SolarTrack: Exploring the Continuous Tracking Capabilities of Wearable Solar Harvesters

Yasien Ghalwash* , Abdelwahed Khamis†*, Moid Sandhu†*, Sara Khalifa*†, and Raja Jurdak*
* Queensland University of Technology, Brisbane, Australia
†Commonwealth Scientific and Industrial Research Organisation (CSIRO), Brisbane, Australia

***Abstract*—Continuous tracking is often thought to require specialised, actively powered sensors. Yet energy harvesters already embedded in commercial devices, such as Garmin solar-powered smartwatches, generate energy signals that inherently carry continuous variations linked to user motion and environment. Prior studies have shown that these signals are sufficient for classification tasks such as human activity and gesture recognition by exploiting class-distinguishing cues. However, whether they can support continuous trajectory tracking has remained an open question-until now.**

In this paper, we present the first fundamental study of continuous hand trajectory tracking with wearable solar harvesters. Through a novel radiometric model, we analytically link photovoltaic (PV) power to the solar cell's geometric configuration, exposing both the promise of energy signals for tracking and their core limitations: the positional ambiguities and distortions that arise when power is used directly for positioning. To resolve this ambiguity, we propose SolarTrack, a framework that embeds a radiometric model as a physical backbone within a sequence-learning pipeline, enforcing cycle consistency between data-driven predictions and physical feasibility. This yields the first standalone solar-based tracker capable of estimating continuous hand motion directly from harvested energy signals.

To validate this, we built a wearable prototype with a solar panel and IMU and collected the first dataset pairing motion-capture ground truth with harvested energy from 15 participants (700k samples). Results show that solar signals alone achieve sub-decimeter tracking accuracy, outperforming purely data-driven baselines and only 1.6 cm worse compared to the IMU tracker despite having access to only 1D power signal. Furthermore, when fused with IMU, it further boosts IMU performance by 13%.

***Index Terms*—Hand tracking, Energy harvesting, Wearables, Solar, Deep learning**

I. Introduction

Energy harvesting–based sensing has emerged as a promising approach in wearable and IoT systems, where the same energy harvester not only powers the embedded device but also serves as a sensor. EH has been utilised in activity recognition [1], gait analysis [2], transport mode detection [3], hotword detection [4] and place identification [5].

Prior work on energy-harvesting sensing [1], [2], [5] has primarily focused on categorical learning (Fig. 1, left), where the objective is to map an energy signal sequence into one of N discrete classes. This formulation suits applications such as gesture or activity recognition [1], [6], where the objective is to identify what action occurred rather than how it unfolded in continuous space. In contrast, continuous mapping aims to regress a continuous sensory state—such as (x, y, z) coordinates—from instantaneous power measurements. This problem remains underexplored, despite its potential: for example, the continuous tracking of a wearable solar cell from EH signals could enable fine-grained applications like hand tracking [7], object reach monitoring in VR/AR [8], and beyond.

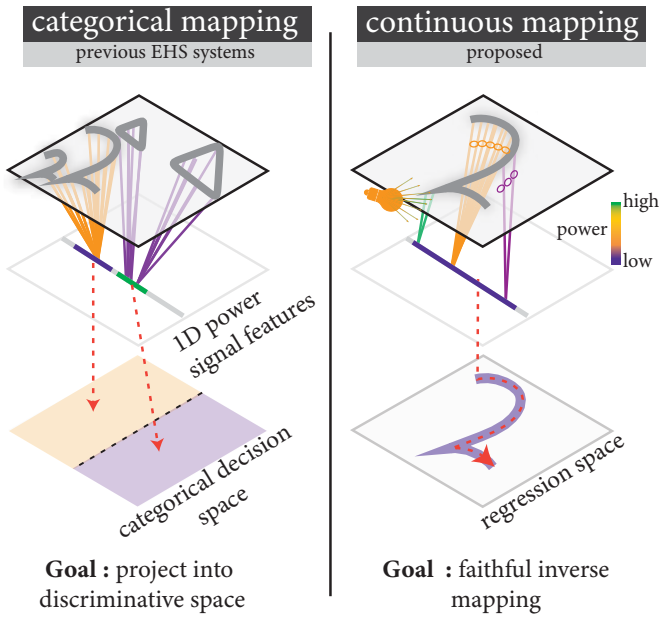


Fig. 1: **Categorical vs. continuous mapping of harvested power signals.** The top plane shows motion trajectories, the middle plane shows corresponding 1D power signals perceived by the solar cell, and the bottom plane the prediction space. (Left) Categorical mapping (e.g., gesture recognition) collapses fine-grained motion variations (e.g., variations in triangular trajectories) into class-separable clusters. (Right) Continuous mapping (our goal) seeks to reconstruct detailed trajectories, but faces ambiguity —different trajectory segments (e.g., all orange points) can produce indistinguishable power values—an ill-posed problem we address in this work.

Fig. 1 illustrates the challenge of continuous mapping. The core difficulty of regressing continuous states from EH signals lies in the intrinsic information bottleneck of the sensing modality: a rich, high-dimensional physical configuration (x, y, z) collapses into a single photocurrent measurement. For categorical tasks, this collapse is not only tolerable but often beneficial, since the goal is merely to separate classes, where

Corresponding author: yasien.ghalwash@qut.edu.au

compression is an asset rather than a liability. In contrast, continuous mapping requires preservation of fine-grained distinctions, yet the bottleneck discards precisely the information needed to disambiguate spatial states. This asymmetry explains why categorical learning has flourished in EH sensing, whereas continuous regression remains far less explored.

In this work, we study the latter regime. Given the extreme information bottleneck, one might ask whether EH signals are fundamentally capable of supporting continuous state estimation. To move beyond the abstract, we ground this inquiry in solar energy harvesting as a representative modality and pose the concrete question: *Can the position of a wearable solar cell be continuously tracked from its output signal?*

Fundamentally, the harvested photocurrent depends nonlinearly on the cell's geometry, orientation, and illumination conditions. While these dependencies are well understood in photovoltaic (PV) modelling [9], their precise characterisation in dynamic, wearable sensing settings (e.g. solar-powered smartwatch) is yet to be realised. Previous modelling effort primarily targeted non-wearable, infrastructure-based settings [6], or focused on relative displacement estimation rather than absolute position recovery [10]. To address this, we build on the fundamentals of radiative heat transfer [11, Chapter 4] and derive a radiometric[1] model that formally connects photocurrent to the solar cell's spatial configuration. Our model states that **harvested power** = **constant**×**geometric coupling** between light source and solar cell. The geometric coupling term encompasses all motion-dependent variations, while the constant term encapsulates fixed hardware properties, such as cell efficiency, cell area, and light intensity.

This formulation has two key implications. First, it demonstrates that continuous tracking is theoretically feasible. Since geometric coupling is a smooth function of the cell's relative geometry to the light source, the harvested power varies smoothly with motion as we show in Sec. III. Second, it provides a forward mapping (geometry $\rightarrow$ power), that can be leveraged in interpretable data synthesis (e.g., synthetic power traces for data-driven model training, Sec. III-B) and sensing system design by anticipating how deployments and placements shape power signatures.

The radiometric model, however, doesn't directly resolve the tracking challenge. That is, the difficulty of recovering continuous position state from 1D (compressed) power signals. While the forward relation is well-defined, the inverse power $\rightarrow$ cell position is non-unique, as many cell poses (offset, distance, orientation) can produce near-identical photocurrents. To address this, we reuse the radiometric model to characterise the cell position ambiguity and to guide mitigation strategies using the levelset concept in Sec. III-C. Notably, it reveals that temporal and physics-based consistencies can significantly alleviate the issue. Based on these findings, we develop a sequence learning pipeline that integrates physical radiometric constraints. Furthermore, we support it with a differentiable Kalman Filter to enable recovering the trajectory even under sampling rates unseen during the training. This work makes the following contributions:

- We derive and experimentally validate a *novel radiomotric model* that characterises the power variations of wearable solar cells as a function of their geometric coupling with the light source. Beyond establishing the feasibility of continuous tracking from PV measurements, we demonstrate the model's utility in simulating power traces for given trajectories and further integrate it as a physical regulator in a data-driven pipeline.
- We introduce SolarTrack, the *first system*, to the best of our knowledge, for an in-air *continuous 2D hand tracking* system solely from PV measurements.
- We present the first dataset that pairs PV power signals with ground-truth motion capture trajectories from 15 subjects ($\sim$700k temporal data points), providing a benchmark for evaluating continuous tracking approaches. SolarTrack achieves 8.5 cm median error using a solar one-dimensional signal (only 1.6 cm worse compared to a six-dimensional IMU). When fusing IMU with solar, it reduces median IMU error by 13%, demonstrating its complementarity and potential in solar-powered IMU wearables [12].
- We release our radiometric forward model code and SolarTrack dataset[2] to the community, to accelerate future efforts in EH-based tracking and sensing.

