Applied Stochastic Models in Business and Industry

WILEY

RESEARCH ARTICLE OPEN ACCESS

Information-Driven Modeling of Energy Markets: An Unbalanced Wasserstein Barycenter Approach

Carlo Mari[1] | Emiliano Mari[2] | Cristiano Baldassari[1]

[1]DEIM – Department of Economics, Engineering, Society, Business Organization, University of Tuscia, Viterbo, Italy | [2]Sydus, Orvieto (TR), Italy

Correspondence: Cristiano Baldassari (cristiano.baldassari@unitus.it)

Received: 30 June 2025 | **Revised:** 18 February 2026 | **Accepted:** 23 February 2026

Keywords: electricity prices | gaussian mixture models | graph machine learning | natural gas prices | stochastic modeling | Wasserstein barycenter

ABSTRACT

A novel methodology is proposed for jointly modeling the price dynamics of natural gas and electricity by integrating graph-based Machine Learning and optimal transport theory. The framework combines visibility graph embeddings with the Wasserstein barycenter to uncover latent structures and asymmetric dependencies between the two interconnected energy markets. Log-return time series are first transformed into visibility graphs and then embedded into high-dimensional vector spaces, where complex temporal and structural patterns become more discernible. In the embedding space, an information-driven Wasserstein barycenter is computed by optimizing the barycenter weights via Shannon entropy maximization. This procedure reveals an asymmetric balance between the two markets, with natural gas exerting a structurally dominant influence. To characterize the joint stochastic dynamics, a Gaussian Mixture Model is fitted to the thus determined unbalanced Wasserstein barycenter using maximum likelihood estimation via the Expectation–Maximization algorithm. An additional Gaussian component is introduced for each commodity to capture market-specific behavior. The resulting model can be calibrated to match the first four moments of the empirical log-return distributions and their observed correlation. Applied to Italian market data from 2019 to 2023, a period marked by extreme volatility and systemic shocks, the methodology accurately reproduces both common dynamics and idiosyncratic deviations. The analysis reveals that the entropy-optimal barycentric weights are $\lambda_{\mathrm{ng}} = 0.65$ for natural gas and $\lambda_{\mathrm{el}} = 0.35$ for electricity, highlighting a dominant role of the natural gas market in the joint representation. Compared with a comprehensive benchmark of GARCH-type models, the proposed framework exhibits markedly superior empirical performance. The approach provides a robust, interpretable, and adaptable tool for risk analysis, derivative pricing, and the study of structural interactions across energy markets.

1 | Introduction

Energy markets exhibit complex interdependencies and dynamic fluctuations that standard statistical models often struggle to capture [1]. These interdependencies arise from structural market dynamics that influence energy pricing, supply–demand relationships, and cross-market interactions. Natural gas and electricity markets, in particular, are highly interconnected, and price movements in these markets influence each other, creating a system in which dependencies emerge over time in a nonlinear way [2].

The interplay between natural gas and electricity prices has become a central issue in economic and energy policy, especially in Europe [3]. Recent historical spikes in electricity prices, particularly in the 2021–2023 period, have underscored the critical role of natural gas in determining electricity costs. The relationship between natural gas and electricity prices is influenced by several factors, including market integration levels, geopolitical tensions, and regulatory frameworks [4]. In particular, the merit order approach, widely used in Europe, assigns the last type of energy needed to meet demand the role of price setter for all power plants. Due to its position on the merit order curve, natural gas has historically played, and continues to play, a central role in determining electricity prices, becoming the main driver of electricity prices for more than 80% of the hours in 2021 in several countries such as Belgium, Britain, Greece, Italy and the Netherlands [5]. This method, while economically viable in stable times, has proven problematic during crises such as the Russian-Ukrainian war [6]. Disruptions in natural gas supplies caused by the war, particularly from Russia, led to unprecedented spikes in gas prices, which significantly increased electricity prices [4]. The increased volatility in the natural gas market has directly translated into increased volatility in the electricity market, revealing the fragility of the current market architecture. Structural reforms are probably necessary, as the impact of the renewable energy transition on electricity prices is difficult to predict [5, 7].

Based on these considerations, it becomes essential to understand the intricate relationship between natural gas and electricity prices in order to develop appropriate models to measure and hedge the risk associated with unpredictable price changes in energy markets [8]. Joint modeling of natural gas and electricity prices that properly accounts for correlation is also an important task in pricing derivatives, such as spark-spread options that are commonly traded in energy markets to hedge price differences between electricity and gas [9, 10].

One influential line of research relies on the family of Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) models and their multivariate extensions (M-GARCH). These models, introduced by Engle [11] and systematically developed in foundational works such as Bollerslev [12], Tsay [13], and Francq and Zakoian [14], have become standard tools for modeling and forecasting time-varying volatility in financial and energy markets. In the modeling of energy commodity prices, multivariate GARCH approaches are commonly employed to study volatility transmission, co-movement patterns, and conditional dependencies [12]. These models remain valuable for quantifying risk and understanding short-term fluctuations, offering well-established estimation procedures ab nd interpretable outputs for practitioners and researchers alike. However, despite their widespread use and strong theoretical foundations, standard GARCH-type models face inherent limitations when it comes to capturing complex nonlinear interactions, latent structural dependencies, and asymmetric influences that often characterize interconnected energy systems.

A second line of research deals more specifically with the construction of stochastic models for derivative securities valuation and financial risk management. Mean-reverting diffusion models under correlated Brownian motions have been proposed in the literature [15, 16]. However, constructing and calibrating stochastic electricity and natural gas price models with realistic correlation is complex, especially when considering the inclusion of (correlated) jumps or regime switching [17, 18].

In a third line of research, machine learning techniques have been used to model the joint behavior of electricity and natural gas prices. Recent studies apply multi-energy demand forecasting [19], and electricity price prediction using macroeconomic data [20], demonstrating the value of data-driven methods in integrated energy systems [21]. However, these approaches focus primarily on making accurate predictions and are not specifically designed to describe how different but interconnected markets interact. In particular, they do not explore the deeper structure of the data, such as how information is shared between markets or how one market may influence the other asymmetrically.

To overcome the limitations of classical approaches, we propose a novel methodology to jointly study the complex dynamics of natural gas and electricity prices through a careful investigation of the daily log-return distributions of both natural gas and electricity. Log-returns are defined as the changes over a given time interval (1 day, in our case) of the stochastic component of log-prices. The use of log-return data is motivated by the typically non-stationary nature of natural gas and electricity market prices, whereas log-returns exhibit more favorable statistical properties [22, 23]. The goal is twofold: (i) to uncover and identify a common stochastic structure affecting the dynamics of natural gas and electricity prices; (ii) investigate commodity-specific stochastic factors of the price dynamics.

We address the first, more challenging goal through a carefully designed multi-stage analytical framework that integrates state-of-the-art techniques from different domains of Machine Learning. These methods offer greater flexibility than econometric approaches and stochastic modeling [24, 25], and effectively capture nonlinearities in the dynamics and dependencies between markets without imposing stringent parametric assumptions. Specifically, we propose a methodology that effectively integrates the Wasserstein barycenter [26, 27] and Graph Machine Learning techniques [28] with the aim to uncover meaningful patterns, relationships, and behaviors within interconnected systems. Using a combination of tools to process and transform data, this methodology allows us to explore the possibility of identifying common factors that drive the dynamics of natural gas and electricity prices.

The Wasserstein barycenter is a weighted mean of probability distributions defined via optimal transport [29]. Two main types exist: balanced and unbalanced barycenters. The balanced barycenter assigns equal weights to all the distributions, ensuring symmetry and interpretability. The unbalanced barycenter adjusts weights to reflect varying importance of the underlying distributions [30, 31]. In the last years, the balanced barycenter has remained dominant [32], but unbalanced solutions are gaining traction with applications ranging from model fusion to multi-domain adaptation [33]. In real-world applications and, in particular, in applications to energy markets, where uncertainty is influenced by structurally different sources, unbalanced barycenters offer greater flexibility in capturing inherent asymmetries between distributions [34]. The relationships between natural gas and electricity log-returns are often asymmetric and nonlinear, meaning that changes in one market may disproportionately affect the other, and not in a predictable or stable way. Capturing these asymmetries is essential to improve both

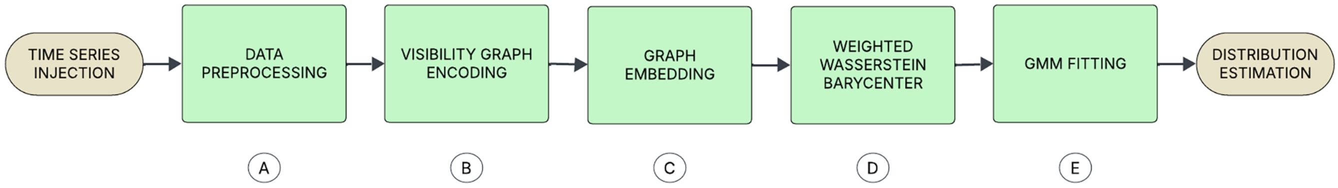


FIGURE 1 | Information-geometric pipeline.

the understanding of the underlying market structure and the description of price behavior.

The construction of the unbalanced Wasserstein barycenter is a process that requires several steps, as illustrated in Figure 1. Each stage plays a distinct but complementary role in the overall modeling architecture. The proposed framework should not be viewed as a sequence of independent operations but as a coherent information–geometric model designed to capture the structural interdependence between energy markets.

In the first step (A), prices are detrended and log-prices are computed. We used the LOESS (LOcally Estimated Scatterplot Smoothing) algorithm which is a flexible technique that removes non-stationarity and isolates long-term structure [35]. After detrending, the resulting series are normalized using a min–max scaling procedure to ensure comparability across markets and stability in the subsequent modeling steps.

In the second step (B), normalized log-return time series are transformed into structured representations through the use of visibility graphs [36]. The choice of the visibility-graph representation is motivated by the fact that this encoding transforms temporal dynamics into relational topologies and effectively captures nonlinear dependencies and long-range structures, including cross-market coupling in financial time series [37, 38]. This method treats each data point as a node and establishes connections based on a visibility criterion, preserving both the temporal order and relationships within the data. The visibility graph transformation encodes individual data points and their interactions that are critical to understanding the price dynamics. Each market price time series is encoded as a complex network of relationships in which connections represent dependencies between different time points. In this way, the graph-based perspective can capture higher-order dependencies and model long-range relationships that would otherwise be overlooked.

In the third step (C), the graphs are embedded into high-dimensional Euclidean vector spaces (128 dimensions in our case), adding a new layer of structural information. This transformation reorganizes the data into a more expressive representation, where meaningful features emerge more clearly. The embedding captures both local and global relationships within the graph, revealing structures such as clusters, recurring market behaviors, and anomalies that may indicate price shocks. In this way, the graph embedding process projects graphs into a common latent space where probability distributions become comparable [39]. To generate these vector representations, we adopt Diff2Vec [40], a graph embedding method based on diffusion processes. Instead of manually specifying features to capture structural dependencies, this approach automatically extracts informative patterns by simulating random walks on the graph. In doing so, it encodes the underlying structure of price dynamics in a high-dimensional space. The embedding into a high-dimensional space does not alter the intrinsic informational content, but redistributes it within a richer geometric framework. This transformation allows us to quantify the asymmetry (barycentric imbalance) in the contribution of each market to the joint representation through the Wasserstein barycenter computed in the embedding space.

In the fourth step (D), the unbalanced Wasserstein barycenter performs the data-fusion step under optimal transport geometry [41]. The weights of the Wasserstein barycenter are determined by an optimization procedure based on Shannon entropy maximization [42] in an attempt to include the most informative content present in market prices. This enables a more accurate and adaptive modeling of the interactions between the two energy markets, reflecting their inherently unbalanced and dynamic nature. To optimize the fusion of market behaviors, a grid of possible pairs of weights $\left(\lambda_{\mathrm{ng}}, \lambda_{\mathrm{el}}\right)$, for natural gas and electricity, respectively, is defined, with $\lambda_{\mathrm{el}} = 1 - \lambda_{\mathrm{ng}}$ and $\lambda_{\mathrm{ng}}, \lambda_{\mathrm{el}} \in [0, 1]$. We systematically explore different weight allocations between natural gas and electricity through an entropy-based optimization approach: for each pair of values $\left(\lambda_{\mathrm{ng}}, \lambda_{\mathrm{el}}\right)$, we compute the Shannon entropy of the resulting barycenter, using it as a criterion to assess the informational richness of the fused distribution. The optimal allocation is identified as the pair of weights that maximizes entropy, ensuring that the resulting barycenter captures the most informative structure of the underlying data.

The fifth step (E) is devoted to identify a common stochastic structure of the dynamics of natural gas and electricity log-returns. To do this, we assume that the Wasserstein barycenter is described by a Gaussian Mixture Model (GMM) [43]. GMMs provide universal approximations of any continuous probability density function and have proven to be very flexible tools for describing the characteristics of observed data, including rare and extreme events that can occur under non-normal market conditions [44]. GMMs provide analytically tractable parametric form that enables inference, simulation, and estimation [45]. We then fit a GMM to the Wasserstein barycenter using maximum likelihood via the Expectation–Maximization algorithm [46]. Estimating a GMM on the barycenter distribution allows us to give an analytical representation of the stochastic structure underlying the dynamics.

Together, all these steps form a unified framework for analyzing how information propagates and concentrates across interconnected energy markets, rather than a chain of unrelated technical steps. The aim is to extract informative geometric representations of market behavior, preserving key temporal structures and highlighting inter-market influences that may not be apparent in raw time series analysis. This methodology allows for a more adaptive and data-driven approach to balancing, enabling the discovery of latent correlations, dependencies and asymmetries between different markets that often escape traditional approaches.

To situate our contribution within the existing literature, Table 1 summarizes the main modeling lines discussed in the introduction and contrasts them with the information–geometric framework proposed in this paper.

The empirical analysis is conducted on the Italian energy market, specifically using time series of natural gas and electricity prices provided by GME (Gestore dei Mercati Energetici SpA). Our dataset consists of daily frequency time series spanning from January 1, 2019, to December 31, 2023. This period is notable for its unique conditions, including the impacts of a pandemic emergency, the onset of economic recovery, and the turbulence caused by the war in Ukraine. These factors make it an excellent test for the proposed methodology. The main results are so summarized. First, we observed that natural gas plays a structurally dominant role in shaping the dynamics of the common market, but the contribution of the electricity market is still significant, albeit to a lesser extent. The imbalance, derived in the high-dimensional embedding space, is then used to compute the unbalanced Wasserstein barycenter in the original one-dimensional time series domain. The analysis reveals that the optimal weights are $\lambda_{\text{ng}} = 0.65$ and $\lambda_{\text{el}} = 0.35$, highlighting a dominant role of natural gas market in the joint representation, a result that is fully consistent with market structure considerations. In European electricity markets, the merit-order mechanism frequently places natural gas–fired power plants as marginal units, making gas the primary price-setting technology during a large share of hours. Consequently, fluctuations in natural gas prices tend to exert a disproportionately strong influence on electricity price dynamics, aligning with the imbalance revealed by the entropy-optimized barycentric weights. Second, a three-factor GMM effectively reproduces the unbalanced Wasserstein barycenter. Third, in addition to common factors, our analysis reveals that the dynamics of both natural gas and electricity prices is influenced by commodity-specific stochastic factors.

