

Fast Times, Slow Times: Timescale Separation in Financial Timeseries Data

Jan Rosenzweig

Abstract

Financial time series exhibit multiscale behavior, with interaction between multiple processes operating on different timescales. This paper introduces a method for separating these processes using variance and tail stationarity criteria, framed as generalized eigenvalue problems. The approach allows for the identification of slow and fast components in asset returns and prices, with applications to parameter drift, mean reversion, and tail risk management. Empirical examples using currencies, equity ETFs and treasury yields illustrate the practical utility of the method.

1 Stationarity in Finance

The study of multiscale processes is well-established in physics, with notable examples ranging from fluid dynamics to protein folding, where fast and slow dynamics interact [1].

The typical physical example involves fast oscillations inside a slowly moving envelope, where the cumulative effect of fast oscillations drives changes in the envelope [1, 2].

In finance, we can see multiple examples of similar multiscale behavior. For example, the S&P 500 exhibits mean reversion over milliseconds (due to low-latency arbitrage between e-mini S&P futures, ETFs and stock baskets) and over years (relative to gold or bonds).

A notable example of feedback from high frequency to macro is of course the Flash Crash of 2010, where multiple feedback loops in high and mid frequency trading lead to the S&P 500 losing 9% of its value, only to recover later [3].

Therefore identification of such processes operating on different timescales, and the nature of their interaction, is of clear importance both to market practitioners and to economists.

The key questions in this context are:

- Can we separate these processes?
- Can we estimate their relaxation timescales?
- Can we assess stationarity in a meaningful way?

This paper addresses these questions by introducing a framework for timescale separation in financial time series, building on methods from signal processing and variational principles.

On the trading side, financial strategies rely on calibrated parameters that drift over time, leading to strategy degradation.

Two types of stationarity are relevant in this context:

- **Stationarity of Returns:** Implies stable distribution of returns, leading to stable parameters.
- **Stationarity of Prices:** Implies mean reversion in price levels.

The same mathematical tools can be applied to both, though their interpretations differ.

2 Types of Stationarity

The textbook definition of stationary processes involves the stationarity of the underlying distribution.

While this definition is attractive, it does not lend itself straightforwardly to data driven analysis. Primarily, this is due to the fact that we only have a finite amount of data available, and that any estimation of the underlying distribution is therefore fuzzy and imprecise.

Rather than aiming for the entire distribution, it is then more practical to focus on particular aspects of the distribution.

The initial promising target is the covariance matrix, i.e. measuring the stationarity of variances and covariances in the basket.

However non-Gaussianity of financial time series means that we have to look further than just to the Gaussian MLE, into the stationarity of the tail behaviour of the relevant timeseries.

We therefore consider two complimentary approaches to timescale separation:

- **Variance Timescales:** Stationarity and drift in the second moment.
- **Tail Timescales:** Stability in higher moments, relevant for tail risk.

2.1 Variance Timescales

Let \mathbf{X}_t be an n -dimensional column vector process (e.g., prices or returns). We seek a matrix \mathbf{W} of n -dimensional column weight vectors \mathbf{w}_i such that the drift of variance is minimized for a fixed unit of variance:

$$\frac{d}{dt} \text{Var}(\mathbf{X}_t \mathbf{w}_i^T) \rightarrow \min, \quad i = 1..n \quad (1)$$

$$\text{Var}(\mathbf{W} \mathbf{X}_t^T) = \mathbf{1} \quad (2)$$

Here, (1) and (2) are both instantaneous at time t and hence they do not contradict each other.

This leads to the generalized eigenvalue problem:

$$\mathbb{E}[(d\mathbf{X}_t)^T \mathbf{X}_t + \mathbf{X}_t^T d\mathbf{X}_t] \mathbf{w} = \beta \mathbb{E}[\mathbf{X}_t^T \mathbf{X}_t] \mathbf{w} \quad (3)$$

$$\mathbb{E}[\mathbf{w}_i \mathbf{X}_t^T \mathbf{X}_t \mathbf{w}_j^T] = \delta_{ij} \quad (4)$$

where β is the Lagrange multiplier.

