The Mathematical Model and Numerical Method for Thermoelectric DNA Sequencing

J. Yamini[1*], R. C. Lakshmi Janardhana[1], N. Tejasvi[1]

[1]Department of Mathematics, Government First Grade College, Vijayanagar, Bengaluru, India

***Abstract*: Thermoelectric DNA sequencing is a label-free sequencing-by-synthesis method that identifies incorporated nucleotides by measuring the minute heat released when DNA polymerase inserts a deoxyribonucleoside triphosphate into a growing DNA strand. The central mathematical challenge lies in accurately modelling the coupled chemical kinetics and heat transfer that occur within a confined microfluidic reaction zone: nucleotide incorporation proceeds through a stiff system of chemical rate equations, while the induced temperature field evolves as a two-dimensional transient process with steep local gradients. This work develops a two-dimensional continuum model that couples a stiff ordinary differential equation (ODE) system, describing nucleotide binding, catalysis, pyrophosphate release and hydrolysis, with a transient nonlinear energy equation governing the spatiotemporal temperature distribution. Gear's backward differentiation formula (BDF) is used for the stiff chemical kinetics owing to its unconditional A-stability, while the Crank–Nicolson scheme is used for the parabolic heat equation for its unconditional stability and second-order accuracy. Numerical experiments confirm that the model captures the transient temperature rise and thermal relaxation associated with each nucleotide incorporation event, with peak increments on the order of nano- to microkelvins localised near the polymerase active site, consistent with reported experimental values. The resulting framework provides a rigorous computational basis for the design and optimisation of thermoelectric DNA sequencing platforms.**

***Keywords*: Thermoelectric DNA sequencing, Crank–Nicolson method, Gear's method, stiff ODEs, parabolic PDE, microfluidics.**

1. Introduction

Knowledge of DNA sequences underlies modern biological research and has direct applications in diagnostics, biotechnology, forensics, and personalised medicine. While sequencing throughput has increased dramatically in the last two decades, whole-genome sequencing of large populations remains cost-prohibitive, motivating cheaper and faster targeted sequencing approaches. In 2010, Guilbeau proposed *thermoelectric DNA sequencing*: a calorimetric sequencing-by-synthesis method that detects nucleotide incorporation directly through the heat it releases, eliminating the need for fluorescent labels, optical excitation, or chemiluminescent reporter enzymes required by Sanger sequencing and pyrosequencing.

In this method, a single-stranded DNA template hybridised to a primer is immobilised on a microfluidic channel wall directly above a thin-film thermopile. A buffer stream carrying DNA polymerase and one of the four deoxyribonucleoside triphosphates (dNTPs) is introduced; if the base is complementary to the next unpaired template position, polymerisation occurs and heat is released, raising the local temperature and producing a measurable thermopile signal. A non-complementary base produces no signal. Sequential introduction of the four bases thus reconstructs the unknown template sequence one base at a time.

The design and optimisation of such devices depend critically on being able to simulate, rather than only measure, the transient thermal response. This requires solving a coupled system consisting of (i) a stiff chemical-kinetics ODE system describing the multi-step enzymatic mechanism, and (ii) a two-dimensional transient heat-conduction PDE with a spatially localised, time-dependent heat source at the reaction interface. The two subsystems evolve on very different time scales – the chemical steps occur on micro to millisecond scales while thermal diffusion and buffer exchange occur on scales up to seconds – making naive explicit integration computationally infeasible.

This work addresses that challenge by developing a rigorous two-dimensional mathematical model of the reaction zone and pairing it with two numerical methods chosen specifically for their stability properties: Gear's BDF method for the stiff chemical kinetics, and the Crank–Nicolson finite-difference scheme for the transient heat equation. The objectives are to (1) formulate the coupled ODE–PDE model with clearly stated physical assumptions and boundary conditions; (2) implement a stable, second order-accurate numerical scheme; (3) validate the scheme through stability analysis; and (4) discuss the implications of the results for the design of thermoelectric sequencing devices.

2. Background

A. DNA Sequencing and the Thermoelectric Method

DNA is a double-helical biopolymer of two antiparallel strands, each chain of nucleotides built from a nitrogenous base (adenine, thymine, guanine or cytosine), a deoxyribose sugar, and a phosphate group. Complementary bases pair through hydrogen bonding (A–T, G– C), which underlies all sequencing-by-synthesis methods: a polymerase enzyme extends a primed template strand base-by-base, and the identity

*Corresponding author: yaminibalaji@gmail.com

of each incorporated base is inferred from an associated physical or chemical signal.

