

Hermite Based Quasi-Likelihood Inference for a Stochastic Thermionic Emission Model with PDE Driven Inputs

R. Pogorielov , D. Ivanenko

Faculty of Radiophysics, Electronics and Computer Systems, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Received 24 September 2026

Abstract. We consider parameter estimation for a stochastic thermionic emission model generated by a diffusion process with multiplicative noise and temperature- and electric field-dependent drift. Since the transition density is not available in closed form, standard likelihood based inference is infeasible. We construct a quasi-likelihood using a Hermite type parametrix expansion, yielding an explicit high order approximation without solving the Kolmogorov equation. This leads to a quasi-maximum likelihood estimator for discretely observed emission currents. Least squares, least-absolute-error, and one step estimators are included for comparison. Monte Carlo experiments under physically structured PDE-generated inputs show that fully iterated estimators, especially Hermite based QML and selected second-order contrast schemes, provide more stable inference than one-step corrections, which are sensitive to the input complexity.

KEY WORDS: Hermite expansion, parameter estimation, PDE-driven inputs, quasi-likelihood, stochastic differential equations, thermionic emission.

1 Introduction

We consider statistical inference for a discretely observed stochastic process that is a solution to a stochastic differential equation (SDE). Let $X(t)$ be a d -dimensional diffusion satisfying

$$dX(t) = a(\theta, X(t)) dt + \sigma(\theta, X(t)) dW(t), \quad X(0) = x, \quad (1)$$

where W is an m -dimensional Brownian motion, a and σ are known coefficient functions, and θ is a vector of unknown parameters. Observations are available at discrete times $t_k = kh$, $k = 0, 1, \dots, n$.

The present work is motivated by thermionic emission experiments, in which the primary stochastic state variable is the emission current $I(t)$ generated by a heated cathode under an applied electric field. According to the classical Richardson–Dushman theory, the stationary emission current density is approximated by

$$J(T, E) \approx A_{\text{eff}} A_{\text{emit}} T^2 \exp\left(-\frac{\phi - \beta_S \sqrt{E}}{k_B T}\right), \quad (2)$$

where T denotes temperature, E the electric field, ϕ the effective work function, A_{eff} the Richardson constant, and A_{emit} the effective emitting area. Thus, $A_{\text{eff}} A_{\text{emit}}$ acts as an overall scale parameter for the equilibrium emission level, while β_S accounts for Schottky-type barrier lowering [1, 2]. This law and its refinements form the basis of thermionic cathode characterization and thermionic energy conversion theory [1, 2, 4–6]. Throughout the paper we use the generic notation $X(t)$ for the diffusion state; in the thermionic setting we identify $X(t) \equiv I(t)$.

In practice, emission parameters are typically extracted from static current temperature characteristics using linearized Richardson plots or related least squares fits of the deterministic law. Such procedures, however, neglect temporal dependence and stochastic fluctuations in the emission current, or treat them heuristically as measurement noise. In many devices, thermal fluctuations, surface processes, and external perturbations induce intrinsic stochastic variability in $I(t)$, making a stochastic dynamical description both natural and necessary for statistically efficient parameter estimation.

In this work we model the emission current $I(t)$ as a component of the diffusion process (1), with temperature $T(t)$ and electric field $E(t)$ entering as time dependent control inputs. The drift term encodes a Richardson–Dushman-type emission law combined with a relaxation mechanism toward a temperature and field dependent equilibrium, while the diffusion coefficient describes multiplicative stochastic fluctuations. The parameter vector θ thus contains both physical emission parameters and dynamical parameters governing the temporal variability of the current.

Statistical inference for discretely observed diffusions is challenging because the transition density is rarely available in closed form. Existing approaches include numerical solution of the Kolmogorov equation, Bayesian and particle methods, simulation based likelihoods, and contrast function estimators [3, 7–25]. In this paper, we avoid direct numerical solution of the Kolmogorov PDE by employing a Hermite type parametrix expansion of the transition density [26–28]. This yields an explicit high order approximation and enables the construction of a quasi-likelihood for thermionic current observations $I(t_k)$ with concomitant operating conditions (T_k, E_k) .

Based on this approximation, we develop a Hermite based quasi-maximum like-

likelihood estimator (QMLE) for the thermionic emission model. In addition, we implement least squares (LSE), least-absolute-error (LAE), and one step update estimators within a unified framework and compare their performance under physically structured PDE generated temperature and electric field inputs. The resulting methodology is likelihood based, accounts explicitly for temporal dependence and multiplicative noise, and provides a flexible and computationally efficient tool for estimating physical emission parameters from discretely sampled thermionic current data.

2 Model Formulation

2.1 Stochastic model for the emission current

In the thermionic–emission application considered in this paper, the primary stochastic state variable is the emission current $I(t)$ generated by a heated cathode under an applied electric field. The temperature $T(t)$ and electric field $E(t)$ are treated as externally generated inputs. In the numerical study, these inputs are produced by a heat-equation model and a Poisson field model driven by Ornstein–Uhlenbeck boundary and forcing processes. The current is modeled as a one–dimensional diffusion process.

Motivated by the Richardson–Dushman law with Schottky correction, we define the equilibrium, or target, emission current by

$$\mu(T, E; \phi, \beta) = A_{\text{eff}} A_{\text{emit}} T^2 \exp \left(-\frac{\phi - \beta \sqrt{E}}{k_B T} \right). \quad (3)$$

Thermal fluctuations and microscopic surface processes introduce stochastic variability into the measured current. To capture these effects, the current dynamics are modeled by the SDE

$$dI(t) = \kappa (\mu(T(t), E(t); \phi, \beta) - I(t)) dt + \xi \sqrt{I(t)} dW(t), \quad I(0) = I_0, \quad (4)$$

where

- $\kappa > 0$ is the relaxation rate toward the equilibrium emission level;
- $\xi > 0$ is the diffusion parameter controlling fluctuation intensity;
- $\sqrt{I(t)}$ reflects shot-noise–like multiplicative fluctuations typical of electron emission;
- $T(t)$ and $E(t)$ are externally prescribed control inputs.

The system is observed at discrete times $t_k = kh$, yielding $I_k = I(t_k)$, $T_k = T(t_k)$, and $E_k = E(t_k)$. The square-root diffusion coefficient is chosen to reflect shot-noise-type fluctuations and to preserve positivity of the current; in the numerical study, Eq. (4) is simulated by a positivity-preserving Milstein-type scheme.

The transition density of the nonlinear diffusion (4) is not available in closed form, and therefore the exact likelihood cannot be computed analytically. In the next section we construct a quasi-likelihood approximation based on a Hermite expansion of the transition density, adapted to the nonlinear drift and multiplicative diffusion present in (4).

2.2 Hermite parametrix expansion of the transition density

A decomposition of the transition probability density in terms of Hermite functions, obtained using the parametrix method, was derived in [26]. This expansion provides an explicit approximation to the density of (4), and therefore allows the construction of an approximate (quasi-)likelihood functional for discretely observed diffusions.

Before introducing the expansion, we recall the notation used in [26]. A scalar function

$$f_t(\theta, x, y), \quad t \in (0, 1], \quad x, y \in \mathbb{R},$$

is said to be of *order at least r* if there exist constants $C, c > 0$, independent of (θ, x, y) and t , such that

$$|f_t(\theta, x, y)| \leq Ct^r \phi_{ct}(x - y),$$

where ϕ_{ct} denotes a Gaussian kernel with variance proportional to t .

We assume throughout that the drift and diffusion coefficients satisfy standard local Lipschitz and linear growth conditions. The accuracy of the Hermite expansion is of order 1, corresponding to the Milstein scheme. We use the first-order Hermite (parametrix) approximation, which is consistent with first-order schemes in the strong sense and matches the Milstein discretization order used to generate trajectories in the Monte Carlo study. Accordingly, the QML functional is built from the first-order transition-density approximation.

To perform the parametrix expansion, we introduce the following notation. Let $b = \sigma^2$, $n = \lfloor T/h \rfloor$ denote the number of observations, and let θ be the vector of unknown parameters. For arbitrary vectors $\mathbf{a}, x, y \in \mathbb{R}^d$ and a strictly positive definite matrix $\mathbf{b} \in \mathbb{R}^{d \times d}$, define

$$\begin{aligned} \Upsilon_t(\mathbf{a}, \mathbf{b}; x, y) &= (2\pi t)^{-d/2} (\det \mathbf{b})^{-1/2} \\ &\quad \times \exp \left(-\frac{1}{2t} (\mathbf{b}^{-1}(y - x - \mathbf{a}t), y - x - \mathbf{a}t) \right), \end{aligned} \quad (5)$$

the zero-order Hermite function (a Gaussian kernel with drift $\mathbf{a}$ and covariance $t\mathbf{b}$).

