (1993). Statistical inference procedures for bivariate archimedean copulas. *Journal of the American Statistical Association 88*(423), 1034–1043.

Griffin, W. L., N. I. Fisher, J. Friedman, C. G. Ryan, and S. Y. O'Reilly (1999). Cr-pyrope garnets in the lithospheric mantle. i. compositional systematics and relations to tectonic setting. *Journal of Petrology 40*(5), 679–704.

Guang, Y., X. Sun, M. Zhang, X. Li, and X. Liu (2014). Study on ways to restrain end effect of hilbert-huang transform". *Journal of Computers 25*(3), 23–311.

Guggenberger, P., F. Kleibergen, S. Mavroeidis, and L. Chen (2012). On the asymptotic sizes of subset anderson–rubin and lagrange multiplier tests in linear instrumental variables regression. *Econometrica 80*(6), 2649–2666.

Hall, B. C. (2013). *Quantum Theory for Mathematicians (Graduate Texts in Mathematics)*. Springer-Verlag New York.

Hansen, L. P. and T. J. Sargent (1981). Exact linear rational expectations models: Specification and estimation. *Federal Reserve Bank of Minneapolis Staff Report 71*.

Hariharan, H., A. Gribok, M. Abidi, and A. Koschan (2006). Image fusion and enhancement via empirical mode decomposition. *Journal of Pattern Recognition Research 11*(1), 16–31.

Hartigan, J. (1975). Printer graphics for clustering. *Journal of Statistical Computation and Simulation 4*(3), 187–213.

Hatcher, A. (2001). *Algebraic Topology*. Cambridge University Press.

Hicks, N. J. (1965). *Notes on Differential Geometry (Van Nostrand mathematical studies)*. Van Nostrand.

Hogg, R. V. and A. T. Craig (1995). *Introduction to Mathematical Statistics*. Prentice Hall.

Holland, P. W. (1973). Covariance stabilizing transformations. *The Annals of Statistics 1*(1), 84–92.

Huang, H. and J. Pan (2006). Speech pitch determination based on hilbert-huang transform. *Signal Processing 86*(4), 792–803.

Huang, N. E. and N. O. Attoh-Okine (Eds.) (2005). *The Hilbert-Huang Transform in Engineering*. CRC Press.

Huang, N. E., S. R. Long, and Z. Shen (1996). The mechanism for frequency downshift in nonlinear wave evolution. *Advances in Applied Mechanics 32*(117A, 117B, 117C), 59–117.

Huang, N. E. and S. S. P. Shen (Eds.) (2005). *Hilbert-Huang Transform and Its Applications*. World Scientific.

Huang, N. E., M.-L. Wu, W. Qu, S. R. Long, and S. S. P. Shen (2003). Applications of hilbert–huang transform to non-stationary financial time series analysis. *Applied Stochastic Models in Business and Industry 19*(n/a), 246–268.

Huang, N. E. and Z. Wu (2008). A review on hilbert-huang transform: Method and its applications to geophysical studies. *Reviews of Geophysics 46*(2), n/a.

Huzurbazar, V. (1950). Probability distributions and orthogonal parameters. *Mathematical Proceedings of the Cambridge Philosophical Society 46*(2), 281–284.

Huzurbazar, V. (1956). Sufficient statistics and orthogonal parameters. *Sankhyā: The Indian Journal of Statistics (1933-1960) 17*(3), 217–220.

Jaworski, P., F. Durante, W. K. Härdle, and T. Rychlik (Eds.) (2010). *Copula Theory and Its Applications: Proceedings of the Workshop Held in Warsaw, 25-26 September 2009*. Springer-Verlag Berlin Heidelberg.

Jeffreys, H. (1998). *Theory of Probability (Oxford Classic Texts in the Physical Sciences)*. Oxford University Press.

Joe, H., H. Li, and A. K. Nikoloulopoulos (2010). Tail dependence functions and vine copulas. *Journal of Multivariate Analysis 101*(1), 252–270.

Jost, J. (2017). *Riemannian Geometry and Geometric Analysis (Universitext)* (7th ed.). Springer-Verlag Berlin Heidelberg.

Kallenberg, W. C. M. and T. Ledwina (1999). Data-driven rank tests for independence. *Journal of the American Statistical Association 94*(445), 285–301.

Kass, R. E. (1980). *The Riemannian Structure of Model Spaces: A Geometrical Approach to Inference*. Ph. D. thesis, The University of Chicago.