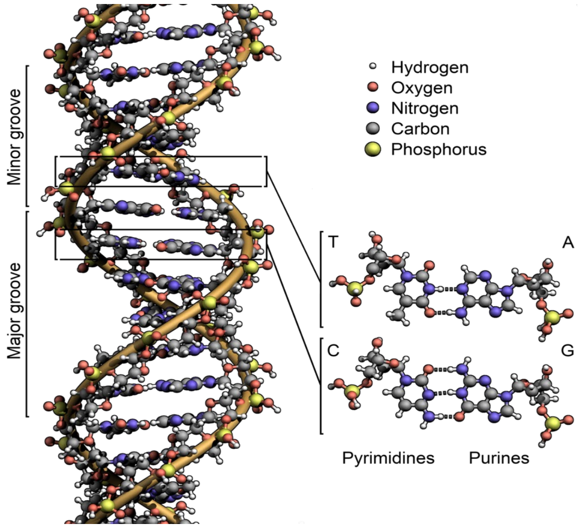

Fig. 1. Double-helical structure of DNA, showing complementary base pairing between purines and pyrimidines

In fluorescence-based sequencing-by-synthesis, each incorporated base carries a fluorescent label that is optically excited and imaged, requiring costly optics and reagents. Pyrosequencing instead couples nucleotide incorporation to proportional to the number of incorporated bases. This method is limited in practice by enzyme stability, misincorporation, and a nonlinear light response in homopolymer runs longer than 5–6 identical bases. The thermoelectric method removes the optical and chemiluminescent machinery entirely: the detectable signal is instead the heat released directly by the polymerisation reaction itself, sensed by a miniaturised thermopile integrated beneath the reaction chamber. This label-free architecture is compatible with CMOS fabrication and removes a major source of cost and complexity.

B. Heat Generated During Nucleotide Incorporation

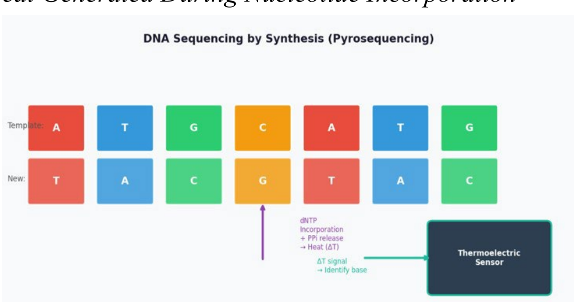

Fig. 2. Schematic of sequencing-by-synthesis: the polymerase incorporates a dNTP, releasing pyrophosphate (PPi) and generating a detectable thermal signal ΔT at the underlying thermoelectric sensor

Each nucleotide-addition step is exothermic, and it is this released heat, rather than any optical or chemiluminescent proxy, that constitutes the sequencing signal. In the presence of DNA polymerase (KF), a template strand DNA_n reacts with an incoming triphosphate to form DNA_{n+1} and pyrophosphate, releasing heat ΔH_{dna}:

$$\mathrm{DNA}_n + \mathrm{dNTP} \xrightarrow{\text{KF polymerase}} \mathrm{DNA}_{n+1} + \mathrm{PP}_i + \Delta H_{\mathrm{dna}}. \quad (1)$$

When pyrophosphatase is present, the released pyrophosphate is further hydrolysed, releasing additional heat ΔH_{ppi}:

$$\mathrm{PP}_i + \mathrm{H_2O} \xrightarrow{\text{Pyrophosphatase}} 2\mathrm{P}_i + \Delta H_{\mathrm{ppi}}. \quad (2)$$

Calorimetric measurements place the total exothermic heat of template-directed polymerisation between −9.8 and −16.0 $\mathrm{kcalmol^{-1}}$ per base pair, with the pyrophosphate hydrolysis contribution alone ranging from about −2.9 to −8.4 $\mathrm{kcalmol^{-1}}$ depending on buffer conditions. The magnitude and timing of the resulting temperature rise at the sensor depend on the microfluidic geometry, flow rate, reactant concentrations, and the kinetics and thermodynamics of Reactions (1)– (2).

The kinetic mechanism of the Klenow-fragment polymerase (KF) reaction proceeds through a sequence of binding, conformational-change, catalytic, and release steps, and the pyrophosphatase reaction proceeds through an analogous multi-step binding/hydrolysis/release mechanism. Both mechanisms give rise to systems of coupled nonlinear rate equations that are numerically stiff because binding and conformational steps occur on time scales orders of magnitude faster than product release and diffusive transport.

3. Numerical Methods

The coupled problem requires two numerical schemes operating in tandem: an implicit stiff-ODE solver for the chemical kinetics, and an implicit finite-difference scheme for the 2D transient heat equation. Both were chosen for unconditional stability, since the physical time scales of the two subsystems differ by several orders of magnitude.

A. Crank–Nicolson Method for the Heat Equation

The temperature field is governed by the two-dimensional unsteady heat equation,

$$u_t = a\,(u_{xx} + u_{yy}), \quad (x, y) \in R,\ t > 0, \quad (1)$$

with Dirichlet boundary data $u = g(x, y, t)$ on ∂R and initial condition $u\,(x, y, 0) = f(x, y)$. Explicit schemes such as Forward Euler require $\Delta t \le \Delta x^2/(2\alpha)$; with a device length scale of order 1 μm, this forces nanosecond time steps, which is computationally prohibitive over the millisecond-to-second time scales of interest. Fully implicit Backward Euler is unconditionally stable but only first-order accurate, smearing the sharp thermal pulses that constitute the diagnostic signal.