Higher-order Hermite functions are defined by

$$\Upsilon_t^{(i_1, \dots, i_n)}(\mathbf{a}, \mathbf{b}; x, y) = \partial_{x_{i_1}} \cdots \partial_{x_{i_n}} \Upsilon_t(\mathbf{a}, \mathbf{b}; x, y), \quad (i_1, \dots, i_n) \in \{1, \dots, d\}^n.$$

These satisfy the factorization

$$\Upsilon_t^{(i_1, \dots, i_n)}(\mathbf{a}, \mathbf{b}; x, y) = H_t^{(i_1, \dots, i_n)}(\mathbf{a}, \mathbf{b}; x, y) \Upsilon_t(\mathbf{a}, \mathbf{b}; x, y), \quad (6)$$

where $H^{(\cdot)}$ are the corresponding Hermite polynomials (the order-zero polynomial is identically 1).

Under the assumptions

- $a \in C^\varrho$, $b \in C^{1+\varrho}$ for some $\varrho \in (0, 1]$,
- a , b , and ∇b are bounded,
- $b(x) = \sigma(x)\sigma(x)^\top$ is uniformly elliptic,

the transition density admits the approximation

$$p_t(x, y) = \Upsilon_t(x, y) + t^2 \sum_{i,j,k=1}^d c_{ijk}(x) \Upsilon_t^{(i,j,k)}(x, y),$$

where

$$c_{ijk}(x) = \frac{1}{4} \sum_{l=1}^d b_{kl}(x) \partial_{x_l} b_{ij}(x), \quad i, j, k = 1, \dots, d,$$

and with

$$\mathbf{a} = a(\theta, x), \quad \mathbf{b} = b(\theta, x).$$

Explicitly,

$$\Upsilon_t^{(i,j,k)}(x, y) = \Upsilon_t(\mathbf{a}, \mathbf{b}; x, y) H_t^{(i,j,k)}(\mathbf{a}, \mathbf{b}; x, y),$$

and

$$\begin{aligned} H_t^{(i,j,k)}(\mathbf{a}, \mathbf{b}; x, y) &= \frac{1}{t^3} (\mathbf{b}^{-1}(y - x - \mathbf{a}t))_i (\mathbf{b}^{-1}(y - x - \mathbf{a}t))_j \\ &\quad \times (\mathbf{b}^{-1}(y - x - \mathbf{a}t))_k - \frac{1}{t^2} \mathbf{b}_{ij}^{-1} (\mathbf{b}^{-1}(y - x - \mathbf{a}t))_k \\ &\quad - \frac{1}{t^2} \mathbf{b}_{ik}^{-1} (\mathbf{b}^{-1}(y - x - \mathbf{a}t))_j - \frac{1}{t^2} \mathbf{b}_{jk}^{-1} (\mathbf{b}^{-1}(y - x - \mathbf{a}t))_i. \end{aligned}$$

This first-order Hermite expansion forms the basis for constructing the quasi-likelihood estimator studied in Section 3.

3 Estimators

We consider three estimator classes: a likelihood-based quasi-maximum likelihood estimator (QML) constructed from the Hermite density approximation, contrast-based estimators (least squares and least absolute error) derived from discrete SDE schemes, and one-step Newton-type refinements based on the QML score and Hessian.

3.1 Quasi-maximum likelihood estimator

Let $\mathfrak{X}^n = (X_0, X_1, \dots, X_n)$ denote the discrete observations of the process at times $t_k = kh$. For a parameter vector $\theta = (\phi, \beta, \kappa, \xi)$, let $p_h(\theta; x, y)$ denote the Hermite-parametrix approximation of the transition density over a single time step of length h . The quasi-log-likelihood is defined by

$$\ell_n(\theta) = \ln QL_n(\theta; \mathfrak{X}^n) = \sum_{k=1}^n \ln p_h(\theta; X_{k-1}, X_k). \quad (7)$$

The quasi-maximum likelihood estimator (QMLE) is obtained by

$$\hat{\theta}_n = \arg\max_{\theta \in \Theta} \ell_n(\theta), \quad \text{equivalently} \quad \hat{\theta}_n = \arg\min_{\theta \in \Theta} [-\ell_n(\theta)]. \quad (8)$$

Since the true transition density of Eq. (4) is unknown, the QMLE $\hat{\theta}_n$ maximizes a likelihood constructed from an explicit, systematically derived approximation. Under standard assumptions for the parametrix expansion ($a \in C^e$, $b \in C^{1+e}$, bounded derivatives, and uniform ellipticity), the resulting estimator is first-order equivalent to the exact maximum likelihood estimator.

3.2 Least-squares and least-absolute-error estimators

In addition to the quasi-likelihood approach, we consider several contrast-type estimators constructed directly from discrete approximations of the SDE. These estimators are easy to compute, require no evaluation of transition densities, and are widely used in parametric diffusion models (see, e.g., [27–33]).

We distinguish four estimators:

- LSE-1: Euler-based least-squares estimator,
- LSE-2: Milstein-based least-squares estimator,
- LAE-1: Euler-based least-absolute-error estimator,
- LAE-2: Milstein-based least-absolute-error estimator.

All of them are constructed from a discrete approximation of the increment

$$\Delta X_k = X_k - X_{k-1}.$$

Euler-based least-squares estimator (LSE-1). Using the Euler scheme

$$\Delta X_k \approx A(X_{k-1}, \theta)h + \sigma(X_{k-1}, \theta)\sqrt{h}Z_k, \quad Z_k \sim N(0, 1),$$

the classical CLS loss function is

$$Q_{\text{LSE}}^{\text{Euler}}(\theta) = \sum_{k=1}^n \frac{(\Delta X_k - A(X_{k-1}, \theta)h)^2}{\sigma(X_{k-1}, \theta)^2 h}. \quad (9)$$

The estimator is

$$\hat{\theta}_{\text{LSE1}} = \arg \min_{\theta \in \Theta} Q_{\text{LSE}}^{\text{Euler}}(\theta).$$

Consistency and asymptotic normality for Euler CLS estimators are established in [31, 33] under standard conditions.

Milstein increment and conditional variance

The Milstein scheme yields

$$\Delta X_k = A(X_{k-1}, \theta)h + \sigma(X_{k-1}, \theta)W_k + \frac{1}{2}\sigma(X_{k-1}, \theta)\sigma'(X_{k-1}, \theta)(W_k^2 - h),$$

$$W_k \sim N(0, h).$$

Define

$$a_k = \sigma(X_{k-1}, \theta), \quad c_k = \frac{1}{2}\sigma(X_{k-1}, \theta)\sigma'(X_{k-1}, \theta),$$

and

$$\eta_k(\theta) = a_k W_k + c_k(W_k^2 - h).$$

Then the increment can be written as

$$\Delta X_k - A(X_{k-1}, \theta)h = \eta_k.$$

Using standard Gaussian identities,

$$\mathbb{E}[W_k^2] = h, \quad \mathbb{E}[W_k(W_k^2 - h)] = 0, \quad \mathbb{E}[(W_k^2 - h)^2] = 2h^2,$$

the conditional variance is

$$\mathbb{E}[\eta_k^2 \mid X_{k-1}] = a_k^2 h + 2c_k^2 h^2 = \sigma(X_{k-1}, \theta)^2 h + \frac{1}{2}\sigma(X_{k-1}, \theta)^2 (\sigma'(X_{k-1}, \theta))^2 h^2.$$

Milstein-based least-squares estimator (LSE-2)

By replacing the Euler variance in (9) with the Milstein conditional variance, we obtain

$$Q_{\text{LSE}}^{\text{Mil}}(\theta) = \sum_{k=1}^n \frac{(\Delta X_k - A(X_{k-1}, \theta)h)^2}{\sigma(X_{k-1}, \theta)^2 h + \frac{1}{2}\sigma(X_{k-1}, \theta)^2 (\sigma'(X_{k-1}, \theta))^2 h^2}. \quad (10)$$

The corresponding estimator is

$$\hat{\theta}_{\text{LSE2}} = \arg \min_{\theta} Q_{\text{LSE}}^{\text{Mil}}(\theta).$$

Euler-based least-absolute-error estimator (LAE-1).

