

Reconstruction and prediction for nonlinear dissipative stochastic systems: A data-driven contact geometry paradigm [version 1]

Junli Wang¹, Tiejian Li^{1,2}, Deyu Zhong^{1,2,✉}

¹State Key Laboratory of Hydrosience and Engineering, Tsinghua University, Beijing 100084, China

²Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China

Received: 17 Apr 2026

© 2026 The Author(s). This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

ABSTRACT

Nonlinear dissipative stochastic systems are widespread but face limitations from three main issues in traditional paradigms: symplectic geometry's incompatibility with dissipation and stochasticity, classical stochastic analysis's failure to capture high-order statistical information (missing path dependence), and underuse of high-dimensional observational data through empirical/semi-empirical parameterisations. To overcome these, we introduce a data-driven contact geometry paradigm that reverses the conventional "model-to-data" approach to "data→probability→geometry→dynamics." Based on stochastic vector bundles, contact geometry, and the least constraint theorem, this framework encodes the evolution of probability, both state-driven distributional changes and distribution-driven state dependencies, into geometric structures. Starting from observational data, we construct infinite-order jet bundles to preserve complete statistical information, derive the system dynamics through variational principles, and inherently incorporate dissipation and path dependence. When considering water cycle dynamics, this paradigm enables parameter-free system reconstruction and stable long-term predictions grounded in global invariants, without relying on phenomenological parameterisations. It offers a unified first-principles framework for characterising nonlinear dissipative stochastic systems.

KEYWORDS

Contact geometry, stochastic vector bundles, least constraint theorem, probability evolution, non-equilibrium systems, water cycle.

1 Introduction

1.1 The fundamental challenge

Nonlinear, dissipative, and stochastic systems are ubiquitous across natural science, engineering, and social science, forming the core of phenomena that shape our understanding of the real world (Zwanzig, 2001; Cross et al., 1993). From the chaotic dynamics of turbulent fluids and the non-equilibrium evolution of climate systems, to the nonlinear water and energy fluxes in the hydrological cycle, the stochastic fluctuations of biochemical reactions in living organisms, and the erratic dynamics of financial markets, these systems are universally characterized by three inseparable traits: inherent nonlinearity that drives emergent behavior and threshold effects, irreversible dissipation that mediates energy and information exchange with the surrounding environment, and intrinsic stochasticity that introduces irreducible uncertainty and path-dependent evolution. Accurately characterising, reconstructing, and predicting the dynamics of these systems is not only a long-standing fundamental problem in theoretical physics and applied mathematics, but also a critical practical necessity for risk management, decision-making, and sustainable regulation of complex real-world systems, from water resource management to climate change adaptation.

For centuries, the study of dynamical systems has been built on two parallel, largely disconnected traditions, creating a fundamental rift in our ability to describe nonlinear dissipative stochastic systems. The first tradition, rooted in classical analytical mechanics, relies on geometric frameworks—most notably symplectic geometry—to describe deterministic, conservative systems (Abraham et al., 1978; Arnold, 2010). This framework has achieved extraordinary success in explaining the motion of celestial bodies, the dynamics of conservative mechanical systems, and integrable Hamiltonian systems, in which energy is conserved, phase space volume is preserved by Liouville's theorem, and evolution is time-reversible. However, the vast majority of real-world systems are open, exchanging energy, mass, and information with their environment; they are inherently dissipative, irreversible, and subject to stochastic fluctuations, features that directly contradict the core mathematical assumptions of symplectic geometry. While decades of work have sought to adapt symplectic frameworks to dissipative and stochastic systems through phenomenological modifications, such as time-dependent Hamiltonians, non-canonical Poisson brackets, or auxiliary phase-space variables, these efforts inevitably compromise the geometric consistency and first-principles rigour that make symplectic geometry powerful. They treat dissipation and stochasticity as external "add-ons" to an otherwise conservative core, rather than intrinsic properties of the system itself, resulting in patchwork models that lack a unified mathematical foundation and often fail to generalise beyond the specific conditions for which they were calibrated.

The second dominant tradition, classical stochastic analysis, centres on the evolution of probability distributions via tools such as the Fokker-Planck equation, Langevin dynamics, and Monte Carlo methods (Van Kampen, 1992). This framework has proven invaluable for quantifying statistical moments and average behaviour of random processes, but it suffers from fundamental limitations that prevent a

✉ Address correspondence to Deyu Zhong, zhongdy@tsinghua.edu.cn

complete description of nonlinear dissipative stochastic systems. By focusing almost exclusively on the evolution of low-order statistical moments (mean, variance, and occasionally skewness and kurtosis), classical stochastic analysis discards the infinitely higher-order statistical information that is essential for characterising non-Gaussian distributions, extreme events, and the tail behaviour that dominates risk in real-world systems. More fundamentally, this approach cannot capture path dependence, the defining feature of non-integrable systems, in which different evolutionary trajectories can lead to distinct final states even with identical initial conditions (Zhong et al., 2020; Zhong et al., 2024). This path dependence, which underpins memory effects, hysteresis, and non-equilibrium irreversibility in complex systems, is not merely a secondary feature of stochastic dynamics, but a core physical property that classical methods cannot encode in a structured way. Most critically, traditional stochastic analysis treats system dynamics and probability evolution as separate, one-way coupled entities: deterministic dynamics drive changes in probability distributions, but there is no coherent framework for probability evolution to act as the fundamental driver of system dynamics, preventing a bidirectional, closed-loop description of the interplay between stochasticity and deterministic motion (Zhong et al., 2025).

Compounding these theoretical and mathematical challenges is a practical, data-centric crisis that has become increasingly acute in the era of big data. Modern observational and sensing technologies, from satellite remote sensing and in-situ sensor networks to high-resolution numerical simulations, generate petabytes of high-fidelity data capturing the behaviour of complex systems across multiple spatial and temporal scales. The hydrological cycle, for example, is monitored by soil moisture sensors, groundwater monitoring wells, eddy covariance towers, stream gauges, and atmospheric sounders, producing continuous, multi-variate observations of the system's state.

Yet traditional modelling paradigms are fundamentally unable to leverage this wealth of data. Because these approaches require the priori specification of governing equations and a small set of key state variables, they are forced to discard the vast majority of available observational data, using only a tiny fraction for model calibration and parameter fitting. The unresolved subgrid-scale processes, multi-scale interactions, and emergent behaviours that dominate complex system dynamics are reduced to empirical parameterisations, thereby introducing massive, unquantifiable uncertainties and systematically biasing model predictions. This creates a paradox: we have unprecedented access to observational data that encodes the full statistical behaviour of real-world systems, but our dominant modelling paradigms lack a framework to convert this data directly into a complete, consistent description of system dynamics, without relying on restrictive, often unrealistic prior assumptions about the underlying physical mechanisms.

Taken together, these challenges converge on a single, fundamental, unresolved question that has stymied the study of nonlinear dissipative stochastic systems for decades: how can we construct a unified, first-principles mathematical framework that simultaneously accommodates nonlinearity, intrinsic dissipation, and stochasticity within a single coherent structure; that encodes the full statistical information of the system, including path dependence and non-equilibrium irreversibility; that enables the complete reconstruction of system dynamics directly from observational data, without ad hoc model assumptions or phenomenological corrections; and that provides stable, long-term predictive capabilities for non-integrable systems, where traditional methods fail due to error accumulation and a lack of global invariants.

Answering this question requires not just incremental technical improvements to existing frameworks, but a fundamental paradigm shift in how we conceptualise the relationship between data, probability, geometry, and dynamics in complex stochastic systems. This paper presents such a paradigm shift, rooted in the contact geometry of stochastic vector bundles and the least constraint theorem, that addresses this fundamental challenge at its core.

1.2 Core bottlenecks of traditional paradigms

Traditional paradigms for studying nonlinear, dissipative, stochastic systems face three fundamental and interconnected limitations that have persisted despite decades of research.

Bottleneck 1: geometric incompatibility. Most classic dynamical theories in physics and applied mathematics are fundamentally grounded in symplectic geometry, the foundation of Hamiltonian and Lagrangian theory. Symplectic geometry, based on closed 2-forms $\omega = dq^i \wedge dp_i$, can describe only conservative, deterministic systems. The requirement that the symplectic form be closed ($d\omega = 0$) is equivalent to energy conservation and phase space volume preservation via Liouville's theorem. This mathematical structure is well-suited to Hamiltonian systems but cannot intrinsically accommodate dissipation or stochastic behaviour. Consequently, researchers modelling dissipative or stochastic systems within the symplectic framework must introduce phenomenological terms as external modifications to the Hamiltonian structure. These additions are not derived from first principles but are imposed to match observed behaviour, and thus violate the theory's geometric consistency. The result is a patchwork approach in which the elegant mathematical structure of symplectic geometry is compromised by ad hoc additions that lack a theoretical foundation.