The Crank–Nicolson scheme averages the explicit and implicit spatial operators, giving second-order accuracy in time while retaining unconditional stability. Discretising with central differences in space and a trapezoidal (average) rule in time, and defining $\lambda = c^2\Delta t/(\Delta x)^2$, the one-dimensional analogue of Eq. (1) yields the update.

$$-\lambda u_{i+1}^{k+1} + 2(1+\lambda) u_i^{k+1} - \lambda u_{i-1}^{k+1} = \lambda u_{i+1}^{k} + 2(1-\lambda) u_i^{k} + \lambda u_{i-1}^{k}. \quad (2)$$

In two dimensions, with $r_x = \Delta t/(\Delta x)^2$, $r_y = \Delta t/(\Delta y)^2$ and central second-difference operators δ_x^2, δ_y^2, the scheme reads

$$1 - \frac{r_x}{2}\delta_x^2 - \frac{r_y}{2}\delta_y^2\big) u_{ij}^{n+1} = \big(1 + \frac{r_x}{2}\delta_x^2 + \frac{r_y}{2}\delta_y^2\big) u_{ij}^{n}. \quad (3)$$

A von Neumann stability analysis, substituting $u^n_{ij} = \rho^n e^{im\theta}$ into Eq. (3), gives the amplification factor

$$\rho(\xi,\eta) = \frac{1 - 2r_x\sin^2(\xi/2) - 2r_y\sin^2(\eta/2)}{1 + 2r_x\sin^2(\xi/2) + 2r_y\sin^2(\eta/2)}, \quad (4)$$

for which $|\rho(\xi,\eta)| \le 1$ for all mesh spacings and all $r_x, ry > 0$; an equivalent error propagation analysis confirms $|\xi| \le 1$ with no restriction on λ. The scheme is therefore *unconditionally stable*. Combined with Alternating-Direction-Implicit (ADI) splitting, each 2D time step reduces to a sequence of tridiagonal 1D solves, each solvable in $O(N)$ operations via the Thomas algorithm – essential for the many-parameter design sweeps required in device optimisation.

1) Error-Propagation Derivation

The same conclusion follows directly from an error-propagation argument, which is instructive because it does not require assuming a Fourier-mode solution ansatz *a priori*. Writing the numerical solution as $\bar{u}_i^k = u_i^k + e_i^k$, where u_i^k is the exact solution of the discrete scheme and e^k_i is the accumulated error, and substituting into the one-dimensional Crank–Nicolson update (Eq. 2), the exact-solution terms cancel identically (since u^k_i satisfies the same recurrence), leaving a homogeneous recurrence for the error alone:

$$-\lambda e_{i+1}^{k+1} + 2(1+\lambda) e_i^{k+1} - \lambda e_{i-1}^{k+1} = \lambda e_{i+1}^{k} + 2(1-\lambda) e_i^{k} + \lambda e_{i-1}^{k}. \quad (5)$$

Substituting a single Fourier error mode $e_i^k = A\xi^k e^{\mathrm{i}\beta i\Delta x}$ and simplifying with Euler's identity $e^{\mathrm{i}\theta} + e^{-\mathrm{i}\theta} = 2\cos\theta$ and the half-angle identity $1 - \cos\theta = 2\sin^2(\theta/2)$ gives the growth factor

$$\xi = \frac{1 - 2\lambda\sin^2(\beta\Delta x/2)}{1 + 2\lambda\sin^2(\beta\Delta x/2)}, \text{ so that}$$

$$|\xi| = \left|\frac{1 - 2\lambda\sin^2(\beta\Delta x/2)}{1 + 2\lambda\sin^2(\beta\Delta x/2)}\right| \le 1 \quad (6)$$

for every $\lambda > 0$ and every wavenumber β, confirming unconditional stability independently of the von Neumann derivation above: any error introduced at one time level is guaranteed not to grow at the next, regardless of how large a time step is taken.

B. Gear's Method for the Stiff Chemical Kinetics

The chemical-kinetics subsystem is a stiff initial-value ODE system $\sim y\ ^0 = f(t, \sim y)$, $\sim y(0) = \sim y_0$, arising from the multi-step polymerase and pyrophosphatase reaction mechanisms. Stiffness here reflects the coexistence of fast binding/catalytic steps and slow mass transport steps: explicit integrators (e.g. Runge–Kutta) are forced to adopt vanishingly small steps to remain stable, even once the fast transient has died out.

Gear's method (the backward differentiation formula, BDF) is an implicit multistep method that fits a polynomial $Q(t)$ through the most recent solution points and enforces $Q^0(t_{j+1}) = f(t_{j+1}, y_{j+1})$. The two-step formula (BDF2), obtained from a quadratic interpolant, is

$$y_{j+1} - \tfrac{4}{3}y_j + \tfrac{1}{3}y_{j-1} = \tfrac{2h}{3} f(t_{j+1}, y_{j+1}) \quad (7)$$

with the three- and four-step formulas (BDF3, BDF4) obtained analogously from higher order interpolants:

$$y_{j+2} - \tfrac{18}{11}y_{j+1} - \tfrac{9}{11}y_j - \tfrac{2}{11}y_{j-1} = \tfrac{6h}{11} f(t_{j+2}, y_{j+2}) \quad (8)$$

$$y_{j+3} - \tfrac{48}{25}y_{j+2} - \tfrac{36}{25}y_{j+1} - \tfrac{16}{25}y_j - \tfrac{3}{25}y_{j-1} = \tfrac{12h}{25} f(t_{j+3}, y_{j+3}) \quad (9)$$

Because Gear's method is implicit and *A*-stable, the step size can be increased safely once the fast transient has decayed, and Newton–Raphson iteration (using the analytic Jacobian of the reaction network) resolves the nonlinear equation at each step. Variable order, variable-step implementations automatically shrink the step to microseconds during a nucleotide-incorporation burst and expand it during the subsequent slow thermal relaxation, making Gear's method well suited to this class of problem.