The Euler LAE contrast is

$$Q_{\text{LAE}}^{\text{Euler}}(\theta) = \sum_{k=1}^n \frac{|\Delta X_k - A(X_{k-1}, \theta)h|}{|\sigma(X_{k-1}, \theta)|\sqrt{h}}. \quad (11)$$

The estimator is

$$\hat{\theta}_{\text{LAE1}} = \arg \min_{\theta} Q_{\text{LAE}}^{\text{Euler}}(\theta).$$

Milstein-based least-absolute-error estimator (LAE-2).

Using the Milstein conditional standard deviation

$$\sigma_k(\theta) = \sqrt{\sigma(X_{k-1}, \theta)^2 h + \frac{1}{2} \sigma(X_{k-1}, \theta)^2 (\sigma'(X_{k-1}, \theta))^2 h^2}.$$

the Milstein LAE contrast becomes

$$Q_{\text{LAE}}^{\text{Mil}}(\theta) = \sum_{k=1}^n \frac{|\Delta X_k - A(X_{k-1}, \theta)h|}{\sigma_k(\theta)}. \quad (12)$$

The estimator is

$$\hat{\theta}_{\text{LAE2}} = \arg \min_{\theta} Q_{\text{LAE}}^{\text{Mil}}(\theta).$$

3.3 One-step estimator

Any of the estimators introduced earlier (QMLE, LSE-1/2, LAE-1/2) can serve as an initial consistent estimator for a one-step refinement. The one-step estimator applies a single Newton-Raphson correction to the initial estimator,

$$\theta_{\text{1step}} = \theta_{\text{start}} - H^{-1}(\theta_{\text{start}}; \mathfrak{X}^n) \nabla \ell_n(\theta_{\text{start}}), \quad (13)$$

where $\theta_{\text{start}} = (\phi_{\text{start}}, \beta_{\text{start}}, \kappa_{\text{start}}, \xi_{\text{start}})$ denotes an initial estimator, $\nabla \ell_n$ is the gradient vector, and H is the full 4×4 Hessian matrix.

Since

$$\ell_n(\theta) = \sum_{k=1}^n \ln p_h(\theta; X_{k-1}, X_k),$$

the gradient takes the form

$$\nabla \ell_n(\theta) = \begin{pmatrix} \partial_{\phi} \ell_n(\theta) \\ \partial_{\beta} \ell_n(\theta) \\ \partial_{\kappa} \ell_n(\theta) \\ \partial_{\xi} \ell_n(\theta) \end{pmatrix}, \quad \partial_{\theta_j} \ell_n(\theta) = \sum_{k=1}^n \frac{\partial_{\theta_j} p_k(\theta)}{p_k(\theta)}.$$

The Hessian matrix is the symmetric 4×4 matrix

$$H(\theta) = \begin{pmatrix} \partial_{\phi\phi} & \partial_{\phi\beta} & \partial_{\phi\kappa} & \partial_{\phi\xi} \\ \partial_{\phi\beta} & \partial_{\beta\beta} & \partial_{\beta\kappa} & \partial_{\beta\xi} \\ \partial_{\phi\kappa} & \partial_{\beta\kappa} & \partial_{\kappa\kappa} & \partial_{\kappa\xi} \\ \partial_{\phi\xi} & \partial_{\beta\xi} & \partial_{\kappa\xi} & \partial_{\xi\xi} \end{pmatrix}, \quad (14)$$

where, for $\theta_i, \theta_j \in \{\phi, \beta, \kappa, \xi\}$,

$$\partial_{\theta_i \theta_j} \ell_n(\theta) = \sum_{k=1}^n \left[\frac{\partial_{\theta_i \theta_j} p_k(\theta)}{p_k(\theta)} - \frac{\partial_{\theta_i} p_k(\theta) \partial_{\theta_j} p_k(\theta)}{p_k(\theta)^2} \right]. \quad (15)$$

Using the parametrix expansion from [26], derivatives of the Hermite functions satisfy,

$$\partial_{\theta_j} \Upsilon_t^{(i_1, \dots, i_n)}(\theta; x, y) = t \Upsilon_t^{(i_1, \dots, i_n, j)}(\theta; x, y), \quad (16)$$

$$\partial_{\xi} \Upsilon_t^{(i_1, \dots, i_n)}(\theta; x, y) = \frac{t}{2} \Upsilon_t^{(i_1, \dots, i_n, \xi, \xi)}(\theta; x, y), \quad (17)$$

where ξ enters through the diffusion coefficient $b = \sigma^2$.

Because higher-order Hermite functions remain explicit, all required first and second derivatives can be computed analytically. Although one-step estimators are asymptotically equivalent to fully iterated procedures under idealized conditions, their finite-sample performance can deteriorate when the score function is distorted by nonlinear or correlated inputs, as demonstrated in Section 5.2.

4 Regularization

4.1 Regularization of L^p estimators

Recall that, for L^p estimators, the objective function has the form

$$Q_{L^p}^{\text{Eul}}(\theta) = \sum_{k=1}^n \frac{|\Delta X_k - A(X_{k-1}, \theta)h|^p}{|\sigma(X_{k-1}, \theta)\sqrt{h}|^p}. \quad (18)$$

or

$$Q_{L^p}^{\text{Mil}}(\theta) = \sum_{k=1}^n \frac{|\Delta X_k - A(X_{k-1}, \theta)h|^p}{\left(\sigma(X_{k-1}, \theta)^2 h + \frac{1}{2} \sigma(X_{k-1}, \theta)^2 (\sigma'(X_{k-1}, \theta))^2 h^2\right)^{p/2}}. \quad (19)$$

The diffusion parameter ξ enters these contrasts only through the normalizing denominator. Consequently, the objective may decrease as ξ increases, and a finite minimizer with respect to ξ is not guaranteed without an additional constraint. This difficulty appears for both LSE and LAE estimators, and the Milstein correction does not remove it. We therefore use the regularized objective

$$\hat{Q}(\theta) = Q_{L^p}(\theta) + \lambda(\xi - \xi_0)^2,$$

where $Q_{L^p}(\theta)$ denotes the corresponding LSE-1, LSE-2, LAE-1, or LAE-2 contrast, and ξ_0 is the initial value used in the Hooke-Jeeves optimization method. The choice of λ is crucial: small values lead to unstable diffusion estimates, while excessively large values produce oversmoothing and bias. We compare six regularization-parameter selection rules and consider residual smoothness as an additional diagnostic. The selection rules and diagnostic are summarized in Table 1.

To quantitatively compare the criteria, we introduce the following diagnostics:

- $C_1 = |\xi^* - \xi_{\text{true}}|$ (parameter error),
- $C_2 = R(\lambda^*)$ (normalized residual level),
- $C_3 = P(\lambda^*)/(R(\lambda^*) + P(\lambda^*))$ (regularization dominance),
- $C_4 = S(\lambda^*)$ (residual smoothness).

All optimization problems in λ were solved by continuous golden-section search. The regularization results are presented separately for LSE-based estimators in Table 2 and LAE-based estimators in Table 3.

Across all four estimators, the discrepancy principle consistently selects λ values of order 10–20, yielding the smallest diffusion parameter errors and balanced residual-penalty ratios. In contrast, UPRE and GCV systematically under-regularize, leading to overfitting and divergence of ξ^* , while L-curve and quasi-optimality over-regularize and bias the estimates. Overall, the discrepancy prin-

Table 1. Regularization parameter selection criteria used for L^p estimators.

Criterion	Definition	Interpretation / Remarks
Discrepancy	$R(\lambda) \approx \tau \hat{\sigma}^2$	Balances data misfit with estimated noise level. Robust when noise variance is correctly normalized; performs well in SDE settings.
UPRE	$(R(\lambda) - \hat{\sigma}^2)^2$	Minimizes deviation from noise variance. Asymptotically equivalent to discrepancy principle under correct noise specification, but sensitive to variance misspecification.
GCV	$\frac{R(\lambda)}{(1 - \text{df}(\lambda))^2}$	Data-driven criterion avoiding explicit noise estimation. In nonlinear diffusion models often under-regularizes and yields unstable diffusion estimates.
L-curve	max curvature of $(\log R(\lambda), \log P(\lambda))$	Geometric balance between data fidelity and regularization. Provides stable intermediate solutions but may over-smooth in strongly nonlinear problems.
Quasi-optimality	$\arg \min_{\lambda} \left \frac{dR(\lambda)}{d\lambda} \right $	Detects plateau regions of the residual curve. Highly sensitive to flat residual behavior and noise; unreliable in monotone regimes.
Hanke–Raus	$\frac{R(\xi_{\lambda})}{ R(\xi_{\lambda}) - R(\xi_{\lambda/2}) }$	Selects λ where residual becomes locally insensitive to further regularization. Often under-regularizes in diffusion estimation problems.
Residual smoothness	$\left \frac{d^2 R(\lambda)}{d\lambda^2} \right $	Penalizes higher-order residual oscillations. Useful as a supplementary diagnostic rather than a standalone selection rule.

