Gaussian STT Tutorial

Ryan T. Grimm, Joel D. Eaves

(Dated: May 20, 2025)

I. INTRODUCTION

In this tutorial, you will learn how to decompose a two-dimensional normal probability density function (PDF) into a spectral tensor train. The PDF is

$$\mathcal{P}(\boldsymbol{x}) = \frac{1}{2\pi\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2} x^\intercal \boldsymbol{\Sigma}^{-1} x\right), \tag{1}$$

where $\boldsymbol{x} = (x, y)$. $\mathcal{P}$ represents the probability for a pair of normally distributed random variables X and Y. A spectral tensor train (STT) is a decomposition for scalar-valued, multivariate functions built from a product of univariate, matrix-valued functions. The STT decomposition for the probability density is $\mathcal{P}(x, y) = \boldsymbol{P}_1(x)\boldsymbol{P}_2(y)$. Because the result must be scalar, the first and last function in the STT, called cores, must have unit left and right dimensions, respectively. Thus, $P_1 \in \mathbb{R}^{1\times\chi}$, and $P_2 \in \mathbb{R}^{\chi\times 1}$, where χ is the bond-dimension. In the language of quantum mechanics, χ encodes the degree of entanglement between x and y. P_1 and P_2 are unknown functions, so they must be parameterized. There are many universal function approximators: polynomials, neural networks, Fourier series, etc. For reasons that will become clear in this tutorial, orthogonal polynomials with matrix-valued coefficients are a good choice. Thus, let $P_n(x) = \sum_k \boldsymbol{A}_n(k)\phi_k(x)$, where $\{\boldsymbol{A}_n(k)\}$ are a set of matrix-valued coefficients, and ϕ_k are a set of orthogonal polynomials.

II. THEORY

To see where the STT decomposition comes from, expand $\mathcal{P}$ in tensor-product basis of orthogonal polynomials

$$\mathcal{P}(x, y) = \sum_{i,j} \mathcal{C}^{i,j}\phi_i(x)\phi_j(y), \tag{2}$$

where $\mathcal{C}$ is the coefficient tensor. Using orthogonality, project out the coefficients

$$\mathcal{C}^{i,j} = \int dw(x)dw(y)\, \frac{\phi_i(x)\phi_j(y)\mathcal{P}(x, y)}{N_i N_j}, \tag{3}$$

where w is the orthogonality weight, or measure, of the polynomials, and N_k is the normalization constant. In quantum mechanics, one needs to take the conjugate of ϕ_i and ϕ_j, but this example takes place in a purely real-valued function space, so $\forall k, \phi_k^* = \phi_k$. To obtain the coefficients computationally, use numerical quadrature to evaluate the integral

$$\mathcal{C}^{i,j} = \sum_{k,l} \frac{W_k W_l \phi_i(z_k)\phi_j(z_l)\mathcal{P}(z_k, z_l)}{N_i N_j}, \tag{4}$$

where $\{W_k\}$ are the quadrature weights, and $\{z_k\}$ are the quadrature points.

The aim of the STT decomposition is to convert $\mathcal{P}$ into a separable function. To do this, use dyadic matrix decomposition, where $\boldsymbol{M} = \boldsymbol{L}\boldsymbol{R}$, such as singular value decomposition, QR, LQ, or Cholesky. SVD is convenient as it allows one to set the accuracy of the decomposition by truncating singular values below some threshold ϵ. By the Eckheart-Young theorem, the distance $||\boldsymbol{M} - \tilde{\boldsymbol{M}}||_2 \leq \epsilon$, where $||\cdot||$ denotes the two-norm, and $\tilde{\boldsymbol{M}}$ is an approximation of $\boldsymbol{M}$ contains no singulars greater than ϵ. To split $\mathcal{P}$, define a matrix $\mathcal{Q}(k,l) = \mathcal{P}(z_k, z_l)$, and decompose it with SVD $\boldsymbol{Q} = \boldsymbol{L}\boldsymbol{R}$, where $\boldsymbol{L} = \boldsymbol{U}\sqrt{\boldsymbol{S}}$ and $\boldsymbol{R} = \sqrt{\boldsymbol{S}}\boldsymbol{V}^\intercal$. Insert this into the expression for $\mathcal{C}$ and rearrange the result to get

$$\mathcal{C}^{i,j} = \sum_n \left[\sum_k \frac{W_k \phi_i(z_k) L^{k,n}}{N_k}\right] \left[\sum_l \frac{W_l \phi_j(z_l) R^{n,l}}{N_l}\right]. \tag{5}$$

Recognizing that this is a dot product between

$$P_1^n(x) = \sum_k \frac{W_k \phi_i(z_k) L^{k,n}}{N_k}, \tag{6}$$

and

$$P_2^n(x) = \sum_l \frac{W_l \phi_j(z_l) R^{n,l}}{N_l}, \tag{7}$$

yields the cores for the STT decomposition of $\mathcal{P}$.