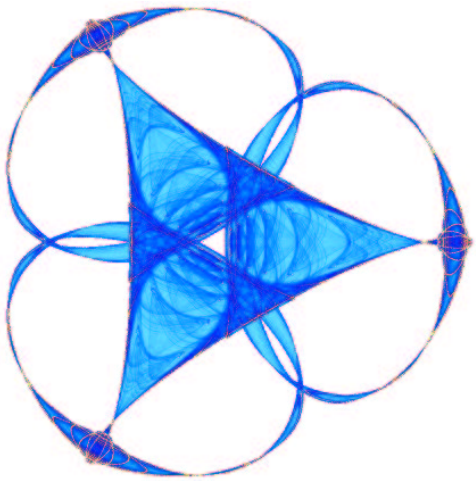
RANDOMIZED VOLATILITY ESTIMATION FROM SEMIMARTINGALES

By

Enrico Capobianco

IMA Preprint Series # 2020

(February 2005)



INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA
514 Vincent Hall
206 Church Street S.E.
Minneapolis, Minnesota 55455–0436

Phone: 612/624-6066 Fax: 612/626-7370
URL: http://www.ima.umn.edu

Randomized Volatility Estimation from Semimartingales

*Enrico Capobianco**
Department of Biomedical Engineering, Boston University
ecapob@bu.edu

Abstract

One established fact in financial economics and mathematics is the convergence of realised to integrated volatility according to the quadratic variation principle. When computed in general semimartingale asset price models, the cumulative squared high frequency returns represent consistent estimators of the integrated volatility. Both time and frequency domain estimators are available for solving what, in an unifying approach, could be considered an inverse problem, the recovery of latent volatility from the realizations of observable return processes. Since the relation between realised and integrated volatility implies that one is transformed into the other with noise, we work in a simulated environment of Brownian motion paths for exemplifying the semimartingale context and produce randomized estimators for the volatility. With the support of experimental evidence, we can show the consistency of time- and frequency-based volatility estimators and their speed of convergence to the quadratic variation limit.

Keywords: *Semimartingales; Signal Processing; Periodogram Analysis; Volatility Recovery*

1 Introduction

By casting volatility processes in inverse problem representation, one may state that the general solution which is needed is recovering the latent volatility from the realizations of the observable return process, or otherwise extracting an informative signal from noisy realised volatilities.

The so-called realised and integrated volatility have been recently proposed by Andersen *et al.* (2001a,b) and Barndorff-Nielsen and Shephard (2001) as means for obtaining a more accurate estimate of the volatility function in the presence of very high frequency observations.

Simply speaking, the computed realised volatility suggests an approximation to its integral counterpart over a certain fixed time interval. They are obtained by averaging a certain number of high frequency values within an established time interval of observation.

The realised volatility estimator of the integrated volatility produces an increasingly more accurate measure in the presence of high frequency observations. In other words, more finely sampled returns within a fixed time interval make the realised volatilities free of measurement error in the limit.

From the theoretical standpoint, it is according to the quadratic variation principle (see Karatzas and Shreve (1988), among many other references) that the convergence of realised to integrated volatility holds. But it remains one aspect which deserves special care; in finite-sample investigations, the noise plays of course a role, detrimental for the accuracy of the estimates.

*The paper has been written while the author was a member of the Mathematical Sciences Research Institute, Berkeley, CA 94720, US.

A different characterization of the general quadratic variation principle and related estimation issues, is offered by some studies (Dzhaparidze and Spreij, 1994; Malliavin and Mancino, 2002) which are based on frequency domain consistent estimators. In our view, this is another aspect which has been not yet evaluated by researchers involved in volatility studies, and that can represent an opportunity for further investigation.

We thus aim to compare the performance of time- and frequency-based volatility estimators in terms of convergence speed from randomized (i.e., simulation-based) experiments. With this goal in mind, we adopt a semimartingale approach and conduct inference from samples obtained by standard Brownian motion paths (and their increment processes) sampled at different grids.

2 Semimartingales

Semimartingales (Shiryaev, 1999) represent a very general class of stochastic processes suitable for asset price models. A stochastic process $\{X(t)_{t\geq 0}\}$ allows for the following classical continuous-time Doob-Meyer decomposition:

$$X_t = X_0 + A_t + M_t \tag{1}$$

with X_0 the baseline process value, A_t a predictable and bounded variation component, i.e., such that $\int_0^t \mid dA_s(w) \mid < \infty$ for $t > 0$ and $\omega \in \Omega$, and with M_t as a local martingale.

The probabilistic framework underlying this decomposition consists of a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$, with the process $\{X(t)_{t\geq 0}\}$ adapted to $(\mathcal{F}_t)_{t\geq 0}$ and *cadlag* (right-continuous and with limits from the left).

Among standard conditions, the σ-algebra $\mathcal{F}$ is P-complete, and the right continuous flow of $\sigma-$algebras (or information) $(\mathcal{F}_t)_{t\geq 0}$ contains all the sets in $\mathcal{F}$ of P-probability zero.

The semimartingale class contains discrete-time sequences $[(X_n)_{n\geq 0}]$ of the given process, but also diffusions and other processes. It is instead excluded the fractional Brownian motion $B^H = \{B^H(t)_{t\geq 0}\}$, except when the Hurst parameter $H \in (0,1)$ is $\frac{1}{2}$, e.g., for which it corresponds to the Brownian motion.

We aim to design a quite general model, by casting a volatility deconvolution problem within this probabilistic framework; this set-up could be thought as typical of a linear inverse problem.