Discretising $d\mathbf{X}_t$ as

$$d\mathbf{X}_t \approx \frac{1}{T} (d\mathbf{X}_{t+T} - d\mathbf{X}_t),$$

(3) becomes

$$\mathbf{C}(t, T) \mathbf{w} = \lambda \mathbf{C}(t, 0) \mathbf{w} \quad (5)$$

where $\mathbf{C}(t, T)$ is the autocovariance of \mathbf{X}_t with lag T ,

$$\mathbf{C}(t, T) = \frac{1}{2} \mathbb{E}[\mathbf{X}_{t+T}^T \mathbf{X}_t + \mathbf{X}_t^T \mathbf{X}_{t+T}]$$

and

$$\lambda = \frac{1 + \beta}{2T}$$

The eigenvalue problem (5) is known as tICA [4], although our derivation is different from the usual tICA derivation; the usual derivation assumes that the discretised process follows a Hidden Markov Model

$$\mathbf{X}_{t+T} = \mathbf{A} \mathbf{X}_t + \epsilon_t \quad (6)$$

for some unknown linear operator \mathbf{A} and noise ϵ_t . Multiplying both sides of (6) by \mathbf{X}_t^T and taking the expectation, this yields

$$\mathbb{E}[\mathbf{X}_t^T \mathbf{X}_{t+T}] = \mathbf{A} \mathbf{C}(t, 0)$$

where the autocovariance term $\mathbb{E}[\mathbf{X}_t^T \mathbf{X}_{t+T}]$ is usually symmetrized to give the eigenvalue problem (5).

Note that the matrix $\mathbf{C}(t, 0)$ on the right hand side of equation (5) is the covariance matrix, and as such it is real, symmetric and positive definite. This makes (5) straightforward to solve numerically, using e.g. LAPACK's gvd driver or its various interfaces such as Python `scipy.linalg.eigh`.

The eigenvectors of (5) correspond to the directions of fastest (slowest) decay in non-stationarity. The eigenvalues λ_i are their autocorrelations, as seen from (5), and are therefore in the range $[-1, 1]$. They are related to decay time scales through

$$t_i = \frac{-2T}{\lambda_i - 1} \quad (7)$$

2.2 Tail Timescales

For tail timescales, we minimize the drift of higher order moments [6]. The moment of order $2k$ is given as

$$M_{2k} = \frac{1}{2k} \mathbb{E}(\mathbf{X}_t \mathbf{w}^T)^{2k}$$

and the minimization problem becomes

$$\frac{d}{dt} \frac{1}{2k} \mathbb{E}(\mathbf{X}_t \mathbf{w}_i^T)^{2k} \rightarrow \min, \quad i = 1..n \quad (8)$$

$$\mathbb{E}[\mathbf{w}_i \mathbf{X}_t^T \mathbf{X}_t \mathbf{w}_j^T] = \delta_{ij} \quad (9)$$

The problem (8), (9) is nonlinear and it does not reduce to a simple eigenvalue problem. The constrained minimum satisfies

$$(2k - 1) \mathbb{E}[(\mathbf{X}_t \mathbf{w}^T)^{2k-2} (\mathbf{d} \mathbf{X}_t \mathbf{w}^T) \mathbf{X}_t] + \mathbb{E}[(\mathbf{X}_t \mathbf{w}^T)^{2k-1} \mathbf{d} \mathbf{X}_t] = \beta \mathbf{w} \quad (10)$$

$$\mathbf{w}^T \mathbf{w} = 1 \quad (11)$$

where β is the Lagrange multiplier as before.