Bottleneck 2: statistical limitations. Classical stochastic analysis, including the Fokker-Planck equation and Langevin approaches, focuses on the evolution of low-order statistical moments (mean, variance) of probability distributions. Although these methods have proven valuable for many applications, yet they suffer from fundamental limitations that restrict their applicability to complex systems. First, by focusing only on low-order moments, these approaches lose all higher-order statistical information, including skewness, kurtosis, and higher-order cumulants, which are essential for characterising non-Gaussian distributions and extreme events. Second, classical stochastic analysis cannot characterise path dependence, the essential phenomenon in which different evolution paths lead to different final states even when starting from the same initial conditions. This path dependence is a hallmark of non-integrable systems and is crucial for understanding memory effects and hysteresis. Third, these methods fail to capture non-equilibrium irreversibility in non-integrable systems, resulting in the loss of essential physics of how systems evolve towards or away from equilibrium. Finally, classical approaches preclude a bidirectional, complete description in which system dynamics and probability evolution mutually influence one another, instead treating dynamics and statistics as separate aspects that do not interact in a structured way.

Beyond the two fundamental bottlenecks identified above, traditional paradigms face a critical practical challenge when dealing with complex systems characterised by a huge number of interacting variables. This third bottleneck is particularly acute in modern scientific

applications where the complexity of systems of interest far exceeds the capacity of traditional modelling approaches.

Bottleneck 3: variable truncation and data underutilisation. Complex real-world systems, such as turbulent flows, climate systems, or hydrological cycles, involve an enormous number of degrees of freedom that interact in nonlinear and often unpredictable ways. The number of variables required to fully describe such systems can range from thousands to millions, or even billions. It is practically impossible to derive dynamical equations for every variable, given computational limitations and the often poorly understood interactions among variables. Consequently, traditional approaches must make drastic simplifications that fundamentally limit their predictive power.

Traditional approaches have to select only a small subset of key variables to model explicitly. For example, the Navier-Stokes equations describe only the mean velocity field, while subgridscale motions, which contain crucial information about energy dissipation and mixing, must be parameterised using empirical relationships that are often poorly constrained. Similarly, general circulation models used for climate prediction track only large-scale atmospheric variables, parameterising cloud microphysics, boundary layer processes, and other small-scale phenomena that significantly influence overall system behaviour. These parameterisations introduce substantial uncertainties and limit models' ability to capture emergent phenomena.

The vast majority of system variables are not explicitly resolved but are instead represented by empirical parameterisations, such as eddy-viscosity models, cloud-fraction schemes, and infiltration formulae. These parameterisations are typically derived from simplified theoretical arguments or fitted to limited observational data, and they often fail to capture the full complexity of the processes they represent. As a result, important physical mechanisms may be misrepresented or neglected entirely, causing systematic errors in model predictions.

Modern observational systems generate massive datasets rich in information about system behaviour across multiple scales. Satellite observations, sensor networks, and high-resolution simulations produce petabytes of data that capture system behaviour in unprecedented detail. However, because traditional models simulate only a limited set of variables, the vast majority of observational data remains unused in the modelling process. The information content of these data is effectively discarded, representing a tremendous waste of potentially valuable information that could constrain and improve model predictions.

This truncation of variables and underutilisation of data fundamentally impose limitations on the predictive power of traditional models. The unresolved variables often contain crucial information about system evolution, particularly for non-equilibrium and path-dependent processes. When these variables are parameterised rather than explicitly modelled, the resulting predictions inevitably accumulate errors and miss important dynamical features. The inability to fully utilise available data leaves models systematically underconstrained, leading to large prediction uncertainties and a limited ability to quantify and reduce them. It is worth noting that the popular notion of a digital twin for catchment remains subject to the same bottleneck as traditional approaches: its model-based nature leaves no room to accommodate more data than nature has.

1.3 The paradigm shift: from model-to-data to data-to-model

The fundamental innovation of this work is a complete reversal of the traditional research logic, which addresses all three bottlenecks simultaneously. The established paradigm follows a clear sequence that has dominated scientific modelling for centuries.

Traditional logic: mechanical model $\rightarrow$ statistical properties $\rightarrow$ data comparison

Researchers first postulate a dynamical equation based on physical considerations, theoretical arguments, or analogies with simpler systems. These equations typically take the form of differential equations that describe the evolution of selected state variables. Next, they derive the statistical consequences of these equations, either analytically or via numerical simulation. Finally, they compare the predicted statistical properties with observed data. If discrepancies arise between model predictions and observations, the model is adjusted phenomenologically by adding new terms, modifying coefficients, or introducing additional parameterisations. This process is iterated until the model achieves acceptable agreement with the data, though the resulting model may have little connection to the underlying physics and may perform poorly when extrapolated to new conditions.

The new paradigm inverts this logic entirely, starting from data rather than from assumed equations.

New logic: observed data statistics $\rightarrow$ probability evolution $\rightarrow$ geometric structure $\rightarrow$ dynamical equations

The formulation is based on our recent studies (Zhong et al., 2024; Zhong et al., 2025). Starting from the statistical regularities of observed data, we first extract the complete probability distribution $P(y)$ and all its spatial derivatives. This distribution encodes all statistical information about the system, including correlations and dependencies among variables that would be difficult or impossible to capture with traditional modelling approaches. Next, we construct the infinite-order jet bundle $J^\infty(E, P)$ (E is the vector space and P is the probability space), which encodes all statistical information in a geometric structure. This jet bundle provides a coordinate system for the infinite-dimensional space of probability distributions and their derivatives, transforming the problem of characterising probability evolution into a problem of differential geometry. We then identify the natural contact structure $\Theta = dP - \rho dy^i$ on this bundle, which encodes the fundamental relationship between probability evolution and state changes. Finally, we derive the complete dynamics via the least-constraint principle $\delta \int \epsilon dt = 0$, which provides a variational criterion for determining the evolution of the system.

This approach eliminates phenomenological assumptions about dissipation and noise which emerge intrinsically from the geometric structure (Zhong et al., 2024). The contact structure naturally accommodates irreversibility through its non-closedness ($d\Theta \neq 0$), while the constraint function captures the interplay between deterministic evolution and stochastic fluctuations without requiring external assumptions about noise statistics.

Addressing the variable truncation problem. The contact geometry paradigm offers a fundamentally different approach to high-dimensional systems, resolving the variable truncation problem that plagues traditional methods. Rather than tracking each variable individually, which becomes computationally infeasible for complex systems, the new paradigm works with the joint probability distribution $P(y)$ over all system variables. This distribution encapsulates the statistical behaviour of the entire system, including correlations and dependencies among all variables, without requiring explicit tracking of each degree of freedom.

The probability distribution $P(y)$ and its derivatives are estimated directly from observational data using kernel density estimation or

other nonparametric methods. Each data point contributes to constructing this distribution, so no observational information is wasted: every available observation is used to build the geometric structure. This stands in stark contrast to traditional approaches that use only a small fraction of available data for model calibration, discarding the vast majority of potentially useful information.

Once the probability distribution and its evolution are characterised through the contact structure, the paradigm can generate new data samples that are statistically consistent with the observed system behaviour. The probability distribution serves as both the input (derived from data) and the output (generating new data), creating a closed loop that enables predictive modelling and scenario analysis. This data-generation capability is particularly valuable for risk assessment and decision-making under uncertainty, where the ability to generate realistic scenarios is essential.

The local connection ϑ encodes all dependencies among variables via derivatives of the probability distribution. Complex behaviours such as turbulence, phase transitions, or threshold effects emerge naturally from the geometry rather than being imposed through external parameterisations. This means the model does not need to know a priori which physical mechanisms are important: the data itself reveals the structure of the dynamics through the geometry of the probability distribution.

In essence, the new paradigm replaces the traditional workflow of selecting key variables, writing equations for them, and parameterising everything else with a unified approach that collects all data, builds the probability distribution, constructs the geometric structure, and derives the dynamics from first principles. This shift enables the full utilisation of available data and captures the full complexity of system behaviour without artificial truncation or ad hoc parameterisations.

The Hi-Fi (high-fidelity) model paradigm. The contact geometry framework naturally embodies what we term the “Hi-Fi model paradigm”—a modelling philosophy that grounds all dynamical descriptions directly in real-world observations rather than simplified theoretical constructs. In this paradigm, the probability distribution $P(y)$ is not an assumed functional form derived from theoretical arguments but is estimated directly from observational data. The local connection ϑ encodes the system’s sensitivity to changes in state variables, quantified by the Fisher information metric derived from the observed probability distribution. The contact probability potential $\mathcal{H} = dP(y)$ represents the observed rate of probability current for configuration changes, incorporating the intrinsic stochasticity of real-world systems.

The constraint function $\varepsilon = \vartheta \dot{y} - dP(y)$ acquires a profound physical interpretation in the Hi-Fi paradigm: it geometrically encodes the instantaneous systematic deviation between deterministic model predictions (based on fixed initial and boundary conditions) and real-world observations of the system. When the accumulated prediction-reality deviation drives ε towards critical values, the smooth predictive dynamics within the current scenario become geometrically constrained, signalling the need to update the model based on new observations. The least constraint theorem mandates that the system must evolve along trajectories that continuously minimise this systematic deviation, ensuring that the model remains faithful to the observed reality.