We, then, addressed the challenge of modeling jointly the probability distribution of natural gas and electricity log-returns. Starting from the estimated GMM representation of the unbalanced Wasserstein barycenter, we include an additional Gaussian component to capture commodity-specific stochastic factors that can influence the dynamics. In the resulting model, the common GMM components are responsible for the correlation between the log-returns of natural gas and electricity while the additional component attempts to explain the commodity-specific behavior of the log-returns. Given the mathematical structure of GMMs, the proposed methodology has features that make it particularly useful for financial applications, such as implementing risk management strategies [49, 50] and pricing energy derivatives [51].

The overall model can calibrated to the first four moments of the empirical distributions of natural gas and electricity log-returns, as well as the empirical correlation between the two distributions. Through this fine-tuning we get a better alignment between the estimated distribution and actual data. This improves the analysis and opens up new possibilities for analyzing and understanding the joint dynamics of natural gas and electricity market prices.

We note that although our approach does not incorporate renewable output as an explicit exogenous variable, the information–geometric representation of price dynamics naturally captures the structural footprint of renewable intermittency through several mechanisms. When projecting electricity embeddings onto gas embeddings, some electricity points remain unmapped or mapped with low-confidence transport mass (see Figure 8). These correspond to price events driven by factors that do not originate from gas, such as: variability in renewable production; short-term supply imbalances; market interventions and system constraints. The geometry of these misaligned regions provides a qualitative signature of renewable-driven fluctuations. Moreover, the estimation reveals an additional electricity-only Gaussian component required to match higher-order moments. This component captures idiosyncratic volatility not explained by gas, which is consistent with the well-documented effects of renewable intermittency (e.g., sudden variations in wind and solar output affecting intraday and daily marginal costs). In this way the method can recognize, isolate, and represent renewable-driven dynamics as part of the electricity market's idiosyncratic structure.

The proposed methodology offers several advantages. The fact that the model captures the first four moments of the empirical distributions of log-returns and the correlation between the two distributions is particularly relevant, especially for financial applications [52]. Notably, it enables the construction of realistic and analytically tractable models that can be easily calibrated to market data. This capability is crucial for all market participants. Identifying common characteristics and underlying factors among energy commodities can help market participants and policymakers better understand market trends and risk factors. Moreover, the ability to reveal the market-specific stochastic factors that drive energy prices could help policymakers predict and mitigate the risks associated with price volatility. These insights could be used to design better strategies for stabilizing energy prices, reducing economic shocks caused by sudden market changes.

The complete code to reproduce our results is available at https://github.com/cbaldassari/unbalanced. LOESS detrending is computed using the implementation provided in the statsmodels library [53]; time series are encoded as visibility networks through the ts2vs package [36]; graph embedding is generated using the diff2vec package, which constructs vector representations via diffusion-based random walks [54, 55]. Computations involving optimal transport and the Wasserstein barycenter rely on the python optimal transport (POT) library [41]; GARCH-type benchmark models used for comparison are estimated using the arch Python library [56]. In addition, Gaussian mixture modelling is performed using the GaussianMixture implementation available in scikit-learn [57]. All software tools are employed with their respective default parameter configurations to ensure methodological consistency and reproducibility. No further hyperparameters are present in the methodology.

TABLE 1 | Comparison between existing approaches in the literature and the proposed information–geometric framework.

Reference(s)	Methodological class	Main objective	NG–EL interaction	Limitations vs. this paper
[47]	Econometric: GARCH-MIDAS	Short-term volatility with mixed-frequency info	Usually univariate or loosely linked; focus on volatility only	Cannot capture full distributional shape; weak for nonlinear and asymmetric effects.
[48]	Econometric: VECM	Long-run cointegration analysis	Explicit cointegration, linear adjustments	Poor for short-run dynamics, extremes, and higher-order dependence.
[12]	Econometric: (M-)GARCH	Volatility spillovers, co-movement	Time-varying covariance via multivariate GARCH	Second-moment focus; struggles with skewness, kurtosis, and full NG–EL joint distribution.
[15, 16]	Stochastic diffusions	Derivative pricing; mean reversion	Correlation via Brownian motions	Hard to calibrate with jumps/regimes; limited joint tail behavior.
[17, 18]	Lévy/regime-switching models	Joint modeling with jumps/regimes	Correlation through Lévy copulas or jumps	Flexible but parametric; limited control on matching all empirical moments.
[19, 20, 24, 25]	Machine learning	Price/demand forecasting	Often multi-output but mainly predictive	Black-box; no explicit joint distribution or structural dependency modeling.
This paper	Information geometric OT + visibility graphs + GMM	Joint modeling of natural gas and electricity log-return distributions; match first 4 moments + correlation	Structural dependencies from graphs, embeddings, unbalanced Wasserstein barycenter	Captures higher-order moments and asymmetries; interpretable and well-calibrated joint structure.

Although the combination of Diff2Vec embeddings, Wasserstein barycenter computation, and Expectation Maximization estimation involves nontrivial computational demands, the overall framework remains scalable. Using our computing cluster, consisting of three nodes each equipped with an NVIDIA RTX 3080 GPU and 64 GB of RAM, the full pipeline required approximately 45 min for both markets. In larger multi market settings computational time would increase linearly with the number of assets, but the method parallelizes efficiently across nodes and across GPUs. This ensures that the proposed approach remains feasible even for higher dimensional systems while preserving its flexibility and interpretability.

The remainder of the paper is structured as follows. After outlining the data preprocessing procedure, Section 2 discusses the Wasserstein barycenter as a data fusion tool operating in an optimal transport-based embedding space and presents an entropy-driven procedure for weight optimization. Then, in the same section, the focus is placed on the parametric representation of the barycenter through a suitable Gaussian Mixture Model, and the corresponding estimation strategy is detailed. Section 3 proposes a joint stochastic model that incorporates both the barycentric representation and the commodity-specific components. In the same section, the ability of the model to replicate empirical distributions and their interdependence is also discussed. Finally, Section 5 concludes by summarizing the main contributions and outlining directions for future research.

2 | Toward an Informative Imbalance in the Wasserstein Barycenter

2.1 | Data Preprocessing

We performed the empirical analysis on the Italian energy market. Specifically, we used the time series of natural gas and electricity prices provided by GME (Gestore dei Mercati Energetici SpA) and freely downloadable from the official website www.mercatoelettrico.org. The electricity price, known as PUN (Prezzo Unico Nazionale), is provided on an hourly basis and expressed in nominal euros per megawatt-hour. Natural gas prices are provided on a daily basis as a volume-weighted average price, also expressed in nominal euros per megawatt-hour. Our analysis was conducted on daily prices, and to obtain a daily electricity price, we performed a volume-weighted average of hourly PUN prices over 24 h.

The period under investigation spans from January 1, 2019, to December 31, 2023, encompassing 1826 daily observations. During this 5-year interval, both natural gas and electricity prices exhibited highly erratic movements due to the pandemic emergency, the onset of economic recovery, and the turbulence caused by the war in Ukraine. Therefore, we seek to identify trends and seasonality over time to highlight the underlying stochastic processes governing the market dynamics. Let us denote by P_t the market price at time t (of one megawatt-hour of natural gas or electricity) and by

TABLE 2 | Stationarity tests for log-return series under different LOESS bandwidth specifications.

Bandwidth	ADF		Phillips–Perron		KPSS	
	NG	EL	NG	EL	NG	EL
0.05	-12.84^{***}	-11.92^{***}	-13.01^{***}	-12.15^{***}	0.041	0.038
0.10	-13.21^{***}	-12.47^{***}	-13.45^{***}	-12.68^{***}	0.035	0.032
0.15	-12.98^{***}	-12.31^{***}	-13.22^{***}	-12.54^{***}	0.039	0.036
0.20	-12.76^{***}	-12.08^{***}	-12.95^{***}	-12.29^{***}	0.043	0.041
0.25	-12.53^{***}	-11.87^{***}	-12.71^{***}	-12.05^{***}	0.048	0.045
0.30	-12.31^{***}	-11.64^{***}	-12.48^{***}	-11.82^{***}	0.052	0.049
Critical values						
1%	−3.43		−3.43		0.739	
5%	−2.86		−2.86		0.463	
10%	−2.57		−2.57		0.347	

Note: ADF and PP test the null hypothesis of a unit root; KPSS tests the null of stationarity. All tests confirm stationarity across the entire bandwidth range [0.05, 0.30]. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

TABLE 3 | Sensitivity of log-return distributional moments to LOESS bandwidth.

Bandwidth	Natural Gas				Electricity			
	Mean	Std	Skew	Kurt	Mean	Std	Skew	Kurt
0.05	4.85e-05	0.0700	0.360	13.69	2.27e-04	0.1465	0.328	5.43
0.10	**2.47e-05**	**0.0705**	**0.404**	**14.02**	**2.49e-04**	**0.1467**	**0.323**	**5.43**
0.15	−8.42e-05	0.0707	0.453	14.17	2.14e-04	0.1467	0.319	5.43
0.20	−12.41e-05	0.0708	0.461	14.19	1.56e-04	0.1467	0.317	5.43
0.25	−13.69e-05	0.0708	0.445	14.17	1.00e-04	0.1467	0.314	5.43
0.30	−15.75e-05	0.0708	0.430	14.15	0.63e-04	0.1467	0.312	5.43

Note: Skewness and kurtosis remain stable across the examined range, confirming robustness of the preprocessing step. The bold row (bandwidth = 0.10) identifies the baseline specification selected for the empirical analysis.

$$s_t = \ln P_t, \tag{1}$$

its natural logarithm. We assumed that s_t is a linear superposition of a deterministic component, f_t, accounting for trend and seasonality, hereafter called trend, and a random component, ξ_t, that is,

$$s_t = f_t + \xi_t. \tag{2}$$

log-returns are defined as daily changes in the random component of log-prices,

$$r_t = \xi_{t+1} - \xi_t. \tag{3}$$

we used the LOESS (LOcally Estimated Scatterplot Smoothing) algorithm to detect f_t, the deterministic component of the dynamics [35]. LOESS is a flexible technique that allows for trend and seasonality removal by fitting simple polynomial models to localized subsets of data. The primary advantage of LOESS over many other methods is that it does not require the specification of a global function or the assumption that the data must fit a particular distribution shape. The LOESS method is based on the notion that any function can be well approximated in a small neighborhood by a low-order polynomial and that simple models can be easily fitted to data [58].

The decomposition of log-prices into a deterministic trend and a stochastic component through LOESS smoothing plays a crucial role in our preprocessing pipeline. To ensure that our results do not depend on an arbitrary choice of the smoothing parameter, we conducted a systematic robustness analysis by varying the LOESS bandwidth within a broad and economically meaningful range, namely from 0.05 to 0.30. For each bandwidth value, we obtained the corresponding residual series, which are subsequently differenced to compute log-returns, and evaluated their statistical properties.

Across the entire bandwidth range, the resulting log-return series consistently exhibited strong evidence of stationarity. This conclusion is supported by a battery of standard unit-root and stationarity tests, including the Augmented Dickey–Fuller (ADF), Phillips–Perron (PP), and KPSS tests. For both natural gas and electricity, the ADF and PP tests rejected the unit-root hypothesis at the 1%, 5%, and 10% significance levels across all bandwidths in the examined range, while the KPSS test did not reject the null of stationarity at the same significance levels for any specification. As shown in Table 2, these results confirm that the stochastic component extracted from the LOESS detrending procedure, and the corresponding log-returns, behave as stationary processes, as required for the subsequent analysis.

Table 3 shows sensitivity of log-return distributional moments to LOESS bandwidth.

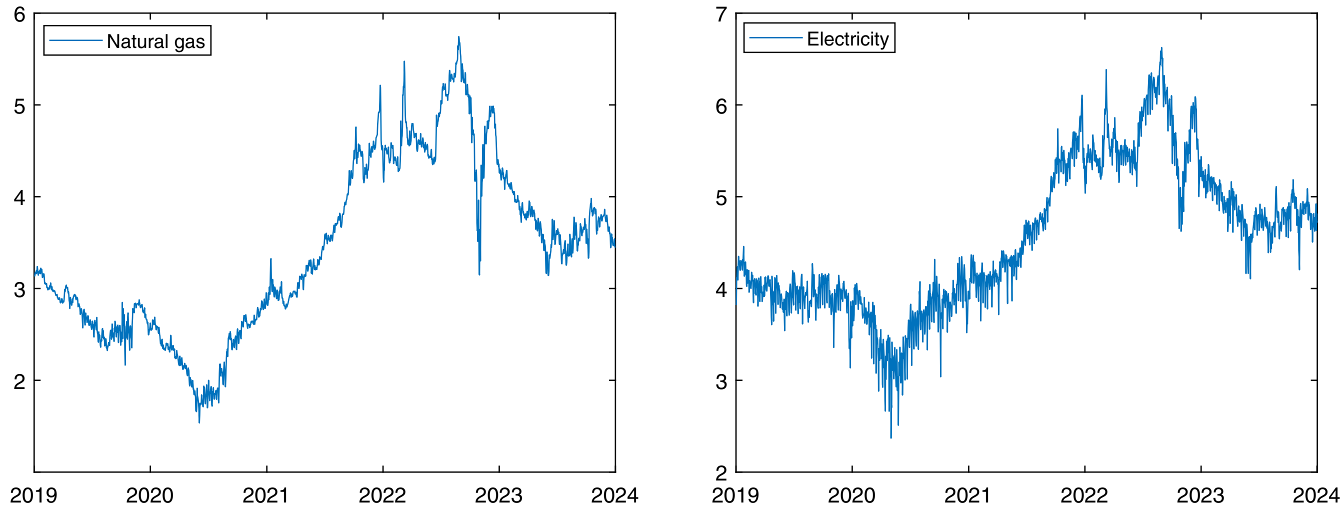


FIGURE 2 | Log-prices.