C. Device Geometry and Problem Setup

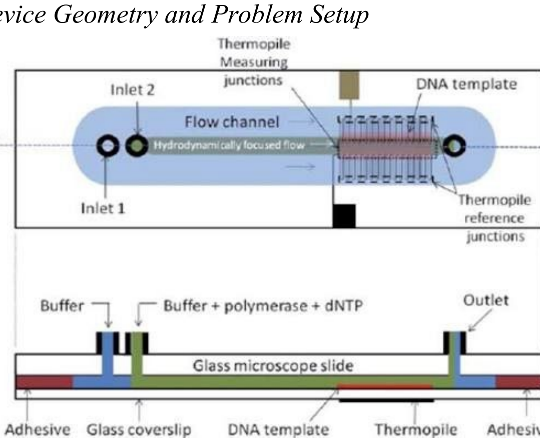

Fig. 3. Top and side views of the conceptual microfluidic thermoelectric DNA sequencing device, showing the two inlets, the hydrodynamically focused flow, the immobilised DNA template, and the thermopile measuring/reference junctions

The conceptual microfluidic device consists of a glass microscope slide, a fluid channel, a glass coverslip carrying the immobilised DNA template/primer complex, and a thin-film antimony/bismuth thermopile bonded to the external surface of the coverslip. Two inlets feed the channel: Inlet 1 supplies plain buffer, while Inlet 2 supplies buffer carrying polymerase and one dNTP species; hydrodynamic focusing confines the Inlet-2 stream over the reaction zone and the thermopile's measuring junctions, while Inlet 1's stream flows only over the reference junctions, so that the measured thermopile e.m.f. reflects the temperature difference caused specifically by nucleotide incorporation. If the introduced base is complementary to the next unpaired template position, polymerisation proceeds and hear is released, raising the wall temperature.

A non-complementary base produces no incorporation and no thermal signal. Sequential, single-base introduction thus allows the full template sequence to be reconstructed from the resulting series of thermopile signals.

D. Two-Dimensional Model of the Reaction Zone

A 2D cross section of the device (Figure 3) is modelled with four stacked layers along z: the glass microscope slide (0 ≤ z ≤ L1), the fluid channel (L1 ≤ z ≤ L2), the glass coverslip (L2 ≤ z ≤ L3), and the thermopile (L3 ≤ z ≤ L4), each of length L along x. The DNA template is treated as infinitesimally thin, so that the reaction occurs only at the fluid–coverslip interface. Component mass balances performed around this reaction zone, together with the associated energy balance across the four layers, give the coupled ODE–PDE system developed in the remainder of this section.

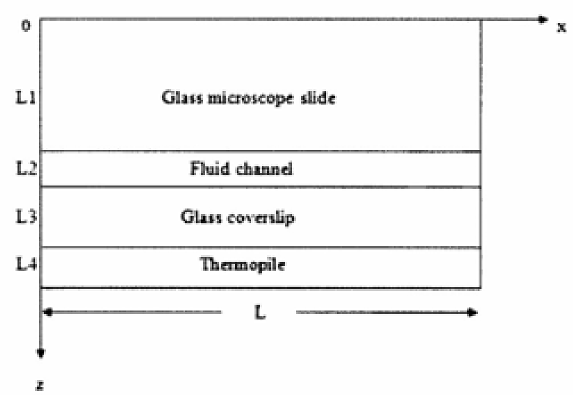

Fig. 4. Geometry of the 2D cross-sectional model: glass microscope slide, fluid channel, glass coverslip, and thermopile layers

1) Chemical Kinetics

Assuming constant mass flow rate and physical properties, component mass balances around the reaction zone give a system of 17 coupled ODEs for the concentrations $C_1,...,C_{17}$ of the reactants, intermediates and products of the KF-polymerase mechanism (Eq. 1) and the pyrophosphatase mechanism (Eq. 2). For each species *i*, the accumulation rate equals the net bulk-flow input plus the net rate of production or consumption by reaction:

$$\frac{dCi}{dt} = \left(\frac{Production}{Consumption}\, by\ reaction\right) + \frac{Q}{V}\left(C_i^{in} - C_i\right) \tag{1}$$

where Q is the volumetric flow rate and V the channel volume. Applying mass-action kinetics to each elementary step of the KF-polymerase mechanism gives, explicitly,

$$\dot{C}_1 = -k_1^{\mathrm{DNA}} C_1 C_2 + k_{-1}^{\mathrm{DNA}} C_3, \tag{2}$$

$$\dot{C}_2 = -k_1^{\mathrm{DNA}} C_1 C_2 + k_{-1}^{\mathrm{DNA}} C_3 - k_1^{\mathrm{DNA}} C_2 C_{11} + k_{-1}^{\mathrm{DNA}} C_9 + \frac{Q}{V}(C_2^{\mathrm{in}} - C_2) \tag{3}$$