Inverse problems arise in many practicl situations, where objects such as images or signals are only indirectly observed through the dynamics of correlated phenomena. The modeling and development of reconstruction algorithms represent the crucial step, and when the addressed context is that of a volatility recovery problem, we show that quite natural instruments exist to accomplish the task with computational efficiency and speed. development of reconstruction

3 Volatility Inverse Problem

Consider a volatility curve as a function f belonging to a certain function space, say $L_2(R)$ or other less homogeneous spaces, for instance Besov classes. A continuous-time model that can be set from sample realizations of return generating stochastic processes, and this model represents the volatility recovery problem by:

$$dy_t = (Kf)_t dt + \epsilon dW_t \tag{2}$$

where y_t is a return realization path, f_t is the unknown volatility function underlying this path and determining its dynamics, and the linear operator K stands for the convolution dynamics which apply from several sources of information, together with the influence of a W_t Wiener process (or Brownian motion).

A discrete-time corresponding model is described as:

$$y_n = (Kz)_n + \xi_n \tag{3}$$

with y_n and z_n being, respectively, the financial return and the volatility discretized sequences, while $\xi_n \sim N(0, var_\xi)$.

When within such model one assing the variables f_t, or z_n, to the volatility σ_t, or σ_n, one can easily establish the relationship between integrated, say σ^I, and realised, say σ^R, volatilities, where the latter quantity represents a noisy measure of the former (at any time point), i.e.:

$$\sigma^R = \sigma^I + \epsilon \tag{4}$$

The volatilities underlying the observed return process result linked to each other through the quadratic variation principle, as later explained in detail. Through its application, one can show the consistency of the realised volatility estimator.

But it is also worth to investigate the fact that given this general setting, one seeks pointwise estimates of the volatility function from indirect measurements, i.e., the noisy observations of returns. This is close to the idea of signal extraction, of course.

In this case we don't know the form of the operator K; thus, one may decide to turn to classical non-parametric estimation techniques like kernel or spline methods, or instead adopt wavelet-type estimation approaches.

One important methodological aspect concerns possible forms of regularization which are usually required when dealing with ill-posed problems. This in turn requires a specification of the functional class to which f belongs, thus implying assumptions on the smoothness involved.

Optimization models based on likelihood-based objective functionals and simulation-based techiques can be employed, particularly in the presence of stochastic volatility, but also backfitting and other iterative algorithms may be preferred.

4 Quadratic Variation and Volatility Recovery

The idea of using realised volatilities as estimators of integrated volatilites leads to an accurate volatility measure based on the now easily available high frequency observations. Therefore, as much information as we almost desire can be used, with positive impact on the precision of the estimates.

Given returns observed at frequencies, say, of 1min or 5min, the realised volatility comes in the following form:

$$\hat{\sigma}_n^R = \sum_{i=1}^{T} y_{in}^2 \tag{5}$$

suggesting that an approximation over a certain time interval is applied to the integral of an unobservable variable, the integrated volatility, by averaging a certain number of squared intraday values y_{in}^2. This leads to the following relationship:

$$\hat{\sigma}_n^R \to \sigma_t^I = \int_0^t \sigma^I(s) ds \tag{6}$$

Due to the quadratic variation principle, it generally holds that given a process $X(t)$ and the partition $T = \{t_0, t_1, \ldots, t_n\}$ of $[0, t]$, the p^{th} variation of $X(t)$ over T is:

$$V_t^{(p)}[T] = \sum_{k=1}^{n} \mid X_{t_k} - X_{t_{k-1}} \mid^p \tag{7}$$

Then, if $|| T ||= max_{1 \leq k \leq n} \mid t_k - t_{k-1} \mid \rightarrow 0$, for the volatility case, and thus looking at $p = 2$, we have:

$$lim_{||T|| \rightarrow 0} V_t^{(2)}[T] =< X >_t \tag{8}$$

where the limit is the quadratic variation of $X(t)$.

Equivalently, the convergence in probability of the intraday to the daily volatility holds for the cumulative squared intraday returns:

$$\sum_{i=1}^{n} y_i^2 \rightarrow_{n \rightarrow \infty} \int_0^1 \sigma^I(t + \delta) d\delta \tag{9}$$

Since a return process described as a semimartingale has an associated quadratic variation process measuring its sample-path variation, the approximation measure supplied by the realised volatility estimator can be evaluated locally, over certain discrete time horizons of defined length, say τ, and from return dynamics such as:

$$y_{t+\tau,\tau} = \kappa + \int_0^t \sigma(t + \delta) dW(\delta) \tag{10}$$

with $W(\delta)$ as the standard Brownian motion, κ the predictable component, and σ the volatility process underlying the observed returns.

5 Preliminary Experiments

Figure 1 and 2 report the experiments of simulating standard Brownian motion paths at different time-grids (or sampling rates); we show (top-located plots) trajectories of 100, 1000 and 10000 points of these paths.

We first recall the definition of standard Brownian motion: it is a stochastic process $\{W(t)_{t \geq 0}\}$ such that **(1)** $W(0) = 0$ and $W(t)$ is a continuous function of time; **(2)** for each $t, s > 0$ then $W(t + s) - W(t) \sim N(0, s)$ and **(3)** $W(t + s) - W(t)$ is independent of $W(t)$.

We then compute the quantity of greatest interest, i.e., the pointwise estimates of the variance for the Brownian motion path simulated in the time interval 0-1.

The pointwise (or running-time) variances are reported at the mid-located plots, while the corresponding cumulative variance appear at the bottom-located plots.

It may be observed, first of all, that achieving the variance bound, and thus a full convergence of the realised to the integrated volatility, is possible for the Brownian motion path sampled quite finely, i.e., with 10^4 sample points. For less sample points, those obtained by sampling at coarser grids, the bound is almost, but not completely, achievable.