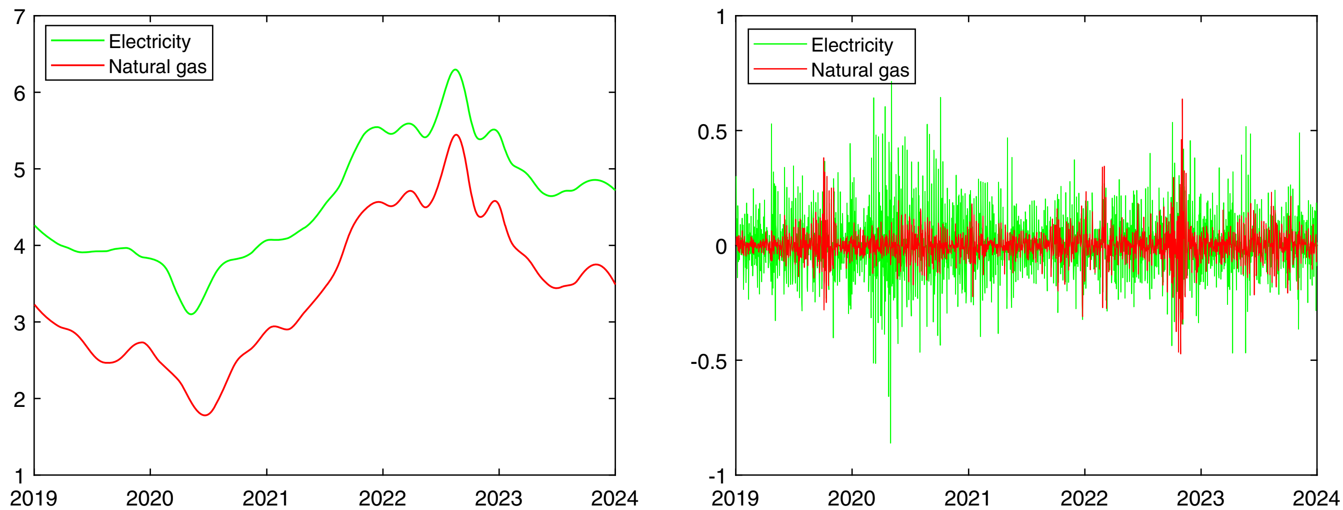


FIGURE 3 | Log-price trends, f_t (left panel). Log-returns, ξ_t (right panel).

Within this validated range of smoothing parameters, we selected a bandwidth of 0.10 as the baseline specification. This choice provides a good balance between removing long-term nonstationary movements and preserving short-term fluctuations that carry relevant information for the visibility-graph construction and the subsequent embedding and fusion steps.

Log-return time series and their empirical distributions are the output of this preprocessing task. Understanding the empirical distribution of log-returns is crucial for developing appropriate models that incorporate the main stylized facts of price dynamics. Figure 2 shows the time series of log-prices over the period under investigation (2019–2023). A visual analysis reveals that the log-prices of natural gas and electricity exhibit very similar dynamic trends. Figure 3 support this observation, with the left panel representing the trend f_t, and the right panel illustrating the stochastic movements of the log-returns ξ_t, for both natural gas and electricity. Table 4 displays the descriptive statistics of log-return distributions, indicating the values of the mean, standard deviation (std), skewness (skew) and kurtosis (kurt), hereafter referred to as the first four moments, and providing a comprehensive summary of their statistical properties. The empirical correlation coefficient between natural gas and electricity log-returns over the period under investigation is $\rho = 0.46$. These data and figures collectively illustrate the dynamic and statistical characteristics of natural gas and electricity log-returns, laying the groundwork for further analysis and modeling.

TABLE 4 | Descriptive statistics.

	Mean	**Std**	**Skew**	**Kurt**
Natural gas	2.47e-05	0.0705	0.4040	14.0248
Electricity	2.49e-04	0.1467	0.3226	5.4253

Analyzing the log-returns reveals that electricity prices exhibit higher volatility than natural gas prices: the standard deviation of electricity log-returns is 0.1467, which is more than twice that of natural gas log-returns, at 0.0705. Conversely,

TABLE 5 | First four moments of normalized log-return distributions.

	Mean	Std	Skew	Kurt
Natural gas	0.4258	0.0635	0.4040	14.0248
Electricity	0.5469	0.0931	0.3226	5.4253

the kurtosis value of natural gas log-returns (14.0248) is significantly higher compared to that of electricity (5.4253). The presence of jumps and spikes in natural gas prices considerably increases kurtosis. These shocks do not produce as pronounced effects on the kurtosis of electricity log-returns, in part because of the inherently higher volatility of electricity prices.

The empirical analysis thus indicates that, in addition to common factors, commodity-specific stochastic factors also influence fluctuations in natural gas and electricity prices. All these aspects will be discussed in depth in the following sections, highlighting the importance of taking into account both common and commodity-specific factors when modeling energy prices.

In the empirical analysis, we normalized the log-returns to account for the heterogeneity of the data and assumed uniform weights. The normalization process was carried out using a min-max scale of the time series values of the log-returns in the range $[0, 1]$ according to the normalization algorithm,

$$\bar{r}_i = \frac{r_i - r_m}{r_M - r_m} \qquad i = 1, 2, \cdots, N, \tag{4}$$

where r_m and r_M are, respectively, the minimum value and the maximum value of the log-return time series $\left\{r_i\right\}_{i=1}^{N}$. The normalization process changes the mean and standard deviation values but leaves skewness and kurtosis unchanged. Table 5 shows the first four moments of normalized log-returns distributions for natural gas and electricity.

2.2 | From Time Series to Graphs

Graphs are particularly well-suited for modeling the complex interdependencies between different markets. To uncover the latent structure of the time series, in this study we transform them into graph representations using the visibility network technique [36]. In this method, each data point in the time series becomes a node, and connections (edges) are established based on a visibility criterion: two nodes are connected if a "line of sight" exists between them, determined by their relative values and temporal positions. The concept behind the Natural Visibility Graph (NVG) is that each observation in the time series is associated with a node in the graph and represented as a vertical bar with a height equal to the numerical value of the observation. The nodes in the graph are serially ordered since each one corresponds to a time stamp of the time series. Two nodes are connected if a line of visibility that is not interrupted by the bar of a node between the two exists between the corresponding data bars. Stated mathematically, two nodes, v_i and v_j, are connected (have visibility) if any additional observation x_k with $i < k < j$ satisfies the following NVG algorithm,

$$x_k < x_j + \frac{j-k}{j-i}\left(x_i - x_j\right). \tag{5}$$

The NVG algorithm is parameter-free, simple to use and computationally fast [59]. This transformation preserves both the temporal order and intrinsic patterns of the data while encoding the dynamics of the gas and electricity markets into structured graphs. It ensures that the sequential nature of the log-returns time series is fully preserved in the graph structure. Each node represents a specific log-return value at a given time step, and edges are established based on the chronological and value-based relationships between these points. This approach guarantees that the progression of price changes over time is accurately reflected in the graph, which is essential for capturing the dynamics of financial markets. The visibility network can highlight key events, such as spikes in gas prices due to supply disruptions, and capture how these propagate into the electricity market, influencing the price dynamics. Figures 4 and 5 show the visibility graph representations of natural gas and electricity log-return time series. In each figure, the top panel displays the graph arranged according to the time series order, preserving its sequential structure. The bottom panel presents the same graph reorganized using the Yifan Hu layout algorithm [60], which enhances visualization by better reflecting structural connections.

By transforming the time series of log-returns into graphs, the interdependencies and specific behaviors of natural gas and electricity markets can be revealed. This approach allows in-depth exploration, identifying (a) correlation patterns; (b) shock propagation; and (c) volatility interactions. Log-returns naturally highlight the relative relationships between price changes, and their graph representations can reveal structural similarities or correlation patterns between the two markets. For example, certain movements in gas prices might be matched by predictable changes in electricity prices, visible as substructures of similar graphs. Abrupt price changes in one market, reflected as anomalies in the graph of log-returns, often spill over to the other market because of their interconnected nature. The graph captures this propagation of shocks, allowing a more detailed analysis of the interactions between the two markets. Finally, log-returns are particularly well suited to capture periods of increased uncertainty or instability. The transformation into graphs highlights how periods of high volatility in one market align with or influence corresponding periods in the other. Dense regions in the graph reveal these interactions, providing a visual and analytical means to study the coupling of markets during periods of turbulence. This structured representation allows us to focus on connections and relationships that are at the core of modern challenges in analyzing complex systems [61].

The adoption of the visibility-graph representation and the optimal transport–based fusion is motivated by previous empirical evidence of their effectiveness in financial and energy domains. In earlier work, the Wasserstein framework was successfully applied to the analysis of interest rate dynamics governed by the Federal Reserve, showing that optimal transport geometry can capture smooth structural transitions and nonlinear adjustments

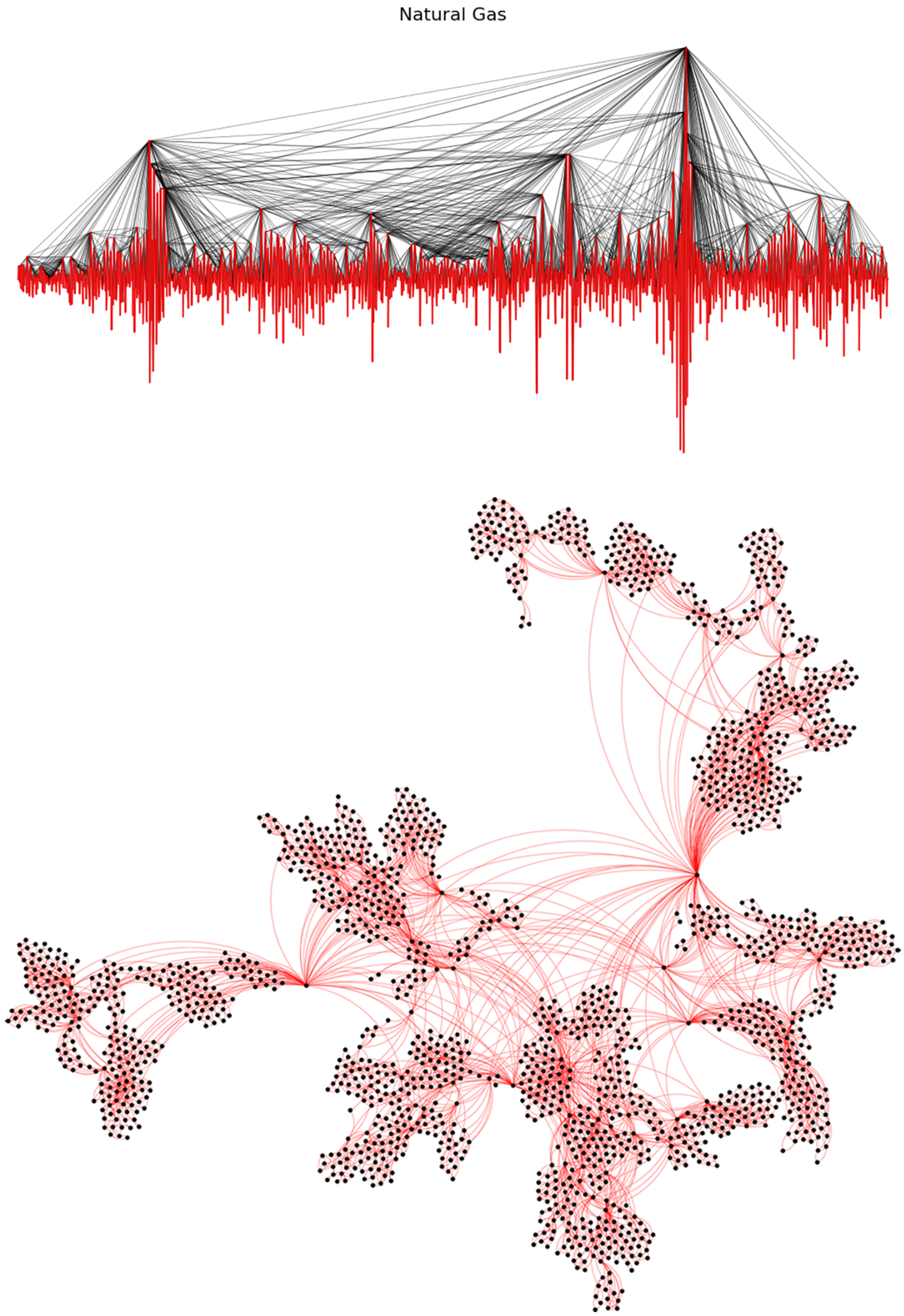


FIGURE 4 | Natural gas: Visibility graph representation of log-returns. Top: Graph arranged by time series order. Bottom: Graph reorganized using the Yifan Hu layout algorithm.

across monetary policy regimes [62]. In addition, recent applications to energy and financial time series have demonstrated that graph-based transport representations are particularly effective in uncovering structural dependencies and cross-market informational patterns [63, 64]. These findings support the use of visibility graphs as a robust intermediate representation for subsequent embedding and barycentric fusion steps in the present framework.

Beyond empirical support, there are theoretical reasons for employing visibility networks rather than working directly with raw time series. Visibility graphs transform temporal sequences into structured networks while preserving ordering, scale relations and monotonicity constraints [36]. This representation has three key advantages. First, it enhances nonlinear dependency detection by converting local geometric patterns—such as sharp changes, clustering of extremes, or long-memory patches-into topological features that are easier to compare and embed [65]. Second, visibility networks offer robustness to noise and non-stationarity, since graph connectivity depends on relative rather than absolute values, improving the stability of downstream geometric computations [66]. Third, they provide a unified structural encoding of both short- and long-range interactions, enabling the extraction of latent regimes and transition patterns that would remain hidden in direct time-domain analysis [67].