This Hi-Fi paradigm represents a fundamental departure from traditional modelling approaches. Rather than constructing models from simplified physical assumptions and then adjusting them to fit data, the contact geometry framework builds the model directly from the data via the probability distribution and its geometric structure. The resulting dynamics are not approximations of some underlying “true” equations but are the natural consequence of the system’s observed statistical regularities. This data-driven, reality-grounded approach eliminates the topological rupture between model predictions and real-world observations that plagues traditional modelling paradigms.

1.4 Three pillars of the new paradigm

The new paradigm rests on three theoretical pillars that together provide a complete framework for understanding and modelling nonlinear dissipative stochastic systems (Zhong et al., 2025).

Pillar 1: probability evolution as core driver. Rather than treating probability and stochasticity as “additional perturbations” to a deterministic mechanical model, probability evolution becomes the fundamental source of dynamics. The two core dimensions of probability evolution, state-driven distribution changes and distribution-driven state dependencies, are explicitly encoded in the geometric structure. This perspective places probability at the centre of the theoretical framework, recognising that the evolution of probability distributions is not a secondary effect but the primary driver of system dynamics.

Pillar 2: contact geometry as mathematical foundation. The non-closed contact 1-form $\Theta = dP - \vartheta dy^i$ (where $d\Theta \neq 0$) intrinsically accommodates irreversibility and dissipation. Unlike symplectic geometry, which requires closed forms and conservative dynamics, contact geometry naturally describes systems that exchange energy and information with their environment. The non-closedness of the contact form directly encodes the irreversibility of dissipative processes, providing a geometric foundation that is fundamentally compatible with the physics of open systems.

Pillar 3: least constraint principle as variational criterion. The variational principle $\delta \int \varepsilon dt = 0$, with $\varepsilon = \vartheta \dot{y} - dP(y)$, provides a universal criterion for dissipative stochastic systems. This extends beyond the energy-conservation requirement ($\delta \int L dt = 0$) of the least-action principle, which applies only to conservative systems. The constraint function ε quantifies the deviation from ideal probability-state coupling, and the least-constraint principle selects trajectories that extremise this deviation, providing a natural generalisation of variational principles to dissipative and stochastic systems.

1.5 Application to hydrological cycle dynamics

The hydrological cycle is an exemplary application domain for demonstrating this paradigm shift. Water cycle dynamics involve multiple interacting compartments, including atmospheric moisture, surface water, soil moisture, and groundwater. These compartments exchange water through complex, nonlinear, stochastic, and highly coupled processes. The transport of water through these compartments exhibits threshold behaviour and saturation effects, in which small changes in conditions can trigger abrupt transitions in system behaviour. For example, soil infiltration capacity can change dramatically when soil moisture exceeds a critical threshold, leading to rapid shifts from absorption-dominated to runoff-dominated behaviour.

The hydrological cycle is subject to stochastic forcing from precipitation events and climate variability. Precipitation events are inherently unpredictable at fine scales, and their timing, intensity, and spatial distribution can vary enormously. Climate variability introduces additional stochasticity through long-term fluctuations in temperature, humidity, and atmospheric circulation patterns. This stochastic forcing creates a complex interplay between deterministic dynamics and random fluctuations that is difficult to capture with traditional modelling approaches.

Dissipative processes, including evaporation, infiltration, runoff, and plant transpiration, remove water from various compartments, convert it into other forms, or transport it to other locations. These processes are fundamentally irreversible and involve the exchange of energy and water with the environment. The rates of these processes depend on local conditions in complex ways that are often poorly understood and difficult to parameterise.

The hydrological cycle exhibits strong path dependence, with drought-heatwave sequences producing different outcomes from uniform precipitation with the same total amount. This path dependence arises from memory effects in the system, where the history of water storage and fluxes influences future behaviour. Groundwater depletion affects future recharge dynamics by altering the hydraulic gradients that drive subsurface flow. Vegetation responses to water stress modify future evapotranspiration patterns through changes in root structure and leaf area. These memory effects create hysteresis and non-equilibrium behaviour that traditional models struggle to capture.

Traditional hydrological models rely on phenomenological parameterisations for each process, including infiltration equations, evapotranspiration formulae, and runoff schemes. These parameterisations are typically derived from simplified theoretical arguments or fitted to limited data, and they often fail to capture the full complexity of the processes they represent. The contact geometry paradigm shows how these complex behaviours emerge naturally from the underlying probability-evolution structure, without requiring ad hoc parameterisations.

2 Theoretical foundations: the three pillars

2.1 Pillar I: jet bundle structure—embedding complete statistical information

2.1.1 The limitation of low-order moments

Traditional stochastic analysis characterises probability distributions using low-order moments such as the mean $\mu = \mathbb{E}[Y]$, the variance $\sigma^2 = \mathbb{E}[(Y - \mu)^2]$, and the skewness $\gamma_1 = \mathbb{E}[(Y - \mu)/\sigma]^3$. These moments provide useful summary statistics for simple distributions and have underpinned statistical analysis for over a century. However, for non-Gaussian, non-equilibrium systems, a complete characterisation requires information of infinite order. The moment-generating function $M(t) = \mathbb{E}[e^{tY}] = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbb{E}[Y^n]$ depends on all moments for reconstruction. Truncating at finite order discards essential information about tail behaviour, extreme events, and path-dependent properties that are crucial for understanding complex systems.

The limitations of low-order moments are particularly severe for systems with heavy-tailed distributions, multiple metastable states, or complex correlation structures. In such systems, the behaviour of extreme events, which are often the most important for risk assessment and decision-making, is determined by the tail of the distribution, which is poorly characterised by low-order moments. Similarly, systems with multiple metastable states may have multimodal or otherwise complex probability distributions that cannot be captured by a finite set of moments. The jet bundle construction addresses these limitations by encoding all moments and their spatial dependencies in a unified geometric structure.

2.1.2 Infinite-order jet bundle construction

Definition 2.1 (Stochastic Vector Bundle). Let $(\Omega, \mathcal{F}, \mathbb{P})$ denote the probability space, where Ω is the sample space, $\mathcal{F}$ is the σ -algebra, and $\mathbb{P}$ is the probability measure. A stochastic process is a measurable map $\gamma: \Omega \hookrightarrow M$ from $(\Omega, \mathcal{F})$ equipped with filtration $\{\mathcal{F}_t\}_{t \in T} \subseteq \mathcal{F}$ to a manifold M of dimension n . If a measurable section $Y: M \rightarrow E$ satisfies $\pi \circ Y = Id_M$, we define a stochastic vector bundle $\pi: E \rightarrow M$.

The space of probability measures $\mathbb{P}$ of stochastic vector bundles possesses an infinite-order jet structure that encodes all statistical information about the system. Let $\pi_{E,0}^{\infty}: J^{\infty}(E, \mathbb{P}) \rightarrow E$ be the infinite-order jet bundle. For $y \in E$ representing realizations of the stochastic variable Y , the local coordinate functions are $(y^i, P, P_{\mu_1}, P_{\mu_1\mu_2}, \dots, P_{\mu_1\cdots\mu_k}, \dots)$ where $y^i(j^{\infty}P) = y^i$ are the state coordinates, $P(j^{\infty}P) = P(y)$ is the probability density, $P_{\mu_1}(j^{\infty}P) = \partial_{\mu_1}P(y)$ are the first derivatives, $P_{\mu_1\mu_2}(j^{\infty}P) = \partial_{\mu_1\mu_2}P(y)$ are the second derivatives, and so on for higher-order derivatives.

This construction transforms the infinite-dimensional probability space into a differentiable manifold with an explicit geometric structure. The key insight is that the mean, variance, and all higher-order cumulants are encoded in these coordinates; no statistical information is lost through truncation, and the geometric structure enables the direct extraction of dynamical information from data. The jet bundle provides a rigorous mathematical framework for working with infinite-dimensional spaces of probability distributions, allowing us to apply the tools of differential geometry to problems in probability and statistics.

2.1.3 From data to jet coordinates

Given observational data $\{y^{(1)}, y^{(2)}, \dots, y^{(N)}\}$, we construct the empirical probability distribution $\hat{P}(y) = \frac{1}{N} \sum_{i=1}^N K_h(y - y^{(i)})$ where K_h is a kernel with bandwidth h . The jet coordinates are then obtained by direct differentiation: $\hat{P}_{\mu_1\cdots\mu_k}(y) = \frac{1}{N} \sum_{i=1}^N \partial_{\mu_1\cdots\mu_k} K_h(y - y^{(i)})$. This data-driven approach completely eliminates the need for prior model assumptions, allowing the data itself to reveal the structure of the

probability distribution and its derivatives.

The choice of kernel and bandwidth is important for obtaining accurate estimates of the probability distribution and its derivatives. The bandwidth h controls the trade-off between bias and variance in the density estimate, with smaller bandwidths yielding more detailed but noisier estimates and larger bandwidths yielding smoother but potentially biased estimates. Cross-validation or other data-driven methods can be used to select an optimal bandwidth that balances these considerations. Once the probability distribution and its derivatives are estimated, they can be used to construct the jet bundle coordinates and proceed with the geometric analysis.