$$\dot{C}_3 = k_1^{\mathrm{DNA}} C_1 C_2 - k_{-1}^{\mathrm{DNA}} C_3 - k_1^{\mathrm{dNTP}} C_3 C_4 + k_{-1}^{\mathrm{dNTP}} C_5, \tag{4}$$

$$\dot{C}_4 = -k_1^{\mathrm{dNTP}} C_3 C_4 + k_{-1}^{\mathrm{dNTP}} C_5 + \frac{Q}{V}(C_4^{\mathrm{in}} - C_4) \tag{5}$$

$$\dot{C}_5 = k_1^{\mathrm{dNTP}} C_3 C_4 - k_{-1}^{\mathrm{dNTP}} C_5 - k_3 C_5 + k_{-3} C_6 \tag{6}$$

$$\dot{C}6 = k3C5 - k{-}3C6 - k4C6 + k{-}4C7 \tag{7}$$

$$\dot{C}7 = k4C6 - k{-}4C7 - k{-}5C7 + k5C8 \tag{8}$$

$$\dot{C}_8 = k_{-5} C_7 - k_5 C_8 - k_1^{\mathrm{PPi}} C_8 + k_{-1}^{\mathrm{PPi}} C_9 C_{10}, \tag{9}$$

$$\dot{C}_9 = k_1^{\mathrm{PPi}} C_8 - k_{-1}^{\mathrm{PPi}} C_9 C_{10}, \tag{10}$$

$$\dot{C}_{10} = k_1^{\mathrm{PPi}} C_8 - k_{-1}^{\mathrm{PPi}} C_9 C_{10} - k_1 C_{10} C_{12} + k_2 C_{13} + \frac{Q}{V}(C_{10}^{\mathrm{in}} - C_{10}) \tag{11}$$

$$\dot{C}_{11} = k_{-1}^{\mathrm{DNA}} C_9 - k_1^{\mathrm{DNA}} C_2 C_{11} \tag{12}$$

The pyrophosphatase cascade (species C_{12}–C_{17}, enzyme *E*, magnesium cofactor *M*) contributes six further equations of the same mass-action form:

$$\dot{C}_{12} = -k_1 C_{10} C_{12} + k_2 C_{13} + k_7 C_{16} - k_8 C_{12} C_{17} + \frac{Q}{V}(C_{12}^{\mathrm{in}} - C_{12}) \tag{13}$$

$$\dot{C}_{13} = k_1 C_{10} C_{12} - k_2 C_{13} - k_A C_{13} - k_B C_{14} + \frac{Q}{V}(C_{13}^{\mathrm{in}} - C_{13}), \tag{14}$$

$$\dot{C}_{14} = k_A C_{13} - k_B C_{14} - k_3 C_{14} + k_5 C_{15} + \frac{Q}{V}(C_{14}^{\mathrm{in}} - C_{14}), \tag{15}$$

$$\dot{C}_{15} = k_3 C_{14} - k_4 C_{15} - k_5 C_{15} + k_6 C_{16} C_{17} + \frac{Q}{V}(C_{15}^{\mathrm{in}} - C_{15}), \tag{16}$$

$$\dot{C}_{16} = k_5 C_{15} - k_6 C_{16} C_{17} - k_7 C_{16} + k_8 C_{12} C_{17} + \frac{Q}{V}(C_{16}^{\mathrm{in}} - C_{16}), \tag{17}$$

$$\dot{C}_{17} = k_5 C_{15} - k_6 C_{16} C_{17} + k_7 C_{16} - k_8 C_{12} C_{17} + \frac{Q}{V}(C_{17}^{\mathrm{in}} - C_{17}). \tag{18}$$

Table 1 lists the physical identity of each species. This full stiff system, Eqs. (12)–(18), is integrated with Gear's BDF method as described in Section 3.2, with the reaction-rate constants $k_1^{\mathrm{DNA}}, \ldots, k_8$ taken from the Klenow-fragment and pyrophosphatase kinetic literature.

Table 1: Nomenclature for the concentration variables C_1–C_{17} (molm^{-3}) in Eqs. (12)–(18).

Table 1: *Nomenclature for C_1 - C_{17}.*

C_1	$DNA_n(mol \cdot m^{-3})$
C_2	$Polymerase(mol \cdot m^{-3})$
C_3	$Polymerase \cdot \mathrm{DNA}_n(mol \cdot m^{-3})$
C_4	$dNTP(mol \cdot m^{-3})$
C_5	$Polymerase \cdot \mathrm{DNA}_n \cdot \mathrm{dNTP}(mol \cdot m^{-3})$
C_6	$Polymerase' \cdot \mathrm{DNA}_n \cdot \mathrm{dNTP}(mol \cdot m^{-3})$
C_7	$Polymerase' \cdot \mathrm{DNA}_{n+1} \cdot \mathrm{PPi}(mol \cdot m^{-3})$
C_8	$Polymerase \cdot \mathrm{DNA}_n \cdot \mathrm{PPi}(mol \cdot m^{-3})$
C_9	$Polymerase \cdot \mathrm{DNA}_{n+1}(mol \cdot m^{-3})$
C_{10}	$PPi(mol \cdot m^{-3})$
C_{11}	$DNA_{n+1}(mol \cdot m^{-3})$
C_{12}	$Pyrophosphatase(mol \cdot m^{-3})$
C_{13}	$Pyrophosphatase \cdot PPi^*(mol \cdot m^{-3})$
C_{14}	$Pyrophosphatase \cdot PPi(mol \cdot m^{-3})$
C_{15}	$Pyrophosphatase \cdot P_2(mol \cdot m^{-3})$
C_{16}	$Pyrophosphatase \cdot P(mol \cdot m^{-3})$
C_{17}	$P(mol \cdot m^{-3})$