Multiplying the right hand side of (10) by \mathbf{w}^T and using (11), we can solve for β as

$$\beta = 2k \mathbb{E}[(\mathbf{X}_t \mathbf{w}^T)^{2k-1} (\mathbf{d} \mathbf{X}_t \mathbf{w}^T)] = 2k \frac{d}{dt} M_{2k} \quad (12)$$

From (12), we get the generalization of the equivalent of λ from the previous section as

$$\lambda = \mathbb{E}[(\mathbf{X}_t \mathbf{w}^T)^{2k-1} (\mathbf{X}_{t+T} \mathbf{w}^T)], \quad (13)$$

so the eigenvalue is now the *tail autocorrelation* of order k , i.e. the tail correlation of order k between \mathbf{X}_t and \mathbf{X}_{t+T} [6].

Plugging (12) into (9), we get

$$(2k - 1) \mathbb{E}[(\mathbf{X}_t \mathbf{w}^T)^{2k-2} (\mathbf{d} \mathbf{X}_t \mathbf{w}^T) \mathbf{X}_t] + \mathbb{E}[(\mathbf{X}_t \mathbf{w}^T)^{2k-1} \mathbf{d} \mathbf{X}_t] = 2k \mathbb{E}[(\mathbf{X}_t \mathbf{w}^T)^{2k-1} (\mathbf{d} \mathbf{X}_t \mathbf{w}^T)] \mathbf{w} \quad (14)$$

Numerically, it is easier to solve (8),(9) using the Fixed Point Iteration of [5], known as FastICA. The iteration procedure is

$$\begin{aligned} \mathbf{w} &\leftarrow \mathbb{E}[\mathbf{X}_t d(\mathbf{w}^T \mathbf{X}_t)^{2k-1})] - (2k - 1) \mathbb{E}[(\mathbf{w}^T \mathbf{X}_t)^{2k-2})] \mathbf{w} \\ \mathbf{w} &\leftarrow \mathbf{w} / |\mathbf{w}| \end{aligned} \quad (15)$$

when iterating a single component, or

$$\begin{aligned} \mathbf{W} &\leftarrow \mathbb{E}[\mathbf{X}_t d(\mathbf{W}^T \mathbf{X}_t)^{2k-1})] - (2k - 1) \mathbb{E}[(\mathbf{W}^T \mathbf{X}_t)^{2k-2})] \mathbf{W} \\ \mathbf{W} &\leftarrow (\mathbf{W} \mathbf{X}_t^T \mathbf{X}_t \mathbf{W}^T)^{-1/2} \mathbf{W} \end{aligned} \quad (16)$$

when iterating all the components in parallel.

3 Empirical Examples

We have applied the time-driven decomposition to three separate baskets, namely the G10 currencies, five multifactor equity ETFs, and US Treasuries.

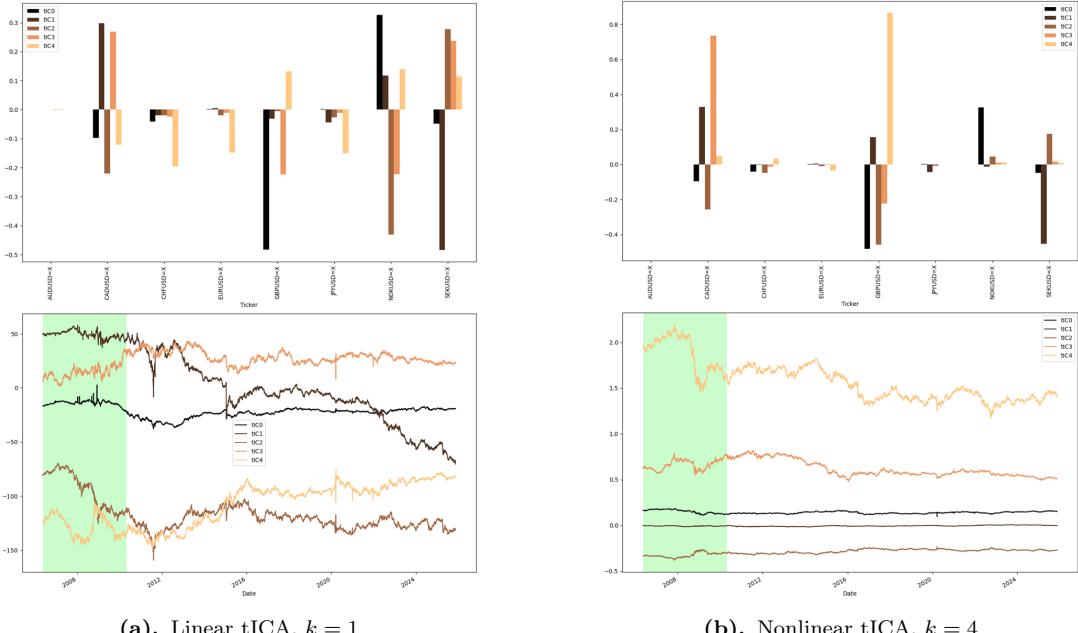
For G10 currencies, we extracted the five slowest components from 2006–2010 and projected them onto 2010–2025.