Table 2. Regularization parameter selection for LSE estimators. Each block reports $(\lambda^*, \xi^*, C_1, C_2, C_3, C_4)$ for the corresponding scheme.

	Method	λ^*	ξ^*	C_1	C_2	C_3	C_4
LSE-1 (Euler)	Discrepancy	10.69	0.491	0.0094	1.000	0.280	6.25
	UPRE	0.0177	2.000	1.500	0.060	0.459	0.38
	GCV	0.0177	2.000	1.500	0.060	0.459	0.38
	L-curve	49.65	0.385	0.1150	1.624	0.181	10.15
	Quasi-opt	99.99	0.354	0.1458	1.919	0.133	11.99
	Hanke-Raus	0.0354	1.695	1.195	0.084	0.451	0.52
LSE-2 (Milstein)	Discrepancy	19.20	0.440	0.0596	1.000	0.275	6.22
	UPRE	0.0061	2.000	1.500	0.0118	0.597	0.071
	GCV	0.0061	2.000	1.500	0.0118	0.597	0.071
	L-curve	46.89	0.386	0.1136	1.358	0.205	8.45
	Quasi-opt	99.99	0.353	0.1466	1.655	0.146	10.38
	Hanke-Raus	0.0121	1.762	1.262	0.018	0.586	0.11

Table 3. Regularization parameter selection for LAE estimators. Each block reports $(\lambda^*, \xi^*, C_1, C_2, C_3, C_4)$ for the corresponding scheme.

	Method	λ^*	ξ^*	C_1	C_2	C_3	C_4
LAE-1 (Euler)	Discrepancy	12.68	0.399	0.101	1.000	0.110	3.08
	UPRE	0.0293	2.000	1.500	0.199	0.298	0.12
	GCV	0.0293	2.000	1.500	0.199	0.298	0.12
	L-curve	7.863	0.434	0.0656	0.918	0.134	2.60
	Quasi-opt	99.99	0.320	0.180	1.248	0.030	4.80
	Hanke-Raus	0.0587	1.611	1.111	0.248	0.289	0.19
LAE-2 (Milstein)	Discrepancy	22.82	0.368	0.132	1.000	0.096	3.08
	UPRE	0.0224	2.000	1.500	0.085	0.431	0.023
	GCV	0.0224	2.000	1.500	0.085	0.431	0.023
	L-curve	5.717	0.480	0.0203	0.732	0.186	1.65
	Quasi-opt	99.99	0.320	0.180	1.170	0.034	4.22
	Hanke-Raus	0.0448	1.664	1.164	0.118	0.414	0.043

ciple emerges as the most reliable and stable criterion for diffusion regularization in the thermionic emission SDE model. In the numerical experiments below we therefore emphasize the discrepancy principle as a primary selection rule, since it yields the most stable diffusion estimates across scenarios.

4.2 Selection of the regularization parameter for QML

In the quasi-maximum likelihood (QML) framework, the estimation of the Schottky parameter β is known to be weakly identifiable due to its nonlinear and ex-

ponentially damped contribution to the drift term of the thermionic emission model. To stabilize the estimation, we consider a Tikhonov-type regularization of the form

$$Q_\lambda(\theta) = L_1(\theta) + \lambda(\log \beta - \log \beta_0)^2, \quad (20)$$

where $L_1(\theta)$ denotes the Hermite based QML objective, β_0 is a reference value, and $\lambda > 0$ is the regularization parameter.

The choice of λ is crucial and is addressed below using two standard criteria: the L-curve method and cross validation.

L-curve criterion

The L-curve method is a deterministic regularization strategy originating from inverse problem theory. For a discrete set of candidate values $\{\lambda_j\}$, one computes the corresponding regularized estimates $\hat{\theta}_{\lambda_j}$ and evaluates:

$$\text{misfit}(\lambda_j) = L_1(\hat{\theta}_{\lambda_j}), \quad \text{penalty}(\lambda_j) = (\log \hat{\beta}_{\lambda_j} - \log \beta_0)^2.$$

The L-curve is defined as the parametric curve

$$(\log \text{misfit}(\lambda), \log \text{penalty}(\lambda)),$$

and the optimal regularization parameter λ_{LC} is chosen at the point of maximum curvature, corresponding to the best compromise between data fidelity and regularization strength. This criterion is purely deterministic and does not rely on resampling or data splitting.

Cross-validation

As an alternative, the regularization parameter can be selected by minimizing a validation likelihood. The data are split into contiguous time segments, and for each candidate λ the model is fitted on the training part, while the unregularized QML objective $L_1(\theta)$ is evaluated on the validation part. The cross-validation choice λ_{CV} is defined as

$$\lambda_{CV} = \arg \min_{\lambda} L_1^{(\text{val})}(\hat{\theta}_{\lambda}).$$

Unlike the L-curve method, cross validation focuses exclusively on predictive performance and does not explicitly control parameter stability.

Quantitative comparison

Table 4 summarizes the results obtained by both criteria for a single simulated dataset. In addition to the selected λ , we report the data misfit, the penalty level, the total regularized objective, and the parameter error $\|\hat{\theta} - \theta_{\text{true}}\|$ (available in simulation experiments).

Table 4. Comparison of L-curve and cross validation criteria for selecting the regularization parameter in QML.

Method	λ	$L_1(\hat{\theta})$	Penalty	$Q_\lambda(\hat{\theta})$	Param. error
L-curve	1.44×10^1	1.147886	6.78×10^{-10}	1.147886	5.56×10^{-2}
Cross-validation	4.28×10^{-2}	1.147898	2.28×10^{-8}	1.147898	6.01×10^{-2}

Theoretical interpretation

Both criteria lead to nearly identical values of the QML misfit, indicating that the regularization does not distort the likelihood fit. However, the L-curve criterion selects a substantially larger value of λ , resulting in a much smaller regularization residual and improved parameter accuracy. This behavior reflects the weak identifiability of β : cross validation prioritizes predictive likelihood and therefore favors weaker regularization, while the L-curve explicitly balances variance reduction against bias. Consequently, for the present thermionic emission model, the L-curve criterion provides a more reliable and stable choice of the regularization parameter.

5 Results

5.1 Software implementation

The numerical experiments were implemented in a dedicated Windows Forms application developed in C#. The software integrates the PDE–OU input generator, the stochastic thermionic-emission simulator, Hermite-based QML, contrast estimators, one-step corrections, and export routines for numerical tables and diagnostic plots. A representative screenshot of the implementation is shown in Figure 1.

In the numerical experiments we use the thermionic-emission model (4) with equilibrium current (3). Unless stated otherwise, we set $h = 1$, $I_0 = 0.5$, and fix $A_{\text{eff}} = 10^6$ and $A_{\text{emit}} = 10^{-6}$, so that $A_{\text{eff}}A_{\text{emit}} = 1$. The true parameters are $\phi = 2.4$, $\beta = 3 \cdot 10^{-5}$, $\kappa = 0.3$, and $\xi = 0.5$. For bounded $E(t)$, the drift $\kappa(\mu(t) - I)$ is mean-reverting, while the square-root diffusion $\xi\sqrt{I(t)}$ yields state-dependent variability. Optimization used a derivative-free search for initialization and log-coordinates for positive parameters (β, κ, ξ) .