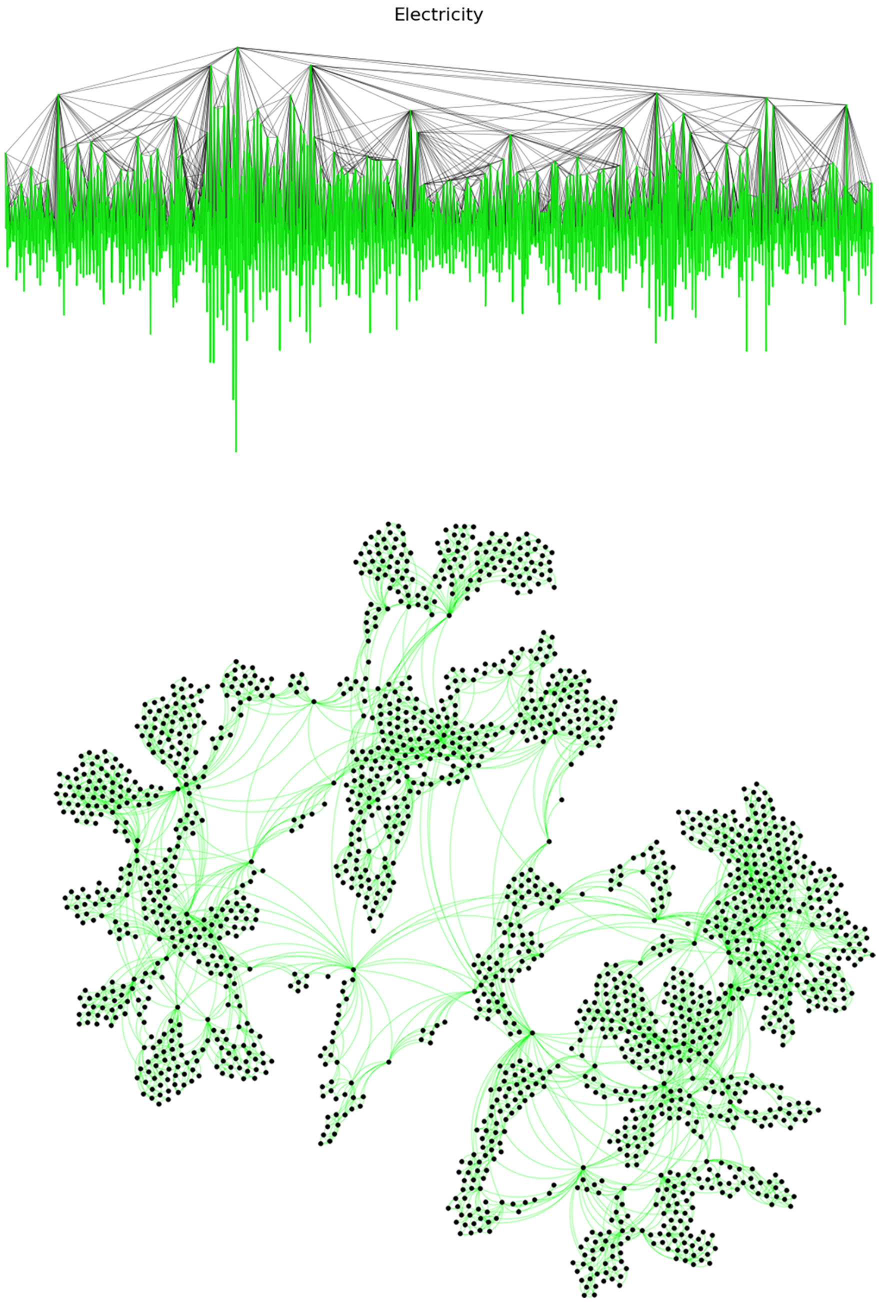
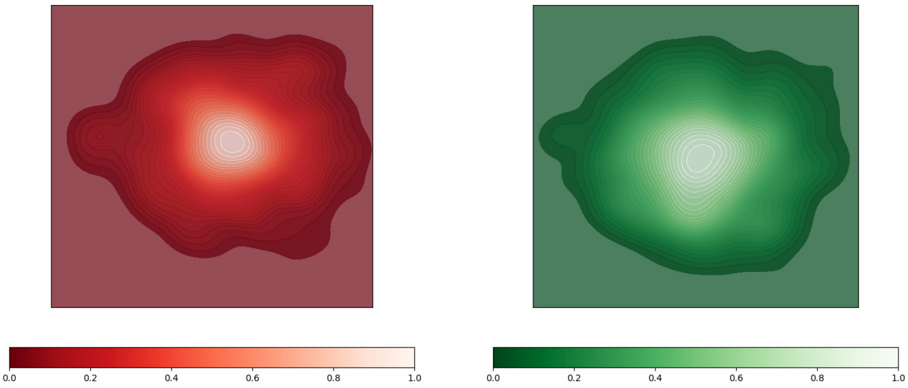


FIGURE 5 | Electricity: Visibility graph representation of log-returns. Top: Graph arranged by time series order. Bottom: Graph reorganized using the Yifan Hu layout algorithm.

For these reasons, visibility graphs are not an aesthetic transformation but a strategically chosen representation: they prepare the data for meaningful geometric comparison in the embedding space and allow the unbalanced Wasserstein barycenter to operate on distributions enriched with structural information rather than raw fluctuations. This ensures that the resulting fusion captures the deeper interdependencies between natural gas and electricity markets rather than artifacts of scale or noise.

2.3 | Graph Embedding

The graphs are then incorporated into 128-dimensional vector spaces through the Diff2Vec algorithm [40]. Diff2Vec leverages diffusion processes to capture both local and global structural properties of the graph. In fact, by simulating diffusion starting from individual nodes, Diff2Vec captures the immediate connectivity patterns of each node, such as how tightly connected it is to its neighbors. This is useful for identifying clusters or communities in the graph that correspond to recurring or similar patterns in the log-returns data. Moreover, diffusion processes also spread information across the graph, capturing long-range relationships between nodes. This helps codify how changes in one part of the time series may be related to distant events, such as the propagation of market shocks over time. The key mechanism involves running random walks on the graph, which simulate how information flows between nodes. In this way, a set of diffusion features that describe how each node relates to others both



(a) Natural gas embedding. (b) Electricity embedding.

FIGURE 6 | Density representations of natural gas (a) and electricity (b) embeddings after PCA-based two-dimensional reduction.

locally (close neighbors) and globally (far-reaching connections) is created. The features are then mapped into a high-dimensional vector space, resulting in a numerical embedding that represents the position and role of each node in the overall graph structure. This embedding step ensures that subtle dependencies and latent patterns within the visibility graphs are preserved, enabling more nuanced analysis of the gas and electricity markets.

The graph embedding generated by Diff2Vec retain the intricate relationships encoded in the visibility graphs, including correlation patterns between price movements. By encoding the diffusion dynamics, Diff2Vec identifies hidden or less apparent structures in the graph, such as weak but consistent connections between specific market events. The embeddings distill the graph complexity into fixed-size vectors and Diff2Vec transforms the graph-based representation of log-returns into a numerical form that is both computationally efficient and information-rich, enabling a deeper and more nuanced understanding of the dynamics of the gas and electricity markets. In particular, the embeddings enable the study to analyze market coupling by comparing the structural similarities and differences between the two graphs. By preserving both local (short-term) and global (long-term) relationships, this embedding process is critical for uncovering hidden relationships and patterns that traditional time series methods might overlook. Figure 6 depicts a two-dimensional view of the probability density representations of natural gas log-returns (left panel) and electricity log-returns (right panel) after PCA-based dimensionality reduction.

The two-dimensional PCA projection is a low-dimensional visualization of the visibility-graph embeddings. Each point corresponds to a high-dimensional embedding vector derived from the visibility graph of a specific time window. By applying PCA, we project these 128-dimensional vectors onto their two leading principal components, which capture the directions of maximum variance. This representation does not affect the analysis itself; instead, it provides an interpretable geometric view of the structural differences between natural-gas and electricity embeddings. Cluster separation and directional patterns in the PCA plane illustrate the asymmetric contribution of each market, visually supporting the entropy-optimized unbalanced barycenter.

2.4 | Fusion of Graph Embeddings via Optimal Transport and the Wasserstein Barycenter

Based on the Wasserstein distance, the Wasserstein barycenter allows an optimal, geometry-aware average, or fusion, to be calculated among multiple distributions. These tools provide a robust mathematical foundation for comparing and aggregating probability distributions, capturing not only pointwise differences but also the underlying geometry of the data [68–70]. In this context, they allow us to represent joint market behavior in a way that preserves the temporal structure, distributional asymmetries and latent dependencies among energy markets.

Given two empirical distributions, P and Q, derived from datasets $\{X_i\}_{i=1}^{N}$ and $\{Y_i\}_{i=1}^{N}$ in a d-dimensional space, the p-Wasserstein distance is defined by

$$W_p(P,Q) = \inf_{\pi}\left(\sum_{i=1}^{N}\left\|X_i - Y_{\pi(i)}\right\|^p\right)^{1/p}, \tag{6}$$

where the infimum is taken over all possible permutations π [71, 72]. In our analysis, we employ the p-Wasserstein distance with $p = 1$.

The Wasserstein barycenter is a probability distribution that minimizes the sum of Wasserstein distances from a given set of input distributions [29]. Given K empirical distributions $\{P_j\}_{j=1}^{K}$ with associated weights $\{\lambda_j\}_{j=1}^{K}$, the Wasserstein barycenter, WB, is computed as the solution to:

$$WB = \min_{P}\sum_{j=1}^{K}\lambda_j W_p\left(P, P_j\right). \tag{7}$$

By minimizing the weighted sum of Wasserstein distances, the barycenter captures the common features of the input distributions, providing a representative probabilistic structure. However, there is an important issue that needs to be carefully considered. Graphs derived from time series data often exhibit structural differences due to variations in scale, trend, or temporal dependencies. As a result, their embeddings may lie in distinct regions of the vector space, complicating direct comparison and integration. From a real-world perspective, alignment is crucial because similar events, such as spikes in volatility or political shocks, can affect gas and electricity at different times or sizes. Without accounting for such temporal mismatches, the fusion process distorts the joint dynamics. Optimal transport resolves this by aligning structurally equivalent patterns, enabling coherent fusion of heterogeneous market behaviors. The Wasserstein barycenter plays a fundamental role in optimal transport theory to compute the optimal matching between different sets of embeddings [29]. This enables a meaningful and information-preserving fusion, minimizing geometric distortion while retaining the statistical and structural content of the original time series. In our study, the Wasserstein barycenter is used to construct a unified embedding in the aligned space that integrates characteristics from both the natural gas and electricity embeddings while preserving essential structural properties. This fusion approach ensures that the resulting representation encapsulates the stochastic behaviors of both data sources without introducing distortions [70]. Optimal transport–based fusion provides a robust framework for analyzing relationships between different time series embeddings. This approach enables the aggregation of shared statistical structures derived from the log-return distributions of natural gas and electricity, yielding a richer and more interpretable representation of their joint dynamics.

To determine the optimal weights for the Wasserstein barycenter in the embedding space, we adopt an information-theoretic criterion based on Shannon entropy [42]. Given two distributions, corresponding to the embeddings of natural gas and electricity time series, and their associated weights λ_{ng} for natural gas and $\lambda_{\mathrm{el}} = 1 - \lambda_{\mathrm{ng}}$ for electricity with $\lambda_{\mathrm{ng}}, \lambda_{\mathrm{el}} \in [0,1]$, the objective is to identify the weight pair that maximizes the Shannon entropy of the resulting Wasserstein barycenter in the embedding space, in accordance with the principle of maximum information. Shannon entropy serves as a proxy for the informational richness and structural complexity of the fused distribution, guiding the selection toward the most representative combination of the two market embeddings.

The barycentric weights $\left(\lambda_{\mathrm{ng}}, \lambda_{\mathrm{el}}\right)$ are not free hyperparameters but are instead obtained through a fully automated procedure. Specifically, we evaluate a dense grid of candidate values $\lambda_{\mathrm{ng}} \in [0,1]$ (with $\lambda_{\mathrm{el}} = 1 - \lambda_{\mathrm{ng}}$) and select the pair that maximizes the Shannon entropy of the resulting fused distribution in the embedding space. Because both embedding distributions are individually normalized before fusion, the selected value $\lambda_{\mathrm{ng}} = 0.65$ does not reflect a scale or variance dominance. Rather, it quantifies the *relative informational contribution* of natural gas required to reconstruct the joint latent geometry in the entropy-maximizing sense. In other words, natural gas provides a structurally richer embedding distribution, and the barycenter places proportionally more weight on it to achieve the most informative and coherent joint representation.

The optimization is carried out over a discrete grid of weight values for λ_{el}, sampled at regular intervals of 0.01. For each pair $\left(\lambda_{\mathrm{ng}}, \lambda_{\mathrm{el}}\right)$, the Wasserstein barycenter is computed and its entropy evaluated. The results are shown in Figure 7, which displays the entropy values across the grid. The maximum entropy is achieved at $\lambda_{\mathrm{ng}} = 0.65$ and $\lambda_{\mathrm{el}} = 0.35$, indicating that the most informative and structurally coherent joint representation reveals a dominant role of natural gas in shaping the joint market dynamics. Figure 8 illustrates the process of aligning the natural gas and electricity embeddings using optimal transport in the maximum entropy case. The alignment reveals distinct mapping patterns that reflect key aspects of natural gas and electricity market prices behavior. In several regions, we observe one-to-one mappings, where individual points from the natural gas embedding align directly with corresponding points in the electricity embedding. This behavior indicates strong local correlations between the two markets. In other regions, we detect many-to-one mappings, where multiple gas states are transported toward a single electricity point. This suggests that different gas configurations may lead to similar electricity outcomes, revealing converging dynamics or dominant structural influence of gas over electricity. Some electricity embedding points remain unmapped, meaning they lack a counterpart in the gas embedding. These market-specific regions likely reflect electricity price variations driven by factors not captured in gas dynamics, such as renewable generation, regulatory interventions, or demand-side behaviors. This is an important point because in describing the dynamics of natural gas and electricity prices we will also have to take commodity-specific factors into account. We will discuss this in Section 3. Each link connects a point from the gas embedding to a point from the electricity embedding. If the link is short, the connected points are similar; if it is long, the connected points are different, and the alignment has to compensate for that. Points without links were not matched during the alignment and reflect specific features of only one market. Figure 9 shows in the rightmost panel a two-dimensional PCA-based reduction of the resulting Wasserstein barycenter that captures a fusion of the two markets. The fused representation captures a coherent joint structure reflecting the dominance of gas in specific regions.

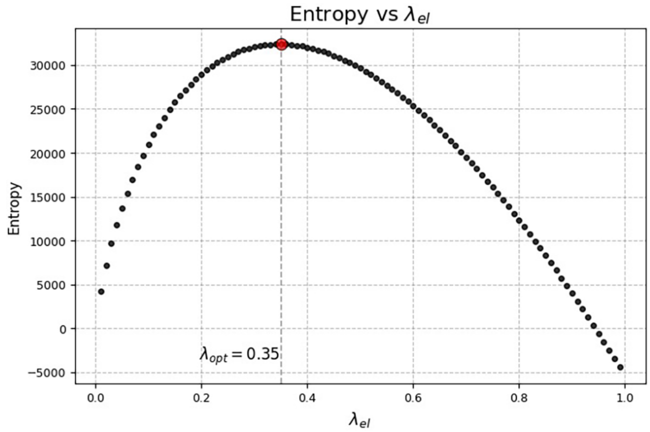


FIGURE 7 | Entropy values for different weight combinations of natural gas and electricity in the Wasserstein barycenter computation. The maximum entropy is obtained at $\lambda_{\mathrm{el}} = 0.35$ and, suggesting that this allocation provides the most structurally informative representation.