II. Related Work

A. Solar Based Sensing

Solar cells' dual nature as sensors and energy harvesters has motivated exploration beyond power generation. Prior work examined classification tasks using discrete photocurrent patterns: SolAR [13] enabled activity recognition, while others addressed gesture classification [6], [14], [15] and place identification [5], consistently treating solar signals categorically.

Researchers have explored spatially aware applications beyond classification. LuxTrace [10] pioneered solar-based indoor positioning, though limited to one-dimensional displacement monitoring. Oosterlee et al. [16] employed wall-mounted solar arrays to track people through shadow analysis, requiring fixed infrastructure. SolarSLAM [17] adopted a hybrid approach using solar photocurrent for loop closure detection to correct IMU drift in SLAM, yet still treated solar data as discrete location markers rather than continuous signals.

Despite these advancements in solar-based sensing, prior studies mostly treated solar signals as discrete cues for classification or event detection. In contrast, our work is the first to leverage wearable solar cells for continuous, high-resolution trajectory estimation by modelling photocurrent variations through a radiometric formulation, revealing their direct geometric relationship to hand position and orientation.

[1]Radiometry is the study of propagation of electromagnetic radiation in the environment, including the transfer of radiant energy between sources, surfaces, and detectors.

[2]https://solar-track.github.io/

B. Continuous Tracking Modalities

Continuous motion tracking has become essential for applications ranging from health monitoring to human-computer interaction, driving the development of diverse sensing modalities. IMUs have been used in that manner to capture the human movement by leveraging accelerometers, gyroscopes, and magnetometers to capture motion dynamics in human body movements [8], [18]–[20]. While IMUs provide high-frequency motion data with good temporal resolution, their fundamental limitation lies in compounded drift errors [21]. This has motivated researchers to explore other approaches, like radio frequency (RF), which leverage the current fixed infrastructure to provide absolute positioning. Previous work has also explored channel state information (CSI) variations to track movement [22], [23]. Despite the advantages of RF tracking, challenges persist with multipath inferences [24]. Acoustic sensing follows similar principles where sound waves encode motion information through Doppler shifts and phase changes [25], [26]. However, acoustic tracking remains vulnerable to ambient noise and requires active sound generation [27]. This opens the idea of investigating the usage of passive optical approaches. Visible light sensing (VLS) has emerged as an alternative tracking approach offering device-free sensing, with photodiodes detecting reflected or interrupted light patterns [28]. While VLS eliminates active signal generation, the photodiodes still consume power for continuous operation.

III. Modelling Solar-based Hand Trajectories

In this section, we model solar-based hand trajectories. Fig. 2 illustrates our setup. The PV element is attached to the back of the user's hand, its active face typically oriented upward toward the desk lamp. As the hand draws mid-air figures—circles, squares, letters—its 2-D position changes, and with it the photocurrent delivered by the cell. Under the assumptions of negligible ambient illumination and no significant interreflections from nearby surfaces, we show that the solar power is a deterministic, smooth function of the cell's position. From that function, we can recover one or more components of the hand trajectory, and in some special cases, the full position.

A. Radiometry-driven Modelling

We introduce a model motivated by the particular wearable setup considered, and base it on radiometric principles [29]. Conceptually, the gap between it and prior works is akin to *device-based versus device-free sensing models*. While the latter model accounts for the blocking/shadowing effect of the hand between the passive (uncontrolled) sender and receiver, the first model accounts for the correlation between the signal observed on the active on-hand receiver and the motion. Previous works [6] rely on the *dynamic occlusion* principle that doesn't carry over to wearable sensing. Similar in spirit to device-free sensing, SolarGest [6] posits that the moving hand modulates the solar cell visible fraction through partial blocking. Standard ambient lighting (i.e. illumination "coming from everywhere") is assumed and, thus, all the emitter (a lamp) properties can be abstracted away. On the contrary, in a wearable setup, the observed power changes systematically with source–receiver *distance* and *orientation*. In this case, to understand how the harvested solar power is modulated by hand motion, we need a *geometry-aware* model that links the observed power to the source strength and relative pose (distance and orientation) between a lamp and the cell. We explicitly model all these aspects.

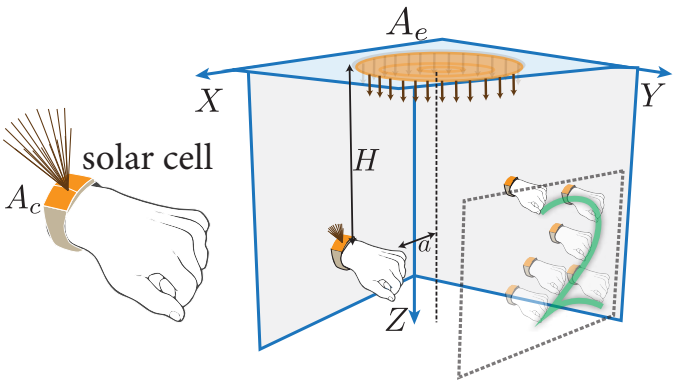


Fig. 2: **Geometric configuration of the SolarTrack.** A light emitter (lamp) of area A_e is mounted above the hand in the xy-plane at vertical separation H. The wearable solar cell of area A_c on the back of the hand is laterally offset by distance a from the emitter's axis. The coordinate system (X, Y, Z) defines displacement and orientation, and the dashed illustration on the right depicts a representative hand trajectory beneath the source.

A central tool in our analysis is the *view factor* (also called configuration/form factor) [11, Chapter 4]; a geometric coupling term long established in radiative heat transfer and later adopted by global-illumination methods in computer graphics. In heat transfer, the view factor quantifies the fraction of diffuse radiative power leaving one surface and reaching the other. *To the best of our knowledge, our work is the first to explicitly apply the view-factor formalism to sensing with on-hand wearable solar cells.* By the end of our analysis, we reach the conclusion:

$$P \;=\; \kappa\, \mathcal{F}(a, H) + \epsilon \tag{1}$$

The harvested power P is a smooth, deterministic function of the cell's lateral displacement a and vertical separation H[a], scaled by a constant κ that captures solar cell and light source properties. ϵ accounts for noise.

[a] $\forall H > 0$

This result establishes a link between the solar cell's relative position with respect to the light source and harvested power. In essence, every small change in a or H induces a predictable, smooth variation in the power trace, with no discontinuities. This property is fundamental: it provides the analytical basis for treating harvested power as a proxy for hand trajectories. While the inverse recovery of pose from power is inherently challenging (Sec. III-C), the smoothness of $\mathcal{F}(a, H)$ renders such trajectory tracking theoretically feasible. Next, we develop the formulation by first considering the idealised point-

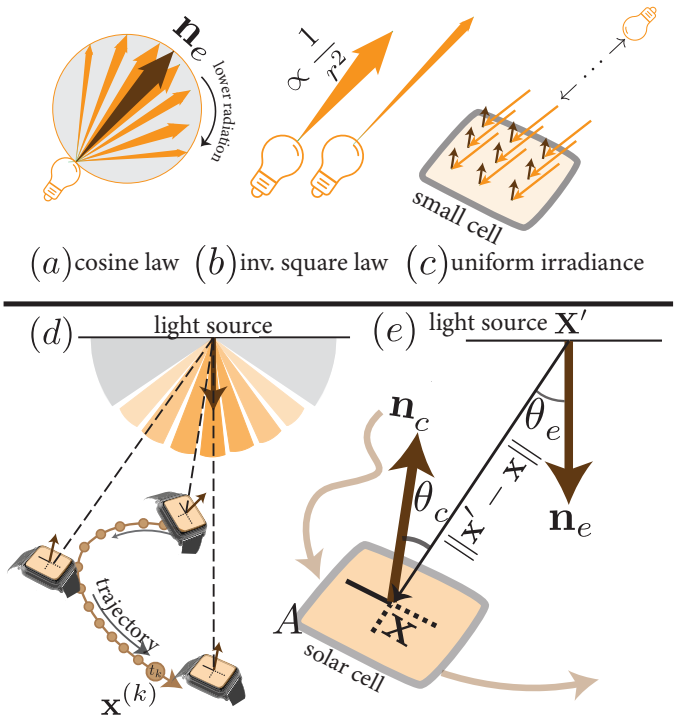


Fig. 3: **SolarTrack Setup. (Top)** Geometric dependencies governing energy received by a PV cell during hand motion. (a) light emission follows a cosine-based angular profile, peaking along the emitter's axis $\mathbf{n}_e$. (b) Irradiance decays with distance squared, following the inverse-square law. (c) A sufficiently small solar cell receives nearly uniform irradiance at any time step. **(Bottom)** A cell follows a hand trajectory, where the received power varies with source–cell distance $\|\mathbf{x}' - \mathbf{x}\|$, emission angle θ_e, and incidence angle θ_c, all tied to the hand's trajectory.

to-point case, where a point source illuminates a point cell. Although simplified, this setting provides analytical clarity. We then extend the formulation progressively to the general case of a lamp surface illuminating a cell area.