2.2 Pillar II: non-closed contact structure—natural framework for dissipation

2.2.1 Why symplectic geometry fails

Symplectic geometry is founded on a closed, non-degenerate 2-form $\omega \in \Omega^2(M)$ satisfying $d\omega = 0$. This closedness is equivalent to energy conservation through Liouville's theorem $\mathcal{L}_{X_H}\omega = 0 \Rightarrow \frac{d}{dt}(dq^i \wedge dp_i) = 0$. For dissipative systems, energy is not conserved, and attempting to force dissipation into the symplectic framework requires breaking the fundamental geometric structure. One approach is to use time-dependent Hamiltonians of the form $H(q, p, t) = H_0(q, p)e^{-\alpha t}$ where α is a decay constant, but this introduces explicit time dependence that complicates the analysis and breaks time-translation invariance. Another approach is to use non-canonical brackets that modify the Poisson structure to incorporate dissipative terms, but this requires ad hoc assumptions about their form. A third approach is to double the phase space dimension by introducing auxiliary variables, but this increases computational cost and introduces additional parameters that must be determined. All these approaches break the fundamental geometric structure that makes symplectic geometry powerful for conservative systems.

The failure of symplectic geometry for dissipative systems is not merely a technical inconvenience but reflects a fundamental mismatch between the mathematical structure and the physical reality of open systems. Dissipative systems exchange energy and entropy with their environment, and this exchange is inherently irreversible. The symplectic framework, with its emphasis on conservation laws and reversible dynamics, provides no natural way to describe this irreversibility. Contact geometry addresses this limitation by providing a geometric framework that naturally accommodates irreversibility through its non-closed structure.

2.2.2 The canonical contact 1-form

Definition 2.2 (contact structure). A contact structure on a $(2n+1)$ -dimensional manifold $\mathcal{E}$ is a maximally non-integrable hyperplane field $\xi \subset T\mathcal{E}$ locally defined as the kernel of a 1-form Θ satisfying $\Theta \wedge (d\Theta)^n \neq 0$. Unlike symplectic forms, contact forms are not closed: $d\Theta \neq 0$.

On the stochastic jet bundle $J^\infty(E, P)$, the canonical contact 1-form is $\Theta = dP - \vartheta dy^i$. The components have precise physical interpretations. The term dP denotes the total differential of probability, describing the evolution of the probability distribution driven by changes in the system state. When the system state changes from y to $y + dy$, the probability changes by $dP = \frac{\partial P}{\partial y^i} dy^i$. The term ϑdy^i represents the intrinsic dependency of probability on system states. The local connection ϑ_i encodes how probability gradients, curvatures, and all higher-order dependencies influence the system through the relation $\vartheta_i = \sum_{k=1}^{\infty} \sum_{\mu_1 < \dots < \mu_k} B_{\mu_1 \dots \mu_k}^i P_{\mu_1 \dots \mu_k}$.

The contact form Θ encodes the fundamental relationship between probability evolution and state changes. The condition $\Theta = 0$ defines a constraint subspace in which probability changes are directly coupled to state changes via the local connection. This coupling is the geometric manifestation of how system dynamics influence probability evolution and vice versa. The non-closedness of Θ ($d\Theta \neq 0$) is the key feature that allows contact geometry to describe irreversible processes, as the failure of Θ to be closed directly encodes the non-conservation of probability fluxes that characterises dissipative systems.

Lemma 2.3 (contact structure on stochastic jet bundles). Let E be a vector bundle over an n -dimensional manifold M , and $J^\infty(E, P)$ be the infinite-order stochastic jet bundle. Given local coordinates (P, y^i, ϑ_i) where ϑ_i encodes all higher-order probability derivatives with coefficients B_i^j satisfying injectivity and surjectivity conditions, the 1-form $\Theta = dP - \vartheta_i dy^i$ defines a contact structure on $J^\infty(E, P)$.

Proof. We verify the contact condition $\Theta \wedge (d\Theta)^n \neq 0$. First, $d\Theta = -d\vartheta_i \wedge dy^i$. Then $\Theta \wedge (d\Theta)^n = (dP - \vartheta_i dy^i) \wedge (-d\vartheta_j \wedge dy^j)^n = (-1)^n dP \wedge d\vartheta_1 \wedge dy^1 \wedge \dots \wedge d\vartheta_n \wedge dy^n \neq 0$. The differentials $dP, dy^i, d\vartheta_i$ are linearly independent by construction, ensuring the contact condition holds.

2.2.3 Tangent space decomposition and path dependence

The contact structure induces a fundamental decomposition of the tangent space $T\mathcal{E} = \ker(\Theta) \oplus \ker(d\Theta)$. This decomposition has profound physical significance that reveals the geometric structure of probability evolution.

The constraint subspace $\ker(\Theta) = \{X \in T\mathcal{E} : \iota_X \Theta = 0\}$ consists of vectors satisfying $dP(X) = \vartheta_i dy^i(X)$, meaning probability changes are directly coupled to state changes through the local connection. This subspace corresponds to probability evolution driven by system states, capturing how changes in the system configuration induce changes in the probability distribution.

The invariant subspace $\ker(d\Theta) = \{X \in T\mathcal{E} : \iota_X d\Theta = 0\}$ captures intrinsic dependency changes of probability on the system, describing how the relationship between probability and states evolves independently of immediate state changes. This subspace encodes the geometric structure of the probability distribution, independent of its temporal evolution.

The defining characteristic of non-integrable systems is $[\ker(\Theta), \ker(\Theta)] \not\subset \ker(\Theta)$, meaning the Lie bracket of two vector fields in $[\ker(\Theta)]$ does not remain in $[\ker(\Theta)]$. Geometrically, this implies that no global foliation of $\mathcal{E}$ by integral submanifolds of $\ker(\Theta)$ exists, different paths between the same endpoints yield different final states, and the evolution is fundamentally history-dependent. For stochastic systems, this means that even starting from identical initial probability distributions $P(y_A)$, following different intermediate

paths leads to different final distributions $P(y_c) \neq P(y_c')$. This path dependence is a hallmark of complex systems and is essential for understanding memory effects, hysteresis, and non-equilibrium behaviour.

2.3 Pillar III: least constraint theorem—universal variational principle

2.3.1 From least action to least constraint

For conservative systems, Hamilton's least action principle states that $\delta \int_{t_1}^{t_2} L(q, \dot{q}, t) dt = 0$ which implies the Euler-Lagrange equations $\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0$. This principle strictly requires energy conservation $\frac{dH}{dt} = 0$, phase space volume preservation $\mathcal{L}_{X_H} \omega = 0$, and time-reversibility.

The principle of least action and the associated Hamiltonian equations stand as the foundational theoretical starting point for deriving a vast array of classical physical equations across disciplines. In fluid mechanics, for instance, the Euler equations governing the motion of inviscid fluids are directly derived from Hamiltonian mechanics by imposing the least action principle on the fluid's Lagrangian density, which encodes the kinetic and potential energy of the fluid flow field. Even the Navier-Stokes equations, the cornerstone of classical fluid dynamics describing viscous fluid motion, retain their Hamiltonian core with phenomenological viscous terms appended to account for dissipation, an ad hoc modification that mirrors the broader limitation of symplectic geometry in accommodating non-conservative effects. Beyond fluid mechanics, this foundational role extends to classical electrodynamics, where Maxwell's equations can be formulated within a Hamiltonian framework, and to continuum mechanics, where the dynamics of elastic solids are derived from the least action principle applied to the material's Lagrangian. In all these cases, Hamiltonian mechanics and the principle of least action serve as the unifying starting point, yet their inherent reliance on energy conservation and symplectic geometric structures forces researchers to introduce external, phenomenological corrections when describing dissipative, stochastic real-world systems.

None of these holds for dissipative, stochastic systems, and the least constraint theorem provides the appropriate generalisation. The least action principle selects trajectories that extremize the action integral, which is proportional to the integral of the Lagrangian over time. For conservative systems, this is equivalent to extremizing the integral of the kinetic minus potential energy, yielding Newton's equations of motion. However, in dissipative systems, energy is not conserved, and the action principle must be modified accordingly. The least constraint theorem provides a variational principle applicable to dissipative and stochastic systems by extremizing a constraint function rather than an energy-based action.

2.3.2 The constraint function

Definition 2.4 (Constraint Function). The constraint function $\varepsilon: \mathcal{E} \rightarrow \mathbb{R}$ is defined as $\varepsilon = \oint \dot{y}^j - dP(\dot{y}) = \oint \frac{\partial \mathcal{H}}{\partial \dot{y}^j} - \mathcal{H}$ where $\mathcal{H} = dP(\dot{y})$ is the contact probability potential.

The constraint function quantifies the deviation from ideal probability-state coupling and plays a central role in the Hi-Fi model paradigm. The term $\oint \dot{y}^j$ represents the work done by the local connection along the trajectory, encoding how the system's sensitivity to parameter changes (quantified by the Fisher information metric) couples with the rate of state evolution. The term $dP(\dot{y})$ denotes the rate of change of probability along the trajectory, reflecting how the observed probability distribution evolves as the system state changes.