5.2 PDE–OU generation of temperature and electric-field inputs

This experiment is designed to mimic physically realistic operating conditions and intentionally departs from the simplifying assumptions underlying many asymptotic one-step approximations, by introducing coupled PDE dynamics and correlated stochastic excitation. In this experiment, both the temperature

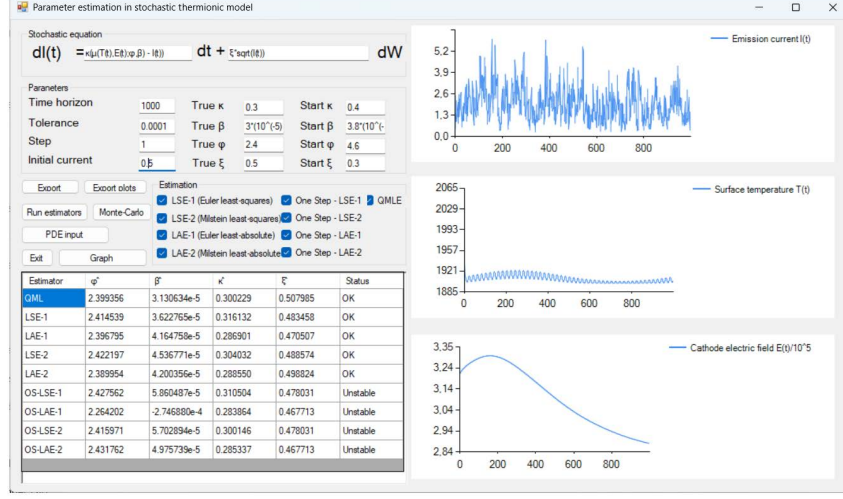


Figure 1. Software interface used for simulation and parameter estimation in the stochastic thermionic emission model. The program generates PDE–OU temperature and electric-field inputs, simulates the emission-current diffusion, applies Hermite-based QML and contrast estimators, and exports parameter estimates and diagnostic plots.

$T(t)$ and the electric field $E(t)$ are not prescribed analytically, but are generated endogenously by coupled PDE solvers driven by Ornstein–Uhlenbeck (OU) boundary/forcing processes. The aim is to mimic more realistic operating conditions where the cathode temperature and gap field evolve dynamically with thermal diffusion, radiative losses, and space-charge effects, while still retaining controlled excitation at multiple time scales (slow drift + fast ripple).

For all multi-panel box plots and QQ plots shown below, panels are arranged row-wise as follows: QML, LSE–1, LSE–2; LAE–1, LAE–2, OS–LSE–1; OS–LSE–2, OS–LAE–1, OS–LAE–2.

Temperature input from a 1D heat PDE with OU forcing

Let $x \in [0, L]$ denote the cathode depth coordinate. The temperature field $T(x, t)$ is modeled by the one-dimensional heat equation with a radiative sink and an additional linear relaxation term:

$$\rho c_p \partial_t T = k \partial_{xx} T + Q(x, t) - \varepsilon \sigma_{\text{SB}} (T^4 - T_{\text{amb}}^4) - \rho c_p \lambda_{\text{lin}} (T - T_{\text{amb}}), \quad x \in (0, L). \quad (21)$$

Here ρ is the density, c_p the specific heat, k the thermal conductivity, ε the emissivity, σ_{SB} the Stefan–Boltzmann constant, and T_{amb} the ambient temperature. Heating is applied near the surface ($x = 0$) in a thin layer:

$$Q(x, t) = Q(t) \mathbf{1}_{[0, x_h]}(x), \quad x_h \text{ (heater width)}.$$

The scalar heater power $Q(t)$ follows an Ornstein–Uhlenbeck dynamics around a deterministic mean schedule:

$$dQ(t) = a_Q(Q_m(t) - Q(t)) dt + \sigma_Q dW_Q(t), \quad (22)$$

with

$$Q_m(t) = Q_0 + A_Q \sin(2\pi f_Q t) + A_f \sin(2\pi f_f t). \quad (23)$$

The boundary condition at $x = 0$ is a prescribed heat flux (Neumann condition), also containing slow and fast excitations:

$$-k \partial_x T(0, t) = q(t), \quad q(t) = q_0 + q_A \sin(2\pi f_q t) + q_{Af} \sin(2\pi f_{qf} t), \quad (24)$$

and a zero-gradient condition at the far boundary $x = L$, $\partial_x T(L, t) = 0$. The temperature input used in the emission SDE is the surface value

$$T(t) \equiv T(0, t). \quad (25)$$

For numerical simulation, (21) is discretized on a uniform grid $M = 101$ over $L = 2 \cdot 10^{-3}$ m and advanced using an implicit finite-difference scheme with $n_{\text{sub}} = 25$ substeps per macro step h (i.e., $\Delta t = h/n_{\text{sub}}$). The nonlinear radiation term is handled by Picard iterations with a local linearization of T^4 using

$$T^4 \approx (T^*)^4 + 4(T^*)^3(T - T^*),$$

which yields a linear “implicit diffusion + linear sink” solve. The resulting tridiagonal system is solved by the Thomas algorithm.

Electric-field input from a 1D Poisson problem with OU voltage

Let $z \in [0, d]$ denote the electrode gap coordinate (d is the gap width). The electric potential $\varphi(z, t)$ is obtained as the solution of a 1D Poisson equation,

$$\partial_{zz}\varphi(z, t) = -\frac{\rho(z, t)}{\varepsilon_0}, \quad z \in (0, d), \quad (26)$$

with Dirichlet boundary conditions

$$\varphi(0, t) = 0, \quad \varphi(d, t) = V(t). \quad (27)$$

The applied voltage $V(t)$ is modeled by an Ornstein–Uhlenbeck process with a two-scale mean schedule:

$$dV(t) = a_V(V_m(t) - V(t)) dt + \sigma_V dW_V(t), \quad (28)$$

where

$$V_m(t) = V_0 + A_{\text{slow}} \cos(2\pi f_{\text{slow}} t + \frac{\pi}{2}) + A_{\text{fast}} \sin(2\pi f_{\text{fast}} t). \quad (29)$$

Here $a_Q > 0$ and $a_V > 0$ are mean reversion rates, while σ_Q and σ_V control the intensity of stochastic fluctuations.

To represent space-charge effects in a simple closure form, the charge density $\rho(z, t)$ is computed from a (bounded) current density J and a drift velocity v obtained from the local potential drop:

$$\rho(z, t) = \frac{J}{v(z, t)}, \quad v(z, t) \approx \sqrt{\frac{2e \max\{V(t) - \varphi(z, t), V_{\min}\}}{m_e}}, \quad (30)$$

where e is the elementary charge and m_e the electron mass; additional floors and caps are used to avoid numerical runaway. The Poisson problem (26)–(30) is solved at each macro time t_n by a fixed-point iteration with SOR sweeps and under-relaxation. Finally, the electric-field input is taken as the cathode field

$$E(t) \equiv |-\partial_z \varphi(0, t)| \approx \left| -\frac{\varphi(z_1, t) - \varphi(0, t)}{\Delta z} \right|. \quad (31)$$

Numerical discretization (fixed-point + SOR)

At each time step, the Poisson problem is solved by a fixed-point iteration. Given a current iterate $\varphi^{(m)}$, the cathode field $E_0^{(m)}$ is computed, the current density $J(T, E_0^{(m)})$ is evaluated, the charge density $\rho^{(m)}$ is constructed, and the discrete Poisson equation is updated by successive over-relaxation (SOR). The SOR relaxation parameter is set to $w_{\text{SOR}} = 1.4$. To stabilize the outer fixed-point iteration, an additional under-relaxation with factor $\omega_{\text{FP}} = 0.2$ is applied:

$$\varphi^{(m+1)} = (1 - \omega_{\text{FP}})\varphi^{(m)} + \omega_{\text{FP}}\varphi_{\text{SOR}}^{(m)}.$$

Parameter settings (PDE–OU case)

In the simulations we use $h = 1$ and a time horizon $T = 1000$. The temperature PDE parameters are $\rho = 8000$, $c_p = 500$, $k = 20$, $T_{\text{amb}} = 300$, $\varepsilon = 0.8$, $\lambda_{\text{lin}} = 0.05$, with OU heater parameters $a_Q = 3.0$, $\sigma_Q = 3.0 \cdot 10^5$, $Q_{\min} = 0$, $Q_{\max} = 3 \cdot 10^7$, and mean schedule components $Q_0 = 2.4 \cdot 10^7$, $A_Q = 0.6 \cdot 10^7$, $f_Q = 10^{-3}$, $A_f = 0.25 \cdot 10^7$, $f_f = 0.05$. The heat-flux boundary forcing is $q_0 = 3 \cdot 10^5$, $q_A = 1.5 \cdot 10^5$, $f_q = 10^{-3}$, $q_{Af} = 10^5$, $f_{qf} = 0.05$. For the electric field, the gap is $d = 0.5 \cdot 10^{-3}$ m with a Poisson grid $M = 201$. The OU voltage parameters are $V(0) = 160$, $a_V = 2.5$, $\sigma_V = 0.25$, $V \in [120, 240]$, and the mean schedule $V_0 = 140$, $A_{\text{slow}} = 35$, $f_{\text{slow}} = 10^{-3}$, $A_{\text{fast}} = 25$, $f_{\text{fast}} = 0.05$. The resulting trajectories $T(t)$ and $E(t)$ are then substituted into the emission diffusion model (4) with equilibrium current (3) to evaluate estimator performance under PDE-driven operating conditions.