Notation. As shown in Fig. 2, we consider a circular emitter on the xy-plane of diameter D (radius $r = \frac{D}{2}$) with surface A_e and peak radiant intensity I_0. The solar cell has an area A_c and efficiency η. A point on the emitter is $\mathbf{x}' \in A_e$ and a point on the cell is $\mathbf{x} \in A_c$, separated by vertical distance H and lateral offset a; the associated emission and incidence angles are θ_e and θ_c. Using polar coordinates (r', ϕ) over A_e, the geometric attenuation between source and cell is expressed as the view factor $\mathcal{F}(a, H)$.

1) Point illuminator to cell's point: Without loss of generality, let's consider a single-point light source $\mathbf{x}'$ illuminating the solar cell according to the geometric arrangement in Fig. 3(e). From radiometry fundamentals [30, Chapter 4] [29], the irradiance (i.e. how much power the receiver surface gets) at the PV cell point $\mathbf{x}$ is:

$$I(\theta_c(\mathbf{x}', \mathbf{x})) = \underbrace{\frac{I_0}{\|\mathbf{x}' - \mathbf{x}\|^2}}_{\text{Inverse-square law}} \underbrace{\cos^m(\theta_e)}_{\text{Emission angle}} \underbrace{\cos(\theta_c)}_{\text{Incidence angle}} \tag{2}$$

where I_0 is the peak radiant intensity emitted in the forward direction $\theta_e = 0$, and $m \geq 0$ controls the beam sharpness ($m = 1$ approximates a Lambertian source). Eq. (2) very intuitively states that irradiance diminishes rapidly with distance between hand and light source (first term) and weakens at oblique emission angles (second term) and when the cell surface tilts away from the incoming direction (third term).

2) Point illuminator to cell's surface: The total power harvested by the cell can then be obtained by integrating over the finite cell surface, weighted by the PV conversion efficiency η:

$$P_{\mathbf{x}' \to A_c} = \eta \int_{\mathcal{A}_c} I(\theta_c(\mathbf{x}', x))\, dA_c(x) \tag{3}$$

for all $x \in A_c$. The cell-specific constant η (e.g., 22%) translates incident optical power into electrical power. For our wearable setting, the small-area cells (4.2cm $\times$ 3.5cm) can be reasonably assumed to have uniform irradiance across the surface, i.e., $I(\theta_c(\mathbf{x}', x)) \approx I(\theta_c(\mathbf{x}', \mathbf{x}))$ as depicted in Fig. 3(c). The integral then simplifies to:

$$P_{\mathbf{x}' \to A_c} = \eta\, A_c\, I(\theta_c(\mathbf{x}', \mathbf{x})) \tag{4}$$

3) Surface illuminator to cell's surface: Now we can obtain the final power by considering the illuminator surface. The total harvested power is obtained by adding up the contributions from every infinitesimal patch of the lamp's emitting surface. Accordingly, Equation. (4) can be updated as:

$$P_{A_e \to A_c} = \eta\, A_c \int_{\mathcal{A}_e} I(\theta_c(x', \mathbf{x}))\, dA_e(x') \tag{5}$$

where $dA_e(x')$ denotes an infinitesimal element of the emitter's surface and accounts for the patch's size. The integral ensures that all rays from the lamp surface are accumulated. After algebraic manipulation (details in Appendix A), the received power reduces to:

$$P_{A_e \to A_c} = \eta\, A_c\, I_0\, \mathcal{F}(a, H) \tag{6}$$

$$\mathcal{F} = \int_0^{D/2} \int_0^{2\pi} \frac{(\mathbf{R} \cdot \mathbf{n}_e)\,(-\mathbf{R} \cdot \mathbf{n}_c)}{R^4} r'\, d\phi\, dr' \tag{7}$$

where $\mathbf{R}$ denotes the vector from a differential lamp surface element to the solar cell, $\mathbf{n}_e$ and $\mathbf{n}_c$ are the outward normals of the emitter and the cell respectively, and the integration is carried out over a circular lamp of diameter D using polar coordinates (r', ϕ), with geometry parameterised by the lateral offset a and vertical separation H (see Appendix A). Eq. (6) recovers exactly the expression introduced in Eq. (1) with $\kappa = \eta\, A_c\, I_0$. This **orientation-aware** formulation, which explicitly accounts for the cell normal $\mathbf{n}_c$ via the view factor $\mathcal{F}$, is the model employed throughout all experiments in this work.

Notably, the radiometric model is hardware-agnostic; it depends only on the geometric relationship between light source and solar cell, which holds regardless of specific cell dimensions or efficiency. The calibration constant κ encapsulates device-specific parameters (cell efficiency, area, light intensity), enabling straightforward adaptation to different solar cell configurations through a simple calibration procedure. This generality means that the model transfers naturally to commercial solar-powered wearables.

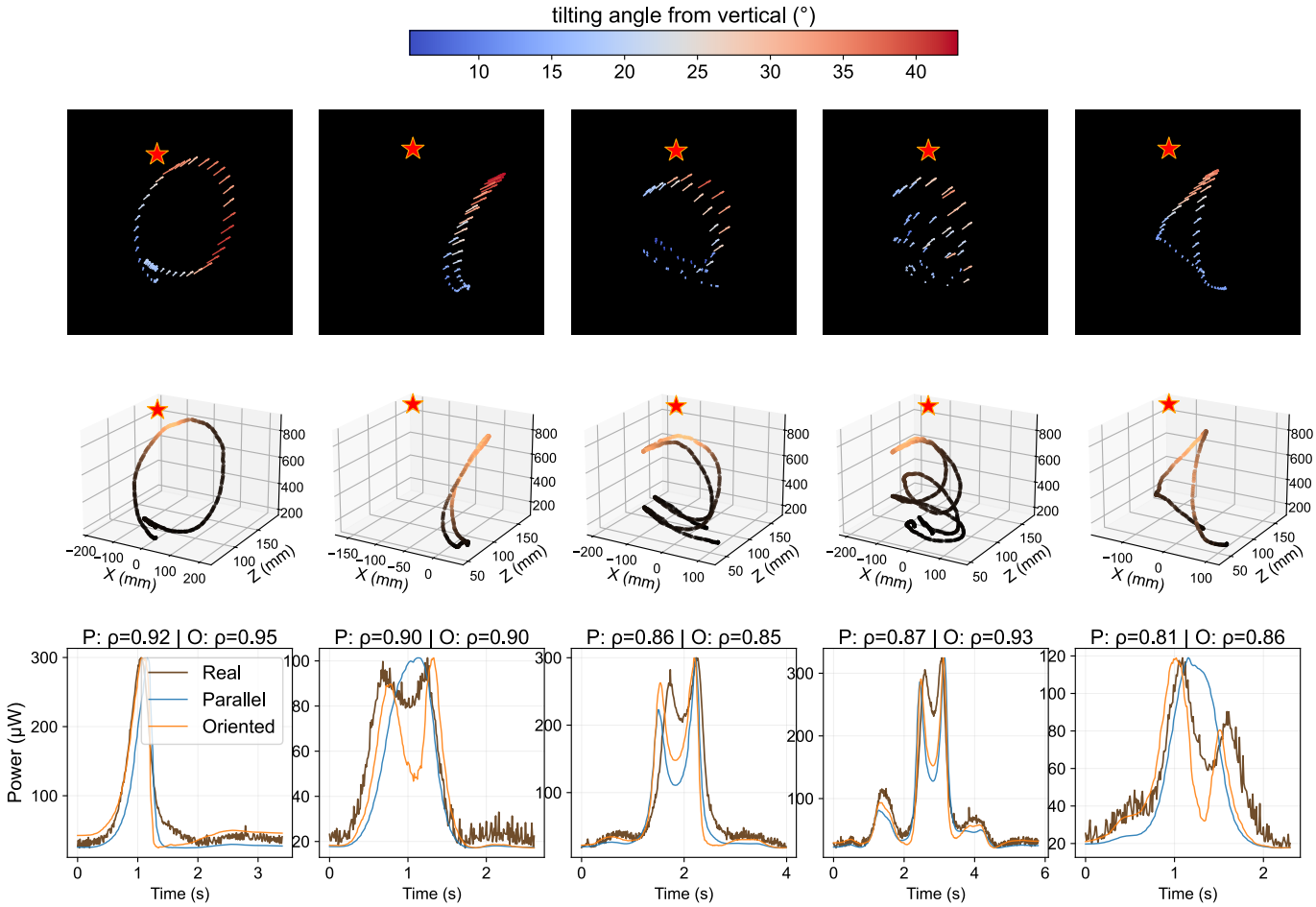


Fig. 4: **Comparison between measured solar power and simulated solar power. Top:** 3D hand trajectories with colour indicating deviation of the solar cell normal from the ideal parallel (upward) configuration (dark blue = smallest deviation, dark red = largest), highlighting substantial wrist rotations. **Middle:** Corresponding 3D trajectories showing hand motion in space. **Bottom:** Measured solar power (*Real*) compared with predictions from the simplified parallel model (*Parallel*) and the full orientation-aware model (*Oriented*); the oriented model consistently follows the real signals very closely.