When $\varepsilon = 0$, there is a perfect balance between connection work and probability evolution, indicating that the model's predictions are in complete agreement with the observed statistical behaviour of the system. When $\varepsilon > 0$, there is excess connection work relative to probability change, signalling that the model is predicting more sensitivity than is observed in reality. When $\varepsilon < 0$, probability evolution exceeds connection work, indicating that the observed system is more variable than the model predicts.

In the Hi-Fi paradigm, the constraint function serves as the core indicator of topological discrepancy between model predictions and real-world observations. It provides a quantitative measure of how far the current model deviates from the observed reality, guiding continuous refinement to maintain fidelity to the data. The least constraint principle requires that the system evolve along trajectories that continuously minimise this systematic deviation, ensuring that the model remains faithful to observed reality at all times.

2.3.3 The least constraint theorem

Theorem 2.5 (Least Constraint Theorem for Stochastic Vector Bundles). Let E be an n -dimensional vector space and $\mathbb{P}$ be the probability measure space. The infinite-order stochastic jet bundle $\pi_{E,0}^\infty: J^\infty(E, \mathbb{P}) \rightarrow E$ exhibits:

(i) **Contact structure:** A natural contact manifold $(\mathcal{E}, \Theta)$ with $\Theta = dP - \oint \dot{y}^j = \mathcal{H} dt - \oint \dot{y}^j$, where $\Theta \wedge (d\Theta)^n \neq 0$.

(ii) **Contact-preserving flow:** A unique vector field $X_{\mathcal{H}} \in \ker(d\Theta)$ satisfying $\iota_{X_{\mathcal{H}}} \Theta = -\varepsilon$, $\iota_{X_{\mathcal{H}}} d\Theta = 0$, and $\mathcal{L}_{X_{\mathcal{H}}} \Theta = 0$.

(iii) **Variational principle:** The flow generated by $X_{\mathcal{H}}$ extremizes the constraint action $\delta \oint \varepsilon dt = \delta \left(- \int \Theta \right) = 0$.

Proof. Part (i) follows from Lemma 2.3. For Part (ii), the conditions $\iota_{X_{\mathcal{H}}} d\Theta = 0$ and $\mathcal{L}_{X_{\mathcal{H}}} \Theta = 0$ are equivalent via Cartan's magic formula $\mathcal{L}_{X_{\mathcal{H}}} \Theta = d(\iota_{X_{\mathcal{H}}} \Theta) + \iota_{X_{\mathcal{H}}} d\Theta = 0$. Setting $\iota_{X_{\mathcal{H}}} \Theta = -\varepsilon$ ensures both conditions are satisfied. Uniqueness follows from the direct sum decomposition $T\mathcal{E} = \ker(\Theta) \oplus \ker(d\Theta)$ and $\ker(\Theta) \cap \ker(d\Theta) = \{0\}$. For Part (iii), the variation of the action is

$$\delta \mathcal{S} = \delta \int (\oint \dot{y}^j - \mathcal{H}) dt = \int \left(\delta \oint \dot{y}^j + \oint \delta \dot{y}^j - \frac{\partial \mathcal{H}}{\partial y^j} \delta y^j - \frac{\partial \mathcal{H}}{\partial \dot{y}^j} \delta \dot{y}^j \right) dt \quad (1)$$

Using $\dot{y}^j = \frac{\partial \mathcal{H}}{\partial \dot{y}^j}$ from the contact structure and integrating by terms, $\delta \mathcal{S} = 0$ yields the contact dynamical equations.

2.3.4 Contact dynamical equations

From the least constraint theorem, we derive the unified contact dynamical equations:

$$\begin{cases} \frac{dy^i}{dt} = \{y^i, \mathcal{H}\} = \frac{\partial \mathcal{H}}{\partial \vartheta_i} \\ \frac{d\vartheta_i}{dt} = \{\vartheta_i, \mathcal{H}\} = -\frac{\partial \mathcal{H}}{\partial y^i} \\ \frac{d\varepsilon}{dt} = -\{\varepsilon, \mathcal{H}\} \end{cases} \quad (2)$$

where the Poisson bracket on the contact manifold is

$$\{f, g\} = \frac{\partial f}{\partial y^i} \frac{\partial g}{\partial \vartheta_i} - \frac{\partial f}{\partial \vartheta_i} \frac{\partial g}{\partial y^i} \quad (3)$$

These equations reveal that system states y^i evolve according to probability gradients ϑ_i , probability gradients ϑ_i evolve according to state gradients of the probability potential, and the constraint function ε evolves to maintain the contact structure. The complete coupling of the system state, probability gradient, and constraint deviation yields a unified dynamical description without switching between model frameworks for different dissipation strengths or nonlinear characteristics. This unified description is a key advantage of the contact geometry approach, as it provides a single mathematical framework that can handle a wide range of physical situations without requiring case-specific modifications.

3 Core capability I: complete system reconstruction from data

3.1 The traditional approach: model-first paradigm

Traditional reconstruction follows a well-established workflow that has been the corner-stone of scientific modelling for centuries. Researchers begin with physical intuition, postulating governing equations based on analogy with simpler systems or fundamental physical principles. These equations are then augmented with phenomenological terms as needed to match observed behaviour. The coefficients of these equations are adjusted through parameter fitting to match available data. The model predictions are validated against observations, and if discrepancies persist, the model is refined by adding additional terms or modifying existing ones.

This traditional approach suffers from several fundamental problems that limit its effectiveness for complex systems. First, the results are only as good as the initial assumptions, and incorrect assumptions can lead to systematically wrong predictions even if the model fits available data well. Second, each new phenomenon requires new parameters, leading to parameter proliferation that makes models increasingly complex and difficult to constrain. Third, multiple models can fit the same data equally well, leading to non-uniqueness that makes it difficult to determine which model is correct. Fourth, path dependence cannot be captured by local fitting, so models may fail to capture important dynamical features that depend on the system's history.

3.2 The new approach: data-first paradigm

The contact geometry paradigm enables a fundamentally different workflow that addresses the limitations of the traditional approach. The process begins with data collection, gathering observations of system states across all available variables and scales. Next, the probability distribution $P(y)$ and all its derivatives are constructed from the data using kernel density estimation or other nonparametric methods. The infinite-order jet bundle $J^\infty(E, P)$ coordinates are then built from these probability distributions. The local connection ϑ_i is computed, and the contact structure Θ is formed. Finally, the dynamics are derived by applying the least constraint theorem.

This approach offers several key advantages over the traditional model-first paradigm. No prior model assumptions are required, as the dynamics emerge directly from the data structure. Dissipation is intrinsic to the geometric structure, so no phenomenological terms need to be added. Complete statistics are preserved, with all-order information encoded in the jet bundle structure. Path dependence is captured geometrically through the non-integrability of the contact structure, enabling the model to capture history-dependent behaviour that traditional approaches miss.

3.3 Reconstruction procedure

The reconstruction procedure consists of six steps that transform raw data into dynamical equations. In Step 1, given observational data $D = \{y^{(1)}, \dots, y^{(N)}\}$, the probability density is estimated as $\hat{P}(y) = \frac{1}{N} \sum_{i=1}^N K_h(y - y^{(i)})$ where K_h is a kernel function with bandwidth h selected by cross-validation. In Step 2, spatial derivatives to arbitrary order are computed as $\hat{P}_{\mu_1 \dots \mu_k}(y) = \frac{1}{N} \sum_{i=1}^N \partial_{\mu_1 \dots \mu_k} K_h(y - y^{(i)})$. For numerical stability, spectral methods or other regularization can be employed. In Step 3, the local connection ϑ_i is computed as a linear combination of all-order derivatives $\vartheta_i(y) = \sum_{k=1}^K \sum_{\mu_1 < \dots < \mu_k} B_{\mu_1 \dots \mu_k}^i(y) \hat{P}_{\mu_1 \dots \mu_k}(y)$ where the coefficients $B_{\mu_1 \dots \mu_k}^i$ are determined from the cumulant structure of the data. In Step 4, the contact 1-form is constructed as $\Theta = d\hat{P} - \vartheta_i dy^i$ and the constraint function is $\varepsilon = \vartheta_i \dot{y}^i - \frac{\partial \hat{P}}{\partial y^i} \dot{y}^i$. In Step 5, the contact probability potential $\mathcal{H} = dP(y)$ is determined by solving $\frac{d\mathcal{H}}{ds} = \mathcal{H} + \varepsilon$ where s parameterizes the characteristic curves. In Step 6, the complete dynamics are given by the contact dynamical equations.

3.4 Intrinsic vs. phenomenological dissipation

The traditional phenomenological approach adds terms to the equations of motion to account for dissipation, treating these terms as external modifications to an otherwise conservative system. In contrast, the contact geometry approach derives dissipation intrinsically

from the geometric structure. The equations $\frac{dy^i}{dt} = \frac{\partial \mathcal{H}}{\partial \mathcal{G}_i}$ and $\frac{d\mathcal{G}_i}{dt} = -\frac{\partial \mathcal{H}}{\partial y^i}$ describe a system where dissipation emerges from the non-closedness of Θ through the constraint function ε . The key difference is that in the traditional approach, dissipation is imposed from outside the geometric structure, while in the contact paradigm, dissipation is intrinsic to the geometry itself.