Tables 5 and 6 summarize the Monte Carlo estimation results for the parameters ϕ , β , κ , and ξ .

Table 5. Estimation results under PDE-generated inputs for ϕ and β . The values for β are reported in units of 10^{-5}

Estimator	ϕ			$\beta (10^{-5})$		
	Mean	MAD	RMSD	Mean	MAD	RMSD
QML	2.3994	0.0054	0.0068	3.0999	0.0999	0.0999
LSE-1	2.4145	0.0150	0.0243	3.6228	0.6228	0.6752
LAE-1	2.3968	0.0419	0.0651	4.1648	1.1648	1.3577
LSE-2	2.4222	0.0225	0.0415	4.5368	1.5367	1.9508
LAE-2	2.3900	0.0484	0.0809	4.2004	1.2004	1.3871
OS-LSE-1	2.4276	0.0310	0.2639	5.8605	3.6483	43.9448
OS-LAE-1	2.2642	0.2398	0.5520	-27.4688	39.2779	996.7132
OS-LSE-2	2.4160	0.0213	0.0472	5.7029	3.4298	7.3996
OS-LAE-2	2.4318	0.0349	0.0803	4.9757	2.6868	5.8546

Table 6. Estimation results under PDE-generated inputs for κ and ξ

Estimator	κ			ξ		
	Mean	MAD	RMSD	Mean	MAD	RMSD
QML	0.3002	0.0172	0.0219	0.5080	0.0124	0.0155
LSE-1	0.3161	0.0305	0.0500	0.4835	0.0208	0.0232
LAE-1	0.2869	0.0273	0.0445	0.4705	0.0295	0.0297
LSE-2	0.3040	0.0212	0.0587	0.4886	0.0138	0.0211
LAE-2	0.2886	0.0250	0.0350	0.4988	0.0035	0.0045
OS-LSE-1	0.3105	0.0327	0.0815	0.4780	0.0261	0.0797
OS-LAE-1	0.2839	0.0293	0.0475	0.4677	0.0339	0.0397
OS-LSE-2	0.3001	0.0223	0.0595	0.4780	0.0261	0.0797
OS-LAE-2	0.2853	0.0265	0.0379	0.4677	0.0339	0.0397

5.2.1 Interpretation

The PDE driven inputs generate a demanding estimation setting because they introduce spatial coupling, nonlinear feedback, and correlated stochastic forcing. This has a pronounced impact on estimator performance.

For parameter ϕ , QML remains the most stable estimator, exhibiting the smallest RMSD. Standard LSE and LAE estimators show moderate bias and dispersion, while the one step (OS) variants, particularly OS-LAE-1, suffer from severe instability and large outliers, indicating sensitivity to the highly structured PDE generated inputs. The estimation of β is particularly challenging in this setting. While QML and non-OS estimators retain reasonable accuracy, OS-based methods display dramatic variance inflation and, in the case of OS-LAE-1, even sign reversal of the mean estimate. This clearly demonstrates that one step corrections can become unreliable when the score function is strongly distorted by coupled stochastic PDE dynamics. For κ , the coupled PDE-OU inputs am-

plify the variability of the estimates, especially for one-step procedures. QML again provides the best overall performance, while OS-LSE methods show systematic variance inflation. LAE-2 yields comparatively balanced bias variance trade-offs among robust estimators. Finally, for ξ , the PDE setting reveals a clear hierarchy: LAE-2 achieves the smallest RMSD, outperforming both QML and LSE variants. In contrast, OS-based estimators display pronounced instability, confirming that robustness to non-Gaussian features does not compensate for numerical sensitivity introduced by one step linearization under PDE driven inputs. Overall, the PDE case highlights that estimator behavior predicted by idealized asymptotic arguments may change substantially under physically structured, correlated input processes. In such regimes, fully iterated estimators (QML, LSE-2, LAE-2) are markedly more reliable than one step corrections.

Figure 2 shows that, for the work-function parameter ϕ , fully iterated procedures are substantially more stable under PDE generated inputs. The QML estimator is the most concentrated around the true value, followed by LSE-2, whereas one-step corrections, especially OS-LAE-1, exhibit large dispersion and are unsuitable in this nonlinear stochastic regime.

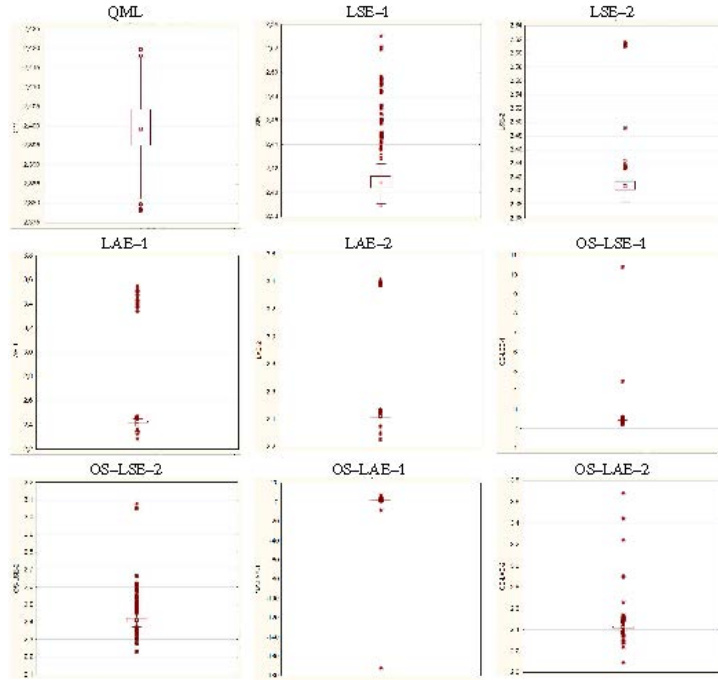


Figure 2. Box plots of estimates of the work-function parameter ϕ under PDE-generated temperature and electric-field inputs.

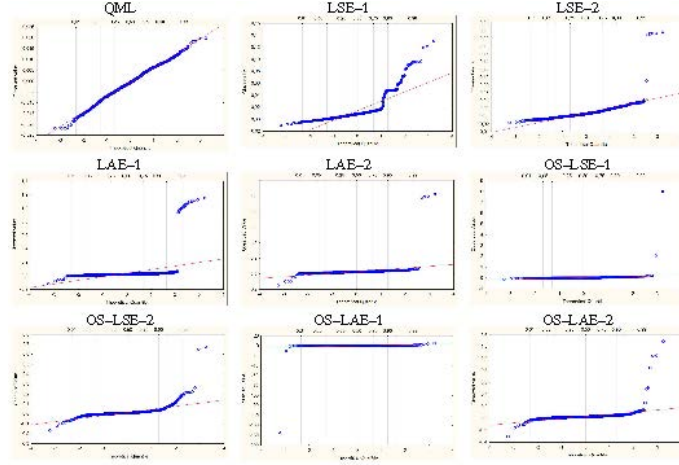


Figure 3. Normal QQ plots of estimation errors for the work-function parameter ϕ under PDE-generated temperature and electric-field inputs.

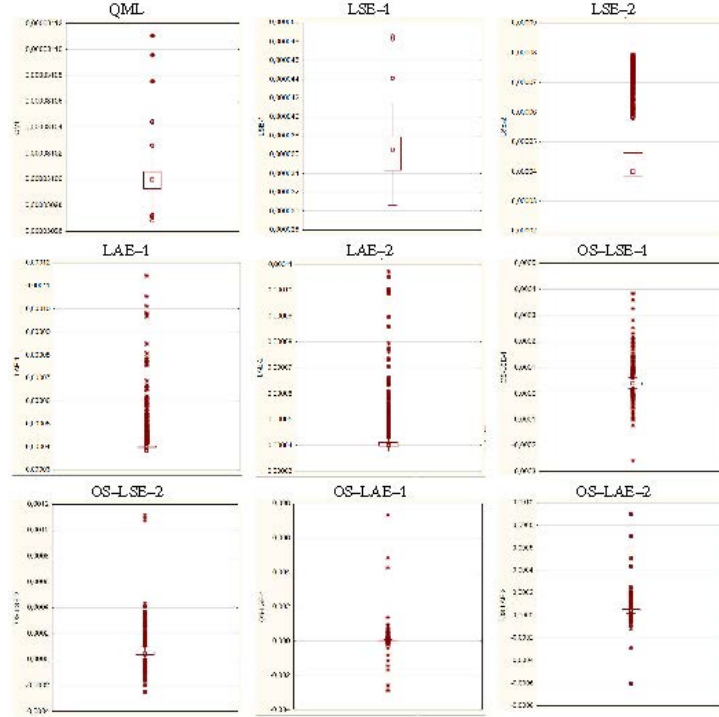


Figure 4. Box plots of estimates of the Schottky parameter β under PDE-generated temperature and electric-field inputs.