What if the illuminator is spherical?: In the spherical case, the model extends straightforwardly from the disk formulation; however, the computation becomes orders of magnitude more expensive due to the need for Monte Carlo integration. The current disk-based formulation in Eq.(7), by contrast, can be evaluated efficiently using numerical quadrature (e.g., via SciPy) with negligible runtime overhead. Our experiments indicate that while the spherical model offers slightly higher accuracy at larger emitter–receiver distances, the disk approximation achieves comparable results at a fraction of the computational cost. Due to space limitations, we omit a detailed comparison.

What if we have multiple illuminators?: A key advantage of the model is that it extends naturally to multi-emitter case and the main findings of Equation (1) hold. Since irradiance contributions from m independent light sources add linearly, the harvested power generalises to $P = \sum_{i=1}^{M} \kappa_i \mathcal{F}_i(a_i, H_i)$. Importantly, the smoothness property of $\mathcal{F}$ is preserved under summation even in multi-emitter configurations. Thus, small changes in hand position/orientation yield predictable changes in the power trace.

B. Radiometric Model Experimental Validation

We collected paired solar and motion capture data for five reference trajectories (check data collection in Section VI). The motion capture stream provides the exact hand position $\mathbf{x}(t)$ and negative cell normal $-\mathbf{n}_c(t)$ (the cell on the back of the hand) at each frame, while the lamp geometry—its center $\mathbf{x}'$ and diameter D —was measured once during setup. These quantities are plugged directly into the radiometric model: the lamp-to-hand vector $\mathbf{r}(t) = \mathbf{x}' - \mathbf{x}(t)$ determines the source distance and lateral offset, the palm normal defines the incidence angle θ_c, and the lamp geometry defines the emission profile. Together, these parameters specify the irradiance received at the cell, which is then converted into predicted power. Similar to [6], a calibration step was applied to scale the simulated traces to the dynamic range of the corresponding real measurements, enabling a direct comparison of temporal patterns.

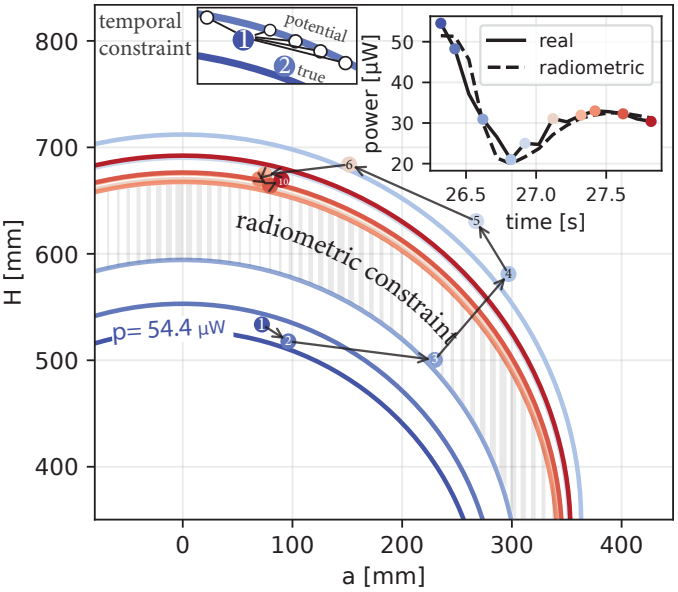


Fig. 5: **Challenge of $p \rightarrow (a, H)$ mapping and how it can be mitigated.** (Top Right) real vs radiometric simulated power traces. Numbered dots are the ground truth trajectory points, while curves are the level sets. Without priors, all the points on the colored curves are possible solutions! Temporal constraint (Top Left) allows guessing for the next point after 1 by projecting it into the next level set within a limited radius (e.g. typical hand velocity). All white circles (potential) are reasonably close to 2. Radiometric constraints steer predictions away from the regions between curves (e.g. shaded area).

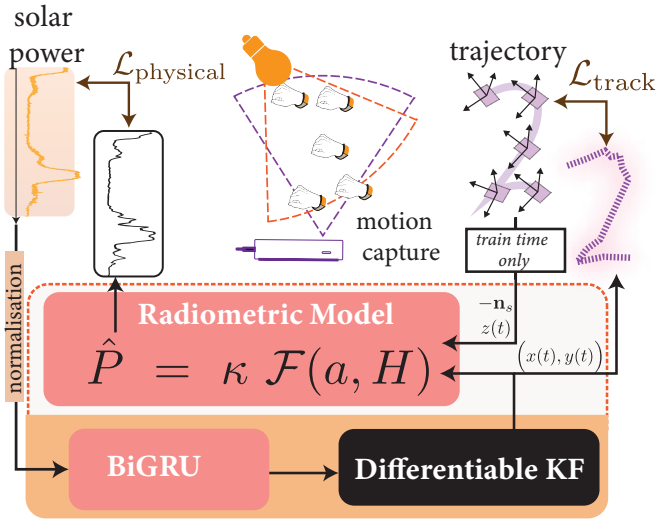


Fig. 6: **SolarTrack Pipeline.** At training, power windows pass through a BiGRU and a differentiable Kalman filter (KF) to produce a trajectory supervised by a tracking loss $\mathcal{L}_{\text{track}}$. A train-only radiometric module maps the estimated $(x(t), y(t))$ (with privileged $z(t), -\mathbf{n}_s$) back to predicted power to enforce physics consistency via $\mathcal{L}_{\text{phys}}$. At test, only the BiGRU+KF path is used (radiometric block disabled) and no privileged signal is used.

Fig. 4 shows a qualitative comparison between measured solar responses and those predicted by both the simplified parallel model and the full orientation-aware radiometric model for the three trajectories. The simulated curves closely follow the overall temporal structure of the real signals, reproducing the number, relative magnitude, and timing of the main peaks. Across all gestures, the oriented model consistently achieves higher correlation with the real signal (O: ρ=0.85–0.95) compared to the parallel model (P: ρ=0.85–0.92). This improvement is most pronounced for trajectories involving significant wrist rotation, where the parallel model's assumption of a fixed upward-facing cell leads to systematic errors. These results validate our use of the orientation-aware formulation Eq. (6) throughout this work. The simulated traces sometimes exaggerate mid-trajectory dips, especially for Two and Three, due to the uniform-irradiance and oriented-disk assumptions that over-penalize misalignment. Still, the overall dynamics closely match the real signals, capturing the key trajectory-dependent variations in harvested power.

Downstream transferability. While not the focus of this work, we further tested the radiometric model on downstream classification tasks. We generated solar measurements for ground truth trajectories and trained the classifier exclusively on the calibrated simulated solar traces. Using a simple k-NN [31] classifier with Dynamic Time Warping (DTW) [32] as the segment distance. The model reaches 83.1% average accuracy.

C. Inverse Mapping (Solar Power → Position)

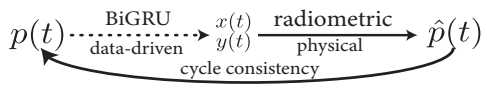
The radiometric model, in its forward form, can be used to analyze position ambiguity. That is, for a given power value, what are all the possible (a, H) pairs that could have produced it? Fig. 5 illustrates this idea. The inset (top right) shows measured and radiometric power over time; for each marked point, the corresponding set of feasible solutions is drawn as a curve of the same colour in the main plot. This curve, formally called a level set, is defined as $\mathcal{S}(p) = \{(a, H) : \kappa\mathcal{F}(a, H) = p\}$. The ground-truth trajectory from Leap Motion is overlaid as dots connected by arrows; while each dot lies close to the corresponding level set, it does not align perfectly due to sensor noise and model imperfections.

If each power value is considered in isolation, any point along its level set is a valid solution. **Temporal consistency** narrows this space by enforcing smooth motion: only nearby candidates across successive level sets remain plausible. In the inset (top left, Fig. 5), the white dots illustrate such possible continuations from point 1, which stay close to the true path and lead naturally toward point 2. **Radiometric consistency** further refines predictions by *snapping* them back to the *correct level set*, eliminating spurious estimates that fall into infeasible regions (e.g., the shaded area in Fig. 5). Both principles will be integrated into the SolarTrack model in the next section.