4 Core capability II: prediction based on global invariants

4.1 The prediction problem for non-integrable systems

Traditional long-term prediction for non-integrable systems fails due to exponential divergence characterised by Lyapunov instability, where $\delta y(t) \sim \delta y(0)e^{\lambda t}$, model error accumulation where small discrepancies grow exponentially over time, and noise amplification where stochastic terms dominate at long times. The root cause is that traditional methods lack global invariants for dissipative, non-integrable systems, making it impossible to constrain predictions over long time scales.

4.2 Contact structure invariance

The contact geometry paradigm provides a fundamental invariant that enables stable long-term prediction. Along the flow generated by $X_{\mathcal{H}}$, the contact structure is preserved: $\mathcal{L}_{X_{\mathcal{H}}} \Theta = 0$. This invariance is more universal than Liouville's theorem, holding even for dissipative, stochastic, non-integrable systems. The proof follows from Theorem 2.5, which states that $X_{\mathcal{H}} \in \ker(d\Theta)$ and $\iota_{X_{\mathcal{H}}} \Theta = -\varepsilon$. By Cartan's formula, $\mathcal{L}_{X_{\mathcal{H}}} \Theta = d(\iota_{X_{\mathcal{H}}} \Theta) + \iota_{X_{\mathcal{H}}} d\Theta = d(-\varepsilon) + 0 = -d\varepsilon$. Since $d\varepsilon = 0$ along the flow (because ε is conserved), we have $\mathcal{L}_{X_{\mathcal{H}}} \Theta = 0$.

4.3 Bidirectional closed-loop prediction

The contact dynamical equations enable bidirectional prediction that couples probability and state evolution. Forward prediction proceeds from probability to state: $P(y, t) \xrightarrow{\text{compute } \mathcal{G}_i} \mathcal{H} \xrightarrow{\text{Eq. (2)}} y(t + \Delta t), P(y, t + \Delta t)$. Reverse prediction proceeds from state to probability: Expected $y(t + \Delta t) \xrightarrow{\text{invert dynamics}} \text{Required } P(y, t), \mathcal{G}_i$. This bidirectional capability enables state forecasting from current probability distributions, probability inference from expected state changes, and optimal control through probability gradient manipulation.

4.4 Long-term stability

The global invariants ensure long-term stability of predictions. For non-integrable dissipative stochastic systems, the contact geometry framework provides stable long-term prediction through three mechanisms. First, the geometric constraint $\mathcal{L}_{X_{\mathcal{H}}} \Theta = 0$ prevents structural divergence by ensuring that the contact structure is preserved along trajectories. Second, constraint conservation $\frac{d\varepsilon}{dt} = 0$ along trajectories ensures that the system evolves in a way that maintains the balance between connection work and probability evolution. Third, path weighting ensures that all possible paths are predicted with their probability weights, providing a complete probabilistic description of future states rather than a single deterministic prediction.

5 Example: water cycle dynamics

5.1 Why the water cycle?

The water cycle serves as an ideal testbed for demonstrating the paradigm's ability to address variable truncation and data underutilization. A comprehensive hydrological model would need to track atmospheric moisture at every point in a 3D grid involving millions of variables, soil moisture across heterogeneous landscapes with varying soil types, groundwater flow through complex geological formations, vegetation states affecting evapotranspiration, river network dynamics with channel geometry variations, and snow and ice processes in cold regions.

Traditional hydrological models address this complexity by selecting only a few state variables (typically 3–5 storage compartments), representing all other processes through empirical parameterisations, and using only a tiny fraction of available observational data for calibration. For example, a typical watershed model might use only daily precipitation and streamflow data, discarding the vast majority of information contained in high-resolution soil moisture measurements from sensor networks, groundwater level records from monitoring wells, evapotranspiration estimates from eddy covariance towers, remote sensing data on vegetation and surface conditions, and atmospheric moisture profiles from weather stations.

The contact geometry paradigm fundamentally changes this approach. All data contribute to the probability structure, with soil moisture sensors, groundwater wells, stream gauges, and weather stations all contributing to construct the joint probability distribution $P(W_a, W_s, W_m, W_g, \dots)$. No artificial variable selection is required, as the probability distribution naturally captures dependencies among all measured variables. Emergent dynamics arise from the contact structure $\Theta = dP - \mathcal{G}_i dW^i$, which encodes how the probability distribution evolves, generating the dynamics of all variables without phenomenological infiltration equations or evapotranspiration formulas. Data generation capability enables the paradigm to generate new hydrological scenarios that are statistically consistent with observed behaviour, enabling predictive modelling and risk assessment.

The water cycle exemplifies all features of nonlinear dissipative stochastic systems. Nonlinearity appears in threshold behaviours and saturation effects. Dissipation appears in evaporation, infiltration, and runoff. Stochasticity appears in precipitation events and climate variability. Path dependence appears in drought-heatwave sequences. Memory effects appear in groundwater depletion and soil moisture dynamics. Non-equilibrium behavior appears in transient states and seasonal cycles.

Traditional hydrological models require extensive phenomenological parameterisations. The contact-geometry paradigm reveals how these behaviours emerge from the underlying evolution of probability.

5.2 Hydrological system as stochastic vector bundle

Hi-Fi hydrological modelling. The application of the Hi-Fi paradigm to hydrological cycle dynamics exemplifies how the contact geometry framework enables a fundamental shift from model-driven to data-driven water resource management. In traditional hydrological modelling, the governing equations (such as the Richards equation for soil moisture or the Saint-Venant equations for channel flow) are postulated based on physical principles and then calibrated against limited observational data. This approach inevitably introduces discrepancies between model predictions and real-world behaviour due to simplifications in the governing equations and uncertainties in parameter estimation.

The Hi-Fi hydrological paradigm takes a radically different approach. The probability distribution $P(W_a, W_s, W_m, W_g, \dots)$ is estimated directly from all available observational data, including precipitation gauges, streamflow stations, soil moisture sensors, groundwater wells, and remote sensing products. This probability distribution encapsulates the complete statistical behaviour of the hydrological system as observed in reality, without the distortions introduced by simplified model assumptions. The local connection $\mathcal{J}_i$ then encodes the sensitivity of the system's water storage to changes in each compartment, derived from the spatial derivatives of the observed probability distribution.

The contact probability potential $\mathcal{H} = dP(y)$ represents the observed rate of change of the hydrological state probability, incorporating the intrinsic stochasticity of precipitation events, evapotranspiration fluctuations, and runoff variability. The constraint function $\varepsilon = \mathcal{J}_i \dot{W}^i - dP(y)$ quantifies the systematic deviation between the predictions of any hydrological model and the observed behaviour of the real watershed. The least constraint principle ensures that the Hi-Fi hydrological model evolves in a way that continuously minimises this deviation, maintaining fidelity to the observed data.

This Hi-Fi approach eliminates the need for phenomenological infiltration equations, empirical evapotranspiration formulas, or calibrated runoff coefficients. Instead, all hydrological processes emerge naturally from the geometric structure of the observed probability distribution. The resulting model is not an approximation to some underlying “true” hydrological equations, but is the natural consequence of the observed statistical regularities of the watershed.

The state vector $y = (y_1, y_2, y_3, y_4) \in E$ is defined with $y_1 = W_a$ (atmospheric moisture content in mm), $y_2 = W_s$ (surface water storage in mm), $y_3 = W_m$ (soil moisture content in mm), and $y_4 = W_g$ (groundwater storage in mm). The base manifold M represents spatial coordinates (x, y, z) and topographic features. The probability distribution $P(y, t)$ describes the likelihood of finding the system in state y at time t , given climate variability and measurement uncertainty. The probability evolution follows

$$\frac{\partial P(y, t)}{\partial t} = \mathcal{L}P(y, t) \quad (4)$$

where the evolution operator

$$\mathcal{L} = -D_i \partial_i + \frac{1}{2} D_{ij} \partial_{ij} - \frac{1}{6} D_{ijk} \partial_{ijk} + \dots \quad (5)$$

encodes the dynamics. The coefficients $D_i, D_{ij}, D_{ijk}, \dots$ are functionals of cumulants of water fluxes including precipitation, evapotranspiration, and runoff.

5.3 Contact structure for water cycle

For the hydrological system, the contact 1-form is $\Theta = dP - \mathcal{J}_a dW_a - \mathcal{J}_s dW_s - \mathcal{J}_m dW_m - \mathcal{J}_g dW_g$. The dP term represents probability changes driven by water state transitions, such as precipitation increasing atmospheric moisture ($dP > 0$ for W_a increase), infiltration transferring water to soil (dP redistributes between W_s and W_m) and groundwater recharge (dP shifts from W_m to W_g). The $\mathcal{J}_i dy^i$ term represents intrinsic dependencies, with $\mathcal{J}_a$ describing how atmospheric moisture probability responds to moisture content, $\mathcal{J}_m$ describing soil moisture dependency, including saturation effects, and $\mathcal{J}_g$ describing groundwater response to storage levels.