Figure 8 visualizes the relative positioning of the natural gas and electricity embeddings in a two–dimensional projection obtained through PCA. Each point represents the high-dimensional embedding of the visibility graph associated with a given time window, while the dashed gray segments link each natural gas embedding to the corresponding electricity embedding for the same date. The plot therefore provides a geometric illustration of how the two markets evolve jointly in the latent space.

Two features emerge clearly from this representation. First, the paired points are systematically close to each other, confirming that natural gas and electricity share a common underlying structure in their temporal dynamics once transformed through the visibility-graph and embedding steps. Second, in several regions of the projection, the electricity points display a greater dispersion relative to their natural gas counterparts. This suggests that electricity prices exhibit stronger idiosyncratic variability, consistent with the presence of market-specific shocks (e.g., renewable intermittency, regulatory effects) that are not fully mirrored in natural gas. The geometry of the figure therefore supports the model's structure: a dominant common component, coupled with additional commodity-specific factors needed to capture asymmetric deviations in electricity.

In this way, the Wasserstein barycenter balances the two markets according to the prescribed weights, respecting the geometry of the data, preserving the structure of both markets and incorporating asymmetries in their behavior.

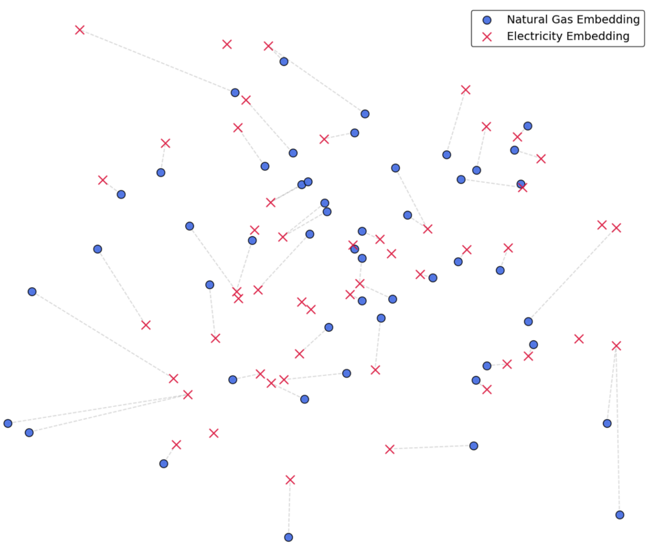


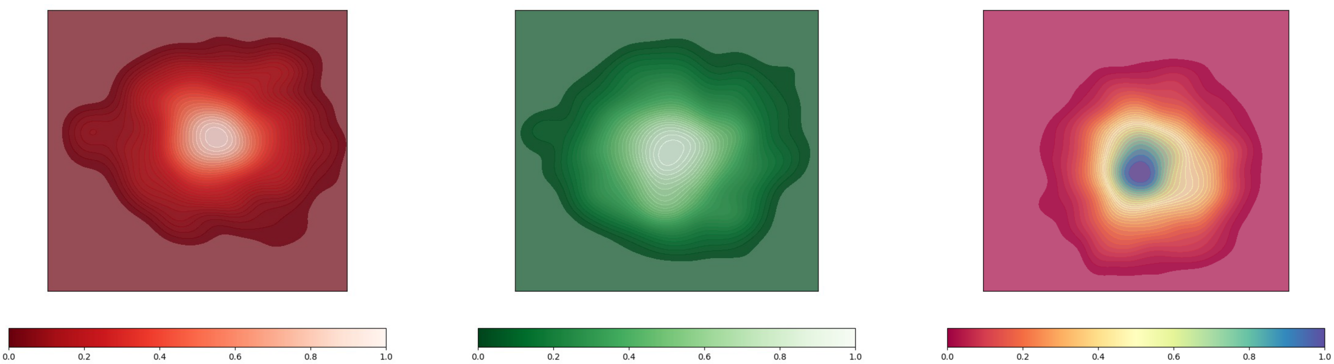
FIGURE 8 | Alignment of natural gas and electricity embeddings via optimal transport. The plot visualizes transport-induced displacements, where sampled electricity embeddings (red crosses) are mapped to their corresponding natural gas embeddings (blue circles), highlighting both shared structures and independent variations in market behavior.

Finally, to compute the Wasserstein barycenter in the original one-dimensional space of log-return distributions, we use the resulting optimal weights obtained by the entropy-based optimization performed in the embedding space. This allows us to maintain consistency with the enhanced information structure and structural insights obtained in the high-dimensional representation, where latent dependencies become more structured and easier to extract, while taking advantage of a tractable and interpretable modeling approach in the statistical domain. Figure 10 depicts, in the left panel, the empirical distribution of the normalized log-returns of natural gas and electricity; in the right panel, the one-dimensional Wasserstein barycenter is superimposed. Table 6 depicts the first four moments of the Wasserstein barycenter. These data and figures provide a detailed comparison between the empirical distributions of normalized natural gas and electricity log-returns and their Wasserstein barycenter.

The proposed method captures meaningful geometric patterns of market activity, maintaining essential temporal dynamics and uncovering relationships between market prices that might be hidden in raw time series data. This approach promotes a more flexible and data-centric strategy for balancing, revealing hidden connections and dependencies across markets that conventional techniques often overlook. In summary, this embedding-driven integration technique provides a more organized and evidence-based framework for examining energy markets, improving our capacity to understand, model, and make informed decisions in environments marked by complexity and mutual influence.

2.5 | Fitting a GMM to the Unbalanced Wasserstein Barycenter

To represent analytically the stochastic factors common in the price dynamics of natural gas and electricity, we assume that the unbalanced Wasserstein barycenter (hereinafter, Wasserstein barycenter) is described by a GMM. Often, the distribution characteristics of empirical data are too complex to be accurately



(a) Natural gas (b) Electricity (c) Wasserstein barycenter

FIGURE 9 | Density representations of the embeddings after PCA-based two-dimensional reduction. From left to right: Natural gas, electricity, and the Wasserstein barycenter. The third panel illustrates how the barycenter captures a fused structure of the first two embeddings.

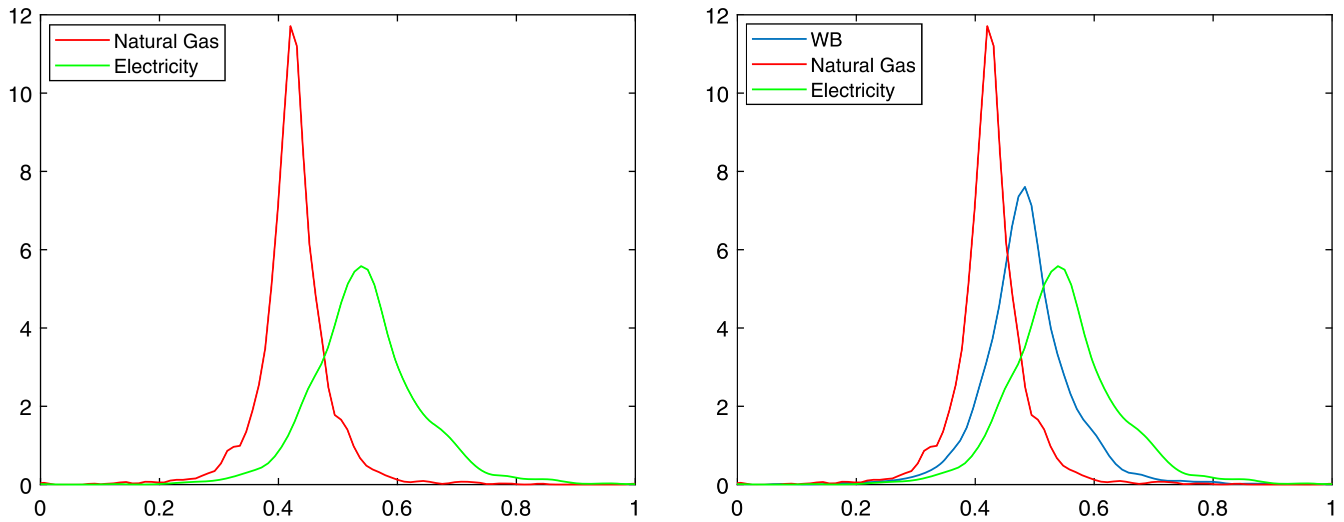


FIGURE 10 | Left Panel: Empirical distributions of normalized log-returns for natural gas and electricity. Right Panel: Wasserstein barycenter superimposed on the empirical distributions.

TABLE 6 | First four moments of the Wasserstein barycenter (WB).

	Mean	Std	Skew	Kurt
WB	0.4682	0.0733	0.3270	9.0895

modeled by a single theoretical distribution [73]. GMMs provide universal approximations of any continuous probability density function and have proven to be very flexible tools for describing the characteristics of observed data, including rare and extreme events that can occur under non-normal market conditions [43, 44] and significantly increase the kurtosis of empirical distributions [45].

A key distinction between the proposed approach and classical factor analysis lies in what is being averaged and which geometry governs the aggregation. A factor model extracts a common latent component by minimizing linear reconstruction error under Euclidean geometry, implicitly assuming that dependence across series is well described by second-order structure (covariances). This approach is powerful but limited when distributions differ in shape, exhibit asymmetries, or display heavy-tailed behavior—features well documented in energy markets.

In contrast, the Wasserstein barycenter performs an aggregation at the level of full probability distributions, not second moments. By relying on optimal transport geometry, it captures nonlinear shifts, distributional asymmetries, and tail behavior that would be invisible to a standard common-factor extraction. This makes it well suited for markets where shocks propagate in non-Gaussian and asymmetric ways.

Empirically, using a Wasserstein barycenter yields a central distribution that matches not only variance but also skewness, kurtosis, and tail dependence across gas and electricity—dimensions that a linear factor model cannot reproduce. This results in a more faithful representation of joint market behavior, particularly during periods of structural breaks or extreme events.

We assume that the Wasserstein barycenter is described by a GMM with a probability density function $p_b(x)$, which is a finite mixture distribution with n univariate components, namely

$$p_b(x) = \sum_{j=1}^{n} w_j p_j(x), \tag{8}$$

where $p_j(x)$ are univariate Gaussian densities,

$$p_j(x) = \frac{1}{\sigma_j\sqrt{2\pi}} \exp\left(\frac{-\left(x-\mu_j\right)^2}{2\sigma_j^2}\right), \tag{9}$$

with mean μ_j and variance σ_j^2. The weights $\left\{w_j\right\}_{j=1}^{n}$ are positive numbers such that

$$\sum_{j=1}^{n} w_j = 1, \tag{10}$$

representing the probability that the value x is generated by the component j. The choice to use GMMs stems from their remarkable flexibility, excellent performance, and ability to maintain a relatively low number of parameters.

GMMs can be estimated on empirical data by maximum likelihood using the Expectation–Maximization (EM) algorithm [46], a commonly recognized training technique for these models [74]. It is well known in the literature that maximum likelihood estimation of GMMs suffers from the problem of frequent local optima. For this reason, the choice of starting values for the EM algorithm is critical to obtaining an accurate solution [75]. To overcome this difficulty, we used a stochastic multiple restart method for initializing the EM algorithm, in which the EM algorithm is executed a thousand times with different random initial conditions. This choice is motivated by the fact that the number of GMM components is low and known a priori, and the number of observations is not too large. Under these circumstances, the random initialization method seems to work well [76, 77]. In our case, independent runs consistently converged to the same maximum-likelihood solution, confirming the numerical stability of the estimation procedure. To verify that the EM

TABLE 7 | Estimated parameters.

	$n = 2$		$n = 3$	
μ_1	0.4649	(0.0187)	0.4683	(0.0196)
σ_1	0.0444	(0.0025)	0.0633	(0.0034)
μ_2	0.4791	(0.0210)	0.4622	(0.0179)
σ_2	0.1280	(0.0102)	0.0208	(0.0022)
μ_3			0.4833	(0.0217)
σ_3			0.1677	(0.0075)
w_1	0.7663	(0.0247)	0.6560	(0.0206)
w_2	0.2337		0.2513	(0.0143)
w_3			0.0927	
BIC	−4703.0		−4733.0	

Note: Standard errors are between parentheses.

TABLE 8 | Stability of GMM estimation under multiple random initializations.

Diagnostic	Result
Number of independent EM runs	1000
Runs converging to global maximum	1000 (100%)
Log-likelihood (all runs)	2396.5
BIC (all runs)	−4733.0
Iterations to convergence (Mean ± Std)	47.3 ± 12.8

estimation is not affected by local minima, we performed 1000 independent runs with different random initializations.

A two-component and a three-component GMM are employed to fit the Wasserstein barycenter. The comparison between the two models is performed on the basis of the Bayesian Information Criterion (BIC) [78]. The BIC criterion addresses the model selection problem by introducing a penalty term for the number of parameters, thus balancing the trade-off between the quality of the fit and the simplicity of the model. BIC is expressed in terms of the log-likelihood in the following way,

$$\mathrm{BIC} = -2\log(\mathcal{L}) + \eta\log(N), \tag{11}$$

where $\mathcal{L}$ is the maximized value of the likelihood function of the model, η is the number of free parameters and N is the sample size. Models with lower BIC are generally preferred. According to Equations (8–10), the number of free parameters is given by $m = 3n - 1$ and the sample size is $N = 1825$. The estimation results are reported in Table 7 for both the two-component GMM ($n = 2$) and the three-component GMM ($n = 3$). To facilitate comparison, the BIC value of each model is also included.

As shown in Table 8 in the case $n = 3$, all runs converged to the same maximum-likelihood solution, confirming the numerical stability of the estimation procedure.

Based on the BIC criterion, the three-component GMM performs better than the two-component GMM. It also reproduces the first four moments of the Wasserstein barycenter well. Recalling that the moments of a univariate Gaussian mixture can be calculated exactly [79], Table 9 compares the values of the first four moments of the Wasserstein barycenter with the (exact) ones obtained from the estimated two- and three-component models.

TABLE 9 | First four moments.

	Mean	Std	Skew	Kurt
WB	0.4682	0.0733	0.3270	9.0895
GMM ($n = 2$)	0.4682	0.0733	0.2808	6.9198
GMM ($n = 3$)	0.4682	0.0733	0.2983	8.8404

Figure 11 compares the Wasserstein barycenter (continuous line) with a simulated probability density function (PDF) obtained by the estimated GMMs (dashed line) through Monte Carlo sampling. On visual inspection, the three-component GMM (right panel) produces a good fit to the observed data, significantly better than that produced by the two-component GMM (left panel). This observation is also supported by the average values of the p-Wasserstein distance ($p = 1$), which are $\overline{W}_p = 0.0035$ for the two-component GMM and $\overline{W}_p = 0.0024$ for the three-component GMM. To determine the average values of the p-Wasserstein distance, we generated by Monte Carlo sampling $N = 1825$ log-return values for each model, corresponding to the number of log-returns in the observed period from January 1, 2019, to December 31, 2023. We then calculated the p-Wasserstein distance between the simulated distribution of log-returns and the Wasserstein barycenter. This procedure was repeated a thousand times, and then the average of the so computed p-Wasserstein distances, $\overline{W}_p$, was calculated. The obtained results confirm that the three-component GMM provides a better fit to the data compared to the two-component GMM.