IV. SolarTrack

Fig. 6 depicts the complete SolarTrack framework. We formalise the tracking task as learning a mapping from observed power traces to hand positions. At each timestep t, the input sequence is a sliding window of normalised power values $p[t - w + 1 : t]$ (and their increments Δp), while the target

output is the 2D wrist location $(x(t), y(t))$. As discussed in the previous section, this setup is inherently ambiguous: a single power value corresponds to an entire level set of feasible positions. Temporal and physical constraints, however, reduce this ambiguity, since natural hand motion evolves smoothly, and measured power must remain physically consistent with the radiometric constraints. We integrate both as follows :



(8)

Intuitively, this formulation forms a cycle: power measurements $p(t)$ are first mapped to hand positions $(x(t), y(t))$ via a data-driven single layer of Bi-directional recurrent neural network (BiGRU) [33]. BiGRU attends to the raw power trace and its increments, thereby emphasising relative changes. Physical consistency is then enforced through the radiometric model, which maps the estimated positions back into the power domain (Eq. (6)) $\hat{p}(t)$. By closing the loop (Eq. (8)), we provide a self-consistency check that constrains the solution space and improves robustness. For this, the model can combine two losses: tracking loss $\mathcal{L}_{\text{track}} = \|(x(t), y(t)) - (x^*(t), y^*(t))\|^2$ and physical loss $\mathcal{L}_{\text{phys}} = \|P_{A_e \to A_c}(t) - p(t)\|^2$. There are, however, two critical considerations to realise a practical model.

1) Radiometric Integration. Leveraging Eq. (6) as a physical constraint is conceptually straightforward: given predicted position and ground-truth $z(t)$ from Leap Motion, one can compute the lateral–vertical separation (a, H) and plug them into the radiometric model to predict the corresponding power. However, evaluating it on the fly during training is computationally prohibitive, since it involves a nested double integral for each sample and update.

Fortunately, the radiometric formulation decomposes neatly into a constant κ and a geometry-dependent view factor $\mathcal{F}(a, H)$. The latter depends only on relative geometry and can be pre-computed offline over a discretized $(\Delta a, \Delta H)$ grid covering the feasible workspace of hand motion. This yields a lookup table (LUT) of view factors, so that at training time the radiometric consistency check reduces to a simple table retrieval. Meanwhile, κ is kept as a learnable parameter, allowing the model to automatically calibrate during training. Finally, it is important to note that $z(t)$ is only used as privileged information during training. At inference time the radiometric module is entirely dropped, and the tracker operates solely from harvested power traces.

2) Variable sampling rate of EH. Fixing a high sampling rate for both training and testing can maximise tracking accuracy. However, during inference, practical energy-harvesting sensing systems can have lower (and often variable) sampling rates [34]. To enable SolarTrack to generalise to unseen sampling rates, we opt to train at a low, coarse sampling rate but integrate mechanisms for tolerating missing samples. Inspired by NeuralPrefix [35], we allow our model to learn trajectory estimation from sparsely sampled input ($\approx$ 6 Hz). Unlike them, we leverage the Differentiable Kalman Filter (DKF) [36] instead of NeuralODE [37]. Specifically, we use DKF to bridge BiGRU and radiometric components. Specifically, we use DKF to bridge the BiGRU and radiometric components by interpolating motion trajectories between sparsely sampled observations, enforcing temporal smoothness and physical consistency. *Note that we validate our system (Sec. VI) on the maximum sampling rate of the dataset (100Hz), which is 16 times higher resolution compared to the training!*

Fusion. SolarTrack can support fusion with other modalities (e.g. IMU) by either appending the additional modality to the input or updating the state transition in DKF. For simplicity, we use the former in our evaluation (Sec. VI) when fusing IMU and solar.

V. Implementation

SolarTrack Hardware. We developed a custom hardware platform shown in Fig. 7 for synchronised multi-modal data acquisition, consisting of an Adafruit Feather nRF52840 microcontroller and a custom energy harvesting interface board that samples voltage/current at 100 Hz. The system integrates: (1) an IXYS SLMD121H10L solar module (4.2×3.5 cm, 4.5 g) for PV sensing, (2) a Movella DOT 9-axis IMU sampling at 120 Hz (downsampled to 100 Hz), and (3) a Leap Motion Controller providing 2D ground truth positions at 100 Hz. The solar module and IMU are co-mounted on the wrist, with participants performing gestures within the Leap Motion's field of view under a fixed lamp.

Synchronisation. We employ event-based synchronisation following the protocol established by Mao et al. [25]. Prior to each recording session, participants perform a distinctive hand-shake gesture that produces a characteristic signal across all sensors. This event serves as a temporal alignment marker, enabling post-hoc synchronisation of the three data streams. All devices, Solar, IMU, and LeapMotion, are set at a sampling rate of 100 Hz.

Training. Models were implemented in PyTorch and trained with Adam (lr = 2×10^{-3}) using gradient clipping. Each recording was temporally split into 70% train / 30% test, ensuring no overlap between splits. Input power signals were windowed into sequences of 64 samples ($\approx$ 0.64s at 100 Hz) with a stride of 16 during training and stride 1 during inference. Each step encoded the normalised power and its temporal derivative. The network backbone is a single-layer BiGRU (hidden = 64) followed by a small MLP head producing 2D position estimates per window. During inference, per-frame predictions are obtained by sliding the window at stride 1 and refined through a Kalman-based smoother to generate dense trajectories. Training was conducted on an NVIDIA RTX 6000 Ada Generation GPU, running for 500 epochs for the general model (across all subjects) and 2000 epochs for the personalised model (single-subject data).

Model Complexity. To assess on-device deployment feasibility, we report model size and computational cost in terms of MACs (multiply–accumulate operations), a standard metric in embedded-AI literature. The complete model comprises only 34,884 trainable parameters (0.13 MB): the WindowObsNet

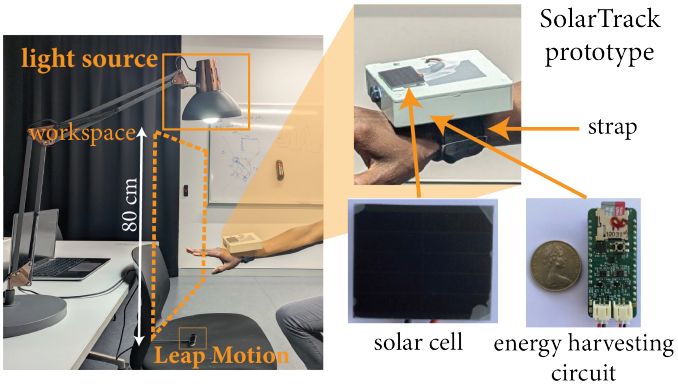


Fig. 7: **SolarTrack data collection setup.** Workspace represents a virtual whiteboard area in which trajectories are performed. Note that LeapMotion isn't exactly underneath the light source.

backbone (single-layer BiGRU with hidden size 64 followed by a two-layer MLP) accounts for 34,882 parameters, while the RadiometricLayer adds just 2 learnable scaling parameters. Per-window inference (64 samples) requires approximately 1.65 MMACs, dominated by the BiGRU (99.5%). The Kalman filter adds negligible overhead (360 MACs/timestep). These figures place our model in the same regime as recent embedded HAR systems; for instance, TinierHAR [38] reports comparable parameter budgets (≈35k) and MAC counts (1–7 MMACs) across standard activity recognition benchmarks.

VI. Experimental Evaluation

Experiment Setup. Fig. 7 shows our data collection setup in the indoor environment using the hardware platform described. Illumination was provided by a 9W LED lamp positioned 80 cm above the workspace, producing approximately 40k lux at the highest point (directly beneath the lamp) and 400 lux at the lowest measured point (above the Leap Motion sensor). Participants were instructed to perform gestures within both the illuminated region and the Leap Motion's field of view to ensure synchronised data capture across all sensors.

Experiments were conducted within an ergonomically defined "virtual whiteboard" region beneath the light source, allowing participants to perform full-range hand motions while remaining within the Leap Motion's field of view. This workspace selection does not remove radiometric symmetry, as the symmetry arises from the underlying lamp–cell geometry rather than from the spatial extent of the interaction region. Consequently, symmetric power responses persist within the chosen workspace, as reflected in the mirrored power profiles observed along circular trajectories (Fig. 4), where opposite sides intersect similar irradiance level sets despite differing spatial positions.