For factor ETFs (IFSU, IUSZ, IUVL, IUMO, IUQA), we extracted the five slowest components over 2021–2022 and projected them through 2022–2025.

Finally, for US Treasuries, we fitted over 1994–1998, and continued over 1998–2025.

We fitted each basket with the liner tICA ($k = 1$, section 2.1) and nonlinear tICA ($k = 4$, section 2.2).

The results are shown in Figures 1, 2, 3.



(a). Linear tICA, $k = 1$

(b). Nonlinear tICA, $k = 4$

Figure 1: G10 currencies: Blind extraction of slowest components over 2006-2010 (green background), continuation over 2010-2025 (white background). Weights are rescaled to unit gross exposure.

While retaining some similarity between the weight profiles in each case, the timeseries for nonlinear tICA results are considerably different from the corresponding linear decompositions. Increasing the exponent k further, and thus increasing the penalty to tails relative to volatility, did not produce further meaningful changes.

This seems to imply that there are overall two choices for timescale separation; namely linear, governed by volatility, and nonlinear, governed by the tails.

We can note that the relaxation timescales, at least for the slowest components, appear to be remarkably robust out of sample, with no visible deterioration accompanying the transition from in sample to out of sample.

Even more remarkably, where higher frequencies visibly enter the slow components such as, say, due to the global financial crisis 2008-2010 or geopolitical events in 2023, the relevant components revert back to their original, slow timescale once the relevant events pass.

This robustness does not carry over to orthogonality of components, which decays as expected out of sample.

4 Conclusion

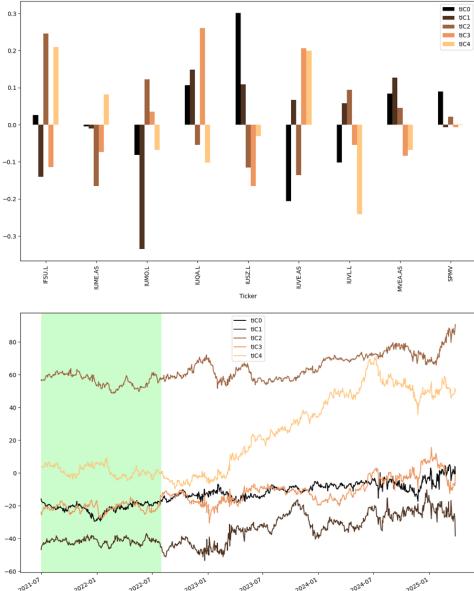
Financial time series contain a hierarchy of processes: stationary, slow, and fast. Using generalized eigenvalue problems, we can separate these components, providing insight into parameter drift, mean reversion regimes and/or tail risk stability.

The method is computationally tractable and can be applied to diverse asset classes, offering a powerful tool for risk and strategy management. The conceptual framework, moving from small to large structures, mirrors approaches used in physical sciences.