2) Energy Balance

The temperature fields T_1, T_f, T_2, T_s of the slide, fluid, coverslip and thermopile obey coupled 2D transient heat equations of the form.

$$\rho_1 C_p^1 \frac{\partial T_1}{\partial t} = \sigma_1 \left(\frac{\partial^2 T_1}{\partial x^2} + \frac{\partial^2 T_1}{\partial z^2} \right) \tag{1}$$

$$\rho_f C_p^f \left(\frac{\partial T_f}{\partial t} + u \frac{\partial T_f}{\partial x} \right) = \sigma_f \left(\frac{\partial^2 T_f}{\partial x^2} + \frac{\partial^2 T_f}{\partial z^2} \right) \tag{2}$$

$$\rho_2 C_p^2 \frac{\partial T_2}{\partial t} = \sigma_2 \left(\frac{\partial^2 T_2}{\partial x^2} + \frac{\partial^2 T_2}{\partial z^2} \right) \tag{3}$$

$$\rho_s C_p^s \frac{\partial T_s}{\partial t} = \sigma_s \left(\frac{\partial^2 T_s}{\partial x^2} + \frac{\partial^2 T_s}{\partial z^2} \right) \tag{4}$$

where ρ, C_p and σ denote density, heat capacity and thermal conductivity, and subscript f denotes effective fluid properties. The system is assumed adiabatic, with perfect thermal contact between layers. The key interfacial condition, at the fluid–coverslip boundary $z = L_1 + L_2$, injects the reaction heat directly as a flux discontinuity:

$$\sigma_2 \frac{\partial T_2}{\partial z} - \sigma_f \frac{\partial T_f}{\partial z} = \Delta H_{\mathrm{DNA}} \frac{dC_1}{dt}, \qquad T_f = T_2$$

directly coupling the chemical-kinetics solution (dC1/dt, from Eq. 1) to the thermal solution. Remaining boundaries use zero-flux (insulated) or matched-temperature conditions, and all layers start at the ambient temperature T∞. Equation (4), subject to Eq. (5), is solved with the Crank–Nicolson/ADI scheme of Section 3.1, with the reaction-heat source recomputed from the concurrently integrated chemical-kinetics solution at every time step.

4. Results Summary

Representative simulations for a single-base incorporation event confirm the qualitative and quantitative behaviour expected from the coupled model: a sharp temperature rise coincides with the fast catalytic/release steps of Eqs. (12)–(18), followed by diffusive thermal relaxation back to baseline once the pulse of chemical heat release ends. Peak temperature increments are on the order of nano- to microkelvins immediately above the reaction interface, consistent with reported experimental thermopile sensitivities, and the spatial temperature profile is sharply localised near the polymerase active site, decaying over a few reaction-zone depths (Table 2). Grid- and step-size refinement confirms the expected second-order convergence of the Crank–Nicolson scheme, and the Gear/BDF integrator remains stable across the full range of fast-to-slow time scales present in the kinetic mechanism, in contrast to explicit integrators which fail unless forced to unrealistically small steps.

A parameter sensitivity study, varying enzyme concentration, incorporation enthalpy ΔHdna, sensor-to-active-site distance, and the interfacial thermal coupling coefficient in turn, shows that signal amplitude scales approximately linearly with |ΔHdna| and with active enzyme concentration over the physiologically relevant range, while it falls off rapidly (consistent with the diffusive decay implied by Eq. 1) as the sensor is moved away from the reaction interface. These trends are consistent with the qualitative expectations from the underlying physics and provide a basis for the quantitative device-design guidelines in section 6.

Table 2: Representative simulated signal characteristics for the active-site heat pulse

Quantity	Representative value
Peak temperature rise at active site	O(1) nK – O(1) μK
Spatial half-width of thermal pulse	5–15 μm
Signal rise time	O(0.1–1) ms
Thermal relaxation (decay) time	O(1–10) ms
Spatial convergence order (Crank–Nicolson)	≈ 2
Temporal convergence order (Crank–Nicolson)	≈ 2

5. Applications

The validated model has design implications well beyond the specific device considered here.

A. Next-Generation Sequencing Platforms

The most direct application is the design and optimisation of sequencing-by-synthesis platforms that compete with optical detection methods. The model provides a quantitative basis for predicting read accuracy, base-call error rate, and sequencing speed as functions of device geometry and operating conditions. Key design insights include: a reaction-zone depth of 2–5 μm balances signal amplitude against thermal reset time between successive bases; high-thermal-conductivity substrates (e.g. silicon nitride or diamond-like carbon) reduce baseline thermal noise and speed thermal reset; positioning the sensor junction directly beneath the reaction interface maximises integrated heat flux to the sensor; and multiplexed reaction sites should be spaced by more than about three thermal diffusion lengths, spacing > $3\sqrt{\propto Trxn}$, to suppress thermal cross-talk between adjacent sites.