3 | Modeling the Joint Dynamics of Natural Gas and Electricity Prices

The Wasserstein barycenter revealed that the dynamics of natural gas and electricity prices is influenced by commodity-specific stochastic factors. In this section, we use the proposed methodology to jointly model the dynamics of natural gas and electricity prices taking into account commodity-specific factors.

Starting from the three-component GMM representation of the Wasserstein barycenter, which better describes the empirical data than the two-component GMM, we incorporate an additional Gaussian component to capture commodity-specific stochastic factors that influence the dynamics. To do this, let us introduce the label c to identify the commodity, equal to c_1 for natural gas and c_2 for electricity. For each commodity, we assume that the stochastic process driving the dynamics of (normalized) log-return, R_t^c, is described by

$$R_t^c = m^c + X_t^c, \tag{12}$$

where m^c is a constant and X_t^c are i.i.d. random variables with probability density function described by the following Gaussian mixture,

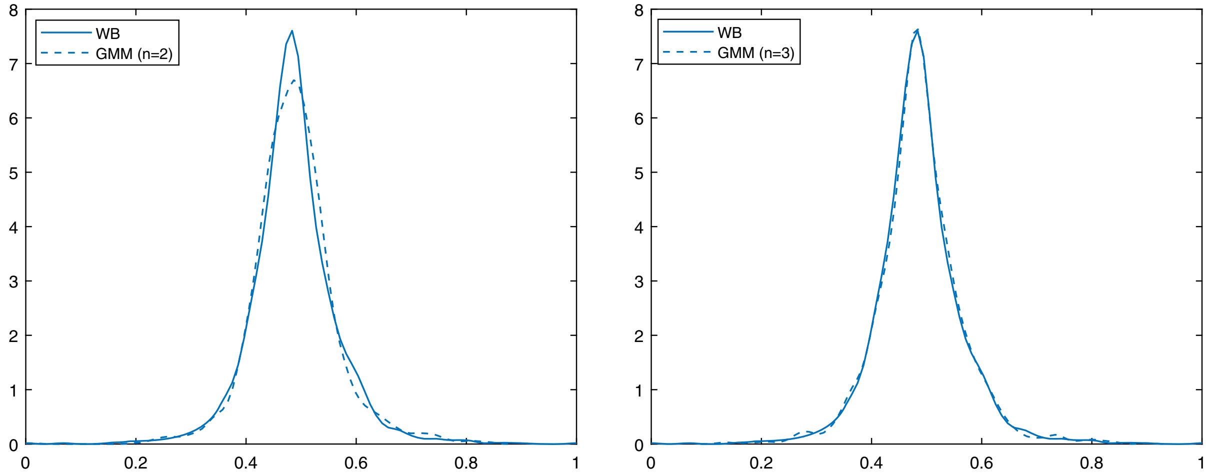


FIGURE 11 | Wasserstein barycenter (continuous line) and a simulated PDF (dashed line) from the estimated two-component GMM (left panel) and three-component GMM (right panel).

$$p^c(x) = \sum_{j=1}^{3} \alpha_j^c p_j^c(x) + \alpha_4^c p_4^c(x), \tag{13}$$

where $p_j^c(x) = p_j(x), j = 1,2,3$, are the Gaussian components of the Wasserstein barycenter, and $p_4^c(x)$ is a fourth Gaussian component, with mean μ_4^c and variance $\left(\sigma_4^c\right)^2$, accounting for a commodity-specific stochastic component. The weights $\left\{\alpha_j^c\right\}_{j=1}^4$ are positive numbers such that

$$\sum_{j=1}^{4} \alpha_j^c = 1. \tag{14}$$

The first three common components are responsible for the correlation between the log-returns of natural gas and electricity. We assume, in fact, that the joint probability density function is given by

$$p\left(x, y \mid Z_{jj} = 1\right) = \alpha_j^{c_1} \alpha_j^{c_2} p_j(x) \delta(x - y) \quad j = 1,2,3, \tag{15}$$

where $\delta(\cdot)$ is the Dirac delta-function, and by

$$p\left(x, y \mid Z_{ij} = 1\right) = \alpha_i^{c_i} \alpha_j^{c_2} p_i^{c_1}(x) p_j^{c_2}(y), \tag{16}$$

for the other cases. In Equations (15) and (16), $\left\{Z_{ij}\right\}_{i,j=1}^4$ is a set of binary random variables that are mutually exclusive and exhaustive (i.e., one and only one of the Z_{ij}'s is equal to 1, while the others are equal to 0). The first index refers to commodity c_1 (natural gas) and the second to commodity c_2 (electricity). The random variable Z_{ij} plays the role of indicator random variable representing the identity of the mixture components responsible for the generation of the outcome x and y. The Z_{ij}'s are not observable and are treated as latent variables. In this model, the covariance between $X_t^{c_1}$ and $X_t^{c_2}$ can be calculated exactly, and reads

$$\text{cov}\left(X_t^{c_1}, X_t^{c_2}\right) = \sum_{j=1}^{3} \alpha_j^{c_1} \alpha_j^{c_2} \sigma_j^2. \tag{17}$$

The four-component GMM features seven new parameters for each commodity: the four coefficients $\alpha_j^c (j = 1,2,3,4)$ subject to Equation (14), the two parameters of the commodity-specific Gaussian component μ_4^c and σ_4^c, and the overall mean m^c. We calibrated the model to the empirical data by requiring the model to reproduce the first four moments of the observed distributions of natural gas and electricity log-returns, as well as the empirical correlation between the two distributions. This is particularly important for financial applications. By ensuring that our probabilistic model accurately reflects these empirical parameters, we can provide a more reliable and robust representation of the underlying stochastic processes that drive natural gas and electricity price dynamics, thereby providing better insights and tools for risk management and derivative pricing [52].

Model calibration proceeds as follows. For each commodity, the set of parameters $\theta = \left\{\alpha_j^c, \mu_4^c, \sigma_4^c, m^c\right\}$ $(j = 1,2,3,4)$ must satisfy five constraints, namely the first four moments of the log-return distribution and Equation (14). We performed the calibration by minimizing the difference between the model correlation coefficient and the empirical correlation coefficient, $\rho = 0.46$, under the above set of five constraints for each commodity. Table 10 shows the results of this fine-tuning process. The correlation coefficient obtained is equal to 0.45. These results indicate that the selected parameters provide a good fit to the empirical data, reproducing the first four moments to the first four decimal places and accurately capturing the correlation between the two log-return distributions.

Figure 12 shows, for each energy commodity time series, a model-simulated PDF of log-returns (dashed line) superimposed on the empirical distribution of log-returns (continuous line). Upon visual examination, the fit looks very promising. This is further confirmed by the values of the Wasserstein distance between the two distribution (the empirical and the simulated), which is equal to 0.0017 for natural gas and 0.0021 for electricity log-returns.

The model offers a good fit to the experimental data. We observe that first component of the Wasserstein barycenter describes more than 80% of the log-return distributions of natural gas and electricity. Overall, the three components of the GMM

TABLE 10 | Results of the fine-tuning process.

	α_1	α_2	α_3	α_4	μ_4	σ_4	m
Natural gas	0.8095	0.1799	0.0055	0.0051	0.5235	0.3219	−0.0418
Electricity	0.8097	0.0137	0.1254	0.0512	0.6635	0.0102	0.0668

Note: The parameters α_1 to α_4 denote the weights of the GMM components; μ_4 and σ_4 refer to the mean and standard deviation of the fourth component; m indicates the mean shift required for calibration.

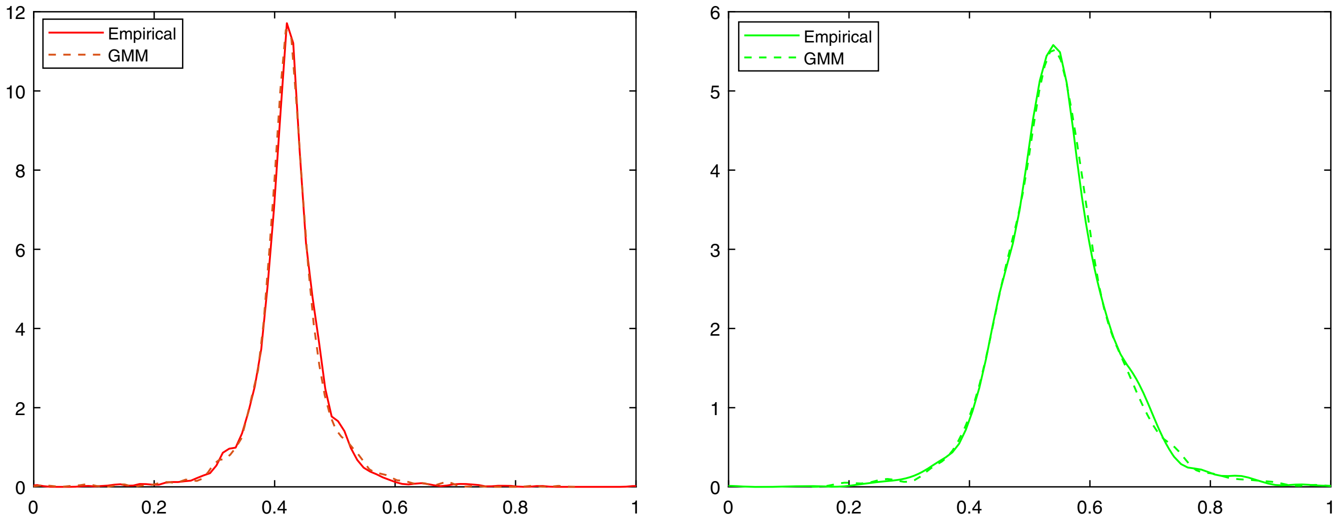


FIGURE 12 | Comparison of the empirical distribution (continuous line) and a model-simulated PDF (dashed line) of log-returns. Left panel: Natural gas. Right panel: Electricity.

representation of the Wasserstein barycenter fully describe the log-return distribution of natural gas, the fourth component being very small. More significant is the fourth component in the case of electricity (about 5%). However, in both cases the presence of this specific component is critical to capturing the first four moments of the log-return distributions as well as the empirical correlation between the two distributions. This supplementary Gaussian component works in synergy with the common model components, allowing for a more granular assessment of the dynamics of both natural gas and electricity prices. By including this additional component, the model can better reflect the underlying stochastic processes and provide a more accurate and detailed description of the price movements of both these energy commodities. In the case of electricity, the additional component is designed to capture not only the variability transmitted from gas prices but also the portion of idiosyncratic volatility that gas alone cannot explain. This aligns with the well-documented impact of renewable intermittency—such as abrupt fluctuations in wind and solar generation that alter intraday and daily marginal costs. Incorporating this term enables the model to identify, disentangle, and explicitly represent the dynamics driven by renewable sources, treating them as an intrinsic element of the electricity market's idiosyncratic structure.

4 | Comparison With GARCH-Type Models

To position the proposed information-geometric methodology within the broader landscape of energy price modeling, we benchmark its performance against a comprehensive set of 25 GARCH-type specifications implemented using the arch Python library [56]. While GARCH models remain the standard econometric approach for capturing time-varying volatility in financial and energy markets, they fundamentally rely on second-moment dynamics and parametric innovation distributions [14]. This comparison is therefore intended not as a like-for-like modeling exercise, but as a stress test evaluating whether traditional volatility-driven frameworks can replicate the full distributional geometry recovered by our Wasserstein-based approach.

4.1 | Benchmark Design

The benchmark set spans a wide range of commonly used GARCH specifications obtained by combining two volatility structures: namely GARCH(1,1) and EGARCH(1,1) [13]; four innovation distributions, Normal, Student-t, Skewed-t, and GED [80]; two allocation strategies, (i) *bivariate* models with identical specifications for both commodities, (ii) *hybrid univariate* models allowing separate specifications for natural gas and electricity [14]. To clarify the composition of every benchmark, we refer to Table 11, which serves as a dictionary, decoding each model label and listing its underlying volatility specification and innovation distribution.

Cross-commodity dependence is introduced via Cholesky decomposition [12, 81]. Here, "innovations" are simply the standardized shocks generated by each univariate GARCH model. On their own, these shocks are independent across commodities. The Cholesky transformation allows us to combine them

TABLE 11 | Dictionary of the 25 benchmark GARCH specifications.

Model label	Specification details (volatility model + innovation law)
GARCH – GED + EGARCH – t	ng: GARCH(1,1) + GED; el: EGARCH(1,1) + Student-t — heavy tails for ng, leverage and fat tails for el.
EGARCH – GED + GARCH – t	ng: EGARCH(1,1) + GED; el: GARCH(1,1) + Student-t.
Heavy – tail ng + Normal el	ng: EGARCH(1,1) + GED; el: GARCH(1,1) + Normal (baseline contrast).
Double EGARCH Extreme	ng: EGARCH(1,1) + GED; el: EGARCH(1,1) + Skewed-t — double leverage and skewness.
GARCH Opposite Distributions	ng: GARCH(1,1) + GED; el: GARCH(1,1) + Skewed-t.
EGARCH + GED	Bivariate EGARCH(1,1) + GED (same spec for both series).
Full GED Hybrid	ng: EGARCH + GED; el: GARCH + GED.
GARCH + GED	Bivariate GARCH(1,1) + GED.
EGARCH + t	Bivariate EGARCH(1,1) + Student-t.
Full t Hybrid	ng: EGARCH + Student-t; el: GARCH + Student-t.
Balanced Symmetric	ng: GARCH + Student-t; el: EGARCH + Student-t.
GARCH + t (heavy)	Bivariate GARCH + Student-t with low degrees-of-freedom (heavier tails).
EGARCH-t + EGARCH – GED	ng: EGARCH + Student-t; el: EGARCH + GED.
EGARCH Opposite Distributions	ng: EGARCH + Student-t; el: EGARCH + GED (same volatility, different tails).
Conservative Mixed	ng: GARCH + Student-t; el: GARCH + GED.
EGARCH + Skewed-t Mixed	ng: EGARCH + Skewed-t; el: EGARCH + Student-t.
GARCH + Skewed-t	Bivariate GARCH + Skewed-t innovations.
Full Skewed-t Hybrid	ng: EGARCH + Skewed-t; el: GARCH + Skewed-t.
EGARCH + Skewed-t	Bivariate EGARCH + Skewed-t.
Alternating Pattern 1	ng: GARCH + Skewed-t; el: EGARCH + Normal.
GARCH + Normal	Bivariate GARCH(1,1) + Normal innovations (baseline).
Alternating Pattern 2	ng: EGARCH + Normal; el: GARCH + Skewed-t.
EGARCH + Normal	Bivariate EGARCH(1,1) + Normal.
Aggressive Mixed	ng: EGARCH + Skewed-t; el: EGARCH + GED — maximum complexity.
Normal ng + Heavy el	ng: GARCH + Normal; el: EGARCH + GED (extreme contrast).

