Supplementary Information for:

Novel Rapid Approach for Adaptive Gaussian Kernel Density Estimation: Gridpoint-wise Propagation of Anisotropic Diffusion Equation

Christian Sustay Martinez[1], Patrick K. Quoika[1], and Martin Zacharias[1]

[1]Center for Functional Protein Assemblies, Technical University of Munich, Ernst-Otto-Fischer-Str. 8, Garching 85748, Germany

Acronyms

VMR variance-to-mean ratio

A Propagation of Mean and Variance of Kernel Evaluation

Let $\hat{f}(x;T)$ denote the kernel density estimate at bandwidth $T \in \mathbb{R}^{n\times n}$ of the true density function $f(x)$. We refer to the initial condition—corresponding to zero bandwidth or "time zero" in the diffusion interpretation—as

$$\hat{f}(x;0) = \frac{1}{N}\sum_{i=1}^{N}\delta(x - X_i), \tag{A.1}$$

a sum of Dirac delta functions centered at the sample points $\{X_i\}$, representing the empirical distribution.

At a later artificial time (i.e., after smoothing with a nonzero bandwidth matrix T), the contribution of a single sample to the density estimate can be expressed as

$$\bar{h}(x;T) = \frac{1}{N}K(x;T) * \hat{f}(x;0), \tag{A.2}$$

where $*$ denotes convolution in the spatial domain and $K(x;T)$ is the kernel function

$$K(x;T) = \frac{1}{\sqrt{(2\pi)^n|T|}}\exp\left(-\frac{1}{2}x^\top T^{-1}x\right). \tag{A.3}$$

In the Fourier domain, convolution becomes a multiplication; therefore, assuming T is diagonal, the Fourier transform of $\bar{h}(x;T)$ becomes

$$\tilde{h}(k;T) = \frac{1}{N}\widetilde{G}_1(k;T)\tilde{f}(k;0), \tag{A.4}$$

where $\widetilde{G}_1$ is the propagator in the Fourier domain, given by

$$\widetilde{G}_1(k;T) = \exp\left(-\frac{1}{2}\sum_{j=1}^{n}k_j^2t_j^2\right). \tag{A.5}$$

Suppose we want to evaluate the effect of total smoothing corresponding to a bandwidth matrix $T_f = T_0 + T_1$. Then, using the semigroup property of the Gaussian propagator, we can write

$$\tilde{h}(k;T_f) = \widetilde{G}_1(k;T_1)\tilde{h}(k;T_0). \tag{A.6}$$

Additionally, since

$$\bar{h}(x;T) = \frac{1}{N^2}\sum_{i=1}^{N}K(x - X_i;T) \tag{A.7}$$

integrates to $\frac{1}{N}$, we can interpret $\bar{h}(x;T)$ as the empirical mean of kernel evaluations centered at $X \sim f$.

Similarly, to compute the empirical variance of the kernel evaluations, we first calculate the empirical second moment. The variance at a point x is given by

$$s^2(x;T) = h_2(x;T) - \bar{h}^2(x;T), \tag{A.8}$$

where $h_2(x;T)$ is the empirical second moment, defined as

$$h_2(x;T) = \frac{1}{N^3} \sum_{i=1}^{N} K^2(x - X_i; T). \tag{A.9}$$

In Fourier space, using the convolution structure of the kernels and the linearity of the Fourier transform, we have

$$\tilde{h}_2(k;T) = \frac{1}{N^3 2\sqrt{\pi \det T}} \exp\left(-\frac{1}{4} \sum_{j=1}^{n} k_j^2 t_j^2 \right) \tilde{f}(k;0), \tag{A.10}$$

which is again a Gaussian (with halved variance).

Then, to evaluate the second moment corresponding to a total bandwidth matrix $T_f = T_0 + T_1$, we can express the propagated second moment in Fourier space as

$$\tilde{h}_2(k;T_f) = \widetilde{G}_2(k;T_1,T_0)\tilde{h}_2(k;T_0), \tag{A.11}$$

where the second-mement propagator $\widetilde{G}_2(k;T,S)$ is given by

$$\widetilde{G}_2(k;T,S) = \left(\frac{\det S}{\det(S+T)} \right)^{1/2} \exp\left(-\frac{1}{4} \sum_{j=1}^{n} k_j t_j \right), \tag{A.12}$$

with T and S being diagonal bandwidth matrices. S is the initial artificial time, whereas T is the artificial time increment.

B VMR Propagation Regimes

The empirical variance-to-mean ratio (VMR) is defined in Equation (2.15). To analyze its variability, we begin by computing its partial derivatives with respect to the first and second empirical moments, as given in

$$\frac{\partial \text{vmr}}{\partial \bar{h}} = -\frac{h_2 + \bar{h}}{\bar{h}^2}, \quad \frac{\partial \text{vmr}}{\partial h_2} = \frac{1}{\bar{h}}. \tag{B.1}$$

We approximate the variance of the empirical VMR using the delta method, which gives

$$\text{Var}\left[\text{vmr}\right] \approx \nabla \text{vmr}^\top \, \text{Cov}\left(\bar{h}, h_2\right) \nabla \text{vmr}, \tag{B.2}$$

where the gradient is evaluated at the true moment values.

The entries of the covariance matrix can be expressed as

$$\begin{aligned}
\text{Var}\left[\bar{h}\right] &= \frac{1}{N}\text{Var}[h], \\
\text{Var}[h_2] &= \frac{1}{N}\text{Var}[h^2], \\
\text{Cov}\left(\bar{h}, h_2\right) &= \frac{1}{N}\text{Cov}[h, h^2],
\end{aligned}$$

where these are written in terms of the true variances and covariance of the single-kernel evaluation (Equation (2.8)).