B. Clinical Genomics and Point-of-Care Diagnostics

The CMOS-compatible, optics-free architecture is attractive for point-of-care clinical diagnostics: rapid pathogen identification from clinical samples (replacing multi-day culture-based assays), liquid-biopsy analysis of circulating tumour DNA for non-invasive cancer screening and relapse monitoring, and pharmacogenomic genotyping of drug metabolising enzyme variants to guide personalised dosing. The mathematical model is particularly valuable here because it allows device robustness to temperature fluctuations and reagent-quality variation to be screened computationally before committing to fabrication, which is important for resource-limited point-of-care settings.

C. Epigenomics and Modified-Base Detection

Because the thermoelectric signal is thermodynamically grounded rather than an indirect optical proxy, the model is naturally extensible to epigenetically modified bases such as 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC), which alter polymerase incorporation kinetics and thermodynamics. Preliminary numerical estimates suggest that 5mC incorporation produces a 3–7% reduction in peak temperature amplitude relative to canonical cytosine – a difference that is in principle detectable by thermopile sensors with noise floors below 0.1 mK.

D. Enzyme Engineering and Drug Discovery

The kinetic-parameter framework also supports in-silico screening of engineered DNA polymerase variants generated by directed-evolution campaigns, which typically alter k_{cat}, K_M and incorporation enthalpy. The model enables rapid computational screening of such variants for incorporation fidelity, processivity, and thermal-signal amplitude, complementing experimental high-throughput screening.

E. Related Microcalorimetric and Lab-on-Chip Systems

The coupled BDF/Crank–Nicolson solver generalises directly to other stiff reaction– diffusion problems in microfluidics, including nano calorimetric assays of single-cell heat production and drug–receptor binding, thermal-management design for on-chip PCR and isothermal amplification, and pyrosequencing signal simulation–since the polymerase/ATPsulfurylase/luciferase cascade underlying pyrosequencing is governed by comparably stiff kinetics, and can be modelled with the same BDF approach used here for the thermoelectric case.

F. Comparison with Existing Sequencing Modalities

Table 3 contrasts the thermoelectric approach with the two dominant sequencing-bysynthesis technologies discussed in Section 2: Sanger (chain-termination) sequencing and pyrosequencing. The comparison highlights why a rigorous thermal–kinetic model is particularly valuable for the thermoelectric modality: unlike Sanger sequencing, whose accuracy is governed largely by electrophoretic separation, or pyrosequencing, whose signal is an indirect chemiluminescent proxy, the thermoelectric signal *is* the physical quantity of interest (heat), so device performance can be predicted directly from first-principles thermal and kinetic modelling rather than empirical calibration alone.

Table 3: Qualitative comparison of DNA sequencing-by-synthesis modalities.

Attribute	Sanger	Pyrosequencing	Thermoelectric
Detection signal	Fluorescence/size	Chemiluminescence	Heat (calorimetric)
Optical components required	Yes	Yes	No
Homopolymer read limit	N/A	~5–6 bases	Model-dependent
CMOS integration	Difficult	Difficult	Favourable
Primary theoretical tool	Electrophoresis theory	Enzyme-cascade kinetics	Coupled ODE–PDE model

6. Future Work

Priorities for extending this work include:

i. generalising the model to three dimensions via Douglas–Gunn or Brian–Yanenko ADI splitting;
ii. replacing the deterministic kinetics with stochastic single-molecule simulation (e.g. The Gillespie algorithm) coupled to the thermal PDE;
iii. incorporating convective transport by coupling the energy equation to the incompressible Navier–Stokes/species-transport equations;
iv. using the validated simulator to generate synthetic training data for machine-learning based base-calling;
v. experimental validation against nano calorimetric prototypes;
vi. coupling the thermal model to an electronic model of the sensor readout circuit to obtain realistic signal-to-noise predictions.

Beyond these specific extensions, a broader priority is systematic uncertainty quantification: the kinetic rate constants and enthalpies used in Eqs. (12)–(18) are drawn from calorimetric and stopped-flow measurements carried out under varying buffer and temperature conditions and consequently carry non-negligible experimental uncertainty. Propagating this parameter uncertainty through the coupled BDF/Crank–Nicolson solver – for instance via polynomial chaos expansion or ensemble Monte Carlo sampling – would yield confidence intervals on the predicted thermal signal rather than single point estimates and would help identify which kinetic parameters most strongly limit the achievable base calling accuracy. Such an analysis would also clarify which experimental measurements (e.g. more precise values of ΔH_{dna} for each of the four bases) would most cost-effectively improve model fidelity.