Abbreviations: el, electricity; ng, natural gas.

so that they exhibit the desired contemporaneous correlation: we specify a target covariance matrix, factorize it, and use it to map independent shocks into correlated ones. This creates a dependence structure between natural gas and electricity while leaving the individual GARCH dynamics unchanged. Each model is calibrated to the empirical log-return series for 2019–2023.

Model performance is evaluated through a weighted score aggregating absolute deviations in kurtosis, standard deviation, skewness, and correlation:

$$\text{Score} = 0.35 \mid \Delta K_{\text{ng}} \mid +0.15 \mid \Delta K_{\text{el}} \mid +0.20 \mid \Delta S_{\text{ng}} \mid +0.20 \mid \Delta S_{\text{el}} \mid +0.10 \mid \Delta \rho \mid, \tag{18}$$

where $\Delta X = X^{\text{emp}} - X^{\text{model}}$. The weights reflect the empirical difficulty of reproducing natural gas kurtosis, the most prominent tail feature observed in the data [82–84].

Table 12 ranks the top five GARCH benchmarks according to the scoring function. Hybrid univariate specifications outperform bivariate ones, yet even the best variant yields a score of 1.032, which is three orders of magnitude larger than the GMM score of 0.001.

TABLE 12 | Top five GARCH benchmarks ranked by total score (lower is better).

Rank	Model	Type	Score
1	GARCH-GED + EGARCH-t	H-Univ.	**1.032**
2	EGARCH-GED + GARCH-t	H-Univ.	1.033
3	Heavy-tail ng + Normal	H-Univ.	1.104
4	Double EGARCH Extreme	Bivariate	1.104
5	GARCH Opposite Distrib.	H-Univ.	1.106
GMM			**0.001**

Note: For comparison purposes, the score of our GMM is also included. The bold value (GMM score = 0.001) indicates the best-performing model according to the scoring function defined in Equation (18).

4.2 | Comparative Performance: GMM Vs. GARCH

Table 13 reports a moment-by-moment comparison between the empirical distributions, the proposed GMM-based model, and the five best-performing GARCH benchmarks.

The proposed GMM representation exactly replicates the empirical first four moments and cross-market correlation.

TABLE 13 | Moment comparison: Empirical data, GMM model, and top five GARCH benchmarks.

Model	Natural gas				Electricity				
	μ	σ	S	K	μ	σ	S	K	ρ
Empirical data	2.47e-05	0.0705	0.4040	14.02	2.49e-04	0.1467	0.323	5.43	0.46
GMM	**0.000**	**0.0705**	**0.4040**	**14.02**	**0.000249**	**0.1467**	**0.323**	**5.43**	**0.45**
GARCH-GED + EGARCH-*t*	0.000	0.0445	0.000	11.23	0.000	0.1227	0.000	5.23	0.31
EGARCH-GED + GARCH-*t*	0.000	0.0465	0.000	11.23	0.000	0.1237	0.000	5.22	0.31
Heavy-tail ng + Normal	0.000	0.0465	0.000	11.23	0.000	0.1347	0.000	4.75	0.28
Double EGARCH extreme	0.000	0.0465	0.000	11.23	0.000	0.1337	0.000	4.75	0.28
GARCH opposite distrib.	0.000	0.0445	0.000	11.23	0.000	0.1337	0.000	4.74	0.28

Note: The bold values in the "GMM" row indicate the results of the proposed model.

In contrast, all GARCH-type models exhibit systematic deviations, with errors particularly concentrated in skewness, kurtosis, and correlation.

Three structural limitations clearly emerge:

4.2.1 | Skewness

All GARCH specifications using symmetric innovations (Normal, Student-*t*, GED) necessarily generate zero skewness. This contradicts the positive empirical skewness of both commodities ($S_{\text{ng}} = 0.404$, $S_{\text{el}} = 0.323$), which reflects the asymmetric nature of energy price shocks. Even Skewed-*t* variants can only partially address this issue and often suffer from estimation instability.

4.2.2 | Kurtosis

Even the best-performing model (GARCH-GED + EGARCH-*t*) underestimates natural gas kurtosis by roughly 20% ($K_{\text{ng}}^{\text{model}} = 11.23$ vs. $K_{\text{ng}}^{\text{emp}} = 14.02$). This bias reflects a fundamental limitation of volatility-driven models: extreme tail events in energy markets are not adequately captured by second-moment dynamics alone.

4.2.3 | Correlation

All GARCH specifications understate the empirical correlation $\rho = 0.46$, with the best model reaching only $\rho = 0.31$. Cholesky-based dependence relies on linear propagation of univariate shocks and cannot represent nonlinear or asymmetric co-movements between natural gas and electricity.

4.3 | Structural Interpretation

The persistent fitting errors of GARCH models are not a consequence of poor calibration, but arise from intrinsic structural constraints.

4.3.1 | Second-Moment Focus

GARCH frameworks are designed to capture conditional variance dynamics. Higher-order moments—particularly skewness and extreme kurtosis—are not directly modeled, but emerge only indirectly through the assumed innovation distribution.

4.3.2 | Linear Dependence Structures

Multivariate GARCH models impose linear dependence via covariance or Cholesky decompositions. These structures cannot replicate the nonlinear and geometry-induced relationships between natural gas and electricity that are revealed in the optimal-transport embedding space.

4.3.3 | Parametric Rigidity

Even flexible innovation laws (Skewed-*t*, GED) impose strong shape constraints that prevent the simultaneous recovery of skewness, kurtosis, and empirical correlation observed in energy markets.

4.4 | Summary

The GARCH benchmark confirms that volatility-based econometric models, despite their widespread use and strong theoretical foundations, are not able to replicate the distributional geometry, latent dependencies, and asymmetric interactions captured by the proposed information–geometric framework.

By operating in an optimal-transport embedding space and leveraging a Gaussian mixture representation, our methodology reconstructs the empirical distributions with much higher accuracy and reveals structural features that remain inaccessible to classical second-moment approaches.

5 | Concluding Remarks

In this study, we presented an innovative methodology to analyze the joint dynamics of natural gas and electricity prices by integrating graph representation learning tools, optimal transport theory and statistical mixture modeling.

The embedding of visibility graphs built from log-return time series of natural gas and electricity into a high-dimensional latent space, allowed us to compute an unbalanced Wasserstein barycenter that would enable us to construct an information-driven fusion of the input distributions, optimized through entropy maximization. This methodology has enabled a more adaptive, data-driven approach to balancing, in which asymmetries are not artifacts but essential features that enhance our understanding of market dynamics. By privileging information over symmetry, we discovered the latent structural dependencies that govern price movements.

Then, we approximated the unbalanced Wasserstein barycenter with a GMM, providing a flexible parametric representation of the common stochastic components. Our findings revealed that the use of the unbalanced Wasserstein barycenter facilitated the derivation of a central distribution that captures a prevailing common underlying stochastic structure driving the joint dynamics of natural gas and electricity prices. This central distribution serves as a foundational component of our analysis. In fact, we then extended this baseline by introducing a commodity-specific Gaussian component, allowing the model to simultaneously capture shared dynamics and market-specific behavior. The introduction of additional commodity-specific components in the unbalanced Wasserstein barycenter allowed for a refined examination of the price dynamics. These components, calibrated to reproduce the first four moments of the observed distributions of the log-returns of natural gas and electricity, as well as the empirical correlation between the two distributions, allowed for a more accurate modeling approach, thus providing a more detailed representation of the price dynamics and a tractable tool for understanding the joint price dynamics.

The empirical analysis, conducted on the Italian energy market, demonstrated the effectiveness of the proposed framework in capturing both distributional features and inter-market dependencies (see Figure 12). The analysis reveals that the entropy-optimal barycentric weights are $\lambda_{\mathrm{ng}} = 0.65$ for natural gas and $\lambda_{\mathrm{el}} = 0.35$ for electricity, highlighting a dominant role of natural gas market in the joint representation, a result that is fully consistent with market structure considerations.

The comparison with standard GARCH models confirmed that, while they provide a useful baseline for modeling energy market volatility, they struggle to match higher-order moments and joint tail dependence. These gaps highlight the need for more flexible solutions. Our proposed framework overcomes these limits, achieving a better fit to extreme features and cross-market asymmetries without relying solely on rigid parametric assumptions. Moreover, the model accurately reproduces the first four moments of the empirical log-return distributions, ensuring a level of distributional fidelity typically unattainable for classical GARCH-type specifications (see Tables 12 and 13).

The proposed methodology is fully market-agnostic and directly transferable to any pair of interconnected energy markets, provided that synchronized price time series are available. All steps of the pipeline are entirely data-driven and do not rely on country-specific assumptions.

The combination of optimal transport and statistical modeling could provide a sound basis for applications such as risk management, derivative pricing and policy evaluation in energy systems characterized by asymmetry and structural change. In this regard, three main directions will guide our future research.

The first direction involves extending the proposed methodology to a broader range of economic sectors beyond the energy sector. We believe that our approach can be useful in exploring the complex interconnections between the dynamics of energy commodity prices, such as oil, gas, and electricity, and the dynamics of macroeconomic variables that reflect the overall economic behavior. These macroeconomic variables include interest rates and stock market indices, which are often used as proxies to describe financial market behavior.

The second direction concerns a more methodological aspect. Specifically, we aim to expand the current framework by incorporating a systematic analysis of the transport plan associated with the unbalanced Wasserstein barycenter, offering a new lens for investigating the directional relationships across markets. By combining optimal transport geometry with network structures derived from market time series, it may be possible to discover more detailed paths of influence and propagation between interconnected systems. We believe that the combined interaction of these techniques will enable a deeper understanding of the common underlying dynamics. This improvement can significantly aid decision making in areas where understanding complex, high-dimensional data is critical.

Third, we note that the analysis developed here focuses on the distributional geometry of joint energy price dynamics, which is where the proposed information-driven framework reveals its full potential. Extending the model to also generate realistic temporal trajectories would require introducing a mechanism of mean reversion consistent with the latent geometric structure identified in this study. Mean-reverting behavior is a well-documented empirical feature of commodity prices [52], and incorporating it in a way that respects the nonlinear, asymmetric structure captured by our approach is therefore essential. Such an extension could rely on standard econometric formulations. Instead, our intention is to investigate mean-reverting dynamics using techniques that are conceptually aligned with those developed in the present work, building future trajectory models on the same information–geometric principles, graph-based representations, and distributional tools that underpin our current framework. This extension will be the object of future research.

Author Contributions

Carlo Mari: conceptualization, methodology, investigation, writing – original draft preparation. **Emiliano Mari:** conceptualization, methodology, software, validation. **Cristiano Baldassari:** conceptualization, methodology, visualization, software, data curation, writing – original draft preparation.

Acknowledgements

This research was carried out using the computational infrastructure of the Rome Technopole at the Unitus Datalab (https://datalab.unitus.it/). Open access publishing facilitated by Universita degli Studi della Tuscia, as part of the Wiley - CRUI-CARE agreement.

Funding

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data is available on www.mercatoelettrico.org.

References

1. M. W. Koczar, F. Jareño, and A. Escribano, "Dynamic Linkages and Contagion Effects: Analyzing the Linkages Between Crude Oil Prices, US Market Sector Indices and Energy Markets," *North American Journal of Economics and Finance* 74 (2024): 102247, https://doi.org/10.1016/j.najef.2024.102247.

2. W. Ding, Y. Xu, F. Gao, B. Wu, and Y. Fu, "Electricity Market Dynamics: The Influence of Fuel Prices and the Spillover Effect," *IOP Conference Series: Earth and Environmental Science* 615 (2020): 012079, https://doi.org/10.1088/1755-1315/615/1/012079.

3. J. M. Uribe, S. Mosquera-López, and O. J. Arenas, "Assessing the Relationship Between Electricity and Natural Gas Prices in European Markets in Times of Distress," *Energy Policy* 166 (2022): 113018, https://doi.org/10.1016/j.enpol.2022.113018.

4. G. Di Bella, M. Flanagan, K. Foda, et al., "Natural Gas in Europe: The Potential Impact of Disruptions to Supply," *Energy Economics* 138 (2024): 107777, https://doi.org/10.1016/j.eneco.2024.107777.

5. B. Zakeri, I. Staffell, P. E. Dodds, et al., "The Role of Natural Gas in Setting Electricity Prices in Europe," *Energy Reports* 10 (2023): 2778–2792, https://doi.org/10.1016/j.egyr.2023.09.069.

6. W. Roeger and P. J. Welfens, "Gas Price Caps and Electricity Production Effects in the Context of the Russo-Ukrainian War: Modeling and New Policy Reforms," *International Economics and Economic Policy* 19, no. 4 (2022): 645–673, https://doi.org/10.1007/s10368-022-00552-7.

7. T. D. Diagoupis, P. E. Andrianesis, and E. N. Dialynas, "A Planning Approach for Reducing the Impact of Natural Gas Network on Electricity Markets," *Applied Energy* 175 (2016): 189–198, https://doi.org/10.1016/j.apenergy.2016.05.006.

8. T. Xia, Q. Ji, and J. B. Geng, "Nonlinear Dependence and Information Spillover Between Electricity and Fuel Source Markets: New Evidence From a Multi-Scale Analysis," *Physica A* 537 (2020): 122298, https://doi.org/10.1016/j.physa.2019.122298.

9. R. Carmona and V. Durrleman, "Pricing and Hedging Spread Options," *SIAM Review* 45, no. 4 (2003): 627–685.

10. N. Frikha and V. Lemaire, "Joint Modelling of Gas and Electricity Spot Prices," *Applied Mathematical Finance* 20, no. 1 (2012): 69–93, https://doi.org/10.1080/1350486X.2012.658220.

11. R. F. Engle, C. W. J. Granger, and D. Kraft, "Combining Competing Forecasts of Inflation Using a Bivariate ARCH Model," *Journal of Economic Dynamics and Control* 8, no. 2 (1984): 151–165.