6 Frequency Analysis

Consider a time discretized realization of a stochastic process, and look at the data $x = x_1, \ldots, x_n$ as values $1, \ldots, n$ of a function with period n, so to define each generic element as a linear combination of harmonics (see Brockwell and Davis, 1991):

$$x_t = n^{\frac{1}{2}} \sum_{-\pi < w_j \leq \pi} \alpha_j e^{itw_j} \tag{11}$$

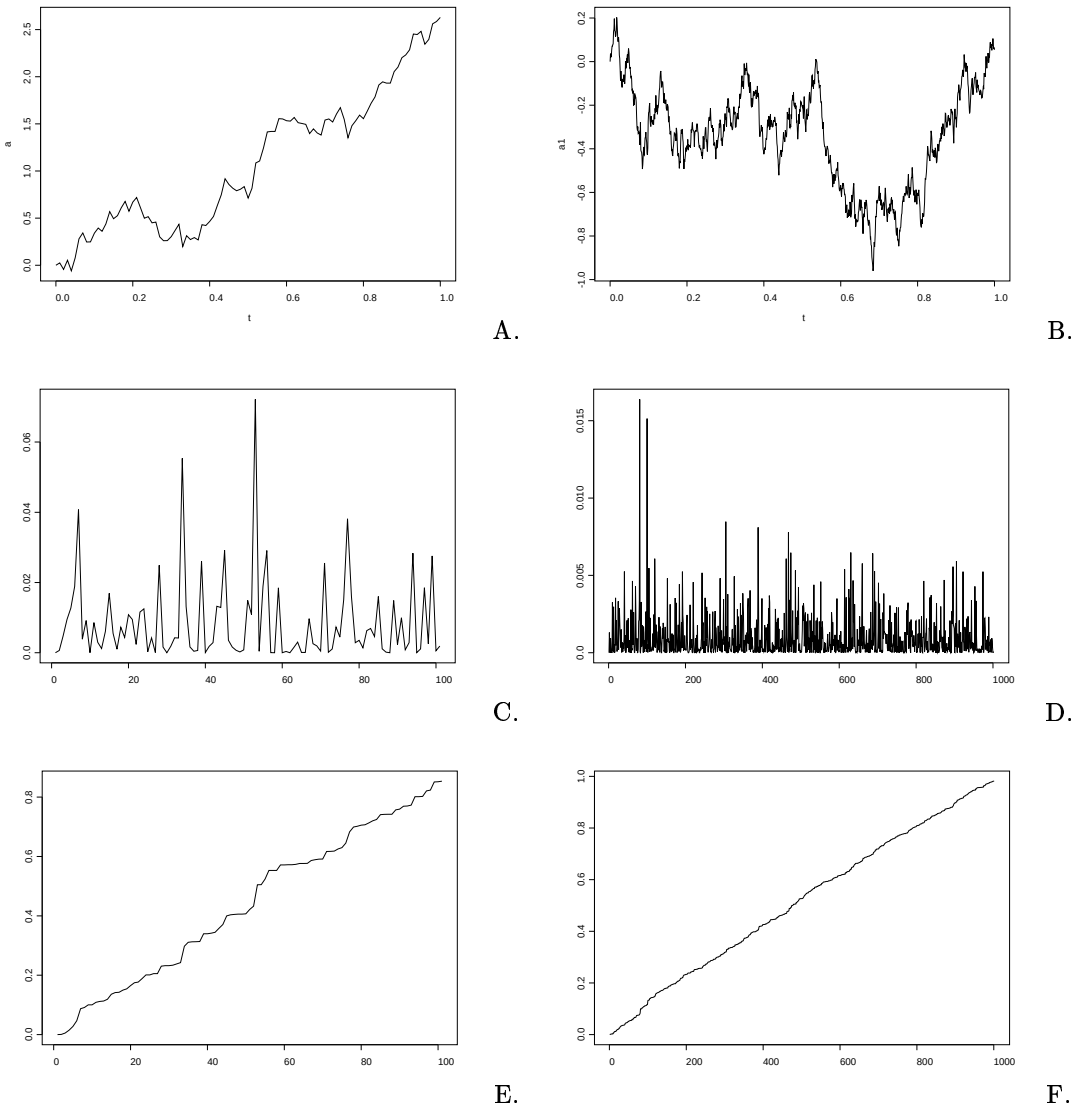
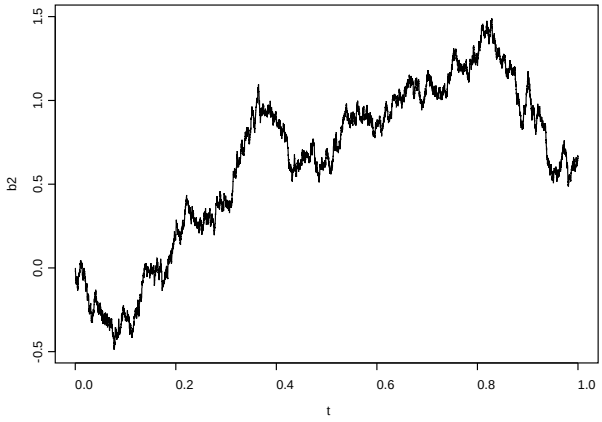
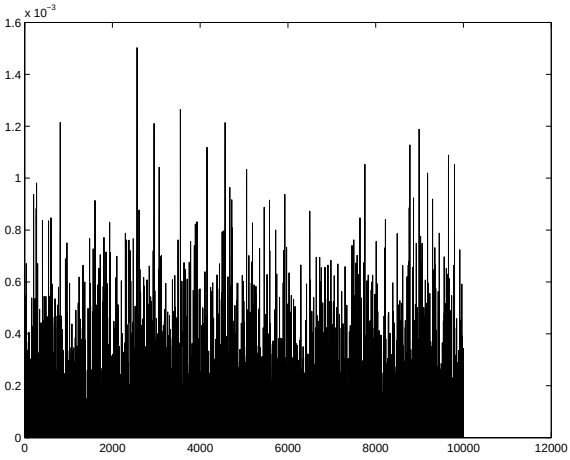


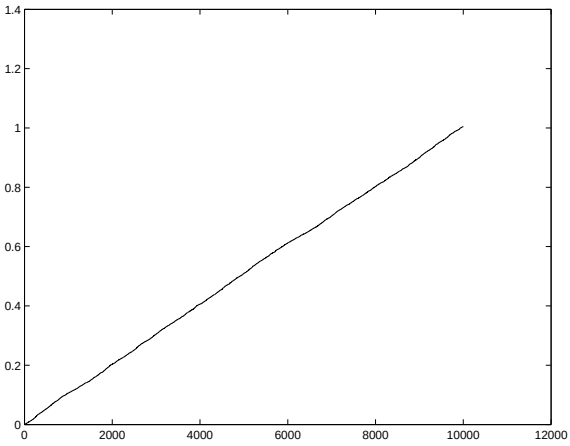
Figure 1: Brownian paths: 100 (A) and 1000 (B) sample points. Corresponding running-time variances in (C) and (D). Cumulative variances respectively at (E) and (F).