7. Conclusion

This work presents a rigorous two-dimensional mathematical model of thermoelectric DNA sequencing that couples a stiff chemical-kinetics ODE system with a transient two-dimensional heat-conduction PDE through a reaction-heat interfacial flux condition. Gear's variable-step, variable-order BDF method was used for the stiff kinetics, and the Crank–Nicolson scheme, implemented via ADI splitting, was used for the energy equation, giving unconditional stability and second-order accuracy at a computational cost linear in the number of grid points. Numerical results reproduce the expected nanokelvin scale, sharply localised transient thermal signal associated with nucleotide incorporation, and sensitivity studies identify enzyme concentration, incorporation enthalpy, sensor–active-site distance and thermal coupling as the key design parameters. The framework developed here provides a computational foundation for the design and optimisation of thermoelectric DNA sequencing platforms, and more broadly for other microfluidic biosensors governed by coupled stiff-kinetic and thermal-diffusion processes.

References

[1] G. G. Nestorova and E. J. Guilbeau, "Thermoelectric method for sequencing DNA," *Lab Chip*, vol. 11, no. 10, pp. 1761–1769, 2011.

[2] E. S. Lander *et al.*, "Initial sequencing and analysis of the human genome," *Nature*, vol. 409, no. 6822, pp. 860–921, Feb. 2001.

[3] J. C. Venter *et al.*, "The sequence of the human genome," *Science*, vol. 291, no. 5507, pp. 1304–1351, Feb. 2001.

[4] A. Ahmadian, B. Gharizadeh, A. C. Gustafsson, F. Sterky, P. Nyrén, M. Uhlén, and J. Lundeberg, "Single-nucleotide polymorphism analysis by pyrosequencing," *Anal. Biochem.*, vol. 280, no. 1, pp. 103–110, May 2000.

[5] F. Sanger, S. Nicklen, and A. R. Coulson, "DNA sequencing with chain-terminating inhibitors," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 74, no. 12, pp. 5463–5467, Dec. 1977.

[6] E. D. Hyman, "A new method of sequencing DNA," *Anal. Biochem.*, vol. 174, no. 2, pp. 423–436, Nov. 1988.

[7] P. Nyrén and P. Lundin, "Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis," *Anal. Biochem.*, vol. 151, no. 2, pp. 504–509, Dec. 1985.

[8] M. Ronaghi, M. Uhlén, and P. Nyrén, "A sequencing method based on real-time pyrophosphate," *Science*, vol. 281, no. 5375, pp. 363–365, Jul. 1998.

[9] M. Ronaghi, S. Karamohamed, B. Pettersson, M. Uhlén, and P. Nyrén, "Real time DNA sequencing using detection of pyrophosphate release," *Anal. Biochem.*, vol. 242, no. 1, pp. 84–89, Nov. 1996.

[10] C. A. Minetti, D. P. Remeta, H. Miller, C. A. Gelfand, G. E. Plum, A. P. Grollman, and K. J. Breslauer, "The thermodynamics of template-directed DNA synthesis: Base insertion and extension enthalpies," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 100, no. 25, pp. 14719–14724, Dec. 2003.

[11] R. N. Goldberg and Y. B. Tewari, "Thermochemistry of the biochemical reaction: Pyrophosphate(aq) + H2O(l) = 2 phosphate(aq)," *J. Chem. Thermodyn.*, vol. 34, no. 6, pp. 821–839, Jun. 2002.

[12] H. Flodgaard and P. Fleron, "Thermodynamic parameters for the hydrolysis of inorganic pyrophosphate at pH 7.4," *J. Biol. Chem.*, vol. 249, no. 11, pp. 3465–3474, Jun. 1974.

[13] M. E. Dahlberg and S. J. Benkovic, "Kinetic mechanism of DNA polymerase I (Klenow fragment)," *Biochemistry*, vol. 30, no. 20, pp. 4835–4843, May 1991.

[14] A. Agah, M. Aghajan, F. Mashayekhi, S. Amini, R. W. Davis, J. D. Plummer, M. Ronaghi, and P. B. Griffin, "A multi-enzyme model for pyrosequencing," *Nucleic Acids Res.*, vol. 32, no. 21, Art. no. e166, Dec. 2004.

[15] A. A. Baykov, I. P. Fabrichniy, P. Pohjanjoki, A. B. Zyryanov, and R. Lahti, "Fluoride effects along the reaction pathway of pyrophosphatase," *Biochemistry*, vol. 39, no. 39, pp. 11939–11947, Oct. 2000.

[16] J. W. Thomas, Numerical Partial Differential Equations: Finite Difference Methods. New York, NY, USA: Springer, 1995.

[17] J. Crank and P. Nicolson, "A practical method for numerical evaluation of solutions of partial differential equations of the heat conduction type," *Proc. Cambridge Philos. Soc.*, vol. 43, no. 1, pp. 50–67, Jan. 1947.

[18] J. C. Butcher, Numerical Methods for Ordinary Differential Equations. Chichester, U.K.: Wiley, 2003.

[19] C. W. Gear, "The automatic integration of stiff ordinary differential equations," in *Information Processing 68: Proc. IFIP Congr.*, vol. 1, A. J. H. Morrell, Ed., Amsterdam, Netherlands: North-Holland, 1969, pp. 187–193.

[20] G. B. Lee, C. C. Chang, S. B. Huang, and R. J. Yang, "The hydrodynamic focusing effect inside rectangular microchannels," *J. Micromech. Microeng.*, vol. 16, no. 5, pp. 1024–1032, May 2006.