The local connection encodes threshold behaviors through $\mathcal{J}_m = \sum_{k=1}^{\infty} B_m^k \frac{\partial^k P}{\partial W_m^k}$. Near field capacity W_m^c , higher derivatives dominate, capturing the rapid transition from infiltration- dominated to runoff- dominated behaviour.

5.4 Constraint function and water balance

The constraint function for the hydrological system is $\varepsilon = \mathcal{J}_i \dot{W}^i - \frac{\partial P}{\partial W^i} \dot{W}^i$ where the first term represents connection work (work done by atmospheric moisture gradients during evaporation, work done by soil moisture gradients during plant uptake, and work done by groundwater gradients during baseflow) and the second term represents the probability rate (rate of probability change due to water state transitions, capturing how likely the system is to transition to a new state). The least constraint principle $\delta \int \varepsilon dt = 0$ implies the hydrological system evolves to extremize the imbalance between connection work and probability evolution. This yields optimal water allocation between compartments, a balance between energy expenditure and state transition likelihood, and the emergence of preferred flow paths.

5.5 Contact dynamical equations for water cycle

The complete dynamics are governed by $\frac{dW_a}{dt} = P - E - Q_{a \rightarrow s}$, $\frac{dW_s}{dt} = Q_{a \rightarrow s} - I - R$, $\frac{dW_m}{dt} = I - E_t - Q_{m \rightarrow g}$, and $\frac{dW_g}{dt} = Q_{m \rightarrow g} - B$ where

P is precipitation, E is evaporation, E_t is transpiration, I is infiltration, R is surface runoff, Q are inter-compartment fluxes, and B is baseflow. The probability gradients evolve as $\frac{d\varphi_i}{dt} = -\frac{\partial \mathcal{H}}{\partial W_i}$, capturing how the relationship between water storage and probability changes over time.

5.6 Non-equilibrium steady states

Non-equilibrium steady states correspond to frozen fluctuation states where $\frac{dP}{dt} = 0$ and $\varphi_i = 0$. In hydrological terms, these represent conditions in which water storage in each compartment is statistically stationary, the probability distribution of water states no longer evolves, water fluxes between compartments are balanced on average, and the system has forgotten its initial transient history.

Through the extremum conditions of ε , we can predict steady state existence by determining under what climate conditions steady states exist, predict relaxation times by estimating how long after a perturbation before a steady state is reached, and regulate steady states by adjusting probability gradients φ_i to achieve desired water allocations.

5.7 Path dependence and climate memory

The non-integrability $[\ker(\Theta), \ker(\Theta)] \not\subset \ker(\Theta)$ manifests in hydrology through drought-heatwave sequences where Path A (Drought $\rightarrow$ Recovery $\rightarrow$ Normal) and Path B (Drought $\rightarrow$ Heatwave $\rightarrow$ Compound event) with the same total precipitation lead to different soil moisture outcomes. Groundwater memory manifests as multi-year depletion affecting recharge dynamics, altered subsurface flow paths persisting, and recovery requiring precipitation patterns different from those for maintenance. Vegetation feedback manifests as water-stress responses that modify future evapotranspiration, root-zone changes that alter infiltration capacity, and ecosystem shifts that alter water partitioning.

The contact geometry framework captures these effects intrinsically, with no ad hoc parameterisations needed, memory effects arising from the noncommutative structure, and path dependence encoded in the jet bundle coordinates.

6 Comparison: traditional vs. contact geometry paradigms

6.1 Fundamental differences

Table 1 summarises the essential differences between paradigms. The traditional paradigm follows a core logic of Model $\rightarrow$ Statistics $\rightarrow$ Data, uses symplectic geometry with closed 2-form ω only for conservative systems, employs the least action principle $\delta \int L dt = 0$ requiring energy conservation, is driven by force/energy with probability as perturbation, applies to deterministic conservative integrable systems, adds dissipation phenomenologically, imposes noise externally, uses only low-order moments, cannot capture path dependence, requires prior model assumptions for reconstruction, diverges in long-term prediction due to lack of invariants, and handles constraints through external Lagrange multipliers.

The contact geometry paradigm follows a core logic of Data $\rightarrow$ Probability $\rightarrow$ Geometry $\rightarrow$ Dynamics, uses contact geometry with non-closed 1-form Θ naturally for dissipative irreversible systems, employs the least constraint principle $\delta \int \varepsilon dt = 0$ universal for dissipative stochastic systems, is driven by probability evolution dual-dimension coupling with stochasticity as intrinsic property, applies to nonlinear dissipative stochastic non-integrable systems, encodes dissipation intrinsically in $d\Theta \neq 0$, has noise emerging from local connection structure, uses all-order information via jet bundle, intrinsically captures path dependence via $[\ker\Theta, \ker\Theta] \not\subset \ker\Theta$, is prior-free with data directly determining geometry, achieves stable long-term prediction via $\mathcal{L}_{X_{\mathcal{H}}} \Theta = 0$, and handles constraints through intrinsic geometric forces.

Table 1 Fundamental comparison of research paradigms.

Dimension	Traditional paradigm	Contact geometry paradigm
Core logic	Model $\rightarrow$ Statistics $\rightarrow$ Data	Data $\rightarrow$ Probability $\rightarrow$ Geometry $\rightarrow$ Dynamics
Geometric foundation	Symplectic geometry: closed 2-form ω , only for conservative systems	Contact geometry: non-closed 1-form Θ , naturally for dissipative irreversible systems
Variational principle	Least action: $\delta \int L dt = 0$ (requires energy conservation)	Least constraint: $\delta \int \varepsilon dt = 0$ (universal for dissipative stochastic systems)
Dynamics driver	Force/energy driven; probability as perturbation	Probability evolution dual-dimension coupling; stochasticity as intrinsic property
Applicable systems	Deterministic, conservative, integrable	Nonlinear, dissipative, stochastic, non-integrable
Dissipation treatment	Phenomenologically added	Intrinsically encoded in $d\Theta \neq 0$
Noise treatment	Externally imposed	Emerges from local connection structure
Statistical information	Low-order moments only (mean, variance)	All-order information via jet bundle
Path dependence	Cannot capture; linearizes	Intrinsically captured via $[\ker\Theta, \ker\Theta] \not\subset \ker\Theta$
Reconstruction	Prior model assumptions required	Prior-free; data directly determines geometry
Long-term prediction	Diverges due to lack of invariants	Stable via $\mathcal{L}_{X_{\mathcal{H}}} \Theta = 0$
Constraint handling	Lagrange multipliers (external)	Geometric forces (intrinsic)

Figure 1 illustrates the foundational epistemological and methodological paradigm shift for modelling nonlinear dissipative stochastic systems, which essentially inverts the traditional model-to-data reasoning logic to a novel data-driven data-to-probability-to-geometry-to-dynamics workflow, and fundamentally resolves the three core bottlenecks of traditional paradigms through intrinsic geometric and statistical design, rather than incremental technical modifications.

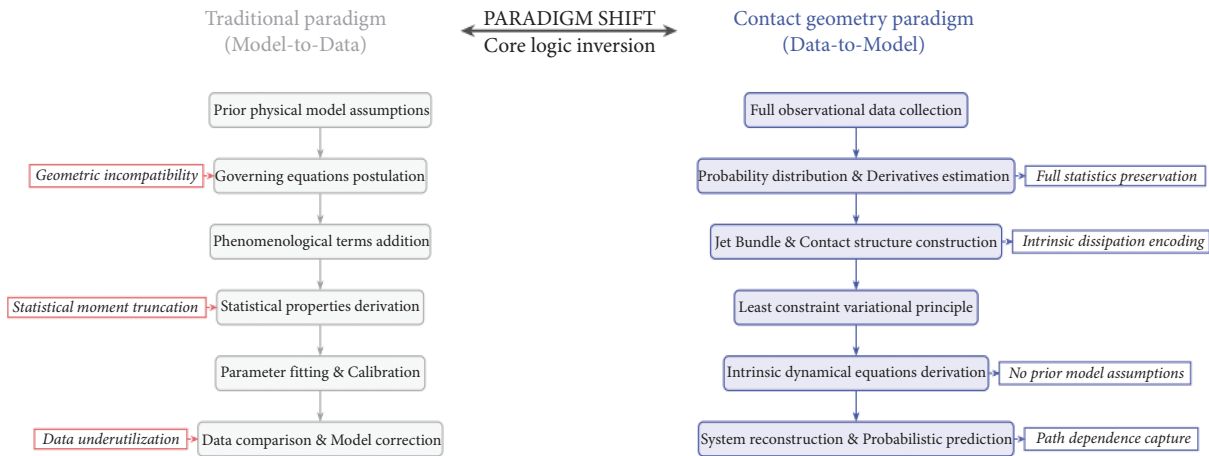


Fig. 1 Core logic inversion between the traditional paradigm and the proposed data-driven contact geometry paradigm.