Kass, R. E. and P. W. Vos (1997). *Geometrical Foundations of Asymptotic Inference*. Wiley-Interscience.

Kent, J. T. (1982). The fisher-bingham distribution on the sphere. *Journal of the Royal Statistical Society. Series B (Methodological) 44*(1), 71–80.

Kizhner, S., K. Blank, T. Flatley, N. E. Huang, D. Patrick, and P. Hestnes (2005). On the hilbert-huang transform theoretical developments. Technical report, National Aeronautics and Space Administration Goddard Space Flight Center.

Klüppelberg, C. and G. Kuhn (2009). Copula structure analysis. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 71*(3), 737–753.

Kobayashi, S. and K. Nomizu (1996a). *Foundations of Differential Geometry, Vol. 1 (Wiley Classics Library)*. Wiley-Interscience.

Kobayashi, S. and K. Nomizu (1996b). *Foundations of Differential Geometry, Vol. 2 (Wiley Classics Library)*. Wiley-Interscience.

Koopman, B. O. (1936). On distributions admitting a sufficient statistic. *Transactions of the American Mathematical Society 39*(3), 399–409.

Kosmidis, I. and D. Karlis (2016). Model-based clustering using copulas with applications. *Statistics and Computing 26*(5), 1079–1099.

Kullback, S. and R. Leibler (1951). On information and sufficiency. *Annals of Mathematical Statistics 22*(1), 79–86.

Lang, S. (2001). *Fundamentals of Differential Geometry (Graduate Texts in Mathematics)* (Corrected ed.). Springer-Verlag New York.

Leão, Jr., D., M. Fragoso, and P. Ruffino (1999). Characterization of Radon spaces. *Statistics and Probability Letters 43*(4), 409–413.

Leão, Jr., D., M. Fragoso, and P. Ruffino (2004). Regular conditional probability, disintegration of probability and Radon spaces. *Proyecciones (Antofagasta) 23*(1), 15–29.

Lehmann, E. L. (2011). *Fisher, Neyman, and the Creation of Classical Statistics*. Springer Science+Business Media.

Lehmannm, E. L. and J. P. Romano (2005). *Testing Statistical Hypothese (Wiley Series in Probability and Statistics)* (3rd ed.). Springer Science+Business Media.

Li, M. and Y. Huang (2014). Hilbert–huang transform based multifractal analysis of china stock market. *Physica A: Statistical Mechanics and its Applications 406*(n/a), 222–229.

Madsen, L. T. (1978). *The Geometry of Statistical Models: A Generalization of Curvature*. Ph. D. thesis, University of Copenhagen.

Madsen, L. T. (1979). The geometry of statistical models: A generalization of curvature. Technical Report 79/1, Danish Social Science Research Council.

Mahalanobis, P. (1936). On the generalised distance in statistics. *Proceedings National Institute of Science, India 2*(1), 49–55.

Manfred, E. (1973). *The Physicist's Conception of Nature*, Chapter The Origin of Biological Information. Springer Netherlands.

Mankiw, N. G. (2015). *Macroeconomics*. Worth Publishers.

Marriott, P. and M. Salmon (Eds.) (2000). *Applications of Differential Geometry to Econometrics*. Cambridge University Press.

McKenna, S., M. Meyer, C. Gregg, and S. Gerber (2016). s-corrplot: An interactive scatterplot for exploring correlation. *Journal of Computational and Graphical Statistics 25*(2),

445–463.

Milnor, J. W. (1997). *Topology from the Differentiable (Princeton Landmarks in Mathematics and Physics)* (Revised ed.). Princeton University Press.

Mitchell, A. F. S. (1962). Sufficient statistics and orthogonal parameters. *Mathematical Proceedings of the Cambridge Philosophical Society 58*(2), 326–337.

Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer-Verlag New York.

Nogales, A. G., J. A. Oyola, and P. Pérez (2000). On conditional independence and the relationship between sufficiency and invariance under the bayesian point of view. *Statistics Probability Letters 46*(1), 75–84.

Oh, D. H. and A. J. Patton (2016). High-dimensional copula-based distributions with mixed frequency data. *Journal of Econometrics 193*(2), 349–366.

Parey, A. and R. Pachori (2012). Variable cosine windowing of intrinsic mode functions: Application to gear fault diagnosis. *Measurement 45*(3), 415–426.