A.



B.



C.

Figure 2: Brownian path of 10000 sampling points (A). Corresponding running-time variance (B) and cumulative variance (C).

with the Fourier frequencies $w_j = \frac{2\pi j}{n}$ being the integer multiples of the fundamental frequency $\frac{2\pi}{n}$ in the interval $(-\pi, \pi]$.

Then, the value of the periodogram of x at the frequencies w_j is given by the discrete Fourier transform of x according to:

$$I(x, w_j) = |< x, e_j >|^2 = \frac{1}{n} \mid \sum_{i=1}^{n} x_t e^{-itw_j} \mid^2 \tag{12}$$

The periodogram, through the sum of squares component in the harmonic decomposition of the discretized process, yields an analysis of variance formula according to $\| x \|^2 = \sum_j I(x, w_j)$.

The covariance function of the process $X(t)$ is defined by $cov[X(t), X(t+\tau)] = C(\tau)$; with a fast decay to 0 and for $\sum_{\infty}^{\infty} \mid C(\tau) \mid < \infty$, its spectral representation is:

$$C(\tau) = \int_{-\pi}^{\pi} e^{iw\tau} d\bar{S}(w) = \int_{-\pi}^{\pi} e^{iw\tau} S(w) dw \tag{13}$$

where the spectrum of $X(t)$ is indicated by $S(w)$ and the corresponding spectral distribution function is $\bar{S}(w)$. The spectrum is obtained by the Fourier transform of $C(\tau)$ as:

$$S(w) = \kappa \sum_{\tau=-\infty}^{\infty} C(\tau) e^{-iw\tau} \tag{14}$$

for $\kappa = \frac{1}{2\pi}$, τ ranging over $(-\infty, \infty)$ and $-\pi \leq w \leq \pi$.

Note that $\bar{S}(-\pi) = 0$ and $\bar{S}(\pi) = \sigma^2$, the latter thus measuring the average total power; thus, given $d\bar{S}(w) = S(w)ds$, we know how the components in $X(t)$ with frequencies in $(w, w + dw)$ contribute to the power of the process.

7 Periodogram in Continuous Time

In Dzhaparidze and Spreij (1994) the periodogram is used to estimate the variance of a continuous time stochastic process where, given a finite stopping time $\bar{T}$ and a real number w, we have:

$$I_t(X, w) = \mid \int_0^{\bar{T}} e^{iwt} dX_t \mid^2 \tag{15}$$

By generalizing Levy's theorem, where the variance of a Brownian motion can be obtained as the limit of the sum of squares of the increments at increasingly finer partitions, other processes can exploit the same principles, and in particular the periodogram of a semimartingale can be useful to estimate the corresponding quadratic variation process.

The periodogram is an asymptotically unbiased estimate of the spectral density of the process, say $f(w)$, and for consistency reasons a smoothed version of the estimator is needed.

The spectral function $f(w)$ is related to the spectral decomposition of the process X into a superposition of harmonics, but such decomposition is not subject to a unique recovery from the discrete observations of a continuous process, i.e., a sampled process ΔX.

The spectrum of such observations may concentrate on the finite frequency band $-\frac{\pi}{\Delta} \leq w < \frac{\pi}{\Delta}$; this represents the main band, large enough that the power distribution belonging to frequencies higher the the Nyquist limits offer just a negligible contribution to the total power of the process.

Consider the spectrum S_X of $X(t)$, the latter having variance function σ^2_{Xt}. The spectrum decomposes the variance with respect to a continuous independent variable w, e.g., the frequency, such that:

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} S_X(w)dw = \sigma^2_{Xt} \tag{16}$$

The spacing Δ between observations is destined to go to 0 for consistency requirements. Accordingly, the number of discrete observations grows, so that the amount of information in the data increases too.

Thus, while looking at the periodogram, we know that it is asymptotically unbiased, i.e., $I_T(w) \approx f(w)$, and its variance $f^2(w)$ is not depending on the sample size. Therefore, again for consistency, we need to smooth the periodogram across frequencies.

8 Other Experiments

We see in Figure 3 and 4 the periodograms computed for the increments of sampled Brownian paths; the spectra are measured in decibels[1]. The reported plots have the frequency range $0 - .5$; thus, half not all of the total power is measured.

The periodograms are shown in tapered and non-tapered versions (Figure 3); this means that a data window is considered or not so to reduce the bias in the spectral estimates due to the periodogram (Priestley, 1981). Then, the smoothed versions of the periodogram (Figure 4) are also computed.

In Figures 5 and 6 the periodograms of the Brownian increments are plotted, varying according to the sampling rate. The x axis accounts for the frequency range.

Considering the sample frequency employed by the estimator, a transformed power spectrum is shown in Figures 7-8, so to refer back to the total power that we try to approximate, and to the plots of the cumulative periodograms.

9 Interpretation of Results and Conclusive Remarks

We aimed to develop methods of volatility estimation within the suitable framework of semimartingale processes of asset prices, and we have supported our results with experimental evidence which showed, in particular, the quality of the convergence to the quadratic variation limit for realised and integrated volatilities.