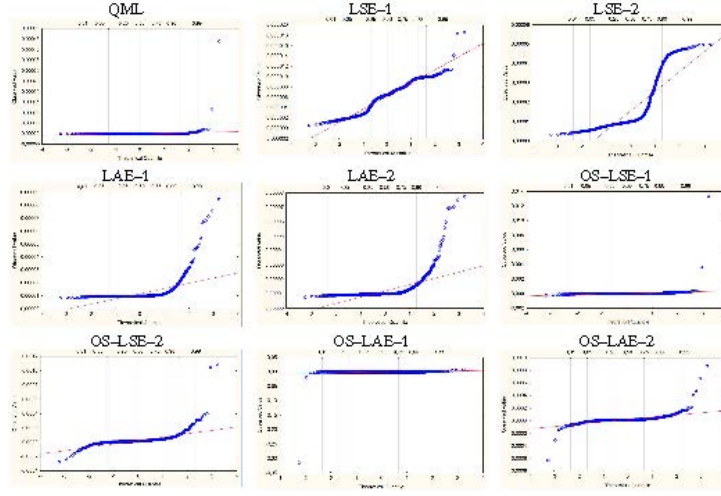


Figure 5. Normal QQ plots of estimation errors for the Schottky parameter β under PDE-generated temperature and electric-field inputs.

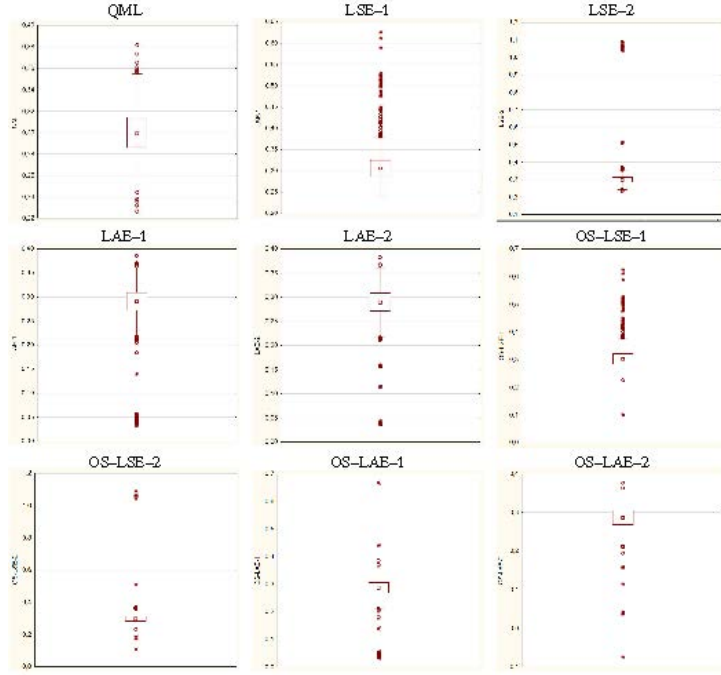


Figure 6. Box plots of estimates of the mean-reversion parameter κ under PDE-generated temperature and electric-field inputs.

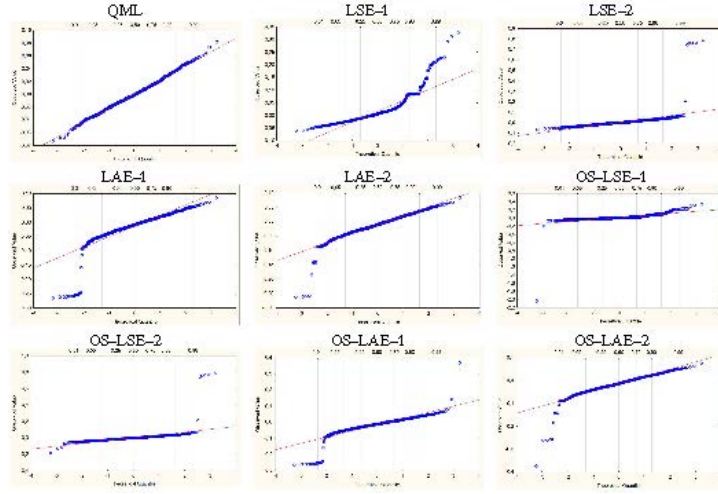


Figure 7. Normal QQ plots of estimation errors for the mean-reversion parameter κ under PDE-generated temperature and electric-field inputs.

The QQ plots in Figure 3 indicate that the Gaussian approximation is most appropriate for QML and, to a lesser extent, for the fully iterated LSE estimators. In contrast, LAE-based and one-step procedures show substantial tail distortions, confirming the loss of normal error behavior under PDE driven excitation.

For the Schottky parameter β , Figure 4 confirms the increased difficulty of estimation caused by the nonlinear field-dependent term $\beta \sqrt{E(t)}$ in the drift. QML gives the most stable estimates, while LAE-based and one-step estimators show pronounced dispersion.

The QQ diagnostics in Figure 5 show that PDE-generated inputs strongly distort the Gaussian approximation for β . QML remains the closest to normality, followed by the fully iterated LSE estimators, whereas LAE and one-step procedures produce visibly non-Gaussian error distributions.

Figure 6 shows that the coupled PDE–OU inputs amplify the variability of the relaxation-rate estimates. QML provides the most concentrated estimates of κ , while one-step procedures remain more sensitive to the nonlinear and correlated input structure.

The QQ plots in Figure 7 confirm that the estimation errors for κ are affected by non-Gaussian tails under PDE-driven forcing. The normal approximation is most reasonable for QML and fully iterated LSE estimators.

For the diffusion parameter ξ , Figure 8 shows that LAE–2 achieves the smallest RMSD, while QML and LSE-based procedures remain reasonably stable. One-step estimators display larger dispersion and outlier sensitivity.

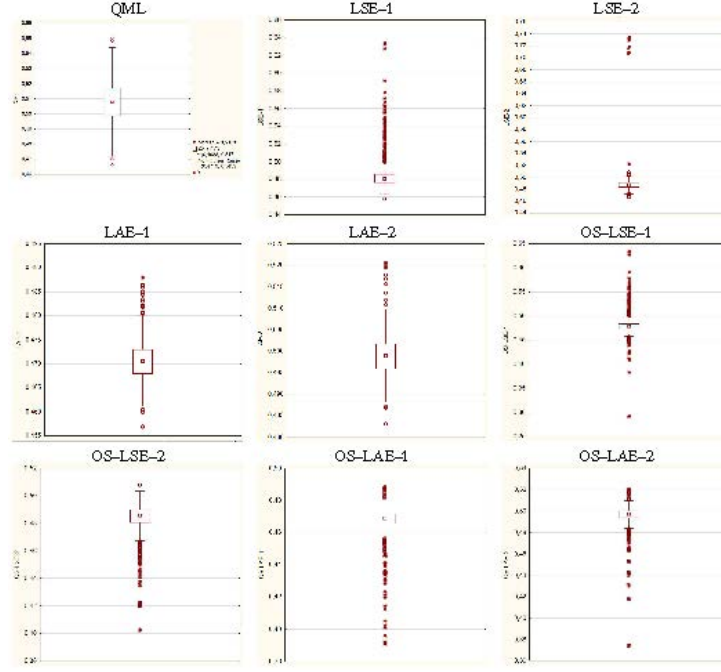


Figure 8. Box plots of estimates of the diffusion parameter ξ under PDE-generated temperature and electric-field inputs.

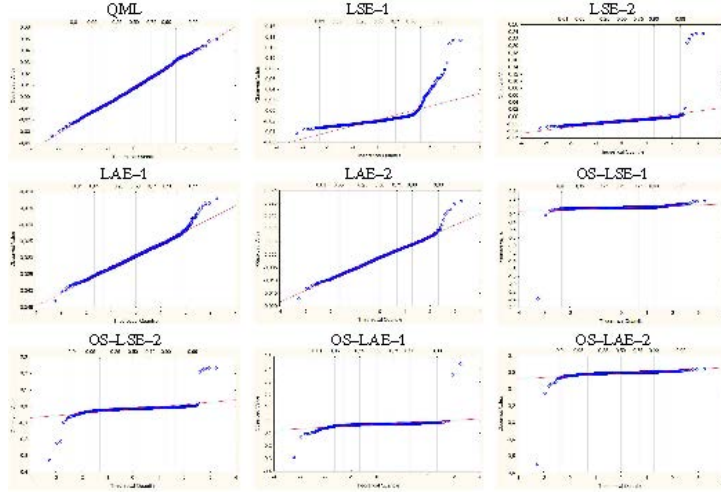


Figure 9. Normal QQ plots of estimation errors for the diffusion parameter ξ under PDE-generated temperature and electric-field inputs.