Pathak, P. (1997). *Breakthroughs in Statistics (Perspectives in Statistics)*, Chapter Introduction to Rao (1945). Springer-Verlag New York.

Pennec, X. (2004). Probabilities and statistics on riemannian manifolds : A geometric approach. Technical Report 5093, Institut national de recherche en informatique et en automatique.

Petersen, P. (2006). *Riemannian Geometry (Graduate Texts in Mathematics, Vol. 171)* (2nd ed.). Springer, New York, NY.

Plancherel, M. (1910). Contribution à l'étude de la représentation d'une fonction arbitraire par les intégrales définies. *Rendiconti del Circolo Matematico di Palermo 30*(289–335).

Prestini, E. (2004). *The Evolution of Applied Harmonic Analysis: Models of the Real World*. Springer.

Rao, C. R. (1945). Introduction to Rao (1945). *Bulletin of Calcutta Mathematical Society 37*, 81–91.

Rao, C. R. (1997). *Breakthroughs in Statistics (Perspectives in Statistics)*, Chapter Information and the Accuracy Attainable in the Estiwmation of Statistical Parameters. Springer-Verlag New York.

Reeds, J. (1975). Discussion on Professor Efron's paper. *The Annals of Statistics 3*(6), 1234–1238.

Rice, J. A. (2006). *Mathematical Statistics and Data Analysis (Duxbury Advanced Series)*. Duxbury Press;.

Ryan, J. A. and J. M. Ulrich (2018). *quantmod: Quantitative Financial Modelling Framework*. R package version 0.4-13.

Said, S., L. Bombrun, Y. Berthoumieu, and J. H. Manton (2017). Riemannian gaussian distributions on the space of symmetric positive definite matrices. *IEEE Transactions on Information Theory 63*(4), 2153–2170.

Sarkar, D. (2008). *Lattice: Multivariate Data Visualization with R*. New York: Springer. ISBN 978-0-387-75968-5.

Sato, Y., K. Sugawa, and M. Kawaguchi (1979). The geometrical structure of the parameter space of the two-dimensional normal distribution. *Reports on Mathematical Physics 16*(1), 111–119.

Scharf, L. L. (1991). *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Pearson.

Schepsmeier, U. (2015). Efficient information based goodness-of-fit tests for vine copula models with fixed margins: A comprehensive review. *Journal of Multivariate Analysis 138*(na), 34–52.

Shankar, R. (2011). *Principles of Quantum Mechanics, 2nd Edition*. Plenum Press.

Skiovgaard, L. T. (1981). A riemannian geometry of the multivariate normal model. Technical Report 167, Stanford University.

Sklar, A. (1959). Fonctions de répartition à *n* dimensions et leurs marges. *Publications de l'Institut de statistique de l'Université de Paris 8*, 229–231.

Smith, M. S. (2015). Copula modelling of dependence in multivariate time series. *International Journal of Forecasting 31*(3), 815–833.

Stein, E. M. and R. Shakarchi (2003). *Fourier Analysis: An Introduction (Princeton Lectures in Analysis, Volume 1)*. Princeton University Press.

Stock, J. H. and J. H. Wright (2000). GMM with weak identification. *Econometrica 68*(5), 1055–1096.

Tsay, R. S. (2013). *Multivariate Time Series Analysis: With R and Financial Applications*. Wiley Series in Probability and Statistics. Wiley.

Tsay, R. S. (2014). *An Introduction to Analysis of Financial Data with R*. Wiley Series in Probability and Statistics. Wiley.

Tsay, R. S. (2015). *MTS: All-Purpose Toolkit for Analyzing Multivariate Time Series (MTS) and Estimating Multivariate Volatility Models*. R package version 0.33.

Wilbert C. M. Kallenberg, T. L. and E. Rafajłowicz (1997). Testing bivariate independence and normality. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002) 59*(1), 42–59.

Wilkinson, L. and M. Friendly (2012). The history of the cluster heat map. *The American Statistician 63*(2), 179–184.

Yan, J. (2007). Enjoy the joy of copulas: With a package copula. *JSS Journal of Statistical Software 21*(4), 1–21.

Yang, Z. and L. Yang (2009). A new definition of the intrinsic mode function. *World Academy of Science, Engineering and Technology 60*, 822–825.

Zamenhof, L. L. (1889). *Dr. Esperanto's International Language: Introduction & Complete Grammar*. Oxford University Press.