Our study has shown that in order to recover the unobserved volatility function both time and frequency domain estimators can be proposed.

Simulation experiments with standard Brownian motion paths sampled at different grids could reveal the speed and the accuracy of convergence of the estimators to the the asymptotic variance bounds.

It is found that a comparable performance in terms of achievable rates characterizes time- and frequency-based estimators. It is also clearly seen that when frequency estimators are employed, the recourse to tapers yields a decrease in the convergence rate.

A final remark concerns which estimator to choose. With noisy returns, the choice should give priority to the more robust method and to the fastest convergence rate which is achievable.

According to our simulation study, the latter aspect doesn't discriminate between estimators computed in time and frequency coordinates, but robustness could be an important issue to investigate.

Some experimental work is currently under study for exploring the role of de-noising techniques with time-frequency and time-scale estimators, following recent work by Capobianco (2002) with wavelets.

[1] One applies the transform $10 * log_{10}(x)$ to x, i.e., the estimated power spectrum.

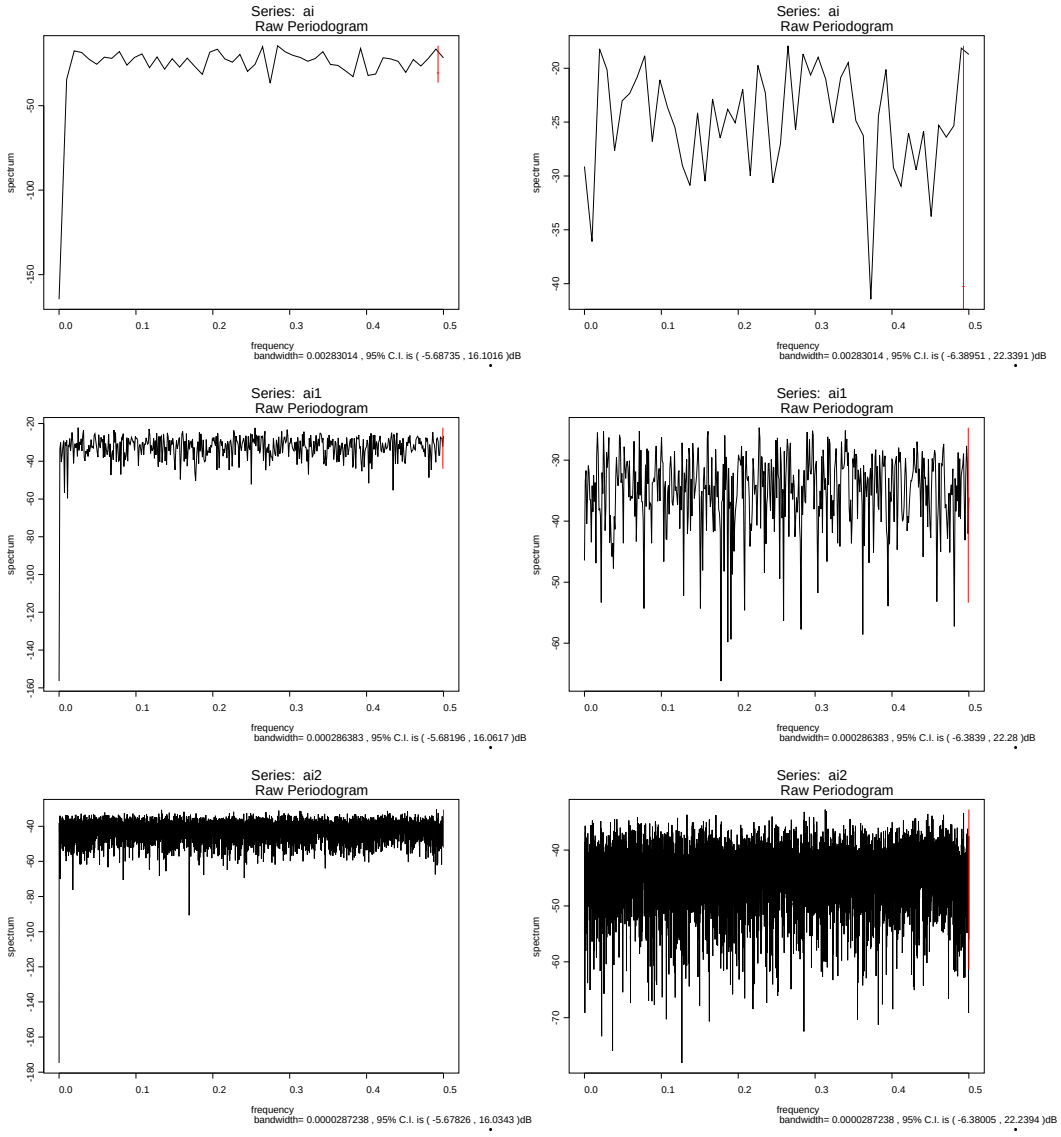


Figure 3: Raw non-tapered (left-hand column) and tapered periodograms for the increments of Brownian paths with 100 (top), 1000 (middle) and 10000 (bottom) sample points. Measures in decibels.