Finally, Figure 9 indicates that estimation errors for ξ are closest to Gaussian for QML and fully iterated LSE-type estimators. The LAE and one-step procedures show stronger tail effects, which is consistent with their larger finite-sample variability under PDE generated inputs.

6 Conclusion

We have developed and analyzed a likelihood based inference framework for a stochastic thermionic emission model with multiplicative noise and time dependent operating conditions. By employing a Hermite type parametrix expansion of the transition density, we constructed a quasi-likelihood estimator that avoids numerical solution of the Kolmogorov equation while retaining high order accuracy. This approach enables unified estimation of both physical emission parameters and dynamical relaxation coefficients from discretely sampled current measurements.

Through Monte Carlo experiments with PDE–OU generated operating conditions, we compared Hermite-based QML with least-squares, least-absolute-error, and one-step update estimators. The coupled heat-equation and Poisson-field inputs reveal substantial differences in estimator stability. Fully iterated procedures remain more reliable in this complex regime, whereas one-step corrections are prone to numerical instability, variance inflation, and non-Gaussian error behavior. A key implication is that estimator hierarchies expected from idealized asymptotic arguments may fail under correlated PDE-driven excitation.

These findings indicate that estimator performance in stochastic emission models depends strongly on the structure of the externally generated inputs. The proposed framework provides a computationally efficient and physically consistent tool for inference in thermionic emission models and may be extended to other applications involving diffusion processes driven by externally generated PDE dynamics.

Acknowledgements

The authors declare that no specific funding was received for this work.

References

- [1] S. Dushman (1930) Thermionic emission. *Rev. Mod. Phys.* **2**(4) 381–476.
- [2] O.W. Richardson (1921) *The Emission of Electricity from Hot Bodies*. Longmans, Green and Company, London, 2 edition.
- [3] Y. Suto, Y. Liu, M. Taniguchi (2016) Asymptotic theory of parameter estimation by a contrast function based on interpolation error. *Stat. Inference Stoch. Process.* **19**(1) 93–110.

- [4] J.-Y. Gao, Y.-F. Yang, X.-K. Zhang, S.-L. Li, P. Hu, J.-S. Wang (2020) A review on recent progress of thermionic cathode. *Tungsten* **2**(3) 289-300.
- [5] D.B. Go, J.R. Haase, J. George, J. Mannhart, R. Wanke, A. Nojeh, R. Nemanich (2017) Thermionic energy conversion in the twenty-first century: Advances and opportunities for space and terrestrial applications. *Front. Mech. Eng.* **3** 13.
- [6] G.N. Hatsopoulos, E.P. Gyftopoulos (1979) *Thermionic Energy Conversion*. MIT Press, Cambridge, MA, Volumes I-II.
- [7] C. Amorino, A. Gloter (2020) Contrast function estimation for the drift parameter of ergodic jump diffusion process. *Scand. J. Stat.* **47**(2) 279-346.
- [8] S. Donnet, J.-L. Foulley, A. Samson (2010) Bayesian analysis of growth curves using mixed models defined by stochastic differential equations. *Biometrics* **66**(3) 733-741.
- [9] D. Duffie, P. Glynn (2004) Estimation of continuous-time markov processes sampled at random time intervals. *Econometrica* **72**(6) 1773-1808.
- [10] B. Eraker (2001) MCMC analysis of diffusion models with application to finance. *J. Bus. Econ. Stat.* **19**(2) 177-191.
- [11] B. Favetto (2014) Parameter estimation by contrast minimization for noisy observations of a diffusion process. *Statistics* **48**(6) 1344-1370.
- [12] A. Golightly, D.J. Wilkinson (2006) Bayesian sequential inference for nonlinear multivariate diffusions. *Stat. Comput.* **16** 323-338.
- [13] A. Golightly, D.J. Wilkinson (2011) Bayesian parameter inference for stochastic biochemical network models using particle markov chain monte carlo. *Interface Focus* **1**(6) 807-820.
- [14] C. Gourieroux, A. Monfort, E. Renault (1993) Indirect inference. *J. Appl. Econometrics* **8**(S1) S85-S118.
- [15] A.S. Hurn, J.I. Jeisman, K.A. Lindsay (2007) Transitional densities of diffusion processes: A new approach to solving the fokker-planck equation. *J. Derivatives* **14**(4) 86.
- [16] A. Hurn, J. Jeisman, K. Lindsay (2005) Ml estimation of the parameters of sde's by numerical solution of the Fokker–Planck equation. In *MODSIM 2005: international congress on modelling and simulation: advances and applications for management and decision making*, pages 849-855. Citeseer.
- [17] B. Jensen, R. Poulsen (2002) Transition densities of diffusion processes: numerical comparison of approximation techniques. *J. Derivatives* **9**(4) 18.
- [18] J.C. Jimenez, R.J. Biscay, T Ozaki (2005) Inference methods for discretely observed continuous-time stochastic volatility models: a commented overview. *Asia-Pacific Financial Markets* **12** 109-141.
- [19] P. Li, W. Gu (2017) Estimation of 1-dimensional nonlinear stochastic differential equations based on higher-order partial differential equation numerical scheme and its application. *Front. Math. China* **12**(6) 1441-1455.
- [20] J.N. Nielsen, H. Madsen (2001) Applying the EKF to stochastic differential equations with level effects. *Automatica* **37**(1) 107-112.
- [21] J.N. Nielsen, H. Madsen, P.C. Young (2000) Parameter estimation in stochastic differential equations: an overview. *Annu. Rev. Control* **24** 83-94.
- [22] L. Pichler, A. Masud, L.A. Bergman (2013) Numerical solution of the fokker–planck equation by finite difference and finite element methods—a comparative

- study. In: *Computational Methods in Stochastic Dynamics*, eds. S.T. Ariaratnam, S. Rajasekar, A.A. Kilbas; Springer, Berlin, Heidelberg, pp. 69-85.
- [23] Y. Shimizu (2009) Quadratic type contrast functions for discretely observed non-ergodic diffusion processes. Technical Report 09-04, Osaka University.
 - [24] K.J. Singleton (2001) Estimation of affine asset pricing models using the empirical characteristic function. *J. Econometrics* **102**(1) 111-141.
 - [25] H. Sørensen (2004) Parametric inference for diffusion processes observed at discrete points in time: a survey. *Int. Stat. Rev.* **72**(3) 337-354.
 - [26] S. Bodnarchuk, D. Ivanenko, A. Kohatsu-Higa, A. Kulik (2020) Improved local approximation for multidimensional diffusions: The g-rates. *Theory Probab. Math. Stat.* **101** 13-38.
 - [27] D. Ivanenko, R. Pogorielov (2021) Parameter estimation in diffusion models with low regularity coefficients. [arXiv:2103.06323](https://arxiv.org/abs/2103.06323) [math.ST].
 - [28] D.O. Ivanenko, R.V. Pogorielov (2023) QML parameter estimation in diffusion models. *AIP Conf. Proc.* **2787** 050026.
 - [29] K.-W. Ahn, K.-S. Chan (2014) Approximate conditional least squares estimation of a nonlinear state-space model via an unscented kalman filter. *Comput. Stat. Data Anal.* **69** 243-254.
 - [30] J.P.N. Bishwal (2007) *Parameter Estimation in Stochastic Differential Equations*, Part of the book series: Lecture Notes in Mathematics (LNM, volume 1923), Springer, Berlin, Heidelberg.
 - [31] R.A. Kasonga (1988) The consistency of a non-linear least squares estimator from diffusion processes. *Stoch. Process. Their Appl.* **30**(2) 263-275.
 - [32] A. Neuenkirch, S. Tindel (2014) A least square-type procedure for parameter estimation in stochastic differential equations with additive fractional noise. *Stat. Inference Stoch. Process.* **17**(1) 99-120.
 - [33] B.L.S. Prakasa Rao (1983) Asymptotic theory for non-linear least squares estimator for diffusion processes. *Statistics* **14**(2) 195-209.