# Best Linear Unbiased Estimation for 2D and 3D Flow with Event-based Cameras

Supplementary Material

## A. BLUE Theorem

**Definition.** *Let $y_1, \dots, y_n$ be $n$ i.i.d. data points drawn from a distribution that depends on a parameter $\theta$. A point estimator $\theta_n$ of $\theta$ is a function of the data, i.e., $\theta_n = g(y_1, \dots, y_n)$. We say that the estimator $\theta_n$ is* unbiased *if $\mathbb{E}[\theta_n] = \theta$, and* consistent *if $\theta_n \to \theta$ in probability as $n \to \infty$.*

**Theorem.** *A linear estimator for the event function $y(x, t; \theta) = \varphi(x) + t\theta$ is unbiased if the coefficients $\alpha_i$ satisfy*

$$\sum_{i=1}^{n} \alpha_i = 0, \quad \sum_{i=1}^{n} \alpha_i t_i = 1.$$

*These conditions are necessary if the matrix $\begin{bmatrix}\mu & \theta\end{bmatrix} \in \mathbb{R}^{d\times 2}$ has full rank, where $\mu = E[\varphi(x)]$. Furthermore, the best linear unbiased estimator is given by*

$$\alpha_i = \frac{t_i - \bar{t}}{\sum_{j=1}^{n}(t_j - \bar{t})^2}, \quad \textit{where} \quad \bar{t} = \frac{1}{n}\sum_{j=1}^{n} t_j.$$

*It is consistent with convergence rate $O\left(\sqrt{\frac{d}{n}}\right)$ if there exists a constant $c > 0$ such that*

$$\frac{1}{n}\sum_{i=1}^{n}(t_i - \bar{t})^2 \geq c$$

*for sufficiently large $n$.*

*Proof.* For $n$ events $y_i$, the expected value of a linear estimator for the $j$-th component is

$$E\left[\sum_{i=1}^{n} \alpha_i y_i^{(j)}\right] = \sum_{i=1}^{n} \alpha_i E[y_i^{(j)}] = \sum_{i=1}^{n} \alpha_i E[\varphi^{(j)}(x_i)] + \sum_{i=1}^{n} \alpha_i t_i \theta^{(j)} + \sum_{i=1}^{n} \alpha_i E[\varepsilon_i^{(j)}].$$

Rearranging,

$$E\left[\sum_{i=1}^{n} \alpha_i y_i^{(j)}\right] = \mu^{(j)} \sum_{i=1}^{n} \alpha_i + \theta^{(j)} \sum_{i=1}^{n} \alpha_i t_i.$$

For unbiasedness, we require

$$\mu^{(j)} \sum_{i=1}^{n} \alpha_i + \theta^{(j)} \sum_{i=1}^{n} \alpha_i t_i = \theta^{(j)}.$$

Setting $a = \sum_{i=1}^{n} \alpha_i$ and $b = \sum_{i=1}^{n} \alpha_i t_i - 1$, it reduces to solve

$$\begin{bmatrix}\mu & \theta\end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \mathbf{0} \in \mathbb{R}^d.$$

This system has $a = 0, b = 0$ as solution and is unique if $\begin{bmatrix}\mu & \theta\end{bmatrix}$ is full rank, yielding

$$\sum_{i=1}^{n} \alpha_i = 0, \quad \sum_{i=1}^{n} \alpha_i t_i = 1.$$

On the other hand, using the independence between the variables, the variance for the $j$-th component is

$$Var\left(\sum_{i=1}^{n} \alpha_i y_i^{(j)}\right) = \sum_{i=1}^{n} \alpha_i^2 Var(y_i^{(j)}) = \sum_{i=1}^{n} \alpha_i^2 \Big(Var(\varphi^{(j)}(x_i)) + Var(\varepsilon_i^{(j)})\Big).$$

Assigning $c_j = Var(\varphi^{(j)}(x_i)) + Var(\varepsilon_i^{(j)})$, we obtain

$$Var\left(\sum_{i=1}^{n} \alpha_i y_i^{(j)}\right) = c_j \sum_{i=1}^{n} \alpha_i^2.$$

To minimize the variance under the unbiasedness constraints, we use Lagrange multipliers

$$L(\alpha_1, \ldots, \alpha_n, \lambda_0, \lambda_1) = \sum_{j=1}^{n} \alpha_j^2 + \lambda_0 \sum_{j=1}^{n} \alpha_j + \lambda_1 \left(\sum_{j=1}^{n} \alpha_j t_j - 1\right).$$

Setting derivatives to zero

$$\frac{\partial L}{\partial \alpha_i} = 2\alpha_i + \lambda_0 + \lambda_1 t_i = 0,$$

summing over all $i$, and using the constraints, we derive

$$\lambda_0 = -\lambda_1 \bar{t}, \quad \lambda_1 = -\frac{2}{\sum_{j=1}^{n}(t_j - \bar{t})^2}.$$

Substituting these into the derivatives gives

$$\alpha_i = \frac{t_i - \bar{t}}{\sum_{j=1}^{n}(t_j - \bar{t})^2}.$$

To analyze consistency, consider

$$\sum_{i=1}^{n} \alpha_i^2 = \frac{\sum_{i=1}^{n}(t_i - \bar{t})^2}{\left(\sum_{j=1}^{n}(t_j - \bar{t})^2\right)^2} = \frac{1}{\sum_{j=1}^{n}(t_j - \bar{t})^2}.$$

Applying the given condition

$$\frac{1}{n}\sum_{i=1}^{n}(t_i - \bar{t})^2 \geq c,$$

we obtain

$$\sum_{i=1}^{n} \alpha_i^2 \leq \frac{c}{n}.$$

By denoting $c' = \max_j c_j$ and using Jensen's inequality we obtain

$$E\left[\left\|\sum_{i=1}^{n} \alpha_i y_i - \theta\right\|\right] \leq \sqrt{\sum_{j=1}^{d} Var\left(\sum_{i=1}^{n} \alpha_i y_i^{(j)}\right)} = \sqrt{\sum_{i=1}^{n} \alpha_i^2 \sum_{j=1}^{d} c_j} \leq \sqrt{\frac{cc'd}{n}}.$$

Thus, the convergence rate is $O\left(\sqrt{\frac{d}{n}}\right)$. $\square$

The consistency condition in the theorem ensures that the variance of the timestamps is not close to zero, preventing events from being overly concentrated around a specific time. A high concentration at a single point in time results in limited motion information.