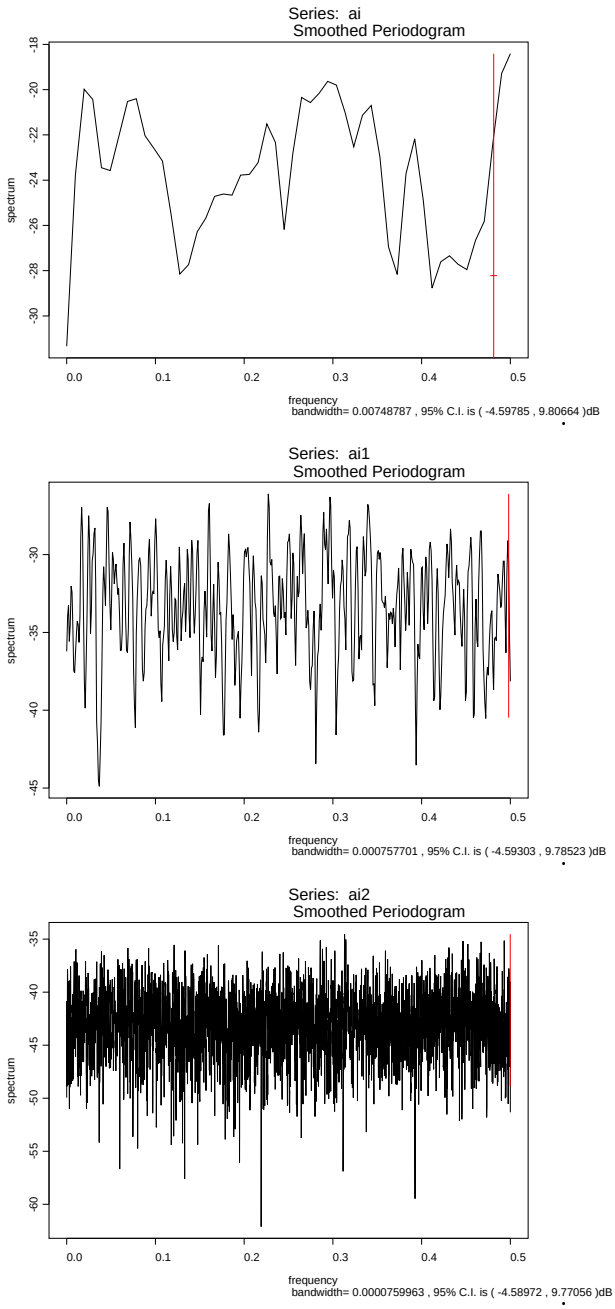


Figure 4: Smoothed periodograms for the increments of Brownian paths with 100 (top), 1000 and 10000 (bottom) sample points. Measures in decibels.

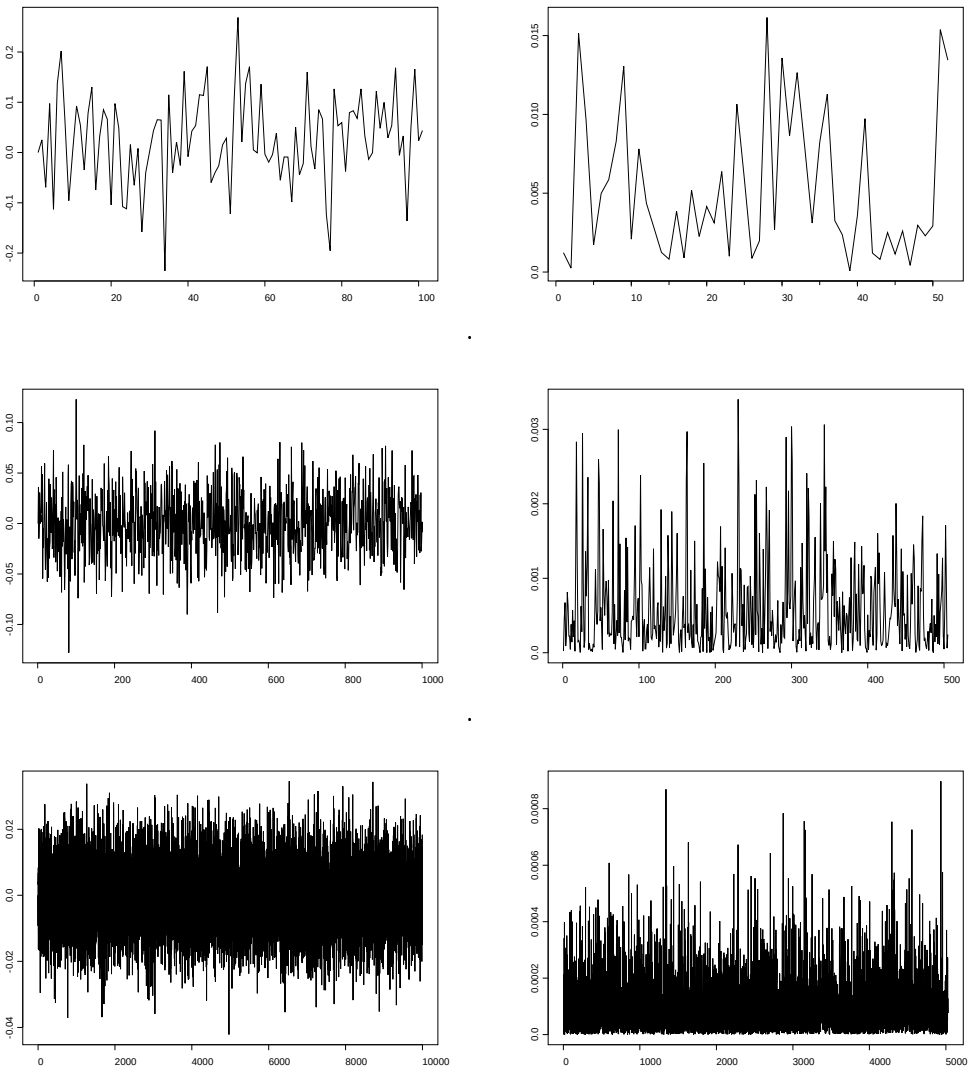


Figure 5: Increments series on the left-hand column: 100 (top), 1000 and 10000 (bottom) sample points, with corresponding raw periodogram transforms (right-hand column).