Substituting these expressions into the delta-method formula yields

$$\mathrm{Var}[\mathrm{vmr}] \approx \frac{1}{N}\left[\left(\frac{\mathbb{E}[h^2]+\mathbb{E}[h]^2}{\mathbb{E}[h]^2}\right)^2 \mathrm{Var}[h] + \frac{1}{\mathbb{E}[h]^2}\mathrm{Var}[h^2] - 2\frac{\mathbb{E}[h^2]+\mathbb{E}[h]^2}{\mathbb{E}[h]^3}\mathrm{Cov}(h,h^2)\right], \tag{B.3}$$

an expression dependent on the true first and second moments, as well as their associated higher-order statistics.

To proceed further, we introduce expressions for the true variances and covariance in terms of the underlying moments of the kernel:

$$\begin{aligned}\mathrm{Var}[h] &= \mathbb{E}[h^2] - \mathbb{E}[h]^2,\\ \mathrm{Var}[h^2] &= \mathbb{E}[h^4] - \mathbb{E}[h^2]^2,\\ \mathrm{Cov}(h,h^2) &= \mathbb{E}[h^3] - \mathbb{E}[h]\mathbb{E}[h^2].\end{aligned}$$

We recall that the m-th moment of the kernel is defined by

$$\mathbb{E}\left[h^m(x;T)\right] = \frac{1}{N^m}\int_\Omega K^m(x-X;T)f(X)dX, \tag{B.4}$$

and, using the local expansion introduced in Equation (2.23), we can write the leading-order approximation

$$\mathbb{E}\left[h^m(x;T)\right] \approx \frac{1}{N^m}A_m f(x),$$

where A_m denotes the integral of the m-th power of the Gaussian kernel:

$$A_m = \int_\Omega K^m(x;T)dX = (2\pi)^{\frac{n(1-m)}{2}} m^{-\frac{n}{2}} (\det T)^{\frac{1-m}{2}}. \tag{B.5}$$

Notably, for a Gaussian kernel, $A_m \propto |T|^{\frac{1-m}{2}}$. Substituting the approximated moments into the VMR expression yields, to leading orders, and using the asymptotic scaling of A_m, we obtain, for small enough bandwidths, the leading order dependencies of the VMR variance on the bandwidth matrix T:

$$\mathrm{Var}[\mathrm{vmr}](x) \propto \frac{|T|^{-3/2}}{N^3 f(x)}. \tag{B.6}$$

For large bandwidths, rather than expanding the moments in Equation (B.4) around the local density $f(x)$, we instead expand them in powers of $|T|^{-1}$. Using the quadratic form $q = (x-X)^\top T^{-1}(x-X)$ and the exponential expansion, this yields

$$\mathbb{E}[h^m(x;T)] = (2\pi)^{-mn/2}|T|^{-m/2}\sum_{s=0}^{\infty}\frac{(-m)^2}{2^s s!}\mathbb{E}[q^s], \tag{B.7}$$

expressed in terms of the moments of the quadratic form q.

Substituting these expressions into Equation (B.3) and retaining the leading-order contributions, we obtain

$$\mathrm{Var}[\mathrm{vmr}](x) \propto \frac{|T|^{-(1+4/n)}}{N^3}, \tag{B.8}$$

which characterizes the asymptotic behavior of the variance in the large-bandwidth regime. In this limit, the dependence on the dimensionality n arises from the scaling of $||T^{-1}||$ with the number of dimensions, reflecting how smoothing distributes mass across higher-dimensional spaces.

This analysis reveals how the variance of the empirical VMR evolves with bandwidth T, allowing us to distinguish between different time regimes. In the small-T regime, the VMR variance becomes sensitive to local variations in the density, particularly in low-density regions where both the mean and second moment decay. In contrast, in the large-T regime, the kernel becomes broader and the effect of local density variations diminishes. The variance of the empirical VMR is then dominated by the kernel's global smoothing effect—in this case a Gaussian curve—, scaling primarily with a different power-law whose behavior becomes independent of the underlying density (see Figures 3 and S1).

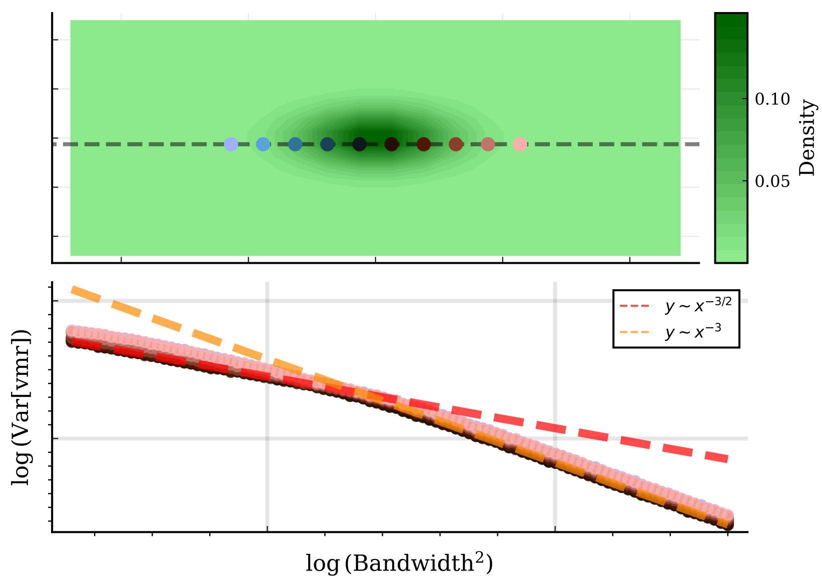

Fig. S1: Empirical demonstration of the asymptotic scaling laws for the variance of the empirical VMR in two dimensions.