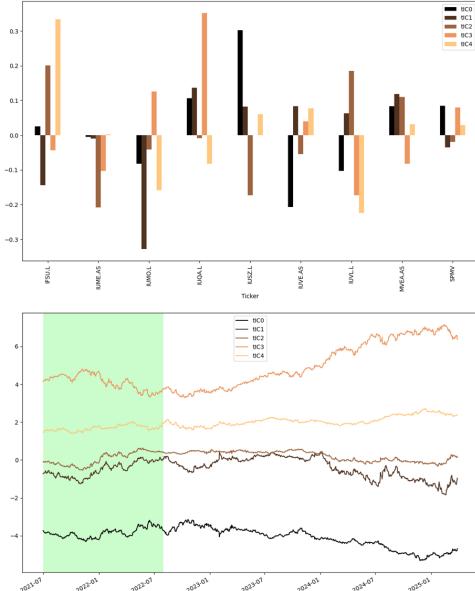
The remarkable property of the methods described in this paper is their ability to persist the in-sample time scales of the selected components out of sample.

References

- [1] Hinch, E.J. (1991). Perturbation Methods. <https://doi.org/10.1017/CBO9781139172189>, Cambridge University Press.



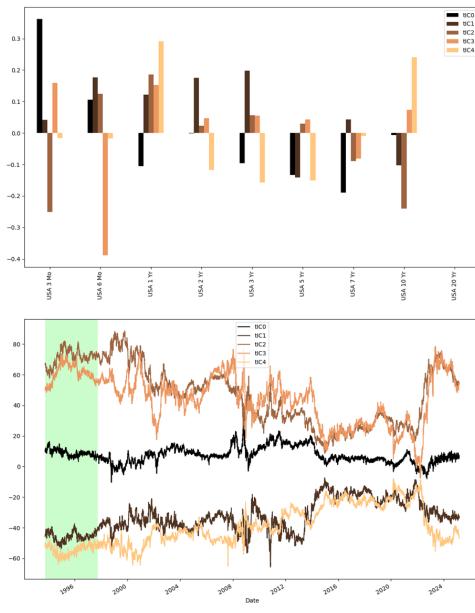
(a). Linear tICA, $k = 1$



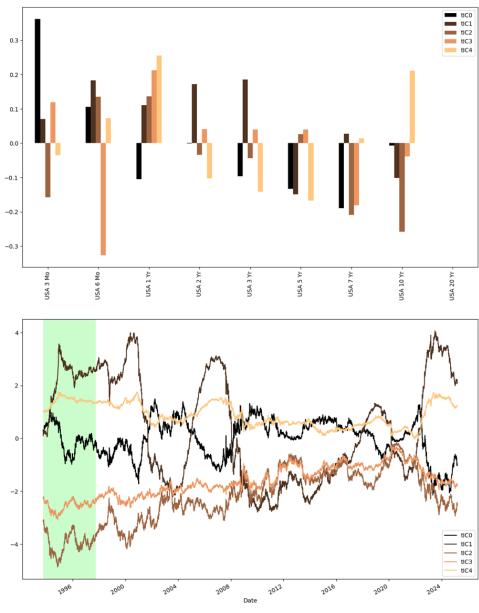
(b). Nonlinear tICA, $k = 4$

Figure 2: Factor ETF: Blind extraction of slowest components over 2021-2022 (green background), continuation over 2022-2025 (white background). Weights are rescaled to unit gross exposure.

- [2] Pérez-Hernández, G., Paul, F., Giorgino, T., De Fabritiis, G., Noé, F. (2013) Identification of slow molecular order parameters for Markov model construction. *J. Chem. Phys* 139, 015102. <https://doi.org/10.1063/1.4811489>
- [3] 2010 flash crash, Wikipedia https://en.wikipedia.org/wiki/2010_flash_crash
- [4] Schmid, P.J. (2010) Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics* 656:5–28.
- [5] Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks* 10 (3): 626–634. <https://www.cs.helsinki.fi/u/ahyvarin/papers/TNN99new.pdf>
- [6] Rosenzweig, J. (2023) A tale of tail covariances. <https://arxiv.org/abs/2302.13646>



(a). Linear tICA, $k = 1$



(b). Nonlinear tICA, $k = 4$

Figure 3: US Treasuries: Blind extraction of slowest components over 1994-1998 (green background), continuation over 1998-2025 (white background). Weights are rescaled to unit gross exposure.