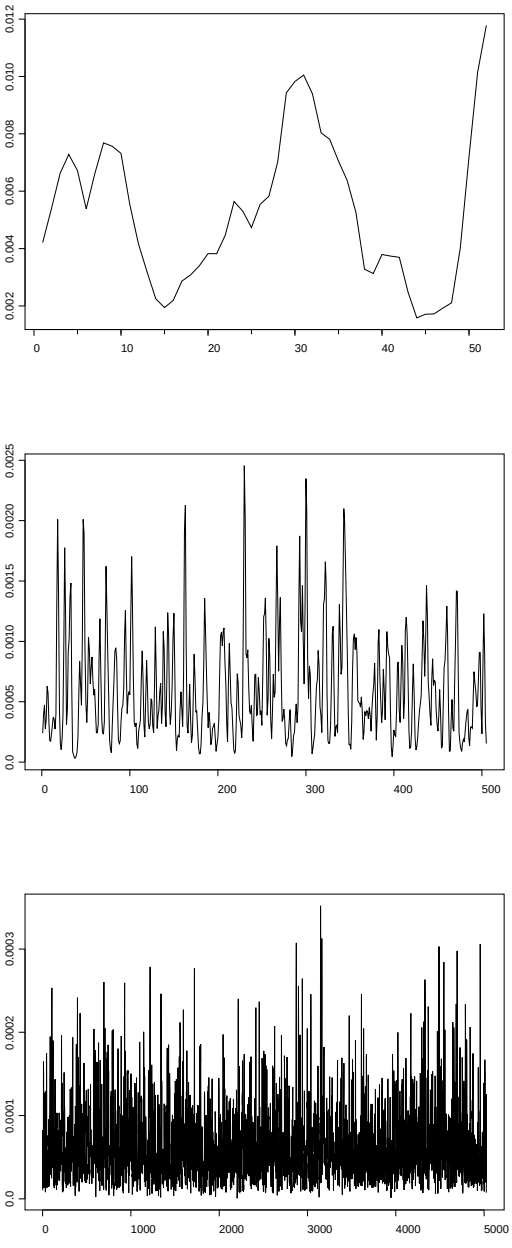


Figure 6: Smoothed periodogram transforms for 100 (top), 1000 and 10000 (bottom) simulated sample points.

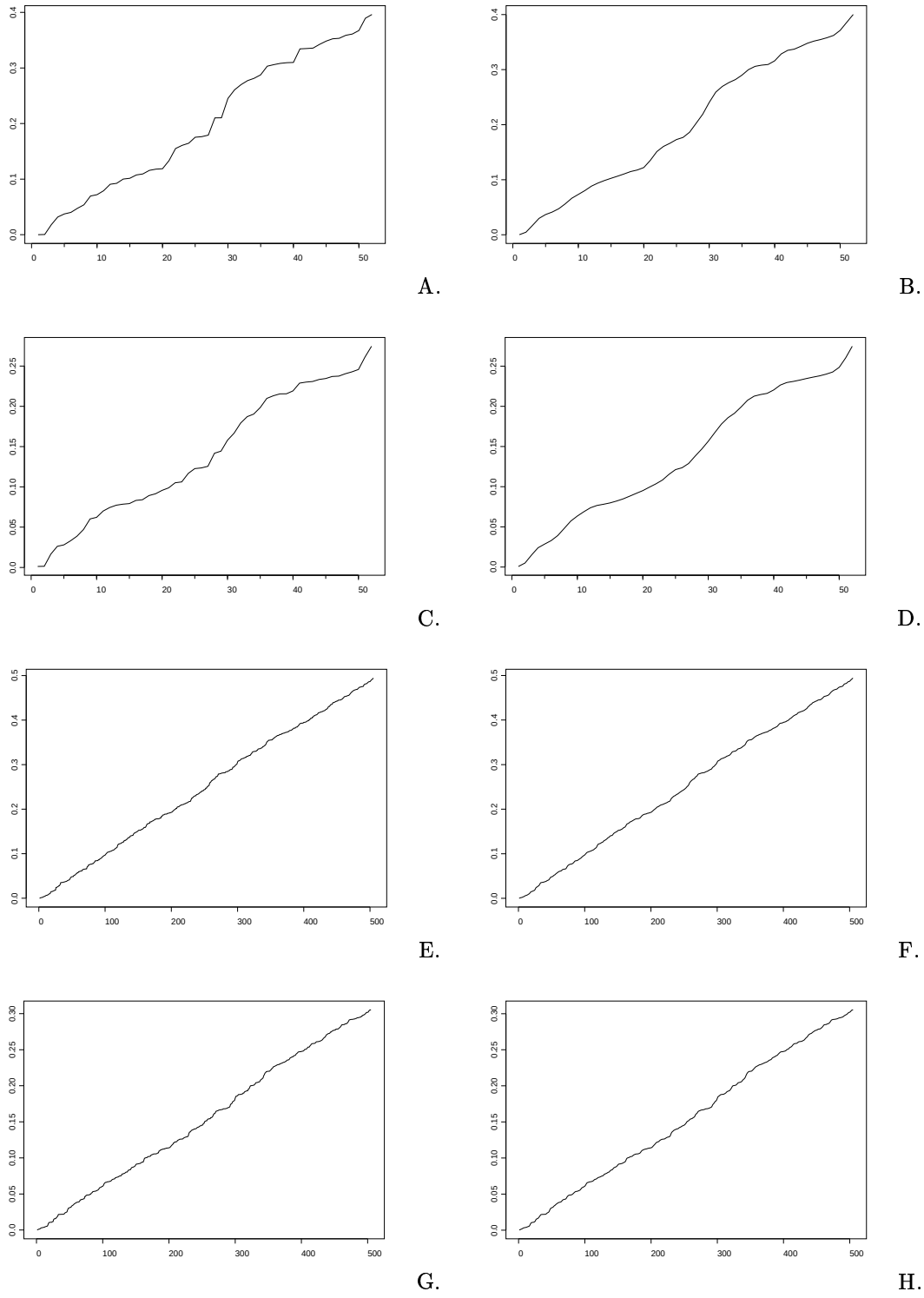


Figure 7: Cumulated power for raw (A and C) and smoothed (B and D) periodogram transforms, with 100 sample points. Non-tapered (A and B) and tapered estimates (C and D). Corresponding plots for 1000 sample points; cumulated power (E-G) and smoothed transforms (F-H). Non-tapered (E-F) and tapered estimates (G-H).