Dataset. We collected data from 15 participants (10 male and 5 female, age 22–35, height 155–185 cm). Ethics approval was obtained from the University's Human Research Ethics Committee[3] before starting data collection. Participants wore the device mounted on their dominant wrist with the solar cell

[3]Ethics approval: QUT Human Research Ethics Committee, Approval #9665

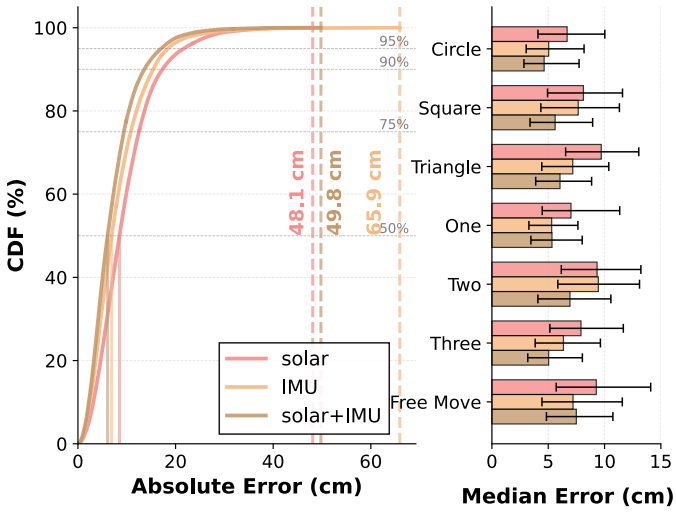


Fig. 8: SolarTrack overall performance. **(Left)** CDF for tracking error. 50%-th percentile for each measurement kind is indicated by solid vertical lines (8.5 cm, 6.9 cm and 6.1 for Solar, IMU and Solar+IMU, respectively). **Vertical dashed lines denote max error. (Right)** Error breakdown by trajectory type.

facing upward. The data collection procedure comprised two main components: 1) *Structured Gesture Set*: six predefined gestures (circle, square, triangle, number one, number two, number three) were performed, with each gesture repeated 10 times. Between repetitions, participants returned to a neutral static position at the gesture starting point and maintained this position for a 5-second rest period before continuing to the next repetition. After completing gestures in each category, the participants were free to take a break as long as they want before starting the next gesture. 2) *Free Movement Session*: A 2-minute unstructured movement period where participants performed spontaneous hand gestures without specific instructions, allowing for natural, unconstrained motion patterns. Participants were explicitly instructed to begin from arbitrary locations of their choosing, resulting in substantial variation in initial heights and lateral positions across the dataset.

Dataset Collection Summary. Each participant's session lasted approximately 20 minutes, including rest periods. The actual recorded movement time averaged one minute per gesture type for each participant. In total, each participant contributed 60 structured gesture instances (10 repetitions × 6 gestures) plus 120 seconds of free movement data. Across all participants, this yielded 900 structured gestures and 30 minutes of free movement, totalling approximately 116 minutes of motion data at 100 Hz sampling rate (∼700,000 temporal samples).

Evaluation Metrics. Unless stated otherwise, we report the absolute error distribution, defined as the per-frame Euclidean distance between predicted and ground-truth 2D positions. For aggregation, we use the Median Absolute Error —the median of these per-frame errors—which provides a robust summary of overall tracking accuracy while mitigating the effect of outliers.

Fig. 9: **Personalised Tracking Performance** . Error distribution with personalised models training (i.e. model trained/tested on the single subject data only). $S1, \cdots, S15$ denote the individual participants.

A. Solar Track Performance

1) Overall Performance: Fig. 8 presents the overall performance of SolarTrack across the modalities Solar, IMU and Solar+IMU. Note that the radiometric component is disabled during training when operating on IMU only. The cumulative distribution function (CDF, left) shows that solar-only signals reach a median error of 8.5 cm, confirming their standalone viability for continuous tracking. IMU-based tracking reduces this error to 6.9 cm, which is comparable to prior IMU-based wrist tracking reports—ranging from 7.3 cm in HOOV [8] to 11.0 cm in SmartPoser [19], with intermediate results (8–9 cm) for full-arm systems [18], [20]. Interestingly, solar reduces to a maximum of 27.1% lower than that of IMU. While it only happens in about 1% of the IMU predictions, it demonstrates the role of the radiometric constraints in bounding the maximum error. *The results are very encouraging, given the fact that we recover the trajectory in the solar case from a 1D power signal while the IMU uses a 6D signal (acceleration + gyroscope).* More importantly, in our setup, we considered only a single light source, which increases the ambiguity (compared to multiple properly positioned emitters). For example, in Fig. 4, in trajectory "Two", large portions of the curve (e.g., part of the upper arc and lower loop) are both mapped to dark brown/black shades. This degeneracy arises from the symmetry of a single emitter setup. Adding a second light source from a different viewpoint would break this tie

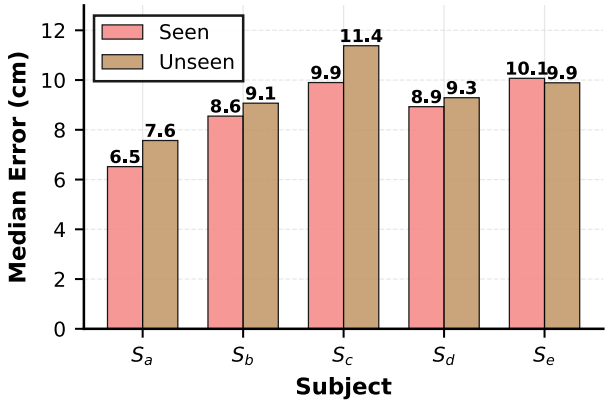


Fig. 10: **Performance on unseen subjects**

by differentiating these otherwise indistinguishable positions. CDF also shows that fusing solar and IMU achieves the best performance at 6.1 cm. This improvement is consistent in both the magnitude of general error and across trajectory types (Fig. 8, right).

2) Personalised Performance: Fig. 9 shows per-participant tracking error depicted as boxplots when we train the model only on the target subject data. Our results show that personalised training, where the model is fine-tuned for individual participants, yields consistent performance gains across all modalities—22.7% for solar-only, 28.1% for IMU-only, and 36.1% for the fused solar + IMU model—reducing median errors to 6.58 cm, 4.97 cm, and 3.88 cm, respectively. Furthermore, the 75th percentile errors show a consistent improvement across all modalities under personalised training—ranging from 12.8% for solar-only to 30.3% for solar+IMU. This indicates that personalisation not only lowers median errors but also effectively suppresses higher-tail deviations, yielding more stable performance across moderately challenging motion segments.

3) Performance on unseen users: To evaluate generalisation to unseen users, we adopt a leave-group-out protocol, where 10 subjects are randomly selected for training and the remaining five subjects are held out entirely for testing. As shown in Fig. 10, across unseen participants, the system maintains consistent generalisation capability with only a modest degradation of +0.65 cm in median error ($\approx$ 7.3 % relative increase) compared to seen subjects, averaging 9.12 cm overall. Unseen-subject performance closely tracks that of seen subjects for most participants, with increases typically within 1–1.5 cm. This demonstrates that the model generalises well to new users without retraining.

B. Sensitivity to multi-emitter Interference

Real-world deployment of SolarTrack inevitably exposes the system to illumination conditions that deviate from the controlled single-emitter environment used during training. To assess robustness under such conditions, we evaluate the pre-trained model in the presence of unseen illumination interference, without any additional calibration, fine-tuning, or retraining. This constitutes a zero-shot generalisation setting,

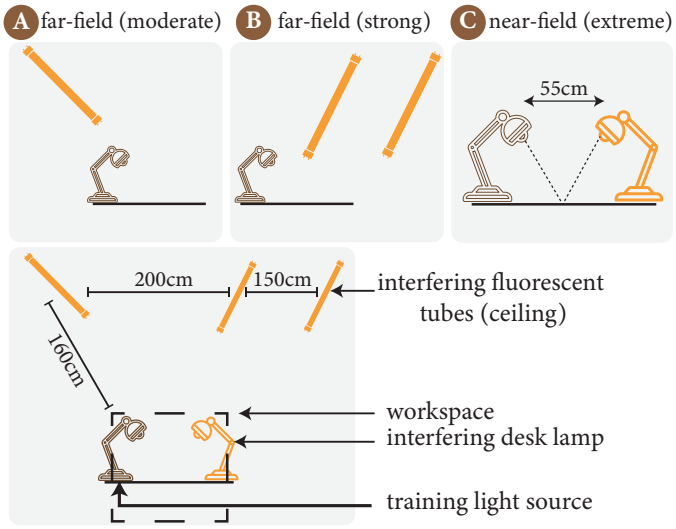


Fig. 11: **Illumination interference.** The model is trained with a single near-field desk lamp (brown). We evaluate it under three unseen interference conditions of increasing severity: (A) moderate far-field interference from one ceiling tube, (B) strong far-field interference from two ceiling tubes, and (C) extreme near-field interference from a competing desk lamp placed 55 cm from the training light source.

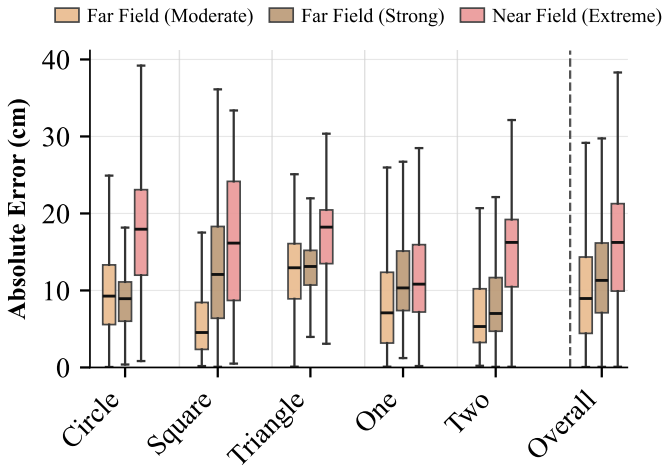


Fig. 12: **Zero-shot evaluation of illumination interference.**

where the model must operate under illumination regimes that violate its training assumptions.