12. R. F. Engle and K. F. Kroner, "Multivariate Simultaneous Generalized ARCH," *Econometric Theory* 11, no. 1 (1995): 122–150.

13. R. S. Tsay, *Analysis of Financial Time Series*, 3rd ed. (Wiley, 2010).

14. C. Francq and J. M. Zakoian, *GARCH Models: Structure, Statistical Inference and Financial Applications*, 2nd ed. (Wiley, 2019).

15. K. M. Maribu, A. Galli, and M. Armstrong, "Valuation of Spark-Spread Options With Mean Reversion and Stochastic Volatility," *International Journal of Electronic Business Management* 5, no. 3 (2007): 173–181.

16. S. Song, X. Wang, and X. Zhang, "Valuation of Spread Options Under Correlated Skew Brownian Motions," *European Journal of Finance* 30, no. 5 (2023): 503–523, https://doi.org/10.1080/1351847X.2023.2202821.

17. T. Meyer-Brandis and M. Morgan, "A Dynamic Lévy Copula Model for the Spark Spread," in *Quantitative Energy Finance*, ed. F. Benth, V. Kholodnyi, and P. Laurence (Springer, 2014).

18. F. Mehrdoust and I. Noorani, "Valuation of Spark-Spread Option Written on Electricity and Gas Forward Contracts Under Two-Factor Models With Non-Gaussian Lévy Processes," *Computational Economics* 61 (2023): 807–853, https://doi.org/10.1007/s10614-021-10232-4.

19. C. Arsene and A. Parisio, "Deep Convolutional Neural Networks for Short-Term Multi-Energy Demand Prediction of Integrated Energy Systems," *International Journal of Electrical Power & Energy Systems* 160 (2024): 110111, https://doi.org/10.1016/j.ijepes.2024.110111.

20. A. Bâra and S. V. Oprea, "Predicting Day-Ahead Electricity Market Prices Through the Integration of Macroeconomic Factors and Machine Learning Techniques," *International Journal of Computational Intelligence Systems* 17 (2024): 17, https://doi.org/10.1007/s44196-023-00387-3.

21. M. Polson and V. Sokolov, "Deep Learning for Energy Markets," *Applied Stochastic Models in Business and Industry* 36 (2020): 195–209, https://doi.org/10.1002/asmb.2518.

22. J. Voit, *The Statistical Mechanics of Financial Markets* (Springer, 2005).

23. A. Bücher and J. Segers, "Inference for Heavy Tailed Stationary Time Series Based on Sliding Blocks," *Electronic Journal of Statistics* 12, no. 1 (2018): 1098–1125, https://doi.org/10.1214/18-EJS1415.

24. C. Mari and E. Mari, "Deep Learning Based Regime-Switching Models of Energy Commodity Prices," *Energy Systems* 14, no. 4 (2023): 913–934, https://doi.org/10.1007/s12667-022-00515-6.

25. Y. Wei, X. Zhang, Y. Shi, et al., "A Review of Data-Driven Approaches for Prediction and Classification of Building Energy Consumption," *Renewable and Sustainable Energy Reviews* 82, no. 1 (2018): 1027–1047, https://doi.org/10.1016/j.rser.2017.09.108.

26. L. V. Kantorovich, "Mathematical Methods of Organizing and Planning Production," *Management Science* 6, no. 4 (1960): 366–422.

27. L. N. Vaserstein, "Markov Processes Over Denumerable Products of Spaces, Describing Large Systems of Automata," *Problemy Peredachi Informatsii* 5, no. 3 (1969): 64–72.

28. F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The Graph Neural Network Model," *IEEE Transactions on Neural Networks* 20, no. 1 (2009): 61–80.

29. M. Agueh and G. Carlier, "Barycenters in the Wasserstein Space," *SIAM Journal on Mathematical Analysis* 43, no. 2 (2011): 904–924.

30. L. Chizat, G. Peyré, B. Schmitzer, and F. X. Vialard, "Unbalanced Optimal Transport: Dynamic and Kantorovich Formulations," *Journal of Functional Analysis* 274, no. 11 (2018): 3090–3123.

31. M. Liero, A. Mielke, and G. Savaré, "Optimal Entropy-Transport Problems and a New Hellinger–Kantorovich Distance Between Positive Measures," *Inventiones Mathematicae* 211 (2018): 969–1117.

32. D. W. Mimouni, P. Malisani, J. Zhu, and W. de Oliveira, "Computing Wasserstein Barycenters via Operator Splitting: The Method of

Averaged Marginals," *SIAM Journal on Mathematics of Data Science* 6 (2024): 1000–1026.

33. L. Chizat, "Doubly Regularized Entropic Wasserstein Barycenters," preprint, arXiv, 2303.11844, 2023.

34. M. Gazdieva, J. Choi, A. Kolesov, J. Choi, P. Mokrov, and A. Korotin, "Robust Barycenter Estimation Using Semi-Unbalanced Neural Optimal Transport," preprint, arXiv:2410.03974 2024.

35. R. B. Cleveland, W. S. Cleveland, J. E. McRae, and I. J. Terpenning, "STL: A Seasonal-Trend Decomposition Procedure Based on LOESS," *Journal of Official Statistics* 6, no. 1 (1990): 3–33.

36. L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, "From Time Series to Complex Networks: The Visibility Graph," *Proceedings of the National Academy of Sciences* 105 (2008): 4972–4975.

37. J. Zhu and D. Wei, "Analysis of Stock Market Based on Visibility Graph and Structure Entropy," *Physica A: Statistical Mechanics and Its Applications* 576 (2021): 126036.

38. C. Mari, T. Laureti, and C. Baldassari, "Group Detection in Energy Commodity Markets Through Manifold-Informed Wasserstein Barycenter," *Quality & Quantity* 59, no. 3 (2025): 2197–2227.

39. H. Cai, V. W. Zheng, and K. C. C. Chang, "A Comprehensive Survey of Graph Embedding: Problems, Techniques, and Applications," *IEEE Transactions on Knowledge and Data Engineering* 30, no. 9 (2018): 1616–1637.

40. B. Rozemberczki and R. Sarkar, "Diffusion Embeddings for Network Analysis," KDD Proceedings 2020.

41. R. Flamary, N. Courty, A. Gramfort, et al., "POT: Python Optimal Transport," *Journal of Machine Learning Research* 22, no. 78 (2021): 1–8.

42. C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal* 27 (1948): 379–423.

43. G. McLachlan and D. Peel, *Finite Mixture Models. Wiley Series in Probability and Statistics* (John Wiley & Sons, 2000).

44. C. Mari and C. Baldassari, "Unsupervised Expectation-Maximization Algorithm Initialization for Mixture Models: A Complex Network-Driven Approach for Modeling Financial Time Series," *Information Sciences* 617 (2022): 1–16, https://doi.org/10.1016/j.ins.2022.10.073.

45. D. Böhning and W. Seidel, "Recent Developments in Mixture Models," *Computational Statistics & Data Analysis* 41 (2003): 349–357, https://doi.org/10.1016/S0167-9473(02)00161-5.

46. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood From Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B, Statistical Methodology* 39, no. 1 (1977): 1–22, https://doi.org/10.1111/j.2517-6161.1977.tb01600.x.

47. Y. Le, J. Wen, Y. Wu, J. Liu, and Y. Zhu, "Investigating Factors Influencing Oil Volatility: A GARCH-MIDAS Model Analysis," *Frontiers in Energy Research* 12 (2024): 1392905, https://doi.org/10.3389/fenrg.2024.1392905.

48. K. Janda and P. Kravec, "VECM Modelling of the Price Dynamics for Fuels," in *Agricultural Commodities and Biofuels* (ZBW-Leibniz Information Centre for Economics, 2022).

49. S. Seyfi, A. Sharifi, and H. Arian, "Portfolio Value-At-Risk and Expected-Shortfall Using an Efficient Simulation Approach Based on Gaussian Mixture Model," *Mathematics and Computers in Simulation* 190 (2021): 1056–1079, https://doi.org/10.1016/j.matcom.2021.05.029.

50. I. Morkūnaitė, D. Celov, and R. Leipus, "Evaluation of Value-At-Risk (VaR) Using the Gaussian Mixture Models," *Research in Statistics* 2, no. 1 (2024): 2346075, https://doi.org/10.1080/27684520.2024.2346075.

51. J. H. Claver, T. Emerson, and S. Felix, "Multivariate Option Pricing With Gaussian Mixture Distributions and Mixed Copulas," *Journal of Mathematics and Statistics* 19, no. 1 (2023): 1–12, https://doi.org/10.3844/jmssp.2023.1.12.

52. H. Geman, *Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy* (Wiley, 2005).

53. S. Seabold and J. Perktold, "Statsmodels: Econometric and Statistical Modeling With Python," *SciPy* 7, no. 1 (2010): 92–96.

54. L. Lacasa, B. Rozemberczki, and R. Sarkar, "Fast Sequence-Based Embedding With Diffusion Graphs," in *International Workshop on Complex Networks* (Springer International Publishing, 2018), 99–107.

55. B. Rozemberczki, O. Kiss, and R. Sarkar, "Karate Club: An API Oriented Open-Source Python Framework for Unsupervised Learning on Graphs," in *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (ACM (Association for Computing Machinery), 2020), 3125–3132.

56. K. Sheppard, "Arch: ARCH Models in Python," https://arch.readthedocs.io.

57. F. Pedregosa, G. Varoquaux, A. Gramfort, et al., "Scikit-Learn: Machine Learning in Python," *Journal of Machine Learning Research* 12 (2011): 2825–2830.

58. E. B. Dagum and S. Bianconcini, *Seasonal Adjustment Methods and Real Time Trend-Cycle Estimation* (Springer, 2016).

59. X. Lan, H. Mo, S. Chen, Q. Liu, and Y. Deng, "Fast Transformation From Time Series to Visibility Graphs," *Chaos (Woodbury, N.Y.)* 25, no. 8 (2015): 083105, https://doi.org/10.1063/1.4927835.

60. Y. Hu, "Algorithms for Visualizing Large Networks," *Combinatorial Scientific Computing* 5 (2011): 180–186.

61. A. L. Barabási, *Network Science* (Cambridge University Press, 2016).

62. C. Mari and C. Baldassari, "Understanding the Complexities of the Fine Structure of Interest Rates: A Wasserstein Barycenter Learning Approach," *Neural Computing and Applications* 36, no. 31 (2024): 19291–19305.

63. C. Mari and C. Baldassari, "Optimization of Mixture Models on Time Series Networks Encoded by Visibility Graphs: An Analysis of the US Electricity Market," *Computational Management Science* 20, no. 1 (2023): 28.

64. C. Mari and C. Baldassari, "A Graph-Based Superframework for Mixture Model Estimation Using EM: An Analysis of US Wholesale Electricity Markets," *Neural Computing and Applications* 35, no. 20 (2023): 14867–14883.

65. B. Luque, L. Lacasa, F. Ballesteros, and J. Luque, "Horizontal Visibility Graphs: Exact Results for Random Time Series," *Physical Review E* 80, no. 4 (2009): 046103.

66. S. Misra, M. W. Fisher, S. Backhaus, R. Bent, M. Chertkov, and F. Pan, "Optimal Compression in Natural Gas Networks: A Geometric Programming Approach," *IEEE Transactions on Control of Network Systems* 2, no. 1 (2014): 47–56.

67. J. Zhang, X. Chen, H. Wang, C. Gu, and H. Yang, "Visibility Graph Approach to Extreme Event Series," *Chinese Physics B* 32, no. 10 (2023): 100505.

68. L. Kantorovitch, "On the Translocation of Masses," *Management Science* 5, no. 1 (1958): 1–4.

69. C. Villani, *Optimal Transport: Old and New* (Springer, 2008).

70. G. Peyré and M. Cuturi, "Computational Optimal Transport: With Applications to Data Science," *Foundations and Trends in Machine Learning* 11 (2019): 355–607.

71. S. Kolouri, S. Park, M. Thorpe, D. Slepčev, and G. K. Rohde, "Optimal Mass Transport: Signal Processing and Machine-Learning Applications," *IEEE Signal Processing Magazine* 4, no. 4 (2017): 43–45, https://doi.org/10.1109/MSP.2017.2695801.

72. J. Bigot, "Statistical Data Analysis in the Wasserstein Space," *ESAIM: Proceedings and Surveys* 68 (2020): 1–19.

73. C. Tofallis, “Selecting the Best Statistical Distribution Using Multiple Criteria,” *Computers & Industrial Engineering* 54, no. 3 (2008): 690–694, https://doi.org/10.1016/j.cie.2007.07.016.

74. M. S. Yang, C. Y. Lai, and C. Y. Lin, “A Robust EM Clustering Algorithm for Gaussian Mixture Models,” *Pattern Recognition* 45 (2012): 3950–3961.

75. E. Shireman, D. Steinley, and M. Brusco, “Examining the Effect of Initialization Strategies on the Performance of Gaussian Mixture Modeling,” *Behavior Research Methods* 49 (2017): 282–293, https://doi.org/10.3758/s13428-015-0697-6.

76. W. Kwedlo, “A New Random Approach for Initialization of the Multiple Restart EM Algorithm for Gaussian Model-Based Clustering,” *Pattern Analysis and Applications* 18, no. 4 (2015): 757–770, https://doi.org/10.1007/s10044-014-0441-3.

77. J. You, Z. Li, and J. Du, “A New Iterative Initialization of EM Algorithm for Gaussian Mixture Models,” *PLoS One* 18, no. 4 (2023): e0284114, https://doi.org/10.1371/journal.pone.0284114.

78. J. Kuha, “AIC and BIC: Comparisons of Assumptions and Performance,” *Sociological Methods & Research* 33, no. 2 (2004): 188–229, https://doi.org/10.1177/0049124103262065.

79. J. Nemec and A. F. Linnell Nemec, “Mixture Models for Studying Stellar Populations. I. Univariate Mixture Models, Parameter Estimation, and the Number of Discrete Population Components,” *Publications of the Astronomical Society of the Pacific* 103, no. 659 (1991): 95–121, https://doi.org/10.1086/132800.

80. D. B. Nelson, “Conditional Heteroskedasticity in Asset Returns: A New Approach,” *Econometrica* 59, no. 2 (1991): 347–370.

81. G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. (Johns Hopkins University Press, 2013).

82. A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management: Concepts, Techniques and Tools*, Revised ed. (Princeton University Press, 2015).

83. P. Jorion, *Value at Risk: The New Benchmark for Managing Financial Risk*, 3rd ed. (McGraw-Hill, 2007).

84. A. J. Patton, “Volatility Forecast Comparison Using Imperfect Volatility Proxies,” *Journal of Econometrics* 160, no. 1 (2011): 246–256.