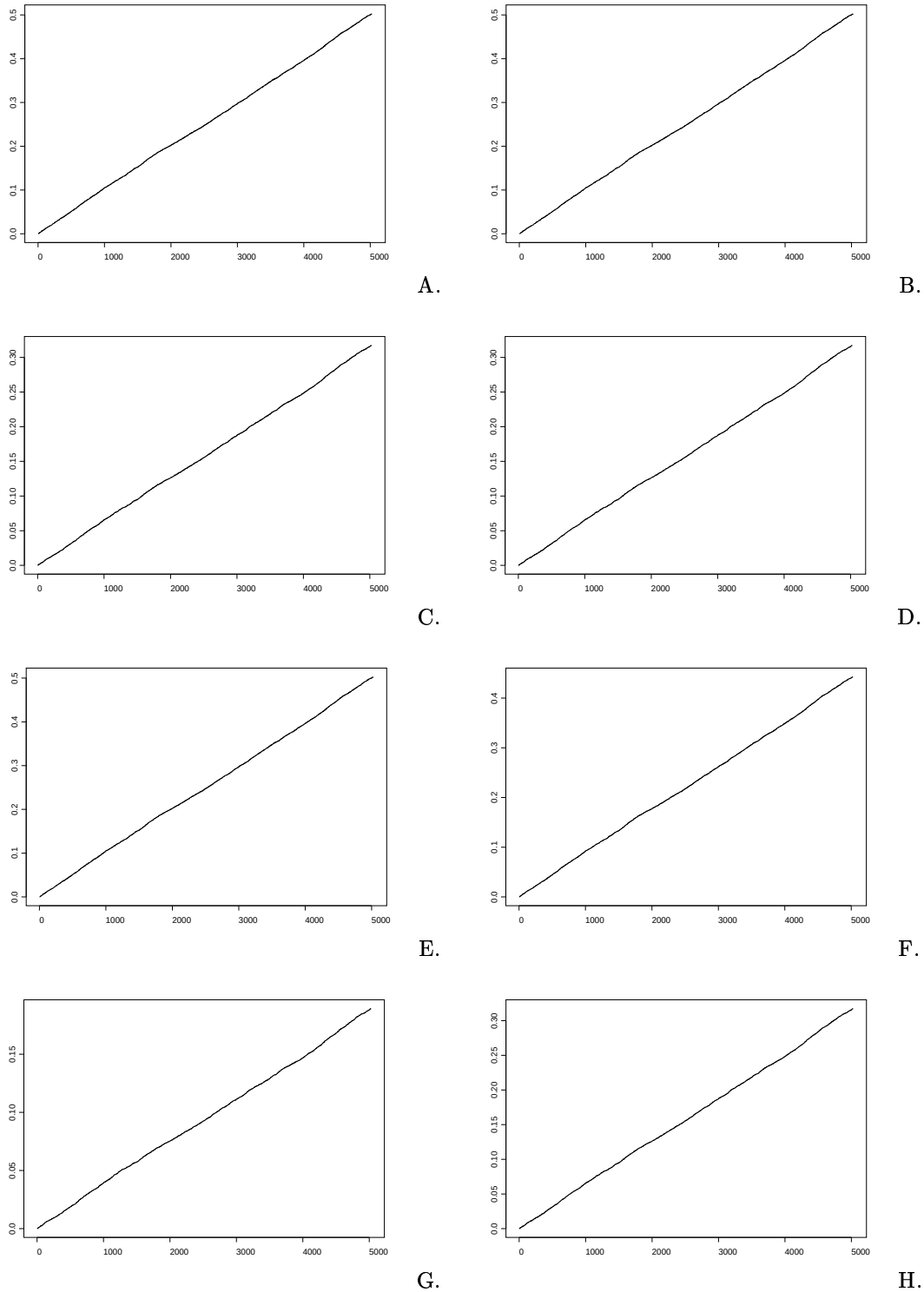


Figure 8: Cumulated power for raw (A and C) and smoothed (B and D) periodogram transforms, with 10000 sample points. non-tapered (A and B) and tapered estimates (C and D). Effect of tapering: no (E) - small (F) - big (G) tapering and combined with smoothing (H).

References

[1] T. Andersen, T. Bollerslev, F.X. Diebold, H. Ebens, The distribution of stock return volatility, Jour. Fin. Econ., 61 (2001a) 43-76.

[2] T. Andersen, T. Bollerslev, F.X. Diebold, P. Labys, The distribution of realized exchange rate volatility, Jour. Amer. Statis. Assoc., 96 (2001b) 42-55.

[3] O.E. Barndorff-Nielsen, N. Shephard N., Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics (with discussion), Jour. Roy. Statist. Soc. Ser. B, 63 (2001) 167-241.

[4] P.J. Brockwell, R.A. Davis, Time Series: Theory and Methods, Springer-Verlag, New York, 1991.

[5] E. Capobianco, Multi-resolution Approximation for Volatility Processes, Quantit. Fin., 2 (2002) 91-110.

[6] K. Dzhaparidze, P. Spreij, Spectral characterization of the optional quadratic variation process, Stoch. Process. Applic., 54 (1994) 165-174.

[7] I. Karatzas, E. Shreve, Brownian Motion and Stochastic Calculus, Springer-Verlag, New York, 1998.

[8] P. Malliavin, M.E. Mancino, Fourier series method for measurement of multivariate volatilities, Fin. Stoch., 6 (2002) 49-61.

[9] M.B. Priestley, Spectral Analysis and Time Series, Academic Press Inc., London, 1981.

[10] A.N. Shiryaev, Essentials of Stochastic Finance: Facts, Models and Theory, World Scientific, Singapore, 1999.