The core of this paradigm shift lies in three essential departures from traditional modelling. First, it abandons the top-down, a priori model assumption-driven framework that has dominated classical dynamical system studies: the traditional paradigm starts with pre-specified physical models and governing equations, and is forced to add ad hoc phenomenological terms and fit empirical parameters to match observational data, which inevitably leads to geometric incompatibility, statistical moment truncation and massive data underutilization (marked by red bottleneck tags). In contrast, the proposed contact geometry paradigm adopts a bottom-up, data-first logic that takes full observational data as the only input, and directly constructs the system's geometric structure from the statistical regularities of data, eliminating the need for prior physical assumptions and artificial variable truncation.

Second, the shift realizes a transition from extrinsic modification to intrinsic embedding of the key characteristics of nonlinear dissipative stochastic systems. Traditional approaches treat dissipation, stochasticity and path dependence as external perturbations to conservative systems (e.g., adding noise terms to Hamiltonian equations or phenomenological dissipation terms to symplectic geometry), which breaks the geometric consistency of the model. The contact geometry paradigm, by virtue of its three theoretical pillars (infinite-order jet bundle, non-closed contact structure, and least constraint theorem), intrinsically embeds complete all-order statistical information, irreversible dissipation and path dependence into the geometric structure of the stochastic vector bundle (marked by blue advantage tags). These core properties become the inherent features of the framework rather than ad hoc add-ons, establishing a rigorous first-principles mathematical foundation for open non-equilibrium systems.

Third, the shift achieves a transition from one-way model-data validation to bidirectional probability-state closed-loop evolution. The traditional paradigm only uses data for parameter fitting and model correction after deriving statistical properties from pre-defined models, forming a one-way and incomplete interaction between model and data. The contact geometry paradigm, however, takes probability evolution as the core driver of system dynamics, and constructs a bidirectional coupling between probability distribution evolution and system state change through the contact structure. Combined with the global contact structure invariance, this closed-loop evolution enables stable long-term prediction for non-integrable systems, which is unattainable for traditional methods suffering from error accumulation and lack of global invariants.

In essence, this paradigm shift is not a trivial technical improvement of traditional models, but a fundamental change in the conceptualization of the relationship between data, probability, geometry and dynamics: it moves from imposing artificial model structures on complex systems to discovering the intrinsic geometric and dynamical structures of systems from observational data, thus providing a unified framework for the reconstruction and prediction of nonlinear dissipative stochastic systems without phenomenological parameterization.

6.2 The paradigm shift in practice

The traditional workflow begins with a priori physical assumptions about soil moisture dynamics, postulating that soil water recession follows a linear storage-discharge relationship as $\frac{d\theta}{dt} = -k(\theta - \theta_r)$, where θ is volumetric soil moisture content, θ_r is residual moisture content, and k is the empirical hydraulic conductivity coefficient. It then adds deterministic infiltration inputs $I(t)$ and stochastic noise terms $\xi(t)$ to form the complete model $\frac{d\theta}{dt} = I(t) - k(\theta - \theta_r) + \xi(t)$. Researchers next fit the empirical parameters k and θ_r , along with noise statistics, to limited observational data. When the model fails to capture threshold behaviours such as saturation-excess runoff during extreme rainfall events or hysteresis in drying-wetting cycles, they iteratively introduce additional phenomenological terms (e.g., nonlinear conductivity functions, hysteresis parameters) and expand the parameter set to improve fit.

The contact geometry workflow, by contrast, starts directly with observed time series of soil moisture, infiltration, and runoff across all monitoring sites. It first estimates the full joint probability distribution $P(\theta)$ of soil moisture states and all its higher-order derivatives $\partial_\theta^k P$

from the observational dataset. Next, it constructs the infinite-order jet bundle coordinates $(\theta, P, P_\theta, P_{\theta\theta}, \dots)$ that embed the complete statistical information of the system. It then forms the canonical contact structure $\Theta = dP - \mathcal{J}d\theta$ that encodes the intrinsic coupling between soil moisture state changes and probability evolution. Finally it derives the full system dynamics via the least constraint principle $\delta \int \varepsilon dt = 0$, and predicts all possible evolution paths of soil moisture and runoff generation with their corresponding probability weights, including extreme events and hysteretic behaviours.

The fundamental difference is that traditional approaches impose rigid model structures based on prior physical assumptions and simplified constitutive relations, whereas the contact geometry framework directly infers the intrinsic dynamical structure from the statistical regularities of observational data. This shift from model-driven to data-driven scientific modelling represents a fundamental change in how we characterise, reconstruct, and predict complex nonlinear dissipative stochastic systems.

7 Conclusions and future directions

7.1 Summary of contributions

This paper presents the probability-evolution-driven contact-geometry paradigm, a fundamental reconceptualisation of how to study nonlinear dissipative stochastic systems. The theoretical contributions include the complete reversal of the traditional “model-to-data” logic to “data-to-dynamics”, the three theoretical pillars of jet bundle structure, non-closed contact structure, and the least constraint theorem, and a unified framework that captures dissipation, stochasticity, and path dependence intrinsically.

The methodological contributions include prior-free reconstruction of system dynamics from observational data, bidirectional closed-loop prediction of probability and state evolution, and long-term stability through global geometric invariants. The applied contributions include demonstrations of hydrological cycle dynamics, the capture of path dependence, memory effects, and non-equilibrium steady states, and the elimination of phenomenological parameterisations.

7.2 Future directions

Theoretical development should focus on extensions to infinite-dimensional systems, such as field theories, quantum generalisations of contact structures, and connections to information geometry and thermodynamic geometry. Computational methods should develop data-driven algorithms for contact structure reconstruction, model reduction techniques preserving contact structure, and real-time prediction systems for high-dimensional applications. Applications should explore climate system dynamics and tipping point prediction, ecosystem modelling with path-dependent transitions, financial market dynamics with memory effects, and biological systems with non-equilibrium steady states.

Hi-Fi Model Paradigm: AI + Scenarios. A particularly promising direction for the Hi-Fi model paradigm lies in its integration with artificial intelligence and scenario generation capabilities. Once a comprehensive probability distribution $P(y)$ is constructed from observational data through the contact geometry framework, this distribution serves as the foundation for generating a wide range of realistic scenarios. Unlike traditional scenario analysis that relies on ad hoc assumptions or limited parameter sweeps, the Hi-Fi approach can systematically explore the full space of possible system behaviours encoded in the probability distribution.

By combining the contact structure with modern AI techniques such as generative models and deep learning, the paradigm can produce high-fidelity scenarios that are statistically consistent with observed reality while spanning the entire range of plausible future states. The local connection $\mathcal{J}$ and the constraint function ε provide natural guidance for scenario generation, ensuring that generated scenarios respect the system’s intrinsic geometric structure and remain physically consistent. This capability is invaluable for risk assessment, decision-making under uncertainty, and long-term planning in complex systems such as climate adaptation, water resource management, and financial risk modelling.

Furthermore, as more data become available, the probability distribution can be continuously updated, and scenario-generation capabilities can be refined in real time. This creates a dynamic, adaptive modelling framework in which the model not only describes the current state of the system but also anticipates future possibilities through data-driven, geometrically grounded scenario exploration. The integration of AI with the Hi-Fi paradigm represents a significant step toward truly intelligent, self-updating models that remain faithful to reality while exploring the full spectrum of possible futures.

Author contribution statement

Junli Wang is responsible for conceptualization, methodology, formal analysis, and writing the original draft. Tiejian Li is responsible for methodology, validation, and reviewing and editing the manuscript. Deyu Zhong is responsible for conceptualization, supervision, funding acquisition, and reviewing and editing the manuscript.

Data availability

No new experimental data were generated in this study. Any data or materials supporting the findings of this work are available from the corresponding author upon reasonable request.

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article. The authors Tiejian Li and Deyu Zhong are the members of Acting Executive Editor of this journal.

References

- Abraham, R., Marsden, J. E. (1978). Foundations of Mechanics. 2nd ed. Reading: Benjamin/Cummings Publishing Co.
- Arnold, V. I. (2010). Mathematical Methods of Classical Mechanics. 2nd ed. New York: Springer.
- Cross, M. C., Hohenberg, P. C. (1993). Pattern formation outside of equilibrium. *Reviews of Modern Physics*, 65: 851–1112
- van Kampen, N. G. (1992). Stochastic Processes in Physics and Chemistry. Amsterdam: North-Holland.
- Zhong, D. Y., Wang, G. Q., Zhang, M. X., Li, T. J. (2020). Kinetic equation for particle transport in turbulent flows. *Physics of Fluids*, 32: 073301
- Zhong, D. Y., Wang, G. Q. (2024). Kinetic equation for stochastic vector bundles. *Journal of Physics A: Mathematical and Theoretical*, 57: 225004
- Zhong, D. Y., Wang, G. Q. (2025). Least constraint and contact dynamics of stochastic vector bundles. *Chaos, Solitons & Fractals*, 201: 117143.
- Zwanzig, R. (2001). Nonequilibrium Statistical Mechanics. New York: Oxford University Press.