We design a set of interference scenarios that progressively increase illumination ambiguity while isolating distinct physical failure modes (Fig. 11). Specifically, we introduce far-field interference using ceiling-mounted fluorescent tubes at two levels of strength: a moderate condition with a single tube (A) and a strong condition with two tubes (B), which primarily introduce global illumination bias. We then introduce an extreme near-field interference condition (C) by activating a second desk lamp positioned 55cm from the primary training light source, creating a competing emitter with comparable proximity and directional influence. This final configuration represents a worst-case scenario in which the single-emitter assumption is fundamentally violated.

Fig. 12 summarises the prediction error under the three zero-shot illumination interference regimes. As expected, performance degrades monotonically with increasing interference severity. Under moderate far-field interference, the median error increases only marginally to 8.95 cm (+5.3%), indicating that the system remains largely robust to moderate, globally additive illumination bias. Increasing the far-field interference strength leads to a more noticeable degradation (11.31 cm, +33.0%), reflecting the growing influence of global illumination on the learned representation, while still preserving stable operation.

In contrast, the near-field extreme interference condition results in a substantial increase in error (16.24 cm, +91.0%). This scenario was intentionally designed to stress the system beyond its training assumptions. In such a regime, correct operation would naturally require re-training using an appropriate multi-emitter loss (discussed in Sec. III-A3), rather than the single-emitter formulation employed during training.

C. Ablation Study

We assess the contribution of each component in SolarTrack through a progressive ablation, incrementally adding modules to a baseline BiGRU network as shown in Table I. The baseline, trained solely on solar power sequences, achieves an MAE of 8.2 cm, indicating that while temporal dependencies are learnable, the absence of physical or motion constraints limits accuracy. Adding the Differentiable Kalman Filter (DKF) improves performance to 7.7 cm (6% gain) by enforcing temporal smoothness and consistent motion dynamics. Replacing the DKF with the radiometric physical model yields 6.9 cm MAE (16% improvement), as the radiometric loss constrains predictions to configurations that reproduce measured power through the view-factor formulation (Eq. 6). Combining all components—BiGRU, DKF, and the radiometric constraint—achieves the best performance of 6.5 cm MAE, representing a 21% reduction over the baseline and underscoring the complementary effects of learned temporal features, motion dynamics, and physical consistency.

TABLE I: Ablation Study: Component Contributions

Configuration	MAE (cm)	Δ MAE	Improv.
BiGRU (baseline)	8.2	–	–
BiGRU + KF	7.7	−0.5	6%
BiGRU + Radiometric	6.9	−1.3	16%
SolarTrack (Full)	**6.5**	−1.7	**21%**

MAE = Mean Absolute Error averaged across 15 participants. Δ represents absolute error reduction from BiGRU baseline.

VII. Discussion and Future Directions

The development and evaluation of SolarTrack provided practical insights into continuous tracking using energy-harvesting signals. In this section, we reflect on SolarTrack's operational boundaries and summarise relevant future directions for extension.

Multi-emitter Deployment. The zero-shot interference results in Sec. VI-B show that the pre-trained single-emitter

system remains stable under far-field interference but requires re-training under extreme near-field interference, marking the limit of validity of the single dominant emitter assumption. Addressing this regime is naturally supported by the radiometric formulation through linear superposition of irradiance contributions. In realistic deployments, however, emitter configurations are rarely known a priori. A key practical extension is therefore the automatic localisation and parametrisation of dominant light sources in a new environment, rather than assuming known emitter geometry. This could be achieved through a brief, structured calibration motion (e.g., a circular or sweeping trajectory), from which emitter directions and relative strengths can be inferred from characteristic temporal power patterns. Such a self-calibrating multi-emitter extension would enable the system to dynamically adapt to unseen illumination regimes, while preserving the same sensing hardware and model structure.

Dominant Emitter's Characteristics. While the radiometric formulation separates geometric effects from emitter-dependent properties via the factor κ in Eq. (1), we do not systematically evaluate the sensitivity of the system to variations in emitter intensity or spectral characteristics. Although operating on normalised solar power and its temporal derivative is expected to reduce sensitivity to brightness changes, this effect has not yet been quantified. Likewise, different lamp shapes can be accommodated by modifying the emitter geometry in the view-factor term, but their impact on robustness remains unexamined. Addressing these open questions will be an important direction for future work.

VIII. Conclusion

We presented SolarTrack, the first framework for continuous hand trajectory tracking using energy-harvesting signals alone from wearable solar cells. We developed a novel radiometric model that analytically links PV power to spatial geometry and integrated it within a sequence-learning pipeline for physically consistent motion estimation. Experiments demonstrate that solar energy harvesting is a viable sensing modality for continuous motion tracking.

IX. Acknowledgment

The authors would like to thank QUT Energy Transition Centre for their support in conducting this research.

Appendix

A. Derivation of SolarTrack radiometric model

Recall from Section III-A3 that the total power accumulated by the cell is:

$$P_{A_e \to A_c} = \eta\ A_c \int_{\mathcal{A}_e} I(\theta_c(x', \mathbf{x}))dA_e(x') \tag{A.1}$$

$$= \underbrace{\eta A_c I_0}_{\kappa} \underbrace{\int_{\mathcal{A}_e} \frac{1}{R^2}\ \cos(\theta_e)\ \cos(\theta_c)\ dA_e}_{\text{View Factor } \mathcal{F}} \tag{A.2}$$

Our goal is to estimate the view factor $\mathcal{F}$ in Equation (A.2) by identifying the terms dA_e, R^2, $\cos(\theta_e)$ and $\cos(\theta_c)$ according to the geometry in Figure A.1. We express all the components in terms of our coordinate system. To integrate over the lamp surface, we define an arbitrary point on its surface at polar coordinates (r', ϕ). In polar coordinates, the differential area element on the disk is $dA_e = r'dr'd\phi$. Substituting in Equation (A.2), the view factor integral can be rewritten as:

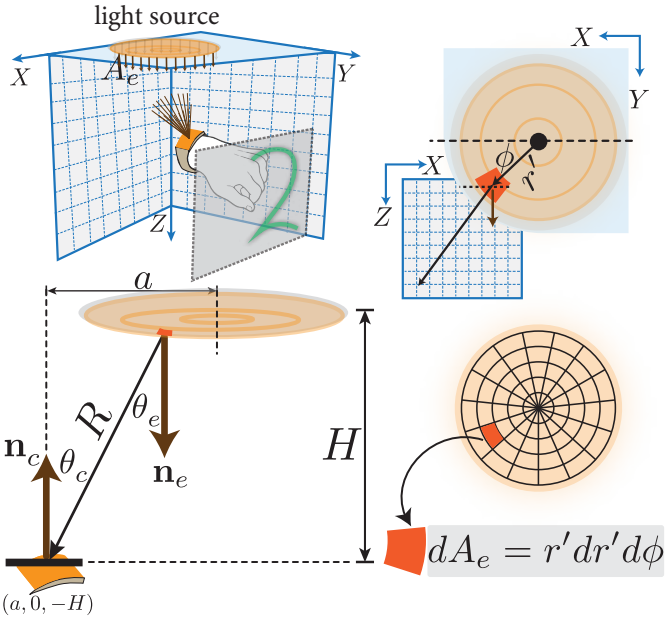


Fig. A.1: **SolarTrack model with a circular disc light source.** The lamp has area of A_e, a diameter D (radius $r = \frac{D}{2}$) and lies on the XY-plane, centred at the origin $(0, 0, 0)$. The cell is located at $(a, 0, -H)$ (top right) Notice the background (solid blue vs dashed grid) that indicates the transition from XY-plane to XZ-plane.

$$\mathcal{F} = \int_0^{\frac{D}{2}} \int_0^{2\pi} \frac{\cos(\theta_e)\ \cos(\theta_c)}{R^2}\ r'\ d\phi\ dr' \tag{A.3}$$

The angles depend on the dot product between the vector $\mathbf{R}$ and the connecting normals; $\cos(\theta_e) = \frac{\mathbf{R}\mathbf{n}_e}{R}$ and $\cos(\theta_c) = \frac{-\mathbf{R}\mathbf{n}_c}{R}$. Substituting all the terms in Eq.(A.3):

$$\mathcal{F} = \int_0^{D/2} \int_0^{2\pi} \frac{(\mathbf{R}\cdot\mathbf{n}_e)\,(-\mathbf{R}\cdot\mathbf{n}_c)}{R^4} r'\ d\phi\ dr'. \tag{A.4}$$

Parallel source/cell case. For the special case where the lamp and cell are parallel, the geometry simplifies as the cell normal becomes constant and opposite to the lamp normal. Also $\cos(\theta_e) = \cos(\theta_c) = \frac{H}{R}$. R^2 is the squared distance from the cell's point at $(a, 0, -H)$ to an arbitrary point on the disk at $(r'\cos\phi, r'\sin\phi, 0)$. Expanding and simplifying using $\cos^2\phi + \sin^2\phi = 1$, we get $R^2 = r'^2 - 2ar'\cos\phi + a^2 + H^2$. Substituting these simplifications into Eq.(A.3) gives:

$$\bar{\mathcal{F}} = H^2 \int_0^{D/2} \int_0^{2\pi} \frac{r'}{\left(r'^2 - 2ar'\cos\phi + a^2 + H^2\right)^2}\, d\phi\, dr'. \tag{A.5}$$

REFERENCES

[1] M. M. Sandhu, S. Khalifa, K. Geissdoerfer, R. Jurdak, M. Portmann, and B. Kusy, "Fusedar: Energy-positive human activity recognition using kinetic and solar signal fusion," *IEEE Sensors Journal*, vol. 23, no. 11, pp. 12 411–12 426, 2023.

[2] W. Xu, G. Lan, Q. Lin, S. Khalifa, M. Hassan, N. Bergmann, and W. Hu, "Keh-gait: Using kinetic energy harvesting for gait-based user authentication systems," *IEEE Transactions on Mobile Computing*, vol. 18, no. 1, pp. 139–152, 2018.

[3] G. Lan, W. Xu, S. Khalifa, M. Hassan, and W. Hu, "Transportation mode detection using kinetic energy harvesting wearables," in *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*. IEEE, 2016, pp. 1–4.

[4] S. Khalifa, M. Hassan, and A. Seneviratne, "Feasibility and accuracy of hotword detection using vibration energy harvester," in *2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*. IEEE, 2016, pp. 1–9.

[5] Y. Umetsu, Y. Nakamura, Y. Arakawa, M. Fujimoto, and H. Suwa, "Ehaas: Energy harvesters as a sensor for place recognition on wearables," in *2019 IEEE International Conference on Pervasive Computing and Communications (PerCom*. IEEE, 2019, pp. 1–10.

[6] D. Ma, G. Lan, M. Hassan, W. Hu, M. B. Upama, A. Uddin, and M. Youssef, "Solargest: Ubiquitous and battery-free gesture recognition using solar cells," in *The 25th annual international conference on mobile computing and networking*, 2019, pp. 1–15.

[7] J. Li, X. Liu, Z. Wang, T. Zhang, S. Qiu, H. Zhao, X. Zhou, H. Cai, R. Ni, and A. Cangelosi, "Real-time hand gesture tracking for human–computer interface based on multi-sensor data fusion," *IEEE Sensors Journal*, vol. 21, no. 23, pp. 26 642–26 654, 2021.

[8] P. Streli, R. Armani, Y. F. Cheng, and C. Holz, "Hoov: Hand out-of-view tracking for proprioceptive interaction using inertial sensing," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–16.

[9] M. A. Green, "Solar cells: operating principles, technology, and system applications," 1981.

[10] J. Randall, O. Amft, J. Bohn, and M. Burri, "Luxtrace: Indoor positioning using building illumination," *Personal and ubiquitous computing*, vol. 11, no. 6, pp. 417–428, 2007.

[11] M. F. Modest and S. Mazumder, *Radiative heat transfer*. Academic press, 2021.

[12] M. Magno, X. Wang, M. Eggimann, L. Cavigelli, L. Benini *et al.*, "Infiniwolf: Energy efficient smart bracelet for edge computing with dual source energy harvesting," in *Proceedings of the 2020 Design, Automation and Test in Europe Conference and Exhibition, DATE 2020*. Institute of Electrical and Electronics Engineers Inc., 2020, pp. 342–345.

[13] M. M. Sandhu, S. Khalifa, K. Geissdoerfer, R. Jurdak, and M. Portmann, "Solar: Energy positive human activity recognition using solar cells," in *2021 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2021, pp. 1–10.

[14] A. Varshney, A. Soleiman, L. Mottola, and T. Voigt, "Battery-free visible light sensing," in *Proceedings of the 4th ACM Workshop on Visible Light Communication Systems*, 2017, pp. 3–8.

[15] Z. Liu, J. Zhu, J. Yang, Y. Sun, Y. Yang, and J. Luo, "Reflexgest: Recognizing hand gestures under vlc-capable lamps," *IEEE Transactions on Mobile Computing*, 2025.

[16] O. Oosterlee, T. Xu, and M. A. Z. Zamalloa, "Inti: Indoor tracking with solar cells." in *EWSN*, 2022, pp. 138–149.

[17] B. Wei, W. Xu, C. Luo, G. Zoppi, D. Ma, and S. Wang, "Solarslam: Battery-free loop closure for indoor localisation," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 4485–4490.

[18] Y. Liu, Z. Li, Z. Liu, and K. Wu, "Real-time arm skeleton tracking and gesture inference tolerant to missing wearable sensors," in *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, 2019, pp. 287–299.

[19] N. Devrio, V. Mollyn, and C. Harrison, "Smartposer: Arm pose estimation with a smartphone and smartwatch using uwb and imu data," in *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. ACM, 2023, pp. 1–14.

[20] S. Shen, M. Gowda, and R. R. Choudhury, "Closing the gaps in inertial motion tracking," in *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*. ACM, 2018, pp. 429–444.

[21] F. Wittmann, O. Lambercy, and R. Gassert, "Magnetometer-based drift correction during rest in imu arm motion tracking," *Sensors*, vol. 19, no. 6, p. 1312, 2019.

[22] D. Wu, Y. Zeng, R. Gao, S. Li, Y. Li, R. C. Shah, H. Lu, and D. Zhang, "Witraj: Robust indoor motion tracking with wifi signals," *IEEE Transactions on Mobile Computing*, vol. 22, no. 5, pp. 3062–3078, 2021.

[23] D. Ma, Z. Sun, Z. Wei, Y. Guo, Y. Lei, Z. Wang, Z. Yu, and B. Guo, "Adapttrack: A robust tracking system for complex environments based on wifi device selection strategy," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 9, no. 3, Sep. 2025.

[24] C. Chen, G. Zhou, and Y. Lin, "Cross-domain wifi sensing with channel state information: A survey," *ACM Comput. Surv.*, vol. 55, no. 11, Feb. 2023.

[25] W. Mao, M. Wang, W. Sun, L. Qiu, S. Pradhan, and Y. C. Chen, "Rnn-based room scale hand motion tracking," in *Proceedings of the 25th Annual International Conference on Mobile Computing and Networking*. ACM, 2019, pp. 1–16.

[26] W. Han and J. Jia, "Advancing real-time 3d hand trajectory tracking using acoustic signals," *Measurement*, p. 118197, 2025.

[27] M. Liu, L. Cheng, K. Qian, J. Wang, J. Wang, and Y. Liu, "Indoor acoustic localization: A survey," *Human-centric Computing and Information Sciences*, vol. 10, no. 1, p. 2, 2020.

[28] N. Faulkner, F. Alam, M. Legg, and S. Demidenko, "Watchers on the wall: Passive visible light-based positioning and tracking with embedded light-sensors on the wall," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 5, pp. 2522–2532, 2019.

[29] F. Hengstberger, *Absolute radiometry: electrically calibrated thermal detectors of optical radiation*. Elsevier, 2012.

[30] M. Pharr, W. Jakob, and G. Humphreys, *Physically based rendering: From theory to implementation*, 2023.

[31] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE transactions on information theory*, vol. 13, no. 1, pp. 21–27, 1967.

[32] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series," in *Proceedings of the 3rd international conference on knowledge discovery and data mining*, 1994, pp. 359–370.

[33] M. Schuster and K. P. Kuldip, "Bidirectional recurrent neural networks. ieee transactions on signal processing," *Bidirectional recurrent neural networks, IEEE Transactions on Signal Processing*, 1997.

[34] M. M. Sandhu, K. Geissdoerfer, S. Khalifa, R. Jurdak, M. Portmann, and B. Kusy, "Towards energy positive sensing using kinetic energy harvesters," in *2020 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2020, pp. 1–10.

[35] A. Khamis and S. Khalifa, "Neuralprefix: A zero-shot sensory data imputation plugin," in *2025 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2025, pp. 171–182.

[36] T. Haarnoja, A. Ajay, S. Levine, and P. Abbeel, "Backprop kf: Learning discriminative deterministic state estimators," *Advances in neural information processing systems*, vol. 29, 2016.

[37] R. T. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, "Neural ordinary differential equations," *Advances in neural information processing systems*, vol. 31, 2018.

[38] S. Bian, M. Liu, V. F. Rey, D. Geissler, and P. Lukowicz, "Tinierhar: Towards ultra-lightweight deep learning models for efficient human activity recognition on edge devices," in *Proceedings of the 2025 ACM International Symposium on Wearable Computers*, 2025, pp. 163–169.